Cours d'algèbre

Roger Godement

COLLECTION ENSEIGNEMENT DES SCIENCES

HERMANN ÉDITEURS

COURS D'ALGÈBRE

COLLECTION

ENSEIGNEMENT DES SCIENCES

ROGER GODEMENT
PROFESSEUR A LA FACULTÉ DES SCIENCES DE PARIS

Cours d'algèbre

TROISIÈME ÉDITION MISE A JOUR

HERMANN
115, Boulevard Saint-Germain, Paris 6

TABLE

PRÉFACE

L'introduction, dans les programmes de Mathématiques des Facultés des Sciences, de notions d'algèbre relativement modernes et étendues, a rendu urgente la rédaction, en français, d'un ouvrage de référence accessible aux débutants; le livre que nous présentons ici est une tentative pour combler cette lacune et nous allons esquisser les principales idées qui nous ont guidé en le rédigeant.

Cet ouvrage, tout d'abord, est basé sur le cours enseigné par l'auteur, à Paris, dans le cadre du certificat de Mathématiques Générales : on a donc fait en sorte que sa lecture ne demande pas d'autres connaissances que celles qu'on pourrait acquérir dans l'Enseignement Secondaire. En même temps, on a ajouté les quelques compléments nécessaires pour couvrir à peu près le programme d'Algèbre de la licence (il s'agit, dans l'état actuel des choses, du certificat de Mathématiques I), qui ne diffère pas sensiblement de celui de Mathématiques Générales. Cet ouvrage devrait donc pouvoir être utilisé pendant plusieurs années par les étudiants en Mathématiques, pures ou appliquées, tout en couvrant approximativement les besoins de ceux qui s'intéresseront à la Physique plus ou moins théorique; après en avoir assimilé le contenu, le lecteur pourra entrer de plain-pied dans les traités spécialisés vraiment sérieux, mais nous n'espérons évidemment pas que le lecteur normal sera dans cette situation un an après son entrée dans une Faculté des Sciences !

Les sujets traités sont ceux que tout le monde s'accorde aujourd'hui à considérer comme indispensables aux futurs mathématiciens ou physiciens : ensembles et fonctions; groupes, anneaux, corps, nombres complexes; espaces vectoriels, applications linéaires, matrices; espaces vectoriels de dimension finie, systèmes d'équations linéaires, déterminants, formules de Cramer; polynômes, fractions rationnelles, équations algébriques; réduction des matrices. Le choix de ces sujets reflète évidemment l'évolution des Mathématiques dans les cinquante dernières années, mais nous avons pensé que cette évolution devait aussi se traduire par l'emploi d'un style qui, jusqu'alors, était réservé aux ouvrages destinés aux mathématiciens professionnels.

Beaucoup de gens, et notamment la plupart de ceux qui se bornent à *utiliser* les Mathématiques, prétendent que lorsqu'on écrit pour les débutants il est inutile, ou

même nuisible, d'essayer de faire preuve d'une trop grande rigueur, de tout démontrer, d'introduire des notions trop générales, d'utiliser une terminologie strictement définie et dépourvue de discours fleuris. S'ils avaient raison, cela voudrait dire que, *contrairement* aux mathématiciens professionnels, et au bon sens, les débutants comprennent d'autant plus facilement un texte mathématique qu'il est plus mal rédigé. Les latinistes professionnels, c'est leur métier, comprennent les inscriptions tronquées qu'on extrait tous les jours du sous-sol de l'Italie, mais il n'est encore venu à l'idée d'aucun professeur de Latin de les utiliser pour enseigner cette langue aux débutants — on préfère avoir recours à des grammaires bien écrites... Il en est de même en Mathématiques, et lorsqu'il s'agit d'interpréter le sens d'une définition obscure, de compléter une démonstration insuffisante, ou de déceler les véritables raisons d'un théorème, on ne peut pas raisonnablement espérer que le débutant fasse preuve du même flair que le professionnel.

Il faut du reste observer que les progrès réalisés depuis le début du siècle donnent maintenant à ceux qui désirent le faire la possibilité de renouveler substantiellement l'enseignement des Mathématiques, soit en apportant de nouvelles notions simples et générales qui permettent d'élargir considérablement la portée des raisonnements traditionnels, soit en mettant à jour des raisonnements qui rendent accessibles aux débutants des résultats considérés autrefois comme difficiles; et le souci de la rigueur dont les grands spécialistes de la Théorie des Nombres avaient toujours fait preuve, après s'être répandu depuis une trentaine d'années dans toutes les branches des Mathématiques, gagne maintenant, encore qu'avec des fortunes très diverses, les auteurs de manuels scolaires, au point que certains d'entre eux sont nettement en avance sur l'ensemble des professionnels eux-mêmes... Ce renouvellement, et les exagérations qui l'accompagnent parfois, ne vont pas sans soulever les protestations de certains utilisateurs, vexés d'avoir de la peine à comprendre les manuels de leurs enfants; et l'on entend parfois reprocher aux mathématiciens d'exagérer l'importance de leurs contributions en détournant l'attention des débutants des problèmes plus concrets. Il y a là, sans aucun doute, un fond de vérité; mais alors que devrait-on dire des spécialistes de la « recherche spatiale », par exemple, qui trouvent tout naturel de réclamer des sommes gigantesques pour aller reconnaître Venus, alors que, sous leurs yeux, des centaines de millions d'hommes en sont encore à tenter de ne pas mourir de faim? Les Mathématiques ont du moins l'avantage d'être bon marché...

Au risque de provoquer, chez certains, les sentiments d'horreur et de consternation que Paolo Uccello a si merveilleusement représentés dans la *Profanation de l'Hostie*, il nous faut bien dire du reste, car la question se pose de plus en plus, notre désaccord avec les nombreuses personnalités qui, actuellement, demandent aux scientifiques en général, et aux mathématiciens en particulier, de former les milliers de techniciens dont nous aurions, paraît-il, besoin de toute urgence pour survivre. Les choses étant ce qu'elles sont, il nous semble que, dans les « grandes » nations sur-développées scientifiquement et techniquement où nous vivons, le premier devoir des mathématiciens, et de beaucoup d'autres, serait plutôt de fournir ce qu'on ne leur demande pas — à savoir des hommes capables de réfléchir par eux-mêmes, de

dépister les arguments faux et les phrases ambiguës, et aux yeux desquels la diffusion de la vérité importerait infiniment plus que, par exemple, la Télévision planétaire en couleurs et en relief : des hommes libres, et non pas des robots pour technocrates. Il est tristement évident que la meilleure façon de former ces hommes qui nous manquent n'est pas de leur enseigner les sciences mathématiques et physiques, ces branches du savoir où la bienséance consiste, en premier lieu, à faire semblant d'ignorer jusqu'à l'existence même de problèmes humains, et auxquelles nos sociétés hautement civilisées accordent, ce qui devrait paraître louche, la première place. Mais même en enseignant des Mathématiques, on peut du moins essayer de donner aux gens le goût de la liberté et de la critique, et les habituer à se voir traités en êtres humains doués de la faculté de comprendre.

Pour en revenir aux débutants auxquels ce livre s'adresse, nous avons donc cherché à leur parler le langage des mathématiciens professionnels, en définissant tous les termes techniques, clairement et une fois pour toutes, en énonçant explicitement tous les théorèmes et. à quelques exceptions près imposées pour rester dans des limites raisonnables, en les démontrant tous complètement (*).

On s'est également efforcé d'établir des théorèmes aussi généraux que possible tout en respectant une règle évidente — à savoir que, pour les débutants, une généralisation est nuisible si elle oblige à compliquer substantiellement la démonstration d'un résultat simple, ou bien si l'on ne s'en sert pas effectivement dans la pratique. C'est ainsi qu'en Algèbre linéaire — question que l'on réduit généralement à l'étude des espaces vectoriels réels de dimension finie — nous avons toujours adopté au minimum le point de vue des espaces vectoriels sur un corps commutatif quelconque, ou même non commutatif là où l'hypothèse de commutativité n'avancerait à rien; et les notions les plus simples, celles qui n'utilisent que des additions et des multiplications, sont même exposées pour les modules sur un anneau quelconque, qui jouent dans toutes les branches des Mathématiques autres que l'Analyse un rôle au moins aussi important que les espaces vectoriels puisque la notion de module contient, entre autres, celle de groupe commutatif. Les simplifications qu'on aurait pu apporter au texte en se limitant aux espaces vectoriels réels seraient beaucoup plus que compensées par la perte de généralité qu'impliquerait une telle limitation. Or c'est justement la possibilité d'apprendre *sans effort supplémentaire* des résultats de plus en plus généraux qui permet aux jeunes de parvenir aussi rapidement au niveau de la recherche qu'il y a cent ans, malgré la formidable accumulation de découvertes à laquelle on a assisté depuis lors.

L'absence d'*Exercices* réduirait à peu de choses l'utilité d'un ouvrage destiné aux débutants; on en trouvera donc ici plusieurs centaines, les uns très faciles,

(*) La quasi-totalité des énoncés non démontrés se trouve dans les §§ 0 à 5; il est évidemment hors de question, pour des débutants, d'exposer la théorie des ensembles et la logique formelle sans admettre de nombreux résultats « évidents ». Le § 0 sur le raisonnement logique a pour but non seulement de signaler au débutant les raisonnements « permis » et ceux qui ne le sont pas (ce que la lecture des copies d'examen rend indispensable), mais aussi de lui montrer que la « philosophie des Mathématiques » ne se réduit pas nécessairement à un verbalisme sans structure.

d'autres plus difficiles (ils sont précédés d'un signe ¶), d'autres encore destinés aux gens vraiment courageux (ils sont précédés du signe ¶¶); certains sont des exercices de calcul pratique ou même numérique, d'autres permettent au lecteur de se familiariser avec les concepts abstraits, d'autres encore apportent au texte des compléments importants et laissent deviner au lecteur les théories plus substantielles de l'Algèbre moderne, comme on disait il y a trente ans. On peut aussi conjecturer que certains énoncés sont faux — on n'a encore jamais vu un traité de Mathématiques échapper totalement à cette regrettable possibilité; les énoncés faux sont du reste souvent les plus instructifs.

Enfin, et contrairement à toutes les traditions en matière d'ouvrages destinés aux débutants, on a cru devoir présenter au lecteur une Bibliographie formée de titres soigneusement choisis, et dont beaucoup sont dus à des mathématiciens de première grandeur. Il nous paraît utile que le lecteur se procure et utilise quelques-uns de ces livres, afin de prendre connaissance d'autres points de vue possibles, et de s'habituer à *consulter* des livres.

Ces ouvrages, pour la plupart étrangers, contribueront peut-être d'autre part à faire prendre conscience à beaucoup de jeunes gens, mystifiés dès l'âge de vingt ans par une propagande écrasante, du fait que même en négligeant les « peuplades inférieures » de nos grands parents, les Français ne forment qu'un ilot de cinquante millions d'hommes au milieu d'un océan de 700 millions de Blancs; or ceux-ci vont comme nous à l'école dès l'âge de six ans, et y restent même, dans certains pays, plus longtemps que nous. Il est facile d'en déduire que les meilleurs ouvrages de Mathématiques (par exemple) ont environ une chance sur quatorze d'être écrits par des gens « bien de chez nous », et c'est justement ce que l'expérience confirme en ce qui concerne l'Algèbre élémentaire. Nous nous en voudrions de ne pas le faire savoir alors que certains jeunes, qui ne sont pourtant pas, eux, responsables des quelques centaines de milliers de cadavres qui encombrent les consciences de leurs pères, se laissent gagner par le nationalisme, le racisme et la xénophobie.

Juillet 1962

Théorie des ensembles

Le but des §§ 0 à 5 est d'introduire les notions d'ensemble et de fonction, sans lesquelles on ne peut *rien* faire en Mathématiques — et avec lesquelles, au contraire, on peut *tout* faire. Ces notions ne se sont pas dégagées avant la fin du siècle dernier, tout au moins sous la forme générale qu'on trouvera ici; auparavant, on ne parlait pas explicitement d'ensembles, et la notion de fonction recouvrait plusieurs espèces différentes d'objets soumis à des limitations imposées par le développement historique des Mathématiques : algébricité, analyticité, dérivabilité, continuité, etc..., fonctions d'une variable, de deux variables, d'une variable complexe, etc.... Toutes ces notions sont aujourd'hui des cas particuliers d'un schéma unique, plus général et conceptuellement plus simple que tous les cas particuliers qu'il gouverne. En même temps, le langage de la théorie des ensembles (dont on modifie encore de temps à autre la terminologie, mais non les concepts de base) s'est universellement répandu, et son utilisation est devenue une condition *sine qua non* de clarté et de rigueur.

L'étude des §§ suivants est donc à peu près indispensable pour aborder la suite de ce livre; les §§ 1 et 2, le n° 1 du § 3 sont particulièrement importants dans l'immédiat; le lecteur pourra attendre d'avoir besoin du § 4 avant de l'étudier sérieusement; le § 5 est d'une utilité pratique assez faible si l'on admet que le lecteur est déjà au courant des principales propriétés des nombres entiers; mais le n° 7 de ce § est très souvent utilisé.

Quant au § 0, c'est une introduction à la Logique mathématique; on a tenté d'y donner une idée approximative de la façon dont les mathématiciens concoivent les objets dont ils s'occupent, et d'y rassembler un certain nombre de modes de raisonnement particulièrement importants. Ce §, comme d'ailleurs les §§ 1, 2 ou 3, n'a pas à être étudié en détail par le débutant, car les notions qu'on y trouvera sont constamment utilisées, et le lecteur se familiarisera forcément avec elles à la longue, et même très rapidement dans la plupart des cas.

On conseille enfin au débutant de ne pas s'effrayer de l'aspect abrupt, et qui lui semblera sans doute formidablement abstrait, de ces premiers §§. Le meilleur conseil qu'on pourrait lui donner serait d'oublier totalement et jusqu'à nouvel ordre les Mathématiques qu'il peut déjà connaître (et en particulier toute la Géométrie Élémentaire qui, mise à part la notion générale de « transformation géométrique », n'a aucun rapport avec les questions traitées ici). Il est également recommandé de prendre les définitions des termes techniques *au pied de la lettre*.

§ 0. — Le raisonnement logique.

1. L'idée de perfection logique

Il y a en Mathématiques trois processus fondamentaux : construire des *objets mathématiques*, former des *relations* entre ces objets et *démontrer* que certaines de ces relations sont *vraies*, ou, comme on dit, sont des *théorèmes*.

Les *objets mathématiques* sont les nombres, les fonctions, les figures géométriques, et d'innombrables autres choses dont s'occupent les mathématiciens : ces objets n'existent pas à proprement parler dans la Nature, mais ce sont des modèles abstraits d'objets physiques plus ou moins compliqués et visibles. Les *relations* sont les assertions (vraies ou non) qu'on peut énoncer concernant ces objets, et qui correspondent à des propriétés hypothétiques des objets naturels dont les objets mathématiques sont les modèles. Quant aux relations *vraies*, ce sont, pour les mathématiciens, celles qu'on peut déduire logiquement d'un petit nombre d'*axiomes* énoncés une fois pour toutes; ces axiomes traduisent en langage mathématique les propriétés les plus « évidentes » des objets concrets auxquels on pense; et la suite de syllogismes par laquelle on passe des axiomes (ou, plus pratiquement, de théorèmes déjà établis) à un théorème donné constitue une *démonstration* de celui-ci.

Les explications de ce genre, qui sembleront peut-être d'une admirable clarté à certains lecteurs débutants, ont depuis longtemps cessé de satisfaire les mathématiciens, non seulement parce que ceux-ci ont peu de goût pour les phrases vagues, mais aussi et surtout parce que les Mathématiques elles-mêmes les ont obligés à réfléchir aux fondements de leur science, et à substituer aux généralités des *formules* dont le sens ne puisse prêter à aucune confusion, et dont il soit possible de décider d'une façon quasi mécanique si elles sont vraies ou non, si elles ont un sens ou non.

Historiquement, la nécessité d'établir sur des bases aussi solides que possible les Mathématiques s'est manifestée à l'occasion du développement de la « théorie des ensembles », et de l'introduction en Mathématiques de nouvelles notions « abstraites » telles que celles d'anneau, de corps, de groupe.

En ce qui concerne la théorie des ensembles, créée par Cantor vers 1870, on s'est assez rapidement aperçu que, dans celle-ci, le recours à « l'intuition géométrique » était inutile dans les bons cas, et franchement nuisible dans les mauvais; au bout d'une

vingtaine d'années, on s'est trouvé en présence de résultats en apparence contraires au bon sens mais solidement démontrés (par exemple l'existence, prouvée par Peano, d'une courbe continue qui passe par *tous* les points d'un carré), et en même temps de véritables contradictions internes dues à l'emploi intempestif de raisonnements ingénieux dont les mathématiciens avaient la conviction, sans pouvoir vraiment le démontrer, que c'étaient de purs sophismes. Or, en Mathématiques, la contradiction interne est l'horreur suprême, car les Grecs savaient déjà que, si l'on dispose d'*une* relation contradictoire (i.e. à la fois vraie et fausse), alors on peut immédiatement démontrer que *toutes* les autres relations le sont également (voir la *Remarque* 5 plus loin).

Quant au développement des premières théories dites « abstraites » ou « axiomatiques », qui date de la même époque à peu près (1890-1910), il avait pour but soit d'englober dans une même théorie générale des théories particulières déjà connues afin de pouvoir appliquer aux unes les méthodes déjà utilisées pour étudier les autres, soit d'établir sur des bases solides des théories laissant à désirer au point de vue logique (l'exemple le plus célèbre de ce dernier cas est l'étude, par Hilbert, des fondements de la Géométrie élémentaire; deux mille ans après les Grecs, on eut enfin, *pour la première fois*, un exposé rigoureux et purement déductif de la Géométrie, dans lequel *tous* les axiomes, sans exception aucune, étaient explicitement énoncés, et où l'on voyait clairement, pour chaque théorème, quels étaient les axiomes strictement nécessaires à sa validité; l'exposé de Hilbert, par la rigueur de son langage et de ses raisonnements, et par son refus de toute concession, est le modèle de l'exposé mathématique moderne, et le restera sans aucun doute pendant de nombreux siècles).

Les efforts accomplis par les mathématiciens dans ces deux voies, du reste étroitement liées l'une à l'autre, les ont habitués à renforcer de beaucoup leurs exigences en matière de rigueur logique, et à raisonner sur des objets de plus en plus éloignés, en apparence, de la « réalité » concrète. On en est ainsi arrivé à la conviction que ce qui compte en Mathématiques, ce sont uniquement les symboles qui, assemblés en observant certaines « règles du jeu » explicitement énoncées, servent à former des objets mathématiques et des relations.

On admet même, aujourd'hui, qu'on pourrait théoriquement écrire toutes les Mathématiques en utilisant exclusivement un petit nombre de *signes fondamentaux* (par exemple les signes représentant les opérations logiques élémentaires, et deux ou trois signes proprement mathématiques — le signe de l'égalité, le signe de l'appartenance qui sera introduit au § suivant, et éventuellement le signe servant à former des « couples » d'objets, et qui sera introduit au § 2) et des *lettres* en quantité non limitée. Dans cette conception des Mathématiques, les objets mathématiques et les relations sont des assemblages (en général d'une complication telle qu'il est hors de question de les écrire effectivement) de signes fondamentaux et de lettres, formés en observant certains critères énoncés une fois pour toutes (le lecteur curieux en trouvera la liste, ou plus exactement une liste possible, au n° 9 de ce §); dans ces assemblages, le rôle des signes fondamentaux est de symboliser certaines opérations élémentaires, logiques ou mathématiques, dont la répétition à l'intérieur d'un même

assemblage conduit à des opérations beaucoup plus complexes; et les lettres, qui sont censées représenter des objets mathématiques totalement indéterminés, servent à introduire des « degrés de liberté » dans les assemblages considérés, i.e. à former des relations et des objets dépendant de « variables arbitraires ».

Une fois établies la liste des signes fondamentaux, et celle des *critères de formation* des objets mathématiques et des relations, il reste à énoncer les axiomes (les uns purement logiques, d'autres de nature mathématique à proprement parler).

A l'heure actuelle, les critères de formation et les axiomes ont été isolés et énoncés avec une précision telle que l'idée d'une machine qui démontrerait des théorèmes et rédigerait des Mathématiques a cessé d'être entièrement utopique. Les Mathématiques telles que les écrirait une telle machine (appliquant *strictement* et *exclusivement* les règles énoncées une fois pour toutes, n'omettant *aucun* intermédiaire de raisonnement si « évident » soit-il, et ne faisant usage que des lettres et des signes fondamentaux *à l'exclusion de toutes les abréviations usuelles*) sont dites *formalisées*; elles n'existent bien entendu que dans l'imagination des mathématiciens. Un texte mathématique écrit en langage formalisé ressemblerait, en infiniment plus compliqué, au compte-rendu d'une partie d'échecs, et donnerait au lecteur, s'il pouvait le comprendre, le sentiment de la *perfection logique*.

2. *Le langage réel des Mathématiques*

On a calculé que, si l'on cherchait à écrire en langage formalisé un objet mathématique aussi simple (en apparence...) que le nombre 1, on trouverait un assemblage comportant plusieurs dizaines de milliers de signes (les signes fondamentaux sont en très petit nombre, mais chacun d'eux peut naturellement être répété un grand nombre de fois dans un même assemblage). Le mathématicien qui essaierait de manipuler de pareils assemblages ressemblerait à l'alpiniste qui, pour choisir ses points d'appui sur une paroi rocheuse, examinerait celle-ci au microscope électronique.

On utilise donc, dans la pratique, une multitude d'*abréviations* (par exemple la lettre grecque π, le signe $\perp$, des mots du langage ordinaire, tels que « nombre », « point », « droite », « fonction », et ainsi de suite); celles-ci sont destinées à représenter par de nouveaux signes simples des assemblages compliqués de lettres et de signes fondamentaux, ou même des assemblages faisant intervenir en outre des signes abréviateurs déjà introduits. Quand il a introduit suffisamment d'abréviations d'assemblages de signes fondamentaux, puis d'abréviations d'assemblages d'abréviations, puis d'abréviations d'assemblages d'abréviations d'assemblages d'abréviations, et ainsi de suite, le mathématicien cesse de penser (*) à la définition complète et détaillée des objets qu'il a ainsi construits; il ne garde présent à l'esprit que la façon de passer d'un échelon de complication à l'échelon *immédiatement* précédent (ce qui constitue la *définition*, au sens usuel du terme, de l'abréviation considérée), et ne cherche pas à redescendre de proche en proche jusqu'au langage formalisé;

(*) On devrait même dire : ne peut plus penser.

à la limite, on en arrive souvent à raisonner sur les abréviations introduites comme si elles constituaient des signes primitifs au même titre que les signes fondamentaux du langage formalisé (c'est ainsi que Hilbert, dans son étude des fondements de la Géométrie, introduisait *a priori* trois notions primitives — celles de point, droite et plan — sans chercher aucunement à les définir, et en se bornant à en dresser le mode d'emploi).

C'est naturellement dans le choix des assemblages qu'il décide de représenter par des abréviations, et auxquels il décide donc de s'intéresser particulièrement, que le mathématicien montre que son activité diffère grandement de celle d'une machine. Une machine mathématique raisonnerait peut-être à la vitesse de la lumière, mais elle se bornerait probablement à accumuler les théorèmes au hasard suivant un processus analogue au mouvement brownien. Le but des mathématiciens est au contraire de démontrer des théorèmes « intéressants », de *résoudre les problèmes* qui sont posés depuis des dizaines d'années et parfois même depuis plusieurs siècles (*), et non pas d'inventer de nouvelles branches des Mathématiques, sans rapport avec les problèmes qui se posent, et purement gratuites. Et en fait c'est justement l'étude de ces problèmes qui oblige les mathématiciens à inventer de nouvelles notions (i.e. à introduire de nouvelles abréviations) et de nouvelles techniques qui, aux yeux du débutant ou de l'utilisateur, semblent parfois les entraîner fort loin du problème initial. Par exemple, en essayant de démontrer le « grand théorème de Fermat », et de résoudre d'autres problèmes arithmétiques, Kummer a été conduit à introduire vers le milieu du siècle dernier une certaine notion de « nombres idéaux »; celle-ci a conduit Dedekind à inventer, vers 1870, les « anneaux d'entiers algébriques » et les « idéaux » de ces anneaux; de là sont sorties les notions « modernes » et « abstraites » d'anneau et d'idéal, qui sont à leur tour en train de fusionner avec la Géométrie Algébrique; or le « grand théorème de Fermat » dit que, pour $n \geqslant 3$, la « courbe plane » dont l'équation est

$$x^n + y^n = 1$$

(*) Le problème de la « quadrature du cercle », et celui de caractériser les constructions géométriques que l'on peut « effectuer à l'aide de la règle et du compas », posés par les Grecs, n'ont été résolus qu'au XIX^e siècle. L'hypothèse de Goldbach (à savoir que tout nombre entier pair peut s'écrire comme somme de deux nombres *premiers* impairs), énoncée au XVIII^e siècle, n'est pas encore démontrée bien qu'on ait fait des progrès très importants dans cette voie depuis trente ans. Le problème de Waring (montrer que, pour tout entier n, il existe un entier p tel que tout nombre entier puisse s'écrire comme somme de p termes qui sont des puissances n^e de nombres entiers — le cas le plus simple est $n = 2$, et on montre alors que tout entier est somme de *quatre* carrés) énoncé lui aussi au XVIII^e siècle, a été résolu au début du XX^e par Hilbert. Le « grand théorème de Fermat » (montrer que, si n est un entier au moins égal à trois, l'équation

$$x^n + y^n = z^n$$

n'admet aucune solution formée d'entiers x, y, z tous strictement positifs), énoncé au début du XVII^e siècle, est encore fort loin de sa solution, qui exigera sans doute la mise en œuvre de tout l'arsenal inventé par les algébristes depuis un siècle, et d'autres techniques plus puissantes. Nous ne citons ici bien entendu que des problèmes dont les énoncés sont suffisamment simples pour pouvoir être compris des débutants.

ne contient aucun autre point du plan à coordonnées *rationnelles* que ses intersections avec les axes de coordonnées, et c'est précisément de ce genre de problèmes que s'occupe la Géométrie Algébrique ! Cet exemple, et beaucoup d'autres dont il n'est pas possible de parler ici, montrent que, malgré ce que peuvent croire les amateurs, les mathématiciens professionnels cherchent à s'engager dans les voies les plus *naturelles* possibles, celles dans lesquelles leur intuition géométrique ou analytique ou arithmétique — car il existe une intuition arithmétique — peut s'exercer.

L'activité des mathématiciens en chair et en os diffère de celle des machines sur un autre point encore : les premiers ne sont pas en mesure de se conformer *strictement* à la rigueur logique absolue dont feraient preuve les secondes si elles existaient. Dans la pratique, les textes mathématiques les mieux écrits comportent une multitude de « trous » en l'absence desquels la lecture de ces textes serait un exercice intolérable. Ces lacunes logiques sont sans importance, parce que chacun est parfaitement convaincu du fait qu'on pourrait les combler si on le désirait — en fait, il est même probable que le lecteur débutant ne les apercevra pas. On estime aujourd'hui qu'un texte mathématique est « parfaitement » correct lorsqu'il a acquis le degré de clarté et de rigueur qu'on a toujours trouvé dans les exposés d'Arithmétique élémentaire (et c'est pourquoi il est fort regrettable que cette branche des Mathématiques n'occupe pas plus de place dans l'enseignement secondaire français) ; l'immense majorité des théories mathématiques peuvent maintenant s'exposer dans ce style, et cette possibilité a pour corollaire le fait qu'on n'admet plus, aujourd'hui, le genre d'exposé décoratif qui permettait encore, il n'y a pas si longtemps, à certains mathématiciens, de briguer à la fois l'Académie des Sciences et l'Académie Française.

Nous allons maintenant essayer de donner au lecteur des informations plus précises sur la façon dont on construit des relations et sur les raisonnements logiques les plus importants. On espère que le lecteur débutant voudra bien ne pas croire que les considérations qui suivent intéressent uniquement les philosophes ou les spécialistes de Logique mathématique; il s'agit en fait des règles de raisonnement que les mathématiciens utilisent *à chaque seconde*, et à propos desquelles les débutants — l'expérience le montre — commettent fréquemment de grossières erreurs.

3. *Opérations logiques élémentaires*

Comme on l'a dit au n° 1, les relations et les objets mathématiques sont, théoriquement, des assemblages de lettres et de signes fondamentaux, formés en observant certains critères, avec lesquels le lecteur entrera en contact progressivement. On va s'intéresser tout d'abord aux deux signes les plus simples servant à former des ***relations***, et aux abréviations qui s'expriment uniquement à l'aide de ces deux signes.

Les logiciens représentent ces signes par $\vee$ et $\neg$, mais nous les désignerons par les mots

ou, non

qu'on emploie toujours en Mathématiques. Si R et S sont des *relations*, alors l'assemblage

R ou S

obtenu en écrivant l'assemblage R, puis le signe ou, puis l'assemblage S, est encore une *relation* qu'on appelle la **disjonction logique** des relations R et S [on verra plus loin, quand on aura défini les relations « vraies », que pour que la relation (R ou S) soit vraie il suffit que *l'une au moins* des deux relations R, S données le soit]. De même, si R est une relation, l'assemblage

non R

obtenu en faisant précéder l'assemblage R du signe non est encore une *relation*, qu'on appelle la **négation** de la relation R [plus loin, une relation sera dite fausse lorsque sa négation est vraie]. Les règles d'emploi de ces deux signes (autrement dit, les axiomes qui les font intervenir) seront énoncées plus tard; pour le moment, il est inutile de se poser ce genre de question.

A partir de ces deux signes, on peut introduire des abréviations d'usage constant, et tout d'abord le mot

et;

étant données des relations R, S, on désigne par

R et S

la relation

non [(non R) ou (non S)],

qu'on appelle la **conjonction logique** de R et S [on verra plus loin que, pour que la relation (R et S) soit vraie, il faut et il suffit que les deux relations R, S données le soient]. On note d'autre part

$$R \Longrightarrow S$$

la relation

S ou (non R);

on l'appelle une **implication logique** et on la lit (*)

R implique S;

[comme on le verra plus loin, dire que celle-ci est vraie signifie que S est conséquence

(*) Dans la pratique, on ne dit « R implique S » que dans le cas où cette relation est vraie au sens qui sera précisé plus loin. Dans ce § par contre, nous écrivons des relations sans nous préoccuper, pour le moment du moins, de savoir si elles sont vraies ou non.

Notons d'autre part que beaucoup de gens ont maintenant tendance à utiliser le signe $\Longrightarrow$ comme abréviation du mot « implique »; la plupart des mathématiciens professionnels rejettent cette façon de ne pas écrire français (ou chinois s'ils sont chinois). Il est du reste fort difficile, dans la pratique, d'utiliser *correctement* le signe $\Longrightarrow$.

logique de R (donc est vraie si R l'est); mais la relation $(R \Longrightarrow S)$ peut fort bien être vraie alors que ni R ni S ne le sont — par exemple, la relation

$$(1 = 2) \Longrightarrow (2 = 3)$$

est évidemment vraie, [de même (en logique nazie) que la relation

les communistes ont incendié le Reichstag donc il faut exterminer les communistes;

la déduction, en logique nazie, était impeccable, mais la prémisse était fausse].

Enfin, on désigne par

$$R \Longleftrightarrow S$$

la relation

$$[(R \Longrightarrow S) \text{ et } (S \Longrightarrow R)];$$

on l'appelle une **équivalence logique** et on la lit

R est équivalente à S.

On verra plus loin d'autres critères pour former des relations.

4. *Axiomes et théorèmes*

Les signes introduits au n° précédent nous permettent déjà de définir les relations **vraies** ou **théorèmes;** ce sont celles qu'on peut obtenir par application répétée des deux règles suivantes :

(RV 1) : *Toute relation obtenue par application d'un axiome est vraie.*

(RV 2) : *Étant données des relations* R *et* S, *si la relation* $(R \Longrightarrow S)$ *est vraie, et si la relation* R *est vraie, alors la relation* S *est vraie.*

Quant aux **axiomes,** ce sont des relations énoncées explicitement et une fois pour toutes, ou bien des règles dans lesquelles interviennent des relations « arbitraires » et qui, appliquées à des relations spécifiques, conduisent à d'autres relations spécifiques (et vraies, d'après la définition même de ce mot); les axiomes qu'on énoncera au n° suivant sont du second type.

Une relation est dite **fausse** lorsque sa négation est vraie.

Remarque 1. On voit donc que ce qui caractérise les relations vraies c'est qu'*on peut les démontrer* et non pas, par exemple, que les lois naturelles qu'elles sont censées schématiser peuvent être vérifiées expérimentalement; on a ici la différence fondamentale entre les vérités mathématiques et les vérités expérimentales. Dans la pratique quotidienne, le débutant aura le plus grand intérêt à garder ce fait présent à l'esprit, tout au moins s'il s'intéresse effectivement aux Mathématiques.

Remarque 2. Il va de soi qu'en Mathématiques la question de savoir si une relation donnée est vraie ou non dépend du système d'axiomes adopté au départ : une relation vraie dans une axiomatique donnée peut cesser de l'être dans un autre système, en apparence aussi « naturel » que le premier. On a par exemple construit des axiomatiques de la théorie des ensembles dans lesquelles l'assertion « il existe des ensembles infinis » n'est pas vraie (i.e. ne peut pas être démontrée — ce qui explique pourquoi cette assertion est l'un des *axiomes* de base de la théorie des ensembles telle qu'elle est utilisée par les mathématiciens).

Noter par ailleurs qu'une relation non vraie (i.e. qui ne peut pas être démontrée) n'est pas forcément fausse : si R est une relation, il se peut fort bien, en théorie, que les axiomes adoptés au départ ne permettent de démontrer ni R ni non R; de telles relations sont dites *indécidables* (dans la théorie axiomatique considérée). On a pu montrer récemment (voir p. 95) qu'il existe de telles relations dans les mathématiques usuelles, fondées sur les axiomes des §§ 0, 1 et 2. Si R est une telle relation, on a le droit, si on le désire, de « compléter » les mathématiques en ajoutant à la liste des axiomes de base soit R, soit non R.

Remarque 3. Outre les relations vraies, les relations fausses et les relations indécidables, on doit encore en principe considérer les relations *contradictoires*, i.e. à la fois vraies et fausses. Tout le monde espère bien entendu que les axiomes de base des mathématiques sont compatibles entre eux, i.e. interdisent l'existence de relations contradictoires — mais on n'a pas pu encore le démontrer. Si des relations contradictoires se manifestaient, il faudrait abandonner ou affaiblir certains des axiomes de base.

5. *Axiomes logiques et tautologies*

Jusqu'à présent nous n'avons énoncé aucun axiome; nous allons, dans ce n°, énoncer les plus simples d'entre eux, ceux qui servent à justifier les raisonnements logiques les plus élémentaires (syllogismes, doubles négations, etc...); ces axiomes sont au nombre de quatre, mais le lecteur aurait tort de croire que la présentation adoptée ici soit la seule possible : il existe de nombreuses autres possibilités de déduire les raisonnements logiques élémentaires d'un petit nombre d'axiomes simples.

(AL. 1) : *Si* R *est une relation, la relation*

$$(\text{R ou R}) \Longrightarrow \text{R}$$

est vraie.

Si donc la relation (R ou R) est vraie, il en sera de même de R d'après la règle (RV 2) du n° précédent.

(AL. 2) : *Si* R *et* S *sont deux relations, la relation*

$$\text{R} \Longrightarrow (\text{R ou S})$$

est vraie.

Si donc R est vraie, il en est de même de (R ou S), conformément au sens intuitif du mot « ou ».

(AL 3) : *Si* R *et* S *sont des relations, la relation*

$$(\text{R ou S}) \Longrightarrow (\text{S ou R})$$

est vraie.

En combinant les deux axiomes précédents, on voit que si R est vraie il en est de même de (S ou R), ou, ce qui revient au même, que si S est vraie il en est de même de (R ou S) ; comme on avait déjà établi que (R ou S) est vrai si R est vraie, on voit que (R ou S) est vraie dès que l'une au moins des deux relations R, S est vraie.

(AL 4) : *Si* R, S *et* T *sont des relations, la relation*

$$(\text{R} \Longrightarrow \text{S}) \Longrightarrow ((\text{R ou T}) \Longrightarrow (\text{S ou T}))$$

est vraie.

Si donc R implique S i.e. si la relation (R $\Longrightarrow$ S) est vraie alors la relation (R ou T) implique la relation (S ou T). Ce genre d'énoncé pourra sembler trivial au lecteur débutant — mais il n'en est que plus remarquable qu'on puisse en tirer des conséquences substantielles. En fait, il faut considérer les quatre axiomes précédents comme les règles d'emploi mécanique des signes « ou » et « non », et non pas comme des découvertes métaphysiques profondes.

Les théorèmes que l'on peut démontrer en utilisant uniquement les axiomes précédents s'appellent des **tautologies** ; ce sont ceux que les mathématiciens utilisent, la plupart du temps sans s'y référer explicitement, dans leurs raisonnements logiques. En voici quelques-uns; nous ne les démontrerons pas tous :

(TL 1) : *Si* R, S *et* T *sont des relations, si* (R $\Longrightarrow$ S) *et si* (S $\Longrightarrow$ T) *sont vraies, alors* (R $\Longrightarrow$ T) *est vraie.*

Voici la démonstration, à titre d'exemple. Appliquons (AL 4) en y remplaçant R, S et T par S, T et (non R) respectivement; on voit que la relation

$$(\text{S} \Longrightarrow \text{T}) \Longrightarrow [(\text{S ou (non R)}) \Longrightarrow (\text{T ou (non R)})]$$

est vraie; vu la définition du signe $\Longrightarrow$, cela signifie que la relation

$$(\text{S} \Longrightarrow \text{T}) \Longrightarrow [(\text{R} \Longrightarrow \text{S}) \Longrightarrow (\text{R} \Longrightarrow \text{T})]$$

est vraie; par hypothèse la relation (S $\Longrightarrow$ T) est vraie; la règle (RV 2) montre donc que la relation

$$(\text{R} \Longrightarrow \text{S}) \Longrightarrow (\text{R} \Longrightarrow \text{T})$$

est vraie; comme la relation (R $\Longrightarrow$ S) est vraie par hypothèse, on voit en appliquant à nouveau (RV 2) que (R $\Longrightarrow$ T) est vraie, ce qui achève la démonstration.

(TL 2) : *Si* R *est une relation, la relation* (R $\Longrightarrow$ R) *est vraie.*

En effet, les relations

$$R \Longrightarrow (R \text{ ou } R), \qquad (R \text{ ou } R) \Longrightarrow R$$

sont vraies d'après (AL 1) et (AL 2) ; il reste donc à appliquer (TL 1).

Remarque 4. Vu la définition du signe $\Longrightarrow$, l'énoncé précédent signifie que, quelle que soit la relation R, la relation

$$R \text{ ou } (\text{non } R)$$

est vraie. *Il ne s'ensuit pas* que l'une au moins des deux relations R, (non R) soit vraie — c'est justement la question de savoir s'il existe des relations indécidables ! En fait, quand on a établi qu'une relation (R ou S) est vraie, on ne peut pas en déduire directement que l'une des relations R, S soit vraie.

(TL 3) : *Si* R *est une relation, la relation*

$$R \Longleftrightarrow \text{non } (\text{non } R)$$

est vraie.

Dire que R est vraie revient donc à dire que la négation de (non R) est vraie, autrement dit que (non R) est fausse.

(TL 4) : *Si* R *et* S *sont des relations, la relation*

$$(R \Longrightarrow S) \Longleftrightarrow [(\text{non } S) \Longrightarrow (\text{non } R)]$$

est vraie.

Pour établir que R implique S, il est donc suffisant (et nécessaire) de prouver que la négation de S implique celle de R. Par contre, l'énoncé

$$(R \Longrightarrow S) \Longrightarrow [(\text{non } R) \Longrightarrow (\text{non } S)]$$

est faux, et est la source de nombreuses erreurs de raisonnements (étant donné que tout homme est mortel, cet énoncé pourrait servir à prouver que tout chien est immortel). On l'utilise cependant très souvent dans la vie courante, le plus souvent à tort bien entendu, et parfois avec raison sur le plan psychologique : « les gens de droite soutiennent l'Algérie Française, donc les gens de gauche soutiennent l'Algérie Indépendante ».

(TL 5) : *Si* R, S *sont des relations, les relations*

$$(R \text{ et } S) \Longrightarrow R, \qquad (R \text{ et } S) \Longrightarrow S$$

sont vraies; si de plus R *et* S *sont vraies, alors* (R et S) *est vraie.*

On déduit de là que, pour que (R et S) soit vraie, *il faut et il suffit* que les relations R et S considérées le soient, conformément au sens intuitif du mot « et ».

Remarque 5. Il est facile de montrer que s'il existait une relation contradictoire R, alors toute autre relation S le serait aussi, comme on l'a dit plus haut.

En effet, (AL 2) et (AL 3) montrent que

$$(\text{non R}) \Longrightarrow (\text{S ou (non R)})$$

est vraie; comme (non R) est vraie par hypothèse, la relation

$$\text{S ou (non R)}, \qquad \text{i.e. } (\text{R} \Longrightarrow \text{S}),$$

est donc vraie; et comme R est vraie par hypothèse, on en déduit que S est vraie — et donc aussi (non S)...

La *Remarque* précédente est à la base du **raisonnement par l'absurde**; celui-ci consiste, pour démontrer qu'une relation R est vraie, à adjoindre temporairement (non R) aux axiomes des Mathématiques, et à établir que les « nouvelles » Mathématiques ainsi obtenues sont contradictoires; d'après la *Remarque* 5, toute relation est alors vraie dans le nouveau système, et en particulier la relation R elle-même. Par suite, R est conséquence logique des axiomes des Mathématiques (usuelles) et de la relation (non R), ce qui signifie, comme on le voit facilement, que *la relation*

$$(\text{non R}) \Longrightarrow \text{R}$$

est vraie (dans les Mathématiques usuelles, auxquelles on est maintenant revenu); il reste à en déduire que R elle-même est vraie; or l'axiome (AL 4), où l'on remplace R, S et T par (non R), R et R respectivement, montre que la relation

$$((\text{non R}) \Longrightarrow \text{R}) \Longrightarrow [((\text{non R}) \text{ ou R}) \Longrightarrow (\text{R ou R})]$$

est vraie; comme ((non R) $\Longrightarrow$ R) est vraie par hypothèse il en est donc de même de

$$((\text{non R}) \text{ ou R}) \Longrightarrow (\text{R ou R});$$

mais ((non R) ou R) est vraie d'après (AL 3) et la *Remarque* 4; par conséquent, (R ou R) est vraie, et (AL 1) montre finalement que R est vraie.

Dans la pratique, le raisonnement par l'absurde s'utilise comme suit : on « suppose la relation R fausse », ce qui revient justement à adjoindre (non R) aux axiomes des Mathématiques; on raisonne à partir de là jusqu'à ce qu'on ait trouvé une relation à la fois vraie et fausse; on termine alors en disant « or ceci est absurde, donc R est vraie ».

Une autre méthode de démonstration fréquemment utilisée dans la pratique est celle de la **disjonction des cas**; elle repose sur l'énoncé suivant :

(11.6) : *Soient* R, S *et* T *trois relations; si les trois relations*

$$\text{R ou S}, \qquad \text{R} \Longrightarrow \text{T}, \qquad \text{S} \Longrightarrow \text{T}$$

sont vraies, alors T *est vraie.*

Comme S implique T, l'axiome (AL 4) montre que (S ou R) implique (T ou R); comme R implique T, on voit de même que (R ou T) implique (T ou T); comme (T ou R) implique (R ou T) d'après (AL 3), on voit donc que (S ou R) implique

(T ou T); or (R ou S) est vraie et implique (S ou R); par suite (T ou T) est vraie, et on conclut la démonstration à l'aide de (AL 1).

Dans la pratique, on utilise surtout l'énoncé précédent en prenant pour S la négation de R; *pour montrer que* T *est vraie, il suffit de faire voir que* R *implique* T, *et que* (non R) *implique* T *également.*

6. Substitutions dans une relation

Soient R une relation, A un objet mathématique, et x une lettre (qui est donc un objet mathématique « totalement indéterminé »); dans l'assemblage de lettres et de signes fondamentaux qui constitue la relation R, remplaçons partout la lettre x par l'assemblage A; l'un des critères de formation des relations est que l'assemblage ainsi obtenu *est encore une relation*, que l'on désigne (*) par la notation

$$(A|x)R$$

et qu'on appelle la relation obtenue en **substituant** A **à** x **dans** R, ou en **donnant à** x **la valeur** A **dans** R; et on dit que l'objet mathématique A **vérifie la relation** R si la relation $(A|x)R$ est vraie. Il va de soi que, si la lettre x ne figure pas effectivement dans l'assemblage R, la relation $(A|x)R$ n'est autre que R, et dans ce cas dire que A vérifie R signifie que R est vraie.

Pour indiquer qu'une lettre x figure dans une relation R, on écrit fréquemment celle-ci sous la forme

$$R\{x\}$$

(analogue, mais non identique, à celle qu'on utilisera au § 2 pour désigner les fonctions); on écrit alors fréquemment

$$R\{A\}$$

au lieu de $(A|x)R$. De même, si x et y sont deux lettres distinctes figurant dans R, et si l'on veut mettre ce fait en évidence, on écrit

$$R\{x, y\}$$

au lieu de R, et ainsi de suite.

(TI, 7) : *Soient* R *une relation,* x *une lettre, et* A *un objet mathématique. Si la relation* R *est vraie, il en est de même de la relation obtenue en substituant* A *à* x *dans* R.

Autrement dit, si R est vraie lorsqu'on y regarde x comme un « objet indéterminé », alors R reste vraie lorsqu'on donne à x une « valeur » spécifique A. Par exemple, si l'on démontre que la relation

$$x = x,$$

(*) **Dans ce § exclusivement.**

où x est une *lettre* au sens technique du terme, est vraie, il s'ensuivra que la relation

$$A = A$$

est vraie quel que soit l'objet mathématique A.

Ce résultat donnera sans doute au lecteur débutant l'impression d'un simple calembour, à cause de la signification intuitive que nous avons attribuée plus haut aux *lettres*, qui sont censées représenter des objets « totalement indéterminés » ou « arbitraires »; *mais cette interprétation n'était jusqu'ici fondée sur aucun résultat mathématique précis*, et n'avait pas encore été effectivement justifiée par des règles gouvernant l'emploi des lettres et conformes à ce que le sens commun attend du comportement d'objets « indéterminés ». *La règle* (TL 7) *a précisément pour but de justifier cette interprétation des lettres comme « objets indéterminés »*, et ce serait ne pas avoir compris la situation véritable que de croire qu'on peut justifier (TL 7) par de simples considérations de « bon sens » : les machines à démontrer ignorent cette notion, et on perdrait son temps à vouloir leur faire comprendre ce qu'est un « objet indéterminé ».

En fait, la démonstration consiste tout d'abord à vérifier (TL 7) lorsque R s'obtient par application directe d'un axiome, et se fait en constatant que, dans ce cas, la relation $(A|x)R$ est soit identique à R, soit obtenue par application directe du même axiome que R — si par exemple R est la relation

$$(S \text{ ou } S) \Longrightarrow S$$

où S est une relation donnée, alors il est clair que $(A|x)R$ n'est autre que la relation

$$(S' \text{ ou } S') \Longrightarrow S'$$

où S' désigne la relation $(A|x)S$. Une fois effectuées ces vérifications, qui supposent bien entendu qu'on a écrit explicitement tous les axiomes, il est immédiat de passer au cas général d'une relation vraie « quelconque ».

7. *Quantificateurs*

Nous avons indiqué jusqu'à présent trois procédés fondamentaux pour former des relations — la disjonction logique, la négation, et la substitution d'un objet à une lettre. Ces procédés sont de nature purement logique (i.e. ne font pas intervenir les signes mathématiques qu'on introduira aux §§ 1 et 2). Dans la pratique, on a encore besoin d'un quatrième procédé purement logique pour former des relations — c'est celui qui exprime l'assertion qu'étant données une relation R et une lettre x, il existe au moins un objet mathématique A tel que la relation $(A|x)R$ soit vraie, i.e. qui vérifie R. Naturellement nous attribuons ici à l'expression « il existe » son sens purement intuitif; les considérations qui suivent ont pour but de la remplacer par un nouveau signe logique, à savoir le signe

$$\exists,$$

et de codifier l'emploi de celui-ci de telle sorte qu'il se comporte conformément à ce

qu'exige le sens commun lorsqu'on utilise l'expression « il existe » en langage courant. Le signe ∃ s'appelle le **quantificateur existentiel.**

Étant données une *relation* R et une *lettre* x, on peut donc former une nouvelle *relation* qui se désigne par

$$(\exists x)\,\mathrm{R} \qquad \text{ou} \qquad (\exists x)\,\mathrm{R}\{x\}$$

et qui se lit (*)

il existe x tel que R.

Pour écrire dans ce langage que, par exemple, l'équation $x^4 + 1 = 0$ possède au moins une racine réelle (ce qui est d'ailleurs faux, mais importe peu), on désigne par **R** l'ensemble des nombres réels et, en utilisant le signe ∈ qui sera introduit au § suivant, on forme la relation

$$(\exists x)\,[(x \in \mathbf{R})\ \text{et}\ (x^4 + 1 = 0)].$$

A partir du signe ∃, on introduit l'abréviation

$$\forall$$

qu'on appelle le **quantificateur universel** ; si R est une relation et x une lettre, on désigne par

$$(\forall x)\,\mathrm{R}$$

la relation

$$\text{non}\,[(\exists x)\ (\text{non R})],$$

et on la lit (*)

pour tout x, R

ou encore

on a R quel que soit x.

Dire que la relation $((\forall x)\,\mathrm{R})$ est fausse signifie donc que l'assertion $((\exists x)$ (non R)) est vraie : l'assertion « tous les hommes sont mortels » est la négation logique de l'assertion « il existe des hommes immortels ». Mais bien entendu, dire que l'assertion « tous les habitants de la Casbah d'Alger ont été soumis à la torture en 1957 » est fausse *ne signifie pas* que l'assertion « aucun habitant de la Casbah d'Alger n'a été torturé en 1957 » soit vraie (**).

(*) Il est fort difficile d'utiliser *correctement* les signes ∃ et ∀ dans la pratique courante; il est donc préférable de se borner à écrire « il existe » et « pour tout », comme on l'a toujours fait.

(**) Le lecteur qui désirerait savoir à quoi s'en tenir pourra consulter la documentation rassemblée dans le livre de Pierre Vidal-Naquet, *La Raison d'État* (Éditions de Minuit, Paris, 1962), où l'on trouvera aussi une bibliographie abondante.

Sur les exploits de la police algérienne voir, du même auteur, *La Question ininterrompue* dans *Le Monde* du 29 septembre 1963.

¶ *Remarque* 6. Il est théoriquement possible d'écrire les Mathématiques en n'utilisant que des lettres, les trois signes logiques

$$\text{ou}, \quad \text{non}, \quad \exists,$$

et les trois signes mathématiques qui, dans les §§ 1 et 2, serviront à former des « égalités », des « appartenances » et des « couples » (on pourrait même se passer du dernier signe).

Dans ce système, une assertion de la forme $(\exists x)\,\mathrm{R}$ peut être vraie sans qu'on ait aucun moyen de « construire effectivement » un objet mathématique vérifiant R — de même, dans la vie de tous les jours, l'assertion « il existe des banquiers honnêtes » ne constitue pas une information très substantielle, car elle ne permet pas, à elle seule, d'*exhiber* un honnête banquier.

On trouvera au n° 9 des indications sur un système plus complet, ne présentant pas l'inconvénient dont on vient de parler, mais moins naturel que le système précédent.

8. *Règles d'emploi des quantificateurs*

La première de ces règles est la suivante :

(AL 5) : *Soient* R *une relation, x une lettre et* A *un objet mathématique. Alors la relation*

$$(\mathrm{A}|x)\,\mathrm{R} \Longrightarrow (\exists x)\,\mathrm{R}$$

est vraie.

Si donc la relation $(\mathrm{A}|x)\mathrm{R}$ est vraie, autrement dit si l'objet A « vérifie » R, alors la relation $(\exists x)\mathrm{R}$ est vraie, conformément à ce qu'on espère. Dans la pratique, c'est presque toujours ainsi qu'on démontre qu'une relation de la forme $(\exists x)\mathrm{R}$ est vraie : on exhibe un objet spécifique A qui vérifie R. On fait de même dans la vie courante : la meilleure façon de prouver qu'il existe des banquiers honnêtes, c'est d'en exhiber un explicitement.

(TL 8) : *Soient* R *une relation, x une lettre et* A *un objet mathématique. Alors la relation*

$$(\forall x)\mathrm{R} \Longrightarrow (\mathrm{A}|x)\mathrm{R}$$

est vraie.

Si donc la relation $(\forall x)\mathrm{R}$ est vraie, on obtient encore une relation vraie en substituant à x, dans R, un objet mathématique quelconque. Par exemple, si l'on démontre que la relation

$$(\forall x)\ (x = x),$$

où x est une lettre au sens technique du terme, est vraie, alors pour tout objet mathématique A la relation

$$\mathrm{A} = \mathrm{A}$$

sera vraie. On procède de même dans la vie courante : l'assertion « tous les intellec-

tuels sont des pédérastes » implique que chaque intellectuel explicitement nommé est un pédéraste.

(TL 9) : *Soient* R *une relation et x une lettre. La relation*

$$\text{non } ((\exists x)\text{R}) \Longleftrightarrow (\forall x)\ (\text{non R})$$

est vraie.

Dire que la relation $(\exists x)$R est fausse signifie donc que la relation $(\forall x)$ (non R) est vraie, et en particulier implique que tout objet mathématique A vérifie (non R).

(TL 10) : *Soient* R *et* S *des relations et x une lettre. La relation*

$$(\forall x)\ (\text{R et S}) \Longleftrightarrow ((\forall x)\text{R et } (\forall x)\text{S})$$

est vraie.

C'est évident intuitivement. On notera par contre que la relation

$$(\forall x)\ (\text{R ou S}) \Longleftrightarrow ((\forall x)\text{R ou } (\forall x)\text{S})$$

n'est pas vraie en général; par exemple, l'assertion « tous les intellectuels sont des traîtres ou des pédérastes » n'implique pas la relation « tous les intellectuels sont des traîtres ou tous les intellectuels sont des pédérastes », car il pourrait exister *à la fois* des intellectuels qui sont des traîtres sans être des pédérastes, *et* des intellectuels qui sont des pédérastes tout en étant patriotes.

(TL 11) : *Soient* R *et* S *des relations et x une lettre. La relation*

$$(\exists x)\ (\text{R ou S}) \Longleftrightarrow ((\exists x)\text{R ou } (\exists x)\text{S})$$

est vraie.

C'est aussi évident intuitivement. Notons d'autre part que la relation

$$(\exists x)\ (\text{R et S}) \Longrightarrow ((\exists x)\text{R et } (\exists x)\text{S})$$

est *vraie*, mais que l'implication opposée

$$((\exists x)\text{R et } (\exists x)\text{S}) \Longrightarrow (\exists x)\ (\text{R et S})$$

est généralement *fausse*. Ainsi l'assertion « il existe des gens riches et honnêtes » implique évidemment l'assertion « il existe des gens riches et il existe des gens honnêtes », mais l'implication opposée n'est pas correcte, car des raisonnements *purement logiques* ne sauraient suffire à exclure l'éventualité dans laquelle les gens riches seraient nécessairement malhonnêtes.

Pour conclure, nous allons énoncer quelques règles faisant intervenir des quantificateurs itérés. Soient R une relation, et x, y deux lettres *distinctes*; on peut alors, dans R, appliquer un quantificateur relativement à la lettre x, puis, dans la relation obtenue, appliquer un quantificateur relativement à la variable y; on peut aussi bien entendu procéder dans l'ordre inverse, et se demander si l'ordre dans

lequel on effectue les opérations a une importance. La réponse est donnée par l'énoncé suivant (qui ne couvre pas tous les cas, pour la simple raison que les énoncés auxquels tout le monde pense et qui ne se trouvent pas ci-dessous sont *faux*) :

(TL 12) : *Soient* R *une relation, x et y des lettres distinctes. Alors les relations*

$$(\forall x)\ (\forall y)\mathrm{R} \iff (\forall y)\ (\forall x)\mathrm{R}$$
$$(\exists x)\ (\exists y)\mathrm{R} \iff (\exists y)\ (\exists x)\mathrm{R}$$
$$(\exists x)\ (\forall y)\mathrm{R} \implies (\forall y)\ (\exists x)\mathrm{R}$$

sont vraies.

Donnons par exemple une démonstration intuitive (i.e. incorrecte...) du second énoncé. Si la relation $(\exists x)(\exists y)\mathrm{R}$ est vraie, cela veut dire qu'on peut trouver un objet A tel qu'en le substituant à x dans la relation $(\exists y)\mathrm{R}$ on trouve une relation vraie — autrement dit tel que la relation $(\exists y)\mathrm{R}\{\mathrm{A}, y\}$ soit vraie; mais ceci veut dire de même qu'il existe un objet B tel qu'en le substituant à y dans $\mathrm{R}\{\mathrm{A}, y\}$ on trouve une relation vraie — autrement dit tel que la relation $\mathrm{R}\{\mathrm{A}, \mathrm{B}\}$ soit vraie. Mais alors la relation $(\exists x)\mathrm{R}\{x, \mathrm{B}\}$ est vraie, donc aussi la relation $(\exists y)\ (\exists x)\mathrm{R}\{x, y\}$ et ceci démontre (sic) la relation considérée.

Remarque 7. Les incorrections de la démonstration précédente sont au nombre de trois au moins: (1) on s'est borné à montrer que la relation $(\exists x)\ (\exists y)\mathrm{R}$ *implique* la relation $(\exists y)(\exists x)\mathrm{R}$ au lieu de prouver que ces deux relations sont équivalentes; ce n'est évidemment pas grave, l'implication opposée se démontrant de la même façon; (2) on s'est borné à montrer que si la relation $(\exists x)(\exists y)\mathrm{R}$ est vraie alors la relation $(\exists y)(\exists x)\mathrm{R}$ l'est aussi — alors qu'une implication peut fort bien être vraie indépendamment de la vérité de ses deux termes; (3) on a admis que, pour qu'une relation de la forme $(\exists x)\mathrm{R}$ soit vraie, *il faut* et il suffit que l'on puisse trouver un objet A tel que la relation $(\mathrm{A}|x)\mathrm{R}$ soit vraie : or la règle (AL 5) indique seulement que cette condition est *suffisante*.

En ce qui concerne le point (3), on peut le justifier à l'aide des considérations du n° 9. Pour ce qui est de (2), on le justifie en faisant usage de la **méthode de l'hypothèse auxiliaire** dont voici l'énoncé: soient R et S deux relations; adjoignons temporairement R aux axiomes des Mathématiques (ce qu'on indique pratiquement en disant « supposons que R soit vraie »), et supposons que S soit vraie dans ces « nouvelles » Mathématiques; alors la relation $(\mathrm{R} \implies \mathrm{S})$ est vraie (dans les Mathématiques usuelles bien entendu !). Cette méthode se justifie à l'aide des critères les plus élémentaires, ceux des n° 4 et 5 : on écrit une suite de syllogismes partant des « nouveaux » axiomes (i.e. des axiomes usuels et de R) et aboutissant à S, et on démontre, en raisonnant de proche en proche, que chaque relation de la chaîne est conséquence logique de R, d'où, à la fin du processus, l'implication $(\mathrm{R} \implies \mathrm{S})$.

Remarque 8. On a fréquemment à écrire la négation d'une relation comportant deux quantificateurs successifs (ou même plus de deux), exercice que les débutants ont généralement quelque peine à accomplir. Soit par exemple à former la négation (ou une relation équivalente à la négation) de la relation $(\forall x)(\exists y)\mathrm{R}$. Désignant provisoirement par S la relation $(\exists y)\mathrm{R}$, on

doit former la négation de $(\forall x)S$; or cette relation n'est autre que

$$\text{non } [(\exists x) \text{ (non S)}];$$

sa négation est donc équivalente, d'après (TL 3), à $(\exists x)$ (non S). Mais comme S n'est autre que $(\exists y)R$, relation équivalente à $(\exists y)$(non non R), on voit que (non S) est équivalente à $(\forall y)$ (non R); par conséquent, on obtient le résultat cherché, à savoir que *la négation de la relation*

$$(\forall x)(\exists y)R$$

est équivalente à la relation

$$(\exists x)(\forall y)(\text{non R}).$$

9. *L'opération de Hilbert. Critères de formation*

Pour conclure ce §, et à l'intention des lecteurs non débutants qui pourraient s'intéresser à la Logique, nous allons donner des précisions supplémentaires sur l'un des systèmes possibles, à savoir celui qui est utilisé par N. Bourbaki dans ses *Éléments de Mathématique*.

Dans ce système, on utilise sept *signes fondamentaux*, et des *lettres*. Quatre des signes fondamentaux, à savoir

$$\text{ou}, \quad \text{non}, \quad \tau, \quad \square$$

sont de nature purement logique; les trois autres sont de nature mathématique à proprement parler, ce sont les signes

$$=, \quad \in, \quad \supset,$$

qui seront introduits aux §§ 1 et 2 (le troisième est le signe du couple; voir le critère (OM 2) ci-dessous).

Un *assemblage* s'obtient en écrivant une succession de signes et de lettres, certains des signes τ figurant dans un assemblage pouvant de plus être joints à certains des signes $\square$ par des *liens* — par exemple, l'expression

$$\tau x \in y = \in \in yx = z \square$$

est un assemblage.

Soit A un assemblage, et soit x une lettre; nous allons indiquer un procédé pour en déduire un nouvel assemblage *qui ne contient plus la lettre x* mais que l'on désigne néanmoins par la notation

$$\tau_x(A);$$

on l'obtient en effectuant les trois opérations suivantes :

a) on écrit l'assemblage τA obtenu en faisant précéder l'assemblage A du signe τ;

b) on joint le signe τ placé devant A à chaque occurence de la lettre x par un lien;

c) dans l'assemblage obtenu, on remplace partout la lettre x par le signe $\square$. Si par exemple A est l'assemblage écrit plus haut, alors $\tau_x(A)$ est l'assemblage

$$\tau\tau\square\in y = \in\in y\square = z\square.$$

L'opération faisant passer de A à $\tau_x(A)$ est essentiellement due à Hilbert; on en donnera plus loin la signification intuitive.

Nous allons maintenant énoncer les *critères de formation* des relations et des objets mathématiques; les voici :

(OM 1) : *Toute lettre est un objet mathématique.*

(OM 2) : *Si* A *et* B *sont des objets mathématiques, l'assemblage*

$$\supset AB,$$

qu'on désigne pratiquement par

$$(A, B),$$

est un objet mathématique.

(OM 3) : *Soient* A *et* T *des objets mathématiques, x une lettre; alors l'assemblage* $(A|x)T$ *déduit de* T *en y remplaçant partout la lettre x par l'assemblage* A, *est un objet mathématique.*

(OM 4) : *Soient* R *une relation et x une lettre. Alors l'assemblage* $\tau_x(R)$ *est un objet mathématique.*

(R 1) : *Si* R *et* S *sont des relations, l'assemblage*

$$\text{ou RS},$$

qu'on écrit pratiquement (R ou S), *est une relation.*

(R 2) : *Si* R *est une relation, l'assemblage*

$$\text{non R}$$

est une relation.

(R 3) : *Soient* R *une relation, x une lettre, et* A *un objet mathématique. L'assemblage* $(A|x)R$ *est une relation.*

(R 4) : *Soient* A *et* B *des objets mathématiques. L'assemblage*

$$= AB,$$

qu'on écrit pratiquement A = B, *est une relation.*

(R 5) : *Soient* A *et* B *des objets mathématiques. L'assemblage*

$$\in AB,$$

qu'on écrit pratiquement $A \in B$, *est une relation.*

Il n'y a pas d'autres méthodes, en Mathématiques, pour former des objets mathématiques et des relations; et, à l'exception de (OM 4), qui ne s'utilise presque jamais directement, tous les critères précédents sont effectivement utilisés à chaque instant dans la pratique.

On remarquera que les quantificateurs $\exists$ et $\forall$ n'interviennent pas dans ce qui précède : c'est parce qu'on va maintenant pouvoir (en utilisant l'opération de Hilbert) les introduire comme simples abréviations.

De façon précise, soient R une relation et x une lettre; alors

$$(\exists x)R$$

sera, par définition, la relation

$$(\tau_x(R)|x)R$$

qu'on déduit de R en y remplaçant partout la lettre x par l'objet mathématique $\tau_x(R)$. Par suite, *pour que la relation* $(\exists x)R$ *soit vraie, il faut et il suffit que l'objet* $\tau_x(R)$ *vérifie la relation* R, et ceci conduit à l'interprétation intuitive de l'opération de Hilbert: celle-ci consiste à choisir une fois pour toutes, pour chaque relation R et chaque lettre x, un objet vérifiant la relation $R\{x\}$ (s'il en « existe »; dans le cas contraire, $\tau_x(R)$ est un objet dont on ne peut rien dire). Il va de soi que ce « choix » est purement fictif : l'intérêt de l'opération de Hilbert est de donner un procédé parfaitement artificiel mais purement mécanique pour construire effectivement un objet dont on sait *seulement* qu'il satisfait à des conditions imposées d'avance (dans le cas où de tels objets existeraient) (*). On l'utilise aussi maintenant à la place de l'axiome du choix (§ 2, *Remarque* 7).

Dans la pratique courante, il est tout à fait exceptionnel d'avoir à utiliser l'opération de Hilbert (voir au § 5, *Remarque* 1 la définition des nombres cardinaux), qui ne peut évidemment conduire à aucun résultat « explicite ». Comme le Dieu des philosophes, l'opération de Hilbert est incompréhensible et ne se voit pas; mais elle gouverne tout, et ses manifestations sensibles éclatent partout.

(*) Le fait que $\tau_x(R)$ vérifie R si l'on peut construire un objet A qui vérifie R n'est bien entendu qu'une reformulation de l'*axiome* (AL 5).

§ 1. Les relations d'égalité et d'appartenance

1. La relation d'égalité

Les signes qui ont été introduits au § précédent sont de nature purement logique : ils servent essentiellement à « formaliser » des modes de raisonnement qui ne sont pas spécifiquement mathématiques. Nous allons par contre, dans ce §, introduire deux « signes fondamentaux » (le signe de l'égalité et le signe d'appartenance) qui servent à construire des relations et des objets ayant une signification à proprement parler mathématique.

Le **signe de l'égalité** se note

$$=$$

et est utilisé pour former des relations de la façon suivante : si a et b sont des *objets mathématiques* (ou **ensembles** — les deux terminologies sont synonymes), on obtient une *relation* en écrivant

$$a = b,$$

i.e. en faisant suivre l'assemblage de signes et de lettres que nous désignons par a du signe $=$, puis de l'assemblage que nous désignons par b. Intuitivement, la relation précédente signifie, lorsqu'elle est vraie, que les objets concrets qui sont censés être représentés par a et b sont « identiques » — notion dont nous ne chercherons pas à approfondir le sens; l'essentiel, pour le mathématicien, n'est pas d'avoir compris, ou cru comprendre, la « signification profonde » du signe $=$; il est de savoir l'utiliser

et pour ce faire, il suffit de se référer à l'énoncé suivant, qui résume les « règles du jeu » qu'on doit observer en écrivant des égalités (*) :

(*) Nous n'utilisons plus à partir de maintenant le langage relativement « formalisé » du § 0; si on voulait l'utiliser il faudrait énoncer sous la forme suivante l'assertion *a*) du Théorème 1 : « soit x une lettre, alors la relation

$$(\forall x)\ (x = x)$$

est vraie » — et sous une forme analogue les autres assertions du Théorème. Rappelons qu'en Mathématiques on n'utilise pas (ce serait impossible) le langage formalisé; on s'arrange simplement pour ne pas s'en éloigner au point de ne plus pouvoir y revenir si l'on avait des raisons sérieuses de le faire.

THÉORÈME 1. *On a les propriétés suivantes:*

a) *La relation* $x = x$ *est vraie pour tout* x.

b) *Les relations* $x = y$ *et* $y = x$ *sont équivalentes quels que soient* x *et* y.

c) *Quels que soient* x, y, z, *les relations* $x = y$ *et* $y = z$ *impliquent la relation* $x = z$.

d) *Soient* u *et* v *des objets tels que* $u = v$, *et* $\mathrm{R}\{x\}$ *une relation contenant une lettre* x; *alors les relations* $\mathrm{R}\{u\}$ *et* $\mathrm{R}\{v\}$, *qui se déduisent de* R *en* y *remplaçant partout la lettre* x *par* u *et* v *respectivement, sont équivalentes.*

Remarque 1. L'assertion *d*) du Théorème signifie que deux objets égaux ont les mêmes propriétés, i.e. que toute assertion valable pour a est aussi valable pour b, et réciproquement. Quant aux assertions *a*), *b*) et *c*) elles traduisent les propriétés les plus « évidentes » de l'égalité.

En ce qui concerne une démonstration du Théorème (sic) précédent, il serait à proprement parler *miraculeux* que nous puissions *démontrer* quoi que ce soit concernant le signe $=$ après l'avoir simplement écrit sur une feuille de papier. En fait, l'assertion *d*) est l'un des *axiomes* de base des Mathématiques, et quant aux assertions « évidentes » *a*), *b*) et *c*) on peut soit les prendre elles-mêmes comme autant d'axiomes (ce que devra faire le débutant), soit les déduire d'un axiome beaucoup plus compliqué (mais unique, et que le lecteur débutant *ne devra pas* chercher à comprendre) — à savoir que si R et S sont deux relations équivalentes et x une lettre, alors la relation

$$\tau_x(\mathrm{R}) = \tau_x(\mathrm{S})$$

est vraie.

Pour le lecteur qui désire réfléchir aux fondements des Mathématiques, ce serait un excellent exercice que d'essayer de « démontrer » le Théorème 1, et de se convaincre de ce que toutes les démonstrations (sic) qu'il en trouvera sont insuffisantes.

Il va de soi, enfin, que dans la pratique on utilise *constamment* le Théorème 1 sans *jamais* s'y référer explicitement.

2. *La relation d'appartenance*

Le second signe fondamental en Mathématiques est le **signe d'appartenance**, qui se note

$$\in$$

et, comme le signe $=$, est utilisé pour construire des relations à partir d'objets mathématiques : si a et b sont des *objets mathématiques*, on obtient une *relation* en écrivant

$$a \in b;$$

cette relation se lit

a appartient à b

ou encore

a est un élément de b.

La négation de la relation $a \in b$ s'écrit

$$a \notin b.$$

Ici encore, la seule chose qui compte ce sont les axiomes gouvernant l'emploi du signe $\in$ — il n'y en a du reste qu'un seul, que voici :

THÉORÈME 2. *Soient* A *et* B *deux ensembles; pour que l'on ait* $A = B$, *il faut et il suffit que les relations*

$$x \in A \qquad et \qquad x \in B$$

soient équivalentes.

Remarque 2. Il est facile de donner du signe $\in$ une interprétation intuitive rendant le Théorème (sic) 2 « évident ». Pour cela, il faut imaginer chaque objet mathématique comme une collection d'autres objets (d'où le mot *ensemble*); la relation $x \in y$ signifie alors (lorsqu'elle est vraie) que x est l'un des objets qui composent l'ensemble y, et le Théorème 2 affirme que, pour que deux collections d'objets soient identiques, il faut et il suffit qu'elles contiennent les mêmes objets (i.e. que tout objet appartenant à la première appartienne à la seconde, et réciproquement).

Dans la pratique, on imagine souvent les objets mathématiques soit comme étant des collections d'autres objets comme on vient de le dire, soit comme étant des objets « individuels » (on verra au n° 6 qu'il n'y a aucune contradiction entre ces deux interprétations); l'interprétation « naturelle » dans chaque cas dépend du contexte, et le lecteur, après un peu d'entraînement, la détectera facilement. En général, quand on pense à un objet mathématique comme à un « ensemble » on le désigne par une lettre majuscule, et quand on le regarde au contraire comme un « élément » d'un ensemble on le désigne souvent par une lettre minuscule — c'est ce qu'on a fait dans l'énoncé du Théorème 2. Cette règle n'est qu'une simple coutume, et souffre de nombreuses exceptions.

Remarque 3. Le fait qu'on puisse interpréter *concrètement* les objets mathématiques comme des collections ou ensembles d'autres objets n'a bien entendu rien à voir avec le problème consistant à définir *mathématiquement* la notion d'ensemble (les machines à démontrer, elles, n'ont pas d'intuition sensible...); ce problème ne peut être résolu que par les méthodes du § 0, n° 9, ou des méthodes analogues.

On trouve, dans certains manuels de Mathématiques à l'usage des élèves des Lycées, la déclaration suivante : « on appelle collection tout ensemble d'objets de même nature ». La première objection à cette « définition » est qu'elle réduit le mot « collection » au mot « ensemble »; or les deux termes sont évidemment synonymes : on est donc en présence d'un simple calembour. La seconde objection est que les auteurs des manuels en question n'éprouvent aucune difficulté à former une « collection » en réunissant deux « collections » quelconques, par exemple une collection de pommes et une collection de poires : il s'ensuit donc que des pommes et des poires sont des objets « de même nature » !

Cet exemple montre bien à quelles absurdités on aboutit en essayant, pour des raisons « pédagogiques », de donner du mot ensemble (ou du mot collection), une définition élémentaire. Il serait assurément bien préférable de dire qu'on regarde la notion d'ensemble comme une notion primitive

qu'on ne définit pas (et que tout le monde comprend intuitivement), et à l'aide de laquelle on peut construire des relations sur lesquelles on peut raisonner logiquement.

3. *Parties d'un ensemble*

A partir du signe d'appartenance, nous allons introduire une abréviation qui se note

$$\subset$$

et s'appelle le **signe d'inclusion**; étant donnés deux ensembles A et B, on représentera par

$$A \subset B$$

la relation suivante :

pour tout x, la relation $x \in A$ implique la relation $x \in B$.

Autrement dit, $A \subset B$ est une relation signifiant que tout élément de A est aussi dans B.

La relation $A \subset B$ se lit A **est contenu dans** B ou B **contient** A ou A **est une partie de** B, et on l'écrit aussi

$$B \supset A.$$

THÉORÈME 3. *La relation d'inclusion possède les propriétés suivantes:*

a) *les relations* $A \subset B$ *et* $B \subset C$ *impliquent la relation* $A \subset C$;

b) *pour que l'on ait* $A = B$ *il faut et il suffit que l'on ait* $A \subset B$ *et* $B \subset A$.

L'assertion (*a*) signifie que si la relation $x \in A$ implique la relation $x \in B$, et si celle-ci implique la relation $x \in C$, alors la première de ces trois relations implique la dernière — ce qui, logiquement, n'est autre que le « principe du syllogisme » ou la règle (TL 1) du § 0. L'assertion (*b*) se réduit évidemment au Théorème 2.

THÉORÈME 4. *Soit* $R\{x\}$ *une relation dans laquelle figure une variable* x*; pour tout ensemble* X, *il existe une partie* A *de* X *et une seule qui possède la propriété suivante: pour que l'on ait* $x \in A$, *il faut et il suffit que les relations* $x \in X$ *et* $R\{x\}$ *soient vraies.*

On dit que A est **l'ensemble des** $x \in X$ **tels que l'on ait la relation** $R\{x\}$.

Exemple 1. Prenons pour X l'ensemble des nombres entiers et pour $R\{x\}$ la relation « x est divisible par 2 »; alors A est l'ensemble des entiers pairs.

Remarque 4. Intuitivement, A est la collection formée par ceux des objets $x \in X$ qui possèdent la propriété exprimée par la relation $R\{x\}$, de sorte que l'existence de A est intuitivement évidente. Mathématiquement, on ne peut établir le Théorème 4 sans utiliser des axiomes qui sont beaucoup moins évidents que lui. Le lecteur débutant devra donc se borner à admettre l'énoncé.

¶ *Remarque* 5. Malgré ce qu'indique le bon sens, *il n'est pas vrai* que, pour toute relation $R\{x\}$, il existe un *ensemble* (au sens précis du § 0) dont les éléments sont *tous* les objets x tels que la relation $R\{x\}$ soit vraie (le Théorème affirme seulement qu'on peut le faire en se bornant à considérer des objets x appartenant à une ensemble X *donné d'avance*), et c'est pour avoir omis de prendre cette précaution que les mathématiciens ont été conduits, à la fin du siècle dernier, à découvrir les célèbres « paradoxes de la théorie des ensembles ».

Prenons par exemple la relation $x \notin x$; supposons qu'il existe un ensemble A tel que les relations $x \in A$ et $x \notin x$ soient équivalentes; substituant A à x on en déduirait, en utilisant la règle (TL 7) du § 0, que la relation $A \in A$ est équivalente à sa négation $A \notin A$, autrement dit que les Mathématiques sont contradictoires ! (A vrai dire, nul n'est certain du contraire à l'heure actuelle, mais toutes les fois qu'on découvre une contradiction on est en tous cas autorisé à en conclure que l'hypothèse d'où on l'a tirée est fausse).

La notion d'**ensemble de tous les ensembles** — il s'agirait d'un ensemble X tel que l'on ait $x \in X$ pour tout x — est non moins contradictoire; car, moyennant le Théorème 4, elle permettrait de parler de l'ensemble de tous les x tels que $x \notin x$, ce qui est impossible comme on vient de le voir.

Ces exemples montrent que l'usage du mot « ensemble » est soumis en Mathématiques à des limitations que l'intuition n'enseigne pas.

Comme illustration du Théorème 4, prenons une ensemble X, une partie M de X, et pour $R\{x\}$ la relation $x \notin M$; la partie A de X ainsi obtenue s'appelle le **complémentaire de M dans X**, et se note

$$X - M \qquad \text{ou} \qquad \complement_X M$$

(nous utiliserons uniquement la notation $X - M$); c'est donc l'ensemble des éléments de X qui n'appartiennent pas à M.

THÉORÈME 5. *Soient* M *et* N *des parties d'un ensemble* X; *alors les relations*

$$M \subset N \qquad \text{et} \qquad X - N \subset X - M$$

sont équivalentes. Pour toute partie M *de* X, *on a*

$$X - (X - M) = M.$$

L'ensemble $X - (X - M)$ se compose des $x \in X$ pour lesquels la relation $x \in X - M$ est fausse; or celle-ci, pour $x \in X$, équivaut à la négation de $x \in M$; donc, $X - (X - M)$ se compose des $x \in X$ pour lesquels la négation de $x \in M$ est fausse, i.e. pour lesquels la relation $x \in M$ est vraie, d'où la formule

$$X - (X - M) = M.$$

Supposons $M \subset N \subset X$; comme la relation $x \in M$ implique la relation $x \in N$, la négation de la seconde implique la négation de la première; donc la relation $x \in X - N$ implique la relation $x \in X - M$, et par suite on a $X - N \subset X - M$. Inversement,

cette relation implique, par le même raisonnement,

$$X - (X - M) \subset X - (X - N),$$

i.e. $M \subset N$, ce qui termine la démonstration.

¶ *Remarque* 6. La « démonstration » précédente est fort insuffisante logiquement car elle fait usage de la signification intuitive du mot « ensemble ». Pour démontrer correctement la relation $X - (X - M) = M$ par exemple, on devrait introduire les relations

$$R : x \in M, \qquad S : x \in X;$$

l'hypothèse que M est une partie de X signifie que R implique S; et la relation $X - (X - M) = M$ signifie donc que R et la relation

$$S \text{ et } [\text{non } (S \text{ et } (\text{non } R))]$$

sont équivalentes si R implique S. On devrait naturellement, pour établir l'équivalence des deux relations précédentes, utiliser les règles de démonstration du § 0.

Ceci montre que des assertions en apparence évidentes cessent d'être simples lorsqu'on veut effectivement les *démontrer*; c'est ce que les Grecs avaient déjà remarqué.

4. *Ensemble vide*

Soit X un ensemble; parmi les parties de X figure X lui-même, ce qui permet de considérer l'ensemble

$$\emptyset = X - X,$$

qu'on appelle la **partie vide de** X; la relation $x \in \emptyset$ est donc équivalente à la conjonction des relations

$$x \in X \qquad \text{et} \qquad x \notin X,$$

de sorte qu'il n'existe évidemment aucun objet x tel que $x \in \emptyset$.

L'ensemble $\emptyset = X - X$ ne dépend pas de l'ensemble X — autrement dit, on a

$$X - X = Y - Y$$

quels que soient les ensembles X et Y; en effet, $X - X$ et $Y - Y$ n'ont pas de mal à posséder exactement les mêmes éléments, puisqu'ils n'en possèdent aucun...

¶ *Remarque* 7. La démonstration précédente n'en est cependant pas une; d'après le Théorème 2 on doit prouver que la relation $x \in X - X$ et la relation $x \in Y - Y$ sont équivalentes; désignant par R la relation $x \in X$ et par S la relation $x \in Y$, tout revient à prouver que la relation

$$R \quad \text{et} \quad (\text{non } R)$$

est équivalente à la relation

$$S \quad \text{et} \quad (\text{non } S),$$

ce qui provient du fait plus général que, quelles que soient les relations R et T, la relation

$$(R \quad \text{et} \quad (\text{non } R)) \Longrightarrow T$$

est vraie (§ 0, *Exercice* 2).

L'ensemble $\emptyset = X - X$, qui est donc toujours le même, s'appelle aussi pour cette raison **ensemble vide**; il ne possède aucun élément, plus exactement la relation $x \in \emptyset$ est fausse quel que soit x.

Il est clair qu'on a

$$\emptyset \subset X$$

pour tout ensemble X, et cette propriété caractérise l'ensemble vide; car si deux ensembles A et B vérifient $A \subset X$ et $B \subset X$ pour *tout* ensemble X, on voit en particulier que $A \subset B$ et $B \subset A$, d'où $A = B$.

5. *Ensembles à un, deux éléments*

Soit x un objet mathématique; il existe alors un ensemble et un seul, noté $\{x\}$, possédant la propriété suivante : la relation

$$y \in \{x\} \qquad \text{équivaut à} \qquad y = x$$

Un ensemble de ce type s'appelle un **ensemble à un élément.** Il est clair que les relations

$$\{x\} = \{y\}, \qquad x = y$$

sont équivalentes quels que soient x et y.

On utilisera très souvent (et sans y référer explicitement) le résultat suivant :

THÉORÈME 6. *Pour qu'un ensemble* X *soit un ensemble à un élément, il faut et il suffit qu'il satisfasse aux conditions suivantes:*

a) X *est non vide;*

b) on a $x = y$ *pour tout* $x \in X$ *et tout* $y \in X$.

Les conditions sont évidemment nécessaires. Supposons-les inversement réalisées et choisissons un élément x de X (c'est possible puisque X est non vide); la relation $y = x$ implique trivialement $y \in X$, et inversement la relation $y \in X$ implique $y = x$ d'après l'hypothèse *b*) de l'énoncé; les conditions $y \in X$ et $y \in \{x\}$ sont donc équivalentes, ce qui prouve que $X = \{x\}$ et achève la démonstration.

Soient maintenant x et y deux objets mathématiques; il existe alors un et un seul ensemble, noté $\{x, y\}$, dont les seuls éléments soient x et y — autrement dit tel que

la relation

$$z \in \{x, y\}$$

soit équivalente à la relation

$$z = x \quad \text{ou} \quad z = y.$$

Un ensemble de ce type s'appelle un **ensemble à deux éléments** si $x \neq y$; si $x = y$, il est clair que

$$\{x, y\} = \{x, x\} = \{x\}$$

est un ensemble à un élément.

On définirait de même les ensembles à trois, quatre,... éléments. Les ensembles qu'on obtient ainsi sont appelés les **ensembles finis**, les autres étant appelés les **ensembles infinis**; ces deux notions seront à nouveau étudiées en détail au § 5.

¶ *Remarque* 8. L'existence d'ensembles à un, deux, trois,... éléments, intuitivement évidente, ne peut pas se démontrer — plus exactement, l'assertion « quels que soient x et y il existe un ensemble dont les seuls éléments soient x et y » est un *axiome* d'où l'on peut déduire facilement l'existence des autres catégories d'ensembles finis.

Ajoutons que l'existence d'ensembles infinis est aussi l'un des axiomes des Mathématiques (l'ensemble des nombres entiers est infini — mais nous n'avons pas donné de définition *mathématique* des nombres entiers, ni démontré *mathématiquement* qu'ils appartiennent à un même ensemble...). Voir § 5.

6. *Ensemble des parties d'un ensemble donné*

Soit X un ensemble; il existe — c'est, ici encore, l'un des axiomes des Mathématiques — un et un seul ensemble, noté

$$\mathfrak{P}(X),$$

qui possède la propriété suivante : *les éléments de* $\mathfrak{P}(X)$ *sont les parties de* X, autrement dit les relations

$$Y \in \mathfrak{P}(X) \quad \text{et} \quad Y \subset X$$

sont équivalentes. On dit que $\mathfrak{P}(X)$ est **l'ensemble des parties de X.**

Remarque 9. L'opération consistant à passer d'un ensemble X à l'ensemble $\mathfrak{P}(X)$ permet de construire des ensembles de plus en plus compliqués. On verra plus loin (§ 5) que si un ensemble X est fini et comporte n éléments, alors $\mathfrak{P}(X)$ est fini et en comporte 2^n. Ainsi, les ensembles

$$\emptyset, \quad \mathfrak{P}(\emptyset), \quad \mathfrak{P}(\mathfrak{P}(\emptyset)), \quad \mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\emptyset))), \quad \mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\emptyset)))), \ \ldots$$

comportent respectivement (voir § 5, *Remarque* 6)

$$0, \quad 1, \quad 2, \quad 2^2 = 4, \quad 2^4 = 16, \quad 2^{16} = 65.536, \quad 2^{65.536}, \ \ldots$$

éléments, ce qui, à partir de l'ensemble vide, permet de former des ensembles si compliqués qu'il devient rapidement impossible d'en énumérer pratiquement les éléments.

On notera d'autre part qu'on a

$$X \in \mathfrak{P}(X)$$

pour tout ensemble X; ceci montre, comme on l'avait annoncé dans la *Remarque* 2, que tout « ensemble » d'objets est lui-même un « élément » d'un autre « ensemble » — la façon la plus simple de le voir serait du reste d'écrire la relation

$$X \in \{X\}.$$

§ 2. La notion de fonction

1. Couples

Nous avons introduit, au § précédent, les signes $=$ et $\in$, qui servent à construire des relations. Nous allons maintenant introduire une opération qui sert à construire des objets mathématiques.

Cette opération consiste à former, à l'aide de deux objets mathématiques x et y énoncés dans cet ordre, un troisième objet, que l'on note

$$(x, y)$$

et qu'on appelle le **couple** (x, y); et l'opération consistant à former des couples est soumise à une seule règle d'emploi, que voici : *pour que l'on ait*

$$(x, y) = (u, v)$$

il faut et il suffit que l'on ait

$$x = u \qquad et \qquad y = v.$$

En particulier, on ne peut avoir $(x, y) = (y, x)$ que si $x = y$, ce qui montre que l'ordre dans lequel on écrit les deux objets figurant dans un couple est essentiel. On aura soin en particulier de ne pas confondre le couple (x, y) avec l'ensemble $\{x, y\}$ défini au § 1.

Remarque 1. On peut considérer la notion de couple comme un *signe fondamental* (§ 0, n° 1) qui, avec les signes

$$\text{ou}, \quad \text{non}, \quad \tau, \quad \square, \quad =, \quad \in$$

et des lettres, permettrait d'écrire les Mathématiques en langage formalisé. Mais on peut aussi exprimer la notion de couple à l'aide des autres signes fondamentaux (ou d'abréviations qui s'y ramènent) : il suffit de prendre comme définition de (x, y) l'ensemble

$$\{\{x\}, \{x, y\}\}$$

dont les *éléments* sont l'ensemble $\{x\}$ et l'ensemble $\{x, y\}$ — il est en effet visible qu'en définissant ainsi un couple on satisfait à l'axiome fondamental donnant la condition d'égalité de deux couples. Toutefois cette seconde méthode met l'accent sur un aspect de la notion de couple qui est parfaitement dénué d'intérêt. Il est donc bien préférable de s'en tenir à la méthode adoptée plus haut, la seule et unique question ayant une importance mathématique étant en effet de connaître les conditions pour que deux couples soient égaux.

On dit qu'un objet z est un couple s'il existe des objets x et y tels que $z = (x, y)$; en vertu de la règle énoncée plus haut, les objets x et y sont alors bien déterminés par la donnée de z; on dit que x est la **première projection** et y la **seconde projection** de z, et on écrit alors

$$x = pr_1(z), \qquad y = pr_2(z).$$

On dit qu'un ensemble G *est un* **graphe** *si tout élément de* G *est un couple.* Lorsqu'il en est ainsi, il existe deux ensembles X et Y caractérisés par les conditions suivantes : la relation $x \in X$ (resp. $y \in Y$) équivaut à l'existence d'un $z \in G$ tel que $x = pr_1(z)$ (resp. $y = pr_1(z)$). On écrit alors

$$X = pr_1(G), \qquad Y = pr_2(G).$$

On peut étendre comme suit la notion de couple. Étant donnés trois objets x, y, z on pose

$$(x, y, z) = ((x, y), z);$$

on dit que (x, y, z) est un **triplet**; pour que l'on ait

$$(x', y', z') = (x'', y'', z'')$$

il faut et il suffit que l'on ait

$$x' = x'', \qquad y' = y'', \qquad z' = z'';$$

en effet la relation considérée s'écrit $((x', y'), z') = ((x'', y''), z'')$, donc équivaut à $(x', y') = (x'', y'')$ et $z' = z''$, donc à $x' = x''$, $y' = y''$ et $z' = z''$.

De même, étant donnés quatre objets x, y, z, t on pose

$$(x, y, z, t) = ((x, y, z), t)$$

et on dit que (x, y, z, t) est un **quadruplet**, etc, etc...

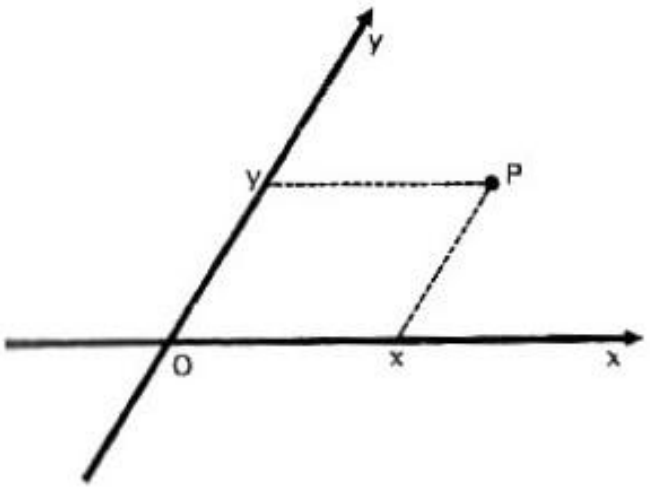
2. *Produit cartésien de deux ensembles*

Soient X et Y deux ensembles; on peut démontrer (à l'aide des méthodes du § 0) qu'il existe un ensemble Z caractérisé par la propriété suivante : *pour que l'on ait $z \in Z$, il faut et il suffit qu'il existe $x \in X$ et $y \in Y$ tels que $z = (x, y)$*. On dit que Z est le **produit cartésien** de X et Y, et on écrit

$$Z = X \times Y.$$

L'ensemble produit $X \times Y$ est donc l'ensemble des couples (x, y) où $x \in X$ et $y \in Y$.

Le fait que l'opération précédente soit nommée en l'honneur de Descartes s'explique comme suit. Dans le « plan » de la Géométrie élémentaire, choisissons deux axes de coordonnées Ox, Oy et des unités de longueur sur ces axes; on peut alors définir l'abscisse et l'ordonnée de tout point P du plan; les désignant par x et y, il est naturel de ne pas faire de différence entre le point P et le couple (x, y); du reste, si P′ est un autre point, de coordonnées x' et y', la relation $P = P'$ équivaut évidemment à $x = x'$ et $y = y'$, i.e. à $(x, y) = (x', y')$. Désignant par $\mathbf{R}$ l'ensemble des nombres réels, on voit donc que le choix d'un système de coordonnées dans le plan permet d'assimiler le plan à l'ensemble



$$\mathbf{R} \times \mathbf{R} = \mathbf{R}^2$$

des couples de nombres réels — c'est ce qui justifie la référence à l'inventeur des systèmes de coordonnées.

Soient A, B, X, Y quatre ensembles; alors les relations

$$A \subset X \quad \text{et} \quad B \subset Y \quad \text{impliquent} \quad A \times B \subset X \times Y;$$

la réciproque de cette assertion évidente est exacte pourvu que A et B soient non vides; en effet, si $A \times B \subset X \times Y$ et si B contient au moins un élément b, alors pour tout $a \in A$ on a $(a, b) \in A \times B$ donc $(a, b) \in X \times Y$ donc $(a, b) = (x, y)$ pour un $x \in X$ et un $y \in Y$ donc $a = x$ pour un $x \in X$, donc $A \subset X$; et on montrerait de même que $B \subset Y$ pourvu que A soit non vide.

Si par contre l'un des ensembles A, B est vide, on a toujours $A \times B \subset X \times Y$, pour la raison triviale que

$$A \times B = \emptyset \quad \text{si} \quad A = \emptyset \quad \text{ou si} \quad B = \emptyset;$$

en effet, si la relation $A \times B = \emptyset$ est fausse, alors $A \times B$ contient un couple (x, y) au moins, on a donc $x \in A$ et $y \in B$, ce qui montre que A et B sont non vides.

La notion de produit cartésien s'étend au cas de plusieurs facteurs; si X, Y, Z, T, ... sont des ensembles, on définit

$$X \times Y \times Z = (X \times Y) \times Z; \qquad X \times Y \times Z \times T = (X \times Y \times Z) \times T; \ \ldots$$

Les éléments de $X \times Y \times Z$ sont visiblement les triplets (x, y, z) définis au n° 1, avec $x \in X$, $y \in Y$, $z \in Z$; de même les éléments de $X \times Y \times Z \times T$ sont les quadruplets (x, y, z, t) avec $x \in X$, $y \in Y$, $z \in Z$ et $t \in T$.

On observera que la relation

$$(X \times Y) \times Z = X \times (Y \times Z)$$

est *fausse*; les éléments du premier membre sont en effet les objets de la forme $((x,y),z)$ avec $x \in X, y \in Y, z \in Z$, tandis que ceux du second sont les objets de la forme $(x,(y,z))$ — or la règle d'égalité de deux couples ne permet pas d'écrire que l'on a $((x,y),z) = (x,(y,z))$ quels que soient x, y, z. Néanmoins, dans la pratique, on convient de ne faire aucune différence entre $((x,y),z)$ et $(x,(y,z))$, et de considérer les ensembles $(X \times Y) \times Z$ et $X \times (Y \times Z)$ comme identiques; cette convention de langage est, à strictement parler, contradictoire — comme beaucoup d'autres conventions que nous introduirons par la suite; cependant, les contradictions auxquelles conduit son emploi sont « sans importance », et le lecteur, après avoir acquis une certaine habitude des raisonnements de la théorie des Ensembles, évitera sans peine de tomber dans ces difficultés.

Enfin, si X est un ensemble, on pose

$$X^2 = X \times X, \qquad X^3 = X \times X \times X, \qquad X^4 = X \times X \times X \times X,$$

et ainsi de suite. Par exemple, et $\mathbf{R}$ désignant l'ensemble des nombres réels, $\mathbf{R}^4$ est l'ensemble des quadruplets (x, y, z, t) formés de quatre nombres réels; c'est ce que les physiciens appelent « l'espace à quatre dimensions » ou « l'espace-temps »; le lecteur débutant fera bien de ne pas se laisser impressionner par cette terminologie, l'expérience montrant qu'il n'est pas plus difficile de raisonner dans $\mathbf{R}^4$ que dans $\mathbf{R}^3$ ou $\mathbf{R}^{100}$ ...

Il va de soi que, comme dans le cas d'un produit de deux facteurs, *un produit d'ensembles est vide dès qu'un des facteurs est vide.*

3. *Graphes et fonctions*

Soient X et Y deux ensembles; on appelle fonction définie sur l'ensemble X et à valeurs dans l'ensemble Y toute opération consistant à faire correspondre, à chaque élément x de X, un élément y de Y, qui dépend de x suivant une loi bien déterminée : par exemple la fonction $y = \sin x$ lorsque $X = Y = \mathbf{R}$.

La définition (sic) précédente contient malheureusement de nombreux mots dont nous n'avons pas donné de définition mathématique — par exemple, que signifie l'expression « faire correspondre »? En prenant pour argent comptant la définition en question, on n'obtient une fois de plus qu'un mauvais calembour.

On a donc été obligé de modifier la définition précédente, et de la remplacer par la suivante (qui, dans le cas classique, reviendrait à définir une fonction en se donnant d'avance sa « courbe représentative » dans le plan, méthode que même un physicien considérerait comme fort raisonnable) : *on appelle* **fonction** *un triplet*

$$f = (G, X, Y)$$

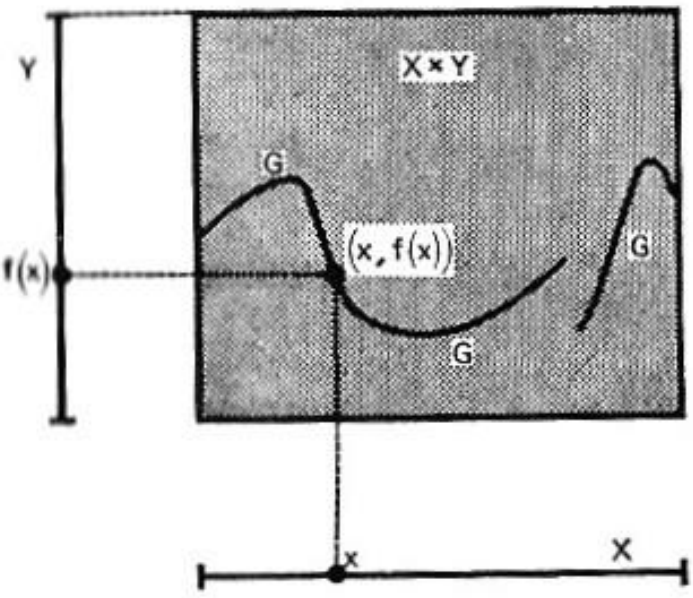
où G, X, Y *sont des ensembles assujettis à vérifier les conditions suivantes:*

(*V* 1) : *on a* $G \subset X \times Y$;

(*V* 2) : *pour tout* $x \in X$ *il existe un et un seul* $y \in Y$ *tel que* $(x, y) \in G$.

La condition (F 1) signifie que G est un graphe (n° 1); on dit que G est le **graphe de la fonction** f; d'après (F 2), pour tout $x \in X$ il existe un $z \in G$ tel que $x = pr_1(z)$; on a donc

$$pr_1(G) = X, \qquad pr_2(G) \subset Y.$$



Soit x un élément de X; l'unique élément y de Y tel que $(x, y) \in G$ s'appelle la **valeur de la fonction** f **en** x, et on utilise pour le désigner la notation

$$y = f(x);$$

il est alors clair que le graphe G de f est l'ensemble des couples de la forme $(x, f(x))$ où $x \in X$, ce qui est conforme à l'idée intuitive qu'on se fait d'une fonction.

Étant donnée une fonction $f = (G, X, Y)$, on dit que X est **l'ensemble de départ** et Y **l'ensemble d'arrivée** de f.

Étant donnés deux ensembles X et Y, on appelle **application de X dans** Y toute fonction ayant X pour ensemble de départ et Y pour ensemble d'arrivée; les mots « fonction » et « application » sont donc synonymes, mais dans la pratique il est souvent plus commode de dire « soit f une application de X dans Y » que de dire « soit f une fonction définie sur X et à valeurs dans Y ». Du reste, au lieu de dire

soit f une application de X dans Y,

on dit souvent

soit une application $f : X \to Y$,

ou encore

soit une application $X \xrightarrow{f} Y$.

Il arrive aussi qu'au lieu de désigner une fonction par une lettre telle que f, g, etc... on la désigne par la « formule » qui permet de calculer $f(x)$ en fonction de x; ainsi, dans le cas où $X = Y = \mathbf{R}$, quand on dit

considérons l'application $x \longmapsto x^3$ de $\mathbf{R}$ dans $\mathbf{R}$,

ou quand on dit

considérons sur $\mathbf{R}$ la fonction x^3,

on doit traduire par

considérons l'application f de $\mathbf{R}$ dans $\mathbf{R}$ telle que $f(x) = x^3$ pour tout $x \in \mathbf{R}$;

on devrait même traduire comme suit :

considérons l'application $f = (G, \mathbf{R}, \mathbf{R})$ où G est l'ensemble des couples $(x, y) \in \mathbf{R} \times \mathbf{R}$ tels que

$$y = x^3.$$

Soient

$$f = (G, X, Y) \qquad \text{et} \qquad f' = (G', X', Y')$$

deux fonctions ou applications. Vu la condition d'égalité de deux triplets, la relation $f = f'$ signifie qu'on a

$$G = G', \qquad X = X', \qquad Y = Y';$$

comme $X = pr_1(G)$ et $X' = pr_1(G')$, la première condition implique du reste la seconde. D'autre part, si les conditions

$$X = X', \qquad Y = Y'$$

sont déjà vérifiées, pour écrire que $G = G'$ (i.e. que $f = f'$) il suffit évidemment d'écrire que l'on a

$$f'(x) = f(x) \quad \text{pour tout} \quad x \in X,$$

puisque G est formé des couples $(x, f(x))$ et G' des couples $(x, f'(x))$, avec $x \in X$.

Remarque 2. Vu le peu d'exemples qu'ils ont eu l'occasion d'étudier, les débutants pensent souvent que toute fonction est définie par une « formule » permettant de calculer $f(x)$ en fonction de x. On ne saurait avoir une idée plus fausse de la notion de fonction.

Tout d'abord le mot « formule » ne veut rien dire aussi longtemps qu'on n'en a pas donné une définition précise. Il est vraisemblable que, pour le débutant, une « formule » est une succession d'opérations algébriques plus ou moins compliquées effectuées sur la variable x. Malheureusement, de telles opérations n'ont aucun sens lorsque x est un élément d'un ensemble arbitraire. Si au contraire la variable x est un nombre réel, auquel cas on peut en effet effectuer des calculs algébriques sur x, les fonctions qu'on obtient de cette façon sont tellement particulières qu'on a abandonné depuis au moins 200 ans l'idée de définir ainsi la notion de fonction; en fait, les besoins de l'Analyse ont obligé les mathématiciens à inventer des catégories de plus en plus générales de fonctions, et l'étude de ces fonctions plus ou moins « arbitraires » a conduit à l'invention de la théorie des Ensembles, puis de la théorie moderne de l'Intégration, puis de celle des espaces topologiques, etc... — autrement dit, c'est en généralisant de plus en plus la notion de fonction qu'on a été amené à construire une bonne partie des Mathématiques contemporaines.

L'étude des fonctions qui peuvent se définir par des « formules algébriques » n'est pas pour autant dépourvue d'intérêt — c'est au contraire l'objet d'une branche des Mathématiques (la Géométrie Algébrique) qui n'a jamais été aussi active qu'à l'heure actuelle. Mais les méthodes qu'on emploie pour étudier ces fonctions, et les problèmes qu'on se pose à leur

sujet, n'ont rien de commun avec ceux auxquels on s'intéresse en Algèbre élémentaire et en Analyse.

Remarque 3. Soient X et Y deux ensembles; alors les applications de X dans Y sont les éléments d'un ensemble, qu'on note

$$Y^X$$

et qu'on appelle l'**ensemble des applications de X dans Y**; comme une fonction est un triplet $f = (G, X, Y)$ avec $G \subset X \times Y$, on voit que l'ensemble des applications de X dans Y est contenu dans $\mathfrak{P}(X \times Y) \times \{X\} \times \{Y\}$.

Dans la pratique, et X et Y étant donnés, on identifie généralement une application de X dans Y à son graphe $G \subset X \times Y$; avec cette convention, l'ensemble des applications de X dans Y apparait comme une partie de l'ensemble $\mathfrak{P}(X \times Y)$.

Remarque 4. On utilise fréquemment une notion voisine de celle de fonction, la notion de **famille** qu'on définit intuitivement comme suit : soit I un ensemble; alors, pour construire une **famille ayant I pour ensemble d'indices** (ou une **famille indexée par** I), on se donne, pour chaque $i \in I$, un objet dépendant de i (*sans* préciser d'avance dans quel ensemble on choisit les objets ainsi associés aux éléments de I); si l'on note x_i l'objet associé à $i \in I$, on désigne généralement la famille considérée par la notation

$$(x_i)_{i \in I}.$$

Mathématiquement, une famille indexée par I est un graphe G possédant les deux propriétés suivantes : on a $pr_1(G) = I$, et pour tout $i \in I$ il existe un seul $z \in G$ tel que $pr_1(z) = i$; écrivant $z = (i, x_i)$ on retrouve la définition intuitive exposée ci-dessus.

Soit $(x_i)_{i \in I}$ une famille; on dit que c'est une **famille d'éléments d'un ensemble** X si l'on a $x_i \in X$ pour tout $i \in I$; il existe toujours un tel ensemble X, par exemple $pr_2(G)$ où G est le graphe de la famille considérée. On dit de même qu'une famille $(A)_{i \in I}$ est une **famille de parties d'un ensemble** X si l'on a $A_i \subset X$ pour tout $i \in I$.

Lorsque l'ensemble d'indices I d'une famille est l'ensemble dont les éléments sont les entiers naturels 1, 2, 3,..., on dit que cette famille est une **suite**; on désigne souvent une suite par une notation telle que

$$(x_n)_{n \geqslant 1};$$

se donner une suite d'éléments d'un ensemble X revient donc à choisir des éléments

$$x_1, x_2, \ldots, x_n, \ldots$$

de X, ou encore à se donner une application $n \to x_n$ de l'ensemble des entiers naturels dans X.

Il va de soi que nous admettons ici l'existence d'un ensemble (qu'on note généralement **N**) dont les éléments sont les nombres 0, 1,... L'existence de cet ensemble est évidente aussi longtemps qu'on donne aux mots « ensemble» et « nombre entier » leur sens intuitif; quant à démontrer *mathématiquement* l'existence de l'ensemble **N**, c'est une toute autre affaire, car non seulement

on doit tout d'abord construire une théorie *mathématique* correcte des nombres entiers, mais en outre admettre l'existence d'ensembles comportant une infinité d'éléments; or l'existence de pareils ensembles, intuitivement claire, ne peut pas se démontrer — c'est l'un des *axiomes* des Mathématiques...

4. *Images directes et images réciproques*

Soit f une application d'un ensemble X dans un ensemble Y; étant donnée une partie A de X, on appelle **image de A par** f l'ensemble des $y \in Y$ possédant la propriété suivante :

$$\text{il existe } x \in A \text{ tel que } y = f(x).$$

Si A se réduit à un seul élément x, son image est évidemment réduite au seul élément $f(x)$. Dans le cas d'une partie A quelconque de X, on désigne l'image de A par la notation $f(A)$ — qui est incorrecte, puisque $f(A)$ n'a de sens que pour $A \in X$ à strictement parler.

On a évidemment

$$f(\emptyset) = \emptyset$$

pour toute application f.

Soit toujours f une application de X dans Y, et considérons une partie B de Y; on appelle **image réciproque de B par** f l'ensemble des $x \in X$ tels que $f(x) \in B$; cet ensemble se désigne par la notation

$$\overset{-1}{f}(B),$$

qu'on abrège souvent (mais à tort) en $f^{-1}(B)$.

On a évidemment les relations

$$A \subset \overset{-1}{f}(f(A)) \quad \text{pour toute partie de A de X,}$$
$$B \supset f(\overset{-1}{f}(B)) \quad \text{pour toute partie B de Y;}$$

mais on n'a pas le droit, dans ces relations, de remplacer les inclusions par des égalités.

On dit qu'une application $f: X \to Y$ est **constante sur une partie A de X si** $f(A)$ se réduit à un seul élément, i.e. (§ 1, Théorème 6) si l'on a

$$f(x') = f(x'') \quad \text{quels que soient} \quad x' \in A \quad \text{et} \quad x'' \in A.$$

On dit que f est une **application constante** si f est constante sur X tout entier.

Soit f une application d'un ensemble X dans lui-même; on dit qu'une partie A de X est **stable par** f si l'on a $f(A) \subset A$, autrement dit si la relation $x \in A$ implique la relation $f(x) \in A$. Lorsque A se réduit à un seul élément x, cela signifie évidemment que

$$f(x) = x;$$

on dit alors que x est un **point fixe de** f.

La notion d'image d'un ensemble par une application apparaît en Géométrie élémentaire lorsqu'on parle par exemple de « la transformée d'une droite par une rotation » : une droite est un ensemble (de points), une rotation est une certaine application (de l'ensemble des points de l'espace dans lui-même), et la transformée en question n'est autre que l'image de l'ensemble considéré par cette application.

5. *Restrictions et prolongements de fonctions*

Soit $f = (G, X, Y)$ une fonction, et considérons un ensemble $X' \subset X$. Soit G' l'ensemble des $z \in G$ tels que $pr_1(z) \in X'$; alors le triplet $f' = (G', X', Y)$ est une fonction — il est clair en effet que $G' \subset X' \times Y$, et que tout pour $x \in X'$ il existe un et un seul $z \in G'$ tel que $pr_1(z) = x$, à savoir $z = (x, f(x))$. La fonction f', application de X' dans Y telle que

$$f'(x) = f(x) \quad \text{pour tout} \quad x \in X',$$

s'appelle la **restriction de f à X'**.

Remarque 5. La restriction de f à X' se désigne souvent par la notation

$$f' = f|X' \qquad \text{ou} \qquad f_{X'}.$$

D'autre part, étant données deux applications f et g dont les ensembles de départ contiennent un même ensemble X, on dit que f et g **coïncident sur** X si l'on a

$$f(x) = g(x) \quad \text{pour tout} \quad x \in X;$$

c'est par exemple le cas si $f|X = g|X$, mais la réciproque n'est pas tout à fait exacte, malgré les apparences...

Soient $f = (G, X, Y)$ et $f' = (G', X', Y')$ deux applications; on dit que f **est un prolongement de f'** si l'on a les relations

$$X' \subset X, \quad Y' \subset Y, \qquad \text{et} \qquad f(x) = f'(x) \quad \text{pour tout} \quad x \in X'.$$

C'est par exemple le cas si l'on prend pour f' la restriction de f à une partie de X. Notons le résultat suivant : *soient* X', X, Y *trois ensembles, avec* $X' \subset X$, *et* f *une application de* X' *dans* Y ; *si* Y *est non vide il existe une application de* X *dans* Y *qui prolonge* f. Pour construire une application $g : X \to Y$ prolongeant f, on choisit une fois pour toutes un élément c de Y, et on pose

$$g(x) = \begin{cases} f(x) & \text{si } x \in X', \\ c & \text{si } x \notin X'. \end{cases}$$

Il existe bien entendu d'autres façons de prolonger f, celle qui précède n'est que la plus simple de toutes.

6. *Applications composées*

Soient X, Y, Z trois ensembles, et

$$f = (G, X, Y), \qquad g = (H, Y, Z)$$

deux applications de X dans Y et de Y dans Z respectivement. On va en déduire une troisième application

$$h = (K, X, Z)$$

de X dans Z, en posant

$$h(x) = g(f(x)) \quad \text{pour tout} \quad x \in X;$$

le graphe K de h est évidemment l'ensemble des couples (x, z) possédant la propriété suivante : il existe un $y \in Y$ tel que l'on ait $(x, y) \in G$ et $(y, z) \in H$.

L'application h se désigne par la notation

$$h = g \circ f;$$

on l'appelle la **composée des applications** f **et** g; l'application composée $g \circ f$ n'est définie que si l'ensemble d'arrivée de f est identique à l'ensemble de départ de g.

Remarque 6. La notion d'application composée remplace aujourd'hui celles de « fonction de fonction » et de « produit de deux transformations » qu'on utilisait autrefois dans certains cas particuliers. Prenons par exemple $X = Y = Z = \mathbf{R}$, ensemble des nombres réels, et $f(x) = x^2$, $g(x) = \sin x$; alors $g \circ f$ est la fonction

$$x \mapsto \sin(x^2),$$

et $f \circ g$ la fonction

$$x \mapsto \sin^2 x,$$

ce qui montre en passant que $f \circ g \neq g \circ f$ en général (quand les deux membres sont définis). Si l'on prend $X = Y = Z =$ l'espace (au sens de la « géométrie dans l'espace » — nous ne chercherons pas à en donner ici une définition correcte), on peut prendre pour f et g des « transformations » au sens géométrique du termes — rotations, translations, homothéties, etc...; l'application composée $f \circ g$ est alors le « produit » des transformations f et g, défini en Géométrie élémentaire.

Nous allons maintenant démontrer un théorème qui jouera par la suite un assez grand rôle :

THÉORÈME 1. *Soient* X, Y, Z *trois ensembles, et considérons deux applications*

$$f : X \to Y, \qquad h : X \to Z;$$

les conditions suivantes sont équivalentes:

a) il existe une application

$$g : Y \to Z$$

telle que l'on ait $h = g \circ f$;

b) quels que soient $x', x'' \in X$, *la relation*

$$f(x') = f(x'')$$

implique la relation

$$h(x') = h(x'').$$

La condition *a*) implique la condition *b*), car si *a*) est satisfaite on a

$$h(x') = g(f(x')) = g(f(x'')) = h(x'').$$

Nous allons maintenant établir que *b*) implique *a*).

Examinons tout d'abord le *cas particulier où* $f(X) = Y$. Pour construire g, on va construire son graphe $G \subset Y \times Z$, comme suit : G *est l'ensemble des couples* (y, z) *tels qu'il existe au moins un* $x \in X$ *vérifiant*

$$y = f(x), \qquad z = h(x).$$

[Cette construction de G est naturelle, car si l'on suppose le problème résolu G est l'ensemble des couples $(y, g(y))$; or, comme $f(X) = Y$, on peut écrire $y = f(x)$, et alors $g(y) = g(f(x)) = h(x)$, de sorte que G se compose bien des couples de la forme $(f(x), h(x))$ où x décrit X]. Montrons que l'ensemble G ainsi obtenu est effectivement le graphe d'une application $g : Y \to Z$ telle que $h = g \circ f$. Pour montrer que G est le graphe d'une fonction, on doit prouver que pour tout $y \in Y$, il existe un et un seul $z \in Z$ tel que $(y, z) \in G$; l'existence d'au moins un tel z est claire : il suffit de choisir un x tel que $y = f(x)$, et de prendre $z = h(x)$; d'autre part, supposons que G contienne (y, z') et (y, z'') ; il existe alors dans X des éléments x' et x'' tels que

$$\begin{aligned} y &= f(x'), \qquad z' = h(x'), \\ y &= f(x''), \qquad z'' = h(x''); \end{aligned}$$

on a alors $f(x') = f(x'')$, donc, d'après l'hypothèse *b*), $h(x') = h(x'')$, i.e. $z' = z''$, et ceci montre bien que G est le graphe d'une application g de Y dans Z. Pour prouver que $h = g \circ f$, considérons un $x \in X$; alors, par construction, G contient le couple $(f(x), h(x))$; par suite on a $h(x) = g(f(x))$, ce qui établit la relation cherchée.

Il reste à montrer que *b*) implique *a*) dans le *cas général*. Posons $Y' = f(X)$ et considérons l'application f' de X dans Y' donnée par $f'(x) = f(x)$ pour tout $x \in X$; remplaçant X, Y, Z, f, h par X, Y', Z, f', h il est clair qu'on est maintenant dans l'hypothèse *b*) avec $f'(X) = Y'$; d'après ce qu'on vient déjà d'établir, il existe donc une application g' de Y' dans Z telle que $h = g' \circ f'$; prolongeons alors g' en une application g de Y dans Z (ce qui est possible comme on l'a vu au n° 5) ; pour tout $x \in X$, on aura

$$h(x) = g'(f'(x)) = g(f'(x)) = g(f(x)),$$

d'où $h = g \circ f$, ce qui achève la démonstration.

L'explication « intuitive » du Théorème 1 est la suivante : comme la relation $f(x') = f(x'')$ implique $h(x') = h(x'')$, il suffit, pour savoir calculer $h(x)$, de connaître $f(x)$, de sorte que $h(x)$ doit être une fonction de $f(x)$. Mais bien entendu ce genre d'explication ne dispense pas de donner une démonstration correcte.

La propriété la plus importante de l'opération consistant à composer des applications est « l'associativité » de cette opération, exprimée par le résultat suivant :

THÉORÈME 2. *Quelles que soient les applications*

$$f : X \to Y, \qquad g : Y \to Z, \qquad h : Z \to T,$$

on a la relation

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Posons $g \circ f = u$ et $h \circ g = v$; en $x \in X$, la valeur du premier membre de la relation à établir est $h(u(x)) = h(g(f(x)))$, et celle du second membre est $v(f(x)) = h(g(f(x)))$, d'où évidemment le Théorème.

Le Théorème 2 permet de définir sans ambiguité $h \circ g \circ f$, et plus généralement des expressions telles que

$$f_1 \circ f_2 \cdots \circ f_n,$$

pourvu naturellement qu'elles aient un sens, i.e. que l'ensemble de départ de chaque facteur soit égal à l'ensemble d'arrivée du facteur suivant. En particulier, pour toute application $u : X \to X$ et tout entier $q \geqslant 1$, on peut définir

$$u^q = u \circ u \cdots \circ u \quad (q \text{ facteurs}).$$

7. *Applications injectives*

On dit qu'une application $f : X \to Y$ est **injective** lorsque, quels que soient x', $x'' \in X$, la relation

$$f(x') = f(x'') \quad \text{implique} \quad x' = x'',$$

ou encore si

$$x' \neq x'' \quad \text{implique} \quad f(x') \neq f(x'').$$

Une application injective s'appelle aussi une **injection.** Au lieu du mot « injectif » on utilisait autrefois (i.e. jusqu'en 1955 environ...) l'adjectif **biunivoque,** qu'on trouve encore dans de nombreux ouvrages, et qui présente l'inconvénient de ne pas correspondre à un substantif.

Exemple 1. Prenons $X = Y = \mathbf{R}$ et $f(x) = x^3$; alors f est injective car, si x et y sont des nombres réels, la relation $x^3 = y^3$ implique $x = y$. Par contre la fonction $f(x) = x^2$ ne l'est pas, car on a par exemple $f(-1) = f(1)$.

Exemple 2. Pour tout ensemble X, on définit l'**application identique de X dans X**

comme étant celle qui, à chaque $x \in X$, fait correspondre x lui-même; on la note souvent

$$j_X,$$

de sorte qu'on a, par définition,

$$j_X(x) = x \quad \text{pour tout} \quad x \in X;$$

on utilise aussi couramment la notation *id* au lieu de j_X. Cela dit, l'application identique de X dans X est évidemment injective.

On notera que le graphe de j_X est, dans $X \times X$, l'ensemble des couples (x, x) avec $x \in X$; cet ensemble s'appelle la **diagonale du produit** $X \times X$.

Lorsque $X = \mathbf{R}$, l'application identique se réduit à la « fonction x » des Lycées et Collèges. Lorsque X est l'ensemble des points du plan, ou de l'espace, l'application identique n'est autre que la « transformation unité » de la Géométrie élémentaire.

Exemple 3. Soient X et Y deux ensembles tels que $X \subset Y$; l'application $j : X \to Y$ donnée par $j(x) = x$ pour tout $x \in X$ est évidemment injective; on l'appelle l'**injection canonique de** X **dans** Y (il y a en général beaucoup d'autres injections de X dans Y; l'adjectif « canonique » utilisé ici signifie que cette injection particulière est obtenue par un procédé « naturel » ne comportant aucun élément d'arbitraire, et ne faisant intervenir que les données intrinsèques de la situation envisagée — à savoir un ensemble Y et une partie X de Y).

THÉORÈME 3. *Soient* X *et* Y *des ensembles non vides et f une application de* X *dans* Y. *Les propriétés suivantes sont équivalentes:*

a) f est injective;

b) il existe une application g: $Y \to X$ *telle que* $g \circ f$ *soit l'application identique de* X *dans* X.

Considérons en effet les deux applications

$$f : X \to Y, \qquad j_X : X \to X;$$

on cherche une condition nécessaire et suffisante pour qu'il existe une application $g : Y \to X$ telle que $j_X = g \circ f$; cette condition est donnée par le Théorème 1 du n° 6 : c'est que la relation

$$f(x') = f(x'') \quad \text{implique} \quad j_X(x') = j_X(x''), \quad \text{i.e.} \quad x' = x'',$$

ce qui signifie précisément que f est injective; d'où le Théorème.

Exemple 4. Prenons $X = \mathbf{R}_+$, ensemble des nombres réels $x \geqslant 0$, et $Y = \mathbf{R}$, ensemble de tous les nombres réels (de signe quelconque); enfin prenons $f(x) = x^2$; cette application est évidemment injective (on notera qu'elle cesserait de l'être si l'on remplaçait $\mathbf{R}_+$ par $\mathbf{R}$); donc il existe une application $g : \mathbf{R} \to \mathbf{R}_+$ telle que $g \circ f$ soit l'identité, i.e. telle que l'on ait $g(x^2) = x$ pour tout $x \in \mathbf{R}_+$. Cette condition exprime évidemment que

$$g(y) = \sqrt{y} \quad \text{si} \quad y \geqslant 0.$$

Autrement dit, on peut prendre pour g *toute* fonction à valeurs positives qui, pour $y \geqslant 0$, coincide avec la fonction $\sqrt{y}$.

Notons à ce sujet qu'on pourrait évidemment démontrer directement le Théorème 3 sans passer par l'intermédiaire du Théorème 1. L'application cherchée g doit vérifier la relation $g(f(x)) = x$; comme c'est la seule condition que doit vérifier g, on peut déjà choisir arbitrairement $g(y)$ lorsque y n'appartient pas à $f(X)$; si au contraire $y \in f(X)$, il existe un $x \in X$, *et un seul* puisque f est injective, tel que $y = f(x)$, et on doit alors choisir $g(y) = x$.

8. *Applications surjectives et bijectives*

On dit qu'une application $f : X \to Y$ *est* **surjective** *si* $f(X) = Y$, autrement dit si pour tout $y \in Y$ il existe *au moins* un $x \in X$ tel que $y = f(x)$. Dire que f est injective signifie, par contre, que pour tout $y \in Y$ il existe *au plus* un $x \in X$ tel que $y = f(x)$.

On dit qu'une application $f : X \to Y$ *est* **bijective** *si elle est à la fois injective et surjective,* autrement dit si, pour tout $y \in Y$, il existe un et un seul $x \in X$ tel que $y = f(x)$.

Il est clair par exemple que, pour tout ensemble X, l'application identique j_X est bijective.

Une application surjective s'appelle encore une **surjection,** une application bijective, une **bijection.** Une bijection d'un ensemble X dans lui-même s'appelle une **permutation** de X. L'ensemble des permutations de X se désigne par la notation

$$\mathfrak{S}(X).$$

Si $X = \{1, 2, \ldots, n\}$ on écrit $\mathfrak{S}_n$ au lieu de $\mathfrak{S}(X)$.

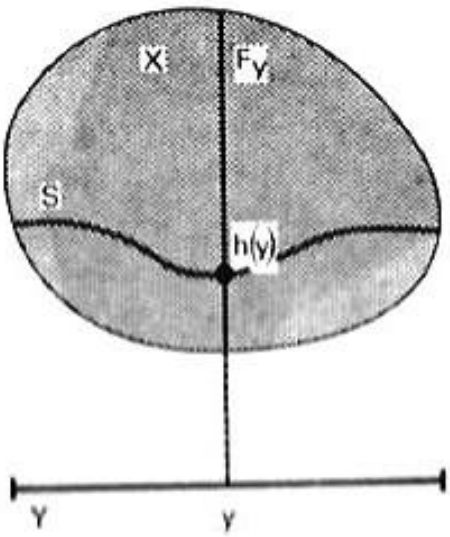
Remarque 6. Dans l'ancienne terminologie, on disait « soit f une application de X **sur** Y » au lieu de dire « soit f une application surjective de X dans Y »; et pour exprimer que f était bijective, on disait fréquemment « biunivoque sur ».

THÉORÈME 4. *Soit* $f : X \to Y$ *une application. Les conditions suivantes sont équivalentes:*

a) f est surjective;

b) il existe une application $h : Y \to X$ *telle que* $f \circ h$ *soit l'application identique de* Y *sur* Y.

Si *b)* est remplie, on a $f(h(y)) = y$ pour tout $y \in Y$, donc tout $y \in Y$ est de la forme $f(x)$, et f est donc bien surjective.



Inversement supposons f surjective, et pour chaque $y \in Y$ désignons par F_y l'ensemble des $x \in X$ tels que $f(x) = y$; c'est une partie *non vide* de X; pour construire l'application cherchée h, il suffit alors de choisir au hasard un élément dans chaque ensemble F_y; notant $h(y)$ cet élément, on définit bien ainsi une application h de Y dans X telle que $f(h(y)) = y$ pour tout $y \in Y$, ce qui achève la démonstration.

¶¶ *Remarque* 7. La seconde partie de la démonstration ci-dessus semblera sans doute correcte, et même évidente, au débutant; mais elle est fort loin d'être mathématiquement complète (on imagine difficilement une machine à démontrer « choisissant au hasard » un élément dans chaque F_y ...). La démonstration correcte s'obtiendrait à l'aide de l'opération de Hilbert du § 0, n° 9 : à défaut de mieux, on obtient une fonction h en posant

$$h(y) = \tau_x(f(x) = y).$$

Notons que le problème revient aussi à construire une partie S de X telle que, pour tout $y \in Y$, l'intersection $S \cap F_y$ soit un ensemble à *un* élément exactement (on peut alors prendre pour $h(y)$ cet élément). La possibilité de construire un tel ensemble (évidente si l'on utilise l'opération de Hilbert) est connue sous le nom d'**axiome du choix.** Certains mathématiciens, jusqu'à une date récente, mettaient cet axiome en doute, mais on a pu démontrer d'abord qu'il est logiquement compatible avec les autres axiomes (K. Gödel, 1939) puis qu'il en est logiquement indépendant (P. Cohen, 1963). Ces « démonstrations » n'ont naturellement de sens que par rapport à un système métamathématique convenable.

Remarque 8. L'application h dont le Théorème 4 assure l'existence n'est généralement pas unique. Prenons par exemple $X = \mathbf{R}$, $Y = \mathbf{R}_+$ et $f(x) = x^2$, de sorte que f est bien surjective (tout nombre réel positif admet une racine carrée). On a alors, entre autres possibilités, les fonctions h suivantes :

$$h_1(y) = \sqrt{y} \quad \text{pour tout} \quad y \geqslant 0;$$
$$h_2(y) = -\sqrt{y} \quad \text{pour tout} \quad y \geqslant 0;$$
$$h_3(y) = \begin{cases} +\sqrt{y} & \text{si } y \geqslant 0 \text{ est rationnel;} \\ -\sqrt{y} & \text{si } y \geqslant 0 \text{ est irrationnel.} \end{cases}$$

La fonction h_3 (qu'il est pratiquement impossible de représenter graphiquement — son graphe n'est pas une « courbe » au sens naïf du terme) pourra sembler étrange au débutant; elle n'est cependant pas moins bonne, du point de vue de la pure théorie des Ensembles, que les deux premières.

THÉORÈME 5. *Soit $f : X \to Y$ une application. Les conditions suivantes sont équivalentes:*

a) f est bijective;

b) il existe des applications $g, h : Y \to X$ telles que l'on ait

$$g \circ f = j_X, \qquad f \circ h = j_Y.$$

De plus, si ces conditions sont vérifiées, les applications g et h sont uniques et coïncident.

L'équivalence des propriétés *a*) et *b*) résulte immédiatement des Théorèmes 3 et

4, puisque « bijective » signifie « injective et surjective ». Pour montrer qu'il existe une seule fonction g, une seule fonction h, et que $g = h$, il suffit de montrer que *toute* fonction g est égale à *toute* fonction h. Or on a

$$(g \circ f) \circ h = g \circ (f \circ h)$$

d'après le Théorème 2, ce qui s'écrit encore $j_X \circ h = g \circ j_Y$; mais il est clair que

$$j_X \circ h = h, \qquad g \circ j_Y = g,$$

ce qui achève la démonstration.

Soit

$$f : X \to Y$$

une bijection; il existe une et une seule application $g : Y \to X$ telle que l'on ait

$$g \circ f = j_X, \qquad f \circ g = j_Y;$$

autrement dit,

$$g(f(x)) = x, \qquad f(g(y)) = y.$$

On dit que g est **l'application réciproque de** f, et on la désigne habituellement par la notation

$$\overset{-1}{f},$$

qu'on écrit souvent (mais à tort) f^{-1}. Il est clair que pout tout $x \in X$ et tout $y \in Y$ les relations

$$y = f(x), \qquad x = \overset{-1}{f}(y)$$

sont équivalentes, et que si $G \subset X \times Y$ est le graphe de f, alors le graphe de $\overset{-1}{f}$ est l'ensemble des couples (y, x) tels que $(x, y) \in G$ (dans le cas classique où $X = Y = \mathbf{R}$, cela signifie que le second graphe se déduit du premier par une *symétrie par rapport à la première bissectrice*).

Remarque 9. Au lieu d'application réciproque on dit souvent « application **inverse** ». Cette terminologie peut entraîner de graves confusions; par exemple (pour $X = Y = \mathbf{R}$), tout le monde pensera que la fonction « inverse » de la fonction x est la fonction $1/x$, alors que la fonction « réciproque » de x est x elle-même (d'une manière générale, une application identique est égale à son application réciproque).

Nous ne saurions trop mettre en garde le débutant contre la tentation de croire que le langage est un détail sans importance; connaître avec précision les définitions de tous les termes techniques, et employer ceux-ci dans leur sens propre, toujours le même, est strictement indispensable à la compréhension des Mathématiques — c'est même parfois suffisant...

THÉORÈME 6. *Soient* $f : X \to Y$ *et* $g : Y \to Z$ *deux applications; si* f *et* g *sont injectives (resp. surjectives) il en est de même de* $g \circ f$*; si* f *et* g *sont bijectives, il en est de même de* $g \circ f$*, et l'on a la relation*

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

Enfin, si f est bijective, il en est de même de f^{-1} et on a la relation

$$(f^{-1})^{-1} = f.$$

Si f et g sont injectives, la relation $g(f(x')) = g(f(x''))$ implique $f(x') = f(x'')$ puisque g est injective, donc implique $x' = x''$ puisque f est injective; ceci montre que $g \circ f$ est injective. Si f et g sont surjectives, pour tout $z \in Z$ il existe un $y \in Y$ tel que $z = g(y)$, puis un $x \in X$ tel que $y = f(x)$, d'où $z = g(f(x))$, ce qui montre que $g \circ f$ est surjective.

Si f et g sont bijectives, il en est donc de même de $g \circ f$; de plus, on a

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ j_Y \circ f = f^{-1} \circ f = j_X,$$

ce qui montre que l'application réciproque de $g \circ f$ est bien donnée par la formule de l'énoncé.

Enfin, si f est bijective, les formules

$$f^{-1} \circ f = j_X, \qquad f \circ f^{-1} = j_Y$$

montrent que f^{-1} est à la fois surjective et injective (Théorèmes 4 et 5 appliqués à f^{-1}), et admet f pour application réciproque; ceci achève la démonstration.

Remarque 10. Non seulement la formule « évidente »

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

est fausse, mais le second membre n'a de sens que si l'on suppose $Y = Z$ (auquel cas la formule est toujours aussi fausse...).

9. *Fonctions de plusieurs variables*

On appelle **fonction de deux variables** toute fonction dont l'ensemble de départ est un produit de deux ensembles, ou contenu dans un tel produit. Si f est une fonction de deux variables, définie sur une partie A d'un produit $X \times Y$, la valeur de f en un point (x, y) de A, que l'on devrait noter $f((x, y))$, se désigne pratiquement par la notation

$$f(x, y).$$

On définirait de façon analogue les fonctions de trois, quatre,... variables, pour lesquelles on emploie les notations $f(x, y, z)$, $f(x, y, z, t)$, etc...

Mathématiquement parlant, il n'existe aucune différence, sinon dans les notations utilisées, entre fonctions d'une variable et fonctions de plusieurs variables — comme en effet nous n'avons fait, au n° 3, aucune espèce d'hypothèse sur les ensembles de départ des fonctions, tout ce qu'on a dit dans ce § s'applique sans aucun changement aux fonctions de « plusieurs » variables. La distinction entre « une » et « plusieurs » variables provient du fait que, jusqu'à une date récente, le mot « variable » désignait ce qu'on appelle aujourd'hui « variable réelle », les

« fonctions d'une variable » étant celles qui sont définies sur une partie de **R**, ensemble des nombres réels, tandis que les « fonctions de trois variables », par exemple, étaient celles dont l'ensemble de départ est une partie de $\mathbf{R}^3$. C'est entre autres pour éviter d'avoir à tenir compte de ces distinctions qu'on a été amené à donner la définition générale du n° 3, laquelle englobe *toutes* les notions de fonction actuellement connues, sans aucune exception.

Dans la pratique, il est souvent utile de considérer des fonctions dont les ensembles de départ *et* d'arrivée sont des produits. Considérons par exemple une application

$$f : X \times Y \to U \times V \times W,$$

où X, Y, U, V, W sont des ensembles quelconques. Considérons les projections

$$\begin{aligned} pr_1 &: U \times V \times W \to U, \\ pr_2 &: U \times V \times W \to V, \\ pr_3 &: U \times V \times W \to W; \end{aligned}$$

il est clair qu'on a

$$z = (pr_1(z),\ pr_2(z),\ pr_3(z)) \quad \text{pour tout} \quad z \in U \times V \times W,$$

par définition même des projections. Considérant les applications

$$\begin{aligned} f_1 &= pr_1 \circ f : X \times Y \to U, \\ f_2 &= pr_2 \circ f : X \times Y \to V, \\ f_3 &= pr_3 \circ f : X \times Y \to W, \end{aligned}$$

on aura donc

$$f(x, y) = (f_1(x, y),\ f_2(x, y),\ f_3(x, y))$$

quels que soient $x \in X$ et $y \in Y$. Ainsi, pour construire f, il est nécessaire et suffisant de connaître les trois fonctions f_1, f_2, f_3 à valeurs dans U, V, W respectivement, et définies sur $X \times Y$. Dans la pratique on écrit

$$f = (f_1, f_2, f_3)$$

(cette notation est en contradiction avec celle qui désigne un triplet, mais cette contradiction est, ici encore, « sans importance »).

Exemple 5. Soient X et Y deux ensembles, et considérons l'application

$$f : X \times Y \to Y \times X$$

donnée par

$$f(x, y) = (y, x);$$

c'est évidemment une bijection, et l'application réciproque n'est autre que $(y, x) \mapsto (x, y)$. On dit que f est la **bijection canonique de $X \times Y$ sur $Y \times X$.** Lorsque $X = Y = \mathbf{R}$, f est la « symétrie par rapport à la première diagonale ».

Exemple 6. Soient X, Y, Z trois ensembles; l'application

$$f : X \times (Y \times Z) \to (X \times Y) \times Z$$

donnée par

$$f[(x, (y, z))] = ((x, y), z)$$

est bijective. On dit, ici encore, que c'est la **bijection canonique de** $X \times (Y \times Z)$ **sur** $(X \times Y) \times Z$. Comme on l'a dit au n° 2, on ne fait pas de différence, dans la pratique, entre $(x, (y, z))$ et $((x, y), z)$.

§ 3. Réunions et intersections

1. Réunion et intersection de deux ensembles

Soient X et Y deux ensembles; on appelle **intersection de X et Y** l'ensemble noté

$$X \cap Y$$

et défini comme suit : la relation $z \in X \cap Y$ est équivalente à la conjonction des relations

$$z \in X \quad \text{et} \quad z \in Y;$$

autrement dit, $X \cap Y$ est formé des objets appartenant à la fois à X et à Y. On appelle d'autre part **réunion de X et Y** l'ensemble noté

$$X \cup Y$$

et défini comme suit : la relation $z \in X \cup Y$ est équivalente à la relation

$$z \in X \quad \text{ou} \quad z \in Y;$$

autrement dit, $X \cup Y$ est formé des objets appartenant soit à X, soit à Y, soit à X et à Y.

¶ *Remarque* 1. L'existence d'ensembles $X \cap Y$ et $X \cup Y$ possédant les propriétés indiquées est évidente intuitivement, mais ne l'est pas du tout mathématiquement. L'existence de $X \cap Y$ s'obtient à l'aide du Théorème 4 du § 1 (qu'on applique à X et à la relation $x \in Y$). Celle de $X \cup Y$ s'obtiendrait de même si l'on savait d'avance qu'il existe un ensemble contenant à la fois X et Y (on appliquerait alors le Théorème 4 du § 1 à cet ensemble et à la relation $z \in X$ ou $z \in Y$); mais l'existence d'un ensemble contenant à la fois X et Y est un axiome (ou résulte d'un axiome plus général servant à former la réunion d'une famille d'ensembles, cf. n° 2), de sorte qu'il est inutile de chercher à la démontrer mathématiquement.

Il est clair qu'on a les relations

$$X \cap Y \subset X, \; Y \subset X \cup Y;$$

en outre, soit Z un ensemble quelconque; pour que Z soit contenu dans X et dans Y, il faut et il suffit qu'on ait $z \in X$ et $z \in Y$ pour tout $z \in Z$, i.e. $z \in X \cap Y$, i.e. $Z \subset X \cap Y$; ainsi, $X \cap Y$ *est le plus grand ensemble contenu à la fois dans* X *et dans* Y. De même, pour que Z contienne X et Y, il faut et il suffit que Z contienne $X \cup Y$, de sorte que $X \cup Y$ *est le plus petit ensemble contenant à la fois* X *et* Y.

On dit que deux ensembles X et Y sont **disjoints** lorsque

$$X \cap Y = \emptyset,$$

i.e. lorsque X et Y n'ont aucun élément commun.

Les règles de calcul gouvernant l'emploi des signes $\cup$ et $\cap$ sont très simples, et nous les utiliserons souvent sans référence; le lecteur établira lui-même ces règles, dont voici les principales :

$$X \cap Y = Y \cap X, \qquad X \cup Y = Y \cup X,$$
$$X \cap (Y \cap Z) = (X \cap Y) \cap Z, \qquad X \cup (Y \cup Z) = (X \cup Y) \cup Z,$$
$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z),$$
$$X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z),$$
$$(X - A) \cap (X - B) = X - (A \cup B) \quad \text{si} \quad A, B \subset X.$$

2. *Réunion d'une famille d'ensembles* (*)

Soit $(A_i)_{i \in I}$ une famille d'ensembles (§ 2, n° 3, *Remarque* 4); on appelle **réunion de cette famille** l'ensemble A défini comme suit : la relation $x \in A$ est équivalente à la relation

$$\text{il existe un } i \in I \text{ tel que l'on ait } x \in A_i.$$

Lorsque I se compose de deux éléments seulement, notés i et j par exemple, il est clair que la réunion n'est autre que l'ensemble $A_i \cup A_j$ défini au n° précédent. Dans le cas d'un ensemble d'indices I quelconque, l'*existence* de la réunion est un axiome des Mathématiques, et on doit donc se borner à l'admettre. L'*unicité* de la réunion (i.e. le fait qu'il existe *au plus un* ensemble A possédant la propriété indiquée) peut par contre se démontrer, à l'aide du § 1, Théorème 2.

Remarque 2. Quand on parle d'une famille d'ensembles $(A_i)_{i \in I}$, on ne suppose pas que les A_i soient des parties d'un même ensemble indépendant de l'indice i; mais l'existence de la réunion montre qu'en fait il en est bien ainsi (il y a même plus : compte-tenu du Théorème 4 du § 1, l'existence d'un ensemble contenant tous les A_i est *équivalente* à l'existence de leur réunion).

Pour désigner la réunion d'une famille d'ensembles $(A_i)_{i \in I}$, on emploie la notation

$$\bigcup_{i \in I} A_i;$$

(*) **En première lecture, on pourra se dispenser d'étudier ce n° et le suivant, ou les considérer comme des exercices.**

bien que la notation en question fasse intervenir une lettre i (qui désigne, intuitivement, un « élément variable » de I), le résultat ne dépend évidemment pas de i, et on peut, dans la notation précédente, utiliser au lieu de la lettre i toute autre lettre non encore employée par ailleurs.

THÉORÈME 1. *Soit* A *la réunion d'une famille d'ensembles* $(A_i)_{i\in I}$. *Pour qu'un ensemble* X *contienne* A_i *quel que soit* $i \in I$, *il faut et il suffit que* X *contienne* A.

Supposons que X contienne tous les A_i; si $x \in A$, il existe un i tel que $x \in A_i$, et comme $A_i \subset X$ on a $x \in X$; donc $A \subset X$. Inversement, si X contient A, pour montrer que X contient tous les A_i il suffit d'établir que $A \supset A_i$ pour tout i, ce qui est clair.

THÉORÈME 2 (associativité de la réunion). *Soient* $(A_i)_{i\in I}$ *et* $(I_\lambda)_{\lambda\in\Lambda}$ *deux familles d'ensembles, et supposons que*

$$I = \bigcup_{\lambda\in\Lambda} I_\lambda ;$$

on a alors

$$\bigcup_{i\in I} A_i = \bigcup_{\lambda\in\Lambda} \Big(\bigcup_{i\in I_\lambda} A_i \Big).$$

Posons en effet

$$B_\lambda = \bigcup_{i\in I_\lambda} A_i ;$$

pour qu'un x appartienne à la réunion de la famille $(A_i)_{i\in I}$ il faut et il suffit qu'il existe un $i \in I$ tel que $x \in A_i$; comme I est réunion des I_λ, cela signifie qu'il existe un $\lambda \in \Lambda$ et un $i \in I_\lambda$ tels que $x \in A_i$, donc qu'il existe un $\lambda \in \Lambda$ tel que $x \in B_\lambda$; par suite, la réunion de la famille $(A_i)_{i\in I}$ est identique à celle de la famille $(B_\lambda)_{\lambda\in\Lambda}$, ce qui achève la démonstration.

Remarque 3. Le Théorème 2 exprime que, pour calculer une réunion, on peut en partager les termes en groupes et remplacer chaque groupe par sa réunion.

THÉORÈME 3. *Soient* $f : X \to Y$ *une application, et* $(A_i)_{i\in I}$ *une famille de parties de* X. *On a alors*

$$f\Big(\bigcup_{i\in I} A_i \Big) = \bigcup_{i\in I} f(A_i).$$

Pour $y \in Y$, la relation

$$y \in \bigcup_{i\in I} f(A_i)$$

équivaut à l'existence d'un $i \in I$ tel que $y \in f(A_i)$, i.e. à l'existence d'un $i \in I$ et

d'un $x \in A$ tels que $y = f(x)$, i.e. à l'existence d'un x vérifiant

$$y = f(x) \qquad \text{et} \qquad x \in \bigcup_{i \in I} A_i ,$$

ce qui achève la démonstration.

3. *Intersection d'une famille d'ensembles*

On appelle **intersection d'une famille** non vide (*) **d'ensembles** $(A_i)_{i \in I}$ l'ensemble A défini comme suit : la relation $x \in A$ est équivalente à la relation

$$x \in A_i \quad \text{pour tout} \quad i \in I.$$

Cette intersection se désigne par la notation

$$\bigcap_{i \in I} A_i .$$

THÉORÈME 4. *Soit* A *l'intersection d'une famille non vide d'ensembles* $(A_i)_{i \in I}$. *Pour qu'un ensemble* X *soit contenu dans* A, *il faut et il suffit qu'il soit contenu dans* A_i *pour tout* $i \in I$.

La démonstration est analogue à celle du Théorème 1, et peut être laissée au lecteur à titre d'exercice.

THÉORÈME 5 (associativité de l'intersection). *Soient* $(A_i)_{i \in I}$ *et* $(I_\lambda)_{\lambda \in \Lambda}$ *deux familles d'ensembles; on suppose* I, Λ *et les* I_λ *non vides, et que*

$$I = \bigcup_{\lambda \in \Lambda} I_\lambda ;$$

on a alors la relation

$$\bigcap_{i \in I} A_i = \bigcap_{\lambda \in \Lambda} \left(\bigcap_{i \in I_\lambda} A_i \right).$$

La démonstration est analogue à celle du Théorème 2.

THÉORÈME 6. *Soient* $f : X \to Y$ *une application et* $(A_i)_{i \in I}$ *une famille non vide de parties de* X. *On a alors*

$$f\left(\bigcap_{i \in I} A_i \right) \subset \bigcap_{i \in I} f(A_i) ;$$

si f *est injective, on a*

$$f\left(\bigcap_{i \in I} A_i \right) = \bigcap_{i \in I} f(A_i).$$

(*) Cette condition signifie que I est non vide. Si I est vide, la relation « on a $x \in A_i$ pour tout $i \in I$ » est vérifiée quel que soit x, et par suite ne définit pas un ensemble (sinon l'ensemble de tous les ensembles...).

Si $x \in A_i$ pour tout i, on a $f(x) \in f(A_i)$ pour tout i, ce qui prouve la première assertion du Théorème. Supposons maintenant f injective, et considérons un élément y de l'intersection des $f(A_i)$; pour tout i il existe donc un élément de A_i, soit x_i, tel que $y = f(x_i)$; mais comme f est injective, il existe un seul x tel que $y = f(x)$, et on a donc nécessairement $x = x_i$ pour tout i; ainsi, $x \in A_i$ pour tout i, et y appartient à l'image par f de l'intersection des A_i; on a donc

$$\bigcap f(A_i) \subset f\left(\bigcap A_i\right),$$

ce qui établit la seconde assertion du Théorème puisqu'on a de toute façon l'inclusion opposée.

Remarque 4. La seconde assertion du Théorème ci-dessus peut être en défaut si f n'est pas injective. Prenons par exemple pour Y un ensemble contenant au moins deux éléments a et b, pour X le produit $Y \times Y$, et pour f l'application pr_2; soient A l'ensemble des couples (a, y), $y \in Y$, et B l'ensemble des couples (b, y), $y \in Y$; on a évidemment $A \cap B = \emptyset$, donc $f(A \cap B) = \emptyset$; par contre, $f(A) = f(B) = Y$, de sorte que $f(A) \cap f(B)$ est non vide.

THÉORÈME 7. *Soient* $f : X \to Y$ *une application et* $(A_i)_{i \in I}$ *une famille non vide de parties de* Y. *On a alors*

$$f^{-1}\left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} f^{-1}(A_i).$$

En effet, pour qu'un $x \in X$ appartienne au premier membre, il faut et il suffit que $f(x)$ appartienne à l'intersection des A_i, i.e. que $f(x) \in A_i$ pour tout i, autrement dit que $x \in f^{-1}(A_i)$ pour tout i, ou enfin que x appartienne au second membre, d'où le Théorème.

On démontrerait de même la formule

$$f^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} f^{-1}(A_i).$$

THÉORÈME 8. *Soit* $(A_i)_{i \in I}$ *une famille non vide de parties d'un ensemble* X. *On a les relations*

$$X - \bigcup_{i \in I} A_i = \bigcap_{i \in I} (X - A_i); \qquad X - \bigcap_{i \in I} A_i = \bigcup_{i \in I} (X - A_i)$$

Soit en effet $x \in X$; la relation

$$x \in X - \bigcap_{i \in I} A_i$$

équivaut à la négation de la relation

$$\text{pour tout } i \in I \text{ on a } x \in A_i,$$

donc à la relation

$$\text{il existe un } i \in I \text{ tel que } x \notin A_i,$$

donc à

$$\text{il existe un } i \in I \text{ tel que } x \in X - A_i,$$

donc à

$$x \in \bigcup_{i \in I} (X - A_i),$$

ce qui établit la première formule. La seconde s'en déduit en tenant compte de la relation $X - (X - A) = A$, valable pour toute partie A de X.

§ 4. Relations d'équivalence

1. Relations d'équivalence

Soit R une relation faisant intervenir deux variables x, y. On dit que c'est une **relation d'équivalence** lorsque les conditions suivantes sont remplies :

a) *la relation* $R\{x, x\}$ *est vraie pour tout* x;

b) *la relation* $R\{x, y\}$ *implique la relation* $R\{y, x\}$;

c) *les relations* $R\{x, y\}$ *et* $R\{y, z\}$ *impliquent la relation* $R\{x, z\}$.

Il est clair par exemple que la relation $x = y$ est une relation d'équivalence.

Une notion voisine de la précédente, et plus utile dans la pratique, est celle de **relation d'équivalence sur un ensemble** E; on appelle ainsi une relation $R\{x, y\}$ faisant intervenir deux variables x, y, et vérifiant les conditions suivantes :

(R 0) : *la relation* $R\{x, y\}$ *implique* $x \in E$ *et* $y \in E$;
(R 1) : *la relation* $R\{x, x\}$ *est vraie pour tout* $x \in E$;
(R 2) : *la relation* $R\{x, y\}$ *implique* $R\{y, x\}$;
(R 3) : *les relations* $R\{x, y\}$ *et* $R\{y, z\}$ *impliquent la relation* $R\{x, z\}$.

Si R est une relation d'équivalence sur un ensemble E, on appelle **graphe de R** l'ensemble $G \subset E \times E$ des couples $(x, y) \in E \times E$ tels que la relation $R\{x, y\}$ soit vraie; la relation $R\{x, y\}$ est alors équivalente à $(x, y) \in G$, et la condition (R 1) signifie que G contient la *diagonale* du produit $E \times E$.

Pour construire un exemple (qui, on le montrera plus loin, conduit à *toutes* les relations d'équivalence sur un ensemble E), considérons une application f de E dans un ensemble quelconque M, et prenons pour $R\{x, y\}$ la relation

$$f(x) = f(y);$$

on obtient évidemment une relation d'équivalence sur E, appelée la **relation d'équivalence associée à l'application** f.

Donnons maintenant quelques autres exemples.

Exemple 1. Pour tout ensemble E, la diagonale de $E \times E$ (§ 2, n° 6, *Exemple* 2) est le graphe d'une relation d'équivalence sur E — à savoir de la relation

$$x = y.$$

Exemple 2. Pour tout ensemble E, l'ensemble $E \times E$ est le graphe d'une relation d'équivalence sur E — à savoir la relation $(x, y) \in E \times E$; dans ce cas la relation $R\{x, y\}$ est donc vraie pour tout $x \in E$ et tout $y \in E$.

Exemple 3. On dit qu'un ensemble X est **équipotent** à un ensemble Y s'il existe une bijection de X sur Y. La relation « X est équipotent à Y » est une relation d'équivalence; en effet X est équipotent à X (considérer la bijection j_X, application identique); si X est équipotent à Y, alors Y est équipotent à X (car si f est une bijection de X sur Y, alors f^{-1} est une bijection de Y sur X); enfin si X est équipotent à Y et si Y est équipotent à Z, alors X est équipotent à Z (car si f est une bijection de X sur Y, et g une bijection de Y sur Z, alors $g \circ f$ est une bijection de X sur Z, en vertu du § 2, Théorème 7). Voir le § suivant.

Exemple 4. Soit **Z** l'ensemble des **entiers rationnels**, i.e. des nombres entiers positifs ou négatifs..., -2, -1, 0, 1, 2,.... Choisissons un entier $p \geqslant 1$ et considérons la relation

$$p \text{ divise } x - y$$

entre deux éléments x et y de **Z**; c'est une relation d'équivalence sur **Z**. En effet, il est clair que la relation « p divise $x - x$ » est toujours vraie; la relation « p divise $x - y$ » implique évidemment la relation « p divise $y - x$ »; enfin, si p divise $x - y$ et $y - z$, il est clair que p divise $x - z = (x - y) + (y - z)$.

La relation d'équivalence ainsi obtenue sur **Z** s'appelle la **congruence modulo** p, et elle se note classiquement

$$x \equiv y \pmod{p},$$

ce qui se lit « x est congru à y modulo p »; cela signifie que x et y ne diffèrent que d'un multiple de p. La théorie des congruences joue un rôle fondamental en Arithmétique depuis deux siècles. es.

Exemple 5. On obtient sur l'ensemble **R** des nombres réels une relation d'équivalence en écrivant

$$\text{il existe un entier } n \text{ tel que } x - y = 2\pi n;$$

on l'appelle la **congruence modulo** 2π. Cette relation d'équivalence est à la base de la définition *mathématique* des angles.

Exemple 6. Prenons pour E l'ensemble des couples (p, q) avec $p, q \in \mathbf{Z}$ et $q \neq 0$, et, pour $x = (p, q), y = (p', q')$, notons $R\{x, y\}$ la relation

$$pq' = p'q;$$

on obtient ainsi une relation d'équivalence sur E (ce n'est pas évident, mais on peut le démontrer à l'aide de quelques calculs élémentaires). Cette relation d'équivalence est à la base de la définition *mathématique* des nombres rationnels à partir des nombres entiers, comme on le verra plus tard (§ 29). On conseille au lecteur de vérifier lui-même, à titre d'exercice, que R est bien une relation d'équivalence sur l'ensemble E considéré.

Exemple 7. Prenons pour E l'ensemble des points de l'espace usuel et, une droite D étant choisie une fois pour toutes, considérons la relation

il existe une droite parallèle à D passant par x et y ;

on obtient alors une relation d'équivalence sur E.

Exemple 8. On prend pour E l'ensemble des triangles dans le plan, et pour $R\{x, y\}$ la relation

les triangles x et y sont égaux,

l'égalité des triangles étant définie (sic) comme en Géométrie élémentaire, i.e. en exigeant qu'il existe un déplacement qui transforme x en y (ce qui suppose qu'on sait ce qu'est un déplacement...). On obtient alors une relation d'équivalence dans E.

Soit R une relation d'équivalence sur un ensemble E; au lieu d'écrire $R\{x, y\}$, nous écrirons le plus souvent dans ce qui suit

$$x \equiv y \pmod{R},$$

notation inspirée de celle de l'Exemple 4. On a donc les propriétés suivantes : la relation

$$x \equiv x \pmod{R}$$

est vraie pour tout $x \in E$; la relation

$$x \equiv y \pmod{R}$$

implique (donc, par symétrie, est équivalente à) la relation

$$y \equiv x \pmod{R};$$

enfin, les relations

$$x \equiv y \pmod{R} \qquad \text{et} \qquad y \equiv z \pmod{R}$$

impliquent la relation

$$x \equiv z \pmod{R}.$$

2. *Quotient d'un ensemble par une relation d'équivalence*

Nous allons démontrer un résultat qui indique comment on obtient toutes les relations d'équivalence sur un ensemble :

THÉORÈME 1. *Soit* R *une relation d'équivalence sur un ensemble* E. *Il existe un ensemble* M *et une application* $f : E \to M$ *tels que les relations*

$$x \equiv y \pmod{R}$$

et

$$f(x) = f(y)$$

soient équivalentes.

On va non seulement démontrer l'existence de f et de M, mais construire explicitement un ensemble M et une application f satisfaisant à la condition énoncée.

Étant donné un $x \in E$, appelons **classe de x modulo** R l'ensemble $F_x \subset E$ formé des $y \in E$ tels que la relation

$$x \equiv y \pmod{R}$$

soit vraie — de sorte que les relations

$$x \equiv y \pmod{R}, \qquad y \in F_x$$

sont équivalentes. On va montrer que, pour $x, y \in E$, *les relations*

$$x \equiv y \pmod{R}$$

et

$$F_x = F_y$$

sont équivalentes.

Supposons en effet $x \equiv y \pmod{R}$; alors la relation $z \in F_y$, qui signifie $y \equiv z \pmod{R}$, implique $x \equiv z \pmod{R}$, donc $z \in F_x$; par suite, la relation $x \equiv y \pmod{R}$ implique $F_y \subset F_x$, donc aussi $F_x \subset F_y$, donc $F_x = F_y$. Inversement, supposons $F_x = F_y$; comme on a toujours $y \equiv y \pmod{R}$, et donc $y \in F_y$, il vient $y \in F_x$, donc $x \equiv y \pmod{R}$, et notre assertion est démontrée.

Nous pouvons maintenant démontrer le Théorème 1. Dans l'ensemble $\mathfrak{P}(E)$ des parties de E, considérons l'ensemble formé des parties F de E telles que l'on ait

$$F = F_x$$

pour au moins un $x \in E$ — c'est donc l'ensemble des classes d'équivalence des divers éléments de E; cet ensemble se note E/R et s'appelle le **quotient de l'ensemble E par la relation d'équivalence** R; définissons maintenant une application

$$f : E \to E/R$$

en posant

$$f(x) = F_x$$

i.e. en associant à chaque $x \in E$ sa classe modulo R; on a vu plus haut que la relation

$$x \equiv y \pmod{R},$$

équivaut à $F_x = F_y$; mais ceci s'écrit encore

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

et le Théorème est démontré.

L'application f qu'on vient de définir s'appelle **l'application canonique de E sur E/R**; elle est évidemment *surjective*, par construction même de E/R.

Notons que les classes F_x possèdent deux propriétés importantes : *la réunion des* F_x *est* E *tout entier* (en vertu du fait qu'on a $x \in F_x$ pour tout $x \in E$); d'autre part, *deux classes* F_x *et* F_y *quelconques sont ou bien identiques ou bien disjointes*, car si $F_x \cap F_y$ contient au moins un élément z, on a les relations

$$x \equiv z \pmod{R} \qquad \text{et} \qquad y \equiv z \pmod{R}$$

d'où, en faisant usage de (R 2) et (R 3), la relation

$$x \equiv y \pmod{R},$$

laquelle entraîne $F_x = F_y$ comme on l'a vu dans la démonstration du Théorème 1.

Remarque 1. La méthode de démonstration du Théorème 1 repose essentiellement sur la construction d'une application de l'ensemble E dans l'ensemble $\mathfrak{P}(E)$ des parties de E, à savoir l'application $x \to F_x$. On ne pourrait évidemment pas effectuer ce genre de construction si l'on ne connaissait que les fonctions classiques (d'une variable réelle, à valeurs réelles), et c'est ce genre de démonstration qui montre la nécessité de considérer des fonctions dont les ensembles de départ et d'arrivée sont arbitraires.

Notons d'autre part que la méthode de construction de E/R, qui pourra sembler étrange au débutant, s'utilise cependant dans la vie de tous les jours, comme le montre l'exemple (non mathématique !) que voici : on prend pour E la collection des hommes, et pour R la relation « x et y sont compatriotes »; on obtient ainsi évidemment une relation d'équivalence sur E. Pour un $x \in E$, la classe F_x est l'ensemble de tous les compatriotes de x; autrement dit c'est la nation à laquelle x appartient; par suite, l'ensemble quotient E/R est ici la collection des diverses nations existantes, et l'application canonique de E sur E/R consiste à associer à chaque homme la nation à laquelle il appartient...

Donnons quelques exemples de construction d'un ensemble quotient.

Exemple 9. Considérons sur $\mathbf{Z}$ la relation d'équivalence de l'*Exemple* 4 (congruence modulo p); pour tout $x \in \mathbf{Z}$, la classe F_x se compose évidemment des entiers de la forme $x + np$, où n est un entier arbitraire; ces classes sont en nombre p exactement (on suppose $p \geqslant 1$). En effet, pour tout $x \in \mathbf{Z}$ on peut écrire (« division avec reste de x par p »)

$$x = np + r \qquad (0 \leqslant r < p),$$

et la connaissance du reste r de la division de x par p détermine évidemment la classe F_x — autrement dit, toute classe modulo p est l'une des classes

$$F_0, \quad F_1, \quad \cdots, \quad F_{p-1};$$

ces p classes sont de plus deux à deux distinctes, car si des entiers r et r' compris entre 0 et $p - 1$ sont congrus modulo p, ils sont évidemment égaux.

On voit donc bien qu'ici l'ensemble E/R, qu'on désigne par la notation

$$\mathbf{Z}/p\mathbf{Z},$$

comporte exactement p éléments, éléments qu'on appelle les **entiers modulo** p. Un entier modulo p est donc un ensemble d'entiers ordinaires, à savoir l'ensemble de tous les entiers se déduisant d'un entier x donné par addition d'un multiple quelconque de p; on dit alors que l'entier x est un **représentant** de l'entier modulo p considéré. Tout entier modulo p admet un et un seul représentant compris entre 0 et $p-1$, ce qui permet de numéroter les entiers modulo p à l'aide des entiers $0, 1, \ldots, p-1$ qui les représentent. Par exemple, on représente les éléments de $\mathbf{Z}/6\mathbf{Z}$ à l'aide des entiers 0, 1, 2, 3, 4, 5; quand on parle de l'élément 4 de $\mathbf{Z}/6\mathbf{Z}$, cela signifie qu'on parle de « la classe de l'entier 4 pour la congruence modulo 6 », ou encore de « l'ensemble des nombres de la forme $6n + 4$ ». Avec ces conventions de langage, *l'application canonique de* $\mathbf{Z}$ *sur* $\mathbf{Z}/p\mathbf{Z}$ *consiste à associer à chaque entier rationnel le reste de sa division par* p.

Exemple 10. Dans le cas de l'*Exemple* 5 du n° 1, l'ensemble quotient se note

$$\mathbf{R}/2\pi\mathbf{Z},$$

et ses éléments s'appellent les **nombres réels modulo** 2π. Un nombre réel modulo 2π est donc un ensemble de nombres réels — à savoir tous ceux qui se déduisent d'un nombre réel donné par addition d'un multiple quelconque de 2π. Il est clair que la mesure d'un angle est précisément un nombre réel modulo 2π (et non pas un nombre réel usuel).

Exemple 11. Dans l'*Exemple* 7 du n° 1, les classes sont les droites parallèles à D, et l'ensemble quotient est l'ensemble des droites parallèles à D.

3. *Fonctions définies sur un ensemble quotient* (*)

Le résultat que voici est très utile :

THÉORÈME 2. *Soient* E *un ensemble,* R *une relation d'équivalence sur* E, p *l'application canonique de* E *sur* E/R, *et* f *une application de* E *dans un ensemble* X. *Les propriétés suivantes sont équivalentes:*

a) la relation $x \equiv y \pmod{R}$ *implique* $f(x) = f(y)$;

b) il existe une application

$$\bar{f}\colon \mathrm{E/R} \to \mathrm{X}$$

telle que $f = \bar{f} \circ p$.

Si ces conditions sont vérifiées, l'application $\bar{f}$ *est unique; pour qu'elle soit injective il faut et il suffit que les relations* $x \equiv y \pmod{R}$ *et* $f(x) = f(y)$ *soient équivalentes; pour qu'elle soit surjective il faut et il suffit que* f *le soit.*

D'après le Théorème 1 du § 2, une condition nécessaire et suffisante pour qu'il

(*) Les résultats de ce n° ne seront pas indispensables avant le § 29.

existe une application $\overline{f}$ satisfaisant à la condition de l'énoncé est que la relation $p(x) = p(y)$ implique la relation $f(x) \equiv f(y)$; or la relation $p(x) = p(y)$ équivaut à la relation $x \equiv y \pmod{R}$; les propriétés *a*) et *b*) sont donc bien équivalentes.

L'unicité de $\overline{f}$ provient de ce que p est surjective; supposons en effet trouvées deux applications g et h telles que $f = g \circ p = h \circ p$; pour tout $x \in E$ on a alors $g(p(x)) = h(p(x))$; donc g et h coïncident sur l'image $p(E)$, i.e. sur E/R, d'où $g = h$ comme annoncé.

On a $f(E) = \overline{f}(p(E)) = \overline{f}(E/R)$, de sorte que f est surjective si et seulement si f l'est. Pour que l'application f soit injective, il faut et il suffit, puisque tout élément de E/R est de la forme $p(z)$, que la relation

$$\overline{f}(p(x)) = \overline{f}(p(y)) \quad \text{implique} \quad p(x) = p(y),$$

autrement dit que $f(x) = f(y)$ implique $x \equiv y \pmod{R}$, ce qui achève la démonstration.

Le résultat suivant, analogue au Théorème précédent mais un peu plus compliqué, est fondamental quand on désire définir des opérations algébriques sur les éléments d'un ensemble quotient :

THÉORÈME 3. *Soient* X, Y, Z *trois ensembles*, R, S, T *des relations d'équivalence sur ces ensembles, et* f *une application de* $X \times Y$ *dans* Z. *Désignons par* $x \to \overline{x}$, $y \to \overline{y}$ *et* $z \to \overline{z}$ *les applications canoniques de* X, Y *et* Z *sur* X/R, Y/S *et* Z/T. *Les assertions suivantes sont équivalentes:*

a) les relations

$$x' \equiv x'' \pmod{R} \qquad \text{et} \qquad y' \equiv y'' \pmod{S}$$

impliquent la relation

$$f(x', y') \equiv f(x'', y'') \pmod{T};$$

b) il existe une application

$$\overline{f} : (X/R) \times (Y/S) \to Z/T$$

telle que l'on ait

$$\overline{f}(\overline{x}, \overline{y}) = \overline{f(x, y)}$$

quels que soient $x \in X$ *et* $y \in Y$.

Si ces conditions sont vérifiées, l'application $\overline{f}$ *est unique.*

Considérons l'application $u : X \times Y \to Z/T$ donnée par

$$u(x, y) = \overline{f(x, y)};$$

considérons d'autre part l'application $v : X \times Y \to (X/R) \times (Y/S)$ donnée par

$$v(x, y) = (\overline{x}, \overline{y});$$

tout revient à construire une application $\overline{f} : (X/R) \times (Y/S) \to Z/T$ telle que l'on

ait $u = \overline{f} \circ v$; pour cela, on applique le Théorème 1 du § 2 : $\overline{f}$ existe si et seulement si la relation

$$v(x', y') = v(x'', y'') \qquad \text{implique} \qquad u(x', y') = u(x'', y'');$$

or la première relation s'écrit $\overline{x}' = \overline{x}''$ et $\overline{y}' = \overline{y}''$, i.e.

$$x' \equiv x'' \pmod{R} \qquad \text{et} \qquad y' \equiv y'' \pmod{S},$$

et la seconde

$$f(x', y') \equiv f(x'', y'') \pmod{T};$$

le Théorème 1 du § 2 montre donc bien l'équivalence des conditions *a*) et *b*) de l'énoncé.

S'il existe une application $\overline{f}$ possédant la propriété cherchée, cette application est unique parce que v est évidemment surjective. Ceci termine la démonstration.

Exemple 12. Prenons $X = Y = Z = \mathbf{Z}$, ensemble des entiers rationnels, et pour R, S et T la relation de congruence modulo p, où p est un entier non nul donné. Considérons les applications f, $g : \mathbf{Z} \times \mathbf{Z} \to \mathbf{Z}$ données par

$$f(x, y) = x + y, \qquad g(x, y) = xy;$$

la condition *a*) du Théorème 3 est vérifiée par f et par g; pour le voir, on doit vérifier que les relations

$$x' \equiv x'' \pmod{p} \qquad \text{et} \qquad y' \equiv y'' \pmod{p}$$

impliquent

$$x' + y' \equiv x'' + y'' \pmod{p} \qquad \text{et} \qquad x'y' \equiv x''y'' \pmod{p},$$

ce qui est clair en vertu des identités

$$(x' + y') - (x'' + y'') = (x' - x'') + (y' - y''),$$
$$x'y' - x''y'' = (x' - x'')y' + x''(y' - y'').$$

On peut donc appliquer le Théorème 3 à f et à g, autrement dit il existe des applications

$$\overline{f}, \overline{g} : (\mathbf{Z}/p\mathbf{Z}) \times (\mathbf{Z}/p\mathbf{Z}) \to \mathbf{Z}/p\mathbf{Z}$$

qui sont caractérisées par les propriétés suivantes : étant donnés des éléments ξ et η de $\mathbf{Z}/p\mathbf{Z}$, représentés par des entiers rationnels x et y (de sorte qu'on a

$$\xi = \overline{x}, \qquad \eta = \overline{y}$$

avec les notations du Théorème 3), alors $\overline{f}(\xi, \eta)$ et $\overline{g}(\xi, \eta)$ sont représentés par $x + y$ et xy, dont les classes modulo p ne dépendent donc que des classes de x et y modulo p, et non du choix de x et y dans les classes ξ et η considérées.

Dans la pratique, on écrit

$$\overline{f}(\xi, \eta) = \xi + \eta, \qquad \overline{g}(\xi, \eta) = \xi\eta,$$

et on dit que $\xi + \eta$ et $\xi\eta$ sont la **somme** et le **produit** des éléments ξ et η de $\mathbf{Z}/p\mathbf{Z}$. Si l'on désigne par θ l'application canonique de $\mathbf{Z}$ sur $\mathbf{Z}/p\mathbf{Z}$, on voit donc que l'addition et la multiplication des entiers modulo p sont définies de telle sorte que l'on ait les relations

$$\theta(x + y) = \theta(x) + \theta(y), \qquad \theta(xy) = \theta(x)\theta(y)$$

quels que soient $x, y \in \mathbf{Z}$.

Pratiquement, l'addition et la multiplication dans $\mathbf{Z}/p\mathbf{Z}$ se calculent comme suit : on représente les entiers modulo p par les entiers ordinaires $0, 1, \ldots, p - 1$ (cf. *Exemple* 9); si des classes modulo p sont représentées par des entiers x et y (compris entre 0 et $p - 1$), la somme et le produit de ces classes seront alors représentées par les restes des divisions de $x + y$ et xy par p.

Voici par exemple les « tables d'addition et de multiplication » des entiers modulo 5 :

	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Une formule telle que $2 \cdot 3 = 1$ signifie bien entendu que

$$2 \cdot 3 \equiv 1 \pmod 5.$$

Le cas le plus simple (à part $p = 1$) est celui où $p = 2$; on a alors *deux* classes, qu'il s'impose d'appeler « pair » et « impair », et les règles de calcul sont alors données par les formules suivantes :

pair + pair = pair	pair × pair = pair
pair + impair = impair	pair × impair = impair
impair + pair = impair	impair × pair = pair
impair + impair = pair	impair × impair = impair.

§ 5. Ensembles finis et nombres entiers

Il est évidemment impossible de faire quoi que ce soit en Mathématiques sans utiliser la théorie des nombres entiers, et avant l'invention de la théorie des ensembles on la considérait comme le point de départ des Mathématiques : « Dieu nous a donné les nombres entiers, tout le reste est l'œuvre de l'homme », disait Kronecker.

On peut en fait aujourd'hui, à l'aide de la théorie des ensembles, construire les nombres entiers sans invoquer aucune divinité; l'idée de base, qu'on a de toute façon toujours enseignée aux enfants, est que les nombres entiers servent à « compter » les éléments des ensembles « finis », et que deux tels ensembles ont le même nombre d'éléments lorsqu'il existe une *bijection* du premier sur le second, et dans ce cas seulement (dans l'enseignement élémentaire on ne parle pas de « bijection » — on aligne les pommes au dessous des poires de façon à associer une pomme à chaque poire et vice-versa, ce qui est une méthode concrète simple pour construire une application bijective). Autrement dit, la notion d'entier est ce qui subsiste de la notion d'ensemble fini lorsque l'on considère comme identiques deux ensembles *équipotents* au sens du § 4, *Exemple* 3.

Lorsqu'il a commencé à édifier la théorie des ensembles, Cantor s'est aussitôt et principalement intéressé à cette question, mais sous un angle beaucoup plus général et difficile, à savoir celui qui consiste à « compter » les éléments d'un ensemble arbitraire, *fini ou non*, ce qui l'a conduit aux *nombres transfinis* qui permettent de distinguer les différentes espèces d'ensembles infinis. Ces « nombres » sont rarement utilisés dans la pratique élémentaire, mais il est difficile d'exposer raisonnablement la théorie des entiers ordinaires sans parler en même temps des autres.

La plupart des résultats énoncés dans ce § le seront sans démonstration, afin de réduire la longueur du texte (et parce que les démonstrations de certains théorèmes aux énoncés fort simples sont malheureusement difficiles, même dans le cadre « élémentaire » de la théorie des entiers ordinaires). Ce § n'est donc qu'un cadre dans lequel on pourrait, si on le désirait, insérer une étude complète et rigoureuse des nombres entiers.

Notons enfin que les entiers dont il s'agit ici sont des *objets mathématiques* au sens du § 0 (leur définition complète ferait intervenir les signes τ et $\square$ du n° 9 du § 0) qui servent de modèles abstraits aux nombres « concrets » avec lesquels on ne doit pas les confondre.

1. Ensembles équipotents

Rappelons (§ 4, *Exemple* 3) qu'un ensemble X est *équipotent* à un ensemble Y s'il existe une bijection de X sur Y. La relation

$$\text{X est équipotent à Y,}$$

qu'on note parfois

$$\text{Eq (X, Y),}$$

est une *relation d'équivalence.*

Cette relation possède un certain nombre de propriétés simples vis-à-vis des opérations de la théorie des ensembles.

Soient X, Y, X′, Y′ quatre ensembles, et supposons X équipotent à Y et X′ équipotent à Y′. Alors l'ensemble $X \times X'$ est équipotent à l'ensemble $Y \times Y'$ (car si f et f' sont des bijections de X sur Y et de X′ sur Y′, il est clair que $f \times f'$ est une bijection de $X \times X'$ sur $Y \times Y'$); de même l'ensemble

$$X^{X'}$$

des applications de X′ dans X (§ 2, *Remarque* 3) est équipotent à l'ensemble $Y^{Y'}$ des applications de Y′ dans Y; enfin, $X \cup X'$ est équipotent à $Y \cup Y'$ *pourvu que* X et X′ soient disjoints, ainsi que Y et Y′ : si f est une bijection de X sur Y, et f' une bijection de X′ sur Y′, on définit une bijection g de $X \cup X'$ sur $Y \cup Y'$ en posant

$$g(x) = \begin{cases} f(x) & \text{si} \quad x \in X \\ f'(x) & \text{si} \quad x \in X'. \end{cases}$$

A côté de ces propriétés d'énoncé simple et de démonstration immédiate, il existe d'autres résultats, d'énoncé également fort simple mais de démonstration incomparablement plus difficile; le plus frappant est le suivant, qui exprime intuitivement qu'étant donnés deux ensembles quelconques X et Y, on peut toujours « comparer » le « nombre » (sic) d'éléments de X au « nombre » d'éléments de Y :

THÉORÈME 1. *Soient* X *et* Y *deux ensembles. L'une au moins des deux assertions suivantes est vraie :*

$$\text{X est équipotent à une partie de Y}$$
$$\text{Y est équipotent à une partie de X.}$$

En outre, si ces deux assertions sont vraies simultanément, X *et* Y *sont équipotents.*

Bien que Cantor ait conjecturé cet énoncé dès ses premières recherches, la seconde partie de celui-ci n'a été démontrée qu'en 1897, par Bernstein, et la première, beaucoup plus difficile, qu'en 1904, par Zermelo. On trouvera le raisonnement de Bernstein dans l'*Exercice* 5 de ce §. Ajoutons que le Théorème 1 est non trivial (bien qu'intuitivement évident) même lorsque X et Y sont des ensembles finis; on conseille vivement au lecteur d'y réfléchir pour s'en convaincre (étant entendu qu'il s'agit de *démontrer* le résultat en question et non pas seulement de le rendre plausible).

2. *Le cardinal d'un ensemble*

Pour étendre au cas général la notion « de nombre » d'éléments d'un ensemble « fini », on a été amené avec Georg Cantor à attacher à chaque ensemble X un nouvel objet mathématique qu'on note

$$\text{Card}(X),$$

qu'on appelle le **cardinal** ou la **puissance** de X, et qui est défini de telle sorte que la condition suivante soit vérifiée : *pour que deux ensembles* X *et* Y *soient équipotents, il faut et il suffit que* $\text{Card}(X) = \text{Card}(Y)$.

¶¶ *Remarque* 1. S'il existait un *ensemble* Ω dont les éléments soient tous les ensembles, il suffirait de prendre pour Card (X) la classe de X pour la relation d'équivalence Eq(X, Y). Mais on sait (§ 1, *Remarque* 5) qu'un tel ensemble Ω n'existe pas, en sorte qu'on ne peut utiliser ici les constructions du § 4. En fait, le cardinal d'un ensemble X est défini par la relation

$$\text{Card}(X) = \tau_Y(\text{Eq}(X, Y))$$

qui utilise les considérations du § 0, n° 9.

On dit qu'un objet mathématique x est un **nombre cardinal** s'il existe un ensemble X tel que $x = \text{Card}(X)$.

Parmi les nombres cardinaux figurent les suivants, qu'on désigne par les symboles 0, 1, 2,... mais qui ne sont naturellement pas les nombres 0, 1, 2,... « usuels » ou « naïfs » (en ce sens que les entiers « usuels » sont des idées métaphysiques déduites de l'expérience concrète, tandis que les entiers « mathématiques » sont des objets théoriquement définissables à l'aide des procédés du § 0) :

(1) $$0 = \text{Card}(\emptyset),$$

cardinal de l'ensemble vide,

(2) $$1 = \text{Card}(\{\emptyset\}),$$

cardinal de l'ensemble $\{\emptyset\}$ réduit au seul élément ∅,

(3) $$2 = \text{Card}(\{\emptyset, \{\emptyset\}\}),$$

cardinal de l'ensemble dont les seuls éléments sont l'ensemble vide ∅ et l'ensemble $\{\emptyset\}$ dont le seul élément est l'ensemble vide ∅, etc... Étant donné un ensemble X, dire que $\text{Card}(X) = 0$ signifie que X est vide; dire que $\text{Card}(X) = 1$ signifie que X est un ensemble à un élément (i.e. que X est non vide et que les relations $x \in X$ et $y \in X$ impliquent $x = y$); dire que $\text{Card}(X) = 2$ signifie que X est un ensemble à deux éléments (i.e. qu'il existe $x \in X$ et $y \in X$ tels que l'on ait $x \neq y$ et tels que la relation $z \in X$ signifie $z = x$ ou $z = y$), etc... On reviendra plus loin (n° 4) sur ces nombres cardinaux particuliers.

Soient x et y deux nombres cardinaux; on écrit

$$x \leqslant y$$

lorsqu'il existe des ensembles X et Y tels que $x = \mathrm{Card}\,(X)$, $y = \mathrm{Card}\,(Y)$, et tels que X soit équipotent à une partie de Y — s'il en est ainsi pour un choix particulier de X et Y il en sera évidemment de même pour tout autre choix. Le Théorème 1 exprime alors que quels que soient x et y, on a toujours

$$x \leqslant y \quad \text{ou} \quad y \leqslant x; \tag{4}$$

et qu'en outre

$$x \leqslant y \quad \text{et} \quad y \leqslant x \quad \text{implique} \quad x = y. \tag{5}$$

Il est clair d'autre part que si x, y, z sont trois cardinaux, alors

$$x \leqslant y \quad \text{et} \quad y \leqslant z \quad \text{implique} \quad x \leqslant z; \tag{6}$$

car s'il existe une injection f d'un ensemble X dans un ensemble Y, et une injection g de Y dans un ensemble Z, il existe une injection de X dans Z, à savoir $g \circ f$.

La principale propriété de la relation $x \leqslant y$ entre cardinaux est exprimée par l'énoncé suivant (que nous admettrons) :

THÉORÈME 2. *Soit* E *un ensemble de cardinaux. Il existe un et un seul cardinal a possédant les propriétés suivantes:*

(i) *on a $x \leqslant a$ (resp. $x \geqslant a$) pour tout $x \in \mathrm{E}$;*

(ii) *si un cardinal b est tel que l'on ait $x \leqslant b$ (resp. $x \geqslant b$) pour tout $x \in \mathrm{E}$, on a $b \geqslant a$ (resp. $b \leqslant a$).*

Ce Théorème exprime que lorsqu'on a un *ensemble* E de cardinaux, il existe des cardinaux supérieurs (resp. inférieurs) à tous les éléments de E, et de plus que, parmi tous les cardinaux qui sont supérieurs (resp. inférieurs) à tous les éléments de E il en existe un qui est inférieur (resp. supérieur) à tous les autres. C'est donc le *plus petit* (resp. plus grand) cardinal a vérifiant $a \geqslant x$ (resp. $a \leqslant x$) pour tout $x \in \mathrm{E}$. On l'appelle la **borne supérieure** (resp. la **borne inférieure**) de l'ensemble E, et on le désigne généralement par la notation

$$\sup\,(\mathrm{E}) \qquad (\text{resp. } \inf\,(\mathrm{E}))$$

ou une notation analogue.

Remarque 2. Le fait que dans l'énoncé précédant la lettre E désigne un *ensemble* de cardinaux est d'autant plus essentiel qu'il n'existe pas d'ensemble contenant tous les cardinaux (pas plus qu'il n'existe d'ensemble contenant tous les ensembles). Du reste, s'il existait un tel ensemble, le Théorème 2 appliqué à cet ensemble montrerait qu'il existe un cardinal a supérieur à *tous* les cardinaux — or c'est impossible, car on peut démontrer que pour tout cardinal a on a l'inégalité stricte

$$a < 2^a.$$

On voit donc, ici encore, que si l'on attribue au mot « ensemble » sa signification intuitive, on s'expose à des contradictions logiques.

Notons d'autre part que si E est un ensemble de cardinaux, les cardinaux sup (E) et inf (E) n'appartiennent pas nécessairement à E. Si par exemple on prend pour E l'ensemble **N** de tous les cardinaux finis (voir plus loin), sup (E) n'est autre que la puissance du dénombrable (n° 5) et par suite n'est pas dans E.

Cependant, lorsque E est un ensemble de cardinaux *finis* i.e. d'entiers naturels (cette notion sera définie au n° 4), on a toujours inf (E) ∈ E. En effet comme inf (E) est inférieur à tout $x \in E$ il est clair que inf (E) est fini; si inf (E) n'appartenait pas à E, on aurait inf (E) $< x$ et donc

$$\inf(E) + 1 \leqslant x$$

pour tout $x \in E$, de sorte que d'après la propriété (ii) du Théorème 2 on aurait la relation

$$\inf(E) + 1 \leqslant \inf(E),$$

ce qui est impossible puisque inf (E) est fini.

En d'autres termes, *si* E *est un ensemble d'entiers naturels il existe un élément de* E *qui est plus petit que tous les autres*, résultat qu'on utilise constamment dans la pratique.

3. *Opérations sur les cardinaux*

Soient x et y deux nombres cardinaux, et posons $x = \mathrm{Card}(X)$, $y = \mathrm{Card}(Y)$; on appelle **produit** de x par y le nombre cardinal

$$xy = \mathrm{Card}(X \times Y); \tag{7}$$

il est clair qu'il ne change pas si l'on remplace X (resp. Y) par un ensemble équipotent à X (resp. Y).

Cette opération vérifie les identités que voici :

$$xy = yx; \qquad x(yz) = (xy)z; \qquad 0x = 0; \qquad 1x = x. \tag{8}$$

Pour démontrer la seconde par exemple, il suffit d'observer que, si X, Y et Z sont trois ensembles, alors $X \times (Y \times Z)$ est équipotent à $(X \times Y) \times Z$, ce qui est clair si l'on associe à chaque élément $(a, (b, c))$ du premier l'élément $((a, b), c)$ du second.

¶ *Remarque* 3. Malgré ce qu'indique l'intuition, *il est faux* que

$$xz = yz \quad \text{implique} \quad x = y$$

même si $z \neq 0$. Pour obtenir un contre-exemple, admettons (voir plus loin) l'existence d'un ensemble **N** dont les éléments sont les entiers 0, 1, 2,...; prenons $z = \mathrm{Card}(\mathbf{N})$, $x = 1$, $y = 2$, en sorte que x est le cardinal d'un ensemble X réduit à un élément a, et y le cardinal d'un ensemble Y réduit à deux éléments

b et c. Tout revient à construire une *bijection* f de $X \times \mathbf{N}$ sur $Y \times \mathbf{N}$; pour cela, on pose

$$f(a, n) = \begin{cases} (b, p) & \text{si } n = 2p \text{ est pair} \\ (c, p) & \text{si } n = 2p + 1 \text{ est impair.} \end{cases}$$

Bien entendu, le fait que

$$xz = yz \quad \text{implique} \quad x = y \quad \text{si} \quad z \neq 0$$

est cependant vrai si x, y, z sont des entiers « naturels » (voir plus loin).

Soient maintenant x et y deux cardinaux, et choisissons deux ensembles *disjoints* X et Y tels que $x = \text{Card}(X)$, $y = \text{Card}(Y)$; on appelle **somme** de x et y le cardinal

$$x + y = \text{Card}(X \cup Y) \qquad (\text{pour } X \cap Y = \emptyset); \tag{9}$$

on vérifie aussitôt qu'il ne dépend pas du choix de X et Y. On a les identités suivantes :

$$x + y = y + x; \qquad x + (y + z) = (x + y) + z; \qquad 0 + x = x. \tag{10}$$

La dernière par exemple exprime que, pour tout ensemble X, l'ensemble X est équipotent à $X \cup \emptyset$, ce qui est d'autant moins surprenant qu'on a même

$$X = X \cup \emptyset \ldots$$

Remarque 4. La définition de $x + y$ donnée ci-dessus suppose établi qu'on peut toujours trouver des ensembles X et Y *disjoints* tels que $x = \text{Card}(X)$, $y = \text{Card}(Y)$. Pour cela, posons $x = \text{Card}(X')$, $y = \text{Card}(Y')$, et prenons

$$X = X' \times \{a\}, \qquad Y = Y' \times \{b\}$$

avec $a \neq b$; alors X et Y sont équipotents à X' et Y', et disjoints car la relation $(x, a) = (y, b)$ exige $a = b$.

¶ *Remarque* 5. Ici encore, *il est faux* que la relation

$$x + z = y + z \qquad \text{implique} \qquad x = y;$$

par exemple, on peut fort bien avoir à la fois

$$x + x = x \qquad \text{et} \qquad x \neq 0;$$

pour cela, prenons $x = \text{Card}(\mathbf{N})$ où $\mathbf{N}$ est l'ensemble (voir plus loin) des entiers 0, 1, 2,...; alors $x + x = \text{Card}(Y)$ où Y est l'ensemble des couples (n, u) avec $n \in \mathbf{N}$ et $u = 0$ ou $u = 1$; on obtient une bijection f de Y sur $\mathbf{N}$ en posant

$$f(n, u) = \begin{cases} 2n & \text{si } u = 0 \\ 2n + 1 & \text{si } u = 1, \end{cases}$$

ce qui montre qu'on a bien $x + x = x$ dans ce cas...

On a d'autre part entre l'addition et la multiplication la relation de « distributivité »

$$x(y+z) = xy + xz; \tag{11}$$

elle signifie qu'étant donnés trois ensembles X, Y, Z, alors $X \times (Y \cup Z)$ est équipotent à $(X \times Y) \cup (X \times Z)$, ce qui est évident « géométriquement ».

Considérons enfin deux cardinaux x et y, et posons $x = \mathrm{Card}(X)$, $y = \mathrm{Card}(Y)$; on pose alors

$$x^y = \mathrm{Card}(X^Y), \tag{12}$$

cardinal de l'ensemble de toutes les applications de Y dans X. Cette opération, appelée **exponentiation** des cardinaux, satisfait aux identités que voici :

$$x^{y+z} = x^y . x^z; \qquad (xy)^z = x^z y^z; \qquad (x^y)^z = x^{yz}; \qquad x^0 = 1; \qquad x^1 = x. \tag{13}$$

Pour établir par exemple la première, posons $x = \mathrm{Card}(X)$, $y = \mathrm{Card}(Y)$, $z = \mathrm{Card}(Z)$ et supposons Y et Z dispoints, de sorte que $y + z = \mathrm{Card}(Y \cup Z)$. Alors tout revient à montrer que les ensembles

$$X^{Y \cup Z} \qquad \text{et} \qquad X^Y \times X^Z$$

sont équipotents, i.e. à construire une bijection f du premier sur le second; pour cela soit u un élément du premier ensemble, i.e. une application de $Y \cup Z$ dans X; soient u_Y et u_Z ses restrictions à Y et Z; on pose alors $f(u) = (u_Y, u_Z)$, et le lecteur vérifiera facilement (en utilisant l'hypothèse que Y et Z sont disjoints) que f est bijective.

Remarque 6. Considérons un ensemble A à deux éléments, par exemple $A = \{0, 1\}$, et soit f une application d'un ensemble X dans A; pour déterminer f, il suffit évidemment de connaître l'ensemble

$$f^{-1}(0) \subset X$$

des x où $f(x) = 0$; on vérifie aussitôt que l'application $f \to f^{-1}(0)$ de A^X dans $\mathfrak{P}(X)$ ainsi définie est *bijective*. Il s'ensuit que

$$\mathrm{Card}(\mathfrak{P}(X)) = 2^{\mathrm{Card}(X)}.$$

Remarque 7. On peut montrer que, pour tout cardinal x, on a

$$x < 2^x$$

(i.e. $x \leq 2^x$ et $x \neq 2^x$). En particulier, si X est un ensemble, $\mathfrak{P}(X)$ *n'est pas* équipotent à X. Voir *Exercice* 1.

¶ *Remarque* 8. On peut définir non seulement la somme ou le produit de *deux* nombres cardinaux, mais plus généralement la somme ou le produit d'une

famille quelconque (même « infinie ») de tels nombres. On procède comme suit.

Tout d'abord soit $(X_i)_{i\in I}$ une famille d'ensembles; on appelle **produit cartésien** des X_i l'ensemble, noté

$$\prod_{i\in I} X_i,$$

dont les éléments sont toutes les familles $(x_i)_{i\in I}$ pour lesquelles on a $x_i \in X_i$ pour tout $i \in I$; cette notion généralise évidemment celle du § 2, n° 2, car si $I = \{1, 2\}$ et si l'on identifie une famille $(x_i)_{i\in I}$ au couple (x_1, x_2) on voit qu'on peut identifier le produit des X_i, $i \in I$, à l'ensemble $X_1 \times X_2$.

Cela dit, si $(x_i)_{i\in I}$ est une famille quelconque de cardinaux, et si l'on choisit pour chaque $i \in I$ un ensemble X_i tel que $x_i = \mathrm{Card}\,(X_i)$, on pose par définition

$$\prod_{i\in I} x_i = \mathrm{Card}\left(\prod_{i\in I} X_i\right).$$

Le cardinal ainsi obtenu s'appelle le **produit de la famille de nombres cardinaux** $(x_i)_{i\in I}$.

On définit de même la **somme de la famille de nombres cardinaux** $(x_i)_{i\in I}$ comme étant le cardinal noté

$$\sum_{i\in I} x_i$$

et défini par la relation

$$\sum_{i\in} x_i = \mathrm{Card}\left(\bigcup_{i\in I} X_i\right)$$

à condition que les ensembles X_i choisis soient *deux à deux disjoints* bien entendu.

Supposons par exemple $x_i = 1$ pour tout $i \in I$; on peut alors prendre $X_i = \{i\}$ pour tout $i \in I$, et la réunion de ces ensembles est I; on a donc

$$(14) \qquad \mathrm{Card}\,(I) = \sum_{i\in I} x_i \qquad \text{avec} \qquad x_i = 1 \quad \text{pour tout} \quad i \in I;$$

intuitivement, cela signifie que tout nombre cardinal est une somme (généralement « infinie » ...) de termes tous égaux à 1, propriété bien connue des entiers usuels.

4. *Ensembles finis et entiers naturels*

Le résultat suivant est fondamental :

THÉORÈME 3. *Soit* X *un ensemble. Les propriétés suivantes sont équivalentes:*

a) *Le seul ensemble contenu dans* X *et équipotent à* X *est* X *lui-même.*

b) *On a* $\mathrm{Card}\,(X) \neq \mathrm{Card}\,(X) + 1$.

Supposons Card (X) = Card (X) + 1; en désignant par a un objet n'appartenant pas à X, il existe donc une bijection f de l'ensemble $X \cup \{a\}$ sur l'ensemble X; l'image de X par f est évidemment équipotente à X et strictement contenue dans X.

Supposons inversement X équipotent à une ensemble X′ strictement contenu dans X. On a alors, puisque $X = X' \cup (X - X')$,

$$\text{Card}(X) = \text{Card}(X) + \text{Card}(X - X');$$

mais $\text{Card}(X - X') \geqslant 1$ puisque $X - X'$ est non vide; il vient donc

$$\text{Card}(X) \geqslant \text{Card}(X) + 1 \geqslant \text{Card}(X)$$

et par suite Card (X) = Card (X) + 1 d'après le Théorème 1, ce qui achève la démonstration du Théorème 3.

On dit qu'un ensemble X est **fini** s'il possède les propriétés *a*) et *b*) de l'énoncé ci-dessus, et **infini** dans le cas contraire. De même, un cardinal x est **fini** si $x \neq x + 1$, et **infini** si $x = x + 1$. Un cardinal fini s'appelle aussi un **entier naturel**, et un cardinal infini un **nombre transfini.**

Les entiers naturels possèdent des propriétés fort simples (et que tout le monde connaît...). Tout d'abord, *si x et y sont des entiers naturels, il en est de même des cardinaux $x + y$, xy et x^y*; plus généralement, si $(x_i)_{i \in I}$ est une famille *finie* d'entiers naturels (on dit qu'une famille $(x_i)_{i \in I}$ est **finie** lorsque l'ensemble d'indices I est fini), alors les cardinaux (*Remarque* 8)

$$\prod_{i \in I} x_i \qquad \text{et} \qquad \sum_{i \in I} x_i$$

sont encore finis.

Si x est un entier naturel, tout cardinal y tel que $y \leqslant x$ est encore fini (autrement dit, *toute partie d'un ensemble fini est un ensemble fini*); et il existe alors un cardinal z et un seul tel que

$$x = y + z;$$

z est fini, s'appelle la **différence** entre x et y, et se note

$$z = x - y;$$

si $x = \text{Card}(X)$ et $y = \text{Card}(Y)$ avec $Y \subset X$, on a évidemment $z = \text{Card}(X - Y)$.

Enfin, quand il s'agit de cardinaux finis, les circonstances pathologiques examinées dans les *Remarques* 3 et 5 ne peuvent pas se produire; de façon précise, la relation

$$x + z = y + z \quad \text{implique} \quad x = y \quad \text{si } z \text{ est fini,}$$

et de même la relation

$$xz = yz \quad \text{implique} \quad x = y \quad \text{si } z \text{ est fini et non nul.}$$

Il va de soi que les cardinaux 0, 1, 2, ... définis plus haut sont finis.

On a fréquemment besoin de la propriété suivante :

THÉORÈME 4. *Soient* X *un ensemble et* f *une application de* X *dans* X. *Les propriétés suivantes sont équivalentes :*

a) f *est injective.*

b) f *est surjective.*

c) f *est bijective.*

Il suffit évidemment de montrer l'équivalence de *a*) et *b*). Si f est injective, f est une bijection de X sur une partie de X, à savoir $f(X)$, qui est donc équipotente à X; si X est fini on a donc $f(X) = X$, en sorte que *a*) implique *b*).

Supposons f surjective; il existe alors (§ 2, Théorème 4) une application h de X dans X telle que $f \circ h = j_X$, application identique; évidemment h est injective, donc surjective puisqu'on a déjà établi que *a*) implique *b*), donc bijective, et par suite f est l'application réciproque de h, et est donc injective, d'où le Théorème.

5. L'ensemble N des entiers naturels

Les considérations précédentes rendent évidente l'existence d'ensembles finis, puisque les ensembles

$$\emptyset, \quad \{\emptyset\}, \quad \{\emptyset, \{\emptyset\}\}, \quad \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}, \ldots$$

sont finis — c'est du reste à ce moment précis qu'on voit les Mathématiques acquérir de la substance, en dépit du fait que tout repose sur la notion d'ensemble vide...

Par contre, l'*existence d'ensembles infinis* n'est pas claire, et l'est même si peu qu'en fait c'est l'un des *axiomes* des Mathématiques. Le lecteur qui trouverait ce point de vue étrange et contraire à l'intuition fera bien de se rappeler du fait qu'en Mathématiques on se propose de *démontrer logiquement* des assertions, et qu'en particulier le mot « existence » n'y a pas du tout le même sens qu'en Physique ou en Théologie.

THÉORÈME 5. *Soit* X *un ensemble infini. Tout ensemble fini est équipotent à une partie de* X.

Tout revient à montrer qu'on a $y \leqslant x$ pour tout cardinal fini y et tout cardinal infini x. Mais s'il n'en était pas ainsi on aurait $y \geqslant x$, et x serait fini comme on l'a vu au n° précédent.

THÉORÈME 6. *Il existe un ensemble* N *et un seul tel que la relation*

$$x \in \mathbf{N}$$

soit équivalente à la relation

x est un entier naturel.

L'ensemble N *est infini.*

L'unicité de N est évidente d'après le Théorème 2 du § 1. L'existence de N provient du fait que, si l'on choisit un cardinal infini a (ce qui est possible en vertu de l'existence d'ensembles infinis), alors tout entier naturel x vérifie la relation $x < a$ comme on vient de le voir; il suffit donc de montrer que, pour tout cardinal a,

les cardinaux x tels que $x < a$ sont les éléments d'un *ensemble*, ce que nous admettrons (*).

Supposons enfin $\mathbf{N}$ fini, et pour chaque $n \in \mathbf{N}$ choisissons un ensemble X_n tel que Card $(X_n) = n$; puisque $\mathbf{N}$ et chaque X_n sont finis, il en est de même de

$$X = \bigcup_{n \in \mathbf{N}} X_n,$$

et comme chaque X_n est contenu dans X on voit donc que, si $\mathbf{N}$ était fini, il existerait un entier naturel $x = \text{Card}(X)$ tel que l'on ait $n \leqslant x$ pour tout n fini; mais alors, puisque x est fini, il en est de même de $x + 1$, on a donc $x + 1 \leqslant x$, mais aussi $x \leqslant x + 1$, donc $x = x + 1$, ce qui contredit le fait que x est fini. On aboutit donc à une contradiction en supposant $\mathbf{N}$ fini, ce qui achève la démonstration (sic).

On voit en résumé que les deux assertions suivantes :

il existe un ensemble infini

. il existe un ensemble dont les éléments sont les entiers naturels

sont *équivalentes*.

Le cardinal

$$\text{Card}(\mathbf{N})$$

s'appelle la **puissance du dénombrable** et on dit qu'un ensemble X est **dénombrable** s'il est équipotent à $\mathbf{N}$, autrement dit s'il existe une *bijection*

$$n \mapsto x_n$$

de l'ensemble des entiers naturels sur l'ensemble des éléments de X.

Exemple 1. L'ensemble des nombres fractionnaires (on suppose cette notion déjà acquise) est dénombrable. Il suffit pour le voir de montrer qu'on peut écrire les nombres fractionnaires (qui *à priori* dépendent de *deux* entiers) sous forme d'une suite illimitée contenant chacun de ces nombres une fois et une seule. On procède comme suit :

$$\frac{0}{1}\,;\quad \frac{1}{1}\,;\quad \frac{1}{2},\ \frac{2}{1}\,;\quad \frac{1}{3},\ \frac{3}{1}\,;\quad \frac{1}{4},\ \frac{2}{3},\ \frac{3}{2},\ \frac{4}{1}\,;\quad \frac{1}{5},\ \frac{5}{1}\,;\quad \ldots$$

on écrit d'abord les nombres p/q tels que $p + q = 1$, puis ceux pour lesquels $p + q = 2$ et qui n'ont pas déjà été écrits, puis ceux pour lesquels $p + q = 3$ et qui n'ont pas déjà été écrits, etc...

L'idée qu'il n'existe « pas plus » de nombres fractionnaires que de nombres entiers est à première vue contraire au bon sens; mais c'est précisément l'une des plus grandes réussites de Cantor que d'avoir pu disqualifier l'emploi du « bon sens » en Mathématiques.

Remarque 8. Le cardinal

$$2^{\text{Card}(\mathbf{N})}$$

(*) Posant $a = \text{Card}(A)$, les $x < a$ sont en « correspondance biunivoque » avec les classes d'équivalence dans $\mathcal{E}(A)$ pour la relation Eq (X, Y) du n° 1 ; c'est ce qui explique (sic) pourquoi ces x restent dans un *ensemble*.

de l'ensemble $\mathfrak{P}(\mathbf{N})$ est strictement plus grand que Card(**N**) d'après la *Remarque* 7; on l'appelle la **puissance du continu**. On peut démontrer que

$$2^{\mathrm{Card}(\mathbf{N})} = \mathrm{Card}(\mathbf{R})$$

où **R** est l'ensemble des nombres réels (ou des points d'une droite).

Depuis le début du siècle, on a cherché en vain à établir l'*hypothèse du continu*, à savoir que toute partie infinie de **R** est équipotente à **N** ou à **R**. On a pu démontrer en 1939 (K. Gödel) que l'hypothèse du continu, comme l'axiome du choix (voir p. 64), était compatible avec les autres axiomes de la théorie des ensembles. Plus récemment (P. Cohen, 1963), on a en outre établi que l'hypothèse du continu et l'axiome du choix sont des assertions indépendantes l'une de l'autre (autrement dit : les axiomes de la théorie des ensembles, axiome du choix inclus, n'impliquent pas l'hypothèse du continu qui, de son côté, n'implique pas non plus l'axiome du choix). On trouvera un excellent exposé de ces questions dans P. Cohen, *Set Theory and the Continuum Hypothesis* (Benjamin, New York, 1966). Noter qu'en dépit de ces résultats impressionnants la question de la non-contradiction de la théorie des ensembles reste toujours ouverte.

6. *Le raisonnement par récurrence*

Le raisonnement par récurrence est fondé sur l'assertion que voici :

THÉORÈME 6. *Soit* $\mathrm{R}\{n\}$ *une relation contenant une variable* $n \in \mathbf{N}$. *Supposons que la relation* $\mathrm{R}\{0\}$ *soit vraie, et que la relation*

$$\mathrm{R}\{n\} \quad \text{implique} \quad \mathrm{R}\{n+1\}$$

soit vraie pour tout $n \in \mathbf{N}$. *Alors la relation* $\mathrm{R}\{n\}$ *est vraie pour tout* $n \in \mathbf{N}$.

Soit en effet E l'ensemble des $n \in \mathbf{N}$ tels que $\mathrm{R}\{n\}$ soit vraie; on doit montrer que $\mathrm{E} = \mathbf{N}$ ou, ce qui revient au même, que $\mathrm{F} = \mathbf{N} - \mathrm{E}$ est vide. Mais supposons F non vide; alors (*Remarque* 2) il existe un $a \in \mathrm{F}$ tel que l'on ait $a \leqslant n$ pour tout $n \in \mathrm{F}$; comme $\mathrm{R}\{0\}$ est vraie par hypothèse, on a $0 \in \mathrm{E}$ et par conséquent $a \geqslant 1$; donc $a = n + 1$ pour un $n \in \mathbf{N}$, et comme $n < a$ on ne peut avoir $n \in \mathrm{F}$, de sorte que $\mathrm{R}\{n\}$ est vraie; mais comme $\mathrm{R}\{n\}$ implique $\mathrm{R}\{n+1\}$ par hypothèse, il s'ensuit que $\mathrm{R}\{n+1\}$, i.e. $\mathrm{R}\{a\}$, est vraie, ce qui contredit le fait que $a \in \mathrm{F}$. D'où le Théorème.

Dans la pratique on utilise fréquemment des variantes du Théorème 6, et en particulier celle qui consiste à remplacer l'hypothèse que $\mathrm{R}\{n\}$ implique $\mathrm{R}\{n+1\}$ par la suivante :

la conjonction des relations $\mathrm{R}\{1\}, \ldots, \mathrm{R}\{n\}$ implique $\mathrm{R}\{n+1\}$

(il peut en effet arriver que, pour établir $\mathrm{R}\{n+1\}$, on ait à se servir non seulement de $\mathrm{R}\{n\}$ mais de *toutes* les assertions précédant $\mathrm{R}\{n+1\}$).

Exemple 2. Pour tout entier n, notons E_n l'ensemble des entiers x tels que $x \leqslant n$; nous allons démontrer, en raisonnant par récurrence sur n, que l'on a

$$\mathrm{Card}(\mathrm{E}_n) = n + 1$$

pour tout n. En effet, pour $n = 0$ on a $\mathrm{E}_n = \{0\}$ et par suite $\mathrm{Card}(\mathrm{E}_0) = 1$.

Il reste donc à établir que l'assertion relative à l'entier n implique l'assertion relative à l'entier $n+1$. Or soit $x \leqslant n+1$; on a soit $x \leqslant n$, et alors $x \in E_n$, soit

$$n < x \leqslant n+1,$$

et alors $x = n+1$ (voir ci-dessous). Donc $E_{n+1} = E_n \cup \{n+1\}$, en sorte que

$$\text{Card}\,(E_{n+1}) = \text{Card}\,(E_n) + 1;$$

cela montre évidemment que la relation $\text{Card}\,(E_n) = n+1$ implique, comme annoncé, la relation $\text{Card}\,(E_{n+1}) = n+2$. On a donc bien $\text{Card}\,(E_n) = n+1$ pour tout n.

On a fait usage ci-dessus du fait que la relation

$$n < x \leqslant n+1 \quad \text{implique} \quad x = n+1.$$

On l'établit par exemple comme suit. Soit A un ensemble tel que

$$\text{Card}\,(A) = n+1;$$

comme $x \leqslant n+1$, il existe un ensemble X tel que $x = \text{Card}\,(X)$ et $X \subset A$; comme $n < x$, il existe un ensemble B tel que

$$n = \text{Card}\,(B), \qquad B \subset X, \qquad B \neq X.$$

On a

$$n+1 = \text{Card}\,(A) = \text{Card}\,(B) + \text{Card}\,(A - B) = n + \text{Card}\,(A - B)$$

et donc $\text{Card}\,(A - B) = 1$, en sorte que le complément de B dans A se réduit à un élément. Comme X contient B et est distinct de B, on a donc $X = A$, et par suite $x = n+1$ comme annoncé.

7. *Analyse combinatoire*

Dans ce qui suit, étant donné un entier naturel n, on appelle **ensemble à n éléments** tout ensemble X tel que $\text{Card}\,(X) = n$.

THÉORÈME 7 (Principe des bergers). *Soit f une application d'un ensemble X dans un ensemble Y. On suppose que Y est un ensemble à q éléments, que pour tout $y \in Y$ l'ensemble $f^{-1}(y) \subset X$ est à p éléments, et que f est surjective. Alors X est un ensemble à pq éléments.*

Désignons par F un ensemble à p éléments choisi une fois pour toutes, et, pour chaque $y \in Y$, choisissons une bijection u_y de l'ensemble F sur l'ensemble $f^{-1}(y)$. Définissons une application

$$u : Y \times F \to X$$

en posant

$$u(y, z) = u_y(z) \quad \text{pour} \quad y \in Y \quad \text{et} \quad z \in F;$$

on vérifie aussitôt (en tenant compte du fait que f est surjective) que u est bijective. Donc

$$\text{Card}\,(X) = \text{Card}\,(Y \times F) = \text{Card}\,(Y).\text{Card}\,(F) = qp,$$

ce qui achève la démonstration.

THÉORÈME 8. *Soient* X *un ensemble à p éléments et* Y *un ensemble à q éléments. Alors l'ensemble des applications de* Y *dans* X *est à* p^q *éléments.*

Ce Théorème est en fait (n° 3) la définition de p^q. Bien entendu, pour lui donner tout son intérêt, il faut vérifier que, lorsqu'il s'agit d'entiers naturels, l'exponentiation des cardinaux se réduit à l'opération que tout le monde connait. Or les formules (13) du n° 3 montrent entre autres que

$$n^1 = n, \qquad n^0 = 1, \qquad n^{p+q} = n^p n^q;$$

on a donc

$$n^2 = n^{1+1} = n^1 n^1 = n.n; \qquad n^3 = n^{2+1} = n^2 n^1 = n.n.n, \quad \text{etc...}$$

Remarque 9. Prenons pour Y l'ensemble des entiers i tels que $1 \leqslant i \leqslant q$; une application de Y dans X est encore une famille $(x_i)_{i \in Y}$ d'éléments de X; une telle famille s'écrit souvent sous la forme

$$(x_i)_{1 \leqslant i \leqslant q},$$

et s'appelait autrefois un **arrangement des éléments de** X **pris** q **à** q. Le Théorème 8 affirme donc que, si X est à p éléments, le nombre de ces arrangements est p^q.

THÉORÈME 9. *Soient* X *un ensemble à p éléments et* Y *un ensemble à q éléments. Supposons* $p \leqslant q$. *Alors le nombre des injections de* X *dans* Y *est*

$$\frac{q!}{(q-p)!}$$

où, pour tout entier naturel n, on pose

$$n! = 1.2.3\ldots n \quad \text{si } n \neq 0, \qquad \text{et} \qquad n! = 1 \quad \text{si } n = 0.$$

Si $p = 0$, l'ensemble X est vide, et il y a une seule injection de X dans Y, de sorte que le Théorème est vrai dans ce cas. Il reste donc (Théorème 6) à montrer que s'il est vrai pour un entier p il est aussi vrai pour $p + 1$.

Supposons donc $\mathrm{Card}(X) = p + 1$ et $\mathrm{Card}(Y) = q \geqslant p + 1$. Choisissons une fois pour toutes un $a \in X$ et posons $X' = X - \{a\}$, de sorte que X' possède p éléments. Soit I l'ensemble des injections de X dans Y. On peut définir une application

$$u : I \to Y$$

en posant

$$u(f) = f(a) \quad \text{pour toute} \quad f \in I,$$

et il est clair que cette application de I dans Y est surjective (autrement dit, pour tout $b \in Y$ il existe une application *injective* f de X dans Y telle que $f(a) = b$). Pour un $b \in Y$ donné, considérons les $f \in I$ telles que $f(a) = b$; posant $Y' = Y - \{b\}$, une telle f induit évidemment une injection f' de X' dans Y', et réciproquement toute injection f' de X' dans Y' peut être complétée en une injection f de X dans Y telle que $f(a) = b$. On voit donc, en utilisant l'hypothèse de récurrence, que le nombre de $f \in I$ telles que $u(f)$ soit donné est

$$\frac{(q-1)!}{[(q-1)-(p-1)]!} = \frac{(q-1)!}{(q-p)!};$$

comme u applique I *sur* l'ensemble Y à q éléments, le principe des bergers montre donc que

$$\text{Card (I)} = q.\frac{(q-1)\,!}{(q-p)\,!},$$

et comme

$$q\,! = q.(q-1)\,!$$

la démonstration est achevée.

COROLLAIRE. *Le nombre des permutations d'un ensemble* X *à n éléments est n* !

Une permutation est une bijection de X dans X (§ 2, n° 8), mais comme X est fini les permutations de X sont aussi les injections de X dans X (Théorème 4). Il reste alors à appliquer le Théorème en prenant Y = X, et $p = q = n$, et à observer que

$$(n-n)\,! = 0\,! = 1.$$

Le nombre n ! se lit **factorielle** n. On a les relations

$$0\,! = 1, \quad 1\,! = 1, \quad 2\,! = 2, \quad 3\,! = 6, \quad 4\,! = 24, \quad 5\,! = 120, \ \ldots$$

THÉORÈME 10. *Soient* X *un ensemble à n éléments et p un entier inférieur ou égal à n. Le nombre d'ensembles à p éléments contenus dans* X *est*

$$\frac{n(n-1)\ldots(n-p+1)}{p\,!} = \frac{n\,!}{p\,!\,(n-p)\,!}.$$

Soit E un ensemble à p éléments choisi une fois pour toutes. Désignons par I l'ensemble des injections de E dans X, et par P l'ensemble des parties à p éléments de X.

Pour toute $f \in$ I, il est clair que f(E) est une partie à p éléments de X. On peut donc définir une application

$$u : \text{I} \to \text{P}$$

en posant

$$u(f) = f(\text{E}).$$

Cette application est *surjective*, car toute partie Y à p éléments de X est équipotente à E, donc est l'image de E par une application injective de E dans X.

Nous allons maintenant compter les $f \in$ I telles que f(E) = Y soit une partie donnée de X. Si f_0 est une telle application, on obtient les autres en composant f_0 avec une permutation arbitraire de l'ensemble E : si f(E) = f_0(E), alors pour tout $x \in$ E il existe un et un seul $s(x) \in$ E tel que $f(x) = f_0(s(x))$, et s est évidemment une permutation de E. Ainsi, en associant à chaque permutation s de E l'application $f = f_0 \circ s$ de E dans X, on obtient une *bijection* de l'ensemble des permutations de E sur l'ensemble des $f \in$ I telles que f(E) = Y.

Les $f \in$ I telles que f(E) = Y soit donné sont donc au nombre de p ! (Corollaire du Théorème 9). D'après le principe des bergers, on a donc

$$\text{Card (I)} = p\,!\ \text{Card (P)};$$

donc

$$\text{Card (P)} = \frac{\text{Card (I)}}{p!} = \frac{n!}{p!\,(n-p)\,!}$$

d'après le Théorème 9, et ceci achève la démonstration.

On pose habituellement

$$\frac{n!}{p!\,(n-p)!} = \binom{n}{p}$$

et on appelle ces entiers les **coefficients du binôme** pour des raisons qui apparaîtront plus loin (§ 8, n° 4).

Remarque 10. Une partie à p éléments d'un ensemble X est aussi ce qu'on appelait autrefois une **combinaison des éléments de X pris p à p**. Les ouvrages traditionnels définissent une telle combinaison comme étant une « suite »

$$x_1, \cdots, x_p$$

d'éléments de X, deux à deux distincts, et en convenant qu'on regarde comme identiques deux telles suites lorsqu'elles ne diffèrent que par l'ordre des facteurs. Il est naturellement beaucoup plus clair d'utiliser le langage de la théorie des ensembles.

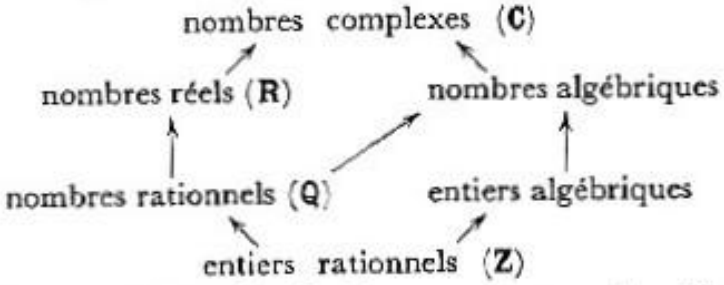
8. *Entiers rationnels* (*)

En plus des entiers naturels, on a besoin en Mathématiques des entiers « de signe quelconque » ou **entiers rationnels**. Nous allons indiquer sommairement comment on peut les définir.

(*) Ce n° a pour but de justifier des résultats avec lesquels le lecteur est déjà familier; il peut donc être négligé en première lecture.

Indiquons en passant que les entiers de signe quelconque (que nous appelons, comme le font tous les mathématiciens, des *entiers rationnels*) sont généralement appelés, dans l'enseignement secondaire français, des « entiers algébriques » (de même qu'on y appelle « nombres algébriques » les nombres de signe quelconque, entiers ou non, que nous appelons, à nouveau comme tous les mathématiciens, des *nombres réels*). Ces divergences de terminologie n'auraient aucune importance si les mathématiciens n'utilisaient déjà dans un tout autre sens les expressions « entier algébrique » et « nombre algébrique » que l'on définira plus loin (§ 11, *Exemple* 11 ainsi que les *Exercices* du § 26).

Les relations d'inclusion entre les diverses notions de nombre qu'on rencontre le plus souvent sont schématisées par le diagramme suivant



où chaque flèche indique que l'ensemble de départ est contenu dans l'ensemble d'arrivée.

L'idée fondamentale est que, si x et y sont deux entiers naturels, il existe un entier rationnel z tel que

$$x + z = y$$

— c'est précisément pour rendre la soustraction possible dans tous les cas qu'on a inventé les entiers négatifs. Si donc on suppose déjà construit l'ensemble $\mathbf{Z}$ des entiers rationnels on peut définir une application

$$\mathrm{D} : \mathbf{N} \times \mathbf{N} \to \mathbf{Z},$$

d'ailleurs *surjective*, par

$$\mathrm{D}(x, y) = x - y.$$

Il est clair de plus que la relation

$$\mathrm{D}(x,y) = \mathrm{D}(x',y') \quad \text{équivaut à} \quad x + y' = x' + y.$$

Ces remarques expliquent la construction qui suit (et dans laquelle nous ne supposerons plus, évidemment, le problème résolu...).

Pour construire l'ensemble $\mathbf{Z}$ des entiers rationnels, on part de l'ensemble $\mathbf{N} \times \mathbf{N}$ des couples d'entiers naturels; et, sur cet ensemble, on définit une *relation d'équivalence* R en déclarant que deux couples (x, y) et (x', y') d'entiers naturels sont équivalents mod R si et seulement si l'on a

$$x + y' = x' + y;$$

la vérification du fait que R est une relation d'équivalence est triviale, et laissée au lecteur. Ceci fait, on pose, par définition,

$$\mathbf{Z} = (\mathbf{N} \times \mathbf{N})/\mathrm{R},$$

et on appelle **entier rationnel** tout élément de l'ensemble $\mathbf{Z}$.

Il faut encore, bien entendu, définir les opérations algébriques sur les entiers rationnels, et montrer comment on peut considérer les entiers naturels comme des entiers rationnels particuliers. Notons pour cela D l'application canonique (§ 4, n° 2) de $\mathbf{N} \times \mathbf{N}$ sur $\mathbf{Z}$; pour définir la **somme** et le **produit** de deux entiers rationnels z et z', on choisit des couples (x, y), $(x', y') \in \mathbf{N} \times \mathbf{N}$ tels que

$$z = \mathrm{D}(x, y), \qquad z' = \mathrm{D}(x', y'),$$

et on pose (*)

$$z + z' = \mathrm{D}(x + x', y + y')$$
$$zz' = \mathrm{D}(xx' + yy', xy' + x'y);$$

on peut aussi, si l'on préfère, appliquer le Théorème 3 du § 4 en prenant, dans les

(*) Cette construction s'explique par le fait qu'on désire parvenir finalement aux relations

$$(x - y) + (x' - y') = (x + x') - (y + y'),$$
$$(x - y) \cdot (x' - y') = (xx' + yy') - (xy' + x'y).$$

notations du Théorème 3 en question,

$$X = Y = Z = \mathbf{N} \times \mathbf{N}, \qquad R = S = T,$$

et pour application $f : X \times Y \to Z$ soit celle qui est donnée par

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

soit celle qui est donnée par

$$f((x, y), (x', y')) = (xx' + yy', xy' + x'y).$$

Il faut vérifier la condition *a*) du Théorème 3; cela veut dire qu'on doit prouver que les relations

$$(x, y) \equiv (u, v) \bmod R \qquad \text{et} \qquad (x', y') \equiv (u', v') \bmod R$$

impliquent

$$(x + x', y + y') \equiv (u + u', v + v') \bmod R$$

et

$$(xx' + yy', xy' + x'y) \equiv (uu' + vv', uv' + u'v) \bmod R$$

Vu la définition de R, on est donc ramené à établir que, quels que soient les entiers naturels $x, y, x', y', u, v, u', v'$, les relations

$$x + v = y + u \qquad \text{et} \qquad x' + v' = y' + u'$$

impliquent les relations

$$x + x' + v + v' = y + y' + u + u'$$

et

$$xx' + yy' + uv' + u'v = xy' + x'y + uu' + vv',$$

ce qui est évidemment facile.

La somme et le produit étant définis dans l'ensemble **Z**, on établit ensuite les propriétés fondamentales de ces opérations, à savoir :

(I) on a $x + y = y + x$, $x + (y + z) = (x + y) + z$ quels que soient $x, y, z \in \mathbf{Z}$, et il existe un élément 0 de **Z** et un seul tel que $x + 0 = x$ pour tout $x \in \mathbf{Z}$.

(II) quels que soient $x, y \in \mathbf{Z}$, il existe un $z \in \mathbf{Z}$ et un seul tel que $x + z = y$.

(III) on a $xy = yx$, $x(yz) = (xy)z$ quels que soient $x, y, z \in \mathbf{Z}$, et il existe un élément 1 de **Z** tel que $1x = x$ pour tout x.

(IV) on a $x(y + z) = xy + xz$ quels que soient $x, y, z \in \mathbf{Z}$.

Montrons par exemple comment on établit (IV). On choisit dans $\mathbf{N} \times \mathbf{N}$ des couples tels que

$$x = D(x', x''), \qquad y = D(y', y''), \qquad z = D(z', z'');$$

on a alors

$$\begin{aligned} x(y+z) &= \mathrm{D}(x', x'').\mathrm{D}(y'+z', y''+z'') \\ &= \mathrm{D}[x'(y'+z') + x''(y''+z''), x'(y''+z'') + x''(y'+z')] \end{aligned}$$

et

$$\begin{aligned} xy + xz &= \mathrm{D}(x'y' + x''y'', x'y'' + x''y') + \mathrm{D}(x'z' + x''z'', x'z'' + x''z') \\ &= \mathrm{D}(x'y' + x''y'' + x'z' + x''z'', x'y'' + x''y' + x'z'' + x''z'), \end{aligned}$$

et on obtient (IV) en comparant les résultats obtenus.

Il reste enfin à identifier chaque entier naturel n à un entier rationnel — à savoir à l'entier rationnel

$$\mathrm{D}(n, 0);$$

on vérifie facilement que l'application $n \to \mathrm{D}(n, 0)$ de **N** dans **Z** ainsi définie est *injective*, et compatible avec les opérations algébriques définies sur les entiers naturels d'une part, et sur les entiers rationnels d'autre part. De cette façon, il n'y a aucun inconvénient à considérer **N** comme une partie de **Z**.

Ceci fait, on observe que, quels que soient les entiers naturels x et y, on a

$$\mathrm{D}(x, y) = \mathrm{D}(x, 0) - \mathrm{D}(y, 0)$$

(la **différence** $a - b$ entre deux entiers rationnels a et b étant, par définition, l'unique entier rationnel c tel que $a = b + c$: voir la propriété (II) ci-dessus); en effet

$$\mathrm{D}(x, y) + \mathrm{D}(y, 0) = \mathrm{D}(x + y, y)$$

en sorte que tout revient à vérifier que $\mathrm{D}(x + y, y) = \mathrm{D}(x, 0)$, i.e. que

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

ce qui est clair. En convenant de poser dorénavant

$$\mathrm{D}(x, 0) = x \quad \text{pour tout } x \in \mathbf{N},$$

la relation précédente s'écrit donc

$$\mathrm{D}(x, y) = x - y$$

et montre que *tout entier rationnel est différence de deux entiers naturels.*

Soit z un entier rationnel; l'entier rationnel $0 - z$ se note

$$-z$$

et s'appelle l'**opposé** de z; il est donc caractérisé par la relation

$$z + (-z) = 0,$$

et il est clair que

$$-z = y - x \qquad \text{si} \qquad z = x - y.$$

Ceci dit, écrivons $z = x - y$ avec $x, y \in \mathbf{N}$; deux cas sont possibles.

Il peut arriver que $x \geqslant y$; alors il existe un $z' \in \mathbf{N}$ tel que $x = y + z'$ et comme cette relation est aussi valable dans **Z** on voit, en vertu de l'unicité de la soustraction dans **Z**, que l'on a dans ce cas $z = z'$, autrement dit $z \in \mathbf{N}$.

Dans le cas où l'on a au contraire $y \geqslant x$, la relation $-z = y - x$ montre que $-z \in \mathbf{N}$.

Autrement dit, *pour tout entier rationnel* z, *on a soit* $z \in \mathbf{N}$, *soit* $-z \in \mathbf{N}$. Les entiers rationnels z tels que $z \in \mathbf{N}$ sont dits **positifs**, les autres sont dits **négatifs**. Si x et y sont deux entiers rationnels, on écrit

$$x \leqslant y$$

lorsque $y - x$ est positif, de sorte que les $z \in \mathbf{Z}$ positifs sont caractérisés par la relation

$$z \geqslant 0.$$

On démontre facilement que, dans l'ensemble **Z**, la relation $x \leqslant y$ possède les propriétés « évidentes », à savoir que les relations

$$x \leqslant y \quad \text{et} \quad y \leqslant z \quad \text{impliquent} \quad x \leqslant z,$$

que la relation

$$x = y \quad \text{équivaut à} \quad x \leqslant y \quad \text{et} \quad y \leqslant x,$$

que quels que soient x et y on a

$$\text{soit } x \leqslant y \text{ soit } y \leqslant x,$$

que la relation

$$x \leqslant y \quad \text{équivaut à} \quad x + z \leqslant y + z,$$

et que la relation

$$x \leqslant y \quad \text{équivaut à} \quad xz \leqslant yz \quad \text{si} \quad z > 0, \quad \text{à} \quad xz \geqslant yz \quad \text{si} \quad z < 0.$$

Il nous arrivera parfois par la suite d'utiliser d'autres propriétés de **Z** que celles que nous venons d'énoncer; nous les démontrerons complètement toutes les fois qu'elles ne seront pas « évidentes ». Pour le moment, on conseille au lecteur de ne pas trop chercher à approfondir les considérations du présent n° : elles n'ont pour but que de justifier des résultats avec lesquels il est déjà familier, et sans lesquels il ne saurait être question de faire des Mathématiques. C'est quand il voudra approfondir sa compréhension des théories exposées dans cet ouvrage que le lecteur aura intérêt à écrire en détail toutes les démonstrations que nous avons escamotées dans le présent n°.

9. *Nombres rationnels*

Après avoir construit des entiers naturels, puis des entiers rationnels, il est nécessaire de construire les **nombres rationnels** i.e. les quotient p/q de deux entiers rationnels p et q, avec $q \neq 0$. L'ensemble de ces entiers rationnels se note

Q.

Toutefois, nous n'exposerons pas ici la construction de **Q** à partir de **Z**; elle s'effectue par des méthodes analogues à celles qu'on a utilisées pour passer de **N** à **Z**; mais surtout, la construction de **Q** à partir de **Z** est un cas particulier d'un procédé beaucoup plus général, qui sera exposé en détail au § 29, et dont nous aurons de toute façon besoin dans d'autres cas.

On conseille donc au lecteur, soit de prendre pour argent comptant les propriétés « évidentes » des nombres rationnels (qui du reste n'interviendront jamais dans le présent ouvrage qu'à titre d'exemple), soit de passer directement au § 29 après avoir lu les §§ 6, 7 et 8 qui sont indispensables à la compréhension du § 29. Il va de soi que, pour le débutant, il sera de beaucoup préférable d'utiliser la première méthode, à condition de ne pas se faire d'illusions (i.e. de ne pas considérer comme triviales des propriétés qu'on ne sait pas démontrer immédiatement).

1. *Lois de composition ; associativité et commutativité*

Étant donné un ensemble X, on appelle **loi de composition sur X** toute application de l'ensemble produit $X \times X$ dans l'ensemble X lui-même. Une loi de composition sur X consiste donc, intuitivement, à faire correspondre, à tout couple (x, y) d'éléments de X, un troisième élément de X qui dépend de x et de y suivant une loi donnée d'avance.

Dans la pratique on emploie pour désigner les lois de composition des notations telles que $(x, y) \mapsto x + y$, ou $(x, y) \mapsto xy$, ou $(x, y) \mapsto x \wedge y$, etc... Dans ce § nous utiliserons souvent le signe $\perp$, qui n'est utilisé nulle part ailleurs en Mathématiques actuellement (et qui par conséquent est susceptible de désigner n'importe quelle loi de composition).

Soit $(x, y) \mapsto x \perp y$ une loi de composition sur un ensemble X. On dit que cette loi de composition est **associative** si l'on a

$$x \perp (y \perp z) = (x \perp y) \perp z \quad \text{quels que soient } x, y, z \in X.$$

Dans ce cas, étant donnés des éléments $x_1, x_2, \ldots, x_n$ de X, en nombre quelconque, on pose, par définition,

$$x_1 \perp x_2 \perp \ldots \perp x_n = (x_1 \perp \ldots \perp x_{n-1}) \perp x_n$$

(récurrence sur n), et on a alors la relation

$$x_1 \perp \ldots \perp x_n = (x_1 \perp \ldots \perp x_p) \perp (x_{p+1} \perp \ldots \perp x_n)$$

pour tout entier p tel que $1 \leqslant p \leqslant n$.

Supposons la loi de composition considérée notée $(x, y) \to xy$, comme une multiplication — on dit alors qu'on utilise la **notation multiplicative**. Pour tout $x \in X$ et tout entier $n \geqslant 1$, on définit alors la **puissance** n^e de x par la formule

$$x^n = x \ldots x \quad (n \text{ facteurs}),$$

et on a alors

$$x^p x^q = x^{p+q}$$

quels que soient les entiers $p, q \geqslant 1$.

Si au contraire on note $(x, y) \mapsto x + y$ la loi de composition considérée, auquel cas on dit qu'on utilise la **notation additive,** on définit

$$nx = x + \cdots + x \quad (n \text{ termes})$$

pour tout $x \in X$ et tout entier $n \geqslant 1$; bien entendu, il n'y a entre cette notion et celle de puissance n^{e} qu'une différence dans les *notations* utilisées; cela dit, la formule de multiplication des puissances écrite plus haut en notation multiplicative se traduit, en notation additive, par la relation

$$px + qx = (p + q)x.$$

Revenons à une loi de composition $(x, y) \mapsto x \perp y$ sur un ensemble X; on dit qu'une telle loi est **commutative** si l'on a

$$x \perp y = y \perp x \quad \text{quels que soient } x, y \in X.$$

On emploie la notation multiplicative aussi bien pour des lois de composition non commutatives que pour des lois de composition commutatives; mais, dans la pratique, *la notation additive s'emploie uniquement pour des lois de composition commutatives.*

Soit $(x, y) \mapsto x \perp y$ une loi de composition *associative et commutative* sur un ensemble X, et soit $(x_i)_{i \in I}$ une famille *finie* d'éléments de X. Soit n le nombre d'éléments de I et écrivons ceux-ci sous forme d'une suite $i_1, \ldots, i_n$; l'élément

$$x_{i_1} \perp x_{i_2} \perp \cdots \perp x_{i_n}$$

de X ne dépend évidemment pas (vu l'associativité et la commutativité de la loi de composition considérée) de la façon dont on a écrit les éléments de I sous la forme d'une suite de n termes. On pose alors, par définition,

$$\bigperp_{i \in I} x_i = x_{i_1} \perp \cdots \perp x_{i_n}.$$

Si la loi de composition considérée est écrite *multiplicativement,* on écrit

$$\prod_{i \in I} x_i = x_{i_1} \cdots x_{i_n};$$

si elle est écrite *additivement,* on utilise la notation

$$\sum_{i \in I} x_i = x_{i_1} + \cdots + x_{i_n}.$$

Remarque 1. Les notations condensées qu'on vient d'introduire subissent dans

la pratique de nombreuses modifications que l'usage enseignera. Par exemple, si I est l'ensemble formé des entiers $1, \cdots, n$, on écrit souvent

$$x_1 + \cdots + x_n = \sum_{i=1}^{i=n} x_i \qquad \text{ou} \qquad \sum_{1 \leqslant i \leqslant n} x_i;$$

si I est l'ensemble des couples (i, j) d'entiers tels que $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant q$, et si l'on note x_{ij} le terme « général » de la famille considérée, on utilise fréquemment la notation

$$\sum_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}} x_{ij} \qquad \text{au lieu de} \qquad \sum_{(i,\, j) \in \mathrm{I}} x_{ij};$$

on notera que, dans ce cas, l'associativité de la loi de composition se traduit par la relation

$$\sum_{\substack{1 \leqslant i \leqslant p \\ 1 \leqslant j \leqslant q}} x_{ij} = \sum_{1 \leqslant i \leqslant p} \left(\sum_{1 \leqslant j \leqslant q} x_{ij} \right),$$

où le second membre désigne la somme

$$(x_{11} + x_{12} + \cdots + x_{1q}) + \cdots + (x_{p1} + x_{p2} + \cdots + x_{pq});$$

l'emploi de ces notations condensées est souvent indispensable pour éviter des formules inextricables.

Notons enfin que dans la notation

$$\sum_{i \in \mathrm{I}} x_i,$$

la lettre i ne joue aucun rôle et n'intervient pas réellement dans le résultat — elle indique simplement une opération à effectuer (à savoir prendre la somme de tous les x_i obtenus en faisant varier i dans I), et on peut la remplacer par toute autre lettre *non encore utilisée par ailleurs* (cette dernière précaution est essentielle pour éviter des erreurs grossières).

Reprenons une loi de composition quelconque $(x, y) \mapsto x \perp y$ sur un ensemble X. On appelle **élément neutre** pour cette loi de composition tout élément $e \in \mathrm{X}$ tel que l'on ait

$$x \perp e = e \perp x = x \qquad \text{pour tout } x \in \mathrm{X}.$$

THÉORÈME 1. *Si une loi de composition admet un élément neutre, elle en admet un seul.*

Supposons en effet que e' et e'' soient des éléments neutres; la formule $e' \perp x = x$ donne en particulier $e' \perp e'' = e''$; la formule $x \perp e'' = x$ donne en particulier $e' \perp e'' = e'$; on a donc $e' = e''$, d'où le Théorème.

Donnons maintenant quelques exemples importants de lois de composition.

Exemple 1. Sur l'ensemble **Z** des entiers rationnels (entiers de signe quelconque),

on a trois lois de composition que tout le monde connaît : l'*addition* $(x, y) \mapsto x + y$, qui est commutative, associative, et admet un élément neutre (à savoir le nombre 0) ; la *multiplication* $(x, y) \mapsto xy$, qui est commutative, associative, et admet un élément neutre (à savoir le nombre 1) ; enfin la *soustraction* $(x, y) \mapsto x - y$, qui n'est pas commutative, ni associative, et n'admet pas d'élément neutre.

On pourrait dans cet Exemple remplacer $\mathbf{Z}$ par l'ensemble $\mathbf{Q}$ des nombres rationnels, ou par l'ensemble $\mathbf{R}$ des nombres réels.

Exemple 2. Soit $\mathbf{Q}^*$ l'ensemble des nombres rationnels *non nuls* ; la multiplication $(x, y) \mapsto xy$ est une loi de composition (associative, commutative et avec élément neutre) sur $\mathbf{Q}^*$; il en est de même de la division $(x, y) \mapsto x/y$ (qui n'est pas associative, ni commutative, et n'admet pas d'élément neutre). On aura soin de remarquer que la division *n'est pas* une loi de composition sur l'ensemble $\mathbf{Q}$ de tous les nombres rationnels, car, le quotient x/y n'étant pas défini pour $y = 0$, l'application $(x, y) \to x/y$ n'est pas définie sur $\mathbf{Q} \times \mathbf{Q}$ tout entier.

Exemple 3. On prend $X = \mathbf{N}$, ensemble des entiers naturels, et les applications $(x, y) \mapsto \text{ppcm}\ (x, y)$ et $(x, y) \mapsto \text{pgcd}\ (x, y)$; ce sont des lois de composition commutatives et associatives ; la première admet un élément neutre, la seconde n'en admet pas (le lecteur devra bien entendu démontrer lui-même ces assertions à titre d'exercice).

Exemple 4. Soient E un ensemble quelconque et X l'ensemble de toutes les applications de E dans E ; l'application $(f, g) \mapsto f \circ g$ de $X \times X$ dans X est une loi de composition sur X, laquelle est associative (§ 2, Théorème 2), admet un élément neutre (l'application identique j_E), mais n'est pas commutative.

Exemple 5. Soient E un ensemble quelconque et $X = \mathfrak{P}(E)$ l'ensemble des parties de E ; alors les applications $(x, y) \mapsto x \cap y$ et $(x, y) \mapsto x \cup y$ sont des lois de composition associatives et commutatives sur X, en vertu des formules du § 3, n° 1. La première de ces lois admet pour élément neutre l'ensemble E, la seconde, l'ensemble vide.

Exemple 6. Soit X l'ensemble des vecteurs d'origine donnée 0 dans l'espace (ce mot étant pris au sens usuel) ; à tout couple de vecteurs x, y d'origine 0, associons leur produit « vectoriel » (noté, suivant les auteurs, $x \times y$ ou $x \wedge y$; nous emploierons ici la notation $x \wedge y$) ; on obtient ainsi une loi de composition qui n'est ni associative, ni commutative.

2. *Éléments symétrisables*

Soit $(x, y) \mapsto x \perp y$ une loi de composition sur un ensemble X, admettant un élément neutre e. Étant donné un élément $x \in X$, on appelle **symétrique à gauche** (resp. **symétrique à droite**) de x tout élément $x' \in X$ tel que l'on ait

$$x' \perp x = e \qquad (\text{resp. } x \perp x' = e) ;$$

et on appelle **symétrique** de x tout élément x' vérifiant

$$x' \perp x = x \perp x' = e.$$

Enfin, on dit que x est **symétrisable** s'il existe un élément symétrique de x.

Lorsqu'on a affaire à une loi de composition écrite en notation *multiplicative*, on emploie le mot **inverse** au lieu du mot symétrique, et le mot **inversible** au lieu du mot symétrisable [par exemple, si l'on considère sur $\mathbf{Q}$ la loi de composition $(x, y) \mapsto x\,y$, les éléments inversibles sont les nombres rationnels *non nuls*, et l'inverse d'un tel nombre x est le nombre $1/x$]; l'inverse d'un élément inversible x de X se note alors généralement

$$x^{-1}.$$

Lorsqu'on a affaire à une loi de composition écrite en notation *additive*, on dit **opposé** au lieu de symétrique, et on note

$$-x$$

l'opposé d'un $x \in X$ (tout au moins si l'on note 0 l'élément neutre de X, ce qui est presque toujours le cas en notation additive).

Notons enfin que la distinction entre symétrique à gauche et symétrique à droite n'a intérêt que pour une loi de composition non commutative.

Donnons un exemple qui montre que, dans le cas non commutatif, les trois notions sont distinctes.

Exemple 7. Prenons la loi de composition de l'*Exemple* 4 ci-dessus. Dire qu'un $f \in X$ admet un *inverse à gauche* signifie qu'il existe une application $g \in X$ telle que $g \circ f = j_E$; pour cela, il faut et il suffit que f soit *injective* (§ 2, Théorème 3). Dire f admet un *inverse à droite* signifie qu'il existe une application g telle que $f \circ g = j_E$; pour cela il faut et il suffit que f soit *surjective* (§ 2, Théorème 4). Enfin dire que que f est inversible signifie évidemment que f est *bijective*, et alors l'inverse de f pour la loi de composition considérée n'est autre que l'application réciproque au sens du § 2.

THÉORÈME 2. *Soit $(x, y) \mapsto x \perp y$ une loi de composition associative et admettant un élément neutre sur un ensemble* E. *Pour qu'un élément x de* E *soit symétrisable, il faut et il suffit qu'il admette un symétrique à gauche et un symétrique à droite; x admet alors un seul symétrique, qui est aussi l'unique symétrique à gauche et l'unique symétrique à droite de x.*

Soient x' un symétrique à gauche et x'' un symétrique à droite de x; on a donc $x' \perp x = x \perp x'' = e$, élément neutre de E; tenant compte de l'associativité, on déduit de là que $x'' = e \perp x'' = (x' \perp x) \perp x'' = x' \perp (x \perp x'') = x' \perp e = x'$; tout symétrique à droite de x est donc égal à tout symétrique à gauche, ce qui montre que x admet *un seul* symétrique à gauche, *un seul* symétrique à droite, et qu'ils sont égaux; notant x' leur valeur commune, on a $x' \perp x = x \perp x' = e$, de sorte que x est symétrisable et admet x' pour symétrique (nécessairement unique, car un symétrique de x est *a fortiori* symétrique à gauche et symétrique à droite, donc égal à x'). Ceci achève la démonstration.

Nous supposerons jusqu'à la fin de ce § que la loi de composition considérée sur l'ensemble E *est associative et admet un élément neutre e.* Pour un élément symétrisable x de E, on peut donc parler *du* symétrique x' de x (ou de l'inverse x^{-1} en notation multiplicative, de l'opposé $-x$ en notation additive).

THÉORÈME 3. *Si un $x \in E$ est symétrisable, il en est de même de son symétrique x', et celui-ci admet x pour symétrique. Si x et y sont symétrisables, il en est de même de $x \perp y$, et on a la relation*

$$(x \perp y)' = y' \perp x'.$$

Les relations

$$x' \perp x = x \perp x' = e$$

rendent triviale la première assertion de l'énoncé. Pour établir la seconde, on calcule

$$(y' \perp x') \perp (x \perp y) = y' \perp (x' \perp x) \perp y = y' \perp e \perp y = y' \perp y = e,$$
$$(x \perp y) \perp (y' \perp x') = x \perp (y \perp y') \perp x' = x \perp e \perp x' = x \perp x' = e,$$

ce qui montre bien que $x \perp y$ est symétrisable et admet $y' \perp x'$ pour symétrique.

En notation multiplicative, le Théorème 3 se traduit comme suit : si x est inversible il en est de même de son inverse, et on a

$$(x^{-1})^{-1} = x;$$

si x et y sont inversibles, il en est de même de xy et on a

$$(xy)^{-1} = y^{-1}x^{-1}.$$

En notation additive, ce qui suppose la loi de composition considérée commutative, le Théorème 3 se traduit ainsi : si x admet un opposé, il en est de même de son opposé, et on a

$$-(-x) = x;$$

si x et y admettent des opposés, il en est de même de $x + y$, et on a

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

ce qu'on écrit d'ailleurs en général sous la forme

$$-(x + y) = -x - y.$$

THÉORÈME 4. *Soit a un élément symétrisable de E; alors, pour tout $b \in E$, il existe un et un seul $x \in E$ tel que*

$$a \perp x = b,$$

à savoir

$$x = a' \perp b.$$

En effet, la relation $a \perp x = b$ implique $a' \perp (a \perp x) = a' \perp b$, i.e.

$$a' \perp b = (a' \perp a) \perp x = e \perp x = x;$$

inversement, de $x = a' \perp b$ résulte

$$a \perp x = a \perp (a' \perp b) = (a \perp a') \perp b = e \perp b = b,$$

ce qui achève la démonstration.

En notation multiplicative: si a est inversible, l'équation

$$ax = b$$

possède une et une seule solution, à savoir

$$x = a^{-1}b;$$

en notation additive: si a admet un opposé, alors l'équation

$$a + x = b$$

possède une et une seule solution, à savoir

$$x = b + (-a),$$

qu'on écrit d'ailleurs

$$x = b - a$$

(**différence** entre b et a).

Indiquons pour terminer ce § que, dans le reste de cet ouvrage, on n'utilisera *jamais plus* le signe $\perp$, ni les mots « symétrisable », « symétrique », etc... On aura *toujours* affaire à des lois de composition notées soit multiplicativement, soit additivement; utiliser pour de telles lois de composition les mots « symétrisable » et « symétrique » (par exemple, parler du « symétrique pour la multiplication » d'un nombre réel non nul, comme le font souvent les débutants), serait parfaitement ridicule.

1. Définition des groupes; exemples

On appelle **groupe** *un couple formé d'un ensemble* G *et d'une loi de composition* $(x, y) \to xy$ *sur l'ensemble* G, *ces données devant vérifier les trois conditions suivantes :*

a) on a $x(yz) = (xy)z$ *quels que soient* $x, y, z \in G$ (associativité) ;

b) il existe un élément e *de* G *tel que* $xe = ex = x$ *pour tout* $x \in G$ (existence d'un élément neutre) ;

c) pour tout $x \in G$, *il existe un élément* $x^{-1} \in G$ *tel que* $x^{-1}x = xx^{-1} = e$ (existence d'un inverse pour tout élément de G).

Pour définir un groupe, il ne suffit pas de se donner un ensemble G; il faut aussi se donner une loi de composition sur l'ensemble G, vérifiant les conditions *a*), *b*), *c*) ci-dessus; néanmoins, on désigne toujours un groupe par la même lettre, G par exemple, que l'ensemble qui en constitue l'une des données.

Le débutant aura soin de *ne pas* dire qu'un groupe « est un ensemble G sur lequel *il existe* une loi de composition vérifiant les conditions *a*), *b*), *c*) ci-dessus », car on peut facilement démontrer que, sur tout ensemble, il existe une telle loi de composition, et même qu'on peut en construire une infinité pour peu que l'ensemble donné soit lui-même infini; en disant qu'un groupe est « un ensemble sur lequel il existe » une loi de composition, on ne dit donc rien d'autre que ceci : « un groupe est un ensemble » — définition dont la stupidité est particulièrement claire...

En fait, en théorie des groupes, on ne s'intéresse pas du tout à l'existence (i.e. à la possibilité de construction) sur un ensemble donné G d'une loi de composition vérifiant les conditions *a*), *b*), *c*); au contraire, on suppose qu'une telle loi *est donnée d'avance* une fois pour toutes, et on se propose de l'utiliser pour démontrer des théorèmes.

Dans la définition donnée plus haut, nous avons utilisé l'écriture multiplicative, ce qu'on fait en effet le plus souvent (signalons en passant que l'élément neutre se note parfois 1 au lieu de e et s'appelle fréquemment l'**élément unité** de G); mais lorsqu'on a un **groupe commutatif** (ou **abélien**), i.e. un groupe dont la loi de composition est commutative, on utilise parfois l'écriture additive $(x, y) \to x + y$; dans ce cas, les conditions *a*), *b*), *c*) se traduisent comme suit :

a') *on a* $x + (y + z) = (x + y) + z$ *quels que soient* $x, y, z \in G$;

b') *il existe un élément* 0 *de* G *tel que* $x + 0 = x$ *pour tout* $x \in G$;

c') *pour tout* $x \in G$, *il existe dans* G *un élément, noté* $-x$, *tel que* $x + (-x) = 0$.

Il faut bien entendu, dans ce cas, ajouter la condition

d') *on a* $x + y = y + x$ *quels que soient* $x, y \in G$.

Exemple 1. L'ensemble **Z** des entiers rationnels et la loi de composition $(x, y) \mapsto x + y$ constituent évidemment un groupe commutatif; on l'appelle le **groupe additif des entiers rationnels.** En remplaçant **Z** par **Q** ou par **R**, on définirait de même le **groupe additif des nombres rationnels** et le **groupe additif des nombres réels.**

Exemple 2. Le couple formé par l'ensemble $\mathbf{Q}^*$ des nombres rationnels non nuls et par la loi de composition $(x, y) \mapsto xy$ sur cet ensemble est un groupe (dont l'élément neutre est le nombre 1); on l'appelle le **groupe multiplicatif des nombres rationnels non nuls.** On définirait de même le **groupe multiplicatif des nombres réels non nuls,** noté $\mathbf{R}^*$.

Exemple 3. On note $\mathbf{Q}_+^*$ le groupe obtenu en considérant l'ensemble de tous les nombres rationnels *strictement positifs*, la loi de composition sur cet ensemble étant la multiplication usuelle. On note de même $\mathbf{R}_+^*$ le groupe multiplicatif des nombres réels strictement positifs.

On notera par contre que le couple formé par l'ensemble I des nombres réels x tels que $0 < x \leqslant 1$, et par la loi de composition $(x, y) \mapsto xy$ sur cet ensemble, n'est pas un groupe : la condition *c*) de la définition des groupes *n'est pas* vérifiée.

Exemple 4. Soit X un ensemble quelconque; rappelons (§ 2, n° 8) qu'on appelle *permutation de* X toute application *bijective* de X dans X. Soit $\mathfrak{S}(X)$ l'ensemble de ces permutations; si f, g sont des permutations de X, il en est de même de l'application composée $f \circ g$ (§ 2, Théorème 6); la formule $(f, g) \mapsto f \circ g$ définit donc une loi de composition sur l'ensemble $\mathfrak{S}(X)$; cette loi de composition est associative (§ 2, Théorème 2); elle admet un élément neutre, à savoir l'application identique j_X (appelée souvent la **permutation unité** de l'ensemble X); enfin, si f est une permutation de X, il en est de même de l'application réciproque f^{-1} en vertu du § 2, Théorème 5, et celle-ci est évidemment inverse de f pour la loi de composition considérée.

Ainsi, le couple formé par l'ensemble $\mathfrak{S}(X)$ et par la loi de composition $(f, g) \to f \circ g$ sur cet ensemble est un groupe; on l'appelle le **groupe des permutations de l'ensemble** X. C'est l'étude de ces groupes par Galois (lorsque X est un ensemble fini) qui a conduit, historiquement, à la notion générale et « abstraite » de groupe.

Prenons par exemple pour X l'ensemble constitué par les entiers 1, 2, 3; alors $\mathfrak{S}(X)$ comporte six éléments, à savoir les permutations

$$\begin{aligned}
s_1 &: \ 1, 2, 3 \mapsto 1, 2, 3 \\
s_2 &: \ 1, 2, 3 \mapsto 2, 3, 1 \\
s_3 &: \ 1, 2, 3 \mapsto 3, 1, 2 \\
s_4 &: \ 1, 2, 3 \mapsto 1, 3, 2 \\
s_5 &: \ 1, 2, 3 \mapsto 2, 1, 3 \\
s_6 &: \ 1, 2, 3 \mapsto 3, 2, 1
\end{aligned}$$

et la loi de composition est donnée par la « table de multiplication » suivante :

	s_1	s_2	s_3	s_4	s_5	s_6
s_1	s_1	s_2	s_3	s_4	s_5	s_6
s_2	s_2	s_3	s_1	s_5	s_6	s_4
s_3	s_3	s_1	s_2	s_6	s_4	s_5
s_4	s_4	s_6	s_5	s_1	s_3	s_2
s_5	s_5	s_4	s_6	s_2	s_1	s_3
s_6	s_6	s_5	s_4	s_3	s_2	s_1

(on a adopté la convention suivante : pour calculer un produit xy à l'aide de cette table, on porte x en *ligne* et y en *colonne*. Exemple : $s_2 s_4 = s_5$, $s_4 s_2 = s_6$).

Cet exemple prouve l'existence de **groupes finis**, i.e. de groupes à un nombre fini d'éléments; il est clair d'ailleurs que $\mathfrak{S}(X)$ est fini dès que (et seulement si) l'ensemble X est fini.

Lorsque X est l'ensemble formé des entiers $1, 2, \cdots, n$ (on a vu ci-dessus ce qui se passe pour $n = 3$), on utilise, au lieu de la notation $\mathfrak{S}(X)$, la notation

$$\mathfrak{S}_n$$

et on appelle $\mathfrak{S}_n$ le **groupe des permutations de n objets** ou encore le **groupe symétrique à n variables** On a vu au § 5 (Corollaire du Théorème 9) que le nombre d'éléments de $\mathfrak{S}_n$ est l'entier

$$n\,! = 1 \cdot 2 \ldots n,$$

produits des n premiers entiers strictement positifs. Étant donné qu'on a

$$6\,! = 720, \quad 7\,! = 5\,040, \quad 8\,! = 40\,320, \quad 9\,! = 362\,880, \quad 10\,! = 3\,628\,800,$$

il serait tout à fait utopique d'espérer déduire les propriétés des groupes $\mathfrak{S}_n$ d'un examen de leurs tables de multiplication...

Exemple 5. On obtient un groupe commutatif (noté additivement) en considérant l'ensemble G des vecteurs d'origine donnée O dans l'espace usuel, et, sur cet ensemble, la loi de composition $(x, y) \to x + y$ donnée par la classique règle du parallélogramme.

2. *Produit direct de groupes*

Soient $G_1, \ldots, G_n$ des groupes notés multiplicativement; sur l'ensemble produit

$$G = G_1 \times \cdots \times G_n$$

(§ 2, n° 2), considérons la loi de composition donnée par la formule

$$(x_1, \ldots, x_n)\ (y_1, \ldots, y_n) = (x_1 y_1, \ldots, x_n y_n);$$

le couple formé par l'ensemble G et cette loi de composition *est un groupe.*

Pour établir l'associativité, considérons trois éléments

$$x = (x_1, \ldots, x_n), \qquad y = (y_1, \ldots, y_n), \qquad z = (z_1, \ldots, z_n)$$

de G; on a par définition

$$xy = (x_1 y_1, \ldots, x_n y_n), \qquad yz = (y_1 z_1, \ldots, y_n z_n)$$

et par suite

$$(xy)z = ((x_1 y_1)z_1, \ldots, (x_n y_n)z_n), \qquad x(yz) = (x_1(y_1 z_1), \ldots, x_n(y_n z_n)),$$

de sorte que l'associativité dans G résulte de l'associativité des lois de composition données sur $G_1, \ldots, G_n$.

Pour montrer que G possède un élément neutre, il suffit de considérer l'élément

$$e = (e_1, \ldots, e_n)$$

où e_i désigne l'élément neutre de G_i pour $1 \leqslant i \leqslant n$; un calcul trivial montre aussitôt que e est élément neutre pour la loi de composition considérée sur G. Enfin, si

$$x = (x_1, \ldots, x_n)$$

est un élément de G, on voit immédiatement que x admet un inverse, donné par la formule

$$x^{-1} = (x_1^{-1}, \ldots, x_n^{-1}).$$

On obtient donc bien un groupe en munissant l'ensemble produit $G_1 \times \cdots \times G_n$ de la loi de composition définie plus haut; le groupe ainsi obtenu s'appelle le **produit direct des groupes** $G_1, \ldots, G_n$.

Lorsque les groupes $G_1, \ldots, G_n$ sont commutatifs et notés additivement, on utilise aussi l'écriture additive sur leur produit direct (ce qui est légitime, un produit direct de groupes commutatifs étant commutatif). La loi de composition sur le produit direct est donc alors donnée par la relation

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

et l'élément neutre du produit direct n'est autre que

$$(0, \ldots, 0)$$

(où l'on désigne par le même symbole 0 les éléments neutres des divers groupes $G_1, \ldots, G_n$).

Enfin, étant donné un groupe G, on définit, pour tout entier $n \geqslant 1$, le *groupe*

$$G^n$$

comme étant le produit direct de n groupes identiques à G :

$$G^n = G \times \cdots \times G \quad (n \text{ facteurs}).$$

Exemple 6. Le **groupe additif** $\mathbf{Z}^n$ est défini comme suit : ses éléments sont les suites $(x_1, \ldots, x_n)$ de n entiers rationnels, et la loi de composition est donnée par la formule

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

On définirait de même le **groupe additif** $\mathbf{Q}^n$ (où $\mathbf{Q}$ est le groupe additif des nombres rationnels, défini dans l'*Exemple* 1), et le **groupe additif** $\mathbf{R}^n$ (où $\mathbf{R}$ est le groupe additif des nombres réels).

Choisissons, dans un plan, un système d'axes de coordonnées Ox, Oy, et à tout élément (x, y) de $\mathbf{R}^2$ associons le vecteur $\overrightarrow{OP}$ d'origine O ayant pour composantes, relativement au système de coordonnées choisi, les nombres x et y. On obtient ainsi une bijection de l'ensemble $\mathbf{R}^2$ sur l'ensemble des vecteurs d'origine O dans le plan. Cette bijection transforme la somme (dans le groupe additif $\mathbf{R}^2$) de deux éléments (x', y') et (x'', y'') de $\mathbf{R}^2$ en la somme des vecteurs $\overrightarrow{OP'}$ et $\overrightarrow{OP''}$ qui leur correspondent : en effet, il est bien connu que pour additionner des vecteurs, on doit additionner leurs composantes.

On peut donc considérer l'addition dans $\mathbf{R}^2$ comme une traduction algébrique de la notion « géométrique » de somme de deux vecteurs dans le plan.

3. *Sous-groupes d'un groupe*

On dit qu'une partie H *d'un groupe* G *est un* **sous-groupe** *de* G *si* H *est non vide et si les relations*

$$x \in H \quad \textit{et} \quad y \in H \quad \textit{impliquent} \quad xy^{-1} \in H.$$

Un groupe G possède toujours au moins deux sous-groupes : G tout entier, et l'ensemble $\{e\}$ réduit à l'élément neutre de G. Il est clair d'autre part que les groupes additifs $\mathbf{Z}$ et $\mathbf{Q}$ sont des sous-groupes du groupe additif $\mathbf{R}$ (*Exemple* 1).

Soit H un sous-groupe d'un groupe G; comme H est non vide, il contient au moins un élément a; des relations $a \in H$ et $a \in H$ résulte alors que H contient

$$aa^{-1} = e;$$

ainsi, *un sous-groupe* H *de* G *contient toujours l'élément neutre de* G. D'autre part, si H contient un élément x, comme il contient e et x il contient aussi

$$ex^{-1} = x^{-1};$$

ainsi, *la relation* $x \in H$ *implique la relation* $x^{-1} \in H$. Soient alors x, y des éléments de H; comme H contient, d'après ce qui précède, x et y^{-1}, il devra contenir

$$x(y^{-1})^{-1} = xy;$$

donc, *les relations* $x \in H$ *et* $y \in H$ *impliquent la relation* $xy \in H$.

Réciproquement, considérons une partie H de G qui possède les trois propriétés suivantes :

a) *les relations* $x \in H$ *et* $y \in H$ *impliquent* $xy \in H$;

b) H *contient l'élément neutre e de* G;

c) *la relation* $x \in H$ *implique la relation* $x^{-1} \in H$.

Alors H est un sous-groupe de G. En effet, H est non vide d'après la condition *b*); soient d'autre part x et y deux éléments de H; d'après *c*), H contient x et y^{-1}; d'après *a*), il contient donc xy^{-1}, et notre assertion est établie.

Ainsi, les conditions *a*), *b*) et *c*) ci-dessus *caractérisent* les sous-groupes; dans la pratique, on les utilise très souvent à la place de la définition initiale.

On notera par contre qu'une partie H de G vérifiant seulement *a*), ou seulement *a*) et *b*), n'est pas nécessairement un sous-groupe de G; si par exemple G est le groupe additif **Z** des entiers rationnels, l'ensemble **N** des entiers $n \geqslant 0$ vérifie *a*) et *b*) mais n'est pas un sous-groupe de G.

Soit H un sous-groupe d'un groupe G; la condition *a*) ci-dessus montre que l'application $(x, y) \to xy$ de $G \times G$ dans G applique $H \times H$ dans H, donc « induit » une loi de composition sur H; cela dit, *l'ensemble* H, *muni de cette loi de composition, est un groupe*. En effet, la loi de composition donnée sur G étant associative l'est *a fortiori* sur H; puisque H contient l'élément neutre de G, il est clair d'autre part que la loi de composition considérée sur H admet bien un élément neutre (à savoir celui de G); enfin, tout $x \in H$ est inversible dans H en vertu de la condition *c*) ci-dessus.

Dorénavant, quand nous considérerons un sous-groupe H d'un groupe G, nous regarderons toujours H comme étant lui-même un groupe en le munissant, comme ci-dessus, de la loi de composition induite par celle du groupe donné G.

Exemple 7. Étant donné un ensemble X quelconque, on appelle **groupe de transformations de l'ensemble** X tout sous-groupe du groupe $\mathfrak{S}(X)$ des permutations de X. Un groupe de transformations de X est donc un ensemble G d'applications de X dans X, possédant les propriétés suivantes : *toute* $s \in G$ *est bijective;* G *contient l'application identique* j_X*; et si* G *contient deux applications s et t, il contient aussi* $s \circ t^{-1}$. On peut alors regarder l'ensemble G comme un groupe en le munissant de la loi de composition

$$(s, t) \mapsto s \circ t.$$

La Géométrie élémentaire fournit de nombreux exemples de groupes de transformations : le groupe des translations sur la droite, ou dans le plan, ou dans l'espace; le groupe des rotations autour d'un point dans le plan, ou dans l'espace; le groupe des déplacements dans le plan, ou dans l'espace; le groupe des homothéties de centre donné et de rapport *non nul* dans le plan ou dans l'espace; etc, etc...

Exemple 8. Prenons pour G le groupe additif $\mathbf{Z}$ des entiers; un sous-groupe de $\mathbf{Z}$ est donc un ensemble I d'entiers vérifiant les conditions suivantes : on a $0 \in \mathrm{I}$, et si I contient deux entiers x et y il contient aussi $x - y$. Pour tout entier n, notons $n\mathbf{Z}$ l'ensemble des multiples de n (i.e. l'ensemble des entiers nx où x parcourt $\mathbf{Z}$); il est clair que c'est un sous-groupe de $\mathbf{Z}$. Inversement, *pour tout sous-groupe* I *de* $\mathbf{Z}$, *il existe un et un seul entier* $n \geqslant 0$ *tel que* $\mathrm{I} = n\mathbf{Z}$. Notons d'abord que cette assertion est triviale si $\mathrm{I} = \{0\}$: il suffit alors de prendre $n = 0$. Supposons donc $\mathrm{I} \neq \{0\}$; il existe dans I des entiers non nuls, et même *strictement positifs* (car si $n \in \mathrm{I}$ on a aussi $-n \in \mathrm{I}$); soit alors n *le plus petit entier strictement positif appartenant à* I (rappelons que, dans tout ensemble d'entiers positifs, il existe un élément plus petit que tous les autres d'après le § 5, *Remarque* 2) ; nous allons montrer que $\mathrm{I} = n\mathbf{Z}$. En effet, comme I contient n et n il contient $n + n = 2n$, donc $n + 2n = 3n$, etc..., donc nx pour tout $x \geqslant 1$; d'autre part I contient $n0 = 0$; enfin, si x est un entier négatif, I contient $n(-x) = -nx$ d'après ce qu'on a déjà vu, donc aussi $-(-nx) = nx$; ainsi, on a déjà l'inclusion $n\mathbf{Z} \subset \mathrm{I}$. Il reste à établir l'inclusion opposée. Pour cela considérons un élément $x \in \mathrm{I}$, et écrivons (division euclidienne)

$$x = nq + r \quad (0 \leqslant r < n);$$

le sous-groupe I contient n, donc nq, et comme il contient x il contient aussi $x - nq = r$; or r est positif ou nul, et strictement inférieur à n; si l'on avait $r \neq 0$, n ne serait pas le *plus petit* entier strictement positif contenu dans I; c'est donc que $r = 0$, et ceci démontre que tout élément de I est un multiple de n, autrement dit que $\mathrm{I} \subset n\mathbf{Z}$; on a donc bien en définitive $\mathrm{I} = n\mathbf{Z}$.

Pour établir l'unicité de n, il est suffisant de montrer que si l'on a $p\mathbf{Z} = q\mathbf{Z}$ avec $p, q \geqslant 0$, alors $p = q$; mais comme $q\mathbf{Z}$ contient q, l'hypothèse faite montre que q est multiple de p — et aussi que p est multiple de q — d'où évidemment $p = q$.

Le fait que tout sous-groupe de $\mathbf{Z}$ soit de la forme $n\mathbf{Z}$ joue un rôle très important en Arithmétique et ailleurs et, dans bien des cas, remplace avantageusement les démonstrations fondées sur la théorie de la « division euclidienne » (ou « division avec reste ») des entiers.

Montrons par exemple comment ce résultat conduit aux principales propriétés des pgcd. Soient $x_1, \ldots, x_n$ des entiers non nuls; désignons par I l'ensemble des $x \in \mathbf{Z}$ tels qu'il existe $u_1, \ldots, u_n \in \mathbf{Z}$ vérifiant

$$x = u_1x_1 + \cdots + u_nx_n;$$

il est évident que I est un sous-groupe de $\mathbf{Z}$, et par suite $\mathrm{I} = d\mathbf{Z}$ où d est un entier positif bien déterminé. Tout élément de I est un multiple de d; en particulier, $x_1, \ldots, x_n$ sont des multiples de d, qui est donc un diviseur commun aux nombres donnés; mais d'autre part, tout diviseur commun d' à $x_1, \ldots, x_n$ divise évidemment $u_1x_1 + \cdots + u_nx_n$ quels que soient les entiers $u_1, \ldots, u_n$, donc divise tout élément de I, et en particulier divise d. Autrement dit, d est le plus grand commun diviseur de $x_1, \ldots, x_n$, et on voit en même temps que celui-ci possède la propriété de pouvoir s'écrire sous la forme

$$d = u_1x_1 + \cdots + u_nx_n$$

pour des entiers u_i $(1 \leqslant i \leqslant n)$ convenablement choisis.

Notons, comme conséquence, le *théorème de Bezout : pour que* $x_1, \ldots, x_n$ *soient premiers entre eux, il faut et il suffit qu'il existe des entiers* $u_1, \ldots, u_n$ *tels que*

$$u_1x_1 + \cdots + u_nx_n = 1.$$

En effet, avec les notations ci-dessus, cette condition exprime que le sous-groupe I contient le nombre 1 ; or pour exprimer que les x_i sont premiers entre eux, i.e. que $d = 1$, il suffit évidemment d'exprimer que 1 est un multiple de d, i.e. que $1 \in I$, d'où le résultat annoncé.

On obtiendrait de même la théorie du ppcm en considérant le sous-groupe

$$x_1\mathbf{Z} \cap \cdots \cap x_n\mathbf{Z}$$

de $\mathbf{Z}$; si l'on note m son générateur positif, il est immédiat de vérifier que m est le ppcm des entiers x_i donnés.

Ces questions seront étudiés d'une façon plus détaillée et plus générale au § 31.

Exemple 9. La construction des sous-groupes $n\mathbf{Z}$ de $\mathbf{Z}$ se généralise comme suit. Soient G un groupe (que nous notons maintenant multiplicativement, car il serait inutile de supposer G commutatif pour ce qui va suivre) et x un élément de G; pour tout entier rationnel p, définissons x^p comme suit :

$$x^p = \begin{cases} x \ldots x \text{ } (p \text{ facteurs}) & \text{si } p \geqslant 1 \\ e \text{ (élément neutre)} & \text{si } p = 0 \\ (x^{-1})^{-p} & \text{si } p < 0. \end{cases}$$

A l'aide de l'associativité de la multiplication dans G, on vérifie facilement les règles de calcul suivantes :

$$x^p x^q = x^{p+q}; \qquad (x^p)^{-1} = x^{-p}; \qquad (x^p)^q = x^{pq}.$$

Il s'ensuit que l'ensemble des x^p (pour x donné, et p variable dans $\mathbf{Z}$), est un sous-groupe de G : en effet, il n'est évidemment pas vide, et s'il contient des éléments $u = x^p$, $v = x^q$, la formule $uv^{-1} = x^{p-q}$ montre qu'il contient aussi uv^{-1}.

Ce sous-groupe s'appelle le **sous-groupe de G engendré par x** (de sorte que, dans le groupe additif $\mathbf{Z}$, $n\mathbf{Z}$ n'est autre que le sous-groupe engendré par n), et ses éléments s'appellent les **puissances de x**.

Lorsque G est écrit additivement, on utilise, au lieu de la notation x^p, la notation px, et on dit que les px sont les **multiples entiers de x**. On a donc par définition

$$px = \begin{cases} x + \cdots + x \text{ } (p \text{ facteurs}) & \text{si } p \geqslant 1 \\ 0 \text{ (élément neutre} & \text{si } p = 0 \\ (-p)\ (-x) & \text{si } p < 0, \end{cases}$$

avec les formules

$$px + qx = (p+q)x, \qquad -(px) = (-p)x, \qquad p(qx) = (pq)x.$$

Remarque 1. Dans un groupe additif on a aussi la relation

$$px + py = p(x+y)$$

quels que soient les éléments x et y du groupe considéré. Dans un groupe quelconque G (donc noté multiplicativement), la formule analogue

$$x^p y^p = (xy)^p$$

est fausse sauf si x et y **commutent** (ou **permutent**, comme on dit encore), i.e. si l'on a

$$xy = yx.$$

Tout d'abord, si $xy = yx$, on a

$$(xy)^2 = xyxy = xxyy = x^2y^2$$

et ainsi de suite. Inversement, si la relation $(xy)^p = x^py^p$ est vraie pour $p = 2$, il vient $xyxy = xxyy$; en multipliant les deux membres à gauche par x^{-1} et à droite par y^{-1}, il vient $x^{-1}xyxyy^{-1} = x^{-1}xxyyy^{-1}$, ce qui s'écrit $xy = yx$ comme prévu.

Remarque 2. On appelle **groupe cyclique** tout groupe G pour lequel il existe un $x \in G$ tel que tout élément de G soit une puissance de x; on dit alors que x est un **générateur** de G. Le groupe additif **Z** est cyclique, et admet pour générateur soit 1, soit -1. Il existe des groupes cycliques qui sont finis (considérer un groupe fini G et le sous-groupe de G engendré par un élément quelconque de G); on verra plus loin qu'on peut décrire entièrement leur structure.

4. *Intersection de sous-groupes; générateurs*

On a le résultat suivant :

THÉORÈME 1. *Soit* $(H_i)_{i \in I}$ *une famille de sous-groupes d'un groupe* G. *Alors l'intersection des* H_i *est encore un sous-groupe de* G. *Pour que la réunion des* H_i *soit aussi un sous-groupe de* G, *il suffit que, quels que soient les indices* $i, j \in I$, *il existe un indice* $k \in I$ *tel que l'on ait*

$$H_i,\ H_j \subset H_k.$$

Soit M l'intersection des H_i; elle n'est pas vide (puisque l'élément neutre de G appartient à tous les H_i, donc aussi à M); si M contient deux éléments x et y, ceux-ci appartiennent à H_i pour tout i, de sorte qu'il en est de même de xy^{-1}, qui appartient donc aussi à M; donc M est un sous-groupe de G.

Soit maintenant U la réunion des H_i; elle est évidemment non vide; soient x, y deux éléments de U; il existe des indices i, $j \in I$ tels que l'on ait $x \in H_i$, $y \in H_j$; d'après l'hypothèse faite dans l'énoncé, il existe donc un indice k tel que H_k contienne *à la fois* x et y, donc aussi xy^{-1}; il s'ensuit que xy^{-1} appartient à la réunion U, ce qui achève la démonstration.

Soit B une partie d'un groupe G; il existe des sous-groupes de G qui contiennent B (par exemple, G lui-même); l'intersection de tous ces sous-groupes est encore un sous-groupe d'après le Théorème 1, et contient encore B, tout en étant contenue, par construction même, dans tout sous-groupe de G contenant B. Ce sous-groupe intersection est donc le « plus petit » de tous les sous-groupes de G contenant B; on dit que c'est le **sous-groupe de G engendré par B**.

Supposons par exemple B réduit à un seul élément x; un sous-groupe contenant x contient évidemment toutes les puissances de x (définies dans l'Exemple 9 ci-dessus);

or celles-ci forment un sous-groupe de G contenant x et donc aussi B. On voit donc qu'ici le plus petit sous-groupe de G contenant B est le sous-groupe formé par les puissances de x, i.e. le sous-groupe de G engendré par x au sens de l'*Exemple* 9 ci-dessus.

Dans le cas d'une partie quelconque B de G, on peut construire le sous-groupe engendré par B par une méthode analogue à celle de l'*Exemple* 9 :

THÉORÈME 2. *Soit* B *une partie d'un groupe* G. *Pour qu'un* $x \in G$ *appartienne au sous-groupe de* G *engendré par* B, *il faut et il suffit qu'il existe un entier* $p \geqslant 0$ *et des éléments* $x_1, \cdots, x_p \in G$ *possédant les propriétés suivantes :*

a) on a la relation

$$x = x_1 \cdots x_p;$$

b) pour chaque i $(1 \leqslant i \leqslant p)$ *on a soit* $x_i \in B$, *soit* $x_i^{-1} \in B$.

Remarque 3. Pour $p = 0$ on doit par convention interpréter la relation figurant dans l'assertion *a)* de l'énoncé comme signifiant $x = e$ (d'une manière générale dans un groupe on convient d'attribuer un sens à la notion de **produit vide** ou **produit de zéro facteur** en déclarant qu'un tel produit n'est autre que l'élément neutre du groupe. Cette convention est nécessaire pour assurer la validité de certains énoncés).

Pour démontrer le Théorème 2, considérons l'ensemble H des $x \in G$ qui satisfont aux conditions de l'énoncé : tout revient à prouver que H est un sous-groupe, contient B, et est contenu dans tout sous-groupe contenant B.

La dernière de ces trois assertions est évidente : si un sous-groupe contient B, il contient évidemment les x_i de l'assertion *b)*, donc l'élément x figurant dans l'assertion *a)* de l'énoncé.

Le fait que H contienne B est non moins clair : un élément x de B vérifie en effet les conditions *a)* et *b)*, comme on le voit en prenant $p = 1$ et $x_1 = x$.

Il reste à prouver que H est un sous-groupe. Tout d'abord, H contient l'élément neutre d'après la *Remarque* 3 ci-dessus. Soient maintenant x et y deux éléments de H; on peut donc écrire

$$x = x_1 \cdots x_p, \qquad y = y_1 \cdots y_q,$$

avec

$$x_i \in B \text{ ou } x_i^{-1} \in B \quad \text{pour tout } i$$
$$y_j \in B \text{ ou } y_j^{-1} \in B \quad \text{pour tout } j;$$

on a alors

$$x y^{-1} = (x_1 \cdots x_p) . (y_1 \cdots y_q)^{-1} = x_1 \cdots x_p y_q^{-1} \cdots y_1^{-1}$$

d'après le § 6, Théorème 3, et on a ainsi décomposé l'élément $x y^{-1}$ de G en un produit

$$x y^{-1} = z_1 \cdots z_{p+q}$$

avec

$$z_k \in B \text{ ou } z_k^{-1} \in B \quad \text{pour tout } k,$$

ce qui prouve que $xy^{-1} \in H$. Par suite, H est un sous-groupe de G, et le Théorème est démontré.

Lorsque le sous-groupe de G engendré par une partie B de G est G tout entier, on dit que B est un **ensemble de générateurs de** G. Si G admet un ensemble fini de générateurs (i.e. s'il existe une partie *finie* B de G qui engendre G), on dit que G est un **groupe à engendrement fini**, ou un **groupe de type fini** — il est clair par exemple que tout groupe cyclique est de type fini.

Soient G un groupe *commutatif de type fini*, et $B = \{a_1, \cdots, a_n\}$ un ensemble fini de générateurs de G. Appliquons le Théorème 2 : tout $x \in G$ admet alors une décomposition

$$x = x_1 \cdots x_p, \quad \text{avec} \quad x_i \in B \quad \text{ou} \quad x_i^{-1} \in B \quad \text{pour tout } i;$$

chacun des facteurs de cette décomposition est donc soit l'un des a_j, soit l'un des éléments a_j^{-1}. Mais comme G est commutatif, on peut grouper ensemble tous ceux des x_i qui, pour un indice j donné, sont égaux soit à a_j soit à a_j^{-1}; le produit de ces x_i est évidemment une puissance de a_j, et finalement on a une décomposition de x de la forme

$$x = a_1^{r_1} \ldots a_n^{r_n}$$

avec des entiers rationnels $r_1, \ldots, r_n$.

Il est clair inversement que si tout $x \in G$ peut s'écrire sous la forme précédente, alors G est de type fini et engendré par les éléments $a_1, \ldots, a_n$.

Exemple 10. *Le groupe additif* $\mathbf{Z}^n$ *est de type fini*, et admet pour ensemble de générateurs les éléments

$$e_1 = (1, 0, \ldots, 0), \qquad e_2 = (0, 1, 0, \ldots, 0), \qquad e_n = (0, \ldots, 0, 1)$$

de ce groupe. En effet, si $r_1, r_2, \ldots, r_n$ sont des entiers quelconques, on a immédiatement les formules

$$\begin{aligned} r_1 e_1 &= (r_1, 0, \ldots, 0) \\ r_2 e_2 &= (0, r_2, \ldots, 0) \\ &\cdots\cdots\cdots\cdots \\ r_n e_n &= (0, 0, \ldots, r_n) \end{aligned}$$

et par suite

$$r_1 e_1 + r_2 e_2 + \cdots + r_n e_n = (r_1, r_2, \ldots, r_n),$$

ce qui montre bien que tout élément de $\mathbf{Z}^n$ est un produit (i.e. une somme) de puissances (i.e. de multiples) des éléments $e_1, \ldots, e_n$.

Par contre, *le groupe additif* $\mathbf{Q}$ *des nombres rationnels n'est pas de type fini*. Supposons-le en effet engendré par des nombres rationnels

$$a_1 = p_1/q_1, \ \ldots, \ a_n = p_n/q_n$$

en nombre fini; cela voudrait dire que, pour *tout* nombre rationnel x, il existe des *entiers* $r_1, \ldots, r_n$ tels que

$$x = r_1 a_1 + \cdots + r_n a_n;$$

mais il est clair qu'alors on pourrait écrire x sous forme d'une fraction ayant pour dénominateur $q_1 \ldots q_n$ (ou plus généralement n'importe quel dénominateur commun à $a_1, \ldots, a_n$); autrement dit, il serait possible de « réduire au même dénominateur » *tous* les nombres rationnels *à la fois*, ce qui est visiblement (*) absurde !

5. *Permutations et transpositions*

Considérons le groupe $\mathfrak{S}_n$ des permutations de l'ensemble

$$\mathrm{I}_n = \{1, 2, \ldots, n\};$$

on dit qu'une permutation $t \in \mathfrak{S}_n$ est une **transposition** s'il existe un entier i, vérifiant $1 \leqslant i \leqslant n-1$, tel qu'on ait les relations suivantes :

$$t(i) = i+1, \qquad t(i+1) = i, \qquad t(k) = k \quad \text{pour} \quad k \neq i,\ i+1.$$

THÉORÈME 3. *Le groupe $\mathfrak{S}_n$ est engendré par les transpositions qu'il contient.*

On va en fait montrer que toute permutation $s \in \mathfrak{S}_n$ est un produit de transpositions, en raisonnant par récurrence sur n (le cas $n = 1$ est trivial puisque le groupe $\mathfrak{S}_n$ se réduit alors à son élément neutre).

Considérons donc une permutation $s \in \mathfrak{S}_n$, et posons $s(n) = i$. Désignant par t_j la transposition qui échange j et $j+1$, il est clair que la permutation

$$u = t_{n-1} \circ \cdots \circ t_i \circ s$$

vérifie $u(n) = n$, et que l'on a

$$s = t_i^{-1} \circ \cdots \circ t_{n-1}^{-1} \circ u = t_i \circ \cdots \circ t_{n-1} \circ u$$

en vertu du fait que

$$t^{-1} = t$$

pour toute transposition. Pour montrer que s est un produit de transpositions, il suffit donc de l'établir pour u, i.e. pour une permutation vérifiant $u(n) = n$.

Mais cette relation montre que u permute les éléments $1, 2, \ldots, n-1$ de I_n, autrement dit que u « induit » dans I_{n-1} une permutation $u' \in \mathfrak{S}_{n-1}$; celle-ci, d'après l'hypothèse de récurrence, peut s'écrire

$$u' = v_1 \circ \cdots \circ v_q$$

où $v_1, \ldots, v_q$ sont des transpositions dans le groupe $\mathfrak{S}_{n-1}$. Définissons alors des permutations $w_1, \ldots, w_q$ de I_n en posant

$$w_j(x) = \begin{cases} v_j(x) & \text{si} \quad x \in \mathrm{I}_{n-1} \\ n & \text{si} \quad x = n; \end{cases}$$

(*) On conseille néanmoins au lecteur de *démontrer* cette assertion.

comme u et u' coïncident sur I_{n-1}, et comme $u(n) = n$, il est clair que

$$u = w_1 \circ \cdots \circ w_q;$$

mais comme les v_j sont des transpositions de I_{n-1}, les w_j sont évidemment des transpositions de I_n. Donc, dans le groupe $\mathfrak{S}_n$, la permutation u est un produit de transpositions, ce qui achève la démonstration.

6. *Classes modulo un sous-groupe*

Soient G un groupe et H un sous-groupe de G; alors *la relation*

$$R\{x, y\} : x^{-1}y \in H$$

est une relation d'équivalence sur l'ensemble G au sens du § 4, n° 1. Il est clair, tout d'abord, que la relation $R\{x, x\}$ est toujours vraie, puisqu'elle signifie que H contient l'élément neutre de G; d'autre part, pour montrer que $R\{x, y\}$ implique $R\{y, x\}$, on observe que, par définition d'un sous-groupe, la relation

$$x^{-1}y \in H \qquad \text{implique} \qquad (x^{-1}y)^{-1} \in H, \quad \text{i.e. } y^{-1}x \in H;$$

enfin, des relations $R\{x, y\}$ et $R\{y, z\}$, i.e. des relations

$$x^{-1}y \in H \qquad \text{et} \qquad y^{-1}z \in H,$$

résulte par définition d'un sous-groupe la relation

$$(x^{-1}y)\,(y^{-1}z) \in H \quad \text{i.e. } x^{-1}z \in H,$$

i.e. la relation $R\{x, z\}$.

La relation considérée est donc bien une relation d'équivalence sur l'ensemble G. Nous allons construire les classes d'équivalence F_x correspondantes (§ 4, n° 2). Pour $x \in G$, l'ensemble F_x est par définition formé des $y \in G$ tels que la relation $R\{x, y\}$ soit vraie, autrement dit des y tels que l'on ait $x^{-1}y \in H$; posant $x^{-1}y = z$ il vient $y = xz$, et dire que $R\{x, y\}$ est vraie signifie que $z \in H$. Ainsi, F_x *est l'ensemble des éléments de* G *de la forme* xz *avec* $z \in H$; pour cette raison, on utilise au lieu de F_x la notation

$$xH$$

et on dit que l'ensemble xH est une **classe à droite modulo** H (on définit de même les **classes à gauche modulo** H : ce sont les parties de G de la forme Hx, où Hx désigne l'ensemble des éléments de la forme zx, avec $z \in H$). L'ensemble des classes xH (resp. Hx) modulo H, i.e. le quotient de l'ensemble G par la relation d'équivalence $x^{-1}y \in H$ (resp. $yx^{-1} \in H$), se note G/H (resp. H\G).

Exemple 11. Prenons pour G le groupe additif $\mathbf{Z}$ des entiers rationnels et pour H le sous-groupe $p\mathbf{Z}$ formé des multiples d'un entier p donné; alors la relation

R $\{x, y\}$ s'écrit (on est maintenant en notation additive)

$$y - x \in p\mathbf{Z} \quad \text{i.e.} \quad x \equiv y \pmod{p},$$

et on retrouve l'*Exemple* 4 du § 4, les classes modulo le sous-groupe $p\mathbf{Z}$ étant ici par conséquent les classes de congruence modulo p définies au § 4, *Exemple* 9.

On notera à ce sujet qu'au § 4 on a défini une « addition » sur l'ensemble $\mathbf{Z}/p\mathbf{Z}$, laquelle vérifie la relation

$$\theta(x + y) = \theta(x) + \theta(y)$$

quels que soient $x, y \in \mathbf{Z}$ (on note θ l'application canonique de $\mathbf{Z}$ sur $\mathbf{Z}/p\mathbf{Z}$). On peut facilement montrer (cf. § 8, n° 3) que l'ensemble $\mathbf{Z}/p\mathbf{Z}$, muni de cette loi de composition, est un *groupe* (groupe additif des entiers modulo p). Voir une construction plus générale dans l'*Exercice* 16 de ce §.

On remarquera que, pour toute classe xH modulo H, il existe une *bijection* de H sur xH, à savoir l'application $z \mapsto xz$ (cette application est surjective par définition des classes, et injective parce que tout $x \in$ G est inversible, de sorte que le Théorème 4 du § 6 s'applique). Si en particulier G est un groupe fini, auquel cas il en est de même de H, on voit donc que chaque classe xH comporte autant d'éléments que H. Or les classes xH forment une partition de l'ensemble G; donc (§ 5, Théorème 7) le nombre d'éléments de G est égal au nombre d'éléments de H multiplié par le nombre de classes xH distinctes. Par suite :

THÉORÈME 4. *Soient* G *un groupe fini et* H *un sous-groupe de* G. *On a alors la relation*

$$\text{Card (G)} = \text{Card (G/H)} . \text{Card (H)}.$$

Remarque 4. Le nombre Card (G) d'éléments d'un groupe fini G s'appelle traditionnellement **l'ordre de** G (donc un « groupe fini d'ordre 5 » n'est autre qu'un groupe possédant 5 éléments). D'autre part, le nombre Card (G/H), noté aussi parfois (G : H), et qui figure dans l'énoncé du Théorème 4 s'appelle **l'indice de** H **dans** G; on peut montrer facilement qu'il est aussi égal au nombre de classes Hx distinctes dans G.

Le Théorème 4 montre, en particulier, que *l'ordre de* H *est un diviseur de l'ordre de* G. On va donner une importante application de ce résultat.

THÉORÈME 5. *Soit* G *un groupe fini d'ordre n. On a alors*

$$x^n = e$$

pour tout $x \in$ G.

Soit en effet H le sous-groupe de G engendré par x, et soit r = Card (H); comme n est un multiple de r, il suffit pour établir le Théorème de prouver que $x^r = e$. Autrement dit, il suffit de prouver le Théorème 5 dans le cas où G est engendré par x, ce que nous supposerons donc dans ce qui suit.

Considérons alors l'application f : $\mathbf{Z} \to G$ donnée par

$$f(q) = x^q;$$

elle est *surjective* par hypothèse. Les règles de calcul sur les puissances (*Exemple* 9) montrent qu'on a les relations

$$f(0) = e, \qquad f(q' - q'') = f(q')f(q'')^{-1};$$

de ces relations résulte aussitôt que les $q \in \mathbf{Z}$ tels que $f(q) = e$ forment un *sous-groupe* de $\mathbf{Z}$, donc de la forme $s\mathbf{Z}$ où s est un entier positif bien déterminé.

De plus, la relation $f(q') = f(q'')$, qui équivaut évidemment à

$$f(q')f(q'')^{-1} = e,$$

s'écrit aussi d'après ce qui précède sous la forme $f(q' - q'') = e$, et équivaut donc à

$$q' - q'' \in s\mathbf{Z}, \quad \text{i.e. à } q' \equiv q'' \pmod{s}.$$

Comme f est *surjective*, G possède donc autant d'éléments qu'il y a de classes modulo s dans $\mathbf{Z}$; autrement dit, s n'est autre que le nombre d'éléments de G, et comme on a $x^s = e$ le Théorème est démontré.

7. *Nombre de permutations de n objets*

On a établi, au § 5 (Corollaire du Théorème 9), le résultat suivant :

THÉORÈME. *Soit* X *un ensemble fini à n éléments. Alors le groupe* $\mathfrak{S}(X)$ *des permutations de* X *est d'ordre*

$$n! = 1.2\ldots n.$$

Nous allons donner ici de ce résultat une démonstration qui, sans différer essentiellement de celle du § 5, fait plus systématiquement usage de la structure de groupe existant sur l'ensemble $\mathfrak{S}(X)$.

Le Théorème est clair si $n = 1$, et on va le démontrer par récurrence sur n, autrement dit prouver que s'il est vrai pour l'entier $n - 1$ il l'est aussi pour l'entier n.

Choisissons pour cela une fois pour toutes un élément a de X, et soit

$$Y = X - \{a\}$$

l'ensemble obtenu en ôtant de X l'élément a; Y est un ensemble à $n - 1$ éléments, auquel le Théorème 6 est donc applicable (hypothèse de récurrence).

D'autre part, on peut considérer le groupe $\mathfrak{S}(Y)$ comme un sous-groupe de $\mathfrak{S}(X)$; il suffit pour cela d'associer à toute permutation s de l'ensemble Y la permutation $\bar{s}$ de X donnée par

$$\bar{s}(x) = \begin{cases} s(x) & \text{si } x \in Y \\ a & \text{si } x = a; \end{cases}$$

de cette façon, $\mathfrak{S}(Y)$ *s'identifie au sous-groupe de* $\mathfrak{S}(X)$ *formé des permutations de* X *qui admettent a pour point fixe.*

D'après l'hypothèse de récurrence, le groupe $\mathfrak{S}(Y)$ possède $(n-1)!$ éléments; pour en déduire que $\mathfrak{S}(X)$ en possède $n!$, i.e. n fois plus, il suffit donc (Théorème 4) de prouver que, dans $\mathfrak{S}(X)$, les classes modulo $\mathfrak{S}(Y)$ sont au nombre de n exactement.

Pour cela, introduisons l'application $f: \mathfrak{S}(X) \to X$ donnée par

$$f(s) = s(a) \quad \text{pour tout} \quad s \in \mathfrak{S}(X).$$

Étant données des permutations s et t de X, la relation $f(s) = f(t)$ s'écrit

$$s(a) = t(a), \quad \text{i.e.} \quad a = s^{-1}t(a),$$

et par suite signifie que

$$s^{-1}t \in \mathfrak{S}(Y),$$

i.e. que s et t appartiennent à la même classe à droite modulo le sous-groupe $\mathfrak{S}(Y)$. Utilisant le Théorème 2 du § 4 on voit donc que le nombre de classes modulo $\mathfrak{S}(Y)$ est égal au nombre d'éléments de l'image de $\mathfrak{S}(X)$ par f, i.e. au nombre d'éléments $x \in X$ pour lesquels il existe une permutation s de X telle que $x = s(a)$; mais il est clair que *tout* $x \in X$ peut s'écrire $x = s(a)$ pour une permutation convenable s; par suite, les classes modulo $\mathfrak{S}(Y)$ dans $\mathfrak{S}(X)$ sont au nombre de n, et ceci termine la démonstration du Théorème.

8. *Homomorphismes de groupes*

Étant donnés des groupes G et H, on appelle **homomorphisme de G dans H** toute application f de G dans H telle que l'on ait

$$f(xy) = f(x)f(y) \quad \text{quels que soient} \quad x, y \in G.$$

Faisant $y = e$ dans la relation précédente, on trouve $f(x) = f(x)f(e)$ et par suite

$$f(e) = e.$$

Enfin, en prenant $y = x^{-1}$, et en tenant compte du résultat qu'on vient d'obtenir, on trouve évidemment que

$$f(x^{-1}) = f(x)^{-1} \quad \text{pour tout} \quad x \in G.$$

Remarque 5. La définition donnée plus haut suppose les groupes G et H écrits multiplicativement, et doit être modifiée en conséquence si G ou H ou G et H sont notés additivement. Par exemple, si G est noté additivement et H multiplicativement, un homomorphisme est une application f de G dans H vérifiant

$$f(x+y) = f(x)f(y) \quad \text{quels que soient} \quad x, y \in G.$$

Exemple 12. Prenons pour G le groupe additif **Z** des entiers rationnels et pour H un groupe (multiplicatif) arbitraire. Pour tout $a \in H$, l'application f donnée

par

$$f(n) = a^n \quad \text{pour tout} \quad n \in \mathbf{Z}$$

est un homomorphisme : cela résulte des formules de l'*Exemple* 9. En outre, *tout* homomorphisme f de $\mathbf{Z}$ dans H s'obtient par la méthode en question. En effet, pour un tel homomorphisme, posons

$$f(1) = a;$$

on a alors

$$f(2) = f(1+1) = f(1)f(1) = aa = a^2,$$
$$f(3) = f(2+1) = f(2)f(1) = a^2a = a^3,$$

etc, d'où $f(n) = a^n$ pour n positif, puis, pour n négatif,

$$f(n) = f(-n)^{-1} = (a^{-n})^{-1} = a^n,$$

de sorte que la relation $f(n) = a^n$ est valable pour tout $n \in \mathbf{Z}$.

Exemple 13. Pour tout entier p, l'application canonique de $\mathbf{Z}$ sur $\mathbf{Z}/p\mathbf{Z}$ est un homomorphisme du groupe additif $\mathbf{Z}$ sur le groupe additif $\mathbf{Z}/p\mathbf{Z}$.

Exemple 14. La formule

$$\log(xy) = \log(x) + \log(y)$$

montre que la fonction logarithmique, définie en Analyse, est un homomorphisme du groupe multiplicatif $\mathbf{R}_+^*$ dans le groupe additif $\mathbf{R}$.

THÉORÈME 6. *Soient* $f : M \to N$ *et* $g : N \to P$ *des homomorphismes de groupes; alors l'application composée* $g \circ f : M \to P$ *est encore un homomorphisme. Si un homomorphisme de groupes* $f : M \to N$ *est bijectif, l'application réciproque* $f^{-1} : N \to M$ *est encore un homomorphisme.*

Pour établir la première assertion, il suffit d'observer que, pour $x, y \in M$, on a

$$g \circ f(xy) = g[f(xy)] = g[f(x)f(y)] = g[f(x)] \cdot g[f(y)].$$

Pour établir la seconde, autrement dit que

$$f^{-1}(uv) = f^{-1}(u)f^{-1}(v)$$

pour $u, v \in N$, on remarque qu'il suffit (puisque f est bijectif, et en particulier injectif) d'établir que les images par f des deux membres de cette relation sont égales. Autrement dit, tout revient à prouver que

$$f[f^{-1}(uv)] = f[f^{-1}(u)f^{-1}(v)];$$

or le premier membre est égal à uv par définition de f^{-1}; le second membre, puisque f est un homomorphisme, est égal à

$$f[f^{-1}(u)] \cdot f[f^{-1}(v)] = uv,$$

ce qui termine la démonstration.

Soient G et H deux groupes; on appelle **isomorphisme de G sur H** tout homomorphisme *bijectif* de G sur H, et on dit que G et H sont **isomorphes** lorsqu'il existe un isomorphisme de G sur H.

Exemple 15. En utilisant comme dans l'*Exemple* 14 la fonction logarithmique, on voit que les groupes $\mathbf{R}_+^*$ et $\mathbf{R}$ sont isomorphes.

La relation

G et H sont isomorphes

est une *relation d'équivalence*; en effet, G et G sont isomorphes quel que soit G, car l'application identique est évidemment un isomorphisme de G sur G; d'autre part, s'il existe un isomorphisme f d'un groupe G sur un groupe H, alors il existe aussi un isomorphisme de H sur G, à savoir f^{-1}; enfin ,s'il existe un isomorphisme f d'un groupe M sur un groupe N, et un isomorphisme g de N sur un troisième groupe P, il existe aussi un isomorphisme de M sur P — à savoir $g \circ f$, qui est un homomorphisme d'après le Théorème 6, et est bijectif puisque f et g le sont.

On appelle **automorphisme d'un groupe** G tout isomorphisme de G sur G.

Exemple 16. Soit G un groupe noté multiplicativement; alors, pour tout $a \in G$, l'application f de G dans G donnée par

$$f(x) = axa^{-1}$$

est un automorphisme de G. On a en effet

$$\begin{aligned} f(x)f(y) &= (axa^{-1})(ay a^{-1}) = (ax)(aa^{-1})(ya^{-1}) \\ &= (ax)e(ya^{-1}) = (ax)(ya^{-1}) = a(xy)a^{-1} = f(xy), \end{aligned}$$

de sorte que f est un homomorphisme; de plus, pour tout $y \in G$, l'équation $axa^{-1} = y$ admet une et une seule solution $x = a^{-1}ya$, ce qui montre que f est bijectif.

Les automorphismes de G obtenus par la méthode qu'on vient de décrire s'appellent les **automorphismes intérieurs** du groupe G. Cette notion n'a évidemment d'intérêt que pour les groupes non commutatifs.

Exemple 17. Considérons le groupe multiplicatif $\mathbf{R}_+^*$ des nombres réels strictement positifs; alors, pour tout nombre réel α non nul la fonction

$$f(x) = x^\alpha,$$

définie en Analyse, est un automorphisme du groupe $\mathbf{R}_+^*$; l'automorphisme réciproque est

$$f^{-1}(x) = x^{1/\alpha}.$$

Remarque 6. Dans la pratique, on considère souvent deux groupes isomorphes G et H comme identiques; plus exactement, et dans la mesure où l'on se place au point de vue de la pure théorie des groupes, G et H possèdent exactement les mêmes propriétés; par exemple, si G est commutatif, il en est de même de

H; si G est engendré par n éléments, il en est de même de H; et d'une manière générale, une fois qu'on a choisi un isomorphisme f de G sur H, on peut « traduire » toute relation entre éléments de G en une relation analogue entre les éléments de H obtenus en appliquant f aux éléments considérés de G.

On notera que la fonction logarithmique, isomorphisme du groupe multiplicatif $\mathbf{R}_+^*$ sur le groupe additif $\mathbf{R}$, a justement été inventée pour transformer toute relation multiplicative entre nombres positifs en une relation additive entre nombres de signe quelconque...

9. *Noyau et image d'un homomorphisme*

Établissons d'abord le résultat suivant :

THÉORÈME 7. *Soit f un homomorphisme d'un groupe* G *dans un groupe* H. *L'image par f de tout sous-groupe de* G *est un sous-groupe de* H. *L'image réciproque par f de tout sous-groupe de* H *est un sous-groupe de* G.

Soient G' un sous-groupe de G et $H' = f(G')$ son image; si $u, v \in H'$, il existe $x, y \in G'$ tels que $u = f(x)$, $v = f(y)$; on a alors

$$uv^{-1} = f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}),$$

et comme $xy^{-1} \in G'$ il s'ensuit que $uv^{-1} \in H'$; ceci établit la première assertion de l'énoncé. La seconde se démontre par des raisonnements analogues, qu'on laisse au lecteur le soin de détailler.

Il résulte du Théorème 7 que, si f est un homomorphisme de G dans H, alors $f(G)$ est un sous-groupe de H; on l'appelle l'**image de** f, et on le note

$$\mathrm{Im}\,(f);$$

de même, l'ensemble $f^{-1}(\{e\})$, formé des $x \in G$ tels que

$$f(x) = e,$$

est un sous-groupe de G; on l'appelle le **noyau de** f, et on le désigne par la notation

$$\mathrm{Ker}\,(f)$$

(le mot « noyau » se traduit par « kernel » en anglais, et par « Kern » en allemand).

Exemple 18. Soient G un groupe multiplicatif, a un élément de G, et considérons l'homomorphisme $f: \mathbf{Z} \to G$ donné par $f(n) = a^n$. Alors l'image de f est le sous-groupe de G engendré par a; et le noyau de f est le sous-groupe de $\mathbf{Z}$ formé des entiers n tels que $a^n = e$ (le fait que ces entiers forment un sous-groupe a déjà été établi et utilisé dans la démonstration du Théorème 5).

Remarque 7. Soit N le noyau d'un homomorphisme $f: G \to H$; pour $x \in N$ et $a \in G$ on a

$$f(axa^{-1}) = f(a)f(x)f(a)^{-1} = f(a)ef(a)^{-1} = f(a)f(a)^{-1} = e;$$

par suite, on a

$$axa^{-1} \in N \text{ quels que soient } a \in G \text{ et } x \in N;$$

un sous-groupe d'un groupe G est dit **invariant** (ou **normal**, ou **distingué**) lorsqu'il possède la propriété précédente; cela signifie évidemment qu'on a

$$s(N) \subset N$$

pour tout automorphisme intérieur s de G.

On voit donc que *le noyau d'un homomorphisme est un sous-groupe invariant* Réciproquement, on peut démontrer que, si N est un sous-groupe invariant d'un groupe G, il existe un groupe H et un homomorphisme f de G dans H tel que N soit le noyau de f; cf. *Exercice* 16.

Lorsque G est commutatif, il va de soi que tout sous-groupe de G est invariant.

THÉORÈME 8. *Soient* G *et* H *deux groupes et* f *un homomorphisme de* G *dans* H. *Pour que* f *soit injectif, il faut et il suffit que son noyau soit réduit à l'élément neutre.*

Comme $f(e) = e$, la relation $f(x) = e$ signifie que $f(x) = f(e)$; si f est injectif elle implique donc $x = e$, autrement dit $\mathrm{Ker}(f) = \{e\}$. Supposons inversement le noyau de f réduit à e; la relation $f(x) = f(y)$ s'écrit encore

$$f(x) f(y)^{-1} = e,$$

ou, puisque f est un homomorphisme, $f(xy^{-1}) = e$, et par suite signifie que

$$xy^{-1} \in \mathrm{Ker}(f);$$

comme $\mathrm{Ker}(f) = \{e\}$, il vient donc $xy^{-1} = e$, autrement dit $x = y$, et f est injectif, ce qui achève la démonstration.

THÉORÈME 9. *Soient* G, H *et* M *trois groupes,* $p : G \to H$ *et* $f : G \to M$ *des homomorphismes; on suppose* p *surjectif. Les conditions suivantes sont alors équivalentes :*

a) il existe un homomorphisme $f' : H \to M$ *tel que* $f = f' \circ p$;

b) on a la relation $\mathrm{Ker}(p) \subset \mathrm{Ker}(f)$.

Si ces conditions sont réalisées, l'homomorphisme f' *est unique; il est injectif si et seulement si* $\mathrm{Ker}(p) = \mathrm{Ker}(f)$, *et surjectif si et seulement si* f *est surjectif.*

Cherchons d'abord à quelle condition il existe une application f' de H dans M telle que $f = f' \circ p$; la réponse est donnée par le Théorème 1 du § 2 : tout revient à vérifier que la relation $p(x) = p(y)$ implique la relation $f(x) = f(y)$. Or, comme p est un homomorphisme, la première s'écrit

$$e = p(x)p(y)^{-1} = p(xy^{-1}),$$

autrement dit $xy^{-1} \in \mathrm{Ker}(p)$; et pour la même raison la seconde relation s'écrit $xy^{-1} \in \mathrm{Ker}(f)$; prenant $y = e$ on voit que la relation $x \in \mathrm{Ker}(p)$ doit impliquer la

relation $x \in \mathrm{Ker}(f)$, ce qui exige $\mathrm{Ker}(p) \subset \mathrm{Ker}(f)$, et cette condition est évidemment suffisante.

Ainsi, la condition *b*) équivaut à l'existence d'une application f' telle que $f = f' \circ p$. Cette application est nécessairement un homomorphisme; soient en effet $u, v \in \mathrm{H}$; puisque p est surjectif, on peut écrire $u = p(x)$, $v = p(y)$ avec $x, y \in \mathrm{G}$; alors

$$f'(uv) = f'(p(x)p(y)) = f'(p(xy)) = f(xy) = f(x)f(y)$$
$$= f'(p(x))\,f'(p(y)) = f'(u)\,f'(v),$$

ce qui établit notre assertion.

L'équivalence des conditions *a*) et *b*) est donc établie.

L'unicité de f' est évidente; car, p étant surjectif, la relation

$$f'_1 \circ p = f'_2 \circ p \quad \text{implique} \quad f'_1 = f'_2.$$

Il est non moins clair que

$$f'(\mathrm{H}) = f'(p(\mathrm{G})) = f(\mathrm{G}),$$

et par suite que f' est surjective si et seulement si f l'est. Enfin, cherchons le noyau de f'; il est formé des $u \in \mathrm{H}$ tels que $f'(u) = e$; posant $u = p(x)$, cela s'écrit encore $f(x) = e$, autrement dit $x \in \mathrm{Ker}(f)$; par suite, on a

$$\mathrm{Ker}(f') = p[\mathrm{Ker}(f)].$$

Pour que f' soit injective, il est donc nécessaire et suffisant (Théorème 8) que $p[\mathrm{Ker}(f)] = \{e\}$, autrement dit que $\mathrm{Ker}(f) \subset \mathrm{Ker}(p)$, autrement dit que

$$\mathrm{Ker}(f) = \mathrm{Ker}(p)$$

puisque la relation $\mathrm{Ker}(p) \subset \mathrm{Ker}(f)$ est déjà vérifiée. Ceci termine la démonstration.

10. *Application aux groupes cycliques*

Soient G un groupe cyclique et x un générateur de G: tout élément de G est donc une puissance de x. Autrement dit l'homomorphisme

$$f : \mathbf{Z} \to \mathrm{G}$$

donné par

$$f(n) = x^n$$

est *surjectif*.

Désignons son noyau par I; c'est un sous-groupe de **Z**, par suite il existe un et un seul entier $p \geqslant 0$ tel que

$$\mathrm{I} = p\mathbf{Z}$$

(cf. *Exemple* 8). Distinguons deux cas.

Tout d'abord, il peut arriver que $p = 0$; alors (Théorème 8) f est *injectif*, donc bijectif, et par suite est un *isomorphisme du groupe additif* **Z** *sur* G.

Supposons maintenant $p \neq 0$; considérons le groupe additif $\mathbf{Z}/p\mathbf{Z}$ (*Exemple* 11) et l'application canonique g de $\mathbf{Z}$ sur $\mathbf{Z}/p\mathbf{Z}$; c'est un homomorphisme surjectif, ayant pour noyau $p\mathbf{Z}$, de sorte que $\mathrm{Ker}(g) = \mathrm{Ker}(f)$. D'après le Théorème 9, il existe donc un et un seul homomorphisme

$$f' : \mathbf{Z}/p\mathbf{Z} \to \mathrm{G}$$

tel que $f = f' \circ g$ [ce qui signifie que, pour tout entier n, x^n est l'image par f' de la classe de n modulo p; l'existence de f' provient du fait que la relation

$$m \equiv n \pmod{p} \quad \text{implique} \quad x^m = x^n,$$

en sorte que x^n dépend, non pas de l'entier n, mais seulement de sa classe modulo p; bien entendu, ce raisonnement n'est autre qu'une traduction, dans le cas particulier qui nous intéresse, du raisonnement général utilisé dans la démonstration du Théorème 9]; comme f est surjectif, f' est surjectif; comme $\mathrm{Ker}(f) = \mathrm{Ker}(g)$, f' est injectif; par suite, f' est bijectif, et G *est isomorphe au groupe additif* $\mathbf{Z}/p\mathbf{Z}$ *des entiers modulo* p. En particulier, G a le même nombre d'éléments que celui-ci, i.e. a p éléments, ce qui caractérise l'entier p : c'est le nombre d'éléments de G. Donc :

THÉORÈME 10. *Tout groupe cyclique infini est isomorphe au groupe additif* $\mathbf{Z}$. *Tout groupe cyclique fini* G *est isomorphe au groupe additif* $\mathbf{Z}/p\mathbf{Z}$, *où* p *est le nombre d'éléments de* G.

On déduit évidemment de là que deux groupes cycliques sont isomorphes si et seulement s'ils ont le même nombre (fini ou non) d'éléments.

Soit x un élément d'un groupe G quelconque; on appelle *ordre de* x l'ordre (ou cardinal) du sous-groupe H de G engendré par x. Comme celui-ci est l'image de $\mathbf{Z}$ par l'homomorphisme $n \to x^n$, on voit qu'une condition nécessaire et suffisante pour que x soit d'ordre fini est qu'il existe un entier p *non nul* tel que

$$x^p = e;$$

l'ordre de x est alors le *plus petit* entier $p \geqslant 1$ vérifiant la relation précédente.

11. *Groupes opérant sur un ensemble*

Soient G un groupe et X un ensemble; on dit que G **opère sur** X si l'on s'est donné une application de $\mathrm{G} \times \mathrm{X}$ dans X, notée

$$(s, x) \mapsto s.x,$$

et vérifiant les deux conditions que voici : on a la relation d'associativité

$$s.(t.x) = (st).x \quad \text{quels que soient} \quad s, t \in \mathrm{G} \quad \text{et} \quad x \in \mathrm{X},$$

et d'autre part

$$e.x = x \quad \text{quel que soit} \quad x \in \mathrm{X},$$

où e désigne bien entendu l'élément neutre de G.

Exemple 19. On peut faire opérer le groupe G sur lui-même de plusieurs

façons, soit à l'aide de l'application

$$(s, x) \to sx,$$

(« translations à gauche »), soit à l'aide de l'application

$$(s, x) \to xs^{-1}$$

(« translations à droite »), soit à l'aide de l'application

$$(s, x) \to sxs^{-1}$$

(« automorphismes intérieurs »).

Exemple 20. Soient G un groupe et H un sous-groupe de G et prenons

$$X = G/H,$$

ensemble des classes xH dans G (n° 6); pour $s \in G$ et $A \in X$, l'ensemble $sA \subset G$ des sa où $a \in A$ est encore une classe modulo H (en effet, si l'on choisit un $x \in A$, alors A est l'ensemble des xh, où $h \in H$, et par suite sA est l'ensemble des sxh; autrement dit, si $A = xH$, on a $sA = (sx)H$, ce qui montre bien que $sA \in X$); ceci permet donc de définir une application de $G \times X$ dans X, à savoir $(s, A) \mapsto sA$; on vérifie alors immédiatement que, grâce à cette construction, G opère sur $X = G/H$.

Exemple 21. Soient E un ensemble et p un entier; prenons

$$X = E^p,$$

ensemble des systèmes $(x_1, \ldots, x_p)$ de p éléments de E, et

$$G = \mathfrak{S}_p,$$

groupe des permutations de l'ensemble $\{1, 2, \ldots, p\}$; pour

$$s \in G, \qquad x = (x_1, \ldots, x_p) \in X,$$

définissons

$$s.x = (x_{s^{-1}(1)}, \ldots, x_{s^{-1}(p)});$$

l'application de $G \times X$ dans X ainsi définie permet de faire opérer G sur X. En effet, soient $s, t \in G$ et $x \in X$, et posons

$$t.x = y = (y_1, \ldots, y_p);$$

on a

$$s.(t.x) = s.y = (y_{s^{-1}(1)}, \ldots, y_{s^{-1}(p)});$$

mais on a d'autre part

$$y = (x_{t^{-1}(1)}, \ldots, x_{t^{-1}(p)})$$

et donc

$$y_i = x_{t^{-1}(i)} \quad \text{pour} \quad 1 \leqslant i \leqslant p;$$

par suite on a

$$s.(t.x) = (z_1, \ldots, z_p)$$

avec

$$z_i = y_{s^{-1}(i)} = x_{t^{-1}(s^{-1}(i))} = x_{(st)^{-1}(i)},$$

ce qui établit la relation $s.(t.x) = (st).x$; la relation $e.x = x$ est évidente.

Exemple 22. Soient X un ensemble et G un groupe de transformations de X (*Exemple* 7), alors l'application $(s, x) \mapsto s(x)$ de $G \times X$ dans X permet de faire opérer G sur X

On notera que, si un groupe G opère sur un ensemble X, alors pour tout $s \in G$ l'application

$$\bar{s} : X \to X$$

donnée par

$$\bar{s}(x) = s.x$$

est *bijective* en vertu du fait que $s^{-1}.(s.x) = (s^{-1}s).x = e.x = x$. Ceci dit, on peut encore interpréter les conditions énoncées au début de ce n° en disant que l'application $s \mapsto \bar{s}$ *est un homomorphisme du groupe* G *sur un groupe de transformations de l'ensemble* X.

Soit G un groupe opérant sur un ensemble X. Pour chaque $x \in X$, les $s \in G$ tels que $s.x = x$ forment évidemment un *sous-groupe* de G; on l'appelle le **stabilisateur de** x **dans** G; d'autre part, on appelle **orbite de** x **par** G l'ensemble (qu'on note fréquemment $G.x$) des éléments de X de la forme $s.x$, $s \in G$.

On trouvera des compléments sur ces notions dans les *Exercices* du présent §.

1. Définition des anneaux, exemples

On appelle **anneau** *un triplet formé d'un ensemble* K *et de deux lois de composition sur* K, *notées* $(x, y) \to x + y$ (« addition » dans K) *et* $(x, y) \mapsto xy$ (« multiplication » dans K), *ces données devant vérifier les conditions suivantes :*
(A 1) : *le couple formé de l'ensemble* K *et de la loi de composition* $(x, y) \mapsto x + y$ *sur* K *est un groupe commutatif;*
(A 2) : *la loi de composition* $(x, y) \to xy$ *est associative et admet un élément neutre* (*);
(A 3) : *quels que soient* $x, y, z \in K$, *on a les relations* (dites de « distributivité de la multiplication par rapport à l'addition »)

$$x(y + z) = xy + xz, \qquad (x + y)z = xz + yz.$$

L'axiome (A 1) des anneaux s'explicite comme suit : on a les identités

$$x + (y + z) = (x + y) + z$$
$$x + y = y + x;$$

il existe d'autre part dans K un élément noté 0, tel que l'on ait

$$x + 0 = x$$

pour tout $x \in K$; enfin, quel que soit $x \in K$, il existe un élément de K, noté $-x$, tel que l'on ait

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

Il s'ensuit que, quels que soient $a, b \in K$, l'équation

$$a + x = b$$

(*) Certains auteurs omettent, dans la définition d'un anneau, d'exiger l'existence d'un élément neutre pour la multiplication; on obtient ainsi une notion d'anneau plus générale que celle du texte; mais la pratique montre, et la théorie le confirme, qu'on peut toujours se limiter, comme nous le ferons dans cet ouvrage, à des anneaux *avec* élément unité.

possède une et une seule solution dans K, à savoir $b + (-a)$; on l'écrit

$$x = b - a.$$

L'axiome (A 2) des anneaux signifie que, dans un anneau, on a l'identité

$$x(yz) = (xy)z,$$

et qu'il existe un élément de K, noté 1, tel que l'on ait

$$x1 = 1x = x$$

pour tout $x \in K$. On dit que 1 est l'*élément unité* de K.

On dit qu'un anneau K est **commutatif** si l'on a

$$xy = yx \quad \text{quels que soient } x, y \in K;$$

on rencontrera dans la théorie des matrices d'importants exemples d'anneaux non commutatifs. On dit que deux éléments x, y d'un anneau quelconque K **commutent** si l'on a $xy = yx$.

Dans un anneau, on a les relations

$$(-1)x = -x, \qquad 0x = 0$$

pour tout x; pour établir la première, il suffit de prouver que $(-1)x + x = 0$; or

$$(-1)x + x = (-1)x + 1x = (-1 + 1)x = 0x,$$

de sorte qu'on est ramené à prouver la seconde relation $0x = 0$. Pour cela on calcule

$$0x + x = 0x + 1x = (0 + 1)x = 1x = x,$$

ce qui montre que $0x = x - x = 0$ comme dans tout groupe additif.

Donnons maintenant des exemples d'anneaux.

Exemple 1. Si l'on munit l'ensemble **Z** des entiers rationnels des deux lois de composition usuelles (addition et multiplication) on obtient évidemment un anneau commutatif; on l'appelle l'**anneau des entiers rationnels**. Les ensembles **Q** (nombres rationnels) et **R** (nombres réels), munis de l'addition et de la multiplication usuelles, sont aussi des anneaux commutatifs.

Exemple 2. Soit K l'ensemble des nombres réels x qui possèdent la propriété suivante : il existe des *entiers rationnels* a, b, c tels que l'on ait

$$x = a + br + cr^2, \qquad \text{où} \qquad r = \sqrt[3]{2}.$$

On vérifie immédiatement que, si $x, y \in K$, alors $x + y$ et xy sont encore dans K, ce qui permet de munir l'ensemble K de deux lois de composition, à savoir l'addition et la multiplication usuelles. Cela dit, l'ensemble K, avec ces deux lois de composition, est un anneau commutatif.

Exemple 3. Soient X un ensemble quelconque, K un anneau quelconque, et désignons par A l'ensemble de toutes les applications de X dans K. Pour $f, g \in A$, définissons comme suit des éléments $f+g$ et fg de A : l'élément $f+g$ sera l'application

$$x \mapsto f(x) + g(x)$$

de X dans K, et fg sera l'application

$$x \mapsto f(x)g(x)$$

de X dans K. Cela dit, l'ensemble A, muni des deux lois de composition

$$(f, g) \mapsto f+g \qquad \text{et} \qquad (f, g) \mapsto fg$$

qu'on vient de définir, est un *anneau* (commutatif si et seulement si K est commutatif). Vérifions par exemple la formule de distributivité

$$f(g+h) = fg + fh;$$

il suffit de montrer que les deux membres (qui sont des applications de X dans K) ont la même valeur en chaque $x \in X$; la valeur du premier membre est $f(x)\,(g(x) + h(x))$, celle du second membre est $f(x)g(x) + f(x)h(x)$; on voit donc que l'axiome de distributivité est vérifié dans A parce qu'il l'est déjà dans K.

Le lecteur débutant fera bien de traiter cet Exemple en détail, et de s'assurer du fait que, pour montrer que A est un anneau, on a effectivement besoin de se servir du fait que K vérifie *tous* les axiomes des anneaux. On comprendra facilement pourquoi en examinant le cas où X est un ensemble à un élément...

L'anneau A qu'on vient de définir s'appelle **l'anneau des applications de l'ensemble X dans l'anneau K.**

Exemple 4. On prend $X = K = \mathbf{R}$ dans l'*Exemple* 3, mais au lieu de considérer toutes les applications de X dans K on se borne à considérer celles qui vérifient certaines conditions de « régularité » données d'avance, par exemple : être continue en un point donné; être continue partout; avoir une dérivée troisième continue partout; etc... Dans chaque cas, on obtient un anneau commutatif.

Exemple 5. On verra plus loin (§ 15) que les matrices carrées à n lignes et n colonnes, à coefficients dans un anneau donné K, forment un nouvel anneau (non commutatif si $n \geqslant 2$, même si K est commutatif) pourvu qu'on définisse l'addition et la multiplication des matrices à l'aide des formules des §§ 14 et 15.

Soit K un anneau; on appelle **sous-anneau de** K toute partie A de K qui vérifie les conditions suivantes : A *est un sous-groupe du groupe additif* K; *les relations* $x \in A$ *et* $y \in A$ *impliquent la relation* $xy \in A$; *on a* $1 \in A$. S'il en est ainsi, il est clair que les applications

$$(x, y) \mapsto x + y \qquad \text{et} \qquad (x, y) \mapsto xy$$

de $K \times K$ dans K appliquent $A \times A$ dans A, donc définissent deux lois de composition sur l'ensemble A. Cela dit, *l'ensemble* A, *muni de ces deux lois de composition, est un anneau.*

En effet, l'ensemble A muni de l'addition est un groupe commutatif en vertu du § 7, n° 3; d'autre part, la multiplication est associative dans K, donc *a fortiori* dans A, et A admet un élément neutre pour la multiplication puisqu'il contient l'élément 1 de K; enfin, les identités de distributivité, étant vérifiées dans K, le sont à plus forte raison dans A.

Il est clair que, dans l'*Exemple* 1 ci-dessus, $\mathbf{Z}$ est un sous-anneau de $\mathbf{Q}$, lui-même sous-anneau de $\mathbf{R}$. D'autre part, les anneaux de l'*Exemple* 4 ci-dessus sont des sous-anneaux de l'anneau de toutes les applications de l'ensemble $X = \mathbf{R}$ dans l'anneau $K = \mathbf{R}$.

On notera que, pour vérifier qu'une partie A d'un anneau K est un sous-anneau de K, il suffit de vérifier les conditions suivantes : *si* A *contient deux éléments* x *et* y *de* K, *il contient aussi leur somme* $x + y$ *et leur produit* xy; *en outre*, A *contient* -1.

En effet, supposons remplies ces conditions; pour $x \in A$, on a alors $-x \in A$ vu que $-x = (-1)x$; mais alors, si A contient x et y, donc aussi x et $-y$, il contient également $x + (-y) = x - y$, ce qui prouve que A est un sous-groupe du groupe additif K, autrement dit vérifie la première condition figurant dans la définition d'un sous-anneau. D'autre part, comme A contient -1, il contient $-(-1) = 1$, donc vérifie la troisième condition figurant dans la définition d'un sous-anneau. La seconde, enfin, est vérifiée par hypothèse.

2. *Anneaux d'intégrité et corps*

Considérons dans un anneau K l'équation

$$ax = b, \tag{1}$$

où a, b sont des éléments donnés, et x un élément « inconnu » de K.

Un premier cas simple est celui où $a = 0$; comme $0x = 0$ pour tout $x \in K$, il est clair qu'alors deux cas seulement sont possibles : ou bien $b = 0$, et alors tout $x \in K$ vérifie l'équation (1); ou bien $b \neq 0$, et alors l'équation (1) n'a aucune solution.

Un second cas très simple est celui où a admet un inverse relativement à la multiplication, i.e. où il existe un (et un seul) élément de K, noté a^{-1}, vérifiant

$$a^{-1}a = aa^{-1} = 1;$$

alors le Théorème 4 du § 6 s'applique : l'équation (1) possède, quel que soit b, une et seule solution

$$x = a^{-1}b.$$

Dans ce cas, on dit que a est un **élément inversible** de K (on dit aussi souvent que a est un élément *unitaire* ou une *unité* de K, mais nous n'utiliserons pas cette terminologie dangereuse).

Reste à examiner, dans la mesure où c'est possible, le cas où a n'est ni nul ni

inversible. Tout d'abord ce cas peut fort bien ne pas se produire — autrement dit, il peut arriver que *tout élément non nul de* K *soit inversible;* on dit alors que K est un **corps** (*). Les anneaux **Q** et **R** sont des corps (**corps des nombres rationnels** et **corps des nombres réels**), par contre l'anneau **Z** n'est pas un corps (pour qu'un $x \in \mathbf{Z}$ soit inversible dans l'anneau **Z**, il faut et il suffit qu'il existe un $y \in \mathbf{Z}$ tel que $xy = 1$; ce n'est évidemment possible que si $x = +1$ ou $x = -1$).

Revenant au cas général, on peut se demander s'il est possible que l'équation (1) admette plusieurs solutions. Si x et y sont deux telles solutions, on aura évidemment $ax = ay$, donc

$$a(x - y) = 0;$$

ceci conduit à introduire la notion suivante : on dit qu'un anneau K est **intègre** ou est un **anneau d'intégrité,** lorsque, pour $u \in \mathrm{K}$ et $v \in \mathrm{K}$, la relation

$$uv = 0 \quad \text{implique} \quad u = 0 \quad \text{ou} \quad v = 0$$

(autrement dit lorsqu'un produit d'éléments de K ne peut être nul sans qu'un au moins des facteurs du produit le soit). L'anneau **Z** est évidemment un anneau d'intégrité. Un corps est nécessairement un anneau d'intégrité, car de la relation $uv = 0$ résulte, si $u \neq 0$, que $v = u^{-1}0 = 0$.

Dans un anneau d'intégrité, l'équation (1), pour $a \neq 0$, possède *au plus* une solution comme le montre le raisonnement ci-dessus. Mais il peut naturellement arriver qu'elle n'en possède aucune — c'est le cas par exemple de l'équation $2x = 3$ dans l'anneau **Z** — et on ne peut rien dire de général sur les conditions de résolubilité de l'équation (1) dans un anneau d'intégrité qui n'est pas un corps.

Il existe des anneaux qui ne sont pas intègres. Prenons par exemple l'anneau de toutes les applications de l'ensemble **R** dans l'anneau **R** (*Exemple* 3 ci-dessus) et considérons les deux éléments f et g de cet anneau définis comme suit :

$$f(x) = \begin{cases} x & \text{pour } x \geqslant 0, \\ 0 & \text{pour } x \leqslant 0, \end{cases} \qquad g(x) = \begin{cases} 0 & \text{pour } x \geqslant 0, \\ x & \text{pour } x \leqslant 0; \end{cases}$$

il est clair que

$$f(x)g(x) = 0 \quad \text{pour tout} \quad x \in \mathbf{R},$$

et par suite que $fg = 0$ dans l'anneau considéré; néanmoins on a $f \neq 0$ et $g \neq 0$ (car l'élément 0 de l'anneau des applications d'un ensemble X dans un anneau K est la fonction qui, en *chaque* $x \in \mathrm{X}$ *sans exception,* prend la valeur 0 — ce qui, ici, n'est le cas ni de f ni de g).

Remarque 1. Soit K un anneau; on désigne habituellement l'ensemble des éléments inversibles de K par la notation

$$\mathrm{K}^*$$

(cf. le passage de **Q** à $\mathbf{Q}^*$ ou de **R** à $\mathbf{R}^*$). D'après le Théorème 3 du § 6, si K^*

(*) En fait, dans un corps, on exige aussi que $1 \neq 0$, de sorte qu'un corps possède toujours au moins *deux* éléments.

contient deux éléments x et y il contient aussi xy; on peut donc munir l'ensemble K^* de la loi de composition $(x, y) \mapsto xy$. Cela dit, l'ensemble K^*, muni de cette loi de composition, est un *groupe*. Il est clair en effet que la loi de composition considérée sur K^* est associative (car elle l'est déjà dans K); d'autre part, on a évidemment $1 \in K^*$, de sorte que la loi de composition considérée sur l'ensemble K^* admet un élément neutre; enfin, si $x \in K^*$, on a aussi $x^{-1} \in K^*$ d'après le Théorème 3 du § 6, et comme on a

$$x^{-1}x = xx^{-1} = 1,$$

élément neutre de K^*, on voit que tout élément de K^* est inversible (dans K^*, et pas seulement dans K !) pour la loi de composition considérée.

L'ensemble K^*, muni de la loi de composition $(x, y) \mapsto xy$, s'appelle le **groupe multiplicatif de l'anneau** K (ou, parfois, le **groupe des unités** de K). Si K est un *corps*, on a

$$K^* = K - \{0\}.$$

Par contre, on a

$$\mathbf{Z}^* = \{1, -1\},$$

avec la table de multiplication suivante :

$$1.1 = 1; \qquad -1.1 = 1.-1 = -1; \qquad -1.-1 = 1.$$

Remarque 2. Soit K un corps. On appelle **sous-corps** de K toute partie A de K vérifiant les conditions suivantes : A *est un sous-anneau de* K, *et pour* $x \neq 0$ *la relation* $x \in A$ *implique* $x^{-1} \in A$. Il est clair qu'alors l'ensemble A, muni des lois de composition « induites » par celles de K, est non seulement un anneau mais un corps.

Ainsi, $\mathbf{Q}$ est un sous-corps de $\mathbf{R}$.

Exemple 6. Soit $K \subset \mathbf{R}$ l'ensemble des nombres réels x possédant la propriété suivante : il existe des nombres *rationnels* a et b tels que l'on ait

$$x = a + br, \qquad \text{où} \qquad r = \sqrt{2}.$$

Le lecteur vérifiera facilement que K est un sous-corps de $\mathbf{R}$.

3. *L'anneau des entiers modulo p*

Au § 4, *Exemple* 9, nous avons défini, pour tout entier rationnel p, l'*ensemble* $\mathbf{Z}/p\mathbf{Z}$ des entiers modulo p, et, dans l'*Exemple* 14 du même §, nous avons défini sur cet ensemble deux lois de composition appelées addition et multiplication; celles-ci sont liées aux lois de composition sur les entiers ordinaires par le fait suivant : *si* θ *désigne l'application canonique de* $\mathbf{Z}$ *sur* $\mathbf{Z}/p\mathbf{Z}$, *on a*

$$\theta(x + y) = \theta(x) + \theta(y), \qquad \theta(xy) = \theta(x)\theta(y)$$

quels que soient $x, y \in \mathbf{Z}$.

On va déduire de là que *l'ensemble* $\mathbf{Z}/p\mathbf{Z}$ *des entiers modulo* p, *muni de l'addition et de la multiplication définies au* § 4, *est un anneau commutatif.*

Montrons d'abord que l'addition dans $\mathbf{Z}/p\mathbf{Z}$ est associative; soient ξ, η, ζ trois éléments de cet ensemble; il existe $x, y, z \in \mathbf{Z}$ tels que $\xi = \theta(x)$, $\eta = \theta(y)$, $\zeta = \theta(z)$; on a alors

$$\xi + \eta = \theta(x) + \theta(y) = \theta(x + y), \qquad \eta + \zeta = \theta(y) + \theta(z) = \theta(y + z),$$

donc

$$(\xi + \eta) + \zeta = \theta(x + y) + \theta(z) = \theta((x + y) + z),$$
$$\xi + (\eta + \zeta) = \theta(x) + \theta(y + z) = \theta(x + (y + z)),$$

et l'associativité de l'addition dans $\mathbf{Z}/p\mathbf{Z}$ résulte donc de l'associativité de l'addition dans $\mathbf{Z}$.

On prouverait de même l'associativité de la multiplication, la commutativité de l'addition et de la multiplication, et la distributivité de la multiplication par rapport à l'addition dans $\mathbf{Z}/p\mathbf{Z}$.

Il est clair, vu la relation

$$\theta(1)\theta(x) = \theta(1x) = \theta(x),$$

que $\theta(1)$ est élément neutre pour la multiplication dans $\mathbf{Z}/p\mathbf{Z}$, et de même que $\theta(0)$ est élément neutre pour l'addition.

Pour établir que $\mathbf{Z}/p\mathbf{Z}$ est un anneau commutatif, il reste donc à montrer que tout élément ξ de $\mathbf{Z}/p\mathbf{Z}$ admet un opposé; pour cela, on écrit $\xi = \theta(x)$ pour un $x \in \mathbf{Z}$ convenablement choisi, et il est alors clair, vu la relation

$$\theta(-x) + \theta(x) = \theta(-x + x) = \theta(0),$$

que ξ admet effectivement un opposé, à savoir $\theta(-x)$.

Les résultats qu'on vient d'établir permettent de parler de **l'anneau $\mathbf{Z}/p\mathbf{Z}$ des entiers modulo p**; cet anneau possède un nombre fini seulement d'éléments si $p \neq 0$, à savoir p (on suppose p positif, ce qui ne restreint pas la généralité).

Pour certaines valeurs de p, l'anneau $\mathbf{Z}/p\mathbf{Z}$ est même un corps (ce qui prouvera l'existence de **corps finis**, i.e. de corps à un nombre fini d'éléments) :

THÉORÈME 1. *Soit $p \geqslant 2$ un entier. Les assertions suivantes sont équivalentes :*

a) l'anneau $\mathbf{Z}/p\mathbf{Z}$ est intègre;

b) l'anneau $\mathbf{Z}/p\mathbf{Z}$ est un corps;

c) le nombre p est premier.

Soient ξ et η deux éléments *non nuls* de $\mathbf{Z}/p\mathbf{Z}$; on a donc $\xi = \theta(x)$, $\eta = \theta(y)$ avec

$$x \not\equiv 0 \pmod{p}, \qquad y \not\equiv 0 \pmod{p};$$

pour en déduire que $\xi\eta$, qui est égal à $\theta(xy)$, est aussi non nul, il faut montrer qu'on a aussi

$$xy \not\equiv 0 \pmod{p};$$

autrement dit, pour que $\mathbf{Z}/p\mathbf{Z}$ soit intègre, il faut et il suffit que la relation

$$xy \equiv 0 \pmod{p} \quad \text{implique} \quad x \equiv 0 \pmod{p} \quad \text{ou} \quad y \equiv 0 \pmod{p},$$

ou encore que si p divise un produit xy, il divise soit x soit y : d'où l'équivalence des propriétés *a*) et *c*).

Il est d'autre part clair que *b*) implique *a*). Pour achever la démonstration, il suffit donc de montrer que *a*) implique *b*), ce qui résultera visiblement du Théorème plus général que voici :

THÉORÈME 2. *Tout anneau d'intégrité fini est un corps.*

Soit K un anneau d'intégrité fini; pour un élément $a \neq 0$ de K, considérons l'application $x \to ax$ de K dans K; comme $ax = ay$ implique $a(x - y) = 0$, donc $x - y = 0$ si K est intègre, on voit que l'application considérée est *injective*; mais comme l'ensemble K est *fini*, cette application est forcément *surjective* (§ 5, Théorème 4), et en particulier on peut résoudre $ax = 1$, ce qui montre que tout élément non nul de K possède un inverse à droite. On montrerait de même, à l'aide de l'application $x \to xa$, que tout élément non nul de K possède un inverse à gauche, ce qui achève la démonstration.

Remarque 3. On peut démontrer que *tout corps fini est commutatif*, mais les techniques nécessaires pour y parvenir dépassent de fort loin le niveau du présent ouvrage.

On peut d'autre part démontrer que le nombre d'éléments d'un corps fini est nécessairement une puissance d'un nombre premier, et que pour tout nombre premier p et tout entier $n \geqslant 1$, il existe essentiellement un seul corps à p^n éléments (ce qui veut dire qu'on sait construire explicitement tous les corps finis). La première étude détaillée des corps finis a été faite par Galois.

4. *Formule du binôme*

Les « identités remarquables » qu'on démontre dans l'enseignement secondaire lorsqu'il s'agit de nombres réels, sont pour la plupart encore valables dans tout anneau (en supposant parfois que les éléments $x, y, \cdots$ figurant dans ces identités commutent deux à deux). Par exemple, soient x, y deux éléments d'un anneau K, et calculons

$$(x + y)^2 = (x + y)(x + y) = x^2 + xy + yx + y^2;$$

on voit que, si x et y commutent, on retrouve l'identité $(x + y)^2 = x^2 + 2xy + y^2$. On a alors

$$(x + y)^3 = (x^2 + 2xy + y^2)(x + y) = x^3 + 3x^2y + 3xy^2 + y^3$$

comme le montre un calcul trivial.

Plus généralement :

THÉORÈME 3. *Soient* K *un anneau, x et y deux éléments de* K, *et supposons que x et y commutent. On a alors, pour tout entier $n \geqslant 1$, la relation*

$$(x + y)^n = \sum_{p=0}^{p=n} \binom{n}{p} x^{n-p} y^p$$

où l'on pose

$$\binom{n}{p} = \frac{n(n-1)\dots(n-p+1)}{1.2\dots p} = \frac{n!}{p!(n-p)!} \quad (0 \leqslant p \leqslant n).$$

Remarque 4. Rappelons (§ 5, Théorème 10) que les nombres $\binom{n}{p}$ sont des *entiers* et pas seulement des nombres rationnels.

Le résultat à établir est trivial pour $n = 1$; il suffit donc de prouver que la formule

$$(x+y)^{n-1} = \sum_{p=0}^{p=n-1} \binom{n-1}{p} x^{n-1-p} y^p,$$

implique la formule analogue pour l'exposant n. Or, en multipliant par $x+y$ la relation précédente, il vient

$$(x+y)^n = \sum_{p=0}^{p=n-1} \binom{n-1}{p} x^{n-p} y^p + \sum_{p=0}^{p=n-1} \binom{n-1}{p} x^{n-1-p} y^{p+1};$$

si r est un entier tel que $0 < r < n$, il y a au second membre de la relation précédente deux termes contenant le monôme

$$x^{n-r} y^r;$$

le premier s'obtient en prenant $p = r$ dans la première somme, ce qui introduit un facteur égal à

$$\binom{n-1}{r},$$

et le second s'obtient en prenant $p = r-1$ dans la seconde somme, ce qui introduit un facteur égal à

$$\binom{n-1}{r-1};$$

pour achever la démonstration il reste donc à vérifier la relation

$$\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}.$$

Or le second membre est égal à

$$\frac{(n-1)(n-2)\dots(n-r)}{1.2\dots r} + \frac{(n-1)(n-2)\dots(n-r+1)}{1.2\dots(r-1)}$$

$$= \frac{(n-1)\dots(n-r+1)(n-r) + (n-1)\dots(n-r+1)r}{r!}$$

$$= \frac{[(n-r)+r](n-1)\dots(n-r+1)}{r!} = \frac{n(n-1)\dots(n-r+1)}{r!} = \binom{n}{r},$$

ce qui achève la démonstration.

La relation

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}$$

permet de calculer facilement les nombres $\binom{n}{r}$, qu'on appelle pour des raisons évidentes les **coefficients du binôme** d'indices n et r. Ils sont donnés par le tableau suivant, appelé **triangle de Pascal** :

$$\begin{array}{ccccccccc}
1 & 1 & & & & & & & \\
1 & 2 & 1 & & & & & & \\
1 & 3 & 3 & 1 & & & & & \\
1 & 4 & 6 & 4 & 1 & & & & \\
1 & 5 & 10 & 10 & 5 & 1 & & & \\
1 & 6 & 15 & 20 & 15 & 6 & 1 & & \\
1 & 7 & 21 & 35 & 35 & 21 & 7 & 1 & \\
\multicolumn{9}{c}{\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots;}
\end{array}$$

la méthode, pour calculer le p^e terme de la n^e ligne, consiste à additionner le p^e et le $(p-1)^e$ termes de la ligne précédente.

L'examen du tableau précédent suggère la formule

$$\binom{n}{r} = \binom{n}{n-r};$$

la vérification de celle-ci est immédiate puisque

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}, \qquad \binom{n}{n-r} = \frac{n!}{(n-r)!\,r!};$$

mais la véritable raison de la formule en question réside dans le fait que l'expression

$$\sum_{p=0}^{p=n} \binom{n}{p} x^{n-p} y^p$$

ne doit pas changer — vu sa signification — si l'on permute x et y; il est donc naturel que les coefficients dans cette expression des « monômes » $x^{n-r}y^r$, $x^r y^{n-r}$ soient égaux, puisque ces monômes se déduisent l'un de l'autre par échange de x et y.

On peut aussi observer que, X désignant un ensemble à n éléments, $\binom{n}{r}$ est le nombre de parties à r éléments de X; en associant à une telle partie Y son complémentaire X — Y, on obtient une *bijection* de l'ensemble des parties à r éléments de X sur l'ensemble des parties à $n-r$ éléments de X; d'où la relation $\binom{n}{r} = \binom{n}{n-r}$.

5. *Développement d'un produit de sommes*

La formule du binôme est un cas particulier d'une formule plus générale que nous allons exposer maintenant.

Soient K un anneau commutatif, I un ensemble *fini*, et $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles d'éléments de K indexées par I. On se propose de « développer » le produit

$$\prod_{i\in I}(x_i + y_i).$$

Pour énoncer le résultat, introduisons d'abord la notation suivante : étant donnée une partie F de I, on posera

$$x_F = \prod_{i\in F} x_i, \qquad y_F = \prod_{i\in F} y_i$$

(étant entendu que $x_F = y_F = 1$ si F est vide). Ceci dit, la formule cherchée s'écrit

$$\prod_{i\in I}(x_i + y_i) = \sum_{F\subset I} x_F y_{I-F},$$

la somme figurant au second membre étant étendue à toutes les parties de l'ensemble I.

En effet, pour multiplier des sommes les unes par les autres, on choisit, de toutes les façons possibles, un terme dans chacune de ces sommes, on effectue le produit des termes ainsi choisis, et on additionne les résultats ainsi obtenus pour tous les choix possibles.

Le produit des sommes $x_i + y_i$ sera donc une somme de termes obtenus en multipliant les x_i figurant dans un certain nombre des sommes données par les y_j figurant dans les autres. Pour un tel produit, notons F l'ensemble des valeurs de i pour lesquelles on décide de choisir x_i, de sorte que $I - F$ est l'ensemble des valeurs de i pour lesquelles on décide de choisir y_i; il est clair que le produit des termes ainsi choisis est $x_F y_{I-F}$; en ajoutant les résultats ainsi obtenus on trouve donc la formule annoncée.

La formule du binôme résulte comme suit de la relation qu'on vient d'établir : on prend pour I un ensemble à n éléments, et on choisit $x_i = x, y_i = y$ pour tout $i \in I$. Le premier membre de la formule générale est donc $(x + y)^n$. Au second membre, il est clair que

$$x_F y_{I-F} = x^r y^{n-r}$$

où r est le nombre d'éléments de F. Il reste donc, pour obtenir la formule du binôme, à tenir compte du fait que, dans un ensemble à n éléments, il y a $\binom{n}{r}$ parties à r éléments.

6. *Homomorphismes d'anneaux*

Étant donnés deux anneaux K et L, on appelle **homomorphisme de K dans L** toute application f de K dans L telle que l'on ait

$$f(x+y)=f(x)+f(y),\quad f(xy)=f(x)f(y)\quad \text{quels que soient}\quad x,y\in K,\quad f(1)=1.$$

Exemple 7. L'application canonique de $\mathbf{Z}$ sur $\mathbf{Z}/p\mathbf{Z}$ est un homomorphisme d'anneaux (et du reste la structure d'anneau de $\mathbf{Z}/p\mathbf{Z}$ a été choisie de telle sorte qu'il en soit ainsi).

Exemple 8. Soient X un ensemble, K un anneau, et L l'anneau des applications de X dans K (*Exemple* 3); alors, pour tout $x\in X$, l'application

$$f\mapsto f(x)$$

de L dans K est un homomorphisme.

Les propriétés des homomorphismes d'anneaux sont analogues à celles des homomorphismes de groupes (§ 7, n° 8 et 9). Étant donnés deux homomorphismes d'anneaux

$$f: K\to L,\qquad g: L\to M,$$

l'application composée $g\circ f$ est encore un homomorphisme.

Si un homomorphisme $f: K\to L$ est bijectif, l'application réciproque est encore un homomorphisme; on dit alors que f est un **isomorphisme**, et on dit que deux anneaux K et L sont **isomorphes** s'il existe un isomorphisme de K sur L; la relation

K et L sont isomorphes

est une relation d'équivalence entre anneaux.

Soit $f: K\to L$ un homomorphisme d'anneaux. On appelle **noyau** de f l'ensemble, noté

$$\mathrm{Ker}(f),$$

des $x\in K$ tels que $f(x)=0$ (c'est donc le noyau de f quand on regarde f comme un homomorphisme du groupe additif K dans le groupe additif L). L'homomorphisme f est injectif si et seulement si $\mathrm{Ker}(f)=\{0\}$.

Les relations

$$f(x-y)=f(x)-f(y),\qquad f(axb)=f(a)f(x)f(b)$$

montrent immédiatement que le noyau I d'un homomorphisme d'un anneau K dans un autre vérifie les deux conditions suivantes :

(i) : I *est un sous-groupe du groupe additif* K;

(ii) : *on a la relation* $axb\in I$ *quels que soient* $a, b\in K$ *et* $x\in I$.

Une partie I d'un anneau K s'appelle un **idéal bilatère** lorsqu'elle vérifie les conditions (i) et (ii) ci-dessus.

Remarque 4. Une partie I d'un anneau K est appelée un **idéal à gauche de K** si c'est un sous-groupe de K et si l'on a $ax \in I$ quels que soient $a \in K$ et $x \in I$ (autrement dit si les **multiples à gauche** de tout $x \in I$ sont tous dans I) ; il revient au même de dire que I est une partie de K, non vide, qui possède la propriété suivante : on a

$$ux + vy \in I \quad \text{quels que soient } u, v \in K \text{ et } x, y \in I.$$

On définirait de même les **idéaux à droite** de K comme étant les parties non vides I de K telles que l'on ait

$$xu + yv \in I \quad \text{quels que soient } u, v \in K \text{ et } x, y \in I.$$

Les idéaux bilatères sont évidemment les parties de K qui sont à la fois des idéaux à gauche et des idéaux à droite.

Lorsque K est *commutatif*, les notions d'idéal à gauche, d'idéal à droite et d'idéal bilatère sont évidemment identiques; on dit alors simplement **idéal** au lieu d'idéal à gauche, ou à droite, ou bilatère.

Exemple 9. Si K est un corps, les seuls idéaux à gauche de K sont $\{0\}$ et K lui-même. Si en effet un idéal à gauche I contient un élément a non nul, donc inversible, il contiendra $a^{-1}a = 1$, donc aussi $u.1 = u$ pour tout $u \in K$, d'où $I = K$.

Le lecteur démontrera, à titre d'exercice, que lorsque $1 \neq 0$ cette propriété caractérise les anneaux qui sont des corps.

Exemple 10. Soit K un anneau commutatif; pour tout $x \in K$, notons xK l'ensemble des **multiples de x dans K**, i.e. des éléments de la forme ux, $u \in K$. Cet ensemble est alors un idéal de K (les idéaux de ce type sont appelés les **idéaux principaux** de K).

On appelle **anneau principal** tout anneau d'intégrité commutatif dont tous les idéaux sont principaux. L'anneau **Z** est principal (un idéal de **Z** est un sous-groupe de **Z**, donc est de la forme $n\mathbf{Z}$ d'après le § 7, *Exemple* 8). On verra au § 31 (que le lecteur peut, s'il le désire, étudier dès maintenant) que les propriétés de divisibilité des entiers rationnels s'étendent aux éléments d'un anneau principal.

Il existe des anneaux non principaux — l'exemple le plus simple est le sous-anneau de **R** formé des nombres de la forme

$$x + y\sqrt{10} \qquad \text{avec} \qquad x, y \in \mathbf{Z}.$$

L'étude de cet anneau et d'anneaux analogues mais plus compliqués (les anneaux d'entiers algébriques) a conduit les mathématiciens du siècle dernier — en premier lieu Dedekind — à inventer la notion d'idéal qui, par la suite, s'est révélée indispensable dans de nombreuses autres branches des Mathématiques.

§ 9. Nombres complexes

1. *Racines carrées*

Soit K un anneau commutatif. On dit qu'un élément d de K est un **carré dans K** s'il existe un $x \in K$ tel que

$$x^2 = d;$$

on dit alors que x est une **racine carrée** de d dans K. Il est clair que si x est une racine carrée de d dans K, il en est de même de $-x$; si de plus l'anneau K est intègre, d ne peut admettre plus de deux racines carrées dans K, car la relation $x^2 = y^2$, qui s'écrit dans tous les cas sous la forme $(x-y)(x+y) = 0$, implique soit $y = x$ soit $y = -x$ lorsque l'anneau K est intègre.

Il peut naturellement arriver qu'un élément d d'un anneau commutatif K ne soit pas un carré dans K : si K est le corps **R** des nombres réels, d est un carré dans K si et seulement si $d \geqslant 0$. Dans le corps **Q** des nombres rationnels, 2 n'est pas un carré (cependant, 2 est un carré dans le corps **R** des nombres réels).

On est ainsi conduit à examiner le problème suivant : *soient* K *un anneau commutatif et* d *un élément de* K *qui n'est pas un carré dans* K*; peut-on construire un anneau commutatif* L *possédant les propriétés suivantes :* K *est un sous-anneau de* L*, et* d *est un carré dans* L*?*

C'est la résolution de ce problème lorsque $K = \mathbf{R}$ et $d = -1$ qui, historiquement, a conduit à l'invention des « nombres complexes » que nous définirons dans la suite de ce §. Pour comprendre vraiment la construction de ces « nombres », il est indispensable (et il n'est pas plus difficile) d'étudier le problème général que nous venons d'énoncer.

2. *Préliminaires*

Soit donc d un élément d'un anneau commutatif K; dans ce n° nous supposerons résolu le problème énoncé au n° précédent, et désignerons par L un anneau commutatif dont K soit un sous-anneau, et par ω une racine carrée de d dans L.

Désignons alors par L' l'ensemble des éléments z de L possédant la propriété

suivante : il existe $x, y \in K$ tels que

$$z = x + \omega y. \tag{1}$$

Alors, L′ *est un sous-anneau de* L *contenant* K, *et dans lequel d possède une racine carrée.*

Il est en effet clair que L′ contient K (faire $y = 0$ dans la relation précédente) ainsi que ω (faire $x = 0$ et $y = 1$) ; il reste donc à faire voir que L′ est un sous-anneau de L; comme L′, contenant K, contient -1, il suffit pour cela de montrer que si L′ contient deux éléments de L, il contient aussi leur somme et leur produit; mais cela résulte aussitôt des formules

$$(x' + \omega y') + (x'' + \omega y'') = (x' + x'') + \omega(y' + y''), \tag{2}$$

$$(x' + \omega y').(x'' + \omega y'') = (x'x'' + dy'y'') + \omega(x'y'' + x''y'), \tag{3}$$

évidentes compte-tenu de la relation

$$\omega^2 = d. \tag{4}$$

Le résultat précédent montre que, si le problème posé admet une solution, autrement dit si l'on a construit un sur-anneau commutatif L de K et un $\omega \in L$ vérifiant (4), alors on peut même construire L de telle sorte que *tout* élément de L puisse se mettre sous la forme (1), avec $x, y \in K$. Autrement dit, si l'on introduit l'application $f : K \times K \to L$ donnée par

$$f(x, y) = x + \omega y,$$

on peut supposer f *surjective*.

Remarque 1. Si K est un corps et si d n'est pas un carré dans K, l'application f est en outre *injective*; en effet, la relation $x' + \omega y' = x'' + \omega y''$ s'écrit

$$x + \omega y = 0 \qquad \text{avec} \qquad x = x' - x'', \qquad y = y' - y'';$$

si $y \neq 0$, alors y est inversible dans K (puisque K est un corps), à *fortiori* dans L, et la relation $x + \omega y = 0$ implique

$$\omega = -y^{-1}x \in K;$$

contrairement à l'hypothèse que d n'est pas un carré dans K. On a donc $y = 0$, et donc $x = 0$, autrement dit $x' = x''$ et $y' = y''$, ce qui montre bien que f est injective.

On remarquera qu'en introduisant l'application f, les formules (2) et (3) s'écrivent

$$f(x', y') + f(x'', y'') = f(x' + x'', y' + y''), \tag{2 bis}$$

$$f(x', y').f(x'', y'') = f(x'x'' + dy'y'', x'y'' + x''y'). \tag{3 bis}$$

Ces formules (obtenues en supposant le problème résolu) vont maintenant nous servir de point de départ pour construire effectivement une solution du problème posé.

3. *L'anneau* $K[\sqrt{d}]$

Soit d un élément d'un anneau commutatif K; nous allons construire un nouvel anneau L, qu'on note traditionnellement

$$K[\sqrt{d}],$$

et qui est défini comme suit : l'ensemble L est le produit cartésien $K \times K$, de sorte qu'un élément de L est un couple (x, y) d'éléments de K; et les deux opérations fondamentales dans L sont *définies* par les formules

$$(2\ ter) \qquad (x,' y') + (x'', y'') = (x' + x'', y' + y''),$$
$$(3\ ter) \qquad (x', y') \cdot (x'', y'') = (x'x'' + dy'y'', x'y'' + x''y'),$$

lesquelles font intervenir à la fois l'élément donné d et les lois de composition dans l'anneau K.

Bien entendu, il n'est nullement évident que les formules (2 *ter*) et (3 *ter*) font de l'ensemble $K \times K$ un *anneau commutatif*, et on doit le démontrer (y compris dans le cas classique où $K = \mathbf{R}$ et $d = -1$; la démonstration dans ce cas particulier n'est ni plus facile ni plus difficile que dans le cas général).

Tout d'abord, l'ensemble $K \times K$ muni de l'addition (2 *ter*) est un groupe commutatif : cela résulte du § 7, n° 2 (produit direct de groupes), et du fait que l'ensemble K, muni de l'addition donnée, est un groupe commutatif.

Montrons maintenant que la multiplication (3 *ter*) est associative. On a en effet par définition

$$\begin{aligned}(x, y)\,[(x', y')\ (x'', y'')] &= (x, y)\ (x'x'' + dy'y'', x'y'' + x''y') \\ &= (x(x'x'' + dy'y'') + dy(x'y'' + x''y'), x(x'y'' + x''y') + y(x'x'' + dy'y''))\end{aligned}$$

et d'autre part

$$\begin{aligned}[(x, y)\ (x', y')](x'', y'') &= (xx' + dyy', xy' + x'y)\ (x'', y'') \\ &= ((xx' + dyy')x'' + d(xy' + x'y)y'', (xx' + dyy')y'' + x''(xy' + x'y));\end{aligned}$$

l'associativité s'obtient trivialement en comparant les résultats obtenus (et, bien entendu, en utilisant les règles de calcul dans K). On établirait par des calculs analogues la commutativité de la multiplication, et sa distributivité par rapport à l'addition.

Enfin, la relation

$$(1, 0)\ (x, y) = (x, y)$$

montre que l'ensemble $K \times K$, muni de la loi de composition (3 *ter*), admet un élément neutre.

L'ensemble $K \times K$, muni des lois de composition (2 *ter*) et (3 *ter*), est donc bien un anneau commutatif.

Montrons maintenant que $K \times K$ *contient un sous-anneau isomorphe à* K; pour cela,

considérons l'application $j : K \to K \times K$ donnée par

$$j(x) = (x, 0);$$

il est clair qu'elle est injective; de plus, un calcul facile montre qu'on a

$$j(x') + j(x'') = j(x' + x''),\ j(x')\,j(x'') = j(x'x''),\ j(-1) = (-1, 0) = -1$$

où, dans le dernier membre de la dernière relation, -1 désigne l'opposé de l'élément unité $1 = (1, 0)$ de l'anneau $K \times K$; cela fait, les formules précédentes montrent évidemment que j est un *isomorphisme de* K *sur un sous-anneau de* $K \times K$.

Comme j transforme les lois de compositions de K en celles du sous-anneau $j(K)$ de $K \times K$, il n'y a aucun inconvénient à identifier chaque élément x de K à l'élément $j(x)$ de $K \times K$; c'est ce que nous ferons désormais (*).

Pour montrer qu'on a ainsi résolu le problème posé au n° 1, il reste à faire voir que *l'élément d de* K *est un carré dans* $K[\sqrt{d}]$. Or considérons l'élément

$$\omega = (0, 1) \tag{5}$$

de L; un calcul trivial montre que

$$\omega^2 = (0, 1)\ (0, 1) = (0.0 + d.1.1,\ 0.1 + 1.0) = (d, 0) = j(d),$$

et comme on a convenu d'une manière générale d'identifier chaque $x \in K$ à l'élément $j(x)$ de L, notre assertion est démontrée.

Nous avons ainsi non seulement résolu le problème posé au n° 1, mais même construit une solution « canonique » de ce problème, à savoir l'anneau $K[\sqrt{d}]$. On dit qu'il s'obtient par **adjonction à K d'une racine carrée de** d; un anneau de la forme $K[\sqrt{d}]$ s'appelle une **extension quadratique de** K.

Pour terminer ces constructions, indiquons une notation commode pour les éléments de $K[\sqrt{d}]$. Tout d'abord l'identification de tout $x \in K$ à l'élément $(x, 0)$ de L permet d'écrire

$$(x, 0) = x \tag{6}$$

(ce qui, à stricternent parler, est aussi faux que possible...). Mais un calcul immédiat montre que, quels que soient $x, y \in K$, on a la formule

$$(x, y) = (x, 0) + (0, 1)\ (y, 0);$$

(*) La méthode utilisée ici est analogue à celle qui permet d'identifier les nombres entiers à des nombres rationnels particuliers (ou les nombres rationnels à des nombres réels particuliers). Quels que soient les procédés utilisés pour définir mathématiquement les nombres entiers et les nombres rationnels, il n'est jamais vrai que le nombre entier n et le nombre rationnel représenté par la fraction $\frac{n}{1}$ soient *identiques*; on peut simplement dire que l'application $n \to \frac{n}{1}$ est injective et transforme les opérations algébriques sur des nombres entiers en les opérations algébriques analogues sur les nombres rationnels correspondants. C'est ce qui explique pourquoi, dans la pratique, on ne fait pas de distinction entre n et $\frac{n}{1}$.

celle-ci s'écrit, compte-tenu de (5) et (6), sous la forme

$$(x, y) = x + \omega y, \tag{7}$$

et on est ainsi ramené à la situation étudiée au n° 2 en supposant le problème résolu.

En conclusion, tout élément de l'anneau $K[\sqrt{d}]$ s'écrit d'une façon et d'une seule sous la forme $x + \omega y$, et ce fait, joint aux axiomes des anneaux commutatifs et à la relation

$$\omega^2 = d, \tag{4}$$

suffit pour effectuer tous les calculs dont on peut avoir besoin : autrement dit le lecteur peut maintenant oublier la *construction* effective de l'anneau $K[\sqrt{d}]$ pour n'en retenir que les *propriétés*.

Exemple 1. Le cas le plus important est celui où $K = \mathbf{R}$, corps des nombres réels, et $d = -1$. L'anneau $\mathbf{R}[\sqrt{-1}]$ ainsi obtenu se note $\mathbf{C}$, et ses éléments s'appelent les **nombres complexes**; un nombre complexe est donc un couple (x, y) de nombres réels, et on calcule sur les nombres complexes à l'aide des formules (2 *ter*) et (3 *ter*) pour $d = -1$, i.e. on pose

$$(x', y') + (x'', y'') = (x' + x'', y' + y'')$$
$$(x', y') \cdot (x'', y'') = (x'x'' - y'y'', x'y'' + x''y').$$

Dans la pratique on n'utilise pas ces formules; on utilise seulement les propriétés suivantes des nombres complexes :

a) *les nombres complexes forment un anneau commutatif* $\mathbf{C}$ (on verra plus loin que $\mathbf{C}$ est même un corps);

b) *le corps* $\mathbf{R}$ *des nombres réels est un sous-anneau de* $\mathbf{C}$;

c) *il existe un nombre complexe* i (cette notation traditionnelle remplace la notation ω utilisée pour les extensions quadratiques générales) *tel que*

$$i^2 = -1;$$

d) *tout nombre complexe* z *s'écrit d'une façon et d'une seule sous la forme*

$$z = x + iy$$

où x *et* y *sont réels.*

Il n'y a rien de plus à savoir pour calculer sur les nombres complexes. Par exemple :

$$(2 + 3i)(5 - 7i) = 2 \cdot 5 - 2 \cdot 7i + 3i \cdot 5 - 3i \cdot 7i$$
$$= 10 - 14i + 15i - 21i^2 = 10 + i + 21 = 31 + i.$$

Étant donné un nombre complexe

$$z = x + iy$$

avec x, y réels, on dit que x est la **partie réelle** et y la **partie imaginaire** de z; on

les désigne par les notations

$$x = \mathrm{Re}(z), \qquad y = \mathrm{Im}(z).$$

On dit que z est **imaginaire pur** si $\mathrm{Re}(z) = 0$; dire par contre que z est **réel** (i.e. appartient au sous-anneau **R** de **C**) se traduit par la relation $\mathrm{Im}(z) = 0$.

4. Éléments inversibles d'une extension quadratique

Soient K un anneau commutatif et d un élément de K; considérons l'extension quadratique

$$\mathrm{L} = \mathrm{K}[\sqrt{d}]$$

construite au n° précédent. Étant donné un élément

$$z = x + \omega y \quad (x, y \in \mathrm{K})$$

de L, on appelle **conjugué de** z l'élément

$$\bar{z} = x - \omega y \tag{8}$$

de L, et **norme de** z l'élément

$$\mathrm{N}(z) = \bar{z}.z = (x - \omega y)(x + \omega y) = x^2 - \omega^2 y^2 = x^2 - dy^2 \tag{9}$$

de K.

Quels que soient les éléments z' et z'' de L, on a les relations

$$\overline{z' + z''} = \bar{z'} + \bar{z''}, \tag{10}$$

$$\overline{z'z''} = \bar{z'}.\bar{z''}, \tag{11}$$

qui montrent comment calculer le conjugué d'une somme ou d'un produit d'éléments de L. Pour établir les relations (10) et (11), le plus simple est de partir des formules (2 *ter*) et (3 *ter*) du n° 3, et d'examiner ce qui se passe lorsqu'on y remplace y' et y'' par leurs opposés, sans modifier x' et x''.

La relation (11) montre qu'on a

$$\mathrm{N}(z'z'') = \mathrm{N}(z')\mathrm{N}(z'') \tag{12}$$

quels que soient z' et z''; en effet

$$\mathrm{N}(z'z'') = \overline{z'z''}.z'z'' = \bar{z'}.\bar{z''}.z'.z'' = \bar{z'}.z'.\bar{z''}.z'' = \mathrm{N}(z')\mathrm{N}(z'')$$

comme annoncé. On observera qu'on a aussi

$$\mathrm{N}(1) = 1. \tag{13}$$

THÉORÈME 1. *Soient* K *un anneau commutatif,* d *un élément de* K, *et* z *un élément de l'anneau* $\mathrm{K}[\sqrt{d}]$. *Pour que* z *soit inversible dans* $\mathrm{K}[\sqrt{d}]$, *il faut et il suffit que* $\mathrm{N}(z)$ *le soit dans* K;

on a alors

$$z^{-1} = \mathrm{N}(z)^{-1}.\bar{z}. \tag{14}$$

Supposons z inversible; alors la relation $z^{-1}.z = 1$ donne, compte-tenu de (12) et (13), la relation

$$\mathrm{N}(z^{-1}).\mathrm{N}(z) = 1,$$

et $\mathrm{N}(z)$ est donc bien un élément inversible de l'anneau K.

Inversement, supposons $\mathrm{N}(z)$ inversible dans K; la relation

$$\bar{z}z = \mathrm{N}(z)$$

implique alors

$$\mathrm{N}(z)^{-1}\bar{z}.z = 1,$$

ce qui montre que z est inversible et que son inverse est donné par la relation (14); d'où le Théorème.

5. *Cas d'un corps commutatif*

Nous pouvons maintenant démontrer le résultat suivant :

THÉORÈME 2. *Soit d un élément d'un corps commutatif* K; *les propriétés suivantes sont équivalentes :*

a) *l'anneau* $\mathrm{K}[\sqrt{d}]$ *est un corps;*

b) *d n'est pas un carré dans* K.

Pour montrer que *a*) implique *b*), supposons qu'il existe un $x \in \mathrm{K}$ tel que

$$x^2 = d;$$

on a alors $x^2 = \omega^2$, donc $(x - \omega)(x + \omega) = 0$; si $\mathrm{K}[\sqrt{d}]$ est un corps, donc un anneau d'intégrité, il s'ensuit qu'on a soit $\omega = x$, soit $\omega = -x$, ce qui est impossible puisque pour $x, y \in \mathrm{K}$ la relation $x + \omega y = 0$ implique $x = y = 0$.

Montrons maintenant que *b*) implique *a*). Soit $z = x + \omega y$ un élément *non nul* de $\mathrm{K}[\sqrt{d}]$; pour montrer qu'il est inversible il suffit (Théorème 1) de montrer que $\mathrm{N}(z)$ est inversible dans K; comme K est un corps, tout revient donc à prouver que la relation $\mathrm{N}(z) = 0$ implique $z = 0$, autrement dit que

$$x^2 - dy^2 = 0 \quad \text{implique} \quad x = y = 0;$$

or si l'on avait $y \neq 0$, l'élément y de K serait inversible, et il viendrait

$$d = (y^2)^{-1}x^2 = (y^{-1}x)^2$$

contrairement à l'hypothèse que d n'est pas le carré d'un élément de K. On a donc $y = 0$, donc $x^2 = 0$, donc $x = 0$, et ceci achève la démonstration du Théorème.

Exemple 2. Prenant $\mathrm{K} = \mathbf{R}$ et $d = -1$ on voit donc que l'anneau $\mathbf{C}$ des nombres complexes est en fait un *corps* commutatif; pour cette raison, on dit que $\mathbf{C}$

est le **corps des nombres complexes**. Étant donné un nombre complexe

$$z = x + iy$$

non nul, son inverse est donné par la relation (14), et comme ici on a

$$N(z) = x^2 + y^2$$

il vient donc

$$z^{-1} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2};$$

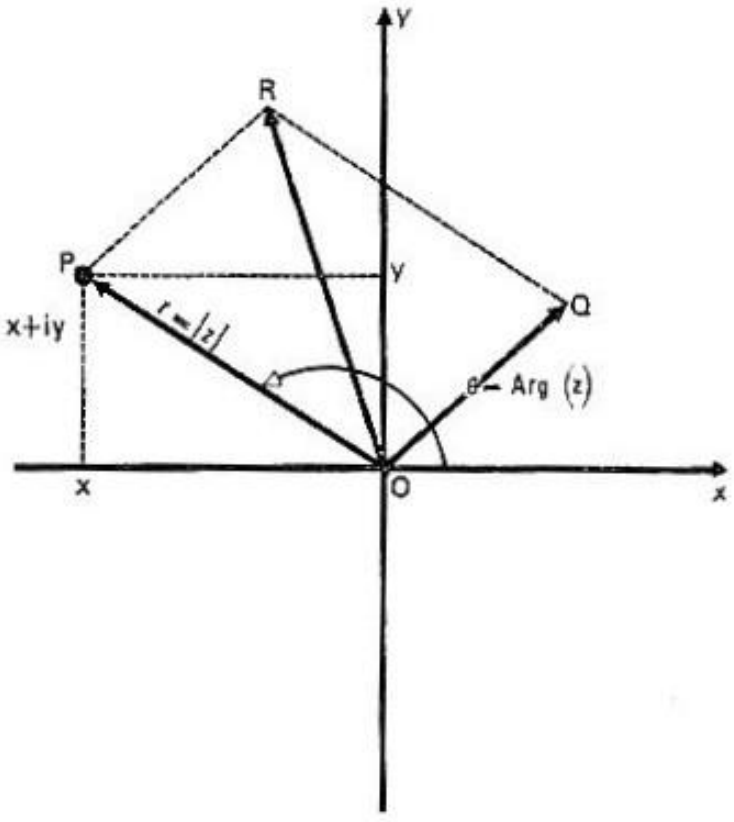
le dénominateur ne peut s'annuler (conformément à la démonstration du Théorème 2) que si $x = y = 0$, i.e. si $z = 0$.

Exemple 3. L'anneau $L = \mathbf{Q}[\sqrt{2}]$ est en fait un *corps* commutatif. On peut d'ailleurs l'identifier à un sous-corps de $\mathbf{R}$: il suffit d'associer à chaque élément $x + y\omega$ de L le nombre réel $x + y\sqrt{2}$ (où $\sqrt{2}$ désigne la racine carrée usuelle du nombre 2); l'application de L dans $\mathbf{R}$ ainsi définie est injective, et compatible avec les lois de composition qui interviennent dans la question. On retrouve ainsi le sous-corps de $\mathbf{R}$ défini au § 8, *Exemple* 6.

6. *Représentation géométrique des nombres complexes*

Considérons dans le plan deux axes de coordonnées rectangulaires Ox et Oy; il est naturel d'associer à tout nombre complexe

$$z = x + iy$$



le point de coordonnées x, y dans le plan; inversement, on associe à tout point P du plan, de coordonnées x et y, le nombre complexe $z = x + iy$, qu'on appelle traditionnellement **l'affixe** du point P. On obtient ainsi une bijection de l'ensemble $\mathbf{C}$ des nombres complexes sur l'ensemble des points du plan (une fois choisis des axes de coordonnées, ce qui a naturellement pour but d'identifier le plan à l'ensemble $\mathbf{R}^2$).

Cette « représentation géométrique » des nombres complexes permet d'interpréter facilement l'addition des nombres complexes : si P et Q sont, dans le plan, des points d'affixes u et v, alors le point R d'affixe $u + v$ est donné par la relation

$$\overrightarrow{OR} = \overrightarrow{OP} + \overrightarrow{OQ},$$

puisque pour additionner des vecteurs d'origine O il suffit d'additionner leurs composantes par rapport aux axes de coordonnées Ox, Oy.

Soit P le point d'affixe $z = x + iy$; le nombre

$$N(z) = \bar{z}z = x^2 + y^2$$

est visiblement donné par

$$N(z) = OP^2,$$

en vertu du théorème de Pythagore. La distance de P au point O, i.e. le nombre

$$r = OP = \sqrt{x^2 + y^2} = \sqrt{\bar{z}z} \tag{15}$$

s'appelle le **module** ou la **valeur absolue** de z, et se désigne par la notation

$$|z|;$$

on a toujours $|z| \geqslant 0$, et on a $|z| = 0$ si et seulement si $z = 0$; de plus, les inégalités classiques entre les côtés d'un triangle montrent, si l'on examine la figure, que l'on a $|u + v| \leqslant |u| + |v|$ quels que soient $u, v \in \mathbf{C}$, et plus généralement

$$\boxed{|z_1 + \cdots + z_n| \leqslant |z_1| + \cdots + |z_n|} \tag{16}$$

quels que soient les nombres complexes $z_1, \ldots, z_n$; ce résultat est souvent cité sous le nom d'**inégalité du triangle**

Supposons $z \neq 0$; l'angle

$$\theta = \widehat{(\overrightarrow{Ox}, \overrightarrow{OP})},$$

qui est un nombre réel modulo 2π (§ 4, *Exemple* 10), s'appelle l'**argument** du nombre complexe $z \neq 0$, et se désigne souvent par la notation

$$\mathrm{Arg}(z).$$

La figure montre que les parties réelle et imaginaire de z sont données, en fonction de la valeur absolue r et de l'argument θ de z, par les formules

$$x = r.\cos\theta, \qquad y = r.\sin\theta, \tag{17}$$

de sorte qu'on peut aussi écrire

$$\boxed{z = r(\cos\theta + i.\sin\theta)}; \tag{18}$$

cette formule s'appelle la **représentation trigonométrique** de z. On notera qu'inversement la relation

$$z = r(\cos\theta + i.\sin\theta) \quad \text{avec} \quad r > 0 \quad \text{implique} \quad r = |z|, \theta = \mathrm{Arg}\,(z);$$

car la relation considérée montre que les parties réelle et imaginaire de z sont

$$x = r.\cos\theta, \qquad y = r.\sin\theta,$$

d'où résulte d'abord que $r^2 = x^2 + y^2$, et, comme r est positif, $r = |z|$; notant θ' l'argument de z il vient alors $\cos\theta = \cos\theta'$, $\sin\theta = \sin\theta'$, d'où $\theta = \theta'$ à un multiple près de 2π, ce qui établit notre assertion.

On va déduire de là les formules

$$\boxed{|z_1 z_2| = |z_1| \cdot |z_2|, \qquad \operatorname{Arg}(z_1 z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2)}, \tag{19}$$

qui permettent de calculer le module et l'argument d'un produit de nombres complexes. Pour cela, écrivons

$$z_1 = r_1(\cos\theta_1 + i.\sin\theta_1), \qquad z_2 = r_2(\cos\theta_2 + i.\sin\theta_2)$$

en mettant en évidence les modules et les arguments de z_1 et z_2; il vient

$$z_1 z_2 = r_1 r_2 (\cos\theta_1 + i.\sin\theta_1)(\cos\theta_2 + i.\sin\theta_2)$$

et comme $r_1 r_2$ est positif, il suffit, pour établir le résultat cherché, de prouver la relation

$$(\cos\theta_1 + i.\sin\theta_1).(\cos\theta_2 + i.\sin\theta_2) = \cos(\theta_1 + \theta_2) + i.\sin(\theta_1 + \theta_2);$$

or, d'après les règles de multiplication des nombres complexes, la partie réelle du premier membre est égale à

$$\cos\theta_1.\cos\theta_2 - \sin\theta_1.\sin\theta_2 = \cos(\theta_1 + \theta_2),$$

et la partie imaginaire à

$$\cos\theta_1.\sin\theta_2 + \sin\theta_1.\cos\theta_2 = \sin(\theta_1 + \theta_2),$$

ce qui établit le résultat annoncé.

Ce résultat s'étend naturellement à un produit de plusieurs facteurs, sous la forme

$$\boxed{\prod_{p=1}^{p=n}(\cos\theta_p + i.\sin\theta_p) = \cos\theta + i.\sin\theta \quad \text{où} \quad \theta = \sum_{p=1}^{p=n}\theta_p}. \tag{20}$$

En particulier, si l'on suppose les θ_p égaux à un même angle θ, il vient la célèbre **formule de De Moivre**

$$\boxed{(\cos\theta + i.\sin\theta)^n = \cos(n\theta) + i.\sin(n\theta)}. \tag{21}$$

Remarque 2. Les formules

$$|z_1 z_2| = |z_1| \cdot |z_2|, \qquad \operatorname{Arg}(z_1 z_2) = \operatorname{Arg}(z_1) + \operatorname{Arg}(z_2)$$

permettent de donner une interprétation « géométrique » de la multiplication

des nombres complexes. Soit

$$z = r.(\cos\theta + i.\sin\theta), \qquad r \geqslant 0,$$

un nombre complexe; à tout point P du plan associons le point P′ dont l'affixe est le produit par z de l'affixe de P (l'application $P \to P'$ correspond donc, dans le plan, à l'application $u \to zu$ de $\mathbf{C}$ dans $\mathbf{C}$); alors la transformation faisant passer de P à P′ est une *similitude* et, de façon précise, c'est le produit de l'homothétie de centre O et de rapport r et de la rotation d'angle θ autour de O.

De même, les relations

$$|z^{-1}| = |z|^{-1}, \qquad \operatorname{Arg}(z^{-1}) = -\operatorname{Arg}(z)$$

(qu'on obtient en faisant $z_1 = z$, $z_2 = z^{-1}$ dans (19)), montrent qu'on passe du point d'affixe $z \neq 0$ au point d'affixe z^{-1} par une inversion de centre O et de puissance 1, suivie d'une symétrie par rapport à l'axe Oy.

Ces interprétations géométriques des opérations sur les nombres complexes sont souvent utiles en Analyse, et Gauss fut le premier, en 1799, à les utiliser systématiquement, notamment pour donner la première démonstration correcte du théorème de d'Alembert — Gauss (§ 33, n° 2).

Quant à l'invention des nombres complexes, elle est de beaucoup antérieure à Gauss, et remonte aux mathématiciens italiens du XVI^e siècle.

7. *Formules de multiplication des fonctions trigonométriques*

La formule de De Moivre

$$(\cos t + i.\sin t)^n = \cos(nt) + i.\sin(nt)$$

établie ci-dessus pour tout nombre réel t permet de calculer $\cos(nt)$ et $\sin(nt)$ en fonction de $\cos t$ et $\sin t$. En effet, le premier membre s'écrit

$$\cos^n t + \binom{n}{1}\cos^{n-1}t.i\,\sin t + \binom{n}{2}\cos^{n-2}t\,(i.\sin t)^2 + \cdots + (i.\sin t)^n$$
$$= \cos^n t + i.\binom{n}{1}\cos^{n-1}t.\sin t - \binom{n}{2}\cos^{n-2}t.\sin^2 t - i.\binom{n}{3}\cos^{n-3}t.\sin^3 t$$
$$+ \binom{n}{4}\cos^{n-4}t.\sin^4 t + i.\binom{n}{5}\cos^{n-5}t.\sin^5 t + \cdots + i^n.\sin^n t;$$

mais puisque les parties réelle et imaginaire de cette expression sont, d'après la formule de De Moivre, égales à $\cos(nt)$ et $\sin(nt)$, il vient

$$(22) \qquad \cos(nt) = \cos^n t - \binom{n}{2}\cos^{n-2}t.\sin^2 t + \binom{n}{4}\cos^{n-4}t.\sin^4 t - \cdots$$

$$(23) \qquad \sin(nt) = \binom{n}{1}\cos^{n-1}t.\sin t - \binom{n}{3}\cos^{n-3}t.\sin^3 t + \binom{n}{5}\cos^{n-5}t.\sin^5 + \cdots$$

d'où les formules cherchées. Dans ces formules, les derniers termes dépendent natu-

rellement de la parité de n puisque i^n est réel si n est pair, et imaginaire pur si n est impair. Par exemple on a

$$\cos(4t) = \cos^4 t - 6\cos^2 t.\sin^2 t + \sin^4 t,$$
$$\cos(5t) = \cos^5 t - 10\cos^3 t.\sin^2 t + 10\cos t.\sin^4 t.$$

Les formules obtenues pour $\cos(nt)$ et $\sin(nt)$ permettent naturellement de calculer

$$(24) \qquad \operatorname{tg}(nt) = \frac{\sin(nt)}{\cos(nt)} = \frac{\binom{n}{1}\cos^{n-1} t.\sin t - \binom{n}{3}\cos^{n-3} t.\sin^3 t + \cdots}{\cos^n t - \binom{n}{2}\cos^{n-2} t.\sin^2 t + \binom{n}{4}\cos^{n-4} t.\sin^4 t - \cdots},$$

et en divisant les deux membres de la dernière fraction par $\cos^n t$ il vient

$$\operatorname{tg}(nt) = \frac{\binom{n}{1}\operatorname{tg} t - \binom{n}{3}\operatorname{tg}^3 t + \binom{n}{5}\operatorname{tg}^5 t - \cdots}{1 - \binom{n}{2}\operatorname{tg}^2 t + \binom{n}{4}\operatorname{tg}^4 t - \cdots}.$$

Par exemple, on a

$$\operatorname{tg}(5t) = \frac{5\operatorname{tg} t - 10\operatorname{tg}^3 t + \operatorname{tg}^5 t}{1 - 10\operatorname{tg}^2 t + 5\operatorname{tg}^4 t}.$$

Appliquons maintenant la formule

$$(25) \qquad 1 + q + q^2 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q}$$

valable sur tout corps pourvu que $q \neq 1$, en prenant

$$q = r^2, \qquad r = \cos t + i.\sin t;$$

il vient d'une part

$$\frac{1 - q^{n+1}}{1 - q} = \frac{r^{2n+2} - 1}{r^2 - 1} = \frac{r^{2n+1} - r^{-1}}{r - r^{-1}} = r^n \frac{r^{n+1} - r^{-n-1}}{r - r^{-1}};$$

or la formule de De Moivre donne

$$r^{n+1} = \cos(n+1)t + i.\sin(n+1)t$$
$$r^{-1} = \cos t - i.\sin t$$
$$r^{-n-1} = \cos(n+1)t - i.\sin(n+1)t$$

d'où

$$\frac{1 - q^{n+1}}{1 - q} = (\cos(nt) + i.\sin(nt))\frac{2i.\sin(n+1)t}{2i.\sin t}$$
$$= (\cos(nt) + i.\sin(nt))\frac{\sin(n+1)t}{\sin t};$$

d'autre part, la formule de De Moivre donne aussi

$$1 + q + \cdots + q^n = \sum_{k=0}^{k=n} (\cos(2kt) + i.\sin(2kt));$$

en égalant les parties réelle et imaginaire de cette expression avec celles de l'expression trouvée auparavant, on obtient donc les formules

$$1 + \cos(2t) + \cos(4t) + \cdots + \cos(2nt) = \cos(nt)\frac{\sin(n+1)t}{\sin t}$$

$$\sin(2t) + \sin(4t) + \cdots + \sin(2nt) = \sin(nt)\frac{\sin(n+1)t}{\sin t}$$

ou, si l'on préfère,

$$(26) \qquad 1 + \cos t + \cos(2t) + \cdots + \cos(nt) = \frac{\cos\left(n\frac{t}{2}\right)\sin(n+1)\frac{t}{2}}{\sin\frac{t}{2}}$$

$$(27) \qquad \sin t + \sin(2t) + \cdots + \sin(nt) = \frac{\sin\left(n\frac{t}{2}\right)\sin(n+1)\frac{t}{2}}{\sin\frac{t}{2}}.$$

Ces formules supposent naturellement $\sin\frac{t}{2}$ non nul, i.e. t non multiple de 2π.

Modules sur un anneau

La notion de module sur un anneau fournit un cadre général et abstrait permettant de traiter les aspects purement algébriques des problèmes « linéaires » qu'on rencontre dans toutes les branches des Mathématiques : théorie des nombres, algèbre linéaire classique, calcul tensoriel, formes différentielles, équations aux dérivées partielles, équations intégrales, géométrie algébrique, fonctions analytiques, topologie algébrique, etc...

Certains résultats de la théorie des modules sur un anneau K ne sont valables que moyennant certaines hypothèses concernant l'anneau K : par exemple, la théorie de la « dimension » suppose que K est un corps.

Par contre, les résultats les plus simples sont valables pour *tout* anneau K; ce sont ces résultats qu'on trouvera dans ce chapitre (§§ 10 à 17). Le lecteur qui trouverait trop général et abstrait le point de vue adopté ici, et qui préférerait supposer que l'anneau de base K, est, par exemple, le corps **R** des nombres réels, ne parviendrait pas à simplifier de façon très substantielle les §§ en question : en faisant l'hypothèse que $K = \mathbf{R}$, on pourra négliger les *Exemples* 5, 6, 9, 10 et le n° 4 du § 10, les *Exemples* 4, 5, 6, 11 du § 11, la *Remarque* 4 du § 16, et l'*Exemple* 2 du § 17, tout le reste demeurant sans aucun autre changement que le remplacement de la lettre K par la lettre **R**.

§ 10. Modules et espaces vectoriels

1. *Définition des modules sur un anneau*

Soit K un anneau; on appelle **module à gauche sur l'anneau** K, ou encore **K-module à gauche**, l'objet formé par un ensemble M, une loi de composition sur M, notée

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

et une application de l'ensemble $K \times M$ dans M, notée

$$(\lambda, x) \mapsto \lambda x,$$

ces données étant assujetties à vérifier les deux conditions que voici :
(M 1) : *L'ensemble* M *muni de la loi de composition* $(x, y) \mapsto x + y$ *est un groupe commutatif;*
(M 2) : *on a les relations*

$$\lambda(\mu x) = (\lambda\mu)x; \qquad 1x = x;$$
$$(\lambda + \mu)x = \lambda x + \mu x; \qquad \lambda(x + y) = \lambda x + \lambda y$$

quels que soient $x, y \in M$ *et* $\lambda, \mu \in K$.

Dans la théorie des modules, l'anneau K est fixé une fois pour toutes, et s'appelle généralement **l'anneau de base**; ses éléments prennent alors le nom de **scalaires** (et on les désignera le plus souvent par des lettres grecques); les éléments des K-modules s'appellent au contraire des **vecteurs** (et on les désigne le plus souvent par des lettres latines, que de nombreuses personnes croient devoir surmonter d'une flèche, ce que l'immense majorité des mathématiciens a cessé de faire depuis longtemps). Mais la distinction entre « scalaires » et « vecteurs » n'a aucun sens mathématique précis (l'*Exemple* 1 ci-dessous montre en effet que les « scalaires » sont des « vecteurs » particuliers...), et son but est plutôt d'aider le lecteur à évoquer des images géométriques familières.

Lorsque l'anneau K est un *corps*, on dit **espace vectoriel à gauche sur** K au lieu de K-module à gauche. En particulier, un espace vectoriel sur le corps **R** des nombres réels s'appele un **espace vectoriel réel**, et un espace vectoriel sur le corps **C** des nombres

complexes un **espace vectoriel complexe**; ces deux notions sont de loin les plus importantes en Analyse et en Physique; par contre, les espaces vectoriels sur des corps arbitraires, et les modules sur l'anneau $\mathbf{Z}$ (cf. *Exemple* 5 ci-dessous) ou sur un « anneau de polynômes », jouent dans beaucoup de branches des Mathématiques un rôle beaucoup plus important que les espaces vectoriels réels ou complexes. Mais même en Physique théorique on utilise des modules sur des anneaux qui ne sont pas des corps, et ne sont pas commutatifs (représentations linéaires du groupe de Lorentz, spineurs, etc...), bien que les physiciens n'utilisent pas encore le *langage* de la théorie des modules.

Notons que les identités figurant dans l'axiome (M 2) des modules impliquent les relations

$$\lambda 0 = 0 \quad \text{pour tout} \quad \lambda \in \mathrm{K}, \qquad 0x = 0 \quad \text{pour tout} \quad x \in \mathrm{M}$$

(dans ces relations le symbole 0 désigne tantôt l'élément nul de K, tantôt l'élément nul du groupe additif M; le lecteur trouvera facilement l'interprétation à choisir pour que les relations écrites aient un sens...). Pour établir la première, on observe que $\lambda 0 + \lambda x = \lambda(0 + x) = \lambda x$ pour tout $x \in \mathrm{M}$; il vient donc $\lambda 0 = \lambda x - \lambda x = 0$ comme annoncé. La seconde relation résulte du fait que $0x + 1x = (0 + 1)x = 1x$, d'où $0x = x - x = 0$.

Nous utiliserons bien entendu sans référence les identités que nous venons de prouver, ainsi que celles qui figurent dans l'axiome (M 2) des modules.

Notons enfin qu'on définit la notion de **K-module à droite** comme suit : on appelle ainsi l'objet formé par un groupe additif M et par une application, notée

$$(x, \lambda) \mapsto x\lambda,$$

de $\mathrm{M} \times \mathrm{K}$ dans M, qui vérifie les conditions exprimées par les identités suivantes :

$$(x\lambda)\,\mu = x(\lambda\mu); \qquad x1 = x;$$
$$x(\lambda + \mu) = x\lambda + x\mu; \qquad (x + y)\,\lambda = x\lambda + y\lambda.$$

On peut montrer facilement (n° 4) que les K-modules *à droite* ne sont autres que les modules *à gauche* sur un anneau déduit de K par un procédé très simple (et du reste identique à K si K est commutatif, de sorte que la distinction entre les deux notions n'a d'intérêt que pour les anneaux non commutatifs). Il nous arrivera d'utiliser tantôt le langage des modules à gauche, tantôt celui des modules à droite; il va de soi qu'on peut passer de l'un à l'autre par des traductions triviales.

Nous allons maintenant donner quelques exemples importants de modules et d'espaces vectoriels.

2. *Exemples de modules*

Exemple 1. Pour tout anneau K et tout entier $n \geqslant 1$, on peut considérer l'ensemble

$$\mathrm{K}^n = \mathrm{K} \times \cdots \times \mathrm{K} \quad (n \text{ facteurs})$$

comme un K-module à gauche, en posant, par définition,

$$(\xi_1, \ldots, \xi_n) + (\eta_1, \ldots, \eta_n) = (\xi_1 + \eta_1, \ldots, \xi_n + \eta_n)$$
$$\lambda.(\xi_1, \ldots, \xi_n) = (\lambda\xi_1, \ldots, \lambda\xi_n);$$

le fait que l'axiome (M 1) soit vérifié résulte du § 7, n° 2 (produit direct de groupes), et le lecteur vérifiera facilement, en utilisant les axiomes des anneaux, les identités figurant dans l'axiome (M 2) des modules.

Par la suite, quand nous parlerons de K^n comme d'un K-module à gauche, ce sera toujours du module ci-dessus qu'il s'agira.

Pour $n = 1$, la construction précédente permet de regarder K lui-même comme un K-module à gauche (ce qui montre que les « scalaires » sont aussi des « vecteurs »...).

On peut naturellement regarder aussi K^n comme un K-module *à droite*; il suffit pour cela de définir l'addition dans K^n comme ci-dessus, et de poser

$$(\xi_1, \ldots, \xi_n).\lambda = (\xi_1\lambda, \ldots, \xi_n\lambda).$$

C'est le K-module *à droite* K^n qui intervient naturellement dans la théorie des équations linéaires comme on le verra (mais il est clair, encore une fois, que la distinction est sans intérêt si K est commutatif !).

Exemple 2. Prenons $K = \mathbf{R}$, corps des nombres réels, et pour M l'ensemble des vecteurs usuels d'origine donnée O dans l'espace usuel; définissons la somme de deux vecteurs par la règle du parallélogramme, et le produit d'un vecteur x d'origine O et d'un nombre réel λ comme le vecteur obtenu en soumettant x à l'homothétie de centre O et de rapport λ; on obtient alors un espace vectoriel réel.

Bien entendu, on devrait prouver ici que les axiomes (M 1) et (M 2) sont vérifiés, ce qui ne peut se faire que dans le cadre de la Géométrie Élémentaire (et pour cause, puisque nous n'avons donné ici aucune définition mathématique rigoureuse de la notion usuelle de « vecteur »).

Cet Exemple est évidemment l'un de ceux qui ont donné naissance à la notion générale d'espace vectoriel ou de module, et explique l'emploi du mot « vecteur » pour désigner les éléments d'un module.

Exemple 3. Dans le plan rapporté à deux axes de coordonnées Ox et Oy, considérons l'ensemble M des vecteurs d'origine O dont les composantes dans le système de coordonnées considéré sont des nombres *rationnels*; il est clair que si $x, y \in M$ on a aussi $x + y \in M$, et que si $x \in M$ et $\lambda \in \mathbf{Q}$ on a aussi $\lambda x \in M$. Ceci permet de regarder M comme un espace vectoriel sur le corps $\mathbf{Q}$ des nombres rationnels.

Exemple 4. Soient K un anneau, M un K-module à gauche (par exemple K lui-même), et X un ensemble quelconque. Désignons par E l'ensemble de toutes les applications

$$f : X \to M;$$

on va en faire un K-module à gauche. Pour cela on doit définir la somme $f + g$ de deux applications de X dans M : ce sera la fonction $f(x) + g(x)$, dont la valeur en chaque $x \in X$ s'obtient en additionnant (dans M) les valeurs de f et g en x; on doit aussi définir le produit λf d'un scalaire $\lambda \in K$ et d'une

application f de X dans M : ce sera la fonction $\lambda f(x)$, dont la valeur en chaque $x \in X$ s'obtient en multipliant par λ la valeur de f en x.

On laisse au lecteur, à titre d'exercice, le soin de vérifier en détail les conditions (M 1) et (M 2) figurant dans la définition des modules.

On notera que si $M = K$ et si l'on prend $X = \{1, 2, \ldots, n\}$, une application f de X dans M n'est autre qu'une suite $(\xi_1, \ldots, \xi_n)$ d'éléments de K — à savoir $\xi_1 = f(1), \ldots, \xi_n = f(n)$; on retrouve alors le module K^n de l'*Exemple* 1.

Exemple 5. Montrons que *tout groupe commutatif G peut être regardé comme un* **Z***-module à gauche*. Pour cela, on écrit G additivement, ce qui permet déjà de définir la somme de deux éléments de G — et l'axiome (M 1) est alors trivialement vérifié. Il reste à définir le produit nx d'un $n \in \mathbf{Z}$ et d'un $x \in G$, ce qu'on fait comme au § 7, i.e. en posant

$$nx = \begin{cases} x + \cdots + x \ (n \text{ facteurs}) & \text{si } n \geqslant 1 \\ 0 & \text{si } n = 0 \\ (-n)\,(-x) & \text{si } n \leqslant -1. \end{cases}$$

L'axiome (M 2) se réduit alors aux règles de calcul établies dans l'*Exemple* 9 et la *Remarque* 1 du § 7.

Si le groupe G était écrit multiplicativement, il faudrait bien entendu définir la « somme » de deux éléments x et y de G comme étant xy, et le « produit » d'un $x \in G$ par un entier rationnel n comme étant x^n ; il n'y a aucune différence avec ce qui précède, si ce n'est dans les *notations* adoptées.

Enfin, on peut facilement vérifier que *tout* **Z**-module s'obtient, par le procédé ci-dessus, à partir d'un groupe additif.

Cet Exemple montre que *la théorie des modules contient, entre autres, celle des groupes commutatifs*, ce qui n'est pas le cas de la théorie des espaces vectoriels (et encore moins si possible de celle des espaces vectoriels réels).

Exemple 6. Soit K un sous-anneau d'un anneau L ; on peut alors regarder L comme un K-module à gauche, en définissant les opérations fondamentales des modules à l'aide de l'addition et de la multiplication données sur L. Autrement dit, si l'on considère deux éléments x et y de L, leur somme en tant que « vecteurs » sera simplement leur somme en tant qu'éléments de l'anneau L ; et pour $\lambda \in K$ et $x \in L$, le produit du « scalaire » λ et du « vecteur » x sera le produit, dans l'anneau L, de λ par x. L'axiome (M 1) est ici vérifié parce que l'anneau L devient un groupe commutatif si l'on fait abstraction de son opération de multiplication ; et quant à (M 2), il se déduit évidemment des règles d'associativité et de distributivité dans l'anneau L.

Par exemple, on peut regarder le corps **R** comme un espace vectoriel sur **Q**, et le corps **C** comme un espace vectoriel sur **R**, ou sur **Q**.

Exemple 7. Prenons $K = \mathbf{R}$ et formons l'ensemble M de toutes les applications

$$f : \mathbf{R} \to \mathbf{R}$$

(fonctions réelles d'une variable réelle) qui sont continues partout. On démontre en Analyse que si f et g sont deux telles fonctions, la fonction $f + g$ est elle aussi partout continue, donc appartient à M ; et que si $f \in M$, alors $\lambda f \in M$ pour tout scalaire $\lambda \in \mathbf{R}$ (les fonctions $f + g$ et λf sont définies comme dans l'*Exemple* 4 ci-dessus). Il est immédiat de voir que le triplet formé par

l'ensemble M, l'application $(f, g) \mapsto f + g$ de $M \times M$ dans M, et l'application $(\lambda, f) \mapsto \lambda f$ de $\mathbf{R} \times M$ dans M, est un espace vectoriel réel.

Cet Exemple (qui est à l'origine de l'introduction des espaces vectoriels en Analyse) est susceptible de nombreuses variantes; au lieu d'imposer aux fonctions f considérées d'être partout continues, on peut exiger qu'elles soient continues en un point donné, ou dérivables en un point donné, ou qu'elles admettent partout une dérivée seconde continue, etc...

3. *Sous-modules, sous-espaces vectoriels*

Soit M un module à gauche sur un anneau K. On appelle **sous-module de M** toute partie M′ de M vérifiant les deux conditions que voici :

(i) : M′ *est un sous-groupe du groupe additif* M;

(ii) : *les relations* $x \in M'$ *et* $\lambda \in K$ *impliquent* $\lambda x \in M'$.

Pour vérifier qu'une partie M′ de M est un sous-module, on doit vérifier que M′ est non vide (pratiquement on vérifie que $0 \in M'$), et que l'on a

$$\lambda x + \mu y \in M' \quad \text{quels que soient} \quad \lambda, \mu \in K \quad \text{et} \quad x, y \in M'.$$

Cette condition est évidemment nécessaire. Inversement, supposons-la vérifiée; faisant $\mu = 0$ on obtient déjà la condition (ii) ci-dessus; pour obtenir la condition (i), il suffit de faire $\lambda = 1$, $\mu = -1$, et de remarquer que dans un module on a

$$-x = (-1)\,x \quad \text{pour tout } x \in M$$

(en effet : $(-1)\,x + x = (-1)\,x + (+1)\,x = (-1+1)\,x = 0x = 0$).

Un module M possède toujours au moins deux sous-modules, à savoir M lui-même, et l'ensemble réduit au seul vecteur 0.

Soit M′ un sous-module d'un module M; la condition (i) ci-dessus permet déjà de regarder M′ comme un groupe additif ; la condition (ii) permet en outre de définir une application $(\lambda, x) \mapsto \lambda x$ de $K \times M'$ dans M′, et les identités qui figurent dans l'axiome (M 2) des modules, étant vérifiées dans M, le sont *a fortiori* dans M′. Par suite, *on peut regarder tout sous-module* M′ *de* M *comme un* K*-module à gauche.*

Lorsque K est un *corps*, on dit **sous-espace vectoriel** au lieu de sous-module.

Exemple 8. Prenons pour M l'espace vectoriel réel de l'*Exemple* 2 ci-dessus; on a alors dans M (en dehors de $\{0\}$ et de M lui-même) deux sortes de sous espaces vectoriels : *a*) l'ensemble des vecteurs d'origine O portés par une *droite* donnée passant par O; *b*) l'ensemble des vecteurs d'origine O contenus dans un *plan* donné passant par O. On voit du reste facilement qu'il n'y a pas d'autres sous-espaces vectoriels de M que ceux qu'on vient de décrire.

Exemple 9. Soit K un anneau et regardons-le (*Exemple* 1) comme un K-module à gauche; ses sous-modules sont donc les parties I de K qui sont non vides et telles que l'on ait

$$ux + vy \in I \quad \text{quels que soient} \quad u, v \in K \quad \text{et} \quad x, y \in I;$$

ce sont donc les *idéaux à gauche* de K définis au § 8, n° 6.

Exemple 10. Soit G un groupe commutatif écrit additivement, et regardons G comme un **Z**-module (*Exemple* 5) ; les sous-modules de G ne sont autres alors que ses sous-groupes, car si H est un sous-groupe de G on a $nx \in H$ pour tout $x \in H$ et tout $n \in \mathbf{Z}$.

Le résultat suivant est souvent utile :

THÉORÈME 1. *Soient* L *un module sur un anneau et* $(M_i)_{i \in I}$ *une famille de sous-modules de* L. *Alors l'intersection des* M_i *est encore un sous-module de* L. *Pour que la réunion des* M_i *soit un sous-module de* L, *il suffit que, quels que soient* $i, j \in I$, *il existe un* $k \in I$ *tel que l'on ait* $M_i \subset M_k$ *et* $M_j \subset M_k$.

Ce résultat se démontre exactement comme le Théorème 1 du § 7, et nous laisserons donc au lecteur le soin d'en rédiger lui-même une démonstration détaillée.

Il va de soi qu'en général une réunion de sous-modules n'est pas un sous-module : par exemple, dans la situation classique (*Exemple* 8), la réunion de deux droites distinctes passant par l'origine n'est pas un plan...

4. *Modules à droite et modules à gauche* (*)

Soit K un anneau. Nous allons construire un nouvel anneau qu'on appelle l'**opposé** de K, et qu'on désigne par la notation

$$K^o;$$

comme on le montrera ensuite, les modules *à droite* sur K ne sont autres que les module *à gauche* sur K^o.

Pour construire K^o, on doit se donner un ensemble, et deux lois de composition sur cet ensemble, une « addition » et une « multiplication ». Par définition, l'*ensemble* K^o sera *l'ensemble* K (les anneaux K et K^o ont donc les mêmes éléments), et l'*addition* sur K^o sera l'*addition* sur K (la somme $x + y$ de deux éléments de K a donc la même valeur, qu'on la calcule dans l'anneau K ou dans l'anneau K^o). Par contre, la *multiplication* dans K^o, au lieu d'être la multiplication $(x, y) \mapsto xy$ donnée sur K, sera l'application $(x, y) \mapsto yx$; autrement dit, si l'on désigne par xy le produit de deux éléments dans l'anneau K, et par $x * y$ leur produit dans l'anneau K^o, on a la relation

$$x * y = yx.$$

Il est facile de voir que l'ensemble K^o (= K !) muni des deux opérations qu'on vient de définir, est un anneau; par exemple la formule

$$(x + y) * z = x * z + y * z$$

se ramène évidemment à la relation $z(x + y) = zx + zy$ dans l'anneau K...

Il va de soi que, si K est un anneau commutatif, l'anneau K^o est identique à l'anneau K. La construction de K^o n'a donc d'intérêt que dans le cas non commutatif.

(*) Ce n° peut être négligé en première lecture; il n'est utilisé qu'à la fin du § 16, et le lecteur débutant pourra supposer à ce moment l'anneau K commutatif.

Soit maintenant M un module *à droite* sur l'anneau K. Définissons une application

$$(\lambda, \ x) \mapsto \lambda * x$$

de $K^o \times M$ dans M en posant

$$\lambda * x = x\lambda \quad \text{pour tout } x \in M \text{ et tout } \lambda \in K.$$

Alors le groupe additif M, muni de l'application qu'on vient de construire, est un module *à gauche* sur l'anneau K^o opposé à K. On a en effet

$$\lambda * (x + y) = (x + y)\lambda = x\lambda + y\lambda = \lambda * x + \lambda * y,$$
$$(\lambda + \mu) * x = x(\lambda + \mu) = x\lambda + x\mu = \lambda * x + \mu * x,$$
$$\lambda * (\mu * x) = (\mu * x)\lambda = (x\mu)\lambda = x(\mu\lambda) = (\mu\lambda) * x = (\lambda * \mu) * x,$$

et enfin

$$1 * x = x1 = x,$$

ce qui établit le résultat annoncé.

Les constructions qu'on vient d'exposer montrent que les résultats établis pour les modules *à gauche* (resp. à droite) s'appliquent automatiquement aux modules *à droite* (resp. à gauche) : il suffit de passer de l'anneau de base donné à l'anneau opposé. On voit aussi que, lorsque l'anneau de base est *commutatif*, il est parfaitement indifférent, dans la théorie des modules, de placer les scalaires à gauche des vecteurs plutôt qu'à droite; c'est une simple question de *convention d'écriture*, qui n'a rien à voir avec la réalité mathématique elle-même, et dont on aurait tout à fait tort d'être l'esclave au point de ne pouvoir passer de l'écriture « droite » à l'écriture « gauche » et vice-versa.

1. *Combinaisons linéaires*

Soient $a_1, \ldots, a_n$ des éléments d'un module à gauche M sur un anneau K; on appelle **combinaison linéaire de** $a_1, \ldots, a_n$ tout vecteur $x \in M$ possédant la propriété suivante : il existe des scalaires $\xi_1, \ldots, \xi_n \in K$ tels que l'on ait

$$x = \xi_1 a_1 + \cdots + \xi_n a_n.$$

On a bien entendu une notion analogue pour les modules à droite.

Exemple 1. Dans K^n (§ 10, *Exemple* 1) considérons les vecteurs

$$\begin{aligned} e_1 &= (1, 0, 0, \ldots, 0, 0) \\ e_2 &= (0, 1, 0, \ldots, 0, 0) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ e_n &= (0, 0, 0, \ldots, 0, 1); \end{aligned}$$

on a visiblement

$$\begin{aligned} \xi_1 e_1 &= (\xi_1, 0, 0, \ldots, 0, 0) \\ \xi_2 e_2 &= (0, \xi_2, 0, \ldots, 0, 0) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \xi_n e_n &= (0, 0, 0, \ldots, 0, \xi_n) \end{aligned}$$

et en ajoutant les résultats obtenus on trouve donc

$$\xi_1 e_1 + \cdots + \xi_n e_n = (\xi_1, \ldots, \xi_n). \tag{1}$$

Autrement dit, *tout* élément de K^n est combinaison linéaire des vecteurs $e_1, \ldots, e_n$; on a même un résultat plus précis : étant donné un vecteur $x \in K^n$, il existe *un et un seul* système de scalaires $\xi_1, \ldots, \xi_n$ tel que l'on ait

$$x = \xi_1 e_1 + \cdots + \xi_n e_n.$$

Exemple 2. Considérons le K-module *à droite* K^p pour un entier $p \geqslant 1$, et soient

$$\begin{aligned} a_1 &= (\alpha_{11}, \alpha_{21}, \ldots, \alpha_{p1}) \\ a_2 &= (\alpha_{12}, \alpha_{22}, \ldots, \alpha_{p2}) \\ &\cdots\cdots\cdots\cdots\cdots \\ a_n &= (\alpha_{1n}, \alpha_{2n}, \ldots, \alpha_{pn}) \end{aligned}$$

des éléments donnés de ce module. Soit

$$b = (\beta_1, \beta_2, \ldots, \beta_p)$$

un élément de K^p; alors la relation

$$b = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n \tag{2}$$

équivaut au **système de p équations linéaires à n inconnues** $\xi_1, \ldots, \xi_n$ que voici :

$$\left\{ \begin{array}{l} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \cdots + \alpha_{1n}\xi_n = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{p1}\xi_1 + \alpha_{p2}\xi_2 + \cdots + \alpha_{pn}\xi_n = \beta_p. \end{array} \right. \tag{3}$$

On a en effet

$$\begin{aligned} a_1\xi_1 &= (\alpha_{11}\xi_1, \ldots, \alpha_{p1}\xi_1) \\ &\cdots\cdots\cdots\cdots\cdots \\ a_n\xi_n &= (\alpha_{1n}\xi_n, \ldots, \alpha_{pn}\xi_n), \end{aligned}$$

de sorte que les premiers membres des relations (3) sont les composantes (*) du second membre de la relation (2).

Cet Exemple est à l'origine de la théorie « géométrique » des systèmes d'équations linéaires.

THÉORÈME 1. *Soient $a_1, \ldots, a_n$ des éléments d'un* K*-module à gauche* M, *et* M′ *l'ensemble des combinaisons linéaires de $a_1, \ldots, a_n$; alors* M′ *est le plus petit sous-module de* M *contenant $a_1, \ldots, a_n$.*

Tout d'abord, la relation

$$a_1 = 1.a_1 + 0.a_2 + \cdots + 0.a_n$$

et des relations analogues pour $a_2, \ldots, a_n$ montrent que M′ contient $a_1, \ldots, a_n$. D'autre part, tout sous-module de M contenant $a_1, \ldots, a_n$ contient aussi $\xi_1a_1, \ldots, \xi_na_n$ quels que soient $\xi_1, \ldots, \xi_n \in K$, donc contient $\xi_1a_1 + \ldots + \xi_na_n$; ainsi, tout sous-module de M contenant les vecteurs a_i $(1 \leqslant i \leqslant n)$ contient M′.

Pour achever la démonstration, il reste donc à faire voir que M′ est effectivement un sous-module de M. Or soient

$$x = \xi_1a_1 + \cdots + \xi_na_n, \qquad y = \eta_1a_1 + \cdots + \eta_na_n$$

deux éléments de M′; un calcul trivial montre que

$$\lambda x + \mu y = \zeta_1a_1 + \cdots + \zeta_na_n$$

(*) Dans K^n on appelle composantes d'un vecteur $(\xi_1, \ldots, \xi_n)$ les scalaires $\xi_1, \ldots, \xi_n$; voir l'*Exemple* 12 ci-dessous.

avec

$$\zeta_1 = \lambda\xi_1 + \mu\eta_1, \ldots, \zeta_n = \lambda\xi_n + \mu\tau_n,$$

de sorte que $\lambda x + \mu y \in M'$ quels que soient les scalaires λ et μ, ce qui termine la démonstration.

Le sous-module M′ du Théorème 1 s'appelle le **sous-module de M engendré par** $a_1, \ldots, a_n$; lorsque $K = \mathbf{Z}$, de sorte que M est simplement un groupe commutatif écrit additivement, M′ n'est autre que le sous-groupe de M engendré par la partie $B = \{a_1, \ldots, a_n\}$ de M (§ 7, n° 4).

2. *Modules de type fini*

Soient M un K-module à gauche et M′ un sous-module de M; on dit que M′ est **de type fini** s'il existe des vecteurs $a_1, \ldots, a_n \in M'$ en nombre fini, tels que M′ soit engendré par ces vecteurs; on dit alors que $a_1, \ldots, a_n$ forment un **système de générateurs** de M′.

Cette définition s'applique en particulier au module M lui-même; autrement dit, on dit qu'un module est de type fini s'il contient des vecteurs $a_1, \ldots, a_n$ en nombre fini tels que tout $x \in M$ soit combinaison linéaire de $a_1, \ldots, a_n$.

Lorsque l'anneau de base K est un corps, on dit **espace vectoriel de dimension finie** au lieu de module de type fini.

Exemple 3. L'*Exemple* 1 montre que K^n est un K-module de type fini.

Exemple 4. Lorsque $K = \mathbf{Z}$, la notion de module de type fini se réduit à celle de groupe (commutatif) de type fini, introduite au § 7, n° 4.

Exemple 5. $\mathbf{Q}$ n'est pas de type fini comme $\mathbf{Z}$-module (§ 7, *Exemple* 10); par contre, $\mathbf{Q}$ est de type fini comme espace vectoriel sur $\mathbf{Q}$, vu l'*Exemple* 3 ci-dessus avec $K = \mathbf{Q}$, $n = 1$.

Exemple 6. Soit K un anneau; cherchons les sous-modules de type fini de K (regardé comme K-module à gauche); ce sont les idéaux à gauche (§ 10, *Exemple* 9) I de K qui possèdent la propriété suivante : *il existe des éléments* $a_1, \ldots, a_n$ *de* I, *en nombre fini, tels que tout* $x \in I$ *puisse se mettre sous la forme*

$$x = u_1a_1 + \ldots + u_na_n$$

pour un choix convenable de $u_1, \ldots, u_n \in K$. Un tel idéal est appelé un **idéal à gauche de type fini** de l'anneau K. Il est clair par exemple que tout idéal principal de K est de type fini, mais la réciproque est en général inexacte.

Étant donnés des éléments $a_1, \ldots, a_n$ de K, l'ensemble des éléments de K qui peuvent se mettre sous la forme $u_1a_1 + \ldots + u_na_n$ (autrement dit, le sous-module de K engendré par $a_1, \ldots, a_n$) s'appelle l'**idéal à gauche de K engendré par** $a_1, \ldots, a_n$.

Exemple 7. Les espaces vectoriels réels décrits au § 10, *Exemple* 7, ne sont pas de dimension finie.

3. Relations linéaires

Soient $a_1, \ldots, a_n$ des éléments d'un K-module à gauche M ; soit x une combinaison linéaire de ces vecteurs. Considérons deux façons d'écrire x comme combinaison linéaire des vecteurs donnés, soient

$$x = \xi_1 a_1 + \cdots + \xi_n a_n = \eta_1 a_1 + \cdots + \eta_n a_n ;$$

par différence, on trouve aussitôt la relation

$$(\xi_1 - \eta_1) a_1 + \ldots + (\xi_n - \eta_n) a_n = 0.$$

Ceci justifie l'introduction des notions suivantes.

On appelle **relation linéaire entre les vecteurs** $a_1, \ldots, a_n$ tout élément $(\lambda_1, \ldots, \lambda_n)$ du module K^n tel que l'on ait

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0.$$

La relation linéaire $(0, \ldots, 0)$ est dite **triviale.** Enfin, on dit que les vecteurs $a_1, \ldots, a_n$ sont **linéairement indépendants,** ou que la famille $(a_i)_{1 \leqslant i \leqslant n}$ est **libre,** s'il n'existe pas d'autre relation linéaire entre $a_1, \ldots, a_n$ que la relation linéaire triviale.

Dire que $a_1, \ldots, a_n$ sont linéairement indépendants signifie donc que la relation

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0 \quad \text{implique} \quad \lambda_1 = \cdots = \lambda_n = 0 ;$$

dire, au contraire, qu'ils ne sont pas linéairement indépendants (on dit alors que les vecteurs $a_1, \ldots, a_n$ sont **liés**) signifie qu'il existe des scalaires $\lambda_1, \ldots, \lambda_n$ *non tous nuls* tels que l'on ait

$$\lambda_1 a_1 + \cdots + \lambda_n a_n = 0.$$

Des vecteurs $a_1, \ldots, a_n$ linéairement indépendants sont nécessairement deux à deux distincts, car si l'on avait par exemple $a_1 = a_2$ l'élément $(1, -1, 0, \ldots, 0)$ de K^n serait évidemment une relation linéaire non triviale entre $a_1, \ldots, a_n$. Mais il ne suffit pas que les a_i soient deux à deux distincts pour qu'ils soient linéairement indépendants.

La façon dont nous avons été conduits à introduire la notion d'indépendance linéaire prouve immédiatement le résultat suivant :

THÉORÈME 2. *Soient $a_1, \ldots, a_n$ des éléments d'un* K*-module à gauche* M, *et x une combinaison linéaire des vecteurs $a_1, \ldots, a_n$. Les propriétés suivantes sont équivalentes :*

a) il existe un seul $(\xi_1, \ldots, \xi_n) \in K^n$ tel que l'on ait

$$x = \xi_1 a_1 + \cdots + \xi_n a_n ;$$

b) les vecteurs $a_1, \ldots, a_n$ sont linéairement indépendants.

Il suffit de remarquer que, si $(\lambda_1, \ldots, \lambda_n)$ est une relation linéaire entre $a_1, \ldots, a_n$,

on a $\lambda_1 a_1 + \cdots + \lambda_n a_n = 0$ et par suite

$$\xi_1 a_1 + \cdots + \xi_n a_n = (\xi_1 + \lambda_1)\, a_1 + \cdots + (\xi_n + \lambda_n)\, a_n;$$

et il est clair en fait que cette propriété *caractérise* les relations linéaires entre les vecteurs $a_1, \ldots, a_n$ donnés.

Exemple 8. Dans K^n, les vecteurs $e_1, \ldots, e_n$ de l'*Exemple* 1 sont linéairement indépendants.

Exemple 9. Prenons $K = \mathbf{R}$ et l'espace vectoriel M formé par les vecteurs d'origine donnée O dans l'espace usuel (§ 10, *Exemple* 2). Pour que des vecteurs $a_1, \ldots, a_n \in M$ soient linéairement indépendants, il faut et il suffit: (1) qu'ils ne soient pas nuls dans le cas $n = 1$; (2) qu'ils ne soient pas portés par une même droite dans le cas $n = 2$; (3) qu'ils ne soient pas contenus dans un même plan dans le cas $n = 3$. Pour $n \geqslant 4$, les vecteurs $a_1, \ldots, a_n$ ne peuvent jamais être linéairement indépendants (car s'ils l'étaient, il en serait déjà ainsi des trois vecteurs a_1, a_2, a_3, et les autres, par exemple a_4, seraient alors des combinaisons linéaires de ces trois vecteurs, ce qui contredit évidemment l'indépendance linéaire).

Exemple 10. Prenons $K = \mathbf{R}$ et pour M l'espace vectoriel de toutes les applications $f: \mathbf{R} \to \mathbf{R}$ (§ 10, *Exemple* 4 où l'on fait $X = M = K = \mathbf{R}$); soient $f_1, \ldots, f_n$ des éléments de M, i.e. des fonctions d'une variable réelle t, à valeurs réelles. Pour $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$ la fonction $f = \lambda_1 f_1 + \cdots + \lambda_n f_n$ est donnée par

$$f(t) = \lambda_1 f_1(t) + \cdots + \lambda_n f_n(t)$$

pour tout $t \in \mathbf{R}$. Une relation linéaire entre $f_1, \ldots, f_n$ est donc une suite $(\lambda_1, \ldots, \lambda_n) \in \mathbf{R}^n$ de n nombres réels tels que l'on ait

$$\lambda_1 f_1(t) + \cdots + \lambda_n f_n(t) = 0 \quad \text{pour tout} \quad t \in \mathbf{R}.$$

Considérons par exemple les $n + 1$ fonctions $1, t, t^2, \ldots, t^n$; une relation linéaire entre ces fonctions est un système de $n + 1$ nombres réels $c_0, c_1, \ldots, c_n$ vérifiant

$$c_0 + c_1 t + \cdots + c_n t^n = 0$$

pour tout $t \in \mathbf{R}$. On verra plus loin, en étudiant le nombre de racines d'une équation algébrique, que la relation précédente implique $c_0 = \cdots = c_n = 0$, de sorte que, *quel que soit n, les fonctions* $1, t, \ldots, t^n$ *sont linéairement indépendantes* en tant qu'éléments de l'espace vectoriel réel M.

Exemple 11. Comme $\mathbf{Q}$ est un sous-corps de $\mathbf{C}$, on peut (§ 10, *Exemple* 6) considérer $\mathbf{C}$ comme un espace vectoriel sur $\mathbf{Q}$. Pour un $z \in \mathbf{C}$ considérons alors les $n + 1$ puissances $1, z, \ldots, z^n$ de z; dire qu'il existe, entre ces $n + 1$ éléments de $\mathbf{C}$, une relation linéaire non triviale (à « coefficients » dans $\mathbf{Q}$) signifie qu'il existe des nombres *rationnels non tous nuls* $c_0, c_1, \ldots, c_n$ tels que z vérifie l'équation

$$c_0 + c_1 z + \ldots + c_n z^n = 0.$$

S'il en est ainsi pour au moins une valeur de n, on dit que z est un **nombre**

algébrique (*). Dans le cas contraire (i.e. si z ne vérifie aucune équation algébrique non triviale à coefficients *rationnels*) on dit que z est un **nombre transcendant**; c'est le cas de $\pi = 3,14159...$

4. *Modules libres, bases*

Soit M un module à gauche sur un anneau K; on dit que M est **libre de type fini** s'il existe des éléments $a_1, \ldots, a_n$ de M en nombre fini qui sont *linéairement indépendants et engendrent* M. On dit alors que $a_1, \ldots, a_n$ forment une **base** de M (une base est donc un système fini de générateurs linéairement indépendants).

Soient $a_1, \ldots, a_n$ des éléments d'un K-module à gauche M; dire qu'ils engendrent M, c'est dire que pour tout $x \in M$ la relation

$$x = \xi_1 a_1 + \cdots + \xi_n a_n$$

est vérifiée pour *au moins* un $(\xi_1, \ldots, \xi_n) \in K^n$; d'autre part, dire que $a_1, \ldots, a_n$ sont linéairement indépendants signifie que, pour chaque $x \in M$, la relation ci-dessus est vérifiée pour un élément de K^n *au plus*.

Par suite, pour que les vecteurs $a_1, \ldots, a_n$ forment une base de M, il faut et il suffit que, pour chaque $x \in M$, il existe *un et un seul* $(\xi_1, \ldots, \xi_n) \in K^n$ tel que

$$x = \xi_1 a_1 + \cdots + \xi_n a_n;$$

les scalaires $\xi_1, \ldots, \xi_n$ s'appellent alors les **coordonnées** ou les **composantes de x par rapport à la base** $a_1, \ldots, a_n$ de M.

Ce qui précède montre que les coordonnées de x sont des fonctions de x, à valeurs dans K; notons-les $f_1, \ldots, f_n$, de sorte qu'on a

$$x = f_1(x) a_1 + \cdots + f_n(x) a_n$$

pour tout $x \in M$. On dit que les applications

$$f_i : M \to K$$

sont les **fonctions coordonnées du module M par rapport à la base** $a_1, \ldots, a_n$. On a les relations

$$f_i(a_j) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j, \end{cases}$$

car la relation

$$a_j = 0.a_1 + \cdots + 0.a_{j-1} + 1.a_j + 0.a_{j+1} + \cdots + 0.a_n$$

permet de calculer immédiatement les coordonnées du vecteur a_j par rapport à la base considérée.

(*) Rappelons que, comme on l'a déjà dit au § 5, n° 8, la notion de nombre algébrique définie ici n'a aucun rapport avec celle qu'on désigne sous ce nom dans l'Enseignement Secondaire (et qui n'est autre que celle de nombre réel).

On peut montrer que les nombres algébriques forment un sous-corps du corps C des nombres complexes (§ 26, n° 2); c'est l'étude de ces nombres au XIX^e^ siècle (surtout par Galois et les grands mathématiciens de l'école allemande : Gauss, Kummer, Jacobi, Lejeune-Dirichlet, Dedekind, Kronecker, Hilbert) qui a donné naissance à toute l'Algèbre moderne.

On a d'autre part les identités

$$f_i(x + y) = f_i(x) + f_i(y), \qquad f_i(\lambda x) = \lambda f_i(x);$$

en effet, les relations

$$x = f_1(x)a_1 + \cdots + f_n(x)a_n, \qquad y = f_1(y)a_1 + \cdots + f_n(y)a_n,$$

additionnées membre à membre, impliquent

$$x + y = [f_1(x) + f_1(y)]a_1 + \cdots + [f_n(x) + f_n(y)]a_n,$$

ce qui montre, comme annoncé, que les coordonnées du vecteur $x + y$ sont les scalaires $f_i(x) + f_i(y)$. La seconde relation s'établit de façon analogue.

Autrement dit, *pour additionner deux vecteurs, on additionne leurs coordonnées de même rang; et pour multiplier un vecteur par un scalaire, on multiplie chacune de ses coordonnées par ce scalaire.*

Exemple 12. Le K-module à gauche K^n est libre de type fini, et l'*Exemple* 1 montre que les vecteurs $e_1, \ldots, e_n$ forment une base de ce module; on l'appelle la **base canonique** de K^n. Étant donné un élément

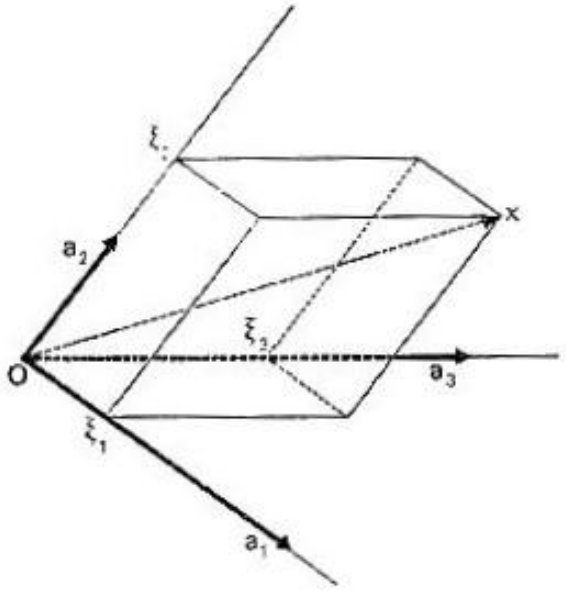
$$x = (\xi_1, \ldots, \xi_n),$$

la relation

$$x = \xi_1 e_1 + \cdots + \xi_n e_n$$

établie au n° 1 montre que les coordonnées de x par rapport à la base canonique de K^n sont les scalaires $\xi_1, \ldots, \xi_n$.

Exemple 13. Reprenons l'*Exemple* 9 ci-dessus; pour que $a_1, \ldots, a_n \in M$ forment une base de M, il faut et il suffit que $n = 3$ et que les vecteurs a_1, a_2, a_3 ne soient pas situés dans un même plan. Les coordonnées d'un vecteur x se calculent alors par la règle du « parallélipipède », que l'on comprendra en examinant la figure ci-dessous : on utilise les vecteurs a_i pour orienter et définir des « unités de longueur » sur les droites qui les portent.



Les définitions qui précèdent s'appliquent notamment aux groupes commutatifs; il suffit de faire $K = \mathbf{Z}$. On peut donc parler de **groupes commutatifs libres de type fini**

et de **bases** d'un groupe commutatif. Étant donné un groupe commutatif G (noté additivement — mais le lecteur aura intérêt à faire la traduction de ce qui va suivre en notation multiplicative), une base de G est une suite $a_1, \ldots, a_n$ d'éléments de G, en nombre fini (*), telle que l'application

$$(r_1, \ldots, r_n) \to r_1a_1 + \cdots + r_na_n$$

de $\mathbf{Z}^n$ dans G soit *bijective*; et on dit que G est libre de type fini s'il admet au moins une base.

Comme l'ensemble $\mathbf{Z}^n$ est infini, il est clair qu'un groupe commutatif libre de type fini est nécessairement infini. Un groupe commutatif fini est donc un **Z**-module de type fini qui n'admet pas de base.

On voit donc qu'un module sur un anneau peut être de type fini *sans* être libre de type fini, i.e. sans possèder de base. Toutefois :

THÉORÈME 3. — *Tout espace vectoriel de dimension finie sur un corps* K *admet une base.*

Comme le présent Chapitre ne groupe que des résultats valables sur un anneau de base arbitraire, nous ne démontrerons pas ici le Théorème 3; le lecteur pourra, s'il le désire, se reporter au § 19, n° 1, attendu que la démonstration du Théorème 3 n'exige rien d'autre que le contenu du présent §.

Le Théorème 3 montre que l'assertion « tout K-module de type fini est libre de type fini », fausse si K est un anneau quelconque, est vraie si K est un *corps*.

Remarque 1. Pour assurer la validité de certains énoncés, on convient de dire que, pour tout anneau K, le K-module réduit à 0 est libre de type fini, et admet une base formée de 0 vecteurs.

Pour comprendre cette convention, on devrait définir comme suit la notion de base d'un K-module M : c'est une famille finie $(a_i)_{i \in I}$ de vecteurs linéairement indépendants qui engendrent M; cette définition autorise l'ensemble d'indices I à être vide, ce qu'on est en effet obligé de faire si l'on veut attribuer une base au module réduit à 0...

Il est de même indispensable de convenir que K^0 est le module réduit à un seul vecteur 0.

5. *Combinaisons linéaires infinies* (**)

Soient K un anneau, I un ensemble quelconque (fini ou non), et $(\lambda_i)_{i \in I}$ une famille indexée par I d'éléments de K; on dit que les scalaires λ_i sont **presque tous nuls** si l'ensemble des $i \in I$ tels que $\lambda_i \neq 0$ est *fini*; cette définition s'étend de façon évidente à une famille $(x_i)_{i \in I}$ d'éléments d'un module quelconque.

(*) Voir cependant le n° suivant.

(**) Le contenu de ce n° ne sera pas utilisé avant le § 26; le lecteur peut donc attendre d'en avoir besoin avant de l'étudier.

Il va de soit que la notion qu'on vient d'introduire n'a d'intérêt que si l'ensemble I est infini.

Soit $(x_i)_{i\in I}$ une famille d'éléments d'un K-module M; supposons les x_i presque tous nuls : il existe donc des parties *finies* J de I telles que l'on ait $x_i = 0$ pour $i \in I - J$; on pose alors, par définition,

$$\sum_{i\in I} x_i = \sum_{i\in J} x_i;$$

il est clair que la valeur du second membre ne dépend pas de J (pourvu que J satisfasse aux conditions énoncées ci-dessus).

Soit $(x_i)_{i\in I}$ une famille quelconque d'éléments d'un K-module M. On appelle **combinaison linéaire** des x_i tout $x \in M$ possédant la propriété suivante : il existe une famille $(\xi_i)_{i\in I}$ de scalaires *presque tous nuls* telle que l'on ait

$$x = \sum_{i\in I} \xi_i x_i,$$

relation qui a un sens puisque les vecteurs $\xi_i x_i (i \in I)$ sont évidemment presque tous nuls.

On démontre facilement que l'ensemble M' des combinaisons linéaires des x_i est le plus petit sous-module de M contenant tous les x_i, $i \in I$; on l'appelle le **sous-module de M engendré par la famille** $(x_i)_{i\in I}$.

D'autre part, on appelle **relation linéaire entre les** $x_i (i \in I)$ toute famille $(\lambda_i)_{i\in I}$ de scalaires *presque tous nuls* tels que l'on ait

$$\sum_{i\in I} \lambda_i x_i = 0;$$

si cette relation implique $\lambda_i = 0$ pour tout $i \in I$, on dit que les $x_i (i \in I)$ sont **linéairement indépendants,** ou que $(x_i)_{i\in I}$ est une **famille libre** d'éléments de M.

On appelle **base de** M toute famille libre $(x_i)_{i\in I}$ d'éléments de M engendrant le module M. Pour tout $x \in M$, il existe alors une et une seule famille $(\xi_i)_{i\in I}$ de scalaires presque tous nuls telle que l'on ait

$$x = \sum_{i\in I} \xi_i x_i;$$

les ξ_i sont appelés les **coordonnées de** x **par rapport à la base** $(x_i)_{i\in I}$ **de** M.

On dit qu'un module est **libre** s'il admet une base.

On peut démontrer que *tout espace vectoriel sur un corps admet une base*; mais la démonstration de ce résultat est nettement plus difficile que celle du Théorème 3, et dépasse le cadre de cet ouvrage.

Exemple 14. L'*Exemple* 10 montre que la famille (infinie) des fonctions

$$t^n (n = 0, 1, 2, \ldots)$$

est *libre* dans l'espace vectoriel en question; les combinaisons linéaires de ces fonctions ne sont autres que les **fonctions polynomiales** d'une variable réelle (qui seront étudiées au § 28).

Exemple 15. Considérons $\mathbf{C}$ comme un espace vectoriel sur $\mathbf{Q}$; dire qu'un nombre $z \in \mathbf{C}$ est transcendant signifie que la famille (infinie) des puissances de z est libre.

1. *Définition des homomorphismes*

Soient L et M des modules à gauche sur un anneau K. On appelle **homomorphisme** ou **application linéaire** de L dans M toute application

$$f : \mathrm{L} \to \mathrm{M}$$

telle que l'on ait

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \quad \text{quels que soient } x, y \in \mathrm{L},\ \lambda, \mu \in \mathrm{K}.$$

On appelle **isomorphisme** de L sur M tout homomorphisme bijectif de L dans M; on dit que L et M sont **isomorphes** s'il existe un isomorphisme de L sur M.

Enfin, étant donné un K-module à gauche M, on appelle **endomorphisme de M** (ou parfois **opérateur linéaire dans M**) tout homomorphisme de M dans M, et **automorphisme** de M tout isomorphisme de M sur lui-même.

Soient L et M deux K-modules à gauche; pour qu'une application f de L dans M soit linéaire, il faut et il suffit qu'on ait les relations

$$f(x + y) = f(x) + f(y) \quad \text{quels que soient } x, y \in \mathrm{L}$$
$$f(\lambda x) = \lambda f(x) \quad \text{quels que soient } \lambda \in \mathrm{K},\ x \in \mathrm{L}.$$

En prenant $\lambda = 0$ dans la seconde relation, on voit donc que

$$f(0) = 0.$$

D'autre part, si f est un homomorphisme de L dans M, on a la relation

$$(1) \qquad f(\lambda_1 x_1 + \cdots + \lambda_n x_n) = \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)$$

quels que soient l'entier n, les vecteurs $x_1, \ldots, x_n \in \mathrm{L}$ et les scalaires $\lambda_1, \ldots, \lambda_n \in \mathrm{K}$; cette relation se réduit pour $n = 2$ à la définition même des homomorphismes, et se démontre dans le cas général par récurrence sur l'entier n :

$$\begin{aligned} f(\lambda_1 x_1 + \cdots + \lambda_n x_n) &= f[(\lambda_1 x_1 + \cdots + \lambda_{n-1} x_{n-1}) + \lambda_n x_n] \\ &= f(\lambda_1 x_1 + \cdots + \lambda_{n-1} x_{n-1}) + f(\lambda_n x_n) \\ &= \lambda_1 f(x_1) + \cdots + \lambda_{n-1} f(x_{n-1}) + \lambda_n f(x_n) \end{aligned}$$

comme annoncé.

Nous utiliserons constamment la relation (1) dans ce qui suit, et le plus souvent sans y référer explicitement.

THÉORÈME 1. *Si* $f : L \to M$ *et* $g : M \to N$ *sont des homomorphismes de modules, l'application composée* $g \circ f$ *est encore un homomorphisme, et c'est un isomorphisme si* f *et* g *sont des isomorphismes. L'application réciproque d'un isomorphisme de modules est un isomorphisme de modules.*

Soit $h = g \circ f$; on a

$$h(\lambda x + \mu y) = g[f(\lambda x + \mu y)] = g[\lambda f(x) + \mu f(y)]$$
$$= \lambda g[f(x)] + \mu g[f(y)] = \lambda h(x) + \mu h(y),$$

ce qui montre que h est un homomorphisme; si de plus f et g sont des isomorphismes, i.e. sont bijectifs, il en est de même de h, qui est donc alors un isomorphisme.

Supposons que f soit un isomorphisme; pour montrer que l'application f^{-1} (qui est bijective) est un isomorphisme il suffit de prouver qu'elle est linéaire, autrement dit qu'on a

$$f^{-1}(\lambda u + \mu v) = \lambda f^{-1}(u) + \mu f^{-1}(v)$$

quels que soient $\lambda, \mu \in K$ et $u, v \in M$, ou enfin qu'on a $\lambda u + \mu v = f[\lambda f^{-1}(u) + \mu f^{-1}(v)]$; comme f est linéaire, cette relation s'écrit $\lambda u + \mu v = \lambda f[f^{-1}(u)] + \mu f[f^{-1}(v)]$, et est trivialement vérifiée, d'où le Théorème.

En raisonnant comme au § 7, n° 8, on déduit du Théorème 1 que « X et Y sont isomorphes » est une relation d'équivalence entre K-modules à gauche.

Dans la pratique, on regarde souvent deux modules isomorphes L et M comme identiques; plus exactement, si l'on choisit un isomorphisme f de L sur M, alors on peut traduire toute relation algébrique entre éléments de L en une relation analogue entre les images par f de ces éléments, et par suite transformer toute propriété de L en une propriété analogue de M. Le lecteur aura intérêt à vérifier ce fait aussi fréquemment que possible.

THÉORÈME 2. *Soit* $f : L \to M$ *un homomorphisme de modules. L'image par* f *d'un sous-module de* L *est un sous-module de* M. *L'image réciproque par* f *d'un sous-module de* M *est un sous-module de* L.

Soit L' un sous-module de L; supposons que $f(L')$ contienne deux éléments u, v de M; on peut donc écrire $u = f(x)$, $v = f(y)$ avec $x, y \in L'$; comme f est linéaire, on a $\lambda u + \mu v = f(\lambda x + \mu y) = f(z)$ avec $z = \lambda x + \mu y \in L'$ puisque L' est un sous-module de L; ainsi $f(L')$ contient $\lambda u + \mu v$ quels que soient les scalaires λ et μ, ce qui établit la première assertion de l'énoncé. La seconde se démontre de façon analogue.

En particulier, le **noyau** de f, i.e. l'ensemble des $x \in L$ tels que $f(x) = 0$, est un sous-module de L, qu'on note

$$\mathrm{Ker}(f)$$

conformément au n° 9 du § 7; et l'image

$$\mathrm{Im}(f) = f(\mathrm{L})$$

de f est un sous-module de M. Rappelons (§ 7, Théorème 8) que *f est injectif si et seulement si son noyau se réduit à* 0.

2. *Homomorphismes d'un module libre de type fini dans un module quelconque*

Le résultat suivant est fondamental :

THÉORÈME 3. *Soient* L *un* K*-module à gauche libre de type fini* (*), $a_1, \ldots, a_p$ *une base de* L, M *un* K*-module à gauche quelconque, et* $c_1, \ldots, c_p$ *des éléments donnés de* M. *Il existe alors un et un seul homomorphisme f de* L *dans* M *vérifiant*

$$f(a_i) = c_i \qquad \textit{pour} \qquad 1 \leqslant i \leqslant p;$$

pour que f soit injectif (resp. surjectif) il faut et il suffit que les vecteurs $c_1, \ldots, c_p$ soient linéairement indépendants (resp. engendrent M).

Vu la relation (1) du n° 1 l'homomorphisme f, s'il existe, est nécessairement donné par la formule

$$f(\xi_1 a_1 + \cdots + \xi_p a_p) = \xi_1 c_1 + \cdots + \xi_p c_p, \tag{2}$$

ce qui montre déjà l'*unicité* de f.

Pour établir l'*existence* de f, notons que, pour tout $x \in \mathrm{L}$, il existe un et un seul système de scalaires $\xi_i (1 \leqslant i \leqslant p)$ tel que

$$x = \xi_1 a_1 + \cdots + \xi_p a_p$$

du reste on a

$$\xi_i = f_i(x) \qquad (1 \leqslant i \leqslant p),$$

les applications $f_i : \mathrm{L} \to \mathrm{K}$ étant (§ 11, n° 4) les fonctions coordonnées du module L par rapport à la base $a_1, \ldots, a_p$. Cela dit, la formule (2) définit effectivement une application f de L dans M, d'ailleurs donnée d'après ce qui précède par la relation

$$f(x) = f_1(x) c_1 + \cdots + f_p(x) c_p,$$

ce tout revient à montrer que f est un *homomorphisme* transformant les vecteurs a_i en les vecteurs c_i.

La seconde assertion résulte aussitôt de la formule

$$f_i(a_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

(*) On peut étendre le Théorème 3 au cas d'un module libre quelconque, comme le lecteur le vérifiera facilement.

établie au § 11, n° 4, et de la relation

$$c_i = 0.c_1 + \cdots + 0.c_{i-1} + 1.c_i + 0.c_{i+1} + \cdots + 0.c_p.$$

Pour établir que f est un homomorphisme, remarquons d'abord qu'on a démontré au § 11, n° 4 que

$$f_i(\lambda x + \mu y) = \lambda f_i(x) + \mu f_i(y),$$

autrement dit que les fonctions coordonnées $f_i : \mathrm{L} \to \mathrm{K}$ sont des homomorphismes de K-modules à gauche. Par suite

$$\begin{aligned} f(\lambda x + \mu y) &= f_1(\lambda x + \mu y)c_1 + \cdots + f_p(\lambda x + \mu y)c_p \\ &= [\lambda f_1(x) + \mu f_1(y)]c_1 + \cdots + [\lambda f_p(x) + \mu f_p(y)]c_p \\ &= \lambda[f_1(x)c_1 + \cdots + f_p(x)c_p] + \mu[f_1(y)c_1 + \cdots + f_p(y)c_p] \\ &= \lambda f(x) + \mu f(y), \end{aligned}$$

ce qui prouve la linéarité de f.

Reste à déterminer les conditions pour que f soit injectif ou surjectif. Tout d'abord il est clair que l'image $f(\mathrm{L})$ est l'ensemble des combinaisons linéaires des vecteurs $c_1, \ldots, c_p$, autrement dit que f *applique* L *sur le sous-module de* M *engendré par* $c_1, \ldots, c_p$: pour que f soit surjectif il est donc nécessaire et suffisant que les vecteurs c_i engendrent M.

D'autre part, pour que f soit injectif il faut et il suffit que la relation $f(x) = 0$ implique $x = 0$, autrement dit que la relation

$$\xi_1 c_1 + \cdots + \xi_p c_p = 0$$

implique $\xi_1 a_1 + \ldots + \xi_p a_p = 0$, i.e. (puisque les a_i sont linéairement indépendants) implique

$$\xi_1 = 0, \ \ldots, \ \xi_p = 0.$$

Le Théorème 3 est donc démontré.

Corollaire 1. *Pour qu'un* K*-module* M *soit libre de type fini, il faut et il suffit qu'il existe un entier* n *tel que* M *soit isomorphe à* K^n. *Plus précisément, soient* $c_1, \ldots, c_n$ *des éléments de* M, *et* $e_1, \ldots, e_n$ *la base canonique de* K^n. *Pour que les vecteurs* c_i *forment une base de* M, *il faut et il suffit qu'il existe un isomorphisme de* K^n *sur* M *appliquant les vecteurs* e_i *sur les vecteurs* c_i.

D'après le Théorème 3, il existe toujours un et un seul homomorphisme

$$f : \mathrm{K}^n \to \mathrm{M}$$

tel que

$$f(e_i) = c_i \qquad (1 \leqslant i \leqslant p).$$

Pour que les c_i forment une base de M, il faut et il suffit qu'ils soient linéairement indépendants et engendrent M, autrement dit que f soit injectif et surjectif, d'où le Corollaire.

COROLLAIRE 2. *Pour qu'un* K-*module* M *soit de type fini, il faut et il suffit qu'il existe un entier n et un homomorphisme* surjectif *de* K^n *dans* M.

Plus précisément, pour que des vecteurs $c_1, \ldots, c_n$ engendrent M, il faut et il suffit que l'homomorphisme $f: K^n \to M$ tel que $f(e_i) = c_i$ soit surjectif, d'où la nécessité de la condition. Le fait qu'elle soit suffisante résulte de ce que K^n est de type fini et de l'observation suivante : *soit*

$$f: \ L \to M,$$

un homomorphisme de K-*modules; si f est surjectif et si* L *est de type fini, alors* M *est de type fini.* Soit en effet $a_1, \ldots, a_n$, un système de générateurs de L, et posons $b_i = f(a_i)$ $(1 \leqslant i \leqslant n)$; pour tout $y \in M$, il existe un $x \in L$ tel que $y = f(x)$; écrivant $x = \xi_1 a_1 + \cdots + \xi_n a_n$, et utilisant la relation (1) du n° 1 on voit que $y = \xi_1 b_1 + \ldots + \xi_n b_n$; les b_i engendrent donc M, qui est par suite de type fini.

3. *Homomorphismes et matrices*

Le Théorème 3 permet de construire tous les homomorphismes d'un module L *libre de type fini* dans un module quelconque M; pour aller plus loin, nous allons supposer que M est lui-même *libre de type fini.* Dans ce qui suit, on désignera par $a_1, \ldots, a_p$ une base de L et par $b_1, \ldots, b_q$, une base de M. D'autre part, nous supposerons que L et M sont des K-modules *à droite* (le lecteur examinera lui-même le cas des modules à gauche, qui ne diffère du reste pas du cas précédent si K est commutatif).

Soit donc un homomorphisme $f: L \to M$. Considérons un vecteur

$$x = a_1\xi_1 + \cdots + a_p\xi_p, \tag{3}$$

dans L, et son image

$$f(x) = y = b_1\eta_1 + \cdots + b_q\eta_q, \tag{4}$$

dans M; on se propose de trouver des formules permettant de *calculer les coordonnées* $\eta_1, \ldots, \eta_q$ *du vecteur* $f(x)$ *par rapport à la base* $b_1, \ldots, b_q$ *de* M *en fonction des coordonnées* $\xi_1, \ldots, \xi_p$ *de* x *par rapport à la base* $a_1, \ldots, a_p$ *de* L.

Pour cela, posons

$$\begin{cases} f(a_1) = c_1 = b_1\alpha_{11} + b_2\alpha_{12} + \cdots + b_q\alpha_{1q} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ f(a_p) = c_p = b_1\alpha_{p1} + b_2\alpha_{p2} + \cdots + b_q\alpha_{pq}, \end{cases} \tag{5}$$

en mettant en évidence les coordonnées, par rapport à la base $b_1, \ldots, b_q$, des vecteurs $f(a_1), \ldots, f(a_p)$: ces coordonnées α_{ij} ne dépendent que de f et des deux bases choisies dans L et M. Cela dit, on a

$$\begin{aligned} f(x) &= c_1\xi_1 + \cdots + c_p\xi_p \\ &= (b_1\alpha_{11} + b_2\alpha_{12} + \cdots + b_q\alpha_{1q})\xi_1 \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &+ (b_1\alpha_{p1} + b_2\alpha_{p2} + \cdots + b_q\alpha_{pq})\xi_p, \end{aligned}$$

et en effectuant colonne par colonne les additions figurant au dernier membre on obtient les valeurs cherchées pour les coordonnées de $f(x)$, à savoir

$$(6)\qquad \begin{cases} \eta_1 = \alpha_{11}\xi_1 + \alpha_{21}\xi_2 + \cdots + \alpha_{p1}\xi_p \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \eta_q = \alpha_{1q}\xi_1 + \alpha_{2q}\xi_2 + \cdots + \alpha_{pq}\xi_p. \end{cases}$$

Ces formules s'appellent les **équations de f par rapport à la base $(a_i)_{1\leqslant i\leqslant p}$ de L et à la base $(b_j)_{1\leqslant j\leqslant q}$ de M**; on peut les écrire comme suit sous forme condensée :

$$(6\ bis)\qquad \eta_j = \sum_{i=1}^{i=p} \alpha_{ij}\xi_i.$$

Donnons-nous inversement des scalaires $\alpha_{ij} \in K$ $(1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q)$ et définissons une application f de L dans M par les formules (6) : f transforme donc chaque vecteur $x \in L$, de coordonnées $\xi_1, \ldots, \xi_p$ par rapport à la base donnée de L, en le vecteur de M dont les coordonnées $\eta_1, \ldots, \eta_q$ par rapport à la base donnée de M se calculent par les relations (6). Alors f est un *homomorphisme* de L dans M. On a en effet

$$\begin{aligned} f(x) &= b_1\eta_1 + \cdots + b_q\eta_q \\ &= b_1(\alpha_{11}\xi_1 + \cdots + \alpha_{p1}\xi_p) \\ &+ \cdots\cdots\cdots\cdots\cdots\cdots \\ &+ b_q(\alpha_{1q}\xi_1 + \cdots + \alpha_{pq}\xi_p) \\ &= c_1\xi_1 + \cdots + c_p\xi_p \end{aligned}$$

où les vecteurs c_i sont donnés par les formules (5), et on voit alors que f est l'homomorphisme dont le Théorème 3 affirme l'existence.

Remarquons enfin que, l'homomorphisme f étant donné, il existe *un seul* système de scalaires α_{ij} tel que f soit donné par les relations (6), car le calcul qu'on vient de faire montre qu'alors les α_{ij} sont nécessairement les coordonnées des vecteurs $f(a_i)$ par rapport à la base $(b_j)_{1\leqslant j\leqslant q}$ de M, ce qui les détermine entièrement.

Revenons maintenant aux formules (6) dont la connaissance permet de définir l'homomorphisme f. Il est clair que pour les retenir, il suffit de connaître le tableau

$$(7)\qquad \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{p1} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1q} & \alpha_{2q} & \cdots & \alpha_{pq} \end{pmatrix}$$

formé par les constantes α_{ij} (qu'on appelle les **coefficients** figurant dans les formules (6)). Un tableau de la forme (7) s'appelle une **matrice à p colonnes et q lignes à coefficients dans l'anneau** K (les α_{ij} s'appellent aussi les **termes** de la matrice en question), et on dit que (7) est la **matrice de l'homomorphisme f par rapport à la base $(a_i)_{1\leqslant i\leqslant p}$ de L et à la base $(b_j)_{1\leqslant j\leqslant q}$ de M**. La notion de matrice joue pour les homomorphismes un rôle analogue à celui que joue, pour les vecteurs, la notion de coordonnées.

Lorsqu'on n'a pas à faire de calculs explicites, on désigne souvent la matrice (7) par la notation condensée

$$(\alpha_{ij})_{1\leqslant i\leqslant p,\ 1\leqslant j\leqslant q},$$

qui a d'ailleurs le mérite de montrer qu'une matrice n'est autre qu'une *famille* d'éléments de K, famille indexée par l'ensemble de tous les couples (i, j) d'entiers tels que $1 \leqslant i \leqslant p$, $1 \leqslant j \leqslant q$.

Lorsque L = M, autrement dit lorsque f est un *endomorphisme* du module L, on utilise le plus souvent la même base $a_1, \ldots, a_p$ dans L et dans M, ce qui permet alors de parler de la **matrice d'un endomorphisme de L par rapport à une base de L.** La matrice en question a évidemment p lignes et p colonnes; on dit que c'est une **matrice carrée d'ordre p à coefficients dans K.**

Remarque 1. Étant donné un homomorphisme $f : L \to M$ de K-modules libres de type fini, on ne peut pas parler de « la » matrice de f; pour donner un sens à cette notion on doit d'abord choisir une base de L et une base de M, et « la » matrice obtenue dépend évidemment du choix des bases (on verra au § 15 ce qui se passe lorsqu'on change de bases dans L et dans M).

Toutefois, lorsque $L = K^p$ et $M = K^q$, il s'impose de choisir la base *canonique* de L et la base *canonique* de M (§ 11, *Exemple* 12); étant donné un homomorphisme

$$f : K^p \to K^q$$

il est donc légitime dans ce cas de parler de *la* matrice de f (sous-entendu : par rapport à la base canonique de K^p et à la base canonique de K^q); si l'on désigne cette matrice par $(\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$, alors f n'est autre que l'application transformant chaque vecteur

$$x = (\xi_1, \ldots, \xi_p) \in K^p,$$

en le vecteur

$$f(x) = (\eta_1, \ldots, \eta_q) \in K^q$$

donné par les relations (6) ci-dessus.

En sens inverse, les mêmes constructions permettent d'attacher à chaque matrice $(\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$ un homomorphisme bien déterminé de K^p dans K^q — à savoir celui dont *la* matrice est la matrice donnée.

Ces considérations montrent donc l'existence d'une bijection « canonique » de l'ensemble de tous les homomorphismes de K^p dans K^q sur l'ensemble de toutes les matrices à p colonnes et q lignes (à coefficients dans K). Nous utiliserons fréquemment cette correspondance, le plus souvent sans y référer explicitement.

Remarque 2. Soient L et M des K-modules à droite libres de type fini, $(a_i)_{1 \leqslant i \leqslant p}$ une base de L et $(b_j)_{1 \leqslant j \leqslant q}$ une base de M. Introduisons les isomorphismes

$$u : K^p \to L, \qquad v : K^q \to M$$

qui appliquent respectivement les bases canoniques de K^p et K^q sur les bases données de L et M (Corollaire 1 du Théorème 3).

Soit un homomorphisme $f : L \to M$; à l'aide de u et v, on en déduit (Théorème 1) un homomorphisme

$$\overline{f} = v^{-1} \circ f \circ u : K^p \to K^q;$$

cela dit, *la matrice de f par rapport à la base (a_i) de L et à la base (b_j) de M est identique à la matrice de $\overline{f}$* (par rapport aux bases canoniques de K^p et K^q). Considérons en effet les vecteurs

$$x = a_1\xi_1 + \cdots + a_p\xi_p, \ f(x) = b_1\eta_1 + \cdots + b_q\eta_q,$$

les τ_{ij} étant donnés en fonction des ξ_i par les formules (6) ci-dessus. Par construction de u on a

$$x = u(\xi_1, \ldots, \xi_p)$$

et de même

$$f(x) = v(\tau_{11}, \ldots, \tau_{1q});$$

donc

$$(\tau_{11}, \ldots, \tau_{1q}) = v^{-1}[f(x)] = v^{-1}\{f[u(\xi_1, \ldots, \xi_p)]\} = \bar{f}(\xi_1, \ldots, \xi_p);$$

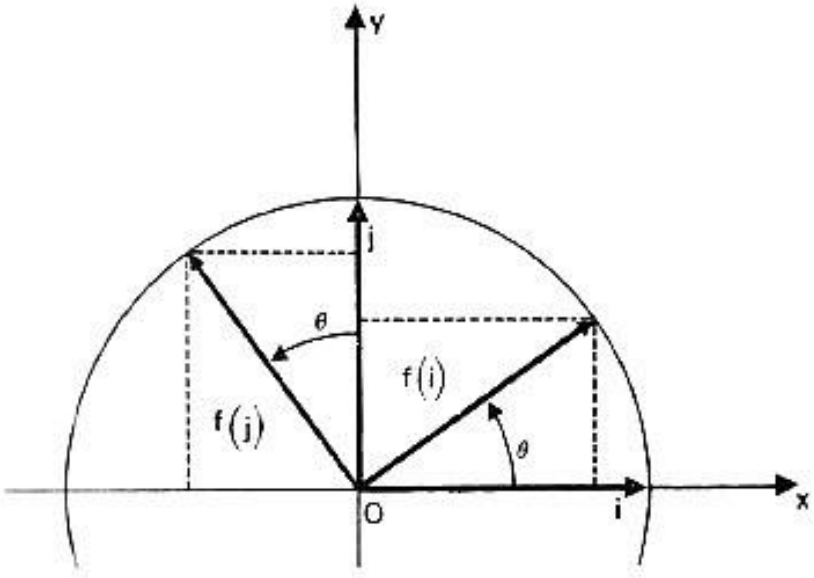
il s'ensuit que les formules (6) sont aussi les équations de $\bar{f}$ (par rapport aux bases canoniques), ce qui établit notre assertion.

On voit donc qu'on aurait encore pu introduire les matrices en montrant tout d'abord que tout homomorphisme de K^p dans K^q est donné par des relations du type (6), puis en observant que si l'on a un homomorphisme $f: L \to M$ de modules libres de type fini, alors le choix d'une base de L et d'une base de M permet d'identifier L à un module K^p, M à un module K^q, et par suite f à un homomorphisme $\bar{f}$ de K^p dans K^q, donc donné par des relations (6), et représenté de façon *canonique* par une matrice. On retrouverait ainsi la possibilité d'attacher une matrice à f une fois choisies des bases de L et M.

4. *Exemples d'homomorphismes et de matrices*

Nous allons indiquer maintenant quelques exemples d'homomorphismes, et de représentation d'un homomorphisme par une matrice.

Exemple 1. Soit L l'espace vectoriel réel formé des vecteurs d'origine donnée O dans le plan usuel, et considérons l'application $f: L \to L$ qui transforme tout vecteur d'origine O en celui qui s'en déduit par la rotation d'angle θ (donné) autour du point O. C'est évidemment une application linéaire (car une rotation transforme un parallélogramme en un parallélogramme, et une rotation de centre O commute à toute homothétie de centre O). Soit Ox, Oy un système de coordonnées *rectangulaires* dans le plan, i et j les vecteurs unité (*) de Ox et Oy; un raisonnement géométrique évident montre que l'on a les relations



$$f(i) = i.\cos\theta + j.\sin\theta$$
$$f(j) = -i.\sin\theta + j.\cos\theta;$$

par suite, la matrice de f par rapport à la base (i, j) de L n'est autre que

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(*) Le vecteur i n'a évidemment aucun rapport avec le nombre complexe i...

(on n'oubliera pas que les *lignes* de la matrice (7) d'un homomorphisme s'obtiennent en considérant les *colonnes* figurant dans les relations (5), ce qui explique le résultat trouvé ici).

Exemple 2. Prenons L = M et supposons K commutatif; alors, pour tout $\lambda \in K$, l'application $x \to \lambda x$ (**homothétie de rapport** λ dans L) est linéaire. Comme elle transforme une base $a_1, \ldots, a_n$ de L en les vecteurs

$$\begin{aligned} \lambda.a_1 &= \lambda.a_1 + 0.a_2 + \cdots + 0.a_n \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \lambda.a_n &= 0.a_1 + 0.a_2 + \cdots + \lambda.a_n, \end{aligned}$$

on voit que la matrice de l'homothétie de rapport λ, par rapport à une base quelconque de L, est

$$\begin{pmatrix} \lambda & 0 & 0 & \ldots & 0 \\ 0 & \lambda & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & \lambda \end{pmatrix};$$

on dit que c'est une **matrice scalaire** à coefficients dans l'anneau K.

On notera qu'ici la matrice de f est indépendante du choix de la base $a_1, \ldots, a_n$ de L; on peut montrer facilement que les homothéties sont les *seuls* endomorphismes de L à posséder cette propriété.

Exemple 3. On appelle **forme linéaire** ou **covecteur** sur un K-module à droite L tout homomorphisme de L dans le K-module à droite, autrement dit toute application

$$f : \mathrm{L} \to \mathrm{K}$$

telle que l'on ait

$$f(x\lambda + y\mu) = f(x)\lambda + f(y)\mu \quad \text{quels que soient } \lambda, \mu \in \mathrm{K},\ x, y \in \mathrm{L}.$$

Par exemple, le § 11, n° 4 montre que *les fonctions coordonnées de* L *par rapport à une base de* L *sont des formes linéaires sur* L.

Soit f une forme linéaire sur L, et choisissons (si c'est possible) une base $a_1, \ldots, a_p$ de L; pour attacher à f, comme à tout homomorphisme de modules libres de type fini, une matrice, il faut encore choisir une base du K-module à droite K : nous choisirons la base canonique de celui-ci, formée du vecteur (sic) 1. Pour calculer la matrice de f par rapport à ces deux bases, on doit écrire

$$\begin{aligned} f(a_1) &= 1.\alpha_{11} \\ &\cdots\cdots\cdots\cdots \\ f(a_p) &= 1.\alpha_{p1}; \end{aligned}$$

il est clair qu'il est plus simple de poser

$$\begin{aligned} f(a_1) &= \alpha_1 \\ &\cdots\cdots\cdots \\ f(a_p) &= \alpha_p, \end{aligned}$$

et alors f est représentée, par rapport aux bases considérées de L et K, par la **matrice ligne**

$$(\alpha_1, \ldots, \alpha_p);$$

la donnée de cette matrice détermine f grâce à la formule

$$f(a_1\xi_1 + \cdots + a_p\xi_p) = \alpha_1\xi_1 + \cdots + \alpha_p\xi_p;$$

on dit souvent que les α_i sont **les coefficients de f par rapport à la base $(a_i)_{1\leqslant i\leqslant p}$ de L.**

Si en particulier $L = K^p$, il s'impose de choisir la base canonique de L; les α_i s'appellent alors simplement les **coefficients de** f (sous-entendu : par rapport à la base canonique de K^p), et f est donnée par la relation

$$f(\xi_1, \ldots, \xi_p) = \alpha_1\xi_1 + \cdots + \alpha_p\xi_p.$$

Exemple 4. Soit M un K-module à droite, et considérons un homomorphisme

$$f: K \to M$$

de K-modules à droite; posant

$$f(1) = c \in M$$

on a

$$f(\xi) = f(1.\xi) = f(1)\xi = c\xi,$$

de sorte que la connaissance de c détermine entièrement f (dans la pratique on ne fait pas de différence entre l'homomorphisme f et l'élément c de M). Supposons que M admette une base $b_1, \ldots, b_q$; dans K, utilisons la base canonique; posant

$$c = f(1) = b_1\alpha_1 + \cdots + b_q\alpha_q$$

on voit que la matrice de f par rapport aux bases considérées de K et M est **la matrice colonne**

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix}$$

formée avec les composantes du vecteur c par rapport à la base considérée de M.

Il nous arrivera parfois dans la suite d'identifier (une fois choisie une base de M) chaque vecteur de M avec la matrice *colonne* dont les termes sont les coordonnées du vecteur considéré par rapport à la base choisie dans M; ce point de vue sera justifié plus loin (§ 14, n° 4).

Les exemples précédents sont de nature purement algébrique (comme tout ce qui se rapporte aux modules *libres de type fini*). L'Analyse fournit par contre des exemples d'homomorphismes qu'il serait tout à fait artificiel de chercher à représenter par des matrices (même infinies). En voici quelques-uns.

Exemple 5. Soient a et b deux nombres réels tels que $a < b$, notons X l'ensemble des $t \in \mathbf{R}$ tels que $a \leqslant t \leqslant b$ (X n'est autre que l'intervalle $[a, b]$), et désignons par L l'espace vectoriel réel formé par toutes les applications $f: X \to \mathbf{R}$ qui

sont continues quel que soit $t \in X$ (cf. § 10, *Exemples* 4 et 7). Soit $N(s, t)$ une fonction à valeurs réelles, définie et continue sur le carré $X \times X$. On peut démontrer (en utilisant le fait que la fonction N est *uniformément* continue sur le carré $X \times X$) que pour toute fonction $f \in L$, la fonction

$$f^*(s) = \int_a^b N(s, t) f(t) dt \qquad (a \leqslant s \leqslant b)$$

est encore continue sur l'intervalle X, autrement dit que $f^* \in L$. Cela fait, l'application $u : f \to f^*$ de L dans L est *linéaire*. Soient en effet f, $g \in L$ et posons $f + g = h$; on a

$$h^*(s) = \int_a^b N(s, t) h(t) dt = \int_a^b N(s, t)[f(t) + g(t)] dt$$
$$= \int_a^b N(s, t) f(t) dt + \int_a^b N(s, t) g(t) dt = f^*(s) + g^*(s),$$

autrement dit $h^* = f^* + g^*$, de sorte qu'on a $u(f + g) = u(f) + u(g)$; on démontrerait de même la relation $u(\lambda f) = \lambda u(f)$ pour tout $\lambda \in \mathbf{R}$.

L'application u de L dans L s'appelle un **opérateur intégral;** l'étude (notamment par Hilbert et F. Riesz) de ces opérateurs a conduit à la création, dans le premier quart du XX^e siècle, de ce qu'on appelle aujourd'hui l'*Analyse Fonctionnelle*. Il va de soi — ce serait trop facile... — que les considérations purement algébriques, et essentiellement triviales, qui sont développées dans cet ouvrage ne sont d'aucun secours sérieux en Analyse Fonctionnelle, sauf pour fournir à celle-ci une terminologie raisonnable et une vague idée des directions dans lesquelles la recherche doit s'effectuer; en fait, les difficultés qu'on rencontre en Analyse Fonctionnelle pour établir des résultats non triviaux sont rarement de nature algébrique; elles sont le plus souvent de nature « analytique » et exigent, pour être surmontées, l'emploi de méthodes « topologiques » (i. e. fondées sur la notion de « continuité »). Il est du reste intéressant de remarquer que le développement de l'Algèbre linéaire élémentaire a été grandement influencé par celui de l'Analyse Fonctionnelle, alors qu'on aurait pu espérer voir plutôt le contraire...

Exemple 6. Prenons le même espace vectoriel réel L que ci-dessus; alors l'application

$$f \mapsto \int_a^b f(t) dt = I(f)$$

de L dans $\mathbf{R}$ est une *forme linéaire* sur L : si en effet f et g sont des fonctions continues sur l'intervalle $[a, b]$, l'intégrale de la fonction $f + g$ est la somme des intégrales des fonctions f et g; et si l'on multiplie f par une constante $\lambda \in \mathbf{R}$, son intégrale est aussi multipliée par λ.

Exemple 7. L'anneau de base étant toujours $\mathbf{R}$, considérons les deux espaces vectoriels réels L et M que voici : les éléments de L sont les applications $f : \mathbf{R} \to \mathbf{R}$ admettant une *dérivée seconde continue* f''; et ceux de M sont toutes les applications *continues* $g : \mathbf{R} \to \mathbf{R}$ (on n'impose aucune condition de dérivabilité). Bien entendu, les opérations vectorielles dans L et M sont définies comme au § 10, *Exemple* 4.

Choisissons une fois pour toutes des fonctions a, b, $c \in \mathrm{M}$ (i. e. des fonctions continues d'une variable réelle), et pour toute fonction $f \in \mathrm{L}$ formons la fonction

$$f^*(t) = a(t)f(t) + b(t)f'(t) + c(t)f''(t);$$

évidemment f^* appartient à M, d'où une application $\mathrm{D} : \mathrm{L} \to \mathrm{M}$ donnée par $\mathrm{D}(f) = f^*$ pour tout $f \in \mathrm{L}$. Cela dit, D est un *homomorphisme*. En effet

$$\begin{aligned} \mathrm{D}(f+g) &= a(f+g) + b(f+g)' + c(f+g)'' \\ &= af + bf' + cf'' + ag + bg' + cg'' = \mathrm{D}(f) + \mathrm{D}(g), \end{aligned}$$

et on montrerait de même que $\mathrm{D}(\lambda f) = \lambda \mathrm{D}(f)$.

Les homomorphismes de ce genre interviennent dans la théorie des *équations différentielles linéaires*.

Notons qu'il est aussi facile de construire des formes linéaires sur l'espace vectoriel L; c'est le cas par exemple de l'application

$$f \mapsto f''(0)$$

de L dans $\mathbf{R}$, qui à chaque $f \in \mathrm{L}$ associe la valeur pour $t = 0$ de sa dérivée seconde.

§ 13. Addition des homomorphismes et matrices

1. *Les groupes additifs* Hom (L, M)

Soient L et M deux K-modules (à gauche par exemple) sur un anneau quelconque K. On désigne par la notation

$$\mathrm{Hom}\ (\mathrm{L},\ \mathrm{M}) \qquad \text{ou} \qquad \mathfrak{L}(\mathrm{L},\ \mathrm{M})$$

l'ensemble de toutes les applications linéaires de L dans M; on utilise aussi (lorsqu'il peut y avoir ambiguïté sur l'anneau de base) la notation

$$\mathrm{Hom}_{\mathrm{K}}(\mathrm{L},\ \mathrm{M}) \qquad \text{ou} \qquad \mathfrak{L}_{\mathrm{K}}(\mathrm{L},\ \mathrm{M}).$$

THÉORÈME 1. *Soient* L *et* M *deux* K*-modules à gauche. Si*

$$f,\ g :\ \mathrm{L} \to \mathrm{M}$$

sont des homomorphismes de L *dans* M, *il en est de même de l'application*

$$f + g :\ x \mapsto f(x) + g(x).$$

L'ensemble Hom (L, M), *muni de la loi de composition* $(f, g) \mapsto f + g$, *est un groupe commutatif.*

Posons $h = f + g$; alors

$$\begin{aligned} h(\lambda x + \mu y) &= f(\lambda x + \mu y) + g(\lambda x + \mu y) \\ &= \lambda f(x) + \mu f(y) + \lambda g(x) + \mu g(y) \\ &= \lambda[f(x) + g(x)] + \mu[f(y) + g(y)] \\ &= \lambda h(x) + \mu h(y), \end{aligned}$$

ce qui établit la première assertion de l'énoncé.

Pour établir la seconde, considérons l'ensemble E de *toutes* les applications (linéaires ou non) de L dans M; muni de la loi de composition $(f, g) \mapsto f + g$, c'est un groupe commutatif (§ 10, *Exemple* 4; le fait que L soit un module n'intervient pas, on regarde simplement L comme un ensemble). Il reste donc à faire voir

que Hom (L, M) est un *sous-groupe* de E; or Hom (L, M) contient évidemment l'élément neutre de E (à savoir l'application de L dans M qui prend partout la valeur 0), et si f, g sont des homomorphismes il en est de même de $f - g$, comme le montre la première partie de la démonstration, d'où le résultat cherché.

Le Théorème 1 permet d'appeler Hom (L, M) le **groupe des homomorphismes** de L dans M.

Lorsque l'anneau de base K est *commutatif*, on peut même considérer Hom (L, M) comme un nouveau K-module à gauche. Tout d'abord, reprenons l'ensemble E de toutes les applications (linéaires ou non) de L dans M; le § 10, *Exemple* 4 permet de considérer E non seulement comme un groupe additif, mais comme un K-module à gauche, le produit λf d'un scalaire $\lambda \in K$ et d'une application $f : L \to M$ étant l'application

$$x \mapsto \lambda f(x)$$

de L dans M (et cela ne suppose pas K commutatif). Or il se trouve que, K étant commutatif, l'ensemble Hom (L, M) est non seulement un sous-*groupe* mais un sous-*module* de E, autrement dit que si f est un homomorphisme de L dans M, il en est encore ainsi de $f' = \lambda f$; en effet, on a

$$\begin{aligned} f'(\alpha x + \beta y) &= \lambda f(\alpha x + \beta y) = \lambda\alpha . f(x) + \lambda\beta . f(y) \\ &= \alpha\lambda f(x) + \beta\lambda . f(y) = \alpha f'(x) + \beta f'(y), \end{aligned}$$

comme annoncé. On peut donc bien, dans ce cas, regarder Hom (L, M) comme un K-module à gauche.

Si par exemple L et M sont des espaces vectoriels réels (resp. complexes), alors on peut regarder Hom (L, M) comme un espace vectoriel réel (resp. complexe).

2. *Addition des matrices*

Dans ce qui précède, supposons que L et M soient des K-modules à droite libres de type fini; choisissons une base $(a_i)_{1 \leqslant i \leqslant p}$ de L et une base $(b_j)_{1 \leqslant j \leqslant q}$ de M; enfin, étant donnés deux homomorphismes f et g de L dans M, soient

$$\begin{aligned} A &= (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q} \\ B &= (\beta_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q} \end{aligned}$$

leurs matrices par rapport aux bases considérées de L et M. On a donc

$$\begin{aligned} f(a_i) &= b_1\alpha_{i1} + \cdots + b_q\alpha_{iq}, \\ g(a_i) &= b_1\beta_{i1} + \cdots + b_q\beta_{iq}. \end{aligned}$$

Posant $h = f + g$, et désignant la matrice de h par rapport aux bases considérées par

$$C = (\gamma_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q},$$

on a

$$h(a_i) = f(a_i) + g(a_i) = b_1(\alpha_{i1} + \beta_{i1}) + \cdots + b_q(\alpha_{iq} + \beta_{iq}),$$

et par suite les termes de C sont donnés par les relations

$$\gamma_{ij} = \alpha_{ij} + \beta_{ij} \qquad (1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q). \tag{1}$$

Étant données deux matrices $A = (\alpha_{ij})$ et $B = (\beta_{ij})$ à coefficients dans K, on est ainsi conduit à appeler **somme des deux matrices** A **et** B la matrice $C = (\gamma_{ij})$ donnée par les relations (1); on la désigne par la notation

$$A + B.$$

On notera que la somme de deux matrices n'est définie que si celles-ci ont le même nombre de lignes, et le même nombre de colonnes.

Si l'on identifie une matrice $(\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$ à un élément de K^{pq}, il est clair que l'addition des matrices se réduit à celle des éléments de K^{pq}. Par suite, l'ensemble des matrices à p colonnes et q lignes, muni de la loi de composition $(A, B) \mapsto A + B$, est un *groupe commutatif*.

On a en outre — et pour cause — le résultat suivant :

THÉORÈME 2. *Soient* L *et* M *des* K*-modules libres de type fini, f et g des homomorphismes de* L *dans* M*; soient* A *et* B *les matrices de f et g par rapport à une base (a_i) de* L *et à une base (b_j) de* M*. Alors la matrice de $f + g$ par rapport à ces bases est* $A + B$.

Notons enfin qu'on peut regarder l'ensemble des matrices à p colonnes et q lignes à coefficients dans K non seulement comme un groupe additif, mais même comme un K-module à gauche (ou à droite); il suffit de définir, pour une matrice

$$A = (\alpha_{ij}),$$

les expressions λA et $A\lambda$ par les formules suivantes :

$$\lambda A = (\lambda \alpha_{ij}), \qquad A\lambda = (\alpha_{ij}\lambda).$$

§ 14. — Produits de matrices

1. *L'anneau des endomorphismes d'un module*

Établissons tout d'abord le résultat suivant :

THÉORÈME 1. *Soient* L, M, N *trois modules; étant donnés des homomorphismes*

$$f, g : \mathrm{L} \to \mathrm{M} \qquad et \qquad h : \mathrm{M} \to \mathrm{N},$$

on a la relation

$$h \circ (f + g) = h \circ f + h \circ g;$$

étant donnés des homomorphismes

$$f : \mathrm{L} \to \mathrm{M} \qquad et \qquad g, h : \mathrm{M} \to \mathrm{N},$$

on a la relation

$$(g + h) \circ f = g \circ f + h \circ f.$$

Établissons par exemple le premier résultat. Posant $u = f + g$, on a

$$h \circ u(x) = h[u(x)] = h[f(x) + g(x)] = h[f(x)] + h[g(x)]$$

ce qui montre que l'application $h \circ u$ est somme des applications $h \circ f$ et $h \circ g$, d'où le Théorème.

COROLLAIRE. *Soit* L *un module sur un anneau; l'ensemble* Hom (L, L) *des endomorphismes du module* L, *muni des lois de composition*

$$(f, g) \mapsto f + g, \qquad (f, g) \mapsto f \circ g,$$

est un anneau.

Le fait que Hom (L, L), muni de l'addition, soit un groupe commutatif résulte du § 13, Théorème 1. L'associativité de la multiplication résulte du § 2, Théorème 2, et l'existence d'un élément neutre du fait que l'application identique j_{L} appartient à Hom (L, L). Enfin, le Théorème 1 montre que les conditions de « distributivité » sont satisfaites, ce qui achève la démonstration.

L'ensemble Hom (L, L), muni des deux lois de composition en question, s'appelle **l'anneau des endomorphismes du module** L. Il est en général non commutatif (même si l'anneau de base K est commutatif), comme on le verra plus loin.

2. *Produit de deux matrices*

Soient L, M, N des modules à droite *libres de type fini*, et choisissons des bases $(a_1, \ldots, a_p)$, $(b_1, \ldots, b_q)$ et $(c_1, \ldots, c_r)$ de ces modules. Considérons des homomorphismes

$$f : \mathrm{M} \to \mathrm{N} \qquad \text{et} \qquad g : \mathrm{L} \to \mathrm{M}$$

et l'homomorphisme composé

$$h = f \circ g : \mathrm{L} \to \mathrm{N}.$$

Désignons la matrice de f par rapport aux bases (b_j) et (c_k) par

$$\mathrm{A} = (\alpha_{jk})_{1 \leqslant j \leqslant q,\ 1 \leqslant k \leqslant r},$$

celle de g par rapport aux bases (a_i) et (b_j) par

$$\mathrm{B} = (\beta_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q},$$

et celle de h par rapport aux bases (a_i) et (c_k) par

$$\mathrm{C} = (\gamma_{ik})_{1 \leqslant i \leqslant p,\ 1 \leqslant k \leqslant r};$$

on se propose de calculer C en fonction de A et B.

Soit

$$x = a_1\xi_1 + \cdots + a_p\xi_p$$

un élément de L; nous poserons

$$\begin{aligned} g(x) &= y = b_1\eta_1 + \cdots + b_q\eta_q \\ h(x) &= f(y) = c_1\zeta_1 + \cdots + c_r\zeta_r. \end{aligned}$$

Comme on connait les matrices A et B de f et g, les formules (6) du § 12, n° 3 montrent qu'on a

$$\begin{aligned} \zeta_k &= \alpha_{1k}\eta_1 + \cdots + \alpha_{qk}\eta_q \quad (1 \leqslant k \leqslant r) \\ \eta_j &= \beta_{1j}\xi_1 + \cdots + \beta_{pj}\xi_p \quad (1 \leqslant j \leqslant q), \end{aligned}$$

et par suite

$$\begin{aligned} \zeta_k = {} & \alpha_{1k}(\beta_{11}\xi_1 + \cdots + \beta_{p1}\xi_p) \\ & + \alpha_{2k}(\beta_{12}\xi_1 + \cdots + \beta_{p2}\xi_p) \\ & + \cdots\cdots\cdots\cdots\cdots\cdots \\ & + \alpha_{qk}(\beta_{1q}\xi_1 + \cdots + \beta_{pq}\xi_p); \end{aligned}$$

mais la matrice $\mathrm{C} = (\gamma_{ik})$ de h est aussi donnée par les formules

$$\zeta_k = \gamma_{1k}\xi_1 + \cdots + \gamma_{pk}\xi_p;$$

comparant les résultats obtenus on trouve donc les relations

$$\gamma_{1k} = \alpha_{1k}\beta_{11} + \alpha_{2k}\beta_{12} + \cdots + \alpha_{qk}\beta_{1q}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\gamma_{pk} = \alpha_{1k}\beta_{p1} + \alpha_{2k}\beta_{p2} + \cdots + \alpha_{qk}\beta_{pq}$$

ou, sous une forme plus condensée,

$$(1) \qquad \gamma_{ik} = \alpha_{1k}\beta_{i1} + \cdots + \alpha_{qk}\beta_{iq} = \sum_{j=1}^{j=q} \alpha_{jk}\beta_{ij}.$$

Ce résultat nous conduit à introduire la définition suivante : étant données des matrices

$$A = (\alpha_{jk})_{1 \leqslant j \leqslant q,\ 1 \leqslant k \leqslant r},\quad B = (\beta_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$$

à coefficients dans l'anneau K, on appelle **produit de A et B** la matrice

$$AB = (\gamma_{ik})_{1 \leqslant i \leqslant p,\ 1 \leqslant k \leqslant r}$$

dont les coefficients sont donnés par les relations (1).

On notera que le produit AB n'est défini que si le nombre de *colonnes* de A est égal au nombre de *lignes* de B; et alors la matrice AB a autant de lignes que A, et autant de colonnes que B.

Il est clair qu'avec la définition précédente nous pouvons énoncer le résultat suivant :

THÉORÈME 2. *Soient* L, M, N *des* K*-modules à droite libres de type fini,* (a_i), (b_j), (c_k) *des bases de* L, M, N *et* $f : M \to N$ *et* $g : L \to M$ *des homomorphismes. Soient* A *la matrice de* f *par rapport aux bases* (b_j) *et* (c_k), *et* B *celle de* g *par rapport aux bases* (a_i) *et* (b_j).

Alors la matrice de $f \circ g$ *par rapport aux bases* (a_i) *et* (c_k) *est* AB.

Donnons maintenant quelques exemples de multiplication de matrices.

Exemple 1. Prenons

$$A = (\alpha_1 \ \ldots \ \alpha_q), \qquad B = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_q \end{pmatrix};$$

le produit AB est défini et est une matrice à une ligne et une colonne, donc de la forme (γ) où γ est un scalaire; la relation (1) donne évidemment

$$\gamma = \alpha_1\beta_1 + \cdots + \alpha_q\beta_q,$$

et on dit, pour des raisons évidentes, que le scalaire γ est le produit de la « ligne » A par la « colonne » B.

Ce résultat permet de retenir facilement la règle générale de multiplication des matrices. En effet la formule (1), avec les conventions qu'on vient d'introduire, s'écrit encore

$$\gamma_{ik} = (\alpha_{1k} \ \ldots \ \alpha_{qk}) \cdot \begin{pmatrix} \beta_{i1} \\ \vdots \\ \beta_{iq} \end{pmatrix};$$

autrement dit, *les termes situés sur la k^e ligne de* AB *s'obtiennent en multipliant la k^e ligne de* A *par les colonnes de* B. C'est cette règle qu'on utilise toujours dans la pratique.

Exemple 2. La multiplication des matrices carrées d'ordre 2 est définie par la formule

$$\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}\begin{pmatrix} a'' & b'' \\ c'' & d'' \end{pmatrix} = \begin{pmatrix} a'a'' + b'c'' & a'b'' + b'd'' \\ c'a'' + d'c'' & c'b'' + d'd'' \end{pmatrix}.$$

Par exemple, si x et y sont des nombres réels, on a

$$\begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}\begin{pmatrix} \cos y & \sin y \\ -\sin y & \cos y \end{pmatrix} = \begin{pmatrix} \cos(x+y) & \sin(x+y) \\ -\sin(x+y) & \cos(x+y) \end{pmatrix};$$

compte-tenu du § 12, *Exemple* 1, ce résultat exprime qu'en composant, dans le plan, les rotations d'angles x et y autour d'un point 0, on trouve la rotation d'angle $x+y$ autour de 0.

Exemple 3. Pour tout anneau K et tout entier $n \geqslant 1$, considérons la matrice

$$1_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

à n lignes et n colonnes; on l'appelle la **matrice unité d'ordre** n. Cette terminologie est justifiée par le fait que, quelles que soient les matrices X et Y à coefficients dans K, les relations

$$1_n . X = X, \qquad Y . 1_n = Y$$

sont vraies (pourvu qu'elles aient un sens, i.e. si X a n lignes, et si Y a n colonnes) Ces relations s'obtiennent facilement sur les formules (1), et s'interprètent géométriquement comme suit. Soit L un K-module à droite libre de type fini, possédant une base $(a_i)_{1 \leqslant i \leqslant n}$ formée de n vecteurs (on peut prendre par exemple $L = K^n$ et la base canonique); alors l'endomorphisme $j : L \to L$ ayant 1_n pour matrice par rapport à la base (a_i) n'est autre que l'*application identique* de L dans L, comme le montrent les formules (6) du § 12, n° 3 et le fait qu'ici on a

$$\alpha_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Cela dit, pour établir par exemple la relation $1_n . X = X$, on introduit un second module M, une base (b_j) de M, et l'homomorphisme $f : M \to L$ ayant X pour matrice par rapport aux bases considérées de L et M; le Théorème 1 montre que $1_n . X$ est la matrice, par rapport à ces bases, de l'homomorphisme $j \circ f$; mais comme $j = j_L$ on a $j \circ f = f$, d'où la relation cherchée.

3. *Anneaux de matrices*

Les opérations d'addition et de multiplication des matrices, que nous avons définies dans ce § et le précédent, obéissent, dans la mesure où elles ont un sens, aux règles

de calcul usuelles. On le sait déjà pour l'addition (§ 13, n° 2). En ce qui concerne la multiplication, on a la relation d'*associativité*

$$A(BC) = (AB)C$$

toutes les fois que les deux membres sont définis; pour le voir, on introduit les homomorphismes

$$f : K^r \to K^s, \qquad g : K^q \to K^r, \qquad h : K^p \to K^q$$

dont les matrices (par rapport aux bases canoniques) sont A, B, C; alors, d'après le Théorème 2, la matrice A(BC) représente $f \circ (g \circ h)$, tandis que (AB)C représente $(f \circ g) \circ h$, d'où le résultat annoncé.

Enfin, on a les relations de *distributivité*

$$A(B + C) = AB + AC, \qquad (A + B)C = AC + BC,$$

qu'on démontre, comme la formule d'associativité, en remplaçant A, B, C par des homomorphismes de modules et en appliquant le Théorème 1.

En particulier, pour tout entier $n \geqslant 1$ et tout anneau K, désignons par la notation

$$M_n(K).$$

l'ensemble des matrices carrées d'ordre n (i.e. à n lignes et n colonnes) à coefficients dans K; on a sur cet ensemble deux lois de composition

$$(A, B) \mapsto A + B, \qquad (A, B) \mapsto AB;$$

cela dit, $M_n(K)$, muni de ces deux lois de composition, est un *anneau*. On sait en effet déjà que, relativement à l'addition, $M_n(K)$ est un groupe commutatif; d'autre part, on vient de démontrer que la multiplication est associative, et il est clair (*Exemple* 3 ci-dessus) qu'elle admet un élément neutre, à savoir 1_n; enfin, les formules précédentes montrent que, dans $M_n(K)$, la multiplication est distributive par rapport à l'addition.

On dit que $M_n(K)$ est **l'anneau des matrices d'ordre n à coefficients dans K**. Il est évidemment isomorphe à l'anneau des endomorphismes du K-module à droite K^n (ou de tout module admettant une base de n vecteurs).

Même lorsque K *est commutatif, l'anneau* $M_n(K)$ *n'est jamais commutatif si* $n \geqslant 2$; pour $n = 2$, il suffit pour le voir d'observer que

$$\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & xy \\ 0 & 1 \end{pmatrix}$$

tandis que

$$\begin{pmatrix} 1 & y \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix};$$

si l'anneau $M_2(K)$ était commutatif on aurait donc $xy = y$ quels que soient $x, y \in K$, ce qui a peu de chances de se produire si K possède au moins deux éléments...

Ce sont naturellement les anneaux $M_n(K)$ qui, historiquement, ont fourni les premiers exemples d'anneaux non commutatifs.

Comme dans tout anneau, on dit que deux matrices X et Y carrées d'ordre n **commutent** ou sont **permutables** lorsqu'on a

$$XY = YX.$$

Exemple 4. Prenons pour X une **matrice diagonale**, i.e. de la forme

$$X = \begin{pmatrix} x_1 & 0 & 0 & \dots & 0 \\ 0 & x_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & x_n \end{pmatrix}$$

et cherchons à quelle condition une matrice $Y = (y_{ij})_{1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n}$ commute à X; un calcul simple montre que

$$XY = (x_i y_{ij}), \qquad YX = (y_{ij} x_j);$$

il est donc nécessaire et suffisant que l'on ait

$$x_i y_{ij} = y_{ij} x_j \quad \text{pour} \quad 1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant n.$$

Supposons que K soit un *anneau d'intégrité commutatif* et que les termes x_i de X soient *deux à deux distincts*; la condition précédente, qui s'écrit

$$(x_i - x_j)\, y_{ij} = 0$$

vu la commutativité de K, montre, puisque K est intègre et que $x_i \neq x_j$ pour $i \neq j$, que l'on doit avoir $y_{ij} = 0$ pour $i \neq j$, autrement dit que Y doit être une *matrice diagonale*; cette condition est du reste suffisante attendu que deux matrices diagonales, à coefficients dans un anneau commutatif K, commutent toujours comme le montre le calcul précédent.

4. *Écriture matricielle des homomorphismes*

Soient L et M des K-modules à droite libres de type fini, et f un homomorphisme de L dans M. Choisissons une base $(a_i)_{1 \leqslant i \leqslant p}$ de L et une base $(b_j)_{1 \leqslant j \leqslant q}$ de M, et soit

$$A = (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$$

la matrice de f par rapport aux basses considérées. Étant donné un vecteur

$$x = a_1 \xi_1 + \dots + a_p \xi_p,$$

le vecteur

$$f(x) = b_1 \eta_1 + \dots + b_q \eta_q$$

est donnée par

$$\begin{aligned} \eta_1 &= \alpha_{11} \xi_1 + \dots + \alpha_{p1} \xi_p \\ &\dots\dots\dots\dots\dots\dots \\ \eta_q &= \alpha_{1q} \xi_1 + \dots + \alpha_{pq} \xi_p : \end{aligned}$$

or ces formules signifient visiblement qu'on a la relation

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots & \cdots & \cdots \\ \alpha_{1q} & \cdots & \alpha_{pq} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_q \end{pmatrix}.$$

Autrement dit, *si l'on identifie chaque* $x \in \mathrm{L}$ *à la matrice colonne formée avec ses coordonnées par rapport à la base* (a_i) *de* L, *et chaque* $y \in \mathrm{M}$ *à la matrice colonne formée avec ses coordonnées par rapport à la base* (b_j) *de* M, *la relation*

$$y = f(x)$$

entre les vecteurs x *et* y *est équivalente à la relation*

$$y = \mathrm{A}x$$

entre les matrices x *et* y. Ce résultat permet de calculer de façon quasi mécanique avec les homomorphismes de modules libres de type fini.

§ 15. Matrices inversibles et changements de bases

1. *Le groupe des automorphismes d'un module*

Rappelons (§ 12, n° 1) qu'on appelle *automorphisme* d'un module M tout homomorphisme bijectif de M dans M, i.e. tout isomorphisme de M sur M. Les automorphismes de M sont donc en particulier des permutations de l'ensemble M; et le Théorème 1 du § 12 montre que, si u et v sont des automorphismes de M, il en est de même de l'application $u \circ v^{-1}$; par suite, l'ensemble

$$\mathrm{GL}(M)$$

des automorphismes du module M est un sous-groupe du groupe $\mathfrak{S}(M)$ des permutations de l'ensemble M; on dit que GL(M) est le **groupe des automorphismes du module** M, ou encore le **groupe linéaire du module** M. Les groupes de la forme GL(M), et leurs sous-groupes, ont joué un rôle de premier plan dans le développement de la théorie générale des groupes.

On remarquera qu'on pourrait aussi définir le groupe GL(M) à partir de l'anneau Hom (M, M) des endomorphismes du module M; on a évidemment

$$\mathrm{GL}(M) \subset \mathrm{Hom}\ (M, M),$$

et les éléments de GL(M) ne sont autres que les éléments inversibles de l'anneau Hom (M, M); si en effet un endomorphisme u est inversible dans l'anneau Hom (M, M), il existe un endomorphisme v tel que $u \circ v = v \circ u = j_M$, et par suite u est bijectif, donc appartient à GL(M); et la réciproque est évidente. En conclusion, GL(M) *n'est autre que le groupe des éléments inversibles de l'anneau* Hom (M, M), au sens du § 8, *Remarque* 1.

2. *Les groupes* GL (n, K)

On dit qu'une matrice $U \in M_n(K)$ est **inversible** s'il existe une matrice $V \in M_n(K)$ telle que

$$UV = VU = I_n.$$

autrement dit si U est un élément inversible de l'anneau $M_n(K)$; la matrice V est alors unique, se note U^{-1}, et s'appelle l'**inverse de U**. On note

$$GL(n, K)$$

l'ensemble des matrices carrées inversibles de degré n à coefficients dans K; muni de la loi de composition $(U, V) \mapsto UV$, cet ensemble est un groupe, appelé le **groupe linéaire à n variables sur l'anneau K**; ce n'est pas autre chose, par conséquent, que le groupe multiplicatif des éléments inversibles de l'anneau $M_n(K)$.

Soit M un module *libre de type fini*, et choisissons une base $a_1, \ldots, a_n$ de M; à chaque endomorphisme f de M on peut alors attacher sa matrice par rapport à la base en question (§ 12, n° 3); en la notant $A(f)$, on obtient une *bijection*

$$f \mapsto A(f)$$

de l'anneau Hom (M, M) sur l'anneau $M_n(K)$ des matrices carrées d'ordre n à coefficients dans K, et la façon même dont on a défini la somme et le produit de deux matrices montre qu'on a les relations

$$A(f+g) = A(f) + A(g), \qquad A(fg) = A(f)A(g), \qquad A(j_M) = 1_n,$$

autrement dit que l'application $f \to A(f)$ est un *isomorphisme* (§ 8, n° 6) de l'anneau Hom (M, M) sur l'anneau $M_n(K)$.

Comme un isomorphisme d'un anneau U sur un anneau V applique évidemment l'ensemble U^* des éléments inversibles de U sur l'ensemble V^* des éléments inversibles de V, on en conclut que, *pour qu'un endomorphisme f de* M *soit un automorphisme de* M *il faut et il suffit que sa matrice* $A(f)$ *soit un élément inversible de l'anneau* $M_n(K)$; autrement dit, les relations

$$f \in GL(M) \qquad \text{et} \qquad A(f) \in GL(n, K)$$

sont équivalentes.

Si l'on identifie $M_n(K)$ à l'anneau des endomorphismes du K-module à droite K^n comme on l'a dit au § 14, n° 3, on voit donc que $GL(n, K)$ *s'identifie au groupe des automorphismes de* K^n, autrement dit au groupe des applications

$$(\xi_1, \ldots, \xi_n) \mapsto (\eta_1, \ldots, \eta_n)$$

de K^n dans K^n qui sont d'une part *bijectives* et d'autre part *linéaires*, i.e. données par des formules du type

$$\eta_j = \alpha_1\xi_1 + \cdots + \alpha_n\xi_n \quad (1 \leqslant j \leqslant n).$$

3. *Exemples : les groupes* GL (1, K) *et* GL (2, K)

Pour $n = 1$, l'anneau $M_n(K)$ est identique à l'anneau K lui-même, et par suite que le groupe GL(1, K) se réduit au groupe multiplicatif K^* des éléments inversibles de K.

Pour étudier le groupe GL(2, K), nous supposerons K *commutatif*. Pour qu'une

matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

admette une inverse

$$A^{-1} = \begin{pmatrix} x & z \\ y & t \end{pmatrix}$$

il faut qu'il existe $x, y, z, t \in K$ tels que

$$(1) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x & z \\ y & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

i.e. tels que

$$(2) \qquad \begin{cases} ax + by = 1 & az + bt = 0 \\ cx + dy = 0 & cz + dt = 1. \end{cases}$$

Pour trouver des conditions nécessaires et suffisantes de résolubilité des équations (2) introduisons la notion de *déterminant* d'une matrice carrée d'ordre 2 à coefficients dans un anneau commutatif K : pour une matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

on appelle ainsi le scalaire $ad - bc$ qu'on désigne par

$$\text{dét}(A) \qquad \text{ou} \qquad \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

Nous étendrons plus loin cette définition aux matrices carrées d'ordre n quelconque.

Pour le moment, il nous suffira de savoir qu'on a

$$\text{dét}(AB) = \text{dét}(A) \cdot \text{dét}(B)$$

quelles que soient $A, B \in M_2(K)$; cette relation est en effet équivalente à l'identité

$$(ad - bc)(xt - yz) = (ax + by)(cz + dt) - (az + bt)(cx + dy),$$

que le lecteur vérifiera sans mal en tenant compte de la commutativité de K.

Cela fait, et comme le déterminant de la matrice unité 1_2 est évidemment égal à 1, la relation

$$\text{dét}(A) \cdot \text{dét}(A^{-1}) = 1$$

montre que, pour que la matrice A soit inversible, il est nécessaire que son déterminant le soit. Inversement, supposons $ad - bc$ inversible ; considérons dans le système (2) les deux relations en x et y ; elles seront visiblement vérifiées si l'on prend

$$x = (ad - bc)^{-1}d, \qquad y = (ad - bc)^{-1}c,$$

et les relations en z et t le seront pour

$$z = -(ad - bc)^{-1}b, \qquad t = (ad - bc)^{-1}a;$$

on vérifie facilement que la matrice

$$\begin{pmatrix} x & z \\ y & t \end{pmatrix}$$

ainsi construire est inverse à droite *et* à gauche de A.

En conclusion, si K est un anneau commutatif, la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *est inversible si et seulement si* $ad - bc$ *est inversible dans* K.

Si par exemple K est un *corps commutatif*, GL(2, K) est formé des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ telles que

$$ad - bc \neq 0.$$

Par contre, le groupe GL(2, **Z**) est formé des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ à coefficients entiers telles que

$$ad - bc = +1 \quad \text{ou} \quad -1.$$

Remarque 1. Soient L un anneau et K un sous-anneau de L; on peut évidemment considérer $M_n(K)$ comme un sous-anneau de $M_n(L)$. Cela dit, il peut arriver qu'une matrice

$$U \in M_n(K)$$

soit inversible dans l'anneau $M_n(L)$ sans l'être dans l'anneau $M_n(K)$: c'est ainsi que la matrice (2), ou la matrice $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, est inversible lorsqu'on la regarde comme matrice à coefficients dans le corps **Q**, mais ne l'est pas en tant que matrice à coefficients dans l'anneau **Z**. La notion de matrice inversible risque donc de prêter à confusion si l'on ne précise pas l'anneau de base choisi.

Cependant ce genre de difficulté ne se présente pas lorsque les anneaux de base considérés sont des *corps*; autrement dit, soient L un corps et K un sous-corps de L; alors si une matrice $U \in M_n(K)$ est inversible dans l'anneau $M_n(L)$, elle l'est déjà dans l'anneau $M_n(K)$. Voir une démonstration au § 20, *Exercice* 20; si L est commutatif, on peut aussi utiliser le Corollaire 1 du Théorème 18 du § 23.

4. *Changements de bases : matrices de passage*

Soit M un K-module à droite libre de type fini, et considérons deux bases $(a_1, \ldots a_n)$ et $(b_1, \ldots, b_n)$ de M ayant le même nombre n d'éléments [cette condition est d'ailleurs toujours réalisée si K est un corps (§ 19, Théorème 6), ou bien est commutatif (§ 23, Corollaire du Théorème 5)]. On se propose, pour chaque $x \in M$, de calculer les coordonnées de x par rapport à la seconde base en fonction de ses coordonnées par rapport à la première.

Considérons pour cela les homomorphismes

$$u, v : K^n \to M$$

donnés par

$$u(e_i) = a_i \qquad v(e_i) = b_i \qquad (1 \leqslant i \leqslant n) \tag{3}$$

où $e_1, \ldots, e_n$ est la base canonique de K^n. Pour que (a_i) et (b_i) soient des bases de M il faut et il suffit que u et v soient des isomorphismes (§ 12, Corollaire 1 du Théorème 3) et si l'on désigne par $\xi_1, \ldots, \xi_n$ les coordonnées de x par rapport à la base (a_i), par $\eta_1, \ldots, \eta_n$ ses coordonnées par rapport à la base (b_i), on a les relations

$$x = u(\xi_1, \ldots, \xi_n) = v(\eta_1, \ldots, \eta_n);$$

en introduisant l'*automorphisme* (§ 12, Théorème 1)

$$w = v^{-1} \circ u \tag{4}$$

de K^n, on a donc

$$(\eta_1, \ldots, \eta_n) = w(\xi_1, \ldots, \xi_n);$$

et en désignant par

$$(\alpha_{ij})_{1 \leqslant i, j \leqslant n}$$

la matrice de w par rapport à la base canonique de K^n, on obtient les formules

$$\eta_j = \alpha_{1j}\xi_1 + \cdots + \alpha_{nj}\xi_n \quad (1 \leqslant j \leqslant n) \tag{5}$$

qui résolvent le problème posé. On les appelle les **formules de changement de coordonnées**, et la matrice (α_{ij}) qui figure dans ces formules s'appelle la **matrice de passage de la base $(a_i)_{1 \leqslant i \leqslant n}$ à la base $(b_i)_{1 \leqslant i \leqslant n}$**; comme c'est la matrice d'un *auto*morphisme de K^n, on a

$$(\alpha_{ij}) \in \mathrm{GL}(n, K).$$

On observera que la relation (4) s'écrit aussi $u = v \circ w$, et implique donc

$$v = u \circ w^{-1}; \tag{6}$$

soit

$$(\beta_{ij}) = (\alpha_{ij})^{-1}$$

la matrice de w^{-1} (par rapport à la base canonique de K^n), matrice qui est nécessairement l'inverse de la matrice de passage (α_{ij}); on a les relations

$$w^{-1}(e_i) = e_1\beta_{i1} + \cdots + e_n\beta_{in},$$

et en tenant compte de (3) il vient

$$\begin{aligned} b_i = v(e_i) &= u[w^{-1}(e_i)] = u(e_1\beta_{i1} + \cdots + e_n\beta_{in}) \\ &= u(e_1)\beta_{i1} + \cdots + u(e_n)\beta_{in} \end{aligned}$$

i.e.

$$b_i = a_1\beta_{i1} + \cdots + a_n\beta_{in} \qquad (1 \leqslant i \leqslant n); \tag{7}$$

on aura soin d'observer que les matrices (α_{ij}) et (β_{ij}) figurant dans les relations (5) et (7) sont non pas identiques mais *inverses l'une de l'autre*.

On démontrerait par un calcul analogue les relations

$$a_i = b_1\alpha_{i1} + \cdots + b_n\alpha_{in}. \tag{7 bis}$$

Remarque 2. Le cas le plus simple est celui où $n = 1$; on a alors deux bases (a) et (b) comprenant chacune un vecteur, avec les relations

$$x = a\xi = b\eta;$$

posant

$$\eta = \alpha\xi$$

on doit avoir $a\xi = b\alpha\xi$ quel que soit ξ, donc $a = b\alpha$, i.e.

$$b = a\alpha^{-1}.$$

Autrement dit, si l'on remplace le vecteur de base a par le vecteur $a\alpha$, la coordonnée ξ de x est remplacée par $\alpha^{-1}\xi$, ce qui est conforme au bon sens puisque le produit $a\xi$ doit rester constant !

Les résultats précédents permettent, à partir d'une base de M, d'en construire toutes les autres :

THÉORÈME 1. *Soit $a_1, \ldots, a_n$ une base d'un* K*-module à droite* M. *Pour que les vecteurs*

$$b_i = a_1\beta_{i1} + \cdots + a_n\beta_{in} \quad (1 \leqslant i \leqslant n)$$

forment une base de M, *il faut et il suffit que la matrice $(\beta_{ij})_{1 \leqslant i, j \leqslant n}$ soit inversible dans l'anneau* $\mathrm{M}_n(\mathrm{K})$.

Considérons en effet les homomorphismes $u, v : \mathrm{K}^n \to \mathrm{M}$ donnés par les formules (3); on sait, puisque les a_i forment une base, que u est bijectif. Pour exprimer que les b_i forment une base, on doit exprimer que v est bijectif, ou, ce qui revient évidemment au même, que $u^{-1} \circ v$ est un *auto*morphisme de K^n; or on a

$$\begin{aligned} u^{-1} \circ v(e_i) &= u^{-1}(b_i) = u^{-1}(a_1\beta_{i1} + \cdots + a_n\beta_{in}) \\ &= u^{-1}(a_1)\beta_{i1} + \cdots + u^{-1}(a_n)\beta_{in} \\ &= e_1\beta_{i1} + \cdots + e_n\beta_{in}; \end{aligned}$$

il s'ensuit que (β_{ij}) est précisément la matrice de l'endomorphisme $u^{-1} \circ v$ de K^n; comme un endomorphisme de K^n est bijectif si et seulement si sa matrice est dans $\mathrm{GL}(n, \mathrm{K})$, le Théorème est démontré.

La condition du Théorème 1 peut encore s'obtenir en introduisant l'endomorphisme f de M tel que

$$f(a_i) = b_i \quad (1 \leqslant i \leqslant n)$$

(son existence résulte du § 12, Théorème 3); pour que les b_i forment une *base* de M, il faut et il suffit que f soit un *automorphisme* de M (§ 12, Corollaire 1 du Théorème 3); or, la matrice (β_{ij}) n'est autre que la matrice de f par rapport à la base (a_i) de M.

Ce résultat est d'ailleurs évident directement, car en introduisant à nouveau les homomorphismes u et v ci-dessus on a

$$f = v \circ u^{-1};$$

comme u est bijectif, dire que v est bijectif revient à dire que f l'est...

5. *Influence d'un changement de bases sur la matrice d'un homomorphisme*

Soient L et M deux K-modules à droite libres de type fini et $f : L \to M$ un homomorphisme. Soient

$$(a'_1, \ldots, a'_p) \qquad \text{et} \qquad (a''_1, \ldots, a''_p)$$

deux bases de L possédant le même nombre p d'éléments, et

$$(b'_1, \ldots, b'_q) \qquad \text{et} \qquad (b''_1, \ldots, b''_q)$$

deux bases de M ayant le même nombre q d'éléments. Soit A' la matrice de f par rapport aux bases (a'_i) et (b'_j), et soit A'' sa matrice par rapport aux bases (a''_i) et (b''_j); on se propose de *calculer* A'' *en fonction de* A', *de la matrice*

$$U \in GL(p, K)$$

qui fait passer de la base (a'_i) *à la base* (a''_i), *et de la matrice*

$$V \in GL(q, K)$$

qui fait passer de la base (b'_j) *à la base* (b''_j).

Pour cela considérons les homomorphismes u'; $u'' : K^p \to L$ qui appliquent la base canonique de K^p sur les bases (a'_i) et (a''_i) respectivement, et les homomorphismes v', $v'' : K^q \to M$ qui appliquent la base canonique de K^q sur les bases (b'_j) et (b''_j) de M respectivement. On a des relations $u' = u'' \circ u$, $v' = v'' \circ v$ où u (resp. v) est un automorphisme de K^p (resp. K^q) dont la matrice par rapport à la base canonique de K^p (resp. K^q) est précisément U (resp. V) en vertu du n° précédent. D'autre part, si l'on introduit les homomorphismes

$$f', f'' : K^p \to K^q$$

donnés par

$$f' = v'^{-1} \circ f \circ u', \qquad f'' = v''^{-1} \circ f \circ u''$$

alors les matrices A' et A'' définies plus haut ne sont autres que les matrices de f' et f'' par rapport aux bases canoniques de K^p et K^q en vertu du § 12, *Remarque* 2. Or en introduisant les applications u et v définies ci-dessus il vient

$$\begin{aligned} f' &= (v'' \circ v)^{-1} \circ f \circ (u'' \circ u) \\ &= v^{-1} \circ v''^{-1} \circ f \circ u'' \circ u = v^{-1} \circ f'' \circ u; \end{aligned}$$

comme la composition de deux homomorphismes se traduit par la multiplication de leurs matrices, on a donc, en prenant les matrices de u, v, f', f'' par rapport aux bases canoniques, la relation

$$A' = V^{-1} . A'' . U,$$

laquelle résoud le problème posé au début de ce n°. Ainsi :

THÉORÈME 2. *Soient* L *et* M *deux* K*-modules à droite libres de type fini,* f *un homomorphisme*

de L *dans* M, A′ *la matrice de f par rapport à une base* $(a'_i)_{1\leqslant i\leqslant p}$ *de* L *et à une base* $(b'_j)_{1\leqslant j\leqslant q}$ *de* M, *et* A″ *sa matrice par rapport à une base* $(a''_i)_{1\leqslant i\leqslant p}$ *de* L *et à une base* $(b''_j)_{1\leqslant j\leqslant q}$ *de* M. *Soient enfin* U *la matrice de passage de la base* (a'_i) *à la base* (a''_i), *et* V *la matrice de passage de la base* (b'_j) *à la base* (b''_j). *On a alors la relation*

$$A' = V^{-1}A''U.$$

Lorsqu'en particulier L = M, on peut supposer dans ce qui précède que la base (b'_j) est identique à la base (a'_i), et la base (b''_j) identique à la base (a''_i), auquel cas on a évidemment U = V; donc :

COROLLAIRE. *Soient* L *un* K-*module à droite libre de type fini, f un endomorphisme de* L, *et*

$$(a'_i)_{1\leqslant i\leqslant p}, \qquad (a''_i)_{1\leqslant i\leqslant p}$$

deux bases de L *ayant le même nombre d'éléments. Soient* A′ *la matrice de f par rapport à la base* (a'_i) *et* A″ *sa matrice par rapport à la base* (a''_i). *On a alors la relation*

$$A'' = UA'U^{-1}$$

où U *est la matrice de passage de la base* (a'_i) *à la base* (a''_i).

Ce dernier résultat conduit à une notion importante : étant données des matrices

$$A', A'' \in M_p(K),$$

on dit que A′ et A″ sont **semblables** (sur l'anneau de base K) s'il existe une matrice

$$U \in GL(p, K)$$

telle que l'on ait la relation

$$A'' = UA'U^{-1}.$$

Remarque 3. Il existe toujours (Théorème 1) un changement de base admettant pour matrice de passage une matrice inversible arbitrairement choisie. On voit donc que le Théorème 2 admet une réciproque : si l'on se donne d'avance l'homomorphisme *f*, et les bases (a'_i) et (b'_j), donc la matrice A′, alors quelles que soient les matrices $U \in GL(p, K)$ et $V \in GL(q, K)$ il existe dans L et M des bases par rapport auxquelles *f* est représenté par la matrice

$$V^{-1}A'U.$$

On a un résultat analogue dans la situation décrite par le Corollaire.

Remarque 4. La démonstration que nous avons donnée du Théorème 2 évite tout calcul explicite, mais oblige par contre à passer par l'intermédiaire des modules « prototypes » K^p et K^q, ce qui risque de gêner le lecteur débutant.

On peut encore démontrer le Théorème 2 comme suit. Soient U et V les matrices de passage, et posons

$$U^{-1} = (u_{ij})_{1\leqslant i,\, j\leqslant p}, \qquad V^{-1} = (v_{kh})_{1\leqslant k,\, h\leqslant q};$$

d'après le n° 4 on a alors les relations

$$(8) \qquad a'_i = \sum_j a'_{j\omega_{ij}}; \qquad b'_k = \sum_k b'_h \rho_{kh};$$

d'autre part, en posant

$$A' = (\alpha'_{jh})_{1 \leqslant j \leqslant p,\ 1 \leqslant h \leqslant q}$$
$$A'' = (\alpha''_{ik})_{1 \leqslant i \leqslant p,\ 1 \leqslant k \leqslant q}$$

il vient

$$(9) \qquad f(a'_j) = \sum_h b'_h \alpha'_{jh}; \qquad f(a''_i) = \sum_k b''_k \alpha''_{ik};$$

tenant compte de (8), la seconde relation (9) s'écrit

$$\sum_j f(a'_j)\omega_{ij} = \sum_k \sum_h b'_h \rho_{kh} \alpha''_{ik},$$

ou encore, d'après la première relation (9),

$$\sum_j \sum_h b'_h \alpha'_{jh} \omega_{ij} = \sum_k \sum_h b'_h \rho_{kh} \alpha''_{ik};$$

comme les vecteurs b'_h sont linéairement indépendants, leurs coefficients dans les deux membres doivent être égaux, ce qui conduit à la relation

$$\sum_j \alpha'_{jh} \omega_{ij} = \sum_k \rho_{kh} \alpha''_{ik},$$

valable quels que soient i et h. Or au premier membre figure le coefficient d'indices h et i de la matrice $A'U^{-1}$, et au second membre le coefficient d'indices h et i de la matrice $V^{-1}A''$; il vient donc $A'U^{-1} = V^{-1}A''$, d'où la relation cherchée $A'' = VA'U^{-1}$.

Les calculs de ce genre étaient autrefois fréquents en Algèbre linéaire et dans la théorie des « tenseurs », et impressionnaient grandement (et à juste titre) les nombreuses personnes qui croyaient qu'Einstein était seul à pouvoir comprendre ses propres travaux. Aujourd'hui, la plupart des mathématiciens préfèrent remplacer les déluges d'indices par des raisonnements géométriques ou, pour mieux dire, conceptuels, qui ont l'avantage d'être beaucoup plus simples. Néanmoins, la plupart des physiciens utilisent encore des méthodes analogues à celle qu'on a exposée dans cette *Remarque* (ce qui est d'autant plus étrange qu'un physicien, encore plus qu'un mathématicien, devrait s'intéresser aux objets « géométriques » ou « physiques » et non pas à leurs coordonnées, tout au moins aussi longtemps qu'il n'a pas en vue des calculs effectifs). Il est donc utile de se familiariser avec les calculs sur les indices et les Σ, même en sachant qu'ils sont théoriquement superflus.

§ 16. Transposée d'une application linéaire

1. *Dual d'un module*

Soit L un module *à droite* sur un anneau quelconque K. Rappelons (§ 12, n° 4, *Exemple* 3) qu'on appelle *forme linéaire sur* L tout homomorphisme de L dans le K-module à droite K, autrement dit toute application

$$f\colon \mathrm{L} \to \mathrm{K}$$

telle que l'on ait

$$f(x\alpha + y\beta) = f(x)\alpha + f(y)\beta \tag{1}$$

quels que soient $x, y \in \mathrm{L}$ et $\alpha, \beta \in \mathrm{K}$. En vertu du § 13, ces formes linéaires sur L sont les éléments du *groupe commutatif* Hom (L, K), la somme $f + g$ de deux formes linéaires sur L étant par définition la fonction $f(x) + g(x)$.

Nous allons voir (en utilisant le fait que K est non seulement un K-module à droite mais aussi un K-module à gauche) qu'en fait on peut regarder l'ensemble Hom (L, K) non seulement comme un groupe commutatif mais même comme un K-*module à gauche*. Soient pour cela une forme linéaire f sur L et un scalaire $\lambda \in \mathrm{K}$; considérons sur L la fonction g donnée par

$$g(x) = \lambda . f(x);$$

en multipliant à gauche par λ la relation (1), et en tenant compte des règles de calcul (associativité, distributivité) dans un anneau, il vient

$$g(x\alpha + y\beta) = g(x)\alpha + g(y)\beta,$$

ce qui prouve que g est encore une forme linéaire sur L. On écrit naturellement

$$g = \lambda f,$$

et on a ainsi défini que l'ensemble Hom (L, K) une seconde opération, consistant à « multiplier » un élément de cet ensemble par un scalaire. Il resterait à voir que

l'ensemble Hom (L, K), muni de l'addition définie au § 13 et de la seconde opération que nous venons de définir, est effectivement un K-module à gauche; on laisse au lecteur le soin de le faire à titre d'exercice (on peut aussi utiliser le fait suivant : soit E l'ensemble de *toutes* les applications de L dans K, linéaires ou non; en regardant L comme un simple ensemble, et K comme un K-module à gauche, l'*Exemple* 4 du § 10 permet de considérer E comme un K-module à gauche; cela dit, pour faire de même avec Hom (L, K), il suffit de montrer que Hom (L, K) est un sous-module de E — ce qui était évidemment le but des considérations qui précèdent).

L'ensemble Hom (L, K), muni de la « structure » de K-module à gauche que nous venons de définir, s'appelle le **dual du K-module à droite** L; on le désigne généralement par la notation

$$L^*,$$

plus commode que Hom (L, K).

Si l'on partait d'un K-module *à gauche* L, on définirait de même son dual; ce serait, cette fois, un K-module *à droite.*

Rappelons à ce sujet que les distinctions entre « droite » et « gauche » n'ont aucun intérêt si l'anneau de base K est commutatif, ce qui, dans la pratique, est presque toujours le cas.

Remarque 1. Soit f une forme linéaire sur un K-module à droite L; pour montrer que λf est encore linéaire, on peut aussi observer qu'elle est composée de f et de l'application $\xi \mapsto \lambda\xi$ de K dans K; il suffit donc de montrer que cette dernière application est un endomorphisme du K-module *à droite* K, autrement dit vérifie

$$\lambda(\xi + \eta_1) = \lambda\xi + \lambda\eta_1, \qquad \lambda(\xi\mu) = (\lambda\xi)\,\mu,$$

ce qui est clair.

Soit L un K-module à droite; puisqu'on a défini sur l'ensemble L^* des formes linéaires sur L une structure de K-module à gauche, on peut appliquer à L^* les définitions et théorèmes de la théorie des modules. En particulier, étant données des formes linéaires $f_1, \ldots, f_p$ sur L, on dira qu'une forme linéaire f sur L est *combinaison linéaire de* $f_1, \ldots, f_p$ s'il existe des scalaires $\lambda_1, \ldots, \lambda_p \in K$ tels que l'on ait

$$f = \lambda_1 f_1 + \cdots + \lambda_p f_p,$$

relation qui, vu la définition des opérations sur les formes linéaires, signifie que

$$f(x) = \lambda_1 f_1(x) + \cdots + \lambda_p f_p(x) \quad \text{pour tout} \quad x \in L.$$

De même, on appellera *relation linéaire entre* $f_1, \ldots, f_p$ tout système

$$(\lambda_1, \ldots, \lambda_p) \in K^p$$

de scalaires tels que l'on ait

$$\lambda_1 f_1 + \cdots + \lambda_p f_p = 0$$

dans L^*, i.e.

$$\lambda_1 . f_1(x) + \cdots + \lambda_p . f_p(x) = 0 \quad \text{pour tout} \quad x \in L,$$

etc, etc...

2. *Dual d'un module libre de type fini*

Dans la pratique élémentaire, le débutant n'aura pas besoin de résultat plus « profond » que le suivant :

THÉORÈME 1. *Soient* L *un* K*-module à droite libre de type fini et* $(a_1, \ldots, a_n)$ *une base de* L. *Considérons l'application* $\theta : L^* \to K^n$, *donnée par*

$$\theta(f) = (f(a_1), \ldots, f(a_n));$$

alors θ *est un isomorphisme de* K*-modules à gauche, et* L^* *possède une base* $(u_1, \ldots, u_n)$ *telle que l'on ait*

$$u_i(a_j) = \begin{cases} 1 & si\ i = j \\ 0 & si\ i \neq j. \end{cases}$$

Soient $\alpha_1, \ldots, \alpha_n$ des éléments arbitraires de K; le Théorème 3 du § 12 montre qu'il existe *une et une seule* $f \in L^*$ telle que l'on ait $f(a_i) = \alpha_i$ pour $1 \leqslant i \leqslant n$, i.e. telle que

$$\theta(f) = (\alpha_1, \ldots, \alpha_n);$$

l'application θ est donc *bijective*. Pour montrer que c'est un isomorphisme, il reste à faire voir qu'elle est *linéaire*; or on a, pour $f, g \in L^*$ et en posant $f + g = h$,

$$\begin{aligned} \theta(f + g) &= (h(a_1), \ldots, h(a_n)) = (f(a_1) + g(a_1), \ldots, f(a_n) + g(a_n)) \\ &= (f(a_1), \ldots, f(a_n)) + (g(a_1), \ldots, g(a_n)) = \theta(f) + \theta(g); \end{aligned}$$

d'autre part, en remplaçant f par λf, on multiplie évidemment les coefficients $f(a_i)$ à gauche par λ, d'où l'identité

$$\theta(\lambda f) = \lambda . \theta(f),$$

de sorte que θ est bien linéaire.

Pour terminer la démonstration du Théorème, il reste à prouver l'existence d'une base $(u_i)_{1 \leqslant i \leqslant n}$ de L^* possédant les propriétés indiquées; or les relations imposées aux u_i signifient que

$$\begin{aligned} \theta(u_1) &= (1, 0, 0, \ldots, 0) \\ \theta(u_2) &= (0, 1, 0, \ldots, 0) \\ &\cdots\cdots\cdots\cdots\cdots \\ \theta(u_n) &= (0, 0, \ldots, 0, 1), \end{aligned}$$

autrement dit que θ applique la base $(u_i)_{1 \leqslant i \leqslant n}$ de L^* sur la base canonique de K^n; comme θ est un isomorphisme de modules, l'existence de la base (u_i) cherchée est claire : il suffit de prendre les images par l'application θ^{-1} des éléments de la base canonique de K^n, ce qui termine la démonstration.

On notera que les relations imposées aux u_i donnent, pour tout élément

$$x = a_1\xi_1 + \cdots + a_n\xi_n$$

de L, la relation

$$u_i(x) = u_i(a_1)\xi_1 + \cdots + u_i(a_n)\xi_n = \xi_i;$$

autrement dit, les u_i ne sont autres que les *fonctions coordonnées du module* L *par rapport à la base* $(a_i)_{1\leqslant i\leqslant n}$ *de* L (§ 11, n° 4). Pour toute $f \in L^*$, soient $\alpha_1, \ldots, \alpha_n$ les coordonnées de f par rapport à la base $(u_i)_{1\leqslant i\leqslant n}$ de L^*; la relation

$$f = \alpha_1 u_1 + \ldots + \alpha_n u_n$$

s'écrit alors

$$f(x) = \alpha_1\xi_1 + \ldots + \alpha_n\xi_n \quad \text{pour tout } x \in L,$$

et par suite il vient

$$\alpha_i = f(a_i);$$

autrement dit, *les coordonnées de* f *par rapport à la base* (u_i) *de* L^* *ne sont autres que les coefficients de* f *par rapport à la base* (a_i) *de* L, définis au § 12, *Exemple* 3.

Le Théorème 1 montre qu'à toute base (a_i) de L, on peut associer une base (u_i) de L^*; on dit que (u_i) est la **base duale de la base** (a_i) de L.

3. *Bidual d'un module*

Soit L un module *à droite* sur un anneau K; nous lui avons attaché un module *à gauche* L^* sur K; celui-ci possède à son tour un dual, qu'on note

$$L^{**} = (L^*)^*,$$

et qu'on appelle le **bidual de** L; comme L, c'est un K-module *à droite*. On définirait de même le **tridual** $L^{***} = (L^{**})^*$, le **quadridual** $L^{****} = (L^{***})^*$, et ainsi de suite indéfiniment.

On peut, dans tous les cas, définir une **application canonique de** L **dans** L^{**}, de la façon suivante. Soit x un élément « fixe » de L, et considérons l'application

$$u : L^* \to K$$

donnée par

$$u(f) = f(x) \qquad \text{pour tout} \quad f \in L^*;$$

elle consiste donc à associer à chaque forme linéaire f sur L sa valeur au point x de L. L'application u est *linéaire*, i.e. vérifie

$$u(\alpha f + \beta g) = \alpha . u(f) + \beta . u(g)$$

quels que soient $f, g \in L^*$ et les scalaires $\alpha, \beta \in K$; posant

$$\alpha f + \beta g = h,$$

la relation en question s'écrit en effet

$$h(x) = \alpha f(x) + \beta g(x),$$

et sous cette forme elle se réduit purement et simplement à la *définition* même de l'élément $h = \alpha f + \beta g$ de L^*.

Ainsi, u est une forme linéaire sur L^*, autrement dit $u \in L^{**}$, et de cette façon nous avons bien attaché à chaque $x \in L$ un $u \in L^{**}$; d'où une application de L dans L^{**}, et c'est, par définition, l'application canonique de L dans son bidual.

Cette application est d'ailleurs *linéaire*; soient en effet $x, y \in L$, $\alpha, \beta \in K$ et posons $z = x\alpha + y\beta$; soient $u, v, w \in L^{**}$ les images de x, y, z par l'application canonique; tout revient à établir la relation

$$w = u\alpha + v\beta;$$

mais comme u, v, w sont des formes linéaires sur L^*, celle-ci signifie

$$w(f) = u(f)\alpha + v(f)\beta \qquad \text{pour tout} \quad f \in L^*;$$

or, par définition de l'application canonique de L dans L^{**}, on a

$$u(f) = f(x), \qquad v(f) = f(y), \qquad w(f) = f(z)$$

et par suite tout revient à montrer que $f(z) = f(x)\alpha + f(y)\beta$, ou, en remplaçant z par sa valeur, que

$$f(x\alpha + y\beta) = f(x)\alpha + f(y)\beta \qquad \text{pour toute} \quad f \in L^*;$$

or cette identité n'est autre que celle qui définit les formes linéaires sur L.

THÉORÈME 2. *Soit* L *un module libre de type fini; l'application canonique de* L *dans son bidual est un isomorphisme.*

Pour établir ce Théorème il nous reste à faire voir que l'application canonique

$$j : L \to L^{**}$$

est bijective si L est libre de type fini. Or soient $(a_i)_{1 \leqslant i \leqslant n}$ une base de L, $(f_i)_{1 \leqslant i \leqslant n}$ la base duale dans L^*, et posons

$$u_i = j(a_i) \in L^{**};$$

il suffit évidemment de montrer que les u_i forment une *base* de L^{**} : nous allons en fait montrer qu'ils forment la base de L^{**} duale de la base (f_i) du module L^*, autrement dit qu'on a les relations

$$u_i(f_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Or, par définition de l'homomorphisme canonique de L dans L**, on a

$$u_i(f) = f(a_i) \quad \text{pour tout } f \in \mathrm{L}^*,$$

et par suite $u_i(f_j) = f_j(a_i)$; mais comme les f_j forment la base duale de la base (a_i) de L, on a les relations

$$f_j(a_i) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j, \end{cases}$$

ce qui fournit les relations cherchées et achève la démonstration.

COROLLAIRE 1. *Soient* L *un* K*-module à droite libre de type fini et* u *une forme linéaire sur le module dual* L*. Alors il existe un et un seul* $x \in \mathrm{L}$ *tel que l'on ait*

$$u(f) = f(x) \quad \textit{pour tout } f \in \mathrm{L}^*.$$

En effet, il existe un et seul $x \in \mathrm{L}$ tel que $u = j(x)$, où j désigne l'isomorphisme canonique de L sur son bidual.

COROLLAIRE 2. *Soient* L *un* K*-module à droite libre de type fini et* $(f_1, \ldots, f_n)$ *une base du module dual* L*. Alors, quels que soient* $\beta_1, \ldots, \beta_n \in \mathrm{K}$, *il existe un et un seul* $x \in \mathrm{L}$ *tel que l'on ait*

$$f_i(x) = \beta_i \quad (1 \leqslant i \leqslant n).$$

En effet, puisque les f_i forment une base de L* il existe une et une seule forme linéaire u sur L* telle que $u(f_i) = \beta_i$ pour tout i (§ 12, Théorème 3), et comme $u(f) = f(x)$ pour un $x \in \mathrm{L}$ entièrement déterminé par u, le Corollaire est démontré.

4. *Transposé d'un homomorphisme*

Soient L et M deux K-modules à droite et $f : \mathrm{L} \to \mathrm{M}$ un homomorphisme. Soit u une forme linéaire sur M; alors l'application composée $u \circ f$ est évidemment (§ 12, Théorème 1) une forme linéaire sur L. On peut donc définir une application

$$^t f : \mathrm{M}^* \to \mathrm{L}^*$$

en posant

$$^t f(u) = u \circ f \quad \text{pour tout } u \in \mathrm{M}^*.$$

L'application $^t f$ s'appelle la **transposée de l'homomorphisme** f.

Cette application est, comme f, un *homomorphisme*. Soient en effet $u, v \in \mathrm{M}^*$; on a

$$^t f(u + v) = (u + v) \circ f = u \circ f + v \circ f$$

en vertu du § 14, Théorème 1, et par suite

$$^t f(u + v) = {}^t f(u) + {}^t f(v);$$

de même, pour tout $\lambda \in K$ notons h_λ l'homothétie $\xi \to \lambda\xi$ dans K; on a alors, pour tout $u \in M^*$,

$$ {}^t f(\lambda u) = {}^t f(h_\lambda \circ u) = (h_\lambda \circ u) \circ f = h_\lambda \circ (u \circ f) = h_\lambda \circ {}^t f(u) = \lambda \cdot {}^t f(u) $$

ce qui établit la linéarité de l'application transposée ${}^t f$.

Remarque 2. La démonstration que nous venons de donner risque de paraître inintelligible au lecteur débutant, à cause de son aspect purement mécanique. Bien entendu on conseille vivement au lecteur de s'exercer à la retraduire en langage clair (en ramenant tout à la *définition* de la structure de module du dual d'un module). Cependant, on doit aussi faire observer que lorsqu'on énonce un théorème (par exemple le Théorème 1 du § 14) c'est dans l'espoir de s'en servir à l'occasion ! Cette Remarque s'applique aussi à la démonstration du résultat suivant.

THÉORÈME 2. *L'opération consistant à passer d'un homomorphisme à son transposé possède les propriétés suivantes :*

a) Soient L, M *deux* K*-modules à droite et* $f, g : L \to M$ *deux homomorphismes; on a*

$$ {}^t(f + g) = {}^t f + {}^t g. $$

b) Soient L, M, N *trois* K*-modules à droite,* $f : L \to M$ *et* $g : M \to N$ *deux homomorphismes; on a*

$$ {}^t(g \circ f) = {}^t f \circ {}^t g. $$

c) Soit L *un* K*-module à droite; le transposé d'un automorphisme (resp. de l'automorphisme identique) de* L *est un automorphisme (resp. l'automorphisme identique) de* L^*.

Pour établir l'assertion *a*), posons $h = f + g$; pour $u \in M^*$ on a

$$ {}^t h(u) = u \circ h = u \circ (f + g) = u \circ f + u \circ g = {}^t f(u) + {}^t g(u), $$

ce qui prouve que ${}^t h = {}^t f + {}^t g$ comme annoncé.

Pour établir *b*), posons $h = g \circ f$; pour $u \in N^*$ on a

$$ {}^t h(u) = u \circ h = u \circ (g \circ f) = (u \circ g) \circ f = {}^t g(u) \circ f = {}^t f[{}^t g(u)], $$

d'où ${}^t h = {}^t f \circ {}^t g$, ce qui prouve *b*).

Pour établir *c*) montrons d'abord que si $f : L \to L$ est l'application identique il en est de même de son transposé; on a en effet, pour toute forme linéaire u sur L,

$$ {}^t f(u) = u \circ f = u $$

puisque f est l'identité, d'où notre assertion. Ceci dit supposons que f soit un automorphisme de L; il y a donc un homomorphisme g de L tel que

$$ f \circ g = g \circ f = j_L; $$

il s'ensuit que

$$ {}^t g \circ {}^t f = {}^t f \circ {}^t g = {}^t(j_L) = j_{L^*}, $$

et par suite ${}^tf$ est bien un automorphisme de L^*. On a du reste, comme le montre le calcul précédent, la relation

$$ {}^t(f^{-1}) = ({}^tf)^{-1} $$

pour tout automorphisme f de L.

Remarque 3. On aura soin de ne pas remplacer l'assertion *b*) du Théorème 3 par la formule « évidente », mais fausse (et même dépourvue de sens), que voici :

$$ {}^t(f \circ g) = {}^tf \circ {}^tg. $$

5. *Transposée d'une matrice*

Soit $f : L \to M$ un homomorphisme de K-modules à droite libres de type fini. Choisissons une base $(a_1, \ldots, a_p)$ de L et une base $(b_1, \ldots, b_q)$ de M; désignons par $(u_1, \ldots, u_p)$ la base de L^* duale de la base $(a_1, \ldots, a_p)$ de L, et par $(v_1, \ldots, v_q)$ la base de M^* duale de la base $(b_1, \ldots, b_q)$ de M. Par rapport aux bases (a_i) et (b_j), l'homomorphisme

$$ f : L \to M $$

est représenté par une matrice de la forme

$$ (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}; $$

et par rapport aux bases (v_j) et (u_i), l'homomorphisme transposé

$$ {}^tf : M^* \to L^* $$

est représenté par une matrice de la forme

$$ (\beta_{ji})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}; $$

on se propose de calculer celle-ci en fonction de la matrice de f.

Par définition de la matrice de ${}^tf$, on a les relations

$$ {}^tf(v_j) = \beta_{j1}u_1 + \cdots + \beta_{jp}u_p \quad (1 \leqslant j \leqslant q) $$

i.e.

$$ v_j \circ f = \beta_{j1}u_1 + \cdots + \beta_{jp}u_p; $$

cela signifie que pour tout $x \in L$ on a

$$ v_j[f(x)] = \beta_{j1}.u_1(x) + \cdots + \beta_{jp}.u_p(x); \tag{γ} $$

or en posant

$$ x = a_1\xi_1 + \cdots + a_p\xi_p, \quad f(x) = b_1\tau_{11} + \cdots + b_q\eta_q $$

on a, d'après le n° 2, les relations

$$ u_i(x) = \xi_i, \qquad v_j[f(x)] = \tau_{ij}; $$

par suite la relation (2) équivaut à

$$\tau_{ij} = \beta_{j1}\xi_1 + \cdots + \beta_{jp}\xi_p,$$

et comme par définition de la matrice (α_{ij}) de f on a aussi les relations

$$\tau_{ij} = \alpha_{1j}\xi_1 + \cdots + \alpha_{pj}\xi_p,$$

on voit en comparant les deux résultats que

$$\beta_{ji} = \alpha_{ij} \qquad \text{pour} \quad 1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q.$$

Ce résultat conduit à introduire la définition suivante. Étant donnée une matrice

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots\cdots\cdots & & \\ \alpha_{1q} & \cdots & \alpha_{pq} \end{pmatrix}$$

à coefficients dans un anneau K, on appelle **transposée de** A la matrice

$${}^tA = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1q} \\ \cdots\cdots\cdots & & \\ \alpha_{p1} & \cdots & \alpha_{pq} \end{pmatrix}$$

obtenue, comme on dit, en « permutant les lignes et les colonnes » de A.

¶ *Remarque* 4. Le passage de A à ${}^tA$ correspond évidemment à celui de f à ${}^tf$ dans ce qui précède. Or, si f est un homomorphisme de K-modules *à droite*, ${}^tf$ est au contraire un homomorphisme de K-modules *à gauche*, i.e. de modules à droite sur l'anneau K^0 opposé à K (§ 10, n° 4). Comme le calcul des matrices est adapté, lorsque K est non commutatif, à la théorie des modules *à droite*, *on doit donc regarder la transposée* ${}^tA$ *d'une matrice* A *à coefficients dans un anneau* K *comme une matrice à coefficients dans l'anneau opposé* K^0. C'est ce qu'on fera, sans y référer à nouveau, dans ce qui suit.

Bien entendu, les considérations qui précèdent sont superflues si l'anneau K est commutatif.

THÉORÈME 4. *Soit* K *un anneau.*

a) Étant données deux matrices A *et* B *à coefficients dans* K *on a la relation*

$${}^t(A + B) = {}^tA + {}^tB$$

pourvu que la somme A + B *soit définie.*

b) Étant données deux matrices A *et* B *à coefficients dans* K, *on a la relation* (*)

$${}^t(AB) = {}^tB.{}^tA$$

pourvu que le produit AB *soit défini.*

(*) Conformément à la *Remarque* 4, le produit ${}^tB.{}^tA$ doit être calculé dans l'anneau K^0 opposé à K (i.e. dans l'anneau K si K est commutatif).

c) Si A *est une matrice carrée à coefficients dans* K, *pour que* A *soit inversible il faut et il suffit que* ^{t}A *le soit.*

d) Pour toute matrice A *à coefficients dans* K *on a*

$$ {}^t({}^t\mathrm{A}) = \mathrm{A}. $$

Pour prouver *a*), posons

$$ \mathrm{A} = (\alpha_{ij}), \qquad \mathrm{B} = (\beta_{ij}); $$

alors $\mathrm{A} + \mathrm{B} = (\alpha_{ij} + \beta_{ji})$, et par suite

$$ {}^t(\mathrm{A} + \mathrm{B}) = (\alpha_{ji} + \beta_{ji}) = (\alpha_{ji}) + (\beta_{ji}) = {}^t\mathrm{A} + {}^t\mathrm{B}. $$

Pour établir maintenant l'assertion *b*), posons

$$ \mathrm{A} = (\alpha_{jk})_{1 \leqslant j \leqslant q,\ 1 \leqslant k \leqslant r}\ \mathrm{B} = (\beta_{ij})_{1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q}; $$

la matrice

$$ \mathrm{AB} = (\gamma_{ik})_{1 \leqslant i \leqslant p,\ 1 \leqslant k \leqslant r} $$

est donnée par les relations

$$ \gamma_{ik} = \sum_j \alpha_{jk}\beta_{ij} \tag{3} $$

(cf. § 14, n° 2), les produits $\alpha_{jk}\beta_{ij}$ étant bien entendu calculés dans l'anneau K donné. D'autre part on a

$$ {}^t\mathrm{A} = (\alpha'_{kj})_{1 \leqslant k \leqslant r,\ 1 \leqslant j \leqslant q} \qquad \text{avec} \qquad \alpha'_{kj} = \alpha_{jk}, $$
$$ {}^t\mathrm{B} = (\beta'_{ji})_{1 \leqslant j \leqslant q,\ 1 \leqslant i \leqslant p} \qquad \text{avec} \qquad \beta'_{ji} = \beta_{ij}; $$

posant

$$ {}^t\mathrm{B}\,{}^t\mathrm{A} = (\gamma'_{ji})_{1 \leqslant k \leqslant r,\ 1 \leqslant i \leqslant p}, $$

on a

$$ \gamma'_{ki} = \sum_j \beta'_{ji}\alpha'_{kj}, $$

où le produit $\beta'_{ji}\alpha'_{kj}$ est calculé dans l'anneau opposé à K. En calculant dans K, et en tenant compte des relations $\beta'_{ji} = \beta_{ij}$, $\alpha'_{kj} = \alpha_{jk}$, il vient donc

$$ \gamma'_{ki} = \sum_j \alpha'_{kj}\beta'_{ji} = \sum_j \alpha_{jk}\beta_{ij} = \gamma_{ik} $$

en vertu de (3); ce qui prouve la relation figurant dans l'énoncé *b*).

L'assertion *d*) est triviale.

Il reste à prouver l'assertion *c*). Soit $\mathrm{A} \in \mathrm{M}_n(\mathrm{K})$; supposons A inversible et soit B son inverse; des relations $\mathrm{AB} = \mathrm{BA} = 1_n$ résultent les relations ${}^t\mathrm{B}\,{}^t\mathrm{A} = {}^t\mathrm{A}\,{}^t\mathrm{B} = {}^t(1_n)$;

or il est clair que

$$ {}^t(1_n) = 1_n; $$

donc ${}^tA$ est inversible, et de plus le calcul fait montre que

$$ ({}^tA)^{-1} = {}^t(A^{-1}). $$

Réciproquement, si ${}^tA$ est inversible, ce qu'on vient d'établir montre qu'il en est de même de la matrice ${}^t({}^tA)$, i.e. de A elle-même. Le Théorème 4 est donc entièrement démontré.

Remarque 5. On pourrait évidemment déduire les assertions *a*) et *b*) du Théorème 4 des assertions analogues du Théorème 3; on laisse au lecteur le soin de le faire à titre d'exercice.

§ 17. Sommes de sous-modules

1. Somme de deux sous-modules

Soient K un anneau, L un K-module à gauche, M et N deux sous-modules de L. On appelle **somme** de M et N le sous-module de L engendré par l'ensemble $M \cup N$, autrement dit le plus petit sous-module de L contenant à la fois M et N.

Il est facile de donner une construction explicite de ce sous-module; c'est *l'ensemble des* $z \in L$ *pour lesquels il existe un* $x \in M$ *et un* $y \in N$ *tels que*

$$z = x + y.$$

En effet, il est clair que tout sous-module de L contenant M et N contient tout vecteur z possédant la propriété en question. Il suffit donc de montrer que ces vecteurs forment un sous-module P contenant M et N. Or il est clair que P contient M et N (faire $x = 0$ ou bien $y = 0$ dans la relation $z = x + y$); d'autre part, si

$$\begin{aligned} z' &= x' + y' \qquad (x' \in M, y' \in N) \\ z'' &= x'' + y'' \qquad (x'' \in M, y'' \in N) \end{aligned}$$

sont deux éléments de P, alors quels que soient les scalaires λ et μ on a

$$\lambda z' + \mu z'' = x + y$$

avec

$$x = \lambda x' + \mu x'' \in M, \qquad y = \lambda y' + \mu y'' \in N,$$

ce qui montre bien que P est un sous-module.

Étant donné ce qu'on vient d'établir, on désigne la somme des sous-modules M et N par la notation (*)

$$M + N.$$

(*) Plus généralement, si G est un groupe multiplicatif (resp. additif) et si A et B sont des parties de G, on note AB (resp. A + B) l'ensemble des $z \in G$ pour lesquels il existe $x \in A$ et $y \in B$ tels que $z = xy$ (resp. $z = x + y$).

Exemple 1. Prenons $K = \mathbf{R}$ et pour L l'ensemble des vecteurs d'origine donnée O dans l'espace usuel; prenons pour M et N des droites passant par O; alors $M + N$ est le plan engendré par ces deux droites si elles sont distinctes, ou bien $M + N = M = N$ si elles sont confondues.

La définition et la construction précédentes se généralisent immédiatement à un nombre quelconque de sous-modules. Soit $M_1, \ldots, M_p$ une famille finie quelconque de sous-modules de L; désignons par P l'ensemble des $z \in L$ qui possèdent la propriété suivante : il existe $x_1 \in M_1, \ldots, x_p \in M_p$ tels que

$$z = x_1 + \cdots + x_p; \tag{1}$$

alors P est le plus petit sous-module de L contenant $M_1, \ldots, M_p$.

Il est d'abord clair que P contient M_i (faire $x_j = 0$ pour $j \neq i$ dans la relation précédente), et qu'un sous-module de L contenant les M_i contient nécessairement les vecteurs (1). Il reste donc à faire voir que P est un sous-module de L; mais si

$$z' = x'_1 + \cdots + x'_p, \qquad z'' = x''_1 + \cdots + x''_p$$

sont des éléments de P, on a quels que soient $\lambda', \lambda'' \in K$ la relation

$$\lambda' z' + \lambda'' z'' = x_1 + \cdots + x_p$$

avec

$$x_i = \lambda' x'_i + \lambda'' x''_i \in M_i \quad (1 \leqslant i \leqslant p),$$

ce qui établit le résultat cherché.

Ici encore, on dit que le sous-module P est la **somme des sous-modules** $M_1, \ldots, M_p$ et on le désigne par la notation

$$M_1 + \cdots + M_p.$$

2. *Produit direct de modules*

Dans ce qui précède, les modules $M_1, \ldots, M_p$ étaient des sous-modules d'un module donné L. On va maintenant partir de K-modules à gauche $M_1, \ldots, M_p$ et, sans les supposer contenus dans un même module, construire un module qui les contient tous à un isomorphisme près.

Considérons pour cela l'ensemble produit

$$L = M_1 \times \cdots \times M_p,$$

formé des familles

$$x = (x_1, \ldots, x_p) \qquad \text{avec} \qquad x_1 \in M_1, \ldots, x_p \in M_p.$$

Nous allons définir sur L une structure de K-module à gauche en posant

$$(x_1, \ldots, x_p) + (y_1, \ldots, y_p) = (x_1 + y_1, \ldots, x_p + y_p)$$
$$\lambda(x_1, \ldots, x_p) = (\lambda x_1, \ldots, \lambda x_p);$$

le fait qu'on obtienne une structure de module sur L de cette façon se vérifie immédiatement, et nous laisserons au lecteur le soin de le faire.

Lorsque $M_1 = \cdots = M_p = K$, on retrouve évidemment le module K^p.

Lorsque $K = \mathbf{Z}$, on retrouve la notion de produit direct de groupes commutatifs définie au § 7, n° 2.

Dans le cas général, on dit que $M_1 \times \cdots \times M_p$, muni de la structure de module qu'on vient de définir, est le **produit direct des modules** $M_1, \ldots, M_p$.

Il est clair que, pour $1 \leqslant i \leqslant p$, l'application

$$\mathrm{pr}_i : M_1 \times \cdots \times M_p \to M_i$$

donnée par $\mathrm{pr}_i(x_1, \ldots, x_p) = x_i$ est un homomorphisme de modules. Il en est de même de l'application

$$u_i : M_i \to M_1 \times \cdots \times M_p$$

donnée par

$$u_i(x) = (0, \ldots, 0, x, 0, \ldots, 0),$$

la lettre $x \in M_i$ étant précédée au second membre de $i - 1$ zéros.

En fait, il est clair que l'homomorphisme u_i est injectif; c'est donc un isomorphisme de M_i sur un sous-module de $M_1 \times \cdots \times M_p$. Dans la pratique, on identifie le plus souvent M_i à son image par u_i. La formule

$$(x_1, x_2, \ldots, x_p) = (x_1, 0, \ldots, 0) + (0, x_2, 0, \ldots, 0) + \cdots + (0, \ldots, 0, x_p)$$

montre alors que tout élément de $M_1 \times \cdots \times M_p$ est somme d'un élément de M_1, d'un élément de M_2, ..., d'un élément de M_p; autrement dit, $M_1 \times \cdots \times M_p$ est somme des sous-modules $M_1, \ldots, M_p$.

3. *Somme directe de sous-modules*

Reprenons comme au n° 1 des sous-modules $M_1, \ldots, M_p$ d'un K-module à gauche L. Considérons l'application

$$f : M_1 \times \cdots \times M_p \to L$$

donnée par

$$f(x_1, \ldots, x_p) = x_1 + \cdots + x_p;$$

c'est évidemment un homomorphisme de modules. De plus ,on a

$$\mathrm{Im}\,(f) = M_1 + \cdots + M_p$$

par définition même de la somme d'une famille de sous-modules.

On dit que les sous-modules $M_1, \ldots, M_p$ sont **linéairement indépendants** lorsque l'homomorphisme f est *injectif*, autrement dit lorsque tout $x \in M_1 + \cdots + M_p$ s'écrit *d'une seule façon* sous la forme $x = x_1 + \cdots + x_p$ avec $x_1 \in M_1, \ldots, x_p \in M_p$. Il revient au même de dire (§ 7, Théorème 8) que $\mathrm{Ker}\,(f)$ est réduit à 0, autrement dit

que les relations

$$x_1 + \cdots + x_p = 0, \qquad x_1 \in M_1, \ldots, x_p \in M_p$$

impliquent

$$x_1 = \cdots = x_p = 0.$$

THÉORÈME 1. *Pour que deux sous-modules* M *et* N *d'un module* L *soient linéairement indépendants il faut et il suffit que* $M \cap N = 0$.

Si $x \in M \cap N$, on a $x + (-x) = 0$ avec $x \in M$ et $-x \in N$, donc $x = 0$ si M et N sont linéairement indépendants. Inversement supposons $M \cap N = 0$; si $x \in M$ et $y \in N$ vérifient $x + y = 0$, on a $x = -y \in M \cap N$, donc $x = y = 0$, ce qui achève la démonstration.

Remarque 1. Considérons l'application $f : M \times N \to L$ donnée par

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

le raisonnement qui précède montre que son noyau est formé des couples $(x, -x)$ avec $x \in M \cap N$; il est clair que l'application $x \to (x, -x)$ est un *isomorphisme de* $M \cap N$ *sur* $\mathrm{Ker}(f)$. Cette Remarque est fort importante pour calculer la dimension d'une somme de sous-espaces vectoriels (cf. § 19, n° 7).

Lorsque des sous-modules $M_1, \ldots, M_p$ d'un module L sont linéairement indépendants, on dit que $M_1 + \cdots + M_p$ est la **somme directe** des sous-modules donnés, et on désigne cette somme directe par la notation

$$M_1 \oplus \cdots \oplus M_p;$$

l'emploi du signe $\oplus$ au lieu du signe $+$ habituel indique donc qu'on a affaire à une somme de sous-modules linéairement indépendants. Il est clair qu'alors l'application f définie plus haut est un *isomorphisme du module* $M_1 \times \cdots \times M_p$ *sur le module* $M_1 \oplus \cdots \oplus M_p$; cet isomorphisme sera qualifié de « canonique ».

Remarque 2. Le module produit $M_1 \times \cdots \times M_p$ est évidemment somme directe des sous-modules auxquels on a identifié $M_1, \ldots, M_p$ au n° précédent.

Remarque 3. Soient $a_1, \ldots, a_p$ des éléments d'un K-module à gauche L, et prenons pour $M_1, \ldots, M_p$ les sous-modules engendrés par $a_1, \ldots, a_p$ respectivement. Alors $M_1 + \cdots + M_p$ est le sous-module de L engendré par $a_1, \ldots, a_p$; les sous-modules $M_1, \ldots, M_p$ sont linéairement indépendants si et seulement si les vecteurs $a_1, \ldots, a_p$ le sont; et la relation

$$L = M_1 \oplus \cdots \oplus M_p$$

signifie que $a_1, \ldots, a_p$ forment une *base* de L. On laisse au lecteur le soin de vérifier lui-même ces assertions.

Soient L un K-module à gauche et M un sous-module de L; on dit que M est **facteur direct** dans L s'il existe un sous-module N de L tel que L soit somme *directe*

de M et de N; on dit alors que N est un **supplémentaire de M dans L**. On verra au § 19 que si L est un espace vectoriel de dimension finie sur un corps, tout sous-espace de L admet un supplémentaire. Mais cette propriété ne s'étend pas aux anneaux quelconques.

Exemple 2. Prenons $K = L = \mathbf{Z}$ et $M = p\mathbf{Z}$ avec $p \neq 0$; soit $N = q\mathbf{Z}$ un sous-module non nul de L; on a alors $M \cap N \neq 0$, par exemple parce que $pq \in M \cap N$; par suite (Théorème 1) M n'admet pas de supplémentaire dans $\mathbf{Z}$ (sauf bien entendu si $M = L$ ou si $M = 0$).

4. *Sommes directes et projecteurs*

Soit L un K-module à gauche, et considérons une **décomposition de L en somme directe** i.e. une relation de la forme

$$L = M_1 \oplus \cdots \oplus M_p$$

où les M_i sont des sous-modules de L, linéairement indépendants. Tout $x \in L$ s'écrit donc d'une façon et d'une seule sous la forme

$$x = x_1 + \cdots + x_p \qquad \text{avec} \qquad x_i \in M_i \quad \text{pour} \quad 1 \leqslant i \leqslant p,$$

de sorte qu'on peut poser

$$x_i = v_i(x)$$

où v_i est une application de L dans L. D'ailleurs, si l'on introduit l'isomorphisme

$$f : M_1 \times \cdots \times M_p \to L$$

donné par

$$f(x_1, \ldots, x_p) = x_1 + \cdots + x_p,$$

et les injections canoniques

$$j_i : M_i \to L,$$

il est clair qu'on a

$$v_i = j_i \circ pr_i \circ f^{-1},$$

ce qui montre que v_i est linéaire (résultat que le lecteur démontrera aussi par un calcul direct).

L'endomorphisme v_i a évidemment pour image le sous-module M_i; pour tout $x \in L$, $v_i(x)$ est le seul et unique vecteur de M_i tel que $x - v_i(x)$ appartienne au sous-module engendré par $M_1, \ldots, M_{i-1}, M_{i+1}, \ldots, M_p$; on dit pour cette raison que $v_i(x)$ est la **projection de x sur M_i**.

Il est clair qu'on a $v_i(x) = x$ si et seulement si $x \in M_i$. Comme $v_i(x) \in M_i$ pour tout $x \in L$, il s'ensuit que $v_i(v_i(x)) = v_i(x)$ pour tout $x \in L$, ce qui veut dire que

$$v_i \circ v_i = v_i;$$

on dit qu'un endomorphisme v d'un module L est un **projecteur** s'il vérifie la condition

précédente, i.e. si

$$v \circ v = v.$$

D'autre part, il est aussi évident que $v_j(x) = 0$ pour $x \in \mathrm{M}_i$, $i \neq j$; donc on a $v_j(v_i(x)) = 0$ pour tout $x \in \mathrm{L}$ si $i \neq j$, autrement dit

$$v_j \circ v_i = 0 \quad \text{si} \quad i \neq j.$$

Enfin, pour tout $x \in \mathrm{L}$ on a $x = x_1 + \cdots + x_p = v_1(x) + \cdots + v_p(x)$, ce qui prouve que

$$v_1 + \cdots + v_p = j_{\mathrm{L}},$$

application identique de L dans lui-même.

Ces propriétés admettent une réciproque :

THÉORÈME 2. *Soient $v_1, \ldots, v_p$ des endomorphismes d'un module* L *vérifiant les conditions suivantes :*

$$v_j \circ v_i = \begin{cases} v_i & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j, \end{cases}$$
$$v_1 + \cdots + v_p = j_{\mathrm{L}}.$$

Alors L *est somme directe des sous-modules* $\mathrm{M}_i = v_i(\mathrm{L})$.

La relation

$$x = v_1(x) + \cdots + v_p(x)$$

montre déjà que L est somme des M_i. Considérons maintenant des $x_i \in \mathrm{M}_i$ $(1 \leqslant i \leqslant p)$ tels que

$$x_1 + \cdots + x_p = 0; \tag{2}$$

il existe des $y_i \in \mathrm{L}$ tels que $x_i = v_i(y_i)$; on a donc

$$v_i(x_j) = v_i[v_j(y_j)] = \begin{cases} 0 & \text{si } i \neq j \\ v_i(y_i) = x_i & \text{si } i = j; \end{cases}$$

il s'ensuit qu'en appliquant v_i à la relation (2) il reste la relation $x_i = 0$, ce qui montre que les sous-modules M_i sont linéairement indépendants, et achève la démonstration.

COROLLAIRE. *Soit* M *un sous-module d'un module* L. *Les propriétés suivantes sont équivalentes :*

(FD 1) M *est facteur direct dans* L;

(FD 2) *il existe un endomorphisme v de* L *tel que*

$$v \circ v = v, \qquad v(\mathrm{L}) = \mathrm{M};$$

(FD 3) *il existe un homomorphisme q de* L *dans* M *tel que $q(x) = x$ pour tout $x \in \mathrm{M}$.*

Il est clair que (FD 1) implique (FD 2) : on écrit $L = M \oplus N$ et on prend pour $v(x)$ la projection de x sur M parallèlement à N.

(FD 2) implique (FD 3) : comme v applique L sur M, on peut définir un homomorphisme q du module L dans le module M en posant $q(x) = v(x)$ pour tout $x \in L$ (la seule différence entre v et q est que v est une application de L dans L, tandis que q est une application de L dans M) ; pour $x \in M$, il existe un $y \in L$ tel que $v(y) = x$, et on a alors $q(x) = v(x) = v(v(y)) = v(y) = x$.

Montrons enfin que (FD 3) implique (FD 1). Soit N le noyau de q. On a

$$M \cap N = 0,$$

car si $x \in M \cap N$ on a d'une part $q(x) = x$, d'autre part $q(x) = 0$, donc $x = 0$. Par ailleurs, pour tout $x \in L$, on a $q(q(x)) = q(x)$, donc $q(x - q(x)) = 0$, donc $x - q(x) \in N$; écrivant $x = q(x) + (x - q(x))$ on voit donc que $L = M + N$. Comme $M \cap N = 0$, ceci prouve (Théorème 1) que M est facteur direct dans L et achève la démonstration.

Espaces vectoriels de dimension finie

Le but des §§ 18 à 20 est d'étudier les espaces vectoriels de dimension finie sur un corps quelconque K, en particulier d'introduire la notion fondamentale de dimension, et d'établir les propriétés les plus importantes des systèmes d'équations linéaires sur un corps.

Bien que, dans la pratique élémentaire, et en Analyse, on ne s'intéresse qu'au cas du corps des nombres réels ou des nombres complexes, il ne servirait *à rien* de supposer $K = \mathbf{R}$ ou $K = \mathbf{C}$ dans ce Chapitre.

Les exposés classiques de l'Algèbre linéaire utilisent d'autre part la théorie des déterminants, mais il y a plus de cinquante ans qu'on sait s'en passer; l'exposé purement « géométrique » ainsi obtenu est non seulement plus simple, il est aussi plus général que ceux reposant sur la théorie des déterminants, car celle-ci suppose le corps ou l'anneau de base commutatif. En fait, la théorie des déterminants (qui sera exposée dans des §§ ultérieurs) a pour principal intérêt de fournir des critères *explicites* d'indépendance linéaire, et des formules *explicites* de résolution des systèmes d'équations linéaires; elle n'est d'aucune utilité pour établir les théorèmes *d'existence* des §§ 19 et 20.

L'un des résultats les plus importants de l'Algèbre linéaire est le Théorème 13 du § 19, qui, étant donné un homomorphisme f d'un espace vectoriel L de dimension finie dans un autre, établit une relation simple entre les dimensions de L, du noyau de f, et de l'image de f. Quand on analyse la démonstration classique de ce Théorème, on constate qu'on peut le formuler de telle sorte qu'il se généralise aux modules sur un anneau quelconque. Le résultat général permet alors de démontrer très simplement des énoncés qui sont le point de départ de la « grande » théorie des anneaux dits noethériens et des anneaux principaux. Ces résultats font l'objet du § 18; l'étude de ce § et des *Exercices* correspondants, quoique peu utile en principe pour comprendre la suite, donnera au lecteur un aperçu de l'une des théories les plus importantes de l'Algèbre actuelle.

§ 18. Théorèmes de finitude

1. *Homomorphismes dont le noyau et l'image sont de type fini* (*)

Les résultats de ce § reposent sur le théorème suivant :

THÉORÈME 1. *Soient* K *un anneau,* L *et* M *deux* K*-modules à gauche, et* f *un homomorphisme de* L *dans* M.

Si Ker (f) *et* Im (f) *sont des* K*-modules de type fini, il en est de même de* L.

Si Ker (f) *est isomorphe à* K^p, *et* Im (f) *isomorphe à* K^q, *alors* L *est isomorphe à* K^{p+q}.

Supposons Ker (f) engendré par des vecteurs $a_1, \ldots, a_p$, et Im (f) engendré par des vecteurs $b_1, \ldots, b_q$; choisissons dans L des vecteurs $a_{p+1}, \ldots, a_{p+q}$ tels que

$$b_1 = f(a_{p+1}), \ \ldots, \ b_q = f(a_{p+q});$$

pour démontrer le Théorème 1, il suffit d'établir d'une part que $a_1, \ldots a_{p+q}$ engendrent L dans tous les cas, d'autre part que ces $p + q$ vecteurs forment une *base* de L dans le cas où $a_1, \ldots, a_p$ forment une *base* de Ker (f) et $b_1, \ldots, b_q$ une *base* de Im (f).

Pour établir la première assertion considérons un vecteur $x \in L$. Comme $f(x)$ appartient au sous-module de M engendré par $b_1, \ldots, b_q$, il existe des scalaires τ_{ij} $(1 \leqslant j \leqslant q)$ tels que

$$f(x) = \tau_{11} b_1 + \cdots + \tau_{1q} b_q,$$

ce qui s'écrit encore

$$f(x) = \tau_{11} f(a_{p+1}) + \cdots + \tau_{1q} f(a_{p+q}) = f(\tau_{11} a_{p+1} + \cdots + \tau_{1q} a_{p+q});$$

on a donc

$$x = \tau_{11} a_{p+1} + \cdots + \tau_{1q} a_{p+q} + y$$

(*) Le lecteur débutant pourra se borner à lire le n° 1 de ce §, les autres n'étant pas utilisés dans la suite. Néanmoins la lecture des n° 2 à 5 sera certainement un exercice très profitable même pour le débutant.

avec $y \in \mathrm{Ker}\ (f)$; mais alors il existe des scalaires $\xi_i (1 \leqslant i \leqslant p)$ tels que

$$y = \xi_1 a_1 + \cdots + \xi_p a_p,$$

et en portant ce résultat dans la relation précédente on voit que x est bien une combinaison linéaire des vecteurs $a_1, \ldots, a_{p+q}$.

Pour établir la seconde assertion, il suffit de montrer que si les vecteurs $a_i (1 \leqslant i \leqslant p)$ sont linéairement indépendants, ainsi que les vecteurs $b_j (1 \leqslant j \leqslant q)$, alors il en est de même de $a_1, \ldots, a_{p+q}$. Or considérons une relation de la forme

$$\lambda_1 a_1 + \cdots + \lambda_{p+q} a_{p+q} = 0;$$

appliquons f au premier membre; comme $a_1, \ldots, a_p \in \mathrm{Ker}\ (f)$, il vient

$$\lambda_{p+1} f(a_{p+1}) + \cdots + \lambda_{p+q} f(a_{p+q}) = 0;$$

or les vecteurs $f(a_{p+1}) = b_1, \ldots, f(a_{p+q}) = b_q$ sont par hypothèse linéairement indépendants; on voit donc que

$$\lambda_{p+1} = \cdots = \lambda_{p+q} = 0;$$

la relation initiale se réduit alors à $\lambda_1 a_1 + \cdots + \lambda_p a_p = 0$, et comme $a_1, \ldots, a_p$ sont linéairement indépendants il s'ensuit que

$$\lambda_1 = \cdots = \lambda_p = 0,$$

ce qui termine la démonstration.

Remarque 1. On verra plus loin (§ 19, Théorème 13) que lorsque K est un *corps*, de sorte qu'on peut parler de la « dimension » d'un espace vectoriel de dimension finie sur K, le Théorème 1 signifie que

$$\dim (\mathrm{L}) = \dim [\mathrm{Ker}\ (f)] + \dim [\mathrm{Im}\ (f)].$$

Il va de soi que, comme toujours, il ne servirait à rien de supposer que K est un corps dans la démonstration du Théorème 1, et la suite de ce § montrera que, si l'on veut exploiter toutes les conséquences de ce Théorème, il est au contraire tout à fait essentiel de l'énoncer en toute généralité.

2. *Modules de type fini sur un anneau noethérien*

Soit I un *idéal à gauche* d'un anneau K (autrement dit, un sous-module de K regardé comme K-module à gauche); rappelons (§ 11, *Exemple* 6) que I est dit *de type fini* s'il est de type fini comme K-module à gauche, autrement dit s'il existe un nombre fini d'éléments $x_1, \ldots, x_n$ de I tels que I soit l'ensemble des éléments de K qui peuvent se mettre sous la forme $u_1 x_1 + \cdots + u_n x_n$.

On dit qu'un anneau K est **noethérien à gauche** lorsque *tout* idéal à gauche de K est de type fini. On définirait de même les anneaux noethériens à droite en considérant les idéaux à droite. Lorsque K est commutatif (cas de beaucoup le plus important dans ce contexte), on dit simplement que K est **noethérien.**

Exemple 1. Un corps est un anneau noethérien (car ses seuls idéaux à gauche sont $\{0\}$ et K, qui sont évidemment de type fini). Un anneau principal (§ 8, *Exemple* 10) est aussi noethérien.

En dehors des anneaux principaux, les exemples les plus importants d'anneaux noethériens sont les anneaux de polynomes à n indéterminées à coefficients dans un corps (cf. § 32. Exercice 27); ces anneaux — qui ne sont pas principaux pour $n \geqslant 2$ — jouent un rôle fondamental dans l'étude des « variétés algébriques », i.e. des « courbes », « surfaces », etc... définies par des équations algébriques.

THÉORÈME 2. *Soit* K *un anneau. Les propriétés suivantes sont équivalentes :*

a) *L'anneau* K *est noethérien à gauche.*

b) *Tout sous-module de tout* K*-module à gauche de type fini est lui-même de type fini.*

Il est immédiat de voir que *b*) implique *a*) : en effet, le K-module à gauche K est de type fini, donc ses sous-modules (i.e. les idéaux à gauche de K) doivent être de type fini si *b*) est vérifiée.

Montrons maintenant que *a*) implique *b*). On va procéder en deux temps : tout d'abord montrer que, pour tout entier $n \geqslant 1$, tout sous-module de K^n est de type fini; puis en déduire *b*) en toute généralité.

Montrons donc que tout sous-module L de K^n est de type fini. Si $n = 1$, cette assertion n'est autre que l'hypothèse *a*); nous allons donc raisonner par récurrence sur n, en supposant la propriété à établir vraie pour $n - 1$. Pour cela, définissons une application $f : L \to K$ en posant $f(\xi_1, \ldots, \xi_n) = \xi_n$; c'est un homomorphisme de L dans K. Pour montrer que L est de type fini, il suffit (Théorème 1) d'établir que Im (f) et Ker (f) sont de type fini. Or Im (f) est un sous-module de K, donc est de type fini d'après l'hypothèse *a*). Quant à Ker (f), c'est un sous-module du sous-module de K^n défini par la relation $\xi_n = 0$; ce sous-module de K^n étant évidemment isomorphe à K^{n-1}, tous ses sous-modules, et en particulier Ker (f), sont de type fini d'après l'hypothèse de récurrence.

Nous avons donc démontré, à l'aide de *a*), que tout sous-module de K^n est de type fini. Prenons maintenant un K-module à gauche M et un sous-module M' de M; on va montrer que, si M est de type fini, il en est de même de M'.

Puisque M est de type fini, il existe (§ 12, Corollaire 2 du Théorème 3) un entier n et un homomorphisme f de K^n *sur* M. Considérons $L = f^{-1}(M')$; comme f est surjectif, f induit un homomorphisme de L *sur* M'; d'autre part L est de type fini d'après ce qu'on a déjà établi; donc M' lui-même est de type fini (de façon précise, si L est engendré par des vecteurs a_i, $1 \leqslant i \leqslant p$, il est clair que M' est engendré par les vecteurs $f(a_i)$, $1 \leqslant i \leqslant p$). Ceci achève la démonstration.

3. *Sous-modules d'un module libre sur un anneau principal*

La méthode utilisée pour démontrer le Théorème 2 conduit aussi au résultat suivant :

THÉORÈME 3. *Soit* K *un anneau. Les propriétés suivantes sont équivalentes :*

a) Pour tout idéal à gauche non nul I *de* K, *il existe un* $a \in I$ *tel que l'application* $x \to xa$ *de* K *dans* I *soit bijective;*

b) Si un K*-module à gauche* M *admet une base composée de n vecteurs, et si* M′ *est un sous-module de* M, *il existe un entier* $p \leqslant n$ *tel que* M′ *admette une base composée de p vecteurs.*

Pour montrer que *b*) implique *a*), on prend M = K; alors M admet une base comprenant *un* vecteur; donc tout sous-module de M (i.e. tout idéal à gauche de K) doit admettre une *base* formée de *zéro* ou de *un* vecteur, ce qui est précisément la propriété *a*) de l'énoncé.

Montrons maintenant que *a*) implique *b*). Il est clair que *b*) est équivalent à l'assertion suivante :

b′) Pour tout entier $n \geqslant 1$ *et tout sous-module* L *de* K^n, *il existe un entier* $p \leqslant n$ *tel que* L *soit isomorphe à* K^p.

Pour $n = 1$, il est clair que *b′*) se réduit à l'hypothèse *a*); on va donc raisonner par récurrence sur n, en supposant *b′*) établi pour $n - 1$. Soit donc L un sous-module de K^n, et considérons, comme au n° précédent, l'homomorphisme $f : L \to K$ donné par $f(\xi_1, \ldots, \xi_n) = \xi_n$; l'image de f est un sous-module de K, donc est isomorphe à K^r pour un entier $r \leqslant 1$; le noyau de f est isomorphe à un sous-module de K^{n-1}, donc, d'après l'hypothèse de récurrence, est isomorphe à K^s pour un entier $s \leqslant n - 1$. Faisant usage du Théorème 1, on en déduit que L est isomorphe à K^{r+s}, et comme

$$r + s \leqslant 1 + (n - 1) = n,$$

la démonstration est achevée.

Remarque 2. L'hypothèse *a*) est vérifiée lorsque K est un *anneau principal* i.e. un anneau intègre, commutatif, dont tous les idéaux sont principaux. En effet, soit I un idéal de K; on a $I = Ka$ pour un $a \in K$; l'application

$$x \to xa$$

de K dans I est donc surjective; de plus, son noyau est formé des x tels que $xa = 0$; mais si I n'est pas nul (auquel cas on a évidemment $a \neq 0$), ceci implique $x = 0$ puisque l'anneau K est intègre; donc l'application considérée est injective, ce qui montre bien que l'hypothèse *a*) est vérifiée.

Il va de soi qu'elle est aussi vérifiée lorsque K est un corps (commutatif ou non), car alors le seul idéal à gauche non nul de K est K, et il suffit de prendre $a = 1$ dans l'énoncé de l'hypothèse *a*).

Ces deux cas sont, dans la pratique, les seuls où l'on fasse usage du Théorème 3.

On notera que, l'anneau $\mathbf{Z}$ étant principal, le Théorème 3 implique le résultat suivant :

COROLLAIRE DU THÉORÈME 3. *Tout sous-groupe de* $\mathbf{Z}^n$ *est isomorphe à* $\mathbf{Z}^p$ *pour un entier* $p \leqslant n$.

Bien entendu on a résultat analogue lorsqu'on remplace $\mathbf{Z}$ par un corps; mais dans ce cas, on trouvera des résultats plus précis au § suivant.

4. *Applications aux systèmes d'équations linéaires*

Soient K un anneau et f un homomorphisme du K-module à droite K^n dans le K-module à droite K^p. Soit $(\alpha_{ij})_{1 \leqslant i \leqslant n,\ 1 \leqslant j \leqslant p}$ la matrice de f (par rapport aux bases canoniques de K^n et K^p); si $x = (\xi_1, \ldots, \xi_n)$ est un vecteur de K^n, son image

$$y = f(x) = (\eta_1, \ldots, \eta_p)$$

est donc donnée par les relations

$$\begin{aligned} \alpha_{11}\xi_1 + \cdots + \alpha_{n1}\xi_n &= \eta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ \alpha_{1p}\xi_1 + \cdots + \alpha_{np}\xi_n &= \eta_p. \end{aligned}$$

Il s'ensuit que le noyau de f est le sous-module de K^n formé des vecteurs qui vérifient les relations

$$(1) \qquad \begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{n1}\xi_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1p}\xi_1 + \cdots + \alpha_{np}\xi_n = 0. \end{cases}$$

On dit que (1) est un **système de p équations linéaires et homogènes à n inconnues** à coefficients dans K.

Soit L le sous-module de K^n formé par les solutions de (1). Nous allons interpréter en langage d'équations linéaires les propriétés suivantes de L : (i) L est de type fini. (ii) Il existe un entier $r \leqslant n$ tel que L soit isomorphe à K^r. La première est vraie pourvu que K soit noethérien à droite (Théorème 2), et la seconde l'est si K est un corps ou bien un anneau principal (Théorème 3).

Supposons L de type fini; alors il existe un nombre fini de vecteurs (*)

$$(2) \qquad \begin{cases} x^1 = (\xi_1^1, \ldots, \xi_n^1), \\ x^2 = (\xi_1^2, \ldots, \xi_n^2), \\ \cdots\cdots\cdots\cdots \\ x^r = (\xi_1^r, \ldots, \xi_n^r) \end{cases}$$

dans L, tels que les éléments de L soient les vecteurs

$$x = (\xi_1, \ldots, \xi_n)$$

(*) Les indices supérieurs figurant dans les formules ci-dessous ne sont naturellement pas des exposants.

qui peuvent se mettre sous la forme

$$x = x^1\tau_1 + \cdots + x^r\tau_r \tag{3}$$

où $\tau_1, \ldots, \tau_r$ sont des scalaires arbitraires. Tenant compte de (2), la relation (3) équivaut évidemment au système de relations que voici :

$$\begin{cases} \xi_1 = \xi_1^1\tau_1 + \cdots + \xi_1^r\tau_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \xi_n = \xi_n^1\tau_1 + \cdots + \xi_n^r\tau_r. \end{cases} \tag{4}$$

Ainsi, *lorsque l'anneau* K *est noethérien à droite, il existe des constantes* $\xi_i^j \in K$ *telles que les solutions du système* (1) *soient les familles de scalaires* $\xi_1, \ldots, \xi_n$ *qui peuvent s'écrire sous la forme* (4), *avec des scalaires arbitraires* $\tau_j \in K$. On exprime ce résultat en disant que les solutions du système (1) *dépendent d'un nombre fini de paramètres arbitraires* (à savoir les variables $\tau_1, \ldots, \tau_r$ figurant dans les formules (4)).

Lorsque K est un corps, ou un anneau principal, on peut même supposer que les vecteurs (2) forment une *base* de L; alors chaque $x \in L$ s'écrit d'une seule façon sous la forme (3). Autrement dit, *si* K *est un corps ou un anneau principal, il existe des constantes*

$$\xi_i^j \in K \quad (1 \leqslant i \leqslant n, \quad 1 \leqslant j \leqslant r \leqslant n)$$

telles que l'application

$$(\tau_1, \ldots, \tau_r) \mapsto (\xi_1, \ldots, \xi_n)$$

donnée par les formules (4) *soit une bijection de* K^n *sur l'ensemble des solutions de* (1).

Lorsque K est un corps, on donnera plus loin des résultats beaucoup plus complets et explicites.

5. *Autres caractérisations des anneaux noethériens*

Dans ce n° nous allons donner quelques propriétés caractéristiques des anneaux noethériens. Introduisons tout d'abord la terminologie suivante.

Soit (A_n) une suite de parties d'un ensemble X; on dit que c'est une **suite croissante** si l'on a

$$A_n \subset A_{n+1} \quad \text{pour tout } n,$$

et que c'est une **suite stationnaire** s'il existe un entier p tel que

$$A_n = A_{n+1} \quad \text{pour tout} \quad n \geqslant p,$$

de sorte qu'on a alors $A_p = A_{p+1} = A_{p+2} = \cdots$.

D'autre part, considérons un ensemble F de parties de X; on dit qu'un $A \in F$ est un **élément maximal** de F si les relations

$$A \subset B \quad \text{et} \quad B \in F \quad \text{impliquent} \quad A = B,$$

autrement dit si F ne contient aucun ensemble strictement plus grand que A (ce qui *ne veut pas dire* que tout $B \in F$ soit contenu dans A...).

THÉORÈME 4. *Soit* K *un anneau. Les propriétés suivantes sont équivalentes :*

(AN 1) K *est noethérien à gauche* (i.e. tout idéal à gauche de K est de type fini);

(AN 2) *toute suite croissante d'idéaux à gauche de* K *est stationnaire;*

(AN 3) *tout ensemble non vide d'idéaux à gauche de* K *possède au moins un élément maximal.*

Montrons que (AN 1) implique (AN 2). Soit (I_n) une suite croissante d'idéaux à gauche de K; la réunion I des I_n est encore un idéal à gauche (§ 10, Théorème 1). Comme K est noethérien à gauche, I est engendré par des éléments $x_1, \ldots, x_r$ en nombre fini; et, par définition d'une réunion, il existe des entiers $p_1, \ldots, p_r$ tels que l'on ait $x_1 \in I_{p_1}, \ldots, x_r \in I_{p_r}$. Posons $p = \text{Max}(p_1, \ldots, p_r)$; alors I_p contient $x_1, \ldots, x_r$ puisqu'il contient $I_{p_1}, \ldots, I_{p_r}$; donc I_p contient l'idéal engendré par $x_1, \ldots, x_r$, autrement dit I; pour $n \geqslant p$ on a alors $I_p \subset I_n \subset I \subset I_p$ donc $I_p = I_n$, ce qui établit (AN 2).

Montrons maintenant que (AN 2) implique (AN 3). Soit F un ensemble non vide d'idéaux à gauche de K. Si F n'admettait aucun élément maximal, alors pour tout $I \in F$ on pourrait trouver un $J \in F$ contenant strictement I. Choisissant un $I_1 \in F$, il existerait un $I_2 \in F$ contenant *strictement* I_1, puis un $I_3 \in F$ contenant *strictement* I_2, et ainsi de suite; de cette façon il est clair qu'on obtiendrait une suite strictement croissante d'idéaux à gauche de K, contrairement à (AN 2).

Il reste à montrer que (AN 3) implique (AN 1). Soit I un idéal à gauche de K; soit F l'ensemble des idéaux à gauche de K qui sont *de type fini et contenus dans* I. L'ensemble F est non vide (il contient par exemple l'idéal 0 de K). D'après (AN 3), l'ensemble F admet donc un élément maximal J. Soit $(x_1, \ldots, x_n)$ un système de générateurs de J. Pour tout $x \in I$, l'idéal à gauche engendré par $x_1, \ldots, x_n$ et x est de type fini et contenu dans I — donc appartient à F — et contient J; comme J est un élément maximal de F, cet idéal ne peut être que J lui-même; on a donc $x \in J$ pour tout $x \in I$, d'où $I = J$, ce qui prouve que I est de type fini et achève la démonstration du Théorème.

Remarque 3. On dit qu'un idéal à gauche (resp. à droite, bilatère) I d'un anneau K est **maximal** si $I \neq K$ et si les seuls idéaux à gauche (resp. à droite, bilatères) de K contenant I sont I et K. On peut démontrer, à l'aide de raisonnements assez compliqués de théorie des ensembles, que si K est un anneau, *tout idéal à gauche (resp. à droite, bilatère)* I *de* K, *distinct de* K, *est contenu dans au moins un idéal à gauche (resp. à droite, bilatère) maximal de* K (Théorème de Krull).

Ce résultat, qui joue maintenant un rôle fondamental en Algèbre, peut se démontrer élémentairement si l'anneau K est noethérien : il suffit pour cela d'appliquer l'assertion (AN 3) du Théorème 4 à l'ensemble des idéaux à gauche (resp. à droite, bilatères) J de K tels que

$$I \subset J, \qquad J \neq K.$$

§ 19. La notion de dimension

1. Existence de bases

Le résultat suivant a déjà été annoncé au § 11, n° 4 :

THÉORÈME 1. *Tout espace vectoriel de dimension finie* (*) *sur un corps admet une base.*

Ce théorème est évidemment une conséquence du résultat plus précis que voici :

THÉORÈME 2. *Soient* M *un espace vectoriel de dimension finie sur un corps,* X *un ensemble fini de générateurs de* M, *et* A *une partie de* X. *Supposons les éléments de* A *linéairement indépendants. Alors il existe une base* B *de* M *telle que l'on ait.*

$$A \subset B \subset X.$$

Considérons en effet toutes les parties de X qui sont *libres* et contiennent A; il en existe, ne serait-ce que A elle-même. Parmi ces parties, considérons celles qui comportent *le plus grand nombre possible d'éléments;* soit B l'une de celles-ci. Pour établir le Théorème 2 il suffit de montrer que B est une base de M, ou même que B engendre M puisque B est libre par construction.

Comme X engendre M, il suffit pour cela d'établir que tout $x \in X$ est combinaison linéaire d'éléments de B. Comme c'est clair si $x \in B$, nous supposerons, pour le démontrer, que $x \notin B$.

L'ensemble $B' = B \cup \{x\}$ est contenu dans X et contient alors strictement plus d'éléments que B; comme on a en outre $A \subset B'$, on voit que B' ne pas peut être libre. Si donc l'on désigne par $x_1, \ldots, x_r$ les divers éléments de B, il existe une relation

$$\lambda_1 x_1 + \cdots + \lambda_r x_r + \lambda x = 0 \tag{1}$$

avec des scalaires $\lambda_1, \ldots, \lambda_r, \lambda$ non tous nuls. On a même $\lambda \neq 0$, car dans le cas contraire (1) se réduirait à une relation linéaire *non triviale* entre les éléments $x_i \in B$, contrairement au fait que B est libre.

(*) Les résultats du présent n° sont en fait valables pour tous les espaces vectoriels même de dimension infinie, mais les démonstrations générales sont trop compliquées pour être reproduites ici.

Puisque λ n'est pas nul, et puisque l'anneau de base K est un *corps*, λ est inversible dans K, et en multipliant le premier membre de (1) par l'inverse de λ on trouve

$$x = -\lambda^{-1}\lambda_1 x_1 - \cdots - \lambda^{-1}\lambda_r x_r;$$

nous avons donc démontré que tout $x \in X$ est combinaison linéaire d'éléments de B, d'où le Théorème 2.

Les théorèmes 1 et 2 (dont la démonstration aurait pu être donnée dès le § 11) ont des conséquences importantes; outre les deux Corollaires ci-dessous, on en trouvera quelques-unes dans les n° 2 et 3.

COROLLAIRE 1. *Soit* M *un espace vectoriel de dimension finie sur un corps* K. *Pour que des éléments donnés de* M *fassent partie d'une base de* M, *il faut et il suffit qu'ils soient linéairement indépendants.*

S'il existe une base de M contenant des vecteurs donnés $x_1, \ldots, x_p$, il est clair que ceux-ci sont linéairement indépendants. Inversement, si cette condition est satisfaite, l'existence d'une base de M contenant les vecteurs donnés s'obtient en choisissant un système fini G de générateurs de M, et en appliquant le Théorème 2 aux ensembles

$$X = G \cup \{x_1, \ldots, x_p\}, \qquad A = \{x_1, \ldots, x_p\}.$$

COROLLAIRE 2. *Soit* M *un espace vectoriel de dimension finie sur un corps. Tout sous-espace vectoriel de* M *est facteur direct dans* M.

Soit en effet M' un sous-espace de M. D'après le § 18, Théorème 2, M' est de dimension finie, donc (Théorème 1) admet une base $(x_i)_{1 \leqslant i \leqslant p}$; appliquant le Corollaire 1 à ces vecteurs et à M on voit qu'il existe des vecteurs $x_{p+1}, \ldots, x_r$ tels que les x_i $(1 \leqslant i \leqslant r)$ forment une *base* de M. Il est alors clair que le sous-espace de M engendré par $x_{p+1}, \ldots, x_r$ est un supplémentaire de M' dans M.

Remarque 1. La démonstration du Corollaire 2 ci-dessus utilise le fait qu'un sous-espace d'un espace vectoriel M de dimension finie sur un corps K est lui-même de dimension finie sur K, ce qui résulte en effet de ce qu'un corps est un anneau noethérien (on utilise alors le Théorème 2 du § 18) ou principal (on raisonne alors à l'aide du Théorème 3 du § 18). Mais il est évidemment possible de donner de ce fait une démonstration directe et élémentaire (qui ne diffère d'ailleurs pas sensiblement de celles des Théorèmes 2 et 3 du § 18); on procède comme suit.

Comme on sait déjà (Théorème 1) que M admet une base, on peut supposer $M = K^n$. Si $n = 1$, les seuls sous-espaces vectoriels de M sont 0 et M puisque K est un corps, et sont donc de dimension finie. Dans le cas général, soit M' un sous-espace vectoriel de K^n, et identifions K^{n-1} au sous-espace de K^n formé des vecteurs $(\xi_1, \ldots, \xi_n)$ tels que $\xi_n = 0$. Si $M' \subset K^{n-1}$ on peut supposer M' de dimension finie (on raisonne par récurrence sur n). Si M' n'est pas

contenu dans K^{n-1}, choisissons dans M′ un vecteur

$$a = (\alpha_1, \ldots, \alpha_n) \quad \text{tel que} \quad \alpha_n \neq 0$$

et posons $M'' = M' \cap K^{n-1}$; pour tout élément

$$x = (\xi_1, \ldots, \xi_n)$$

de M′, on peut écrire, puisque α_n est inversible dans K,

$$x = \xi_n \alpha_n^{-1} a + y$$

avec un vecteur y dont la dernière composante est nulle, donc appartenant à K^{n-1} et évidemment à M′, donc à M″; ainsi, on voit aussitôt que M′ est somme directe du sous-espace Ka engendré par a et de M″; mais celui-ci, étant contenu dans K^{n-1}, est de dimension finie d'après l'hypothèse de récurrence; il en est donc de même de M′, ce qui achève la démonstration.

On trouvera dans la Bibliographie beaucoup d'autres méthodes de démonstration. Par exemple, on peut établir d'abord les Théorèmes 2 et 6, puis les Théorèmes 10, 11 et 12; le résultat établi dans cette *Remarque* se déduit alors aussitôt du Théorème 12.

2. *Définition d'un sous-espace vectoriel par des équations linéaires*

Soient L un espace vectoriel de dimension finie sur un corps K, et M un sous-espace vectoriel de L. Dans le dual L* de L, considérons l'ensemble, noté

$$M^0,$$

des formes linéaires f sur L qui vérifient

$$f(x) = 0 \quad \text{pour tout} \quad x \in M.$$

Cet ensemble est un sous-espace vectoriel de L*, car s'il contient deux formes f et g, et si l'on pose $h = \alpha f + \beta g$ où α et β sont des scalaires arbitraires, on a

$$h(x) = \alpha . f(x) + \beta . g(x) = 0 \quad \text{pour tout} \quad x \in M,$$

donc $h \in M^0$, ce qui prouve notre assertion.

On dit que M^0 est l'**orthogonal de M dans L***. Le résultat suivant montre que la connaissance du sous-espace M^0 permet de reconstituer le sous-espace M :

THÉORÈME 3. *Soient* L *un espace vectoriel de dimension finie,* M *un sous-espace de* L, *et* M^0 *l'orthogonal de* M *dans* L*. *Pour qu'un* $x \in$ L *soit dans* M, *il faut et il suffit que l'on ait*

$$f(x) = 0 \tag{2}$$

pour toute forme linéaire $f \in M^0$.

La condition étant trivialement nécessaire, nous allons montrer qu'elle est suffisante. Comme on l'a vu en démontrant le Corollaire 2 du Théorème 2, il existe une base $(a_i)_{1 \leqslant i \leqslant r}$ de L et un entier $p \leqslant r$ tels que M soit engendré par $a_1, \ldots, a_p$. Soient $f_1, \ldots, f_r$ les fonctions coordonnées par rapport à la base $a_1, \ldots, a_r$ de L; il est clair que M est défini par les relations

$$f_{p+1}(x) = \cdots = f_r(x) = 0, \tag{3}$$

autrement dit que M^0 contient les formes $f_{p+1}, \ldots, f_r$ et que les relations (3) *caractérisent* les éléments de M. Comme les relations (3) sont évidemment vérifiées si la condition (2) est remplie pour tout $f \in M^0$, le Théorème est démontré.

Remarque 2. En fait les formes $f_{p+1}, \ldots, f_r$ construites ci-dessus forment une *base* de M^0. Il est en effet clair qu'elles sont linéairement indépendantes (du reste les formes $f_1, \ldots, f_r$ constituent la base de L^* duale de la base $a_1, \ldots, a_r$ de L); il suffit donc d'établir qu'elles engendrent M^0. Or considérons une forme $f \in L^*$; posant $f(a_i) = \alpha_i$, on a

$$f(x) = f(a_1\xi_1 + \cdots + a_r\xi_r) = \alpha_1\xi_1 + \cdots + \alpha_r\xi_r$$

pour $x = a_1\xi_1 + \cdots + a_r\xi_r$, et comme $\xi_i = f_i(x)$ il vient donc

$$f = \alpha_1 f_1 + \cdots + \alpha_r f_r \qquad \text{avec} \qquad \alpha_i = f(a_i);$$

si $f \in M^0$ on a en particulier $f(a_1) = \cdots = f(a_p) = 0$, donc

$$f = \alpha_{p+1} f_{p+1} + \cdots + \alpha_r f_r,$$

ce qui montre bien que $f_{p+1}, \ldots, f_r$ engendrent M^0.

On notera inversement que, si des formes f_i $(1 \leqslant i \leqslant m)$ engendrent M^0, alors les éléments de M sont caractérisés par les relations

$$f_1(x) = \ldots = f_m(x) = 0.$$

En effet toute $f \in M^0$ peut s'écrire sous la forme $f = \alpha_1 f_1 + \cdots + \alpha_m f_m$ avec des coefficients $\alpha_i \in K$, et il est clair que les équations considérées impliquent alors $f(x) = 0$, donc $x \in M$ d'après le Théorème 3.

Le principal intérêt du Théorème 3 réside du reste dans la propriété que nous venons d'établir, et qui mérite un énoncé explicite :

COROLLAIRE 1. *Soit $(f_i)_{1 \leqslant i \leqslant m}$ un système de générateurs de M^0. Alors les éléments de M sont caractérisés par les équations*

$$f_1(x) = \ldots = f_m(x) = 0.$$

Voici une autre conséquence importante du Théorème 3 :

COROLLAIRE 2. *Soient L un espace vectoriel de dimension finie sur un corps et M un sous-*

espace vectoriel de L. *Pour que l'on ait* $M \neq L$, *il faut et il suffit qu'il existe sur* L *une forme linéaire* f *non nulle vérifiant*

$$f(x) = 0 \quad \text{pour tout} \quad x \in M.$$

Si $M = L$ il est clair qu'on a $M^0 = \{0\}$. Inversement, si $M^0 = \{0\}$, le Théorème 3 montre que $M = L$. Les relations $M = L$ et $M^0 = \{0\}$ sont donc équivalentes; par suite, la relation $M \neq L$ équivaut à la relation $M^0 \neq \{0\}$, ce qui établit le Corollaire.

3. *Conditions de compatibilité d'un système d'équations linéaires*

La théorie des systèmes d'équations linéaires sera étudiée en détail au § 20; mais nous pouvons dès maintenant résoudre la question de savoir à quelles conditions un système d'équations linéaires possède des solutions. Ce sera l'objet des deux Théorèmes établis dans ce n°.

THÉORÈME 4. *Soient* $f_1, \ldots, f_r$ *des formes linéaires sur un espace vectoriel* M *sur un corps* K. *Les propriétés suivantes sont équivalentes :*

a) Les formes $f_1, \ldots, f_r$ *sont linéairement indépendantes.*

b) Quels que soient les scalaires $\beta_1, \ldots, \beta_r \in K$, *il existe au moins un* $x \in M$ *qui vérifie les relations*

$$\begin{cases} f_1(x) = \beta_1 \\ \cdots\cdots\cdots \\ f_r(x) = \beta_r. \end{cases} \tag{4}$$

Considérons en effet l'homomorphisme $f : M \to K^r$ donné par

$$f(x) = (f_1(x), \ldots, f_r(x));$$

la propriété *b*) exprime que f est *surjectif*. Or $f(M)$ est un sous-espace de K^r; en vertu du Corollaire 2 du Théorème 3, tout revient donc à exprimer que, si une forme linéaire u sur K^r est nulle sur le sous-espace $f(M)$, on a $u = 0$.

Mais soit

$$u(\xi_1, \ldots, \xi_r) = \lambda_1\xi_1 + \cdots + \lambda_r\xi_r$$

une telle forme; on a

$$u(f(x)) = \lambda_1 f_1(x) + \cdots + \lambda_r f_r(x)$$

pour tout $x \in M$; dire que u est nulle sur $f(M)$ signifie donc que les formes f_i satisfont à la relation linéaire

$$\lambda_1 f_1 + \cdots + \lambda_r f_r = 0.$$

Ainsi, la propriété *a*) de l'énoncé signifie bien que la seule forme u qui soit nulle sur le sous-espace $f(M)$ est $u = 0$, et le Théorème 4 est démontré.

Considérons maintenant un système d'équations

$$\text{(5)} \qquad \begin{cases} f_1(x) = \beta_1 \\ \cdots\cdots\cdots \\ f_n(x) = \beta_n \end{cases}$$

où les formes linéaires $f_1, \ldots, f_n$ sur l'espace vectoriel M ne sont plus nécessairement linéairement indépendantes. Dans le dual M* de M, ces formes engendrent un sous-espace vectoriel F de dimension finie, et d'après le Théorème 2 on peut extraire de la famille $(f_i)_{1 \leqslant i \leqslant n}$ une base de F; en modifiant au besoin la numérotation des f_i, on peut donc supposer que les formes $f_1, \ldots, f_r$ forment une base de F. Il est clair qu'on est alors dans les conditions d'application du résultat suivant :

THÉORÈME 5. *Soient $f_1, \ldots, f_n$ des formes linéaires sur un espace vectoriel* M. *Supposons que $f_1, \ldots, f_r$ soient linéairement indépendantes, et qu'on ait des relations de la forme*

$$\text{(6)} \qquad f_j = \rho_{j1} f_1 + \cdots + \rho_{jr} f_r \quad \textit{pour} \quad r + 1 \leqslant j \leqslant n,$$

avec des constantes $\rho_{jk} \in \mathrm{K}$. Alors, pour que le système d'équations (5) *possède au moins une solution $x \in \mathrm{M}$, il faut et il suffit que l'on ait*

$$\text{(7)} \qquad \beta_j = \rho_{j1}\beta_1 + \cdots + \rho_{jr}\beta_r \quad \textit{pour} \quad r + 1 \leqslant j \leqslant n;$$

le système (5) *possède alors les mêmes solutions que le système*

$$\text{(8)} \qquad \begin{cases} f_1(x) = \beta_1 \\ \cdots\cdots\cdots \\ f_r(x) = \beta_r \end{cases}$$

Les relations (6) signifient qu'on a

$$\text{(7 \textit{bis})} \qquad f_j(x) = \rho_{j1} f_1(x) + \cdots + \rho_{jr} f_r(x)$$

pour *tout* $x \in \mathrm{M}$; il est donc clair que, s'il existe un x vérifiant les relations (5), les relations (7) seront nécessairement vérifiées.

Inversement, supposons les relations (7) vérifiées et considérons une solution de (8); on a alors, d'après (7 *bis*),

$$f_j(x) = \rho_{j1}\beta_1 + \cdots + \rho_{jr}\beta_r,$$

et en tenant compte de (7) on voit donc que les systèmes (5) et (8) ont les mêmes solutions. Comme (8) possède effectivement des solutions en vertu du Théorème précédent, le Théorème 5 est donc démontré.

Remarque 3. Les relations (7), qui sont nécessaires et suffisantes pour que le système d'équations linéaires (6) admette au moins une solution, sont appelées les **conditions de compatibilité** du système (6). Le Théorème 5 montre que, pour trouver les solutions d'un système d'équations linéaires, on peut toujours se

ramener (si les conditions de compatibilité sont vérifiées) à un système d'équations dont les premiers membres sont des équations *linéairement indépendantes*. Ce résultat doit naturellement être complété par le Théorème 4 — à savoir qu'un système d'équations linéaires linéairement indépendantes possède *toujours* des solutions.

On notera également que nous n'avons pas supposé M de dimension finie dans les énoncés précédents; cette hypothèse est superflue dans le contexte actuel, et ne sera utilisée que pour établir les résultats plus précis du § 20.

4. *Existence de relations linéaires*

Nous allons maintenant établir l'un des résultats fondamentaux de l'Algèbre linéaire.

THÉORÈME 6. *Soit* M *un espace vectoriel de dimension finie sur un corps* K. *Toutes les bases de* M *ont le même nombre d'éléments.*

Soient $(a_i)_{1 \leqslant i \leqslant p}$ et $(b_j)_{1 \leqslant j \leqslant q}$ deux bases de M; pour montrer que $p = q$, il suffit de prouver qu'on a $p \leqslant q$ et $q \leqslant p$; par raison de symétrie, il suffit même d'établir que $q \leqslant p$. Cela va évidemment résulter de l'énoncé suivant :

THÉORÈME 7. *Soient* M *un espace vectoriel de dimension finie sur un corps* K *et* p *un entier tel que* K *possède une base formée de* p *vecteurs. Pour que* q *éléments de* M *soient linéairement indépendants il faut que* $q \leqslant p$.

Il va de soi que la réciproque du Théorème 7 est fausse.

Soient $(a_i)_{1 \leqslant i \leqslant p}$ une base de M et $b_1, \ldots, b_q$ des éléments de M; nous allons montrer qu'il existe une relation linéaire non triviale entre les vecteurs b_j dès que l'on a $q > p$, ce qui établira le Théorème 7.

L'assertion à établir est triviale si $p = 0$: on a en effet alors $M = \{0\}$, et comme $q > 1$ il est clair que les b_j vérifient par exemple la relation

$$1 . b_1 + 0 . b_2 + \cdots + 0 . b_q = 0.$$

Nous supposerons maintenant le Théorème établi pour $p - 1$, et allons en déduire qu'il est vrai pour p. Soit M' le sous-espace de M engendré par les vecteurs $a_1, \ldots, a_{p-1}$; puisque $a_1, \ldots, a_p$ engendrent M, on a des relations

$$b_j = b'_j + \alpha_j a_p \quad (1 \leqslant j \leqslant q) \tag{9}$$

avec des vecteurs $b'_j \in M'$ et des scalaires $\alpha_j \in K$. Si tous les α_j sont nuls, les b_j sont dans M'; mais comme M' admet une base formée de $p - 1$ vecteurs, le Théorème s'applique à M', et comme la relation $q > p$ implique *a fortiori* $q > p - 1$, on voit que l'existence d'une relation linéaire non triviale entre les b_j s'obtient immédiatement dans ce cas.

Reste donc à examiner le cas où les α_j ne sont pas tous nuls; supposons par exemple

$$\alpha_q \neq 0,$$

donc α_q inversible puisque K est un *corps*. Alors la dernière relation (9) donne

$$a_p = \alpha_q^{-1}(b_q - b'_q)$$

et en portant dans les autres relations (9) on obtient, après un calcul facile, les relations

$$b_j - \nu_j b_q = b'_j - \nu_j b'_q \quad (1 \leqslant j \leqslant q-1) \tag{10}$$

avec $\nu_j = \alpha_j \alpha_q^{-1}$.

Les relations (10) montrent que les $q-1$ vecteurs $b_j - \nu_j b_q$ sont dans le sous-espace M'; comme $q > p$ implique $q-1 > p-1$, l'hypothèse de récurrence montre donc l'existence d'une relation

$$\lambda_1(b_1 - \nu_1 b_q) + \cdots + \lambda_{q-1}(b_{q-1} - \nu_{q-1} b_q) = 0$$

avec des scalaires $\lambda_1, \ldots, \lambda_{q-1}$ *non tous nuls*; mais cette relation s'écrit aussi

$$\lambda_1 b_1 + \cdots + \lambda_q b_q = 0 \qquad \text{avec} \qquad \lambda_q = -(\lambda_1 \nu_1 + \cdots + \lambda_{q-1}\nu_{q-1}),$$

et l'existence d'une relation linéaire non triviale entre $b_1, \ldots, b_q$ est ainsi démontrée.

Remarque 4. Le Théorème 7 est en fait valable non seulement pour les espaces vectoriels sur un *corps* (commutatif ou non), mais aussi pour les modules sur un *anneau commutatif*; on peut le montrer à l'aide de la théorie des déterminants (§ 23, Corollaire du théorème 5).

COROLLAIRE. *Soit*

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{p1}\xi_p = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}\xi_1 + \cdots + \alpha_{pn}\xi_p = 0 \end{cases} \tag{11}$$

un système de n équations linéaires et homogènes à p inconnues $\xi_1, \ldots, \xi_p$ *à coefficients dans un corps* K. *Si l'on a* $p > n$ (i.e. si le nombre des inconnues dépasse celui des équations), *le système* (11) *admet au moins une solution non triviale* (i.e. pour laquelle les inconnues ξ_i ne sont pas toutes nulles).

Considérons en effet les p éléments

$$a_i = (\alpha_{i1}, \ldots, \alpha_{in})$$

de K^n; les relations (11) sont évidemment équivalentes à la relation

$$a_1\xi_1 + \cdots + a_p\xi_p = 0,$$

et pour que le système (11) possède une solution autre que $\xi_1 = \cdots = \xi_p = 0$ il est donc nécessaire et suffisant que les vecteurs a_i soient liés; c'est toujours le cas si $p > n$ en vertu du Théorème 7.

Remarque 5. L'explication intuitive du Corollaire précédent est la suivante :

comme le système (11) admet toujours la solution triviale

$$\xi_1 = \cdots = \xi_p = 0,$$

tout revient à trouver des hypothèses assurant que (11) admet *au moins deux* solutions, autrement dit assurant que les relations (11) ne déterminent pas entièrement les inconnues ξ_i; or, si l'on impose au moins p conditions à p inconnues ξ_i, on a de fortes chances de les déterminer entièrement, alors qu'au contraire il est peu probable qu'on arrive à les déterminer à l'aide de moins de p conditions...

5. *La notion de dimension*

Soit M un espace vectoriel de dimension finie sur K; il existe, d'après le Théorème 6, un entier n bien déterminé tel que toutes les bases de M possèdent n éléments. On dit que n est la **dimension de M sur le corps** K (ou la **dimension de** M tout court si aucune ambiguité n'est possible sur le corps de base), ou que M **est de dimension** n **sur** K, et la dimension de M sur K se désigne par la notation

$$\dim_K(M)$$

ou simplement par dim (M) s'il ne peut y avoir aucune confusion quant au corps de base choisi. On convient de prendre dim (M) = 0 lorsque $M = \{0\}$.

Remarque 6. Un espace vectoriel *complexe* M peut aussi être regardé comme un espace vectoriel *réel*; on a alors la relation

$$\dim_{\mathbf{R}}(M) = 2.\dim_{\mathbf{C}}(M)$$

(cf. *Exercice* 15), ce qui montre bien qu'il est indispensable de préciser le corps de base choisi.

Remarque 7. On peut aussi définir la notion de dimension pour les espaces vectoriels de dimension infinie, mais on doit, pour ce faire, utiliser la théorie des nombres cardinaux (§ 5); la méthode élémentaire consistant à poser dim (M) = $+\infty$ lorsque M n'est pas de dimension finie ne présente aucun intérêt, parce qu'avec cette définition trop simple de la notion de dimension aucun des énoncés qu'on a en vue (et en premier lieu une généralisation du Théorème 8 ci-dessous) n'est vrai; or l'intérêt d'une définition est avant tout, pour ne pas dire exclusivement, de conduire à des théorèmes (*).

Le principal intérêt de la notion de dimension est concentré dans l'énoncé suivant :

(*) C'est aussi parfois le cas dans la vie de tous les jours. Si l'on a par exemple pour but de démontrer que « l'armée poldève ne tire pas sur les Poldèves », il suffit d'introduire la définition suivante : en Poldévie, on appelle Poldèves les gens contre lesquels l'armée poldève ne tire pas. Ce procédé, malgré son efficacité certaine, présente un grave défaut : on peut en effet le retourner en une définition de l'armée poldève elle-même.

THÉORÈME 8. *Soient* L *et* M *deux espaces vectoriels de dimension finie sur un corps* K. *Pour que* L *et* M *soient isomorphes il faut et il suffit que* dim (L) = dim (M).

S'il existe un isomorphisme f de L sur M, celui-ci applique une base de L sur une base de M, de sorte que dim (L) = dim (M). Si inversement cette condition est réalisée, et si n est la dimension commune de L et M, alors L et M sont isomorphes à K^n en vertu du § 12, Corollaire 1 du théorème 3, donc isomorphes entre eux.

Remarque 8. *Il n'est pas vrai* que deux espaces vectoriels de dimension infinie soient toujours isomorphes.

Exemple 1. On a évidemment

$$\dim (K^n) = n$$

quel que soit n.

Exemple 2. Prenons K = **R** et pour M l'espace vectoriel formé des vecteurs d'origine donnée O dans l'espace usuel (§ 10, *Exemple* 2). On a alors dim (M) = 3. Si l'on prend pour M l'ensemble des vecteurs portés par un plan donné passant par O, on trouve dim (M) = 2. Si enfin on prend pour M l'ensemble des vecteurs portés par une droite donnée passant par O, on a dim (M) = 1.

Soient L et M des espaces vectoriels de dimension finie sur un corps K; on a alors

$$\dim (L \times M) = \dim (L) + \dim (M);$$

plus précisément : si $(a_i)_{1 \leqslant i \leqslant p}$ et $(b_j)_{1 \leqslant j \leqslant q}$ sont des bases de L et M respectivement, les $p + q$ vecteurs

$$(a_1, 0), \ldots, (a_p, 0), (0, b_1), \ldots, (0, b_q)$$

forment une base de L × M.

Exemple 3. L'**espace-temps** des physiciens est l'ensemble des couples (x, t) où x est un vecteur d'origine donnée O dans l'espace usuel, et t un nombre réel qu'on appelle le temps. C'est donc le produit cartésien M × **R** où M est l'espace usuel. Par suite, regardé comme espace vectoriel réel, l'espace-temps est de dimension 3 + 1 = 4. Le savant Cosinus a beaucoup travaillé pour pas grand'chose.

Soit L un espace vectoriel de dimension n; on a vu au § 16, Théorème 1, que son dual admet une base formée de n vecteurs; par suite, on a

$$\dim (L) = \dim (L^*).$$

THÉORÈME 9. *Soient* L *un espace vectoriel de dimension finie sur un corps,* M *un sous-espace de* L, *et* M° *l'orthogonal de* M *dans* L^*. *On a alors*

$$\dim (M) + \dim (M^\circ) = \dim (L).$$

Posons dim (M) $= p$ et dim (L) $= r$; comme on l'a vu en démontrant le Corollaire 2 du Théorème 2, il existe une base $(x_i)_{1 \leqslant i \leqslant r}$ de L telle que $(x_i)_{1 \leqslant i \leqslant p}$ soit une base de M. La *Remarque* 2 du n° 2 montre alors que le sous-espace M^0 de L^* admet une base de $r - p$ vecteurs, donc est de dimension $r - p$, ce qui achève la démonstration.

COROLLAIRE. *Soit* M *un sous-espace d'un espace vectoriel* L *de dimension finie sur un corps. On a alors*

$$\dim (M) \leqslant \dim (L),$$

et on ne peut avoir dim (M) = dim (L) *que si* M = L.

La première assertion résulte immédiatement du Théorème 9. Celui-ci montre aussi que la relation dim (L) = dim (M) équivaut à

$$\dim (M^0) = 0,$$

i.e. à $M^0 = \{0\}$; or on sait (Corollaire 2 du Théorème 3) que cette dernière relation équivaut à L = M.

6. *Caractérisations des bases et de la dimension*

L'énoncé suivant fournit plusieurs caractérisations utiles des bases d'un espace vectoriel de dimension finie :

THÉORÈME 10. *Soient* $x_1, \ldots, x_n$ *des éléments d'un espace vectoriel* M *sur un corps* K. *Les propriétés suivantes sont équivalentes :*

a) *Les vecteurs* $x_i (1 \leqslant i \leqslant n)$ *forment une base de* M.

b) *Les vecteurs* x_i *sont linéairement indépendants et* M *est de dimension* n.

c) *Les vecteurs* x_i *sont linéairement indépendants et toute partie libre de* M *comporte au plus* n *éléments.*

d) *Les vecteurs* x_i *engendrent* M *et* M *est de dimension* n.

e) *Les vecteurs* x_i *engendrent* M *et tout système de générateurs de* M *comporte au moins* n *éléments.*

Il est clair que *a*) implique *b*).

Pour montrer que *b*) implique *c*), on remarque qu'une famille libre fait partie d'une base de M (Corollaire 1 du Théorème 2), donc (Théorème 6) comporte au plus n éléments si M est de dimension n.

Pour montrer que *c*) implique *d*), considérons un $x \in M$; d'après *c*) les vecteurs $x_1, \ldots, x_n, x$, en nombre $n + 1$, sont liés par une relation *non triviale*

$$\lambda_1 x_1 + \cdots + \lambda_n x_n + \lambda x = 0;$$

on a du reste $\lambda \neq 0$, car dans le cas contraire on trouverait une relation *non triviale* entre les x_i, contrairement à *c*); donc λ est inversible et il vient

$$x = -\lambda^{-1}\lambda_1 x_1 - \cdots - \lambda^{-1}\lambda_n x_n,$$

ce qui prouve que les x_i engendrent M — et donc forment une base de M, qui est par suite de dimension n.

La propriété *d*) implique la propriété *e*), car tout système de générateurs de M contient une base de M (Théorème 2), donc comporte au moins n éléments si $\dim(M) = n$.

Pour achever la démonstration, il reste à montrer que *e*) implique *a*); or d'après le Théorème 2 on peut extraire de la famille $(x_i)_{1 \leqslant i \leqslant n}$ une base de M; celle-ci, étant un système de générateurs, comporte au moins n éléments d'après *e*); c'est donc nécessairement la famille $(x_i)_{1 \leqslant i \leqslant n}$ toute entière. Ainsi, celle-ci est une base de M, et le Théorème est démontré.

THÉORÈME 11. *Soient* M *un espace vectoriel sur un corps* K *et* n *un entier. Les propriétés suivantes sont équivalentes :*

a) M *est de dimension* n.

b) n *est le plus grand entier tel qu'on puisse extraire de* M *une famille de* n *vecteurs linéairement indépendants.*

c) n *est le plus petit entier tel qu'on puisse extraire de* M *une famille de* n *vecteurs engendrant* M.

L'équivalence des propriétés *a*) et *b*) provient de l'équivalence entre les propriétés *a*) et *c*) du Théorème 10. De même, l'équivalence entre les propriétés *a*) et *c*) du Théorème 11 provient de l'équivalence entre les propriétés *a*) et *e*) du Théorème 10.

THÉORÈME 12. *Pour qu'un espace vectoriel* M *sur un corps* K *soit de dimension finie, il faut et il suffit qu'il existe un entier* n *tel que toute partie libre de* M *possède au plus* n *éléments. On a alors* $\dim(M) \leqslant n$.

La condition est évidemment nécessaire d'après le résultat précédent. Inversement, si elle est vérifiée, on peut considérer le plus grand entier r tel que M contienne une famille de r vecteurs linéairement indépendants; on a $r \leqslant n$ par hypothèse, et M est de dimension r d'après le Théorème 11, *c*); donc M est de dimension finie et $\dim(M) \leqslant n$.

7. *Dimensions du noyau et de l'image d'un homomorphisme*

Lorsque l'anneau de base est un corps, le Théorème 1 du § 18 se traduit comme suit :

THÉORÈME 13. *Soient* L *et* M *deux espaces vectoriels sur un corps et* f *un homomorphisme de* L *dans* M. *Si* L *est de dimension finie, on a*

$$\dim(L) = \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)).$$

En effet, $\mathrm{Ker}(f)$ est de dimension finie, donc isomorphe à K^p où $p = \dim(\mathrm{Ker}(f))$, et $\mathrm{Im}(f)$ est aussi de dimension finie (l'image d'un module

de type fini par un homomorphisme est évidemment de type fini), donc isomorphe à K^q où $q = \dim(\mathrm{Im}(f))$; pour obtenir le Théorème 13 il reste donc à utiliser le Théorème 1 du § 18.

Remarque 9. Le Théorème 13, bien qu'étant un simple cas particulier du résultat beaucoup plus général établi au § 18, *ne peut pas* se démontrer plus simplement que celui-ci; il est par contre très facile d'en donner des démonstrations plus compliquées que celle du Théorème 1 du § 18...

Les Corollaires du Théorème 13 que nous allons maintenant énoncer sont au moins aussi importants dans la pratique que le Théorème 13 lui-même, tout particulièrement celui que voici :

COROLLAIRE 1. *Soient* L *et* M *deux espaces vectoriels de dimension finie sur un corps et* f *un homomorphisme de* L *dans* M. *On suppose* $\dim(L) = \dim(M)$. *Alors les propriétés suivantes sont équivalentes :*

a) f *est bijectif.*

b) f *est surjectif.*

c) f *est injectif.*

d) on a $\mathrm{Ker}(f) = \{0\}$.

On sait depuis longtemps (§ 7, Théorème 8) que les assertions *c)* et *d)* sont équivalentes. Comme d'autre part *a)* est la conjonction de *b)* et *c)*, tout revient à prouver l'équivalence de *b)* et *d)*. Or *b)* équivaut (Corollaire du Théorème 9) à la relation

$$\dim(\mathrm{Im}(f)) = \dim(M), \tag{12}$$

et *d)* équivaut à

$$\dim(\mathrm{Ker}(f)) = 0; \tag{13}$$

l'équivalence de (12) et (13) résulte donc du Théorème 13 et de l'hypothèse $\dim(L) = \dim(M)$, d'où le Corollaire.

On utilise le plus souvent le Corollaire 1 lorsque $L = M$. On verra au § suivant qu'il peut aussi se traduire dans le langage de la théorie des systèmes d'équations linéaires.

COROLLAIRE 2. *Soient* E *et* F *des sous-espaces vectoriels d'un espace vectoriel* M *de dimension finie sur un corps. On a alors*

$$\dim(E + F) = \dim(E) + \dim(F) - \dim(E \cap F).$$

Considérons en effet l'homomorphisme $f : E \times F \to M$ donné par $f(x, y) = x + y$ (cf. § 17, n° 3); on a $E + F = \mathrm{Im}(f)$, et d'autre part $\mathrm{Ker}(f)$ est isomorphe à $E \cap F$ (§ 17, *Remarque* 1); donc

$$\begin{aligned}\dim(E) + \dim(F) = \dim(E \times F) &= \dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) \\ &= \dim(E \cap F) + \dim(E + F),\end{aligned}$$

ce qui est la relation annoncée.

COROLLAIRE 3. *Soient* $E_1, \ldots, E_r$ *des sous-espaces vectoriels d'un espace vectoriel* E *de dimension finie. On a alors*

$$\dim (E_1 + \cdots + E_r) \leqslant \dim (E_1) + \cdots + \dim (E_r)$$

et pour que les deux membres de cette relation soient égaux, il faut et il suffit que les sous-espaces $E_1, \ldots, E_r$ *soient linéairement indépendants.*

Considérons en effet l'homomorphisme

$$f : E_1 \times \cdots \times E_r \to E_1 + \cdots + E_r$$

donné par

$$f(x_1, \ldots, x_r) = x_1 + \cdots + x_r;$$

il est surjectif, et par suite il résulte du Théorème 13 que

$$\begin{aligned} \dim (E_1 + \cdots + E_r) &= \dim (E_1 \times \cdots \times E_r) - \dim (\mathrm{Ker}\,(f)) \\ &= \dim (E_1) + \cdots + \dim (E_r) - \dim (\mathrm{Ker}\,(f)) \end{aligned}$$

l'inégalité cherchée résulte de là, et il y a égalité si et seulement si $\mathrm{Ker}\,(f) = 0$, ce qui signifie précisément (§ 17, n° 3) que les sous-espaces E_i de E sont linéairement indépendants.

8. Rang d'un homomorphisme, d'une famille de vecteurs, d'une matrice

Soient L et M deux espaces vectoriels de dimension finie et f un homomorphisme de L dans M. On appelle **rang de l'homomorphisme** f la dimension du sous-espace $\mathrm{Im}\,(f) = f(L)$ de M.

Soit $(a_i)_{1 \leqslant i \leqslant n}$ un système de générateurs de L; alors $\mathrm{Im}\,(f)$ est engendré par les vecteurs $x_i = f(a_i)$. On est ainsi conduit à introduire la notion suivante : étant donnée une famille $(x_i)_{1 \leqslant i \leqslant n}$ d'éléments d'un espace vectoriel M (de dimension finie ou non), on appelle **rang de la famille** (x_i) la dimension du sous-espace de M engendré par les x_i (on notera que, même si M est de dimension infinie, ce sous-espace est toujours de dimension finie).

Soit r le rang d'une famille $(x_i)_{1 \leqslant i \leqslant n}$ d'éléments d'un espace vectoriel M, et désignons par M' le sous-espace de M engendré par les x_i. D'après le Théorème 2, on peut extraire de la famille $(x_i)_{1 \leqslant i \leqslant n}$ une base de M', base comportant nécessairement r éléments. Au besoin en modifiant la numérotation des x_i, on peut donc supposer que $x_1, \ldots, x_r$ forment une base de M'. Il s'ensuit qu'alors les vecteurs $x_1, \ldots, x_r$ sont *linéairement indépendants* et que tout élément de M' est combinaison linéaire de ces vecteurs; en particulier, on a des relations

$$x_j = \rho_{j1} x_1 + \cdots + \rho_{jr} x_r \qquad \text{pour} \qquad r + 1 \leqslant j \leqslant n. \tag{14}$$

Inversement, si l'on a des relations (14), le sous-espace engendré par $x_1, \ldots, x_r$ contient non seulement ces vecteurs eux-mêmes mais aussi $x_{r+1}, \ldots, x_n$, et par suite $x_1, \ldots, x_r$ engendrent M' ; si de plus $x_1, \ldots, x_r$ sont linéairement indépendants,

ils forment une base de M', qui est donc de dimension r, et par suite la famille $(x_i)_{1 \leqslant i \leqslant n}$ est de rang r.

Ces considérations s'appliquent naturellement lorsque M est le dual d'un espace vectoriel L; les x_i sont alors des formes linéaires f_i sur L, ce qui permet de définir le **rang d'une famille de formes linéaires** sur L; c'est évidemment cette notion qui figure implicitement dans l'énoncé du Théorème 5.

Soit maintenant

$$A = (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n}$$

une matrice à p colonnes et n lignes, à coefficients dans un corps K. On appelle **rang de la matrice** A le rang de l'homomorphisme de K^p dans K^n représenté par A (§ 12, n° 3, *Remarque* 1). L'image de cet homomorphisme est le sous-espace vectoriel de K^n engendré par les images des éléments de la base canonique de K^p, i.e. par les colonnes de la matrice A. Donc, *le rang de la matrice* A *est la dimension du sous-espace vectoriel de* K^n *engendré par les p colonnes de* A.

THÉORÈME 14. *Soient* L *et* M *des espaces vectoriels de dimension finie, f un homomorphisme de* L *dans* M, *et* A *la matrice de f par rapport à une base* $(a_i)_{1 \leqslant i \leqslant p}$ *de* L *et à une base* $(b_j)_{1 \leqslant j \leqslant n}$ *de* M. *Alors le rang de f est égal au rang de* A.

Considérons en effet les isomorphismes

$$u : K^p \to L, \qquad v : K^n \to M$$

qui appliquent les bases canoniques de K^p et K^n sur les bases données de L et M; alors l'homomorphisme

$$v^{-1} \circ f \circ u : \ K^p \to K^n$$

a précisément A pour matrice par rapport aux bases canoniques de K^p et K^n (§ 12, n° 3, *Remarque* 2), et le rang de cet homomorphisme est donc égal à celui de A par définition. Mais comme u et v sont bijectifs, il est clair que les images de f et de $v^{-1} \circ f \circ u$ sont isomorphes, de sorte que ces deux homomorphismes ont même rang. D'où le Théorème.

THÉORÈME 15. *Soient* M *un espace vectoriel de dimension finie,* $(b_j)_{1 \leqslant j \leqslant n}$ *une base de* M, *et*

$$x_i = \alpha_{i1} b_1 + \cdots + \alpha_{in} b_n \quad (1 \leqslant i \leqslant p)$$

une famille d'éléments de M. *Alors le rang de la famille* $(x_i)_{1 \leqslant i \leqslant p}$ *est égal au rang de la matrice* $(\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n}$.

Il suffit pour le voir d'appliquer le Théorème précédent en prenant pour L l'espace K^p, pour base de L la base canonique, et pour f l'homomorphisme qui applique les éléments de la base canonique sur les vecteurs x_i donnés; ceux-ci engendrent le sous-espace Im (f), de sorte que le rang de la famille (x_i) est égal au rang de f, i.e. au rang de la matrice de f, laquelle est justement la matrice (α_{ij}).

9. Calcul effectif du rang d'une matrice

Les résultats du n° précédent montrent que le calcul du rang d'un homomorphisme ou d'un système de vecteurs peut se ramener à celui du rang d'une matrice. Le résultat fondamental à ce sujet est l'énoncé suivant :

THÉORÈME 16. *Soit* $A = (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n}$ *une matrice à coefficients dans un corps* K. *Le rang de* A *est le plus grand entier* r *tel que l'on puisse extraire de* A *une matrice carrée* inversible *d'ordre* r.

Soient L et M des espaces vectoriels de dimension p et n respectivement sur K; choisissons une base $(a_i)_{1 \leqslant i \leqslant p}$ de L et une base $(b_j)_{1 \leqslant j \leqslant n}$ de M, et considérons l'homomorphisme f de L dans M admettant A pour matrice par rapport aux bases considérées; le rang de A est égal à celui de f, i.e. à la dimension de Im (f). Soit r ce rang; nous allons montrer tout d'abord qu'on peut effectivement extraire de la matrice A une matrice carrée *inversible* d'ordre r.

En vertu du Théorème 13, on a dim (Ker (f)) $= p - r$, et par suite Ker (f) admet une base formée de $p - r$ vecteurs $c_1, \ldots, c_{p-r}$; comme les vecteurs

$$c_1, \ldots, c_{p-r}, a_1, \ldots, a_p$$

engendrent L, et comme les $p - r$ premiers d'entre eux sont linéairement indépendants, il est possible (Théorème 2) de former une base de L en adjoignant aux $p - r$ vecteurs $c_1, \ldots, c_{p-r}$ des vecteurs en nombre r, choisis parmi $a_1, \ldots, a_p$; en modifiant au besoin la numérotation des vecteurs $a_1, \ldots, a_p$ (ce qui revient à permuter entre elles les *colonnes* de la matrice A), on peut donc supposer que les vecteurs

$$c_1, \ldots, c_{p-r}, a_1, \ldots, a_r$$

forment une base de L. Désignant par L′ le sous-espace de L engendré par $a_1, \ldots, a_r$ on a donc

$$L = \text{Ker}\,(f) \oplus L'.$$

Si $x = y + z$ avec $y \in \text{Ker}\,(f)$, $z \in L'$, il est clair que $f(x) = f(z)$; donc $f(L') = \text{Im}(f)$ et comme Ker $(f) \cap L' = \{0\}$, la restriction de f à L′ est un *isomorphisme* de L′ sur le sous-espace Im (f) de M. Il en résulte que les vecteurs

$$f(a_1), \ldots, f(a_r)$$

forment une base de Im (f).

En utilisant à nouveau, comme ci-dessus, le Théorème 2, on voit donc qu'au besoin en modifiant la numérotation des vecteurs b_j (ce qui revient à permuter les *lignes* de la matrice A), on peut supposer que les vecteurs

$$f(a_1), \ldots, f(a_r), b_{r+1}, \ldots, b_n$$

forment une base de M. On a alors

$$M = \text{Im}\,(f) \oplus M'' \tag{15}$$

où M'' est le sous-espace engendré par $b_{r+1}, \ldots, b_n$. On a aussi du reste

$$M = M' \oplus M'' \tag{16}$$

en désignant par M' le sous-espace engendré par $b_1, \ldots, b_r$. Nous désignerons par u l'homomorphisme de M sur M' correspondant à la décomposition (16) — autrement dit, si

$$x = x' + x'' \qquad \text{avec} \qquad x' \in M', \quad x'' \in M'',$$

on pose $u(x) = x'$ (cf. § 17, n° 4).

On a évidemment $\mathrm{Ker}\,(u) = M''$, et comme la somme (15) est directe on a donc

$$\mathrm{Ker}(u) \cap \mathrm{Im}(f) = 0;$$

donc l'application de Im (f) dans M' induite par u est *injective*; mais comme

$$\dim\,(\mathrm{Im}\,(f)) = \dim\,(M') = r,$$

cette application est en fait *bijective* (Corollaire 1 du Théorème 13); et comme on a déjà montré que f induit une *bijection* de L' sur Im (f), il s'ensuit que l'application

$$f' : L' \to M'$$

donnée par

$$f'(x) = u(f(x)) \quad \text{pour tout} \quad x \in L'$$

est elle-même bijective, autrement dit est un *isomorphisme*, en sorte que la matrice de f' par rapport à la base $(a_1, \ldots, a_r)$ de L' et à la base $(b_1, \ldots, b_r)$ de M' est *inversible* (§ 15, n° 1). Or, si $A = (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n}$, on a

$$f(a_i) = \alpha_{i1} b_1 + \cdots + \alpha_{in} b_n$$

et par suite

$$u(f(a_i)) = \alpha_{i1} b_1 + \cdots + \alpha_{ir} b_r;$$

la matrice de f' est donc la matrice $(\alpha_{ij})_{1 \leqslant i,\ j \leqslant r}$ formée avec les r premières lignes et colonnes de A, et ceci démontre comme annoncé que, si A est de rang r, on peut extraire de A une matrice carrée inversible d'ordre r.

Remarque 10. En général ce n'est pas la matrice $(\alpha_{ij})_{1 \leqslant i,j \leqslant r}$ qui sera inversible; on n'oubliera pas en effet que, dans la démonstration précédente, on a dû admettre la possibilité de permuter les lignes, et les colonnes, de la matrice A.

Pour achever la démonstration du Théorème, il reste à faire voir que si A est de rang r, et si l'on peut extraire de A une matrice carrée inversible d'ordre s, alors on a nécessairement $s \leqslant r$.

Or supposons par exemple que la matrice $(\alpha_{ij})_{1 \leqslant i,j \leqslant s}$ soit inversible. Désignons maintenant par L' le sous-espace de L engendré par $a_1, \ldots, a_s$, par M' le

sous-espace de M engendré par $b_1, \ldots, b_s$, par M'' le sous-espace engendré par $b_{s+1}, \ldots, b_n$, et par u l'homomorphisme de M sur M' consistant à projeter parallèlement à M''. Alors $(\alpha_{ij})_{1 \leqslant i,\ j \leqslant s}$ est la matrice, par rapport aux bases $(a_i)_{1 \leqslant i \leqslant r}$ et $(b_j)_{1 \leqslant j \leqslant s}$, de l'homomorphisme f' de L' dans M' donné par $f'(x) = u(f(x))$ pour $x \in L'$. Cette matrice étant inversible par hypothèse, f' est bijectif, donc injectif, et la restriction de f à L' est *à fortiori* injective; donc f applique L' sur un sous-espace de même dimension s que L', ce qui montre que Im (f) est de dimension s au moins; autrement dit, on a $s \leqslant r$, ce qui termine la démonstration.

Remarque 11. Si le corps K est *commutatif*, on dispose en outre d'un critère théoriquement très simple pour décider si une matrice carrée est inversible ou non : il suffit d'examiner son déterminant (§ 23, Corollaire 1 du Théorème 8).

Exemple 4. Considérons dans K^3 deux vecteurs

$$x' = (a', b', c'), \qquad x'' = (a'', b'', c'');$$

dire qu'ils sont linéairement indépendants signifie que le rang de la matrice

$$\begin{pmatrix} a' & a'' \\ b' & b'' \\ c' & c'' \end{pmatrix}$$

est égal à 2, autrement dit que l'une au moins des matrices

$$\begin{pmatrix} b' & b'' \\ c' & c'' \end{pmatrix}, \qquad \begin{pmatrix} a' & a'' \\ c' & c'' \end{pmatrix}, \qquad \begin{pmatrix} a' & a'' \\ b' & b'' \end{pmatrix}$$

est inversible. Si K est commutatif, cela signifie que les scalaires

$$b'c'' - b''c', \qquad a'c'' - a''c', \qquad a'b'' - a''b'$$

ne sont pas tous nuls (§ 15, n° 3).

COROLLAIRE. *Le rang d'une matrice* A *à coefficients dans un corps est égal au rang de la matrice transposée* ${}^tA$.

Il est en effet clair que les matrices carrées extraites de ${}^tA$ sont les transposées des matrices carrées extraites de A; or, si une matrice carrée est inversible, il en est de même de sa transposée, et réciproquement.

10. Calcul de la dimension d'un sous-espace vectoriel à partir de ses équations

Soient L un espace vectoriel de dimension finie n sur un corps K et M un sous-espace vectoriel de L; on a vu plus haut (Corollaire 1 du Théorème 3) qu'il existe des formes linéaires f_i $(1 \leqslant i \leqslant m)$ sur L telles que M soit l'ensemble des $x \in L$

satisfaisant aux relations

$$f_1(x) = \cdots = f_m(x) = 0. \tag{17}$$

Il est clair inversement que, quelles que soient les formes linéaires f_i, les relations (17) définissent un sous-espace vectoriel M de L.

Ceci dit, le calcul de la dimension de M s'effectue comme suit :

THÉORÈME 17. *Soient $f_1, \ldots, f_m$ des formes linéaires sur un espace vectoriel* L *de dimension finie n sur un corps* K. *La dimension du sous-espace vectoriel* M *de* L *défini par les relations* (17) *est égale à $n - r$, où r est le rang de la famille $(f_i)_{1 \leqslant i \leqslant m}$.*

On peut supposer $f_1, \ldots, f_r$ linéairement indépendantes, auquel cas on a des relations

$$f_j = \rho_{j1} f_1 + \cdots + \rho_{jr} f_r \quad \text{pour} \quad r + 1 \leqslant j \leqslant m;$$

il est alors clair que les solutions de (17) sont aussi les solutions de

$$f_1(x) = \cdots = f_r(x) = 0, \tag{18}$$

ce qui nous ramène au cas où les formes f_i données sont linéairement indépendantes.

Considérons alors l'homomorphisme $f : \mathrm{L} \to \mathrm{K}^r$ donné par

$$f(x) = (f_1(x), \ldots, f_r(x));$$

il est clair que $\mathrm{M} = \mathrm{Ker}\,(f)$, et par suite (Théorème 13) que

$$\dim(\mathrm{M}) = n - \dim(\mathrm{Im}\,(f));$$

tout revient donc à montrer que

$$\dim(\mathrm{Im}\,(f)) = r = \dim(\mathrm{K}^r),$$

autrement dit que $\mathrm{Im}\,(f) = \mathrm{K}^r$; mais cela résulte du Théorème 4.

1. Notations et traductions

Étant donné un anneau K, on appelle **système de n équations linéaires à p inconnues à coefficients dans** K tout système de relations

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{p1}\xi_p = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}\xi_1 + \cdots + \alpha_{pn}\xi_p = \beta_n \end{cases} \tag{1}$$

où les **coefficients** α_{ij} et les **seconds membres** β_j sont des éléments donnés de K. On appelle **solution** de (1), ou **solution dans l'anneau** K lorsqu'on veut écarter toute ambiguïté, tout vecteur

$$x = (\xi_1, \ldots, \xi_p) \in \mathrm{K}^p$$

dont les coordonnées satisfont aux relations (1).

Pour étudier le système (1), il est utile, et même indispensable, de le considérer sous les trois points de vue que voici.

Tout d'abord, on doit introduire l'homomorphisme

$$f : \mathrm{K}^p \to \mathrm{K}^n$$

dont la matrice, par rapport aux bases canoniques, est

$$\mathrm{A} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots\cdots & \cdots & \cdots\cdots \\ \alpha_{1n} & \cdots & \alpha_{pn} \end{pmatrix}; \tag{2}$$

en considérant le vecteur

$$b = (\beta_1, \ldots, \beta_n) \in \mathrm{K}^n,$$

il est clair, d'après le § 12, n° 3, que les solutions de (1) ne sont autres que les vecteurs $x \in \mathrm{K}^p$ qui vérifient la relation

$$f(x) = b. \tag{3}$$

En second lieu, identifions les vecteurs x et b aux matrices colonnes

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_p \end{pmatrix}, \qquad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix};$$

alors le système (1) prend la forme

$$Ax = b. \tag{4}$$

Enfin, introduisons sur K^p les formes linéaires

$$f_j(\xi_1, \ldots, \xi_p) = \alpha_{1j}\xi_1 + \cdots + \alpha_{pj}\xi_p;$$

alors le système (1) s'écrit

$$\begin{cases} f_1(x) = \beta_1 \\ \cdots\cdots\cdots \\ f_n(x) = \beta_n. \end{cases} \tag{5}$$

Dans toute la suite de ce § nous utiliserons sans explication les notations introduites ci-dessus, et nous supposerons que K est un *corps*

2. *Rang d'un système d'équations linéaires. Conditions d'existence de solutions*

On appelle **rang du système** (1) le rang r de la famille de formes linéaires $(f_j)_{1 \leqslant j \leqslant n}$; comme les coordonnées de f_j par rapport à la base de $(K^p)^*$ duale de la base canonique de K^p sont les scalaires α_{ij}, on voit (§ 19, Théorème 15) que le rang cherché est aussi celui de la matrice

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \cdots\cdots & & \cdots\cdots \\ \alpha_{p1} & \cdots & \alpha_{pn} \end{pmatrix}$$

i.e. de la transposée de A; donc (§ 19, Corollaire du Théorème 16) *le rang du système* (1) *est égal au rang de la matrice* A.

Soit r ce rang; au besoin en modifiant la numérotation des équations (1), on peut supposer que les formes $f_1, \ldots, f_r$ sont linéairement indépendantes; on a alors des relations de la forme

$$f_j = \rho_{j1}f_1 + \cdots + \rho_{jr}f_r \quad (r+1 \leqslant j \leqslant n),$$

et comme on l'a établi au § 19, n° 3, le système (15) possède des solutions si et seulement si les *conditions de compatibilité*

$$\beta_j = \rho_{j1}\beta_1 + \cdots + \rho_{jr}\beta_r \quad (r+1 \leqslant j \leqslant n)$$

sont vérifiées. S'il en est ainsi, les solutions de (5) sont les solutions de

$$\begin{cases} f_1(x) = \beta_1 \\ \cdots\cdots\cdots \\ f_r(x) = \beta_r. \end{cases} \tag{6}$$

On est ainsi ramené à étudier les systèmes dont les premiers membres sont des formes linéaires *linéairement indépendantes*; un tel système, comme on l'a montré au § 19 (Théorème 4), *possède toujours au moins une solution.*

3. Système homogène associé

Si l'équation (2) possède au moins une solution x, les autres sont évidemment les vecteurs $x + y$ où $f(y) = 0$, i.e. où $y \in \mathrm{Ker}(f)$. La question du nombre de solutions de (2) dépend donc de l'équation $f(y) = 0$; cf. § 7, Théorème 8.

Si l'on pose $y = (\eta_{i1}, \ldots, \eta_{ip})$, il est clair que la relation $f(y) = 0$ s'écrit

$$(1\ bis) \qquad \begin{cases} \alpha_{11}\eta_{i1} + \cdots + \alpha_{p1}\eta_{ip} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}\eta_{i1} + \cdots + \alpha_{pn}\eta_{ip} = 0; \end{cases}$$

on dit que (1 *bis*) est le **système d'équations linéaires et homogènes associé au système** (1) donné; on appelle **solution triviale** de ce système la solution pour laquelle

$$\eta_{i1} = \cdots = \eta_{ip} = 0.$$

On peut encore écrire (1 *bis*) sous la forme

$$(5\ bis) \qquad \begin{cases} f_1(y) = 0 \\ \cdots\cdots\cdots \\ f_n(y) = 0. \end{cases}$$

Les solutions y de ce système d'équations forment un sous-espace vectoriel de K^p dont la dimension, d'après le Théorème 17 du § précédent, est $p - r$, où r est le rang de la famille $(f_i)_{1 \leqslant i \leqslant n}$.

4. Systèmes de Cramer

Rappelons que, jusqu'à la fin de ce §, l'anneau de base K est supposé être un *corps*.

THÉORÈME 1. *Les propriétés suivantes sont équivalentes :*

a) Quels que soient les scalaires $\beta_1, \ldots, \beta_n \in \mathrm{K}$, *le système* (1) *possède une et une seule solution.*

b) On a $p = n$ *et le système homogène associé au système* (1) *n'admet que la solution triviale.*

c) On a $p = n$ *et les formes linéaires* $f_1, \ldots, f_n$ *sont linéairement indépendantes.*

La propriété *a*) exprime que l'homomorphisme

$$f : \mathrm{K}^p \to \mathrm{K}^n$$

est *bijectif*; s'il en est ainsi, f est un isomorphisme, de sorte que $\mathrm{Ker}(f) = \{0\}$ et que $\dim(\mathrm{K}^p) = \dim(\mathrm{K}^n)$, donc que $p = n$; ainsi *a*) implique *b*). Réciproquement, si f est injectif et si $p = n$ alors f est bijectif en vertu du § 19, Corollaire 1 du Théorème 13; donc *b*) implique *a*).

Montrons que *a*) implique *c*); on a déjà vu que *a*) implique $p = n$; si

$$\lambda_1 f_1 + \cdots + \lambda_n f_n = 0$$

est une relation linéaire entre les formes f_i, on a

$$\lambda_1 f_1(x) + \cdots + \lambda_n f_n(x) = 0 \quad \text{pour tout } x \in \mathrm{K}^p;$$

prenant pour x une solution de (1) il vient donc

$$\lambda_1 \beta_1 + \cdots + \lambda_n \beta_n = 0; \tag{7}$$

si (1) admet une solution quelles que soient les valeurs des seconds membres, la relation (7) est donc vérifiée quels que soient les $\beta_j \in \mathrm{K}$, ce qui implique évidemment

$$\lambda_1 = \cdots = \lambda_n = 0;$$

donc *a*) implique *c*) (on pourrait aussi remarquer que, si les f_i ne sont pas linéairement indépendantes, le système (5) n'admet de solutions que si les β_i vérifient des conditions non triviales, à savoir les conditions de compatibilité du n° 2).

Il reste à établir que *c*) implique *a*). D'après le n° 2, ou le Théorème 4 du § 19, le système (1) possède toujours *au moins une* solution; donc l'homomorphisme f est surjectif; mais comme K^p et K^n ont par hypothèse même dimension, il est aussi injectif, et par suite (1) possède toujours *au plus une* solution, ce qui termine la démonstration.

On appelle **système de Cramer** tout système d'équations linéaires vérifiant les conditions *a*), *b*), *c*) du Théorème 1. Le but du présent § est essentiellement de montrer que *la résolution d'un système quelconque d'équations linéaires peut toujours se ramener à celle d'un système de Cramer.*

Lorsqu'on sait d'avance que $p = n$ (autrement dit, lorsqu'on a un système comportant autant d'équations que d'inconnues), on peut utiliser le Théorème suivant, qui donne plusieurs caractérisations importantes des systèmes de Cramer :

THÉORÈME 2. *Étant donné un système*

$$\mathrm{A}x = b \tag{4}$$

de n équations linéaires à n inconnues, les propriétés suivantes sont équivalentes :

a) *Quel que soit b, le système* (4) *admet une et une seule solution.*

b) *Quel que soit b, le système* (4) *possède au moins une solution.*

c) *Quel que soit b, le système* (4) *possède au plus une solution.*

d) *Il existe une valeur de b pour laquelle le système* (4) *possède une et une seule solution.*

e) *Le système homogène* $\mathrm{A}y = 0$ *associé au système* (4) *n'admet que la solution triviale* $y = 0$.

f) *La matrice* A *est inversible.*

En outre, si ces conditions sont vérifiées, la solution du système (4) *est*

$$x = \mathrm{A}^{-1} b. \tag{8}$$

En écrivant (4) sous la forme $f(x) = b$, où f est l'endomorphisme de K^n dont la matrice par rapport à la base canonique est A, on voit que *a*) signifie que f est bijectif, *b*) que f est surjectif, *c*) que f est injectif, *e*) que $\mathrm{Ker}\,(f) = 0$, et *f*) que f est inversible dans l'anneau des endomorphismes de K^n (cf. § 15, n° 2); l'équivalence des conditions *a*), *b*), *c*), *e*) et *f*) résulte donc du Corollaire 1 du Théroème 13 du § 19.

Il est clair par ailleurs que *a*) implique *d*); et que *d*) implique *e*) comme il résulte des raisonnements du n° 2.

Pour achever la démonstration, il reste donc à établir la formule (8), ce qui est trivial.

Exemple 1. Supposons $n = 2$ et K commutatif; on a donc un système

$$\begin{cases} ax + by = u \\ cx + dy = v \end{cases}$$

de deux équations linéaires à deux inconnues x et y. D'après le § 15, n° 3, la matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible si et seulement si

$$ad - bc \neq 0,$$

de sorte que cette condition caractérise donc les systèmes de Cramer dans ce cas. Si elle est remplie, il est facile de voir que la solution du système considéré est donnée par les formules

$$x = \frac{du - bv}{ad - bc}, \qquad y = \frac{av - cu}{ad - bc}.$$

La théorie des déterminants permet, comme on le verra au § 24, d'étendre ces formules aux systèmes de Cramer à un nombre quelconque d'équations, à coefficients dans un corps K *commutatif*. Toute la question est évidemment de trouver des formules explicites pour calculer l'inverse d'une matrice carrée inversible.

5. *Systèmes d'équations indépendantes : réduction à un système de Cramer*

Nous allons maintenant examiner le système (1) dans le cas où les formes linéaires $f_i (1 \leqslant i \leqslant n)$ sont linéairement indépendantes, cas auquel on peut toujours se ramener comme on l'a vu n° 2; mais nous ne supposerons plus que $p = n$. On a alors nécessairement

$$n \leqslant p,$$

puisque $f_1, \ldots, f_n$ sont par hypothèse des éléments linéairement indépendants d'un espace vectoriel de dimension p (à savoir le dual de K^p).

La matrice $A = (a_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n}$ est alors de rang n (n° 2), et par suite (§ 19, Théorème 16) on peut en extraire une matrice carrée inversible d'ordre n; nous supposerons que la matrice

$$U = (a_{ij})_{1 \leqslant i,\, j \leqslant n}$$

est inversible dans ce qui suit — on peut toujours s'y ramener, au besoin en modifiant la numérotation des inconnues ξ_i.

Écrivons alors le système (1) sous la forme

$$(9)\qquad \begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{n1}\xi_n = \gamma_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}\xi_1 + \cdots + \alpha_{nn}\xi_n = \gamma_n \end{cases}$$

où l'on a posé

$$(10)\qquad \gamma_j = \beta_j - (\alpha_{n+1,\,j}\xi_{n+1} + \cdots + \alpha_{p,j}\xi_p);$$

comme U est inversible, (9) est un *système de Cramer* par rapport aux inconnues $\xi_1, \ldots, \xi_n$ et par suite admet une et une seule solution quelles que soient les valeurs des seconds membres γ_i : or ceux-ci dépendent uniquement des inconnues $\xi_{n+1}, \ldots, \xi_p$; il s'ensuit donc que le système (1) possède une et une seule solution pour laquelle les inconnues $\xi_{n+1}, \ldots, \xi_p$ ont des valeurs données d'avance.

D'ailleurs, la solution de (9) est donnée par la relation

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \mathrm{U}^{-1}.\begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix};$$

posant

$$\mathrm{U}^{-1} = (\nu_{ij})_{1 \leqslant i,\, j \leqslant n}$$

il vient donc

$$(11)\qquad \xi_i = \nu_{1i}\gamma_1 + \cdots + \nu_{ni}\gamma_n,$$

et en remplaçant les γ_j par leurs valeurs (10) il vient évidemment des formules du type

$$(12)\qquad \xi_i = \nu_{1i}\beta_1 + \cdots + \nu_{ni}\beta_n + \lambda_{n+1,i}\xi_{n+1} + \cdots + \lambda_{pi}\xi_p \quad (1 \leqslant i \leqslant n)$$

où les λ_{ki} sont de nouvelles constantes, qui ne dépendent que des coefficients α_{ij} du système (1).

Comme (1) est équivalent à la conjonction de (9) et (10), et comme (9) est équivalent à (11), il est clair que l'ensemble des relations (1) est équivalent à l'ensemble des relations (12). Autrement dit :

THÉORÈME 3. *Supposons que les formes linéaires $f_1, \ldots, f_n$ soient linéairement indépendantes. On peut alors numéroter les inconnues ξ_i de telle sorte que le système* (1) *possède une et une seule solution pour laquelle les inconnues $\xi_{n+1}, \ldots, \xi_p$ aient des valeurs arbitrairement données. S'il en est ainsi, il existe des constantes ν_{ij} et λ_{ki}, ne dépendant que des coefficients α_{ij} du système* (1), *et telles que les solutions de* (1) *soient toutes les suites $(\xi_1, \ldots, \xi_p)$ vérifiant les relations* (12).

Exemple 2. Prenons un système

$$\begin{aligned} a'x + b'y + c'z &= u' \\ a''x + b''y + c''z &= u'' \end{aligned}$$

de deux équations à trois inconnues. Le rang du système est 0, 1 ou 2.

Si le rang est 0, on a

$$a' = b' = c' = a'' = b'' = c'' = 0,$$

et le système n'a de solutions que si $u' = u'' = 0$ (auquel cas toutes les suites (x, y, z) sont des solutions).

Si le rang est 1, cela veut dire que les coefficients $a, \ldots, c''$ ne sont pas tous nuls, mais que les deux formes linéaires figurant dans les premiers membres du système donné sont proportionnelles; si K est commutatif, cela veut dire aussi (§ 19, Exemple 4) que

$$b'c'' - b''c' = a'c'' - a''c' = a'b'' - a''b' = 0. \tag{13}$$

Supposons alors par exemple $a' \neq 0$; on a nécessairement une relation

$$a''x + b''y + c''z = \rho(a'x + b'y + c'z),$$

ce qui donne en particulier $a'' = \rho a'$, donc

$$\rho = a''/a'.$$

Dans ce cas, le système admet une solution si et seulement si $u'' = \rho u'$, i.e. si $u''a' - u'a'' = 0$. S'il en est ainsi on est ramené à résoudre

$$a'x + b'y + c'z = u'$$

et puisque $a' \neq 0$ la solution « générale » est donnée par

$$x = \frac{u' - b'y - c'z}{a'},$$

en sorte qu'on peut choisir arbitrairement y et z.

Si enfin le système est de rang 2, et si K est commutatif, les relations (13) ne sont pas toutes vérifiées; supposons par exemple que

$$a'c'' - a''c' \neq 0;$$

en écrivant le système donné sous la forme

$$\begin{aligned} a'x + c'z &= u' - b'y \\ a''x + c''z &= u'' - b''y \end{aligned}$$

on est ramené à un système de Cramer en x et z; autrement dit, on peut attribuer à y une valeur arbitraire, et alors la solution du système donné est fournie par les formules

$$x = \frac{c''(u' - b'y) - c'(u'' - b''y)}{a'c'' - a''c'} = \frac{c''u' - c'u''}{a'c'' - a''c'} + \frac{c'b'' - c''b'}{a'c'' - a''c'}y,$$

$$z = \frac{a'u'' - a''u'}{a'c'' - a''c'} + \frac{a''b' - a'b''}{a'c'' - a''c'}y.$$

Déterminants.

Comme on l'a dit dans l'introduction aux §§ 18, 19 et 20, la théorie des déterminants a notamment pour but de fournir des critères *explicites* d'indépendance linéaire, et des formules *explicites* de résolution des systèmes d'équations linéaires.

Au lieu de la méthode traditionnelle qui consiste à définir un déterminant à l'aide de la formule du § 23, n° 5 et à en déduire ensuite les propriétés des déterminants, nous avons adopté, pour exposer la théorie des déterminants, la méthode « géométrique » fondée sur la théorie des formes multilinéaires alternées consistant, au contraire, à retrouver la règle de calcul des déterminants à partir des propriétés fondamentales de ceux-ci. Cette méthode était déjà enseignée par Kronecker il y a quatre-vingts ans, et se répand de plus en plus depuis une quinzaine d'années.

Alors que la théorie des modules développée au Chapitre III était valable sur un anneau de base K quelconque, celle des déterminants suppose K *commutatif*. Quelques-uns de ses résultats supposent même que K est un *corps* commutatif, mais nous n'avons fait cette hypothèse que lorsqu'elle était indispensable. Dans la plupart des cas, on n'a besoin que de calculs mécaniques pour établir les formules qu'on a en vue, et il est alors aussi simple (et plus utile) de les établir pour un anneau commutatif quelconque que pour un corps, ou même que pour le corps des nombres réels.

Le lecteur débutant qui trouvera trop difficiles les calculs du § 23 pourra passer d'abord au § 24 (en admettant le Théorème 1) et s'entraîner, à l'aide des *Exercices*, à calculer des déterminants. Il pourra revenir alors au § 23.

§ 21. Fonctions multilinéaires

1. *Définition des applications multilinéaires*

Soient X, Y et M des modules sur un anneau commutatif K. On dit qu'une application

$$f : X \times Y \to M$$

est **bilinéaire** si $f(x, b)$ est fonction linéaire de $x \in X$ pour tout $b \in Y$, et si $f(a, y)$ est fonction linéaire de $y \in Y$ pour tout $a \in X$; cela signifie donc qu'on a les identités

$$(1) \qquad \begin{aligned} f(x' + x'', y) &= f(x', y) + f(x'', y), \qquad f(\lambda x, y) = \lambda f(x, y) \\ f(x, y' + y'') &= f(x, y') + f(x, y''), \qquad f(x, \lambda y) = \lambda f(x, y). \end{aligned}$$

Soient maintenant X, Y, Z et M des K-modules; on dit qu'une application

$$f : X \times Y \times Z \to M$$

est **trilinéaire** si $f(x, b, c)$ est fonction linéaire de $x \in X$ pour $b \in Y$ et $c \in Z$ donnés, si $f(a, y, c)$ est fonction linéaire de $y \in Y$ pour $a \in X$ et $c \in Z$ donnés, et si $f(a, b, z)$ est fonction linéaire de $z \in Z$ pour $a \in X$ et $b \in Y$ donnés.

Plus généralement, soient $X_1, \ldots, X_p$ et M des K-modules; on dit qu'une application de la forme

$$f : X_1 \times \cdots \times X_p \to M$$

est **multilinéaire** (et, plus précisément, **p-linéaire**) si, pour tout indice i tel que $1 \leqslant i \leqslant p$ et quels que soient les vecteurs

$$a_1 \in X_1, \ldots, a_{i-1} \in X_{i-1}, \qquad a_{i+1} \in X_{i+1}, \ldots, a_p \in X_p,$$

l'application

$$(2) \qquad x \mapsto f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_p)$$

de X_i dans M est linéaire. Autrement dit, f est multilinéaire si, en donnant à $p - 1$ des variables des valeurs fixes, on obtient une fonction linéaire de la variable non

fixée. On pourrait exprimer cette condition par des formules généralisant (1), à savoir

$$(3)\quad f(x_1, \ldots, x_{i-1}, x_i' + x_i'', x_{i+1}, \ldots, x_p) = f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_p) + f(x_1, \ldots, x_{i-1}, x_i'', x_{i+1}, \ldots, x_p)$$

$$(4)\quad f(x_1, \ldots, x_{i-1}, \lambda x_i, x_{i+1}, \ldots, x_p) = \lambda . f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_p).$$

On désigne généralement l'ensemble de toutes les applications multilinéaires de $X_1 \times \cdots \times X_p$ dans M par la notation

$$\mathfrak{L}(X_1, \ldots, X_p; M).$$

C'est un *sous-module* du module (§ 10, *Exemple* 4) de toutes les applications (multilinéaires ou non) de l'ensemble $X = X_1 \times \cdots \times X_p$ dans M; autrement dit, si f et g sont multilinéaires, il en est de même de $\lambda f + \mu g$ quels que soient les scalaires λ et μ. Ce résultat (qui provient immédiatement du fait qu'une combinaison linéaire d'applications linéaires, par exemple d'applications de la forme (2), est encore une application linéaire) permet de considérer l'ensemble $\mathfrak{L}(X_1, \ldots, X_p; M)$ comme un *module sur l'anneau* K.

Donnons maintenant quelques exemples importants.

Exemple 1. Prenons $X_1 = \cdots = X_p = M = K$ et posons

$$f(x_1, \ldots, x_p) = x_1 \ldots x_p;$$

alors f est multilinéaire en vertu des axiomes des anneaux commutatifs. On notera du reste l'analogie de la relation (1) ou (3) avec celle qui exprime la distributivité de la multiplication par rapport à l'addition.

Exemple 2. Soient X, Y deux K-modules; on appelle **forme bilinéaire sur** $X \times Y$ toute application bilinéaire de $X \times Y$ dans l'anneau de base K. Considérons par exemple une forme linéaire u sur X et une forme linéaire v sur Y; alors

$$f(x, y) = u(x)v(y)$$

est une forme bilinéaire sur $X \times Y$, car, si l'on donne par exemple à y une valeur fixe b, on obtient l'expression $u(x)v(b)$ qui est proportionnelle à $u(x)$, et est donc fonction linéaire de x.

Plus généralement, si $X_1, \ldots, X_p$ sont des modules sur l'anneau K, on appelle **forme multilinéaire sur** $X_1 \times \cdots \times X_p$ toute application multilinéaire de $X_1 \times \cdots \times X_p$ dans K. Si l'on choisit une forme linéaire u_i sur X_i pour tout i, et si l'on pose

$$f(x_1, \ldots, x_p) = u_1(x_1) \ldots u_p(x_p),$$

on obtient une forme multilinéaire sur $X_1 \times \cdots \times X_p$. On l'appelle le **produit tensoriel des formes linéaires** $u_1, \ldots, u_p$ et on la désigne généralement par la notation

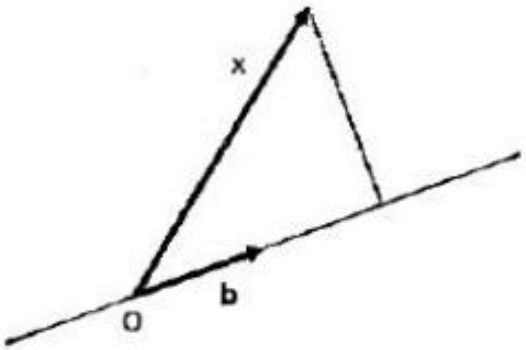
$$f = u_1 \otimes \cdots \otimes u_p.$$

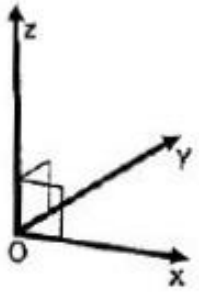
On montrera plus loin (Théorème 3) que si $X_1, \ldots, X_p$ sont des K-modules libres de type fini, alors *toute* forme multilinéaire sur $X_1 \times \cdots \times X_p$ est une somme de produits tensoriels de formes linéaires.

Exemple 3. Prenons $K = R$ et désignons par E l'espace vectoriel formé par les vecteurs d'origine donnée O dans l'espace usuel à trois dimensions. Supposant choisie une unité de longueur, on peut définir le **produit scalaire de deux vecteurs** x et y : c'est le *nombre* produit des longueurs de x et y et du cosinus de l'angle amenant x sur y. On désigne ce produit scalaire par l'une ou l'autre des notations suivantes :

$$x.y, \qquad (x,y), \qquad (x|y);$$

nous utiliserons exclusivement la troisième (la seconde est exclue puisqu'elle désigne déjà le couple formé par x et y). Ceci dit, l'expression $(x|y)$ est une





forme bilinéaire sur E × E; pour montrer par exemple que $(x|b)$ est fonction linéaire de x, on observe que ce nombre est proportionnel à la mesure orientée de la projection orthogonale du vecteur x sur la droite portant le vecteur b; or l'opération consistant à projeter sur une droite fixe est linéaire (*).

Exemple 4. K et E étant comme dans l'*Exemple* précédent, considérons le **produit vectoriel de deux vecteurs** x et y : c'est le *vecteur* z qui a pour longueur le produit des longueurs de x et y et du sinus de leur angle, autrement dit l'aire du parallélogramme de côtés x et y, qui est orthogonal au plan défini par x et y, et qui est orienté de telle sorte que le trièdre xyz soit direct, i.e. conforme aux injonctions d'Ampère et de Maxwell (la notion de trièdre direct n'a pas de sens *mathématique*; la seule notion correcte est celle de deux trièdres

(*) On laisse au lecteur le soin de développer en détail les considérations de géométrie élémentaire qui sont à la base de cet Exemple et du suivant.

de même orientation, qu'on peut définir à l'aide de la théorie des déterminants comme on le verra au § 23, *Remarque* 4). On le note

$$x \times y \qquad \text{ou} \qquad x \wedge y;$$

nous utiliserons uniquement la seconde notation. Cela dit, il est facile de voir (et le lecteur devra montrer à titre d'exercice) que l'application

$$(x, y) \mapsto x \wedge y$$

de $E \times E$ dans E est bilinéaire.

On notera les formules de commutation suivantes :

$$(x|y) = (y|x); \qquad x \wedge y = -y \wedge x.$$

Exemple 5. K et E restant comme ci-dessus, on appelle **produit mixte de trois vecteurs** x, y, z le nombre

$$(x|y|z) = (x|\, y \wedge z),$$

produit scalaire de x et du produit vectoriel de y et z. C'est une fonction trilinéaire de x, y, z. En valeur absolue, $(x|y|z)$ est égal au volume du parallélipipède construit sur les vecteurs x, y, z, et le signe de $(x|y|z)$ est positif si le trièdre x, y, z est direct, négatif s'il est rétrograde. Il s'ensuit donc que l'on a les relations

$$(x|y|z) = (y|z|x) = (z|x|y) = -(x|z|y) = -(y|x|z) = -(z|y|x).$$

¶ *Exemple* 6. Soient K un anneau commutatif quelconque, X un K-module, et X^* le module dual (§ 16, n° 1). Étant donnés des entiers $p, q \geqslant 0$, on appelle **tenseur p fois covariant et q fois contravariant**, ou encore **tenseur d'espèce** $\binom{p}{q}$, toute application $(p+q)$-linéaire de

$$(X^*)^p \times X^q = \underbrace{X^* \times \cdots \times X^*}_{p} \times \underbrace{X \times \cdots \times X}_{q}$$

dans l'anneau de base K. C'est donc une fonction $f(u_1, \ldots, u_p, x_1, \ldots, x_q)$ à valeurs dans K, définie lorsque $u_1, \ldots, u_p$ sont des formes linéaires sur X et $x_1, \ldots, x_q$ des vecteurs de X, et qui dépend linéairement de chacune des variables $u_1, \ldots, x_q$. En particulier, on appelle **forme q-linéaire sur X** tout tenseur d'espèce $\binom{0}{q}$, i.e. toute application multilinéaire de X^q dans K.

Parmi les tenseurs sur X figurent les formes linéaires sur X : ce sont les tenseurs d'espèce $\binom{0}{1}$. D'autre part, chaque vecteur x de X définit aussi un tenseur sur X, d'espèce $\binom{1}{0}$, i.e. une forme linéaire sur X^*, à savoir la forme linéaire $u \to u(x)$ déjà utilisée au § 16, n° 3 pour plonger un module dans son bidual. Lorsque X est libre de type fini, la correspondance entre éléments de X et tenseurs d'espèce $\binom{1}{0}$ est bijective (§ 16, Théorème 2), ce qui permet d'identifier ces deux notions; c'est ce qu'on fait en Physique notamment.

La plupart des physiciens donnent des tenseurs (dans le cas des espaces vectoriels de petite dimension sur **R**) une définition nettement plus compliquée que la précédente, et qu'on trouvera plus loin (n° 5, *Remarque* 3). En outre, en Physique, on s'intéresse exclusivement aux **champs de tenseurs**; on appelle ainsi toute application de l'espace vectoriel X considéré dans l'ensemble des tenseurs d'espèce donnée sur X — autrement dit, toute fonction sur X dont la valeur en chaque point de X est un tenseur d'espèce donnée sur X. L'utilisation des tenseurs en Physique a été pendant longtemps rendue difficilement compréhensible par le fait qu'on ne distinguait pas entre un tenseur (au sens donné ici à ce mot) et une fonction dont les valeurs sont des tenseurs; on obtenait ainsi une situation analogue à celle qui consisterait à parler de fonctions à valeurs vectorielles sans avoir d'abord défini ce qu'est un vecteur... Il faut préciser, à la décharge des physiciens, que la notion de forme linéaire sur un espace vectoriel (sans laquelle il est pratiquement impossible de donner une définition simple des tenseurs) ne s'est vraiment répandue en Mathématiques qu'à partir de 1930 environ.

2. *Produit tensoriel d'applications multilinéaires*

Soient

$$f : \mathrm{X}_1 \times \cdots \times \mathrm{X}_p \to \mathrm{K}, \qquad g : \mathrm{Y}_1 \times \cdots \times \mathrm{Y}_q \to \mathrm{K}$$

des formes multilinéaires. On appelle **produit tensoriel** de f et g l'application

$$f \otimes g : \mathrm{X}_1 \times \cdots \times \mathrm{X}_p \times \mathrm{Y}_1 \times \cdots \times \mathrm{Y}_q \to \mathrm{K}$$

donnée par

$$(5) \qquad f \otimes g(x_1, \ldots, x_p, y_1, \ldots, y_q) = f(x_1, \ldots, x_p) g(y_1, \ldots, y_q);$$

c'est encore une application *multilinéaire*, car si l'on donne par exemple à $x_1, \ldots, y_q$ des valeurs fixes $a_2, \ldots, b_q$, il reste l'expression

$$f(x_1, a_2, \ldots, a_p) g(b_1, \ldots, b_q)$$

proportionnelle, comme fonction de x_1, à $f(x_1, a_2, \ldots, a_p)$, donc fonction linéaire de x_1 comme cette dernière.

Si l'on a trois formes multilinéaires f, g, h on a la relation d'associativité

$$(f \otimes g) \otimes h = f \otimes (g \otimes h),$$

et du reste la valeur commune des deux membres est la fonction

$$f(x_1, \ldots, x_p) g(y_1, \ldots, y_q) h(z_1, \ldots, z_r)$$

comme on le voit aussitôt.

Ceci permet de définir des produits tensoriels d'un nombre quelconque de formes multilinéaires, et de généraliser la notion introduite dans l'*Exemple* 2 ci-dessus pour les formes linéaires.

On notera que la formule

$$f \otimes g = g \otimes f,$$

même lorsque les deux membres sont de même espèce, est *fausse*. Si par exemple f et g

sont des formes linéaires sur un module X, on a

$$f \otimes g(x, y) = f(x)g(y)$$
$$g \otimes f(x, y) = f(y)g(x),$$

et ces expressions n'ont aucune raison d'être identiques.

¶ *Exemple* 7. Considérons, sur un module X, un tenseur f d'espèce $\binom{p}{q}$ et un tenseur g d'espèce $\binom{r}{s}$; alors $f \otimes g$ est une forme multilinéaire sur le produit cartésien

$$(X^*)^p \times X^q \times (X^*)^r \times X^s;$$

on identifie en général celui-ci à

$$(X^*)^{p+r} \times X^{q+s},$$

ce qui permet de considérer $f \otimes g$ comme un tenseur d'espèce $\binom{p+r}{q+s}$ sur X, donné par

$$f \otimes g(u_1, \ldots, u_{p+r}, x_1, \ldots, x_{q+s}) = f(u_1, \ldots, u_p, x_1, \ldots, x_q)\, g(u_{p+1}, \ldots, u_{p+r}, x_{q+1}, \ldots, x_{q+s})$$

quels que soient les $u_i \in X^*$ et les $x_j \in X$. On dit que $f \otimes g$ est le **produit tensoriel des tenseurs** f et g.

Considérons par exemple deux vecteurs a, $b \in X$ et trois formes linéaires f, g, $h \in X^*$; identifiant a et b à des tenseurs (*Exemple* 6) on peut définir le tenseur $a \otimes b \otimes f \otimes g \otimes h = \varphi$; il est deux fois covariant et trois contravariant, i.e. c'est une fonction multilinéaire de deux formes linéaires u, $v \in X^*$ et de trois vecteurs x, y, $z \in X$; on vérifie facilement qu'on a en fait

$$\varphi(u, v, x, y, z) = u(a)v(b)f(x)g(y)h(z).$$

Remarque 1. Dans la formule (5) définissant $f \otimes g$, il est essentiel, si l'on veut obtenir une fonction multilinéaire, que les variables y_j soient indépendantes des variables x_i. Par exemple, si f et g sont deux formes linéaires sur un module L, l'expression $f(x)g(y)$ est une forme bilinéaire sur L × L, mais la fonction $f(x)g(x)$ n'est pas, en général, une forme linéaire sur L.

3. *Quelques identités algébriques*

Lorsqu'on effectue des calculs algébriques dans un anneau, on a souvent à calculer un produit dont chaque terme est une somme d'autres termes, et à le « développer » à l'aide de la règle suivante : pour multiplier des sommes les unes par les autres, on choisit arbitrairement un terme dans chaque somme, on multiplie les termes choisis les uns par les autres, et on ajoute les résultats ainsi obtenus. Cette règle peut se traduire par des formules; par exemple si $(x_i)_{i \in I}$ et $(y_j)_{j \in J}$ sont deux familles finies d'éléments d'un anneau K, on a

$$\sum_{i \in I} x_i \cdot \sum_{j \in J} y_j = \sum_{(i,j) \in I \times J} x_i y_j$$

si l'on a trois familles finies $(x_i)_{i\in I}$, $(y_j)_{j\in J}$ et $(z_k)_{k\in K}$ d'éléments d'un anneau, on a la relation

$$\sum_{i\in I} x_i \cdot \sum_{j\in J} y_j \cdot \sum_{k\in K} z_k = \sum_{\substack{i\in I\\ j\in J\\ k\in K}} x_i y_j z_k; \tag{7}$$

plus généralement encore, supposons données p familles finies d'éléments d'un anneau, familles que nous noterons $(x_{1i_1})_{i_1\in I_1}, \ldots, (x_{pi_p})_{i_p\in I_p}$; alors on a

$$\sum_{i_1\in I_1} x_{1i_1} \cdots \sum_{i_p\in I_p} x_{pi_p} = \sum_{i_1\in I_1,\ldots,i_p\in I_p} x_{1i_1} \cdots x_{pi_p}. \tag{8}$$

Remarque 2. Dans la formule (6) il arrive fréquemment que $I = J$ et qu'on désigne les familles données par $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$; on fera attention dans ce cas à *ne pas* écrire la relation (6) sous la forme

$$\sum_{i\in I} x_i \cdot \sum_{i\in I} y_i = \sum_{i\in I} x_i y_i$$

car cette relation est évidemment fausse — par exemple, il serait déraisonnable de croire qu'on a toujours

$$(x' + x'')(y' + y'') = x'y' + x''y''\ldots$$

La seule méthode sûre pour éviter des erreurs de ce genre est de ne jamais désigner par la même lettre des indices de sommation figurant dans deux sommes distinctes. On fera bien aussi de garder présent à l'esprit le fait qu'un indice de sommation ne figure pas effectivement dans le résultat de la sommation, et ne joue que le rôle d'une abréviation destinée à désigner une opération à effectuer; on peut donc toujours le remplacer par toute autre lettre non encore utilisée. Si par exemple on a des éléments $x_1, \ldots, x_n$ d'un groupe additif, les expressions

$$\sum_{1\leqslant i\leqslant n} x_i, \qquad \sum_{1\leqslant h\leqslant n} x_h \qquad \sum_{1\leqslant \lambda\leqslant n} x_\lambda$$

sont identiques.

On peut démontrer pour les applications multilinéaires des formules analogues à (6), (7) et (8), la raison en étant que les identités de « distributivité de la multiplication par rapport à l'addition » qui impliquent (6), (7) et (8) sont encore vérifiées, par définition, dans le cas des applications multilinéaires.

Considérons par exemple des modules X, Y, M sur un anneau commutatif K, et une application bilinéaire f de $X \times Y$ dans M. Étant données une famille finie $(x_i)_{i\in I}$ d'éléments de X, et une famille finie $(y_j)_{j\in J}$ d'éléments de Y, on a alors la relation

$$f\Big(\sum_{i\in I} x_i, \sum_{j\in J} y_j\Big) = \sum_{(i,j)} f(x_i, y_j). \tag{6 bis}$$

Pour le voir, posons

$$a = \sum_{i \in I} x_i \qquad \text{et} \qquad f_a(y) = f(a, y);$$

comme f est bilinéaire, f_a est une application linéaire de Y dans M, et par suite

$$f\left(a, \sum_{j \in J} y_j\right) = f_a\left(\sum_{j \in J} y_j\right) = \sum_{j \in J} f_a(y_j) = \sum_{j \in J} f(a, y_j);$$

mais en posant $f_j(x) = f(x, y_j)$ on obtient, pour la même raison, une application linéaire de X dans M, en sorte que

$$f(a, y_j) = f_j(a) = f_j\left(\sum_{i \in I} x_i\right) = \sum_{i \in I} f_j(x_i) = \sum_{i \in I} f(x_i, y_j);$$

en portant ce résultat dans le précédent, on obtient évidemment (6 *bis*).

Soient de même X, Y, Z et M des K-modules, et f une application trilinéaire de $X \times Y \times Z$ dans M. Alors, si $(x_i)_{i \in I}$, $(y_j)_{j \in J}$ et $(z_k)_{k \in K}$ sont des familles finies d'éléments de X, Y et Z respectivement, on a la relation

(7 *bis*)
$$f\left(\sum_{i \in I} x_i, \sum_{j \in J} y_j, \sum_{k \in K} z_k\right) = \sum_{\substack{i \in I \\ j \in J \\ k \in K}} f(x_i, y_j, z_k).$$

Posons en effet

$$c = \sum z_k \qquad \text{et} \qquad f_c(x, y) = f(x, y, c);$$

il est clair que f_c est bilinéaire, et en appliquant (6 *bis*) à f_c on trouve donc que le premier membre de (7 *bis*) est égal à

$$\sum_{\substack{i \in I \\ j \in J}} f_c(x_i, y_j) = \sum_{\substack{i \in I \\ j \in J}} f(x_i, y_j, \sum_{k \in K} z_k);$$

mais l'application

$$f_{ij}(z) = f(x_i, y_j, z)$$

de Z dans M est linéaire, et comme le terme général de la somme qu'on vient d'écrire est la valeur de f_{ij} sur le vecteur $\sum z_k$ on voit que ce terme est égal à

$$\sum_{k \in K} f_{ij}(z_k) = \sum_{k \in K} f(x_i, y_j, z_k);$$

en définitive, le premier membre de (7 *bis*) est donc égal à

$$\sum_{\substack{i \in I \\ j \in J}} \sum_{k \in K} f(x_i, y_j, z_k) = \sum_{\substack{i \in I \\ j \in J \\ k \in K}} f(x_i, y_j, z_k),$$

ce qui prouve (7 *bis*).

Enfin, pour généraliser (8), on considère une application p-linéaire

$$f : \mathrm{X}_1 \times \cdots \times \mathrm{X}_p \to \mathrm{M}$$

et l'on choisit dans chaque module X_h ($1 \leqslant h \leqslant p$) une famille finie de vecteurs, soit $(x_{hi_h})_{i_h \in \mathrm{I}_h}$; on a alors l'identité

$$(8\ bis) \qquad f\left(\sum_{i_1 \in \mathrm{I}_1} x_{1i_1}, \ldots, \sum_{i_p \in \mathrm{I}_p} x_{pi_p}\right) = \sum_{\substack{i_1 \in \mathrm{I}_1 \\ \cdots\cdots \\ i_p \in \mathrm{I}_p}} f(x_{1i_1}, \ldots, x_{pi_p}).$$

Cette identité se démontre par récurrence sur p; posant

$$c = \sum_{i_p \in \mathrm{I}_p} x_{pi_p} \qquad \text{et} \qquad f_c(x_1, \ldots, x_{p-1}) = f(x_1, \ldots, x_{p-1}, c)$$

on obtient une application $(p-1)$-linéaire f_c; pour celle-ci la formule analogue à (8 *bis*) donne

$$f_c\left(\sum_{i_1 \in \mathrm{I}_1} x_{1i_1}, \ldots, \sum_{i_{p-1} \in \mathrm{I}_{p-1}} x_{p-1, i_{p-1}}\right) = \sum_{\substack{i_1 \in \mathrm{I}_1 \\ \cdots\cdots \\ i_{p-1} \in \mathrm{I}_{p-1}}} f_c(x_{1i_1}, \ldots, x_{p-1, i_{p-1}})$$

$$= \sum_{\substack{i_1 \in \mathrm{I}_1 \\ \cdots\cdots \\ i_{p-1} \in \mathrm{I}_{p-1}}} f_{i_1 \ldots i_{p-1}}\left(\sum_{i_p \in \mathrm{I}_p} x_{pi_p}\right)$$

où l'on a posé

$$f_{i_1 \ldots i_{p-1}}(x) = f(x_{1i_1}, \ldots, x_{p-1, i_{p-1}}, x) \quad \text{pour } x \in \mathrm{X}_p;$$

or cette dernière expression est fonction linéaire de x, et on voit donc que le premier membre de (8 *bis*) est égal à

$$\sum_{\substack{i_1 \in \mathrm{I}_1 \\ \cdots\cdots \\ i_{p-1} \in \mathrm{I}_{p-1}}} \sum_{i_p \in \mathrm{I}_p} f_{i_1 \ldots i_{p-1}}(x_{pi_p}),$$

et comme

$$f_{i_1 \ldots i_{p-1}}(x_{pi_p}) = f(x_{1i_1}, \ldots, x_{pi_p})$$

on obtient bien (8 *bis*).

On observera que dans ces formules l'anneau de base K n'est jamais intervenu, autrement dit on n'a utilisé que l'identité (3) du n° 1, et non pas l'identité (4). Pour faire intervenir celle-ci dans (8 *bis*), choisissons en outre des familles

$$(\lambda_{1i_1})_{i_1 \in \mathrm{I}_1}, \quad \ldots, \quad (\lambda_{pi_p})_{i_p \in \mathrm{I}_p}$$

d'éléments de K, et dans (8 *bis*) remplaçons chaque vecteur x_{hi_h} par $\lambda_{hi_h} x_{hi_h}$; le terme général du second membre de (8 *bis*) est alors remplacé par

$$\lambda_{1i_1} \ldots \lambda_{pi_p} f(x_{1i_1}, \ldots, x_{pi_p})$$

en raison du fait qu'on a d'une manière générale

$$f(\xi_1 x_1, \ldots, \xi_p x_p) = \xi_1 \ldots \xi_p f(x_1, \ldots, x_p). \tag{9}$$

Ceci fait, on voit que (8 *bis*) conduit à la formule plus générale que voici :

$$f\left(\sum_{i_1 \in I_1} \lambda_{1i_1} x_{1i_1}, \ldots, \sum_{i_p \in I_p} \lambda_{pi_p} x_{pi_p}\right) = \sum_{\substack{i_1 \in I_1 \\ \ldots\ldots \\ i_p \in I_p}} \lambda_{1i_1} \ldots \lambda_{pi_p} f(x_{1i_1}, \ldots, x_{pi_p}) \tag{10}$$

qui permet de calculer les valeurs de f sur des combinaisons linéaires de vecteurs. Pour $p = 2$ cette formule s'écrit encore

$$f\left(\sum_{i \in I} \lambda_i x_i, \sum_{j \in J} \mu_j y_j\right) = \sum_{\substack{i \in I \\ j \in J}} \lambda_i \mu_j f(x_i, y_j), \tag{11}$$

et pour $p = 3$ on obtient

$$f\left(\sum_{i \in I} \lambda_i x_i, \sum_{j \in J} \mu_j y_j, \sum_{k \in K} \nu_k z_k\right) = \sum_{\substack{i \in I \\ j \in J \\ k \in K}} \lambda_i \mu_j \nu_k f(x_i, y_j, z_k). \tag{12}$$

4. *Cas des modules libres de type fini*

Les formules du n° précédent permettent de déterminer toutes les applications multilinéaires lorsque les modules de départ $X_1, \ldots, X_p$ sont libres de type fini (par exemple lorsqu'il s'agit d'espaces vectoriels de dimension finie sur un corps, cas de loin le plus important dans la pratique).

Examinons d'abord le cas le plus simple, celui où $p = 2$:

THÉORÈME 1. *Soient* X, Y, *et* M *des modules sur un anneau commutatif* K. *Supposons que* X *et* Y *soient libres de type fini, et soient* $(a_i)_{1 \leqslant i \leqslant m}$ *une base de* X, *et* $(b_j)_{1 \leqslant j \leqslant n}$ *une base de* Y. *Pour qu'une application* f *de* $X \times Y$ *dans* M *soit bilinéaire, il faut et il suffit qu'il existe des* $c_{ij} \in M$ *tels que l'on ait*

$$f(x, y) = \sum_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \xi_i \eta_j c_{ij} \tag{13}$$

quels que soient les vecteurs

$$x = \sum_{1 \leqslant i \leqslant m} \xi_i a_i \in X \qquad et \qquad y = \sum_{1 \leqslant j \leqslant n} \eta_j b_j \in Y;$$

s'il en est ainsi on a nécessairement

$$c_{ij} = f(a_i, b_j). \tag{14}$$

La formule (11) montre que si f est bilinéaire on a nécessairement

$$f(x, y) = \sum_{ij} \xi_i \eta_j f(a_i, b_j),$$

de sorte que f est bien donnée par une formule du type (13). Inversement supposons f donnée par (13); pour montrer que f est bilinéaire il suffit de montrer que le terme général $\xi_i\eta_j c_{ij}$ de la somme (13) est fonction bilinéaire de x et y; comme c_{ij} est indépendant de x et y, il suffit même de montrer que $\xi_i\eta_j$ est bilinéaire; or en désignant par u_i les fonctions coordonnées de X par rapport à la base (a_i), et par v_j celles de Y par rapport à la base (b_j), on a

$$\xi_i\eta_j = u_i(x)v_j(y),$$

ce qui montre bien (*Exemple* 2) que cette expression est une fonction bilinéaire de x et y.

Il reste à montrer que la relation (13), i.e.

$$f(x, y) = \sum u_i(x)v_j(y)c_{ij},$$

implique nécessairement $c_{ij} = f(a_i, b_j)$. Or on a

$$f(a_i, b_j) = \sum_{k,h} u_k(a_i)v_h(b_j)c_{kh}$$

et d'autre part

$$u_k(a_i) = \begin{cases} 0 & \text{si} \quad k \neq i, \\ 1 & \text{si} \quad k = i, \end{cases} \qquad v_h(b_j) = \begin{cases} 0 & \text{si} \quad h \neq j, \\ 1 & \text{si} \quad h = j, \end{cases}$$

(cf. par exemple § 16, n° 2, ou observer directement que

$$a_i = 0.a_1 + \cdots + 0.a_{i-1} + 1.a_i + 0.a_{i+1} + \cdots + 0.a_m);$$

le seul terme éventuellement non nul de la somme définissant $f(a_i, b_j)$ est donc le terme pour lequel $k = i$ et $h = j$, lequel se réduit visiblement à c_{ij}, ce qui achève la démonstration du Théorème.

Les éléments c_{ij} de M s'appellent les **coefficients** de f par rapport aux bases (a_i) de X et (b_j) de Y; lorsque X = Y, on prend habituellement la même base (a_i) dans X et dans Y, et alors les $c_{ij} = f(a_i, a_j)$ s'appellent les coefficients de f par rapport à *la* base (a_i) de X.

Lorsque M = K, on écrit généralement (13) sous la forme

$$f(x, y) = \sum\gamma_{ij}\xi_i\eta_j, \qquad \text{où} \qquad \gamma_{ij} = f(a_i, b_j);$$

comme

$$\xi_i\eta_j = u_i \otimes v_j(x, y)$$

la formule précédente s'écrit encore, dans le module $\mathfrak{L}(X, Y; K)$ des formes bilinéaires sur $X \times Y$, sous la forme

$$f = \sum\gamma_{ij}.u_i \otimes v_j,$$

et comme cette décomposition de f est unique on en déduit que les mn formes $u_i \otimes v_j$ constituent une *base* de $\mathfrak{L}(X, Y; K)$.

Exemple 8. Considérons la forme bilinéaire $(x|y)$ de l'*Exemple* 3. Étant donnée une base a_1, a_2, a_3 de E, on a donc

$$(x|y) = \sum (a_i|a_j) \cdot \xi_i \eta_j.$$

Cette formule se simplifie lorsque la base a_1, a_2, a_3 est **orthonormale**, i.e. formée de vecteurs de longueur 1 et deux à deux orthogonaux, autrement dit lorsqu'on calcule en **coordonnées rectangulaires**; il est en effet clair que dans ce cas on a

$$(a_i|a_j) = \begin{cases} 0 & \text{si} \quad i \neq j, \\ 1 & \text{si} \quad i = j, \end{cases}$$

et par suite il reste

$$(x|y) = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3,$$

formule qui permet de calculer le produit scalaire de deux vecteurs en coordonnées rectangulaires.

Cette formule permet aussi de calculer, en coordonnées rectangulaires, la distance de deux points P et Q dans l'espace. C'est en effet la longueur du vecteur

$$\overrightarrow{PQ} = \overrightarrow{OQ} - \overrightarrow{OP};$$

si P a pour coordonnées (x, y, z) et Q pour coordonnées (x', y', z'), le vecteur PQ (ou, plus exactement, le vecteur d'origine O équipollent à PQ) a pour composantes $x' - x$, $y' - y$, $z' - z$; sa longueur, i.e. la racine carrée de son produit scalaire par lui-même, est donc

$$\sqrt{(x' - x)^2 + (y' - y)^2 + (z' - z)^2},$$

et c'est l'expression cherchée de la distance des points P et Q en coordonnées *rectangulaires*.

Exemple 9. Dans l'*Exemple* précédent remplaçons le produit scalaire $(x|y)$ par le produit vectoriel $x \wedge y$ de l'*Exemple* 4. On a alors

$$x \wedge y = \sum \xi_i \eta_j a_i \wedge a_j;$$

or, quelle que soit la base choisie, on a les relations

$$a_i \wedge a_i = 0, \qquad a_i \wedge a_j = -a_j \wedge a_i;$$

il reste donc en fait la formule

$$x \wedge y = (\xi_2\eta_3 - \xi_3\eta_2)a_2 \wedge a_3 + (\xi_3\eta_1 - \xi_1\eta_3)a_3 \wedge a_1 + (\xi_1\eta_2 - \xi_2\eta_1)a_1 \wedge a_2;$$

si la base est orthonormale, on a en outre

$$a_2 \wedge a_3 = a_1, \qquad a_3 \wedge a_1 = a_2, \qquad a_1 \wedge a_2 = a_3,$$

et il reste la formule suivante, valable en coordonnées rectangulaires :

$$x \wedge y = (\xi_2\eta_3 - \xi_3\eta_2)a_1 + (\xi_3\eta_1 - \xi_1\eta_3)a_2 + (\xi_1\eta_2 - \xi_2\eta_1)a_3.$$

Ces formules, ainsi que celles de l'*Exemple* précédent, sont tout à fait fondamentales dans les applications pratiques (Géométrie analytique à trois dimensions, Mécanique, Physique, etc...), et reposent exclusivement sur la bilinéarité des produits scalaire et vectoriel.

On a un résultat analogue au Théorème 1 pour les applications trilinéaires :

THÉORÈME 2. *Soient* X, Y, Z *et* M *des modules sur un anneau commutatif* K. *Supposons que* X, Y *et* Z *soient libres de type fini, et soient* $(a_i)_{1\leqslant i\leqslant m}$ *une base de* X, $(b_j)_{1\leqslant j\leqslant n}$ *une base de* Y, *et* $(c_k)_{1\leqslant k\leqslant p}$ *une base de* Z. *Pour qu'une application* f *de* $X\times Y\times Z$ *dans* M *soit trilinéaire, il faut et il suffit qu'il existe des* $c_{ijk}\in M$ *tels que l'on ait*

$$f(x,y,z)=\sum_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n\\ 1\leqslant k\leqslant p}}\xi_i\eta_j\zeta_k c_{ijk} \tag{15}$$

quels que soient les vecteurs

$$x=\sum\xi_i a_i\in X,\qquad y=\sum\eta_j b_j\in Y,\qquad z=\sum\zeta_k c_k\in Z;$$

s'il en est ainsi, on a nécessairement

$$c_{ijk}=f(a_i,b_j,c_k). \tag{16}$$

La formule (12) montre que si f est trilinéaire on a effectivement la relation (15) avec des c_{ijk} donnés par (16). Inversement, si f est donnée par une relation (15), il suffit, pour établir que f est trilinéaire, de montrer qu'il en est ainsi de la fonction $\xi_i\eta_j\zeta_k c_{ijk}$, ou même, puisque c_{ijk} est indépendant de x, y, z, de la fonction $\xi_i\eta_j\zeta_k$; or en introduisant les fonctions coordonnées u_i, v_j, w_k des modules X, Y, Z par rapport aux bases considérées, il est clair que

$$\xi_i\eta_j\zeta_k=u_i(x)v_j(y)w_k(z),$$

expression qui est bien une forme trilinéaire sur $X\times Y\times Z$, à savoir la forme

$$u_i\otimes v_j\otimes w_k.$$

Pour terminer la démonstration il reste à montrer que (15) implique (16); or de (15) résulte

$$f(a_i,b_j,c_k)=\sum_{\lambda,\mu,\nu}u_\lambda(a_i)v_\mu(b_j)w_\nu(c_k)c_{\lambda\mu\nu};$$

le seul terme éventuellement non nul du second membre est celui pour lequel on a $\lambda=i$, $\mu=j$ et $\nu=k$, et alors on a $u_\lambda(a_i)=v_\mu(b_j)=w_\nu(c_k)=1$, d'où (16).

Les éléments c_{ijk} de M s'appellent, ici encore, les **coefficients** de f par rapport aux bases (a_i), (b_j) et (c_k) de X, Y et Z; lorsque $X=Y=Z$, on utilise généralement la même base (a_i) dans X, Y et Z, et on dit alors que les $c_{ijk}=f(a_i,a_j,a_k)$ sont les coefficients de f par rapport à *la* base (a_i) de X.

Lorsque $M=K$, on écrit habituellement (15) sous la forme

$$f(x,y,z)=\sum\gamma_{ijk}\xi_i\eta_j\zeta_k,\quad\text{où}\quad\gamma_{ijk}=f(a_i,b_j,c_k);$$

on a alors

$$f = \sum \gamma_{ijk}\, u_i \otimes v_j \otimes w_k$$

dans le module $\mathfrak{L}(X, Y, Z; K)$ des formes trilinéaires sur $X \times Y \times Z$, et comme la décomposition précédente est unique on voit que les mnp formes $u_i \otimes v_j \otimes w_k$ constituent une *base* du module $\mathfrak{L}(X, Y, Z; K)$.

Exemple 10. Considérons le produit mixte $(x|y|z)$ de l'*Exemple* 5; on a

$$(x|y|z) = \sum (a_i|a_j|a_k)\xi_i\eta_j\zeta_k;$$

mais on a $(a_i|a_j|a_k) = 0$ si les indices ne sont pas deux à deux distincts, et de plus

$$(a_1|a_2|a_3) = (a_2|a_3|a_1) = (a_3|a_1|a_2) = -(a_1|a_3|a_2) = -(a_2|a_1|a_3) = -(a_3|a_2|a_1);$$

par suite il reste

$$(x|y|z) = (a_1|a_2|a_3)\cdot(\xi_1\eta_2\zeta_3 + \xi_2\eta_3\zeta_1 + \xi_3\eta_1\zeta_2 - \xi_1\eta_3\zeta_2 - \xi_2\eta_1\zeta_3 - \xi_3\eta_2\zeta_1).$$

Si la base a_1, a_2, a_3 est *orthonormale*, il reste

$$(x|y|z) = \xi_1\eta_2\zeta_3 + \xi_2\eta_3\zeta_1 + \xi_3\eta_1\zeta_2 - \xi_1\eta_3\zeta_2 - \xi_2\eta_1\zeta_3 - \xi_3\eta_2\zeta_1.$$

On désigne habituellement le second membre de cette relation par la notation condensée

$$\begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}.$$

On reviendra sur cette expression dans les § suivants.

¶ *Exemple* 11. Soit T un tenseur deux fois covariant et une fois contravariant sur un module X, i.e. une forme trilinéaire sur

$$X^* \times X^* \times X;$$

supposons que X possède une base $(a_1, \ldots, a_n)$, et désignons par la notation (*) $(a^1, \ldots, a^n)$ la base duale dans le module X^* (§ 16, n° 2) — autrement dit, les formes linéaires a^i ne sont autres que les fonctions coordonnées du module X par rapport à la base (a_i) considérée. Pour calculer $T(u, v, x)$ pour, $u, v \in X^*$ et $x \in X$, on pose

$$\begin{aligned} u &= \sum \alpha_i . a^i \qquad \text{d'où} \qquad \alpha_i = u(a_i), \\ v &= \sum \beta_i . a^i \qquad \text{d'où} \qquad \beta_i = v(a_i), \\ x &= \sum \xi^i . a_i \end{aligned}$$

(*) Les indices supérieurs dans les formules qui suivent ne sont naturellement pas des exposants. Il est conforme aux traditions du Calcul Tensoriel de noter les composantes des *vecteurs* avec des indices *supérieurs*, celles des *formes linéaires* avec des indices *inférieurs*, et d'observer des conventions analogues pour les composantes des tenseurs de telle sorte qu'une simple inspection de la position des indices indique aussitôt l'espèce du tenseur considéré.

et alors il vient

$$T(u, v, x) = \sum T^{ijk} \alpha_i \beta_j \xi^k$$

avec des constantes

$$T^{ij}_k = T(a^i, a^j, a_k)$$

qu'on appelle les coefficients (ou les **composantes**, ou les **coordonnées**) du tenseur T par rapport à la base (a_i) de X.

Pour terminer, il nous reste à étendre les Théorèmes 1 et 2 aux applications p-linéaires avec p quelconque :

THÉORÈME 3. *Soient* $X_1, \ldots, X_p$ *et* M *des modules sur un anneau commutatif* K. *Supposons* $X_1, \ldots, X_p$ *libres de type fini, et, pour tout indice* h *tel que* $1 \leqslant h \leqslant p$, *soit* $(a_{h1}, a_{h2}, \ldots, a_{hn_h})$ *une base de* X_h. *Pour qu'une application* f *de* $X_1 \times \cdots \times X_p$ *dans* M *soit multilinéaire, il faut et il suffit qu'il existe des constantes*

$$c_{i_1 \ldots i_p} \in M \quad (1 \leqslant i_1 \leqslant n_1, \ldots, 1 \leqslant i_p \leqslant n_p)$$

telles que l'on ait

$$f(x_1, \ldots, x_p) = \sum_{\substack{1 \leqslant i_1 \leqslant n_1 \\ \cdots\cdots\cdots \\ 1 \leqslant i_p \leqslant n_p}} \xi_{1 i_1} \cdots \xi_{p i_p} c_{i_1 \ldots i_p} \tag{17}$$

quels que soient les vecteurs

$$x_h = \sum_{1 \leqslant i_h \leqslant n_h} \xi_{h i_h} a_{h i_h} \in X_h \quad (1 \leqslant h \leqslant p);$$

s'il en est ainsi on a nécessairement

$$c_{i_1 \ldots i_p} = f(a_{1 i_1}, \ldots, a_{p i_p}). \tag{18}$$

La formule (17) lorsque f est multilinéaire résulte évidemment de la relation (10) du n° 3. Pour montrer qu'inversement (17) représente toujours une application multilinéaire, il suffit de montrer que la fonction

$$u_{i_1 \ldots i_p}(x_1, \ldots, x_p) = \xi_{1 i_1} \cdots \xi_{p i_p}$$

est toujours p-linéaire; or en désignant par $u_{h1}, u_{h2}, \ldots, u_{hn_h}$ les fonctions coordonnées du module X_h par rapport à la base $a_{h1}, \ldots, a_{hn_h}$, il est clair que

$$u_{i_1 \ldots i_p}(x_1, \ldots, x_p) = u_{1 i_1}(x_1) \cdots u_{p i_p}(x_p),$$

en sorte que la fonction

$$u_{i_1 \ldots i_p} = u_{1 i_1} \otimes \cdots \otimes u_{p i_p}$$

est bien multilinéaire (*Exemple* 2). Enfin, on laisse au lecteur le soin de montrer que (17) implique (18) en procédant comme dans les démonstrations des Théorèmes 1 et 2.

Les éléments $c_{i_1 \ldots i_p}$ de M s'appellent les **coefficients de** f par rapport aux bases considérées dans $X_1, \ldots, X_p$; lorsque $X_1 = \cdots = X_p = X$, on utilise habituelle-

ment la même base (a_i) dans $X_1, \ldots, X_p$; la formule (17) ne change pas et (18) s'écrit

$$c_{i_1 \ldots i_p} = f(a_{i_1}, \ldots, a_{i_p}); \tag{18 bis}$$

ces éléments de M sont alors appelés les coefficients de f par rapport à *la* base (a_i) de X.

¶ *Exemple* 12. Soit T un tenseur d'espèce $\begin{pmatrix} p \\ q \end{pmatrix}$ sur un module X libre de type fini; choisi dans X une base (a_i), et dans le module dual X* la base duale (a^i); pour calculer $T(u_1, \ldots, u_p, x_1, \ldots, x_q)$ pour des vecteurs $x_j \in X$ et des formes linéaires (ou **covecteurs**) $u_i \in X^*$, on pose

$$u_h = \sum_{i_h} \alpha_{h i_h} . a^{i_h} \qquad \text{d'où} \qquad \alpha_{h i_h} = u_h(a_{i_h})$$

et

$$x_k = \sum_{j_k} \xi^{j_k}_k . a_{j_k};$$

on a alors

$$T(u_1, \ldots, u_p, x_1, \ldots, x_q) = \sum T^{i_1 \ldots i_p}_{j_1 \ldots j_q} \alpha_{1 i_1} \ldots \alpha_{p i_p} \xi^{j_1}_1 \ldots \xi^{j_q}_q$$

avec des constantes

$$T^{i_1 \ldots i_p}_{j_1 \ldots j_q} = T(a^{i_1}, \ldots, a^{i_p}, a_{j_1}, \ldots, a_{j_q})$$

qu'on appelle les **coefficients** (ou les **composantes**, ou les **coordonnées**) du tenseur T par rapport à la base (a_i) de X.

Supposons par exemple $p = 2$ et $q = 3$; on a alors

$$T(u, v, x, y, z) = \sum T^{ij}_{klm} \alpha_i \beta_j \xi^k \eta^l \zeta^m$$

pour

$$u = \sum \alpha_i a^i, \qquad v = \sum \beta_i a^i, \qquad x = \sum \xi^i a_i, \qquad y = \sum \eta^i a_i, \qquad z = \sum \zeta^i a_i$$

et les composantes T^{ij}_{klm} de T sont données par

$$T^{ij}_{klm} = T(a^i, a^j, a_k, a_l, a_m).$$

On notera qu'on peut facilement calculer les composantes d'un produit tensoriel (*Exemple* 7). Soient par exemple U un tenseur d'espèce $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ et V un tenseur d'espèce $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$; alors $T = U \otimes V$ est le tenseur d'espèce $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ donné par

$$T(u, v, w, x, y) = U(u, v, x) V(w, y);$$

les composantes

$$T^{ijk}_{lm} = T(a^i, a^j, a^k, a_l, a_m)$$

de T sont donc données par la relation

$$T_{hl}^{ijk} = U_h^{ij} . V_l^k.$$

On a évidemment des formules analogues dans le cas général.

Le lecteur débutant fera bien de ne pas se laisser impressionner par ces formules et les « débauches d'indices »; les résultats de ce § sont de simples identités algébriques ni plus ni moins profondes que la formule

$$x(y + z) = xy + xz;$$

par conséquent, toutes ces formules sont essentiellement *triviales* malgré leur aspect imposant, la seule difficulté en l'occurence étant de choisir des *notations* commodes, et non pas de construire des *raisonnements* ingénieux.

5. *Effet d'un changement de base sur les composantes d'un tenseur*

Soit X un K-module libre de type fini, et considérons deux bases

$$(a_i)_{1 \leqslant i \leqslant n}, \qquad (b_\lambda)_{1 \leqslant \lambda \leqslant n}$$

de X; on utilisera des indices latins pour tout ce qui se rapporte à la première base, et des indices grecs pour tout ce qui se rapporte à la seconde. Étant donné un tenseur T sur X, on se propose de calculer ses composantes par rapport à la base (b_λ) en fonction de ses composantes par rapport à la base (a_i).

Posons pour cela

$$b_\lambda = \sum_i \theta_\lambda^i . a_i, \qquad a_i = \sum_\lambda \rho_i^\lambda . b_\lambda,$$

de sorte que les matrices de passage (θ_λ^i) et (ρ_i^λ) sont inverses l'une de l'autre (§ 15, n° 4). Nous aurons besoin des formules faisant passer de la base duale (a^i) de X^* à la base duale (b^λ). Observons que pour tout $f \in X^*$ on a

$$f = \sum f(a_i) . a^i = \sum f(b_\lambda) . b^\lambda$$

en vertu du § 16, n° 2. Donc

$$b^\lambda = \sum_i b^\lambda(a_i) . a^i;$$

or

$$b^\lambda(a_i) = b^\lambda\left(\sum_\mu \rho_i^\mu . b_\mu\right) = \sum_\mu \rho_i^\mu . b^\lambda(b_\mu)$$

et comme

$$b^\lambda(b_\mu) = \begin{cases} 0 \text{ si } \lambda \neq \mu \\ 1 \text{ si } \lambda = \mu \end{cases}$$

par définition d'une base duale, il reste

$$b^\lambda(a_i) = \rho_i^\lambda.$$

par conséquent, il vient

$$b^\lambda = \sum_i \rho_i^\lambda . a^i$$

et de même

$$a^i = \sum_\lambda \theta_\lambda^i . b^\lambda$$

Cela dit, considérons par exemple un tenseur T d'espèce $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$; ses composantes par rapport à la première base sont les scalaires

$$T_{kl}^{ijk} = T(a^i, a^j, a^k, a_h, a_l)$$

et ses composantes par rapport à la seconde sont les scalaires

$$T_{\alpha\beta}^{\lambda\mu\nu} = T(b^\lambda, b^\mu, b^\nu, b_\alpha, b_\beta);$$

or on a vu que

$$T(u, v, w, x, y) = \sum T_{kl}^{ijk} \alpha_i \beta_j \gamma_k \xi^h \eta^l$$

pour

$$u = \sum \alpha_i a^i, \qquad v = \sum \beta_i a^i, \qquad w = \sum \gamma_i a^i$$
$$x = \sum \xi^i a_i, \qquad y = \sum \eta^i a_i,$$

remplaçant u, v, w, x, y par $b^\lambda, b^\mu, b^\nu, b_\alpha, b_\beta$ il vient donc

$$T_{\alpha\beta}^{\lambda\mu\nu} = \sum_{i,j,k,h,l} \rho_i^\lambda \rho_j^\mu \rho_k^\nu \theta_\alpha^h \theta_\beta^l T_{kl}^{ijk}, \tag{19}$$

ce qui est la relation cherchée. On aurait évidemment des formules analogues pour les autres espèces de tenseurs.

Remarque 3. La formule (19) est très souvent utilisée comme *définition* des tenseurs, et l'a même été exclusivement jusqu'à une date récente. On procède alors comme suit : on appelle tenseur trois fois covariant et deux fois contravariant un « objet géométrique » dont on ne précise pas la nature « concrète », mais dont on convient que, par rapport à chaque base (a_i) de X, il possède des « composantes » T_{kl}^{ijk}, ces composantes étant assujetties à varier avec la base choisie conformément à la formule (19). Pour montrer que cette définition équivaut à celle du texte, tout revient à prouver que si l'on associe à chaque base (a_i) de X des scalaires

$$T_{kl}^{ijk}$$

de façon que les relations (19) soient vérifiées quelles que soient les bases (a_i) et (b_λ) de X, alors il existe un et un seul tenseur T d'espèce $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ sur X dont les composantes par rapport à toute base (a_i) de X sont justement les T_{kl}^{ijk} associés à cette base. Pour cela, choisissons une base (a_i) une fois pour toutes, et à l'aide de cette base particulière construisons le tenseur

$$T(u, v, w, x, y) = \sum T_{kl}^{ijk} \alpha_i \beta_j \gamma_k \xi^h \eta^l$$

tout revient à montrer que ses composantes par rapport à toute autre base (b_λ) de X sont les scalaires

$$T^{\lambda\mu\nu}_{\alpha\beta\gamma}$$

attachés à celle-ci; or, d'après les calculs précédents, les composantes de T par rapport à la base (b_λ) sont les scalaires

$$\sum_{i,j,k,h,l} \rho^\lambda_i \rho^\mu_j \rho^\nu_k \theta^h_\alpha \theta^l_\beta \theta^\gamma_l T^{ijk}_{hl},$$

et comme par hypothèse ces expressions sont justement les $T^{\lambda\mu\nu}_{\alpha\beta\gamma}$ attachés à la base (b_λ), notre assertion est établie.

On peut aussi regarder ce raisonnement sous la forme suivante : les formules (19) expriment qu'étant donnés des vecteurs $x, y \in X$ et des covecteurs $u, v, w \in X^*$, l'expression

$$\sum T^{ijk}_{hl} \alpha_i \beta_j \gamma_k \xi^h \eta^l,$$

calculée à l'aide des coordonnées de x, y, u, v, w par rapport à cette base, est en fait *indépendante du système de coordonnées choisi pour la définir*. C'est précisément parce que les formules analogues à (19) permettent de définir des objets ayant une signification indépendante des systèmes de coordonnées utilisés pour les construire que la théorie des tenseurs s'est finalement introduite en Physique et en Mathématiques. L'idée fondamentale est que les systèmes de coordonnées ne sont que des instruments pour étudier des objets ayant une signification intrinsèque, et que seuls ces objets présentent un intérêt quelconque.

§ 22. Applications bilinéaires alternées

1. *Applications bilinéaires alternées*

Soient X et M des modules sur un anneau commutatif K. On dit qu'une application bilinéaire

$$f : X \times X \to M$$

est **alternée** si l'on a

(1) $$f(x, x) = 0 \quad \text{pour tout } x \in X.$$

Quels que soient $x, y \in X$ on a alors

$$0 = f(x + y, x + y) = f(x, x) + f(x, y) + f(y, x) + f(y, y) = f(x, y) + f(y, x)$$

et par suite une application bilinéaire alternée satisfait à l'identité

(2) $$f(y, x) = -f(x, y) \quad \text{quels que soient } x, y \in X.$$

Remarque 1. Il est clair que réciproquement (2) implique

2. $f(x, x) = 0$ pour tout $x \in X$.

Si, pour $m \in M$, la relation $2m = 0$ implique $m = 0$ (c'est par exemple le cas si K est un corps de « caractéristique » différente de 2, car alors 2 est inversible dans K : voir § 30, n° 6), on voit donc que la relation (2) caractérise les applications bilinéaires alternées. Par contre, si $M = K$ est l'anneau des entiers modulo 2, la relation (2), qui s'écrit alors $f(y, x) = f(x, y)$, ne suffit pas à impliquer que f soit alternée. Cette situation se rencontre rarement dans la pratique élémentaire.

Il est clair que toute combinaison linéaire d'applications bilinéaires alternées est encore une application bilinéaire alternée; autrement dit, les applications bilinéaires *alternées* forment un *sous-module* du module $\mathfrak{L}(X, X; M)$ de toutes les applications bilinéaires de $X \times X$ dans M.

Exemple 1. Prenons $K = \mathbf{R}$ et pour X l'espace usuel à trois dimensions; alors l'application f de $X \times X$ dans X donnée par

$$f(x, y) = x \wedge y$$

(produit vectoriel; cf. § 21, *Exemple* 4) est bilinéaire alternée.

Exemple 2. Si f est une application bilinéaire de $X \times X$ dans M, l'application g donnée par

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

est bilinéaire alternée. En particulier, si u et v sont des formes linéaires sur X, l'application

$$g(x, y) = u(x)v(y) - u(y)v(x)$$

de $X \times X$ dans K est une forme bilinéaire alternée sur $X \times X$; on l'appelle le **produit extérieur des formes linéaires** u **et** v et on la désigne par la notation

$$u \wedge v.$$

Prenons par exemple $K = \mathbf{R}$ et X comme dans l'*Exemple* précédent, et

$$u(x) = (a|x), \qquad v(x) = (b|x)$$

où a et b sont des vecteurs fixes. On a alors, pour $g = u \wedge v$, la formule

$$g(x, y) = (a \wedge b|x|y),$$

ce qui veut dire encore qu'on a l'identité

$$(a|x)\ (b|y) - (a|y)\ (b|x) = (a \wedge b|x|y) = (a|b|x \wedge y).$$

On laisse au lecteur le soin d'établir ce résultat soit par des raisonnements géométriques directs, soit en calculant en coordonnées rectangulaires (on peut même calculer dans un système de coordonnées quelconque).

2. *Cas des modules libres de type fini*

On a pour les applications bilinéaires alternées un résultat analogue au Théorème 1 du § 21 :

THÉORÈME 1. *Soient* X *et* M *des modules sur un anneau commutatif* K, *et supposons* X *libre de type fini. Soit* $(a_i)_{1 \leqslant i \leqslant n}$ *une base de* X. *Pour qu'une application bilinéaire* f *de* $X \times X$ *dans* M *soit alternée, il faut et il suffit que ses coefficients*

$$e_{ij} = f(a_i, a_j)$$

par rapport à la base (a_i) *vérifient les relations*

$$e_{ii} = 0, \qquad e_{ij} + e_{ji} = 0; \tag{3}$$

on a alors

$$f(x, y) = \sum_{i<j} c_{ij}(\xi_i\eta_j - \xi_j\eta_i) \tag{4}$$

quels que soient les vecteurs

$$x = \sum_i \xi_i a_i, \qquad y = \sum \eta_i a_i.$$

En exprimant (1) pour $x = a_i$ et (2) pour $x = a_i$, $y = a_j$ il est clair qu'on obtient les relations (3). Supposons-les inversement vérifiées; dans la formule

$$f(x, y) = \sum_{1 \leqslant i,\, j \leqslant n} \xi_i\eta_j c_{ij}$$

qui résulte du § 21, Théorème 1, les termes pour lesquels $i = j$ sont nuls, et en groupant les autres d'après les grandeurs relatives de i et j on trouve

$$f(x, y) = \sum_{i<j} \xi_i\eta_j c_{ij} + \sum_{i>j} \xi_i\eta_j c_{ij};$$

dans la seconde somme, remplaçons les lettres i et j par les lettres j et i (il s'agit donc d'un simple changement de *notations*); il vient

$$f(x, y) = \sum_{i<j} \xi_i\eta_j c_{ij} + \sum_{i<j} \xi_j\eta_i c_{ji} = \sum_{i<j} \xi_i\eta_j c_{ij} - \sum_{i<j} \xi_j\eta_i c_{ij}$$

d'après (3); on obtient finalement la relation (4), et celle-ci montre que f est alternée, car il est clair que la différence $\xi_i\eta_j - \xi_j\eta_i$ est nulle lorsque $x = y$, ce qui achève la démonstration.

Remarque 2. En notant u_i les fonctions coordonnées de X par rapport à la base (a_i), on voit d'après l'*Exemple* 2 que

$$\xi_i\eta_j - \xi_j\eta_i = u_i(x)u_j(y) - u_j(x)u_i(y) = u_i \wedge u_j\,(x, y);$$

par suite (4) s'écrit encore

$$f(x, y) = \sum_{i<j} u_{ij}(x, y)c_{ij}$$

où l'on pose $u_{ij} = u_i \wedge u_j$, et cette décomposition de f est unique car elle implique $c_{ij} = f(a_i, a_j)$.

Si en particulier M = K, la formule précédente s'écrit

$$f = \sum_{i<j} \gamma_{ij}u_{ij} \qquad \text{où} \qquad \gamma_{ij} = f(a_i, a_j);$$

on en déduit que les $n(n-1)/2$ formes bilinéaires alternées $u_{ij} = u_i \wedge u_j$ $(1 \leqslant i < j \leqslant n)$ forment une *base* du module des formes bilinéaires alternées sur $X \times X$.

En utilisant la notation

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

déjà introduite au § 15, n° 3, on voit que si f est une forme bilinéaire alternée sur X on a

$$f(x, y) = \sum_{i<j} \gamma_{ij} \begin{vmatrix} \xi_i & \eta_i \\ \xi_j & \eta_j \end{vmatrix} \quad \text{où} \quad \gamma_{ij} = f(a_i, a_j). \tag{5}$$

Lorsque $n = 1$, il est clair que f est identiquement nulle, parce que

$$f(x, y) = f(\xi_1 a_1, \eta_1 a_1) = \xi_1 \eta_1 f(a_1, a_1) = 0.$$

Lorsque $n = 2$, la formule (5) se réduit à

$$f(x, y) = \gamma_{12} \cdot \begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix} \quad \text{où} \quad \gamma_{12} = f(a_1, a_2); \tag{6}$$

la fonction

$$\mathrm{D}(x, y) = \begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix} = \xi_1 \eta_2 - \xi_2 \eta_1$$

s'appelle le **déterminant des vecteurs x et y par rapport à la base (a_1, a_2) de X**; c'est une forme bilinéaire alternée sur X telle que

$$\mathrm{D}(a_1, a_2) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1,$$

et cette propriété *caractérise* D, car (5) s'écrit

$$f = f(a_1, a_2) \cdot \mathrm{D}$$

et implique donc $f = \mathrm{D}$ si $f(a_1, a_2) = 1$.

Lorsque $n = 3$, la relation (5) s'écrit

$$f(x, y) = \gamma_{23} \cdot \begin{vmatrix} \xi_2 & \eta_2 \\ \xi_3 & \eta_3 \end{vmatrix} + \gamma_{31} \cdot \begin{vmatrix} \xi_3 & \eta_3 \\ \xi_1 & \eta_1 \end{vmatrix} + \gamma_{12} \cdot \begin{vmatrix} \xi_1 & \eta_1 \\ \xi_2 & \eta_2 \end{vmatrix}, \tag{7}$$

et dans ce cas les formes bilinéaires alternées sur $X \times X$ dépendent de trois constantes arbitraires

$$\gamma_{23} = f(a_2, a_3), \qquad \gamma_{31} = f(a_3, a_1), \qquad \gamma_{12} = f(a_1, a_2).$$

Remarque 3. Lorsque $K = R$ et que X est l'espace usuel à trois dimensions, supposons la base $(a_i)_{1 \leq i \leq 3}$ de X orthonormale; considérant le vecteur

$$u = \gamma_{23} a_1 + \gamma_{31} a_2 + \gamma_{12} a_3 \tag{8}$$

la relation (7) signifie alors que $f(x, y)$ est le produit scalaire de u par le vecteur $x \wedge y$, autrement dit (§ 21, *Exemple* 10) que

$$f(x, y) = (u | x | y).$$

Réciproquement il est clair que toute fonction f donnée par une formule de ce type est une forme bilinéaire alternée sur X.

On peut ainsi identifier f au vecteur u, ce qui explique pourquoi les formes bilinéaires alternées n'interviennent généralement pas en Géométrie élémentaire ou en Physique. Toutefois il faut observer que l'identification de f au vecteur (8) n'a de sens qu'en coordonnées *rectangulaires* (et que la notion de coordonnées rectangulaires suppose le choix d'une unité de longueur); plus précisément, le vecteur (8) est indépendant de la base (a_i) dans la mesure où l'on passe d'une base *orthonormale* à une base *orthonormale*; mais si l'on autorise (a_i) à varier dans l'ensemble de toutes les bases de X, alors le vecteur (8) dépend non seulement de f mais aussi de la base (a_i) considérée. Cela provient du fait que les formules de changement de coordonnées pour un tenseur d'espèce $\binom{0}{2}$ ne sont pas les mêmes que pour un vecteur, i.e. pour un tenseur d'espèce $\binom{1}{0}$.

3. *Applications trilinéaires alternées*

Étant donnés des modules X et M sur un anneau commutatif K, on dit qu'une application trilinéaire

$$f : \mathrm{X} \times \mathrm{X} \times \mathrm{X} \to \mathrm{M}$$

est alternée si l'expression $f(x, y, z)$ est nulle dès que deux des vecteurs x, y, z sont égaux, autrement dit si l'on a

$$f(x, y, y) = f(x, y, x) = f(x, x, z) = 0 \tag{8}$$

quels que soient x, y, z. Il s'ensuit alors que pour tout $a \in \mathrm{X}$, les fonctions bilinéaires $f(a, y, z)$, $f(x, a, z)$ et $f(x, y, a)$ sont alternées; par suite on a les relations

$$f(x, y, z) = -f(x, z, y), \qquad f(x, y, z) = -f(z, y, x), \qquad f(x, y, z) = -f(y, x, z),$$

ou, ce qui revient visiblement au même, les relations

$$f(x, y, z) = f(y, z, x) = f(z, x, y) = -f(x, z, y) = -f(y, x, z) = -f(z, y, x) \tag{9}$$

Si la relation $2m = 0$, pour un $m \in \mathrm{M}$, implique $m = 0$, alors les relations (9) impliquent les relations (8), et caractérisent donc les fonctions trilinéaires alternées.

Exemple 3. Si l'on prend $\mathrm{K} = \mathbf{R}$ et pour X l'espace usuel à trois dimensions, le produit mixte $(x|y|z)$ est une forme trilinéaire alternée sur X.

Exemple 4. X et K étant arbitraires, choisissons une application trilinéaire g et posons

$$\begin{aligned} f(x, y, z) = {} & g(x, y, z) + g(y, z, x) + g(z, x, y) \\ & - g(x, z, y) - g(y, x, z) - g(z, y, x); \end{aligned}$$

alors f est une application trilinéaire alternée comme on le voit aussitôt.

En particulier, soient u, v, w des formes linéaires sur X; alors

$$\begin{aligned} f(x, y, z) = {} & u(x)v(y)w(z) + u(y)v(z)w(x) + u(z)v(x)w(y) \\ & - u(x)v(z)w(y) - u(y)v(x)w(z) - u(z)v(y)w(x) \end{aligned}$$

est une forme trilinéaire alternée sur $X \times X \times X$; on l'appelle le produit extérieur des formes linéaires u, v, w et on la désigne par la notation

$$u \wedge v \wedge w.$$

On a $u \wedge v \wedge w = 0$ si u, v, w ne sont pas deux à deux distinctes, et de plus

$$u \wedge v \wedge w = v \wedge w \wedge u = w \wedge u \wedge v$$
$$= -u \wedge w \wedge v = -v \wedge u \wedge w = -w \wedge v \wedge u.$$

Exemple 5. Soient u une forme linéaire et f une forme bilinéaire alternée sur un module X; alors la fonction

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

est une forme trilinéaire alternée sur $X \times X \times X$; on l'appelle le produit extérieur de la forme linéaire u par la forme bilinéaire alternée f, et on la désigne par la notation

$$u \wedge f.$$

Il est clair que si u, v, w sont des formes linéaires sur X on a

$$u \wedge v \wedge w = u \wedge (v \wedge w).$$

De même, on désigne par

$$f \wedge u$$

la forme trilinéaire alternée

$$f(x, y)u(z) + f(y, z)u(x) + f(z, x)u(y);$$

on a

$$u \wedge f = f \wedge u$$

(alors qu'on avait $u \wedge v = -v \wedge u$ lorsque u et v sont deux formes linéaires)

4. *Développement par rapport à une base*

On a pour les applications trilinéaires alternées le résultat suivant, analogue au Théorème 1 :

THÉORÈME 2. *Soient* X *et* M *des modules sur un anneau commutatif. Supposons* X *libre de type fini et soit* $(a_i)_{1 \leq i \leq n}$ *une base de* X. *Pour qu'une application trilinéaire* f *de* $X \times X \times X$ *dans* M *soit alternée, il faut et il suffit que ses coefficients*

$$c_{ijk} = f(a_i, a_j, a_k)$$

par rapport à la base considérée vérifient les relations

$$c_{ij} = c_{ji} = c_{ik} = 0 \tag{10}$$

$$c_{ijk} = c_{jki} = c_{kij} = -c_{kji} = -c_{jik} = -c_{ikj} \tag{11}$$

on a alors

$$f(x, y, z) = \sum_{i<j<k} c_{ijk} \cdot \begin{vmatrix} \xi_i & \eta_i & \zeta_i \\ \xi_j & \eta_j & \zeta_j \\ \xi_k & \eta_k & \zeta_k \end{vmatrix} \tag{12}$$

quels que soient les vecteurs

$$x = \sum \xi_i a_i, \qquad y = \sum \eta_i a_i, \qquad z = \sum \zeta_i a_i.$$

Dans la formule (12), on utilise la notation

$$\begin{vmatrix} \xi_i & \eta_i & \zeta_i \\ \xi_j & \eta_j & \zeta_j \\ \xi_k & \eta_k & \zeta_k \end{vmatrix} = \xi_i\eta_j\zeta_k + \xi_j\eta_k\zeta_i + \xi_k\eta_i\zeta_j - \xi_i\eta_k\zeta_j - \xi_j\eta_i\zeta_k - \xi_k\eta_j\zeta_i, \tag{13}$$

analogue à celle des déterminants d'ordre 2, et déjà introduite au § 21, Exemple 10.

Pour établir le Théorème 2, on doit d'abord montrer que l'on a (10) et (11) pour toute application trilinéaire alternée f ; ce qu'on peut faire en écrivant les relations (8) et (9) pour $x = a_i, y = a_j, z = a_k$.

Supposons inversement (10) et (11) vérifiées; dans la formule

$$f(x, y, z) = \sum_{1 \leqslant i, j, k \leqslant n} c_{ijk}\xi_i\eta_j\zeta_k$$

qui résulte du § 21, Théorème 2, les seuls termes éventuellement non nuls sont ceux pour lesquels les indices i, j, k sont deux à deux distincts. En classant les triplets (i, j, k) d'après les grandeurs relatives de i, j, k on obtient donc une décomposition en six sommes partielles, à savoir

$$f(x, y, z) = \sum_{i<j<k} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{j<k<i} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{k<i<j} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{i<k<j} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{j<i<k} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{k<j<i} c_{ijk}\xi_i\eta_j\zeta_k.$$

Pour chacune de ces sommes partielles, il est possible de changer les notations de telle sorte que la somme en question soit étendue aux triplets (i, j, k) tels que $i < j < k$ (par exemple, dans la seconde somme, on doit pour ce faire remplacer j par i, k par j et i par k) ; on trouve ainsi

$$f(x, y, z) = \sum_{i<j<k} c_{ijk}\xi_i\eta_j\zeta_k + \sum_{i<j<k} c_{kij}\xi_k\eta_i\zeta_j + \sum_{i<j<k} c_{jki}\xi_j\eta_k\zeta_i + \sum_{i<j<k} c_{ikj}\xi_i\eta_k\zeta_j + \sum_{i<j<k} c_{jik}\xi_j\eta_i\zeta_k + \sum_{i<j<k} c_{kji}\xi_k\eta_j\zeta_i;$$

tenant compte des relations (11) et groupant les termes d'indices i, j, k des six sommes précédentes, on trouve

$$f(x, y, z) = \sum_{i<j<k} c_{ijk}(\xi_i\eta_j\zeta_k + \xi_j\eta_k\zeta_i + \xi_k\eta_i\zeta_j - \xi_i\eta_k\zeta_j - \xi_j\eta_i\zeta_k - \xi_k\eta_j\zeta_i),$$

ce qui est précisément la formule (12).

Il reste à vérifier que f est effectivement alternée; pour cela il suffit évidemment de prouver que la forme trilinéaire

$$u_{ijk}(x, y, z) = \xi_i\eta_j\zeta_k + \xi_j\eta_k\zeta_i + \xi_k\eta_i\zeta_j - \xi_i\eta_k\zeta_j - \xi_j\eta_i\zeta_k - \xi_k\eta_j\zeta_i$$

est alternée; mais en notant u_i la i^e fonction coordonnée du module X par rapport à la base (a_i), on a

$$\begin{aligned} u_{ijk}(x, y, z) = {} & u_i(x)u_j(y)u_k(z) + u_i(y)u_j(z)u_k(x) + u_i(z)u_j(x)u_k(y) \\ & - u_i(x)u_j(z)u_k(y) - u_i(y)u_j(x)u_k(z) - u_i(z)u_j(y)u_k(x), \end{aligned}$$

autrement dit

$$u_{ijk} = u_i \wedge u_j \wedge u_k,$$

ce qui prouve (*Exemple* 5) que l'expression en question est alternée. Le Théorème 2 est donc démontré.

Étant donnée une matrice carrée d'ordre trois

$$\mathrm{A} = \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}$$

à coefficients dans un anneau commutatif K, on appelle **déterminant de** A l'élément

$$ab'c'' + bc'a'' + ca'b'' - ac'b'' - ba'c'' - cb'a''$$

de K; on le désigne soit par la notation

$$\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}$$

utilisée dans l'énoncé du Théorème 2, soit par la notation

$$\text{dét(A)}.$$

On observera que

$$\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} = a \cdot \begin{vmatrix} b' & b'' \\ c' & c'' \end{vmatrix} - a' \cdot \begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} + a'' \cdot \begin{vmatrix} b & b' \\ c & c' \end{vmatrix}.$$

Exemple 6. Lorsque $n < 2$, les indices i, j, k ne peuvent jamais être deux à deux distincts; les coefficients c_{ijk} sont donc *tous* nuls, et par suite la seule application trilinéaire alternée dans ce cas est $f = 0$.

Exemple 7. Supposons $n = 3$; alors (12) se réduit à

$$f(x, y, z) = c_{123} \cdot \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix} \quad \text{où} \quad c_{123} = f(a_1, a_2, a_3).$$

L'expression

$$D(x, y, z) = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}$$

s'appelle le **déterminant des vecteurs** x, y, z **par rapport à la base** (a_1, a_2, a_3) — sa valeur dépend effectivement du choix de la base. La fonction $D(x, y, z)$ est une forme trilinéaire sur X, pour laquelle on a

$$D(a_1, a_2, a_3) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1;$$

et cette relation caractérise D parmi les formes trilinéaires alternées sur X; en effet, pour une telle forme f, on a d'après le Théorème 2

$$f = f(a_1, a_2, a_3) \cdot D$$

et par suite $f = D$ si f est égale à 1 sur les vecteurs de base.

Exemple 8. Prenons $K = \mathbf{R}$ et pour X l'espace usuel à trois dimensions; considérons le produit mixte $(x|y|z)$; c'est une forme trilinéaire alternée sur X, on a donc

$$(x|y|z) = (a_1|a_2|a_3) \cdot \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix}$$

par rapport à n'importe quelle base de X. On retrouve ainsi le résultat du § 21, *Exemple* 10.

L'étude des formes trilinéaires alternées et des déterminants d'ordre trois comporte beaucoup d'autres résultats que les précédents; mais nous ne les énoncerons pas ici attendu qu'ils se généralisent tous aux formes p-linéaires alternées et aux déterminants d'ordre quelconque dont traiteront les deux prochains §§.

§ 23. Applications multilinéaires alternées

1. *La signature d'une permutation*

Rappelons (§ 7, *Exemple* 4) que pour tout entier p on désigne par $\mathfrak{S}_p$ le groupe des permutations de l'ensemble $\{1, 2, \ldots, p\}$. Parmi ces permutations figurent les transpositions (§ 7, n° 5), qui consistent à échanger deux entiers consécutifs i et $i+1$ sans modifier les autres; et on a vu au § 7, n° 5 que toute permutation $\sigma \in \mathfrak{S}_p$ peut se mettre sous la forme

$$\sigma = \tau_1 \circ \cdots \circ \tau_r \tag{1}$$

d'un produit de transpositions; bien entendu la décomposition (1) de σ n'est pas unique : si par exemple τ est une transposition, on a $\tau \circ \tau = e$ et par suite

$$\tau = \tau \circ \tau \circ \tau = \tau \circ \tau \circ \tau \circ \tau \circ \tau = \cdots$$

Toutefois nous allons montrer dans ce n° que, lorsqu'on passe d'une décomposition (1) de σ à une autre, la *parité* de l'entier r ne change pas — autrement dit que si σ peut s'écrire comme produit d'un nombre pair (resp. impair) de transpositions, alors *toute* décomposition de σ en produit de transpositions comporte un nombre pair (resp. impair) de facteurs.

Supposons provisoirement ce résultat établi. Alors, dans la décomposition (1), l'entier

$$(-1)^r$$

dépend uniquement de σ, et non de la façon dont on a écrit σ comme produit de transpositions; on peut donc définir sur le groupe $\mathfrak{S}_p$ une fonction ε, dont les valeurs sont les entiers $+1$ et -1, et telle que

$$\varepsilon(\sigma) = (-1)^r \tag{2}$$

lorsque σ est produit de r transpositions. On a évidemment

$$\varepsilon(\sigma) = -1 \text{ si } \sigma \text{ est une transposition;} \tag{3}$$

d'autre part, si deux permutations σ' et σ'' peuvent s'écrire comme produits de r' et s

transpositions respectivement, il est clair que $\sigma' \circ \sigma''$ s'écrit comme produit de $r+s$ transpositions; on a donc

$$\mathfrak{p}(\sigma')\mathfrak{p}(\sigma'') = \mathfrak{p}(\sigma \circ \sigma'') \tag{4}$$

quelles que soient les permutations σ' et σ''.

On notera que l'ensemble $\{-1, +1\}$ n'est autre que le groupe multiplicatif $\mathbf{Z}^*$ des éléments inversibles de l'anneau $\mathbf{Z}$ des entiers (§ 8, *Remarque* 1); les formules (3) et (4) expriment donc que l'application

$$\mathfrak{p} : \mathfrak{S}_p \to \mathbf{Z}^* \tag{5}$$

est un *homomorphisme de groupes, prenant la valeur* -1 *sur les transpositions.*

Inversement, supposons construit un tel homomorphisme; la relation (1) donne alors

$$\mathfrak{p}(\sigma) = \mathfrak{p}(\tau_1)\dots\mathfrak{p}(\tau_r) = (-1)\dots(-1) = (-1)^r,$$

et comme le premier membre ne dépend que de σ on en déduit que la parité de r est indépendante de la décomposition (1) de σ en produit de transpositions.

Nous voyons donc que, pour montrer que la parité de r dans la décomposition (1) est toujours la même, *tout revient à construire un homomorphisme* (5) *égal à* -1 *sur chaque transposition.* Nous allons y parvenir à l'aide de la notion suivante.

Soient X un ensemble arbitraire, M un groupe additif, p un entier strictement positif, et considérons une application

$$f : \mathrm{X}^p \to \mathrm{M},$$

i.e. une fonction $f(x_1, \dots, x_p)$ de p variables $x_i \in \mathrm{X}$, à valeurs dans M. On dit que f est **antisymétrique** si l'on a

$$f(x_1, \dots, x_{i-1}, x_{i+1}, x_i, x_{i+2}, \dots, x_p) = -f(x_1, \dots, x_p), \tag{6}$$

autrement dit si

$$f(x_{\tau(1)}, \dots, x_{\tau(p)}) = -f(x_1, \dots, x_p) \tag{7}$$

pour toute *transposition* τ.

Considérons donc une telle fonction; nous allons montrer qu'on a alors

$$f(x_{\sigma(1)}, \dots, x_{\sigma(p)}) = (-1)^r \cdot f(x_1, \dots, x_p) \tag{8}$$

toutes les fois que la permutation $\sigma \in \mathfrak{S}_p$ peut s'écrire comme produit de r transpositions.

Le cas $r = 1$ se réduisant à la définition des fonctions antisymétriques, nous allons raisonner par récurrence sur r. Si σ est produit de r transpositions, on peut écrire

$$\sigma = \tau \circ \omega$$

où τ est une transposition, et où ω est produit de $r-1$ transpositions; d'après l'hypo-

thèse de récurrence, on aura donc

$$f(y_{\omega(1)}, \ldots, y_{\omega(p)}) = (-1)^{r-1} f(y_1, \ldots, y_p) \tag{9}$$

quels que soient les $y_i \in X$. Écrivons cette relation lorsque

$$y_1 = x_{\tau(1)}, \ldots, y_p = x_{\tau(p)};$$

on a $y_i = x_{\tau(i)}$, et en remplaçant i par $\omega(i)$ il vient donc

$$y_{\omega(i)} = x_{\tau(\omega(i))} = x_{\sigma(i)};$$

donc (9) s'écrit

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) = (-1)^{r-1} f(x_{\tau(1)}, \ldots, x_{\tau(p)});$$

mais comme τ est une transposition, le second membre de cette relation est égal d'après (7) à celui de la relation (8), qui est donc établie.

La relation (8) va nous permettre de montrer que la parité de r ne dépend que de σ. Comme le premier membre de (8) ne dépend que de σ, si l'on a deux décompositions de σ en produits de r et s transpositions, on voit que l'on a

$$(-1)^r f(x_1, \ldots, x_p) = (-1)^s f(x_1, \ldots, x_p)$$

pour *toute* fonction antisymétrique f; pour en déduire que $(-1)^r = (-1)^s$, il suffit d'être dans une situation où l'on pourra « simplifier » par $f(x_1, \ldots, x_p)$, ce qui sera par exemple le cas si f est à valeurs dans $\mathbf{Z}$ et s'il existe $x_1, \ldots, x_p \in X$ tels que $f(x_1, \ldots, x_p) \neq 0$. Autrement dit, tout revient maintenant à construire, sur un ensemble X convenablement choisi, une fonction antisymétrique *non identiquement nulle* à valeurs dans $\mathbf{Z}$.

Prenons pour cela $X = \mathbf{Z}$ et

$$f(x_1, \ldots, x_p) = \prod_{1 \leq i < j \leq p} (x_i - x_j); \tag{10}$$

si les x_i sont deux à deux distincts on a évidemment $f(x_1, \ldots, x_p) \neq 0$; il reste donc à faire voir que f est antisymétrique.

Or supposons qu'on échange x_k et x_{k+1} pour un indice donné k. Dans le second membre de (10), les termes $x_i - x_j$ ne seront modifiés que si l'on a $i = k$, ou $i = k+1$, ou $j = k$, ou $j = k+1$; la contribution de ces termes à f est le produit

$$\begin{aligned}(x_k - x_{k+1}) \cdot &[(x_k - x_{k+2}) \ldots (x_k - x_p)] \cdot [(x_{k+1} - x_{k+2}) \ldots (x_{k+1} - x_p)] \\ \cdot &[(x_1 - x_k) \ldots (x_{k-1} - x_k)] \cdot [(x_1 - x_{k+1}) \ldots (x_{k-1} - x_{k+1})];\end{aligned}$$

cela dit, lorsqu'on échange x_k et x_{k+1}, le premier facteur est multiplié par -1, le second et le troisième produits partiels s'échangent, de même que la quatrième et le cinquième, de sorte que le produit total est simplement multiplié par -1.

La fonction (10) est donc bien antisymétrique, et en comparant (9) et (8) on voit qu'on a en définitive démontré le résultat suivant :

THÉORÈME 1. *Pour tout entier $p \geqslant 1$, il existe un et un seul homomorphisme*

$$\mathfrak{p} : \mathfrak{S}_p \to \mathbf{Z}^*$$

tel que l'on ait $\mathfrak{p}(\sigma) = -1$ pour toute transposition σ. On a $\mathfrak{p}(\sigma) = (-r)^r$ si la permutation σ est produit de r transpositions.

En outre, étant donnés un ensemble X, *un groupe additif* M, *et une application antisymétrique*

$$f : \mathrm{X}^p \to \mathrm{M},$$

on a la relation

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) = \mathfrak{p}(\sigma) \cdot f(x_1, \ldots, x_p)$$

quels que soient les $x_i \in \mathrm{X}$ et la permutation $\sigma \in \mathfrak{S}_p$.

Le nombre $\mathfrak{p}(\sigma)$ s'appelle la **signature** de la permutation σ; on utilise aussi fréquemment le mot **parité** au lieu du mot signature et la notation ε_σ au lieu de $\mathfrak{p}(\sigma)$. On dit qu'une permutation est **paire** si sa signature est $+1$, et **impaire** si sa signature est -1. Les permutations paires forment un sous-groupe invariant de $\mathfrak{S}_p$, puisque leur ensemble est le noyau de l'homomorphisme $\mathfrak{p}$ (cf. § 7, *Remarque* 7).

Remarque 1. Reprenons la forme (10); on a

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) = \prod_{i<j} (x_{\sigma(i)} - x_{\sigma(j)});$$

les différences $x_{\sigma(i)} - x_{\sigma(j)}$ qu'on trouve ici sont les mêmes que les différences $x_h - x_k$ $(k < h)$ à ceci près qu'on n'a pas toujours $\sigma(i) < \sigma(j)$; on voit donc que chaque couple i, j tel que l'on ait

$$i < j, \qquad \sigma(i) > \sigma(j)$$

contribue un facteur -1 au calcul de la signature de σ. Le nombre de ces couples s'appelle **le nombre d'inversions** de σ; en le notant $\mathrm{I}(\sigma)$, on a donc

$$\mathfrak{p}(\sigma) = (-1)^{\mathrm{I}(\sigma)}.$$

Par exemple prenons $p = 6$ et la permutation qui applique 1, 2, 3, 4, 5, 6 sur

$$2, 4, 3, 6, 5, 1;$$

les couples i, j tels que $i < j$, $\sigma(i) > \sigma(j)$ sont alors

$$1,6 ; \quad 2,3 ; \quad 2,6 ; \quad 3,5 ; \quad 4,5 ; \quad 4,6 ; \quad 5,6 ;$$

de sorte que le nombre d'inversions est 7, et la signature -1.

Pour p quelconque, considérons une **permutation circulaire** des entiers $1, 2, \ldots, p$, i.e. une permutation σ donnée par

$$\sigma(1) = k, \ldots, \sigma(p - k + 1) = p, \qquad \sigma(p - k + 2) = 1, \ldots, \sigma(p) = k - 1$$

où k est un entier compris entre 1 et p. Les couples i, j tels que l'on ait $i < j$

et $\sigma(i) > \sigma(j)$ sont ceux pour lesquels on a

$$1 \leqslant i \leqslant p-k+1, \qquad p-k+2 \leqslant j \leqslant p;$$

pour ces couples, l'entier i peut prendre $p-k+1$ valeurs distinctes, et l'entier j peut en prendre $k-1$; donc on a

$$I(\sigma) = (k-1)(p-k+1) = (k-1)(p+1) - k(k-1)$$

et comme $k(k-1)$ est un entier pair dans tous les cas il vient

$$\varepsilon(\sigma) = (-1)^{(k-1)(p+1)}.$$

Par exemple, pour $p = 3$ (et plus généralement pour p impair), les permutations circulaires sont paires.

2. *Antisymétrisation d'une fonction de plusieurs variables*

Soient X un ensemble, M un groupe additif, $p \geqslant 1$ un entier, et f une application de X^p dans M. On appelle **antisymétrisée** de f l'application g de X^p dans M donnée par

$$g(x_1, \ldots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) \cdot f(x_{\sigma(1)}, \ldots, x_{\sigma(p)}), \tag{11}$$

la somme figurant au second membre étant étendue à toutes les permutations $\sigma \in \mathfrak{S}_p$. Nous allons montrer d'une part que g *est antisymétrique*, d'autre part que l'on a

$$g(x_1, \ldots, x_p) = 0 \quad \text{si} \quad x_1, \ldots, x_p \text{ ne sont pas tous distincts.} \tag{12}$$

Exemple 1. Si $p = 3$, la fonction g est donnée par

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

et les résultats annoncés sont à peu près évidents.

Pour établir ces résultats, faisons opérer le groupe $\mathfrak{S}_p$ sur l'ensemble X^p en posant (cf. § 7, n° 11, *Exemple* 21).

$$\sigma(x) = (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(p)}) \quad \text{pour} \quad x = (x_1, \ldots, x_p) \in X^p.$$

La formule (11) s'écrit alors

$$g(x) = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) \cdot f(\sigma^{-1}(x)).$$

Soit ω une permutation quelconque; on a alors

$$g(\omega(x)) = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma) \cdot f(\sigma^{-1}(\omega(x))); \tag{13}$$

or, comme l'application $\sigma \to \omega \circ \sigma$ de $\mathfrak{S}_p$ dans $\mathfrak{S}_p$ est, pour ω donné, une *bijection*,

il est clair que pour toute fonction φ définie sur le groupe $\mathfrak{S}_p$ on a la relation

$$\sum_{\sigma\in\mathfrak{S}_p}\varphi(\sigma)=\sum_{\sigma\in\mathfrak{S}_p}\varphi(\omega\circ\sigma);$$

appliquant ce résultat à

$$\varphi(\sigma)=\mathfrak{p}(\sigma).f(\sigma^{-1}(\omega(x))$$

et observant que

$$\varphi(\omega\circ\sigma)=\mathfrak{p}(\omega\circ\sigma).f(\sigma^{-1}(\omega^{-1}(\omega(x))))=\mathfrak{p}(\omega)\mathfrak{p}(\sigma)f(\sigma^{-1}(x)),$$

on voit que (13) s'écrit aussi

$$g(\omega(x))=\sum_{\sigma\in\mathfrak{S}_p}\mathfrak{p}(\omega)\mathfrak{p}(\sigma)f(\sigma^{-1}(x))=\mathfrak{p}(\omega).\sum_{\sigma\in\mathfrak{S}_p}\mathfrak{p}(\sigma)f(\sigma^{-1}(x));$$

ceci montre que

$$g(\omega(x))=\mathfrak{p}(\omega)g(x)$$

et prouve comme annoncé que la fonction $g(x)=g(x_1,\ldots,x_p)$ est antisymétrique.

Pour démontrer (12) supposons par exemple $x_i=x_j$ pour des entiers i et j tels que $i<j$; désignons par τ la permutation définie comme suit :

$$\tau(k)=k\quad\text{si}\quad k\neq i,\ j;\qquad\tau(i)=j;\qquad\tau(j)=i;$$

il est clair que $\mathfrak{p}(\tau)=-1$ et que, pour l'élément $x=(x_1,\ldots,x_p)\in X^p$ considéré on a

$$\tau^{-1}(x)=\tau(x)=x. \tag{14}$$

Dans l'expression

$$g(x)=\sum_{\sigma\in\mathfrak{S}_p}\mathfrak{p}(\sigma).f(\sigma^{-1}(x)),$$

on peut grouper les permutations σ en deux classes : celles pour lesquelles on a $\sigma(i)<\sigma(j)$, et celles pour lesquelles on a $\sigma(i)>\sigma(j)$; notant $\mathfrak{S}'_p$ et $\mathfrak{S}''_p$ les deux parties de $\mathfrak{S}_p$ ainsi obtenues, lesquelles sont disjointes et ont pour réunion tout $\mathfrak{S}_p$, on obtient donc

$$g(x)=\sum_{\sigma\in\mathfrak{S}'_p}\mathfrak{p}(\sigma).f(\sigma^{-1}(x))+\sum_{\omega\in\mathfrak{S}''_p}\mathfrak{p}(\omega).f(\omega^{-1}(x)); \tag{15}$$

or l'application $\sigma\to\sigma\circ\tau$ est une *bijection* de $\mathfrak{S}'_p$ sur $\mathfrak{S}''_p$; on peut donc grouper les termes de (15) deux par deux, en associant le terme relatif à σ de la première somme au terme de la seconde pour lequel $\omega=\tau\circ\sigma$; la somme de ces deux termes est

$$\mathfrak{p}(\sigma).f(\sigma^{-1}(x))+\mathfrak{p}(\tau\circ\sigma)f(\sigma^{-1}(\tau^{-1}(x))$$

i.e., en vertu de (14),

$$\mathfrak{p}(\sigma).f(\sigma^{-1}(x))+\mathfrak{p}(\tau)\mathfrak{p}(\sigma).f(\sigma^{-1}(x)),$$

et comme $\mathfrak{p}(\tau) = -1$ cette somme est nulle; ainsi les termes des deux sommes figurant dans (15) se détruisent deux à deux, ce qui achève la démonstration de (12).

3. *Applications multilinéaires alternées*

Soient X et M des *modules sur un anneau commutatif* K, et $p \geqslant 1$ un entier. On dit qu'une application p-linéaire

$$f : X^p \to M$$

est **alternée** si l'on a $f(x_1, \ldots, x_p) = 0$ toutes les fois qu'il existe des indices i et j distincts tels que $x_i = x_j$. Pour $p = 1$, cette notion se réduit à celle d'application linéaire de X dans M; pour $p = 2$ et $p = 3$ on retrouve les définitions du § précédent.

THÉORÈME 2. *Toute application multilinéaire alternée est antisymétrique.*

Pour établie la relation (6) du n° 1, donnons des valeurs fixes aux variables autres que x_i et x_{i+1}, et regardons f comme fonction de x_i et x_{i+1}; on obtient alors une fonction *bilinéaire* de x_i et x_{i+1} parce que f est multilinéaire, et *alternée* car f s'annule pour $x_i = x_{i+1}$; on en déduit donc (§ 22, n° 1, relation (2)) que l'expression $f(x_1, \ldots, x_p)$ est multipliée par -1 lorsqu'on échange x_i et x_{i+1}, ce qui prouve le Théorème.

D'après les résultats du n° 1, on voit donc qu'une application p-linéaire alternée satisfait à l'identité

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(p)}) = \mathfrak{p}(\sigma) \cdot f(x_1, \ldots, x_p); \tag{16}$$

pour $p = 2$ on retrouve la relation (2) du § 22, et pour $p = 3$ les relations (9).

THÉORÈME 3. *Soit f une application p-linéaire de X^p dans M; alors l'application g donnée par*

$$g(x_1, \ldots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot f(x_{\sigma(1)}, \ldots, x_{\sigma(p)})$$

est p-linéaire alternée.

Le terme général du second membre est évidemment une fonction p-linéaire de $x_1, \ldots, x_p$, de sorte qu'il en est de même de g; il reste donc à montrer que $g(x_1, \ldots, x_p)$ est nul si les vecteurs $x_1, \ldots, x_p$ ne sont pas deux à deux distincts, ce qui n'est autre qu'un cas particulier de l'assertion (12) du n° 2.

Exemple 2. Soient $u_1, \ldots, u_p$ des formes linéaires sur X; alors la forme p-linéaire

$$\sum \mathfrak{p}(\sigma) u_1(x_{\sigma(1)}) \ldots u_p(x_{\sigma(p)})$$

est alternée : il suffit pour le voir d'appliquer le Théorème 3 à la forme $u_1 \otimes \ldots \otimes u_p$ définie au § 21, *Exemple* 2.

La forme p-linéaire alternée ainsi obtenue s'appelle le **produit extérieur** des formes linéaires $u_1, \ldots, u_p$, et se désigne par la notation

$$u_1 \wedge \ldots \wedge u_p.$$

THÉORÈME 4. *Soit f une application p-linéaire alternée de X^p dans M. On a*

$$f(a_1, \ldots, a_p) = 0$$

toutes les fois que les vecteurs $a_1, \ldots, a_p$ sont des combinaisons linéaires de $p - 1$ vecteurs au plus.

Supposons en effet qu'on ait des relations

$$a_i = \sum_{1 \leqslant j \leqslant q} \lambda_{ij} b_j;$$

la formule (10) du § 21, n° 3, montre qu'alors

$$f(a_1, \ldots, a_p) = \sum_{j_1, \ldots, j_p} \lambda_{1,j_1} \ldots \lambda_{p,j_p} f(b_{j_1}, \ldots, b_{j_p});$$

cela dit, supposons $q < p$; alors les p entiers $j_1, \ldots, j_p$, compris entre 1 et q, ne sont jamais deux à deux distincts; comme f est alternée, on a donc

$$f(b_{j_1}, \ldots, b_{j_p}) = 0$$

quels que soient $j_1, \ldots, j_p$, d'où le résultat cherché.

COROLLAIRE 1. *Soient X et M des espaces vectoriels sur un corps commutatif et f une application multilinéaire alternée de X^p dans M. On a*

$$f(a_1, \ldots, a_p) = 0$$

lorsque les vecteurs $a_1, \ldots, a_p$ ne sont pas linéairement indépendants.

S'il existe en effet une relation linéaire

$$\lambda_1 a_1 + \cdots + \lambda_p a_p = 0$$

non triviale, avec par exemple $\lambda_p \neq 0$, on peut puisque K est ici un corps en déduire que a_p est combinaison linéaire de $a_1, \ldots, a_{p-1}$; donc les p vecteurs $a_1, \ldots, a_p$ sont des combinaisons linéaires de $p - 1$ d'entre eux, et il reste à appliquer le Théorème 4.

COROLLAIRE 2. *Soit X un K-module libre ayant une base formée de r vecteurs. Toute application p-linéaire alternée de X^p dans M est nulle pour $p \geqslant r + 1$.*

Car quels que soient $x_1, \ldots, x_p \in X$ on peut alors exprimer linéairement les x_i à l'aide de $r < p$ vecteurs.

Exemple 3. Sur un espace vectoriel de dimension r il est inutile d'étudier les formes p-linéaires alternées pour $p > r$; il suffit de se borner aux entiers $p = 1, 2, \ldots, r$.

Si en particulier X est l'espace usuel à trois dimensions sur $K = \mathbf{R}$, il n'existe aucune forme p-linéaire alternée non identiquement nulle sur X si $p \geqslant 4$.

Le n° suivant nous permettra de préciser comme suit le Corollaire précédent :

si X possède une base formée de r vecteurs, il existe effectivement sur X des formes r-linéaires alternées non identiquement nulles.

4. *Fonctions p-linéaires alternées sur un module isomorphe à* K^p

Nous allons maintenant étudier les applications p-linéaires alternées de X^p dans M lorsque le module X est libre de type fini. Dans ce n° nous étudierons le cas particulier où X est isomorphe à K^p, i.e. admet une base formée de p vecteurs $a_1, \ldots, a_p$; le cas général fera l'objet du n° 7.

Posant

$$x_i = \sum_{1 \leqslant j \leqslant p} \xi_{ij} a_j$$

le Théorème 3 du § 2 1 montre que (*)

$$f(x_1, \ldots, x_p) = \sum c_{i_1 \ldots i_p} \xi_{1 i_1} \ldots \xi_{p i_p} \tag{17}$$

où

$$c_{i_1 \ldots i_p} = f(a_{i_1}, \ldots, a_{i_p}); \tag{18}$$

puisque f est alternée, on a tout d'abord

$$c_{i_1 \ldots i_p} = 0 \quad \text{si } i_1, \ldots, i_p \text{ ne sont pas deux à deux distincts;} \tag{19}$$

on peut donc se borner, dans (17), aux termes pour lesquels les p entiers $i_1, \ldots, i_p$ sont deux à deux distincts; mais comme ces entiers sont compris entre 1 et p, ils forment alors une permutation de $1, \ldots, p$ — autrement dit il existe une et une seule permutation $\sigma \in \mathfrak{S}_p$ telle que

$$i_1 = \sigma(1), \ldots, i_p = \sigma(p);$$

mais alors

$$c_{i_1 \ldots i_p} = f(a_{\sigma(1)}, \ldots, a_{\sigma(p)}) = \mathfrak{p}(\sigma) f(a_1, \ldots, a_p)$$

puisque f est alternée.

On voit donc que la somme (17) s'écrit

$$f(x_1, \ldots, x_p) = f(a_1, \ldots, a_p) \cdot \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot \xi_{1, \sigma(1)} \ldots \xi_{p, \sigma(p)}. \tag{20}$$

Inversement, toute application f de X^p dans M satisfaisant à cette relation est multilinéaire alternée. Pour le voir, il suffit évidemment de montrer que l'expression

$$D(x_1, \ldots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \cdot \xi_{1, \sigma(1)} \ldots \xi_{p, \sigma(p)} \tag{21}$$

est une forme p-linéaire alternée sur X^p. Or désignons par $u_1, \ldots, u_p$ les fonctions

(*) Dans les calculs qui suivent on écrit les scalaires indifféremment à droite ou à gauche des éléments de M, ce qui n'a aucune importance puisque K est commutatif.

coordonnées du module X par rapport à la base $a_1, \ldots, a_p$; on a

$$\xi_{ij} = u_j(x_i)$$

et par suite

$$\mathrm{D}(x_1, \ldots, x_p) = \sum \mathfrak{p}(\sigma) . u_{\sigma(1)}(x_1) \ldots u_{\sigma(1)}(x_p); \tag{22}$$

or au lieu d'écrire le produit figurant au second membre dans l'ordre $1, \ldots, p$, on peut, puisque K est commutatif, l'écrire dans l'ordre

$$\sigma^{-1}(1), \ldots, \sigma^{-1}(p);$$

on voit alors que

$$u_{\sigma(1)}(x_1) \ldots u_{\sigma(p)}(x_p) = u_1(x_{\sigma^{-1}(1)}) \ldots u_p(x_{\sigma^{-1}(p)}),$$

en sorte que (22) s'écrit aussi

$$\mathrm{D}(x_1, \ldots, x_p) = \sum \mathfrak{p}(\sigma) . u_1(x_{\sigma^{-1}(1)}) \ldots u_p(x_{\sigma^{-1}(p)});$$

mais comme $\mathfrak{S}_p$ est un groupe l'application $\sigma \mapsto \sigma^{-1}$ de $\mathfrak{S}_p$ dans $\mathfrak{S}_p$ est bijective; en remplaçant σ par σ^{-1} dans la somme précédente on modifie donc simplement l'ordre des termes, et par suite

$$\mathrm{D}(x_1, \ldots, x_p) = \sum \mathfrak{p}(\sigma^{-1}) u_1(x_{\sigma(1)}) \ldots u_p(x_{\sigma(p)});$$

en tenant compte du fait évident que

$$\mathfrak{p}(\sigma^{-1}) = \mathfrak{p}(\sigma)^{-1} = \mathfrak{p}(\sigma),$$

il reste en définitive

$$\mathrm{D}(x_1, \ldots, x_p) = \sum \mathfrak{p}(\sigma) . u_1(x_{\sigma(1)}) \ldots u_p(x_{\sigma(p)}) \tag{23}$$

ce qui montre (*Exemple* 2) que D n'est autre que le produit extérieur des formes linéaires $u_1, \ldots, u_p$:

$$\mathrm{D} = u_1 \wedge \ldots \wedge u_p; \tag{24}$$

par suite D est une forme p-linéaire alternée sur X^p comme annoncé, et la formule (20) caractérise les applications p-linéaires alternées de X^p dans un K-module M.

On remarquera que l'on a

$$\mathrm{D}(a_1, \ldots, a_p) = 1; \tag{25}$$

en effet, on a

$$u_i(a_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

en sorte que si l'on calcule $\mathrm{D}(a_1, \ldots, a_p)$ à l'aide de la formule (22) le seul terme éventuellement non nul est celui pour lequel on a $\sigma(1) = 1, \ldots, \sigma(p) = p$; on a alors $u_{\sigma(i)}(a_i) = 1$, et évidemment $\mathfrak{p}(\sigma) = +1$, de sorte qu'on trouve bien (25).

De plus, la relation (25) caractérise D; en effet, d'après (20), on a

$$f = f(a_1, \ldots, a_p).\mathrm{D}$$

pour toute forme p-linéaire alternée f sur X^p; donc la relation $f(a_1, \ldots, a_p) = 1$ implique $f = \mathrm{D}$.

En définitive, on a démontré le résultat suivant :

THÉORÈME 5. *Soit* X *un module libre de type fini sur un anneau commutatif* K. *Soit* $(a_1, \ldots, a_p)$ *une base de* X. *Il existe alors une et une seule forme p-linéaire alternée* D *sur* X^p *telle que l'on ait*

$$\mathrm{D}(a_1, \ldots, a_p) = 1;$$

on a

$$\mathrm{D}(x_1, \ldots, x_p) = \sum_{\sigma \in \mathfrak{S}_p} \varepsilon(\sigma).\xi_{1,\sigma(1)} \cdots \xi_{p,\sigma(p)}$$

quels que soient les vecteurs

$$x_i = \sum \xi_{ij} a_j \quad (1 \leqslant i \leqslant p);$$

enfin, pour toute application p-linéaire alternée f de X^p *dans un* K*-module* M, *on a*

$$f(x_1, \ldots, x_p) = \mathrm{D}(x_1, \ldots, x_p).f(a_1, \ldots, a_p)$$

quels que soient $x_1, \ldots, x_p \in \mathrm{X}$.

Voici une conséquence intéressante de ce résultat :

COROLLAIRE. *Soit* X *un module libre de type fini sur un anneau* commutatif K. *Toutes les bases de* X *ont le même nombre d'éléments.*

Soient en effet $(a_1, \ldots, a_p)$ et $(b_1, \ldots, b_q)$ des bases de X; d'après le Théorème précédent il existe sur X une forme q-linéaire alternée f telle que

$$f(b_1, \ldots, b_q) = 1;$$

donc f n'est pas identiquement nulle; comme X possède une base formée de p vecteurs il s'ensuit que $q \leqslant p$ d'après le Corollaire du Théorème 4. Mais on a $p \leqslant q$ par un raisonnement identique, et finalement $p = q$ comme annoncé.

5. *Déterminant d'un système de vecteurs, d'une matrice, d'un endomorphisme*

Étant donnés un K-module X possédant une base $(a_1, \ldots, a_p)$ et p vecteurs $x_1, \ldots, x_p \in \mathrm{X}$, on appelle **déterminant** de $x_1, \ldots, x_p$ par **rapport à la base** $a_1, \ldots, a_p$ le scalaire $\mathrm{D}(x_1, \ldots, x_p)$ donné par la relation (21) du n° précédent.

D'autre part, étant donnée une matrice carrée

$$\Lambda = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots\cdots\cdots & & \\ \alpha_{1p} & \cdots & \alpha_{pp} \end{pmatrix}$$

à coefficients dans K, on appelle **déterminant** de A le scalaire

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots & \cdots & \cdots \\ \alpha_{1p} & \cdots & \alpha_{pp} \end{vmatrix} = \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \alpha_{1,\sigma(1)} \cdots \alpha_{p,\sigma(p)};$$

on le désigne aussi par la notation

$$\text{dét}(A).$$

Il est clair qu'avec cette définition la formule (21) s'écrit encore

$$D(x_1, \ldots, x_p) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{p1} \\ \cdots & \cdots & \cdots \\ \xi_{1p} & \cdots & \xi_{pp} \end{vmatrix};$$

or on a vu au n° précédent que ce déterminant est aussi donné par la formule (23), i.e. est aussi égal à l'expression

$$\sum \mathfrak{p}(\sigma) \cdot \xi_{\sigma(1),1} \cdots \xi_{\sigma(p),p};$$

celle-ci se déduit de (21) en y remplaçant partout ξ_{ij} par ξ_{ji}; donc :

THÉORÈME 6. *Le déterminant d'une matrice carrée à coefficients dans un anneau commutatif est égal à celui de la matrice transposée.*

Voici maintenant un autre résultat important :

THÉORÈME 7. *Soient* A *et* B *deux matrices carrées d'ordre p à coefficients dans un anneau commutatif* K. *On a alors*

$$\text{dét}(AB) = \text{dét}(A)\text{dét}(B).$$

Soit X un K-module admettant une base $a_1, \ldots, a_p$; soient u et v les endomorphismes de X admettant pour matrices, par rapport à la base en question, les matrices A et B données; la matrice AB correspond alors à l'application composée $u \circ v$.

Soit $D(x_1, \ldots, x_p)$ le déterminant de $x_1, \ldots, x_p \in X$ par rapport à la base $a_1, \ldots, a_p$. Définissons une nouvelle application D_u de X^p dans K en posant

$$D_u(x_1, \ldots, x_p) = D(u(x_1), \ldots, u(x_p));$$

alors D_u est encore une application p-linéaire alternée. Tout d'abord D_u est multilinéaire; si en effet l'on donne à $x_2, \ldots, x_p$ par exemple des valeurs fixes $b_2, \ldots, b_p$, et si l'on pose $c_i = u(b_i)$, il reste l'expression

$$D(u(x_1), b_2, \ldots, b_p);$$

comme fonction de x_1, celle-ci s'obtient en composant avec l'application *linéaire* u l'application *linéaire* $x_1 \to D(x_1, b_2, \ldots, b_p)$, en sorte que le résultat est bien fonction linéaire de x_1. Ainsi la fonction D_u est multilinéaire, et il est évident qu'elle est

alternée car la relation $x_i = x_j$ implique $u(x_i) = u(x_j)$ et donc $D(u(x_1), \ldots, u(x_p)) = 0$.

Puisque D_u est p-linéaire alternée, le Théorème 5 montre que l'on a

$$D_u(x_1, \ldots, x_p) = D_u(a_1, \ldots, a_p)D(x_1, \ldots, x_p) \tag{26}$$

quels que soient les x_i; or comme

$$u(a_i) = \sum_j \alpha_{ij} a_j \qquad \text{si} \qquad A = (\alpha_{ij})_{1 \leqslant i, j \leqslant p}$$

on voit que

$$D_u(a_1, \ldots, a_p) = D(u(a_1), \ldots, u(a_p)) = \begin{vmatrix} \alpha_{11} & \ldots & \alpha_{p1} \\ \ldots\ldots\ldots\ldots \\ \alpha_{1p} & \ldots & \alpha_{pp} \end{vmatrix} = \text{dét}(A),$$

en sorte que (26) s'écrit

$$D(u(x_1), \ldots, u(x_p)) = \text{dét}(A)D(x_1, \ldots, x_p); \tag{27}$$

on a de même

$$D(v(x_1), \ldots, v(x_p)) = \text{dét}(B)D(x_1, \ldots, x_p),$$

et en posant $w = u \circ v$ on a aussi

$$D(w(x_1), \ldots, w(x_p)) = \text{dét}(AB)D(x_1, \ldots, x_p); \tag{28}$$

mais

$$\begin{aligned} D(w(x_1), \ldots, w(x_p)) = D(u(v(x_1)), \ldots, u(v(x_p))) &= \text{dét}(A) \cdot D(v(x_1), \ldots, v(x_p)) \\ &= \text{dét}(A)\,\text{dét}(B)D(x_1, \ldots, x_p), \end{aligned}$$

et en comparant avec (28) on voit que le Théorème est démontré.

La formule (27) conduit à la notion suivante. Soit u un endomorphisme de X; la forme p-linéaire alternée D_u étant proportionnelle à D, il existe un scalaire noté

$$\text{dét}(u)$$

tel que l'on ait

$$D(u(x_1), \ldots, u(x_p)) = \text{dét}(u)D(x_1, \ldots, x_p)$$

quels que soient les $x_i \in X$; comme du reste toute forme p-linéaire alternée f sur X^p est, d'après le Théorème 5, proportionnelle à D, on a aussi

$$f(u(x_1), \ldots, u(x_p)) = \text{dét}(u) \cdot f(x_1, \ldots, x_p). \tag{29}$$

On dit que le scalaire dét(u) est le **déterminant de l'endomorphisme** u. Si A est la matrice de u par rapport à une base quelconque de X, les calculs ci-dessus (en appliquant (29) au déterminant par rapport à la dite base) montrent que

$$\text{dét}(u) = \text{dét}(A).$$

Il est clair que, si u et v sont deux endomorphismes de X, on a

$$\text{dét}(u \circ v) = \text{dét}(u)\text{dét}(v); \tag{30}$$

il est d'autre part évident sur la formule (29) que le déterminant de l'endomorphisme identique de X est égal à 1.

Remarque 2. Ce qui précède montre que, si u est un endomorphisme de X, le déterminant de la matrice de u par rapport à une base de X est indépendant de cette base. On peut le voir aussi à l'aide du raisonnement suivant. Soit A la matrice de u par rapport à une base de X; alors les matrices de u par rapport aux autres bases de X sont de la forme

$$UAU^{-1}$$

avec $U \in GL(p, K)$ comme on l'a vu au § 15. Or on a

$$\text{dét}(UAU^{-1}) = \text{dét}(U)\text{dét}(A)\text{dét}(U^{-1}),$$

d'autre part

$$\text{dét}(U)\text{dét}(U^{-1}) = \text{dét}(1_p) = 1,$$

donc

$$\text{dét}(U^{-1}) = \text{dét}(U)^{-1},$$

et il vient finalement

$$\text{dét}(UAU^{-1}) = \text{dét}(A)$$

comme annoncé.

6. *Caractérisation des bases d'un espace vectoriel de dimension finie*

Les résultats des n° précédents impliquent le théorème suivant :

THÉORÈME 8. *Soient* X *un espace vectoriel de dimension p sur un corps commutatif, $a_1, \ldots, a_p$ une base de* X, *et*

$$x_i = \sum \xi_{ij} a_j \quad (1 \leqslant i \leqslant p)$$

p éléments de X. *Les propriétés suivantes sont équivalentes :*

a) Les vecteurs $x_1, \ldots, x_p$ sont linéairement indépendants.

b) Les vecteurs $x_1, \ldots, x_p$ forment une base de X.

c) La matrice

$$\begin{pmatrix} \xi_{11} & \cdots & \xi_{p1} \\ \cdots & \cdots & \cdots \\ \xi_{1p} & \cdots & \xi_{pp} \end{pmatrix}$$

est inversible;

d) On a

$$\begin{vmatrix} \xi_{11} & \cdots & \xi_{p1} \\ \cdots & \cdots & \cdots \\ \xi_{1p} & \cdots & \xi_{pp} \end{vmatrix} \neq 0.$$

L'équivalence des conditions *a*) et *b*) résulte du § 19, Théorème 10. L'équivalence de *b*) et *c*) a été établie au § 15, Théorème 1.

En désignant par D la forme multilinéaire alternée déterminant par rapport à la base $a_1, \ldots, a_p$, la condition *d*) s'écrit $D(x_1, \ldots, x_p) \neq 0$; elle implique *a*) en vertu du Corollaire 1 du Théorème 4. Il reste à montrer que *b*) implique *d*); or si les x_i forment une base, il existe (Théorème 5) une forme p-linéaire alternée f sur X^p telle que

$$f(x_1, \ldots, x_p) \neq 0;$$

comme toute forme p-linéaire alternée sur X^p est proportionnelle à D, on en déduit qu'on a *a fortiori*

$$D(x_1, \ldots, x_p) \neq 0,$$

ce qui est la condition *d*) et achève la démonstration.

COROLLAIRE 1. *Pour qu'une matrice carrée à coefficients dans un corps commutatif soit inversible, il faut et il suffit que son déterminant soit non nul.*

Cela résulte de l'équivalence entre les propriétés *c*) et *d*) dans l'énoncé du Théorème 8.

COROLLAIRE 2. *Soient* L *un espace vectoriel de dimension finie sur un corps commutatif* K *et u un endomorphisme de* L. *Les propriétés suivantes sont équivalentes :*

a) *u est bijectif.*

b) *u est surjectif.*

c) *u est injectif.*

d) *On a* Ker $(u) = 0$, i.e. la relation $u(x) = 0$ implique la relation $x = 0$.

e) *Le déterminant de u n'est pas nul.*

L'équivalence des quatre premières conditions a déjà été établie pour K commutatif ou non (§ 19, Corollaire 1 du Théorème 13); d'autre part, pour que u soit bijectif il faut et il suffit (§ 15, n° 2) que sa matrice A par rapport à une base de L soit inversible, i.e. que dét (A) soit non nul d'après le Corollaire précédent; mais comme

$$\text{dét}(A) = \text{dét}(u),$$

on voit donc que les propriétés *a*) et *e*) sont équivalentes.

COROLLAIRE 3. *Pour qu'un système de n équations linéaires et homogènes à n inconnues*

$$\begin{cases} a_{11}\xi_1 + \cdots + a_{n1}\xi_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1n}\xi_1 + \cdots + a_{nn}\xi_n = 0 \end{cases}$$

à coefficients dans un corps commutatif K, *possède une solution non triviale, il faut et il suffit que*

$$\text{dét}((a_{ij})) = 0.$$

En effet, d'après le Théorème 2 du § 20 (équivalence entre les conditions *e*) et *f*) dans l'énoncé du Théorème), l'existence d'une solution non triviale signifie que la matrice (α_{ij}) n'est pas inversible.

COROLLAIRE 4. *Pour qu'un système de n équations linéaires à n inconnues*

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{n1}\xi_n = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}\xi_1 + \cdots + \alpha_{nn}\xi_n = \beta_n \end{cases}$$

à coefficients dans un corps commutatif, admette une et une seule solution (i.e. soit un système de Cramer) *il faut et il suffit que*

$$\text{dét}\,((\alpha_{ij})) \neq 0.$$

Cela résulte de l'équivalence entre les propriétés *a*), *d*) et *f*) dans l'énoncé du Théorème 2 du § 20.

Nous verrons au § suivant que la théorie des déterminants fournit en outre une formule explicite pour calculer la solution d'un système de Cramer.

Remarque 3. On verra au § suivant que le Corollaire 1 s'étend aux *anneaux* commutatifs, la condition dét (A) $\neq$ 0 étant remplacée par la condition que le déterminant de la matrice A soit un élément inversible de l'anneau K.

Remarque 4. Supposons que K soit le corps **R** des nombres réels, et soit X un espace vectoriel réel de dimension finie p. Étant données deux bases $(a_1, \ldots, a_p)$ et $(b_1, \ldots, b_p)$ de X, désignons par la notation

$$D(b_1, \ldots, b_p; a_1, \ldots, a_p)$$

le déterminant des vecteurs $b_1, \ldots, b_p$ par rapport à la base $(a_1, \ldots, a_p)$; c'est donc le déterminant de la matrice de passage de la base (a_i) à la base (b_i). Si l'on a trois bases (a_i), (b_i) et (c_i) de X, la matrice de passage de la première à la troisième est évidemment le produit de la matrice de passage de la première à la seconde par la matrice de passage de la seconde à la troisième; on en déduit (Théorème 7) que

$$D(c_1, \ldots, c_p; a_1, \ldots, a_p) = D(c_1, \ldots, c_p; b_1, \ldots, b_p) \cdot D(b_1, \ldots, b_p; a_1, \ldots, a_p).$$

Comme on a

$$D(a_1, \ldots, a_p; a_1, \ldots, a_p) = 1,$$

il s'ensuit que

$$D(a_1, \ldots, a_p; b_1, \ldots, b_p) = D(b_1, \ldots, b_p; a_1, \ldots, a_p)^{-1}.$$

Cela étant, on dit que deux bases $(a_1, \ldots, a_p)$ et $(b_1, \ldots, b_p)$ de X ont **même orientation** si

$$D(b_1, \ldots, b_p; a_1, \ldots, a_p) > 0,$$

et sont **d'orientations opposées** dans le cas contraire. Les trois relations qu'on

vient d'établir montrent que la propriété pour deux bases d'avoir la même orientation est une *relation d'équivalence* dans l'ensemble des bases de X; de plus, l'ensemble de ces bases se décompose, pour cette relation d'équivalence, en *deux* classes exactement [pour le voir, choisissons une base (a_i) une fois pour toutes; les bases orientées comme (a_i) forment une première classe; celles qui sont d'orientation opposée forment la seconde classe, car si (b_i) et (c_i) sont d'orientations opposées à (a_i), alors (b_i) et (c_i) sont de même orientation puisque le produit de deux nombres négatifs est positif].

Par définition, on appelle **orientation de X** chacune de ces deux classes d'équivalence : X possède donc deux orientations possibles. Toujours par définition, **orienter** X consiste à choisir une orientation de X, i.e. une de ces deux classes de bases. Pour orienter X, la façon la plus simple de procéder est de choisir une base (a_i) de X, et de déclarer qu'on choisit pour orientation celle des deux classes à laquelle la base (a_i) appartient.

Lorsqu'on a orienté l'espace vectoriel X, les bases de X qui appartiennent à l'orientation choisie sur X sont qualifiées de **directes** ou de **positivement orientés**, les autres étant qualifiées de **rétrogrades** ou de **négativement orientées.**

Étant donné un espace vectoriel X quelconque, il n'existe aucun moyen « naturel » ou « canonique » ou « intrinsèque » de choisir une orientation dans X — autrement dit, la notion de « base directe » suppose toujours un choix arbitraire, et n'a aucun sens absolu. Dans l'espace physique, la règle dite du « tire-bouchon de Maxwell » ou du « bonhomme d'Ampère » semble fournir un procédé « naturel » pour distinguer les trièdres « directs » des trièdres « rétrogrades »; mais la notion de tire-bouchon de Maxwell, comme celles de « gauche » et de « droite » sur lesquelles elle repose, n'a aucun sens *mathématique*.

Les seuls espaces où il soit possible de choisir *canoniquement* une orientation positive sont les espaces $\mathbf{R}^n$: il est en effet naturel de déclarer alors qu'on qualifiera de directe toute base orientée comme la base *canonique* de $\mathbf{R}^n$. Mais l'espace physique n'est qu'isomorphe à $\mathbf{R}^3$, il ne lui est pas identique, et on ne peut définir un isomorphisme du premier sur le second sans d'abord choisir une base du premier...

7. *Applications multilinéaires alternées : cas général*

Jusqu'à présent on a étudié les applications p-linéaires alternées sur un K-module isomorphe à K^p. Dans le cas général, on a des résultats analogues mais un peu plus compliqués :

THÉORÈME 9. *Soient* X *et* M *des modules sur un anneau commutatif* K, *et supposons* X *libre de type fini; soit* $(a_i)_{1 \leqslant i \leqslant n}$ *une base de* X. *Pour qu'une application* p*-linéaire* f *de* X^p *dans* M *soit alternée, il faut et il suffit que ses coefficients*

$$c_{i_1 \dots i_p} = f(a_{i_1}, \dots, a_{i_p})$$

par rapport à la base considérée vérifient les conditions suivantes :

(31) $c_{i_1 \dots i_p} = 0$ *si* $i_1, \dots, i_p$ *ne sont pas deux à deux distincts;*

(32) $c_{i_{\sigma(1)} \dots i_{\sigma(p)}} = \varepsilon(\sigma) c_{i_1 \dots i_p}$ *pour toute permutation* $\sigma \in \mathfrak{S}_p$.

Si ces conditions sont remplies, on a

$$(33)\qquad f(x_1, \ldots, x_p) = \sum_{1 \leq i_1 < \cdots < i_p \leq n} c_{i_1 \ldots i_p} \cdot \begin{vmatrix} \xi_{1,i_1} & \cdots & \xi_{p,i_1} \\ \cdots & \cdots & \cdots \\ \xi_{1,i_p} & \cdots & \xi_{p,i_p} \end{vmatrix}$$

quels que soient les vecteurs

$$x_i = \sum_{1 \leq j \leq n} \xi_{ij} a_j \in X.$$

Posons $a_{i_1} = b_1, \ldots, a_{i_p} = b_p$; si les indices $i_1, \ldots, i_p$ ne sont pas deux à deux distincts, deux au moins des vecteurs $b_1, \ldots, b_p$ sont égaux, de sorte que si f est alternée il vient $f(b_1, \ldots, b_p) = 0$, d'où (31). En écrivant

$$f(b_{\sigma(1)}, \ldots, b_{\sigma(p)}) = \mathfrak{p}(\sigma) f(b_1, \ldots, b_p)$$

et en remarquant que

$$b_{\sigma(k)} = a_{i_{\sigma(k)}}$$

on obtient de même la relation (32).

Supposons inversement (31) et (32) vérifiées; dans la formule

$$(34)\qquad f(x_1, \ldots, x_p) = \sum c_{i_1 \ldots i_p} \xi_{1,i_1} \ldots \xi_{p,i_p}$$

on peut se borner à étendre la sommation aux suites $i_1, \ldots, i_p$ formées d'entiers deux à deux distincts, compris entre 1 et n.

Désignons provisoirement par S l'ensemble de ces suites, et soit $S^+ \subset S$ l'ensemble des suites $i_1, \ldots, i_p$ telles que

$$(35)\qquad i_1 < \cdots < i_p.$$

Il est clair que toute suite appartenant à S s'obtient, d'une façon et d'une seule, en faisant subir une permutation convenable à une suite vérifiant (35); autrement dit, si à toute suite $(i_1, \ldots, i_p)$ vérifiant (35) et à toute permutation $\sigma \in \mathfrak{S}_p$ on associe la suite $(i_{\sigma(1)}, \ldots, i_{\sigma(p)})$ on définit une *bijection* de l'ensemble produit $S^+ \times \mathfrak{S}_p$ sur S.

La formule (34) peut donc s'écrire encore

$$f(x_1, \ldots, x_p) = \sum_{i_1 < \cdots < i_p} \sum_{\sigma \in \mathfrak{S}_p} c_{i_{\sigma(1)} \ldots i_{\sigma(p)}} \xi_{1,i_{\sigma(1)}} \ldots \xi_{p,i_{\sigma(p)}}$$

en tenant compte de (32) on trouve donc

$$f(x_1, \ldots, x_p) = \sum_{i_1 < \cdots < i_p} c_{i_1 \ldots i_p} \sum_{\sigma \in \mathfrak{S}_p} \mathfrak{p}(\sigma) \xi_{1,i_{\sigma(1)}} \ldots \xi_{p,i_{\sigma(p)}};$$

mais en posant $a_{kh} = \xi_{k,i_h}$ la somme partielle étendue à $\mathfrak{S}_p$ s'écrit

$$\sum \mathfrak{p}(\sigma) a_{1,\sigma(1)} \ldots a_{p,\sigma(p)},$$

i.e. n'est autre que le déterminant de la matrice

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots & \cdots & \cdots \\ \alpha_{1p} & \cdots & \alpha_{pp} \end{vmatrix} = \begin{vmatrix} \xi_{1,i_1} & \cdots & \xi_{p,i_1} \\ \cdots & \cdots & \cdots \\ \xi_{1,i_p} & \cdots & \xi_{p,i_p} \end{vmatrix}, \tag{36}$$

ce qui établit (33). Il reste à montrer que f est alternée.

En désignant par u_i $(1 \leqslant i \leqslant n)$ les fonctions coordonnées du module X par rapport à la base $(a_i)_{1 \leqslant i \leqslant n}$, l'expression $\xi_{1,i_1} \cdots \xi_{p,i_p}$ est la valeur sur les vecteurs $x_1, \ldots, x_p$ de la forme p-linéaire $u_{i_1} \otimes \cdots \otimes u_{i_p}$ puisqu'on a d'une manière générale

$$\xi_{ij} = u_j(x_i);$$

il s'ensuit, par un calcul analogue à celui qu'on a développé en détail au n° 4 (voir le passage de (22) à (23)), que le déterminant (36) est la valeur sur les vecteurs $x_1, \ldots, x_p$ du produit extérieur $u_{i_1} \wedge \cdots \wedge u_{i_p}$ défini au n° 3, *Exemple* 2. Par suite (33) s'écrit

$$f(x_1, \ldots, x_p) = \sum_{i_1 < \cdots < i_p} c_{i_1 \ldots i_p} u_{i_1 \ldots i_p}(x_1, \ldots, x_p) \tag{37}$$

où l'on a posé

$$u_{i_1 \ldots i_p} = u_{i_1} \wedge \cdots \wedge u_{i_p}, \tag{38}$$

et comme les formes $u_{i_1} \wedge \cdots \wedge u_{i_p}$ sont alternées (*Exemple* 2), il en est donc de même de f, ce qui achève la démonstration du Théorème.

Lorsque M = K, on déduit facilement de ce qui précède que les formes (38) pour $i_1 < \cdots < i_p$ constituent une *base* du module des formes p-linéaires alternées sur X^p. Les formes (38) sont en nombre

$$\binom{n}{p} = \frac{n!}{p!\,(n-p)!},$$

car ce coefficient binomial est aussi le nombre des suites strictement croissantes de p entiers compris entre 1 et n (ces suites correspondent en effet *biunivoquement* aux parties à p éléments de l'ensemble des entiers compris entre 1 et n).

Il est facile et utile de calculer les valeurs des formes p-linéaires alternées $u_{i_1 \ldots i_p}$ sur les vecteurs de base. Le résultat est le suivant : on a

(39) $u_{i_1 \ldots i_p}(a_{j_1}, \ldots, a_{j_p}) = \mathfrak{p}(\sigma)$ s'il existe une permutation $\sigma \in \mathfrak{S}_p$ telle que $j_1 = i_{\sigma(1)}, \ldots, j_p = i_{\sigma(p)}$

(40) $u_{i_1 \ldots i_p}(a_{j_1}, \ldots, a_{j_p}) = 0$ dans le cas contraire.

En effet, le premier membre est (d'après le Théorème 3 du § 21) le coefficient de $\xi_{1j_1} \cdots \xi_{pj_p}$ dans le développement de $u_{i_1 \ldots i_p}(x_1, \ldots, x_p)$ en fonction des coordonnées des vecteurs $x_1, \ldots, x_p$; or ce développement est

$$u_{i_1 \ldots i_p}(x_1, \ldots, x_p) = \sum \mathfrak{p}(\sigma) \xi_{1,i_{\sigma(1)}} \cdots \xi_{p,i_{\sigma(p)}},$$

d'où immédiatement les formules (39) et (40).

8. Le critère d'indépendance linéaire

Le Théorème 8 peut se généraliser comme suit :

THÉORÈME 10. *Soient* X *un espace vectoriel de dimension n sur un corps commutatif,* $a_1, \ldots, a_n$ *une base de* X, *et*

$$x_i = \sum_{1 \leqslant j \leqslant n} \alpha_{ij} a_j \quad (1 \leqslant i \leqslant p)$$

des éléments de X. *Les conditions suivantes sont équivalentes :*

a) Les vecteurs $x_1, \ldots, x_p$ *sont linéairement indépendants.*

b) Il existe sur X *une forme p-linéaire alternée f telle que*

$$f(x_1, \ldots, x_p) \neq 0.$$

c) On peut extraire de la matrice

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots\cdots\cdots\cdots \\ \alpha_{1n} & \cdots & \alpha_{pn} \end{pmatrix}$$

une matrice carrée d'ordre p de déterminant non nul.

Avec les notations du n° précédent, les déterminants d'ordre p extraits de cette matrice sont les scalaires $u_{i_1 \ldots i_p}(x_1, \ldots, x_p)$; si l'un d'eux est non nul, il est clair que la condition *b*) sera remplie — et même avec $f = u_{i_1 \ldots i_p}$ pour un choix convenable de $i_1, \ldots, i_p$. Donc *c*) implique *b*). D'autre part *b*) implique *a*) en vertu du Théorème 4.

Si la condition *a*) est vérifiée, il existe une base de X qui commence par $x_1, \ldots, x_p$; la relation (39) montre alors l'existence d'une forme p-linéaire alternée qui prend sur $x_1, \ldots, x_p$ une valeur non nulle. Donc *a*) implique *b*).

Il reste à montrer que *b*) implique *c*). Or la formule (37) montre que si $f(x_1, \ldots, x_p)$ n'est pas nul, l'un au moins des scalaires $u_{i_1 \ldots i_p}(x_1, \ldots, x_p)$ n'est pas nul, ce qui est précisément la propriété *c*). Le Théorème est donc démontré.

Remarque 5. Soit

$$A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{p1} \\ \cdots\cdots\cdots\cdots \\ \alpha_{1n} & \cdots & \alpha_{pn} \end{pmatrix}$$

une matrice à coefficients dans un corps commutatif K. D'après le Théorème 16 du § 19, le rang de A est le plus grand entier r tel qu'on puisse extraire de A une matrice carrée d'ordre r et *inversible*, autrement dit *de déterminant non nul* (Corollaire 1 du Théorème 8). La condition *c*) dans l'énoncé ci-dessus signifie donc que la matrice (α_{ij}) est de rang p, et l'équivalence avec la condition *a*) résulte alors du Théorème 15 du § 19.

Ce raisonnement permet bien entendu — c'est son principal intérêt — de calculer le rang d'une matrice (sur un corps commutatif) en examinant les déterminants qu'on peut en extraire, ce qui est fort utile dans la pratique.

On voit d'autre part que, pour exprimer que p vecteurs $x_i = \sum_j \alpha_{ij} a_j$

sont *liés*, il suffit d'écrire $\binom{n}{p}$ relations « algébriques » entre leurs coordonnées, à savoir

$$\begin{vmatrix} \xi_{1i_1} & \cdots & \xi_{pi_1} \\ \cdots\cdots\cdots & & \\ \xi_{1i_p} & \cdots & \xi_{pi_p} \end{vmatrix} = 0 \text{ quels que soient } i_1 < \cdots < i_p.$$

Pour exprimer par exemple que trois vecteurs

$$x = (\xi_1, \xi_2, \xi_3, \xi_4)$$
$$y = (\eta_1, \eta_2, \eta_3, \eta_4)$$
$$z = (\zeta_1, \zeta_2, \zeta_3, \zeta_4)$$

de $\mathbf{R}^4$ sont liés par une relation non triviale, on écrit que

$$\begin{vmatrix} \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \\ \xi_4 & \eta_4 & \zeta_4 \end{vmatrix} = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_3 & \eta_3 & \zeta_3 \\ \xi_4 & \eta_4 & \zeta_4 \end{vmatrix} = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_4 & \eta_4 & \zeta_4 \end{vmatrix} = \begin{vmatrix} \xi_1 & \eta_1 & \zeta_1 \\ \xi_2 & \eta_2 & \zeta_2 \\ \xi_3 & \eta_3 & \zeta_3 \end{vmatrix} = 0$$

9. *Conditions de compatibilité d'un système d'équations linéaires*

La théorie des déterminants permet de mettre sous une forme commode les conditions de compatibilité d'un système d'équations linéaires (§ 19, Théorème 5) lorsque le corps de base K est *commutatif*.

Soit

$$f_j(x) = \alpha_{1j}\xi_1 + \cdots + \alpha_{pj}\xi_p = \beta_j \quad (1 \leqslant j \leqslant n) \tag{41}$$

un système de n équations linéaires à p inconnues à coefficients dans K; nous noterons r le rang du système, i.e. (§ 20, n° 2) le rang de la famille des formes linéaires $f_1, \ldots, f_n$ sur K^p; ou, ce qui revient au même, le rang de la matrice (α_{ij}) formée avec les coefficients des f_j. On peut alors extraire de celle-ci une matrice carrée d'ordre r inversible, i.e. de déterminant non nul; nous supposons donc dans ce qui suit qu'on a

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{r1} \\ \cdots\cdots\cdots & & \\ \alpha_{1r} & \cdots & \alpha_{rr} \end{vmatrix} \neq 0, \tag{42}$$

en sorte que $f_1, \ldots, f_r$ sont linéairement indépendantes, et que $f_{r+1}, \ldots, f_n$ en sont des combinaisons linéaires.

THÉORÈME 11. *La relation* (42) *étant supposée vérifiée, pour que le système d'équations linéaires* (41) *possède au moins une solution il faut et il suffit qu'on ait*

$$\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{r1} & \beta_1 \\ \cdots\cdots\cdots & & & \\ \alpha_{1r} & \cdots & \alpha_{rr} & \beta_r \\ \alpha_{1j} & \cdots & \alpha_{rj} & \beta_j \end{vmatrix} = 0$$

pour tout entier j tel que $r+1 \leqslant j \leqslant n$.

Le Théorème 5 du § 19 montre tout d'abord que, pour que le système (41) possède une solution, il est nécessaire *et suffisant* qu'il en soit de même du système

$$\begin{aligned} f_1(x) &= \beta_1 \\ &\cdots\cdots\cdots \\ f_r(x) &= \beta_r \\ f_j(x) &= \beta_j \end{aligned}$$

pour tout j tel que $r+1 \leqslant j \leqslant n$; on peut donc se borner à établir le Théorème 11 dans le cas particulier où $n = r+1$, ce que nous supposerons donc dans ce qui suit.

Si le système (41) possède des solutions, on obtient alors celles-ci en résolvant le système formé par les r premières équations, et comme l'hypothèse (42) montre que le système d'équations linéaires

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{r1}\xi_r = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1r}\xi_1 + \cdots + \alpha_{rr}\xi_r = \beta_r \end{cases}$$

est de Cramer (§ 20, Théorème 2 ou bien Corollaire 4 du Théorème 8 du présent §), on voit qu'on peut alors attribuer aux inconnues $\xi_{r+1}, \ldots, \xi_p$ figurant dans le système (41) des valeurs arbitraires (§ 20, n° 5); en particulier, si le système (41) admet une solution, il en admet une pour laquelle

$$\xi_{r+1} = \cdots = \xi_p = 0,$$

et la réciproque est bien entendu triviale. Dans le cas $n = r+1$ qui nous intéresse, tout revient donc à exprimer que le système

$$\text{(43)} \qquad \begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{r1}\xi_r = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1,r+1}\xi_1 + \cdots + \alpha_{r,r+1}\xi_r = \beta_{r+1} \end{cases}$$

de $r+1$ équations à r inconnues, et de rang r, possède au moins une solution, et à montrer que pour qu'il en soit ainsi il faut et il suffit que l'on ait

$$\text{(44)} \qquad \begin{vmatrix} \alpha_{11} & \cdots & \alpha_{r1} & \beta_1 \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1,r+1} & \cdots & \alpha_{r,r+1} & \beta_{r+1} \end{vmatrix} = 0.$$

Or la condition (44) exprime aussi que le système

$$\text{(45)} \qquad \begin{cases} \alpha_{11}\eta_1 + \cdots + \alpha_{r1}\eta_r + \beta_1\eta_{r+1} = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1,r+1}\eta_1 + \cdots + \alpha_{r,r+1}\eta_r + \beta_{r+1}\eta_{r+1} = 0 \end{cases}$$

possède une solution non triviale (Corollaire 3 du Théorème 8); on est donc ramené à montrer que, dans l'hypothèse (42), *l'existence d'une solution de* (43) *équivaut à l'existence d'une solution non triviale de* (45).

Il est tout d'abord clair que la première propriété implique la seconde, car si

$(\xi_1, \ldots, \xi_r)$ est une solution de (43), alors $(\xi_1, \ldots, \xi_r, -1)$ est une solution non triviale de (45).

Considérons inversement une solution non triviale $(\eta_1, \ldots, \eta_{r+1})$ de (45); on a alors

$$\eta_{r+1} \neq 0, \tag{46}$$

car si l'on avait $\eta_{r+1} = 0$ il est clair que $(\eta_1, \ldots, \eta_r)$ serait une solution *non triviale* du système homogène associé à (43), et à fortiori du système

$$\begin{cases} a_{11}\eta_1 + \cdots + a_{r1}\eta_1 = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{1r}\eta_1 + \cdots + a_{rr}\eta_r = 0, \end{cases}$$

ce qui contredirait l'hypothèse (42) et le Corollaire 3 du Théorème 8. Cela dit, la relation (46) permet de diviser par η_{r+1} les relations (45), et on voit alors que les expressions

$$\xi_i = -\eta_i/\eta_{r+1} \quad (1 \leqslant i \leqslant r)$$

vérifient (43), ce qui achève la démonstration.

Exemple 4. Prenons $K = \mathbf{R}$ et considérons le système de 3 équations linéaires à 3 inconnues

$$\begin{cases} x + 2y + 3z = a \\ 4x + 5y + 6z = b \\ 7x + 8y + 9z = c; \end{cases}$$

son déterminant

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2 \cdot \begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3 \cdot \begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix}$$

est nul comme on le voit immédiatement, de sorte que le système considéré *n'est pas* un système de Cramer. Comme

$$\begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix} = -3$$

n'est pas nul, le système est de rang $r = 2$; il y a une seule condition de compatibilité, à savoir

$$\begin{vmatrix} 1 & 2 & a \\ 4 & 5 & b \\ 7 & 8 & c \end{vmatrix} = 0,$$

ce qui s'écrit encore, comme on le voit facilement,

$$3(2b - a - c) = 0$$

ou, si l'on préfère (*),

$$2b = a + c.$$

Comme il est évident que

$$(x + 2y + 3z) + (7x + 8y + 9z) = 2(4x + 5y + 6z),$$

la *nécessité* de la condition trouvée pouvait être prévue *a priori*.

Remarque 6. Le lecteur aura intérêt à comparer le Théorème 11 à l'*Exercice* 23 du § 19.

(*) Le lecteur qui est déjà au courant de la notion de caractéristique d'un corps (§ 30, n° 6) fera bien d'étudier complètement le système considéré lorsque K est un corps commutatif quelconque; il est clair par exemple que le raisonnement du texte cesse d'être valable en caractéristique 3.

§ 24. Déterminants

1. *Propriétés fondamentales des déterminants*

Étant donnée une matrice carrée

$$X = \begin{vmatrix} \xi_{11} & \dots & \xi_{n1} \\ \dots\dots\dots\dots \\ \xi_{1n} & \dots & \xi_{nn} \end{vmatrix}$$

à coefficients dans un anneau commutatif K, on a défini son déterminant au § 23, n° 5 comme étant le scalaire

$$\det(X) = \sum_{\sigma \in \mathfrak{S}_n} p(\sigma)\, \xi_{1,\sigma(1)} \dots \xi_{n,\sigma(n)}.$$

Si l'on introduit, dans le module K^n, l'endomorphisme u dont la matrice (par rapport à la base canonique de K^n) est X, et les vecteurs

$$x_i = u(e_i) = (\xi_{i1}, \dots, \xi_{in})$$

représentés par les *colonnes* de la matrice X, on a aussi

$$\det(X) = \det(u) = D(x_1, \dots, x_n)$$

où $D(x_1, \dots, x_n)$ désigne le déterminant des vecteurs x_i par rapport à la base canonique $e_1, \dots, e_n$ de K^n.

Il résulte évidemment de là que *le déterminant de* X *est une fonction multilinéaire alternée des colonnes de* X. De là résultent les règles de calcul suivantes, importantes dans la pratique :

a) *Un déterminant qui a deux colonnes égales est nul.*

Car une forme multilinéaire alternée est nulle lorsque deux des vecteurs variables qu'elle contient sont égaux.

b) ***Si l'on fait subir aux colonnes d'un déterminant une permutation*** σ, ***la valeur du déterminant considéré est multipliée par la signature de*** σ; en particulier, ***un déterminant est multiplié par*** -1 ***lorsqu'on permute deux de ses colonnes.***

Cela provient de l'identité

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma) f(x_1, \ldots, x_n)$$

valable pour toute fonction multilinéaire alternée.

Par exemple :

$$\begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix} = - \begin{vmatrix} b & a & c \\ b' & a' & c' \\ b'' & a'' & c'' \end{vmatrix} = + \begin{vmatrix} c & a & b \\ c' & a' & b' \\ c'' & a'' & b'' \end{vmatrix}.$$

c) Si, dans un déterminant, on multiplie tous les termes d'une colonne donnée par le même scalaire λ, *le déterminant considéré est multiplié par* λ.

d) Supposons que, pour un entier donné i, les termes de la i^{e} *colonne de la matrice* X *soient de la forme*

$$\xi_{ij} = \xi'_{ij} + \xi''_{ij} \quad (1 \leqslant j \leqslant n);$$

soit X' (*resp.* X'') *la matrice déduite de* X *en y substituant* ξ'_{ij} (*resp.* ξ''_{ij}) *à* ξ_{ij} *pour* $1 \leqslant j \leqslant n$; *on a alors*

$$\text{dét}(X) = \text{dét}(X') + \text{dét}(X'').$$

Les propriétés *c*) et *d*) traduisent les relations (3) et (4) du § 22.

Par exemple on a

$$\begin{vmatrix} a & u + 2v & c \\ a' & u' + 2v' & c' \\ a'' & u'' + 2v'' & c'' \end{vmatrix} = \begin{vmatrix} a & u & c \\ a' & u' & c' \\ a'' & u'' & c'' \end{vmatrix} + 2 \cdot \begin{vmatrix} a & v & c \\ a' & v' & c' \\ a'' & v'' & c'' \end{vmatrix}$$

En combinant les propriétés *a*), *c*) et *d*) on obtient

e) La valeur d'un déterminant ne change pas si l'on ajoute à l'une de ses colonnes une combinaison linéaire quelconque des autres colonnes.

Par exemple considérons le déterminant

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix};$$

en retranchant la seconde colonne de la troisième, puis la première de la seconde, on trouve le déterminant

$$\begin{vmatrix} 1 & 1 & 1 \\ 4 & 1 & 1 \\ 7 & 1 & 1 \end{vmatrix},$$

qui est nul puisqu'il a deux colonnes égales.

D'autre part, la relation

$$\text{dét}({}^{t}X) = \text{dét}(X)$$

du § 23, **Théorème 6**, montre que :

f) La valeur d'un déterminant ne change pas lorsqu'on échange ses lignes et ses colonnes.

Il résulte de là qu'un déterminant est fonction multilinéaire alternée de ses *lignes* aussi bien que de ses colonnes. Par suite :

g) Les règles a), ..., e) demeurent valables si l'on y remplace partout le mot colonne par le mot ligne.

2. *Développement suivant les éléments d'une ligne ou d'une colonne*

Soient X un K-module admettant une base $(a_1, \ldots, a_n)$ à n éléments, et f une forme $(n-1)$-linéaire alternée sur X. Le Théorème 9 du § 23 montre que

$$f(x_1, \ldots, x_{n-1}) = \sum_{1 \leqslant i_1 < \cdots < i_{n-1} \leqslant n} \gamma_{i_1 \ldots i_{n-1}} \cdot \begin{vmatrix} \xi_{1,i_1} & \cdots & \xi_{n-1,i_1} \\ \cdots & \cdots & \cdots \\ \xi_{1,i_{n-1}} & \cdots & \xi_{n-1,i_{n-1}} \end{vmatrix}$$

où

$$x_i = \sum \xi_{ij} a_j, \qquad \gamma_{i_1 \ldots i_{n-1}} = f(a_{i_1}, \ldots, a_{i_{n-1}}).$$

Mais les $n-1$ entiers $i_1, \ldots, i_{n-1}$ étant deux à deux distincts et compris entre 1 et n, on voit que les suites $i_1 < \cdots < i_{n-1}$ qui figurent dans la formule précédente sont au nombre de n, et sont les suivantes :

$$(2, \ldots, n); \quad (1, 3, \ldots, n); \quad \ldots; \quad (1, \ldots, n-1),$$

autrement dit ce sont les suites de la forme

$$1, \ldots, j-1, j+1, \ldots, n$$

avec $1 \leqslant j \leqslant n$. Donc la formule précédente s'écrit encore

$$f(x_1, \ldots, x_{n-1}) = \sum_{1 \leqslant j \leqslant n} \gamma_j \cdot \mathrm{D}_j(x_1, \ldots, x_{n-1}) \tag{1}$$

où l'on pose

$$\mathrm{D}_j(x_1, \ldots, x_{n-1}) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{n-1,1} \\ \cdots & \cdots & \cdots \\ \xi_{1,j-1} & \cdots & \xi_{n-1,j-1} \\ \xi_{1,j+1} & \cdots & \xi_{n-1,j+1} \\ \cdots & \cdots & \cdots \\ \xi_{1n} & \cdots & \xi_{n-1,n} \end{vmatrix} \tag{2}$$

et

$$\gamma_j = f(a_1, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n). \tag{3}$$

On notera que le scalaire $\mathrm{D}_j(x_1, \ldots, x_{n-1})$ est le déterminant de la matrice obtenue en supprimant la j^e ligne de la matrice

$$\begin{pmatrix} \xi_{11} & \cdots & \xi_{n-1,1} \\ \cdots & \cdots & \cdots \\ \xi_{1n} & \cdots & \xi_{n-1,n} \end{pmatrix} \tag{4}$$

formée avec les composantes des vecteurs x_i.

Cela établi, prenons pour X le module K^n, pour base de X la base canonique de K^n et pour f l'expression

$$f(x_1, \ldots, x_{n-1}) = D(x_1, \ldots, x_{i-1}, u, x_i, \ldots, x_{n-1}) \tag{5}$$

où u est un élément fixe de K^n et où D désigne le déterminant par rapport à la base canonique ; il est clair, puisque D est n-linéaire alternée, que f est bien une forme $(n-1)$-linéaire alternée sur K^n. Posant

$$u = \sum \alpha_j e_j,$$

on a d'après le n° 1

$$f(x_1, \ldots, x_{n-1}) = \begin{vmatrix} \xi_{11} & \cdots & \xi_{i-1,1} & \alpha_1 & \xi_{i1} & \cdots & \xi_{n-1,1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \xi_{1n} & \cdots & \xi_{i-1,n} & \alpha_n & \xi_{in} & \cdots & \xi_{n-1,n} \end{vmatrix} \tag{6}$$

D'autre part l'expression (2) est le déterminant d'ordre $n-1$ formé avec les coordonnées d'indice $\neq j$ des vecteurs $x_1, \ldots, x_{n-1}$. Il reste à calculer

$$\gamma_j = f(e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n) = \begin{cases} D(e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_{i-1}, u, e_i, \ldots, e_n) & \text{si } j \leqslant i \\ D(e_1, \ldots, e_{i-1}, u, e_i, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n) & \text{si } j > i; \end{cases}$$

comme u est somme du vecteur $\alpha_j e_j$ et d'une combinaison linéaire des vecteurs

$$e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n$$

qui figurent dans le déterminant à calculer, on peut remplacer u par $\alpha_j e_j$; il vient donc

$$\gamma_j = \begin{cases} \alpha_j . D(e_1, \ldots, e_{j-1}, e_{j+1}, \ldots, e_{i-1}, e_j, e_i, \ldots, e_n) & \text{si } j \leqslant i \\ \alpha_j . D(e_1, \ldots, e_{i-1}, e_j, e_i, \ldots, e_{j-1}, e_{j+1}, \ldots, e_n) & \text{si } j > i, \end{cases}$$

d'où immédiatement

$$\gamma_j = (-1)^{i+j} \alpha_j . D(e_1, \ldots, e_n) = (-1)^{i+j} \alpha_j.$$

Portant les résultats obtenus dans la relation (1) il vient donc l'identité

$$f(x_1, \ldots, x_{n-1}) = \sum_{j=1}^{j=n} (-1)^{i+j} \alpha_j . \begin{vmatrix} \xi_{11} & \cdots & \xi_{n-1,1} \\ \cdots & \cdots & \cdots \\ \xi_{1,j-1} & \cdots & \xi_{n-1,j-1} \\ \xi_{1,j+1} & \cdots & \xi_{n-1,j+1} \\ \cdots & \cdots & \cdots \\ \xi_{1,n} & \cdots & \xi_{n-1,n} \end{vmatrix}$$

En comparant avec (6), et en adoptant des notations plus symétriques, on obtient donc le résultat suivant :

THÉORÈME 1. *Soit $X = (\xi_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice carrée d'ordre n à coefficients dans un anneau commutatif K. Désignons par X_{ij} la matrice obtenue en supprimant la i^e colonne et la*

j^e ligne de X. *On a alors*

$$\det(X) = \sum_{1 \leqslant j \leqslant n} (-1)^{i+j} \xi_{ij} . \det(X_{ij}) \tag{7}$$

pour tout entier i tel que $1 \leqslant i \leqslant n$.

Comme les scalaires $\det(X_{ij})$ sont indépendants des termes $\xi_{i1}, \ldots, \xi_{in}$ de la i^e colonne de X, la formule (7) met en évidence le fait (évident par définition des formes multilinéaires) que le déterminant de X est fonction linéaire des termes qui figurent sur la i^e colonne de X. Pour cette raison on dit que la formule (7) est le **développement de** $\det(X)$ **suivant la i^e colonne de** X.

Comme $\det({}^tX) = \det(X)$ on a aussi bien entendu la formule

$$\det(X) = \sum_{1 \leqslant i \leqslant n} (-1)^{i+j} \xi_{ij} . \det(X_{ij}) ; \tag{8}$$

celle-ci s'appelle le **développement** de $\det(X)$ **suivant la j^e ligne de** X.

Exemple 1. On a l'identité

$$\begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{vmatrix} = a . \begin{vmatrix} b' & c' & d' \\ b'' & c'' & d'' \\ b''' & c''' & d''' \end{vmatrix} - b . \begin{vmatrix} a' & c' & d' \\ a'' & c'' & d'' \\ a''' & c''' & d''' \end{vmatrix} + c . \begin{vmatrix} a' & b' & d' \\ a'' & b'' & d'' \\ a''' & b''' & d''' \end{vmatrix} - d . \begin{vmatrix} a' & b' & c' \\ a'' & b'' & c'' \\ a''' & b''' & c''' \end{vmatrix} .$$

Exemple 2. Prenons pour X une matrice **triangulaire**, i.e. de la forme

$$X = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \ldots & \alpha_{n1} \\ 0 & \alpha_{22} & \alpha_{32} & \ldots & \alpha_{n2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \alpha_{nn} \end{pmatrix} ;$$

en la développant suivant sa première colonne on trouve

$$\det(X) = \alpha_{11} . \begin{pmatrix} \alpha_{22} & \alpha_{32} & \ldots & \alpha_{n2} \\ 0 & \alpha_{33} & \ldots & \alpha_{n3} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \alpha_{nn} \end{pmatrix}$$

d'où résulte, par récurrence sur n, qu'on a

$$\det(X) = \alpha_{11} \alpha_{22} \ldots \alpha_{nn},$$

produit des termes diagonaux de X.

Cette formule est un cas particulier de la suivante. Soient $n_1, \ldots, n_p$ des entiers strictement positifs, et considérons une matrice de la forme

$$X = \begin{pmatrix} A_{11} & A_{21} & \ldots & A_{p1} \\ 0 & A_{22} & \ldots & A_{p2} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & A_{pp} \end{pmatrix}$$

où A_{ij} est une matrice à n_i lignes et n_j colonnes quels que soient i et j. On a alors

$$\text{dét}\,(X) = \text{dét}\,(A_{11})\ \text{dét}\,(A_{22})\ \ldots\ \text{dét}\,(A_{pp}).$$

¶ Pour établir ce résultat, il suffit, en raisonnant par récurrence sur p, de le prouver pour $p = 2$, autrement dit de montrer que

$$\text{dét}\begin{pmatrix} A & B \\ O & D \end{pmatrix} = \text{dét}(A).\,\text{dét}(D)$$

si A est une matrice carrée d'ordre p, D une matrice carrée d'ordre q, et B une matrice à p colonnes et q lignes. Pour cela, plaçons-nous dans K^{p+q} et soit u l'endomorphisme de K^{p+q} ayant X pour matrice par rapport à la base canonique $(e_1, \ldots, e_{p+q})$; soient E le sous-espace engendré par $e_1, \ldots, e_p$ et F le sous-espace engendré par $e_{p+1}, \ldots, e_q$, de sorte que

$$K^{p+q} = E \oplus F;$$

soient enfin v l'endomorphisme de E admettant A pour matrice par rapport à la base $(e_1, \ldots, e_p)$ de E, w l'endomorphisme de F de matrice D par rapport à la base $(e_{p+1}, \ldots, e_{p+q})$ de F, et $D(x_1, \ldots, x_{p+q})$ le déterminant des vecteurs variables $x_1, \ldots, x_{p+q}$ par rapport à la base $e_1, \ldots, e_{p+q}$ de K^{p+q} en sorte que

$$\text{dét}\,(X) = D(u(e_1), \ldots, u(e_{p+q})).$$

On a visiblement

$$u(e_1) = v(e_1), \ldots, u(e_p) = v(e_p),$$

$$\text{dét}\,(X) = D(v(e_1), \ldots, v(e_p), b_{p+1}, \ldots, b_{p+q})$$

où l'on pose provisoirement

$$b_{p+j} = u(e_{p+j});$$

or, $b_{p+1}, \ldots, b_{p+q}$ étant donnés, il est clair que l'expression

$$D(x_1, \ldots, x_p, b_{p+1}, \ldots, b_{p+q}),$$

où les x_i varient dans E, est une forme p-linéaire alternée sur E; donc (§ 23, n° 5, formule (29)) on a

$$D(v(e_1), \ldots, v(e_p), b_{p+1}, \ldots, b_{p+q}) = \text{dét}\,(v).\,D(e_1, \ldots, e_p, b_{p+1}, \ldots)$$

en sorte qu'il vient déjà

$$\text{dét}\,(X) = \text{dét}\,(A).\,D(e_1, \ldots, e_p, b_{p+1}, \ldots, b_{p+q});$$

or

$$u(e_{p+1}) = w(e_{p+1}) + a_{p+1}, \ldots, u(e_{p+q}) = w(e_{p+q}) + a_{p+q}$$

avec des vecteurs $a_{p+j} \in E$; donc

$$\begin{aligned} D(e_1, \ldots, b_{p+q}) &= D(e_1, \ldots, e_p, w(e_{p+1}) + a_{p+1}, \ldots, w(e_{p+q}) + a_{p+q}) \\ &= D(e_1, \ldots, e_p, w(e_{p+1}), \ldots, w(e_{p+q})) \end{aligned}$$

puisque les $a_{p+j} \in E$ sont des combinaisons linéaires de $e_1, \ldots, e_p$ (utiliser la règle (e) du n° 1 par exemple). Mais un raisonnement analogue à celui qu'on a utilisé plus haut montre évidemment que

$$D(e_1, \ldots, e_p, w(e_{p+1}), \ldots, w(e_{p+q})) = \text{dét } (w) . D(e_1, \ldots, e_{p+q}),$$

et comme dét $(w) =$ dét (D) on trouve en définitive la formule

$$\text{dét } (X) = \text{dét } (A) \text{ dét } (D)$$

cherchée.

3. *Matrices complémentaires*

Soit

$$X = \begin{pmatrix} \xi_{11} & \cdots & \xi_{n1} \\ \cdots & \cdots & \cdots \\ \xi_{1n} & \cdots & \xi_{nn} \end{pmatrix}$$

une matrice carrée à coefficients dans l'anneau commutatif K. On appelle **complémentaire** de X la matrice

$$\tilde{X} = \begin{pmatrix} \tilde{\xi}_{11} & \cdots & \tilde{\xi}_{n1} \\ \cdots & \cdots & \cdots \\ \tilde{\xi}_{1n} & \cdots & \tilde{\xi}_{nn} \end{pmatrix}$$

dont les termes sont donnés par la relation

$$\tilde{\xi}_{ij} = (-1)^{i+j} . \text{dét } (X_{ji}) \tag{9}$$

où X_{ji}, rappelons-le, désigne la matrice déduite de X par suppression de la i^e ligne et de la j^e colonne.

THÉORÈME 2. *Soit* X *une matrice carrée d'ordre n à coefficients dans un anneau commutatif. On a alors*

$$\tilde{X} . X = X . \tilde{X} = \text{dét } (X) . 1_n.$$

Pour montrer par exemple que $\tilde{X} . X = \text{dét } (X) . 1_n$, i.e. que $\tilde{X} . X$ est la matrice diagonale dont tous les termes diagonaux sont égaux à dét (X), tout revient à prouver que

$$\sum_i \tilde{\xi}_{ij} . \xi_{ik} = \begin{cases} \text{dét } (X) & \text{si } j = k \\ 0 & \text{si } j \neq k. \end{cases} \tag{10}$$

Or le premier membre s'écrit

$$\sum_i (-1)^{i+j} \xi_{ik} \text{ dét } (X_{ij}) ;$$

d'après la formule (8), cette expression n'est autre que le déterminant de la matrice obtenue en remplaçant les termes $\xi_{1k}, \ldots, \xi_{nk}$ de la k^e ligne de X par les scalaires

$\xi_{1j}, \ldots, \xi_{nj}$; si $j \neq k$ la matrice ainsi obtenue a deux lignes identiques, de sorte que son déterminant est nul; si $j = k$, la matrice obtenue n'est autre que X; d'où les relations (10). La formule $\tilde{X}.X = \text{dét}\,(X).1_n$ se démontre de façon analogue : il faut utiliser (7) au lieu de (8).

COROLLAIRE 1. *Soit* X *une matrice carrée à coefficients dans un anneau commutatif* K. *Pour que* X *soit inversible, il faut et il suffit que le déterminant de* X *soit un élément inversible de* K; *on a alors*

$$X^{-1} = \text{dét}\,(X)^{-1}.\tilde{X} \tag{11}$$

Si X est inversible, le Théorème 7 du § 24 montre que

$$\text{dét}(X.X^{-1}) = \text{dét}(X)\ \text{dét}(X^{-1}) = \text{dét}(1_n) = 1,$$

de sorte que le déterminant de X est inversible dans K. Si inversement cette condition est réalisée, le Théorème 2 montre évidemment que la matrice

$$\text{dét}\,(X)^{-1}.\tilde{X}$$

est inverse de X.

Remarque 1. Le cas d'un *corps* commutatif K avait déjà été traité au § 23 (Corollaire 1 du Théorème 8).

COROLLAIRE 2. *Soient* X *un module libre de type fini sur un anneau commutatif* K, *et* $(a_1, \ldots, a_n)$ *une base de* X. *Pour que des vecteurs* $x_1, \ldots, x_n \in X$ *forment une base de* X *il faut et il suffit que leur déterminant par rapport à la base* $(a_1, \ldots, a_n)$ *soit un élément inversible de* K.

On doit en effet exprimer que les coordonnées des x_i par rapport à la base considérée forment une matrice inversible dans l'anneau $M_n(K)$.

Exemple 3. Prenons $K = \mathbf{Z}$ et $X = \mathbf{Z}^n$, la base (a_i) étant la base canonique; on obtient alors le résultat suivant : *pour que des vecteurs*

$$x_i = (\xi_{i1}, \ldots, \xi_{in}) \quad (1 \leqslant i \leqslant n)$$

forment une base de $\mathbf{Z}^n$ *il faut et il suffit que*

$$\text{dét}\,((\xi_{ij})) = \pm 1.$$

On voit aussi que le groupe $GL(n, \mathbf{Z})$ est formé des matrices carrées d'ordre n à coefficients dans $\mathbf{Z}$ et de déterminant $+1$ ou -1.

4. *Formules de Cramer*

On a vu que la résolution d'un système d'équations linéaires à coefficients dans un corps K pouvait toujours se ramener à celle d'un système de Cramer, i.e. d'un système

$$\begin{cases} \alpha_{11}\xi_1 + \cdots + \alpha_{n1}\xi_n = \beta_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{1n}\xi_1 + \cdots + \alpha_{nn}\xi_n = \beta_n \end{cases} \tag{12}$$

où la matrice

$$A = (\alpha_{ij})_{1 \leqslant i,j \leqslant n}$$

est inversible, i.e. de déterminant non nul. Posant

$$x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \qquad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$$

le système donné s'écrit

$$Ax = b$$

et a pour solution

$$x = A^{-1}b.$$

Or, on a donné au n° précédent (Corollaire 1 du Théorème 2) une formule explicite pour calculer l'inverse de A; il vient donc, si l'on en fait usage, la relation

$$x = \text{dét } (A)^{-1}.\tilde{A}b;$$

en explicitant cette formule on trouve évidemment

$$\text{dét } (A).\xi_i = \sum_j (-1)^{i+j} \text{ dét } (A_{ij})\beta_j$$

où A_{ij} se déduit de A par suppression de la j^e ligne et de la i^e colonne. Mais d'après le Théorème 1, le second membre de la relation ci-dessus n'est autre que le déterminant de la matrice obtenue, à partir de A, en remplaçant les termes $\alpha_{i1}, \ldots, \alpha_{in}$ de la i^e colonne de A par les seconds membres $\beta_1, \ldots, \beta_n$ du système (12). On voit par conséquent que la solution du système de Cramer (12) est donnée par les formules

$$\xi_i = \frac{\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{i-1,1} & \beta_1 & \alpha_{i+1,1} & \cdots & \alpha_{n1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{1n} & \cdots & \alpha_{i-1,n} & \beta_n & \alpha_{i+1,n} & \cdots & \alpha_{nn} \end{vmatrix}}{\begin{vmatrix} \alpha_{11} & \cdots & \alpha_{n1} \\ \cdots & \cdots & \cdots \\ \alpha_{1n} & \cdots & \alpha_{nn} \end{vmatrix}}$$

Celles-ci sont connues sous le nom de **formules de Cramer.**

Exemple 4. Prenons $K = \mathbf{R}$ et le système

$$\begin{cases} 2x + 3y + 4z = a \\ 5x + 6y + 7z = b \\ 8x + 9y + 9z = c; \end{cases}$$

le déterminant du système est

$$\begin{vmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 9 & 9 \end{vmatrix} = \begin{vmatrix} 2 & 3 & 1 \\ 5 & 6 & 1 \\ 8 & 9 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 1 & 1 \\ 5 & 1 & 1 \\ 8 & 1 & 0 \end{vmatrix} = \begin{vmatrix} -3 & 0 & 0 \\ -3 & 0 & 1 \\ 8 & 1 & 0 \end{vmatrix} = 3$$

de sorte qu'on a bien un système de Cramer. L'inconnue x par exemple est donnée par

$$x = \frac{1}{3}\begin{vmatrix} a & 3 & 4 \\ b & 6 & 7 \\ c & 9 & 9 \end{vmatrix} = \frac{1}{3}\begin{vmatrix} a & 3 & 1 \\ b & 6 & 1 \\ c & 9 & 0 \end{vmatrix} = -3a - c + 3b,$$

et on calculerait de même les autres inconnues.

Remarque 2. Si K est un anneau commutatif, les formules de Cramer sont bien entendu encore valables pourvu que le déterminant du système soit inversible dans K.

§ 25. Espaces affines

1. L'espace vectoriel des translations

Dans ce n° nous utiliserons les mots « espace », « point », « vecteur », « équipollent », « translation », etc...; le lecteur devra leur attribuer la même signification qu'en Géométrie Élémentaire. Nous n'en donnerons pas de définitions précises, attendu qu'un concept tel que celui d'espace n'est pas à proprement parler un objet mathématique (on ne peut pas le décrire à l'aide des signes fondamentaux de la théorie des Ensembles...). Mais on peut en dégager, en quelque sorte expérimentalement, des propriétés qui servent de base à la construction d'objets mathématiques analogues, les *espaces affines*, qui seront définis au n° suivant.

Désignons donc par E l'espace usuel, dont les éléments sont les points usuels. Nous avons vu (§ 10, *Exemple* 2) que si l'on choisit une fois pour toutes un point O dans E, on peut regarder l'ensemble des vecteurs d'origine O dans E comme un espace vectoriel réel, de dimension 3. Mais la définition de cet espace vectoriel comporte un élément d'arbitraire — à savoir le choix du point O — et si l'on remplace O par un autre point O', le nouvel espace vectoriel obtenu est isomorphe, mais non identique, au premier (on obtient un isomorphisme *canonique* du premier espace vectoriel sur le second en associant à tout vecteur d'origine O le vecteur d'origine O' qui lui est équipollent). Nous allons maintenant montrer qu'en modifiant cette construction, on peut attacher à l'espace usuel un espace vectoriel *canonique*, i.e. dont la définition ne comporte aucun choix arbitraire.

Pour cela désignons par T l'ensemble de toutes les translations dans E : une translation est une application, d'ailleurs bijective, de E dans E qui transforme chaque point $P \in E$ en le point $P' \in E$ tel que le vecteur PP' soit équipollent à un vecteur donné. Nous allons montrer que l'ensemble T peut être muni d'une structure d'espace vectoriel réel.

Pour cela on doit définir la somme $s + t$ de deux translations, et le produit λs d'une translation par un scalaire $\lambda \in \mathbf{R}$. En ce qui concerne la somme, on posera

$$s + t = s \circ t,$$

de sorte que $s + t$ sera ce qu'on appelle, en Géométrie Élémentaire, la translation

produit de s et t. Quant à λs, ce sera la translation dont le vecteur s'obtient en multipliant par λ celui de s.

Choisissons un point $O \in E$ et associons, à tout vecteur x d'origine O, la translation s_x de vecteur x : pour tout point $P \in E$, le vecteur d'origine P et d'extrémité $s_x(P)$ est donc équipollent à x. Ceci dit, il est clair qu'avec les définitions précédentes on a

$$(1) \qquad s_x + s_y = s_{x+y}, \qquad \lambda . s_x = s_{\lambda x};$$

comme l'ensemble des vecteurs d'origine O, muni des opérations algébriques évidentes, est un espace vectoriel réel, il en est donc de même de l'ensemble T muni des opérations définies plus haut, et en fait l'application $x \mapsto s_x$ est un *isomorphisme* du premier espace vectoriel sur le second.

Nous dirons que T est l'*espace vectoriel des translations* dans l'espace usuel.

Étant donnés un point $P \in E$ et une translation $s \in T$, nous poserons, par définition,

$$s + P = s(P);$$

on obtient ainsi une application $(s, P) \mapsto s + P$ de $T \times E$ dans E, et il est clair qu'on a les propriétés suivantes :

(EA 1) : *on a la relation*

$$s + (t + P) = (s + t) + P$$

quels que soient $s, t \in T$ *et* $P \in E$;

(EA 2) : *on a la relation*

$$0 + P = P$$

pour tout $P \in E$ (dans cette formule, 0 désigne bien entendu l'élément neutre de T);

(EA 3) : *quels que soient* $P, Q \in E$, *il existe un et un seul* $s \in T$ *tel que l'on ait*

$$s + P = Q.$$

La dernière propriété signifie qu'il existe toujours une et une seule translation transformant un point donné en un autre point donné, ce qui est en effet bien connu.

Lorsqu'on a

$$s + P = Q,$$

on écrit souvent, par définition,

$$s = Q - P;$$

la « différence » entre deux points P et Q est donc la translation qui amène P sur Q.

Remarque 1. On fera attention au fait que, si nous venons de définir la « différence » de deux points, laquelle est une translation et non pas un point ou un vecteur, nous n'avons par contre pas défini la « somme » de deux points, attendu que cette opération n'a aucun sens géométrique ou physique.

2. *Espaces affines associés à un espace vectoriel*

Soit T un espace vectoriel sur un corps K; on appelle **espace affine associé à T** tout objet formé par un ensemble E, et par une application de $T \times E$ dans E, notée

$$(s, P) \mapsto s + P,$$

de telle sorte que les propriétés (EA 1), (EA 2) et (EA 3) du n° précédent soient vérifiées. On appelle alors **points** les éléments de E, et tout $s \in T$ permet de définir une application de E dans lui-même, à savoir l'application, que nous noterons $\bar{s}$, donnée par

$$\bar{s}(P) = s + P \quad \text{pour tout } P \in E.$$

Ces applications sont bijectives, et prennent le nom de **translations** dans E.

Exemple 1. Soit T un espace vectoriel sur un corps K. On peut alors regarder T lui-même comme un espace affine associé à T, attendu que l'addition dans T est une application de $T \times T$ dans T qui vérifie évidemment les conditions (EA 1) à (EA 3) — celles-ci traduisent simplement, dans ce cas, le fait que T, muni de l'addition, est un groupe.

Exemple 2. Soient X un ensemble, A une partie de X, et u une application donnée de A dans un corps K. Désignons par T l'espace vectoriel sur K formé des applications s de X dans K telles que $s(a) = 0$ pour tout $a \in A$, et par E l'ensemble des applications f de X dans K telles que l'on ait $f(a) = u(a)$ pour tout $a \in A$ (autrement dit, E est l'ensemble des prolongements de u à X). Pour $s \in T$ et $f \in E$ définissons $s + f$ comme la fonction $s(x) + f(x)$; les conditions du n° 1 sont alors vérifiées, de sorte qu'on peut regarder E comme un espace affine associé à T.

Cet *Exemple*, comme le lecteur le vérifiera, est en fait un cas particulier de l'*Exemple* que voici.

Exemple 3. Soit M un espace vectoriel sur K, et prenons pour T un sous-espace vectoriel de M. Puisque T est un sous-groupe du groupe additif M, on peut considérer dans M les classes modulo T (§ 7, n° 6); pour tout $a \in M$ nous désignerons par $T + a$ la classe de a modulo T : c'est donc une partie de M, à savoir l'ensemble des $x \in M$ tels que $x - a \in T$.

Soit E une telle classe modulo T; ce n'est pas un sous-espace vectoriel de M en général; mais pour $s \in T$ et $u \in E$, il est clair que le vecteur somme $s + u$ est encore dans E; d'où une application de $T \times E$ dans E, et celle-ci permet de regarder E comme un espace affine associé à T comme on le vérifie aussitôt.

Exemple 4. Prenons E et T comme au n° 1 (espace usuel et espace vectoriel des translations usuel). Soit $E' \subset E$ un plan (resp. une droite), et soit $T' \subset T$ l'ensemble des translations de vecteur parallèle à E', autrement dit des translations qui appliquent E' dans E'. Il est clair que T' est un sous-espace vectoriel de T, et comme on a $s + P \in E'$ pour $s \in T'$ et $P \in E'$ on peut définir canoniquement une application de $T' \times E'$ dans E'. Cela permet, comme on le voit facilement, de considérer E' comme un espace affine associé à T'.

Soient T un espace vectoriel sur K et E un espace affine. Choisissons un point O de E; nous allons montrer que l'on peut regarder E comme un espace vectoriel sur K (mais la structure d'espace vectoriel sur K ainsi obtenue sur E *dépendra du choix de* O, i.e. ne sera pas canonique). Pour cela, étant donnés des points P et Q de E, on posera $P + Q = R$ où R est le point tel que

$$R - O = (P - O) + (Q - O) \tag{2}$$

(cela signifie que la translation amenant O sur R est composée de la translation amenant O sur P et de la translation amenant O sur Q); et étant donnés un point P de E et un scalaire λ de K, le point $P' = \lambda P$ sera défini par la relation

$$P' - O = \lambda.(P - O). \tag{3}$$

Nous allons montrer que E, muni de ces opérations, est un espace vectoriel isomorphe à T.

Pour cela, considérons l'application $f : T \to E$ donnée par

$$f(s) = s + O;$$

elle est bijective d'après la condition (EA 3). Soient $s, t \in T$, posons $P = f(s)$ et $Q = f(t)$, et calculons le point $R = P + Q$ donné par (2); on a $P = s + O$, donc $P - O = s$, et de même $Q - O = t$, en sorte que d'après (2) on voit que $R - O = s + t$; mais cela veut dire que $R = f(s + t)$, en sorte qu'on a

$$f(s + t) = f(s) + f(t);$$

on voit de même à l'aide de (3) que

$$f(\lambda s) = \lambda f(s).$$

Cela dit, puisque les axiomes des espaces vectoriels sont vérifiés dans T et puisque f est une bijection de T sur E, il est clair que les axiomes des espaces vectoriels sont aussi vérifiés dans E, et que f est un isomorphisme d'espaces vectoriels.

L'espace vectoriel obtenu en munissant l'ensemble E des lois de composition (2) et (3) sera noté E_O.

Exemple 5. Prenons E et T comme au n° 1; un point « origine » O étant choisi, l'addition $P + Q = R$ dans E est alors définie par la condition que le vecteur OR soit la somme des vecteurs OP et OQ, et le point $\lambda P = P'$ par la condition que le vecteur OP' soit égal au vecteur OP multiplié par λ (autrement dit P' est l'image de P par l'homothétie de centre O et de rapport λ).

3. *Barycentres dans un espace affine*

Les hypothèses et les notations restant celles du n° 2, nous allons examiner la façon dont la structure d'espace vectoriel définie sur E à l'aide du choix d'un point « origine » O dépend de ce choix.

Pour cela considérons des points $P_1, \ldots, P_n \in E$ et des scalaires $\lambda_1, \ldots, \lambda_n \in K$; *une fois choisi un point* $O \in E$, on peut définir dans l'espace vectoriel EO la combinaison linéaire

$$P = \lambda_1 P_1 + \cdots + \lambda_n P_n;$$

vu les relations (2) et (3), on a évidemment

$$P - O = \lambda_1(P_1 - O) + \cdots + \lambda_n(P_n - O), \tag{4}$$

cette relation étant une relation linéaire entre éléments de l'espace vectoriel T.

Si l'on remplace O par un autre point O', le point P est remplacé par le point P' donné par

$$P' - O' = \lambda_1(P_1 - O') + \cdots + \lambda_n(P_n - O'); \tag{5}$$

or, étant donnés des points A, B, C de E, on a d'une manière générale la relation

$$A - C = (A - B) + (B - C), \tag{6}$$

car en posant $s = A - B$ et $t = B - C$ on a $A = s + B$ et $B = t + C$, donc $A = s + (t + C)$ et par suite $A = (s + t) + C$ d'après l'axiome (EA 1), en sorte que $A - C = s + t$ comme annoncé. Cela dit, il vient

$$P_i - O' = (P_i - O) + (O - O')$$

et par suite (5) s'écrit

$$\begin{aligned} P' - O' &= \lambda_1[(P_1 - O) + (O - O')] + \cdots + \lambda_n[(P_n - O) + (O - O')] \\ &= \lambda_1(P_1 - O) + \cdots + \lambda_n(P_n - O) + (\lambda_1 + \cdots + \lambda_n).(O - O') \\ &= P - O + (\lambda_1 + \cdots + \lambda_n).(O - O'); \end{aligned}$$

comme $P' - O' = (P' - O) + (O - O')$ cette relation s'écrit encore

$$P' - O = P - O + (\lambda_1 + \cdots + \lambda_n - 1).(O - O'),$$

et comme $P' - O = (P' - P) + (P - O)$ il vient en définitive

$$P' - P = (\lambda_1 + \ldots + \lambda_n - 1).(O - O'). \tag{7}$$

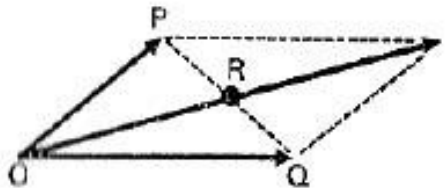
Ceci montre bien qu'on a en général $P \neq P'$, autrement dit que la structure d'espace vectoriel de E dépend effectivement du choix de l'origine O dans E.

On notera cependant que l'on a $P = P'$ dès que

$$\lambda_1 + \cdots + \lambda_n = 1; \tag{8}$$

autrement dit, *lorsque cette condition est remplie, le point* $\lambda_1 P_1 + \cdots + \lambda_n P_n$ défini tout d'abord en choisissant une origine O dans E, *est en fait indépendant du choix de* O, et a par conséquent un sens « intrinsèque » ou « canonique ». On dit que c'est le **barycentre** (ou le **centre de gravité**) **des points** $P_1, \ldots, P_n$ **affectés des masses** $\lambda_1, \ldots, \lambda_n$.

Exemple 6. Prenons E et T comme au n° 1. Étant donnés deux points P et Q, on peut donc définir le point



$$R = \frac{1}{2}P + \frac{1}{2}Q,$$

qu'en pratique on écrit

$$\frac{P+Q}{2};$$

il est donné par la relation

$$\overrightarrow{OR} = \frac{1}{2}(\overrightarrow{OP} + \overrightarrow{OQ})$$

où O est un point quelconque de E. Comme $\overrightarrow{OP} + \overrightarrow{OQ}$ est donné par la « règle du parallélogramme », il est clair que R est le milieu du segment de droite joignant P à Q.

Plus généralement considérons le point

$$R = tP + (1-t)Q$$

où t est un nombre réel quelconque; il est donné par

$$\overrightarrow{OR} = t.\overrightarrow{OP} + (1-t).\overrightarrow{OQ} = \overrightarrow{OQ} + t(\overrightarrow{OP} - \overrightarrow{OQ});$$

or $\overrightarrow{OP} - \overrightarrow{OQ}$ est le vecteur d'origine O équipollent au vecteur $\overrightarrow{PQ}$; on en déduit immédiatement, en examinant la figure ci-dessus, que R se trouve sur la droite passant par P et Q — et même que cette droite n'est autre que l'ensemble des points de la forme $tP + (1-t)Q$ où $t \in \mathbf{R}$.

Exemple 7. Prenons toujours E et T comme ci-dessus, et trois points P, Q, R; on peut alors définir le point

$$M = uP + vQ + wR \quad \text{pourvu que } u + v + w = 1; \tag{9}$$

si O est un point quelconque, on a

$$\begin{aligned}\overrightarrow{OM} = u.\overrightarrow{OP} + v.\overrightarrow{OQ} + w.\overrightarrow{OR} &= u.\overrightarrow{OP} + v.\overrightarrow{OQ} + (1-u-v).\overrightarrow{OR} \\ &= \overrightarrow{OR} + u.(\overrightarrow{OP} - \overrightarrow{OR}) + v.(\overrightarrow{OQ} - \overrightarrow{OR});\end{aligned}$$

or $\overrightarrow{OM} - \overrightarrow{OR}$ est équipollent à $\overrightarrow{RM}$, $\overrightarrow{OP} - \overrightarrow{OR}$ à $\overrightarrow{RP}$, et $\overrightarrow{OQ} - \overrightarrow{OR}$ à $\overrightarrow{RQ}$; on voit donc que le vecteur $\overrightarrow{RM}$ est une combinaison linéaire des vecteurs $\overrightarrow{RP}$ et $\overrightarrow{RQ}$, plus précisément que

$$\overrightarrow{RM} = u.\overrightarrow{RP} + v.\overrightarrow{RQ},$$

et par suite que M se trouve dans le plan déterminé par les trois points P, Q, R (en supposant que ces trois points ne sont pas situés sur une même droite). Il est clair inversement que tout point M de ce plan peut se mettre sous la forme (9) pour un choix convenable de u et v.

En particulier, si l'on prend $u = v = w = \frac{1}{3}$, le point M obtenu est donné par

$$\mathrm{RM} = \frac{1}{3}(\mathrm{RP} + \mathrm{RQ}) = \frac{2}{3} \cdot \frac{\mathrm{RP} + \mathrm{RQ}}{2}$$

et n'est autre, par suite, que le *centre de gravité du triangle* PQR au sens habituel.

Lorsqu'on calcule des barycentres dans un espace affine, on a souvent besoin de la *formule de distributivité* que voici :

$$\sum_i \mu_i \sum_j \lambda_{ij} \cdot \mathrm{P}_{ij} = \sum_{i,j} \mu_i \lambda_{ij} \cdot \mathrm{P}_{ij}; \tag{10}$$

cette formule est valable dès qu'elle a un sens, i.e. dès que les relations

$$\sum_j \lambda_{ij} = 1 \quad \text{pour tout } i, \quad \sum_i \mu_i = 1$$

sont vérifiées. (On notera que ces relations impliquent

$$\sum_{i,j} \mu_i \lambda_{ij} = 1,$$

de sorte que le second membre de (10) a effectivement un sens lorsqu'elles sont vérifiées.) Pour prouver (10) il suffit de montrer que, si O est un point de E, on a

$$\sum_i \mu_i \left(\sum_j \lambda_{ij} \cdot \mathrm{P}_{ij} - \mathrm{O} \right) = \sum_{i,j} \mu_i \lambda_{ij} (\mathrm{P}_{ij} - \mathrm{O});$$

mais on a par définition

$$\sum_j \lambda_{ij} \cdot \mathrm{P}_{ij} - \mathrm{O} = \sum_j \lambda_{ij} (\mathrm{P}_{ij} - \mathrm{O}),$$

et par conséquent la relation à établir se réduit à la distributivité dans l'espace vectoriel T.

4. *Variétés linéaires dans un espace affine*

Les notions de « droite » et de « plan » de la Géométrie Élémentaire peuvent se généraliser comme suit. Soient T un espace vectoriel sur un corps K, E un espace affine associé à T, et V une partie de E. On dit que V est une **variété linéaire** dans E si, quels que soient les points $\mathrm{P}_1, \ldots, \mathrm{P}_n \in \mathrm{V}$ et les scalaires $\lambda_1, \ldots, \lambda_n \in \mathrm{K}$ tels que

$$\lambda_1 + \cdots + \lambda_n = 1,$$

on a

$$\lambda_1 \mathrm{P}_1 + \cdots + \lambda_n \mathrm{P}_n \in \mathrm{V}.$$

Exemple 8. L'ensemble vide est une variété linéaire.

Exemple 9. Pour tout point $P \in E$, l'ensemble réduit à P est une variété linéaire; cela provient de ce que l'on a

$$\lambda_1 P_1 + \cdots + \lambda_n P_n = P \quad \text{si} \quad P_1 = \cdots = P_n = P.$$

Dans la pratique on ne fait naturellement aucune différence entre le point P et la variété linéaire $\{P\}$.

Exemple 10. Soient $P_1, \ldots, P_n$ des points donnés dans E; alors l'ensemble V des $P \in E$ qui peuvent se mettre sous la forme

$$P = \lambda_1 P_1 + \cdots + \lambda_n P_n$$

avec des scalaires $\lambda_i \in K$ de somme égale à 1 est une variété linéaire.
En effet, étant donnés des points

$$Q_j = \sum_i \lambda_{ij} P_i \qquad \text{avec} \qquad \sum_i \lambda_{ij} = 1$$

de V et des scalaires μ_j de somme totale 1, on a d'après (10)

$$\sum_j \mu_j Q_j = \sum_j \mu_j \sum_i \lambda_{ij} P_i = \sum_{i,j} \mu_j \lambda_{ij} P_i = \sum_i \nu_i P_i$$

avec

$$\nu_i = \sum_j \mu_j \lambda_{ij},$$

ce qui montre bien que V satisfait à la définition des variétés linéaires.

On notera que V contient les points donnés P_i (par exemple la relation $P_1 \in V$ s'obtient en prenant $\lambda_1 = 1$, et $\lambda_2 = \cdots = \lambda_n = 0$), et que toute variété linéaire contenant les P_i contient V par définition. Par suite, V est la plus petite variété linéaire contenant les P_i. On dit que c'est la **variété linéaire engendrée par les points** $P_1, \ldots, P_n$.

Exemple 11. Soient P et Q deux points de E; la variété linéaire qu'ils engendrent est formée des points $\lambda P + (1 - \lambda)Q$, où $\lambda \in K$ est arbitraire. Si $P = Q$ elle se réduit à P. Si $P \neq Q$, on l'appelle la **droite joignant P et Q**. Cette terminologie est justifiée par l'*Exemple* 6.

On notera que, si K est le corps des entiers modulo 2, corps ne possédant que les éléments 0 et 1, la droite joignant P et Q se réduit à l'ensemble $\{P, Q\}$. Cette circonstance, qui paraîtra sans doute étrange au débutant, complique la géométrie sur ce corps (et plus généralement sur les **corps de caractéristique** 2; on appelle ainsi tout corps K pour lequel $1 + 1 = 0$, où les symboles 0 et 1 désignent bien entendu l'élément neutre et l'élément unité de K; voir § 30, n° 6).

THÉORÈME 1. *Soient* T *un espace vectoriel sur un corps* K, E *un espace affine associé à* T, *et* V *une partie non vide de* E. *Les propriétés suivantes sont équivalentes :*

a) V *est une variété linéaire.*

b) Pour tout point $P_0 \in V$, *l'ensemble des vecteurs* $P - P_0$ *où* $P \in V$ *est un sous-espace vectoriel de* T.

c) Il existe un point $P_0 \in V$ *tel que l'ensemble des vecteurs* $P - P_0$ *où* $P \in V$ *soit un sous-espace vectoriel de* T.

Si ces conditions équivalentes sont remplies, le sous-espace vectoriel de T *formé des vecteurs de la forme* $P - P_0$ (où P_0 est un point donné et P un point variable de V) *est indépendant du choix de* P_0 *dans* V, *et c'est l'ensemble des* $s \in T$ *tels que l'on ait*

$$s + P \in V \quad \textit{pour tout} \quad P \in V.$$

Montrons que *a)* implique *b)*; il suffit d'établir que, quels que soient $Q, R \in V$ et les scalaires $\lambda, \mu \in K$, il existe un $S \in V$ tel que

$$S - P_0 = \lambda(Q - P_0) + \mu(R - P_0);$$

or le point S défini par cette relation vérifie aussi

$$S - P_0 = \lambda(Q - P_0) + \mu(R - P_0) + (1 - \lambda - \mu)(P_0 - P_0),$$

en sorte que

$$S = \lambda Q + \mu R + (1 - \lambda - \mu)P_0,$$

ce qui montre que $S \in V$ si V est une variété linéaire.

L'implication *b)* $\Longrightarrow$ *c)* étant triviale, montrons que *c)* implique *a)*. Considérons le point

$$P = \sum_{i=1}^{i=n} \lambda_i P_i \qquad \text{avec} \qquad \sum \lambda_i = 1;$$

on doit montrer que $P \in V$ si V contient $P_1, \ldots, P_n$. Or on a

$$P - P_0 = \sum \lambda_i (P_i - P_0);$$

si V contient les P_i l'hypothèse *c)* montre donc que $P - P_0 = P' - P_0$ pour un $P' \in V$, ce qui implique évidemment $P = P'$, donc $P \in V$ comme annoncé.

Montrons maintenant que, si V est une variété linéaire non vide, le sous-espace de T formé par les vecteurs $P - P_0$ ($P \in V$) ne dépend pas du choix de P_0 dans V.

Soient en effet H_0 le sous-espace vectoriel associé à P_0 et H_1 le sous-espace obtenu pour un autre choix $P_1 \in V$; pour tout $P \in V$ on a

$$P - P_0 = (P - P_1) + (P_1 - P_0) = (P - P_1) - (P_0 - P_1);$$

comme, par définition, H_1 contient $P - P_1$ et $P_0 - P_1$, on a donc $P - P_0 \in H_1$; ceci montre que $H_0 \subset H_1$, d'où $H_0 = H_1$ par raison de symétrie.

On voit donc que le sous-espace vectoriel formé par les vecteurs $P - P_0$ ($P \in V$) est indépendant du choix de P_0 dans V; notons-le H. Il est clair que c'est aussi l'ensemble des vecteurs de la forme $P - Q$, où P et Q sont dans V. Soient $s \in H$ et $P \in V$; en choisissant $P_0 = P$, on voit qu'il existe un $Q \in V$ tel que $s = Q - P$; on a alors $Q = s + P$, ce qui montre qu'on a $s + P \in V$ pour tout $s \in H$ et tout $P \in V$.

Inversement, considérons un $s \in T$ vérifiant cette condition; choisissant un $P \in V$ et posant $s + P = Q$, on a d'une part $Q \in V$, d'autre part $s = Q - P$, et par suite $s \in H$, ce qui achève la démonstration.

Étant donnée une variété linéaire non vide V dans E, le sous-espace H de T formé des vecteurs de la forme $Q - P$, avec $P, Q \in V$, s'appelle le **sous-espace directeur** de V.

La connaissance de H ne suffit pas à déterminer V; mais si deux variétés linéaires non vides V' et V'' ont le même sous-espace directeur H, alors il existe, dans E, une translation qui applique V' sur V''. En effet, choisissons un point P' de V' et un point P'' de V'' une fois pour toutes, et considérons le vecteur $a = P'' - P'$. Pour qu'un point P soit dans V' il faut et il suffit que $P - P' \in H$; comme on a évidemment

$$P - P' = (P + a) - P'',$$

et comme H est aussi le sous-espace directeur de V'', on voit que les relations $P \in V'$ et $P + a \in V''$ sont équivalentes; autrement dit, la translation $P \to P + a$ applique V' sur V''.

Exemple 12. Soit T un espace vectoriel sur K, et prenons pour E l'ensemble T lui-même comme dans l'*Exemple* 1 ci-dessus. Soient V une variété linéaire et H son sous-espace directeur; si $P_0 \in V$, on voit que V est l'ensemble des vecteurs (ici on n'a pas à distinguer les points des vecteurs) de la forme $P_0 + P$ où P décrit H, autrement dit c'est la *classe* de P_0 *modulo* H au sens du § 7, n° 6 appliqué au groupe additif T.

Comme certains auteurs appellent *variété linéaire* ce que nous appelons *sous-espace vectoriel*, il est utile, dans le cas d'un espace vectoriel regardé comme espace affine, d'appeler **variété linéaire affine** ce que nous devrions appeler variété linéaire. Dans un espace vectoriel T, une variété linéaire affine est donc, soit l'ensemble vide, soit une classe modulo un sous-espace vectoriel de T.

Remarque 2. Soit V une variété linéaire non vide dans un espace affine E associé à un espace vectoriel T, et soit $H \subset T$ le sous-espace directeur de V. On a $s + P \in V$ pour tout $s \in H$ et tout $P \in V$, d'où une application $(s, P) \to s + P$ de $H \times V$ dans V; cette application vérifie les axiomes (EA 1), (EA 2) et (EA 3) des espaces affines comme on le voit aussitôt. Par suite, *on peut considérer* V *comme un espace affine associé à* H, et appliquer à V toutes les définitions de la théorie des espaces affines. Par exemple, étant donnés des points $P_i \in V$ et des scalaires $\lambda_i \in K$ de somme 1, on peut définir dans l'espace affine V le barycentre des points P_i affectés des masses λ_i; le point de V ainsi obtenu est évidemment le même que celui qu'on obtiendrait en se plaçant dans l'espace affine E.

Remarque 3. Supposons que V soit la variété linéaire engendrée par des points $P_0, \ldots, P_n \in E$; le sous-espace directeur H est l'ensemble des vecteurs de la forme $P - M$, où P varie dans V et où M est un point choisi une fois pour toutes dans V. Or les éléments de V ne sont autres que les points

$$P = \sum_{0 \leq i \leq n} t_i P_i \quad \text{avec} \quad \sum_{0 \leq i \leq n} t_i = 1;$$

pour un tel point on a, d'après la relation (4) ci-dessus,

$$P - M = \sum_{0 \leqslant i \leqslant n} \xi_i (P_i - M).$$

En particulier, prenons $M = P_0$; on voit alors que H est l'ensemble des vecteurs

$$\xi_1(P_1 - P_0) + \cdots + \xi_n(P_n - P_0)$$

où $\xi_1, \ldots, \xi_n$ sont des scalaires arbitraires (car la relation

$$\xi_0 + \xi_1 + \cdots + \xi_n = 1$$

permet évidemment de choisir arbitrairement n des $n + 1$ scalaires $\xi_0, \ldots, \xi_n$). Autrement dit, H *est le sous-espace vectoriel de* T *engendré par les vecteurs*

$$P_i - P_0 \quad (1 \leqslant i \leqslant n).$$

Remarque 4. Soit E un espace affine. On a vu au n° 2 que, si l'on choisit dans E un point « origine » O, on peut regarder E comme un espace vectoriel dont O est l'élément neutre, en définissant par exemple la somme $P + Q = R$ de deux éléments de E par la relation

$$R - O = (P - O) + (Q - O).$$

Notons E_0 l'espace vectoriel ainsi obtenu. Soit V une partie non vide de E et prenons pour O un point de V; alors, *pour que* V *soit une variété linéaire dans* E, *il faut et il suffit que* V *soit un sous-espace vectoriel de* E_0. On laisse au lecteur le soin d'établir ce fait à titre d'exercice.

5. *Génération d'une variété linéaire par des droites*

En Géométrie Élémentaire, on définit un plan par la condition que, s'il contient deux points, il contient aussi la droite qui les joint. Comme nous avons déjà adopté une autre définition des variétés linéaires (et en particulier des plans), on peut conjecturer que, dans l'optique adoptée ici, la « définition » élémentaire va devenir un théorème (ce qui compense le fait que, dans la situation classique, notre définition des plans est en fait un théorème...). C'est en effet ce qui se passe pourvu que le corps de base K ne soit pas de caractéristique 2 (autrement dit pourvu que l'on n'ait pas $1 + 1 = 0$ dans K) :

Théorème 2. *Soient* T *un espace vectoriel sur un corps* K *qui n'est pas de caractéristique* 2, E *un espace affine associé à* T, *et* V *une partie de* E. *Pour que* V *soit une variété linéaire, il faut et il suffit que, quels que soient les points distincts* P, Q ∈ V, *la droite joignant* P *et* Q *soit contenue dans* V.

Supposons que V soit une variété linéaire. Si V contient deux points distincts P et Q, elle contient aussi la variété linéaire engendrée par P et Q (*Exemple* 10), i.e. la droite joignant P et Q.

Inversement supposons cette condition remplie. Comme l'ensemble vide est une variété linéaire, on peut supposer V non vide; choisissant un point $P_0 \in V$, tout revient (Théorème 1) à faire voir que l'ensemble des vecteurs $Q - P_0$, pour $Q \in V$, est un sous-espace vectoriel de T; autrement dit que, quels que soient les scalaires λ, $\mu \in K$ et les points $Q, R \in V$, il existe un $S \in V$ tel que

$$S - P_0 = \lambda(Q - P_0) + \mu(R - P_0).$$

Mais le point S défini par cette relation n'est autre, évidemment, que

$$S = \lambda Q + \mu R + (1 - \lambda - \mu)P_0.$$

On est donc ramené à montrer, en changeant les notations, que quels que soient $P, Q, R \in V$ et les scalaires $\lambda, \mu, \nu \in K$ tels que

$$\lambda + \mu + \nu = 1, \tag{11}$$

on a

$$\lambda P + \mu Q + \nu R \in V.$$

Or comme K n'est pas de caractéristique 2, on a $2 \neq 0$ et donc $3 \neq 1$ dans K; par suite on ne peut pas avoir $\lambda = \mu = \nu = 1$; on peut donc supposer par exemple $\lambda \neq 1$. Alors $1 - \lambda$ est inversible, et la formule (10) du n° 4 permet d'écrire

$$\lambda P + \mu R + \nu S = \lambda P + (1 - \lambda)M \qquad \text{où} \qquad M = \frac{\mu}{1 - \lambda} Q + \frac{\nu}{1 - \lambda} R.$$

Si $Q = R$ on a $M = Q = R$ et M est dans V; si $Q \neq R$, le point M se trouve sur la droite joignant Q et R, donc appartient par hypothèse à V. Ainsi on a $M \in V$ dans tous les cas. Un raisonnement identique montre alors que V contient aussi le point $\lambda P + (1 - \lambda)M$, ce qui achève la démonstration.

La conclusion du Théorème 2 est *évidemment* en défaut si K est le corps des entiers modulo 2, car dans ce cas une partie *arbitraire* de E, dès qu'elle contient deux points P et Q distincts, n'a aucun mérite à contenir la droite joignant P et Q : celle-ci se réduit en effet (*Exemple* 11) aux deux points P et Q eux-mêmes !

6. *Espaces affines de dimension finie. Bases affines*

Soient T un espace vectoriel sur un corps K et E un espace affine associé à T; on dit que E est de dimension finie s'il en est ainsi de T; on appelle alors **dimension** de E (sur le corps de base K) le nombre

$$\dim(E) = \dim(T).$$

Supposons E de dimension finie n. Choisissons une base $(a_1, \ldots, a_n)$ de T, un point P_0 de E, et posons

$$P_i = a_i + P_0 \quad (1 \leqslant i \leqslant n).$$

Pour tout $P \in E$, il existe un et un seul système de scalaires $\xi_1, \ldots, \xi_n \in K$ tels que l'on ait

$$P - P_0 = \xi_1 a_1 + \cdots + \xi_n a_n;$$

posant

$$\xi_0 = 1 - \xi_1 - \cdots - \xi_n$$

il est clair que cette relation s'écrit aussi

$$P = \xi_0 P_0 + \xi_1 P_1 + \cdots + \xi_n P_n, \tag{12}$$

et l'on voit immédiatement que, *pour tout* $P \in E$ *il existe un et un seul système de scalaires* $\xi_0, \ldots, \xi_n$ *vérifiant* (12) *et*

$$\xi_0 + \xi_1 + \ldots + \xi_n = 1. \tag{13}$$

Inversement, pour que cette propriété soit vérifiée, il faut que les vecteurs $a_i = P_i - P_0$ $(1 \leqslant i \leqslant n)$ forment une base de T.

Lorsqu'on est dans la situation qu'on vient de décrire, on dit que la famille $(P_0, \ldots, P_n)$ est une **base affine** de E, et les scalaires ξ_i vérifiant (12) et (13) s'appellent les **coordonnées affines** du point P par rapport à la base affine considérée.

Exemple 13. Prenons pour E l'espace usuel de la Géométrie Élémentaire; une base affine est alors un système de quatre points (A, B, C, D) non situés dans un même plan, i.e. constituant les sommets d'un véritable tétraèdre. Les coordonnées affines d'un point $P \in E$ sont alors les nombres x, y, z, t tels que l'on ait

$$P = xA + yB + zC + tD, \qquad x + y + z + t = 1.$$

On a évidemment alors

$$\overrightarrow{AP} = y.\overrightarrow{AB} + z.\overrightarrow{AC} + t.\overrightarrow{AD}$$

en sorte que y, z, t sont les composantes du vecteur $\overrightarrow{AP}$ par rapport au système des trois vecteurs linéairement indépendants $\overrightarrow{AB}$, $\overrightarrow{AC}$, $\overrightarrow{AD}$, lesquels forment une base de l'espace vectoriel E_A.

Remarque 5. Soit V une variété linéaire non vide dans E; on a vu (*Remarque* 2) qu'on peut regarder V comme un espace affine associé à son sous-espace directeur H. On appellera donc **dimension de la variété linéaire** V la dimension de H. Si V est engendrée par des points $P_0, \ldots, P_n$, la *Remarque* 3 montre que *la dimension de* V *est égale au rang de la famille des vecteurs* $P_i - P_0$ $(1 \leqslant i \leqslant n)$.

Étant données deux variétés linéaires V et W dans E, telles que $V \subset W$, on a $\dim(V) \leqslant \dim(W)$, et on ne peut avoir $\dim(V) = \dim(W)$ que si $V = W$. On laisse au lecteur le soin d'établir ce résultat à titre d'exercice.

Soit E un espace affine associé à un espace vectoriel T de dimension finie; on a vu plus haut que, pour que des points $P_0, \ldots, P_n \in E$ forment une base affine de E, il faut et il suffit que les vecteurs $P_i - P_0$ $(1 \leqslant i \leqslant n)$ forment une base de T. On en déduit évidemment que toutes les bases affines de E comportent le même nombre de points, à savoir $n + 1$ où $n = \dim(E)$.

D'autre part, pour que les vecteurs $P_i - P_0$, en nombre $n = \dim(E)$, forment une base de T, il faut et il suffit (§ 19, Théorème 10) qu'ils engendrent T, ce qui signifie évidemment que tout $P \in E$ peut s'écrire d'*au moins* une façon sous la forme

$$P = \sum_{0 \leqslant i \leqslant n} \xi_i P_i \qquad \text{avec} \qquad \sum_{0 \leqslant i \leqslant n} \xi_i = 1,$$

ou encore que la variété linéaire engendrée par les P_i est E tout entier; comme cette variété est la plus petite qui contienne les P_i, on a donc le résultat suivant :

THÉORÈME 3. *Soit* E *un espace affine de dimension* n *sur un corps* K. *Pour que* $n + 1$ *points de* E *forment une base affine de* E, *il faut et il suffit qu'ils ne soient contenus dans aucune variété linéaire distincte de* E *tout entier.*

Par exemple, pour que quatre points forment une base affine de l'espace usuel, il est nécessaire et suffisant qu'ils ne soient pas contenus dans un même plan.

Remarque 6. Soient $Q_0, \ldots, Q_m$ des points de E. On dit qu'ils sont **indépendants** s'ils forment une base de la variété linéaire V qu'ils engendrent, autrement dit si tout $P \in E$ peut s'écrire d'*au plus* une façon sous la forme

$$P = \sum_{0 \leqslant j \leqslant m} \xi_j Q_j \qquad \text{avec} \qquad \sum_{0 \leqslant j \leqslant m} \xi_j = 1.$$

Il revient évidemment au même de dire que, dans T, les vecteurs $Q_j - Q_0$ $(1 \leqslant j \leqslant m)$ sont linéairement indépendants.

Par exemple, pour que trois points de l'espace usuel soient indépendants il faut et il suffit qu'ils ne soient pas situés sur une même droite, autrement dit qu'ils engendrent effectivement un plan.

7. *Calcul de la dimension d'une variété linéaire*

Dans la pratique on a fréquemment à calculer la dimension d'une variété linéaire engendrée par des points d'un espace affine. On utilise alors le résultat que voici :

THÉORÈME 4. *Soient* E *un espace affine sur un corps,* $(P_0, \ldots, P_n)$ *une base affine de* E, *et*

$$Q_i = \sum_{j=0}^{j=n} \alpha_{ij} P_j \quad (0 \leqslant i \leqslant m)$$

des points de E. *Soit* r *la dimension de la variété linéaire engendrée par* $Q_0, \ldots, Q_m$. *Alors la matrice*

$$A = (\alpha_{ij})_{0 \leqslant i \leqslant m,\, 0 \leqslant j \leqslant n}$$

est de rang $r + 1$.

D'après la Remarque 5, la dimension r cherchée est égale au rang de la famille des vecteurs $Q_i - Q_0$. Or on a

$$Q_i - Q_0 = (Q_i - P_0) - (Q_0 - P_0) = \sum_j \alpha_{ij}(P_j - P_0) - \sum_j \alpha_{0j}(P_j - P_0) = \sum_{j=1}^{j=n} (\alpha_{ij} - \alpha_{0j}) \cdot a_j$$

où les vecteurs $a_j = P_j - P_0$ $(1 \leqslant j \leqslant n)$ forment une base de T. Par suite (§ 19, Théorème 15) le nombre r est égal au rang de la matrice

$$A' = (\alpha_{ij} - \alpha_{0j})_{1 \leqslant i \leqslant m,\, 1 \leqslant j \leqslant n}$$

Il reste donc à montrer, en utilisant bien entendu les relations

$$\sum_{j=0}^{j=n} \alpha_{ij} = 1 \quad (0 \leqslant i \leqslant m) \tag{14}$$

existant entre les coordonnées affines des divers points Q_i, que le rang de la matrice A surpasse celui de A′ d'une unité.

Considérons pour cela d'une part le système d'équations linéaires et homogènes

$$\sum_{i=1}^{i=m} \xi_i \alpha_{ij} = 0 \quad (0 \leqslant j \leqslant n) \tag{15}$$

et d'autre part le système

$$\sum_{i=0}^{i=n} \eta_i (\alpha_{ij} - \alpha_{0j}) = 0. \tag{16}$$

Les solutions du premier forment un sous-espace M de K^{m+1}, et celles du second un sous-espace N de K^n; la matrice A′ étant de rang r, on a

$$\dim (N) = n - r$$

d'après le § 19, Théorème 17. Si on désigne provisoirement par s le rang de A, on voit de même que

$$\dim (M) = n + 1 - s;$$

pour montrer que $s = r + 1$ il suffit donc de montrer que M et N ont même dimension, autrement dit sont isomorphes.

Or, en ajoutant membre à membre les équations (15) et en tenant compte de (14), il vient évidemment $\xi_0 + \xi_1 + \cdots + \xi_n = 0$, en sorte que (15) implique

$$\sum_{i=1}^{i=n} \xi_i \alpha_{ij} - (\xi_1 + \cdots + \xi_n) \alpha_{0j} = 0$$

i.e.

$$\sum_{i=1}^{i=n} \xi_i (\alpha_{ij} - \alpha_{0j}) = 0.$$

On obtient donc une application u de M dans N en posant

$$u(\xi_0, \xi_1, \ldots, \xi_n) = (\xi_1, \ldots, \xi_n);$$

un raisonnement analogue montrerait que l'on peut construire une application v de N dans M en posant

$$v(\eta_{|1}, \ldots, \eta_{|n}) = (-\eta_1 - \cdots - \eta_{|n}, \eta_{|1}, \ldots, \eta_{|n});$$

ceci dit, il est clair que u et v sont des homomorphismes réciproques l'un de l'autre de M dans N et de N dans M, autrement dit des isomorphismes de M dans N et de N dans M, et ceci achève la démonstration du Théorème.

Remarque 7. Le rang de A ne change pas si l'on ajoute à la première ligne de A la somme des autres, i.e. si l'on remplace la première ligne de A par $(1, 1, \ldots, 1)$. On peut alors retrancher la première colonne de la matrice ainsi obtenue des colonnes suivantes; on obtient alors une matrice de même rang que A, dont la première ligne est $(1, 0, \ldots, 0)$, et dont les termes situés en dessous et à droite de la première ligne et de la première colonne ne sont autres que ceux de A'; d'où une autre démonstration du fait que $rg(A) = 1 + rg(A')$.

8. Équations d'une variété linéaire en coordonnées affines

Soient E un espace affine de dimension finie n sur un corps *commutatif* K et $(P_0, \ldots, P_n)$ une base affine de E. Considérons r points *indépendants*

$$Q_i = \sum_{j=0}^{j=n} \alpha_{ij} P_j \quad (1 \leqslant i \leqslant r)$$

dans E et soit V la variété linéaire, de dimension $r - 1$, qu'ils engendrent. Soit

$$P = \sum \xi_j P_j$$

un point de E. Pour que P appartienne à V, il faut et il suffit que la variété linéaire W engendrée par $P, Q_1, \ldots, Q_r$ soit de dimension $r - 1$; en effet, si $P \in V$ il est clair qu'on a $W = V$, d'où la nécessité de la condition; inversement, supposons-la vérifiée; comme on a de toute façon $V \subset W$, la relation dim (V) = dim (W) montre que $V = W$, et donc que P appartient à V.

Pour exprimer que $P \in V$, tout revient donc, d'après le Théorème 4, à écrire que la matrice

$$\begin{pmatrix} \xi_0 & \alpha_{10} & \cdots & \alpha_{r0} \\ \xi_1 & \alpha_{11} & \cdots & \alpha_{r1} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_n & \alpha_{1n} & \cdots & \alpha_{rn} \end{pmatrix}$$

est de rang r, autrement dit, puisque K est commutatif, que tout déterminant d'ordre $s > r$ extrait de celle-ci est nul (§ 23, n° 8, *Remarque* 5), et il suffit évidemment d'exprimer cette condition pour $s = r + 1$

En exprimant la condition en question pour $s = r + 1$, on voit qu'on doit avoir

$$(17) \qquad \begin{vmatrix} \xi_{i_0} & \alpha_{1i_0} & \cdots & \alpha_{ri_0} \\ \xi_{i_1} & \alpha_{1i_1} & \cdots & \alpha_{ri_1} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_{i_r} & \alpha_{1i_r} & \cdots & \alpha_{ri_r} \end{vmatrix} = 0 \quad \text{pour } 0 \leqslant i_0 < i_1 < \cdots < i_r \leqslant n;$$

Ces relations, qui caractérisent les points P de V, s'appellent les **équations de la variété linéaire** V par rapport à la base affine considérée.

Le cas le plus simple est celui où $r = n$, de sorte que V est de dimension $n - 1$ (on dit alors que V est un **hyperplan** dans l'espace affine E); les relations (17) se réduisent alors à la seule et unique équation

$$\begin{vmatrix} \xi_0 & \alpha_{10} & \cdots & \alpha_{n0} \\ \xi_1 & \alpha_{11} & \cdots & \alpha_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_n & \alpha_{1n} & \cdots & \alpha_{nn} \end{vmatrix} = 0;$$

comme on ne modifie pas la valeur de ce déterminant en ajoutant à la première ligne la somme des autres (§ 24, n° 1, propriété (*e*)), et comme la somme des coordonnées affines d'un point est toujours égale à 1, on voit qu'on peut remplacer la relation précédente par l'équation

$$(18) \qquad \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi_1 & \alpha_{11} & \cdots & \alpha_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ \xi_n & \alpha_{1n} & \cdots & \alpha_{nn} \end{vmatrix} = 0;$$

on dit que (18) est l'équation de l'hyperplan engendré par les points (indépendants) $Q_1, \ldots, Q_n$.

Polynomes et équations algébriques

Les méthodes développées dans les Chapitres précédents avaient principalement pour but l'étude des systèmes d'équations *linéaires*. Mais tout le monde sait qu'il est aussi important — et beaucoup plus difficile — d'étudier les systèmes d'équations *algébriques*, i.e. ceux dont les premiers membres sont des combinaisons linéaires de monômes en les inconnues considérées. L'étude générale de ses systèmes est l'un des objectifs fondamentaux de l'Algèbre, et conduit à la Géométrie Algébrique, actuellement l'une des branches les plus actives des Mathématiques.

Le but du présent Chapitre est d'introduire des notions très élémentaires et de toute façon indispensables dans toutes les Mathématiques, par exemple celles de relation algébrique, de polynôme, de fraction rationnelle, d'équation algébrique, etc...

Comme dans les §§ précédents, nous avons cherché à introduire ces notions avec le degré de généralité maximum, en sorte que, par exemple, nous définissons des polynômes à plusieurs variables à coefficients dans un anneau commutatif K arbitraire. Ici encore il ne s'agit pas d'une généralisation séduisante mais par ailleurs gratuite; la notion de polynôme à coefficients dans un anneau qui n'est pas un corps est indispensable si l'on veut pouvoir, dans la théorie des polynômes à n variables, raisonner par récurrence sur le nombre n (les polynômes à n variables sont des polynômes à une variable à coefficients dans l'anneau, *qui n'est pas un corps*, des polynômes à $n-1$ variables), et il y a naturellement d'autres raisons plus sérieuses de se placer à ce niveau de généralité — sans parler du fait que, comme toujours, on ne simplifierait pratiquement pas l'exposé en supposant que l'anneau de base est le corps des nombres réels par exemple. (On pourrait alors se dispenser de faire la distinction entre « polynômes » et « fonctions polynomiales », mais non de prouver qu'une fonction polynomiale qui est « identiquement nulle » a tous ses coefficients nuls.)

Notons que les §§ 26 à 33 peuvent être étudiés par le lecteur qui n'a lu que les §§ 0 à 12, à l'exception du n° 3 du § 26 qui de toute façon n'est pas pour les lecteurs débutants.

§ 26. Relations algébriques

1. *Monômes et polynômes en les éléments d'un anneau*

Soit L un anneau commutatif. Étant donnés un sous-anneau K de L et une partie B de L, considérons les sous-anneaux de L qui contiennent à la fois K et B; l'intersection de ces sous-anneaux est encore un sous-anneau contenant K et B, et c'est évidemment le *plus petit* sous-anneau de L contenant à la fois K et B. On dit que c'est le **sous-anneau de L engendré par K et B**, et on le désigne par la notation

$$K[B].$$

Considérons par exemple le cas où B se compose de n éléments $x_1, \ldots, x_n$ — on utilise alors la notation

$$K[x_1, \ldots, x_n]$$

pour désigner le sous-anneau engendré par K et les x_i. Il est clair que ce sous-anneau contient tout **monôme** en $x_1, \ldots, x_n$, i.e. tout élément de L qui peut s'écrire

$$x_1^{r_1} \ldots x_n^{r_n}$$

avec des exposants entiers r_i positifs ou nuls. Il contient aussi le produit d'un tel monôme par un élément de K, donc toute somme d'un nombre fini de tels produits, autrement dit tout **polynôme en $x_1, \ldots, x_n$ à coefficients dans K**; on appelle ainsi tout $y \in L$ qui peut se mettre sous la forme

$$(1) \qquad y = \sum_{r_1, \ldots, r_n \geq 0} a_{r_1 \ldots r_n} x_1^{r_1} \ldots x_n^{r_n},$$

avec des coefficients $a_{r_1 \ldots r_n} \in K$ presque tous nuls (*). Un tel polynôme s'écrit encore, si l'on préfère, sous la forme

$$(1\ bis) \qquad y = a + \sum_{1 \leq i \leq n} a_i x_i + \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j + \sum_{1 \leq i,j,k \leq n} a_{ijk} x_i x_j x_k + \cdots$$

(*) Cela signifie (§ 11, n° 5) que les systèmes $(r_1, \ldots, r_n)$ tels que l'on ait $a_{r_1 \ldots r_n} \neq 0$ sont en nombre fini. Cette condition est indispensable pour donner un sens au second membre de (1) en tant que somme d'éléments de L.

avec des « coefficients » a, a_i, a_{ij}, a_{ijk}, ... dans K et presque tous nuls; il suffit pour le voir de considérer dans (1) les termes pour lesquels $r_1 + \cdots + r_n = 0$, puis ceux pour lesquels $r_1 + \cdots + r_n = 1$, puis ceux pour lesquels $r_1 + \cdots + r_n = 2$, et ainsi de suite.

Non seulement l'anneau $K[x_1, \ldots, x_n]$ contient tous les polynômes (1), mais tous les éléments de cet anneau sont de tels polynômes — autrement dit $K[x_1, \ldots, x_n]$ *est exactement l'ensemble des* $y \in L$ *qui peuvent s'écrire sous la forme* (1). En effet, comme le produit de deux monômes en les x_i est encore un monôme en les x_i, il est immédiat de voir que l'ensemble L′ des polynômes (1) est un sous-anneau de L contenant K et les x_i, donc contenant $K[x_1, \ldots, x_n]$; mais comme on a aussi, d'après ce qui précède, l'inclusion opposée, on voit finalement que $L' = K[x_1, \ldots, x_n]$ comme annoncé.

Le cas le plus simple est celui où $n = 1$, i.e. où B contient un seul élément x; les polynômes en x à coefficients dans K, qui constituent le sous-anneau $K[x]$ de L engendré par K et x, sont alors les $y \in L$ pour lesquels il existe un entier $p \geqslant 0$ et des $a_0, \ldots, a_p \in K$ tels que l'on ait

$$y = a_0 + a_1 x + \cdots + a_p x^p.$$

Remarque 1. Le lecteur débutant, en dépit des idées préconçues qu'il pourrait avoir sur les polynômes, fera bien d'observer qu'ici les lettres x ou $x_1, \ldots, x_n$ désignent des éléments *fixes* de l'anneau L, et non pas des « variables ».

Exemple 1. On a $\mathbf{C} = \mathbf{R}[i]$ puisque tout nombre complexe s'écrit sous la forme $a + bi$ avec $a, b \in \mathbf{R}$.

Exemple 2. Prenons $K = \mathbf{Q}$, $L = \mathbf{R}$ et considérons le sous-anneau $K[x]$ où

$$x = \sqrt[3]{2};$$

les puissances successives de x sont

$$1, x, x^2, 2, 2x, 2x^2, 4, 4x, 4x^2, 8, \ldots$$

et par suite le sous-anneau $\mathbf{Q}[x]$ de $\mathbf{R}$ est l'ensemble des nombres réels qui peuvent s'écrire sous la forme

$$a + bx + cx^2 \qquad \text{avec} \qquad a, b, c \in \mathbf{Q}.$$

Exemple 3. Soient K un anneau commutatif et L l'anneau des applications de K dans K (§ 8, *Exemple* 3); on regarde K comme un sous-anneau de L en associant à chaque élément a de K la fonction constante donnée par

$$f(t) = a \quad \text{pour tout } t \in K.$$

Désignons alors par x l'application identique de K dans K, donnée par

$$x(t) = t \quad \text{pour tout } t \in K.$$

Les éléments du sous-anneau K[x] de L sont appelés les **fonctions polynomiales sur l'anneau K**. Ce sont évidemment les applications f de K dans K pour lesquelles il existe un entier $r \geqslant 0$ et des éléments $a_0, \ldots, a_r$ de K tels que l'on ait

$$f(t) = a + a_1 t + \cdots + a_r t^r \quad \text{pour tout } t \in \mathrm{K}.$$

Cette notion sera généralisée au § 28.

2. *Relations algébriques*

Soient L un anneau commutatif, K un sous-anneau de L, et $x_1, \ldots, x_n$ des éléments de L en nombre fini. On appelle **relation algébrique entre $x_1, \ldots, x_n$ à coefficients dans K** toute relation linéaire à coefficients dans K entre les monômes en $x_1, \ldots, x_n$, autrement dit toute famille $(a_{r_1 \ldots r_n})_{r_1, \ldots, r_n \geqslant 0}$ d'éléments presque tous nuls de K tels que l'on ait

$$\sum_{r_1, \ldots, r_n \geqslant 0} a_{r_1 \ldots r_n} x_1^{r_1} \ldots x_n^{r_n} = 0. \tag{2}$$

Lorsque la relation (2) n'a lieu que si *tous* les coefficients $a_{r_1 \ldots r_n}$ sont nuls, on dit que $x_1, \ldots, x_n$ sont **algébriquement indépendants sur** K; dans le cas contraire, i.e. s'il existe au moins une relation (2) non triviale, on dit que $x_1, \ldots, x_n$ sont **algébriquement liés sur K.**

Prenons en particulier $n = 1$, i.e. un seul $x \in \mathrm{L}$. Si x est algébriquement lié sur K, i.e. s'il existe une relation de la forme

$$a_0 + a_1 x + \cdots + a_p x^p = 0 \tag{3}$$

pour au moins un entier $p \geqslant 0$ et des coefficients $a_i \in \mathrm{K}$ non tous nuls, on dit que x est **algébrique sur** K. Dans le cas contraire, on dit que x est **transcendant sur K.**

Exemple 4. Le nombre complexe i est algébrique sur **R** et même sur **Q** puisqu'il vérifie la relation

$$i^2 + 1 = 0.$$

Exemple 5. Prenons $\mathrm{K} = \mathbf{Q}$, $\mathrm{L} = \mathbf{C}$ et

$$x = \sqrt[3]{2} - \sqrt{3};$$

alors x est algébrique sur **Q**; on a en effet $x^3 = 2 - \sqrt{3}$ et par suite

$$(x^3 - 2)^2 = 3,$$

ce qui s'écrit encore

$$x^6 - 4x^3 + 1 = 0.$$

On appelle **nombres algébriques** (§ 11, *Exemple* 11) les éléments du corps C des nombres complexes qui sont algébriques sur le corps **Q** des nombres ration-

nels, i.e. les $z \in \mathbf{C}$ qui vérifient au moins une relation de la forme

$$a_r z^r + \cdots + a_1 z + a_0 = 0$$

où les a_i sont des nombres *rationnels* non tous nuls (on peut du reste les supposer *entiers* en chassant les dénominateurs).

Exemple 6. On appelle **nombre transcendant** tout nombre complexe qui n'est pas algébrique. Les premiers exemples de tels nombres ont été donnés par Liouville en 1844. En 1882, Lindemann a démontré que le nombre $\pi = 3,14159...$ est transcendant, résultat qui implique l'impossibilité de résoudre par l'affirmative le problème dit de la « quadrature du cercle », contrairement à ce que la plupart des mathématiciens avaient conjecturé depuis l'Antiquité.

On peut également démontrer (Hermite, 1873) que le nombre $e = 2,71828...$ utilisé en Analyse est transcendant.

Exemple 7. Soient L l'anneau de toutes les applications de **R** dans **R** et K le sous-anneau de L formé des fonctions polynomiales (*Exemple* 3), i.e. le sous-anneau de L engendré par les fonctions constantes et la fonction x donnée par $x(t) = t$ pour tout $t \in \mathbf{R}$. On dit qu'une fonction $f \in \mathrm{L}$ est **algébrique** si les éléments x et f de L sont liés algébriquement sur le sous-anneau **R** de L, autrement dit s'il existe des constantes a_{pq} presque toutes nulles telles que l'on ait

$$\sum_{p,\, q \geqslant 0} a_{pq} t^p f(t)^q = 0 \quad \text{pour tout } t \in \mathbf{R}. \tag{4}$$

Il est clair par exemple que la fonction f donnée par

$$f(t) = \sqrt[3]{t^2 - 1}$$

est algébrique. Une fonction qui n'est pas algébrique est dite **transcendante**; c'est par exemple le cas de la fonction $f(t) = \sin t$; supposons en effet que celle-ci vérifie la relation (4); posant

$$f_p(t) = \sum_{q \geqslant 0} a_{pq} . \sin^q t$$

la relation en question s'écrit

$$\sum_p f_p(t) . t^p = 0,$$

et vu la périodicité de la fonction sinus on aura aussi

$$\sum_p f_p(t) . (t + 2\pi n)^q = 0$$

pour tout entier n; pour $t \in \mathbf{R}$ donné, le polynôme $\sum f_p(t) . x^q$ s'annule donc pour une infinité de valeurs de x, ce qui entraîne, comme on le montrera plus loin (§ 32, n° 4), que ses coefficients $f_p(t)$ sont tous nuls. Ainsi la relation (4) implique

$$\sum_q a_{pq} . \sin^q t = 0$$

quels que soient $p \geqslant 0$ et $t \in \mathbf{R}$; mais alors, pour tout entier $p \geqslant 0$, le polynôme $\sum a_{pq}x^q$ s'annule pour une infinité de valeurs de x (à savoir dès qu'on peut écrire $x = \sin t$, i.e. pour tout x tel que $-1 \leqslant x \leqslant +1$), donc a tous ses coefficients nuls, et on voit en définitive que la relation (4) implique $a_{pq} = 0$ quels que soient p et q, ce qui montre comme annoncé que sin t est une fonction transcendante.

3. *Cas des corps commutatifs*

Démontrons d'abord le résultat suivant :

THÉORÈME 1. *Soient* L *un corps commutatif,* K *un sous-corps de* L, *et* x *un élément de* L. *Les propriétés suivantes sont équivalentes :*

a) x *est algébrique sur* K;

b) K$[x]$ *est de dimension finie en tant qu'espace vectoriel sur* K;

c) K$[x]$ *est un sous-corps de* L.

Supposons x algébrique sur K; on a alors une relation

$$c_0 + c_1x + \cdots + c_px^p = 0$$

avec des coefficients $c_i \in$ K non tous nuls. On peut supposer $c_p \neq 0$, et comme K est un corps on déduit de la relation précédente une relation de la forme

$$x^p = a_0 + a_1x + \cdots + a_{p-1}x^{p-1} \tag{5}$$

avec des coefficients

$$a_i = -c_i/c_p$$

dans K. On va en déduire plus généralement que, pour tout $n \geqslant 0$, le monôme x^n est une combinaison linéaire à coefficients dans K des éléments $1, x, \ldots, x^{p-1}$ de K$[x]$. C'est évident pour $n = 0$, ce qui permet de raisonner par récurrence sur n; supposant établie l'existence d'une relation de la forme

$$x^{n-1} = d_0 + d_1x + \cdots + d_{p-1}x^{p-1}$$

à coefficients $d_i \in$ K, il vient, en tenant compte de (5),

$$x^n = x.x^{n-1} = d_0x + \cdots + d_{p-2}x^{p-1} + d_{p-1}x^p = d_0x + \cdots + d_{p-2}x^{p-1} + d_{p-1}(a_0 + a_1x + \cdots + a_{p-1}x^{p-1}),$$

ce qui prouve évidemment notre assertion. On voit donc que, si x est algébrique, il existe un entier p tel que *toute* puissance de x soit combinaison linéaire, à coefficients dans K, des puissances

$$1, x, \ldots, x^{p-1};$$

il s'ensuit évidemment que ces p éléments engendrent K$[x]$ regardé comme espace vectoriel sur K, et ceci montre que l'assertion *a*) de l'énoncé implique l'assertion *b*).

Montrons maintenant que *b*) implique *c*). Posons K$[x] =$ F et, pour un $a \in$ F

non nul, considérons l'application

$$f : \mathbf{F} \to \mathbf{F}$$

donnée par

$$f(u) = au \quad \text{pour tout } u \in \mathbf{F};$$

regardant F comme un espace vectoriel sur K, il est clair que f est un endomorphisme de F. Le noyau de f est formé des $u \in \mathbf{F}$ tels que $au = 0$; comme F est un anneau d'intégrité (comme sous-anneau du *corps* L) et comme $a \neq 0$, on voit que Ker (f) se réduit à zéro, donc que f est *injective*. Si F est de dimension finie, on en déduit (§ 19, Corollaire 1 du Théorème 13) que f est *surjective*, et en particulier qu'il existe un $u \in \mathbf{F}$ tel que $au = 1$; ceci montre que tout élément non nul de l'anneau F est inversible dans F, donc que F est un sous-corps de L.

Il reste à établir que *c*) implique *a*). Or si K[x] est un sous-corps de L, l'inverse (dans L) de tout élément non nul de K[x] est dans K[x]; en particulier K[x] contient l'inverse de x; on a donc une relation de la forme

$$x^{-1} = a_0 + a_1 x + \cdots + a_r x^r$$

avec des coefficients $a_j \in \mathbf{K}$, et comme cette relation s'écrit aussi

$$a_r x^{r+1} + \cdots + a_0 x - 1 = 0,$$

on voit que x est algébrique sur K, ce qui achève la démonstration.

Exemple 8. Prenons $\mathbf{K} = \mathbf{Q}$ et $\mathbf{L} = \mathbf{C}$; soit x un nombre algébrique, i.e. un élément de $\mathbf{C}$ vérifiant une relation de la forme

$$a_n x^n + \cdots + a_1 x + a_0 = 0$$

à coefficients rationnels $a_0, \ldots, a_n$ non tous nuls (*Exemple* 5); alors l'ensemble des nombres complexes qui peuvent s'écrire sous la forme

$$c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}$$

avec des coefficients rationnels $c_0, \ldots, c_{n-1}$, est un *sous-corps* de $\mathbf{C}$. On comparera ce résultat à l'*Exemple* 2 du § 8.

Théorème 2. *Soient* L *un corps commutatif,* K *un sous-corps de* L, *et* $\mathbf{K}$ *l'ensemble des éléments de* L *qui sont algébriques sur* K. *Alors* $\mathbf{K}$ *est un sous-corps de* L.

Supposons x et y algébriques sur K. Il existe des entiers p et q tels que toute puissance de x (resp. y) soit combinaison linéaire, à coefficients dans K, des puissances

$$1, x, \ldots, x^{p-1} \quad (\text{resp. } 1, y, \ldots, y^{q-1});$$

il s'ensuit évidemment, par multiplication, que *tous* les monômes $x^i y^j$ sont des combinaisons linéaires, à coefficients dans K, des pq monômes

$$x^i y^j \quad (0 \leqslant i \leqslant p-1,\ 0 \leqslant j \leqslant q-1).$$

Donc le sous-anneau $K[x, y]$ de L est de dimension finie sur K. Pour tout $z \in K[x, y]$ le sous-anneau $K[z]$ est contenu dans $K[x, y]$, donc est *a fortiori* de dimension finie sur K; par suite, tout élément de $K[x, y]$ est algébrique sur K; en particulier, $x - y$ et xy sont algébriques sur K, ce qui montre déjà que $\overline{K}$ est un sous-anneau de L.

Pour achever la démonstration, il reste à prouver que si un $x \neq 0$ est algébrique sur K, il en est de même de x^{-1}; or si l'on a une relation

$$a_n x^n + \cdots + a_1 x + a_0 = 0$$

à coefficients dans K, il est clair que l'élément $x^{-1} = y$ vérifie

$$a_n + \cdots + a_1 y^{n-1} + a_0 y^n = 0,$$

ce qui achève la démonstration.

Exemple 9. L'ensemble des nombres algébriques est un sous-corps de $\mathbf{C}$.

Le lecteur désireux de s'informer sur ces questions, qui jouent un rôle essentiel en Algèbre « supérieure », trouvera de nombreux résultats complémentaires dans les exercices de ce § et des §§ suivants, ainsi que dans certains des ouvrages cités dans la Bibliographie (Van der Waerden, Samuel-Zariski, Lang).

§ 27. Anneaux de polynômes

Soient K un anneau commutatif et n un entier positif. On se propose dans ce § de construire un anneau commutatif L et n éléments $X_1, \ldots, X_n$ de L tels que les conditions suivantes soient remplies :

(AP 1) K *est un sous-anneau de* L;

(AP 2) *les éléments* $X_1, \ldots, X_n$ *sont algébriquement indépendants sur* K;

(AP 3) L *est engendré par* K *et* $X_1, \ldots, X_n$.

Avant d'étudier le cas général, nous résoudrons ce problème dans le cas particulier où $n = 1$: il s'agit alors de construire un sur-anneau commutatif L de K et un $X \in L$ transcendant sur K (§ 26, n° 2) tel que $L = K[X]$.

1. Préliminaires sur le cas d'une variable

Supposons construit un sur-anneau commutatif L de K et un $X \in L$ transcendant sur K, tel que $L = K[X]$. Tout $f \in L$ s'écrit alors *d'une façon et d'une seule*

$$f = a_0 + a_1X + \cdots + a_nX^n + \cdots \tag{1}$$

avec des $a_n \in K$ presque tous nuls; en effet, comme $L = K[X]$, les puissances

$$1, X, \ldots, X^n, \ldots$$

engendrent L regardé comme module sur K, et comme X est transcendant sur K ces puissances sont linéairement indépendantes sur K (autrement dit, forment une *base* de L sur K au sens du § 11, n° 5). Réciproquement, il est clair que toute suite $(a_n)_{n \geq 0}$ d'éléments presque tous nuls de K définit, par la formule (1), un élément de L.

Soient

$$f = a_0 + a_1X + \cdots, \qquad g = b_0 + b_1X + \cdots$$

deux éléments de L, et posons

$$f + g = c_0 + c_1X + \cdots$$
$$fg = d_0 + d_1X + \cdots;$$

il est clair qu'on a

$$c_n = a_n + b_n \tag{2}$$

pour tout $n \geqslant 0$. Pour calculer fg, on multiplie chaque terme $a_p X^p$ de f par chaque terme $b_q X^q$ de g, et on ajoute les résultats obtenus; pour obtenir ainsi un terme en X^n, il faut prendre $p + q = n$, et on obtient alors la contribution $a_p b_q$ pour calculer le coefficient d_n; ainsi

$$d_n = \sum_{p+q=n} a_p b_q = a_n b_0 + a_{n-1} b_1 + \cdots + a_1 b_{n-1} + a_0 b_n. \tag{3}$$

L'intérêt des formules (2) et (3) est qu'elles permettent de calculer $f + g$ et fg sans faire intervenir X; elles vont nous servir de point de départ pour la construction effective de l'anneau K[X] cherché.

2. *Polynômes à une indéterminée*

A partir de K, nous allons maintenant construire effectivement un anneau L satisfaisant aux conditions imposées (ce dernier point sera établi au n° suivant; dans le présent n° on va simplement définir les éléments de L, et les opérations algébriques sur ces éléments).

Par définition, L sera l'ensemble des suites

$$f = (a_n)_{n \geqslant 0} = (a_0, a_1, \ldots)$$

d'éléments *presque tous nuls* de K; une telle suite sera appelée un **polynôme à une indéterminée à coefficients dans** K, et les éléments a_n seront appelés les **coefficients** du polynôme f.

Étant donné un polynôme $f = (a_n)_{n \geqslant 0}$ à coefficients dans K, on appelle **degré** de f le plus grand entier $d \geqslant 0$ tel que l'on ait $a_d \neq 0$; cette définition, toutefois, est en défaut dans le cas où l'on a $a_n = 0$ pour tout $n \geqslant 0$; dans ce cas, on appelle degré de f le symbole $-\infty$ [ce symbole ne désigne naturellement pas un entier ordinaire; c'est un nouvel objet mathématique que nous utiliserons en observant les deux conventions suivantes : on posera, par définition

$$-\infty + n = -\infty \quad \text{si} \quad n \in \mathbf{Z} \quad \text{ou si} \quad n = -\infty,$$

et, d'autre part, on conviendra que

$$-\infty \leqslant n \quad \text{si} \quad n \in \mathbf{Z} \quad \text{ou si} \quad n = -\infty;$$

toute autre opération faisant intervenir le symbole $-\infty$, par exemple l'expression

$$-\infty - (-\infty),$$

sera considérée comme dépourvue de sens].

Le degré d'un polynôme f se désigne par la notation

$$d^0(f).$$

Dans l'ensemble des polynômes à une indéterminée à coefficients dans K, on va définir une **addition** et une **multiplication** comme suit : étant donnés deux polynômes

$$f = (a_n)_{n\geqslant 0}, \qquad g = (b_n)_{n\geqslant 0},$$

les polynômes

$$f + g = (c_n)_{n\geqslant 0}, \qquad fg = (d_n)_{n\geqslant 0}$$

sont donnés par les relations

$$c_n = a_n + b_n \tag{2}$$

$$d_n = a_n b_0 + a_{n-1} b_1 + \cdots + a_1 b_{n-1} + a_0 b_n = \sum_{r+s=n} a_r b_s \tag{3}$$

du n° précédent. Pour justifier ces définitions, on doit montrer que les suites (c_n) et (d_n) définies par (2) et (3) sont encore des polynômes, i.e. qu'on a $c_n = d_n = 0$ lorsque l'entier n est suffisamment grand. Or soient p et q les degrés de f et g; si l'on a $n > \mathrm{Max}\,(p, q)$, il vient $a_n = b_n = 0$, donc $c_n = 0$, ce qui montre bien que $f + g$ est un polynôme, et même que

$$d^0(f + g) \leqslant \mathrm{Max}\,(d^0(f), d^0(g)); \tag{4}$$

supposons d'autre part $n > p + q$; alors, dans le terme général $a_r b_s$ de l'expression (3), on a soit $r > p$ (et donc $a_r = 0$), soit $s > q$ (et donc $b_s = 0$); donc tous les termes de (3) sont nuls; ce raisonnement montre non seulement que fg est aussi un polynôme, mais en outre que

$$d^0(fg) \leqslant d^0(f) + d^0(g). \tag{5}$$

Nous devons maintenant vérifier que l'ensemble L, muni des lois de composition qu'on vient de définir, est un *anneau commutatif*. Les calculs à effectuer pour ce faire sont en principe triviaux, et du même ordre que ceux du § 9, n° 3; nous laisserons donc au lecteur le soin de vérifier lui-même la plupart des axiomes, et nous bornerons ici à établir que la multiplication est associative.

Soient pour cela

$$f = (a_n)_{n\geqslant 0}, \qquad g = (b_n)_{n\geqslant 0}, \qquad h = (c_n)_{n\geqslant 0}$$

trois polynômes à coefficients dans K. Posons

$$fg = (u_n)_{n\geqslant 0}, \qquad gh = (v_n)_{n\geqslant 0}.$$

Le coefficient d'indice n de $(fg)h$ est alors égal à

$$u_n c_0 + \cdots + u_0 c_n \tag{6}$$

et celui de $f(gh)$ à

$$a_n v_0 + \cdots + a_0 v_n, \tag{7}$$

de sorte que tout revient à établir que les expressions (6) et (7) sont égales quel que soit n. Or u_p est la somme des produits $a_i b_j$ tels que $i + j = p$; par suite (6) est la somme des expressions

$$(a_i b_j) c_k \quad \text{telles que} \quad (i + j) + k = n;$$

d'autre part, v_q est la somme des produits $b_j c_k$ tels que $j + k = q$; donc (7) est la somme des expressions

$$a_i (b_j c_k) \quad \text{telles que} \quad i + (j + k) = n;$$

l'égalité de (6) et (7) résulte évidemment de ces considérations.

On vérifie d'autre part aisément que les éléments 0 et 1 de L sont donnés par les relations

$$0 = (0, 0, \ldots)$$
$$1 = (1, 0, 0, \ldots).$$

3. *La notation polynomiale*

Montrons maintenant qu'*on peut identifier* K *à un sous-anneau de* L. Pour cela, définissons une application

$$j : K \to L$$

en posant

$$j(a) = (a, 0, 0, \ldots)$$

pour tout $a \in K$. Évidemment j est injective; et on vérifie trivialement, à l'aide des formules (2) et (3), que

$$j(a + b) = j(a) + j(b), \qquad j(ab) = j(a)\, j(b),$$
$$j(0) = 0, \qquad j(1) = 1.$$

Il s'ensuit que j est un isomorphisme de K sur un sous-anneau de L et qu'il revient au même, pour effectuer des calculs algébriques sur des éléments de K, de remplacer ceux-ci par leurs images par j, et d'effectuer les calculs en question sur les éléments de L ainsi obtenus. Par suite, il est naturel de considérer comme *identiques* un élément a de K et l'élément correspondant de L; autrement dit, nous écrirons dorénavant

$$(8) \qquad a = (a, 0, 0, \ldots) \quad \text{pour tout } a \in K.$$

Les éléments de L ainsi obtenus sont évidemment les polynômes de degré 0 à coefficients dans K (à ceci près que l'élément de L qui correspond à l'élément 0 de K est de degré $-\infty$, et non pas 0). Ces éléments de L s'appellent souvent des **constantes**.

Remarque 1. La notion de « constante » qu'on vient de définir est relative à un anneau de base K : c'est un polynôme, nul ou de degré 0, à coefficients dans K, ou, si l'on veut, un élément de K (regardé comme polynôme à coefficients dans K). La terminologie utilisée ici recevra au § suivant son explication intuitive.

On notera que puisque K est un sous-anneau de L, on peut regarder L comme un module sur K, le produit d'un $a \in K$ et d'un $f \in L$ étant le produit de f et de l'élément (8) de L. En utilisant (3), on trouve facilement que

$$a.(b_0, b_1, \ldots) = (ab_0, ab_1, \ldots). \tag{9}$$

Construisons maintenant un élément de L transcendant sur K. Nous poserons

$$X = (0, 1, 0, 0, \ldots). \tag{10}$$

En utilisant la formule (3), on trouve facilement que

$$X^2 = (0, 0, 1, 0, 0, \ldots)$$
$$X^3 = (0, 0, 0, 1, 0, \ldots),$$

et ainsi de suite; d'où, en utilisant (9), et si $a_0, a_1, \ldots$ sont des éléments de K, les formules

$$a_0 = (a_0, 0, 0, 0, \ldots)$$
$$a_1X = (0, a_1, 0, 0, \ldots)$$
$$a_2X^2 = (0, 0, a_2, 0, 0, \ldots)$$
$$a_3X^3 = (0, 0, 0, a_3, 0, \ldots)$$

et ainsi de suite. Si les a_i sont presque tous nuls, on trouve donc

$$a_0 + a_1X + a_2X^2 + \cdots = (a_0, a_1, a_2, \ldots). \tag{11}$$

Ce résultat, compte tenu de la relation

$$0 = (0, 0, 0, \ldots),$$

montre que le premier membre de (11) ne peut être nul que si $a_0 = a_1 = a_2 = \cdots = 0$: par suite, X est transcendant sur K.

En outre, la relation (11) montre que *tout* élément de L est un polynôme en X à coefficients dans K (au sens du § 26, n° 1), autrement dit que

$$L = K[X].$$

Ceci montre que l'anneau L satisfait bien aux conditions énoncées au début de ce §.

Dans la pratique, on dit que L est l'**anneau des polynômes à une indéterminée à coefficients dans** K, et pour désigner un élément

$$f = (a_0, a_1, \ldots)$$

de L on utilise *exclusivement* l'écriture

$$f = a_0 + a_1X + a_2X^2 + \cdots, \tag{12}$$

justifiée par la relation (11). Autrement dit, à partir de maintenant, le lecteur peut négliger les considérations du n° précédent, qui ne servent qu'à *définir* les polynômes;

pour tout ce qui suit (et tout ce qu'on peut faire des polynômes en Mathématiques), il n'y a rien de plus à retenir que les conditions (AP 1), (AP 2) et (AP 3) énoncées plus haut — autrement dit : *un polynôme à une indéterminée à coefficients dans* K *est un objet qui peut s'écrire d'une façon et d'une seule sous la forme* (12), *avec des coefficients* $a_n \in K$ *presque tous nuls, et on calcule sur les polynômes en les considérant comme des éléments d'un anneau commutatif, i.e. en utilisant les règles de calcul « évidentes ».*

Remarque 2. Le lecteur se demandera peut-être pourquoi nous n'avons pas défini *a priori* un polynôme comme une expression de la forme (12) sur laquelle on effectue des calculs conformément aux règles « évidentes ». La raison en est qu'en procédant ainsi on n'aurait pas pu donner de signification mathématique précise à la lettre X figurant dans (12), en sorte que la « définition » des polynômes qu'on aurait obtenue ainsi n'en serait pas une en réalité.

Remarque 3. Contrairement à des traditions qui ont parfois encore cours, on aura soin de ne pas considérer la lettre X comme représentant un « élément variable » de K; la lettre X désigne le polynôme particulier (10), dont la définition comporte aussi peu d'arbitraire et d'indétermination que possible...

L'idée que X représente un élément variable de K provient de la confusion, qu'on effectue souvent, entre un polynôme à coefficients dans K et une fonction polynomiale (§ 26, *Exemple* 3) sur l'anneau K. Les relations entre ces deux notions seront discutées au § suivant.

Il va de soi par ailleurs qu'au lieu de désigner le polynôme (10) par X, on peut le désigner par toute autre lettre (en pratique on utilise fréquemment les lettres Y, Z, T, etc...), pourvu que la lettre choisie n'ait pas déjà été utilisée par ailleurs à d'autres fins.

4. *Polynômes à plusieurs indéterminées*

Nous allons maintenant construire, par récurrence sur l'entier n, une solution au problème posé au début du présent §.

Désignons par K' un anneau commutatif contenant K comme sous-anneau, et engendré par K et $n-1$ éléments $X_1, \ldots, X_{n-1}$ algébriquement indépendants sur K. Désignons par L l'anneau des polynômes à une indéterminée à coefficients dans K', et par X_n l'indéterminée en question : on a donc

$$L = K'[X_n] = K[X_1, \ldots, X_{n-1}][X_n].$$

Nous allons montrer que L répond aux conditions (AP 1), (AP 2) et (AP 3).

Comme K est un sous-anneau de K', lui-même sous-anneau de L, la vérification de (AP 1) est triviale.

D'autre part, soit L' le sous-anneau de L engendré par K et $X_1, \ldots, X_n$; comme il contient K et $X_1, \ldots, X_{n-1}$, il contient K' ; comme il contient K' et X_n, il est identique à L; par suite L est engendré par K et $X_1, \ldots, X_n$, ce qui est la condition (AP 3).

Considérons enfin une relation

$$\sum a_{r_1 \ldots r_n} X_1^{r_1} \ldots X_n^{r_n} = 0$$

où les éléments $a_{r_1 \dots r_n}$ de K sont presque tous nuls. Introduisons les éléments

$$f_s = \sum_{r_1, \dots, r_{n-1}} a_{r_1 \dots r_{n-1} s} X_1^{r_1} \dots X_{n-1}^{r_{n-1}}$$

de K′; la relation considérée s'écrit

$$\sum_s f_s X_n^s = 0,$$

et comme X_n est transcendant sur K′, il s'ensuit que $f_s = 0$ pour tout $s \geqslant 0$. Mais comme $X_1, \dots, X_{n-1}$ sont algébriquement indépendants sur K, ceci implique

$$a_{r_1 \dots r_{n-1} s} = 0$$

quels que soient $r_1, \dots, r_{n-1}$ et s, ce qui établit (AP 2).

L'anneau

$$L = K[X_1, \dots, X_{n-1}][X_n] = K[X_1, \dots, X_n]$$

que nous venons de définir s'appelle **l'anneau des polynômes à n indéterminées à coefficients dans** K, et ses éléments s'appellent les **polynômes à n indéterminées à coefficients dans** K. Un tel polynôme f s'écrit donc *d'une façon et d'une seule* sous la forme

$$f = \sum a_{r_1 \dots r_n} X_1^{r_1} \dots X_n^{r_n}, \tag{19}$$

avec des **coefficients** $a_{r_1 \dots r_n} \in K$ presque tous nuls; et les calculs sur ces polynômes s'effectuent de façon « évidente », i.e. en les regardant comme des éléments d'un anneau commutatif, ce qu'ils sont en effet !

Si par exemple $n = 2$, auquel cas on désigne le plus souvent les indéterminées X_1 et X_2 par X et Y, un polynôme à deux indéterminées à coefficients dans K est une expression de la forme

$$f = \sum a_{rs} X^r Y^s$$

avec des coefficients $a_{rs} \in K$ presque tous nuls; la sommation est étendue à tous les couples (r, s) d'entiers naturels. En groupant les termes pour lesquels $r + s$ possède une valeur donnée, on voit qu'un tel polynôme peut encore s'écrire sous la forme

$$f = a_{00} + (a_{10}X + a_{01}Y) + (a_{20}X^2 + a_{11}XY + a_{02}Y^2) + (a_{30}X^3 + a_{21}X^2Y + a_{12}XY^2 + a_{03}Y^3) + \cdots,$$

avec bien entendu un nombre fini seulement de termes non nuls.

De même, si $n = 3$, on désigne les indéterminées par X, Y, Z et un polynôme à trois indéterminées à coefficients dans K est une expression

$$f = \sum_{i, j, k \geqslant 0} a_{ijk} X^i Y^j Z^k$$

avec des $a_{ijk} \in \mathbf{K}$ presque tous nuls; un tel polynôme s'écrit aussi

$$f = a_{000} + (a_{100}X + a_{010}Y + a_{001}Z) + (a_{200}X^2 + a_{020}Y^2 + a_{002}Z^2 + a_{011}YZ + a_{101}ZX + a_{110}XY) + \cdots$$

Il va de soi qu'on n'est pas obligé, pour désigner les coefficients d'un polynôme, d'utiliser les notations ci-dessus : en Mathématiques, chacun est libre de choisir ses notations (pourvu qu'elles soient cohérentes); on peut par exemple, au lieu de distinguer les coefficients les uns des autres à l'aide d'indices multiples, utiliser pour les désigner des lettres différentes, et écrire par exemple un polynôme à deux indéterminées sous la forme

$$f = a + bX + cY + dX^2 + eXY + fY^2 + gX^3 + \cdots;$$

l'inconvénient de cette méthode est que l'alphabet latin n'offre qu'un nombre limité de possibilités. On espère en outre que le lecteur détectera la contradiction interne dans les notations que nous venons d'exhiber à l'instant...

5. *Degrés partiels et degré total*

Si, dans la formule (13), on groupe ensemble tous les termes pour lesquels l'exposant r_i a une valeur donnée, on voit que l'on peut regarder f comme un polynôme à une indéterminée X_i, à coefficients dans l'anneau

$$K_i = K[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n].$$

Considérant f comme élément de l'anneau de polynômes $K_i[X_i]$, on peut définir le degré de f; celui-ci s'appelle le **degré de f par rapport à X_i**. Il est donc défini comme suit : on écrit

$$f = \sum_{n \geqslant 0} u_n . X_i^n$$

où les u_n sont des polynômes en les indéterminées autres que X_i, à coefficients dans K; cela fait, le degré de f par rapport à X_i est le plus grand entier n tel que $u_n \neq 0$, ou bien le symbole $-\infty$ si $f = 0$.

On appelle d'autre part **degré total**, ou simplement **degré**, de f le plus grand entier d tel qu'il existe des entiers $r_1, \ldots, r_n$ vérifiant

$$a_{r_1 \ldots r_n} \neq 0, \qquad r_1 + \cdots + r_n = d.$$

On peut encore définir le degré total comme suit. Disons qu'un polynôme à coefficients dans K est **homogène de degré** d si chacun des monômes qu'il contient effectivement (i.e. avec un coefficient non nul) est de degré total d. En groupant ensemble, dans (13), les termes pour lesquels $r_1 + \cdots + r_n$ a une valeur donnée, on voit que *tout polynôme f à coefficients dans* K *s'écrit d'une façon et d'une seule sous la forme*

$$f = f_0 + f_1 + \cdots + f_r + \cdots$$

où f_r est homogène de degré r pour tout $r \geqslant 0$, et nul pour presque tout r. Ceci fait, le degré total de f est aussi le plus grand entier d tel que $f_d \neq 0$ (on convient toutefois d'attribuer au polynôme 0 le degré total $-\infty$).

Le degré total de f se désigne par la notation $d^\circ(f)$, comme dans le cas d'une seule indéterminée. On a aussi les inégalités

$$d^\circ(f+g) \leqslant \mathrm{Max}\,[d^\circ(f), d^\circ(g)] \tag{4}$$

$$d^\circ(fg) \leqslant d^\circ(f) + d^\circ(g). \tag{5}$$

En effet, soient p et q les degrés de f et g; on a donc

$$f = f_0 + \cdots + f_p, \qquad g = g_0 + \cdots + g_q$$

(on désigne d'une façon générale par u_r la somme des monômes de degré total r d'un polynôme u). En ajoutant terme à terme, on voit que $f+g$ ne fait pas intervenir de monômes de degré supérieur au plus grand des entiers p, q, d'où la première relation. D'autre part, fg est somme des produits $f_i g_j$; mais il est clair que $f_i g_j$ est homogène et de degré total $i+j$; par suite, les monômes qui interviennent effectivement dans le produit fg sont de degré total au plus égal à $p+q$, ce qui prouve la seconde inégalité.

6. *Polynômes à coefficients dans un anneau d'intégrité*

Nous allons établir le résultat suivant :

THÉORÈME 1. *Soit* K *un anneau d'intégrité commutatif; alors, pour tout entier n, l'anneau* $K[X_1, \ldots, X_n]$ *est intègre. En outre, quels que soient* $f, g \in K[X_1, \ldots, X_n]$, *on a*

$$d^\circ(fg) = d^\circ(f) + d^\circ(g).$$

Pour montrer que $K[X_1, \ldots, X_n]$ est un anneau d'intégrité, on tient compte de la relation

$$K[X_1, \ldots, X_n] = K[X_1, \ldots, X_{n-1}][X_n];$$

celle-ci permet, en raisonnant par récurrence (*) sur n, de se ramener au cas où $n = 1$.

Soient alors

$$f = a_0 + \cdots + a_p X^p, \qquad a_p \neq 0$$
$$g = b_0 + \cdots + b_q X^q, \qquad b_q \neq 0$$

deux polynômes non nuls à une variable, à coefficients dans K. Il est clair que fg contient un seul terme de degré $p+q$, à savoir

$$a_p b_q X^{p+q};$$

(*) Il est fréquent de procéder ainsi dans la théorie des polynômes à plusieurs indéterminées — mais on ne peut le faire que si l'on admet des polynômes à coefficients dans un *anneau* quelconque, car, même si K est un corps, l'anneau $K[X_1, \ldots, X_{n-1}]$ *n'est pas* un corps.

comme K est intègre, on a $a_p b_q \neq 0$, et par suite $fg \neq 0$. Ceci établit la première assertion du Théorème.

Pour établir la seconde (dans le cas général de plusieurs variables) on écrit

$$f = f_0 + \cdots + f_p, \qquad f_p \neq 0$$
$$g = g_0 + \cdots + g_q, \qquad g_q \neq 0;$$

il est clair que la partie homogène de degré total $p + q$ de fg est $f_p g_q$; or comme on sait déjà que $K[X_1, \ldots, X_n]$ est un anneau d'intégrité, on a $f_p g_q \neq 0$; donc f est de degré total $p + q$, ce qui achève la démonstration.

Remarque 4. A vrai dire nous n'avons démontré la relation

$$d^\circ(fg) = d^\circ(f) + d^\circ(g)$$

que dans le cas où f et g sont non nuls. Si $f = 0$, cette relation s'écrit $-\infty = -\infty + d^\circ(g)$, et résulte des règles de calcul dont on a convenu de se servir au n° 2 en ce qui concerne le symbole $-\infty$.

Remarque 5. Il est clair que, si K n'est pas intègre, $K[X_1, \ldots, X_n]$ ne peut pas l'être non plus. En outre, dans ce cas, la relation

$$d^\circ(fg) = d^\circ(f) + d^\circ(g)$$

peut aussi être en défaut : choisir dans K deux éléments a, b non nuls tels que $ab = 0$, et prendre

$$f = aX, \qquad g = bX;$$

on a

$$d^\circ(fg) = -\infty, \qquad d^\circ(f) + d^\circ(g) = 2.$$

§ 28. Fonctions polynomiales

1. *Valeurs d'un polynôme*

Soient K un anneau commutatif et

$$f = \sum a_{r_1 \dots r_n} X_1^{r_1} \dots X_n^{r_n} \tag{1}$$

un polynôme à n indéterminées à coefficients dans K. Étant donné un sur-anneau commutatif L de K, on appelle **valeur de** f **en un point**

$$u = (u_1, \dots, u_n) \in L^n$$

l'élément

$$\sum a_{r_1 \dots r_n} u_1^{r_1} \dots u_n^{r_n} \tag{2}$$

de L obtenu en remplaçant les lettres $X_1, \dots, X_n$ dans l'expression (1) de f par les éléments $u_1, \dots, u_n$ de L. La valeur de f en u se désigne par l'une ou l'autre des notations

$$f(u), \qquad f(u_1, \dots, u_n).$$

On dit que u est un **zéro** ou, si $n = 1$, une **racine** de f si l'on a

$$f(u) = 0.$$

Étant donnés un anneau commutatif L et un sous-anneau K de L, on appelle **fonction polynomiale à coefficients dans** K **sur** L^n toute application $\varphi : L^n \to L$ telle qu'il existe un polynôme f à n indéterminées, à coefficients dans K, tel que l'on ait

$$\varphi(u) = f(u) \quad \text{pour tout } u \in L^n.$$

Plus généralement, on dit qu'une application $\varphi : L^n \to L'$ est polynomiale à coefficients dans K si l'on a (§ 2, n° 9)

$$\varphi = (\varphi_1, \dots, \varphi_r)$$

où les φ_j sont des fonctions polynomiales à coefficients dans K sur L^n.

Exemple 1. Si l'on prend $K = L$ et $n = 1$, on retrouve évidemment les fonctions du § 26, *Exemple* 3, i.e. les applications de K dans K qui sont de la forme

$$a_0 + a_1 t + \cdots + a_n t^n$$

où $a_0, \ldots, a_n$ sont des éléments « fixes » de K et où t désigne l'élément « variable » de K.

Exemple 2. K et n étant quelconques, prenons pour L un anneau de polynômes à coefficients dans K, autrement dit

$$L = K[Y_1, \ldots, Y_p].$$

Pour tout polynôme

$$f \in K[X_1, \ldots, X_n]$$

et quels que soient

$$u_1, \ldots, u_n \in K[Y_1, \ldots, Y_p],$$

on peut donc définir $f(u_1, \ldots, u_n)$; pour des raisons évidentes, on dit que c'est le **polynôme en** $Y_1, \ldots, Y_p$ **obtenu en substituant les polynômes** $u_1, \ldots, u_n$ **aux indéterminées** $X_1, \ldots, X_n$ **dans** f.

On notera que si l'on prend en particulier $L = K[X_1, \ldots, X_n]$ et

$$u_1 = X_1, \ldots, u_n = X_n,$$

le polynôme $f(u_1, \ldots, u_n)$ ainsi obtenu est évidemment f lui-même, résultat que l'on exprime à l'aide de la relation

$$f = f(X_1, \ldots, X_n);$$

dans la pratique, on utilise fréquemment la notation $f(X_1, \ldots, X_n)$ au lieu de f.

L'emploi de cette notation (qui permet de mettre en évidence les notations utilisées pour désigner les indéterminées figurant dans f) ne devra pas faire oublier au lecteur que les lettres X_i *ne désignent pas* des éléments variables de K; voir la *Remarque* 3 du § précédent.

2. *Somme et produit de fonctions polynomiales*

Soient K un anneau commutatif, L un sur-anneau commutatif de K, et n un entier au moins égal à 1. Étant donnés des éléments $u_1, \ldots, u_n$ de L, considérons l'application

$$v : K[X_1, \ldots, X_n] \to L$$

donnée par

$$v(f) = f(u_1, \ldots, u_n)$$

pour tout polynôme $f \in K[X_1, \ldots, X_n]$. On a évidemment

(3) $\quad v(f) = f$ si $f \in K$ (i.e. si f est une « constante »)

(4) $\quad v(f) = u_i$ si $f = X_i$.

D'autre part, l'application v est un *homomorphisme d'anneaux*; autrement dit, et puisque (3) montre déjà que $v(1) = 1$, on a les relations

$$v(f+g) = v(f) + v(g) \tag{5}$$
$$v(fg) = v(f)v(g) \tag{6}$$

quels que soient f et g. En effet, en posant

$$f = \sum a_{r_1 \dots r_n} X_1^{r_1} \dots X_n^{r_n}, \qquad g = \sum b_{r_1 \dots r_n} X_1^{r_1} \dots X_n^{r_n},$$

on a

$$f + g = h = \sum c_{r_1 \dots r_n} X_1^{r_1} \dots X_n^{r_n}$$

avec

$$c_{r_1 \dots r_n} = a_{r_1 \dots r_n} + b_{r_1 \dots r_n},$$

et par suite

$$h(u_1, \dots, u_n) = \sum c_{r_1 \dots r_n} u_1^{r_1} \dots u_n^{r_n} = \sum a_{r_1 \dots r_n} u_1^{r_1} \dots u_n^{r_n} + \sum b_{r_1 \dots r_n} u_1^{r_1} \dots u_n^{r_n}$$
$$= f(u_1, \dots, u_n) + g(u_1, \dots, u_n),$$

ce qui est la relation (5). La relation (6) s'obtient par des calculs analogues, mais un peu plus compliqués.

Les propriétés (3), (4), (5) et (6) *caractérisent* du reste l'application v comme on le voit immédiatement, et il est par ailleurs clair que *l'image par v de l'anneau de polynômes* $K[X_1, \dots, X_n]$ *n'est autre que le sous-anneau* $K[u_1, \dots, u_n]$ *de* L *engendré par* K *et les éléments* $u_1, \dots, u_n$ *de* L.

Remarque 1. Le noyau de l'homomorphisme v est formé des polynômes f tels que l'on ait $f(u_1, \dots, u_n) = 0$, i.e. des familles $(a_{r_1 \dots r_n})$ d'éléments presque tous nuls de K tels que l'on ait

$$\sum a_{r_1 \dots r_n} u_1^{r_1} \dots u_n^{r_n} = 0;$$

autrement dit, *le noyau de v est formé des relations algébriques entre $u_1, \dots, u_n$ à coefficients dans* K, définies au § 26, n° 2. Dans la pratique, on ne fait aucune différence entre ces relations algébriques et les polynômes f à coefficients dans K tels que $f(u_1, \dots, u_n) = 0$.

Ce qui précède montre en passant que, si $u_1, \dots, u_n$ sont *algébriquement indépendants sur* K, le noyau de v est réduit à 0, en sorte qu'alors v est un *isomorphisme* de l'anneau $K[X_1, \dots, X_n]$ sur le sous-anneau $K[u_1, \dots, u_n]$ de L.

Les formules (5) et (6) peuvent encore s'interpréter comme suit. Pour chaque polynôme $f \in K[X_1, \dots, X_n]$, considérons l'application polynomiale correspondante

$$f^* : L^n \to L,$$

définie par

$$f^*(u_1, \dots, u_n) = f(u_1, \dots, u_n)$$

quels que soient les $u_i \in L$. Les relations (5) et (6) s'écrivent alors

$$(f+g)^* = f^* + g^*,$$
$$(fg)^* = f^* g^*,$$

et montrent que *la somme et le produit de deux fonctions polynomiales sur* L^n, *à coefficients dans* K, *sont encore des fonctions polynomiales à coefficients dans* K. La relation (4) montre évidemment que parmi ces fonctions polynomiales figurent les fonctions coordonnées par rapport à la base canonique de L^n, et (3) que parmi ces fonctions figurent les applications constantes de L^n dans K.

En fait, et si l'on désigne par M l'anneau de toutes les applications de L^n dans L, il est clair que *l'ensemble des applications polynomiales considérées n'est autre que le sous-anneau de* M *engendré d'une part par les fonctions coordonnées, d'autre part par les applications constantes de* L^n *dans* K (applications que l'on identifie généralement aux éléments de K eux-mêmes).

3. Cas d'un corps infini

Soient K un anneau commutatif, f et g deux polynômes à n indéterminées à coefficients dans K, et

$$f^*, g^* : K^n \to K$$

les fonctions polynomiales correspondantes sur K^n. Il se peut que les applications f^* et g^* coïncident bien que les polynômes f et g soient distincts (et c'est précisément la raison pour laquelle on ne peut généralement pas identifier un *polynôme* à coefficients dans K avec une *fonction polynomiale* sur K^n).

Exemple 3. Prenons $n = 1$ et pour K un corps fini à q éléments. Comme le groupe multiplicatif K^* est à $q - 1$ éléments, le Théorème 5 du § 7 montre que l'on a

$$x^{q-1} = 1 \quad \text{pour tout } x \in K^*$$

et par conséquent

$$x^q = x \quad \text{pour tout } x \in K;$$

les polynômes X et X^q, bien qu'évidemment distincts, définissent donc la même fonction polynomiale sur K.

Cette difficulté ne se présente cependant pas dans les cas classiques ($K = \mathbf{Q}$, $\mathbf{R}$ ou $\mathbf{C}$) en vertu du résultat suivant :

THÉORÈME 1. *Soient f et g deux polynômes à n indéterminées à coefficients dans un anneau d'intégrité infini* K. *Pour que l'on ait $f = g$ il faut et il suffit que l'on ait*

$$f(x) = g(x) \quad \textit{pour tout } x \in K^n.$$

En considérant le polynôme $f - g$, il revient évidemment au même de montrer que, pour tout polynôme $h \in K[X_1, \ldots, X_n]$, la relation

$$h(x) = 0 \quad \text{pour tout } x \in K^n$$

implique $h = 0$. Nous allons l'établir en plusieurs étapes.

LEMME 1. *Soit f un polynôme à une indéterminée à coefficients dans un anneau commutatif* K,

Pour qu'un élément a de K *soit racine de f, il faut et il suffit qu'il existe un polynôme $g \in K[X]$ tel que*

$$f(X) = (X - a)g(X).$$

(Cette relation exprime que f est un multiple de $X - a$ dans l'anneau K[X].)

Soit en effet Y une indéterminée distincte de X, et substituons à X le polynôme $a + Y$; on obtient un polynôme $f(a + Y)$ en Y, qui peut donc s'écrire

$$f(a + Y) = u_0 + u_1 Y + \cdots + u_m Y^m$$

avec des $u_j \in K$. Substituant 0 à Y dans cette relation on trouve évidemment

$$f(a) = u_0$$

et par suite on peut écrire

$$f(a + Y) = f(a) + Y.h(Y)$$

pour un certain polynôme h en Y. Substituant $X - a$ à Y dans le résultat obtenu, et posant $h(X - a) = g(X)$, il vient

$$f(X) = f(a) + (X - a)g(X).$$

Ceci montre évidemment que $f(X) = (X - a)g(X)$ si a est racine de f; la réciproque est triviale.

LEMME 2. *Soit f un polynôme de degré $n \geqslant 0$ à une indéterminée, à coefficients dans un anneau commutatif* K. *Supposons* K *intègre; alors f possède au plus n racines dans* K.

Si f est de degré 0, alors f est une constante non nulle et ne possède aucune racine, en sorte que le lemme est vrai pour $n = 0$. On va maintenant considérer le cas général en raisonnant par récurrence sur le degré n de f.

Soit $a \in K$ une racine de f; on peut écrire $f(X) = (X - a)g(X)$, où g est de degré $n - 1$ puisque K est intègre (§ 27, Théorème 1); si b est une autre racine de f dans K, on doit avoir $0 = (b - a)g(b)$, et comme K est intègre on en conclut que les racines de f dans K autres que a sont celles de g; comme g est de degré $n - 1$, il a au plus $n - 1$ racines dans K d'après l'hypothèse de récurrence, et par suite f possède lui-même au plus n racines dans K, ce qui démontre le Lemme.

Le Lemme 2 prouve évidemment le Théorème 1 dans le cas des polynômes à une variable puisqu'il montre que, si l'anneau de base K est intègre, un polynôme $h \in K[X]$ ne peut avoir une infinité de racines dans K que s'il est nul.

Il reste à établir le Théorème 1 dans le cas d'un polynôme h à n variables, en raisonnant par récurrence sur n. On peut écrire

$$h(X_1, \ldots, X_n) = \sum h_r(X_1, \ldots, X_{n-1}) X_n^r$$

avec des polynômes $h_r \in K[X_1, \ldots, X_{n-1}]$. Supposons $h(x) = 0$ pour tout $x \in K^n$; il vient alors évidemment

$$\sum h_r(y) t^r = 0$$

pour tout $y \in \mathrm{K}^{n-1}$ et tout $t \in \mathrm{K}$. Pour $y \in \mathrm{K}^{n-1}$ donné, on voit donc que le polynôme à une indéterminée

$$\sum h_r(y)\mathrm{T}^r$$

est nul en tout point de K, et comme le Théorème 1 est déjà établi pour les polynômes à une indéterminée on en déduit donc que

$$h_r(y) = 0 \quad \text{pour tout } y \in \mathrm{K}^{n-1}$$

et tout r; mais alors l'hypothèse de récurrence montre que $h_r = 0$ pour tout r, et on a bien finalement la relation $h = 0$, ce qui termine la démonstration.

On peut en fait améliorer le Théorème 1; cf. *Exercice* 1.

§ 29. Fractions rationnelles

1. *Corps des fractions d'un anneau d'intégrité : préliminaires*

Il est clair que tout sous-anneau K d'un corps est un anneau d'intégrité. Inversement, peut-on considérer tout anneau d'intégrité comme un sous-anneau d'un corps ? Nous allons montrer dans ce § que ce problème admet une réponse affirmative, dans le cas d'un anneau commutatif tout au moins.

Pour construire un corps contenant un anneau d'intégrité commutatif K donné, supposons d'abord le problème résolu ; autrement dit, supposons construit un corps L dont K soit un sous-anneau. Alors tout élément non nul a de K admet un inverse dans L ; plus généralement, étant donnés deux éléments a, $b \in K$, avec $b \neq 0$, on peut considérer dans L la fraction

$$a/b = ab^{-1}.$$

L'ensemble de ces fractions est un sous-corps commutatif de L *contenant* K. Soit en effet L' cet ensemble, et considérons deux éléments x, y de L' ; on peut donc écrire

$$x = ab^{-1}, \qquad y = cd^{-1}$$

avec des éléments a, b, c, d de K, et $b \neq 0$, $d \neq 0$; un calcul trivial (compte tenu de la *commutativité* de K) montre alors qu'on a les relations

$$x + y = (ad + bc)/bd \tag{1}$$

$$yx = xy = ac/bd, \tag{2}$$

et comme évidemment on a

$$a/1 = a \tag{3}$$

pour tout $a \in K$, on voit déjà que L' est un sous-anneau de L contenant K. Pour établir que L' est un sous-corps de L, on remarque qu'un élément $x = a/b$ de L' est non nul si et seulement si $a \neq 0$; on a alors

$$x^{-1} = (ab^{-1})^{-1} = (b^{-1})^{-1}a^{-1} = ba^{-1} = b/a,$$

ce qui prouve évidemment que x est inversible dans L'.

S'il est possible de « plonger » K dans un corps L, on peut donc supposer (au besoin en utilisant la construction ci-dessus pour remplacer L par L'), que *tout* élément de L peut s'écrire sous la forme

$$x = a/b \quad \text{avec} \quad a, b \in \mathrm{K}, \qquad b \neq 0. \tag{4}$$

Désignons alors par F l'ensemble des couples

$$(a, b) \quad \text{tels que} \quad a, b \in \mathrm{K}, \qquad b \neq 0;$$

l'application

$$\nu : \mathrm{F} \to \mathrm{L}$$

donnée par

$$\nu(a, b) = a/b = ab^{-1}$$

est donc *surjective*; désignons alors par R la relation d'équivalence sur l'ensemble F associée à l'application ν (§ 4, n° 1), autrement dit la relation

$$\nu(a, b) = \nu(c, d),$$

qui s'écrit encore $a/b = c/d$, i.e.

$$ad = bc;$$

alors ν est composée de l'application canonique de F sur F/R, et d'une application ν' de F/R sur L (§ 4, Théorème 2), et il est clair que ν' est *bijective*. Si l'on utilise ν' pour identifier les éléments de L aux éléments de F/R qui leur correspondent, on voit donc qu'on peut regarder F/R comme un *corps* dans lequel les opérations sont définies par les formules (1) et (2) — plus exactement : si des éléments x et y de F/R sont représentés dans F par des couples (a, b) et (c, d), alors $x + y$ sera l'élément de F/R représenté par le couple $(ad + bc, bd)$, et xy l'élément représenté par (ac, bd).

Ces considérations, qui supposaient le problème résolu, vont nous permettre, à partir de l'anneau K, de construire effectivement un corps L contenant K.

2. *Construction du corps des fractions*

Soit donc K un *anneau d'intégrité commutatif*. Nous désignerons par F l'ensemble des couples (a, b) avec $a, b \in \mathrm{K}$, $b \neq 0$. Étant donnés deux éléments (a, b) et (c, d) de F, on désignera par

$$(a, b) \equiv (c, d) \mod \mathrm{R}$$

la relation

$$ad = bc.$$

Nous allons d'abord montrer que R est une *relation d'équivalence* sur F.

Tout d'abord, la relation

$$(a, b) \equiv (a, b) \mod \mathrm{R}$$

est toujours vraie puisqu'elle s'écrit

$$ab = ba.$$

D'autre part, la relation

$$(a, b) \equiv (c, d) \mod R$$

implique la relation

$$(c, d) \equiv (a, b) \mod R,$$

car la première s'écrit $ad = bc$, et la seconde $cb = da$.

Considérons enfin les relations

$$\begin{aligned} (a, b) &\equiv (c, d) \mod R, \\ (c, d) &\equiv (e, f) \mod R, \\ (a, b) &\equiv (e, f) \mod R; \end{aligned}$$

elles s'écrivent respectivement $ad = bc$, $cf = de$ et $af = be$; en multipliant la première par f et la seconde par b (de façon à faire apparaître le terme bcf dans les deux relations considérées), on déduit des deux premières relations la relation $adf = bde$, ou encore $(af - be)d = 0$; *comme* $d \neq 0$ *et comme* K *est un anneau d'intégrité*, ceci implique $af - be = 0$, i.e. $af = be$, et par suite les deux premières relations considérées impliquent la troisième.

Nous avons donc montré que R est une relation d'équivalence sur F, ce qui permet de construire un ensemble quotient F/R; nous désignerons par θ l'application canonique de F sur F/R.

Montrons maintenant qu'*il existe sur l'ensemble* F/R *deux lois de compositions*

$$(x, y) \mapsto x + y \qquad \text{et} \qquad (x, y) \mapsto xy$$

telles que l'on ait

$$\theta(a, b) + \theta(c, d) = \theta(ad + bc, bd) \tag{5}$$

$$\theta(a, b) \cdot \theta(c, d) = \theta(ac, bd) \tag{6}$$

quels que soient (a, b), $(c, d) \in$ F.

Remarquons d'abord que les seconds membres des relations (5) et (6) ont un sens, autrement dit qu'on a $bd \neq 0$: cela provient des inégalités $b \neq 0$, $d \neq 0$, et du fait que K est un *anneau d'intégrité*.

Pour prouver maintenant l'existence d'applications de (F/R) × (F/R) dans F/R vérifiant les conditions (5) et (6), nous utiliserons le Théorème 3 du § 4, en prenant X = Y = Z = F, R = S = T, et pour application

$$f : F \times F \to F$$

soit l'application donnée par

$$f[(a, b), (c, d)] = (ad + bc, bd),$$

soit l'application donnée par

$$f[(a, b), (c, d)] = (ac, bd).$$

En vertu du Théorème en question, tout revient à établir le résultat suivant : *les relations*

$$(a', b') \equiv (a'', b'') \mod R$$

et

$$(c', d') \equiv (c'', d'') \mod R$$

impliquent les relations

$$(7) \qquad (a'd' + b'c', b'd') \equiv (a''d'' + b''c'', b''d'') \mod R$$

et

$$(8) \qquad (a'c', b'd') \equiv (a''c'', b''d'') \mod R$$

Prouvons d'abord la relation (8) ; celle-ci s'écrit

$$a'c'b''d'' = b'd'a''c'';$$

comme on a par hypothèse

$$(9) \qquad a'b'' = b'a'' \qquad \text{et} \qquad c'd'' = d'c'',$$

la relation cherchée s'obtient en multipliant membre à membre les deux relations qu'on vient d'écrire. Quant à la relation (7), qui s'écrit encore

$$(a'd' + b'c')b''d'' = (a''d'' + b''c'')b'd',$$

elle est visiblement équivalente à

$$(a'b'' - a''b')d'd'' + (c'd'' - c''d')b'b'' = 0,$$

et résulte évidemment de (9).

Nous sommes donc bien dans les conditions d'application du Théorème 3 du § 4, et il existe donc sur l'ensemble quotient F/R deux lois de composition vérifiant les conditions (5) et (6), qui du reste déterminent sans ambiguïté les lois de composition en question comme le montre le même Théorème.

Remarque 1. Le lecteur évitera de croire qu'on pourrait simplifier ces raisonnements dans le cas « élémentaire » où il s'agit de construire les nombres rationnels à partir des nombres entiers. Un nombre rationnel peut s'écrire d'une infinité de façons différentes sous la forme d'une fraction, et on ne peut pas définir la somme (par exemple) de deux nombres rationnels en se bornant à poser

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd};$$

pour que cette formule définisse la somme de deux *nombres rationnels* (et non pas seulement de deux *fractions*, notion totalement dépourvue d'intérêt en soi), on doit montrer que, si l'on remplace les fractions a/b et c/d par des fractions équivalentes (i.e. définissant les mêmes nombres rationnels), le second membre est de même remplacé par une fraction équivalente — par exemple, on

doit démontrer que les fractions

$$1/2 + 4/6 = 14/12$$

et

$$4/8 + 2/3 = 28/24$$

sont équivalentes. Le fait qu'on ne se donne généralement pas la peine, dans l'enseignement élémentaire, de faire ces démonstrations ôte toute espèce de valeur mathématique aux définitions (sic) ainsi obtenues de la somme et du produit de deux nombres rationnels, et constitue une escroquerie majeure, destinée à masquer à des enfants innocents et sans défense la difficulté réelle du problème.

En fait, et quelle que soit la méthode adoptée, on est obligé d'utiliser les raisonnements du § 4, Théorème 3, ainsi que les calculs, d'ailleurs fort simples, qui démontrent (7) et (8); une construction qui parviendrait à s'en passer serait immanquablement erronée.

3. *Vérification des axiomes des corps*

Dans ce n° on va montrer que l'ensemble F/R, muni des deux lois de composition définies au n° précédent, *est un corps commutatif*.

Montrons d'abord que l'addition est commutative; il suffit, compte tenu de (5), de remarquer que l'élément $(ad + bc, bd)$ de F ne change pas si l'on permute (a, b) et (c, d), autrement dit si l'on remplace a, b, c, d par c, d, a, b respectivement.

Montrons maintenant que l'addition est associative. Soient

$$x = 0(a, b), \qquad y = 0(c, d), \qquad z = 0(e, f)$$

des éléments de F/R; on a

$$(x + y) + z = 0(ad + bc, bd) + 0(e, f) = 0[(ad + bc)f + bde, bdf]$$
$$x + (y + z) = 0(a, b) + 0(cf + de, df) = 0[adf + b(cf + de), bdf];$$

il suffit donc d'établir que

$$(ad + bc)f + bde = adf + b(cf + de),$$

ce qui est facile...

L'addition admet un élément neutre, à savoir

$$0 = 0(0, 1);$$

on a en effet d'après (5).

$$0(0, 1) + 0(a, b) = 0(0.b + 1.a, 1.a) = 0(a, b).$$

Tout élément

$$x = 0(a, b)$$

de F/R admet enfin un opposé, à savoir

$$-x = 0(-a, b).$$

En effet on a

$$\theta(a, b) + \theta(-a, b) = \theta(0, b^2)$$

en sorte qu'il suffit de montrer que

$$\theta(0, b) = \theta(0, 1) \quad \text{pour tout } b \in K,\ b \neq 0;$$

or cette relation s'écrit $0.1 = b.0$, et est donc trivialement vraie.

Ainsi, nous avons déjà établi que F/R, muni de l'addition, est un groupe commutatif.

Montrons maintenant que dans F/R la multiplication est commutative, associative, et admet un élément neutre. La commutativité est évidente sur la formule (6). L'associativité s'obtient en observant que

$$[\theta(a, b)\theta(c, d)]\theta(e, f) = \theta(ac, bd)\theta(e, f) = \theta[(ac)e, (bd)f]$$

tandis que

$$\theta(a, b)[\theta(c, d)\theta(e, f)] = \theta(a, b)\theta(ce, df) = \theta[a(ce), b(df)];$$

enfin, l'élément

$$1 = \theta(1, 1)$$

de F/R est élément neutre pour la multiplication, comme le montre un calcul trivial.

Pour achever la démonstration du fait que F/R est un corps, il resterait d'une part à vérifier l'identité de distributivité $(x + y)z = xz + yz$ dans F/R, d'autre part à montrer que tout élément non nul de F/R est inversible. On laisse au lecteur le soin d'établir le premier point à titre d'exercice. Quant au second, remarquons d'abord que la relation

$$\theta(a, b) = 0$$

équivaut à $a = 0$, car étant donné que $0 = \theta(0, 1)$, la relation en question s'écrit encore $a.1 = b.0$. Ceci dit soit x un élément non nul de F/R; on a donc

$$x = \theta(a, b) \qquad \text{avec} \qquad a \neq 0, b \neq 0;$$

on a alors le droit de considérer l'élément

$$x^{-1} = \theta(b, a)$$

de F/R, et celui-ci est effectivement inverse de x; on a en effet

$$x.x^{-1} = \theta(ab, ab)$$

de sorte que tout revient à montrer que

$$\theta(e, e) = 1 = \theta(1, 1)$$

pour tout $e \neq 0$; mais c'est clair, puisque la relation considérée s'écrit $e.1 = 1.e$.

4. *Immersion de l'anneau* K *dans son corps des fractions*

Il nous reste, pour résoudre entièrement le problème posé au n° 1, à montrer comment on peut considérer K comme un sous-anneau du corps F/R ; ici comme en beaucoup d'autres circonstances, nous ne montrerons pas que K est à proprement parler un sous-anneau (ni même un sous-ensemble) de F/R, mais nous construirons un isomorphisme « canonique » de K sur un sous-anneau de F/R (le lecteur aura intérêt à se souvenir de la façon dont on identifie les nombres réels à des nombres complexes ou les éléments d'un anneau commutatif à des polynômes à coefficients dans cet anneau, etc...).

Pour cela, considérons l'application

$$j : \mathrm{K} \to \mathrm{F/R}$$

donnée par

$$j(a) = \theta(a, 1)$$

pour tout $a \in \mathrm{K}$. Elle est *injective*, car la relation $\theta(a', 1) = \theta(a'', 1)$ s'écrit $a' \cdot 1 = 1 \cdot a''$. C'est de plus un *homomorphisme d'anneaux*. On a en effet

$$\begin{aligned} j(a') + j(a'') &= \theta(a', 1) + \theta(a'', 1) = \theta(a' \cdot 1 + 1 \cdot a'', 1 \cdot 1) \\ &= \theta(a' + a'', 1) = j(a' + a''), \\ j(a') \cdot j(a'') &= \theta(a', 1)\theta(a'', 1) = \theta(a'a'', 1 \cdot 1) = \theta(a'a'', 1) = j(a'a''), \\ j(1) &= \theta(1, 1) = 1. \end{aligned}$$

L'application j est donc bien un isomorphisme de K sur un sous-anneau de F/R, à savoir le sous-anneau formé des éléments de la forme $\theta(a, b)$ avec $b = 1$. Dorénavant nous ne ferons aucune différence entre un élément a de K et l'élément $\theta(a, 1)$ de F/R ; autrement dit, nous écrirons

$$\theta(a, 1) = a, \tag{10}$$

pour tout $a \in \mathrm{K}$.

On a alors le résultat suivant : *tout élément de* F/R *est un quotient de deux éléments de* K ; de façon précise, on a

$$\theta(a, b) = ab^{-1}, \tag{11}$$

ou, si l'on préfère,

$$\theta(a, b) = \theta(a, 1) \cdot \theta(b, 1)^{-1} \tag{12}$$

on notera que $\theta(b, 1)$ n'est pas nul puisque $b \neq 0$, donc est inversible dans F/R d'après le n° précédent.

Pour prouver (12), il suffit de montrer que

$$\theta(a, b) \cdot \theta(b, 1) = \theta(a, 1),$$

ou que

$$\theta(ab, b) = \theta(a, 1);$$

mais cette relation s'écrit $ab1 = ba$, et est donc évidente.

Le corps F/R que nous avons construit dans ce qui précède s'appelle le **corps des fractions** de l'**anneau d'intégrité commutatif** K; le lecteur pourra par la suite oublier la façon dont nous l'avons obtenu; il suffit, pour toutes les applications, d'en connaître les propriétés suivantes : K *est un sous-anneau de son corps des fractions, et tout élément de celui-ci est quotient de deux éléments de* K. En particulier, nous n'utiliserons plus jamais la notation $0(a, b)$ pour désigner les éléments du corps des fractions de K; nous les désignerons par la notation a/b ou $\frac{a}{b}$ ou ab^{-1}. Mais bien entendu le lecteur devra se garder de commettre l'erreur grossière qui consisterait à croire que, pour que deux fractions a/b et c/d soient égales, il faut que $a = c$ et $b = d$; une fraction *n'est pas* un couple (a, b) d'éléments de K, avec $b \neq 0$; c'est une classe de tels couples.

En partant de $K = \mathbf{Z}$, anneau des entiers rationnels, il est clair que les considérations précédentes conduiraient au corps $\mathbf{Q}$ des nombres rationnels. D'ailleurs, si l'on connaît de nombreuses méthodes « simples » ou « géométriques » pour définir les nombres rationnels à partir des nombres entiers, on n'en connaît qu'une seule qui soit mathématiquement correcte — à savoir celle que nous venons d'exposer; l'hypothèse que $K = \mathbf{Z}$ ne permet pas de la simplifier si peu que ce soit; voir la *Remarque* 1 ci-dessus.

5. *Fractions rationnelles à coefficients dans un corps*

Soit K un corps commutatif; on a vu (§ 27, Théorème 1) que pour tout entier $n \geqslant 1$ l'anneau $K[X_1, \ldots, X_n]$ des polynômes à n indéterminées à coefficients dans K est intègre. On peut donc lui appliquer les considérations des n° précédents. Le corps des fractions de l'anneau $K[X_1, \ldots, X_n]$ se désigne par la notation

$$K(X_1, \ldots, X_n)$$

et s'appelle le **corps des fractions rationnelles à** n **variables** sur le corps K. Les éléments de $K(X_1, \ldots, X_n)$ s'appellent eux-mêmes des **fractions rationnelles à** n **variables à coefficients dans** K.

Pour manipuler pratiquement des fractions rationnelles, il est inutile, encore une fois, d'avoir recours aux considérations du n° 2 : celles-ci servent uniquement à prouver l'*existence* du corps $K(X_1, \ldots, X_n)$; mais ce qui compte dans la pratique, ce sont les *propriétés* de ce corps. Autrement dit, le lecteur devra retenir les assertions suivantes, *et rien d'autre* :

(FR 1) : *les fractions rationnelles à n variables à coefficients dans* K *sont les éléments d'un corps commutatif noté* $K(X_1, \ldots, X_n)$;

(FR 2) : *parmi les fractions rationnelles à n variables à coefficients dans* K *figurent les polynômes à n indéterminées à coefficients dans* K; plus précisément, l'anneau $K[X_1, \ldots, X_n]$ des polynômes à n indéterminées à coefficients dans K est un sous-anneau du corps $K(X_1, \ldots, X_n)$;

(FR 3) : *pour toute fraction rationnelle* $f \in K(X_1, \ldots, X_n)$, *il existe deux polynômes*

$p, q \in \mathrm{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$, *avec* $q \neq 0$, *tels que l'on ait*

$$f = pq^{-1} = p/q.$$

Bien entendu, il existe plusieurs façons d'écrire f comme quotient de deux polynômes, et pour que deux fractions rationnelles p'/q' et p''/q'' soient égales, il faut et il suffit que $p'q'' = p''q'$.

Exemple 1. Une fraction rationnelle à une variable à coefficients dans un corps K est une expression de la forme

$$f = \frac{p(\mathrm{X})}{q(\mathrm{X})},$$

où p et q sont des polynômes à une indéterminée X, à coefficients dans K, avec $q \neq 0$. Exemple :

$$\frac{\mathrm{X}^3 + \mathrm{X}^2 - 1}{\mathrm{X}^2 - 1}.$$

On notera que la condition $q \neq 0$ *n'exclut pas* la possibilité que l'on ait $q(x) = 0$ pour certains $x \in \mathrm{K}$; elle signifie simplement que q n'est pas l'élément 0 de l'anneau K[X], i.e. que ses *coefficients* ne sont pas tous nuls.

Exemple 2. L'expression

$$f = \frac{\mathrm{X} + \mathrm{Y}}{\mathrm{X} - \mathrm{Y}}$$

est une fraction rationnelle à deux variables X et Y, à coefficients dans **Q**.

6. *Valeurs d'une fraction rationnelle*

Soient L un corps commutatif, K un sous-corps de L, et f une fraction rationnelle à n variables à coefficients dans K. Soit

$$u = (u_1, \ldots, u_n)$$

un élément de L^n. On dit que f **est définie en** u, ou que $u_1, \ldots, u_n$ **sont substituables dans** f, **s'il** existe des polynômes p et q vérifiant

$$f = p/q, \qquad q(u_1, \ldots, u_n) \neq 0;$$

cela ne veut pas dire que, *quels que soient* les polynômes p et q (avec $q \neq 0$) tels que $f = p/q$, on aura nécessairement $q(u_1, \ldots, u_n) \neq 0$.

Exemple 3. $u = 0$ est substituable dans la fraction rationnelle

$$\frac{\mathrm{X}^9}{\mathrm{X}^8 + \mathrm{X}^7}$$

car celle-ci s'écrit aussi

$$\frac{X}{X+1}$$

et sous cette forme on voit que le dénominateur ne s'annulle pas pour $u = 0$.

Supposons f définie en $u = (u_1, \ldots, u_n)$; on peut alors définir la **valeur de f en u**, de la façon suivante. Écrivons

$$f = p/q \qquad \text{avec} \qquad q(u) \neq 0;$$

l'élément $p(u)/q(u)$ de L ne dépend alors que de f, et non du choix de p et q; en effet, si l'on a une autre représentation de f, soit

$$f = p'/q' \qquad \text{avec} \qquad q'(u) \neq 0,$$

il vient

$$pq' = p'q,$$

donc (§ 28, formule (6))

$$p(u)q'(u) = p'(u)q(u),$$

et par suite

$$p(u)/q(u) = p'(u)/q'(u),$$

ce qui établit notre assertion. Cela dit, c'est l'élément $p(u)/q(u)$ de L qu'on appelle valeur de f en u; et on le désigne par la notation

$$f(u) \qquad \text{ou} \qquad f(u_1, \ldots, u_n);$$

on a donc

$$f(u) = p(u)/q(u) \qquad \text{si} \qquad f = p/q \qquad \text{avec} \qquad q(u) \neq 0.$$

Il est clair que, si f est un polynôme, alors f est définie en u quel que soit u, et que la valeur $f(u)$ qui résulte de la définition précédente coïncide avec celle qu'on a définie au § 28, n° 1 : il suffit pour le voir d'écrire $f = f/1$.

Montrons que *l'ensemble des fractions rationnelles qui sont définies en un point donné $u \in \mathrm{L}^n$ est un sous-anneau de* $\mathrm{K}(X_1, \ldots, X_n)$. Supposons en effet f' et f'' définies en u; on peut alors écrire

$$f' = p'/q' \qquad \text{avec} \qquad q'(u) \neq 0,$$
$$f'' = p''/q'' \qquad \text{avec} \qquad q''(u) \neq 0,$$

d'où

$$f' + f'' = \frac{p'q'' + p''q'}{q'q''}, \qquad f'f'' = \frac{p'p''}{q'q''},$$

et comme on a $q'(u)q''(u) \neq 0$ on voit que $f' + f''$ et $f'f''$ sont définies en u; comme les polynômes (et en particulier -1) sont définis en u, il s'ensuit bien que les fractions rationnelles définies en u forment un sous-anneau de $\mathrm{K}(X_1, \ldots, X_n)$.

Supposons f' et f'' définies en $u \in \mathrm{L}^n$; alors les valeurs en u des fractions $f' + f''$ et $f'f''$ sont égales respectivement à $f'(u) + f''(u)$ et $f'(u)f''(u)$. Conservons en effet les

notations utilisées ci-dessus, et soit $g = f' + f''$; on a donc

$$g = p/q \qquad \text{avec} \qquad p = p'q'' + p''q', \qquad q = q'q'';$$

donc

$$g(u) = p(u)/q(u) = \frac{p'(u)q''(u) + p''(u)q'(u)}{q'(u)q''(u)} = \frac{p'(u)}{q'(u)} + \frac{p''(u)}{q''(u)},$$

ce qui prouve le premier résultat annoncé; le second se démontre de même.

Supposons f définie en u; pour que f^{-1} soit définie en u, il faut et il suffit que $f(u) \neq 0$; on a alors

$$f^{-1}(u) = f(u)^{-1}.$$

En effet supposons f et f^{-1} définies en u; on a $1 = f \cdot f^{-1}$, d'où, en prenant les valeurs en u,

$$1 = f(u) \cdot f^{-1}(u),$$

ce qui prouve que $f(u) \neq 0$ et que $f^{-1}(u) = f(u)^{-1}$. Inversement supposons f définie en u et $f(u) \neq 0$; on a $f = p/q$ avec $q(u) \neq 0$, et aussi $p(u) \neq 0$ puisque $f(u) \neq 0$; écrivant $f^{-1} = q/p$ on voit donc que la fraction rationnelle f^{-1} est, elle aussi, définie en u.

Exemple 4. Prenons $n = 2$ et la fraction rationnelle

$$f = \frac{X + Y}{X - Y};$$

elle est définie en tout point $(u, v) \in L^2$ tel que $u \neq v$ (i.e. en dehors de la diagonale); il n'existe aucun autre point de L^2 où f soit définie. En effet, soient p et q des polynômes à coefficients dans K tels que $f = p/q$; on a donc

$$(X + Y)q(X, Y) = (X - Y)p(X, Y);$$

posant $p = \sum p_n$, $q = \sum q_n$ où p_n et q_n sont homogènes de degré total n, on déduit immédiatement de la relation précédente que l'on a

$$(X + Y)q_n(X, Y) = (X - Y)p_n(X, Y);$$

on déduit de là (le lecteur le démontrera à titre d'exercice, en mettant en évidence les coefficients de p_n et q_n, et en établissant des relations entre ces coefficients) que pour tout n il existe un polynôme $r_n(X, Y)$ tel que

$$q_n(X, Y) = (X - Y)r_n(X, Y);$$

il s'ensuit évidemment que $q(u, v) = 0$ pour $u = v$, de sorte que f n'est définie en aucun point de la diagonale de L^2, comme annoncé.

Exemple 5. K étant un corps commutatif arbitraire, prenons

$$L = K(X_1, \ldots, X_n),$$

et pour u le point

$$u = (X_1, \ldots, X_n) \in L^n;$$

alors *toute* fraction rationnelle $f \in K(X_1, \ldots, X_n)$ est définie en u, car en écrivant $f = p/q$ avec $q \neq 0$, on a $q(X_1, \ldots, X_n) \neq 0$ pour la raison que

$$q(X_1, \ldots, X_n) = q$$

comme on l'a vu au § 28, *Remarque* 1. On peut donc définir $f(X_1, \ldots, X_n)$; cet élément de L est donné par

$$f(X_1, \ldots, X_n) = p(X_1, \ldots, X_n)/q(X_1, \ldots, X_n) = p/q = f.$$

Cela explique pourquoi, dans la pratique, on désigne souvent une fraction rationnelle par la notation $f(X_1, \ldots, X_n)$; mais les considérations développées ici permettent de *démontrer* que $f(X_1, \ldots, X_n) = f$, alors que dans les manuels classiques cette relation n'est considérée que comme une simple convention d'écriture.

Cet *Exemple* (et beaucoup d'autres) explique aussi pourquoi il est nécessaire de définir la valeur d'une fraction rationnelle à coefficients dans un corps K en un point dont les coordonnées appartiennent non à K, mais à un sur-corps arbitraire de K. Dans la pratique la plus élémentaire, il est évidemment indispensable de savoir attribuer une valeur à une fraction rationnelle à coefficients *réels* en un point à coordonnées *complexes*.

Remarque 2. Lorsqu'une fraction rationnelle $f \in K(X_1, \ldots, X_n)$ n'est pas définie en un point $u \in L^n$, où L est un surcorps commutatif de K, deux cas sont possibles : il peut arriver que l'inverse $1/f$ de f soit définie en u (on dit alors que u est un **pôle de** f), et il peut arriver que $1/f$ ne soit pas non plus définie en u (on dit alors que u est un **point d'indétermination de** f). Le premier cas se produit si et seulement si l'on peut écrire

$$f = p/q \qquad \text{avec} \qquad p(u) \neq 0, \qquad q(u) = 0;$$

s'il en est ainsi, il est en effet clair que $f^{-1} = q/p$ est définie en u, et y a pour valeur 0, en sorte que f ne peut être définie en u; inversement, si u est un pôle de f, on peut écrire, puisque f^{-1} est définie en u,

$$f^{-1} = q/p \qquad \text{avec} \qquad p(u) \neq 0;$$

si l'on avait $f^{-1}(u) \neq 0$, f serait aussi définie en u comme on l'a vu plus haut, ce qui par hypothèse n'est pas le cas; on a donc $f^{-1}(u) = 0$, i.e. $q(u) = 0$, et par suite

$$f = p/q \qquad \text{avec} \qquad p(u) \neq 0, \qquad q(u) = 0$$

comme annoncé.

Les considérations qui précèdent permettent d'autre part de caractériser les points d'indétermination de f : u est un point d'indétermination de f **si**, *quels que soient* les polynômes p et q tels que

$$f = p/q, \quad q \neq 0,$$

on a

$$p(u) = q(u) = 0.$$

Prenons par exemple $K = L = C$ et la fraction rationnelle à deux indéterminées

$$f = \frac{X + Y}{X - Y};$$

elle est évidemment définie en tout point $(u, v) \in C^2$ tel que $u \neq v$; les points de la diagonale $u = v$ autres que $(0, 0)$ sont évidemment des *pôles* de f; enfin, le point $(0, 0)$ est un *point d'indétermination* de f. Pour établir ce dernier résultat, il faut prouver que, si deux polynômes p, $q \in C[X, Y]$ vérifient

$$\frac{p}{q} = \frac{X + Y}{X - Y},$$

alors on a nécessairement

$$p(0, 0) = q(0, 0) = 0.$$

Or posons

$$p = a + a'X + a''Y + \cdots$$
$$q = b + b'X + b''Y + \cdots,$$

les termes non écrits étant tous de degré deux au moins; on a par hypothèse

$$(X - Y)p = (X + Y)q$$

i.e.

$$(X - Y)(a + a'X + a''Y + \cdots) = (X + Y)(b + b'X + b''Y + \cdots)$$

et par suite

$$a(X - Y) = b(X + Y), \qquad \text{i.e. } a = b = -b,$$

d'où résulte évidemment $a = b = 0$; comme $p(0, 0) = a$ et $q(0, 0) = b$ notre assertion est établie.

L'étude détaillée des fractions rationnelles, et d'objets similaires (les « fonctions algébriques de plusieurs variables ») est l'un des principaux buts de la Géométrie Algébrique.

§ 30. Dérivations, formule de Taylor

Dans la théorie des fonctions d'une variable réelle, la dérivée $f'(t)$ d'une fonction $f(t)$ est définie à l'aide d'un « passage à la limite », i.e. à l'aide d'un processus aussi peu algébrique que possible. Cependant, lorsque f est une fonction polynomiale, soit

$$f(t) = \sum a_r t^r,$$

on sait qu'il en est de même de la dérivée de f, et que celle-ci est

$$f'(t) = \sum r a_r t^{r-1}.$$

Il est clair que, si l'on se bornait à étudier des fonctions polynomiales, on pourrait *définir* la dérivée d'une telle fonction à l'aide des formules précédentes, dans lesquelles ne figure aucune opération de passage à la limite.

Ces remarques suggèrent la possibilité d'étendre la notion de dérivée aux polynômes à coefficients dans un anneau commutatif quelconque, et à démontrer dans ce cas, par des procédés purement algébriques, des propriétés qui, dans le cas classique, s'obtiennent par des raisonnements « analytiques » applicables à des fonctions beaucoup plus générales que les fonctions polynomiales. C'est ce qu'on va faire dans ce §. Les résultats ainsi obtenus ne sont pas de simples généralisations élégantes mais dénuées d'intérêt pratique de la théorie classique; on les utilisera effectivement plus loin pour distinguer les racines simples des racines multiples d'une équation algébrique, et on s'en sert aujourd'hui à propos d'autres problèmes importants (par exemple pour distinguer les « points simples » des « points multiples » d'une variété algébrique).

1. Dérivations dans un anneau

Soit K un anneau; on appelle **dérivation de l'anneau** K toute application

$$D : K \to K$$

qui vérifie les relations

$$D(x + y) = D(x) + D(y) \tag{1}$$

$$D(xy) = D(x)\,y + xD(y) \tag{2}$$

quels que soient $x, y \in K$. Cette définition est évidemment inspirée par les règles de calcul classiques des dérivées; du reste :

Exemple 1. Prenons pour K l'anneau des fonctions polynomiales sur **R**; si

$$x(t) = \sum a_r t^r$$

est une telle fonction, définissons $D(x)$ comme étant la fonction

$$x'(t) = \sum r a_r t^{r-1};$$

les règles classiques de dérivation d'une somme et d'un produit montrent alors que l'application D ainsi définie est une dérivation de l'anneau K.

Les relations (1) et (2) montrent qu'on a

$$D(1) = 0 \tag{3}$$

car en faisant $y = 1$ dans (2) il vient $x.D(1) = 0$ pour tout x, en particulier pour $x = 1$. Plus généralement, on a, *si* K *est commutatif*,

$$D(x^n) = nx^{n-1}D(x) \tag{4}$$

pour tout $x \in K$ et tout entier $n \geqslant 0$; pour $n = 0$ ce résultat se réduit en effet à (3); et s'il est établi pour $n - 1$, alors (2) montre que

$$D(x^n) = D(x^{n-1}.x) = D(x^{n-1})x + x^{n-1}D(x) = (n-1)x^{n-2}D(x)x + x^{n-1}D(x) = nx^{n-1}D(x)$$

si K est commutatif.

2. *Dérivations d'un anneau de polynômes*

Nous allons démontrer le résultat suivant :

THÉORÈME 1. *Soit* D *une dérivation d'un anneau commutatif* K. *Étant donné un polynôme* $u \in K[X]$, *il existe dans l'anneau* K[X] *une et une seule dérivation qui coïncide avec* D *sur* K *et qui applique le polynôme* X *sur le polynôme donné* u.

Supposons trouvée une telle dérivation D'; étant donné un polynôme

$$f = \sum a_r X^r \quad (a_r \in K),$$

les règles de calcul (1), (2), (4) donnent

$$D'(f) = \sum D'(a_r X^r) = \sum [D'(a_r)X^r + r a_r X^{r-1} D'(X)]$$

et vu les conditions imposées à D' il reste donc

$$D'(f) = \sum D(a_r)X^r + u(X)\sum r a_r X^{r-1}. \tag{5}$$

Pour exprimer le résultat, il est commode d'introduire d'une part le polynôme

$$f^D(X) = \sum D(a_r)X^r \tag{6}$$

obtenu en appliquant D aux coefficients de f, d'autre part le **polynôme dérivé**

$$f'(X) = \sum r a_r X^{r-1} \tag{7}$$

de f; ceci fait, la relation (5) s'écrit

$$D'(f) = f^D + u.f' \tag{8}$$

et prouve l'unicité de D'.

Il reste à montrer que l'application

$$D' : K[X] \to K[X]$$

donnée par la formule (8) est effectivement une dérivation satisfaisant aux conditions de l'énoncé. Les formules évidentes

$$(f+g)^D = f^D + g^D, \qquad (f+g)' = f' + g'$$

montrent déjà que D' vérifie (1). Pour établir (2), remarquons d'abord que si D_1 et D_2 sont des dérivations d'un anneau L, alors quels que soient $u_1, u_2 \in L$ l'application

$$x \mapsto u_1.D_1(x) + u_2.D_2(x)$$

de L dans L est encore une dérivation de L. Pour montrer que D' satisfait à (2), il suffit donc d'après (8) de prouver que, dans K[X], les deux applications

$$f \mapsto f^D, \qquad f \mapsto f'$$

sont des dérivations, autrement dit qu'on a les formules

$$(fg)^D = f^D.g + f.g^D, \qquad (fg)' = f'g + fg'.$$

Supposons d'abord f et g réduits à des monômes,

$$f = aX^p, \qquad g = bX^q;$$

on a alors

$$(fg)^D = (abX^{p+q})^D = D(ab)X^{p+q} = D(a)X^p.bX^q + aX^p.D(b)X^q$$
$$(fg)' = (abX^{p+q})' = (p+q)abX^{p+q-1} = paX^{p-1}.bX^q + aX^p.qbX^{q-1},$$

ce qui établit évidemment les formules cherchées dans ce cas particulier. Dans le cas général, on décompose f et g en sommes de monômes, ce qui permet de se ramener facilement au cas particulier qu'on vient d'étudier.

Nous avons déjà établi que l'application (8) est bien une dérivation; il reste à vérifier d'une part qu'elle coïncide avec D sur K, autrement dit que

$$f = a \in K \quad \text{implique} \quad D'(f) = D(a),$$

ce qui est clair d'après (5); et d'autre part, que

$$D'(X) = u,$$

ce qui est aussi évident si l'on écrit $X = 1.X$ et si l'on remarque que $D(1) = 0$. Le Théorème 1 est donc entièrement démontré.

3. *Dérivées partielles*

Le Théorème 1 s'étend comme suit aux polynômes à plusieurs indéterminées :

THÉORÈME 2. *Soient* K *un anneau commutatif,* D *une dérivation de* K, *et* $u_1, \ldots, u_n$ *des polynômes à* n *indéterminées à coefficients dans* K. *Il existe alors une et une seule dérivation* D' *de l'anneau* $K[X_1, \ldots, X_n]$ *qui se réduit à* D *sur* K *et qui vérifie*

$$D'(X_i) = u_i \quad \textit{pour} \quad 1 \leqslant i \leqslant n.$$

La démonstration est évidemment analogue à celle du Théorème 1, et nous nous bornerons à indiquer comment on calcule $D'(f)$ pour un polynôme

$$f = \sum a_{r_1 \ldots r_n} X_1^{r_1} \ldots X_n^{r_n}.$$

Comme D' est une dérivation, on a

$$D'(f) = \sum D'(a_{r_1 \ldots r_n} X_1^{r_1} \ldots X_n^{r_n}) = \sum D'(a_{r_1 \ldots r_n}) X_1^{r_1} \ldots X_n^{r_n} + \sum a_{r_1 \ldots r_n} D'(X_1^{r_1}) X_2^{r_2} \ldots X_n^{r_n} + \cdots + \sum a_{r_1 \ldots r_n} X_1^{r_1} \ldots X_{n-1}^{r_{n-1}} D'(X_n^{r_n}),$$

et comme $D'(a) = D(a)$ pour $a \in K$, $D'(X_i) = u_i$, il vient

$$(9) \quad D'(f) = \sum D(a_{r_1 \ldots r_n}) X_1^{r_1} \ldots X_n^{r_n} + u_1 \sum r_1 a_{r_1 \ldots r_n} X_1^{r_1 - 1} X_2^{r_2} \ldots X_n^{r_n} + \cdots + u_n \sum r_n a_{r_1 \ldots r_n} X_1^{r_1} \ldots X_{n-1}^{r_{n-1}} X_n^{r_n - 1}.$$

On est ainsi conduit à introduire, comme au n° précédent, les notations suivantes. Tout d'abord on posera

$$f^D = \sum D(a_{r_1 \ldots r_n}) X_1^{r_1} \ldots X_n^{r_n};$$

c'est le polynôme obtenu en appliquant D aux coefficients de f. D'autre part, on appellera **dérivée partielle de** f **par rapport à** X_i le polynôme

$$(10) \quad f_i' = \sum r_i a_{r_1 \ldots r_n} X_1^{r_1} \ldots X_{i-1}^{r_{i-1}} X_i^{r_i - 1} X_{i+1}^{r_{i+1}} \ldots X_n^{r_n};$$

si l'on regarde f comme un polynôme en X_i, à coefficients dans l'anneau

$$K[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n],$$

il est clair que f_i' n'est autre que le polynôme dérivé de f au sens du n° précédent (ce qui correspond bien à la notion classique de dérivée partielle d'une fonction de plusieurs variables : on dérive par rapport à une variable, en regardant les autres comme des « constantes »). Ceci dit, la relation (9) s'écrit encore

$$(11) \quad D'(f) = f^D + \sum_{i=1}^{i=n} u_i f_i'.$$

Dans la pratique, au lieu de la notation f'_i, on utilise souvent les notations

$$f'_{X_i} \qquad \frac{\partial f}{\partial X_i}$$

qui sont d'usage courant en Analyse. Si l'on pose

$$D_i(f) = f'_i,$$

il est clair que l'application D_i de l'anneau $K[X_1, \ldots, X_n]$ dans lui-même vérifie les conditions suivantes : c'est une *dérivation*, et on a

$$D_i(a) = 0 \quad \text{si} \quad a \in K$$
$$D_i(X_j) = \begin{cases} 0 & \text{si} \quad i \neq j \\ 1 & \text{si} \quad i = j. \end{cases}$$

Et, d'après le Théorème 2, ou d'après le Théorème 1 appliqué à l'anneau

$$K[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n]$$

sur lequel D_i est nulle, ces propriétés *caractérisent* l'application D_i.

4. *Dérivation des fonctions composées*

Le classique théorème de « dérivation des fonctions composées » est, lorsqu'il s'agit de polynômes, une conséquence du résultat suivant :

THÉORÈME 3. *Soient* L *un anneau commutatif,* K *un sous-anneau de* L, D *une dérivation de* L *nulle sur* K, *et* f *un polynôme à* n *indéterminées à coefficients dans* K. *On a alors*

$$D[f(u_1, \ldots, u_n)] = \sum_{i=1}^{i=n} f'_i(u_1, \ldots, u_n) \cdot D(u_i)$$

quels que soient $u_1, \ldots, u_n \in L$.

Il suffit évidemment d'établir ce résultat lorsque f est un monôme, soit

$$f = aX_1^{r_1} \ldots X_n^{r_n};$$

on a alors

$$\begin{aligned} D[f(u_1, \ldots, u_n)] &= D(a.u_1^{r_1} \ldots u_n^{r_n}) \\ &= D(a)u_1^{r_1} \ldots u_n^{r_n} + \sum_{i=1}^{i=n} a.u_1^{r_1} \ldots u_i^{r_i - 1} D(u_i^{r_i}) u_{i+1}^{r_{i+1}} \ldots u_n^{r_n} \\ &= D(a)u_1^{r_1} \ldots u_n^{r_n} + \sum_{i=1}^{i=n} D(u_i) \cdot r_i a u_1^{r_1} \ldots u_i^{r_i - 1} u_{i+1}^{r_{i+1}} \ldots u_n^{r_n}; \end{aligned}$$

si $D = 0$ sur K, le premier terme disparaît, et il reste visiblement la formule cherchée.

COROLLAIRE. *Soient* K *un anneau commutatif, f un polynôme à n indéterminées à coefficients dans* K, *et*

$$u_1, \ldots, u_n \in \mathrm{K}[\mathrm{Y}_1, \ldots, \mathrm{Y}_p]$$

des polynômes à p indéterminées à coefficients dans K. *Les dérivées partielles du polynôme*

$$g(\mathrm{Y}_1, \ldots, \mathrm{Y}_p) = f[u_1(\mathrm{Y}_1, \ldots, \mathrm{Y}_p), \ldots, u_n(\mathrm{Y}_1, \ldots, \mathrm{Y}_p)]$$

sont données par les relations

$$\frac{\partial g}{\partial \mathrm{Y}_j} = \sum_{i=1}^{i=n} f'_i(u_1, \ldots, u_p) \cdot \frac{\partial u_i}{\partial \mathrm{Y}_j} \quad (1 \leqslant j \leqslant p).$$

Il suffit pour le voir d'appliquer le Théorème 3 en prenant $\mathrm{L} = \mathrm{K}[\mathrm{Y}_1, \ldots, \mathrm{Y}_p]$ et pour D la dérivation partielle par rapport à Y_j.

Le Corollaire précédent est à proprement parler le *théorème des fonctions composées pour les polynômes*.

On ne saurait en attendre des conséquences très profondes, étant donné qu'il résulte à peu près trivialement des *définitions* posées dans ce §.

5. *Formule de Taylor*

Soient K un anneau commutatif, f un polynôme à *une* indéterminée et à coefficients dans K, et considérons le polynôme $\mathrm{X} + \mathrm{Y}$ à *deux* indéterminées X et Y et à coefficients dans K. En le substituant à la variable qui figure dans f, on obtient un polynôme

$$f(\mathrm{X} + \mathrm{Y}) \in \mathrm{K}[\mathrm{X}, \mathrm{Y}] = \mathrm{K}[\mathrm{X}][\mathrm{Y}],$$

qu'on peut donc écrire comme polynôme en Y à coefficients dans l'anneau K[X], i.e. sous la forme

$$f(\mathrm{X} + \mathrm{Y}) = f_0(\mathrm{X}) + f_1(\mathrm{X})\mathrm{Y} + \cdots + f_n(\mathrm{X})\mathrm{Y}^n \tag{12}$$

si f est de degré n. On se propose de calculer les polynômes $f_p(\mathrm{X})$ à l'aide des **dérivées successives** de f, à savoir les polynômes

$$f' = (f')', \qquad f''' = (f'')', \ldots$$

Pour cela, dérivons par rapport à Y les deux membres de la relation (12); d'après le Corollaire du Théorème 3, le premier membre a pour dérivée $f'(\mathrm{X} + \mathrm{Y})$; il vient donc

$$f'(\mathrm{X} + \mathrm{Y}) = f_1(\mathrm{X}) + 2f_2(\mathrm{X})\mathrm{Y} + \cdots + nf_n(\mathrm{X})\mathrm{Y}^{n-1};$$

en dérivant à nouveau ce résultat par rapport à Y, on trouve

$$f''(\mathrm{X} + \mathrm{Y}) = 2f_2(\mathrm{X}) + 3 \cdot 2f_3(\mathrm{X})\mathrm{Y} + \cdots + n(n-1)f_n(\mathrm{X})\mathrm{Y}^{n-2},$$

et en poursuivant ainsi de suite on obtient évidemment

$$f^{(k)}(\mathrm{X} + \mathrm{Y}) = k!f_k(\mathrm{X}) + (k+1)k \ldots 2f_{k+1}(\mathrm{X})\mathrm{Y} + \ldots + n(n-1)\ldots(n-k+1)f_n(\mathrm{X})\mathrm{Y}^{n-k},$$

Il est clair que la relation précédente, étant une égalité entre polynômes en X et Y à coefficients dans K, reste vraie si on y remplace X et Y par u et v, où u et v sont des éléments arbitraires d'un sur-anneau commutatif L de K. En particulier, on peut remplacer X et Y par X et 0; il reste alors la relation

$$f^{(k)}(X) = k!\,f_k(X), \qquad 0 \leqslant k \leqslant n.$$

Donc :

THÉORÈME 4. *Soit f un polynôme de degré n à une indéterminée à coefficients dans un anneau commutatif* K, *et soient* X *et* Y *deux indéterminées sur* K. *On a alors*

$$f(X+Y) = f(X) + f'(X)Y + f_2(X)Y^2 + \cdots + f_n(X)Y^n \tag{12}$$

avec

$$k!\,f_k(X) = f^{(k)}(X) \quad \textit{pour} \quad 2 \leqslant k \leqslant n. \tag{13}$$

Ce résultat est connu sous le nom de **formule de Taylor** pour la raison suivante. Tout d'abord, comme c'est une égalité entre polynômes, on peut y remplacer les indéterminées X et Y par des éléments quelconques d'un sur-anneau commutatif arbitraire de K, en particulier par des éléments x et h de K; on a donc

$$f(x+h) = f(x) + f'(x)h + f_2(x)h^2 + \cdots + f_n(x)h^n \quad \text{avec} \quad k!\,f_k(x) = f^{(k)}(x),$$

quels que soient x, $h \in$ K. Supposons alors K = **R**, corps des nombres réels; il vient évidemment

$$f_k(x) = \frac{f^{(k)}(x)}{k!},$$

et le résultat obtenu s'écrit donc

$$f(x+h) = f(x) + h\frac{f'(x)}{1!} + h^2\frac{f''(x)}{2!} + \cdots + h^n\frac{f^{(n)}(x)}{n!},$$

ce qui est justement la formule de Taylor classique (celle-ci, dans le cas des fonctions *polynomiales*, est donc un résultat de nature purement algébrique).

EXEMPLE 2. Prenons K = **Z** et

$$f(X) = X^n;$$

les formules (12) et (13) s'écrivent

$$(X+Y)^n = X^n + nX^{n-1}Y + f_2(X)Y^2 + \cdots + f_n(X)Y^n$$

avec

$$k!\,f_k(X) = n(n-1)\ldots(n-k+1)X^{n-k},$$

ou encore

$$k!\left[f_k(X) - \binom{n}{k}X^{n-k}\right] = 0;$$

puisque l'anneau **Z** est intègre, il s'ensuit que

$$f_k(X) = \binom{n}{k} X^{n-k};$$

ainsi, on obtient dans ce cas la relation

$$(X + Y)^n = \sum_{k=0}^{k=n} \binom{n}{k} X^{n-k} Y^k.$$

Pour en déduire la formule du binôme établie au § 8, n° 4, il suffit de remarquer que si deux polynômes $f(X, Y)$ et $g(X, Y)$ à coefficients *entiers rationnels* sont identiques, alors la relation

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

est vraie lorsque x et y sont des éléments *arbitraires* d'un anneau commutatif *arbitraire* L.

6. *Caractéristique d'un corps commutatif*

Revenons au Théorème 4. Pour que la formule (12) soit réellement intéressante, il est nécessaire qu'on puisse calculer complètement les polynômes f_k, et pour cela il s'impose d'essayer de les déduire de la formule (13). Autrement dit, nous avons à résoudre le problème suivant : étant donnés un entier rationnel $r \neq 0$ (en l'occurence, $k!$) et un élément b de K [en l'occurence, l'un quelconque des coefficients du polynôme $f^{(k)}(X)$], trouver tous les $x \in K$ tels que

$$r.x = b.$$

Étant donné que

$$rx = (r.1)x$$

où 1 est l'élément unité de K, le problème aura une et une seule solution dès que $r.1$ est *inversible*. Si K est un *corps commutatif*, cela signifie que $r.1 \neq 0$.

Dans l'hypothèse où K est un corps commutatif, on est donc amené à considérer, dans l'anneau **Z** des entiers rationnels, l'ensemble I des entiers r tels que

$$r.1 = 0;$$

I contient 0, et s'il contient deux entiers r et s il contient visiblement $r - s$: c'est donc un sous-groupe du groupe additif **Z**, et par suite (§ 7, *Exemple* 8) *il existe un entier* $p \geqslant 0$ *et un seul tel que* $I = p\mathbf{Z}$; on dit que p est la **caractéristique** du corps K.

La caractéristique p d'un corps commutatif K peut évidemment se définir comme suit : si la relation $r.1 = 0$ implique $r = 0$, alors $p = 0$; sinon, p est le plus petit entier strictement positif tel que $p.1 = 0$. Dans tous les cas, la relation $r.1 = 0$ signifie que r est multiple de p, ou encore : *pour que* $r.1 \neq 0$ *il faut et il suffit que* r *ne soit pas divisible par la caractéristique de* K.

Il est important de noter que *la caractéristique d'un corps commutatif est toujours*

soit 0, soit un nombre premier. Supposons en effet le corps K de caractéristique $p \neq 0$, et considérons deux entiers r et s tels que

$$p = rs;$$

on a alors

$$0 = p.1 = (r.1)\,(s.1);$$

comme un corps est un anneau d'intégrité, il en résulte soit $r.1 = 0$ (auquel cas p divise r) soit $s.1 = 0$ (auquel cas p divise s), ce qui établit notre assertion.

Si K est un *corps de caractéristique* 0, alors pour tout $b \in K$ et tout entier $r \neq 0$, l'équation

$$rx = b$$

possède dans K une et une seule solution, à savoir

$$x = (r.1)^{-1}b,$$

que l'on écrit plus simplement sous la forme

$$x = \frac{b}{r} \qquad \text{ou} \qquad b/r.$$

Exemple 3. Les corps **Q**, **R**, **C** et plus généralement tout sur-corps commutatif de **Q**, sont évidemment de caractéristique 0.

Exemple 4. Prenons $K = \mathbf{Z}/p\mathbf{Z}$ où p est un nombre premier (§ 8, Théorème 1); pour tout entier rationnel r, la relation $r.1 = 0$ dans $\mathbf{Z}/p\mathbf{Z}$ signifie évidemment que r est nul comme entier modulo p, autrement dit que r est multiple de p; donc ici on a $I = p\mathbf{Z}$, autrement dit *le corps* $\mathbf{Z}/p\mathbf{Z}$, *pour* p *premier, est de caractéristique* p.

Remarque 1. Soit K un corps de caractéristique 0; l'application

$$n \longmapsto n.1$$

de **Z** dans K est donc injective, et par suite est un isomorphisme de **Z** sur un sous-anneau de K. En fait, cette application peut même se prolonger en un isomorphisme du corps **Q** sur un sous-corps de K; soit en effet $x \in \mathbf{Q}$ et écrivons $x = a/b$ avec $a, b \in \mathbf{Z}$, $b \neq 0$; alors l'élément

$$j(x) = (a.1)\,(b.1)^{-1}$$

de K ne dépend que de x, et non de la représentation de x sous forme de fraction; en effet, la relation $a'/b' = a''/b''$ s'écrit $a'b'' = a''b'$, donc implique

$$(a'.1)\,(b''.1) = (a''.1)\,(b'.1),$$

donc

$$(a'.1)\,(b'.1)^{-1} = (a''.1)\,(b''.1)^{-1},$$

ce qui prouve notre assertion. Ceci dit, il est immédiat de vérifier que l'application j de $\mathbf{Q}$ dans K ainsi obtenue est un isomorphisme de $\mathbf{Q}$ sur un sous-corps de K.

Dans la pratique, on identifie le plus souvent chaque nombre rationnel $x \in \mathbf{Q}$ à son image $j(x)$ dans K, de sorte que $\mathbf{Q}$ *est un sous-corps de tout corps de caractéristique* 0.

On peut également montrer (cf. *Exercice* 8) que tout corps K de caractéristique $p \neq 0$ contient un sous-corps isomorphe à $\mathbf{Z}/p\mathbf{Z}$, à savoir l'ensemble des multiples entiers $r.1$ ($r \in \mathbf{Z}$) de l'élément unité de K.

Si K est un *corps de caractéristique* 0, il est clair que la formule de Taylor du n° précédent prend la forme

$$f(X+Y) = \sum_{k=0}^{k=n} f^{(k)}(X) \cdot \frac{Y^k}{k!}. \tag{14}$$

Ce résultat est encore valable en caractéristique $p \neq 0$ pourvu que l'on ait $k!.1 \neq 0$ pour $k \leqslant n$, autrement dit pourvu que p ne divise aucun entier inférieur à n, autrement dit pourvu que

$$n < p.$$

7. *Ordre de multiplicité des racines d'une équation*

Soient K un anneau commutatif, f un polynôme à une indéterminée à coefficients dans K, et dans la formule de Taylor

$$f(X+Y) = f(X) + f'(X)Y + f_2(X)Y^2 + \cdots$$

remplaçons X et Y par a et $T - a$ respectivement, où a est un élément donné de K et où T est une indéterminée sur K; il vient

$$\begin{aligned} f(T) &= f(a) + f'(a).(T-a) + f_2(a).(T-a)^2 + \cdots \\ &= f(a) + (T-a)q(T) \end{aligned}$$

où $q \in K[T]$. Étant donnés des polynômes g, h à coefficients dans K, disons que h **est divisible par** g s'il existe un troisième polynôme q, à coefficients dans K, tel que $h = gq$. Le résultat que nous venons d'obtenir conduit immédiatement à l'énoncé que voici (déjà établi au § 28, Lemme 1) :

THÉORÈME 5. *Soit f un polynôme à une indéterminée à coefficients dans un anneau commutatif K; pour qu'un élément a de K soit racine de f, il faut et il suffit que le polynôme $f(X)$ soit divisible par le polynôme $X - a$.*

En effet, si a est racine de f, on a $f(a) = 0$, et la formule de Taylor montre donc que

$$f(T) = (T-a)q(T).$$

Inversement, de cette relation résulte

$$f(a) = (a-a)q(a) = 0,$$

d'où le Théorème.

Étant donnée une racine $a \in K$ du polynôme f, on appelle **ordre de multiplicité de a** le plus grand entier r tel que le polynôme $f(T)$ soit divisible par le polynôme

$$(T - a)^r.$$

Si $r = 1$, on dit que a est une **racine simple**; si $r = 2$, que a est une **racine double**, etc.

THÉORÈME 6. *Soit f un polynôme à une indéterminée à coefficients dans un anneau commutatif quelconque* K; *pour qu'un élément $a \in K$ soit racine simple de f, il faut et il suffit que l'on ait*

$$f(a) = 0, \qquad f'(a) \neq 0.$$

En effet, si a est racine, on a

$$f(T) = (T - a)q(T) \quad \text{avec} \quad q(T) = f'(a) + f_2(a).(T - a) + \cdots,$$

d'où

$$q(a) = f'(a).$$

Supposons $f'(a) = 0$; alors (Théorème 5) le polynôme $q(T)$ est divisible par $T - a$, donc f est divisible par $(T - a)^2$, et a n'est pas racine simple. Si inversement a n'est pas racine simple, on a une relation

$$f(T) = (T - a)^2 g(T),$$

d'où

$$f'(T) = 2(T - a)g(T) + (T - a)^2 g'(T),$$

ce qui prouve évidemment que $f'(a) = 0$ et achève la démonstration.

THÉORÈME 7. *Soit f un polynôme à une indéterminée à coefficients dans un corps commutatif* K *de caractéristique* 0. *Pour qu'un élément a de* K *soit racine multiple d'ordre r de f, il faut et il suffit qu'il vérifie les relations*

$$f(a) = f'(a) = \cdots = f^{(r-1)}(a) = 0, \qquad f^{(r)}(a) \neq 0. \tag{15}$$

Supposons que a soit racine multiple d'ordre r; on a alors

$$f(T) = (T - a)^r g(T), \tag{16}$$

avec en outre $g(a) \neq 0$, sinon (Théorème 5) $g(T)$ serait divisible par $T - a$, et $f(T)$ par $(T - a)^{r+1}$. Or, en dérivant k fois la relation (16) on trouve évidemment une relation de la forme

$$f^{(k)}(T) = r(r - 1)\ldots(r - k + 1)(T - a)^{r-k} g(T) + (T - a)^{r-k+1} q_k(T),$$

où q_k est un polynôme dont la forme exacte importe peu; de cette relation on tire $f^{(k)}(a) = 0$ pour $k \leqslant r - 1$, et en outre

$$f^{(r)}(a) = r!\, g(a);$$

comme $g(a) \neq 0$ et comme K est de caractéristique 0, on a bien $f^{(r)}(a) \neq 0$.

Inversement, supposons réalisées les conditions (15); comme K est de caractéristique 0, on peut écrire

$$f(T) = \sum_{k=0}^{k=n} (T-a)^k \cdot \frac{f^{(k)}(a)}{k!} = (T-a)^r \left[\frac{f^{(r)}(a)}{r!} + (T-a)\frac{f^{(r+1)}(a)}{(r+1)!} + \cdots \right];$$

autrement dit, on a une relation

$$f(T) = (T-a)^r g(T)$$

avec $g(a) \neq 0$; si f était divisible par $(T-a)^{r+1}$ il est clair, puisque K[T] est un anneau d'intégrité, que g serait divisible par $T-a$, ce qui est impossible puisque $g(a)$ n'est pas nul; par conséquent, $(T-a)^r$ est la plus haute puissance de $T-a$ qui divise f, et a est donc racine multiple d'ordre r, ce qui achève la démonstration.

Exemple 5. K étant un corps de caractéristique 0, cherchons à quelle condition l'équation

$$x^3 + px + q = 0$$

à coefficients p, q dans K admet une racine multiple a. Celle-ci doit annuler le premier membre *et* sa dérivée (Théorème 6) i.e. vérifier

$$\begin{aligned} a^3 + pa + q &= 0 \\ 3a^2 + p &= 0; \end{aligned}$$

multipliant la première relation par 3, la seconde par a, et retranchant membre à membre, on en déduit que $2pa + 3q = 0$, i.e. que

$$a = -3q/2p;$$

et en portant ce résultat dans la relation $3a^2 + p = 0$, on en conclut aussitôt que les coefficients p et q doivent vérifier

$$4p^3 + 27q^2 = 0.$$

Inversement, si cette relation est vérifiée, il est immédiat de voir que $-3q/2p$ est racine double de l'équation considérée.

Le lecteur pourra vérifier qu'en fait ces résultats supposent seulement que la caractéristique de K est différente de 2 et 3, et non pas que K est de caractéristique 0.

§ 31. Anneaux principaux

Rappelons (§ 8, *Exemple* 10) qu'on appelle *anneau principal* tout anneau d'intégrité commutatif dont tous les idéaux sont principaux. Nous verrons au § suivant que, si K est un *corps*, l'anneau K[X] des polynômes à *une* variable à coefficients dans K est principal. Dans ce §, on va établir un certain nombre de propriétés arithmétiques des anneaux principaux; ces propriétés généralisent celles des nombres entiers, et seront appliquées aux polynômes au § suivant. *Dans tout ce § on désigne par* K *un anneau principal.*

1. *Plus grand commun diviseur*

Soient K un anneau principal et $x_1, \ldots, x_n$ des éléments de K; soit

$$\mathrm{I} = (x_1, \ldots, x_n)$$

l'idéal (*) de K engendré par $x_1, \ldots, x_n$. Comme K est principal, cet idéal est engendré par un élément d, unique à ceci près qu'on peut le remplacer par ud, où u est un élément inversible (une « unité ») quelconque de l'anneau K (**).

On appelle **plus grand commun diviseur** (p.g.c.d.) de $x_1, \ldots, x_n$ tout élément d de K tel que

$$(1) \qquad (x_1, \ldots, x_n) = (d);$$

comme le premier membre est l'ensemble des $y \in \mathrm{K}$ tels qu'il existe $u_1, \ldots, u_n \in \mathrm{K}$ vérifiant

$$(2) \qquad y = u_1x_1 + \cdots + u_nx_n,$$

(*) On espère que le lecteur débutant ne confondra pas $(x_1, \ldots, x_n)$ avec l'élément de K^n qu'on désigne par la même notation ! La notation utilisée ici pour désigner l'idéal engendré par $x_1, \ldots, x_n$ est traditionnelle en Arithmétique.

(**) Supposons en effet $(d) = (d')$; il existe $u, v \in \mathrm{K}$ tels que $d' = ud$, $d = vd'$, d'où $vud = d$; si $d \neq 0$ on en déduit (puisque K est intègre) que $uv = 1$, en sorte que u est inversible.

on voit que d est encore caractérisé (à une « unité » près) par le fait que l'ensemble des éléments de la forme (2) est identique à l'ensemble des multiples de d dans K. En particulier, comme $d \in (d)$, on voit qu'*il existe* $u_1, \ldots, u_n \in K$ *tels que*

$$d = u_1x_1 + \cdots + u_nx_n, \tag{3}$$

résultat connu sous le nom de *théorème de Bezout*.

La terminologie utilisée pour désigner d est justifiée par le résultat suivant : *pour qu'un élément de* K *divise simultanément* $x_1, \ldots, x_n$, *il faut et il suffit qu'il divise* d.

La relation (1) montre en effet que l'idéal (d) contient les x_i, qui sont donc des multiples de d; il est par suite clair que tout diviseur de d divise les x_i.

Soit inversement $m \in K$ un diviseur de $x_1, \ldots, x_n$; écrivons

$$x_i = my_i \quad (1 \leqslant i \leqslant n)$$

et portons dans (3); il vient

$$d = m(u_1y_1 + \cdots + u_ny_n),$$

ce qui prouve que m divise d, et achève la démonstration.

Exemple 1. Prenons pour K l'anneau **Z**, qui est effectivement principal (§ 10, *Exemple* 9). Étant donnés des entiers $x_1, \ldots, x_n$, il existe alors un moyen « naturel » ou « canonique » de choisir un générateur de l'idéal $(x_1, \ldots, x_n)$, c'est de prendre *le générateur positif* de cet idéal. On peut alors parler *du* pgcd de $x_1, \ldots, x_n$, comme on le fait classiquement. Les diviseurs communs à $x_1, \ldots, x_n$ étant aussi les diviseurs de ce pgcd, il est alors clair que celui-ci (choisi positif) est le plus grand de tous les diviseurs communs à $x_1, \ldots, x_n$. On retrouve donc bien la notion classique, complétée par le théorème de Bezout qu'on ne démontre pas dans l'enseignement élémentaire de l'Arithmétique. Voir l'*Exemple* 8 du § 7.

2. *Éléments premiers entre eux*

On dit que les éléments $x_1, \ldots, x_n$ de K sont **premiers entre eux** s'ils admettent 1 pour p.g.c.d.

THÉORÈME 1. *Soient* $x_1, \ldots, x_n$ *des éléments d'un anneau principal* K. *Les propriétés suivantes sont équivalentes:*

a) $x_1, \ldots, x_n$ *sont premiers entre eux;*

b) les seuls diviseurs communs à $x_1, \ldots, x_n$ *sont les éléments inversibles de* K;

c) il existe des éléments $u_1, \ldots, u_n$ *de* K *tels que*

$$u_1x_1 + \cdots + u_nx_n = 1;$$

d) pour tout $y \in K$, *il existe des éléments* $u_1, \ldots, u_n$ *de* K *tels que*

$$y = u_1x_1 + \cdots + u_nx_n.$$

a) signifie que

$$(x_1, \ldots, x_n) = (1) = \mathrm{K},$$

d'où l'équivalence de *a*) et *d*). Si les x_i sont premiers entre eux, leurs diviseurs communs sont les diviseurs de 1, i.e. les éléments inversibles de K; inversement, si tout diviseur commun aux x_i est inversible, alors le (ou, plus correctement, un) p.g.c.d. des x_i est inversible, et comme les multiples d'un élément inversible constituent l'anneau K tout entier on a donc $(x_1, \ldots, x_n) = \mathrm{K}$, de sorte que les propriétés *a*) et *b*) sont équivalentes. Enfin, *a*) implique *c*) en vertu du théorème de Bezout; et *c*) implique *d*), car *c*) signifie que l'idéal $(x_1, \ldots, x_n)$ contient 1, et est donc K tout entier. Ceci achève la démonstration.

La propriété *c*) du Théorème 1 est fort utile pour démontrer certaines propriétés « classiques » mais peu évidentes à première vue. Par exemple :

THÉORÈME 2. *Soient x, y des éléments non nuls de* K *et d un diviseur du produit xy; si d est premier à x, alors d divise y.*

Comme d et x sont premiers entre eux, il existe $u, v \in \mathrm{K}$ tels que

$$ud + vx = 1;$$

multipliant le résultat par y il vient

$$y = yud + vxy;$$

comme d divise xy et yud, il divise évidemment le second membre, donc divise aussi y, d'où le Théorème.

3. *Plus petit commun multiple*

Soient $x_1, \ldots, x_n$ des éléments non nuls de K. Les multiples de x_i sont les éléments de l'idéal (x_i); par suite, les multiples communs aux x_i sont les éléments de l'idéal $(x_1) \cap \cdots \cap (x_n)$. Celui-ci étant principal, on peut poser la définition suivante : on appelle **plus petit commun multiple** (p.p.c.m.) de $x_1, \ldots, x_n$ tout élément m de K tel que

$$(4) \qquad (x_1) \cap \cdots \cap (x_n) = (m).$$

L'élément m est unique à une unité près (autrement dit, les p.p.c.m. des x_i sont obtenus en multipliant m par un élément inversible quelconque de K), et la relation (4) montre que *les multiples communs aux x_i ne sont autres que les multiples de m.*

THÉORÈME 3. *Soient x et y des éléments non nuls de* K. *On a alors*

$$xy = md$$

où m est un p.p.c.m. *et où d est un* p.g.c.d. *de x et y.*

Remarque 1. Comme on a le droit de remplacer d par ud et m par vm où u, v sont des éléments inversibles arbitraires de K, il est clair que la relation

$$xy = md$$

ne peut être valable que si m et d sont convenablement choisis; c'est sous cette forme qu'on doit interpréter l'énoncé.

Pour démontrer le Théorème 3, posons

$$x = x'd, \qquad y = y'd,$$

et soit m' un p.p.c.m. de x' et y'; pour qu'un élément z de K soit multiple de x et de y il faut et il suffit, évidemment, que l'on puisse écrire

$$z = z'd$$

où z' est multiple commun à x' et y', autrement dit est multiple de m'. Ainsi, les multiples communs à x et y sont les multiples de $m'd$, qui est donc un p.p.c.m. de x et y. La relation à établir s'écrit donc $xy = m'd.d$ ou, en simplifiant par d^2,

$$x'y' = m'.$$

Le Théorème 3 sera donc une conséquence des deux lemmes que voici :

Lemme 1. *Soient x, y deux éléments non nuls de* K *et d un* p.g.c.d. *de x et y; posons $x = x'd$, $y = y'd$: alors x' et y' sont premiers entre eux.*

En effet il existe u, $v \in$ K tels que $ux + vy = d$, ce qui, en simplifiant par d, s'écrit $ux' + vy' = 1$, et prouve le lemme vu le Théorème 1.

Lemme 2. *Si x et y sont premiers entre eux, xy est un* p.p.c.m. *de x et y* (ou encore : les multiples communs à x et y sont les multiples de xy).

Soit m un multiple commun à x et y; posons $m = xz$; l'élément y divise donc xz; comme il est premier à x, il divise z (Théorème 2), d'où le Lemme.

Le Théorème 3 est maintenant démontré.

Le lemme 2 peut se généraliser comme suit :

Théorème 4. *Soient $x_1, \ldots, x_n$ des éléments non nuls de* K *deux à deux premiers entre eux; alors $x_1 \ldots x_n$ est un* p.p.c.m. *de $x_1, \ldots, x_n$.*

Pour $n = 2$, c'est le Lemme 2. On va donc montrer que si le Théorème est vrai pour un produit de $n - 1$ facteurs, il est vrai pour un produit de n facteurs.

Montrons d'abord par récurrence sur n que x_n *est premier au produit* $x_1 \ldots x_{n-1}$. Si $n = 2$, c'est l'hypothèse que les x_i sont deux à deux premiers entre eux. Supposons alors prouvé que x_n est premier à $x_1 \ldots x_{n-2}$, et soit d un diviseur commun à x_n et $x_1 \ldots x_{n-1}$; comme d divise x_n, qui est premier à x_{n-1}, il est clair que d est premier à x_{n-1}; comme d divise $x_n \ldots x_{n-2}x_{n-1}$, on voit donc que d divise $x_1 \ldots x_{n-2}$ (Théorème 2); mais comme x_n et $x_1 \ldots x_{n-2}$ sont premiers entre eux d'après l'hypothèse de récurrence, on voit que d est inversible, d'où notre assertion.

Nous pouvons maintenant démontrer le Théorème 4 par récurrence sur n. Soit m un multiple commun aux x_i; le théorème étant supposé établi pour un produit de $n-1$ facteurs, on voit que m est un multiple de $x_1 \dots x_{n-1}$, et aussi de x_n; ces deux éléments de K étant, comme on vient de le voir, premiers entre eux, m est donc, d'après le Lemme 2, un multiple de leur produit, ce qui achève la démonstration.

4. *Existence de diviseurs premiers*

On dit qu'un élément p de K est **premier** ou **irréductible** ou **extrémal** (*) s'il n'est pas inversible et si ses seuls diviseurs sont ceux qui sont évidents *a priori*, à savoir les éléments inversibles de K, et les éléments pu où u est inversible.

Lorsque $K = \mathbf{Z}$ on retrouve évidemment la notion classique de nombre premier; or on sait que, dans ce cas, tout $x \in K$ peut s'écrire sous forme d'un produit de nombres premiers; nous allons montrer que ce résultat s'étend aux anneaux principaux :

THÉORÈME 5. *Soit* K *un anneau principal; tout élément non nul de* K *est produit d'un élément inversible et d'éléments extrémaux de* K.

Ou encore : *tout* $x \in K$ *qui n'est ni nul ni inversible est un produit d'éléments extrémaux de* K (un élément inversible ne peut évidemment pas se décomposer en produits d'éléments extrémaux, car les diviseurs d'un élément inversible sont inversibles, et ne peuvent donc jamais être extrémaux).

Pour démontrer le Théorème 5, observons d'abord qu'un anneau principal est *a fortiori* noetherien : tous ses idéaux sont évidemment de type fini. Par conséquent (§ 18, Théorème 4 ou Remarque 2 ci-dessous) on a le résultat suivant :

LEMME 3. *Soit* X *un ensemble non vide d'idéaux d'un anneau principal* K; *alors* X *possède au moins un élément maximal*, i.e. il existe un $I \in X$ qui n'est contenu dans aucun autre $J \in X$.

Ceci étant rappelé, on va démontrer le Théorème 5 en raisonnant par l'absurde. Soit X l'ensemble des idéaux $I = (x)$ de K pour lesquels x est un élément non nul de K qui *ne peut pas* s'écrire comme produit d'un élément inversible et d'éléments extrémaux de K : pour établir le Théorème, tout revient à montrer que l'ensemble X est *vide*. S'il ne l'était pas, il contiendrait au moins un élément maximal, soit (a). L'élément a ne pouvant pas se décomposer en produit d'un élément inversible et d'éléments extrémaux ne serait ni inversible, ni extrémal. Par suite on pourait écrire $a = bc$ où ni b ni c ne seraient inversibles; il est clair qu'alors les idéaux (b) et (c) contiendraient *strictement* (a); comme (a) n'est strictement contenu dans aucun idéal appartenant à l'ensemble X, on voit donc que ni (b) ni (c) n'appartiendrait à X; donc le Théorème 5 serait vrai pour b et pour c; mais alors il serait évidemment vrai pour leur produit a, contrairement à l'hypothèse que $(a) \in X$, ce qui termine la démonstration.

(*) La première terminologie s'emploie plutôt pour l'anneau **Z** et les anneaux analogues, la seconde pour les anneaux de polynômes, et la troisième pour les anneaux principaux généraux.

Remarque 2. On peut naturellement se passer du Lemme 3 lorsque $K = \mathbf{Z}$ (à vrai dire, les démonstrations élémentaires dans ce cas ne réfèrent pas explicitement au Lemme 3, mais en font néanmoins usage implicitement; la démonstration classique consisterait à introduire le *plus petit* entier $a > 0$ *non* décomposable en facteurs premiers, s'il en existe, puis à observer que a ne peut être premier, donc qu'on peut écrire $a = bc$ avec $0 < b < a$ et $0 < c < a$, auquel cas b et c sont décomposables en facteurs premiers et a aussi par conséquent; il est clair que la démonstration générale est directement calquée sur ce raisonnement traditionnel). On verra au § suivant qu'on peut aussi se passer du Lemme 3 (ou, si l'on préfère, le démontrer trivialement) lorsque $K = L[X]$ où L est un corps.

Rappelons au lecteur que, de toute façon, la démonstration du Lemme 3 est fort simple, et s'obtient comme suit. Si le Lemme 3 était faux, on pourrait, en partant d'un $I_1 \in X$ quelconque, construire un $I_2 \in X$ contenant *strictement* I_1, puis un $I_3 \in X$ contenant *strictement* I_2, et ainsi de suite indéfiniment; pour tirer de là une contradiction tout revient à montrer que *toute suite croissante*

$$I_1 \subset I_2 \subset \cdots \subset I_n \subset \cdots,$$

d'idéaux de K *est stationnaire*, autrement dit qu'il existe un entier r tel que

$$I_n = I_r \quad \text{pour tout} \quad n \geqslant r.$$

Or soit I la réunion des I_n, qui est un idéal; comme K est principal, on a $I = (x)$ pour un $x \in K$; comme x appartient à la réunion des I_n, on a $x \in I_r$ pour au moins un indice r; mais alors il est clair que I_r contient les multiples de x, donc que $I_r \supset I$, et pour $n \geqslant r$ il vient alors

$$I_r \subset I_n \subset I \subset I_r,$$

d'où $I_n = I_r$, ce qui prouve le Lemme 3.

5. *Propriétés des éléments extrémaux*

Le résultat suivant contient plusieurs caractérisations utiles des éléments extrémaux (y compris leur définition, qu'on a inclus dans ce résultat pour obtenir un énoncé aussi complet que possible) :

THÉORÈME 6. *Soit p un élément non nul et non inversible d'un anneau principal* K; *les propriétés suivantes sont équivalentes :*

a) p est extrémal;

b) tout diviseur de p est soit inversible, soit de la forme pu où u est inversible;

c) l'idéal $I = (p)$ est maximal (autrement dit on a $I \neq K$ et les seuls idéaux de K contenant I sont I et K);

d) si p divise un produit d'éléments de K, *p divise l'un au moins des facteurs de ce produit.*

L'équivalence des assertions *a*) et *b*) n'est autre que la définition des éléments extrémaux. Pour montrer l'équivalence de *b*) et *c*), considérons un idéal J contenant I; comme K est principal on a $J = (x)$ pour un $x \in K$, et dire que J contient I signifie que x divise p; de plus, la relation $J = K$ signifie que x est inversible, et la

relation J = I que $x = pu$ avec u inversible : l'équivalence de b) et c) résulte aussitôt de ces remarques.

Montrons enfin que la propriété d) caractérise aussi les éléments extrémaux. Supposons p extrémal, et soient $x_1, \ldots, x_n$ des éléments non nuls de K tels que p divise le produit $x_1 \ldots x_n$; pour montrer que p divise l'un au moins des x_i on écrit

$$x_1 \ldots x_n = (x_1 \ldots x_{n-1})x_n,$$

et un raisonnement par récurrence montre évidemment qu'il suffit d'examiner le cas où $n = 2$, autrement dit de prouver que *si p divise xy, alors p divise soit x soit y.* En vertu du Théorème 2, tout revient à prouver que, *quel que soit $x \neq 0$, ou bien p divise x ou bien p et x sont premiers entre eux*; or considérons l'idéal (p, x) engendré par p et x; il contient évidemment l'idéal (p); d'après l'assertion c) du Théorème 6, qui est déjà établie, deux cas seulement sont donc possibles : ou bien

$$(p, x) = \mathrm{K},$$

et alors p et x sont premiers entre eux, ou bien

$$(p, x) = (p),$$

auquel cas x appartient à (p), donc est multiple de p.

Nous avons démontré que tout élément extrémal vérifie d). Inversement considérons un élément non nul et non inversible vérifiant d), et montrons qu'il est extrémal. En vertu du Théorème 5, l'élément p considéré peut s'écrire

$$p = p_1 \ldots p_n$$

où les p_i sont extrémaux; d'après d), p divise l'un au moins des p_i; mais comme p_i est extrémal, et p non inversible, il s'ensuit que $p = up_i$ avec u inversible, et par suite p est extrémal, ce qui termine la démonstration du Théorème.

COROLLAIRE. *Soient $x, y_1, \ldots, y_n$ des éléments non nuls d'un anneau principal* K, *et supposons x et y_i premiers entre eux quel que soit i. Alors x est premier au produit $y_1 \ldots y_n$.*

Supposons en effet qu'il existe un diviseur commun non inversible d à x et $y_1 \ldots y_n$; comme d n'est pas inversible, il est produit d'éléments extrémaux, donc multiple d'au moins un élément extrémal; par suite il existe un élément extrémal p qui divise à la fois x et $y_1 \ldots y_n$; d'après le Théorème 6, d), il existe un i tel que p divise y_i, ce qui contredit l'hypothèse que x et y_i sont premiers entre eux, et démontre le Corollaire.

Il va de soi qu'en supposant $\mathrm{K} = \mathbf{Z}$ on ne parviendrait pas à simplifier si peu que ce soit les démonstrations précédentes.

6. *Unicité de la décomposition en facteurs premiers*

On dit que deux éléments extrémaux p' et p'' de K sont **associés** s'il existe un élément inversible u de K tel que $p'' = up'$; cela signifie évidemment que les idéaux (p') et (p'') sont égaux. Lorsque $\mathrm{K} = \mathbf{Z}$, cela veut dire que $p'' = \pm p'$.

THÉORÈME 7. *Soit x un élément non nul et non inversible d'un anneau principal* K. *Soient*

$$x = p'_1 \dots p'_m = p''_1 \dots p''_n$$

deux décompositions de x en produit d'éléments extrémaux de K. *On a alors $m = n$ et il existe une permutation σ des indices* 1, ..., n *telle que p'_i et $p''_{\sigma(i)}$ soient associés pour tout i tel que* $1 \leqslant i \leqslant n$.

Autrement dit, une décomposition étant donnée, on obtient *toutes* les autres en effectuant sur celle-ci les opérations « triviales » que voici : (1) modification de l'ordre des termes (2) multiplication de chaque facteur extrémal par un élément inversible de K (ces éléments étant choisis de telle sorte que leur produit soit égal à 1, de façon à ne pas modifier le produit des éléments extrémaux considérés). Lorsque $K = \mathbf{Z}$, on convient souvent de se borner à des nombres premiers *positifs*, ce qui élimine la possibilité (2) puisque, dans ce cas, le seul nombre premier associé à p est $-p$.

Pour établir le Théorème 7, considérons la relation

$$p'_1 \dots p'_m = p''_1 \dots p''_n; \tag{5}$$

p'_1 divise le produit $p''_1 \dots p''_n$ et est extrémal; il divise donc l'un des p''_j; en permutant l'ordre des facteurs du second produit, on peut donc supposer que p'_1 divise p''_1; mais comme p''_1 est premier, il s'ensuit que

$$p''_1 = u_1 p'_1$$

avec u_1 inversible; simplifiant la relation (5) par p'_1, il reste alors

$$p'_2 \dots p'_m = u_1 p''_2 \dots p''_n;$$

en raisonnant comme ci-dessus, on en déduit que p'_2 est associé à l'un des p''_j $(2 \leqslant j \leqslant n)$ et en poursuivant ainsi le raisonnement, il est clair qu'on parvient au Théorème 7 (le lecteur désireux de terminer correctement la démonstration devra raisonner par récurrence sur l'entier m).

Dans la pratique, on écrit souvent de la façon suivante les décompositions en facteurs extrémaux. On choisit une fois pour toutes un ensemble P d'éléments extrémaux de K possédant la propriété suivante : pour tout élément extrémal p de K, il existe *un et un seul* $p' \in P$ qui soit associé à p (pour construire un tel ensemble P, il suffit, en vertu du Théorème 6, *c*), de considérer l'ensemble des idéaux maximaux de K, et de choisir une fois pour toutes un générateur de chacun de ces idéaux; si $K = \mathbf{Z}$ on choisit, pour chaque idéal maximal I, *le* nombre premier *positif* qui engendre I). Soit alors

$$x = p'_1 \dots p'_n$$

une décomposition d'un $x \in K$ non inversible en produit d'éléments extrémaux; pour chaque i on peut écrire $p'_i = u_i p_i$ avec $p_i \in P$ et u_i inversible; d'où évidemment

$$x = u p_1 \dots p_n \quad \text{avec } u \text{ inversible, } p_1, \dots, p_n \in P; \tag{6}$$

la décomposition (6) est alors unique à l'ordre près des facteurs (car deux éléments de P ne peuvent être associés sans être égaux).

Il peut naturellement arriver que, dans la décomposition (6), un élément de P soit répété plusieurs fois; en bloquant ensemble les facteurs égaux, on trouve donc aussi une décomposition de la forme

$$x = up_1^{n_1} \dots p_r^{n_r} \tag{7}$$

où $p_1, \dots, p_r$ sont des éléments de P deux à deux distincts, et les n_i des exposants positifs. Pour écrire sous une forme plus frappante ce résultat, considérons, pour chaque $p \in P$, le nombre (éventuellement nul) de facteurs de la décomposition (6) qui sont égaux à p, et notons-le

$$v_p(x);$$

il est clair qu'on a

$$v_p(x) = 0 \quad \text{pour presque tout } p \in P, \tag{8}$$

autrement dit qu'on n'a $v_p(x) \geqslant 1$ que pour un nombre fini d'éléments de P. Ceci dit, la décomposition (7), obtenue en groupant ensemble les facteurs égaux dans la décomposition (6), s'écrit encore sous la forme

$$x = u \cdot \prod_{p \in P} p^{v_p(x)}; \tag{9}$$

le produit figurant au second membre comporte en apparence une infinité de facteurs; mais d'après (8) on a

$$p^{v_p(x)} = 1 \quad \text{pour presque tout } p \in P,$$

de sorte que le produit en question ne comporte qu'un nombre fini de termes autres que 1.

7. *Calcul du pgcd et du ppcm à l'aide de la décomposition en facteurs premiers*

La décomposition (9) permet d'exprimer très simplement les propriétés de divisibilité dans l'anneau K. Tout repose sur le résultat suivant :

LEMME 4. *Soient x et y deux éléments non nuls de* K; *pour que x divise y il faut et il suffit que l'on ait*

$$v_p(x) \leqslant v_p(y)$$

pour tout $p \in P$.

La condition est suffisante car, si elle est remplie, on a

$$x = u' \prod p^{v_p(x)}$$
$$y = u'' \prod p^{v_p(x)+v_p}$$

avec des entiers

$$n_p = v_p(y) - v_p(x)$$

tous positifs et presque tous nuls, d'où résulte que $y = xz$ avec

$$z = u'^{-1}u'' \prod p^{n_p}.$$

Inversement, supposons $y = xz$; alors, en utilisant les décompositions en facteurs premiers de x et z, il vient

$$y = u \cdot \prod p^{v_p(x)+v_p(z)},$$

avec un $u \in K$ inversible, et comme la décomposition de y en produit d'éléments de P est unique on voit que $v_p(y) = v_p(x) + v_p(z)$, donc que $v_p(y) \geqslant v_p(x)$ pour tout $p \in P$, ce qui termine la démonstration du Lemme 4.

Étant donnés deux éléments non nuls x et y de K, il est facile, à l'aide du lemme 4, de construire leur pgcd et leur ppcm. Soit en effet d un pgcd de x et y; on doit exprimer que les diviseurs de x et y sont les diviseurs de d; or les diviseurs de x et y sont, d'après le lemme 4, les $z \in K$ vérifiant

$$v_p(z) \leqslant v_p(x) \qquad \text{et} \qquad v_p(z) \leqslant v_p(y),$$

autrement dit

$$v_p(z) \leqslant \text{Min}[v_p(x),\ v_p(y)] \tag{10}$$

pour tout $p \in P$; et les diviseurs de d sont les $z \in K$ qui vérifient

$$v_p(z) \leqslant v_p(d) \tag{11}$$

pour tout $p \in P$. Pour que d soit un pgcd, il est donc nécessaire et suffisant que les conditions (10) et (11) soient équivalentes, autrement dit que

$$v_p(d) = \text{Min}\,[v_p(x), v_p(y)] \quad \text{pour tout } p \in P, \tag{12}$$

et on retrouve ainsi la règle classique : l'exposant de p dans la décomposition de d est égal à celui des exposants de p dans x et y qui est le plus petit.

Un raisonnement analogue montrerait qu'un ppcm m de x et y est donné par

$$v_p(m) = \text{Max}[v_p(x), v_p(y)] \quad \text{pour tout } y \in P. \tag{13}$$

Autrement dit, les règles connues dans le cas de l'anneau **Z** s'étendent à tous les anneaux principaux — et s'établissent à l'aide des mêmes raisonnements que dans le cas classique.

Remarque 3. Il va de soi qu'un produit de la forme

$$\prod_{p \in P} p^{n_p}$$

n'a de sens dans K que si les exposants n_p sont des entiers *tous* positifs et *presque*

tous nuls. Si l'on voulait utiliser les relations (12) et (13) pour définir le pgcd et le ppcm de x et y, il faudrait vérifier que les seconds membres de ces relations vérifient les conditions en question. On l'établit comme suit. Soit X (resp. Y) l'ensemble des $p \in P$ tels que $v_p(x) \neq 0$ (resp. $v_p(y) \neq 0$) — autrement dit, l'ensemble des $p \in P$ qui divisent x (resp. y); X et Y sont des ensembles finis, donc aussi $Z = X \cup Y$; pour $p \notin Z$, on a $v_p(x) = v_p(y) = 0$, donc

$$\text{Max}\,[v_p(x), v_p(y)] = \text{Min}\,[v_p(x), v_p(y)] = 0,$$

ce qui est le résultat cherché.

8. *Décomposition en éléments simples des fractions sur un anneau principal*

Soient K un anneau principal et F le corps des fractions de K, défini au § 29. Soit

$$x = a/b, \qquad a, b \in K, \qquad b \neq 0$$

un élément de F; conservant les notations du n° précédent, on peut écrire

$$a = u \prod_{p \in P} p^{v_p(a)}, \qquad b = v \prod_{p \in P} p^{v_p(b)},$$

d'où

$$x = w \cdot \prod_{p \in P} p^{v_p(x)} \tag{14}$$

avec un élément inversible w de K, et des entiers

$$v_p(x) = v_p(a) - v_p(b)$$

qui sont encore presque tous nuls, mais peuvent être de signe quelconque. On vérifie facilement — le lecteur l'établira à titre d'exercice — que la décomposition (14) de x est unique.

Nous allons maintenant établir, pour les éléments de F, une décomposition d'une toute autre nature; le résultat qui suit est utile en Analyse (calcul des primitives des fractions rationnelles à coefficients réels ou complexes), et à peu près complètement dépourvu d'intérêt par ailleurs.

THÉORÈME 8. *Soit*

$$x = \frac{a}{p_1^{r_1} \dots p_n^{r_n}}$$

un élément de F; *on suppose* $a \in K$, *les* r_i *positifs, et les* $p_i \in K$ *extrémaux et deux à deux non associés. Alors il existe des éléments* $a_1, \dots, a_n$ *de* K *tels que*

$$x = \frac{a_1}{p_1^{r_1}} + \dots + \frac{a_n}{p_n^{r_n}}.$$

Autrement dit, tout élément de F est somme de fractions de la forme a/p^n avec $a \in K$, p extrémal et $n \geqslant 0$.

Exemple 2. Prenons $K = \mathbf{Z}$ et

$$x = 5/18 = 5/2.3^2;$$

on a alors

$$x = 1/2 - 2/9,$$

ce qui est le Théorème 8 dans ce cas.

Le Théorème sera visiblement une conséquence des deux lemmes que voici :

LEMME 5. *Supposons*

$$x = a/b_1 \dots b_n$$

où les éléments b_i de K sont deux à deux premiers entre eux. Alors il existe des $a_i \in K$ tels que

$$x = a_1/b_1 + \cdots + a_n/b_n.$$

LEMME 6. *Soient $p_1, \dots, p_n$ des éléments extrémaux deux à deux non associés de K. Alors, quels que soient les entiers positifs $r_1, \dots, r_n$, les éléments*

$$b_1 = p_1^{r_1}, \dots, b_n = p_n^{r_n}$$

sont deux à deux premiers entre eux.

Démontrons d'abord le lemme 6; soit d un diviseur commun à b_i et b_j $(i \neq j)$; si d n'est pas inversible, il admet un diviseur premier p, qui divise donc b_i et b_j; mais si p divise

$$p_i^{r_i} = p_i p_i \dots p_i,$$

il divise p_i et est donc associé à p_i; si donc b_i et b_j n'étaient pas premiers entre eux, il existerait un élément extrémal p de K associé à la fois à p_i et p_j, contrairement à l'hypothèse que p_j et p_i sont non associés pour $i \neq j$; d'où le lemme 6.

Démontrons maintenant le lemme 5, tout d'abord pour $n = 2$. Comme b_1 et b_2 sont premiers entre eux, il existe $u_1, u_2 \in K$ tels que

$$u_1 b_1 + u_2 b_2 = 1;$$

on a alors

$$\frac{a}{b_1 b_2} = \frac{a(u_1 b_1 + u_2 b_2)}{b_1 b_2} = \frac{a u_2}{b_1} + \frac{a u_1}{b_2}$$

ce qui établit le lemme 5 dans ce cas. Pour l'établir dans le cas général, on va raisonner par récurrence sur n; puisque (Corollaire du Théorème 6) b_n est premier à $b_1 \dots b_{n-1}$ et puisque le lemme est déjà établi pour un produit de deux facteurs, on peut écrire

$$x = \frac{a'}{b_1 \dots b_{n-1}} + \frac{a_n}{b_n};$$

il suffit alors d'appliquer l'hypothèse de récurrence à la première fraction du second membre pour obtenir la décomposition cherchée de x.

§ 32. Division des polynômes

1. *Division des polynômes à une variable*

Soit

$$f(X) = a_0 + a_1X + \cdots + a_nX^n + \cdots$$

un polynôme à une indéterminée, à coefficients dans un anneau commutatif K; on appelle **coefficient dominant** de f le coefficient du terme de plus haut degré de f (ce qui suppose $f \neq 0$); si f est de degré n, le coefficient dominant de f est donc le coefficient de X^n. On dit que f est **unitaire** si son coefficient dominant est un élément *inversible* de K; si K est un corps, tout polynôme $f \neq 0$ est donc unitaire.

Le résultat suivant est analogue à celui qui est à la base de la théorie de la division (avec reste) des entiers rationnels :

THÉORÈME 1. *Soient* K *un anneau commutatif et* g *un polynôme unitaire à une indéterminée à coefficients dans* K. *Pour tout polynôme* $f \in K[X]$, *il existe des polynômes* q, $r \in K[X]$ *vérifiant les relations*

$$f = gq + r, \qquad d^\circ(r) < d^\circ(g) \quad ; \tag{1}$$

et les polynômes q *et* r *sont uniques.*

Soit

$$g(X) = b_0 + b_1X + \cdots + b_nX^n$$

avec b_n inversible; en multipliant les deux membres de (1) par la constante $b_n^{-1} \in K$, on se ramène évidemment au cas où g est de la forme

$$g(X) = X^n + b_{n-1}X^{n-1} + \cdots + b_0, \tag{2}$$

ce que nous supposerons dans ce qui suit.

Considérons alors (pour f donné et q variable) tous les polynômes de la forme $f - gq$; leurs degrés sont des entiers positifs ou le symbole $-\infty$ (ce qui se produit si f est multiple de g — dans ce cas le Théorème 1 est bien entendu trivial!); par conséquent, il est possible de choisir q de telle sorte que le degré de $f - gq$ soit mini-

mum; on a alors

$$d^\circ(f - gq) \leqslant d^\circ(f - gq') \tag{3}$$

pour *tout* polynôme $q' \in \mathrm{K}[\mathrm{X}]$. Posant

$$f = gq + r, \tag{4}$$

tout revient à faire voir que

$$d^\circ(r) < d^\circ(g). \tag{5}$$

Or supposons qu'il n'en soit pas ainsi, et posons

$$r(\mathrm{X}) = c_{n+k}\mathrm{X}^{n+k} + c_{n+k-1}\mathrm{X}^{n+k-1} + \cdots$$

avec $c_{n+k} \neq 0$ et $k \geqslant 1$ puisqu'on suppose la relation (5) fausse; il est clair que r et le polynôme

$$c_{n+k}\mathrm{X}^k g(\mathrm{X})$$

ont le même coefficient dominant c_{n+k}; par suite on peut écrire

$$r(\mathrm{X}) = c_{n+k}\mathrm{X}^k g(\mathrm{X}) + r'(\mathrm{X}) \quad \text{avec } d^\circ(r') < d^\circ(r); \tag{6}$$

(4) s'écrit alors

$$f(\mathrm{X}) = [q(\mathrm{X}) + c_{n+k}\mathrm{X}^k]g(\mathrm{X}) + r'(\mathrm{X}),$$

et en posant

$$q'(\mathrm{X}) = q(\mathrm{X}) + c_{n+k}\mathrm{X}^k$$

il vient donc

$$f = gq' + r' \qquad \text{avec} \qquad d^\circ(r') < d^\circ(r);$$

cette dernière inégalité s'écrit encore $d^\circ(f - gq') < d^\circ(f - gq)$, et contredit (3), ce qui établit (5).

Nous avons donc démontré qu'il existe *au moins* un couple de polynômes q, r vérifiant (1); il reste à montrer qu'il en existe *au plus* un. Or considérons deux relations de la forme

$$f = gq_1 + r_1 = gq_2 + r_2$$

avec

$$d^\circ(r_1) < d^\circ(g), \qquad d^\circ(r_2) < d^\circ(g); \tag{7}$$

on en tire

$$g(q_1 - q_2) = r_2 - r_1; \tag{8}$$

tout revient évidemment à montrer que $q_1 = q_2$; or, d'après (7), on a $d^\circ(r_2 - r_1) < d^\circ(g)$; la relation $q_2 - q_1 = 0$ sera donc une conséquence du résultat suivant :

LEMME 1. *Soient g un polynôme unitaire et q un polynôme non nul; on a*

$$d^0(gq) = d^0(g) + d^0(q) \geqslant d^0(g).$$

Posant

$$g(X) = b_n X^n + \cdots + b_0, \qquad q(X) = c_m X^m + \cdots + c_0$$

avec b_n inversible et c_m non nul, on voit que le coefficient de X^{m+n} dans gq est égal à $b_n c_m$; comme b_n est *inversible*, ce produit ne peut être nul que si $c_m = 0$, ce qui n'est pas le cas; d'où le lemme.

Le Théorème 1 est donc démontré.

Dans les hypothèses du Théorème 1 (si K est un *corps*, le Théorème 1 s'applique pour peu que $g \neq 0$), on dit que q est le **quotient** et r le **reste** de la **division** de f par g. La démonstration du Théorème 1 conduit à une méthode pratique pour les calculer. En effet, posons

$$f(X) = a_m X^m + \cdots + a_0, \qquad g(X) = b_n X^n + \cdots + b_0$$

avec a_m non nul et b_n inversible; si $m < n$, la division s'effectue trivialement : on prend $q = 0$, $r = f$; si au contraire $m \geqslant n$, on a visiblement

$$f(X) = a_m b_n^{-1} X^{m-n} g(X) + f_1(X) = c_k X^k g(X) + f_1(X) \tag{9}$$

avec

$$d^0(f_1) \leqslant d^0(f) - 1;$$

si f_1 est de degré inférieur à g, on a terminé et de façon précise on a

$$q(X) = a_m b_n^{-1} X^{m-n}, \qquad r(X) = f_1(X)$$

dans ce cas; si au contraire $d^0(f_1) \geqslant d^0(g)$, on effectue sur f_1 la même opération que sur f, en écrivant

$$f_1(X) = c_h X^h g(X) + f_2(X) \tag{10}$$

avec

$$d^0(f_2) < d^0(f_1);$$

en combinant (9) et (10); il vient

$$f(X) = (c_k X^k + c_h X^h) g(X) + f_2(X), \qquad d^0(f_2) \leqslant d^0(f) - 2;$$

si le degré de f_2 est inférieur à celui de g, le problème est résolu; sinon, on écrit

$$f_2(X) = c_l X^l g(X) + f_3(X)$$

avec

$$d^0(f_3) < d^0(f_2),$$

d'où

$$f(X) = (c_k X^k + c_h X^h + c_l X^l) g(X) + f_3(X), \quad d^0(f_3) \leqslant d^0(f) - 3;$$

il est clair qu'en procédant ainsi on parvient à former le quotient et le reste de la division de f par g.

Exemple 1. Prenons $f(X) = X^6 - X^4 - X^2 + 1$, $g(X) = X^3 - 1$; on dispose alors les opérations comme suit :

$$\begin{array}{rrrrr|l} X^6 - X^4 & & - X^2 & + 1 & X^3 - 1 \\ \cline{5-5} - X^4 + X^3 & - X^2 & + 1 & & X^3 - X + 1 \\ X^3 - X^2 & - X & + 1 & & \\ - X^2 & - X & + 2 & & \end{array}$$

ici le quotient est $X^3 - X + 1$, et le reste $-X^2 - X + 2$. La méthode adoptée pour les calculs est la même que pour les divisions de nombres entiers.

2. *Idéaux d'un anneau de polynômes à une indéterminée*

L'une des conséquences les plus importantes du Théorème 1 est le résultat que voici :

Théorème 2. *Soit* K *un* corps *commutatif; l'anneau* K[X] *des polynômes à une indéterminée à coefficients dans* K *est principal.*

L'anneau K[X] est commutatif et intègre (§ 27, Théorème 1). Soit I un idéal de K[X]; pour montrer qu'il est principal, on peut supposer qu'il ne se réduit pas à 0. Parmi les éléments *non nuls* de I, choisissons un polynôme f *de degré minimum* : il est clair qu'alors les relations

$$r \in I \quad \text{et} \quad d^\circ(r) < d^\circ(f) \quad \text{impliquent } r = 0. \tag{11}$$

Cela dit, soit g un élément de I; comme K est un corps et f non nul, le Théorème 1 montre que

$$g = fq + r \quad \text{avec } d^\circ(r) < d^\circ(f);$$

comme I contient f et g, il contient aussi $g - fq$, donc r; par suite on peut faire usage de (11), et on voit que $r = 0$. Ainsi tout élément de I est multiple de f, et tout multiple de f est évidemment dans I; l'idéal I est donc engendré par f, ce qui achève la démonstration.

Le lecteur devra rapprocher la démonstration précédente du § 7, *Exemple* 8, les méthodes utilisées dans ces deux cas étant analogues.

Soit $I = (f)$ un idéal de l'anneau K[X]; le polynôme f est unique à ceci près qu'on peut le multiplier par un élément inversible de K[X]; notons à ce sujet le résultat suivant :

Lemme 2. *Soit* K *un anneau d'intégrité commutatif; les éléments inversibles de* K[X] *sont ceux de* K.

Évidemment, tout élément inversible de K est inversible dans K[X]. Soit réciproquement

$$f(X) = a_n X^n + \cdots + a_0, \quad a_n \neq 0$$

un élément inversible de K, et

$$g(X) = b_m X^m + \cdots + b_0, \quad b_m \neq 0,$$

son inverse; on a donc, en effectuant le produit de f par g,

$$1 = a_n b_m X^{n+m} + \cdots,$$

les termes non écrits étant de degrés inférieurs à $m + n$; comme K est un anneau d'intégrité on a $a_n b_m \neq 0$; donc $m + n = 0$, et f se réduit à l'élément a_n de K, qui est inversible dans K vu la relation $a_n b_m = 1$; ceci achève la démonstration.

On aurait pu aussi (§ 27, Théorème 1) écrire directement que

$$d^\circ(f) + d^\circ(g) = d^\circ(fg) = 0,$$

ce qui exige que f et g se réduisent à des éléments de K.

Le lemme 2 montre que, si f est un générateur d'un idéal I de K[X] (on suppose à nouveau que K soit un corps), les autres générateurs de I s'obtiennent en multipliant f par une constante (i.e. un élément de K) non nulle. En particulier, en multipliant f par l'inverse de son coefficient dominant, *on peut s'arranger pour que le coefficient dominant de f soit égal à 1, et alors f est le seul générateur de* I *possédant cette propriété.* On voit qu'ici comme dans le cas de l'anneau **Z**, il existe un moyen « canonique » de choisir un générateur pour chaque idéal non nul de l'anneau considéré.

Le Théorème 2 est très important et très utile pour résoudre des problèmes relativement élémentaires; mais dès qu'on aborde la « Géométrie Algébrique » (voir la *Remarque* 3 plus loin) on a besoin de résultats beaucoup plus généraux que le Théorème 2; dans cet ordre d'idées, l'énoncé fondamental est le suivant : *soit* K *un anneau commutatif noethérien* (i.e. dont tous les idéaux sont de type fini); *alors l'anneau* K[X] *est encore noethérien.* On trouvera dans l'*Exercice* 27 de ce § une démonstration simple de ce résultat. En raisonnant par récurrence sur n, on en déduit plus généralement que *l'anneau* $K[X_1, \ldots, X_n]$ *est noethérien si* K *est noethérien*; en particulier, *si* K *est un corps, l'anneau* $K[X_1, \ldots, X_n]$ *est noethérien quel que soit n.* Ces résultats, dus à Hilbert, expliquent pourquoi les calculs concernant des polynômes à coefficients dans un corps peuvent toujours s'effectuer « en un nombre fini de pas »; mais leur importance réelle dépasse de loin cette remarque d'ordre plus ou moins métaphysique.

3. *Pgcd et ppcm de plusieurs polynômes; polynômes irréductibles*

Le Théorème 2 permet d'appliquer les résultats du § précédent à l'anneau K[X] lorsque K est un *corps*.

Soient en particulier $f_1, \ldots, f_n$ des polynômes non nuls à une indéterminée à coefficients dans K; alors ces polynômes admettent un **plus grand commun diviseur** d; et on a le résultat suivant :

THÉORÈME 3. *Soient $d, f_1, \ldots, f_n$ des polynômes non nuls à une indéterminée à coefficients dans un corps commutatif* K. *Les propriétés suivantes sont équivalentes :*

a) d est un p.g.c.d. *de* $f_1, \ldots, f_n$;

b) les diviseurs communs à $f_1, \ldots, f_n$ *sont les diviseurs de d;*

c) pour qu'un polynôme puisse s'écrire sous la forme

$$u_1(X) f_1(X) + \cdots + u_n(X) f_n(X)$$

avec des $u_i \in K[X]$, *il faut et il suffit qu'il soit multiple de d;*

d) $d(X)$ *peut se mettre sous la forme*

$$u_1(X) f_1(X) + \cdots + u_n(X) f_n(X)$$

et est de degré minimum dans l'ensemble des polynômes non nuls possédant cette propriété.

L'équivalence de *a)* et *b)* est claire; *c)* signifie que *d* engendre l'idéal engendré par $f_1, \ldots, f_n$, ce qui est précisément la définition du p.g.c.d. donnée au § précédent; enfin, *d)* s'obtient en observant — voir la démonstration du théorème 2 — que les générateurs d'un idéal sont les éléments de degré minimum parmi les éléments non nuls de cet idéal. D'où le Théorème 3.

Corollaire. *Soient* $f_1, \ldots, f_n$ *des polynômes non nuls à une indéterminée à coefficients dans un corps* K. *Les propriétés suivantes sont équivalentes :*

a) les seuls diviseurs communs à $f_1, \ldots, f_n$ *sont les éléments non nuls de* K;

b) il existe des polynômes $u_i \in K[X]$ *tels que*

$$u_1(X) f_1(X) + \cdots + u_n(X) f_n(X) = 1.$$

Ce Corollaire est d'ailleurs un cas particulier du Théorème 1 du § 31. Lorsqu'il est vérifié, on dit bien entendu que les polynômes f_i sont **premiers entre eux.**

Il va de soi qu'on peut aussi définir un **plus petit commun multiple** de $f_1, \ldots, f_n$; c'est un multiple commun m des f_i possédant la propriété que les autres multiples communs aux f_i sont les multiples de m; ou encore, c'est un multiple commun aux f_i non nul et *de degré minimum.*

Les Théorèmes 2, 3, 4 du § 31 s'appliquent mot pour mot à l'anneau K[X], et il est parfaitement inutile de les énoncer ici à nouveau.

Les éléments extrémaux de l'anneau K[X] s'appellent généralement les **polynômes irréductibles** à une indéterminée à coefficients dans le corps K. *Tout polynôme* $f \in K[X]$ *peut se mettre, et essentiellement d'une seule façon, sous forme d'un produit de polynômes irréductibles* : cela résulte des Théorèmes 5 et 7 du § précédent.

Remarque 1. En multipliant un polynôme irréductible par l'inverse de son coefficient dominant (ce qui le remplace par un polynôme associé, au sens du § 31, n° 6), on voit que pour effectuer la décomposition d'un polynôme f en facteurs irréductibles on peut se borner à choisir des polynômes irréductibles dont le coefficient dominant est égal à 1; en vertu du Lemme 1, deux tels polynômes ne peuvent être associés que s'ils sont égaux; autrement dit, si l'on veut appliquer à K[X] la décomposition (9) du § 31, n° 6, il s'impose de prendre pour P l'ensemble des polynômes irréductibles dont le coefficient dominant est égal à 1.

Remarque 2. Soient p_1 et p_2 des polynômes à une indéterminée à coefficients dans le corps K. Pour calculer leur p.g.c.d., on peut procéder par divisions successives : supposant $d^0(p_1) > d^0(p_2)$, on écrit

$$p_1 = p_2 v_2 + p_3 \quad \text{avec } d^0(p_3) < d^0(p_2)$$
$$p_2 = p_3 v_3 + p_4 \quad \text{avec } d^0(p_4) < d^0(p_3)$$

et ainsi de suite; comme les degrés des p_i vont en diminuant, on aboutit finalement à une relation de la forme

$$p_{n-1} = p_n v_n.$$

Il est clair que le pgcd de p_1 et p_2 est aussi celui de p_2 et p_3, donc aussi celui de p_3 et p_4, etc..., donc aussi celui de p_{n-1} et p_n — autrement dit c'est p_n, i.e. le dernier reste *non nul* dans les divisions successives.

Cette méthode permet aussi d'expliciter le théorème de Bezout, i.e. de construire deux polynômes u_1 et u_2 tels que $u_1 p_1 + u_2 p_2 = p_n$; on a en effet

$$\begin{aligned} p_n &= p_{n-2} - p_{n-1}v_{n-1} = (p_{n-4} - p_{n-3}v_{n-3}) - (p_{n-3} - p_{n-2}v_{n-2})v_{n-1} \\ &= p_{n-4} - p_{n-3}(v_{n-3} + v_{n-1}) + p_{n-2}v_{n-2}v_{n-1} \\ &= p_{n-6} - p_{n-5}v_{n-5} - (p_{n-5} - p_{n-4}v_{n-4})\,(v_{n-3} + v_{n-1}) \\ &\qquad + (p_{n-4} - p_{n-3}v_{n-3})v_{n-3}v_{n-1} \end{aligned}$$

et en « remontant » ainsi les calculs on parvient évidemment à une relation de la forme cherchée. Voir *Exercice* 3.

4. *Application aux fractions rationnelles*

Soient K un corps commutatif et

$$f(X) = \frac{p(X)}{q(X)}$$

une fraction rationnelle à une indéterminée à coefficients dans K. Écrivons

$$q(X) = q_1(X)^{r_1} \dots q_n(X)^{r_n}$$

où les q_i sont des polynômes irréductibles deux à deux non proportionnels. Le Théorème 8 du § 31 montre alors qu'il existe des polynômes $p_i(X)$ $(1 \leqslant i \leqslant n)$ tels que l'on ait

$$f(X) = \frac{p_1(X)}{q_1(X)^{r_1}} + \dots + \frac{p_n(X)}{q_n(X)^{r_n}}. \tag{11}$$

Or on a le lemme suivant (*) :

LEMME 3. *Soient p et q des polynômes à une indéterminée à coefficients dans un anneau commutatif* K. *Supposons q unitaire; il existe alors des polynômes h_i $(i \geqslant 0)$ presque tous nuls tels que l'on ait*

$$p = h_0 + h_1 q + h_2 q^2 + \dots, \qquad d^0(h_i) < d^0(q) \text{ pour tout } i.$$

(*) On notera l'analogie existant entre ce Lemme et les raisonnements classiques conduisant, pour tout entier $q \neq 0$, à la « numération de base q ».

On écrit pour cela, à l'aide du Théorème 1, que

$$p = h_0 + p_1 q \quad \text{avec } d^0(h_0) < d^0(q),$$

puis que

$$p_1 = h_1 + p_2 q \quad \text{avec } d^0(h_1) < d^0(q),$$

et ainsi de suite; les degrés des polynômes p_i décroissent strictement, de sorte que l'on a $p_i = 0$ pour i suffisamment grand, et en portant chacune des relations obtenues dans la précédente on obtient évidemment le Lemme.

Le Lemme 3 montre que, pour tout entier $r > 0$, on peut écrire

$$\frac{p}{q^r} = \frac{h_0}{q^r} + \frac{h_1}{q^{r-1}} + \cdots + \frac{h_r}{q} + p_0$$

où $h_0, \ldots, h_r$ et p_0 sont des polynômes, avec

$$d^0(h_i) < d^0(q) \quad \text{pour } 0 \leqslant i \leqslant r.$$

Appliquant ce résultat à chacune des fractions qui figurent dans le second membre de (11), et en groupant ensemble les termes p_0 obtenus pour chacune de ces fractions, on obtient donc une relation de la forme

$$f(\mathrm{X}) = g(\mathrm{X}) + \sum_{i=1}^{i=n} \sum_{0 \leqslant r \leqslant r_i} \frac{h_{ir}(\mathrm{X})}{q_i(\mathrm{X})^r} \tag{12}$$

où g et les h_{ir} sont des polynômes, avec

$$d^0(h_{ir}) < d^0(q_i) \tag{13}$$

quels que soient i et r.

La formule (12) s'appelle **la décomposition en éléments simples sur le corps K** de la fraction rationnelle donnée f. On verra plus loin que, si $\mathrm{K} = \mathbf{C}$, les h_{ir} sont nécessairement des constantes, et que si $\mathrm{K} = \mathbf{R}$ ce sont des polynômes de degré 1 au plus.

Exemple 2. Prenons

$$f(\mathrm{X}) = \frac{1}{\mathrm{X}^2(\mathrm{X}-1)^3};$$

les polynômes X et X − 1, étant de degré 1, sont irréductibles. On a

$$(\mathrm{X}-1)^3 = \mathrm{X}^3 - 3\mathrm{X}^2 + 3\mathrm{X} - 1 = \mathrm{X}^2(\mathrm{X}-3) + 3\mathrm{X} - 1$$

puis

$$\mathrm{X}^2 = (3\mathrm{X}-1)\frac{\mathrm{X}}{3} + \frac{\mathrm{X}}{3},$$

puis

$$3\mathrm{X} - 1 = \frac{\mathrm{X}}{3} \cdot 9 - 1;$$

il vient donc

$$1 = \frac{X}{3} \cdot 9 - (3X - 1) = 9X^2 - (3X + 1)(3X - 1)$$
$$= 9X^2 - (3X + 1)\,((X - 1)^3 - X^2(X - 3))$$

de sorte que le théorème de Bezout pour X^3 et $(X - 1)^3$ s'écrit

$$1 = (3X^2 - 8X + 6)X^2 - (3X + 1)(X - 1)^3.$$

On a donc

$$f(X) = \frac{(3X^2 - 8X + 6)X^2 - (3X + 1)(X - 1)^3}{X^2(X - 1)^3}$$
$$= \frac{3X^2 - 8X + 6}{(X - 1)^3} - \frac{3X + 1}{X^2};$$

or

$$3X^2 - 8X + 6 = (X - 1)(3X - 5) + 1 = (X - 1)[3(X - 1) - 2] + 1$$
$$= 3(X - 1)^2 - 2(X - 1) + 1;$$

il vient donc

$$f(X) = \frac{3}{X - 1} - \frac{2}{(X - 1)^2} + \frac{1}{(X - 1)^3} - \frac{1}{X^2} - \frac{3}{X},$$

ce qui est la décomposition cherchée de f en éléments simples.

§ 33. Racines d'une équation algébrique

1. *Nombre maximum de racines*

Nous avons déjà démontré (§ 28, Lemme 2) qu'un polynôme de degré n à une indéterminée, à coefficients dans un anneau d'intégrité commutatif K, possède au plus n racines dans K. On peut, comme le montrera le Corollaire du Théorème ci-dessous, améliorer ce résultat.

THÉORÈME 1. *Soit f un polynôme de degré n à une indéterminée à coefficients dans un corps commutatif* K. *Soient $a_1, \ldots, a_p$ les racines de f dans* K, *et $r_1, \ldots, r_p$ leurs ordres de multiplicité. On a alors*

$$f(X) = (X - a_1)^{r_1} \ldots (X - a_p)^{r_p} g(X) \tag{1}$$

où g est un polynôme qui n'a aucune racine dans K.

Par définition de l'ordre de multiplicité d'une racine, f est divisible par chacun des polynômes

$$(X - a_1)^{r_1}, \ldots, (X - a_p)^{r_p}. \tag{2}$$

D'autre part, *tout polynôme de la forme $X - a$ est irréductible* car si

$$X - a = p(X)q(X)$$

on a $1 = d^o(p) + d^o(q)$, de sorte que l'un des polynômes p, q est de degré 0, i.e. est un élément inversible de l'anneau K[X]. Enfin, il est clair que les polynômes $X - a$ et $X - b$ sont non associés (i.e. non proportionnels) pour $a \neq b$.

Il résulte de là que les polynômes (2) sont deux à deux premiers entre eux (§ 31, Lemme 6), et par suite (§ 31, Théorème 4) que leur produit divise f; d'où l'existence d'une relation (1).

La relation (1) montre de plus que toute racine a de g dans K, étant racine de f dans K, est l'un des éléments $a_1, \ldots, a_p$; mais si a_1 par exemple était racine de g, le polynôme $g(X)$ serait divisible par $X - a_1$, et $f(X)$ serait par suite divisible par $(X - a_1)^{r_1+1}$, contrairement à la définition d'un ordre de multiplicité. Donc g n'a aucune racine dans K, ce qui achève la démonstration.

COROLLAIRE. *Avec les notations du Théorème* 1, *on a*

$$r_1 + \cdots + r_p \leqslant n = d^\circ(f).$$

C'est évident car

$$d^\circ(f) = r_1 + \cdots + r_p + d^\circ(g). \tag{3}$$

Le Corollaire exprime que *le nombre des racines de f dans* K *est au plus n même si l'on convient de compter une racine multiple d'ordre r comme l'équivalent de r racines simples.*

Supposons par exemple f de degré 3; alors les seules possibilités en ce qui concerne le nombre de racines de f dans K sont les suivantes : *a*) f n'a que des racines simples, il y en a alors au plus trois; *b*) f possède des racines doubles; alors f admet soit une seule racine dans K, laquelle est racine double, soit une racine double et une racine simple; *c*) f admet une racine triple; celle-ci est alors unique, et f n'admet aucune autre racine. En fait, ces possibilités ne se présentent pas toutes, en vertu du fait que si f admet deux racines simples distinctes, ou bien une racine double, alors f admet nécessairement une autre racine dans K; en effet, si l'on a

$$f(X) = (X - a)\,(X - b)\,g(X),$$

le polynôme g est nécessairement de degré 1, donc proportionnel à $X - c$ pour un $c \in K$ convenable, et c est racine de f.

Autrement dit, pour un polynôme du troisième degré, les éventualités suivantes peuvent se produire :

aucune racine
une racine simple
trois racines simples
une racine simple et une racine double
une racine triple.

Il est facile de donner des exemples de chacun de ces cas; pour le premier on prend (*)

$$K = \mathbf{Q} \qquad \text{et} \qquad f(X) = X^3 - 2;$$

pour les autres, il suffit de prendre $K = \mathbf{R}$ et les polynômes

$$(X - 1)\,(X^2 + 1), \quad (X - 1)\,(X - 2)\,(X - 3), \quad (X - 1)\,(X - 2)^2, \quad (X - 1)^3.$$

(*) Il n'est pas possible, pour donner un exemple du premier cas, de prendre $K = \mathbf{R}$, car on démontre en Analyse que tout polynôme de degré *impair* à coefficients réels possède au moins une racine réelle (rappelons la démonstration brièvement : on montre d'abord que, pour les valeurs très grandes de la variable, un polynôme est un « infiniment grand » équivalent à son terme de plus haut degré, donc du même signe que celui-ci; on en déduit qu'un polynôme de degré impair ne peut pas, sur $\mathbf{R}$, garder un signe constant; le théorème des valeurs intermédiaires pour les fonctions continues montre alors qu'un tel polynôme s'annule nécessairement quelque part).

Les résultats de ce genre, bien qu'ils concernent des polynômes, appartiennent en réalité à l'Analyse. C'est pourquoi nous ne les exposons pas dans ce volume.

Revenons au cas général, le corps K et le polynôme f étant quelconques. Il peut arriver que, dans la décomposition du Théorème 1, le polynôme g soit constant, i.e. de degré 0, autrement dit qu'on ait

$$f(X) = c(X - a_1)^{r_1} \dots (X - a_p)^{r_p}$$

où c est nécessairement le coefficient dominant de f; ou, ce qui revient au même, que f puisse s'écrire comme produit de polynômes du premier degré à coefficients dans K. Lorsqu'il en est ainsi on dit que f a **toutes ses racines dans** K; la raison de cette terminologie est que, si L est un surcorps de K, alors les racines de f dans L, i.e. les $x \in$ L tels que

$$c(x - a_1)^{r_1} \dots (x - a_p)^{r_p} = 0,$$

sont $a_1, \dots, a_p$, autrement dit les racines de f dans K. On verra plus loin par contre que si f n'a pas toutes ses racines dans K, il existe un surcorps L de K (par exemple un corps algébriquement clos contenant K — voir le n° 2) et des racines de f dans L qui ne sont pas des éléments de K.

Pour que f ait toutes ses racines dans K, il est évidemment nécessaire et suffisant que les ordres de multiplicité $r_1, \dots, r_p$ des racines de f dans K vérifient la relation

$$r_1 + \dots + r_p = d^0(f),$$

autrement dit que le nombre de racines de f dans K (chaque racine multiple d'ordre r étant considérée comme l'équivalent de r racines simples) soit *égal* au degré de f.

Notons enfin que *si f a toutes ses racines dans* K *il en est de même de tout diviseur g de f;* en effet, g est produit de polynômes irréductibles, lesquels, divisant g, divisent aussi f; mais comme f est produit de polynômes du premier degré, qui sont évidemment irréductibles, les seuls facteurs irréductibles de f sont ces facteurs du premier degré, et par suite les diviseurs irréductibles de g sont eux aussi de degré 1; par conséquent, g est produit de facteurs du premier degré, ce qui établit notre assertion. (On pourrait aussi appliquer à f et g le Lemme 4 du n° 7 du § 31, lequel montre comment trouver tous les diviseurs de f à partir de la décomposition de f en facteurs irréductibles.)

2. *Corps algébriquement clos*

Soient K un *corps commutatif* et f un polynôme à une indéterminée, à coefficients dans K, et de degré $n \geq 1$. D'après le n° précédent, f possède *au plus* n racines dans K; mais nous n'avons encore énoncé aucun théorème affirmant l'*existence* de racines de f dans K (et pour cause, le polynôme $X^2 + 1$ n'ayant pas de racines dans le corps **R**).

On dit qu'un corps commutatif K est **algébriquement clos** si *tout* polynôme à une indéterminée, à coefficients dans K, et non constant, possède au moins une racine dans K. En ce qui concerne l'existence de tels corps, les deux résultats fondamentaux sont les suivants :

THÉORÈME 2 (d'Alembert-Gauss). *Le corps des nombres complexes est algébriquement clos.*

THÉORÈME 3 (Steinitz). *Tout corps commutatif peut être plongé dans un corps algébriquement clos.*

Le Théorème 2 signifie que **toute équation algébrique**

$$a_0 + a_1 x + \cdots + a_n x^n = 0$$

à coefficients complexes, de degré $n \geqslant 1$, possède au moins une racine complexe — résultat d'autant plus extraordinaire que les nombres complexes ont été inventés pour attribuer des racines aux équations du *second degré* seulement. Quant au théorème de Steinitz, il signifie que, pour tout corps commutatif K, on peut construire un corps algébriquement clos L qui contient un sous-corps isomorphe à K; si $K = \mathbf{R}$ on peut par exemple prendre $L = \mathbf{C}$ (dans le cas général, la construction de L à partir de K est beaucoup plus compliquée que celle de C à partir de R, mais peut néanmoins s'effectuer à l'aide de méthodes analogues).

On ne peut démontrer le Théorème 2 sans recourir, d'une façon ou d'une autre, à des considérations qui relèvent de l'Analyse, et non de l'Algèbre. Voir une démonstration simple dans l'*Exercice* 25 de ce §. Quant au Théorème 3, sa démonstration dépasserait de beaucoup le cadre du présent ouvrage; cf. *Exercice* 20.

On peut se demander si l'existence de racines pour toutes les équations algébriques à *une* inconnue entraîne une propriété analogue pour les systèmes d'équations algébriques à *plusieurs* inconnues; la réponse à cette question est l'un des plus célèbres théorèmes que l'on doive à Hilbert :

THÉORÈME 4 (Nullstellensatz de Hilbert). *Soient* K *un corps commutatif algébriquement clos et* I *un idéal de l'anneau* $K[X_1, \ldots, X_n]$. *Les propriétés suivantes sont équivalentes :*

a) il existe au moins un point $x \in K^n$ *tel que l'on ait*

$$f(x) = 0 \quad \textit{pour tout} \quad f \in I; \tag{4}$$

b) on a $I \neq K[X_1, \ldots, X_n]$.

Il est trivial de démontrer que *a*) implique *b*), attendu que le polynôme 1 ne saurait s'annuler en un point de K^n; mais la démonstration du fait que *b*) implique *a*) est trop compliquée pour être exposée ici; voir l'*Exercice* 33 de ce §.

Pour comprendre l'énoncé du Théorème 4, donnons-nous des polynômes

$$f_1, \ldots, f_p \in K[X_1, \ldots, X_n]$$

et cherchons à résoudre, dans K^n, le **système d'équations algébriques**

$$f_1(x) = \cdots = f_p(x) = 0; \tag{5}$$

soit I l'idéal de $K[X_1, \ldots, X_n]$ engendré par $f_1, \ldots, f_p$, i.e. l'ensemble des polynômes de la forme

$$f = u_1 f_1 + \cdots + u_p f_p;$$

il est *évident* que les solutions de (5) sont les mêmes que les solutions de (4); donc, dire que (5) possède au moins une solution signifie que I n'est pas l'anneau des

polynômes tout entier, ou, ce qui revient au même, que

$$1 \notin \mathrm{I}.$$

En d'autres termes :

COROLLAIRE DU THÉORÈME 4. *Étant donnés un corps commutatif algébriquement clos* K *et des polynômes*

$$f_1, \ldots, f_p \in \mathrm{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n],$$

les propriétés suivantes sont équivalentes :

a) le système d'équations algébriques

$$f_1(x) = \cdots = f_p(x) = 0$$

n'a aucune solution $x \in \mathrm{K}^n$;

b) il existe des polynômes

$$u_1, \ldots, u_p \in \mathrm{K}[\mathrm{X}_1, \ldots, \mathrm{X}_n]$$

tels que l'on ait

$$u_1 f_1 + \cdots + u_p f_p = 1.$$

Remarque 1. Soit K un corps algébriquement clos; une partie de K^n est appelée une **variété algébrique affine** si c'est l'ensemble des solutions d'un système d'équations algébriques $f_1(x) = \ldots = f_p(x) = 0$. L'étude de ces variétés algébriques affines (et d'objets analogues) est le but de la Géométrie Algébrique. Une variété algébrique définie par une seule équation $f(x) = 0$, où f est un polynôme non nul, s'appelle une **hypersurface** (ou une **surface** lorsque $n = 3$, ou une **courbe plane** lorsque $n = 2$). Le Corollaire du Théorème 4 est donc une condition nécessaire et suffisante pour qu'une intersection d'hypersurfaces algébriques soit vide.

3. *Nombre de racines d'une équation à coefficients dans un corps algébriquement clos*

Sur un corps algébriquement clos, on peut améliorer grandement le Théorème 1 :

THÉORÈME 5. *Soit f un polynôme à une indéterminée, à coefficients dans un corps algébriquement clos* K, *et de degré $n \geqslant 1$. Soient $a_1, \ldots, a_p$ les diverses racines de f dans* K, *et $r_1, \ldots, r_p$ leurs ordres de multiplicité. On a alors*

$$f(\mathrm{X}) = c(\mathrm{X} - a_1)^{r_1}(\mathrm{X} - a_2)^{r_2} \ldots (\mathrm{X} - a_p)^{r_p} \tag{6}$$

où c est le coefficient dominant de f. En outre on a

$$r_1 + \cdots + r_p = n. \tag{7}$$

Il suffit d'appliquer le Théorème 1; comme le polynôme g ne s'annule jamais dans K, il est constant par définition d'un corps algébriquement clos, d'où la pre-

mière assertion de l'énoncé. La seconde s'obtient en calculant le degré du second membre de la relation (6).

La relation (7) s'exprime généralement de la façon suivante : *dans un corps algébriquement clos, une équation algébrique de degré $n \geqslant 1$ possède exactement n racines* pourvu que l'on considère une racine multiple d'ordre r comme l'équivalent de r racines distinctes.

Exemple 1. Considérons l'équation

$$x^n = 1 \tag{8}$$

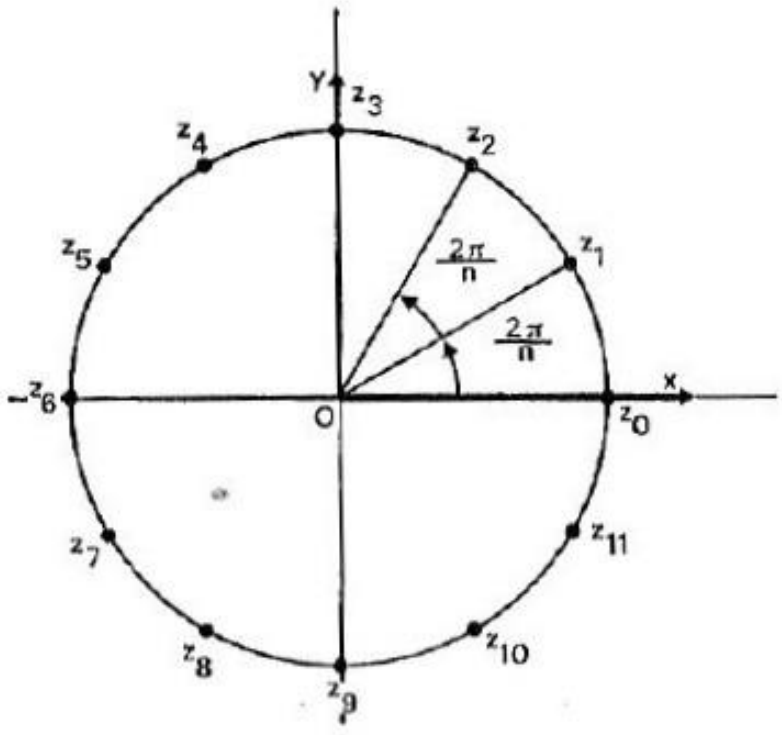
dans un corps commutatif K (ses racines sont les **racines n^e de l'unité dans K**); posant

$$f(X) = X^n - 1, \qquad \text{d'où} \qquad f'(X) = nX^{n-1},$$

on voit que, *si n n'est pas multiple de la caractéristique p de* K, la seule racine de f' est 0, qui n'est évidemment pas racine n^e de l'unité; donc (§ 30, Théorème 6) *les racines de f sont toutes simples* dans cette hypothèse, et si K est de plus algébriquement clos il s'ensuit qu'il existe dans K exactement n racines *distinctes* de l'équation considérée; c'est le cas dans C.

Lorsque $K = C$, les racines n^e de l'unité sont données par la formule

$$z_k = \cos(2k\pi/n) + i.\sin(2k\pi/n) \qquad (0 \leqslant k \leqslant n-1),$$



autrement dit sont représentées, dans le plan complexe, par les sommets d'un polygone régulier de n côtés inscrit dans le cercle unité (cf. la figure ci-contre). En effet, la formule de Moivre (§ 9, n° 6) montre que

$$z_k^n = \cos(2k\pi n/n) + i.\sin(2k\pi n/n) = 1,$$

de sorte que la formule ci-dessus représente n racines, deux à deux distinctes, de l'équation (8); celle-ci étant de degré n ne saurait en posséder d'autres.

Supposant toujours $K = C$, considérons plus généralement l'équation

$$z^n = a$$

où a est un nombre complexe non nul donné (**ses racines s'appellent les racines n^e de a**). Écrivant

$$z = \rho(\cos\theta + i.\sin\theta), \qquad a = r(\cos\varphi + i.\sin\varphi)$$

avec r et ρ réels positifs, θ et φ réels, tout revient à écrire que

$$\rho^n[\cos(n\theta) + i.\sin(n\theta)] = r(\cos\varphi + i.\sin\varphi);$$

le premier membre a pour valeur absolue ρ^n et pour argument $n\theta$ (à un multiple près de 2π : un argument est un nombre réel défini modulo 2π), le second pour valeur absolue r et pour argument φ; l'équation $x^n = a$ équivaut donc aux conditions

$$\rho^n = r, \qquad n\theta \equiv \varphi \pmod{2\pi};$$

il s'ensuit qu'elle a pour racines les nombres complexes

$$z_k = \sqrt[n]{r}\left[\cos\frac{\varphi + 2k\pi}{n} + i.\sin\frac{\varphi + 2k\pi}{n}\right]$$

comme on a évidemment $z_k = z_{k+n}$, il suffit du reste de donner n valeurs consécutives à l'entier k pour obtenir toutes les racines (en nombre n et deux à deux distinctes) de l'équation considérée.

Remarque 2. Le Théorème 5 peut évidemment être en défaut si le corps K n'est pas algébriquement clos; mais dans ce cas il redevient vrai si l'on considère les racines de f dans un corps *algébriquement clos* contenant K, corps dont l'existence résulte du Théorème de Steinitz. Par exemple, soit f un polynôme de degré n à coefficients réels; il est en général faux que f possède n racines réelles; mais il est toujours vrai que f possède n racines *réelles ou complexes* (le nombre de racines de f étant bien entendu calculé en tenant compte des ordres de multiplicité). Par exemple, le polynôme $X^2 + 1$ possède deux racines réelles ou complexes, à savoir i et $-i$. De même, l'équation

$$x^n = 1,$$

qui ne possède dans **R** qu'une racine (si n est impair) ou deux racines (si n est pair), en possède n dans **C**.

4. *Polynômes irréductibles à coefficients dans un corps algébriquement clos*

Soit $f(X)$ un polynôme à une indéterminée à coefficients dans un corps algébriquement clos K; alors *pour que f soit irréductible il faut et il suffit que f soit de degré* 1, i.e. de la forme

$$f(X) = aX + b, \qquad a \neq 0.$$

La condition est évidemment suffisante (même si K n'est pas algébriquement clos) puisqu'alors les seuls diviseurs de f sont de degré 0 (i.e. constants) ou 1 (i.e. proportionnels à f). Inversement, supposons f irréductible; f n'est pas inversible dans l'anneau K[X], donc est de degré $n \geqslant 1$ (§ 32, Lemme 2); comme K est algébriquement clos, f possède au moins une racine $a \in K$; mais alors f est multiple de $X - a$, donc proportionnel à $X - a$ puisqu'irréductible, et notre assertion est démontrée.

Lorsqu'on effectue, en utilisant la formule (9) du § 31, fin du n° 6, la décom-

position d'un polynôme $f \in K[X]$ en facteurs irréductibles, on peut donc prendre pour P l'ensemble des polynômes de la forme

$$X - a, \qquad a \in K;$$

la formule (9) du § 31 se réduit alors évidemment à la formule (6) du Théorème 5.

Celle-ci constituant la décomposition de f en facteurs irréductibles dans l'anneau principal K[X], on peut l'utiliser par exemple pour former le pgcd de deux polynômes f et g; désignons les racines distinctes de f par

$$a_1, \ldots, a_m, b_1, \ldots, b_n$$

et celles de g par

$$a_1, \ldots, a_m, c_1, \ldots, c_p,$$

en mettant en évidence les racines $a_1, \ldots, a_m$ communes à f et g, s'il y en a; écrivant

$$f(X) = u(X - a_1)^{r_1'} \ldots (X - a_m)^{r_m'}(X - b_1)^{s_1} \ldots (X - b_n)^{s_n}$$
$$g(X) = v(X - a_1)^{r_1''} \ldots (X - a_m)^{r_m''}(X - c_1)^{t_1} \ldots (X - c_p)^{t_p},$$

où u et v sont des constantes, il est clair que d'après le § 31, n° 7, un pgcd de f et g sera donné par la formule

$$d(X) = (X - a_1)^{r_1} \ldots (X - a_m)^{r_m} \qquad \text{où} \qquad r_i = \operatorname{Min}(r_i', r_i'').$$

Autrement dit, *les racines du* pgcd *sont les racines communes à f et g, et si une racine commune est d'ordre r' pour f et r'' pour g, elle est d'ordre* $\operatorname{Min}(r', r'')$ *pour le pgcd de f et g.*

En particulier, *pour que f et g soient premiers entre eux il faut et il suffit qu'ils n'admettent aucune racine commune*, résultat qui n'est autre que le Corollaire du Théorème 4 dans le cas particulier où $n = 1$.

Le Théorème 5 permet également d'améliorer, dans le cas d'un corps algébriquement clos, la décomposition d'une fraction rationnelle en éléments simples du § 32, n° 4. En effet, dans ce cas les polynômes irréductibles q_i du cas général sont donnés par

$$q_i(X) = X - a_i$$

et sont de degré 1; les polynômes h_{ir} figurant dans la formule (12) du § 32 sont donc de degré 0 au plus, autrement dit ce sont des *constantes*, et on voit en définitive que, sur un corps algébriquement clos, toute fraction rationnelle peut se mettre sous la forme

$$f(X) = g(X) + \sum_{i=1}^{i=n} \sum_{1 \leqslant r \leqslant r_i} \frac{c_{ir}}{(X - a_i)^r} \tag{9}$$

où les c_{ir} sont des éléments de K, les a_i les diverses racines du dénominateur de f, les r_i les multiplicités de ces racines, et g un polynôme (qui est du reste, comme on le vérifie facilement, le quotient du numérateur de f par son dénominateur).

5. *Polynômes irréductibles à coefficients réels*

Les résultats du n° 4 s'appliquent au corps **C** des nombres complexes, mais non au corps **R** des nombres réels puisque celui-ci n'est pas algébriquement clos. On peut toutefois, dans ce cas, obtenir encore des résultats complets :

THÉORÈME 6. *Les éléments irréductibles de l'anneau* $\mathbf{R}[X]$ *des polynômes à une indéterminée à coefficients réels sont d'une part les polynômes*

$$aX + b \quad avec \quad a \neq 0,$$

d'autre part les polynômes

$$aX^2 + bX + c \quad avec \quad b^2 - 4ac < 0.$$

Il est clair que, pour tout corps commutatif K, les polynômes de degré 1 sont des éléments irréductibles de K[X]. Il en est de même des polynômes de degré 2 *qui n'ont aucune racine dans* K, car un diviseur non trivial d'un tel polynôme f est nécessairement de degré 1, donc de la forme $aX + b$, et si f est divisible par $aX + b$ alors $-b/a$ est une racine de f dans K.

Il reste à faire voir que, pour $K = \mathbf{R}$, il n'existe pas d'autres polynômes irréductibles que ceux qu'on vient d'énumérer.

Soit f un élément irréductible (donc de degré 1 au moins) de $\mathbf{R}[X]$. Si f admet dans **R** une racine a, alors f est divisible par $X - a$, donc proportionnel à $X - a$ et par suite f est de degré 1, la réciproque étant claire.

Supposons maintenant que f ne possède aucune racine dans **R**. Comme **C** est algébriquement clos, le polynôme

$$f(X) = a_0 + a_1X + \cdots + a_nX^n$$

considéré admet au moins une racine complexe

$$w = u + iv \quad (u, \ v \in \mathbf{R});$$

mais il admet alors aussi pour racine le nombre

$$\overline{w} = u - iv$$

conjugué de w, car, les coefficients a_i étant réels, on a

$$f(w) = a_0 + a_1w + \cdots + a_n\overline{w}^n = \overline{a_0} + \overline{a_1w} + \cdots + \overline{a_nw^n} = \overline{f(\overline{w})},$$

d'où notre assertion. Par conséquent, dans l'anneau $\mathbf{C}[X]$, le polynôme f est divisible par $X - w$ et $X - \overline{w}$; comme ces polynômes sont premiers entre eux puisque $w \neq \overline{w}$, on voit que dans $\mathbf{C}[X]$ il est même divisible par

$$(X - w)(X - \overline{w}) = (X - u - iv)(X - u + iv) = (X - u)^2 + v^2,$$

polynôme du second degré à coefficients *réels* et sans racine réelle. Pour en déduire

que le polynôme irréductible f est du second degré, il suffit de montrer que f est divisible par $(X - u)^2 + v^2$ non seulement dans $\mathbf{C}[X]$, mais aussi dans $\mathbf{R}[X]$.

Or considérons d'une manière générale un corps K, un surcorps L de K, et deux polynômes f et g à coefficients dans K. Soient q et r le quotient et le reste de la division de f par g dans K[X]; ce sont des polynômes à coefficients dans K vérifiant

$$f = gq + r, \qquad d^\circ(r) < d^\circ(g);$$

évidemment ces relations subsistent si l'on regarde f, g, q, r comme des polynômes à coefficients dans L; par suite, *le quotient et le reste de la division de f par g dans l'anneau* K[X] *sont les mêmes que dans l'anneau* L[X].

Si en particulier g divise f dans L[X], on voit que g divise aussi f dans K[X], et ceci achève la démonstration du Théorème.

Il résulte du Théorème 6 et du § 31, Théorème 5, que tout polynôme f à coefficients *réels* peut s'écrire sous la forme

$$f(X) = u.(X - a_1)^{r_1} \dots (X - a_p)^{r_p} (X^2 + b_1X + c_1)^{s_1} \dots (X^2 + b_qX + c_q)^{s_q}$$

où u est le coefficient dominant de f, où les a_i sont les diverses racines de f dans $\mathbf{R}$, et les r_i leurs ordres de multiplicités, et où les polynômes $X^2 + b_jX + c_j$ sont sans racine réelle, i.e. vérifient $b_j^2 - 4c_j < 0$.

On peut obtenir encore comme suit cette décomposition. Étant donné que la conjuguée d'une racine complexe de f est encore une racine de f comme il résulte de la relation

$$f(\bar{w}) = \overline{f(w)}$$

établie plus haut, le polynôme f possède un nombre *pair* de racines non réelles, soit $2q$; désignant par

$$w_j = u_j + iv_j \quad (1 \leqslant j \leqslant q)$$

celles dont la partie imaginaire est positive, on voit que les autres sont les nombres

$$\bar{w}_j = u_j - iv_j$$

conjugués des précédents. Si l'on se place dans l'anneau $\mathbf{C}[X]$, on a donc une formule $f(X) = u.(X - a_1)^{r_1} \dots (X - a_p)^{r_p} (X - w_1)^{s_1} \dots (X - w_q)^{s_q} (X - \bar{w}_1)^{s'_1} \dots (X - \bar{w}_q)^{s'_q}$; en fait, les ordres de multiplicité s_j et s'_j de deux racines imaginaires conjuguées sont les mêmes comme il résulte du § 30, Théorème 7, et en groupant les termes correspondants dans le décomposition on trouve un facteur

$$[(X - w_j)(X - \bar{w}_j)]^{s_j} = [(X - u_j)^2 + v_j^2]^{s_j},$$

d'où la décomposition en facteurs irréductibles dans l'anneau $\mathbf{R}[X]$.

Comme au n° précédent, ces résultats permettent d'améliorer la décomposition d'une fraction rationnelle en éléments simples donnée au § 32, n° 4. En effet, les polynômes q_i du cas général sont ici de deux espèces : ceux de la forme

$$q_i(X) = X - a_i$$

qui correspondent aux racines réelles du dénominateur (les numérateurs h_{ir} des éléments simples sont alors de degré < 1, autrement dit sont des *constantes*); et ceux de la forme

$$q_j(X) = X^2 + b_jX + c_j \quad \text{avec} \quad b_j^2 - 4c_j < 0,$$

qui correspondent aux couples de racines imaginaires conjuguées (les polynômes h_{jr} sont alors de degré < 2, i.e. de la forme $u_{jr}X + v_{jr}$ avec des constantes u_{jr}, v_{jr}). Par suite, la formule (12) du § 32 s'écrit, dans le cas du corps des nombres réels, sous la forme

$$f(X) = g(X) + \sum_{i=1}^{i=m} \sum_{0 \leqslant r \leqslant r_i} \frac{c_{ir}}{(X - a_i)^r} + \sum_{j=1}^{j=s} \sum_{0 \leqslant r \leqslant s_j} \frac{u_{jr}X + v_{jr}}{(X^2 + b_jX + c_j)^r};$$

cette formule joue un grand rôle dans le calcul des primitives des fractions rationnelles, comme le lecteur s'en persuadera aisément en consultant la liste des questions posées, depuis plus de 150 ans, aux épreuves orales du concours d'entrée à l'École Polytechnique.

6. *Relations entre les coefficients et les racines d'une équation*

Soit

$$f(X) = u_nX^n + u_{n-1}X^{n-1} + \cdots + u_0 \tag{10}$$

un polynôme de degré $n \geqslant 1$ à coefficients dans un corps commutatif K, et supposons que f possède n racines (compte tenu des ordres de multiplicité de ces racines), ce qui sera par exemple toujours le cas si K est algébriquement clos. Désignons ces racines par

$$a_1, \ \ldots, \ a_n$$

en écrivant r fois chaque racine multiple d'ordre r. On a alors, d'après le Théorème 1, la relation

$$f(X) = u_n(X - a_1)\,(X - a_2) \ldots (X - a_n). \tag{11}$$

Posons

$$(X - a_1)\,(X - a_2) \ldots (X - a_n) = X^n + v_{n-1}X^{n-1} + \cdots + v_0; \tag{12}$$

en utilisant la formule

$$\prod_{i \in I} (x_i + y_i) = \sum_{V \subset I} x_V y_{I-V}$$

du § 8, n° 5, on trouve évidemment

$$v_{n-k} = (-1)^k \sum a_{i_1}a_{i_2} \ldots a_{i_k},$$

la somme du second membre étant étendue à toutes les *parties* $\{i_1, \ldots, i_k\}$ de l'en-

semble $\{1, \ldots, n\}$; comme

$$u_{n-k} = u_n v_{n-k}$$

en vertu de (11) et (12), on obtient en définitive les relations

$$\sum a_{i_1} a_{i_2} \ldots a_{i_k} = (-1)^k u_{n-k}/u_n ; \tag{13}$$

on les appelle les **relations entre les coefficients et les racines** de l'équation $f(x) = 0$; les premiers membres des relations (13) s'appellent les **fonctions symétriques élémentaires** des racines de l'équation $f(x) = 0$. Voir l'*Exercice* 13 de ce §.

Exemple 2. Si u et v sont les racines (distinctes ou non) d'une équation du second degré

$$ax^2 + bx + c = 0,$$

on a

$$u + v = -b/a, \qquad uv = c/a.$$

Exemple 3. Si u, v et w sont les trois racines (distinctes ou non) d'une équation

$$ax^3 + bx^2 + cx + d = 0,$$

on a

$$u + v + w = -b/a, \qquad vw + wu + uv = c/a, \qquad uvw = -d/a.$$

On observera que, pour $k = 1$, la relation (13) s'écrit

$$a_1 + \cdots + a_n = -u_{n-1}/u_n, \tag{14}$$

et que pour $k = n$ elle s'écrit

$$a_1 \ldots a_n = (-1)^n u_0/u_n. \tag{15}$$

Réduction des matrices

Etant donné un endomorphisme u d'un espace vectoriel E de dimension finie, il est très souvent indispensable de trouver une base de E par rapport à laquelle la matrice de u soit aussi simple que possible — le degré maximum de « simplicité » qu'on peut espérer étant fourni par les matrices diagonales. L'outil principal pour résoudre ce problème est la théorie des vecteurs propres exposée au § 34, qui suffit déjà dans beaucoup de cas. Pour des endomorphismes u tout à fait généraux, on doit utiliser les raisonnements nettement plus difficiles du § 35, que le lecteur débutant pourra négliger, mais qui sont néanmoins aussi utiles que ceux du § 34, dans les applications de la théorie (équations différentielles linéaires à coefficients constants par exemple).

Le § 36 a pour but de fournir des classes de matrices dont on peut affirmer d'avance qu'elles sont réductibles à la forme diagonale. Ici l'outil principal consiste à utiliser sur l'espace E un « produit scalaire » analogue au produit scalaire classique dans l'espace usuel, et permettant de donner un sens à la notion de vecteurs « orthogonaux ». Les considérations du § 36 sont aussi à la base de la classification des « quadriques » (surfaces définies par des équations algébriques du second degré), dont nous n'avons pas parlé dans le texte.

§ 34. Valeurs propres

1. *Définition des vecteurs propres et valeurs propres*

Soient E un espace vectoriel sur un corps commutatif K, et u un endomorphisme de E. On appelle **vecteur propre** de u tout vecteur *non nul* $x \in E$ tel que $u(x)$ soit proportionnel à x; il est clair qu'alors le sous-espace vectoriel D de E engendré par x (i.e. l'ensemble des multiples de x) vérifie $u(D) \subset D$; réciproquement, si une droite D de E, passant par l'origine, vérifie $u(D) \subset D$, tout vecteur non nul porté par D est un vecteur propre de u.

On dit qu'un scalaire $\lambda \in K$ est une **valeur propre** de u s'il existe un vecteur *non nul* $x \in E$ tel que

$$u(x) = \lambda x; \tag{1}$$

x est alors un vecteur propre de u; on dit que x est un vecteur propre **associé** à la valeur propre λ.

On remarquera que la relation (1) signifie que x est annulé par l'endomorphisme

$$u - \lambda . 1$$

de E (où 1 désigne l'endomorphisme unité de E); donc, pour que λ soit valeur propre de u, il faut et il suffit que

$$\mathrm{Ker}(u - \lambda . 1) \neq 0.$$

Supposons E de dimension finie sur K; on peut alors appliquer à $u - \lambda . 1$ le Corollaire 2 du Théorème 8 du § 23, et on obtient donc le résultat suivant :

THÉORÈME 1. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie sur un corps commutatif K. Pour qu'un scalaire $\lambda \in K$ soit une valeur propre de u, il faut et il suffit que l'on ait*

$$\det(u - \lambda . 1) = 0. \tag{2}$$

Ce résultat va nous permettre de montrer que les valeurs propres de u sont les racines d'une équation algébrique à coefficients dans K.

2. *Polynôme caractéristique d'une matrice*

Soient E un espace vectoriel de dimension finie n sur le corps K et u un endomorphisme de E; choisissons une base $(a_i)_{1 \leqslant i \leqslant n}$ de E, et soit

$$U = (\alpha_{ij})_{1 \leqslant i,j \leqslant n}$$

la matrice de u par rapport à cette base (§ 12, n° 3). Comme celle de l'endomorphisme unité de E est la matrice unité

$$1_n = (\delta_{ij})_{1 \leqslant i,j \leqslant n} \quad \text{avec } \delta_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases},$$

on voit que par rapport à la base considérée l'endomorphisme $u - \lambda . 1$ est représenté par la matrice

$$U - \lambda . 1_n = (\alpha_{ij} - \lambda \delta_{ij}) = \begin{pmatrix} \alpha_{11} - \lambda & \alpha_{21} & \dots & \alpha_{n1} \\ \dots\dots\dots\dots & \dots\dots\dots & \dots & \dots\dots\dots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} - \lambda \end{pmatrix}, \tag{3}$$

qu'on obtient en retranchant λ à chaque terme diagonal α_{ii} de U. D'après le Théorème 1, les valeurs propres de u s'obtiennent donc en écrivant la relation

$$\begin{vmatrix} \alpha_{11} - \lambda & \alpha_{21} & \dots & \alpha_{n1} \\ \dots\dots\dots\dots & \dots\dots\dots & \dots & \dots\dots\dots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} - \lambda \end{vmatrix} = 0. \tag{4}$$

Considérons alors K comme plongé dans l'anneau K[X] des polynômes à une indéterminée à coefficients dans K, et formons la matrice $U - X.1_n$, à coefficients dans l'anneau commutatif K[X]; son déterminant

$$p_U(X) = \det(U - X.1_n) = \begin{vmatrix} \alpha_{11} - X & \alpha_{21} & \dots & \alpha_{n1} \\ \dots\dots\dots\dots & \dots\dots\dots & \dots & \dots\dots\dots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} - X \end{vmatrix} \tag{5}$$

est un élément de K[X], i.e. un polynôme à une indéterminée à coefficients dans K; on l'appelle le **polynôme caractéristique de la matrice** U, et la relation (4) montre que *les valeurs propres de u sont les racines dans* K *de l'équation*

$$p_U(\lambda) = 0. \tag{6}$$

Il faut remarquer que le polynôme p_U ne change pas si l'on remplace U par la matrice U′ de u par rapport à une autre base de E; en effet, on a alors (§ 15, Corollaire du Théorème 2)

$$U' = PUP^{-1}$$

pour une matrice $P \in GL(n, K)$; alors, et comme $PI_nP^{-1} = I_n$, on a

$$U' - X.1_n = PUP^{-1} - X.PI_nP^{-1} = P(U - X.1_n)P^{-1}$$

et par suite

$$p_{U'}(X) = \det(P) . p_U(X) . \det(P)^{-1} = p_U(X)$$

en vertu du théorème de multiplication des déterminants; ce qui prouve notre assertion.

Il est donc naturel d'appeler **polynôme caractéristique de l'endomorphisme** u le polynôme $p_U(X)$ où U est la matrice de u par rapport à une base quelconque de E. On notera ce polynôme

$$p_u(X),$$

et on a évidemment

$$p_u(\lambda) = \det(u - \lambda . 1) \quad \text{pour tout } \lambda \in K \tag{7}$$

(si K est un corps infini, cette relation suffit à caractériser p_u d'après le Théorème 1 du § 28). Le Théorème 1 s'énonce alors en disant que *les valeurs propres de u sont les racines de son polynôme caractéristique.*

D'autre part, les considérations précédentes rendent naturelles la définition suivante : on appelle **valeur propre d'une matrice carrée** $U = (\alpha_{ij})_{1 \leqslant i,j \leqslant n}$ à coefficients dans K tout élément de K (ou, plus généralement, d'un sur-corps commutatif de K) qui vérifie l'équation (6).

3. *Forme du polynôme caractéristique*

Conservant les notations ci-dessus, nous allons chercher à obtenir quelques renseignements sur la forme du polynôme

$$p_U(X) = \begin{vmatrix} \alpha_{11} - X & \alpha_{21} & \cdots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} - X & \cdots & \alpha_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} - X \end{vmatrix}.$$

Lorsqu'on développe ce déterminant, on trouve une somme de $n!$ termes; le « terme principal » est le produit

$$(\alpha_{11} - X)(\alpha_{22} - X) \ldots (\alpha_{nn} - X) \tag{8}$$

des termes diagonaux du déterminant considéré; chacun des autres termes est lui aussi un produit de n termes du déterminant considéré, mais $n - 2$ au plus de ces n termes se trouvent sur la diagonale principale; par suite, $p_U(X)$ *est somme de* (8) *et d'un polynôme de degré* $n - 2$ *au plus en* X. Il s'ensuit que les monômes de degré $> n - 2$ de p_U sont les mêmes que ceux du polynôme (8), et par conséquent on a une relation de la forme

$$(-1)^n p_U(X) = X^n - (\alpha_{11} + \alpha_{22} + \cdots + \alpha_{nn}) X^{n-1} + \ldots,$$

les termes non écrits étant de degré $n - 2$ au plus.

On peut donc écrire

$$(-1)^n p_U(X) = X^n - \tau_1(U) X^{n-1} + \tau_2(U) X^{n-2} - \cdots + (-1)^n \tau_n(U); \tag{9}$$

les coefficients $\tau_i(U) \in K$ sont évidemment des fonctions polynômiales des coefficients a_{ij} de la matrice U, à coefficients entiers rationnels; on a vu que

$$\tau_1(U) = a_{11} + a_{22} + \cdots + a_{nn} \tag{9}$$

est la somme des coefficients diagonaux de la matrice U; on appelle ce scalaire la **trace** (*) de la matrice U, et on le désigne le plus souvent par la notation

$$\mathrm{Tr}(U).$$

D'autre part, en faisant $X = 0$ dans (9), et en tenant compte du fait évident *a priori* que $p_U(0) = \det(U)$, on voit que

$$\tau_n(U) = \det(U). \tag{10}$$

Exemple 1. Si $n = 2$ et

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

on a

$$p_U(X) = \begin{vmatrix} a - X & b \\ c & d - X \end{vmatrix} = (a-X)(d-X) - bc = X^2 - \mathrm{Tr}(U)X + \det(U).$$

Si $n = 3$ et si

$$U = \begin{pmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{pmatrix}$$

il vient

$$p_U(X) = \begin{vmatrix} a - X & b & c \\ a' & b' - X & c' \\ a'' & b'' & c'' - X \end{vmatrix} = -X^3 + \mathrm{Tr}(U)X^2 - \tau_2(U)X + \det(U)$$

on laisse au lecteur le soin de calculer $\tau_2(U)$; cf. *Exercice* 40.

4. *Existence de valeurs propres*

Etant donné qu'une équation algébrique à coefficients dans un corps algébriquement clos K possède toujours au moins une racine dans K (par définition même des corps algébriquement clos...), on a évidemment le résultat suivant :

THÉORÈME 2. *Tout endomorphisme d'un espace vectoriel non nul de dimension finie sur un corps* algébriquement clos *possède au moins une valeur propre.*

Il est clair de même que toute matrice carrée à coefficients dans un corps algébriquement clos K possède au moins une valeur propre dans K.

Dans la pratique élémentaire, ces résultats s'appliquent surtout lorsque $K = \mathbf{C}$.

(*) Voir § 12, *Exercice* 8, § 16, *Exercice* 5, § 26, *Exercices* 4 et 5, § 34, *Exercices* 18 et 27.

Remarque 1. Une matrice à coefficients dans un corps K non algébriquement clos peut fort bien n'avoir aucune valeur propre dans K. Prenons par exemple $K = \mathbf{R}$ et la matrice

$$U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

qui, en coordonnées rectangulaires, représente la rotation d'angle θ autour de l'origine. Ses valeurs propres sont les solutions de l'équation

$$\begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{vmatrix} = (\cos\theta - \lambda)^2 + \sin^2\theta = 0;$$

pour que cette équation possède une racine réelle il faut et il suffit que θ soit multiple de π; dans le cas contraire, les racines sont les nombres complexes

$$\cos\theta \pm i.\sin\theta,$$

qui ne sont pas réels.

Soient E un espace vectoriel de dimension finie sur un corps commutatif K, et u un endomorphisme de E. On dit que u a **toutes ses valeurs propres** dans K si le polynôme p_u a toutes ses racines dans K, i.e. (§ 33, n° 1) si l'on peut écrire

$$(-1)^n p_u(X) = (X - \lambda_1)^{r_1} \dots (X - \lambda_q)^{r_q}$$

où les λ_i sont les diverses racines de p_u dans K, et les r_i leurs ordres de multiplicité. De même on dit qu'une matrice carrée U à coefficients dans K a toutes ses valeurs propres dans K si son polynôme caractéristique a toutes ses racines dans K. C'est toujours le cas si le corps de base est algébriquement clos.

Exemple 2. — Prenons $K = \mathbf{R}$ et une matrice U d'ordre 2; on a

$$p_U(X) = X^2 - \mathrm{Tr}(U)X + \mathrm{dét}(U),$$

donc U a toutes ses valeurs propres réelles si et seulement si l'on a

$$\mathrm{Tr}(U)^2 - 4.\mathrm{dét}(U) \geqslant 0.$$

5. *Réduction à la forme triangulaire*

Une matrice carrée U à coefficients dans un anneau est dite **triangulaire si** elle est de la forme

$$\begin{pmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \dots & \alpha_{n1} \\ 0 & \alpha_{22} & \alpha_{32} & \dots & \alpha_{n2} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \alpha_{nn} \end{pmatrix},$$

autrement dit si ceux de ses termes qui sont situés en-dessous de la diagonale sont nuls. D'autre part, un endomorphisme u d'un espace vectoriel E de dimension finie sur un corps K est dit **trigonalisable** s'il existe une base de E par rapport à laquelle

la matrice de u soit triangulaire, autrement dit s'il existe une base $x_1, \ldots, x_n$ de E telle que l'on ait des relations de la forme

$$(10)\qquad \begin{cases} u(x_1) = \alpha_{11}x_1 \\ u(x_2) = \alpha_{21}x_1 + \alpha_{22}x_2 \\ \cdots\cdots\cdots\cdots\cdots \\ u(x_n) = \alpha_{n1}x_1 + \alpha_{n2}x_2 + \cdots + \alpha_{nn}x_n; \end{cases}$$

s'il en est ainsi il est clair que u possède au moins une valeur propre dans K, à savoir α_{11}; de plus on a alors

$$p_u(X) = \begin{vmatrix} \alpha_{11} - X & \alpha_{21} & \cdots & \alpha_{n1} \\ 0 & \alpha_{22} - X & \cdots & \alpha_{n2} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 0 & 0 & \cdots & \alpha_{nn} - X \end{vmatrix}$$

et donc

$$(11)\qquad p_u(X) = (\alpha_{11} - X)(\alpha_{22} - X)\ldots(\alpha_{nn} - X)$$

en vertu du § 24, *Exemple* 2. Comme les α_{ii} sont dans K, on voit donc qu'alors u a toutes ses valeurs propres dans K. Cette condition, nécessaire pour que u soit trigonalisable (et toujours vérifiée si K est algébriquement clos), est aussi suffisante — autrement dit, on a le résultat suivant :

THÉORÈME 3. *Soient* E *un espace vectoriel de dimension finie sur un corps commutatif* K, *et* u *un endomorphisme de* E. *Pour que* u *soit trigonalisable il faut et il suffit que* u *ait toutes ses valeurs propres dans* K.

Tout revient à montrer que la condition est suffisante; pour la commodité du lecteur débutant, nous allons d'abord le faire lorsque K est *algébriquement clos*, le cas général étant un peu plus difficile à traiter.

Nous devons donc montrer que, si K est algébriquement clos, *tout* endomorphisme de E est trigonalisable. Soit u un tel endomorphisme; comme K est algébriquement clos, u possède au moins une valeur propre, autrement dit il existe un $\alpha_{11} \in K$ et un vecteur *non nul* $x_1 \in E$ tels que l'on ait

$$(12)\qquad u(x_1) = \alpha_{11}x_1.$$

Soit D le sous-espace vectoriel de dimension 1 de E engendré par x_1, de sorte que D est stable par u, et soit F un supplémentaire de D dans E, de sorte que $E = D \oplus F$ (somme directe); l'existence de F résulte du § 19, Corollaire 2 du Théorème 2. Désignons par p l'homomorphisme de E sur F obtenu en projetant chaque $x \in E$ sur F parallèlement à D (§ 17, n° 4), et construisons un endomorphisme v de F en posant

$$(13)\qquad v(x) = p(u(x)) \quad \text{pour tout } x \in F;$$

comme F est de dimension $n - 1$, on peut, en raisonnant par récurrence, supposer le Théorème déjà établi pour F et v, donc construire une base $x_2, \ldots, x_n$ de F telle

que l'on ait des relations de la forme

$$(14)\qquad \begin{cases} v(x_2) = \alpha_{22}x_2 \\ v(x_3) = \alpha_{32}x_2 + \alpha_{33}x_3 \\ \cdots\cdots\cdots\cdots\cdots \\ v(x_n) = \alpha_{n2}x_2 + \alpha_{n3}x_3 + \ldots + \alpha_{nn}x_n; \end{cases}$$

mais la relation (13) montre que, pour tout $x \in F$, le vecteur $u(x)$ ne diffère du vecteur $v(x)$ que par un multiple scalaire de x_1, et en particulier on aura des relations de la forme

$$(15)\qquad \begin{cases} u(x_2) = \alpha_{21}x_1 + v(x_2) \\ u(x_3) = \alpha_{31}x_1 + v(x_3) \\ \cdots\cdots\cdots\cdots\cdots \\ u(x_n) = \alpha_{n1}x_1 + v(x_n); \end{cases}$$

cela dit, puisque $x_2, \ldots, x_n$ forment une base de F il est clair que $x_1, x_2, \ldots, x_n$ forment une base de E, et les relations (12), (14) et (15) montrent que la matrice de u par rapport à cette base est triangulaire. Le Théorème est donc établi pour un corps K algébriquement clos.

Passons maintenant au *cas d'un corps commutatif* K *quelconque*; on va naturellement s'inpsirer de la démonstration précédente, et raisonner par récurrence sur la dimension n de E. Choisissons d'abord une valeur propre $\alpha_{11} \in K$ de u et un vecteur propre correspondant $x_1 \in E$; on a donc à nouveau la relation (12). Comme ci-dessus, notons D la droite engendrée par x_1, F un sous-espace de dimension $n - 1$ supplémentaire de D dans E, et v l'endomorphisme de F donné par (13) : tout revient évidemment à faire voir que v est trigonalisable. Or, si l'on raisonne par récurrence sur n, en sorte qu'on peut supposer le théorème déjà démontré pour $n - 1 = \dim(F)$, tout revient, pour montrer que v est trigonalisable, à prouver que v *a toutes ses valeurs propres dans* K.

Pour cela choisissons une base quelconque $y_2, \ldots, y_n$ de F et soit V la matrice de v par rapport à cette base; comme on a des relations de la forme

$$u(y_i) = \alpha_{i1}x_1 + v(y_i) \quad (2 \leqslant i \leqslant n)$$

il est clair que, par rapport à la base $x_1, y_2, \ldots, y_n$ de E, l'endomorphisme u admet pour matrice

$$U = \begin{pmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{n1} \\ 0 & \beta_{22} & \cdots & \beta_{n2} \\ 0 & \beta_{32} & \cdots & \beta_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \beta_{2n} & \cdots & \beta_{nn} \end{pmatrix}$$

où l'on a posé $V = (\beta_{ij})_{2 \leqslant i,j \leqslant n}$. Retranchant X aux termes diagonaux de U et calculant à l'aide du § 24, *Exemple* 2 le déterminant de la matrice ainsi obtenue on trouve

$$P_U(X) = (\alpha_{11} - X)\,P_V(X).$$

Par suite, le polynôme $p_v = p_u$ divise le polynôme $p_U = p_u$, et comme celui-ci a toutes ses racines dans K par hypothèse, il en est donc de même de p_u (§ 33, fin du n° 1); autrement dit, u a bien toutes ses valeurs propres dans K, et la démonstration est achevée.

COROLLAIRE 1. *Soit* E *un espace vectoriel de dimension finie sur un corps algébriquement clos; tout endomorphisme de* E *est trigonalisable.*

COROLLAIRE 2. *Soit* U *une matrice carrée d'ordre* $n \geqslant 1$ *à coefficients dans un corps algébriquement clos* K. *Il existe une matrice* $P \in GL(n, K)$ *telle que la matrice*

$$PUP^{-1}$$

soit triangulaire.

Il suffit pour le voir d'appliquer le Corollaire 1 à l'endomorphisme de K^n défini par U, et de tenir compte du § 15, Corollaire du Théorème 2.

COROLLAIRE 3. *Soit* U *une matrice carrée d'ordre* $n \geqslant 1$ *à coefficients dans un corps commutatif* K. *Les propriétés suivantes sont équivalentes :*

a) il existe une matrice $P \in GL(n, K)$ *telle que* PUP^{-1} *soit triangulaire ;*

b) la matrice U *a toutes ses valeurs propres dans* K.

En considérant l'endomorphisme u de K^n défini par U, la propriété *a*) signifie que u est trigonalisable; comme $p_u = p_U$, le Corollaire 3 résulte donc aussitôt du Théorème.

Une matrice $U \in M_n(K)$ qui vérifie la propriété *a*) est dite **trigonalisable sur K.** Bien entendu, il existe toujours une matrice inversible P telle que PUP^{-1} soit triangulaire, mais en général P est à coefficients non dans K, mais dans un corps algébriquement clos contenant K (par exemple, si $K = \mathbf{R}$, on est en général obligé de prendre pour P une matrice inversible complexe) — il suffit pour le voir d'appliquer le Corollaire 2. Dire que U est trigonalisable sur K signifie qu'on peut supposer P à coefficients dans K.

Comme on le verra au § suivant, on peut grandement améliorer le Théorème 3 en montrant qu'il existe une base de E par rapport à laquelle la matrice de u comporte beaucoup plus de zéros qu'une matrice triangulaire; mais la démonstration du résultat complet est nettement plus difficile que celle du Théorème 3, lequel suffit dans beaucoup d'applications.

6. *Cas où toutes les valeurs propres sont simples*

Le Théorème 3 ne repose sur aucune hypothèse concernant les multiplicités des valeurs propres de u. Il arrive fréquemment dans la pratique que non seulement le polynôme p_u ait toutes ses racines dans K mais en outre que celles-ci soient toutes simples. On a alors un résultat beaucoup plus précis que le Théorème 3.

Avant de l'établir, nous allons tout d'abord démontrer le résultat suivant :

THÉORÈME 4. *Soient u un endomorphisme d'un espace vectoriel* E *sur un corps commutatif et* $x_1, \ldots, x_n \in E$ *des vecteurs propres de u associés à des valeurs propres* deux à deux distinctes $\lambda_1, \ldots, \lambda_n$. *Alors les vecteurs $x_1, \ldots, x_n$ sont linéairement indépendants.*

Soit F le sous-espace vectoriel de E engendré par $x_1, \ldots, x_n$; les relations

$$u(x_i) = \lambda_i x_i \quad (1 \leqslant i \leqslant n) \tag{16}$$

montrent que, si

$$x = \sum \xi_i x_i$$

est un vecteur de F, le vecteur

$$u(x) = \sum \xi_i u(x_i) = \sum \lambda_i \xi_i x_i$$

est encore dans F. Par suite F est stable par u, et on peut considérer l'endomorphisme v de F induit par u. On a $v(x_i) = \lambda_i x_i$ et par suite les λ_i sont des valeurs propres de v; les λ_i étant supposés deux à deux distincts on voit que v possède *au moins* n valeurs propres. Mais celles-ci sont racines d'une équation de degré égal à dim (F); on a donc

$$n \leqslant \dim(F);$$

comme $x_1, \ldots, x_n$ engendrent F, on en déduit (§ 19, Théorème 10) que ces vecteurs forment en fait une base de F, et sont donc linéairement indépendants, ce qui achève la démonstration. (Voir aussi les *Exercices* 38 et 39 de ce §).

Remarque 2. L'hypothèse que les valeurs propres λ_i sont *deux à deux distinctes* est essentielle pour assurer la validité du Théorème 4. Par exemple prenons pour u l'application nulle, ou l'application identique; alors, quels que soient $x_1, \ldots, x_n$ *non nuls*, il est clair que $x_1, \ldots, x_n$ sont des vecteurs propres de u, mais il est évidemment hors de question de démontrer que n vecteurs non nuls dans un espace vectoriel quelconque sont toujours linéairement indépendants !

Nous pouvons maintenant énoncer et démontrer le résultat annoncé au début de ce n° :

THÉORÈME 5. *Soit u un endomorphisme d'un espace vectoriel* E *de dimension finie $n \geqslant 1$ sur un corps commutatif* K. *Supposons que le polynôme caractéristique p_u de u admette n racines simples $\lambda_1, \ldots, \lambda_n$ dans* K. *Alors il existe une base de* E *par rapport à laquelle la matrice de u est*

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Pour chaque i, choisissons dans E un vecteur propre x_i de u appartenant à la valeur propre λ_i; d'après le Théorème précédent, les n vecteurs ainsi obtenus sont linéairement indépendants, et comme ils sont en nombre $n = \dim(E)$, ils forment donc une base de E. Il est clair que la matrice de u par rapport à cette base n'est autre que celle de l'énoncé.

COROLLAIRE. *Soit* U *une matrice carrée d'ordre* n *à coefficients dans un corps commutatif* K. *Supposons que son polynôme caractéristique* p_U *possède* n *racines simples* $\lambda_1, \ldots, \lambda_n$ *dans* K. *Il existe alors une matrice* $P \in GL(n, K)$ *telle que l'on ait*

$$PUP^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{pmatrix}.$$

Il suffit pour le voir d'appliquer le Théorème 5 à l'endomorphisme de K^n défini par U.

Exemple 3. Prenons $n = 2$ et $K = \mathbf{C}$; l'équation caractéristique de U est

$$X^2 - \mathrm{Tr}(U)X + \mathrm{dét}(U) = 0,$$

elle a donc deux racines simples (i.e. distinctes) dans K si et seulement si

$$\mathrm{Tr}(U) - 4.\mathrm{dét}(U) \neq 0;$$

s'il en est ainsi, U est donc semblable (§ 15, n° 5) à une matrice diagonale.

Il n'en est plus nécessairement de même si $\mathrm{Tr}(U)^2 - 4.\mathrm{dét}(U) = 0$, par exemple si

$$U = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix};$$

pour que U soit semblable à une matrice diagonale, il faut que U admette dans K^2 deux vecteurs propres linéairement indépendants, i.e. non proportionnels; or, la matrice considérée a pour seule valeur propre 1, et les vecteurs propres correspondants sont les $x = (\xi_1, \xi_2)$ tels que l'on ait $Ux = x$, i.e.

$$\begin{aligned} \xi_1 + \xi_2 &= \xi_1 \\ \xi_2 &= \xi_2, \end{aligned}$$

autrement dit ce sont les multiples scalaires du premier vecteur de la base canonique; on ne peut donc pas trouver dans K^2 deux vecteurs propres non proportionnels de U.

Exemple 4. Prenons $K = \mathbf{R}$ et $n = 2$; alors le Corollaire s'applique si et seulement si

$$\mathrm{Tr}(U)^2 - 4.\mathrm{dét}(U) > 0.$$

Dans ce cas, il existe une matrice inversible *réelle* P telle que PUP^{-1} soit diagonale. Lorsqu'on a au contraire

$$\mathrm{Tr}(U)^2 - 4.\mathrm{dét}(U) < 0,$$

il existe une matrice inversible P *complexe* telle que PUP^{-1} soit diagonale (appliquer l'*Exemple* 3), mais on ne peut pas prendre pour P une matrice réelle.

Enfin, supposons

$$\mathrm{Tr}(U)^2 - 4.\mathrm{dét}(U) = 0,$$

et soit $\lambda \in \mathbf{C}$ l'unique valeur propre de U; celle-ci est du reste réelle, et l'on a même

$$\lambda = \frac{\mathrm{Tr}(U)}{2}$$

vu la théorie des équations du second degré à discriminant nul. Il existe une matrice inversible P (réelle ou complexe) telle que

$$PUP^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

il est clair que p_U est aussi le polynôme caractéristique du second membre, i.e. $(X - \lambda_1)(X - \lambda_2)$; comme p_U admet par hypothèse λ pour racine double il vient donc

$$\lambda_1 = \lambda_2 = \lambda,$$

et par suite

$$PUP^{-1} = \lambda . 1_2,$$

d'où résulte que

$$U = \lambda . 1_2.$$

Lorsque $\mathrm{Tr}(U)^2 - 4.\mathrm{dét}(U) = 0$, il est donc impossible de « diagonaliser » U sauf dans le cas où U est déjà une matrice diagonale.

7. *Caractérisation des endomorphismes diagonalisables*

Soit u un endomorphisme d'un espace vectoriel E de dimension finie n sur un corps commutatif K. On dit que u est **diagonalisable** s'il existe une base de E formée de vecteurs propres de u, autrement dit par rapport à laquelle la matrice de u soit diagonale. Le Théorème 5 donne une condition *suffisante* pour qu'il en soit ainsi : à savoir, que le polynôme caractéristique de u ait toutes ses racines dans K, et qu'elles soient toutes simples (ou, si l'on préfère, que ce polynôme admettre n racines deux à deux distinctes dans K). Mais cette condition n'est évidemment pas nécessaire (exemple trivial : l'endomorphisme unité; sa matrice par rapport à n'importe quelle base de E est diagonale, mais son polynôme caractéristique, à savoir

$$(1 - X)^n,$$

ne possède pas de racines simples !).

On va donc énoncer une condition *nécessaire et suffisante* pour qu'un endomorphisme u de E soit diagonalisable. Pour cela, nous aurons besoin des notions suivantes. Étant donnée une valeur propre λ de u, nous appellerons **multiplicité de λ** la multiplicité de λ comme racine du polynôme p_u. D'autre part, nous appellerons **sous-espace propre de E associé à λ** l'ensemble (qui est évidemment un sous-espace vectoriel de E) des vecteurs propres de u associés à λ; on notera $E(\lambda)$ ce sous-espace; on a donc

$$E(\lambda) = \mathrm{Ker}(u - \lambda . 1).$$

THÉORÈME 6. *Soit u un endomorphisme d'un espace vectoriel* **E** *de dimension finie sur un corps* K. *Pour que u soit diagonalisable, il faut et il suffit que les deux conditions suivantes soient remplies* :

a) le polynôme caractéristique p_u a toutes ses racines dans K;

b) pour toute valeur propre λ, la dimension du sous-espace propre $\mathbf{E}(\lambda)$ *est égale à la multiplicité de λ comme racine de p_u.*

Supposons ces conditions vérifiées, et soient $\lambda_1, \ldots, \lambda_q$ les racines de p_u, et $r_1, \ldots, r_q$ leurs ordres de multiplicité. En vertu du Théorème 4, les sous-espaces vectoriels $\mathrm{E}(\lambda_1), \ldots, \mathrm{E}(\lambda_q)$ sont linéairement indépendants au sens du § 17, n° 3; comme on a en outre

$$\dim \mathrm{E}(\lambda_1) + \cdots + \dim \mathrm{E}(\lambda_q) = n = \dim \mathrm{E}$$

en vertu des relations (*)

$$\dim\ \mathrm{E}(\lambda_i) = r_i, \quad r_1 + \cdots + r_q = n,$$

on voit (§ 19, Corollaire 3 du Théorème 13) que

$$\mathrm{E} = \mathrm{E}(\lambda_1) \oplus \cdots \oplus \mathrm{E}(\lambda_q) \quad \text{(somme directe)}; \tag{17}$$

on obtient alors une base de **E** en réunissant des bases de $\mathrm{E}(\lambda_1), \ldots, \mathrm{E}(\lambda_q)$; cette base est formée de vecteurs propres de u, et u est donc bien diagonalisable.

Supposons inversement que u soit diagonalisable; notons encore $\lambda_1, \ldots, \lambda_q$ ses diverses valeurs propres et posons

$$s_i = \dim \mathrm{E}(\lambda_i).$$

Les sous-espaces $\mathrm{E}(\lambda_i)$ sont linéairement indépendants (Théorème 4); comme E admet une base formée de vecteurs propres de u, i.e. d'éléments de ces sous-espaces, on voit que ces sous-espaces engendrent E. On a donc comme plus haut la relation (17), et on obtient une base de E en réunissant des bases de $\mathrm{E}(\lambda_1), \ldots, \mathrm{E}(\lambda_q)$; par rapport à cette base, la matrice de u est évidemment

$$\mathrm{U} = \begin{pmatrix} \lambda_1 & & & & 0 \\ & \ddots & & & \\ & & \lambda_1 & & \\ & & & \lambda_q & \\ & & & & \ddots \\ 0 & & & & \lambda_q \end{pmatrix}$$

avec s_i termes diagonaux égaux à λ_i; le polynôme caractéristique de **U**, i.e. de u, est donc

$$p_u(\mathrm{X}) = (\lambda_1 - \mathrm{X})^{s_1} \ldots (\lambda_q - \mathrm{X})^{s_q};$$

(*) La première exprime l'hypothèse (*b*), et la seconde l'hypothèse (*a*) de l'énoncé.

cela prouve d'une part que toutes ses racines sont dans K, d'autre part que la multiplicité de λ_i est égale à $s_i = \dim E(\lambda_i)$, et ceci achève la démonstration.

Remarque 3. La condition *a*) est toujours remplie si K est algébriquement clos, par exemple si $K = \mathbf{C}$.

Remarque 4. La condition *b*) signifie que $E(\lambda)$ *possède la plus grande dimension possible* (ce qui est conforme au bon sens, puisqu'on désire trouver suffisamment de vecteurs propres pour pouvoir en extraire une base de E); autrement dit, quel que soit l'endomorphisme u, on a l'inégalité

$$\dim E(\lambda) \leqslant \text{multiplicité de } \lambda$$

pour toute valeur propre de u dans K.

Pour établir ce point posons $E(\lambda) = F$; il est clair que $u(F) \subset F$ et que l'endomorphisme v de F induit par u est l'homothétie de rapport λ; le polynôme caractéristique de v est donc

$$p_v(X) = (\lambda - X)^{\dim E(\lambda)};$$

pour établir le résultat cherché il suffit donc de faire voir que $p_v(X)$ divise $p_u(X)$. Autrement dit, on est ramené à établir le résultat général que voici : *soit* F *un sous-espace de* E *stable par* u, i.e. tel que $u(F) \subset F$; *soit* v *l'endomorphisme de* F *induit par* u, i.e. tel que $v(x) = u(x)$ pour tout $x \in F$; *alors le polynôme* p_v *divise le polynôme* p_u.

Pour cela, choisissons arbitrairement une base $x_1, \ldots, x_r$ de F, et complétons-la en une base de E (ce qui est toujours possible : § 19, Corollaire 2 du Théorème 2). La matrice de u par rapport à la base obtenue est évidemment de la forme

$$U = \begin{pmatrix} V & T \\ 0 & W \end{pmatrix}$$

où V est la matrice de v par rapport à la base $x_1, \ldots, x_r$, où T est une matrice à r lignes et $n - r$ colonnes, et où W est une matrice carrée d'ordre $n - r$. Par suite

$$U - X.1_n = \begin{pmatrix} V - X.1_r & T \\ 0 & W - X.1_{n-r} \end{pmatrix},$$

et le § 24, *Exemple* 2 montre que

$$\det (V - X.1_n) = \det (V - X.1_r) \det (W - X.1_{n-r}),$$

ce qui s'écrit encore

$$p_u(X) = p_v(X) p_w(X)$$

et démontre le résultat annoncé.

A la notion d'endomorphisme diagonalisable correspond celle de matrice **diagonalisable**. Une matrice $U \in M_n(K)$ est dite **diagonalisable** s'il existe une matrice $P \in GL(n, K)$ telle que la matrice PUP^{-1} soit diagonale; en raison du § 15, n° 5, cela veut dire que l'endomorphisme de K^n défini par U est diagonalisable. C'est

toujours le cas si le polynôme caractéristique p_u a toutes ses racines dans K, et si elles sont toutes simples.

Une notion analogue mais un peu plus subtile est celle de **matrice semi-simple**. On appelle ainsi toute matrice $U \in M_n(K)$ possédant la propriété suivante : il existe un sur-corps commutatif L de K tel que U soit diagonalisable en tant que matrice à coefficients dans L (on démontre qu'on peut alors prendre pour L n'importe quel sur-corps algébriquement clos de K). Si l'on regarde U comme la matrice d'un endomorphisme u de K^n, cela veut dire, intuitivement, que u possède « suffisamment » de vecteurs propres pourvu qu'on autorise les vecteurs propres à avoir leurs coordonnées dans un sur-corps de K.

Exemple 3. Prenons $K = \mathbf{R}$ et la matrice

$$U = \begin{pmatrix} \cos\theta, & -\sin\theta \\ \sin\theta, & \cos\theta \end{pmatrix};$$

ses valeurs propres dans $\mathbf{C}$ sont (*Remarque* 1) les nombres

$$\cos\theta \pm i.\sin\theta;$$

si θ n'est pas multiple de π, ces valeurs propres sont distinctes, en sorte que U est diagonalisable en tant que matrice à coefficients dans $\mathbf{C}$; mais elle ne l'est pas comme matrice à coefficients dans $\mathbf{R}$, puisque ses valeurs propres ne sont pas réelles — autrement dit, sur le corps $\mathbf{R}$, la matrice U est seulement semi-simple (y compris évidemment pour θ multiple de π).

§ 35. Forme canonique d'une matrice

Le but de ce § est de démontrer le Théorème de Jordan dont on trouvera l'énoncé plus loin, et qui remplace la réduction à la forme diagonale pour les matrices non diagonalisables. Quoique ce théorème joue un grand rôle dans certaines parties des Mathématiques (notamment dans la théorie des équations différentielles linéaires), sa connaissance n'est pas indispensable au lecteur débutant. Celui-ci pourra donc, soit négliger purement et simplement ce §, soit ne l'étudier qu'à titre d'exercice.

1. Le théorème de Hamilton-Cayley

Soient K un anneau commutatif et f un polynôme à une indéterminée à coefficients dans K, à savoir

$$f(X) = a_0 + a_1X + \cdots + a_rX^r.$$

Étant donnés un sur-anneau L de K tel que K soit dans le *centre* de L (i.e. tel que $au = ua$ pour $a \in K$, $u \in L$) et un élément u de L, nous poserons

$$f(u) = a_0 + a_1u + \cdots + a_ru^r;$$

lorsque L est commutatif on retrouve la définition du § 28, n° 1 (dans le cas général, on pourrait se ramener à un anneau commutatif en remplaçant L par le sous-anneau K[u] formé des combinaisons linéaires, à coefficients dans K, des puissances de u).

Si en particulier on prend $L = M_n(K)$, anneau des matrices carrées d'ordre n à coefficients dans K, on voit qu'on peut définir $f(U)$ pour tout polynôme $f \in K[X]$ et toute matrice carrée U à coefficients dans K (ou même dans un sur-anneau commutatif de K).

Ceci dit :

THÉORÈME 1. *Soient* U *une matrice carrée à coefficients dans un anneau commutatif* K *et*

$$p_U(X) = \det(U - X.1)$$

son polynôme caractéristique. On a alors

$$p_U(U) = 0.$$

Posons, si U est d'ordre n,

$$(-1)^n p_U(X) = X^n - \tau_1(U)X^{n-1} + \cdots + (-1)^n \tau_n(U)$$

comme au § 34, n° 3. Les coefficients $\tau_i(U)$ sont des polynômes à coefficients entiers rationnels en les éléments de U, polynômes dont les coefficients sont évidemment indépendants de U comme de l'anneau de base K : par exemple la formule

$$\tau_1(U) = \alpha_{11} + \cdots + \alpha_{nn}$$

permettant de calculer $\tau_1(U)$ en fonction des coefficients α_{ij} de U est la même pour toutes les matrices U de degré n et tous les anneaux commutatifs. Comme

$$(-1)^n p_U(U) = U^n - \tau_1(U)U^{n-1} + \cdots + (-1)^n \tau_n(U).1_n,$$

on voit donc qu'il existe des polynômes f_{ij} $(1 \leqslant i, j \leqslant n)$ à coefficients *entiers rationnels*, à n^2 indéterminées X_{ij} $(1 \leqslant i, j \leqslant n)$, tels que, pour tout anneau commutatif K et toute matrice

$$U = (\alpha_{ij})_{1 \leqslant i, j \leqslant n}$$

de degré n à coefficients dans K, les coefficients de la matrice $p_U(U)$ s'obtiennent en substituant les coefficients de U aux indéterminées dans les polynômes f_{ij}. Pour montrer que $p_U(U) = 0$, il suffit donc de montrer que les polynômes f_{ij} sont nuls, i.e. que leurs coefficients sont nuls, et pour cela (§ 28, Théorème 1) que chaque f_{ij} s'annule toutes les fois qu'on y remplace les indéterminées par des *entiers rationnels* arbitraires. Mais par construction même des f_{ij} cela signifie qu'on a $p_U(U) = 0$ pour toute matrice carrée à coefficients dans l'anneau **Z**.

Ainsi, pour établir le Théorème 1 pour *tout* anneau commutatif K, il suffit de l'établir pour l'anneau **Z**. Pour cela nous allons l'établir pour un corps K *algébriquement clos* quelconque; le Théorème sera alors établi pour **C**, donc pour le sous-anneau **Z** de **C**, donc dans tous les cas !

Pour cela considérons l'endomorphisme u de K^n ayant U pour matrice par rapport à la base canonique. En vertu du § 34, Théorème 3, il existe une base $(x_i)_{1 \leqslant i \leqslant n}$ de K^n telle que l'on ait des relations

$$u(x_i) = \rho_{ii}x_i + \cdots + \rho_{i1}x_1 \quad (1 \leqslant i \leqslant n); \tag{1}$$

comme le déterminant d'une matrice triangulaire est le produit de ses termes diagonaux on a alors

$$p_U(X) = p_u(X) = (\rho_{11} - X) \ldots (\rho_{nn} - X),$$

et tout revient par conséquent, puisque $p_U(U)$ est évidemment la matrice (par rapport à la base canonique de K^n) de l'endomorphisme

$$p_u(u) = (\rho_{11} - u) \ldots (\rho_{nn} - u),$$

à montrer que celui-ci est nul. Or posons

$$u_i = \rho_{ii} - u;$$

il résulte de (1) que

$$u_j(x_i) = (\rho_{ii} - \rho_{jj})x_i + y_{ij}$$

où y_{ij} appartient au sous-espace vectoriel F_{i-1} de K^n engendré par les vecteurs $x_1, \ldots, x_{i-1}$, et il résulte aussitôt de ces formules que l'on a

$$u_i(F_i) \subset F_{i-1}$$

pour tout i. Pour en déduire que l'endomorphisme

$$v = p_u(u) = u_1 \circ \cdots \circ u_n$$

est nul, il suffit alors de remarquer que

$$v(K^n) = v(F_n) = u_1 \circ \cdots \circ u_{n-1} \circ u_n(F_n) \subset u_1 \circ \cdots \circ u_{n-1}(F_{n-1}) \subset \cdots \subset u_1(F_1)$$

et comme on a évidemment $u_1(F_1) = \{0\}$, le Théorème est établi.

Exemple 1. Si U est une matrice carrée d'ordre 2 à coefficients dans un anneau commutatif quelconque, on a

$$U^2 - \mathrm{Tr}(U).U + \mathrm{dét}\,(U).1_2 = 0.$$

On conseille au lecteur de vérifier ce résultat par un calcul direct.

Remarque 1. Comme $p_U(X) = \mathrm{dét}\,(U - X.1_n)$, le lecteur débutant et ingénieux aura sans doute l'idée de démontrer le Théorème 1 en écrivant que

$$p_U(U) = \mathrm{dét}\,(U - U.1_n) = \mathrm{dét}\,(U - U) = \mathrm{dét}\,(0) = 0.$$

Cette démonstration séduisante repose malheureusement sur l'erreur qui consiste à croire que, lorsqu'on remplace dans $U - X.1_n$ l'indéterminée X par la matrice U, on trouve la matrice $U - U.1_n = 0$, ce qui n'est pas le cas; en effet, $U - X.1_n$ s'obtient en retranchant X aux termes diagonaux de U, et si l'on remplace X par U dans le résultat obtenu on trouve la matrice

$$\begin{pmatrix} \alpha_{11} - U & \alpha_{21} & \cdots & \alpha_{n1} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{1n} & \alpha_{2n} & \cdots & \alpha_{nn} - U \end{pmatrix}$$

à coefficients dans le sous-anneau commutatif de $M_n(K)$ engendré par K et U; or cette matrice n'est visiblement pas nulle en général ! Le théorème de Hamilton-Cayley affirme seulement que le *déterminant* de cette matrice est nul.

Cette *Remarque* montre qu'on peut être parfois induit en erreur par les notations qu'on utilise.

2. *Décomposition en endomorphismes nilpotents*

Le premier pas dans la démonstration du Théorème de Jordan consiste à établir le résultat suivant :

THÉORÈME 2. *Soit u un endomorphisme d'un espace vectoriel* E *de dimension finie sur un corps commutatif* K, *et supposons que le polynôme p_u ait toutes ses racines dans* K. *Posons*

$$p_u(X) = (\lambda_1 - X)^{r_1} \dots (\lambda_q - X)^{r_q}$$

où $\lambda_1, \dots, \lambda_q \in K$ sont les diverses racines de p_u et $r_1, \dots, r_q$ leurs ordres de multiplicité. Enfin, posons

$$E_i = \operatorname{Ker}[(u - \lambda_i)^{r_i}] \quad \text{pour} \quad 1 \leqslant i \leqslant q.$$

Alors E *est somme directe des sous-espaces E_i, et on a*

$$\dim(E_i) = r_i \quad \text{pour} \quad 1 \leqslant i \leqslant q.$$

Pour simplifier les notations, posons

$$p(X) = p_u(X), \qquad f_i(X) = (\lambda_i - X)^{r_i}$$
$$g_i(X) = f_1(X) \dots f_{i-1}(X) f_{i+1}(X) \dots f_q(X).$$

Comme les λ_i sont deux à deux distincts, les polynômes g_i ($1 \leqslant i \leqslant q$) sont premiers entre eux, et par suite il existe des polynômes $h_i \in K[X]$ tels que l'on ait

$$\sum_{1 \leqslant i \leqslant q} h_i(X) g_i(X) = 1.$$

Posant

$$u_i = f_i(u), \qquad v_i = g_i(u), \qquad w_i = h_i(u)$$

de sorte que les $3q$ endomorphismes de E ainsi obtenus commutent deux à deux (car ce sont tous des polynômes en u), on a donc

$$w_1 \circ v_1 + \dots + w_q \circ v_q = j_E,$$

endomorphisme identique de E. Ceci montre qu'on a

$$x = w_1(v_1(x)) + \dots + w_q(v_q(x)) \quad \text{pour tout } x \in E; \tag{2}$$

mais d'après le Théorème 1 on a

$$0 = p(u) = f_i(u) g_i(u) = u_i \circ v_i$$

pour tout i, et puisque u_i et w_i commutent il vient *a fortiori* $u_i \circ w_i \circ v_i = 0$. Ceci montre que

$$w_i(v_i(x)) \in E_i$$

pour tout i et tout x, et (2) prouve donc que

$$E = E_1 + \dots + E_q.$$

Il reste à montrer que la somme est directe et que dim $(E_i) = r_i$; comme

$$r_1 + \cdots + r_q = d^0(p_u) = \dim(E),$$

le Corollaire 3 du Théorème 13 du § 19 montre d'ailleurs qu'il suffit d'établir les relations

$$\dim(E_i) \leqslant r_i. \tag{3}$$

Or comme $E_i = \mathrm{Ker}\,(u_i)$ et comme u commute à u_i, il est clair que $u(E_i) \subset E_i$; désignant par u_i' l'endomorphisme de E_i induit par u, on voit donc (§ 34, *Remarque* 4) que le polynôme p_u est divisible par le polynôme $p_{u_i'}$. Comme le polynôme p_u a toutes ses racines dans K, il en est donc de même de $p_{u_i'}$, et par suite il existe une base de E_i par rapport à laquelle la matrice de u_i' est triangulaire (§ 34, Théorème 3); mais comme on a

$$(u_i' - \lambda_i)^{r_i} = 0$$

par construction même de E_i, les coefficients diagonaux de la matrice de u_i' par rapport à la base en question sont nécessairement égaux à λ_i, en sorte que

$$p_{u_i'}(X) = (\lambda_i - X)^{\dim(E_i)}.$$

Pour que ce polynôme divise p_u il est évidemment nécessaire que la relation (3) soit vérifiée, ce qui achève la démonstration.

Les scalaires $\lambda_1, \ldots, \lambda_q$ sont naturellement les diverses *valeurs propres de u*. Le Théorème 2 montre que E est somme directe de sous-espaces E_i tels que l'on ait

$$u(E_i) \subset E_i, \qquad (u - \lambda_i)^{r_i} = 0 \ \text{ dans } \ E_i.$$

Comme on obtient une base de E en réunissant des bases des divers E_i, on voit que, pour mettre la matrice de u par rapport à une base de E sous une forme aussi simple que possible il suffit de se placer dans E_i et de résoudre le même problème pour l'endomorphisme de E_i induit par u ou, ce qui revient au même, par $u - \lambda_i$, lequel est **nilpotent**, ce qui veut dire que l'une de ses puissances est nulle.

Ainsi tout revient maintenant, étant donné un endomorphisme nilpotent d'un espace vectoriel, à choisir une base de celui-ci par rapport à laquelle la matrice de l'endomorphisme donné soit aussi simple que possible. C'est ce qu'on va faire dans le n° suivant.

3. *Structure des endomorphismes nilpotents*

Voici maintenant le second pas dans la démonstration du Théorème de Jordan :

THÉORÈME 3. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie $n \geqslant 1$ sur un corps commutatif* K. *Supposons qu'il existe un entier $p \geqslant 0$ tel que*

$$u^p = 0$$

Il reste à montrer que la somme est directe et que $\dim(E_i) = r_i$; comme

$$r_1 + \cdots + r_q = d^0(p_u) = \dim(E),$$

le Corollaire 3 du Théorème 13 du § 19 montre d'ailleurs qu'il suffit d'établir les relations

$$\dim(E_i) \leqslant r_i. \tag{3}$$

Or comme $E_i = \mathrm{Ker}(u_i)$ et comme u commute à u_i, il est clair que $u(E_i) \subset E_i$; désignant par u_i' l'endomorphisme de E_i induit par u, on voit donc (§ 34, *Remarque* 4) que le polynôme p_u est divisible par le polynôme $p_{u_i'}$. Comme le polynôme p_u a toutes ses racines dans K, il en est donc de même de $p_{u_i'}$, et par suite il existe une base de E_i par rapport à laquelle la matrice de u_i' est triangulaire (§ 34, Théorème 3); mais comme on a

$$(u_i' - \lambda_i)^{r_i} = 0$$

par construction même de E_i, les coefficients diagonaux de la matrice de u_i' par rapport à la base en question sont nécessairement égaux à λ_i, en sorte que

$$p_{u_i'}(X) = (\lambda_i - X)^{\dim(E_i)}.$$

Pour que ce polynôme divise p_u il est évidemment nécessaire que la relation (3) soit vérifiée, ce qui achève la démonstration.

Les scalaires $\lambda_1, \ldots, \lambda_q$ sont naturellement les diverses *valeurs propres de u*. Le Théorème 2 montre que E est somme directe de sous-espaces E_i tels que l'on ait

$$u(E_i) \subset E_i, \qquad (u - \lambda_i)^{r_i} = 0 \text{ dans } E_i.$$

Comme on obtient une base de E en réunissant des bases des divers E_i, on voit que, pour mettre la matrice de u par rapport à une base de E sous une forme aussi simple que possible il suffit de se placer dans E_i et de résoudre le même problème pour l'endomorphisme de E_i induit par u ou, ce qui revient au même, par $u - \lambda_i$, lequel est **nilpotent**, ce qui veut dire que l'une de ses puissances est nulle.

Ainsi tout revient maintenant, étant donné un endomorphisme nilpotent d'un espace vectoriel, à choisir une base de celui-ci par rapport à laquelle la matrice de l'endomorphisme donné soit aussi simple que possible. C'est ce qu'on va faire dans le n° suivant.

3. *Structure des endomorphismes nilpotents*

Voici maintenant le second pas dans la démonstration du Théorème de Jordan :

THÉORÈME 3. *Soit u un endomorphisme d'un espace vectoriel E de dimension finie $n \geqslant 1$ sur un corps commutatif* K. *Supposons qu'il existe un entier $p > 0$ tel que*

$$u^p = 0$$

(i.e. que u soit nilpotent). *Il existe alors une base de* E *par rapport à laquelle la matrice de* u *est de la forme*

$$\begin{pmatrix} 0 & \nu_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & \nu_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & \nu_n \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

où chaque scalaire ν_i *est égal à* 0 *ou à* 1.

On peut évidemment supposer $u \neq 0$, le Théorème étant trivial si $u = 0$. Il existe alors un entier $q \geqslant 1$ tel que

$$u^q \neq 0,\ u^{q+1} = 0;$$

pour tout entier $r \geqslant 0$, nous poserons d'autre part

$$\mathrm{E}_r = \mathrm{Ker}(u^r);$$

on a $\mathrm{E}_0 = \{0\}$ et $\mathrm{E}_{q+1} = \mathrm{E}$.

LEMME 1. *La suite de sous-espaces*

$$0 = \mathrm{E}_0 \subset \mathrm{E}_1 \subset \dots \subset \mathrm{E}_q \subset \mathrm{E}_{q+1} = \mathrm{E}$$

est strictement croissante, et on a $u(\mathrm{E}_{i+1}) \subset \mathrm{E}_i$ *pour tout* $i \geqslant 0$.

Si un vecteur x est annulé par l'endomorphisme u^{i+1}, il est clair que $u(x)$ est annulé par u^i, ce qui établit la seconde assertion de l'énoncé du Lemme 1. Il reste à établir que E_{i+1} contient strictement E_i pour $0 \leqslant i \leqslant q$. Tout d'abord, la relation

$$\mathrm{E}_{i+1} \supset \mathrm{E}_i$$

est triviale pour tout i. Supposons maintenant qu'on ait $\mathrm{E}_{i+1} = \mathrm{E}_i$ pour un indice i tel que $0 < i \leqslant q$; pour tout $x \in \mathrm{E}$, on a

$$0 = u^{q+1}(x) = u^{i+1}(u^{q-i}(x)),$$

donc $u^{q-i}(x) \in \mathrm{E}_{i+1}$; de $\mathrm{E}_{i+1} = \mathrm{E}_i$ résulterait donc que $u^{q-i}(x) \in \mathrm{E}_i$ pour tout $x \in \mathrm{E}$, donc que $u^q(x) = 0$ pour tout $x \in \mathrm{E}$, contrairement à la définition de q.

LEMME 2. *Soient* i *un entier tel que* $1 \leqslant i \leqslant q$ *et* F *un sous-espace vectoriel de* E *tel que* $\mathrm{F} \cap \mathrm{E}_i = \{0\}$; *on a alors* $u(\mathrm{F}) \cap \mathrm{E}_{i-1} = \{0\}$, *et* u *induit un isomorphisme de* F *sur* $u(\mathrm{F})$.

Considérons un vecteur $x \in u(\mathrm{F}) \cap \mathrm{E}_{i-1}$; il existe un $y \in \mathrm{F}$ tel que $x = u(y)$, et comme $u^{i-1}(x) = u^i(y)$ on voit que la relation $x \in u(\mathrm{F}) \cap \mathrm{E}_{i-1}$ implique $y \in \mathrm{F} \cap \mathrm{E}_i$, donc $y = 0$, donc $x = 0$. L'application de F dans $u(\mathrm{F})$ induite par u est évidemment linéaire et surjective; elle est injective, car si $y \in \mathrm{F}$ vérifie $u(y) = 0$, ce qui implique $u(y) \in \mathrm{E}_{i-1}$, le raisonnement précédent montre aussi que $y = 0$. Le lemme est donc établi.

LEMME 3. *Il existe des sous-espaces vectoriels* $F_1, \ldots, F_{q+1}$ *de* E *qui possèdent les propriétés suivantes* :

a) E_i *est somme directe de* E_{i-1} *et de* F_i *pour tout* i *tel que* $1 \leqslant i \leqslant q+1$;

b) u *applique injectivement* F_i *dans* F_{i-1} *pour tout* i *tel que* $2 \leqslant i \leqslant q+1$.

On prend tout d'abord pour F_{q+1} un supplémentaire de E_q dans $E_{q+1} = E$; le sous-espace $u(F_{q+1})$ est alors contenu dans E_q, et ne rencontre E_{q-1} qu'en 0 d'après le lemme 2; par suite, il existe un supplémentaire F_q de E_{q-1} dans E_q qui contient $u(F_{q+1})$; d'après le lemme 1, le sous-espace $u(F_q)$ est contenu dans E_{q-1}, et ne rencontre E_{q-2} qu'en 0 d'après le lemme 2; on peut donc construire un supplémentaire F_{q-1} de E_{q-2} dans E_{q-1} qui contienne $u(F_q)$; en poursuivant la construction ainsi de suite, on forme évidemment des sous-espaces F_i vérifiant la condition *a*) et tels que $u(F_i) \subset F_{i-1}$; le fait que u applique *injectivement* F_i dans F_{i-1} résulte alors de la seconde assertion du Lemme 2.

Nous pouvons maintenant achever la *démonstration du Théorème* 3. Pour cela, construisons les sous-espaces F_i ($1 \leqslant i \leqslant q+1$) du Lemme 3. Désignons par

$$x_{11}, x_{12}, \ldots, x_{1,r_1}$$

une base de F_{q+1}. Comme ces vecteurs sont linéairement indépendants, et que u applique *injectivement* F_{q+1} dans F_q, leurs images par u sont linéairement indépendantes, et font donc partie d'une base de F_q (§ 19, Théorème 2); autrement dit, il existe une base de F_q de la forme

$$x_{21}, x_{22}, \ldots, x_{2,r_1}, x_{2,r_1+1}, \ldots, x_{2,r_2}$$

avec

$$u(x_{1j}) = x_{2j} \quad \text{pour} \quad 1 \leqslant j \leqslant r_1.$$

En raisonnant de même à partir des vecteurs x_{2j}, on voit qu'il existe une base

$$x_{31}, \ldots, x_{3,r_2}$$

de F_{q-1} telle que l'on ait

$$u(x_{2j}) = x_{3j} \quad \text{pour} \quad 1 \leqslant j \leqslant r_2.$$

En poursuivant ainsi de suite on parvient finalement à une base

$$x_{q+1,1}, \ldots, x_{q+1,r_{q+1}}$$

de $F_1 = E_1$, telle que l'on ait des relations

$$u(x_{q,j}) = x_{q+1,j} \quad \text{pour } 1 \leqslant j \leqslant r_q;$$

et comme $E_1 = \mathrm{Ker}\,(u)$ on a en outre

$$u(x_{q+1,j}) = 0 \quad \text{pour } 1 \leqslant j \leqslant r_{q+1}.$$

Ceci dit, et comme E est évidemment somme directe des sous-espaces $F_1, \ldots, F_{q+1}$, on voit que les $r_1 + r_2 + \cdots + r_{q+1}$ vecteurs x_{ij} ainsi construits forment une base de E. On peut écrire cette base sous la forme du tableau que voici :

$$\begin{array}{llll} x_{11} & \ldots & x_{1,r_1} & \\ u(x_{11}) & \ldots & u(x_{1,r_1}), & x_{2,r_1+1} \quad \ldots \quad x_{2,r_2} \\ \ldots & \ldots & \ldots & \ldots \\ u^q(x_{11}) & \ldots & u^q(x_{1,r_1}), u^{q-1}(x_{2,r_1+1}) & \ldots \; u^{q-1}(x_{2,r_2}) \quad \ldots \quad x_{q+1,r_{q+1}}. \end{array}$$

En écrivant ces vecteurs colonne par colonne, en commençant par le bas, on trouve donc une base $(x_i)_{1 \leqslant i \leqslant n}$ de E avec la propriété que, pour chaque i, on a

$$\text{soit} \quad u(x_i) = 0, \qquad \text{soit} \quad u(x_i) = x_{i-1};$$

la matrice de u par rapport à cette base possède alors la forme indiquée dans l'énoncé du Théorème 3.

4. *Le théorème de Jordan*

Avant d'énoncer le résultat final de ce §, introduisons la définition suivante : on appelle **matrice réduite** (ou **matrice de Jordan**) à coefficients dans un corps commutatif K toute matrice carrée à coefficients dans K qui est de la forme

$$\begin{pmatrix} \lambda & 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \lambda & 1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}$$

autrement dit dont tous les termes diagonaux sont égaux, qui comporte des 1 immédiatement au-dessus de la diagonale, et des 0 partout ailleurs.

Par exemple,

$$(\lambda), \qquad \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \qquad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}$$

sont des matrices réduites d'ordres 1, 2, 3 respectivement.

Cela dit :

THÉORÈME 4 (Jordan). *Soient* E *un espace vectoriel de dimension finie sur un corps commutatif* K *et* u *un endomorphisme de* E. *Les propriétés suivantes sont équivalentes :*

a) Toutes les valeurs propres de u *sont dans* K, i.e. le polynôme p_u a toutes ses racines dans K.

b) il existe une base de E *par rapport à laquelle la matrice de u est de la forme*

$$\begin{pmatrix} U_1 & 0 & 0 & \dots & 0 \\ 0 & U_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & U_r \end{pmatrix}$$

où chaque U_i *est une matrice réduite à coefficients dans* K.

Supposons que toutes les valeurs propres de u soient dans K, et appliquons tout d'abord le Théorème 2; soit u_i la restriction de u à chaque E_i. Pour chaque i, supposons construite une base $(x_{ij})_{1 \leqslant j \leqslant n_i}$ de E_i par rapport à laquelle la matrice de u_i ait la forme indiquée dans l'énoncé du Théorème; alors, en superposant les bases des divers E_i ainsi obtenues, on trouve évidemment une base de E par rapport à laquelle la matrice de u a la forme cherchée. Tout revient donc à examiner u_i. Or on a

$$u_i = \lambda_i + v_i$$

où v_i est un endomorphisme *nilpotent* de E_i; il suffit donc de montrer qu'il existe une base de E_i par rapport à laquelle la matrice de v_i est un tableau diagonal de matrices réduites; pour cela, il suffit d'appliquer le Théorème 3 à v_i et à E_i, en formant une matrice réduite toutes les fois qu'on a une « chaîne » de coefficients v_i égaux à 1. Par exemple, la matrice

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

s'écrit

$$\begin{pmatrix} U_1 & 0 & 0 \\ 0 & U_2 & 0 \\ 0 & 0 & U_3 \end{pmatrix}$$

avec

$$U_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad U_2 = (0),$$

$$U_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Il est clair qu'on procédant ainsi on montre que *a)* implique *b)*.

Pour montrer que *b)* implique *a)*, on observe que, d'après le § 3+, *Remarque* 4, la propriété *b)* implique

$$p_u(X) = p_{U_1}(X) \dots p_{U_r}(X);$$

pour montrer que p_u a toutes ses racines dans K, il suffit donc de montrer qu'il en est ainsi de p_U lorsque U est une matrice réduite, ce qui est clair comme on l'a vu au § 34, n° 5. Le Théorème est donc démontré.

L'hypothèse *a*) du Théorème de Jordan, et donc aussi la propriété *b*), est toujours vérifiée lorsque le corps K est *algébriquement clos*. Dans le cas général, la condition *a*) signifie aussi (§ 34, Théorème 3) que u est *trigonalisable*.

Bien entendu, le Théorème de Jordan s'applique aussi aux matrices : si U est une matrice carrée d'ordre n à coefficients dans K, et si U a toutes ses valeurs propres dans K, il existe une matrice $P \in GL(n, K)$ telle que

$$PUP^{-1} = \begin{pmatrix} U_1 & 0 & 0 & \dots & 0 \\ 0 & U_2 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & U_r \end{pmatrix}$$

où les U_i sont des matrices de Jordan.

Exemple 2. Soit U une matrice carrée d'ordre 4 sur un corps K algébriquement clos. Alors U est semblable à une matrice ayant l'une des formes que voici :

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix}; \quad \begin{pmatrix} & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda_4 \end{pmatrix};$$

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \mu & 1 \\ 0 & 0 & 0 & \mu \end{pmatrix}; \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

Il existe à première vue d'autres possibilités, mais elles se ramènent aux précédentes — par exemple la matrice

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix} \quad \text{est semblable à} \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda_1 \end{pmatrix}$$

comme on le voit en effectuant une permutation des axes de coordonnées.

§ 36. Formes hermitiennes

Dans tout ce §, on désigne par K un *corps commutatif*, et on suppose qu'on s'est donné une fois pour toutes une application de K dans K, notée

$$\lambda \mapsto \lambda^*,$$

et possédant les propriétés suivantes : c'est un homomorphisme, autrement dit on a les identités

$$(1) \qquad (\lambda + \mu)^* = \lambda^* + \mu^*, \qquad (\lambda\mu)^* = \lambda^*\mu^*, \qquad 1^* = 1,$$

et en outre on a

$$(2) \qquad (\lambda^*)^* = \lambda;$$

une telle application s'appelle une **involution** dans K. Ce n'est autre qu'un automorphisme du corps K dont le carré est l'automorphisme identique.

Les exemples élémentaires les plus importants sont les suivants :

Exemple 1 (**cas orthogonal réel**). On prend $K = \mathbf{R}$ et $\lambda^* = \lambda$ pour tout $\lambda \in K$.

Exemple 2 (**cas orthogonal complexe**). On prend $K = \mathbf{C}$ et $\lambda^* = \lambda$ pour tout $\lambda \in K$.

Exemple 3 (**cas hermitien complexe**). On prend $K = \mathbf{C}$ et $\lambda^* = \bar{\lambda}$ pout tout $\lambda \in K$.

Mais il y a beaucoup d'autres situations possibles.

Exemple 4. On prend $K = \mathbf{Q}[\sqrt{d}]$ où d est un entier rationnel qui n'est pas un carré parfait (§ 9) et on pose

$$(a + b\sqrt{d})^* = a - b\sqrt{d}$$

quels que soient $a, b \in \mathbf{Q}$.

1. *Formes sesquilinéaires, formes hermitiennes*

Soit L un espace vectoriel sur le corps K. On appelle **forme sesquilinéaire sur L** (relativement à l'involution donnée sur K) toute application

$$f : L \times L \to K$$

vérifiant les deux conditions que voici :

(SQ 1) : *Pour tout* $y \in L$, *l'application* $x \to f(x, y)$ *de* L *dans* K *est linéaire.*

(SQ 2) : *Pour tout* $x \in L$, *l'application* $y \to f(x, y)^*$ *de* L *dans* K *est linéaire.*

On peut remplacer ces axiomes par les identités que voici :

$$f(x' + x'', y) = f(x', y) + f(x'', y), \qquad f(\lambda x, y) = \lambda f(x, y)$$
$$f(x, y' + y'') = f(x, y') + f(x, y''), \qquad f(x, \lambda y) = \lambda^* . f(x, y)$$

Lorsque l'involution choisie sur K est l'identité, on retrouve évidemment les formes bilinéaires du § 21.

On dit qu'une forme sesquilinéaire f est **hermitienne** si l'on a

$$f(x, y) = f(y, x)^* \tag{3}$$

quels que soient $x, y \in L$; on a alors, en particulier,

$$f(x, x)^* = f(x, x); \tag{4}$$

dans le cas hermitien complexe (*Exemple* 3), le nombre $f(x, x)$ est donc toujours *réel* puisqu'il est égal à son conjugué.

Si l'involution choisie sur K est l'identité, la relation (3) s'écrit

$$f(x, y) = f(y, x); \tag{5}$$

on dit alors que f est une forme bilinéaire **symétrique sur L.**

Exemple 5. Il est facile de construire toutes les formes sesquilinéaires sur L lorsque L est de dimension finie sur K. Choisissons pour cela une base $(a_i)_{1 \leqslant i \leqslant n}$ de L et soient

$$x = \sum \xi_i a_i, \quad y = \sum \eta_i a_i$$

deux vecteurs de L. Comme l'expression

$$f_y(x) = f(x, y)$$

est une forme linéaire en x, d'après (SQ 1), il vient

$$f(x, y) = f_y(\sum \xi_i a_i) = \sum \xi_i f_y(a_i) = \sum \xi_i f(a_i, y);$$

mais comme

$$f_i(y) = f(a_i, y)^*$$

est une forme linéaire en y d'après (SQ_2), il vient

$$f(a_i, y)^* = f_i(\sum \eta_j a_j) = \sum_j \eta_j f_i(a_j) = \sum_j \eta_j f(a_i, a_j)^*$$

et en prenant le * de chaque membre on voit, en tenant compte de (1) et (2), que

$$f(a_i, y) = \sum_j f(a_i, a_j)\eta_j^*;$$

portant dans l'expression trouvée plus haut pour $f(x, y)$ il vient évidemment

$$f(x, y) = \sum_{i,j} \alpha_{ij}\xi_i\eta_j^* \qquad \text{où} \qquad \alpha_{ij} = f(a_i, a_j). \tag{6}$$

Il est immédiat de vérifier qu'inversement, quels que soient les $\alpha_{ij} \in K$, la fonction f définie par (6) est une forme sesquilinéaire sur L.

Exemple 6. Supposons f hermitienne dans l'*Exemple* 5; il vient alors

$$\alpha_{ji} = f(a_j, a_i) = f(a_i, a_j)^* = \alpha_{ij}^*;$$

inversement, la relation

$$\alpha_{ji} = \alpha_{ij}^* \tag{7}$$

implique

$$f(x, y) = \sum \alpha_{ij}\xi_i\eta_j^* = \sum \alpha_{ji}^*\eta_j^*\xi_i = (\sum \alpha_{ji}\eta_j\xi_i^*)^* = f(y, x)^*$$

et caractérise donc les formes hermitiennes.

Lorsqu'on a choisi l'involution identique sur K, la relation (7) se réduit à

$$\alpha_{ij} = \alpha_{ji}, \tag{8}$$

et caractérise alors les formes bilinéaires *symétriques* sur L.

Exemple 7. Plaçons-nous dans le cas orthogonal réel (Exemple 3) et prenons pour L l'espace vectoriel des vecteurs d'origine donnée 0 dans l'espace usuel; le produit scalaire (§ 21, *Exemple* 3)

$$f(x, y) = (x|y)$$

est une forme bilinéaire symétrique sur L.

Exemple 8. Dans la théorie de la Relativité, on utilise la forme bilinéaire symétrique sur $\mathbf{R}^4$ donnée par

$$f(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \xi_3\eta_3 - c^2\xi_4\eta_4$$

où c est la vitesse de la lumière. On l'appelle la **forme de Lorentz;** il s'agit bien entendu du cas orthogonal réel.

Exemple 9. Dans le cas hermitien complexe, prenons pour L l'espace vectoriel formé par les fonctions $x(t)$ à valeurs complexes qui sont définies et continues dans l'intervalle $0 \leqslant t \leqslant 1$; alors la formule

$$f(x, y) = \int_0^1 x(t)\, y(t)\, dt$$

définit une forme sesquilinéaire hermitienne sur L (laquelle joue un rôle important en Analyse, notamment dans la théorie des séries de Fourier). On pourrait prendre aussi

$$f(x, y) = \int_0^1 \int_0^1 x(s)\, \overline{y(t)}\, \mathrm{K}(s, t)\, ds\, dt$$

où $\mathrm{K}(s, t)$ est une fonction définie et continue dans le carré $0 \leqslant s, t \leqslant 1$, et choisie une fois pour toutes; l'expression obtenue est une forme sesquilinéaire sur L, et est hermitienne si et seulement si la fonction K vérifie

$$\mathrm{K}(t, s) = \overline{\mathrm{K}(s, t)}$$

quels que soient s, t.

Les méthodes purement algébriques développées dans ce § ne permettent pas d'établir des propriétés non triviales des formes définies dans cet *Exemple :* leur étude détaillée exige l'emploi de méthodes analytiques inspirées des considérations du présent §, mais beaucoup plus compliquées que celles-ci, et est en grande partie à l'origine de l'Analyse Fonctionnelle (théorie des espaces de Hilbert).

Soient L un espace vectoriel de dimension finie sur K et f une forme sesquilinéaire sur L. Étant donnée une base $(a_i)_{1 \leqslant i \leqslant n}$ de L, les scalaires

$$\alpha_{ij} = f(a_i, a_j)$$

figurant dans la formule (6) s'appellent les **coefficients de** f par rapport à la base considérée, et la matrice carrée

$$(\alpha_{ij})_{1 \leqslant i, j \leqslant n}$$

s'appelle la **matrice de** f **par rapport à la base** en question.

On dit qu'une matrice carrée

$$\mathrm{A} = (\alpha_{ij})_{1 \leqslant i, j \leqslant n}$$

à coefficients dans K est **hermitienne** (relativement à l'involution considérée sur K) lorsque ses coefficients vérifient la relation (7) ci-dessus. Lorsque l'involution choisie sur K est l'identité, i.e. lorsque

$$\alpha_{ji} = \alpha_{ij},$$

on dit que la matrice A est **symétrique.**

Remarque 1. On appelle **matrice hermitienne complexe** une matrice hermitienne dans le cas hermitien complexe (*Exemple* 9), autrement dit une matrice carrée

(α_{ij}) à coefficients complexes telle que

$$\alpha_{ji} = \overline{\alpha_{ij}}$$

quels que soient i, j. D'une manière générale, lorsque $K = \mathbf{C}$ et qu'on emploie la terminologie « hermitienne » sans préciser davantage, il est toujours sous-entendu qu'on choisit sur $\mathbf{C}$ l'involution

$$\lambda \mapsto \bar{\lambda},$$

passage à l'imaginaire conjugué.

2. *Formes non dégénérées*

Soient L un espace vectoriel sur K et f une forme sesquilinéaire sur L. Pour tout $y \in L$, la fonction

$$f_y(x) = f(x, y) \tag{9}$$

est une forme linéaire sur L d'après (SQ 1) : par suite, f définit une application

$$\hat{f} : y \mapsto f_y \tag{10}$$

de L dans son dual L* (§ 16, n° 1). Quels que soient $x, y, z \in L$ on a

$$f_{y+z}(x) = f(x, y + z) = f(x, y) + f(x, z) = f_y(x) + f_z(x),$$

donc

$$f_{y+z} = f_y + f_z$$

ou, si l'on préfère,

$$\hat{f}(y + z) = \hat{f}(y) + \hat{f}(z); \tag{11}$$

on a d'autre part

$$f_{\lambda y}(x) = f(x, \lambda y) = \lambda^* . f(x, y) = \lambda^* . f_y(x)$$

et par suite

$$\hat{f}(\lambda y) = \lambda^* . \hat{f}(y) \tag{12}$$

quels que soient $\lambda \in K$ et $y \in L$. Si l'involution choisie sur K est l'application identique (cas des formes bilinéaires), on voit donc que $\hat{f}$ est un homomorphisme de l'espace vectoriel L dans l'espace vectoriel L*. Dans le cas général il n'en est plus de même, et on exprime les relations (11) et (12) en disant que l'application $\hat{f}$ de L dans son dual est **semi-linéaire**. En dimension finie, les applications semi-linéaires possèdent, à des modifications triviales près, les mêmes propriétés que les applications linéaires.

Le noyau de $\hat{f}$, ensemble des $y \in L$ tels que $\hat{f}(y) = 0$, est aussi l'ensemble des

$y \in \mathrm{L}$ tels que l'on ait

$$f(x, y) = 0 \quad \text{pour tout } x \in \mathrm{L}. \tag{13}$$

On dit que la forme sesquilinéaire f est **non dégénérée** si (13) implique $y = 0$, autrement dit si l'application $\hat{f}$ de L dans son dual est injective.

THÉORÈME 1. *Soient* L *un espace vectoriel de dimension finie sur* K, f *une forme sesquilinéaire sur* L, *et* $\mathrm{A} = (\alpha_{ij})$ *la matrice de* f *par rapport à une base de* L. *Les propriétés suivantes sont équivalentes :*

a) : *La relation*

$$f(x, y) = 0 \quad \text{pour tout } x \in \mathrm{L}$$

implique $y = 0$.

b) : *La relation*

$$f(x, y) = 0 \quad \text{pour tout } y \in \mathrm{L}$$

implique $x = 0$.

c) : *La matrice* A *est inversible*, i.e. on a

$$\det(\alpha_{ij}) \neq 0.$$

d) : *L'application* $\hat{f}$ *de* L *dans son dual est bijective*, autrement dit, *pour toute forme linéaire* u *sur* L, *il existe un et un seul* $y \in \mathrm{L}$ *tel que l'on ait*

$$u(x) = f(x, y) \quad \text{pour tout } x \in \mathrm{L}.$$

Utilisant les notations de l'*Exemple* 5, les y tels que $f(x, y) = 0$ pour tout x sont évidemment les solutions du système d'équations

$$\sum_j \alpha_{ij} \eta_j^* = 0 \quad (1 \leqslant i \leqslant n), \tag{14}$$

et les x tels que $f(x, y) = 0$ pour tout y sont les solutions du système d'équations

$$\sum_i \alpha_{ij} \xi_i = 0 \quad (1 \leqslant j \leqslant n); \tag{15}$$

dans la propriété *d*), si l'on pose $u(a_i) = \nu_i$, tout revient à déterminer y de telle sorte que $f(a_i, y) = \nu_i$ pour tout i, i.e. à résoudre le système d'équations

$$\sum_i \alpha_{ij} \eta_j^* = \nu_i \quad (1 \leqslant i \leqslant n). \tag{16}$$

Ceci dit, *d*) signifie que le système (16) possède une et une seule solution quels que soient les seconds membres; l'équivalence avec *c*), et avec *a*) qui exprime que le système (14) n'admet que la solution triviale, résulte donc du § 20, Théorème 2. Par ailleurs, *b*) signifie que (15) n'a que la solution triviale, donc que la matrice (α_{ij}) est inversible, et comme celle-ci est la transposée de la matrice (α_{ij}) on voit que *b*) et *c*) sont équivalentes, ce qui achève la démonstration.

Remarque 2. Comme *a*) signifie que le noyau de $\hat{f}$ est nul, l'équivalence des propriétés *a*) et *d*) résulte aussi du § 19, Corollaire 1 du Théorème 13 (lequel s'étend trivialement aux applications semi-linéaires).

Exemple 10. La forme de Lorentz (*Exemple* 8) est non dégénérée car sa matrice par rapport à la base canonique de $\mathbf{R}^4$ est

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

et est donc inversible puisque c n'est pas nulle.

3. *Adjoint d'un homomorphisme*

Soient L et M deux espaces vectoriels de dimension finie sur K, et f et g deux formes sesquilinéaires *non dégénérées* sur L et M respectivement.

Soit u un homomorphisme de L dans M, et considérons l'expression

$$g(u(x), y) \quad \text{pour } x \in \mathrm{L},\ y \in \mathrm{M};$$

pour $y \in \mathrm{M}$ donné, c'est évidemment une forme linéaire en $x \in \mathrm{L}$, et par suite le Théorème 1, *d*) montre que pour chaque $y \in \mathrm{M}$ il existe un et un seul $y' \in \mathrm{L}$ tel que l'on ait

$$g(u(x), y) = f(x, y') \quad \text{pour tout } x \in \mathrm{L};$$

en posant $y' = u^*(y)$ on définit une application

$$u^* : \mathrm{M} \to \mathrm{L}$$

et on a, par conséquent, la relation

$$f(x, u^*(y)) = g(u(x), y) \quad \text{quels que soient } x \in \mathrm{L},\ y \in \mathrm{M}. \tag{17}$$

L'application u^* est *linéaire* comme u; en effet, quels que soient $y, z \in \mathrm{M}$, on a

$$\begin{aligned} f(x, u^*(y+z)) = g(u(x), y+z) &= g(u(x), y) + g(u(x), z) \\ &= f(x, u^*(y)) + f(x, u^*(z)) = f(x, u^*(y) + u^*(z)) \end{aligned}$$

d'où la relation $u^*(y+z) = u^*(y) + u^*(z)$; de même, on a

$$f(x, u^*(\lambda y)) = g(u(x), \lambda y) = \lambda^* g(u(x), y) = \lambda^* f(x, u^*(y)) = f(x, \lambda u^*(y))$$

ce qui montre que $u^*(\lambda y) = \lambda u^*(y)$ et prouve notre assertion.

L'homomorphisme

$$u^* : \mathrm{M} \to \mathrm{L}$$

défini par (17) s'appelle **l'adjoint de l'homomorphisme**

$$u : \mathrm{L} \to \mathrm{M}$$

par rapport aux formes f et g. Dans le cas particulier où $\mathrm{L} = \mathrm{M}, f = g$, on dit que u^* est l'adjoint de u par rapport à f.

A la notion d'adjoint d'un homomorphisme correspond celle d'adjointe d'une matrice. Dans ce qui précède, *supposons* que L admette une base $(a_i)_{1 \leqslant i \leqslant p}$ par rapport à laquelle l'expression de f soit

$$f(x, y) = \sum_{1 \leqslant i \leqslant p} \xi_i \eta_i^*,$$

et que M admette de même une base $(b_j)_{1 \leqslant j \leqslant q}$ par rapport à laquelle l'expression de g soit

$$g(z, t) = \sum_{1 \leqslant j \leqslant q} \zeta_j \tau_j^*,$$

enfin soient

$$\mathrm{A} = (\alpha_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q}$$
$$\mathrm{B} = (\beta_{ji})_{1 \leqslant j \leqslant q,\ 1 \leqslant i \leqslant p}$$

les matrices de u et u^* par rapport aux bases considérées. Prenons des vecteurs

$$x = \sum \xi_i a_i, \qquad y = \sum \eta_j b_j;$$

on aura

$$u(x) = \sum \alpha_{ij} \xi_i b_j, \qquad u^*(y) = \sum \beta_{ji} \eta_j a_i$$

et par suite

$$g(u(x), y) = \sum \alpha_{ij} \xi_i \eta_j^*$$
$$f(x, u^*(y)) = \sum \xi_i \beta_{ji}^* \eta_j^*;$$

écrivant que les résultats obtenus sont égaux quels que soient x et y il vient donc la relation

$$\beta_{ji}^* = \alpha_{ij}$$

ou encore

$$\beta_{ji} = \alpha_{ij}^*,$$

en sorte que B *s'obtient en transformant par l'involution de* K *les termes de la matrice* ${}^t\mathrm{A}$ *transposée de* A. On dit que B est l'**adjointe de la matrice** A par rapport à l'involution donnée, et on écrit

$$\mathrm{B} = \mathrm{A}^*$$

(toutefois, lorsque l'involution choisie sur K est l'application identique, par exemple dans le cas orthogonal réel ou orthogonal complexe, on dit transposée au lieu d'adjointe, et on utilise la notation habituelle

$${}^t\mathrm{A}$$

au lieu de A^*).

Par exemple, dans le cas hermitien complexe, l'**adjointe d'une matrice complexe** A n'est autre que

$$\mathrm{A}^* = {}^t\bar{\mathrm{A}},$$

transposée de la matrice imaginaire conjuguée de A.

Remarque 3. Il est possible de calculer la matrice de u^* en fonction de la matrice de u sans faire d'hypothèses sur les bases choisies dans L et M (le calcul précédent suppose que les bases sont « orthonormales » pour f et g comme on le verra plus loin). Voir *Exercice* 11.

Les règles de calcul sur les matrices adjointes sont analogues à celles concernant les matrices transposées — de façon précise, on a les relations

$$(A + B)^* = A^* + B^*, \qquad (\lambda A)^* = \lambda^* A^*,$$
$$(AB)^* = B^* A^*, \qquad 1_n^* = 1_n,$$
$$(A^*)^* = A;$$

pour les obtenir, on écrit les formules relatives aux transposées (§ 16, Théorème 4), puis on transforme les résultats obtenus par l'automorphisme

$$\lambda \mapsto \lambda^*$$

de K.

On peut aussi prouver des formules analogues concernant les homomorphismes eux-mêmes. Si par exemple u et v sont deux homomorphismes de L dans M, on a évidemment

$$(\lambda u + \mu v)^* = \lambda^* u^* + \mu^* v^*$$

quels que soient les scalaires λ et μ. Si L, M, N sont trois espaces vectoriels de dimension finie sur K, munis de formes sesquilinéaires non dégénérées f, g, h, et si l'on a des homomorphismes $u : L \to M$ et $v : M \to N$, alors on a de même la formule

$$(v \circ u)^* = u^* \circ v^*;$$

en effet, posant $v \circ u = w$ il vient

$$f(x, w^*(z)) = h(w(x), z) = h(v(u(x)), z) = g(u(x), v^*(z)) = f(x, u^*(v^*(z))),$$

d'où le résultat cherché.

La formule « évidente »

$$(u^*)^* = u,$$

où u est un homomorphisme de L dans M, n'est par contre valable que si f et g sont *hermitiennes*. Posons en effet $u^* = v$; l'adjoint de cet homomorphisme de M dans L est alors donné par la relation (17) dans laquelle on doit remplacer f par g, g par f, u par v, autrement dit par la relation

$$g(y, v^*(x)) = f(v(y), x);$$

si f est hermitienne, le second membre s'écrit encore

$$f(x, v(y))^* = f(x, u^*(y))^* = g(u(x), y)^*$$

d'après (17); si de plus g est hermitienne, il vient donc finalement

$$g(y, v^*(x)) = g(y, u(x))$$

d'où le résultat cherché, à savoir que $v^* = u$ si $v = u^*$.

4. *Orthogonalité par rapport à une forme hermitienne non dégénérée*

Soit f une forme *hermitienne* sur un espace vectoriel **L** sur **K**. On dit que deux vecteurs $x, y \in \mathbf{L}$ sont **orthogonaux** par rapport à f si

$$f(x, y) = 0;$$

dans le cas de l'*Exemple* 7 on retrouve la notion usuelle, puisque $(x|y)$, étant le produit des longueurs de x et y par le cosinus de leur angle, ne peut être nul que si l'un des vecteurs est nul, ou bien si cet angle est multiple de $\pi/2$.

Si x et y sont orthogonaux, on a

$$f(x+y, x+y) = f(x, x) + f(y, y) \tag{17}$$

car dans tous les cas on a

$$\begin{aligned} f(x+y, x+y) &= f(x, x) + f(x, y) + f(y, x) + f(y, y) \\ &= f(x, x) + f(x, y) + f(x, y)^* + f(y, y) \end{aligned}$$

puisque f est hermitienne. Dans le cas de l'*Exemple* 7, la relation précédente signifie que le carré de la longueur de $x+y$ est la somme des carrés des longueurs de x et de y, autrement dit se réduit au *théorème de Pythagore*.

Soit maintenant M un sous-espace vectoriel de **L**; on appelle **orthogonal de** M par rapport à f l'ensemble, noté généralement

$$\mathrm{M}^{\perp},$$

des $x \in \mathrm{L}$ qui sont orthogonaux à *tout* $y \in \mathrm{M}$. D'après (SQ_1), c'est un sous-espace vectoriel de L. Dans le cas de l'*Exemple* 7, si M est une droite (resp. un plan) passant par l'origine, alors $\mathrm{M}^{\perp}$ est le plan (resp. la droite) perpendiculaire à M et passant par l'origine.

On a toujours

$$\mathrm{M} \subset (\mathrm{M}^{\perp})^{\perp}, \tag{18}$$

car le second membre est formé des $z \in \mathbf{L}$ orthogonaux à tous les $y \in \mathrm{M}^{\perp}$, et comme ceux-ci sont orthogonaux aux $x \in \mathrm{M}$ il est clair que tout $x \in \mathrm{M}$ est dans le second membre de la relation (18).

Théorème 2. *Soit f une forme hermitienne non dégénérée sur un espace vectoriel* L *de dimension finie sur* K; *on a alors*

$$(\mathrm{M}^{\perp})^{\perp} = \mathrm{M}$$

pour tout sous-espace vectoriel M *de* L.

Il suffit d'établir l'inclusion opposée à (18), i.e. de prouver que tout $z \in (\mathrm{M}^{\perp})^{\perp}$ appartient à M; pour cela, il suffit (§ 19, Théorème 3) de montrer que si une forme linéaire u sur L est nulle sur M, alors $u(z) = 0$. Or, d'après le Théorème 1, *d*), on

peut écrire

$$u(x) = f(x, y)$$

pour un $y \in L$; l'hypothèse que u est nulle sur M signifie que $y \in M^\perp$, et le fait que $u(z) = 0$, i.e. que $f(z, y) = 0$, résulte alors trivialement de l'hypothèse que z est orthogonal à $M^\perp$, ce qui achève la démonstration.

THÉORÈME 3. *Soient* M *et* N *des sous-espaces vectoriels de* L. *Dans les hypothèses du Théorème 2, on a les relations*

$$(M + N)^\perp = M^\perp \cap N^\perp; \qquad (M \cap N)^\perp = M^\perp + N^\perp.$$

Comme $M + N$ se compose des $x + y$ où $x \in M$ et $y \in N$, et comme

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

la première relation est immédiate. Pour obtenir la seconde, on remplace M et N par $M^\perp$ et $N^\perp$ dans la première; il vient alors

$$(M^\perp + N^\perp)^\perp = (M^\perp)^\perp \cap (N^\perp)^\perp = M \cap N$$

d'après le Théorème 2, et en appliquant le Théorème 2 à nouveau on obtient la relation cherchée.

THÉORÈME 4. *Soit f une forme hermitienne non dégénérée sur un espace vectoriel* L *de dimension finie sur* K. *On a alors*

$$\dim(M) + \dim(M^\perp) = \dim(L)$$

pour tout sous-espace vectoriel M *de* L.

On sait que, si M^0 désigne l'orthogonal de M dans le dual L^* de L, i.e. l'ensemble des formes linéaires u sur L qui vérifient

$$u(x) = 0 \quad \text{pour tout } x \in M, \tag{19}$$

on a

$$\dim(M) + \dim(M^0) = \dim(L)$$

(§ 19, Théorème 9). Tout revient donc à montrer que $M^\perp$ et M^0 ont la même dimension. Or, étant donnée une forme linéaire u sur L, il existe un $y \in L$ et un seul (Théorème 1) tel que l'on ait $u(x) = f(x, y)$ pour tout $x \in L$, et évidemment la relation (19) équivaut à $y \in M^\perp$; il s'ensuit que l'application f du n° 2 applique $M^\perp$ sur M^0, et par suite qu'il existe une application semi-linéaire et bijective de $M^\perp$ sur M^0, ces deux espaces vectoriels ont donc bien même dimension, ce qui achève la démonstration.

THÉORÈME 5. *Soit f une forme hermitienne sur un espace vectoriel* L *de dimension finie sur* K. *Étant donné un sous-espace* M *de* L, *les propriétés suivantes sont équivalentes :*

a) On a $M \cap M^\perp = 0$

b) *La restriction de f à* M *est une forme hermitienne non dégénérée sur* M.

c) *L'espace* L *est somme directe des sous-espaces* M *et* $M^{\perp}$, i.e. tout $x \in L$ s'écrit d'une façon et d'une seule sous la forme

$$x = y + z \quad \text{avec} \quad y \in M \text{ et } z \in M^{\perp}.$$

Si de plus f est non dégénérée, les conditions précédentes sont encore équivalentes à la suivante :

d) *L'espace* L *est somme des sous-espace* M *et* $M^{\perp}$, i.e. tout $x \in L$ s'écrit d'une façon au moins sous la forme

$$x = y + z \quad \text{avec} \quad y \in M \text{ et } z \in M.$$

Pour démontrer ce Théorème, notons d'abord que les $x \in M \cap M^{\perp}$ sont les $x \in M$ qui vérifient

$$f(x, y) = 0 \quad \text{pour tout } y \in M;$$

l'équivalence des conditions *a*) et *b*) est donc claire. Il est non moins évident que la condition *c*), i.e. la relation

$$L = M \oplus M^{\perp},$$

implique *a*). Pour achever de montrer l'équivalence des conditions *a*), *b*) et *c*), il suffit donc de montrer que *b*) implique *c*). Comme d'ailleurs *c*) n'est autre (§ 17, Théorème 1) que la conjonction de *a*) et *d*), et comme on sait déjà que *b*) implique *a*), il suffit d'établir que *b*) implique *d*).

Or soit $x \in L$, et considérons sur M la fonction

$$u(y) = f(y, x);$$

c'est évidemment une forme linéaire sur M, et comme la restriction de f à M est non dégénérée le Théorème 1 montre qu'il existe un $x' \in M$ tel que l'on ait

$$u(y) = f(y, x') \quad \text{pour tout } y \in M;$$

on a donc

$$f(y, x) = f(y, x')$$

ou encore

$$f(y, x - x') = 0$$

pour tout $y \in M$, et par suite $x - x' \in M^{\perp}$; comme $x' \in M$ la condition *d*) est bien vérifiée.

Supposons maintenant f non dégénérée, et *d*) vérifiée; écrivons (§ 19, Corollaire 2 du Théorème 13) la relation

$$\dim(M + M^{\perp}) = \dim(M) + \dim(M^{\perp}) - \dim(M \cap M^{\perp});$$

par hypothèse le premier membre est égal à dim (L); tenant compte du Théorème 4, on voit donc que $\dim(M \cap M^{\perp}) = 0$. Ainsi *d*) implique *a*), et ceci termine la démonstration.

On notera que, lorsque f est non dégénérée, la relation ci-dessus s'écrit

$$\dim (M + M^\perp) = \dim (L) - \dim (M \cap M^\perp),$$

en sorte que dans cette hypothèse l'équivalence des propriétés *a*), *c*) et *d*) est immédiate. Mais on a parfois besoin dans la pratique du Théorème 5 pour des formes dégénérées, et de plus le raisonnement que nous avons utilisé pour établir que *b*) implique *d*) s'étend aux espaces de Hilbert (ce sont certains espaces vectoriels complexes de dimension infinie munis d'une forme hermitienne définie positive, et qui jouent un rôle fondamental en Analyse).

Considérons à nouveau une forme hermitienne f *non dégénérée* sur un espace vectoriel L de dimension finie sur K.

On dit qu'un sous-espace vectoriel M de L est **isotrope** lorsque

$$M \cap M^\perp \neq 0,$$

et **non isotrope** lorsque

$$M \cap M^\perp = 0,$$

Le Théorème 5 signifie donc que l'on a

$$L = M \oplus M^\perp \quad \text{(somme directe)}$$

si et seulement si M est non isotrope. Dans ce cas, le n° 4 du § 17 montre qu'il existe un et un seul endomorphisme p_M de l'espace vectoriel L vérifiant les conditions suivantes :

$$p_M \circ p_M = p_M;$$
$$p_M(L) = M;$$
$$p_M(M^\perp) = 0;$$

pour tout $x \in L$, on a

$$x = p_M(x) + y \quad \text{avec } y \in M^\perp,$$

et cette relation caractérise p_M — plus précisément, $p_M(x)$ est l'unique élément de M tel que $x - p_M(x)$ soit orthogonal à M. On dit que $p_M(x)$ est **la projection orthogonale** de x sur M, et que p_M est l'**opérateur de projection orthogonale** sur M.

Exemple 11. Si M est la droite engendrée par un vecteur $a \in L$, il est clair que M est isotrope si et seulement si

$$f(a, a) = 0;$$

on dit alors que a est un **vecteur isotrope** pour f, et l'ensemble de ces vecteurs (qui est évidemment une réunion de droites passant par 0) s'appelle **le cône isotrope** de f. Lorsque f est par exemple la forme de Lorentz, celui-ci est l'ensemble des $(x, y, z, t) \in \mathbf{R}^4$ tels que

$$x^2 + y^2 + z^2 - c^2t^2 = 0.$$

Par contre, dans le cas de l'*Exemple* 7 (produit scalaire dans l'espace usuel), le seul vecteur isotrope est 0 — du reste il est clair que si trois nombres *réels*

x, y, z vérifient la relation

$$x^2 + y^2 + z^2 = 0,$$

on a $x = y = z = 0$. (Par contre cette équation a des solutions *complexes* non triviales.)

Revenant au cas général, soit a un vecteur non isotrope pour f. Alors la projection orthogonale de tout $x \in L$ sur la droite M engendrée par a est donnée par la formule

$$p_M(x) = \frac{f(x, a)}{f(a, a)} a;$$

le second membre est en effet dans M, et on a

$$f(x - p_M(x), a) = f(x, a) - \frac{f(x, a)}{f(a, a)} f(a, a) = 0,$$

ce qui montre que $x - p_M(x)$ est orthogonal à M.

Corollaire du Théorème 5. *Soit f une forme hermitienne sur un espace vectoriel* L *de dimension finie sur* K. *Les propriétés suivantes sont équivalentes :*

a) : *Pour $x \in L$, la relation $f(x, x) = 0$ implique $x = 0$* (autrement dit f ne possède aucun vecteur isotrope non nul).

b) : *On a*

$$L = M \oplus M^\perp$$

pour tout sous-espace vectoriel M *de* L.

Il est clair que, pour tout sous-espace M de L, le sous-espace $M \cap M^\perp$ se compose de vecteurs isotropes; donc la propriété *a*) implique

$$M \cap M^\perp = 0$$

et par suite implique *b*) d'après le Théorème 5. Le fait que *b*) implique *a*) s'obtient en écrivant qu'on a

$$M \cap M^\perp = 0$$

pour tout sous-espace M de dimension un de L.

Remarque 4. Supposons K algébriquement clos et $\lambda^* = \lambda$ pour tout $\lambda \in K$. Si f est une forme hermitienne (i.e. bilinéaire symétrique, vu l'hypothèse faite sur l'involution de K) et si L est de dimension au moins égale à 2, il existe toujours des vecteurs isotropes *non nuls* pour f.

En effet, comme $\dim(L) \geq 2$ on peut choisir dans L deux vecteurs a et b non proportionnels. Pour $\lambda \in K$, on a

$$f(a\lambda + b, a\lambda + b) = f(a, a)\lambda^2 + 2f(a, b)\lambda + f(b, b);$$

si a est isotrope, il n'y a rien à démontrer; si a n'est pas isotrope, on voit que les $\lambda \in K$ pour lesquels $a\lambda + b$ est isotrope sont les racines d'une équation du second degré. Comme K est supposé algébriquement clos, l'équation

considérée a au moins une racine, et le vecteur isotrope $a\lambda + b$ correspondant n'est pas nul puisque a et b sont non proportionnels.

Cette *Remarque* s'applique notamment dans le cas *orthogonal complexe.*

5. *Bases orthogonales*

Soient L un espace vectoriel de dimension finie sur K et f une forme hermitienne sur L. On appelle **base orthogonale de L** (par rapport à f) toute base $(a_i)_{1 \leqslant i \leqslant n}$ de L telle que l'on ait

$$f(a_i, a_j) = 0 \quad \text{pour } i \neq j;$$

cela signifie que la matrice de f par rapport à la base en question est *diagonale,* ou encore que l'expression de f par rapport à cette base est de la forme

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \alpha_i \xi_i \eta_i^*. \tag{20}$$

Par exemple, la base canonique de $\mathbf{R}^4$ est orthogonale par rapport à la forme de Lorentz.

THÉORÈME 6. *Soit f une forme hermitienne sur un espace vectoriel* L *de dimension finie sur* K. *Si* K *est de caractéristique différente de* 2, *il existe une base de* L *orthogonale relativement à f.*

Le Théorème est trivial si $f = 0$, de sorte que nous supposerons $f \neq 0$. Comme il n'y a rien à démontrer si L est de dimension 1, nous raisonnerons par récurrence sur $n = \dim(\mathrm{L})$.

Supposons trouvé un vecteur $a_1 \in \mathrm{L}$ *non isotrope* pour f. Alors (Théorème 5) L est somme directe de la droite engendrée par a_1 et de l'hyperplan L′ orthogonal à celle-ci. Comme $\dim(\mathrm{L}') = n - 1$, l'hypothèse de récurrence montre que L′ admet une base $(a_2, \ldots, a_n)$ orthogonale par rapport à la restriction f' de f à L′. Il est clair alors que $(a_1, a_2, \ldots, a_n)$ est une base de L orthogonale relativement à L.

Pour achever la démonstration, il reste donc à établir le résultat suivant :

LEMME. *Soit f une forme hermitienne sur un espace vectoriel* L *sur* K, *et supposons la caractéristique de* K *différente de* 2. *Si f n'est pas nulle, il existe dans* L *des vecteurs non isotropes pour f.*

Autrement dit, si $f \neq 0$ et si $f(x, x) = 0$ pour tout $x \in \mathrm{L}$, alors K est de caractéristique 2.

En effet, supposons $f(x, x) = 0$ pour tout $x \in \mathrm{L}$; la relation

$$f(x + y, x + y) = f(x, x) + f(x, y) + f(x, y)^* + f(y, y)$$

montre qu'on a alors

$$f(x, y) + f(x, y)^* = 0$$

quels que soient $x, y \in \mathrm{L}$. Remplaçant x par tx où $t \in \mathrm{K}$, on voit que le scalaire $u = f(x, y)$ vérifie

$$tu + (tu)^* = 0 \quad \text{pour tout } t \in \mathrm{K}.$$

Si $f \neq 0$, on peut choisir x et y de telle sorte que $u \neq 0$; prenant $t = u^{-1}$ il vient alors

$$0 = 1 + 1^* = 1 + 1,$$

et ceci montre que K est de caractéristique 2.

¶ *Remarque* 5. Si K est de caractéristique 2 et si l'involution de K est l'identité, alors toute forme alternée est aussi symétrique, et tout vecteur $x \in L$ est isotrope par rapport à une telle forme. L'hypothèse que K est de caractéristique $\neq 2$ est donc essentielle pour assurer la validité de l'énoncé précédent, et est bien entendu toujours vérifiée dans la « pratique ».

Corollaire 1. *Soit f une forme bilinéaire symétrique sur un espace vectoriel* L *de dimension finie n sur* K. *Supposons* K *algébriquement clos et de caractéristique différente de* 2 (par exemple $K = C$). *Alors il existe une base de* L *telle que l'expression de f par rapport à cette base soit*

$$f(x, y) = \xi_1\eta_1 + \xi_2\eta_2 + \cdots + \xi_r\eta_r$$

où r est un entier inférieur à n. Pour que f soit non dégénérée il faut et il suffit que $r = n$.

Choisissons en effet une base $(b_i)_{1 \leqslant i \leqslant n}$ orthogonale pour f; on peut supposer

$$f(b_i, b_i) = \beta_i \neq 0 \quad \text{pour } 1 \leqslant i \leqslant r$$
$$f(b_i, b_i) = 0 \quad \text{pour } r + 1 \leqslant i \leqslant n.$$

Comme K est algébriquement clos il existe des $\lambda_i \in K$ $(1 \leqslant i \leqslant r)$ vérifiant

$$\lambda_i^2 f(b_i, b_i) = 1 \quad \text{pour } 1 \leqslant i \leqslant r;$$

cela veut dire, puisque f est bilinéaire, que les vecteurs $a_i = \lambda_i b_i$ vérifient

$$f(a_i, a_i) = 1 \quad \text{pour } 1 \leqslant i \leqslant r;$$

la base $(a_1, \ldots, a_r, b_{r+1}, \ldots, b_n)$ satisfait alors aux conditions de l'énoncé.

Si $r < n$, il est clair que b_n est orthogonal à tous les b_i $(1 \leqslant i \leqslant n)$ donc à tout $x \in L$, et par suite que f est dégénérée; on a donc $r = n$ si f est non dégénérée. Si inversement $r = n$, la matrice de f par rapport à la base qu'on vient de former est la matrice unité, donc est inversible, et f est non dégénérée.

Remarque 6. Plus généralement, pour que la forme (20) soit non dégénérée il faut et il suffit que $\alpha_i \neq 0$ pour $1 \leqslant i \leqslant n$, car on doit exprimer que la matrice de f par rapport à la base considérée est inversible.

Corollaire 2. *Soit f une forme bilinéaire symétrique sur un espace vectoriel* réel L *de dimension finie n. Il existe des entiers p et q tels que $p + q \leqslant n$, et une base de* L *telle que l'expression de f par rapport à cette base soit*

$$f(x, y) = \xi_1\eta_1 + \cdots + \xi_p\eta_p - \xi_{p+1}\eta_{p+1} - \cdots - \xi_{p+q}\eta_{p+q}$$

f est non dégénérée si et seulement si $p + q = n$.

La démonstration est analogue à celle du Corollaire précédent; on choisit une base (b_i) de L orthogonale pour f; on peut supposer

$$\begin{aligned} f(b_i, b_i) &> 0 \quad \text{pour } 1 \leqslant i \leqslant p \\ f(b_i, b_i) &< 0 \quad \text{pour } p+1 \leqslant i \leqslant p+q \\ f(b_i, b_i) &= 0 \quad \text{pour } p+q+1 \leqslant i \leqslant n. \end{aligned}$$

En prenant

$$a_i = \begin{cases} \dfrac{b_i}{\sqrt{f(b_i, b_i)}} & \text{pour } 1 \leqslant i \leqslant p \\ \dfrac{b_i}{\sqrt{-f(b_i, b_i)}} & \text{pour } p+1 \leqslant i \leqslant p+q \\ b_i & \text{pour } p+q+1 \leqslant i \leqslant n \end{cases}$$

on trouve une base de L par rapport à laquelle l'expression de f est évidemment celle de l'énoncé. Le fait que f soit non dégénérée si et seulement si $n = p + q$ résulte de la *Remarque* 6.

Corollaire 3. *Soit f une forme hermitienne sur un espace vectoriel* complexe L *de dimension finie n. Il existe des entiers p et q tels que $p + q \leqslant n$, et une base de* L *telle que l'expression de f par rapport à cette base soit*

$$f(x, y) = \xi_1\bar{\eta}_1 + \cdots + \xi_p\bar{\eta}_p - \xi_{p+1}\bar{\eta}_{p+1} - \cdots - \xi_{p+q}\bar{\eta}_{p+q};$$

f est non dégénérée si et seulement si $n = p + q$.

La démonstration est identique à celle du Corollaire précédent, compte tenu du fait que $f(x, x)$ est réel quel que soit $x \in$ L comme on l'a vu au n° 1, relation (4).

On peut démontrer (*Loi d'inertie*) que les entiers p et q figurant dans ces énoncés ne dépendent que de f, et non du choix de la base; cf. *Exercice* 24.

6. *Bases orthonormales*

Soit f une forme hermitienne sur un espace vectoriel L de dimension finie sur K. On dit qu'une base (a_i) de L est **orthonormale** par rapport à f si l'on a

$$f(a_i, a_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j, \end{cases}$$

autrement dit si la matrice de f par rapport à la base en question est la matrice unité, ou enfin si l'expression de f par rapport à cette base est

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \xi_i\eta_i^*$$

où $n = \dim(\mathrm{L})$.

L'existence d'une base orthonormale exige évidemment que f soit *non dégénérée*, et le Corollaire 1 du Théorème 6 montre que cette condition est aussi *suffisante dans le cas orthogonal complexe* par exemple (plus généralement, si K est algébrique-

ment clos, de caractéristique $\neq 2$, et si l'involution choisie sur K est l'identité).

Plaçons-nous maintenant dans le cas *orthogonal réel* ou dans le cas *hermitien complexe*. Si f admet une base orthonormale, on a pour tout $x \in \mathrm{L}$

$$f(x, x) = \sum_{1 \leqslant i \leqslant n} \xi \xi_i^* = \sum_{1 \leqslant i \leqslant n} |\xi_i|^2$$

et par suite

$$f(x, x) > 0 \quad \text{pour tout } x \neq 0.$$

Une forme satisfaisant à cette condition (dans le cas orthogonal réel ou hermitien complexe — on notera qu'alors $f(x, x)$ est toujours réel) est dite **définie positive.** Si inversement f est définie positive, la démonstration du Corollaire 2 montre qu'on a $p = n$ et par suite que L admet une base orthonormale. Donc :

THÉORÈME 7. *Soient* L *un espace vectoriel réel (resp. complexe) de dimension finie et f une forme bilinéaire symétrique (resp. sesquilinéaire hermitienne) sur* L. *Pour que* L *admette une base orthonormale relativement à f, il faut et il suffit que f soit définie positive.*

Exemple 12. La forme $(x|y)$ de l'*Exemple* 7 est évidemment définie positive; une base orthonormale pour cette forme n'est autre qu'un système de trois vecteurs orthogonaux et de longueur 1; une telle base s'appelle aussi un **système de coordonnées rectangulaires** dans l'espace usuel.

Remarque 7. Si (a_i) est une base orthonormale pour f, et si

$$x = \sum \xi_i a_i \in \mathrm{L},$$

alors les coordonnées de x sont données par la relation

$$\xi_i = f(x, a_i).$$

On a en effet

$$f(x, a_i) = \sum \xi_j f(a_j, a_i) = \xi_i f(a_i, a_i) = \xi_i.$$

Remarque 8. Soient L un espace vectoriel réel (resp. complexe) de dimension n et f une forme bilinéaire symétrique (resp. sesquilinéaire hermitienne) sur L; il existe donc une base de L par rapport à laquelle l'expression de f est

$$f(x, y) = \xi_1 \overline{\eta}_1 + \cdots + \xi_p \overline{\eta}_p - \xi_{p+1} \overline{\eta}_{p+1} - \cdots - \xi_{p+q} \overline{\eta}_{p+q}$$

avec des entiers $p, q \geqslant 0$ tels que $p + q \leqslant n$.

Pour que f soit *définie positive* il faut et il suffit que $p = n$, $q = 0$. L'autre cas extrême, celui où $p = 0$, $q = n$, est celui des formes **définies négatives,, i.e.** telles que l'on ait

$$f(x, x) < 0 \quad \text{pour tout } x \neq 0.$$

On dit plus généralement que f est **positive** (resp. **négative** si l'on a $f(x, x) \geqslant 0$ (resp. $f(x, x) \leqslant 0$) pour tout x; cela signifie évidemment qu'on a $q = 0$ (resp. $p = 0$), en sorte que les formes définies positives (resp. définies négatives) sont les formes positives (resp. négatives) *non dégénérées.*

Une forme qui n'est ni positive ni négative est dite **indéfinie**; cela signifie donc qu'on a $p \geqslant 1$ et $q \geqslant 1$, ou encore qu'on peut trouver des vecteurs x et y tels que

$$f(x, x) > 0, \qquad f(y, y) < 0$$

C'est par exemple le cas de la forme de Lorentz.

7. *Automorphismes d'une forme hermitienne*

Soient L un espace vectoriel sur K et f une forme hermitienne non dégénérée sur L. On appelle **automorphisme de** f tout automorphisme u de l'espace vectoriel L tel que

$$f(u(x), u(y)) = f(x, y) \quad \text{quels que soient } x, y \in \mathrm{L}.$$

Comme on a

$$f(u(x), u(y)) = f(x, u^*(u(y)))$$

où u^* est l'adjoint de u par rapport à f (n° 3), on voit que les automorphismes de f ne sont autres que les automorphismes de L vérifiant la relation

$$u^* \circ u = j_{\mathrm{L}},$$

application identique. De là et des relations

$$(j_{\mathrm{L}})^* = j_{\mathrm{L}}, \qquad (v \circ u)^* = u^* \circ v^*, \qquad (u^{-1})^* = (u^*)^{-1}$$

on déduit aussitôt que les automorphismes de f forment un *sous-groupe* du groupe GL(L) des automorphismes de L; ce sous-groupe se note

$$\mathrm{GL}(f)$$

et s'appelle le **groupe des automorphismes** de f.

Que u soit ou non un automorphisme de f, la fonction

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

est une forme hermitienne sur L. Pour exprimer que u est un automorphisme de f, i.e. que $g = f$, il suffit d'écrire que les coefficients de f et g par rapport à une base donnée $(a_i)_{1 \leqslant i \leqslant n}$ de L sont les mêmes (on suppose bien entendu L de dimension finie). Par suite les automorphismes de f sont caractérisés par le fait que l'on a

$$f(u(a_i), u(a_j)) = f(a_i, a_j)$$

quels que soient i et j. En particulier, *s'il existe une base de* L *orthonormale relativement à* f, *pour que* u *soit un automorphisme de* f *il faut et il suffit que* u *la transforme en une autre base orthonormale relativement à* f.

Lorsque l'involution de K est l'identité, et que l'on prend pour f la forme bilinéaire

$$f(x, y) = \sum \xi_i \eta_i$$

sur K^n, le groupe $GL(f)$ se note

$$O(n, K)$$

et s'appelle le **groupe orthogonal à n variables sur le corps K**; dans ce cas, si A est la matrice d'un automorphisme u de L, la matrice de u^* n'est autre comme on l'a vu au n° 3 que ${}^tA$, transposée de A; donc $O(n, K)$ *est le sous-groupe de* $GL(n, K)$ *formé des matrices* $A \in M_n(K)$ *telles que*

$${}^tA.A = 1_n;$$

une telle matrice est dite **orthogonale.** On notera que la relation précédente implique

$$1 = \text{dét } ({}^tA) \text{ dét } (A) = \text{dét } (A)^2$$

et par suite

$$\text{dét } (A) = +1 \quad \text{ou} \quad -1;$$

les matrices orthogonales de déterminant $+1$ forment un sous-groupe de $O(n, K)$, qu'on note

$$O^+(n, K) \quad \text{ou} \quad SO(n, K),$$

et qu'on appelle **le groupe des rotations à n variables sur le corps K.**

Si $K = R$ (resp. $K = C$) on parle du **groupe orthogonal réel** (resp. **complexe**) et du **groupe des rotations réel** (resp. **complexe**).

Dans le cas d'un corps K et d'une involution quelconques, et où l'on prend pour f la forme **hermitienne**

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \xi_i \eta_i^*$$

sur K^n, le groupe $GL(f)$ se note

$$U(n, K)$$

et s'appelle le **groupe unitaire à n variables sur le corps** K relativement à l'involution considérée sur K; c'est donc aussi le groupe des matrices $A \in M_n(K)$ telles que

$$A^*A = 1_n$$

(une telle matrice est dite **unitaire** relativement à l'involution considérée). En particulier, pour $K = C$, l'involution étant $\lambda \to \bar{\lambda}$, on obtient le **groupe unitaire complexe à n variables**; il est formé des matrices complexes **unitaires**, i.e. telles que

$${}^t\bar{A}.A = 1_n.$$

Exemple 13. Prenons pour L l'espace usuel et pour f la forme $(x|y)$ de l'*Exemple* 7; alors le groupe des rotations correspondant est formé des rotations (au sens usuel) autour de l'origine, i.e. des déplacements laissant fixe le point O.

Exemple 14. On appelle **groupe de Lorentz** le groupe des automorphismes de la forme de Lorentz — ou, plus précisément, le sous-groupe du groupe des automorphismes de la forme de Lorentz f formé par ceux de ces automorphismes

qui sont de déterminant + 1 et qui appliquent dans lui-même l'ensemble des vecteurs $(x, y, z, t) \in \mathbf{R}^4$ tels que $t > 0$. Il joue un rôle fondamental en Physique.

8. *Automorphismes d'une forme hermitienne positive : réduction à la forme diagonale*

Nous allons démontrer le résultat suivant :

THÉORÈME 8. *Soient* L *un espace vectoriel complexe de dimension finie, f une forme hermitienne définie positive sur* L, *et u un automorphisme de f. Il existe alors des vecteurs propres de u formant une base orthonormale de* L *par rapport à f.*

Désignant d'une manière générale par u^* l'adjoint relativement à f d'un endomorphisme u de L, les automorphismes de f sont caractérisés par la relation

$$u^* = u^{-1}$$

et par suite satisfont à la condition

$$u^* \circ u = u \circ u^*;$$

un endomorphisme u de L qui la vérifie est dit **normal** relativement à f. Ceci dit, le Théorème 8 est évidemment un cas particulier du suivant :

THÉORÈME 9. *Soient* L *un espace vectoriel complexe de dimension finie, f une forme hermitienne définie positive sur* L, *et u un endomorphisme de* L *normal relativement à f. Il existe alors une base de* L *formée de vecteurs propres de u et orthonormale par rapport à f.*

La démonstration de ce Théorème repose sur plusieurs lemmes.

LEMME 1. *Supposons u normal; alors, pour tout $x \in$ L, la relation $u(x) = 0$ implique $u^*(x) = 0$* autrement dit on a $\mathrm{Ker}\,(u) = \mathrm{Ker}\,(u^*)$.

Notons d'abord que, quels que soient $x, y \in$ L, on a

$$f(u(x), u(y)) = f(u^* \circ u(x), y) = f(u \circ u^*(x), y) = f(u^*(x), u^*(y));$$

par suite

$$f(u(x), u(x)) = f(u^*(x), u^*(x));$$

si $u(x) = 0$, le second membre est donc nul, ce qui montre que $u^*(x) = 0$ puisque f est définie positive.

LEMME 2. *Supposons u normal et soit λ une valeur propre de u; alors $\overline{\lambda}$ est une valeur propre de u^*, et pour $x \in$ L la relation $u(x) = \lambda x$ équivaut à la relation $u^*(x) = \overline{\lambda} x$.*

Posons $v = u - \lambda . j_{\mathrm{L}}$; comme λ commute à tout endomorphisme de L, un calcul trivial montre que v est normal comme u; comme $v^* = u^* - \overline{\lambda} . j_{\mathrm{L}}$, le Lemme 2 s'obtient alors en appliquant le Lemme 1 à v.

Dans ce qui suit on désigne par $\lambda_1, \ldots, \lambda_r$ les diverses valeurs propres de u, et par L_i $(1 \leqslant i \leqslant r)$ le sous-espace de L formé des solutions de $u(x) = \lambda_i x$.

LEMME 3. *Supposons u normal; alors les sous-espaces propres $L_1, \ldots, L_r$ sont deux à deux orthogonaux pour f.*

Soient en effet $x \in L_i$ et $y \in L_j$; on a, d'après le lemme 2,

$$\lambda_i . f(x, y) = f(u(x), y) = f(x, u^*(y)) = f(x, \lambda_j^* y) = \lambda_j f(x, y),$$

et puisque $\lambda_i \neq \lambda_j$ pour $i \neq j$ on voit que dans ce cas on trouve $f(x, y) = 0$, d'où le Lemme.

LEMME 4. *Soit M un sous-espace de L stable par u et par u^*; alors l'orthogonal $M^\perp$ de M relativement à f est aussi stable par u et u^*.*

Il suffit de montrer que, pour $x \in L$, la relation $f(x, y) = 0$ pour tout $y \in M$ implique la même propriété pour $u(x)$ et $u^*(x)$; or on a $f(u(x), y) = f(x, u^*(y))$, et comme par hypothèse la relation $y \in M$ implique $u^*(y) \in M$, le premier point est évident; le second se démontre de même.

Pour achever la démonstration du Théorème 9, considérons le sous-espace

$$M = L_1 + \cdots + L_r$$

de L engendré par les vecteurs propres de u. D'après le lemme 2, il est stable par u et u^*; il en est donc de même de $M^\perp$; si l'on avait $M^\perp \neq 0$, l'endomorphisme u aurait au moins un vecteur propre dans $M^\perp$, et l'on aurait donc

$$M \cap M^\perp \neq 0$$

contrairement au fait que f est définie positive; on a donc $M^\perp = \{0\}$, et comme L est somme directe de M et de $M^\perp$ (Corollaire du Théorème 5) il vient

$$L = M.$$

Ainsi L est somme des L_i, et même somme *directe* d'après le Théorème 4 du § 34.

Ceci fait, et en vertu du Lemme 3, il suffit, pour construire une base orthonormale de L formée de valeurs propres de u, de choisir dans chaque L_i une base orthonormale pour la restriction de f à L_i ce qui est possible (Théorème 7) en vertu du fait évident que la restriction d'une forme définie positive à un sous-espace est encore définie positive. Ceci achève la démonstration des Théorèmes 8 et 9.

Lorsque u est un *automorphisme* de f, le lemme 2 montre que $u(x) = \lambda x$ implique

$$u^{-1}(x) = \overline{\lambda} x,$$

d'où résulte que chaque valeur propre λ de u vérifie

$$\lambda^{-1} = \overline{\lambda},$$

i.e. est un *nombre complexe de module égal à* 1.

Un autre cas particulier important est celui d'un endomorphisme u autoadjoint pour f, i.e. tel que

$$u^* = u;$$

le lemme 2 montre alors que

$$\lambda = \bar{\lambda},$$

autrement dit que *toute valeur propre de u est réelle.*

Les résultats précédents peuvent évidemment s'interpréter en langage de matrices. Prenons $L = \mathbf{C}^n$ et

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \xi_i \overline{\eta_i}$$

ce qui ne restreint d'ailleurs pas la généralité en vertu du Théorème 7. Soit A la matrice de u par rapport à la base canonique, et soit U la matrice de passage de la base canonique à une base orthonormale formée de vecteurs propres de u; la matrice de u par rapport à celle-ci est alors $U^{-1}AU$ (§ 15, Corollaire du Théorème 2), en sorte que

$$A = UDU^{-1}$$

où D est diagonale; mais comme U fait passer de la base canonique (qui est *orthonormale* par rapport à la forme considérée) à une autre base *orthonormale*, la matrice U est *unitaire* comme on l'a vu au n° 7. Donc :

Corollaire du Théorème 9. *Soit* A *une matrice carrée d'ordre n à coefficients complexes telle que*

$$A^*A = AA^*.$$

Il existe alors une matrice unitaire complexe U *d'ordre n et une matrice diagonale* D *d'ordre n telles que l'on ait*

$$A = UDU^{-1};$$

pour que A *soit unitaire (resp. hermitienne) il faut et il suffit que les termes diagonaux de* D *soient de module un (resp. réels).*

En ce qui concerne la dernière assertion, on observe que de $A = UDU^{-1}$ résulte

$$A^* = (U^{-1})^*D^*U^* = UD^*U^{-1} = U\overline{D}U^{-1}$$

où $\overline{D}$ est la matrice imaginaire conjuguée de D; pour que A soit hermitienne il est donc nécessaire et suffisant que

$$UDU^{-1} = U\overline{D}U^{-1}$$

i.e. que $D = \overline{D}$, i.e. que D soit réelle; et pour que A soit unitaire, il est nécessaire et suffisant que

$$1_n = A^*A = U\overline{D}U^{-1}UDU^{-1} = U\overline{D}DU^{-1}$$

i.e. que $\overline{D}.D = 1_n$, ce qui signifie évidemment que les termes diagonaux de D sont de valeur absolue égale à 1.

Une matrice $A \in M_n(\mathbf{C})$ telle que $A^*A = AA^*$ est dite **normale**; une telle matrice est donc diagonalisable.

Le Corollaire ci-dessus montre naturellement que les valeurs propres d'une ma-

trice *hermitienne* (donc aussi d'une matrice *symétrique réelle*) sont toutes *réelles*. Ce résultat implique le suivant :

THÉORÈME 10. *Soient* L *un espace vectoriel réel de dimension finie, f une forme bilinéaire symétrie définie positive sur* L, *et u un endomorphisme de* L *autoadjoint pour f. Il existe alors une base de* L *orthonormale pour f et composée de vecteurs propres de u.*

On procède comme dans la démonstration du Théorème 9; les lemmes 1 et 2 sont ici triviaux puisque $u^* = u$, et les lemmes 3 et 4 sont encore valables avec les mêmes démonstrations. Formant comme plus haut le sous-espace M de L engendré par les vecteurs propres de u, tout revient à montrer que $M = L$, i.e. (Corollaire du Théorème 5) que $M^\perp = 0$, et pour cela tout revient, comme plus haut, à montrer que *tout sous-espace* N *de* L *stable par u et non nul contient au moins un vecteur propre de u.* En remplaçant f par sa restriction à N (laquelle est encore symétrique et définie positive), et u par sa restriction à N (laquelle est encore autoadjointe relativement à la restriction de f à N), on est finalement ramené à montrer que, dans les hypothèses de l'énoncé, u possède au moins un vecteur propre dans L, i.e. admet au moins une valeur propre *réelle*.

Or soit (a_i) une base de L orthonormale pour f (Théorème 7); la matrice A de u par rapport à cette base est *symétrique réelle* comme il résulte des calculs du n° 3, et toutes ses valeurs propres sont donc réelles (Corollaire du Théorème 9). Comme les valeurs propres de u sont aussi celles de A, la démonstration est achevée.

COROLLAIRE. *Soit* A *une matrice symétrique réelle; il existe une matrice orthogonale réelle* U *et une matrice diagonale réelle* D *telles que*

$$A = UDU^{-1}.$$

Il suffit pour le voir de considérer sur $\mathbf{R}^n$ la forme bilinéaire

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \xi_i \eta_i$$

et l'endomorphisme u admettant A pour matrice par rapport à la base canonique; celui-ci est autoadjoint pour f, et on peut alors prendre pour U la matrice de passage de la base canonique à une base de $\mathbf{R}^n$ composée de vecteurs propres de u et orthonormale par rapport à f.

Remarque 9. Dans les ouvrages classiques de Mathématiques Spéciales (sic), le fait qu'une matrice symétrique réelle A d'ordre $n = 3$ ait toutes ses valeurs propres réelles est fréquemment démontré en invoquant le fait que ces valeurs propres sont racines d'une équation algébrique à coefficients réels de degré 3, et qu'une telle équation (plus généralement, une équation algébrique de degré *impair* sur R) admet toujours au moins une racine réelle.

L'inconvénient majeur de cette démonstration est évidemment qu'on ne peut pas l'étendre aux matrices symétriques réelles d'ordre 4 (ni même, à vrai dire, à celles d'ordre 2 ...), attendu qu'une équation algébrique réelle de degré pair peut fort bien ne posséder aucune racine réelle.

Une démonstration fort élémentaire du résultat général s'obtient comme suit. Soit (α_{ij}) une matrice symétrique réelle (ou même hermitienne complexe) et soit $\lambda \in \mathbf{C}$ une valeur propre de celle-ci; alors le système d'équations linéaires et homogènes

$$\sum_i \alpha_{ij}\xi_i = \lambda\xi_j$$

possède dans $\mathbf{C}^n$ une solution *non triviale*; pour cette solution on a

$$\sum_{i,j} \alpha_{ij}\xi_i\bar{\xi}_j = \lambda \sum_j \xi_j\bar{\xi}_j;$$

or le premier membre est réel puisque (α_{ij}) est hermitienne, et d'autre part l'expression

$$\sum \xi_j\bar{\xi}_j = \sum |\xi_j|^2$$

est réelle et même > 0 puisque les ξ_j ne sont pas tous nuls; la valeur propre λ est donc nécessairement réelle.

Le raisonnement « géométrique » correspondant consiste à écrire, avec les notations du Théorème 9, que

$$\lambda . f(x, x) = f(\lambda x, x) = f(u(x), x) = f(x, u(x)) = f(x, \lambda x) = \bar{\lambda} . f(x, x)$$

et à observer que $f(x, x) \neq 0$ puisque f est définie positive, d'où résulte que

$$\lambda = \bar{\lambda}.$$

9. *Vecteurs isotropes et formes indéfinies*

Dans le cas orthogonal réel ou hermitien complexe, une forme hermitienne f sur un espace vectoriel L est dite définie si elle soit définie positive, soit définie négative, autrement dit si, pour $x \in \mathrm{L}$, l'expression $f(x, x)$ garde un signe constant et ne devient nulle qu'en $x = 0$.

Une forme définie ne possède évidemment aucun vecteur isotrope non nul; en fait, cette propriété caractérise même les formes définies; autrement dit, *pour que f soit définie il faut et il suffit que la relation $f(x, x) = 0$ implique $x = 0$.*

Il revient au même de montrer qu'une forme f non définie (ce qui ne veut pas dire indéfinie...) admet des vecteurs isotropes non nuls. Or puisque f n'est pas définie, il existe des vecteurs non nuls $a, b \in \mathrm{L}$ tels que l'on ait

$$f(a, a) \geqslant 0, \qquad f(b, b) \leqslant 0,$$

et comme il n'y a rien à démontrer si a ou b est isotrope on peut même supposer que

$$f(a, a) > 0, \qquad f(b, b) < 0.$$

On va alors montrer qu'il existe un scalaire λ tel que $x = a\lambda + b$ soit isotrope et non nul.

Le dernier point est clair, car si $b + \lambda a = 0$ alors

$$f(b, b) = f(-\lambda a, -\lambda a) = |\lambda|^2 . f(a, a)$$

a le même signe que $f(a, a)$, contrairement à l'hypothèse. Tout revient donc à prouver l'existence d'un λ tel que

$$0 = f(a\lambda + b, a\lambda + b) = f(a\lambda, a\lambda) + f(a\lambda, b) + f(b, a\lambda) + f(b, b)$$
$$= \lambda\bar{\lambda}u + v\lambda + \bar{v}\bar{\lambda} + w$$

où l'on a posé

$$u = f(a, a), \qquad v = f(a, b), \qquad w = f(b, b).$$

Or on a visiblement (cf. la réduction d'un trinôme du second degré à une somme de carrés)

$$u\lambda\bar{\lambda} + v\lambda + \bar{v}\bar{\lambda} + w = u\left[\left(\lambda + \frac{\bar{v}}{u}\right)\left(\bar{\lambda} + \frac{v}{u}\right) - \frac{\bar{v}v - uw}{u^2}\right]$$
$$= u\left[\left|\lambda + \frac{\bar{v}}{u}\right|^2 - \frac{|v|^2 - uw}{u^2}\right];$$

pour que cette équation en λ possède une solution dans le corps de base K (qui est R dans le cas orthogonal réel, et C dans le cas orthogonal complexe) il faut et il suffit que

$$|v|^2 - uw \geqslant 0;$$

comme $u = f(a, a)$ est positif et $w = f(b, b)$ négatif, cette condition est vérifiée, et notre assertion est établie.

Remarque 10. On pourrait aussi, dans la démonstration précédente, remplacer L par le sous-espace engendré par a et b, et appliquer le Corollaire 2 ou le Corollaire 3 du Théorème 6 (avec les notations de ces corollaires, il est *évident* qu'il existe des vecteurs isotropes non nuls en dehors des deux cas $p = n$, $q = 0$ et $p = 0$, $q = n$, lesquels caractérisent justement les formes définies).

10. L'inégalité de Cauchy-Schwarz

Pour terminer ce §, nous allons établir un résultat fort utile en Analyse, malgré sa simplicité :

THÉORÈME 11. *Soient* L *un espace vectoriel réel (resp. complexe) et* f *une forme bilinéaire symétrique (resp. sesquilinéaire hermitienne) sur* L. *On suppose*

$$f(x, x) \geqslant 0 \quad \text{pour tout } x \in \mathrm{L}.$$

On a alors

$$|f(x, y)|^2 \leqslant f(x, x) f(y, y)$$

quels que soient $x, y \in \mathrm{L}$.

En effet, pour tout scalaire λ, l'expression

$$f(x\lambda + y, x\lambda + y) = u\lambda\bar{\lambda} + v\lambda + \bar{v}\bar{\lambda} + w, \tag{21}$$

où

$$u = f(x, x), \qquad v = f(x, y), \qquad w = f(y, y),$$

est positive; de là et du fait que u et w sont positifs on doit déduire que

$$|v|^2 - uw \leqslant 0.$$

Si $u = 0$, il est clair que l'expression (21) ne peut être constamment positive que si $v = 0$ (car si $v \neq 0$, on peut toujours choisir λ de façon que $v\lambda$ soit un scalaire arbitraire), en sorte que l'inégalité à établir est évidente dans ce cas.

Si $u \neq 0$, on peut par exemple exprimer que (21) est positive pour

$$\lambda = -\bar{v}/u;$$

la valeur de (21) pour ce choix de λ est

$$-\frac{|v|^2 - uw}{u},$$

et comme $u > 0$ le fait que ce résultat soit positif établit évidemment le Théorème.

COROLLAIRE 1. *Supposons $f(x, x) \geqslant 0$ pour tout x. Pour que f soit non dégénérée il faut et il suffit que f soit définie positive.*

En effet le Théorème 11 montre que la relation $f(x, x) = 0$ implique $f(x, y) = 0$ pour tout $y \in \mathbf{L}$, donc $x = 0$ si f est non dégénérée.

Exemple 13. Prenons $\mathbf{L} = \mathbf{C}^n$ et

$$f(x, y) = \sum_{1 \leqslant i \leqslant n} \xi_i \overline{\eta_i};$$

on trouve alors l'inégalité

$$|\xi_1\eta_1 + \cdots + \xi_n\eta_n| \leqslant \sqrt{|\xi_1|^2 + \cdots + |\xi_n|^2}\cdot\sqrt{|\eta_1|^2 + \cdots + |\eta_n|^2}$$

valable quels que soient les nombres complexes ξ_i, η_i.

Exemple 14. Prenons la forme

$$f(x, y) = \int_0^1 x(t)\,\overline{y(t)}\,dt$$

de l'*Exemple* 9; on trouve alors l'inégalité

$$\left|\int_0^1 x(t)\,\overline{y(t)}\,dt\right| \leqslant \sqrt{\int_0^1 |x(t)|^2\,dt}\cdot\sqrt{\int_0^1 |y(t)|^2\,dt},$$

fréquemment utilisée en Analyse.

Remarque 11. Dans le cas du produit scalaire $(x|y)$ dans l'espace usuel, le Théorème 11 signifie qu'en valeur absolue le produit scalaire de x et y est inférieur ou égal au produit des longueurs de x et y; l'explication géométrique de ce fait est évidente, puisque $(x|y)$ est égal au produit des longueurs de x et y et du cosinus de l'angle de x et y : or, en valeur absolue, le cosinus d'un angle est toujours inférieur ou égal à 1.

COROLLAIRE 2. *Supposons* $f(x, x) \geqslant 0$ *pour tout* x, *et posons*

$$\|x\| = \sqrt{f(x, x)};$$

on a alors

$$\|x + y\| \leqslant \|x\| + \|y\|$$

quels que soient $x, y \in \mathrm{L}$.

On a en effet

$$\|x + y\|^2 = f(x + y, x + y) = f(x, x) + f(x, y) + \overline{f(x, y)} + f(y, y)$$
$$= \|x\|^2 + \|y\|^2 + 2 \cdot \mathrm{Re}(f(x, y));$$

comme l'inégalité de Cauchy-Schwarz s'écrit encore

$$|f(x, y)| \leqslant \|x\| \cdot \|y\|,$$

il vient *à fortiori*

$$\mathrm{Re}(f(x, y)) \leqslant \|x\| \cdot \|y\|$$

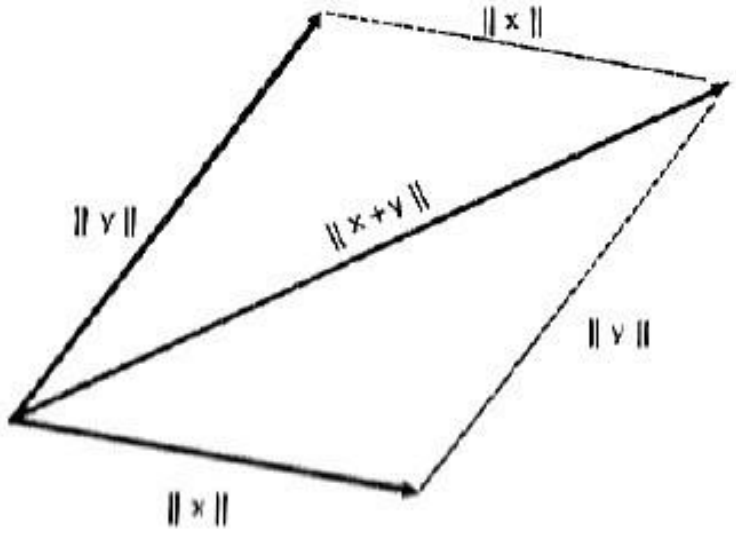
et par suite

$$\|x + y\|^2 \leqslant \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2,$$

ce qui conduit aussitôt au résultat cherché.

Dans le cas du produit scalaire usuel $(x|y)$, il est clair que $\|x\|$ n'est autre que la *longueur* du vecteur x. Le Corollaire 2 démontre donc le résultat suivant :

COROLLAIRE 3. *Dans un triangle, la longueur de chaque côté est inférieure à la somme des longueurs des deux autres.*



EXERCICES

Il est parfaitement utopique d'espérer apprendre des Mathématiques, si élémentaires ou si supérieures soient-elles, sans résoudre des Exercices.

Les Exercices qu'on trouvera dans ce livre sont de trois sortes. Certains sont des illustrations pratiques ou même numériques des théories exposées dans le texte; le lecteur débutant ne pourra pas acquérir la technique du calcul sans résoudre une partie appréciable des Exercices de ce genre. D'autres apportent au texte des compléments théoriques élémentaires; en les étudiant, le lecteur s'habituera à manipuler le langage et les modes de raisonnements utilisés dans le texte; ceux de ces Exercices qui ne sont pas *très* faciles sont précédés d'un signe ¶. Enfin, la dernière catégorie est constituée par des Exercices qui apportent au texte des compléments importants et difficiles; ils sont destinés uniquement aux étudiants déjà avancés qui s'intéressent vraiment aux Mathématiques; ces Exercices sont précédés de deux ou même trois signes ¶.

Nous ne saurions trop insister enfin sur le fait que résoudre un Exercice ne consiste pas seulement à se convaincre, à l'aide d'un « brouillon » fait à la hâte, du fait qu'on en a à peu près compris la solution; si cette méthode est admissible pour les Exercices de calcul numérique, il faut par contre s'efforcer de *rédiger intégralement* les Exercices plus théoriques, où l'on doit construire de véritables démonstrations. De cette façon, et uniquement de cette façon, l'étudiant parviendra à acquérir un langage clair et correct, et à utiliser les termes techniques dans leur sens propre, ce qui, en Mathématiques, est le signe le plus certain de la compréhension d'un sujet.

1. Soient R et S deux relations. Montrer que, si R est fausse, la relation $(R \Longrightarrow S)$ est vraie. Peut-on déduire de là que S est vraie?

2. Soient R et S deux relations. Montrer que la relation

$$[R \text{ et } (\text{non } R)] \Longrightarrow S$$

est vraie.

3. Montrer à l'aide d'un exemple que la relation

$$(\forall x)\ (\exists y)R \Longrightarrow (\exists y)\ (\forall x)R$$

n'est généralement pas vraie.

4. Soient R et S deux relations équivalentes, et T une relation quelconque. Montrer que chacune des relations suivantes est vraie :

$$\begin{gathered}(\text{non } R) \Longleftrightarrow (\text{non } S)\\ (R \Longrightarrow T) \Longleftrightarrow (S \Longrightarrow T)\\ (T \Longrightarrow R) \Longleftrightarrow (T \Longrightarrow S)\\ (R \text{ et } T) \Longleftrightarrow (S \text{ et } T)\\ (R \text{ ou } T) \Longleftrightarrow (S \text{ ou } T).\end{gathered}$$

qq 5. Démontrer les relations suivantes, où R, S et T désignent des relations quelconques :

$$\begin{gathered}R \Longrightarrow (S \Longrightarrow R)\\ (R \Longrightarrow S) \Longrightarrow [(S \Longrightarrow T) \Longrightarrow (R \Longrightarrow T)]\\ R \Longrightarrow [(\text{non } R) \Longrightarrow S]\\ (R \text{ ou } S) \Longleftrightarrow [(R \Longrightarrow S) \Longrightarrow S]\\ (R \Longleftrightarrow S) \Longleftrightarrow \{(R \text{ et } S) \text{ ou } [(\text{non } R) \text{ et } (\text{non } S)]\}\\ (R \Longleftrightarrow S) \Longleftrightarrow \text{non } [(\text{non } R) \Longleftrightarrow S]\\ \{R \Longrightarrow [S \text{ ou } (\text{non } T)]\} \Longleftrightarrow [(T \text{ et } R) \Longrightarrow S]\\ [R \Longrightarrow (S \text{ ou } T)] \Longleftrightarrow [S \text{ ou } (R \Longrightarrow T)\\ (R \Longrightarrow S) \Longrightarrow \{(R \Longrightarrow T) \Longrightarrow [R \Longrightarrow (S \text{ et } T)]\}\\ (R \Longrightarrow T) \Longrightarrow \{(S \Longrightarrow T) \Longrightarrow [(R \text{ ou } S) \Longrightarrow T]\},\\ (R \Longrightarrow S) \Longrightarrow [(R \text{ et } T) \Longrightarrow (S \text{ et } T)]\\ (R \Longrightarrow S) \Longrightarrow [(R \text{ ou } T) \Longrightarrow (S \text{ ou } T)].\end{gathered}$$

6. Soient R et S deux relations et x une lettre ne figurant pas dans R. Montrer que les relations

$$(\forall x)\ (\text{R ou S}) \Longleftrightarrow (\text{R ou } (\forall x)\text{S})$$
$$(\exists x)\ (\text{R et S}) \Longleftrightarrow (\text{R et } (\exists x)\text{S})$$

sont vraies.

7. Soient R et S deux relations, x une lettre. Démontrer les relations

$$[(\forall x)(\text{R ou S})] \Longrightarrow [(\forall x)\text{R ou } (\exists x)\text{S}]$$
$$[(\exists x)\text{R et } (\exists x)\text{S}] \Longrightarrow [(\exists x)\ (\text{R et S})].$$

8. Les cannibales d'une tribu se préparent à manger un missionnaire. Désirant lui prouver une dernière fois leur respect de la dignité et de la liberté humaines, les cannibales proposent au missionnaire de décider lui-même de son sort en faisant une courte déclaration : si celle-ci est vraie, le missionnaire sera roti, et il sera bouilli dans le cas contraire. Que doit dire le missionnaire pour sauver sa vie ? (d'après Cervantès).

9. Le Colonel X traite le Professeur Y d'assassin. Deux semaines plus tard, le Colonel est l'objet d'une tentative d'assassinat inspirée par le Professeur. Le Colonel avait-il raison?

10. Enoncer des assertions équivalentes aux négations des assertions suivantes (*) :

a) Tout triangle rectangle possède un angle droit.

b) Dans toutes les prisons tous les détenus détestent tous les gardiens.

c) Pour tout entier x il existe un entier y tel que pour tout entier z la relation $z < y$ implique la relation $z = x + 1$.

11. Examiner les relations logiques existant entre les assertions suivantes :

A : Tous les hommes sont mortels.
B : Tous les hommes sont immortels.
C : Aucun homme n'est mortel.
D : Aucun homme n'est immortel.
E : Il existe des hommes immortels.
F : Il existe des hommes mortels.

12. Montrer, à l'aide de l'opération de Hilbert, que si R est une relation et x une lettre figurant dans R, la lettre *x ne figure plus* dans les relations $(\forall x)$R et $(\exists x)$R, en dépit des notations utilisées pour désigner ces deux relations.
Ce résultat fort simple montre que, dans la notation $(\forall x)$R, la lettre x ne figure que pour indiquer une opération à effectuer sur la relation R, opération ayant pour résultat, entre autres, d'éliminer x de la relation R. Un phénomène analogue se retrouve dans la notation traditionnelle

$$\int_0^{+1} f(x)dx,$$

où la lettre x ne joue évidemment aucun rôle et, en particulier, ne figure pas dans le résultat final.

(*) La façon la plus simple (et pour cause) d'écrire la négation d'une relation est de faire précéder celle-ci d'un signe « non ». Ce n'est évidemment pas ce qu'on demande au lecteur de faire dans cet *Exercice*...

Cet Exercice explique aussi pourquoi, dans la relation $(\forall x)R$, on peut si on le désire remplacer la lettre x par toute autre lettre ne figurant pas dans R; par exemple, les relations

$$(\forall x)\ (x \times y = x - z) \qquad \text{et} \qquad (\forall t)\ (t \times y = t - z)$$

sont non seulement équivalentes mais en fait identiques. Il n'en est par contre pas de même des relations

$$(\forall x)\ (x \times y = x - z) \qquad \text{et} \qquad (\forall y)\ (y \times y = y - z).]$$

¶ **13.** Sur la planète Mars, on distingue en première approximation deux sortes d'opinions politiques : celles de droite et celles de gauche. D'autre part, les étudiants martiens se répartissent en deux associations : l'Union Planétaire des Étudiants Martiens (UPEM) et la Fédération Planétaire des Étudiants Martiens (FPEM). Sachant que les étudiants de gauche adhèrent à l'UPEM, démontrer que la FPEM est apolitique.

14. On considère quatre nombres entiers m, n, p, q sur lesquels on fait les hypothèses suivantes : *a*) les entiers m, p et q sont premiers entre eux; *b*) le reste de la division de m par pq est égal à 12; *c*) le reste de la division de $2n - 3$ par n est égal à 3; *d*) il n'existe aucun couple d'entiers x, y vérifiant la relation

$$x^4 - y^5 = p^2 - q^2 + m^6.$$

Démontrer que n est pair.

15. On considère les deux assertions suivantes :

a) : « En plein accord avec M. Robert Lacoste, ministre résidant en Algérie, nous confions la responsabilité de ramener la paix et la sécurité à Alger à la 10e division parachutiste. Cette unité gagnera en trois mois la bataille d'Alger sans tirer sur les immeubles avec des mitrailleuses lourdes, et sans qu'un seul avion français arrose de balles la Casbah. » (Extrait de la déclaration faite par le Général Salan à son procès).

b) : « The result was that the « battle of Algiers » became, for the paratroopers who fought it, and for France itself, a pyrrhic victory : it is estimated that out of the Kasbah's total population of 80 000 between 30 and 40 per cent of its active male population was, at one stage or another of the « battle », arrested for questioning and questioning came to involve the use of torture as a basic instrument, as a time-saving device to obtain quick results. » (Edward Behr, correspondant de *Time* à Alger, dans *The Algerian Problem*, W. W. Norton, New York, 1961).

Ces assertions sont-elles logiquement incompatibles? (On ne demande pas de décider si elles sont vraies ou fausses).

16. Utiliser la règle non (non A) $\Longleftrightarrow$ A pour simplifier la phrase suivante (extraite d'un compte rendu de match de football) :

« ... il ne se trouvera aucun sportif pour nier que le contraire n'eût été immérité... ».

Même question avec le texte suivant :

« Je vous envoie encore une note sur les examens. J'ai tenu à rappeler quelques principes fondamentaux. Je fais allusion à certaines « décisions » ou certains comportements qui sont encore, heureusement, peu nombreux. Mais, même s'ils restent peu nombreux et s'ils devaient être confirmés légalement, il ne manquera pas de gens pour dénier toute valeur à la qualité du travail que la très grande majorité des enseignants de la Faculté s'efforce de mener à bien. »

17. A la suite d'une représentation de *Pelléas et Mélisande*, un journaliste hésite entre les deux rédactions suivantes :

A) Jamais le rôle de Mélisande n'a été si bien chanté.

B) Jamais si jeune cantatrice, aux si beaux cheveux, n'a si bien chanté Mélisande.

Lequel de ces compliments est le plus fort ? (Expliciter non A et non B.)

¶¶ 1. La définition mathématique complète de l'ensemble vide, utilisant l'opération de Hilbert, est que $\varnothing$ désigne l'objet mathématique

$$\tau_X[(\forall x)\ (x \notin X)];$$

en déduire la définition de $\varnothing$ en langage formalisé (i.e. écrire $\varnothing$ sous la forme d'un assemblage ne comportant que des signes fondamentaux, à l'exclusion de toute abréviation).

¶¶ 2. Construire une démonstration logiquement complète du Théorème 5 (cf. *Remarque* 6).

3. Écrire tous les éléments de l'ensemble

$$\mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\mathfrak{P}(\varnothing)))).$$

4. Soient X et Y deux ensembles. Montrer que les relations $X \subset Y$ et $\mathfrak{P}(X) \subset \mathfrak{P}(Y)$ sont équivalentes.

¶ 5. Montrer qu'il n'existe aucun ensemble X pour lequel la relation

$$\mathfrak{P}(X) \subset X$$

soit vraie.

6. Soient X l'ensemble des nombres x tels que $0 \leqslant x < 1/100\,000\,000$, et Y l'ensemble des nombres y tels que $0 < y \leqslant 100\,000\,000$; démontrer que $X \subset Y$.

§ 2

Soient I l'ensemble des nombres réels θ tels que $0 \leqslant \theta \leqslant 2\pi$, et G l'ensemble des rotations autour d'un point donné O dans le plan. On considère l'application f de I dans G qui, à chaque nombre $\theta \in I$, associe la rotation d'angle θ autour du point O. L'application f est-elle surjective? injective? bijective? Que se passe-t-il lorsqu'on prend pour I l'ensemble des nombres réels θ tels que $0 < \theta \leqslant 2\pi$?

Soient X et Y des ensembles; pour qu'une partie G de $X \times Y$ soit le graphe d'une application de X dans Y, il faut et il suffit que l'application pr_1 de G dans X soit bijective.

Soient $f : X \to Y$ et $g : Y \to Z$ deux applications, et $h = g \circ f$ l'application composée. *a*) h est injective, f est injective; si de plus f est surjective, alors g est injective. *b*) si h est surjective, g est surjective; si en outre g est injective, alors f est surjective.

On considère des ensembles X, Y, Z et des applications

$$f : X \to Y, \qquad g : Y \to Z, \qquad h : Z \to X;$$

forme les applications composées

$$h \circ g \circ f, \qquad g \circ f \circ h, \qquad f \circ h \circ g$$

on suppose soit que deux d'entre elles sont injectives et la troisième surjective, soit que deux entre elles sont surjectives et la troisième injective. Montrer qu'alors f, g et h sont bijectives.

Soient X, Y, Z trois ensembles, E l'ensemble de toutes les applications de $X \times Y$ dans Z, F l'ensemble de toutes les applications de X dans l'ensemble

$$Z^Y$$

toutes les applications de Y dans Z. Construire une bijection de E sur F.

On appelle **correspondance entre deux ensembles** X **et** Y tout triplet

$$f = (G, X, Y) \quad \text{avec} \quad G \subset X \times Y;$$

tte notion généralise donc celle d'application de X dans Y [une correspondance entre X Y ou, comme on dit aussi, de X à Y, est aussi appelée fréquemment une « fonction multi-

forme non partout définie » pour des raisons qui apparaîtront plus bas; cette terminologie, utilisée jusqu'à une date très récente, présente l'inconvénient majeur de laisser supposer que la notion de correspondance est un cas particulier de celle de fonction, alors que c'est l'opposé qui est vrai.] L'ensemble G s'appelle le **graphe** de f. On dit que f est **définie** en un élément x de X si $x \in pr_1(G)$; il existe alors au moins un $y \in Y$ tel que $(x, y) \in G$, et on dit que x et y **se correspondent par** f (il peut naturellement arriver que f ne soit pas définie pour tous les $x \in X$, et que, si f est définie en x, il existe plusieurs $y \in Y$ qui correspondent à x par f; ce sont ces deux circonstances qui expliquent la terminologie « fonction multiforme non partout définie »).

a) Étudier (pour $X = Y = \mathbf{R}$, ensemble des nombres réels) les correspondances dont les graphes sont les ensembles définis par les équations suivantes :

$$xy = 1; \qquad axy + bx + cy + d = 0; \qquad x^2 + y^2 = 1; \qquad x = \sin y$$

(dans le second exemple, a, b, c, d sont des constantes réelles données). Dans chaque cas on déterminera les valeurs de x pour lesquelles la correspondance est définie, et les y qui correspondent à un tel x.

b) Soient Γ et Γ' deux cercles distincts dans un plan; on prend pour X et Y l'ensemble des points de Γ, et pour G l'ensemble des couples $(x, y) \in X \times X$ qui sont situés sur une même tangente à Γ'. Quels sont les points de Γ où f est définie? Combien de points correspondent-ils à un point où f est définie?

¶ **7.** Soit $f = (G, X, Y)$ une correspondance entre deux ensembles X et Y (cf. Exercice 6); pour toute partie A de X, on note $f(A)$ l'ensemble des $y \in Y$ qui correspondent par f à au moins un $x \in A$, et pour toute partie B de Y on note

$$f^{-1}(B)$$

l'ensemble des $x \in X$ tels qu'il corresponde à x au moins un $y \in B$. Démontrer les relations

$$A \subset f^{-1}(f(A)), \qquad B \supset f(f^{-1}(B)).$$

¶ **8.** Soient $f = (G, X, Y)$ et $g = (H, Y, Z)$ deux correspondances; on appelle **composée** de g et f la correspondance

$$g \circ f = (K, X, Z) = h$$

définie comme suit : on a $(x, z) \in K$ si et seulement s'il existe un $y \in Y$ tel que l'on ait $(x, y) \in G$ et $(y, z) \in H$. Montrer que cette définition généralise celle de composée de deux applications; étendre le Théorème 2 du § 2 aux correspondances. Étant donnée une partie A de X, a-t-on nécessairement la relation $h(A) = g(f(A))$?

Si $f = (G, X, Y)$ est une correspondance, on appelle correspondance **réciproque** de f la correspondance

$$f^{-1} = (G', Y, X)$$

où $G' \subset Y \times X$ est l'ensemble des couples (y, x) tels que $(x, y) \in G$. La correspondance composée $f^{-1} \circ f$ est-elle l'application identique de X dans X? Montrer que, si f est une application de X dans Y, pour que la correspondance réciproque soit elle-même une application de Y dans X il faut et il suffit que f soit bijective; la correspondance f^{-1} est alors identique à l'application réciproque de f.

Soient $f = (G, X, Y)$ et $g = (H, Y, Z)$ deux correspondances. La formule

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

est-elle correcte?

Soit f une application d'un ensemble X dans un ensemble Y. On suppose que la relation

$$f(x') \neq f(x'') \quad \text{implique} \quad x' \neq x'';$$

démontrer que f est injective.

§ 3

Soient $(A_i)_{i\in I}$ une famille de parties d'un ensemble X, et $(B_j)_{j\in J}$ une famille de parties ın ensemble Y. A l'aide des intersections et des réunions de ces familles, calculer l'inter- :tion et la réunion de la famille $(A_i \times B_j)_{i\in I,\ j\in J}$ de parties de $X \times Y$.

Soit $(A_i)_{i\in I}$ une famille de parties d'un ensemble X; on a donc $\mathfrak{F}(A_i) \subset \mathfrak{F}(X)$ pour tout I. Les relations

$$\mathfrak{F}\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} \mathfrak{F}(A_i)$$
$$\mathfrak{F}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i} \mathfrak{F}(A_i)$$

ıt-elles vraies?

Soient X et Y deux ensembles et $(X_i)_{i\in I}$ une famille de parties de X, ayant pour réunion nsemble X tout entier. On suppose donnée, pour chaque $i \in I$, une application f_i de X_i ns Y. Montrer que les deux conditions suivantes sont équivalentes : *a*) quels que soient $\in I$, on a $f_i(x) = f_j(x)$ pour tout $x \in X_i \cap X_j$; *b*) il existe une application f de X dans Y i, pour tout $i \in I$, coïncide avec f_i dans X_i. L'application f est alors unique (on dit qu'elle btient par **recollement** des f_i).

On appelle **recouvrement** d'un ensemble X toute famille $(U_i)_{i\in I}$ de parties de X ayant ur réunion l'ensemble X tout entier. Étant donnés deux recouvrements $(U_i)_{i\in I}$ et $(V_j)_{j\in J}$ X, on dit que le second est **plus fin** que le premier si, pour tout indice $j \in J$, il existe un lice $i \in I$ tel que l'on ait $V_j \subset U_i$.
ontrer qu'étant donnés deux recouvrements quelconques de X, il en existe un troisième ıs fin que les deux premiers.

On appelle **schéma simplicial** tout objet formé d'un ensemble K et d'un ensemble de ·ties *finies et non vides* de K (appelées les **simplexes** du schéma simplicial considéré) vérifiant :ondition suivante : toute partie non vide d'un simplexe de K est encore un simplexe de K. n schéma simplicial n'est donc pas seulement un ensemble K, c'est un couple formé de K l'un ensemble de parties de K; néanmoins, on désigne toujours un schéma simplicial par la me lettre que l'ensemble correspondant. Dans un schéma simplicial K, on appelle **simplexe dimension** n tout simplexe comportant $n + 1$ éléments; les simplexes de dimension 0 ppellent les **sommets** de K, ceux de dimension 1 les **arêtes** de K, ceux de dimension 2 les es de K, etc...]. La notion de schéma simplicial a été introduite en particulier pour étudier

les propriétés « topologiques » des polyèdres de dimension quelconque, et apparaît déj chez Euler, qui a démontré le résultat suivant : étant donnée dans l'espace usuel une surfac polyédrale, comportant a sommets, b arêtes et c faces, le nombre $a - b + c$ ne dépend pas de l façon dont on a décomposé cette surface en triangles. A partir d'une surface polyédra décomposée en triangles, on construit un schéma simplicial comme suit : l'ensemble K e l'ensemble des sommets du polyèdre considéré P; toute partie à un élément de K est u simplexe de K; une partie $\{a, b\}$ à deux éléments de K est une arête si le segment de droi joignant le point a au point b est une arête de P; enfin une partie $\{a, b, c\}$ à trois élémen de K est une face si le triangle de sommets a, b, c est une face du polyèdre P. La généralisatic à un nombre quelconque de dimensions est facile.]

a) Soient X un ensemble quelconque et $(U_i)_{i\in I}$ un recouvrement de X. On convient de di qu'une partie S de l'ensemble d'indices I est un simplexe si elle est finie, non vide, et si l'inte section

$$\bigcap_{i\in S} U_i$$

est non vide. Montrer que le couple formé par l'ensemble I et les simplexes qu'on vient définir est un schéma simplicial (appelé le **nerf** du recouvrement donné).

b) Soit K un schéma simplicial. On désigne par P(K) l'ensemble des fonctions f défir sur l'ensemble K, à valeurs réelles, et possédant les trois propriétés suivantes : l'ensemble d $x\in K$ tels que $f(x) \neq 0$ est un simplexe de K; on a $f(x) \geqslant 0$ pour tout $x\in K$; enfin, la somr

$$\sum_{x\in K} f(x)$$

des valeurs de f aux divers points de K (somme qui ne comporte qu'un nombre fini de term non nuls) est égale à 1. Enfin, pour tout $x\in K$, on note U_x l'ensemble des $f\in P(K)$ telles q $f(x) \neq 0$. Montrer que la famille $(U_x)_{x\in K}$ est un recouvrement de l'ensemble P(K), et q le nerf de ce recouvrement est précisément le schéma simplicial donné K.

[Si le schéma simplicial K est fini et comporte n éléments $x_1, \ldots, x_n$, on peut représen chaque $f\in P(K)$ par le point de $\mathbf{R}^n$ dont les coordonnées sont les n nombres $f(x_1), \ldots, f(x$ on obtient alors une bijection de P(K) sur un polyèdre de l'espace $\mathbf{R}^n$, polyèdre dont la « form dépend précisément des relations combinatoires existant entre les divers simplexes du schér simplicial donné. C'est l'opération fondamentale qui permet de transformer un problèr en apparence purement « qualitatif », l'étude de la « forme » des polyèdres, en un problèr purement algébrique, l'étude des schémas simpliciaux finis.

§ 4

1. On appelle **partition** d'un ensemble X toute famille $(A_i)_{i\in I}$ d'ensembles non vides, deux à deux disjoints, ayant l'ensemble X pour réunion. Étant donnée une telle partition, on considère la relation

$$\text{il existe un } i\in I \text{ tel que } x\in A_i \text{ et } y\in A_i$$

entre éléments x, y de X. Montrer que celle-ci est une relation d'équivalence, dont on construira les classes et l'ensemble quotient. Montrer que toute relation d'équivalence sur X peut s'obtenir de la façon précédente.

¶ **2.** Soient R et S des relations d'équivalence sur des ensembles X et Y. Si (x', y') et (x'', y'') sont des éléments de $X \times Y$, on désigne par $T\{(x', y'), (x'', y'')\}$ la conjonction des relations $R\{x', x''\}$ et $S\{y', y''\}$; montrer que T est une relation d'équivalence sur $X \times Y$. Construire le graphe de T en fonction des graphes de R et S. Définir une bijection « canonique » du quotient de $X \times Y$ par T sur l'ensemble produit $(X/R) \times (Y/S)$.
Montrer, en utilisant ces résultats, que le Théorème 3 du § 4 est un cas particulier du Théorème 2.

3. Étant donnés, dans un plan rapporté à deux axes de coordonnées rectangulaires, deux points P′ et P″ de coordonnées (x', y') et (x'', y'') respectivement, on note $R\{P', P''\}$ la relation $x'y' = x''y''$. Montrer que c'est une relation d'équivalence dans le plan, et en construire les classes d'équivalence.
On désigne maintenant par $S\{P', P''\}$ la relation

$$(x'y' = x''y'') \qquad \text{et} \qquad (x'x'' \geqslant 0).$$

Est-ce encore une relation d'équivalence ?

4. Soient A un ensemble et B une partie de A. On note $R\{X, Y\}$ la relation $X \cap B = Y \cap B$. Montrer que c'est une relation d'équivalence sur l'ensemble $\mathfrak{P}(A)$ et construire une bijection de l'ensemble $\mathfrak{P}(A)/R$ sur l'ensemble $\mathfrak{P}(B)$.

5. Construire les tables d'addition et de multiplication des entiers modulo 17.

6. Soit E l'espace usuel (considéré comme ensemble de points). On choisit un point O une fois pour toutes, et étant donnés des points P′, P″ on note $R\{P', P''\}$ la relation

$$\text{les points O, P′ et P″ sont alignés.}$$

Est-ce une relation d'équivalence sur E? On note E^* l'ensemble des points $P \in E$ autres que O, de sorte que $E^* = E - \{O\}$; montrer que R est une relation d'équivalence sur E^* et déterminer les classes d'équivalence correspondantes (l'ensemble quotient E^*/R s'appelle le **plan projectif**).

7. Soit X l'ensemble de toutes les applications de **R** dans **R** (fonctions d'une variable réelle t, définies quel que soit t, et à valeurs réelles). Étant donnés deux éléments x, y de X, on désigne par $R\{x, y\}$ la relation

$$\text{il existe un nombre } c > 0 \text{ tel que l'on ait } \quad x(t) = y(t) \quad \text{pour} \quad |t| < c.$$

Montrer que R est une relation d'équivalence sur X.

8. Soit X l'ensemble de toutes les applications de **R** dans **R**; on choisit dans ce qui suit un entier $n \geqslant 0$. Étant données des fonctions $x, y \in X$, on désigne par $R\{x, y\}$ la relation

$$\lim_{t=0} \frac{x(t) - y(t)}{t^n} = 0$$

(qu'on écrit habituellement sous la forme

$$x(t) - y(t) = o(t^n) \qquad \text{pour} \qquad t \to 0).$$

Montrer que R est une relation d'équivalence sur X.

9. Soient X et Y deux ensembles, R et S des relations d'équivalence sur X et Y, et f une application de X dans Y. On considère le diagramme

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ \downarrow p & & \downarrow q \\ X/R & & Y/S \end{array}$$

où p et q désignent les applications canoniques de X et Y sur leurs quotients. Montrer que les deux propriétés suivantes sont équivalentes : (i) il existe une application

$$\overline{f} : X/R \to Y/S$$

telle que

$$\overline{f} \circ p = q \circ f;$$

(ii) quels que soient $x', x'' \in X$, la relation

$$x' \equiv x'' \pmod{R} \qquad \text{implique} \qquad f(x') \equiv f(x'') \pmod{S}.$$

Montrer de plus que, si la condition (ii) est vérifiée, il existe une seule application $\overline{f}$ satisfaisant à (i).

Exemple : on prend $X = Y = \mathbf{Z}$, ensemble des entiers rationnels; on prend pour R la relation de congruence modulo r et pour S la relation de congruence modulo s (où r et s sont des entiers non nuls donnés); enfin on prend pour f l'application identique de X dans Y. Dans quel cas le résultat précédent s'applique-t-il?

10. Soit K un schéma simplicial (§ 3, *Exercice* 5); on suppose que toute partie à un élément de K soit un simplexe de K. Étant donnés deux éléments x et y de K, on désigne par $R\{x, y\}$

la relation suivante : il existe un entier $n \geqslant 0$ et des sommets

$$z_0 = x, \quad z_1, \ldots, z_n = y$$

de K tels que l'ensemble $\{z_i, z_{i+1}\}$ soit un simplexe de K pour tout i tel que $0 \leqslant i < n$. Montrer que R est une relation d'équivalence sur l'ensemble K. [Les classes modulo R s'appellent les **composantes connexes** du schéma simplicial K; on dit que K est **connexe s'il** possède une seule composante connexe.]

¶ **1.** Soient X un ensemble et f une application de X dans l'ensemble $\mathfrak{P}(X)$ des parties de X. On note A l'ensemble des $x \in X$ vérifiant la relation $x \notin f(x)$. Montrer qu'il n'existe aucun $x \in X$ tel que $A = f(x)$. En déduire qu'il n'existe aucune application *surjective* de X dans $\mathfrak{P}(X)$, et que par suite on a

$$x < 2^x$$

pour tout nombre cardinal x (G. Cantor).

¶ **2.** Soit $(f_n)_{n \in \mathbf{N}}$ une suite d'applications de l'ensemble **N** dans lui-même; on définit une application f de **N** dans **N** en posant

$$f(n) = f_n(n) + 1 \quad \text{pour tout } n \in \mathbf{N}.$$

Montrer qu'il n'existe aucun entier $p \in \mathbf{N}$ tel que $f = f_p$. En déduire que l'ensemble de toutes les applications de **N** dans **N** est non dénombrable (G. Cantor).

¶ **3.** La réunion d'une infinité dénombrable d'ensembles dénombrables est un ensemble dénombrable (s'inspirer de la méthode utilisée dans l'*Exemple* 1 du n° 5).

4. Soit I l'ensemble des nombres réels compris entre 0 et 1; on représentera chaque $x \in I$ par un développement décimal illimité (pouvant se terminer par une infinité de chiffres 0). Soit $(x_n)_{n \in \mathbf{N}}$ une suite d'éléments de I; on forme un nombre $x \in I$ comme suit : la n^e décimale de x est égale à 1 si la n^e décimale de x_n est différente de 1, et est égale à 2 si la n^e décimale de x_n est égale à 1. Montrer qu'on a $x \neq x_n$ pour tout n. En conclure que l'ensemble I (et *à fortiori* l'ensemble **R** de tous les nombres réels) est non dénombrable (G. Cantor).
Préciser l'analogie entre ce raisonnement et celui de l'*Exercice* 2 ci-dessus.

¶ **5.** Soient f une bijection d'un ensemble X sur une partie Y_1 d'un ensemble Y, et g une bijection de Y sur une partie X_1 de X; on se propose de montrer que X et Y sont équipotents (Bernstein). Pour cela, on définit des parties A_n de X et B_n de Y en posant

$$A_0 = X - X_1, \qquad B_1 = f(A_0), \qquad A_1 = g(B_1), \qquad B_2 = f(A_1), \qquad A_2 = g(B_2), \ldots$$

puis on définit une application h de X dans Y comme suit : étant donné un $x \in X$, on prend

$$h(x) = f(x) \quad \text{si} \quad x \in \bigcup_{n \in \mathbf{N}} A_n,$$

et dans le cas contraire (de sorte qu'alors $x \in X_1$) on prend

$$h(x) = g^{-1}(x).$$

Montrer que h est une bijection de X sur Y.

6. Soient E un ensemble fini à h éléments, et n un entier naturel. On considère les applications f de E dans **N** telles que la somme des h nombres $f(x)$, $x \in E$, soit au plus égale à n. Montrer que les applications f considérées sont en nombre égal à

$$\binom{n+h}{h}.$$

7. Montrer que, pour entier n, la somme des coefficients du binôme $\binom{n}{p}$ est égale à 2^n.

¶ **8.** Démontrer la relation

$$\binom{n}{0}\cdot\binom{n}{p} + \binom{n}{1}\cdot\binom{n-1}{p-1} + \binom{n}{2}\cdot\binom{n-2}{p-2} + \cdots + \binom{n}{p}\cdot\binom{n-p}{0} = 2^p\cdot\binom{n}{p}.$$

(On cherchera d'abord, dans un ensemble X à n éléments, combien il existe de parties à p éléments qui contiennent un ensemble à k éléments donné d'avance).

¶ **9.** Soient p et n des entiers tels que $1 \leqslant p \leqslant n$, et $S_{n,p}$ le nombre des applications surjectives de l'ensemble $\{1, 2, \ldots, n\}$ dans l'ensemble $\{1, 2, \ldots, p\}$. Montrer qu'on a

$$p^n = S_{n,p} + \binom{p}{1}S_{n,p-1} + \binom{p}{2}\cdot S_{n,p-2} + \cdots + \binom{p}{p-1}.$$

En déduire que

$$S_{n,p} = p^n - \binom{p}{1}\cdot(p-1)^n + \binom{p}{2}\cdot(p-2)^n - \cdots + (-1)^{p-1}\binom{p}{p-1}.$$

Simplifier ce résultat pour $p = 2, 3$.

¶¶¶ **10.** Soient E et F deux ensembles finis, et $x \mapsto A(x)$ une application de E dans l'ensemble des parties de F. Pour qu'il existe une *injection* f de E dans F vérifiant

$$f(x) \in A(x) \qquad \text{pour tout } x \in E,$$

il faut et il suffit qu'on ait

$$\text{Card}\left(\bigcup_{x\in H} A(x)\right) \geqslant \text{Card}\,(H)$$

pout toute partie H de E. (Ce résultat est généralement connu sous le nom de *lemme des mariages*).

11. En utilisant le fait que, dans tout ensemble (fini ou infini) d'entiers naturels, il existe un entier plus petit que tous les autres, démontrer les résultats classiques que voici :

a) Tout entier $n \geqslant 2$ possède au moins un diviseur premier (on rappelle qu'un **nombre premier** est un entier $p \geqslant 2$ n'admettant pas d'autres diviseurs positifs que 1 et p).

b) Étant donnés deux entiers naturels a et b, avec $b \geqslant 1$, il existe des entiers naturels q et r tels que

$$a = bq + r, \qquad r < b,$$

et les entiers q et r sont uniques (**division euclidienne,** ou division avec reste, de a par b).

12. Démontrer que l'ensemble des nombres premiers est infini.

13. L'ensemble des nombres premiers de la forme $4n - 1$ (resp. $6n - 1$) est infini.
[Ce résultat est un cas particulier du *théorème de la progression arithmétique* de Dirichlet, à savoir que si a et b sont des entiers premiers entre eux, il existe une infinité de nombres premiers de la forme $an - b$. La démonstration générale du théorème de Dirichlet, l'un des plus célèbres de toute la Théorie des Nombres, ne peut pas se faire à l'heure actuelle par des procédés purement arithmétiques; toutes les démonstrations connues (à commencer par celle de Dirichlet lui-même, qu'on n'a pas pu substantiellement simplifier) utilisent des méthodes appartenant à l'Analyse, ou qui s'en inspirent directement.]

14. (Numération de base q). On choisit un entier naturel $q \geqslant 2$.

a) Soit x un entier naturel non nul. Montrer qu'il existe un et un seul entier $n \geqslant 0$ tel que

$$q^n \leqslant x < q^{n+1},$$

puis un et un seul entier a_n vérifiant

$$0 \leqslant a_n \leqslant q - 1, \qquad a_n q^n \leqslant x < (a_n + 1)q^n.$$

b) Montrer que, pour tout entier naturel x, il existe une et une seule suite d'entiers $a_0, a_1, \ldots, a_r, \ldots$ vérifiant les conditions suivantes :

1) on a $0 \leqslant a_r \leqslant q - 1$ pour tout $r \geqslant 0$,

2) les entiers r tels que $a_r \neq 0$ sont en nombre fini,

3) on a

$$x = a_0 + a_1 q + a_2 q^2 + \cdots + a_r q^r + \cdots$$

[la dernière somme, qui comporte en apparence une infinité de termes, a un sens à cause de la condition (2) imposée aux a_r]. Montrer que l'entier n de la question *a*) est le plus grand entier tel que $a_n \neq 0$, et que le nombre a_n défini à la question *a*) est égal au nombre a_n de la question *b*). La suite

$$a_n a_{n-1} \cdots a_0$$

(il ne s'agit pas d'un produit !) s'appelle le **développement de x dans le système de numération de base q**.

c) Trouver le développement du nombre 718 dans le système de numération binaire (i.e. à base $q = 2$).

d) Soit x un nombre rationnel positif, non nécessairement entier. Montrer qu'il existe une et une seule suite de nombres entiers

$$\ldots, a_r, a_{r-1}, \ldots, a_0, a_{-1}, a_{-2}, \ldots$$

vérifiant les conditions suivantes :

1) on a $0 \leqslant a_r \leqslant q - 1$ pour tout $r \in \mathbf{Z}$.

2) les entiers *positifs* r tels que $a_r \neq 0$ sont en nombre fini.

3) pour tout $r \in \mathbf{Z}$, on a

$$a_r q^r + a_{r+1} q^{r+1} + \cdots \leqslant x < q^r + a_r q^r + a_{r+1} q^{r+1} + \cdots$$

On dit alors que

$$a_n a_{n-1} \ldots a_0, \; a_{-1} a_{-2} \ldots a_{-r} \ldots$$

(où n est le plus petit entier naturel tel que l'on ait $a_r = 0$ pour tout $r > n$) est le développement de x dans le système de numération de base q.

e) Pour qu'une famille d'entiers a_n ($n \in \mathbf{Z}$) vérifiant les conditions 1) et 2) de la question précédente constitue le développement d'un nombre rationnel dans le système de numération de base q, il faut et il suffit qu'il existe un entier rationnel r et un entier naturel $k \geqslant 1$ tels que l'on ait

$$a_{n-k} = a_n \qquad \text{pour tout } n \leqslant r$$

(périodicité des développements des nombres rationnels)

15. Le développement de la Recherche Spatiale purement pacifique lui ayant permis de réaliser quelques différences intéressantes, le Président-Directeur-Général de la *Société Anonyme pour l'Exploitation Financière de la Physique Purement Théorique*, Commandeur de la Légion d'Honneur, achète en Basse-Normandie un domaine de 200 hectares à raison de 8 000 F l'hectare. Inspiré par cet exemple, un ouvrier plombier-zingueur attaché à l'établissement en question, et gagnant 800 F par mois, décide de placer chaque année un dixième de son salaire en Bons du Trésor rapportant 4 % l'an. Combien d'années devra-t-il travailler avant d'être en mesure d'acquérir en Basse-Normandie un domaine de 200 hectares pour y finir paisiblement ses jours ? (On tiendra compte des intérêts composés mais on négligera l'effet des dévaluations possibles de la monnaie).

16. D'après le journal *Le Monde* du 21 Juillet 1954, les dépenses occasionnées par la guerre d'Indochine sont fournies par le tableau suivant (il s'agit de milliards d'anciens francs) :

1946 :	101,8	1949 :	177,3	1952 :	427,6
1947 :	131,3	1950 :	258,3	1953 :	403,5
1948 :	136,3	1951 :	321	1954 :	428.

Vérifier sur cet exemple les propriétés fondamentales de l'addition (associativité et commutativité).

17. En Novembre 1954, on comptait en Algérie 1 230 000 Européens et 8 300 000 indigènes. A la même date, l'Université d'Alger comptait 4 548 étudiants européens et 557 indigènes. Calculer, à une unité près par défaut, le rapport entre les chances d'un Européen et celles d'un indigène d'accéder à l'Enseignement Supérieur.

§ 7

1. Trouver tous les groupes à 1, 2 ou 3 éléments.

2. On munit un ensemble à quatre éléments (notés e, a, b, c dans ce qui suit) de la loi de composition commutative donnée par la table de multiplication suivante :

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Montrer que l'on obtient ainsi un groupe commutatif. Trouver tous ses automorphismes. (Ce groupe est connu sous le nom de **Vierergruppe** de Klein). Interpréter géométriquement ce groupe (considérer, dans l'espace, les symétries par rapport aux arêtes d'un trièdre trirectangle).

3. Montrer que le groupe $\mathfrak{S}_4$ des permutations de l'ensemble $\{1, 2, 3, 4\}$ possède un sous groupe invariant isomorphe au Vierergruppe de Klein.

4. On munit l'ensemble $\mathbf{R}$ des nombres réels de la loi de composition

$$(x, y) \to \sqrt[3]{x^3 + y^3};$$

montrer qu'on obtient ainsi un groupe, isomorphe au groupe additif $\mathbf{R}$.

5. Soient $G_1, \ldots, G_n$ des groupes, et $H_1, \ldots, H_n$ des sous-groupes de $G_1, \ldots, G_n$; montrer que $H_1 \times \cdots \times H_n$ est un sous-groupe du groupe produit $G_1 \times \cdots \times G_n$.

6. Soit G un groupe noté multiplicativement. Pour tout $a \in G$, on définit une application s_a de G dans G en posant

$$s_a(x) = ax \quad \text{pour tout } x \in G$$

(**translation à gauche** d'amplitude a dans G; le lecteur comprendra l'origine de cette terminologie en examinant le cas où G est le groupe additif des vecteurs d'origine donnée O de l'espace usuel). Montrer que l'application $a \to s_a$ est un isomorphisme du groupe G sur un groupe de permutations de l'ensemble G.

7. Soient G un groupe cyclique à m éléments et x un générateur de G. Pour que x^k soit un générateur de G, il faut et il suffit que les entiers m et k soient premiers entre eux (utiliser le théorème de Bezout). Dans le cas général, quel est l'ordre du sous-groupe de G engendré par x^k?

8. Soient m et n des entiers rationnels. Pour qu'il existe un entier r tel que l'on ait

$$r \equiv 0 \pmod{m} \qquad \text{et} \qquad r \equiv 1 \pmod{n},$$

il faut et il suffit que m et n soient premiers entre eux.

q Déduire de là le résultat suivant. Soient G un groupe commutatif, x et y des éléments de G d'ordres m et n premiers entre eux; alors $z = xy$ est d'ordre mn, et le sous-groupe engendré par z contient x et y. (On appelle **ordre** d'un élément x d'un groupe l'ordre, i.e. le nombre d'éléments, du sous-groupe engendré par x; cet ordre est fini si et seulement s'il existe un entier $n \neq 0$ tel que

$$x^n = e;$$

dans ce cas, l'ordre de x est le *plus petit* $n \geqslant 1$ vérifiant cette relation, comme le lecteur le démontrera).

q 9. Soient G et H des groupes cycliques à m et n éléments. Pour que $G \times H$ soit cyclique il faut et il suffit que m et n soient premiers entre eux. Si x et y sont des générateurs de G et H, le couple (x, y) est alors un générateur de $G \times H$.

q 10. Tout groupe fini d'ordre premier est cyclique, et admet pour générateur chacun de ses éléments autre que l'élément neutre (utiliser le Théorème 4 du § 7, ou l'*Exercice* 7).

11. Soit A une partie d'un groupe G. On appelle **centralisateur** de A dans G l'ensemble Z(A) des $x \in G$ tels que $xa = ax$ pour tout $a \in A$. Montrer que Z(A) est un sous-groupe de G. Montrer que Z(G) (qu'on appelle le **centre** de G) est un sous-groupe commutatif et invariant de G.

12. Deux éléments x et y d'un groupe G sont dits **conjugués** s'il existe un $s \in G$ tel que

$$y = sxs^{-1}.$$

Montrer que

$$x \text{ et } y \text{ sont conjugués}$$

est une relation d'équivalence sur l'ensemble G. On prend pour G le groupe des rotations autour d'un point donné O dans l'espace, et on choisit une droite D passant par O; montrer que tout élément de G est conjugué d'une rotation autour de D.

13. Étant donnée une partie A d'un groupe G, on note sAs^{-1} (pour $s \in G$ donné) l'ensemble des éléments de G de la forme sxs^{-1}, avec $x \in A$. Montrer que si A est un sous-groupe il en est de même de sAs^{-1} (on dit alors que c'est un sous-groupe **conjugué** de A dans G). On appelle **normalisateur** d'un sous-groupe A de G l'ensemble N(A) des $s \in G$ tels que $sAs^{-1} = A$. Montrer que le centralisateur de A (*Exercice* **11**) est un sous-groupe invariant du normalisateur de A.

14. Soit G un groupe opérant sur un ensemble X.

a) Montrer que la relation

$$\text{il existe un } s \in G \text{ tel que } y = sx$$

est une relation d'équivalence sur l'ensemble X (la classe pour cette relation d'un $x \in X$ s'appelle l'**orbite** de x par G). Montrer que, pour tout $x \in X$, l'ensemble des $s \in G$ tels que $sx = x$ est un sous-groupe H_x de G (appelé **stabilisateur** de x dans G), et que les stabilisateurs des divers points d'une même orbite sont deux à deux conjugués dans G au sens de l'*Exercice* **13**.

b) On considère, pour un $x \in X$ donné, l'application f de G dans X donnée par

$$f(s) = sx;$$

montrer qu'elle est composée de l'application canonique de G sur G/H_x et d'une application de G/H_x dans X; montrer que celle-ci induit une bijection de G/H_x sur l'orbite M de x par G, et que Card (G) = Card (M) . Card (H_x) si G est fini.

c) Décrire les orbites et les stabilisateurs lorsqu'on prend pour X l'espace usuel et pour G le groupe des rotations autour d'un point donné O dans X.

¶ *d*) On suppose que G est fini, d'ordre une puissance d'un nombre premier p, et que X est fini, le nombre d'éléments de X n'étant pas multiple de p. Montrer qu'alors G admet au moins un **point fixe** dans X (i.e. qu'il existe un $x \in X$ tel que $sx = x$ pour tout $s \in G$).

¶ *e*) Soit G un **p-groupe** i.e. un groupe fini dont l'ordre est une puissance d'un nombre premier p. En faisant opérer G sur lui-même par les automorphismes intérieurs (cf. *Exemple* **19**), montrer que le centre de G (*Exercice* **11**) n'est pas réduit à l'élément neutre.

¶ **15.** Soient G un groupe et H un sous-groupe de G; on fait opérer G sur G/H (*Exemple* **20**). Montrer que les éléments de G/H dont le stabilisateur contient H sont les images, par l'application canonique de G dans G/H, des éléments du sous-groupe N(H), normalisateur de H dans G, défini dans l'*Exercice* **13** ci-dessus.

¶ **16.** Soit H un sous-groupe *invariant* d'un groupe G. Montrer qu'il existe sur l'ensemble G/H une et une seule loi de composition faisant de G/H un groupe et telle que l'application canonique de G dans G/H soit un homomorphisme de groupes (utiliser le Théorème 3 du § 4); le groupe ainsi obtenu s'appelle le **groupe quotient** de G par H. Que se passe-t-il lorsqu'on prend pour G le groupe additif **Z** des entiers rationnels et pour H un sous-groupe de G?

Soit p l'application canonique de G sur G/H; montrer que, pour tout sous-groupe A de G/H, il existe un et un seul sous-groupe K de G contenant H tel que $A = p(K)$, et que l'on a du reste $K = p^{-1}(A)$.

On appelle **sous-groupe dérivé** de G le sous-groupe, noté G' ou D(G), engendré par les éléments de la forme $xyx^{-1}y^{-1}$. Montrer que D(G) est un sous-groupe invariant de G, et que, si H est un sous-groupe invariant de G, pour que le groupe quotient G/H soit commutatif il faut et il suffit que $H \supset D(G)$.

qq **17.** Étant donnés des sous-groupes A et B d'un groupe G, on note (A, B) le sous-groupe de G engendré par les éléments $xyx^{-1}y^{-1}$ où $x \in A$ et $y \in B$. On pose

$$D(G) = (G, G), \qquad D^2(G) = D(D(G)), \qquad D^3(G) = D(D^2(G)), \text{ etc...}$$

Montrer que les conditions suivantes sont équivalentes :

a) Il existe un entier r tel que $D^{r+1}(G) = \{e\}$;

b) On peut construire des sous-groupes

$$\{e\} = H_0 \subset H_1 \subset \cdots \subset H_s = G$$

de G tels que, pour chaque indice i vérifiant $0 \leqslant i \leqslant s-1$, le sous-groupe H_i soit invariant dans H_{i+1} et le groupe quotient H_{i+1}/H_i soit commutatif (la notion de groupe quotient est définie dans l'*Exercice* précédent) ;

c) On peut construire des sous-groupes invariants

$$\{e\} = K_0 \subset K \subset \cdots \subset K_s = G$$

de G tels que tous les groupes quotients K_{j+1}/K_j soient commutatifs.

Un groupe G vérifiant ces conditions est dit **résoluble.** Montrer que tout sous-groupe d'un groupe résoluble est résoluble. Soient G un groupe et H un sous-groupe invariant de G ; les groupes H et G/H sont résolubles, il en est de même de G.

q **18.** Soit G un p-groupe (*Exercice* 14). Montrer que tout sous-groupe et tout groupe quotient de G est un p-groupe. En utilisant l'*Exercice* 14, *e*), montrer que G contient des sous-groupes invariants

$$\{e\} = H_0 \subset H_1 \subset \cdots \subset H_s = G$$

tels que les quotients H_i/H_{i-1} soient isomorphes au groupe additif $\mathbf{Z}/p\mathbf{Z}$ (i.e. soient cycliques d'ordre p) pour $1 \leqslant i \leqslant r$. En particulier, un p-groupe est résoluble.

19. Soit G un groupe cyclique d'ordre fini n.

a) Montrer que pour tout diviseur d de n, les $x \in G$ tels $x^d = e$ sont en nombre d.

b) Soit d un diviseur de n. Pour qu'un $x \in G$ puisse s'écrire sous la forme y^d pour un $y \in G$ convenablement choisi, il faut et il suffit que

$$x^{n/d} = e.$$

qq **20.** Soit G un groupe commutatif fini d'ordre n. On suppose que, pour tout diviseur d de n, les $x \in G$ tels que

$$x^d = e$$

soient en nombre d *au plus* ; on se propose d'en déduire que G est *cyclique* (résultat indispensable pour l'étude des corps finis : voir § 33, *Exercice* 2). Dans ce qui suit, on désigne par

$$n = p_1^{r_1} \cdots p_h^{r_h}$$

la décomposition de n en facteurs premiers.

a) Prouver que, pour tout i tel que $1 \leqslant i \leqslant h$, il existe un $a_i \in G$ vérifiant

$$a_i^{p_i^{r_i}} = e, \qquad a_i^{p_i^{r_i-1}} \neq e$$

et que a_i est d'ordre

$$q_i = p_i^{r_i}$$

exactement.

b) Montrer, en utilisant le fait que les q_i sont deux à deux premiers entre eux, que l'élément $a_1 \dots a_h$ est d'ordre $q_1 \dots q_h = n$, et en conclure que G est cyclique comme annoncé.

c) Montrer qu'on parviendrait à la même conclusion en faisant l'hypothèse moins forte que voici : pour $1 \leqslant i \leqslant h$, les $x \in G$ tels que

$$x^{p_i} = e$$

sont en nombre p_i au plus.

21. Soit f un homomorphisme d'un groupe fini G dans un groupe H. Montrer que

$$\mathrm{Card}(G) = \mathrm{Card}(\mathrm{Ker}(f)) \,.\, \mathrm{Card}(\mathrm{Im}(f)).$$

22. Soient G un groupe commutatif fini et n un entier tel que (*)

$$x^n = e \qquad \text{pour tout } x \in G.$$

a) On suppose $n = rs$ avec r et s premiers entre eux; soit M (resp. N) l'ensemble des $x \in G$ tels que

$$x^r = e \text{ (resp. } x^s = e).$$

Montrer que M et N sont des sous-groupes de G. En écrivant l'identité de Bezout pour r et s, montrer que l'application

$$f : M \times N \to G$$

donnée par $f(x, y) = xy$ est un isomorphisme de groupes.

b) Soit

$$n = p_1^{r_1} \dots p_h^{r_h} = q_1 \dots q_h \qquad \text{où} \qquad q_i = p_i^{r_i}$$

la décomposition de n en facteurs premiers; pour tout i tel que $1 \leqslant i \leqslant h$, soit M_i le sous-groupe des $x \in G$ tels que

$$x^{q_i} = e.$$

Montrer que G est isomorphe au produit direct des groupes $M_1, \dots, M_h$.

c) Soient M un groupe commutatif fini, p un nombre premier et r un entier naturel; on suppose que

$$x^{p^r} = e$$

pour tout $x \in M$; montrer que Card (M) est une puissance de p (observer que, si $M \neq \{e\}$, on peut trouver dans M un sous-groupe M' d'ordre p; l'introduction du groupe quotient M/M', cf. *Exercice* 16, permet alors de raisonner par récurrence sur le nombre d'éléments de M).

d) Démontrer le théorème suivant : *soit* G *un groupe commutatif fini d'ordre*

$$n = p_1^{r_1} \dots p_h^{r_h};$$

alors G *est isomorphe au produit direct de* h *groupes d'ordres* $p_1^{r_1}, \dots, p_h^{r_h}$ (ce résultat, que Gauss avait

(*) L'énoncé de cet *Exercice* est rédigé en notation multiplicative, mais le lecteur aura intérêt, pour des généralisations ultérieures, à le traduire en notation additive.

déjà plus ou moins démontré en 1801, sera complété dans les *Exercices* du § 31 : un groupe commutatif dont l'ordre est une puissance de p est isomorphe à un produit direct de groupes *cycliques* dont les ordres sont des puissances de p. Il en résultera que tout groupe commutatif fini est isomorphe à un produit direct de groupes cycliques. L'étude complète des groupes commutatifs à un nombre fini de générateurs a été faite par Kronecker en 1870; un tel groupe est produit direct d'un groupe commutatif fini et d'un groupe $\mathbf{Z}^n$).

¶ **23.** On reprend la question *a*) de l'*Exercice* précédent. Soit A (resp. B) le sous-groupe de G formé des x tels que l'on ait

$$x = y^s \text{ (resp. } x = y^r\text{)}$$

pour au moins un $y \in G$. Montrer que $A = M$ et $B = N$ (on prouvera qu'on a les relations

$$\mathrm{Card}(G) = \mathrm{Card}(M).\mathrm{Card}(N) = \mathrm{Card}(A).\mathrm{Card}(N) = \mathrm{Card}(B).\mathrm{Card}(M)$$

et on observera que $A \subset M$, $B \subset N$).

On suppose que $n = \mathrm{Card}(G)$. Montrer que $\mathrm{Card}(M) = r$, $\mathrm{Card}(N) = s$.

24. Soit $s \in \mathfrak{S}_n$ une permutation de l'ensemble $X = \{1, 2, \ldots, n\}$ et soit G le sous-groupe de $\mathfrak{S}_n$ formé par les puissances de s.

a) Montrer qu'on peut trouver des parties non vides $I_1, \ldots, I_r$ de X vérifiant les conditions suivantes : on a $g(I_k) = I_k$ pour $1 \leqslant k \leqslant r$ et tout $g \in G$; les ensembles I_k sont deux à deux disjoints et leur réunion est X tout entier; pour que $p, q \in X$ appartiennent à un même I_k, il faut et il suffit qu'il existe un $g \in G$ tel que $q = g(p)$. Relation avec l'*Exercice* 14, *a*)?

b) Montrer que les conditions précédentes caractérisent les ensembles I_k (à ceci près qu'on peut naturellement modifier l'ordre dans lequel on les écrit).

c) Montrer que, pour chaque k, on peut écrire les éléments de I_k sous forme d'une suite $i_0, \ldots, i_p$ de telle sorte que l'on ait

$$s(i_0) = i_1, \quad s(i_1) = i_2, \quad \ldots, \quad s(i_{p-1}) = i_p, \quad s(i_p) = i_0$$

(*décomposition d'une permutation en cycles*; on appelle **cycle** pour s toute suite d'entiers $i_0, \ldots, i_p$ écrits dans l'ordre naturel, deux à deux distincts, et vérifiant les relations précédentes).

d) Trouver les cycles des permutations suivantes :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 1 & 3 & 6 & 5 & 7 & 4 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 5 & 8 & 9 & 2 & 1 & 4 & 3 & 6 & 7 \end{pmatrix}.$$

(NB. — On utilise ici la notation standard pour représenter une permutation s; celle-ci consiste à écrire sur la seconde ligne les images par s des éléments de la première).

e) Étant donnée une permutation $s \in \mathfrak{S}_n$, soient $n_1, \ldots, n_r$ les nombres de termes des divers cycles de s. Montrer que l'ordre de s (i.e. le plus petit entier $q \geqslant 1$ tel que $s^q = e$, ou l'ordre du groupe cyclique engendré par s) est le ppcm des entiers $n_1, \ldots, n_r$.

f) On considère la permutation

$$s : \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 5 & 4 & 1 & 7 & 10 & 2 & 6 & 9 & 8 \end{pmatrix};$$

calculer l'ordre du sous-groupe de $\mathfrak{S}_{10}$ engendré par s. Calculer la permutation

s^{100}.

qq 25. Dans cet Exercice, on utilise la terminologie suivante. Étant donnée une suite

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \longrightarrow \cdots G_n \xrightarrow{f_n} G_{n+1}$$

formée de groupes additifs G_i et d'homomorphismes $f_i : G_i \to G_{i+1}$, on dit que cette suite est **exacte** si, pour chaque entier i tel que $1 \leqslant i < n$, l'image de l'homomorphisme f_i est égale au noyau de l'homomorphisme suivant f_{i+1}. D'autre part, on dit qu'un diagramme, par exemple

$$\begin{array}{ccccccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C & \xrightarrow{h} & D & \xrightarrow{k} & E \\ \downarrow p & & \downarrow q & & \downarrow r & & \downarrow s & & \downarrow t \\ A' & \xrightarrow{f'} & B' & \xrightarrow{g'} & C' & \xrightarrow{h'} & D' & \xrightarrow{k'} & E' \end{array} \qquad (1)$$

formé d'ensembles et d'applications de ces ensembles les uns dans les autres, est **commutatif** si, quels que soient les «sommets» X et Y du diagramme, toutes les applications de X dans Y qu'on obtient en composant, de toutes les façons possibles, les applications figurant dans le diagramme donné, sont égales. Dans le cas du diagramme (1), la commutativité se traduirait par la relation $f' \circ p = q \circ f$ et de nombreuses autres relations analogues que le lecteur écrira.

On considère un diagramme *commutatif* (1) dans lequel A, B, ... sont des groupes additifs, et les applications des homomorphismes de groupes. On suppose que les deux lignes horizontales du diagramme sont des suites *exactes* — de sorte qu'on a

$$\mathrm{Im}(f) = \mathrm{Ker}(g), \qquad \mathrm{Im}(f') = \mathrm{Ker}(g'),$$

etc... Établir les résultats suivants (connus sous le nom de **lemme des cinq**) :

a) Si p est surjectif, et si q et s sont injectifs, alors r est injectif.

b) Si q et s sont surjectifs, et si t est injectif, alors r est surjectif.

c) Si p est surjectif, si q et s sont bijectifs, et si t est injectif, alors r est bijectif.

1. Soit K un anneau (qu'on *ne* suppose *pas* commutatif).

a) Montrer que, si deux éléments x et y de K commutent (i.e. si $xy = yx$), on a

$$x^n - y^n = (x - y)\,(x^{n-1} + x^{n-2}y + \cdots + xy^{n-2} + y^{n-1})$$

pour tout entier $n \geqslant 1$.

b) On dit qu'un élément x de K est **nilpotent** s'il existe un entier $n \geqslant 1$ tel que

$$x^n = 0.$$

Montrer qu'alors $1 - x$ est inversible.

c) Si deux éléments nilpotents x et y de K commutent, alors $x + y$ est nilpotent (utiliser l formule du binôme pour un exposant assez élevé), ainsi que xy.

d) Soit a un élément de K; on considère l'application u de K dans K donnée par

$$u(x) = ax - xa \quad \text{pour tout } x \in \mathrm{K}.$$

Montrer que, si $a^2 = 0$, on a $u^3(x) = 0$ pour tout $x \in \mathrm{K}$, et que si $a^3 = 0$ on a $u^5(x) = 0$ po tout $x \in \mathrm{K}$. Montrer d'une manière générale que, si a est nilpotent, il existe un entier q que

$$u^q(x) = 0 \quad \text{pour tout } x \in \mathrm{K}.$$

Montrer que l'on a

$$u^p(x) = \sum_{k=0}^{p} (-1)^k \binom{p}{k} a^{p-k}\, x a^k.$$

e) On dit qu'un élément u de K est **unipotent** si $1 - u$ est nilpotent. Montrer que si $u, v \in$ sont unipotents et commutent, alors uv est aussi unipotent. Montrer que tout élément unip tent de K est inversible, et a pour inverse un élément unipotent.

[Pour des exemples d'éléments unipotents et nilpotents d'un anneau, voir l'*Exercice* 10 d §§ 12, 13 et 14 et l'*Exercice* 19 du § 19. En Analyse, la théorie des *développements limités* four aussi des exemples d'éléments nilpotents : considérer l'anneau des fonctions $f(t)$ d'une variab réelle t, définie au voisinage de $t = 0$, et, un entier $n \geqslant 1$ étant choisi, passer au quotient cf. *Exercice* 7, *c*) ci-dessous — par l'idéal des fonctions qui sont $o(t^n)$ quand t tend vers 0; l'anne quotient a évidemment des éléments nilpotents non nuls — par exemple l'image dans ce qu tient de la fonction t].

qq **2.** Soit K un anneau; on suppose que le corps **Q** des nombres rationnels est un sous-anne de K (ce qui permet de multiplier tout $x \in \mathrm{K}$ par tout nombre rationnel, et en particulier diviser tout $x \in \mathrm{K}$ par tout entier rationnel non nul).

a) Soit x un élément nilpotent de K (*Exercice* 1); on définit (*)

$$\exp(x) = 1 + \frac{x}{1} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

Montrer, à l'aide de la formule du binôme, que si $x, y \in K$ sont nilpotents et commutent on a

$$\exp(x + y) = \exp(x).\exp(y).$$

b) Soit u un élément unipotent de K (*Exercice* 1); on définit

$$\log(u) = -\frac{1-u}{1} - \frac{(1-u)^2}{2} - \cdots - \frac{(1-u)^n}{n} - \cdots;$$

montrer que si $u, v \in K$ sont unipotents et commutent on a

$$\log(uv) = \log(u) + \log(v).$$

c) Soit x un élément nilpotent de K. Montrer que $\exp(x)$ est unipotent et que

$$\log(\exp(x)) = x.$$

d) Soit u un élément unipotent de K. Montrer que $\log(u)$ est nilpotent et que

$$\exp(\log(u)) = u.$$

e) Pour tout élément nilpotent x de K, on définit

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots$$

Démontrer que si $x, y \in K$ sont nilpotents et commutent on a

$$\cos(x+y) = \cos(x).\cos(y) - \sin(x).\sin(y)$$
$$\sin(x+y) = \sin(x).\cos(y) + \sin(y).\cos(x).$$

Montrer que

$$\cos^2(x) + \sin^2(x) = 1$$

pour tout élément nilpotent de K; on pose bien entendu $\cos^2(x) = \cos(x).\cos(x)$ et $\sin^2(x) = \sin(x).\sin(x)$.

(*) Ces définitions ont évidemment leur origine dans les développements en série entière des fonctions

$$e^t, \log(1+t), \cos t \text{ et } \sin t$$

étudiées en Analyse. Il n'y a ici aucun problème de convergence puisque les « séries » sont en réalité des sommes finies. L'*Exercice* consiste à transposer sur un plan purement algébrique la relation existant entre, par exemple, la propriété bien connue

$$e^{x+y} = e^x e^y$$

de la fonction exponentielle usuelle, et la nature du développement en série entière de celle-ci. Il arrive souvent que l'on puisse ainsi trouver des analogues purement algébriques de phénomènes faisant intervenir des considérations d'Analyse, i.e. des passages à la limite.

3. Soit K un anneau; quels que soient $x, y \in K$, on pose

$$[x, y] = xy - yx.$$

Démontrer l'**identité de Jacobi**

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0.$$

4. Dans un anneau K, on considère des éléments x, y, h vérifiant les relations

$$[h, x] = 2x, \qquad [h, y] = -2y, \qquad [x, y] = h.$$

a) Établir les formules

$$[h, x^n] = 2n.x^n, \qquad [h, y^n] = -2n.y^n$$

b) Montrer que l'élément

$$4xy - h^2 - 2h$$

de K commute à x, y et h.

c) Montrer que le plus petit sous-anneau de K contenant x, y et h est l'ensemble des éléments qui peuvent s'écrire sous la forme d'une somme d'un nombre fini de termes de la forme

$$a.x^i y^j h^k$$

où i, j, k sont des entiers naturels et a un entier rationnel.

5. Soit p un nombre premier. On désigne par $\mathbf{Z}_p$ l'ensemble des $x \in \mathbf{Q}$ qu'on peut écrire sous la forme d'une fraction dont le dénominateur n'est pas divisible par p.

a) Montrer que $\mathbf{Z}_p$ est un sous-anneau de $\mathbf{Q}$.

b) Pour tout $x \in \mathbf{Q}$ on a soit $x \in \mathbf{Z}_p$ soit $x^{-1} \in \mathbf{Z}_p$.

c) Les seuls sous-anneaux de $\mathbf{Q}$ contenant $\mathbf{Z}_p$ sont $\mathbf{Z}_p$ et $\mathbf{Q}$.

d) Pour tout idéal I de l'anneau $\mathbf{Z}_p$ il existe un entier $n \geq 0$ et un seul tel que I soit engendré par p^n (i.e. formé des $p^n u$, $u \in \mathbf{Z}_p$).

e) Pour tout $x \in \mathbf{Q}$ non nul il existe un $n \in \mathbf{Z}$ et un seul tel que

$$x = p^n.u$$

où u est un élément inversible de l'anneau $\mathbf{Z}_p$.

f) Pour tout $x \in \mathbf{Q}$ non nul on pose $v_p(x) = n$ où n est l'entier de la question précédente; en outre on définit (*)

$$v_p(0) = +\infty.$$

(*) On désigne par le symbole $+\infty$ un objet soumis aux règles de calcul que voici, et à celles-ci uniquement (autrement dit, les opérations non définies ci-dessous *n'ont aucun sens*) :

$$n + (+\infty) = +\infty \quad \text{pour tout} \quad n \in \mathbf{Z}; \qquad (+\infty) + (+\infty) = +\infty;$$

enfin on convient que

$$+\infty > n \quad \text{pour tout} \quad n \in \mathbf{Z}; \qquad +\infty \geq +\infty.$$

Il s'ensuit par exemple que $\mathrm{Max}\,(2, -\infty) = +\infty$. Bien entendu on pourrait se passer pour définir v_p du symbole $+\infty$: il suffit de ne pas attribuer de sens à $v_p(0)$, et d'énoncer alors les relations à démontrer de telle sorte qu'on n'ait jamais à écrire $v_p(0)$. Cette méthode compliquerait beaucoup la situation.

Montrer que l'on a

$$v_p(xy) = v_p(x) + v_p(y)$$
$$v_p(x + y) \geqslant \text{Min } [v_p(x), v_p(y)]$$

quels que soient $x, y \in \mathbf{Q}$; et que $\mathbf{Z}_p$ est l'ensemble des $x \in \mathbf{Q}$ tels que $v_p(x) \geqslant 0$.

g) Montrer que l'intersection des sous-anneaux $\mathbf{Z}_p$ de $\mathbf{Q}$ associés à tous les nombres premiers p est l'anneau $\mathbf{Z}$ des entiers rationnels.

6. Soient K un corps commutatif et A un sous-anneau de K. On dit que A est un **anneau de valuation** de K si $A \neq K$ et si l'on a

$$x \in A \quad \text{ou} \quad x^{-1} \in A \quad \text{pour tout } x \in K \text{ non nul.}$$

Montrer qu'alors les éléments non inversibles de l'anneau A forment un idéal $\mathfrak{m}$ de A, et que tout idéal de A, distinct de A tout entier, est contenu dans $\mathfrak{m}$ [de sorte que $\mathfrak{m}$ est l'unique idéal maximal de A, dans la terminologie de l'*Exercice* 7, *d*) ci-dessous].

On appelle **valuation discrète** de K toute fonction v définie sur K, dont les valeurs sont des entiers rationnels ou le symbole $+\infty$, et possédant les propriétés suivantes :

$$v(0) = +\infty; \quad v(x) \in \mathbf{Z} \text{ si } x \neq 0;$$
$$v(xy) = v(x) + v(y) \text{ quels que soient } x, y \in K;$$
$$v(x + y) \geqslant \text{Min } [v(x), v(y)] \text{ quels que soient } x, y \in K.$$

On suppose v non triviale (i.e. que $v(K)$ ne se réduit pas à 0 et $+\infty$). Montrer que l'ensemble A des $x \in K$ tels que $v(x) \geqslant 0$ est un anneau de valuation de K, et que l'idéal maximal $\mathfrak{m}$ de A est l'ensemble des $x \in K$ tels que $v(x) > 0$. On choisit un élément $\pi \in \mathfrak{m}$ tel que $v(\pi)$ soit minimum; montrer que $\mathfrak{m} = A\pi$ et que tout idéal de A est de la forme $A\pi^n$, pour un entier $n \geqslant 0$.

Montrer que les seuls anneaux de valuation du corps $\mathbf{Q}$ sont les anneaux $\mathbf{Z}_p$ de l'*Exercice* précédent. Trouver toutes les valuations discrètes de $\mathbf{Q}$.

7. Soit I un idéal bilatère d'un anneau K; on note

$$x \equiv y \pmod{I}$$

la relation $x - y \in I$ (**congruence** modulo I).

a) Montrer que c'est une relation d'équivalence sur l'ensemble K. Que se passe-t-il si $K = \mathbf{Z}$ et $I = p\mathbf{Z}$?

b) Montrer que les relations

$$x' \equiv y' \pmod{I} \quad \text{et} \quad x'' \equiv y'' \pmod{I}$$

impliquent les relations

$$x' + x'' \equiv y' + y'' \pmod{I} \quad \text{et} \quad x'x'' \equiv y'y'' \pmod{I}.$$

c) On note K/I l'ensemble quotient de K par la relation d'équivalence considérée, et θ l'application canonique de K sur K/I; montrer qu'il existe sur l'ensemble K/I une et une seule structure d'anneau telle que l'application θ soit un homomorphisme (imiter la construction donnée pour les anneaux $\mathbf{Z}/p\mathbf{Z}$). On dit que K/I est l'**anneau quotient** de K par l'idéal bilatère I.

d) On suppose K commutatif. On dit qu'un idéal I de K est **maximal** si $I \neq K$ et si les seuls idéaux de K contenant I sont I et K. Montrer que, pour que I soit maximal, il faut et il suffit

que l'anneau quotient K/I soit un *corps* (on notera qu'un corps ne possède aucun idéal autre que lui-même et $\{0\}$, et réciproquement). Quels sont les idéaux maximaux de l'anneau $\mathbf{Z}$?

e) Un idéal I d'un anneau commutatif K est dit **premier** si $I \neq K$ et si, pour $x, y \in K$, la relation

$$xy \in I \quad \text{implique} \quad x \in I \quad \text{ou} \quad y \in I.$$

Montrer que cette condition signifie que l'anneau quotient K/I est *intègre*. Quels sont les idéaux premiers de l'anneau $\mathbf{Z}$?

f) Montrer que tout idéal maximal est premier. [NB — La réciproque n'est vraie que pour des catégories d'anneaux très particulières.]

g) Soient K un corps commutatif et A un sous-anneau de K; on suppose que tout $x \in K$ puisse se mettre sous la forme u/v avec $u, v \in A$ et $v \neq 0$ (ceci signifie que K est le corps des fractions de A, cf. § 29). Soit (*) $\mathfrak{p}$ un idéal premier de A; on note $A_{\mathfrak{p}}$ (**anneau local de** $\mathfrak{p}$) l'ensemble des éléments de K qui peuvent se mettre sous la forme

$$u/v \quad \text{avec} \quad u, v \in A \quad \text{et} \quad v \notin \mathfrak{p}.$$

Montrer que $A_{\mathfrak{p}}$ est un sous-anneau de K possédant un seul idéal maximal, et que si l'on associe à chaque idéal $I \neq A_{\mathfrak{p}}$ de l'anneau $A_{\mathfrak{p}}$ son intersection $I \cap A$ avec A, on définit une *bijection* de l'ensemble des idéaux $I \neq A_{\mathfrak{p}}$ de $A_{\mathfrak{p}}$ sur l'ensemble des idéaux de A contenus dans $\mathfrak{p}$.

8. Soient A et B deux anneaux. Montrer qu'on obtient un anneau (**composé direct** de A et B) en munissant l'ensemble $A \times B$ des lois de compositions données par les formules

$$(x', y') + (x'', y'') = (x' + x'', y' + y''), \qquad (x', y') \cdot (x'', y'') = (x'x'', y'y'').$$

Le composé direct $A \times B$ peut-il être un anneau d'intégrité?

9. Soient m et n des entiers rationnels premiers entre eux.

a) Montrer que, quels que soient $a, b \in \mathbf{Z}$, il existe $x \in \mathbf{Z}$ tel que

$$x \equiv a \pmod{m} \qquad \text{et} \qquad x \equiv b \pmod{n}$$

et que la classe de x modulo mn est entièrement déterminée par la classe de a modulo m et celle de b modulo n. Exemple : trouver toutes les solutions du système de congruences

$$x \equiv 4 \pmod{7}, \qquad x \equiv 9 \pmod{11}.$$

b) On considère les anneaux $A = \mathbf{Z}/m\mathbf{Z}$, $B = \mathbf{Z}/n\mathbf{Z}$ et $C = \mathbf{Z}/mn\mathbf{Z}$. A l'aide de la question précédente, construire un isomorphisme du composé direct $A \times B$ (*Exercice* 8) sur l'anneau C. (On pourra utiliser le Théorème 3 du § 4).

c) Soient $q_1, \ldots, q_h$ des entiers deux à deux premiers entre eux. Montrer par récurrence sur h que, quels que soient $a_1, \ldots, a_h \in \mathbf{Z}$, il existe un $x \in \mathbf{Z}$ qui vérifie les h relations

$$x \equiv a_i \pmod{q_i} \quad (1 \leqslant i \leqslant h).$$

(Ce résultat est connu sous le nom de **théorème chinois**, attendu que les Chinois en utilisaient des cas particuliers pour choisir les dates d'événements liés aux périodes de certains phénomènes astronomiques ou autres.)

(*) La tradition pour désigner des idéaux est d'utiliser des lettres gothiques; on ne s'y est pas conformé dans le texte du § 8 pour éviter de troubler les débutants.

d) Soit

$$n = p_1^{r_1} \dots p_h^{r_h}$$

la décomposition d'un entier n en facteurs premiers. Montrer que l'anneau $\mathbf{Z}/n\mathbf{Z}$ est isomorphe au composé direct des h anneaux $\mathbf{Z}/q_i\mathbf{Z}$, où l'on pose

$$q_i = p_i^{r_i} \quad (1 \leqslant i \leqslant h).$$

e) Soient $q_1, \dots, q_h$ des entiers rationnels quelconques; pour qu'on puisse résoudre le système de congruences

$$x \equiv a_i \pmod{q_i} \quad 1 \leqslant i \leqslant h,$$

il faut et il suffit qu'on ait

$$a_i \equiv a_j \pmod{d_{ij}} \qquad \text{pour} \qquad 1 \leqslant i < j \leqslant h,$$

où d_{ij} désigne le pgcd de q_i et q_j.

10. Soient I et J des idéaux d'un anneau commutatif K. On note $\mathrm{I} + \mathrm{J}$ (**somme** des idéaux I et J) l'ensemble des éléments de K qui peuvent se mettre sous la forme $x + y$ avec $x \in \mathrm{I}$ et $y \in \mathrm{J}$, et IJ (**produit** des deux idéaux I et J) l'ensemble des $z \in \mathrm{K}$ possédant la propriété suivante : il existe un entier $n \geqslant 1$, des éléments $x_1, \dots, x_n$ de I, et des éléments $y_1, \dots, y_n$ de J, tels que $z = x_1 y_1 + \dots + x_n y_n$.

a) Montrer que $\mathrm{I} + \mathrm{J}$ est le plus petit idéal de K contenant I et J. Montrer que IJ est aussi un idéal de K, contenu dans $\mathrm{I} \cap \mathrm{J}$. Établir les relations

$$\mathrm{I} + \mathrm{J} = \mathrm{J} + \mathrm{I}, \qquad \mathrm{I} + (\mathrm{I}' + \mathrm{I}'') = (\mathrm{I} + \mathrm{I}') + \mathrm{I}'',$$
$$\mathrm{IJ} = \mathrm{JI}, \qquad \mathrm{I}(\mathrm{I}'\mathrm{I}'') = (\mathrm{II}')\mathrm{I}'', \qquad \mathrm{I}(\mathrm{J}' + \mathrm{J}'') = \mathrm{IJ}' + \mathrm{IJ}''$$

où I, I', etc... désignent des idéaux de K. Interprétation de $\mathrm{I} + \mathrm{J}$ et de IJ lorsque I et J sont principaux?

b) On dit que deux idéaux I et J de K sont **étrangers** lorsque $\mathrm{I} + \mathrm{J} = \mathrm{K}$. Quelle est la signification de cette propriété lorsque $\mathrm{K} = \mathbf{Z}$? Montrer que, si I et J sont étrangers, on a $\mathrm{I} \cap \mathrm{J} = \mathrm{IJ}$.

c) Pour que deux idéaux I et J de K soient étrangers, il faut et il suffit que, quels que soient $a, b \in \mathrm{K}$, il existe $x \in \mathrm{K}$ tel que l'on ait

$$x \equiv a \pmod{\mathrm{I}} \qquad \text{et} \qquad x \equiv b \pmod{\mathrm{J}}.$$

¶ En déduire que l'anneau quotient K/IJ (*Exercice* 7) est isomorphe au composé direct (*Exercice* 8) des anneaux K/I et K/J.

d) Soient $\mathrm{I}, \mathrm{J}_1, \dots, \mathrm{J}_r$ des idéaux de K; on suppose I et J_k étrangers pour $1 \leqslant k \leqslant r$. Montrer que I est étranger au produit $\mathrm{J}_1 \dots \mathrm{J}_r$.

e) Soient $\mathrm{J}_1, \dots, \mathrm{J}_r$ des idéaux deux à deux étrangers. Montrer que

$$\mathrm{J}_1 \dots \mathrm{J}_r = \mathrm{J}_1 \cap \dots \cap \mathrm{J}_r.$$

f) Soient $\mathrm{J}_1, \dots, \mathrm{J}_r$ des idéaux deux à deux étrangers. Montrer que, quels que soient $a_1, \dots, a_r \in \mathrm{K}$ on peut trouver un $x \in \mathrm{K}$ tel que l'on ait

$$x \equiv a_k \pmod{\mathrm{J}_k} \quad \text{pour } 1 \leqslant k \leqslant r.$$

g) Deux idéaux maximaux (*Exercice* 7, (*d*)) de K sont étrangers dès qu'ils sont distincts.

¶ **11.** Soient $I_1, \ldots, I_r$ des idéaux d'un anneau commutatif K. Si un idéal **premier** (*Exercice* 7, (*e*)) de K contient le produit $I_1 \ldots I_r$ il contient l'un au moins des idéaux $I_1, \ldots, I_r$.
Soit I un idéal non premier de K. Montrer qu'il existe des idéaux J′ et J″ de K possédant les propriétés suivantes : J′ et J″ contiennent I et sont distincts de I, et I contient l'idéal produit J′J″.

12. Étant donné un idéal I d'un anneau commutatif K, on appelle **radical** de I l'ensemble des $x \in K$ tels que l'on ait

$$x^n \in I$$

pour un entier $n \geqslant 1$ au moins. Dans cet *Exercice*, on désigne le radical d'un idéal I par la notation

$$\sqrt{I}$$

(laquelle est aussi mauvaise que possible comme le montreront les formules qui vont suivre...).
a) Montrer que le radical d'un idéal I est encore un idéal. Que se passe-t-il si $I = \{0\}$? Quel est le radical d'un idéal premier (*Exercice* 7, (*e*)) de K?
b) Démontrer les formules suivantes, où I et J désignent deux idéaux quelconques de l'anneau K :

$$\sqrt{I.J} = \sqrt{I \cap J} = \sqrt{I} \cap \sqrt{J}$$
$$\sqrt{I + J} = \sqrt{\sqrt{I} + \sqrt{J}}$$
$$\sqrt{\sqrt{I}} = \sqrt{I}.$$

c) Déterminer complètement le radical d'un idéal de l'anneau **Z** des entiers rationnels.

13. On dit qu'un idéal I d'un anneau commutatif K est **primaire** si $I \neq K$ et si, quels que soient les éléments x, y de K tels que l'on ait

$$xy \in I, \qquad x \notin I,$$

il existe un entier $n \geqslant 1$ tel que

$$y^n \in I.$$

Montrer que le radical d'un idéal primaire est un idéal premier.
¶ Pour que I (supposé distinct de K) soit primaire, il faut et il suffit que, dans l'anneau quotient K/I, tout diviseur de zéro soit nilpotent.
Quels sont les idéaux primaires de l'anneau **Z**?

14. Soit $\mathfrak{m}$ un idéal maximal d'un anneau commutatif K. Montrer que les puissances

$$\mathfrak{m}^n = \mathfrak{m} \ldots \mathfrak{m} \ (n \text{ facteurs})$$

de $\mathfrak{m}$ sont des idéaux primaires, ayant $\mathfrak{m}$ pour radical.

15. Soient $\mathfrak{m}$ un idéal maximal d'un anneau commutatif K, et $\mathfrak{a}$ un idéal de K contenu dans $\mathfrak{m}$. On suppose que chaque élément de $\mathfrak{m}$ possède une puissance dans $\mathfrak{a}$. Montrer que $\mathfrak{a}$ est primaire et que son radical est $\mathfrak{m}$.

16. Pour qu'un élément d'un anneau commutatif K soit inversible, il faut et il suffit qu'il n'appartienne à aucun autre idéal de K que K lui-même.

On admet le **théorème de Krull** que voici : étant donné un anneau commutatif (*) K, tout idéal de K, distinct de K, est contenu dans au moins un idéal maximal de K [on rappelle, *Exercice* 7, *d*), qu'un idéal I de K est dit maximal si $I \neq K$ et si les seuls idéaux de K contenant I sont I et K].
Montrer que, pour qu'un élément de K soit inversible, il faut et il suffit qu'il n'appartienne à aucun idéal maximal de K.

17. Soit I l'intersection de tous les idéaux maximaux d'un anneau commutatif K. Montrer qu'un élément a de K appartient à I si et seulement si $1 + ax$ est inversible pour tout $x \in K$ (utiliser l'*Exercice* précédent).
[Ce résultat (Jacobson) s'étend aux anneaux non commutatifs : dans un tel anneau, l'intersection des idéaux à gauche maximaux est identique à l'intersection des idéaux à droite maximaux; et les éléments a de cette intersection sont caractérisés par le fait que $1 + xay$ est inversible quels que soient $x, y \in K$. Les démonstrations de ces résultats sont parfaitement élémentaires.]

¶ **18.** On désigne par $\mathbf{F}_p$ le corps $\mathbf{Z}/p\mathbf{Z}$ pour p premier. On munit l'ensemble $\mathbf{F}_{11} \times \mathbf{F}_{11}$ des deux lois de composition données par les formules suivantes :

$$(u, v) + (x, y) = (u + x, v + y)$$
$$(u, v) \cdot (x, y) = (ux + 7vy, uy + vx)$$

(où 7 désigne naturellement la classe modulo 11 de l'entier naturel 7). Montrer qu'on obtient de cette façon un corps commutatif à 121 éléments.

19. Montrer que, si p est un nombre premier, le coefficient du binôme $\binom{n}{p}$ est multiple de p pour $1 \leqslant n \leqslant p - 1$. En déduire que, si p est un nombre premier impair, on a

$$(1 + p)^{p^k} \equiv 1 + p^{k+1} \pmod{p^{k+2}}$$

pour tout entier $k \geqslant 0$ (raisonner par récurrence).

(*) Le théorème de Krull s'étend en fait à tout anneau K, commutatif ou non, de la façon suivante. Dans un tel anneau, un idéal à gauche (resp. à droite, bilatère) I est dit **maximal** si $I \neq K$ et si les seuls idéaux à gauche (resp. à droite, bilatère) de K contenant I sont I et K. Ceci dit, tout idéal à gauche (resp. à droite, bilatère) de K, autre que K lui-même, est contenu dans au moins un idéal à gauche (resp. à droite, bilatère) maximal de K.

Lorsque $K = \mathbf{Z}$, le théorème de Krull signifie que tout entier $n \geqslant 2$ possède au moins un diviseur premier. Le théorème de Krull peut donc être considéré comme une extension de ce résultat à *tous les anneaux sans exception*; à ce titre, et en dépit de la simplicité de son énoncé, c'est l'un des résultats les plus utiles de toute l'Algèbre, et on l'a même utilisé (Gelfand) depuis une vingtaine d'années pour démontrer des théorèmes difficiles d'Analyse, en l'appliquant à des anneaux dont les éléments sont des fonctions d'une ou plusieurs variables réelles vérifiant certaines conditions.

La démonstration générale du théorème de Krull est facile pourvu qu'on connaisse suffisamment la théorie des Ensembles et des nombres transfinis (c'est même l'un des points précis où l'on voit la « grande » théorie des nombres transfinis servir à démontrer des résultats « concrets » très peu évidents). On verra au § 18 une démonstration élémentaire du théorème de Krull pour les anneaux *noethériens* (mais ceux qu'on étudie en Analyse le sont rarement). L'idée de la démonstration générale est que, si l'anneau K ne contenait aucun idéal maximal, on pourrait construire dans K une chaîne croissante infinie, et même « transfinie », d'idéaux (autrement dit, attacher à chaque cardinal α un idéal I_α de telle sorte que, pour $\alpha < \beta$, l'idéal I_α soit *strictement* contenu dans I_β — la construction est évidemment facile pour les α finis, et toute la difficulté est de prolonger la récurrence au delà des entiers naturels), ce qui serait en contradiction avec le fait qu'il n'existe pas d'injection de l'ensemble (sic) de tous les cardinaux dans un ensemble donné (en espèce, l'ensemble des idéaux de K), pour la raison que les cardinaux ne forment pas un ensemble...

20. On pose $r = \sqrt[3]{2}$. Montrer que l'ensemble des nombres de la forme

$$a + br + cr^2,$$

où a, b, c sont des nombres *rationnels* arbitraires, est un sous-corps du corps **R** des nombres réels.
(Cet *Exercice* sera aussi, pour le lecteur débutant, une occasion de *démontrer* que 2 n'est pas le cube d'un nombre rationnel).

§ 9

1 *a*) Montrer que tout nombre complexe $d \neq 0$ possède exactement deux racines carrées dans le corps **C** (on cherchera leur module et leur argument en fonction de ceux de d).

b) En déduire que toute équation du second degré

$$az^2 + bz + c = 0$$

à coefficients a, b, c complexes, possède au moins une racine dans **C**, et en possède deux si

$$b^2 - 4ac \neq 0.$$

c) Résoudre dans **C** les équations

$$z^2 + (5 - 2i)z + 5 - 5i = 0$$
$$z^2 + (1 - 2i)z - 2i = 0$$
$$z^4 - 30z^2 + 289 = 0$$

2. Trouver les parties réelle et imaginaire des nombres complexes suivants :

$$\frac{(1+2i)^3 - (1-i)^3}{(3+2i)^3 - (2+i)^2}; \qquad \frac{(1+i)^9}{(1-i)^7}; \qquad \left(-\frac{1}{2} + \frac{i\sqrt{3}}{2}\right)^3; \qquad \sqrt[4]{2 - i\sqrt{12}};$$
$$(az^2 + bz)(bz^2 + az) \quad \text{où} \quad z = -\frac{1}{2} + \frac{i\sqrt{3}}{2}.$$

3. Calculer le module et l'argument des nombres complexes suivants :

$$-1 + i; \qquad -1 - i\sqrt{3}; \qquad 2 + \sqrt{3} + i; \qquad \left(\frac{1 + i\sqrt{3}}{1 - i}\right)^{20};$$
$$(1 + \cos\theta + i.\sin\theta)^n.$$

4. Trouver toutes les solutions complexes des équations suivantes (on utilisera la formule de Moivre) :

$$z^3 + 1 = 0; \qquad z^5 = 1; \qquad z^3 + 27 = 0; \qquad z^8 = \frac{1+i}{\sqrt{3} - i}.$$

5. *a*) Calculer

$$\cos 5t, \quad \cos 8t, \quad \sin 6t, \quad \sin 9t, \quad \operatorname{tg} 6t$$

en fonction des lignes trigonométriques de l'angle t.

b) Calculer

$$\sin^3 t, \quad \sin^4 t, \quad \cos^5 t, \quad \cos^8 t$$

à l'aide des lignes trigonométriques des multiples entiers de l'angle t.

6. Soient 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 ?

7. Démontrer les relations suivantes :

$$2^{2n}\cos^{2n}t = 2\cos 2nt + 2\binom{2n}{1}\cos(2n-2)t + \cdots + 2\binom{2n}{n-1}\cos 2t + \binom{2n}{n},$$

$$2^{2n}\cos^{2n+1}t = \cos(2n+1)t + \binom{2n+1}{1}\cos(2n-1)t + \cdots + \binom{2n+1}{n}\cos t,$$

$$2^{2n}\sin^{2n}t = 2\sum_{k=0}^{k=n-1}(-1)^{n+k}\binom{k}{2n}\cos 2(n-k)t + \binom{n}{2n},$$

$$2^{2n}\sin^{2n+1}t = \sum_{k=0}^{k=n}(-1)^{n+k}\binom{k}{2n+1}\sin(2n-2k+1)t,$$

$$\binom{1}{n} - \frac{1}{3}\binom{3}{n} + \frac{1}{9}\binom{5}{n} - \frac{1}{27}\binom{7}{n} + \cdots = \frac{2^n}{3^{\frac{n-1}{2}}}\sin\frac{n\pi}{6},$$

$$\cos\frac{\pi}{11} + \cos\frac{3\pi}{11} + \cos\frac{5\pi}{11} + \cos\frac{7\pi}{11} + \cos\frac{9\pi}{11} = \frac{1}{2}.$$

8. Soit $\mathbf{U}$ l'ensemble des nombres complexes z tels que $|z| = 1$; montrer que c'est un sous-groupe du groupe multiplicatif $\mathbf{C}^*$ des nombres complexes non nuls. Construire un isomorphisme entre les groupes $\mathbf{U} \times \mathbf{R}_+^*$ et $\mathbf{C}^*$. Montrer que $\mathbf{U}$ est isomorphe au groupe des rotations autour d'un point dans le plan.

9. On appelle **demi-plan de Poincaré** l'ensemble P des nombres complexes z tels que

$$\operatorname{Im}(z) > 0,$$

et **disque unité** l'ensemble D des nombres complexes z tels que

$$|z| < 1.$$

Montrer que

$$z \mapsto \frac{z-i}{z+i}$$

est une application bijective de P sur D (on désignera par A et B les points d'affixes i et $-i$,

par M le point d'affixe z, et on interprétera le module et l'argument de $z - i/z + i$ à l'aide des éléments du triangle MAB).

10. On désigne par a, b, c, d des nombres réels tels que $ad - bc = 1$.

a) Montrer que, pour tout nombre complexe z non réel, on a

$$\operatorname{Im}\left(\frac{az+b}{cz+d}\right) = \frac{\operatorname{Im}(z)}{|cz+d|^2}.$$

b) Soit P le demi-plan de Poincaré (*Exercice* 9). Montrer qu'il existe une permutation s de P telle que

$$s(z) = \frac{az+b}{cz+d}$$

pour tout $z \in P$, et que les permutations de P obtenues de cette façon (en faisant varier a, b, c, d) forment un groupe de transformations de l'ensemble P.

c) Montrer que l'application s de P dans P définie dans la question *b*) peut se décomposer en un produit de transformations appartenant à l'un des trois types suivants :

$$z \mapsto z + u \quad (u \text{ réel}); \qquad z \mapsto vz \quad (v \text{ réel strictement positif}); \qquad z \mapsto -1/z.$$

Interpréter géométriquement ces applications.

d) Quelle figure décrit $s(z)$ lorsque le point d'affixe z décrit soit un cercle contenu dans P, soit un demi-cercle contenu dans P et centré sur l'axe réel, soit une demi-droite contenue dans P, limitée à l'axe réel, et orthogonale à celui-ci?

11. Soit G l'ensemble des permutations du demi-plan de Poincaré P données par

$$s(z) = \frac{az+b}{cz+d}$$

où a, b, c, d sont des *entiers rationnels* tels que $ad - bc = 1$ (cf. *Exercice* **10**).

a) Montrer que G est un groupe de permutations de P (on l'appelle le **groupe modulaire arithmétique**) qui contient les applications u et v données par

$$u(z) = z + 1, \qquad v(z) = -1/z.$$

On se propose dans ce qui suit de démontrer que G *est engendré par u et v*. On note G_0 le sous-groupe de G engendré par u et v.

b) Pour tout $z \in P$, soit I_z l'ensemble des nombres $y > 0$ possédant la propriété suivante : il existe un $s \in G_0$ tel que $y = \operatorname{Im}(s(z))$. En utilisant la question *a*) de l'*Exercice* 10, montrer que pour tout $m > 0$ les éléments de I_z tels que $y > m$ sont en nombre fini.

c) Montrer que pour tout $z \in P$ il existe un $s \in G_0$ tel que le nombre $z_0 = s(z)$ vérifie

$$\operatorname{Im}(t(z_0)) \leqslant \operatorname{Im}(z_0) \qquad \text{pour tout } t \in G_0.$$

Montrer qu'on a les relations

$$|z_0| \geqslant 1, \qquad -\frac{1}{2} \leqslant \operatorname{Re}(z_0) \leqslant +\frac{1}{2}.$$

d) Soit D la partie de P définie par les inégalités précédentes (de sorte que l'orbite par G_0

de tout élément de P rencontre D). Soit z un point « intérieur » à D, i.e. tel que

$$|z| < 1, \qquad -\frac{1}{2} < \mathrm{Re}\,(z) < +\frac{1}{2};$$

montrer que le seul $s \in G$ tel que $s(z) \in D$ est l'élément neutre. Pour tout $t \in G$, montrer qu'il existe un $t' \in G_0$ tel que $t'(t(z)) \in D$. En déduire que $t \in G_0$, et donc que $G = G_0$ comme annoncé.

e) Montrer que, pour tout $s \in G$, l'ensemble $s(D)$ est un triangle (sic) limité par des demi-cercles orthogonaux à l'axe réel (certains de ces demi-cercles pouvant dégénérer en demi-droites), que le demi-plan P est réunion de $s(D)$, $s \in G$, et que deux domaines $s(D)$, $t(D)$ distincts (i.e. pour lesquels $s \neq t$) ne peuvent avoir en commun qu'un de leurs côtés. Tracer à la règle et au compas, *avec la plus grande exactitude possible*, la figure formée par ce « pavage » du demi-plan P — on partira de D, puis on construira les ensembles $u^n(D)$, puis les ensembles $vu^n(D)$, puis les ensembles $u^p v u^n(D)$, etc... [La figure obtenue, excessivement compliquée si on en poursuit la construction suffisamment loin, a servi de point de départ aux travaux de F. Klein et de H. Poincaré sur la théorie des « fonctions automorphes ».]

12. Soit K le sous-anneau de $\mathbf{C}$ formé des nombres de la forme $x + iy$ avec $x, y \in \mathbf{Z}$ (on les appelle des **entiers de Gauss**).

a) Soient $u, v \in K$ avec $v \neq 0$; on pose

$$u/v = x + iy$$

avec x, y rationnels, puis on détermine des entiers rationnels m et n tels que

$$|x - m| \leqslant \frac{1}{2}, \qquad |y - n| \leqslant \frac{1}{2};$$

enfin, on pose

$$q = m + in, \qquad r = u - qv.$$

Montrer qu'on a $N(r) < N(v)$ ou, si l'on préfère, $|r| < |v|$.

b) Déduire de là que tout idéal de l'anneau K est principal (imiter les raisonnements utilisés dans l'Exemple 8 du § 7).

13. Pour tout nombre premier p, on désigne par $\mathbf{F}_p$ le corps $\mathbf{Z}/p\mathbf{Z}$. Trouver les $d \in \mathbf{F}_p$ tels que l'anneau $\mathbf{F}_p[\sqrt{d}]$ soit un corps pour

$$p = 2, 3, 5, 7, 11.$$

Voyez-vous un rapport avec l'*Exercice* 11 du § 8?

14. Soient K un corps commutatif, u et v des éléments de K qui ne sont pas des carrés dans K. Pour qu'il existe un isomorphisme f du corps $K[\sqrt{u}]$ sur le corps $K[\sqrt{v}]$ tel que l'on ait $f(x) = x$ pour tout $x \in K$, il faut et il suffit que u/v soit le carré d'un élément de K.

Application : en considérant les corps construits dans l'*Exercice* 13, montrer que deux quelconques de ces corps sont isomorphes dès qu'ils ont le même nombre d'éléments.

(Ce résultat est un cas particulier du fait que deux corps finis ayant le même nombre d'éléments sont toujours isomorphes).

15. Soient K un corps commutatif et d un élément de K, qui n'est pas un carré dans K. On suppose que l'élément $2 = 1 + 1$ de K n'est pas nul (ce qui exclut par exemple le corps

$\mathbf{Z}/2\mathbf{Z}$). Montrer que, pour qu'un élément x du corps $K[\sqrt{d}]$, n'appartenant pas à K, soit un carré dans $K[\sqrt{d}]$, il faut et il suffit que $N(x)$ soit un carré dans K.
Application : on prend $K = \mathbf{Z}/11\mathbf{Z}$ et $d = 7$. Trouver tous les éléments $u + v\sqrt{d}$ ($u, v \in K$) qui sont des carrés dans $K[\sqrt{d}]$ (on pourra rassembler dans un tableau à double entrée les valeurs de $u^2 - 7v^2$ lorsque u et v décrivent K). Déduire de là des exemples de corps finis à 14 641 éléments... et vérifiez que tous les corps que vous aurez obtenus sont deux à deux isomorphes.

¶ **16.** Soit K un corps commutatif dans lequel -1 est un carré. Soit d un élément de K qui n'est pas un carré dans K. Montrer que $\sqrt{d}$ n'est pas un carré dans $K[\sqrt{d}]$.
Application : construire un corps à 17^4 éléments.

¶¶ **17.** Soit K un corps commutatif fini à q éléments; on suppose q impair.

a) Montrer que les éléments 1 et -1 de K sont distincts (observer que dans le cas contraire on aurait $2x = qx = 0$ pour tout $x \in K$, et noter que 2 et q sont premiers entre eux).

b) Montrer que $x \to x^2$ est un homomorphisme du groupe multiplicatif K* des éléments non nuls de K dans lui-même, dont le noyau se compose de 1 et -1. En déduire que les carrés non nuls dans K forment un sous-groupe H à

$$\frac{q-1}{2}$$

éléments de K* (utiliser l'*Exercice* 21 du § 7), et que le groupe quotient K*/H a deux éléments.

c) Montrer que, pour tout $x \in K^*$, on a

$$x^{\frac{q-1}{2}} = \begin{cases} 1 & \text{si } x \text{ est un carré} \\ -1 & \text{si } x \text{ n'est pas un carré.} \end{cases}$$

d) Soit p un nombre premier impair. On dit qu'un entier n non divisible par p est un **reste quadratique** modulo p s'il existe un $x \in \mathbf{Z}$ tel que

$$x^2 \equiv n \pmod{p}.$$

Montrer que les restes quadratiques se répartissent en $\frac{p-1}{2}$ classes modulo p, et que pour tout entier n non divisible par p on a

$$n^{\frac{p-1}{2}} \equiv \begin{cases} +1 \pmod{p} & \text{si } n \text{ est reste quadratique mod } p \\ -1 \pmod{p} & \text{si } n \text{ n'est pas reste quadratique mod } p. \end{cases}$$

(critère d'Euler). Montrer que -1 est reste quadratique mod p si et seulement si

$$p \equiv 1 \pmod{4}.$$

[L'étude, au XVIII^e^ siècle et au début du XIX^e^, de la théorie des restes quadratiques a été l'un des points de départ de toute l'Arithmétique et de toute l'Algèbre « modernes ». Le résultat le plus célèbre dans cette voie est la **loi de réciprocité quadratique** conjecturée par Euler, et démontrée par Gauss à l'âge de 19 ans — un beau début. Pour l'énoncer, on doit d'abord introduire le **symbole de Legendre**

$$\left(\frac{n}{p}\right) = \begin{cases} +1 & \text{si } n \text{ est reste quadratique mod } p \\ -1 & \text{sinon.} \end{cases}$$

La loi de réciprocité est alors que, si p et q sont des nombres premiers impairs et distincts, on a

$$\left(\frac{p}{q}\right)\cdot\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}$$

Pour une démonstration de ce résultat, et plus généralement pour tout ce qui concerne l'Arithmétique élémentaire, voir par exemple G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Oxford, 1938).

Après Gauss, les mathématiciens se sont efforcés d'étendre la loi de réciprocité quadratique aux anneaux d'entiers algébriques (§ 34, *Exercice* 48), ce qui est incomparablement plus difficile que le cas des entiers rationnels traité par Gauss. Les résultats complets, qui constituent la « théorie du corps de classe », après avoir été conjecturés par Hilbert, ont été obtenus vers 1920-1925 par Takagi, Artin et Hasse; la théorie a été grandement améliorée une dizaine d'années plus tard par Chevalley, et est encore actuellement l'objet de recherches très actives.

1. Montrer que, dans $\mathbf{R}^3$, le vecteur $x = (6, 2, -7)$ est combinaison linéaire des vecteurs

$$a = (2, 1, -3), \qquad b = (3, 2, -5), \qquad c = (1, -1, 1);$$

les vecteurs a, b, c forment-ils une base de $\mathbf{R}^3$?
Mêmes questions dans $\mathbf{R}^4$ pour les vecteurs $x = (7, 14, -1, 2)$ et

$$a = (1, 2, -1, -2), \qquad b = (2, 3, 0, -1), \qquad c = (1, 2, 1, 3), \qquad d = (1, 3, -1, 0).$$

Montrer que les vecteurs a, b, c forment une base de $\mathbf{R}^3$, et trouver les coordonnées du vecteur x par rapport à cette base, dans chacun des cas suivants :

$a = (1, 1, 1)$, $b = (1, 1, 2)$, $c = (1, 2, 3)$, $x = (6, 9, 14)$,
$a = (2, 1, -3)$, $b = (3, 2, -5)$, $c = (1, -1, 1)$, $x = (6, 2, -7)$.

Même question dans $\mathbf{R}^4$ pour les vecteurs

$$a = (1, 2, -1, -2), \qquad b = (2, 3, 0, -1), \qquad c = (1, 2, 1, 4), \qquad d = (1, 3, -1, 0)$$

et

$$x = (7, 14, -1, 2).$$

2. Pour que deux vecteurs (a, b) et (c, d) de $\mathbf{R}^2$ forment une base de $\mathbf{R}^2$ il faut et il suffit que

$$ad - bc \neq 0.$$

3. Soit M l'espace vectoriel réel formé des vecteurs d'origine donnée O dans l'espace usuel. Soit $M' \subset M$ l'ensemble des vecteurs d'origine O dont l'extrémité est située sur un plan donné dans l'espace. M' est-il un sous-espace vectoriel de M?

4. Soit K un anneau. On considère dans K^n l'ensemble M des vecteurs $(x_1, \ldots, x_n)$ vérifiant la relation

$$x_1 + \cdots + x_n = 0;$$

est-ce un sous-module de K^n? Même question en remplaçant la relation précédente par

$$x_1 + \cdots + x_n = 1.$$

5. On considère dans $\mathbf{C}^n$ des vecteurs

$$x_k = (\xi_{k1}, \ldots, \xi_{kn}) \quad (1 \leqslant k \leqslant r)$$

dont les composantes vérifient les inégalités

$$|\xi_{kk}| > \sum_{\substack{1 \leqslant j \leqslant r \\ j \neq k}} |\xi_{kj}| \quad (1 \leqslant k \leqslant r);$$

montrer que ces vecteurs sont linéairement indépendants sur $\mathbf{C}$.

6. Quel est le sous-espace vectoriel de $\mathbf{R}^4$ engendré par les vecteurs

$$a = (1, -1, 1, 0), \qquad b = (1, 1, 0, 1), \qquad c = (2, 0, 1, 1)?$$

7. Traduire les notions de module de type fini, de module libre de type fini, et de base d'un tel module, dans le langage des groupes commutatifs écrits multiplicativement (un tel groupe étant regardé comme un $\mathbf{Z}$-module conformément à l'*Exemple* 5 du § 10).

¶ **8.** Soit V un espace vectoriel de dimension finie sur le corps $\mathbf{Q}$ des nombres rationnels. On dit qu'une partie M de V est un **réseau** dans V si M est un sous-groupe de type fini du groupe additif V, et si M contient un système de générateurs de V.

a) Soit M un sous-groupe de type fini de V. Pour que M soit un réseau de V, il faut et il suffit que, pour tout $x \in V$, il existe un entier rationnel $r \neq 0$ tel que $rx \in M$. Montrer qu'alors l'ensemble des $r \in \mathbf{Z}$ tels que $rx \in M$ est un idéal (non nul) de l'anneau $\mathbf{Z}$.

b) Soient p un nombre premier et M un réseau de V. On note M_p l'ensemble des $x \in V$ vérifiant la condition suivante : il existe un entier r non multiple de p tel que $rx \in M$. Montrer que M_p est un sous-module de V regardé comme module sur l'anneau $\mathbf{Z}_p$ du § 8, *Exercice* 5.

c) Montrer que, si M est un réseau de V, on a

$$M = \bigcap_{p \text{ premier}} M_p.$$

d) Soient M et N deux réseaux de V. Montrer que les nombres premiers p tels que

$$M_p \neq N_p$$

sont en nombre fini.
(On trouvera des propriétés supplémentaires des réseaux dans l'*Exercice* 1 du § 18.)

9. Soit A un sous-anneau d'un corps commutatif K ; on suppose que tout élément de K puisse se mettre sous la forme uv^{-1} avec $u, v \in A$, $v \neq 0$, et que K soit un A-module de type fini. Montrer qu'alors $A = K$.

¶ **10.** Soient K un anneau, M un K-module à gauche et M′ un sous-module de M.

a) Montrer que

$$x - y \in M'$$

est une relation d'équivalence sur l'ensemble M [on l'appelle la **congruence modulo** M′ et on l'écrit souvent sous la forme

$$x \equiv y \pmod{M'};$$

voir § 7, n° 6].

b) On note M/M′ l'ensemble quotient de M par cette relation d'équivalence, et p l'application canonique de M sur M/M′. Montrer qu'il existe sur M/M′ une et une seule structure de K-module à gauche telle que l'on ait

$$p(x+y) = p(x) + p(y), \qquad p(\lambda x) = \lambda p(x)$$

quels que soient $x, y \in M$ et $\lambda \in K$ (utiliser le § 4, n° 3; voir aussi § 7, *Exercice* 16 et § 8, *Exercice* 7 pour des constructions analogues). On dit que M/M′, muni de cette structure de module, est le **module quotient** de M par le sous-module M′.

c) A chaque sous-module de M/M′ on associe son image réciproque par l'application p; montrer qu'on obtient ainsi une bijection de l'ensemble des sous-modules de M/M′ sur l'ensemble des sous-modules de M contenant M′.

d) Montrer que si M est de type fini il en est de même de M/M′ quel que soit M′. Si M est libre de type fini, en est-il de même de M/M′?

§ **11.** Soit M un module à gauche sur un anneau K.

a) Pour tout $x \in M$, on appelle **annulateur** de x l'ensemble des $\lambda \in K$ tels que $\lambda x = 0$. Montrer que c'est un idéal à gauche de K.

b) On suppose K intègre. Montrer que les $x \in M$ dont l'annulateur ne se réduit pas à 0 forment un sous-module T de M (on dit que T est le **sous-module de torsion** de M, et que M est **sans torsion** si $T = \{0\}$).

c) Montrer que le module quotient M/T (*Exercice* 10) est sans torsion.

d) Calculer T lorsque $K = \mathbf{Z}$ et $M = \mathbf{Z}^2/L$, où L est le sous-groupe de $\mathbf{Z}^2$ engendré par le vecteur (4,6).

§ **12.** On dit qu'un module à gauche M sur un anneau K est **de torsion** si, pour tout $x \in M$, il existe un scalaire *non nul* $\lambda \in K$ tel que $\lambda x = 0$.

a) Soient M un K-module à gauche et M′ un sous-module de K. On suppose que M′ et M/M′ sont des modules de torsion. Montrer que, si K est intègre, M est alors un module de torsion.

b) On suppose $K = \mathbf{Z}$. Montrer que, pour qu'un K-module de type fini soit de torsion, il faut et il suffit qu'il soit fini.

13. Soit M un module sur un anneau K. Une partie B de M (finie ou non) est appelée un **ensemble de générateurs** de M si le seul sous-module de M contenant B est M tout entier ou, ce qui revient au même, si chaque élément de M est combinaison linéaire d'éléments de B en nombre fini.
On suppose M de type fini. Montrer que l'on peut alors extraire de B un ensemble *fini* de générateurs de M (choisir dans M un système fini de générateurs x_i, et exprimer chaque x_i à l'aide d'éléments de B).

§§ **14.** Soient K un corps commutatif et A un sous-anneau de K; on suppose que K est le corps des fractions de A i.e. que, pour tout $x \in K$, il existe des éléments u et v de A tels que $v \neq 0$ et

$$x = uv^{-1}.$$

Dans ce qui suit on regarde K comme un A-module, et on dit qu'une partie I de K est un **idéal fractionnaire** de l'anneau A si elle ne se réduit pas à 0, si c'est un sous-module de K, et si, enfin, il existe un élément $d \neq 0$ de K tel que l'on ait

$$dI \subset A$$

(où dI désigne l'ensemble des produits dx avec $x \in I$).

a) Pour qu'une partie I de K soit un idéal fractionnaire il faut et il suffit qu'il existe un idéal non nul J de l'anneau A et un $d \in A$ non nul tels que

$$I = d^{-1}J.$$

b) Soient I et J des idéaux fractionnaires, et soit (I : J) l'ensemble des $x \in K$ tels que $xJ \subset I$; montrer que c'est un idéal fractionnaire (souvent appelé le **transporteur** de J dans I).

c) Soient I et J deux idéaux fractionnaires; on note I + J l'ensemble des sommes $x + y$ avec $x \in I$ et $y \in J$, et IJ l'ensemble des éléments de K qui peuvent s'écrire comme somme (finie) de produits xy avec $x \in I$ et $y \in J$ (cf. § 8, *Exercice* 10 pour le cas où $I, J \subset A$). Montrer que I + J, IJ et $I \cap J$ sont des idéaux fractionnaires de l'anneau A. Étendre les formules du § 8, *Exercice* 10, (*a*).

d) On dit qu'un idéal fractionnaire I est **inversible** s'il existe un idéal fractionnaire J tel que

$$I.J = A;$$

montrer qu'alors J est unique, et donné par la relation

$$J = (A : I)$$

(on dit alors que J est l'**inverse** de I, et on le note I^{-1}). Autrement dit, pour que I soit inversible, il faut et il suffit que

$$(A : I).I = A.$$

e) Pour qu'un idéal I soit inversible il faut et il suffit qu'il existe des éléments $x_k \in I$ et $y_k \in (A : I)$ en nombre fini, tels que

$$1 = \sum x_k y_k;$$

les x_k forment alors un système de générateurs du A-module I (un idéal fractionnaire inversible est donc de type fini — on convient de dire qu'un idéal fractionnaire est **de type fini** s'il est de type fini comme A-module).

f) On dit que A est un **anneau de Dedekind** si tout idéal fractionnaire de A est inversible. Montrer que l'ensemble des idéaux fractionnaires de A, muni de la loi de composition $(I, J) \to I.J$, est alors un groupe.

g) Si A est un anneau de Dedekind, tout idéal *premier* non nul de A est *maximal*.

[Les anneaux de Dedekind se sont introduits d'abord dans la théorie des nombres algébriques — cf. § 34, *Exercice* — et on n'en a donné une définition générale et abstraite que beaucoup plus tard. La principale propriété de ces anneaux, est que *dans un anneau de Dedekind tout idéal s'écrit, d'une façon unique à des permutations près, sous la forme d'un produit d'idéaux premiers*; cf. § 18, *Exercice* 7. Inversement, cette propriété caractérise les anneaux de Dedekind. Une troisième caractérisation, beaucoup plus maniable dans la pratique, sera donnée au § 34, *Exercice* 50. Notons enfin que les anneaux de Dedekind interviennent non seulement dans la théorie des nombres algébriques, mais aussi dans l'étude des courbes algébriques et en beaucoup d'autres questions de Géométrie algébrique; c'est ce qui justifie l'introduction et l'étude des anneaux de Dedekind « abstraits ».

Comme exemple aussi élémentaire que possible d'un anneau de Dedekind, mis à part bien entendu l'anneau **Z**, citons les anneaux $\mathbf{Z}[\sqrt{d}]$ où

$$d \equiv 2 \text{ ou } 3 \pmod 4.]$$

¶ **15.** Soient K un anneau et X un ensemble. On désigne par

$$K^{(X)}$$

l'ensemble de toutes les applications

$$u : X \to K$$

telles que les $x \in X$ vérifiant $u(x) \neq 0$ soient en nombre *fini*. Montrer que $K^{(X)}$ est un sous-module du K-module à gauche K^X (formé de toutes les applications de K dans X). Pour chaque $x \in X$, on considère l'élément e_x de $K^{(X)}$ défini par

$$e_x(y) = \begin{cases} 1 & \text{si } y = x \\ 0 & \text{si } y \neq x; \end{cases}$$

montrer que la famille $(e_x)_{x \in X}$ est une *base* du K-module à gauche $K^{(X)}$, et que les composantes par rapport à cette base de tout $u \in K^{(X)}$ sont les scalaires $u(x)$, autrement dit qu'on a

$$u = \sum_{x \in X} u(x) . e_x$$

pour tout $u \in K^{(X)}$.
Soit f une application de X dans un K-module à gauche M; montrer qu'il existe un et un seul homomorphisme

$$\bar{f} : K^{(X)} \to M$$

tel que l'on ait

$$\bar{f}(e_x) = f(x) \quad \text{pour tout } x \in X.$$

(Les éléments du module $K^{(X)}$ sont généralement appelés les **combinaisons linéaires formelles** d'éléments de X à coefficients dans K, et on identifie le plus souvent l'élément de base e_x de ce module à l'élément $x \in X$ correspondant).

¶ **16.** Soient K et L deux anneaux commutatifs, et $j_1, \ldots, j_n$ des homomorphismes deux à deux distincts de K dans L. Montrer que $j_1, \ldots, j_n$ sont linéairement indépendants dans le L-module de toutes les applications de K dans L (théorème de Dedekind) (écrire une relation linéaire entre $j_1, \ldots, j_n$ et utiliser l'identité $j_k(xy) = j_k(x) j_k(y)$ pour se ramener au cas de $n - 1$ homomorphismes).

¶ **17.** Soient G un groupe commutatif, K un anneau commutatif, et $\rho_1, \ldots, \rho_n$ des homomorphismes deux à deux distincts de G dans le groupe multiplicatif K^*. Montrer que $\rho_1, \ldots, \rho_n$ sont linéairement indépendants sur K, i.e. que si $a_1, \ldots, a_n \in K$ vérifient

$$a_1\rho_1(s) + \ldots + a_n\rho_n(s) = 0 \quad \text{pour tout } s \in G,$$

alors $a_1 = \ldots = a_n = 0$ (raisonner par récurrence sur n).
Exemple : soient $c_1, \ldots, c_n$ des nombres complexes deux à deux distincts; on considère les fonctions

$$e^{c_1 t}, \ldots, e^{c_n t}$$

pour $t \in \mathbf{R}$; montrer qu'elles sont linéairement indépendantes sur $\mathbf{C}$.

1. On considère les matrices (à coefficients dans $\mathbf{C}$)

$$I_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad I_3 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$I_4 = \begin{pmatrix} 0 & i \\ 0 & 0 \end{pmatrix}, \quad I_5 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad I_6 = \begin{pmatrix} 0 & 0 \\ i & 0 \end{pmatrix}.$$

Établir les quinze relations suivantes :

$$\begin{array}{lll} [I_1, I_3] = 2I_3, & [I_1, I_4] = 2I_4, & [I_2, I_3] = 2I_4 \\ [I_1, I_5] = -2I_5, & [I_1, I_6] = -2I_6, & [I_2, I_5] = -2I_6 \\ [I_3, I_5] = I_1, & [I_3, I_6] = I_2, & [I_4, I_5] = I_2 \\ [I_2, I_4] = -2I_3, & [I_2, I_6] = 2I_5, & [I_4, I_6] = -I_1 \\ [I_1, I_2] = 0, & [I_3, I_4] = 0, & [I_5, I_6] = 0, \end{array}$$

où l'on pose d'une façon générale

$$[X, Y] = XY - YX.$$

Trouver toutes les matrices carrées d'ordre 2 qui commutent aux six matrices $I_1, \ldots, I_6$.

2. Établir les formules de l'*Exercice* précédent pour les matrices

$$I_1 = 2\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad I_2 = 2\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix},$$
$$I_4 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad I_5 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad I_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

3. Trouver toutes les matrices carrées d'ordre 3 commutant à la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 1 & 2 \end{pmatrix}$$

(on prendra comme anneau de base soit le corps **C**, soit un corps commutatif quelconque, soit un anneau arbitraire — au choix...).

4. Soit K un anneau commutatif. Montrer que l'application de $M_2(K)$ dans $M_4(K)$ qui transforme chaque matrice $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ en la matrice

$$\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix}$$

est un isomorphisme de l'anneau $M_2(K)$ sur un sous-anneau de $M_4(K)$. Interprétation en termes de modules?

5. Calculer le produit des trois matrices

$$\begin{pmatrix} 0 & 2 & -1 \\ -2 & -1 & 2 \\ 3 & -2 & -1 \end{pmatrix}, \quad \begin{pmatrix} 70 & 34 & -107 \\ 52 & 26 & -68 \\ 101 & 50 & -140 \end{pmatrix}, \quad \begin{pmatrix} 27 & -18 & 10 \\ -46 & 31 & -17 \\ 3 & 2 & 1 \end{pmatrix}.$$

Effectuer le même calcul en prenant pour anneau de base l'anneau $\mathbf{Z}/7\mathbf{Z}$ des entiers modulo 7.

6. Calculer le cube de la matrice carrée d'ordre n

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

7. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice carrée d'ordre 2 à coefficients dans un anneau commutatif quelconque. Montrer qu'on a la relation

$$A^2 - (a + d)A + (ad - bc)1_2 = 0.$$

8. Étant donnée une matrice carrée

$$A = (x_{ij})_{1 \leqslant i,j \leqslant n}$$

à coefficients dans un anneau *commutatif* K, on appelle **trace** de A le scalaire

$$\mathrm{Tr}(A) = x_{11} + x_{22} + \dots + x_{nn},$$

somme des éléments diagonaux de A. Montrer qu'on a

$$\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B), \qquad \mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

quelles que soient les matrices A et B.
On suppose $K = \mathbf{C}$. Déduire de ce qui précède qu'il est impossible de trouver des matrices carrées X et Y d'ordre n telles que

$$XY - YX = 1_n.$$

9. Soient K un anneau commutatif et d un élément de K. Montrer que les matrices

$$\begin{pmatrix} x & dy \\ y & x \end{pmatrix}$$

(où x et y sont des éléments arbitraires de K) forment un sous-anneau L de $M_2(K)$, et que L est isomorphe à l'anneau $K[\sqrt{d}]$ du § 9. Cas où $K = \mathbf{R}$, $d = -1$?

10. On dit qu'une matrice carrée X d'ordre n à coefficients dans un anneau K est **nilpotente** s'il existe un entier $r \geqslant 1$ tel que $X^r = 0$, et **unipotente** si la matrice $1_n - X$ est nilpotente. On suppose $K = \mathbf{C}$ dans ce qui suit. Étant données une matrice carrée nilpotente N, et une matrice carrée unipotente U, on pose (cf. § 8, *Exercice* 2)

$$\exp(N) = 1 + \frac{N}{1!} + \frac{N^2}{2!} + \dots \frac{N^k}{k!} + \dots,$$

$$\log(U) = -\frac{1-U}{1} - \frac{(1-U)^2}{2} - \dots - \frac{(1-U)^k}{k} - \dots.$$

On prend

$$N = \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix};$$

vérifier que N est nilpotente, que U est unipotente, et qu'on a les relations

$$\exp(\log(U)) = U, \qquad \log(\exp(N)) = N$$

[on calculera effectivement les matrices $\exp(\log(U))$ et $\log(\exp(N))$, sans utiliser le résultat général de l'*Exercice* 2 du § 8].

11. Pour tout nombre complexe t, on pose

$$U(t) = \begin{pmatrix} 1 & t & 2t + 2t^2 & 3t + \frac{17}{2}t^3 + 4t^4 \\ 0 & 1 & 4t & 5t + 12t^3 + 3t^4 \\ 0 & 0 & 1 & 6t \\ 0 & 0 & 0 & 1 \end{pmatrix};$$

montrer qu'on a

$$U(s)\,U(t) = U(s+t)$$

quels que soient $s, t \in \mathbf{C}$ et que $U(t) = \exp(tN)$ où N est une matrice nilpotente qu'on calculera.

¶ **12.** Soit z un nombre algébrique, i.e. (§ 11, *Exemple* 11) un nombre complexe racine d'une équation algébrique à coefficients rationnels non tous nuls.

a) Montrer qu'il existe un entier $n \geqslant 1$ et des nombres rationnels $a_0, \dots, a_{n-1}$ tels que l'on ait une relation de la forme

$$z^n = a_0 + a_1 z + \dots + a_{n-1} z^{n-1}.$$

b) Soit K le sous-anneau de $\mathbf{C}$ engendré par $\mathbf{Q}$ et z; montrer que K, considéré comme espace vectoriel sur $\mathbf{Q}$, est engendré par les éléments $1, z, \dots, z^{n-1}$.

c) On suppose n minimum dans ce qui précède. Montrer qu'alors $1, z, \dots, z^{n-1}$ forment une base de K regardé comme espace vectoriel sur $\mathbf{Q}$.

d) Soit f l'application de K dans K donnée par

$$f(u) = zu \qquad \text{pour tout } u \in K.$$

Montrer que c'est un endomorphisme de K regardé comme espace vectoriel sur $\mathbf{Q}$. Calculer la matrice de f par rapport à la base de K définie dans la question *c*).

¶ **13.** Soient L et M deux modules à gauche sur un anneau K ; on suppose donnés un sous-module L′ de L et un homomorphisme f de L dans M. Prouver que les deux conditions suivantes sont équivalentes : *a*) L′ est contenu dans le noyau de f, i.e. on a $f(x) = 0$ pour tout $x \in L'$; *b*) f est composé de l'application canonique de L sur le module quotient L/L′ (§ 10, *Exercice* 10) et d'un homomorphisme de L/L′ dans M.

14. Soient K un anneau, L, M et N trois K-modules à gauche, f un homomorphisme de L dans M, et p un homomorphisme de L dans N ; on suppose p surjectif. Montrer que les deux conditions suivantes sont équivalentes : *a*) on a $\mathrm{Ker}(f) \supset \mathrm{Ker}(p)$; *b*) f est composé de p et d'un homomorphisme de N dans M.

15. Soient L l'espace vectoriel réel formé des vecteurs d'origine donnée O dans l'espace usuel, et L′ le sous-espace vectoriel de L formé des vecteurs portés par une droite donnée (resp. un plan donné) passant par O. Pour tout $x \in L$, on note $f(x)$ le vecteur projection orthogonale de x sur L′. Montrer que l'application f de L dans L′ est linéaire. Quel est son noyau ?

¶¶ **16.** Soit K un anneau. On dit qu'un K-module à gauche M est **simple** ou **irréductible** s'il n'est pas réduit à 0 et si les seuls sous-modules de M sont $\{0\}$ et M tout entier.

a) Pour que M, supposé non nul, soit simple, il faut et il suffit que, quels que soient $a, b \in M$ avec $a \neq 0$, il existe un $\lambda \in K$ tel que $b = \lambda a$. En déduire que, si K est un *corps*, tout K-module simple est isomorphe à K.

b) Soit M un K-module simple. On choisit dans M un élément $a \neq 0$, et on note I l'ensemble des $\lambda \in K$ tels que $\lambda a = 0$. Montrer que I est un idéal à gauche *maximal* de K (§ 8, *Exercice* 7). Montrer que l'application $\lambda \to \lambda a$ de K dans M est composée de l'application canonique de K sur K/I, et d'un isomorphisme du K-module à gauche K/I sur M.
Montrer inversement que, pour tout idéal à gauche maximal I de K, le K-module à gauche K/I est simple.

c) Soient L et M deux K-modules à gauche simples, et f un homomorphisme de L dans M. Montrer que, si f n'est pas nul, c'est un isomorphisme de L sur M (lemme de Schur). (On examinera le noyau et l'image de f.)

d) Soit L un K-module à gauche simple. Montrer que l'anneau des endomorphismes de L est un *corps* (en général non commutatif).

1. L'anneau de base étant **R**, trouver les inverses des matrices suivantes (*) :

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}; \quad \begin{pmatrix} 2 & 5 & 7 \\ 6 & 3 & 4 \\ 5 & -2 & -3 \end{pmatrix}; \quad \begin{pmatrix} 1 & 2 & -3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}; \quad \begin{pmatrix} 2 & 2 & 3 \\ 1 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}.$$

2. Calculer l'inverse de la matrice

$$\begin{pmatrix} 1 & a & a^2 & \dots & a^n \\ 0 & 1 & a & \dots & a^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}.$$

3. Soit N une matrice carrée nilpotente (i.e. dont une puissance est nulle) à coefficients dans un anneau. Montrer que la matrice $1 - N$ est inversible, et que

$$(1 - N)^{-1} = 1 + N + N^2 + \cdots$$

Application : calculer l'inverse de la matrice

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

4. Calculer l'inverse de la matrice

$$\begin{pmatrix} 1 + a_1 & 1 & 1 & \dots & 1 \\ 1 & 1 + a_2 & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 + a_n \end{pmatrix}$$

(*) Le lecteur plus avancé pourra aussi résoudre cet *Exercice* en utilisant les formules de Cramer.

¶ 5. Soit $\omega = \cos(2\pi/n) + i.\sin(2\pi/n)$. Montrer que l'inverse de la matrice

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \dots & \omega^{(n-1)^2} \end{pmatrix}$$

s'obtient en remplaçant ω par ω^{-1} dans cette matrice, et en divisant par n la matrice ainsi obtenue.

6. Trouver une matrice carrée X d'ordre 3 telle que (*)

$$\begin{pmatrix} 2 & -3 & 1 \\ 4 & -5 & 2 \\ 5 & -7 & 3 \end{pmatrix} X \begin{pmatrix} 9 & 7 & 6 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 18 & 12 & 9 \\ 23 & 15 & 11 \end{pmatrix}.$$

7. Montrer que les vecteurs

$$(1, 2, 1), \quad (2, 3, 3), \quad (3, 7, 1)$$

forment une base de $\mathbf{R}^3$, ainsi que les vecteurs

$$(3, 1, 4), \quad (5, 2, 3), \quad (1, 1, -6),$$

et calculer la matrice de passage de la première base à la seconde.
Même problème, dans $\mathbf{R}^4$, pour les vecteurs

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

et

$$(1, 0, 3, 3), \quad (-2, -3, -5, -4), \quad (2, 2, 5, 4), \quad (-2, -3, -4, -4).$$

8. Les matrices

$$\begin{pmatrix} x+y & 4y \\ -y & x-y \end{pmatrix}$$

où x et y sont des nombres rationnels arbitraires, forment un sous-corps de l'anneau $M_2(\mathbf{Q})$.

9. Les matrices

$$\begin{pmatrix} x & y & z \\ 2z & x & y \\ 2y & 2z & x \end{pmatrix}$$

où x, y, z sont des nombres rationnels, forment un sous-corps de $M_3(\mathbf{Q})$.

¶ 10. Soient K un anneau commutatif, p et q deux éléments donnés de K. On considère l'anneau

$$L = K[\sqrt{p}]$$

et l'ensemble $M \subset M_2(L)$ des matrices de la forme

$$z = \begin{pmatrix} x & qy \\ \bar{y} & \bar{x} \end{pmatrix} \quad \text{avec } x, y \in L$$

(pour les notations, voir le § 9).

(*) Voir note page précédente.

a) Montrer que M est un sous-anneau de $M_2(L)$.

b) Pour toute matrice

$$z = \begin{pmatrix} x & qy \\ \bar{y} & \bar{x} \end{pmatrix} \tag{1}$$

de M on pose

$$z^* = \begin{pmatrix} \bar{x} & -qy \\ -\bar{y} & x \end{pmatrix};$$

montrer qu'on a $(z_1z_2)^* = z_2^*z_1^*$ quels que soient $z_1, z_2 \in M$.

c) Pour la matrice (1), calculer le produit z^*z, et montrer que z est un élément inversible de l'anneau M si et seulement si

$$N(z) = \bar{x}x - q\bar{y}y$$

est un élément inversible de l'anneau L.

d) On suppose que K soit un corps commutatif et que p ne soit pas un carré dans K. Montrer que les assertions suivantes sont équivalentes : (*i*) l'anneau M est un corps (*ii*) il n'existe aucun couple d'éléments x, y de K tels que

$$q = x^2 - py^2.$$

e) On suppose $K = \mathbf{R}$. Montrer que M est un corps si et seulement si l'on a

$$p < 0, \qquad q < 0.$$

Montrer que le corps M ainsi obtenu est isomorphe à celui qu'on obtiendrait en prenant $p = q = -1$ (et qu'on appelle le **corps des quaternions**, premier exemple, historiquement, d'un corps non commutatif).

f) On suppose $K = \mathbf{Q}$ et p et q entiers. Montrer que, pour que M soit un corps, il faut et il suffit que l'équation

$$px^2 + qy^2 = z^2$$

ne possède aucune solution (x, y, z) *entière* autre que $(0, 0, 0)$. Montrer que cette condition est satisfaite dans les cas suivants par exemple :

$$p = 5, \quad q \equiv 2 \pmod 5; \qquad p = 5, \quad q \equiv 3 \pmod 5;$$
$$p = 11, \quad q \equiv 2, 6, 7, 8 \text{ ou } 10 \pmod{11}.$$

11. Montrer que les matrices de la forme

$$\begin{pmatrix} x & -y & -z & -t \\ y & x & -t & z \\ z & t & x & -y \\ t & -z & y & x \end{pmatrix},$$

où x, y, z, t sont des nombres réels arbitraires, forment un sous-corps de $M_4(\mathbf{R})$, et que ce sous-corps est isomorphe au corps des quaternions défini dans l'*Exercice* précédent. Montrer que, considéré comme espace vectoriel réel, ce corps admet une base formée de quatre éléments e, i, j, k vérifiant les formules suivantes :

$$e^2 = e, \qquad i^2 = j^2 = k^2 = -e,$$
$$ei = ie = i, \qquad ej = je = j, \qquad ek = ke = k,$$
$$ij = -ji = k; \qquad jk = -kj = i; \qquad ki = -ik = j.$$

Obtiendrait-on encore un corps si l'on autorisait les variables x, y, z, t à prendre des valeurs complexes quelconques ?

12. L'anneau de base étant $\mathbf{C}$, calculer l'inverse de la matrice

$$\begin{pmatrix} 1 & 1+i & -i \\ 0 & i & 1-2i \\ 1 & 1 & i \end{pmatrix},$$

13. On considère la matrice $U(t)$ de l'*Exercice* 11 des §§ 12, 13, 14. Calculer son inverse en effectuant le moins possible de calculs.

14. Soit K un anneau commutatif. Montrer que les matrices de la forme

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

$(x, y, z \in \mathrm{K})$ forment un sous-groupe de $\mathrm{GL}(3, \mathrm{K})$. Déterminer le centre de ce sous-groupe.

¶ 15. Soient K un anneau commutatif et n un entier. Montrer que les matrices carrées d'ordre n, à coefficients dans K, et de la forme

$$\begin{pmatrix} 1 & * & * & \dots & * \\ 0 & 1 & * & \dots & * \\ 0 & 0 & 1 & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

(où les signes * désignent des éléments arbitraires de K) forment un sous-groupe de $\mathrm{GL}(n, \mathrm{K})$, dont on déterminera le centre.

16. Soit K un corps. On désigne par H le sous-groupe de $\mathrm{GL}(n, \mathrm{K})$ formé des matrices diagonales de $\mathrm{GL}(n, \mathrm{K})$. Trouver le normalisateur (§ 7, *Exercice* 13) de H dans $\mathrm{GL}(n, \mathrm{K})$.

¶ 17. Pour que des éléments (a, b) et (c, d) de $\mathbf{Z}^2$ forment une base de $\mathbf{Z}^2$ il faut et il suffit que $ad - bc = +1$ ou -1.

18. Soient K un anneau commutatif et I un idéal de K. Soit H l'ensemble des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, à coefficients dans K, qui vérifient

$$ad - bc = 1, \qquad a \equiv d \equiv 1 \pmod{\mathrm{I}}, \qquad b \equiv c \equiv 0 \pmod{\mathrm{I}}.$$

Montrer que H est un sous-groupe invariant de $\mathrm{GL}(2, \mathrm{K})$.

¶¶ 19. Soit K un corps fini à q éléments. Calculer le nombre d'éléments du groupe $\mathrm{GL}(n, \mathrm{K})$.

¶ 20. Pour tout entier $n \geqslant 1$, on note G_n l'ensemble des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec

$$a, b, c, d \in \mathbf{Z}, \qquad ad - bc = n,$$

et on pose $\mathrm{G}_1 = \mathrm{G}$.

a) Montrer que G est un sous-groupe de $\mathrm{GL}(2, \mathbf{Z})$. En est-il de même de G_n pour $n \geqslant 2$?

b) Montrer que si $\mathrm{X} \in \mathrm{G}_n$ on a $\mathrm{UXV} \in \mathrm{G}_n$ quelles que soient $\mathrm{U}, \mathrm{V} \in \mathrm{G}$.

c) Montrer que, pour toute $\mathrm{X} \in \mathrm{G}_n$, il existe $\mathrm{U}, \mathrm{V} \in \mathrm{G}$ telles que UXV soit diagonale.

¶ **17.** Soit $u = (a, b)$ un élément du $\mathbf{Z}$-module $\mathbf{Z}^2$.

a) Montrer que s'il existe une base de $\mathbf{Z}^2$ qui contient u, alors il existe une forme linéaire sur le module considéré telle que $f(u) = 1$. En déduire qu'alors les entiers a et b sont premiers entre eux.

b) On suppose inversement a et b premiers entre eux. Montrer qu'il existe une forme linéaire f sur $\mathbf{Z}^2$ telle que $f(u) = 1$. Montrer qu'il existe un vecteur v tel que $\mathrm{Ker}(f)$ soit l'ensemble des multiples entiers de v. Prouver que les vecteurs u et v forment une base de $\mathbf{Z}^2$.

c) On prend $u = (6, 35)$; trouver un vecteur v tel que u et v forment une base de $\mathbf{Z}^2$.

18. Soient L et M deux modules à gauche sur un anneau quelconque K, et

$$f : \mathrm{M} \to \mathrm{L}$$

un homomorphisme *surjectif*. On suppose L libre de type fini. Montrer qu'il existe un homomorphisme

$$g : \mathrm{L} \to \mathrm{M}$$

tel que

$$f \circ g = id.$$

¶ **19.** Soient L et M deux modules à gauche sur un anneau, L′ et M′ des sous-modules de L et M et u un homomorphisme de L dans M tel que $u(\mathrm{L}') \subset \mathrm{M}'$. On note p et q les applications canoniques de L sur L/L′ et de M sur M/M′ (§§ 10, 11, *Exercice* 10). Montrer qu'il existe un et un seul homomorphisme

$$\bar{u} : \mathrm{L}/\mathrm{L}' \to \mathrm{M}/\mathrm{M}'$$

tel que l'on ait

$$q \circ u = \bar{u} \circ p$$

(on dit que $\bar{u}$ est l'homomorphisme déduit de u par passage aux quotients). A quelles conditions $\bar{u}$ est-il injectif, ou surjectif, ou bijectif?

20. Soit k un corps commutatif. Dans le groupe $\mathrm{SL}(2, k)$ des matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ à coefficients dans k et telles que $ad - bc = 1$, on considère les matrices

$$x_-(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \qquad x_-(t) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$$

où $t \in k$. Désignant par α l'un des deux symboles $+$ ou $-$, et par $-\alpha$ le symbole opposé, on définit pour $t \neq 0$ les matrices

$$w_\alpha(t) = x_\alpha(t) x_{-\alpha}(-t^{-1}) x_\alpha(t), \qquad h_\alpha(t) = w_\alpha(t) w_\alpha(1)^{-1}.$$

a) Calculer ces matrices, et montrer que les matrices $x_+(t)$ et $x_-(t)$ engendrent $\mathrm{SL}(2, k)$.

b) Établir les relations suivantes :

(R 1)	$x_\alpha(t + u) = x_\alpha(t)\, x_\alpha(u)$	pour	$t, u \in k$;
(R 2)	$w_\alpha(t)\, x_\alpha(u)\, w_\alpha(t)^{-1} = x_{-\alpha}(-u t^2)$	pour	$t, u \in k$, $t \neq 0$;
(R 3)	$h_\alpha(tu) = h_\alpha(t)\, h_\alpha(u)$	pour	$t, u \in k$, $t \neq 0$, $u \neq 0$.

c) Montrer que toute matrice $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, k)$ peut se mettre d'une façon et d'une seule soit sous la forme $g = h_+(t) x_-(u)$, soit sous la forme $g = h_+(t)\, x_+(u)\, w x_+(v)$, où l'on pose $w = w_+(1)$ (distinguer deux cas, suivant que $c = 0$ ou que $c \neq 0$).

d) On appelle *groupe dérivé* d'un groupe G le sous-groupe G′ de G engendré par les « commu

lateurs » $(x, y) = xyx^{-1}y^{-1}$ de ses éléments. Montrer que $\mathrm{SL}(2, k)$ est égal à son groupe dérivé pourvu que k contienne au moins 4 éléments.

¶ **21.** On considère dans $\mathrm{SL}(2, k) = \mathrm{G}$ le sous-groupe U formé des $x_+(t)$, et le sous-groupe H formé des $h_-(t)$; on a donc $\mathrm{G} = \mathrm{HU} \cup \mathrm{HU}u\mathrm{U}$ d'après la question *c*) de l'*exercice* précédent. On se propose de montrer que *si k possède au moins 4 éléments le groupe* G *contient un seul sous-groupe invariant* M *autre que* $\{e\}$ *et* G, *à savoir le sous-groupe* Z *formé des matrices* 1 *et* -1 (qui se réduit d'ailleurs à l'élément neutre en caractéristique 2; noter que, dans tous les cas, ce sous-groupe est le centre de G, comme on le vérifiera facilement). On pose $\mathrm{B} = \mathrm{HU}$.

a) Montrer que $\mathrm{B} \cap u\mathrm{B}u^{-1} = \mathrm{H}$, et en déduire que si $\mathrm{M} \subset \mathrm{B}$ alors on a $\mathrm{M} \subset \mathrm{H}$.

b) On suppose $\mathrm{M} \subset \mathrm{H}$. Supposons $h = h_+(t) \in \mathrm{M}$. En calculant le commutateur de h et de $x_-(u)$ montrer que M contient $x_-(u - u/t^2)$, et en déduire que $\mathrm{M} \subset \mathrm{Z}$.

c) On suppose M non contenu dans B. Montrer, à l'aide de la question *c*) de l'*exercice* précédent, que $\mathrm{G} = \mathrm{MB}$, et en déduire (§ 7, exercice 16) que le groupe quotient G/M est isomorphe à $\mathrm{B}/\mathrm{B} \cap \mathrm{M}$.

d) On suppose $\mathrm{Card}(k) \geqslant 4$, et donc $\mathrm{G} = \mathrm{G}'$ d'après l'*exercice* précédent. Montrer que le groupe G/M est égal à son groupe dérivé. Déduire de là et de la question précédente que $\mathrm{M} = \mathrm{G}$ si M n'est pas contenu dans B (remarquer que $\mathrm{B}' \subset \mathrm{U}$ et que $\mathrm{U}' = \{e\}$, et observer que le seul sous-groupe invariant L de B tel que B/L soit son propre groupe dérivé est B lui-même).

¶ **22.** Dans un groupe G, on considère deux familles d'éléments que l'on notera $\hat{x}_+(t)$ et $\hat{x}_-(t)$ et qui dépendent d'un paramètre t qui varie dans un corps commutatif k donné. A partir des éléments $\hat{x}_\alpha(t)$ de G on définit des éléments $\hat{w}_\alpha(t)$ et $\hat{h}_\alpha(t)$ de G par les formules de l'*exercice* 20, i.e. en posant

$$\hat{w}_\alpha(t) = \hat{x}_\alpha(t)\hat{x}_{-\alpha}(-1/t)\hat{x}_\alpha(t), \qquad \hat{h}_\alpha(t) = \hat{w}_\alpha(t)\hat{w}_\alpha(1)^{-1},$$

et on suppose vérifiées les relations

$$(\mathrm{R}\,1)\quad \hat{x}_\alpha(t+u) = \hat{x}_\alpha(t)\hat{x}_\alpha(u), \qquad (\mathrm{R}\,2)\quad \hat{w}_\alpha(t)\hat{x}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{x}_{-\alpha}(-u/t^2)$$

de l'*exercice* 20.

a) Démontrer les formules suivantes :

$$\hat{w}_\alpha(t)\hat{w}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{w}_{-\alpha}(-u/t^2), \qquad \hat{w}_\alpha(t)\hat{h}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{h}_{-\alpha}(-u/t^2)\hat{h}_{-\alpha}(-1/t^2)^{-1},$$
$$\hat{w}_\alpha(t)\hat{x}_\alpha(u)\hat{w}_\alpha(t)^{-1} = \hat{x}_{-\alpha}(-u/t^2), \qquad \hat{h}_\alpha(t)\hat{x}_\alpha(u)\hat{h}_\alpha(t)^{-1} = \hat{x}_\alpha(t^2u),$$
$$\hat{h}_\alpha(t)\hat{w}_\alpha(u)\hat{h}_\alpha(t)^{-1} = \hat{w}_\alpha(t^2u), \qquad \hat{h}_\alpha(t)\hat{h}_\alpha(u)\hat{h}_\alpha(t)^{-1} = \hat{h}_\alpha(t^2u)\hat{h}_\alpha(t^2)^{-1},$$
$$\hat{w}_\alpha(-1/t) = \hat{w}_{-\alpha}(t), \qquad \hat{w}_\alpha(t)^{-1} = \hat{w}_\alpha(-t),$$
$$\hat{h}_\alpha(-1/t) = \hat{h}_{-\alpha}(t), \qquad \hat{h}_\alpha(t)\hat{h}_\alpha(-1/t) = \hat{h}_\alpha(-1),$$
$$\hat{w}_\alpha(1)\hat{h}_\alpha(t)\hat{w}_\alpha(1)^{-1} = \hat{h}_\alpha(1/t), \qquad \hat{w}_\alpha(1)^{-2} = \hat{h}_\alpha(-1).$$

b) Soient U le sous-groupe de G formé par les $\hat{x}_+(t)$, et H le sous-groupe engendré par les $\hat{h}_+(t)$. On pose $\hat{w} = \hat{w}_+(1)$ et $\mathrm{N} = \mathrm{H} \cup \hat{w}\mathrm{H}$. Montrer que N est un sous-groupe de G et que H est invariant dans N. Montrer que l'on a $\hat{w}\mathrm{U}\hat{w} \subset \mathrm{UNU}$ (ensemble des produits $u'nu''$ avec $u' \in \mathrm{U}$, $u'' \in \mathrm{U}$ et $n \in \mathrm{N}$), et que l'ensemble $\mathrm{UH} \cup \mathrm{UH}\hat{w}\mathrm{U}$ est le sous-groupe de G engendré par les $x_\alpha(t)$.

c) On reprend le groupe $\mathrm{SL}(2, k)$ de l'*exercice* 20 et les éléments $x_\alpha(t)$, $w_\alpha(t)$ et $h_\alpha(t)$ de ce groupe. Montrer, en utilisant la question *c*) de l'*exercice* 20, qu'il existe une application π de $\mathrm{SL}(2, k)$ dans G, et une seule, telle que

$$\pi(h_+(t)x_+(u)) = \hat{h}_+(t)\hat{x}_+(u), \qquad \pi(h_+(t)x_+(u)wx_+(v)) = \hat{h}_+(t)\hat{x}_+(u)\hat{w}\hat{x}_+(v).$$

Montrer que, pour que π soit un homomorphisme, il faut et il suffit que la relation

$$\hat{h}_\alpha(tu) = \hat{h}_\alpha(t)\hat{h}_\alpha(u) \tag{R 3}$$

soit vérifiée. Autrement dit, *pour construire un homomorphisme de* $\mathrm{SL}(2, k)$ *dans un groupe quelconque* G, *il suffit de se donner des éléments* $\hat{x}_+(t)$ *et* $\hat{x}_-(t)$ *de* G *vérifiant les relations* (R 1), (R 2) *et* (R 3) (« définition de $\mathrm{SL}(2, k)$ par générateurs et relations »).

qq 23. On considère le groupe $G = \mathrm{GL}(n, k)$ sur un corps commutatif k, le sous-groupe T de G formé des matrices diagonales, le sous-groupe B des matrices dont les termes situés en dessous de la diagonale sont tous nuls, et le sous-groupe U de B formé des matrices de B dont tous les termes diagonaux sont égaux à 1. Enfin on note N l'ensemble des $n \in G$ tels que $nTn^{-1} = T$ (normalisateur de T dans G). On identifie les éléments de G aux automorphismes de l'espace vectoriel k^n, dont la base canonique sera notée $e_1, \ldots, e_n$.

a) Montrer que $g \in N$ si et seulement s'il existe une permutation $w \in \mathfrak{S}_n$ et des scalaires $t_i \neq 0$ tels que l'on ait $g(e_i) = t_i e_{w(i)}$ pour $1 \leqslant i \leqslant n$. En déduire que le groupe quotient $N/T = W$ est isomorphe au groupe symétrique $\mathfrak{S}_n$. Pour $1 \leqslant i \leqslant n-1$, soit $\omega_i \in G$ la matrice qui permute les vecteurs de base e_i et e_{i+1} et laisse fixe e_j pour tout $j \neq i, i+1$. Montrer que N est engendré par T et les ω_i.

b) Pour $i \neq j$ et $t \in k$ on note $x_{ij}(t)$ l'élément de G défini par les formules suivantes :

$$x_{ij}(t)e_j = e_j - te_i, \qquad x_{ij}(t)e_k = e_k \quad \text{si} \quad k \neq j.$$

Quelle est la matrice de $x_{ij}(t)$? Montrer que l'on a $x_{ij}(t-u) = x_{ij}(t)x_{ij}(u)$ quels que soient $t, u \in k$, et que les $x_{ij}(t)$, pour i et j donnés et t variable, forment un sous-groupe U_{ij} de G. En posant d'une manière générale $(a, b) = aba^{-1}b^{-1}$ montrer qu'on a

$$\begin{aligned} &(x_{ij}(t), x_{jk}(u)) = x_{ik}(tu) && \text{si } i, j, k \text{ sont deux à deux distincts.} \\ &(x_{ij}(t), x_{kl}(u)) = 1 && \text{si } j \neq k \text{ et } i \neq l. \end{aligned}$$

Montrer que U est engendré par les matrices $x_{i,i+1}(t)$ $(1 \leqslant i \leqslant n-1, t \in k)$. Calculer $nx_{ij}(t)n^{-1}$ pour $n \in N$.

c) Soit B' le sous-groupe de G formé des matrices dont les termes situés au-dessus de la diagonale sont tous nuls. Montrer qu'il existe un $n \in N$ tel que $B' = nBn^{-1}$, et que B' est engendré par T et les $x_{i-1,i}(t)$. Soit g un élément quelconque de G ; montrer qu'il existe un $b \in B$ et $b' \in B'$ tels qu'en posant $g = bb'g_1$ la matrice g_1 soit de la forme

$$g_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & * & \ldots & * \\ \ldots & \ldots & \ldots & \ldots \\ 0 & * & \ldots & * \end{pmatrix}.$$

En déduire que G est engendré par B et B' ou, ce qui revient au même, par B et N (raisonner par récurrence sur n).

Montrer, en utilisant les formules de l'exercice 20, que le sous-groupe de G engendré par $x_{i,i+1}(t)$ et les $x_{i+1,i}(t)$ est $\mathrm{SL}(n, k)$, ensemble des $g \in \mathrm{GL}(n, k)$ tels que $\det(g) = 1$ (cette question suppose le lecteur au courant de la théorie des déterminants, et ne sera pas utilisée par la suite).

d) On considère le sous-groupe $B_i = B \cap \omega_i^{-1}B\omega_i$. Montrer que B_i est invariant dans B, et que tout $b \in B$ s'écrit, de façon unique, comme produit d'un élément de $U_{i,i+1}$ et de B_i. En déduire que la double classe $B\omega_iB$, ensemble des $b'\omega_ib''$ avec $b', b'' \in B$, est réunion des classes $B\omega_ix_{i,i+1}(t)$, $t \in k$.

e) En posant $n\,{}^1U_{i,i+1}n = U_{j,k}$ montrer que l'ensemble $B\omega_i BnB$ (pour $n \in N$) est réunion des classes $B\omega_i n x_{jk}(t)B$, où t varie. En déduire que l'on a

$$B\omega_i BnB = B\omega_i nB \quad \text{si } j < k, \qquad \text{et} \qquad B\omega_i BnB \quad BnB \cup B\omega_i nB \quad \text{si } j > k$$

(on utilisera le fait, qu'il suffit de vérifier dans $GL(2, k)$, que $x_{i+1,i}(t) \in B\omega_i B$ si $t \neq 0$).

f) Soit G_0 la réunion des doubles classes BnB (lesquelles sont en nombre fini puisque N/T est fini). En utilisant le fait que N est engendré par T et les ω_i, montrer à l'aide de la question précédente que $nG_0 \subset G_0$ pour tout $n \in N$. Montrer que G_0 est un sous-groupe de G, et en déduire que

$$G = \bigcup BnB$$

(théorème de Bruhat pour le groupe linéaire).

24. Soit k un corps fini à q éléments, et soit V un espace vectoriel de dimension finie n sur k; soit m un entier compris entre 0 et n.

a) Soit X_m l'ensemble des familles $(x_1, \ldots, x_m)$ d'éléments de V linéairement indépendants. Montrer que l'on a :

$$\mathrm{Card}(X_m) = (q^n - 1)(q^n - q) \ldots (q^n - q^{m-1}).$$

(Raisonner par récurrence sur m.)

b) Montrer que l'ordre du groupe GL(V) des automorphismes de V est donné par la formule :

$$\mathrm{Card}\ GL(V) = (q^n - 1)(q^n - q) \ldots (q^n - q^{n-1}) = q^{n^2} \prod_{i=1}^{i=n} \left(1 - \frac{1}{q^i}\right).$$

c) Soit $G_{n,m}$ l'ensemble des sous-espaces vectoriels de V de dimension m (« grassmannienne »). Montrer que l'on a :

$$\mathrm{Card}\ G_{n,m} = q^{m(n-m)} \prod_{i=n-m+1}^{i=n} \left(1 - \frac{1}{q^i}\right).$$

25. Soit V un espace vectoriel de dimension 2 sur le corps à 2 éléments; soient x, y, z les trois éléments non nuls de V.

a) Montrer que l'on a $x + x = y + y = z + z = 0$ et $x + y = z$, $y + z = x$, $z + x = y$.

b) En déduire que le groupe GL(V) des automorphismes de V est isomorphe au groupe des permutations de x, y, z.

26. Soit V un espace vectoriel de dimension 2 sur le corps à 3 éléments.

a) Montrer que V contient 4 sous-espaces de dimension 1, soient D_1, D_2, D_3, D_4.

b) Tout élément du groupe d'automorphismes GL(V) de V permute les D_i entre eux. On déduit de là un homomorphisme

$$\varepsilon : GL(V) \to S_4,$$

où S_4 désigne le groupe des permutations de $\{1, 2, 3, 4\}$. Montrer que le noyau de ε est $\{\pm 1\}$; en déduire que ε est surjectif (comparer les ordres des deux groupes).

c) Soit (LSV) le sous-groupe de GL(V) formé des éléments de déterminant 1. Montrer que ε définit un isomorphisme de $SL_2(V)$ sur le groupe alterné A_4, formé des permutations paires, de S_4. [Cette question suppose le lecteur au courant de la théorie des déterminants.]

1. On considère les trois formes linéaires

$$2x - y + 3z, \qquad 3x - 5y + z, \qquad 4x - 7y + z$$

sur $\mathbf{R}^3$, Forment-elles une base du dual de $\mathbf{R}^3$?

2. Montrer que les formes linéaires

$$x + 2y + z, \qquad 2x + 3y + 3z, \qquad 3x + 7y + z$$

forment une base du dual de $\mathbf{R}^3$, et trouver la base de $\mathbf{R}^3$ duale de celle-ci.

3. Soient K un anneau et $f_1, \ldots, f_n$ des formes linéaires sur le K-module à droite K^n. Pour que celles-ci forment une base du dual de K^n, il faut et il suffit qu'il existe des vecteurs $x_1, \ldots, x_n \in K^n$ tels que l'on ait

$$f_i(x_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

4. Soient K un anneau commutatif et U une matrice carrée d'ordre n à coefficients dans K.

a) Montrer que les relations

$${}^t\mathrm{U}.\mathrm{U} = 1_n, \qquad \mathrm{U}.{}^t\mathrm{U} = 1_n$$

sont équivalentes.

b) Montrer que les matrices $\mathrm{U} \in \mathrm{M}_n(\mathrm{K})$ vérifiant les conditions précédentes forment un sous-groupe de $\mathrm{GL}(n, \mathrm{K})$ (**groupe orthogonal à n variables** sur l'anneau K).

c) On dit qu'une matrice $\mathrm{S} \in \mathrm{M}_n(\mathrm{K})$ est **symétrique** si ${}^t\mathrm{S} = \mathrm{S}$. Soient X et Y deux matrices symétriques; pour que XY soit symétrique, il faut et il suffit que $\mathrm{XY} = \mathrm{YX}$.

d) Montrer (en prenant $\mathrm{K} = \mathbf{Q}$ ou un surcorps quelconque de $\mathbf{Q}$) que la matrice

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

est à la fois orthogonale et symétrique.

e) Trouver toutes les matrices orthogonales d'ordre 3 à coefficients entiers rationnels.

5. Soit $M_n(K)$ l'anneau des matrices carrées d'ordre n sur un anneau quelconque K; on regarde $M_n(K)$ comme un K-module à droite. Montrer que, pour toute forme linéaire f sur $M_n(K)$, il existe une et une seule matrice $A \in M_n(K)$ telle que

$$f(X) = \mathrm{Tr}(AX) \quad \text{pour tout } X \in M_n(K)$$

(voir l'*Exercice* 8 du § 12 pour une définition du symbole Tr). Pour que l'on ait

$$f(XY) = f(YX)$$

quelles que soient $X, Y \in M_n(K)$, il faut et il suffit, lorsque l'anneau K est commutatif, que la matrice A soit proportionnelle à 1_n.

1. Soient K un anneau, L un K-module et

$$L = M_1 \oplus \cdots \oplus M_n$$

une décomposition de L en somme directe de sous-modules.

a) Pour tout endomorphisme u de L, montrer qu'il existe un et un seul système d'homomorphismes

$$u_{ij} : M_i \to M_j \quad (1 \leqslant i, j \leqslant n)$$

tels que l'on ait

$$u(x) = u_{ij}(x) + \cdots + u_{nj}(x) \quad \text{pour tout } x \in M_j$$

et tout j tel que $1 \leqslant j \leqslant n$. Montrer inversement qu'étant donnés de tels homomorphismes u_{ij}, il existe un et un seul endomorphisme u de L vérifiant les conditions précédentes.

b) A tout endomorphisme u de L on associe la « matrice »

$$\begin{pmatrix} u_{11} & \cdots & u_{n1} \\ \cdots & \cdots & \cdots \\ u_{1n} & \cdots & u_{nn} \end{pmatrix}$$

formée avec les homomorphismes définis dans la question *a*). Étant donnés deux endomorphismes u et v de L, comment calcule-t-on la « matrice » de $v \circ u$ en fonction de celles de u et v?

2. Soient r un entier $\geqslant 1$ donné et $r_1, \ldots, r_n$ des entiers $\geqslant 1$ tels que

$$r = r_1 + \cdots + r_n.$$

Étant donnée une matrice carrée

$$U = \begin{pmatrix} a_{11} & \cdots & a_{1r} \\ \cdots & \cdots & \cdots \\ a_{1r} & \cdots & a_{rr} \end{pmatrix}$$

à coefficients dans un anneau quelconque K, on désigne par U_{ij} $(1 \leqslant i, j \leqslant n)$ la matrice (à r_i colonnes et r_j lignes) formée avec ceux des termes a_{pq} de la matrice U pour lesquels on a à la fois

$$r_1 + \cdots + r_{i-1} < p \leqslant r_1 + \cdots + r_i$$
$$r_1 + \cdots + r_{j-1} < q \leqslant r_1 + \cdots + r_j,$$

ce qui permet, avec des conventions évidentes, d'écrire U sous la form

$$U = \begin{pmatrix} U_{11} & \cdots & U_{n1} \\ \cdots & \cdots & \cdots \\ U_{1n} & \cdots & U_{nn} \end{pmatrix}.$$

Soit V une autre matrice carrée d'ordre n à coefficients dans K, et soit $W = VU$ la matrice produit. Montrer que les « blocs » W_{ij} qui composent W sont donnés par la formule

$$W_{ij} = V_{1j}U_{i1} + \cdots + V_{nj}U_{in}$$

analogue à la règle de calcul usuelle des matrices (**formule de multiplication par blocs** des matrices). Peut-on étendre ce résultat à des produits de matrices rectangulaires?

8. Montrer qu'avec les notations de l'Exercice précédent, la transposée d'une matrice U est donnée par

$${}^tU = \begin{pmatrix} {}^tU_{11} & \cdots & {}^tU_{1n} \\ \cdots & \cdots & \cdots \\ {}^tU_{n1} & \cdots & {}^tU_{nn} \end{pmatrix}.$$

4. Soit K un corps commutatif. On considère la matrice carrée

$$J = \begin{pmatrix} 0 & -1_n \\ 1_n & 0 \end{pmatrix}$$

(où 0 est la matrice nulle à n lignes et n colonnes), et on cherche les matrices carrées

$$U = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

d'ordre $2n$, à coefficients dans K, telles que

$${}^tU.J.U = J$$

(où A, B, C, D sont des matrices carrées d'ordre n à coefficients dans K). Écrire les relations que doivent vérifier A, B, C, D. Trouver les matrices U pour lesquelles $C = 0$.

5. Soit $S = {}^tS$ une matrice symétrique carrée d'ordre n à coefficients dans un corps commutatif K. On pose

$$J = \begin{pmatrix} 0 & 0 & 1_p \\ 0 & S & 0 \\ 1_p & 0 & 0 \end{pmatrix}$$

(de sorte que J est une matrice carrée d'ordre $n + 2p$). A quelles conditions la matrice

$$M = \begin{pmatrix} U & X & Z \\ 0 & V & Y \\ 0 & 0 & W \end{pmatrix}$$

(où la décomposition en blocs de M est de même nature que celle de J) vérifie-t-elle la relation

$${}^tM.J.M = J?$$

6. Soient K un anneau et $r_1, \ldots, r_n$ des entiers $\geqslant 1$. On pose $r = r_1 + \ldots + r_n$, et on considère les matrices

$$U = \begin{pmatrix} U_{11} & \ldots & U_{n1} \\ \ldots\ldots\ldots\ldots \\ U_{1n} & \ldots & U_{nn} \end{pmatrix}$$

(avec U_{ij} à r_i colonnes et r_j lignes) qui vérifient les conditions suivantes : on a

$$U_{ij} = 0 \quad \text{pour } i < j,$$

et de plus les matrices U_{ii} sont inversibles. Montrer que les matrices U forment un sous-groupe du groupe linéaire GL(r, K). Même question pour les matrices telles que

$$U_{ij} = 0 \text{ si } i < j, \qquad U_{ii} = 1.$$

Montrer que le second sous-groupe est invariant dans le premier, et formé de matrices unipotentes.

7. Soit $L = M_1 \oplus \ldots \oplus M_p$ une décomposition en somme directe de sous-modules d'un module à gauche L sur un anneau K. Dans le module à droite L*, dual de L, on considère pour chaque i tel que $1 \leqslant i \leqslant p$ l'ensemble M_i' des formes linéaires f sur L telles que

$$f(M_j) = 0 \quad \text{pour tout } j \neq i.$$

Montrer que les M_i' sont des sous-modules de L* et que $L^* = M_1' \oplus \ldots \oplus M_p'$.

8. Soient M un module à gauche sur un anneau K et M' un sous-module M. On suppose que le module quotient M/M' (§§ 10, 11, *Exercice* 10) est libre de type fini. Montrer qu'alors M' est facteur direct dans M (considérer le sous-module de M engendré par des éléments dont les images dans M/M' forment une base de M/M'). Pour une application importante de ce résultat, voir § 29, *Exercice* 11, g).

¶ 9. Soient M un module à gauche sur un anneau K et a un élément de M tel que $\lambda a = 0$ implique $\lambda = 0$; pour que le sous-module Ka engendré par a soit facteur direct dans M, il faut et il suffit qu'il existe sur M une forme linéaire f telle que $f(a) = 1$; on a alors

$$M = Ka \oplus \mathrm{Ker}(f).$$

§ 18

1. Montrer que tout sous-groupe *de type fini* du groupe additif $\mathbf{Q}^n$ possède une *base* d'au plus n éléments (imiter la démonstration du Théorème 1).

2. Pour qu'il existe une base de $\mathbf{Z}^n$ contenant un élément donné $(a_1, \ldots, a_n)$ de $\mathbf{Z}^n$, il faut et il suffit que les entiers a_i soient premiers entre eux (choisir des $u_i \in \mathbf{Z}$ tels que $\sum u_i a_i = 1$ et considérer le sous-groupe de $\mathbf{Z}^n$ défini par l'équation $\sum u_i x_i = 0$).
Plus généralement, soit K un anneau; pour qu'il existe une base de K^n contenant un $a \in K^n$ donné, il est nécessaire qu'il existe une forme linéaire f sur K^n telle que $f(a) = 1$, et cette condition est suffisante si K est principal (cf. § 17, *Exercice* 9).

3. Pour qu'il existe une matrice $U \in GL(n, \mathbf{Z})$ ayant une première ligne donnée

$$a_{11} \quad a_{21} \quad \cdots \quad a_{n1},$$

il faut et il suffit que les entiers $a_{11}, a_{21}, \ldots, a_{n1}$ soient premiers entre eux.

4. Soient L et M deux modules de type fini sur un anneau commutatif noethérien K. Montrer que le K-module $\mathrm{Hom}_K(L, M)$ est de type fini (construire un homomorphisme injectif de ce module dans une puissance convenable de M).

5. Soit M un module de type fini sur un anneau noethérien K. Montrer que toute suite croissante de sous-modules de M est stationnaire, et que tout ensemble de sous-modules de M possède au moins un élément maximal.

6. Soit a un idéal d'un anneau commutatif noethérien K. Montrer qu'il existe une suite finie d'idéaux premiers $\mathfrak{p}_1, \ldots, \mathfrak{p}_n$ (pas nécessairement deux à deux distincts) de K telle que l'on ait

$$\mathfrak{p}_1 \ldots \mathfrak{p}_n \subset a$$

(raisonner par l'absurde, considérer un élément maximal de l'ensemble des idéaux ne possédant pas cette propriété, et lui appliquer l'*Exercice* 11 du § 8).

7. Soient K un corps commutatif et A un sous-anneau de K admettant K pour corps des fractions (i.e. tel que tout $x \in K$ soit quotient de deux éléments de A). On utilise dans ce qui suit les définitions et résultats de l'*Exercice* 14 des §§ 10, 11.

a) Si A est un anneau de Dedekind, A est noethérien.

b) On suppose A noethérien, et que tout idéal *maximal* de l'anneau A est inversible; montrer que tout idéal de A est produit d'un nombre fini d'idéaux maximaux (méthode analogue à celle de l'*Exercice* précédent). En déduire que A est un anneau de Dedekind.

c) Soit A un anneau de Dedekind. Montrer que tout idéal fractionnaire de A s'écrit sous la forme d'un produit fini de puissances (positives ou négatives) d'idéaux premiers de A, et que cette décomposition est unique à l'ordre près des facteurs.

d) Soit $\mathfrak{p}$ un idéal premier de l'anneau de Dedekind A. Pour tout $x \in K$ non nul, on désigne par $v_{\mathfrak{p}}(x)$ l'exposant (qui peut être nul) de $\mathfrak{p}$ dans la décomposition de l'idéal fractionnaire Ax de A en produit de facteurs premiers; et on définit $v_{\mathfrak{p}}(0) = +\infty$. Montrer que la fonction $v_{\mathfrak{p}}$ est une valuation discrète (§ 8, *Exercice* 6) du corps K.

¶ **8.** Soient K un anneau commutatif noethérien, M un K-module de type fini, et u l'homothétie de rapport $a \in K$ dans M, donnée par

$$u(x) = ax \quad \text{pour tout } x \in M.$$

a) Montrer qu'il existe un entier $p \geqslant 0$ tel que l'on ait $\mathrm{Ker}(u^n) = \mathrm{Ker}(u^{n+1})$ pour tout $n \geqslant p$ (utiliser l'*Exercice* 5).

b) Montrer qu'on a $\mathrm{Im}(u^n) \cap \mathrm{Ker}(u) = \{0\}$ pour tout $n \geqslant p$.

c) On dit que le module M est **primaire** si, dans M, toute homothétie est soit injective soit nilpotente. Montrer qu'il en est ainsi lorsque le module M possède la propriété suivante : l'intersection de deux sous-modules non nuls de M n'est jamais nulle.

d) Soit $\mathfrak{q}$ un idéal de l'anneau K. Pour que $\mathfrak{q}$ soit primaire (§ 8, *Exercice* 13) il faut et il suffit que le quotient $K/\mathfrak{q}$, regardé comme K-module, soit primaire.

e) Un idéal $\mathfrak{q}$ d'un anneau K est dit **irréductible** si $\mathfrak{q} \neq K$ et si, quels que soient les idéaux $\mathfrak{a}$ et $\mathfrak{b}$ de K, la relation

$$\mathfrak{a} \cap \mathfrak{b} = \mathfrak{q} \text{ implique } \mathfrak{a} = \mathfrak{q} \text{ ou } \mathfrak{b} = \mathfrak{q}.$$

Montrer que tout idéal irréductible d'un anneau noethérien est primaire.

NB. — La réciproque est fausse.

¶¶ **9.** Soit K un anneau commutatif noethérien.

a) Montrer que tout idéal $\mathfrak{a} \neq K$ est intersection finie d'idéaux irréductibles (méthode de l'*Exercice* 6).

b) En déduire que *tout idéal d'un anneau commutatif noethérien est intersection finie d'idéaux primaires* (Emmy Noether).

¶¶ **10.** Soit M un module de type fini sur un anneau commutatif noethérien K. On dit qu'un idéal premier $\mathfrak{p}$ de K est **associé à** M s'il existe un $x \in M$ tel que $\mathfrak{p}$ soit l'annulateur de x dans M (i.e. tel que la relation $a \in \mathfrak{p}$ soit équivalente à la relation $ax = 0$).

a) Si $M \neq \{0\}$, il existe au moins un idéal premier associé à M (considérer les annulateurs des éléments non nuls de M et en prendre un maximal).

b) Il existe une suite croissante

$$0 = M_0 \subset M_1 \subset \cdots \subset M_r = M$$

de sous-modules de M et des idéaux premiers $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ de K tels que M_i/M_{i-1} soit isomorphe au module quotient $K/\mathfrak{p}_i$ pour tout i tel que $1 \leqslant i \leqslant r$ [observer que si l'annulateur d'un élément x d'un module est un idéal $\mathfrak{a}$ de K, alors le sous-module Kx engendré par x est isomorphe à $K/\mathfrak{a}$].

c) Avec les notations de la question précédente, tout idéal premier associé à M est l'un des $\mathfrak{p}_i$.

d) Soit $u(x) = ax (a \in K)$ une homothétie dans M. Pour que u soit injective, il faut et il suffit que a n'appartienne à aucun des idéaux premiers associés à M. Pour que u soit nilpotente, il faut et il suffit que a appartienne à tous les idéaux premiers associés à M.

e) On prend dans ce qui précède M = K et on note $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ les idéaux premiers associés à M. Montrer que

$$\mathfrak{p}_1 \cup \cdots \cup \mathfrak{p}_r$$

est l'ensemble des diviseurs de zéro dans K, et que

$$\mathfrak{p}_1 \cap \cdots \cap \mathfrak{p}_r$$

est l'ensemble des éléments nilpotents de K.
Montrer que tout idéal premier de K contient l'un des $\mathfrak{p}_i$. En déduire que, dans un anneau noethérien, l'intersection de tous les idéaux premiers est l'ensemble des éléments nilpotents (ce résultat s'étend en fait à tous les anneaux commutatifs : §§ 27, 28, *Exercice* 9).

11. Soit $\mathfrak{a}$ un idéal d'un anneau commutatif noethérien K; on suppose $\mathfrak{a} \neq K$. On appelle **idéal premier minimal de** $\mathfrak{a}$ tout idéal premier $\mathfrak{p}$ de K qui contient $\mathfrak{a}$, et qui ne contient aucun idéal premier contenant $\mathfrak{a}$ autre que lui-même.
Montrer que les idéaux premiers minimaux de $\mathfrak{a}$ sont en nombre fini, et que leur intersection est le radical (§ 8, *Exercice* 12) de $\mathfrak{a}$ (appliquer à $K/\mathfrak{a}$ la question *e*) de l'*Exercice* précédent).
(Les anneaux noethériens ont été inventés vers 1920 par Emmy Noether et ont été l'un des principaux points de départ de l'Algèbre « abstraite » moderne. Leur théorie, aujourd'hui extraordinairement développée, est à la base de la Géométrie Algébrique; mais on les utilise aussi ailleurs, en particulier dans la théorie des fonctions analytiques de plusieurs variables complexes; en fait, l'invention de ces anneaux est certainement l'une des découvertes mathématiques les plus utiles des temps modernes. Avant l'introduction des anneaux noethériens, on se bornait à en démontrer certaines propriétés sur les anneaux de polynômes — ce qui conduisait fréquemment à des démonstrations beaucoup plus compliquées que celles qu'on connaît aujourd'hui puisqu'on n'avait pas encore isolé l'idée fondamentale qui permet de les simplifier, à savoir les Théorèmes 3 et 4 de ce §.)

Résoudre les systèmes d'équations linéaires suivants (*) par la méthode des éliminations successives (celle-ci consiste à utiliser une équation pour calculer une inconnue en fonction des autres, puis à reporter le résultat ainsi obtenu dans les autres équations, de façon à se ramener à un système comportant une inconnue et une équation de moins que le système initial).

1.
$$\begin{aligned} 2x - y + 3z &= 9 \\ 3x - 5y + z &= -4 \\ 4x - 7y + z &= 5 \end{aligned}$$

2.
$$\begin{aligned} 2x + 3y + 5z &= 10 \\ 3x + 7y + 4z &= 3 \\ x + 2y + 2z &= 3 \end{aligned}$$

3.
$$\begin{aligned} 5x + 2y + 3z &= -2 \\ 2x - 2y + 5z &= 0 \\ 3x + 4y + 2z &= -10 \end{aligned}$$

4.
$$\begin{aligned} 4bcx + acy - 2abz &= 0 \\ 5bcx + 3acy - 4abz &= -abc \\ 3bcx + 2acy - abz &= 4abc \end{aligned}$$
(on suppose $abc \neq 0$)

5.
$$\begin{aligned} x + y + z &= a \\ x + \omega y + \omega^2 z &= b \\ x + \omega^2 y + \omega z &= c \end{aligned}$$
(où ω est une racine cubique de l'unité)

6.
$$\begin{aligned} ax - 3y + 5z &= 4 \\ x - ay + 3z &= 2 \\ 9x - 7y + 8az &= 0 \end{aligned}$$
(discuter suivant les valeurs de a)

(*) Dans les *Exercices* 1 à 17 (extraits du recueil de Proskurjakov, où l'on en trouvera beaucoup d'autres) le corps de base est **C**. Toutefois, le lecteur désireux d'introduire plus de variété dans les calculs pourra se placer sur un corps commutatif K quelconque et tenir compte de la caractéristique de K, ou bien chercher les solutions dans l'anneau **Z** des entiers rationnels lorsque la question a un sens. Il va de soi d'autre part qu'après avoir étudié la théorie des déterminants, le lecteur devra l'appliquer à la résolution de ces *Exercices*.

7.
$$\begin{aligned} ax + 2z &= 2 \\ 5x + 2y &= 1 \\ x - 2y + bz &= 3 \end{aligned}$$
(discuter suivant les valeurs de a et b)

8.
$$\begin{aligned} 2x + 2y - z + t &= 4 \\ 4x + 3y - z + 2t &= 6 \\ 8x + 5y - 3z + 4t &= 12 \\ 3x + 3y - 2z + 2t &= 6 \end{aligned}$$

9.
$$\begin{aligned} 2x - y - 6z + 3t &= -1 \\ 7x - 4y + 2z - 15t &= -32 \\ x - 2y - 4z + 9t &= 5 \\ x - y + 2z - 6t &= -8 \end{aligned}$$

10.
$$\begin{aligned} 2x - 5y + 3z + t &= 5 \\ 3x - 7y + 3z - t &= -1 \\ 5x - 9y + 6z + 2t &= 7 \\ 4x - 6y + 3z + t &= 8 \end{aligned}$$

11.
$$\begin{aligned} 6x + 6y + 5z + 18t + 20u &= 14 \\ 10x + 9y + 7z + 24t + 30u &= 18 \\ 12x + 12y + 13z + 27t + 35u &= 32 \\ 8x + 6y + 6z + 15t + 20u &= 16 \\ 4x + 5y + 4z + 15t + 15u &= 11 \end{aligned}$$

12.
$$\begin{aligned} 2x + 7y + 3z + t &= 5 \\ x + 3y + 5z - 2t &= 3 \\ x + 5y - 9z + 8t &= 1 \\ 5x + 18y + 4z + 5t &= 12 \end{aligned}$$

13.
$$\begin{aligned} 2x + 5y - 8z &= 8 \\ 4x + 3y - 9z &= 9 \\ 2x + 3y - 5z &= 7 \\ x + 8y - 7z &= 12 \end{aligned}$$

14.
$$\begin{aligned} 6x + 3y + 2z + 3t + 4u &= 5 \\ 4x + 2y + z + 2t + 3u &= 4 \\ 4x + 2y + 3z + 2t + u &= 0 \\ 2x + y + 7z + 3t + 2u &= 1 \end{aligned}$$

(on déterminera en outre toutes les solutions *entières* de ce dernier système).

15.
$$\begin{aligned} 2x + 3y + z + 2t &= 3 \\ 4x + 6y + 3z + 4t &= 5 \\ 6x + 9y + 5z + 6t &= 7 \\ 8x + 12y + 7z + \lambda t &= 9 \end{aligned}$$
(discuter suivant les valeurs de λ)

16.
$$\begin{aligned} ax + y + z &= 1 \\ x + ay + z &= 1 \\ x + y + az &= 1 \end{aligned}$$
(discuter suivant les valeurs de a)

17.
$$\begin{aligned}(a+1)x+y+z &= a^2+3a\\ x+(a+1)y+z &= a^3+3a^2\\ x+y+(a+1)z &= a^4+3a^3\end{aligned}$$
(discuter suivant les valeurs de a).

18. Soit K un corps. On considère d'une part le système

$$(i)\qquad \begin{cases} a_{11}x_1+\cdots+a_{n1}x_n = b_1\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ a_{1n}x_1+\cdots+a_{nn}x_n = b_n\end{cases}$$

de n équations linéaires à n inconnues, à coefficients dans K, et d'autre part le système

$$(ii)\qquad \begin{cases} a_{11}y_1+\cdots+a_{n1}y_n-b_1y_{n+1} = 0\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ a_{1n}y_1+\cdots+a_{nn}y_n-b_ny_{n+1} = 0\end{cases}$$

de n équations linéaires et homogènes à $n+1$ inconnues.

a) Montrer que si une solution de (ii) vérifie $y_{n+1}\neq 0$, alors les

$$x_i = y_i y_{n+1}^{-1}$$

forment une solution de (i), et que réciproquement toute solution de (i) permet de construire une solution de (ii) telle que $y_{n+1}\neq 0$.

b) Montrer que les deux propriétés suivantes sont équivalentes : le système homogène associé à (i) ne possède aucune solution non triviale; toute solution non triviale de (ii) vérifie $y_{n+1}\neq 0$.

c) En utilisant le Corollaire du Théorème 7 du § 19, déduire de là une démonstration du fait que les conditions *a*) et *e*) du Théorème 2 du § 20 sont équivalentes.

19. Si un système d'équations linéaires à coefficients réels admet au moins une solution complexe, il admet une solution réelle.

¶¶ **20.** Soient L un corps et K un sous-corps de L.

a) On considère un système de n équations linéaires et homogènes à n inconnues, à coefficients dans K. Montrer que si le système admet une solution non triviale dans L, il admet une solution non triviale dans K (raisonner par récurrence sur n en utilisant par exemple la méthode des éliminations successives).

b) Montrer que si une matrice $A\in M_n(K)$ est inversible dans l'anneau $M_n(L)$, elle est inversible dans l'anneau $M_n(K)$.

c) Si des éléments de K^n (resp. des formes linéaires sur K^n) sont linéairement indépendants sur K, ils sont aussi linéairement indépendants sur L, et réciproquement.

d) Pour qu'un système d'équations linéaires, à coefficients et seconds membres dans K, admette une solution dans K, il faut et il suffit qu'il admette une solution dans L.

e) Les solutions dans L d'un système d'équations linéaires et homogènes à coefficients dans K, sont les combinaisons linéaires, à coefficients dans L, des solutions dans K du système considéré.

f) Si un système d'équations linéaires et homogènes à coefficients dans **Z** admet une solution non triviale dans **C**, il admet une solution non triviale dans **Z**.

§ 21

1. Soit V un espace vectoriel de dimension finie n sur un corps commutatif K ; quels que soient les entiers p, $q \geqslant 0$, on désigne par

$$\mathrm{T}_q^p(\mathrm{V})$$

l'espace vectoriel formé par les tenseurs p fois covariants et q fois contravariants sur V.

a) Soient $(a_i)_{1 \leqslant i \leqslant n}$ une base de V et $(a^i)_{1 \leqslant i \leqslant n}$ la base duale de V*. Montrer que les éléments

$$(*) \qquad a_{i_1} \otimes \cdots \otimes a_{i_p} \otimes a^{j_1} \otimes \cdots \otimes a^{j_q}$$

(où $i_1, \ldots, j_q$ prennent toutes les valeurs comprises entre 1 et n) forment une base de l'espace vectoriel $\mathrm{T}_q^p(\mathrm{V})$ [NB — On trouvera dans l'*Exemple* 6 du § 21 la règle pour identifier chaque $a \in \mathrm{V}$ à un tenseur d'espèce $\binom{1}{0}$; elle est essentielle pour donner un sens à l'expression (*) ci-dessus].

En déduire que $\mathrm{T}_q^p(\mathrm{V})$ est de dimension n^{p+q} sur K.

b) Soit u un *auto*morphisme de V. Étant donné un tenseur $f \in \mathrm{T}_q^p(\mathrm{V})$, on considère sur $(\mathrm{V}^*)^p \times \mathrm{V}^q$ la fonction f' donnée par

$$f'(y_1, \ldots, y_p, x_1, \ldots, x_q) = f[{}^t u(y_1), \ldots, {}^t u(y_p), u^{-1}(x_1), \ldots, u^{-1}(x_q)]$$

quels que soient les $y_i \in \mathrm{V}^*$ et les $x_i \in \mathrm{V}$. Montrer que $f' \in \mathrm{T}_q^p(\mathrm{V})$ et que l'application $f \to f'$ ainsi définie dans $\mathrm{T}_q^p(\mathrm{V})$ est linéaire. On désigne dans ce qui suit cette application par

$$\mathrm{T}_q^p(u).$$

Montrer que l'on a

$$\mathrm{T}_q^p(v \circ u) = \mathrm{T}_q^p(v) \circ \mathrm{T}_q^p(u)$$

quels que soient $u, v \in \mathrm{GL}(\mathrm{V})$. En déduire que l'application $u \to \mathrm{T}_q^p(u)$ est un homomorphisme du groupe des automorphismes de V dans le groupe des automorphismes de $\mathrm{T}_q^p(\mathrm{V})$.

On pose

$$u(a_i) = \sum_j a_i^j a_j,$$

de sorte que $(a_i^j)_{1 \leqslant i,j \leqslant n}$ est la matrice de u par rapport à la base (a_i) de V. Calculer la matrice de $\mathrm{T}_q^p(u)$ par rapport à la base de $\mathrm{T}_q^p(\mathrm{V})$ définie dans la question *a*).

c) Soit u un *endo*morphisme de V. Étant donné un tenseur $f \in \mathrm{T}^p_q(\mathrm{V})$, on définit une nouvelle fonction f'' sur $(\mathrm{V}^*)^p \times \mathrm{V}^q$ en posant

$$f''(y_1, \ldots, y_p, x_1, \ldots, x_q) = \sum_{1 \leqslant l \leqslant p} f[y_1, \ldots, y_{l-1}, {}^t u(y_l), y_{l+1}, \ldots, y_p, x_1, \ldots, x_q] - \sum_{1 \leqslant j \leqslant q} f[y_1, \ldots, y_p, x_1, \ldots, x_{j-1}, u(x_j), x_{j+1}, \ldots, x_q].$$

Montrer qu'on a encore $f'' \in \mathrm{T}^p_q(\mathrm{V})$ et que l'application $f \to f''$ de $\mathrm{T}^p_q(\mathrm{V})$ dans lui-même ainsi définie est linéaire. On la note dans ce qui suit $\mathrm{D}^p_q(u)$. Montrer qu'on a

$$\mathrm{D}^p_q(u + v) = \mathrm{D}^p_q(u) + \mathrm{D}^p_q(v)$$
$$\mathrm{D}^p_q(u \circ v - v \circ u) = \mathrm{D}^p_q(u) \circ \mathrm{D}^p_q(v) - \mathrm{D}^p_q(v) \circ \mathrm{D}^p_q(u)$$

quels que soient les endomorphismes u et v de V. Connaissant la matrice de u par rapport à la base (a_i) de V, calculer celle de $\mathrm{D}^p_q(u)$ par rapport à la base (*) de la question *a*).

¶ **2.** Soient V un espace vectoriel de dimension finie sur un corps commutatif K, et T un tenseur deux fois covariant et trois fois contravariant sur V. Montrer qu'il existe un et un seul tenseur U une fois covariant et deux fois contravariant sur V dont les composantes par rapport à toute base de V sont données, en fonction de celles de T, par la relation

$$\mathrm{U}^i_{jk} = \sum_{1 \leqslant h \leqslant n} \mathrm{T}^{ih}_{jkh}.$$

Interpréter cette opération (**contraction** du tenseur T par rapport au second indice covariant et au troisième indice contravariant) en regardant U et T comme des formes multilinéaires.

¶ **3.** Soit T un tenseur deux fois covariant et deux fois contravariant. Montrer que le scalaire

$$\sum_{1 \leqslant i, j \leqslant n} \mathrm{T}^{ij}_{ji}$$

est indépendant de la base choisie pour le définir.

¶¶ **4.** Soient L et M deux modules sur un anneau commutatif K. On considère le module

$$\mathrm{N} = \mathrm{K}^{(\mathrm{L} \times \mathrm{M})}$$

admettant pour base l'ensemble $\mathrm{L} \times \mathrm{M}$ (§§ 10, 11, *Exercice* 15); identifiant chaque élément de $\mathrm{L} \times \mathrm{M}$ à l'élément correspondant de N, on considère dans N le sous-module N' engendré par les éléments de N qui sont de la forme

$$(\lambda' x' + \lambda'' x'', \mu' y' + \mu'' y'') - \lambda'\mu'(x', y') - \lambda'\mu''(x', y'') - \lambda''\mu'(x'', y') - \lambda''\mu''(x'', y'').$$

On appelle **produit tensoriel des modules** L et M le module quotient N/N'; on le désigne par la notation

$$\mathrm{L} \otimes \mathrm{M}.$$

Étant donnés des éléments $x \in \mathrm{L}$ et $y \in \mathrm{M}$, on désigne par

$$x \otimes y$$

l'élément de $\mathrm{L} \otimes \mathrm{M} = \mathrm{N}/\mathrm{N}'$ représenté par l'élément (x, y) de N.

a) Montrer que l'application

$$(x, y) \to x \otimes y$$

de $L \times M$ dans $L \otimes M$ est *bilinéaire*, et que les « produits » $x \otimes y$ engendrent le module $L \otimes M$.

b) Soit f une application de $L \times M$ dans un K-module quelconque E. Montrer que, pour que f soit *bilinéaire*, il faut et il suffit qu'il existe une application *linéaire*

$$\bar{f} : L \otimes M \to E$$

telle que l'on ait

$$f(x, y) = \bar{f}(x \otimes y)$$

quels que soient $x \in L$ et $y \in M$; l'application $\bar{f}$ est alors unique. (Ce résultat est la propriété fondamentale des produits tensoriels de modules : ils servent à ramener l'étude des applications bilinéaires à celle des applications linéaires).

c) On suppose L et M libres de type fini; soient $(a_i)_{1 \leqslant i \leqslant p}$ une base de L et $(b_j)_{1 \leqslant j \leqslant q}$ une base de M; montrer que les produits

$$a_i \otimes b_j \qquad (1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q)$$

forment une base de $L \otimes M$. En déduire que

$$\dim(L \otimes M) = \dim(L) \cdot \dim(M)$$

si K est un corps.

d) Montrer qu'il existe un isomorphisme et un seul de $L \otimes M$ sur $M \otimes L$ qui applique $x \otimes y$ sur $y \otimes x$ quels que soient $x \in L$ et $y \in M$.

e) Soient L, M, L' et M' quatre modules sur K; on considère des homomorphismes

$$u : L \to L' \qquad \text{et} \qquad v : M \to M';$$

montrer qu'il existe un et un seul homomorphisme

$$f : L \otimes M \to L' \otimes M'$$

tel que l'on ait

$$f(x \otimes y) = u(x) \otimes v(y)$$

quels que soient $x \in L$ et $y \in M$ (observer que le second membre est fonction bilinéaire de x et y). On dit que f est le produit tensoriel des homomorphismes u et v, et on le note généralement $u \otimes v$. [Cette notation traditionnelle peut prêter à confusion, car elle désigne aussi un élément du module

$$\mathrm{Hom}(L, L') \otimes \mathrm{Hom}(M, M');$$

en pratique, on n'a presque jamais à considérer ce dernier produit tensoriel, et $u \otimes v$ a toujours la signification définie plus haut.]

f) On suppose L, ..., M' libres de type fini; on choisit des bases de ces modules, donc (question *c*) ci-dessus) de $L \otimes M$ et $L' \otimes M'$; calculer la matrice de $u \otimes v$ par rapport à ces bases en fonction des matrices de u et v par rapport aux bases choisies dans L, ..., M'. [Le résultat obtenu conduit à la notion de **produit tensoriel de deux matrices**; soient

$$A = (a_{ij})_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant q} \qquad \text{et} \qquad B = (b_{kh})_{1 \leqslant k \leqslant m,\ 1 \leqslant h \leqslant n}$$

deux matrices à coefficients dans un anneau commutatif K. On appelle produit tensoriel de A et B la matrice à pm colonnes et qn lignes définie comme suit : on numérote les colonnes du produit tensoriel à l'aide des couples (i, k) tels que $1 \leqslant i \leqslant p$, $1 \leqslant k \leqslant m$, et les lignes à l'aide

des couples (i, h) tels que $1 \leqslant j \leqslant q$, $1 \leqslant h \leqslant n$; cela dit, le terme de la matrice produit tensoriel

$$A \otimes B$$

situé à l'intersection de la colonne d'indice (i, k) et de la ligne d'indice (j, h) est par définition

$$a_{ij}b_{kh}.$$

Par exemple :

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes \begin{pmatrix} u & v \\ u' & v' \\ u'' & v'' \end{pmatrix} = \begin{pmatrix} au & av & bu & bv \\ cu & cv & du & dv \\ au' & av' & bu' & bv' \\ cu' & cv' & du' & dv' \\ au'' & av'' & bu'' & bv'' \\ cu'' & cv'' & du'' & dv'' \end{pmatrix};$$

on utilise, pour ordonner les couples d'entiers (i, k), l'ordre lexicographique, qui consiste à convenir que (i, j) précède (k, h) si $i \leqslant k$, ou bien si $i = k$ et $j \leqslant h$.]

g) Soient L, L′, L″, M, M′ et M″ six modules sur l'anneau K; on prend des homomorphismes

$$u' : L \to L', \qquad u'' : L' \to L'', \qquad v' : M \to M', \qquad v'' : M' \to M'';$$

montrer qu'on a

$$(u'' \circ u') \otimes (v'' \circ v') = (u'' \otimes v'') \circ (u' \otimes v').$$

En déduire que si A′, A″, B′, B″ sont des matrices à coefficients dans K, la formule

$$(A''A') \otimes (B''B') = (A'' \otimes B'').(A' \otimes B')$$

est vraie pourvu qu'elle ait un sens.

h) Soient L, M et N trois modules. Montrer qu'il existe un et un seul isomorphisme

$$(L \otimes M) \otimes N \to L \otimes (M \otimes N)$$

qui, quels que soient $x \in L$, $y \in M$ et $z \in N$, applique $(x \otimes y) \otimes z$ sur $x \otimes (y \otimes z)$. [« Associativité » du produit tensoriel; dans la pratique on ne fait aucune différence entre $(x \otimes y) \otimes z$ et $x \otimes (y \otimes z)$, qu'on écrit $x \otimes y \otimes z$].

i) Soient $M_1, \ldots, M_p$ des modules sur K; on forme le module

$$M_1 \otimes M_2 \otimes \cdots \otimes M_p = M_1 \otimes (M_2 \otimes \cdots \otimes M_p)$$

(définition par récurence sur p). Soit f une application de $M_1 \times M_2 \times \cdots \times M_p$ dans un K-module N. Montrer que, pour que f soit p-linéaire, il faut et il suffit qu'il existe un *homomorphisme* de modules

$$\bar{f} : M_1 \otimes \cdots \otimes M_p \to N$$

tel que l'on ait

$$f(x_1, \ldots, x_p) = \bar{f}(x_1 \otimes \cdots \otimes x_p)$$

quels que soient les $x_i \in M_i$; l'application linéaire $\bar{f}$ est alors entièrement déterminée par f (raisonner par récurrence sur p en attribuant une valeur fixe à l'une des variables figurant dans f).

j) Soit M un K-module; on considère le module

$$M \otimes M \otimes M^*$$

où M^* est le dual de M; à l'aide de la question précédente, montrer qu'il existe un et un seul homomorphisme

$$j : M \otimes M \otimes M^* \to T_1^2(M)$$

de $M \otimes M \otimes M^*$ dans le module des tenseurs 2 fois covariants et une fois contravariants sur M qui, quels que soient $x, y \in M$ et $u \in M^*$, applique l'élément

$$x \otimes y \otimes u \in M \otimes M \otimes M^*$$

sur l'élément

$$x \otimes y \otimes u \in T_1^2(M)$$

[on rappelle, *Exemple* 7 du § 21, que cette dernière expression est la forme trilinéaire sur $M^* \times M^* \times M$ dont la valeur en $(f, g, z) \in M^* \times M^* \times M$ est l'élément

$$f(x)\, g(y)\, u(z)$$

de K]. Montrer que *j est bijectif si* M *est libre de type fini.* Généraliser ce résultat en remplaçant

$$M \otimes M \otimes M^* \quad \text{par} \quad M \otimes \cdots \otimes M \otimes M^* \otimes \cdots \otimes M^*$$

(p facteurs M et q facteurs M^*) et

$$T_1^2(M) \quad \text{par} \quad T_q^p(M),$$

défini au début de l'*Exercice* 1 ci-dessus.

k) On prend $K = \mathbf{Z}$ et $M = \mathbf{Z}/p\mathbf{Z}$; montrer que $T_0^2(M)$ est réduit à 0, mais qu'il n'en est pas ainsi de $M \otimes M$ [considérer l'application $(x, y) \to xy$ de $M \times M$ dans M, multiplication des entiers modulo p]. En conclure que dans ce cas l'homomorphisme j de la question précédente n'est pas bijectif (et est même nul...).

[La notion de produit tensoriel de deux modules définie dans cet *Exercice* est beaucoup plus utile que celle de tenseur définie au § 21, sauf lorsqu'il s'agit de modules libres de type fini. La raison en est que les tenseurs du § 21 sont adaptés à l'étude des applications multilinéaires *dans l'anneau de base* K lui-même, tandis que les produits tensoriels de l'*Exercice* 21 servent à étudier les applications multilinéaires *dans des* K*-modules quelconques*. Or il peut arriver, cf. la question (*k*) de l'*Exercice* 21, que les premières soient toutes identiquement nulles, sans qu'il en soit de même des secondes. Notons enfin que le produit tensoriel, lorsqu'il s'agit de matrices ou d'espaces vectoriels de dimension finie, remonte essentiellement à Kronecker; pour cette raison, certains auteurs l'appellent le **produit kroneckerien.**]

§ 22

Calculer les déterminants suivants.

1. $\begin{vmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{vmatrix}$

2. $\begin{vmatrix} 2\sin t\cos t & 2\sin^2 t - 1 \\ 2\cos^2 t - 1 & 2\sin t\cos t \end{vmatrix}$

3. $\begin{vmatrix} 4 & -3 & 5 \\ 3 & -2 & 8 \\ 1 & -7 & -5 \end{vmatrix}$

4. $\begin{vmatrix} 3 & 4 & -5 \\ 8 & 7 & -2 \\ 2 & -1 & 8 \end{vmatrix}$

5. $\begin{vmatrix} x^2+1 & xy & xz \\ xy & y^2+1 & yz \\ xz & yz & z^2+1 \end{vmatrix}$

6. $\begin{vmatrix} \cos a & \sin a\cos b & \sin a\sin b \\ -\sin a & \cos a\cos b & \cos a\sin b \\ 0 & -\sin b & \cos b \end{vmatrix}$

7. $\begin{vmatrix} 1 & 0 & 1+i \\ 0 & 1 & i \\ 1-i & -i & 1 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 1 & 1 \\ 1 & z & z^2 \\ 1 & z^2 & z \end{vmatrix}$ où $z = \cos(4\pi/3) + i\sin(4\pi/3)$

9. $\begin{vmatrix} \sin^2 a & \cos 2a & \cos^2 a \\ \sin^2 b & \cos 2b & \cos^2 b \\ \sin^2 c & \cos 2c & \cos^2 c \end{vmatrix}$

10. $\begin{vmatrix} 1 & a & a^4 \\ 1 & b & b^4 \\ 1 & c & c^4 \end{vmatrix}$

11. $\begin{vmatrix} 1 & x & x^3 \\ 1 & y & y^3 \\ 1 & z & z^3 \end{vmatrix}$

12. $\begin{vmatrix} x+a & b & c \\ a & x+b & c \\ a & b & x+c \end{vmatrix}$

13. $\begin{vmatrix} a & b & c \\ 0 & d & e \\ 0 & 0 & f \end{vmatrix}$

14. $\begin{vmatrix} a & b & c \\ a & x & c \\ a & b & x \end{vmatrix}$

15. $\begin{vmatrix} ab & ab' & ab'' \\ a'b & a'b' & a'b'' \\ a''b & a''b' & a''b'' \end{vmatrix}$

16. $\begin{vmatrix} 1 & x & x^2 \\ a & 1 & x \\ b & c & 1 \end{vmatrix}$

¶ **17.** Soient V un espace vectoriel de dimension n sur un corps commutatif K, et f une forme bilinéaire *alternée* sur V. On suppose que le seul vecteur $a \in V$ tel que

$$f(a, x) = 0 \quad \text{pour tout } x \in V$$

est $a = 0$ (on dit alors que f est non dégénérée).

a) Soit A la matrice formée par les coefficients $a_{ij} = f(a_i, a_j)$ de f par rapport à une base (a_i) de V. Montrer que, pour que f soit non dégénérée, il faut et il suffit que A soit inversible (utiliser le Théorème 2 du § 20).

b) Soient $a, b \in V$ tels que $f(a, b) \neq 0$. On pose $f_a(x) = f(a, x)$ et $f_b(x) = f(b, x)$; montrer que f_a et f_b sont des formes linéaires non proportionnelles sur V, et que les $x \in V$ tels que

$$(*) \qquad f(a, x) = f(b, x) = 0$$

forment un sous-espace vectoriel de dimension $n - 2$ de V.

c) Les hypothèses restant celles de *b*), montrer que V est somme directe du plan engendré par a et b et du sous-espace V′ des solutions de (*). Montrer que la restriction de f à V′ est non dégénérée.

d) En raisonnant par récurrence sur n, montrer *i*) que s'il existe sur V une forme bilinéaire alternée non dégénérée alors la dimension de V est paire *ii*) que si f est une forme bilinéaire alternée non dégénérée sur un espace vectoriel V de dimension $2p$, il existe une base de V par rapport à laquelle la matrice de f est

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

où 0_p désigne la matrice nulle à p lignes et p colonnes.

e) Une matrice carrée $A = (a_{ij})_{1 \leq i, j \leq n}$ à coefficients dans K est dite **alternée ou antisymétrique** si elle vérifie les relations

$$a_{ii} = 0, \qquad a_{ij} + a_{ji} = 0.$$

Montrer qu'une telle matrice ne peut être inversible si n est impair. Si A est inversible et si $n = 2p$, il existe une matrice $U \in GL(n, K)$ telle que

$$UA{}^tU = \begin{pmatrix} 0_p & 1_p \\ -1_p & 0_p \end{pmatrix}.$$

f) Montrer que

$$\begin{vmatrix} 0 & x & z \\ -x & 0 & y \\ -z & -y & 0 \end{vmatrix} = 0$$

quels que soient $x, y, z \in K$.

18. Soient f, g, h trois formes linéaires sur un espace vectoriel V sur un corps commutatif K. Montrer que, pour que f, g, h soient linéairement indépendantes, il faut et il suffit que

$$f \wedge g \wedge h \neq 0.$$

19. Soient V un espace vectoriel de dimension n sur un corps commutatif K, $(a_i)_{1 \leq i \leq n}$ une base de V, et f, g, h trois formes linéaires sur V. Montrer que les coefficients de $f \wedge g \wedge h$ par rapport à la base (a_i) sont les scalaires

$$a_{ijk} = \begin{vmatrix} f(a_i) & f(a_j) & f(a_k) \\ g(a_i) & g(a_j) & g(a_k) \\ h(a_i) & h(a_j) & h(a_k) \end{vmatrix}$$

20. Soient V un espace vectoriel de dimension finie sur un corps commutatif, (a_i) une base de V, f une forme bilinéaire alternée sur V, et g une forme linéaire sur V. Montrer que les coefficients par rapport à la base (a_i) de la forme trilinéaire alternée $f \wedge g$ sont les scalaires

$$a_{ijk} = f(a_i, a_j)g(a_k) + f(a_j, a_k)g(a_i) + f(a_k, a_i)g(a_j).$$

21. Soit V un espace vectoriel de dimension 3 sur un corps commutatif, et soient x, y, z trois éléments de V.

a) Si x, y, z sont linéairement dépendants, on a $f(x, y, z) = 0$ pour toute forme trilinéaire alternée f sur V (exprimer l'un des vecteurs à l'aide des deux autres).

b) Si x, y, z sont linéairement indépendants, on a $f(x, y, z) \neq 0$ pour toute forme trilinéaire alternée $f \neq 0$ sur V (observer que x, y, z forment une base de V).

c) Soit a, b, c une base de V; pour que x, y, z soient linéairement indépendants, il faut et il suffit que le déterminant de leurs coordonnées par rapport à la base a, b, c soit non nul (utiliser les questions *a*) et *b*) et l'*Exemple* 7 du § 22).

(Les résultats de cet *Exercice* seront généralisés au § suivant, mais on conseille au lecteur d'examiner tout d'abord en détail le cas des espaces à trois dimensions, du reste fort important dans la pratique).

22. Les vecteurs

$$(2, -3, 1), \qquad (3, -1, 5), \qquad (1, -4, 3)$$

sont-ils linéairement indépendants dans $\mathbf{R}^3$? Même question pour les vecteurs

$$(5, 4, 3), \qquad (3, 3, 2), \qquad (8, 1, 3).$$

§ 23

1. Soit K un anneau commutatif. Montrer que les matrices $U \in M_n(K)$ telles que

$$\text{dét}(U) = 1$$

forment un sous-groupe de $GL(n, K)$ — on le désigne généralement par la notation $SL(n, K)$ et on l'appelle le **groupe spécial linéaire à n variables** sur l'anneau K.

2. Le déterminant d'une matrice nilpotente à coefficients dans un corps commutatif est nul.

3. Soit U une matrice carrée à coefficients entiers, de déterminant non nul. Montrer que les seuls nombres premiers p qui figurent dans les dénominateurs des coefficients (rationnels) de U^{-1} sont ceux qui divisent le déterminant de U.

4. Soit U une matrice carrée orthogonale (i.e. telle que ${}^tU.U = 1$) à coefficients dans un corps commutatif. Montrer que $\text{dét}(U) = +1$ ou -1.

5. Trouver le nombre d'inversions de la permutation

$$\begin{pmatrix} 1 & 2 & 3 & \dots & n & n+1 & n+2 & & \dots & 2n \\ 1 & 3 & 5 & \dots & 2n-1 & 2 & 4 & 6 & \dots & 2n \end{pmatrix}.$$

6. De toutes les permutations des entiers $1, 2, \dots, n$, quelle est celle dont le nombre d'inversions est maximum?

7. Montrer que pour tout entier k tel que $0 \leqslant k \leqslant \binom{n}{2}$ il existe une permutation des entiers $1, 2, \dots, n$ dont le nombre d'inversions est k.

8. On considère un déterminant d'ordre 6, dont on désigne les termes par a_{ij} $(1 \leqslant i, j \leqslant 6)$. Quel est le signe dont on doit faire précéder le produit

$$a_{61}a_{23}a_{45}a_{36}a_{12}a_{54}$$

dans le développement de ce déterminant?

9. Dans le groupe $\mathfrak{S}_n$ des permutations de l'ensemble $|1, 2, \dots, n|$ on désigne par $\mathfrak{A}_n$ l'ensemble des permutations paires.

a) Montrer que $\mathfrak{A}_n$ est un sous-groupe invariant de $\mathfrak{S}_n$ (**groupe alterné** de n objets) et que le groupe quotient $\mathfrak{S}_n/\mathfrak{A}_n$ est isomorphe à $\mathbf{Z}/2\mathbf{Z}$.

b) Pour $3 \leqslant i \leqslant n$ (on suppose $n \geqslant 3$) on désigne par s_i la permutation

$$\begin{pmatrix} 1 & 2 & 3 & \dots & i-1 & i & i+1 & \dots & n \\ i & 1 & 3 & \dots & i-1 & 2 & i+1 & \dots & n \end{pmatrix};$$

montrer que les s_i engendrent $\mathfrak{A}_n$.

c) Montrer que, pour $n \geqslant 5$, les seuls sous-groupes invariants de $\mathfrak{A}_n$ sont $\mathfrak{A}_n$ lui-même et le sous-groupe réduit à l'identité (ce qu'on exprime en disant que $\mathfrak{A}_n$ est un groupe **simple** pour $n \geqslant 5$). Quels sont les sous-groupes invariants de $\mathfrak{A}_n$ pour $n = 2$, 3 ou 4?

d) Montrer que, pour $n \neq 4$, le seul sous-groupe invariant non trivial de $\mathfrak{S}_n$ est $\mathfrak{A}_n$.

¶ **10.** Soit $A = (a_{ij})$ une matrice carrée inversible d'ordre n à coefficients dans un corps commutatif K. On cherche des matrices X et Y (carrées d'ordre n) à coefficients dans K, vérifiant la relation

$$A = X.Y,$$

et de la forme

$$X = \begin{pmatrix} * & 0 & 0 & \dots & 0 \\ * & * & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ * & * & * & \dots & * \end{pmatrix}, \qquad Y = \begin{pmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & * \end{pmatrix}$$

(où les signes $*$ désignent des éléments arbitraires de K). Montrer que, pour que X et Y existent, il faut et il suffit qu'on ait

$$\begin{vmatrix} a_{11} & \dots & a_{1p} \\ \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} \end{vmatrix} \neq 0 \qquad \text{pour} \qquad 1 \leqslant p \leqslant n-1.$$

Dans ce cas, on peut imposer aux coefficients diagonaux de X (ou de Y) d'être tous égaux à 1, et cette condition détermine entièrement X et Y.

11. Soit K un anneau commutatif. On appelle **dérivation** de K toute application D de K dans K telle que l'on ait

$$D(x+y) = D(x) + D(y), \qquad D(xy) = D(x).y + x.D(y)$$

quels que soient $x, y \in K$ (cf. § 30, n° 1).

Soient D une dérivation de K et $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice carrée d'ordre n à coefficients dans K. Pour chaque entier i tel que $1 \leqslant i \leqslant n$, on note A_i la matrice obtenue en appliquant D aux termes situés sur la i^e colonne de A. Montrer que

$$D(\text{dét}(A)) = \text{dét}(A_1) + \dots + \text{dét}(A_n)$$

(formule de dérivation des déterminants).

¶¶ **12.** Soient M un module sur un anneau commutatif, f une forme p-linéaire alternée sur M, et g une forme q-linéaire alternée sur M. On définit une application

$$h : M^{p+q} \to K$$

(où K est l'anneau de base) en posant

$$(*) \qquad h(x_1, \ldots, x_{p+q}) = \sum_{\substack{s \in \mathfrak{S}_{p+q} \\ s(1) < \cdots < s(p) \\ s(p+1) < \cdots < s(p+q)}} \varepsilon(s) \cdot f(x_{s(1)}, \ldots, x_{s(p)}) \cdot g(x_{s(p+1)}, \ldots, x_{s(p+q)})$$

où la sommation est étendue à toutes les permutations s des entiers $1, \ldots, p+q$ qui respectent l'ordre des p premiers, ainsi que des q derniers, de ces entiers.

a) Montrer que h est une forme $(p+q)$-linéaire alternée sur M. On l'appelle le **produit extérieur** des formes f et g, et on la désigne par la notation

$$h = f \wedge g.$$

b) Montrer que

$$g \wedge f = (-1)^{pq} f \wedge g.$$

c) Montrer que, si f, g, h sont trois formes multilinéaires alternées sur M, on a

$$f \wedge (g \wedge h) = (f \wedge g) \wedge h$$

(« associativité » du produit extérieur).

d) Soient $u_1, \ldots, u_p, v_1, \ldots, v_q$ des formes linéaires sur M. Dans la formule $(*)$ on prend

$$f = u_1 \wedge \cdots \wedge u_p, \qquad g = v_1 \wedge \cdots \wedge v_q$$

(cf. § 23, n° 3, *Exemple* 2). Montrer qu'alors

$$h = u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge \cdots \wedge v_q.$$

18. (Généralisation du théorème de multiplication des déterminants.) Soit

$$\mathrm{A} = (a_{ij})_{1 \leqslant i \leqslant p,\, 1 \leqslant j \leqslant q}$$

une matrice rectangulaire à coefficients dans un anneau commutatif K; on choisit un entier r tel que $1 \leqslant r \leqslant p$ et $1 \leqslant r \leqslant q$.

Étant données une partie I à r éléments de l'ensemble $\{1, \ldots, p\}$ et une partie J à r éléments de l'ensemble $\{1, \ldots, q\}$, on désigne par a_{IJ} la matrice carrée d'ordre r formée avec les a_{ij} tels que $i \in \mathrm{I}$ et $j \in \mathrm{J}$; enfin, on désigne par

$$\Lambda^r(\mathrm{A})$$

la matrice

$$(\det(a_{IJ}))_{I \subset \{1, \ldots, p\},\, J \subset \{1, \ldots, q\},\, \mathrm{card}(I) = \mathrm{card}(J) = r}$$

on peut par exemple ordonner lexicographiquement les parties I à r éléments de $\{1, \ldots, p\}$ en convenant qu'une partie

$$\{i_1, \ldots, i_r\} \qquad \text{avec} \qquad i_1 < \cdots < i_r$$

précède une partie

$$\{j_1, \ldots, j_r\} \qquad \text{avec} \qquad j_1 < \cdots < j_r$$

s'il existe un entier h $(1 \leqslant h \leqslant r)$ tel que l'on ait

$$i_1 = j_1, \ \ldots, \ i_{h-1} = j_{h-1}, \ i_h < j_h$$

(cf. la méthode de numérotation des mots figurant dans un dictionnaire...) Les scalaires $\det(a_{ij})$ s'appellent les **mineurs d'ordre r** de la matrice A.
Cela dit, montrer que si A et B sont deux matrices à coefficients dans K, telles que le produit AB ait un sens, on a

$$\Lambda^r(AB) = \Lambda^r(A).\Lambda^r(B).$$

¶ **14.** Soient A et B deux matrices à p lignes et q colonnes à coefficients dans un anneau commutatif K. On dit que A et B sont **équivalentes** (sur l'anneau de base K) s'il existe des matrices

$$U \in GL(p, K), \qquad V \in GL(q, K)$$

telles que $B = UAV$.
S'il en est ainsi, montrer que pour tout $r \leqslant p, q$ l'idéal de K engendré par les mineurs d'ordre r de A est égal à l'idéal engendré par les mineurs d'ordre r de B.
[NB — La réciproque est vraie si K est *principal*; cf. § 31, *Exercice* 11, (*e*).]

¶ **15.** Soient K un anneau commutatif et $A \in M_p(K)$, $B \in M_q(K)$; on considère [§ 21, *Exercice* 4, *f*)] la matrice $A \otimes B \in M_{pq}(K)$. Montrer qu'on a

$$\det(A \otimes B) = \det(A)^q . \det(B)^p.$$

¶¶ **16.** Soit A une matrice carrée d'ordre n à coefficients dans un anneau commutatif K. On considère, pour $1 \leqslant r \leqslant n$, la matrice $\Lambda^r(A)$ de l'*Exercice* 14. Montrer que

$$\det(\Lambda^r(A)) = \det(A)^{\binom{n-1}{r-1}}.$$

¶¶ **17.** Soient M un module libre de type fini sur un anneau commutatif, et u un automorphisme de M. Calculer le déterminant de l'automorphisme $T_q^p(u)$ de $T_q^p(M)$ (§ 21, *Exercice* 1) en fonction de celui de u.

¶ **18.** Soit A une matrice carrée d'ordre *impair* à coefficients dans un anneau commutatif K. On suppose A *antisymétrique*, i.e. que ${}^tA = -A$. Montrer que $\det(A) = 0$. (Utiliser l'*Exercice* 17 du § 22).

Calculer les déterminants suivants

1. $\begin{vmatrix} a & 3 & 0 & 5 \\ 0 & b & 0 & 2 \\ 1 & 2 & c & 3 \\ 0 & 0 & 0 & d \end{vmatrix}$

2. $\begin{vmatrix} x & a & b & 0 & c \\ 0 & y & 0 & 0 & d \\ 0 & e & z & 0 & f \\ g & h & k & u & l \\ 0 & 0 & 0 & 0 & v \end{vmatrix}$

3. $\begin{vmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{vmatrix}$

4. $\begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{vmatrix}$

5. $\begin{vmatrix} 6 & -5 & 8 & 4 \\ 9 & 7 & 5 & 2 \\ 7 & 5 & 3 & 7 \\ -4 & 8 & -8 & -3 \end{vmatrix}$

6. $\begin{vmatrix} 24 & 11 & 13 & 17 & 19 \\ 51 & 13 & 32 & 40 & 46 \\ 61 & 11 & 14 & 50 & 56 \\ 62 & 20 & 7 & 13 & 52 \\ 80 & 24 & 45 & 57 & 70 \end{vmatrix}$

7. $\begin{vmatrix} 1 & 2 & 3 & \dots & n \\ -1 & 0 & 3 & \dots & n \\ -1 & -2 & 0 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ -1 & -2 & -3 & \dots & 0 \end{vmatrix}$

8. $\begin{vmatrix} a_0 & a_1 & a_2 & \dots & a_n \\ -x & x & 0 & \dots & 0 \\ 0 & -x & x & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & x \end{vmatrix}$

9. $\begin{vmatrix} 3 & 2 & 0 & 0 & \dots & 0 \\ 1 & 3 & 2 & 0 & \dots & 0 \\ 0 & 1 & 3 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 3 \end{vmatrix}$

Notant D_n ce déterminant d'ordre n, on établira une relation simple entre D_n, D_{n-1} et D_{n-2} et on utilisera le résultat suivant. Soit $(u_n)_{n\leq 1}$ une suite de nombres complexes telle que l'on ait la relation de récurrence

$$u_{n+2} = au_{n+1} + bu_n;$$

soient ε_1 et ε_2 les racines (distinctes ou non) de l'équation

$$z^2 - az - b = 0;$$

alors, si $\varepsilon_1 \neq \varepsilon_2$, il existe des constantes ℓ_1 et ℓ_2 telles que l'on ait

$$u_n = \ell_1 \varepsilon_1^n + \ell_2 \varepsilon_2^n \text{ pour tout } n,$$

et si $z_1 = z_2$ il existe des constantes c_1 et c_2 telles que l'on ait

$$u_n = (c_1 n + c_2) z_1^n$$

pour tout n. Pour des résultats beaucoup plus généraux, voir § 35, *Exercice* 16.

¶ **10.** $\begin{vmatrix} 1 & 2 & 0 & 0 & 0 & \dots & 0 & 0 \\ 3 & 4 & 3 & 0 & 0 & \dots & 0 & 0 \\ 0 & 2 & 5 & 3 & 0 & \dots & 0 & 0 \\ 0 & 0 & 2 & 5 & 3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 5 & 3 \\ 0 & 0 & 0 & 0 & 0 & \dots & 2 & 5 \end{vmatrix}$ (Même méthode que ci-dessus).

11. $\begin{vmatrix} 1-n & 1 & 1 & \dots & 1 \\ 1 & 1-n & 1 & \dots & 1 \\ 1 & 1 & 1-n & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1-n \end{vmatrix}$ (déterminant à n lignes et n colonnes)

12. $\begin{vmatrix} 1 & n & n & \dots & n \\ n & 2 & n & \dots & n \\ n & n & 3 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ n & n & n & \dots & n \end{vmatrix}$

13. $\begin{vmatrix} x & a_1 & a_2 & \dots & a_{n-1} & 1 \\ a_1 & x & a_2 & \dots & a_{n-1} & 1 \\ a_1 & a_2 & x & \dots & a_{n-1} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_1 & a_2 & a_3 & \dots & x & 1 \\ a_1 & a_2 & a_3 & \dots & a_n & 1 \end{vmatrix}$

(Chercher les racines de cette fonction polynomiale de x.)

¶ **14.** Montrer que

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \leqslant i < j \leqslant n} (x_j - x_i)$$

(déterminant de VanderMonde; on remarque par exemple que le premier membre est un polynôme de degré $n - 1$ au plus en x_1, dont $x_2, \dots, x_n$ sont des racines évidentes).

¶ **15.** $\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & x_1^n \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & x_n^n \end{vmatrix}$

¶ **16.** $\begin{vmatrix} 1 & f_1(x_1) & f_2(x_1) & \dots & f_{n-1}(x_1) \\ \dots & \dots & \dots & \dots & \dots \\ 1 & f_1(x_n) & f_2(x_n) & \dots & f_{n-1}(x_n) \end{vmatrix}$ où $f_k(x) = x^k + a_{k1}x^{k-1} + \dots + a_{kk}$.

¶ **17.** $\begin{vmatrix} 1 & \binom{1}{x_1} & \binom{2}{x_1} & \dots & \binom{n-1}{x_1} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & \binom{1}{x_n} & \binom{2}{x_n} & \dots & \binom{n-1}{x_n} \end{vmatrix}$ où l'on pose $\binom{x}{k} = \dfrac{x(x-1)\dots(x-k+1)}{k!}$.

18. Montrer que

$$\begin{vmatrix} a & b & c & d & e & f & g & h \\ b & a & d & c & f & e & h & g \\ c & d & a & b & g & h & e & f \\ d & c & b & a & h & g & f & e \\ e & f & g & h & a & b & c & d \\ f & e & h & g & b & a & d & c \\ g & h & e & f & c & d & a & b \\ h & g & f & e & d & c & b & a \end{vmatrix}$$

$$= \begin{array}{l} (a+b+c+d+e+f+g+h)\,(a+b+c+d-e-f-g-h)\times \\ \times\,(a+b-c-d+e+f-g-h)\,(a+b-c-d-e-f+g+h)\times \\ \times\,(a-b+c-d+e-f+g-h)\,(a-b+c-d-e+f-g+h)\times \\ \times\,(a-b-c+d+e-f-g+h)\,(a-b-c+d-e+f+g-h) \end{array}$$

19. Montrer que

$$\begin{vmatrix} 1+x_1y_1 & 1+x_1y_2 & \dots & 1+x_1y_n \\ \dots\dots\dots & \dots\dots\dots & \dots & \dots\dots\dots \\ 1+x_ny_1 & 1+x_ny_2 & \dots & 1+x_ny_n \end{vmatrix} = 0 \quad \text{si } n \geqslant 3.$$

20. Calculer, par récurrence sur le nombre n de lignes, le déterminant

$$\begin{vmatrix} 1 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 1 \end{vmatrix}$$

21. Montrer que

$$\begin{vmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & \binom{2}{1} & \binom{2}{2} & 0 & \dots & 0 \\ 1 & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \binom{n}{1} & \binom{n}{2} & \binom{n}{3} & \dots & \binom{n}{n-1} \end{vmatrix} = 1.$$

22. 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

$$(a^2+b^2+c^2+d^2)(x^2+y^2+z^2+t^2) = (ax+by+cz+dt)^2+(ay-bx+ct-dz)^2 + (az-bt-cx+dy)^2+(at+bz-cy+dx)^2.$$

Voyez-vous un rapport avec les *Exercices* 10 et 11 du § 15 ?
Calculer les inverses des matrices suivantes :

23. $\begin{pmatrix} 2 & 7 & 3 \\ 3 & 9 & 4 \\ 1 & 5 & 3 \end{pmatrix}$ **24.** $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 2 \\ 1 & 1 & 1 & -1 \\ 1 & 0 & -2 & -6 \end{pmatrix}$ **25.** $\begin{pmatrix} 3 & -2 & -5 & 1 \\ 2 & -3 & 1 & 5 \\ 1 & 2 & 0 & -4 \\ 1 & -1 & -4 & 9 \end{pmatrix}$

26. Calculer le rang des matrices

$$\begin{pmatrix} 17 & -28 & 45 & 11 & 39 \\ 24 & -37 & 61 & 13 & 50 \\ 25 & -7 & 32 & -18 & -11 \\ 31 & 12 & 19 & -43 & -55 \\ 42 & 13 & 29 & -55 & -68 \end{pmatrix}, \qquad \begin{pmatrix} 2 & -1 & 1 & 3 & 4 \\ 2 & -1 & 2 & 1 & -2 \\ 2 & -3 & 1 & 2 & -2 \\ 1 & 0 & 1 & -2 & -6 \\ 1 & 2 & 1 & -1 & 0 \\ 4 & -1 & 3 & -1 & -8 \end{pmatrix}.$$

27. Les vecteurs

$$(1, 0, 0, 2, 5), \qquad (0, 1, 0, 3, 4), \qquad (0, 0, 1, 4, 7), \qquad (2, -3, 4, 11, 12)$$

sont-ils linéairement indépendants dans $\mathbf{R}^5$?

qq 28. Montrer que

$$\begin{vmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 2 \end{vmatrix} = 1.$$

29. Calculer

$$\begin{vmatrix} 1 & 4 & 9 & 16 \\ 4 & 9 & 16 & 25 \\ 9 & 16 & 25 & 36 \\ 16 & 25 & 36 & 49 \end{vmatrix}.$$

30. Combien d'additions et de multiplications doit-on en principe effectuer pour calculer un déterminant d'ordre 10?

31. Résoudre par la théorie des déterminants les *Exercices* 1 à 17 du § 20.

32. Résoudre et discuter les systèmes d'équations linéaires suivants :

a)
$$\begin{aligned} ax + by + z &= 1 \\ x + aby + z &= b \\ x + by + az &= 1 \end{aligned}$$

b)
$$\begin{aligned} ax + by + 2z &= 1 \\ ax + (2b - 1)y + 3z &= 1 \\ ax + by + (b + 3)z &= 2b - 1 \end{aligned}$$

c)
$$\begin{aligned} 2(a + 1)x + 3y + az &= a + 4 \\ (4a - 1)x + (a + 1)y + (2a - 1)z &= 2a + 2 \\ (5a - 4)x + (a + 1)y + (3a - 4)z &= a - 1 \end{aligned}$$

¶ 33. (Développement d'un déterminant suivant la règle de Laplace.) Soient K un anneau commutatif et n un entier > 1. On désigne par

$$D(x_1, \ldots, x_n)$$

le déterminant de n vecteurs $x_i \in K^n$ par rapport à la base canonique (e_i) de K^n. On pose

$$x_i = \xi_{i1}e_1 + \cdots + \xi_{in}e_n.$$

Enfin on choisit un entier p tel que $1 \leqslant p \leqslant n$.

a) Montrer qu'on a la relation

$$D(x_1, \ldots, x_n) = \sum_{i_1 < \cdots < i_p} \begin{vmatrix} \xi_{1i_1} & \cdots & \xi_{pi_1} \\ \cdots\cdots & \cdots & \cdots\cdots \\ \xi_{1i_p} & \cdots & \xi_{pi_p} \end{vmatrix} . D(e_{i_1}, \ldots, e_{i_p}, x_{p+1}, \ldots, x_n)$$

(regarder $D(x_1, \ldots, x_n)$ comme une fonction multilinéaire alternée de $x_1, \ldots, x_p$).

b) Montrer que, pour $1 \leqslant i_1 < \cdots < i_p \leqslant n$, on a

$$D(e_{i_1}, \ldots, e_{i_p}, x_{p+1}, \ldots, x_n) = \begin{vmatrix} \xi_{p+1,\,j_1} & \cdots & \xi_{n,\,j_1} \\ \cdots\cdots & \cdots & \cdots\cdots \\ \xi_{p+1,\,j_{n-p}} & \cdots & \xi_{n,\,j_{n-p}} \end{vmatrix}$$

où l'on désigne par $j_1, \ldots, j_{n-p}$ ceux des entiers $\{1, 2, \ldots, n\}$ qui n'appartiennent pas à l'ensemble $\{i_1, \ldots, i_p\}$, rangés de telle sorte que la permutation $(i_1, \ldots, i_p j_1, \ldots, j_{n-p})$ de $\{1, \ldots, n\}$ soit *paire*.

c) Soit

$$X = (\xi_{ij})_{1 \leqslant i,\, j \leqslant n}$$

une matrice carrée d'ordre n à coefficients dans K. Pour toute partie I de l'ensemble $\{1, 2, \ldots, n\}$ comprenant p éléments, on désigne par X_I la matrice formée avec les ξ_{ij} tels que l'on ait $i \in I$ et $1 \leqslant j \leqslant p$, et par X'_I la matrice « complémentaire », formée avec les ξ_{ij} tels que l'on ait $i \notin I$ et $p + 1 \leqslant j \leqslant n$. Enfin on désigne par $n(I)$ le nombre de couples (i, j) tels que l'on ait $i \in I$, $j \notin I$ et $i > j$. Montrer que l'on a

$$\det(X) = \sum_{\mathrm{Card}(I) = p} (-1)^{n(I)} . \det(X_I) . \det(X'_I)$$

(Formule de Laplace), où la somme est étendue à toutes les *parties* I à p éléments de l'ensemble $\{1, \ldots, n\}$.

d) Déduire ce résultat de la formule d'associativité du produit extérieur [§ 23, *Exercice* 13, *d*)] Appliquer la règle de Laplace au calcul des déterminants suivants :

34. $\begin{vmatrix} 1 & 1 & 3 & 4 \\ 2 & 0 & 0 & 8 \\ 3 & 0 & 0 & 2 \\ 4 & 4 & 7 & 5 \end{vmatrix}$ 35. $\begin{vmatrix} 2 & 1 & 4 & 3 & 5 \\ 3 & 4 & 0 & 5 & 0 \\ 3 & 4 & 5 & 2 & 1 \\ 1 & 5 & 2 & 4 & 3 \\ 4 & 6 & 0 & 7 & 0 \end{vmatrix}$ 36. $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 & 3 \\ 6 & 5 & 7 & 8 & 4 & 2 \\ 9 & 8 & 6 & 7 & 0 & 0 \\ 3 & 2 & 4 & 5 & 0 & 0 \\ 3 & 4 & 0 & 0 & 0 & 0 \\ 5 & 6 & 0 & 0 & 0 & 0 \end{vmatrix}$

37. $\begin{vmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & a & b & c & d \\ 0 & a^3 & b^3 & c^3 & d^3 \end{vmatrix}$ 38. $\begin{vmatrix} 3 & 4 & -3 & -1 & 2 \\ -5 & 6 & 5 & 2 & 3 \\ 4 & -0 & -3 & 7 & -5 \\ -1 & -4 & 1 & 1 & -2 \\ -3 & 7 & 5 & 2 & 3 \end{vmatrix}$

1. Montrer que les nombres complexes suivants sont algébriques et former pour chacun d'entre eux une équation algébrique à coefficients rationnels :

$$\sqrt{2}+\sqrt{3}; \qquad \sqrt[3]{2}+\sqrt{3}; \qquad \sqrt[4]{2}+\sqrt[3]{3}; \qquad \sqrt{2}+\sqrt{3}+\sqrt{5}.$$

2. Soient L un corps commutatif, K un sous-corps de L, et x un élément de L transcendant sur K. Trouver toutes les relations algébriques à coefficients dans K existant entre les éléments

$$x^3+1 \qquad \text{et} \qquad x^3$$

de L. Même question pour

$$x^3+x+1 \qquad \text{et} \qquad x^5.$$

¶ 3. Soient A un anneau d'intégrité commutatif et K un sous-corps de A. On suppose A de dimension finie en tant qu'espace vectoriel sur K; montrer que A est un corps.
Soient L un corps commutatif, K un sous-corps de L, et $x_1, \ldots, x_n$ des éléments de L algébriques sur K. Montrer que le sous-anneau $K[x_1, \ldots, x_n]$ de L est un corps.

¶ ¶ **4.** Soit K un corps commutatif. On appelle **extension** de K tout corps L admettant K pour sous-corps (par exemple, **C** est une extension de **R**, qui est une extension de **Q**). On peut alors regarder L comme un espace vectoriel sur K; si L est de dimension finie sur K, i.e. s'il existe des éléments $a_1, \ldots, a_r \in L$ en nombre fini tels que tout élément de L puisse s'écrire sous la forme

$$x_1a_1 + \cdots + x_ra_r$$

avec des $x_i \in K$, on dit que L est une **extension de degré fini de K**; la dimension de L comme espace vectoriel sur K 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 espace vectoriel sur K. Lorsque $K = Q$, les extensions de degré fini de K sont, par définition, les **corps de nombres algébriques** [historiquement, on imposait aux corps de nombres algébriques d'être des extensions de degré fini de **Q** *contenues dans* **C**, mais il est facile de voir que toute extention de degré fini de **Q** peut se « plonger » dans **C**, de sorte que cette condition est superflue]. On désigne dans ce qui suit

par K un corps commutatif et par L une extention de degré fini n de K. Pour tout $a \in \mathbf{L}$, on note u_a l'application de L dans L donnée par

$$u_a(x) = ax \quad \text{pour tout } x \in \mathbf{L}.$$

a) Montrer que u_a est un endomorphisme de L considéré comme espace vectoriel sur K, et qu'on a les relations

$$u_a + u_b = u_{a+b}, \qquad u_a \circ u_b = u_{ab}$$

quels que soient a, $b \in \mathbf{L}$. Quels sont les endomorphismes de L (regardé comme espace vectoriel sur K) qui commutent à tous les u_a?

b) Pour tout $a \in \mathbf{L}$, on pose

$$\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a) = \mathrm{Tr}(u_a), \qquad \mathrm{N}_{\mathbf{L}/\mathbf{K}}(a) = \det(u_a)$$

(on regarde u_a comme un endomorphisme d'un espace vectoriel de dimension finie sur K; le déterminant de u_a est défini au § 23, n° 5, et la trace au § 19, *Exercice* 22). On dit que $\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a)$ est la **trace** et $\mathrm{N}_{\mathbf{L}/\mathbf{K}}(a)$ la **norme de** a (relativement au sous-corps K); ce sont donc des éléments de K. Montrer qu'on a

$$\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a+b) = \mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a) + \mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(b), \qquad \mathrm{N}_{\mathbf{L}/\mathbf{K}}(ab) = \mathrm{N}_{\mathbf{L}/\mathbf{K}}(a)\mathrm{N}_{\mathbf{L}/\mathbf{K}}(b)$$

quels que soient a, $b \in \mathbf{L}$. Si L est de degré n sur K, on a de plus

$$\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a) = na, \qquad \mathrm{N}_{\mathbf{L}/\mathbf{K}}(a) = a^n$$

pour tout $a \in \mathbf{K}$.

c) Soit $(a_i)_{1 \leqslant i \leqslant n}$ une base de L regardé comme espace vectoriel sur K; tout $x \in \mathbf{L}$ s'écrit donc d'une façon et d'une seule sous la forme

$$x = \xi_1 a_1 + \cdots + \xi_n a_n \qquad \text{avec} \qquad \xi_1, \ldots, \xi_n \in \mathbf{K}.$$

On pose

$$xa_i = \sum_{1 \leqslant j \leqslant n} \lambda_{ij} a_j$$

où les λ_{ij} sont dans K. Calculer $\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(x)$ et $\mathrm{N}_{\mathbf{L}/\mathbf{K}}(x)$ en fonction des λ_{ij}.

d) On suppose K de caractéristique 0 (i.e. que si $x \in \mathbf{K}$ et $r \in \mathbf{Z}$ vérifient $rx = 0$, on a soit $r = 0$ soit $x = 0$; cf. § 30, n° 6). Montrer que si un $a \in \mathbf{L}$ vérifie

$$\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(ax) = 0 \quad \text{pour tout } x \in \mathbf{L}$$

on a $a = 0$. En déduire que, si $(a_i)_{1 \leqslant i \leqslant n}$ est une base de L sur K, on a

$$\det(\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a_i a_j)_{1 \leqslant i,j \leqslant n}) \neq 0.$$

e) Soit $(a_i)_{1 \leqslant i \leqslant n}$ une base de L sur K; on considère n éléments

$$b_i = \sum_{1 \leqslant j \leqslant n} \rho_{ij} a_j \quad (\rho_{ij} \in \mathbf{K}, \quad 1 \leqslant i \leqslant n)$$

de L. Montrer que, en introduisant les matrices

$$\mathrm{A} = (\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(a_i a_j))_{1 \leqslant i,j \leqslant n}$$
$$\mathrm{B} = (\mathrm{Tr}_{\mathbf{L}/\mathbf{K}}(b_i b_j))_{1 \leqslant i,j \leqslant n},$$

on a

$$\det(\mathrm{B}) = \det(\mathrm{A}) \cdot \det(\rho_{ij})^2.$$

En déduire (si K est de caractéristique 0) le résultat suivant : pour que n éléments $x_1, \ldots, x_n$

de L forment une base de L sur K, il faut et il suffit que le déterminant de la matrice

$$(\mathrm{Tr}_{L/K}(x_i x_j))_{1 \leqslant i,j \leqslant n}$$

soit non nul. Ce déterminant s'appelle le **discriminant** des n éléments $x_1, \ldots, x_n$ de L et se note généralement

$$D_{L/K}(x_1, \ldots, x_n).$$

f) K étant supposé de caractéristique 0, soit $(u_i)_{1 \leqslant i \leqslant n}$ une base de L sur K; montrer qu'il existe une autre base $(v_i)_{1 \leqslant i \leqslant n}$ de L sur K telle que l'on ait

$$\mathrm{Tr}_{L/K}(u_i v_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

(on montrera que les coordonnées des v_i par rapport à la base (u_i) sont données par un système de Cramer). On dit que les bases (u_i) et (v_i) sont **complémentaires**.

g) Les hypothèses et notations étant celles de la question (*f*), montrer que les coordonnées de tout $x \in L$ par rapport à la base (v_i) sont les éléments

$$\mathrm{Tr}_{L/K}(xu_i)$$

de K.

h) On ne suppose plus K de caractéristique 0. On dit que L est une extension **séparable** de K s'il existe un $x \in L$ vérifiant

$$\mathrm{Tr}_{L/K}(x) \neq 0.$$

Montrer que les résultats des questions *d*), *e*), *f*) et *g*) sont encore valables dans ce cas.

¶¶ 5. Soient L un corps commutatif et K un sous-corps de L; on suppose que L est extension de degré fini de K (*Exercice* 4), et on désigne par E un sous-corps de L contenant K.

a) Montrer que L est extension de degré fini de E, et que E est extension de degré fini de K.

b) Soit $(a_i)_{1 \leqslant i \leqslant r}$ une base de L sur E, et soit $(b_j)_{1 \leqslant j \leqslant s}$ une base de E sur K. Montrer que les rs éléments $a_i b_j$ forment une base de L sur K. En déduire que, si l'on note $[L : K]$ le degré de L sur K (i.e. la dimension de L comme espace vectoriel sur K) on a la relation

$$[L : K] = [L : E]\ [E : K]$$

c) Montrer que, pour tout $x \in L$, on a

$$\mathrm{Tr}_{L/K}(x) = \mathrm{Tr}_{E/K}(\mathrm{Tr}_{L/E}(x))$$
$$\mathrm{N}_{L/K}(x) = \mathrm{N}_{E/K}(\mathrm{N}_{L/E}(x))$$

(voir l'*Exercice* 4 en ce qui concerne les notations utilisées).

d) Soient x un élément de L et

$$x^s - a_{s-1}x^{s-1} + \cdots + (-1)^s a_0 = 0$$

une équation algébrique à coefficients dans K vérifiée par x, et de degré s *minimum*. Montrer que les éléments

$$1, x, \ldots, x^{s-1}$$

forment une base du corps $K[x]$ sur K. En utilisant la question *e*) de l'*Exercice* 4, et en posant $K[x] = E$, montrer qu'on a

$$\mathrm{Tr}_{E/K}(x) = a_{s-1}, \qquad \mathrm{N}_{E/K}(x) = a_0.$$

En conclure que

$$\mathrm{Tr}_{L/K}(x) = \frac{n}{s} a_{s-1}, \qquad \mathrm{N}_{L/K}(x) = (a_0)^{n/s}$$

où $n = [L : K]$.

1. Soit K un anneau d'intégrité infini. On dit qu'une partie A de K^n est un ouvert de Zariski dans K^n s'il existe des polynômes $f_1, \ldots, f_r \in K[X_1, \ldots, X_n]$, en nombre fini, tels que le complémentaire de l'ensemble A dans K^n soit l'ensemble des $x \in K^n$ qui vérifient les relations

$$f_1(x) = \cdots = f_r(x) = 0.$$

Ceci dit, soient f et g deux polynômes à n indéterminées, à coefficients dans K; on suppose qu'il existe dans K^n un ouvert de Zariski non vide A tel que l'on ait

$$f(x) = g(x) \quad \text{pour tout } x \in A;$$

montrer qu'alors $f = g$ (principe de prolongement des identités algébriques)

2. Montrer que si trois polynômes $f, g, h \in \mathbf{R}[X]$ vérifient l'une quelconque des trois relations suivantes, on a $f = g = h = 0$:

$$f(X)^2 - Xg(X)^2 = Xh(X)^2$$
$$f(X)^2 - Xg(X)^2 + h(X)^2 = 0$$
$$f(X)^2 + g(X)^2 + (X+2)h(X)^2 = 0.$$

Peut-on dans ce qui précède remplacer **R** par un corps commutatif quelconque?

3. Soient K un corps commutatif, f un polynôme à une indéterminée à coefficients dans K, et $a_1, \ldots, a_r$ des racines deux à deux distinctes de f dans K. Montrer, à l'aide du lemme 1 du § 28, qu'il existe un polynôme g à coefficients dans K tel que

$$f(X) = (X - a_1) \ldots (X - a_r) g(X).$$

Application : calculer (sans calculs !) le déterminant

$$\begin{vmatrix} 1 & 1 & 2 & 3 \\ 1 & 2 - x^2 & 2 & 3 \\ 2 & 3 & 1 & 5 \\ 2 & 3 & 1 & 9 - x^2 \end{vmatrix}$$

4. Même question pour le déterminant

$$\begin{vmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 1 - x & 1 & \ldots & 1 \\ 1 & 1 & 2 - x & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 1 & 1 & \ldots & n - x \end{vmatrix}.$$

5. Soient $f_1, \ldots, f_n$ des polynômes à une variable à coefficients dans un anneau commutatif K, et de degrés $n-2$ au plus. Montrer qu'on a

$$\begin{vmatrix} f_1(x_1) & f_1(x_2) & \ldots & f_1(x_n) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_n(x_1) & f_n(x_2) & \ldots & f_n(x_n) \end{vmatrix} = 0$$

quels que soient les $x_i \in K$.

¶ **6.** Soient K un corps commutatif infini et $a_0, \ldots, a_n$ des éléments donnés, deux à deux distincts, de K. Montrer qu'il existe un et un seul polynome $f \in K[X]$ de degré n au plus vérifiant

$$f(a_i) = b_i \quad (0 \leqslant i \leqslant n),$$

où les b_i sont des éléments donnés de K, et que f est fourni par la **formule d'interpolation de Lagrange**

$$f(X) = \sum_{i=0}^{i=n} b_i \frac{(X-a_0)\ldots(X-a_{i-1})(X-a_{i+1})\ldots(X-a_n)}{(a_i-a_0)\ldots(a_i-a_{i-1})(a_i-a_{i+1})\ldots(a_i-a_n)}.$$

Exemple : trouver un polynôme f de degré 3 tel que

$$f(1) = 2, \quad f(2) = 1, \quad f(3) = 4, \quad f(4) = 3.$$

¶ **7.** On pose

$$z_k = \cos\frac{2k\pi}{n} + i.\sin\frac{2k\pi}{n};$$

trouver un polynôme f de degré $n-1$, à coefficients complexes, tel que l'on ait

$$f(z_k) = k+1 \quad \text{pour } 0 \leqslant k \leqslant n-1$$

Réponse :

$$f(X) = \frac{n+1}{2} - \frac{1}{2}\sum_{k=1}^{k=n-1}\left(1 - i.\operatorname{cotg}\frac{k\pi}{n}\right)X^k.$$

8. Soit f une fonction définie sur l'ensemble N des entiers naturels, et à valeurs complexes On définit une nouvelle fonction Δf par

$$\Delta f(n) = f(n+1) - f(n),$$

et on définit successivement

$$\Delta^2 f = \Delta(\Delta f), \qquad \Delta^3 f = \Delta(\Delta^2 f), \ \ldots$$

Enfin, on dit que f est *polynomiale de degré* r s'il existe des constantes $a_0, \ldots, a_r$ telles que

$$f(n) = a_r n^r + a_{r-1}n^{r-1} + \cdots + a_0 \quad \text{pour tout } n \in \mathbf{N},$$

avec de plus $a_r \neq 0$.

a) Montrer que si f est polynomiale de degré r on a

$$\Delta^{r+1} f = 0, \qquad \Delta^r f \neq 0.$$

b) Calculer Δf lorsque

$$f(n) = \frac{n(n-1)\ldots(n-r+1)}{r!} = \binom{n}{r} \quad \text{pour tout } n \in \mathbf{N};$$

en déduire que, si f est une fonction polynomiale quelconque de degré r, on a

$$f(n) = c_0 + c_1\binom{n}{1} + \cdots + c_r\binom{n}{r} \quad \text{avec } c_k = \Delta^k f(0).$$

c) Montrer que si une fonction $f(n)$ vérifie

$$\Delta^{r+1} f = 0, \qquad \Delta^r f \neq 0,$$

alors f est polynomiale de degré r.

d) Soit g une fonction polynomiale de degré $r - 1$ sur $\mathbf{N}$. Montrer qu'il existe une et une seule fonction polynomiale f sur $\mathbf{N}$, de degré r, telle que

$$\Delta f = g, \qquad f(0) = 0.$$

En calculant f par la formule de la question *b*), en déduire une expression de la somme

$$g(0) + g(1) + \cdots + g(n).$$

Application : calculer les sommes

$$1^2 + 2^2 + \cdots + n^2; \qquad 1^3 + 2^3 + \cdots + n^3.$$

¶ **9.** (On rappelle que si A est un anneau commutatif, un idéal I de A est dit *premier* si $I \neq A$ et si, pour $x, y \in A$, la relation $xy \in I$ implique $x \in I$ ou $y \in I$; qu'un idéal I de A est dit *maximal* si $I \neq A$ et si les seuls idéaux de A contenant I sont I et A; et qu'enfin tout idéal de A, autre que A tout entier, est contenu dans au moins un idéal maximal). On se propose de prouver que, *si* K *est un anneau commutatif, l'intersection de tous les idéaux premiers de* K *est l'ensemble des éléments nilpotents de* K.

a) Montrer que, si un idéal I de K vérifie $I \neq K$, alors l'idéal I′ de l'anneau de polynômes K[X] engendré par I vérifie $I' \neq K[X]$. En déduire que, pour tout idéal premier de K, il existe un idéal maximal de K[X] qui le contient.

b) On suppose que $u \in K$ appartient à tous les idéaux premiers de K. Montrer que le polynôme $1 - uX$ n'appartient à aucun idéal maximal de l'anneau K[X]; en déduire qu'il est inversible dans l'anneau K[X].

c) Montrer que le polynôme $1 - uX$ ($u \in K$) est inversible dans K[X] si et seulement si u est nilpotent; en déduire le théorème annoncé.

d) Soit I un idéal d'un anneau commutatif K, avec $I \neq K$. Montrer que l'intersection des idéaux premiers de K contenant I est formée des $x \in K$ tels que l'on ait

$$x^n \in I$$

pour au moins un entier n.

10. Soit K un corps commutatif. Pour que le sous-anneau $K[f]$ de K[X] engendré par un polynôme $f \in K[X]$ soit K[X] tout entier, il faut et il suffit que

$$f(X) = aX + b, \qquad a \neq 0.$$

¶ **11.** Soit K un anneau commutatif. On appelle **série formelle** à une indéterminée à coefficients dans K toute suite

$$f = (a_0, a_1, \ldots, a_n, \ldots)$$

d'éléments de K (on ne suppose pas les a_i presque tous nuls). On définit la somme et le produit

de deux telles séries formelles à l'aide des formules (2) et (3) du § 27, n° 2, utilisées pour définir la somme et le produit de deux polynômes. Montrer qu'avec ces définitions on obtient un *anneau commutatif* contenant un sous-anneau isomorphe à K[X].
L'anneau ainsi obtenu se note habituellement K[[X]]; au lieu de la notation initiale $f = (a_0, a_1, \ldots, a_n, \ldots)$, on représente les séries formelles par l'écriture

$$(*) \qquad f = a_0 + a_1 X + \cdots + a_n X^n + \cdots = \sum_{n=0}^{\infty} a_n X^n,$$

qui permet de retenir plus facilement les formules définissant les opérations fondamentales : pour multiplier deux séries formelles, on les multiplie « terme à terme » puis on groupe ensemble les termes de même degré dans le résultat obtenu. Bien entendu, la formule (*) n'a théoriquement aucun sens, puisqu'elle peut contenir une infinité de termes non nuls; on ne doit la considérer que comme une simple *notation* commode pour représenter la suite des $a_i \in K$.
Démontrer les résultats suivants :

a) Pour que l'anneau K[[X]] soit intègre, il faut et il suffit que K le soit.

b) Pour qu'un élément (*) de K[[X]] soit inversible dans K[[X]], il faut et il suffit que son « terme constant » a_0 soit inversible dans K (différence majeure avec les anneaux de polynômes...).

c) Calculer l'inverse de $1 - X$ dans K[[X]].

¶¶ **12.** Soient K un anneau commutatif et

$$p(X, Y) = p_0(X) + p_1(X)Y + \cdots + p_n(X)Y^n$$

un polynôme à deux variables à coefficients dans K. On suppose $p_0(0) = 0$ et $p_1(0)$ inversible dans K [i.e. $p(0, 0) = 0$ et $p'_Y(0, 0) \neq 0$ si K est un *corps*.]
Montrer qu'il existe une série formelle et une seule

$$y = a_1 X + a_2 X^2 + \cdots$$

à coefficients dans K, sans terme constant, qui vérifie la relation $p(X, y) = 0$.
[On pourra procéder comme suit : supposant trouvées des constantes $a_1, \ldots, a_r \in K$ telles que le polynôme

$$p(X, a_1 X + a_2 X^2 + \cdots + a_r X^r)$$

ne contienne aucun terme de degré $\leqslant r$, on montrera qu'il existe $a_{r+1} \in K$ tel que le polynôme

$$p(X, a_1 X + a_2 X^2 + \cdots + a_{r+1} X^{r+1})$$

ne contienne aucun terme de degré $\leqslant r + 1$.]
Calculer y si $K = C$ et $f(X, Y) = (X - 1)^p - Y^q$ où p et q sont des entiers positifs. Voyez-vous un rapport entre la série formelle obtenue et le développement en série entière, établi en Analyse, de la fonction

$$(z - 1)^{p/q}?$$

¶ **13.** Soient p un nombre premier, f et g deux polynômes à coefficients entiers rationnels.
a) Montrer que si p divise tous les coefficients de fg, il divise tous les coefficients de f, ou bien tous les coefficients de g.

b) On dit qu'un polynôme à coefficients entiers rationnels est **primitif** si le *pgcd* de ses coefficients est égal à 1. Montrer que si f, g sont primitifs, il en est de même de leur produit.

c) Étant donné un polynôme h à coefficients entiers rationnels, on appelle **contenu** de h le *pgcd*, noté $c(h)$, de ses coefficients. Montrer qu'on a

$$c(fg) = c(f)\,c(g) \qquad \text{quels que soient} \qquad f, g \in \mathbf{Z}[X]$$

(lemme de Gauss).

¶ 14. Soient K un anneau commutatif, $\mathfrak{p}$ un idéal premier de K, et f, g deux polynômes à coefficients dans K. On suppose que tous les coefficients de fg sont dans $\mathfrak{p}$. Montrer que $\mathfrak{p}$ contient alors tous les coefficients de f, ou tous ceux de g.

¶ 15. Soit K un anneau d'intégrité commutatif.

a) Soient f et g deux polynômes non constants à une indéterminée, à coefficients dans l'anneau K. Dans l'anneau K[X, Y] des polynômes à deux indéterminées à coefficients dans K, on considère l'idéal I engendré par les polynômes $f(X)$ et $g(Y)$. Montrer qu'on a

$$I \neq K[X, Y]$$

(on supposera qu'on a une relation de la forme

$$u(X, Y)\,f(X) + v(X, Y)\,g(Y) = 1$$

et on examinera les termes homogènes de degré maximum du premier membre).

b) Soient $f_1, \ldots, f_n$ des polynômes à une indéterminée, à coefficients dans K. Montrer que l'idéal de l'anneau $K[X_1, \ldots, X_n]$ engendré par $f_1(X_1), \ldots, f_n(X_n)$ n'est pas l'anneau $K[X_1, \ldots, X_n]$ tout entier si les f_i ne sont pas constants.

c) Montrer que, pour tout entier k tel que $1 \leqslant k \leqslant n$, l'idéal de $K[X_1, \ldots, X_n]$ engendré par $X_1, \ldots, X_k$ est premier. Ces n idéaux sont-ils deux à deux distincts?

¶ 16. Soit K un anneau commutatif. Montrer que les propriétés suivantes de K sont équivalentes : (*i*) K n'a pas d'élément nilpotent non nul (*ii*) tout élément inversible de l'anneau K[X] est constant (cf. *Exercice* 9 de ce §, et l'*Exercice* 1 du § 8; on pourra aussi examiner les rapports avec l'*Exercice* 11 de ce §).

17. Soient V et W deux espaces vectoriels de dimension finie sur un corps commutatif K. On dit qu'une application f de V dans W est **polynomiale** s'il existe une base de V et une base de W telles que les coordonnées du vecteur $y = f(x) \in W$ soient données, en fonction de celles du vecteur $x \in V$, par des formules de la forme

$$\eta_j = p_j(\xi_1, \ldots, \xi_m) \quad (1 \leqslant j \leqslant n)$$

où les p_j sont des polynômes à $m = \dim(V)$ indéterminées, à coefficients dans K. On dit en outre que f est **homogène de degré** r si les p_j sont homogènes de degré r.

Montrer que ces définitions sont indépendantes des bases choisies dans V et W (i.e. que si les conditions énoncées sont satisfaites pour un choix particulier de ces bases, elles le sont pour tout autre choix). Montrer que, si le corps K est fini, toute application de V dans W est polynomiale (ce qui ôte beaucoup de son intérêt à cette notion dans ce cas...). On suppose K infini dans ce qui suit.

On note S(V, W) l'ensemble des applications polynomiales de V dans W, et $S_r(V, W)$ l'ensemble de celles qui sont homogènes de degré r. On pose enfin

$$S(V) = S(V, K), \qquad S_r(V) = S_r(V, K);$$

les éléments de S(V) [resp. $S_r(V)$] sont appelés les **fonctions polynomiales** [resp. les **fonctions polynomiales homogènes de degré** r] sur V.
Montrer que toute $f \in S(V, W)$ s'écrit d'une façon et d'une seule sous la forme

$$f = f_0 + f_1 + \cdots$$

où f_r est polynomiale et homogène de degré r, avec $f_r = 0$ pour presque tout r.
Montrer que S(V) est un sous-anneau de l'anneau de toutes les applications de l'ensemble V dans le corps K, que S(V) contient les applications linéaires et les applications constantes (qu'on identifie habituellement aux éléments de K, de sorte que K s'identifie canoniquement à un sous-corps de l'anneau S(V)). Soient $f_1, \ldots, f_n$ les fonctions coordonnées de V par rapport à une base de V; montrer que

$$S(V) = K[f_1, \ldots, f_n]$$

et que les éléments $f_1, \ldots, f_n$ sont algébriquement indépendants sur K.
Montrer que S(V, W) est un sous-espace vectoriel de l'espace vectoriel de toutes les applications de l'ensemble V dans l'espace vectoriel W. Montrer que, si $f \in S(V)$ et si $g \in S(V, W)$, l'application $h = fg$ de V dans W définie par

$$h(x) = f(x)g(x) \quad \text{pour tout } x \in V$$

est encore polynomiale. En déduire qu'on peut regarder S(V, W) comme un module sur l'anneau S(V). Soient (a_i) une base de V, (b_j) une base de W, et notons f_{ij} l'application *linéaire* de V dans W qui vérifie

$$f_{ij}(a_k) = \begin{cases} b_j & \text{si } k = i \\ 0 & \text{si } k \neq i \end{cases};$$

montrer que les f_{ij} forment une *base* du S(V)-module S(V, W).
Soient U, V, W trois espaces vectoriels de dimension finie sur K, et

$$f : U \to V, \qquad g : V \to W$$

deux applications polynomiales. Montrer que l'application composée $g \circ f$ est polynomiale Si f et g sont homogènes de degrés r et s, alors $g \circ f$ est homogène de degré rs.

18. Soit K un corps commutatif infini. On considère l'application polynomiale f de K dan K^3 donnée par

$$f(t) = (t^2 + t + 1, t^3 + t + 1, t^4 + t + 1);$$

trouver toutes les fonctions polynomiales sur K^3 qui sont nulles en tout point de $f(K)$. Que sont les points de K^3 où toutes ces fonctions sont nulles?
Même question pour l'application de K^* dans K^3 donnée par

$$f(t) = \left(\frac{t+1}{t}, \frac{t^2+1}{t}, \frac{t^3+1}{t}\right).$$

¶ **19.** Soient K un anneau commutatif et M un K-module; on se propose de « plonger » M dans un module sur l'anneau K[X] des polynômes à une indéterminée à coefficients dans K. Pour cela, on considère l'ensemble, noté M[X], dont les éléments sont les suites

$$(m_0, m_1, \ldots)$$

d'éléments presque tous nuls de M; on définit une addition dans M[X] par la formule

$$(m'_0, m'_1, \ldots) + (m''_0, m''_1, \ldots) = (m'_0 + m''_0, m'_1 + m''_1, \ldots),$$

enfin, on définit le produit d'un élément de M[X] par un élément de K[X] en posant

$$(a_0, a_1, a_2, \ldots).(m_0, m_1, m_2, \ldots) = (a_0m_0, a_0m_1 + a_1m_0, \ldots)$$

Montrer qu'avec ces définitions l'ensemble M[X] est effectivement un K[X]-module; on l'appelle le module des **polynômes à une indéterminée, à coefficients** dans M. On identifie chaque $m \in M$ à l'élément $(m, 0, \ldots)$ de M[X]; montrer qu'on a alors

$$(m_0, m_1, m_2, \ldots) = m_0 + m_1X + m_2X^2 + \cdots$$

dans le K[X]-module M[X] (NB — On écrit ici les scalaires, i.e. les éléments de K[X], à droite des éléments de M[X] pour se conformer à la tradition suivant laquelle, dans un polynôme, on écrit les coefficients à gauche des monômes).
Soient M et N deux K-modules et u un homomorphisme de M dans N; montrer qu'il existe un et un seul homomorphisme

$$\bar{u} : M[X] \to N[X]$$

de K[X]-modules qui coïncide avec u sur M.
On suppose M libre de type fini; montrer qu'alors M[X] est un K[X]-module libre de type fini, et que toute base de M sur K est aussi une base de M[X] sur K[X].
(Pour une application des constructions précédentes, voir § 35, *Exercice* 10)

¶ 20. Soient K un anneau commutatif, E un K-module, et u un endomorphisme de E. Étant donné un polynôme

$$f(X) = a_0 + a_1X + \cdots + a_nX^n$$

à coefficients dans K, on pose

$$f(u) = a_0 \cdot j_E + a_1u + \cdots + a_nu^n,$$

d'où un nouvel endomorphisme de E. Ceci fait, on considère l'application de l'ensemble produit $K[X] \times E$ dans l'ensemble E donnée par

$$(f, x) \mapsto f(u)(x);$$

montrer que l'ensemble E, muni de la loi de composition $(x, y) \mapsto x + y$ et de l'application qu'on vient de définir, est un *module sur l'anneau* K[X]; on le note E_u.
Réciproquement, soit M un K[X]-module; soit E le K-module déduit de M par restriction des scalaires (*) à l'anneau K; on considère l'homothétie de rapport X dans M comme une application u de E dans E. Montrer que u est un endomorphisme du K-module E, et que $M = E_u$. Autrement dit : *un module sur l'anneau* K[X] *s'identifie à un couple formé d'un module sur l'anneau* K *et d'un endomorphisme u de ce* K*-module.* (Ce résultat montre que l'étude des endomorphismes des K-modules revient à celle des K[X]-modules, ce qui sera confirmé dans les *Exercices* du § 35).
On se donne E et u comme ci-dessus; quels sont les sous-modules du K[X]-module E_u?
On considère deux K-modules E et F, un endomorphisme u de E, et un endomorphisme v de F. Quels sont les homomorphismes du K[X]-module E_u dans le K[X]-module F_v?

(*) Soient L un anneau, K un sous-anneau de L, et M un L-module à gauche. Le module déduit de M par restriction à K des scalaires s'obtient en considérant le groupe additif M et l'application $(a, m) \to am$ de $K \times M$ dans M qui coïncide, sur $K \times M$, avec l'application de $L \times M$ dans M donnée dans la structure de L-module de M. Autrement dit, on garde les mêmes « vecteurs », l'addition reste la même, mais on ne regarde que les homothéties dont le rapport appartient à K. Par exemple, tout espace vectoriel *complexe* définit aussi un espace vectoriel *réel*.

¶ 21. En imitant les constructions de l'*Exercice* précédent, montrer que, si K est un anneau commutatif, les modules sur l'anneau de polynômes K[X, Y] s'identifient aux triplets (E, u, v) formés d'un K-module E et de deux endomorphismes u et v de E tels que $u \circ v = v \circ u$. Généralisation à n indéterminées?

¶¶ 22. Soient L un corps commutatif et K un sous-corps de L.

a) On considère un système d'équations linéaires

$$(*) \qquad \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

à coefficients et seconds membres dans K, et un polynôme p à coefficients dans K et à n indéterminées. On suppose qu'il existe une solution

$$(x_1, \ldots, x_n) \in \mathrm{L}^n$$

telle que $p(x_1, \ldots, x_n) \neq 0$; montrer que, si K est infini, il existe alors dans K^n une solution de $(*)$ qui n'annule pas p (exemple **R** et **C**).

b) On considère deux matrices $A, B \in \mathrm{M}_n(\mathrm{K})$. On suppose qu'il existe une matrice $V \in \mathrm{GL}(n, \mathrm{L})$ telle que $B = VAV^{-1}$; montrer qu'alors il existe $U \in \mathrm{GL}(n, \mathrm{K})$ telle que $B = UAU^{-1}$ (considérer l'équation $UA = BU$ avec la condition $\det(U) \neq 0$ et appliquer la question précédente). [NB — Ce résultat subsiste si K est un corps fini, mais la démonstration est nettement plus difficile; cf. § 35, *Exercice* 13].

¶ 23. Soient K un anneau commutatif et I l'idéal de K[X] engendré par X^{n+1}. Décrire l'anneau quotient $\mathrm{L} = \mathrm{K}[X]/\mathrm{I}$ (on montrera que L est engendré par K et un élément ε assujetti à vérifier la relation $\varepsilon^{n+1} = 0$).

Trouver ses éléments inversibles. [L'anneau L intervient, pour $\mathrm{K} = \mathbf{C}$, dans la théorie des développements limités d'ordre n].

§ 29

¶ **1.** On considère deux fractions rationnelles $f, g \in K(X_1, \ldots, X_n)$ où K est un anneau d'intégrité commutatif *infini*. On suppose qu'il existe dans l'ensemble K^n un ouvert de Zariski (§§ 27, 28, *Exercice* 1) non vide A tel que f et g soient définies et aient la même valeur en tout point $x \in A$. Montrer qu'alors $f = g$. (Ce résultat, notamment dans le cas classique où $K = \mathbf{R}$ ou $\mathbf{C}$, explique pourquoi on a le droit d'identifier toute fraction rationnelle à coefficients dans K à la fonction qu'elle définit sur une partie de K^n.)

2. Une fonction rationnelle à *une* variable ne possède aucun point d'indétermination (si $f = p/q$ où p et q s'annulent simultanément en a, mettre en facteur dans p et q les plus hautes puissances possibles de $X - a$).

3. Soient K un corps commutatif, a un élément de K, et A l'ensemble des fractions rationnelles $f \in K(X)$ qui sont définies en a. Montrer que A est un anneau de valuation du corps $K(X)$ (§ 8, *Exercice* 6) et que l'idéal des éléments non inversibles de A est formé des $f \in A$ telles que $f(a) = 0$.
Montrer que les $f \in K(X)$ qui peuvent s'écrire sous la forme $f = p/q$, où p et q sont des polynômes tels que

$$d^o(p) \leqslant d^o(q),$$

forment également un anneau de valuation de $K(X)$.

4. Soient L un corps commutatif, K un sous-corps de L, et $x_1, \ldots, x_n$ des éléments de L. On désigne par $K(x_1, \ldots, x_n)$ le plus petit sous-corps de L contenant K et les x_i (sous-corps de L engendré par K et les x_i; un corps contenant K comme sous-corps et engendré par K et un nombre fini d'éléments s'appelle une extension de type fini de K, ou un corps de fonctions algébriques sur K). Montrer que c'est l'ensemble des éléments de L qui peuvent s'écrire sous la forme

$$f(x_1, \ldots, x_n)$$

où $f \in K(X_1, \ldots, X_n)$ est définie en $(x_1, \ldots, x_n)$. Montrer que $K(x_1, \ldots, x_n)$ est isomorphe à $K(X_1, \ldots, X_n)$ si les x_i sont algébriquement indépendants sur K.

¶ **5.** Soit K un corps commutatif. Étant donnée une fraction rationnelle $f \in K(X)$ non dans K, montrer que l'élément X du corps $K(X)$ est algébrique sur le sous-corps $K(f)$ engendré par K et f. En déduire qu'il en est de même de tout $g \in K(X)$.
Montrer qu'*étant donnés deux polynômes $p, q \in K[X]$, il existe une relation algébrique non triviale, à coefficients dans* K, *entre p et q.*

qq **6.** Soient L un corps commutatif et K un sous-corps de L.

a) Soient $x_1, \ldots, x_r, y, z$ des éléments de L; on suppose que z est algébrique sur le sous-corps $K(x_1, \ldots, x_r, y)$ mais non sur $K(x_1, \ldots, x_r)$; montrer qu'alors y est algébrique sur

$$K(x_1, \ldots, x_r, z).$$

b) On suppose que L est de **degré de transcendance fini** sur K, autrement dit qu'il existe un entier n tel que $n+1$ éléments quelconques de L vérifient une relation algébrique non triviale à coefficients dans K. Montrer qu'on peut alors trouver des éléments $x_1, \ldots, x_r$ de L, en nombre fini, algébriquement indépendants sur K, et tels que tout élément de L soit algébrique sur le sous-corps $K(x_1, \ldots, x_r)$ (on dit alors que les x_i forment une **base de transcendance** de L sur K).

c) Soient $x_1, \ldots, x_r$ et $y_1, \ldots, y_s$ deux bases de transcendance de L sur K. Montrer qu'il existe un indice j tel que y_j ne soit pas algébrique sur $K(x_1, \ldots, x_{r-1})$ (observer que dans le cas contraire tout élément de L serait algébrique sur ce sous-corps, et en particulier x_r). En déduire, à l'aide de la question *a*), que $x_1, \ldots, x_{r-1}, y_j$ forment une base de transcendance de L sur K.

d) Déduire de là que deux bases de transcendance quelconques de L sur K ont le même nombre d'éléments (qu'on appelle le **degré de transcendance** de L sur K). Montrer que ce nombre est le plus grand entier n tel qu'on puisse trouver n éléments de L algébriquement indépendants sur K.

e) Montrer que *si* $f_1, \ldots, f_{n+1}$ *sont* $n+1$ *fractions rationnelles à* n *indéterminées, à coefficients dans un corps commutatif* K, *il existe une relation algébrique non triviale, à coefficients dans* K, *entre* $f_1, \ldots, f_{n+1}$.

f) On suppose le corps commutatif K *infini*. Soit A un ouvert de Zariski (§§ 27, 28, *Exercice* 1) non vide dans K^p; on dit qu'une application f de A dans K^q est **rationnelle** s'il existe des fractions rationnelles

$$f_1, \ldots, f_q \in K(X_1, \ldots, X_p)$$

telles que $f_1, \ldots, f_q$ soient définies en tout $x \in A$ et que l'on ait

$$f(x) = (f_1(x), \ldots, f_q(x)) \quad \text{pour tout } x \in A$$

(Cette notion généralise celle d'application polynomiale des §§ 27, 28, *Exercice* 17.) Cela dit, montrer que si une application rationnelle de A dans K^q est *surjective*, on a $p \geqslant q$. (Ce résultat montre que les phénomènes « pathologiques » du type de la courbe de Peano — existence d'une application *continue* d'une droite *sur* un plan, par exemple — ne peuvent pas se produire lorsqu'on se limite à des applications définies par des fonctions polynomiales ou rationnelles.)

[La notion de degré de transcendance exposée dans cet *Exercice* est à la base de la définition de la dimension d'une variété algébrique.

Soit V une **variété algébrique** dans C^n, i.e. une partie de C^n définie par un nombre fini d'équations

$$f_1(x) = \cdots = f_r(x) = 0$$

où $f_1, \ldots, f_r$ sont des polynômes à n variables à coefficients dans C. On appelle *fonction polynomiale* sur V toute application de V dans C qui est la restriction à V d'une fonction polynomiale sur C^n. Ces fonctions polynomiales sur V forment évidemment un anneau A contenant C (fonctions constantes), et d'ailleurs engendré sur C par n éléments convenablement choisis (par exemple les restrictions à V des fonctions coordonnées de C^n). On dit que V est **irréductible** si l'anneau A est intègre; il revient au même, comme on peut le démontrer, d'exiger que V

ne peut pas s'écrire comme réunion de deux autres variétés algébriques distinctes de V. Si V est irréductible, on peut former le corps L des fractions de A; en notant $f_1, \ldots, f_n$ les restrictions à V des fonctions coordonnées de $\mathbf{C}^n$, il est clair que

$$L = \mathbf{C}(f_1, \ldots, f_n).$$

On dit que L est **le corps des fonctions rationnelles de la variété** V. Cela fait, L est de degré de transcendance fini sur **C**, et on appelle alors **dimension** de V le degré de transcendance de L sur **C**. Une « courbe » est de dimension 1, une « surface » de dimension 2, etc...
On démontre que la dimension p d'une variété algébrique irréductible V est aussi le plus grand entier tel que l'on puisse construire une chaîne croissante

$$V_0 \subset V_1 \subset \cdots \subset V_p = V$$

de variétés algébriques irréductibles non vides et deux à deux distinctes. Une « surface » contient une « courbe » qui contient un « point », ce qui explique (sic) pourquoi une « surface » est de dimension 2.
Les variétés algébriques dans $\mathbf{C}^n$ sont utilisées en Analyse, notamment pour étudier les systèmes d'équations aux dérivées partielles linéaires à coefficients constants. Considérons par exemple l'équation

$$\sum_{i_1, \ldots, i_n \geqslant 0} a_{i_1 \ldots i_n} \frac{\partial^{i_1 + \cdots + i_n} f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}} = 0$$

où f est une fonction inconnue de n variables réelles, et où les coefficients $a_{i_1 \ldots i_n}$ sont des constantes complexes presque toutes nulles. Si l'on cherche les solutions de la forme

$$f(x_1, \ldots, x_n) = e^{u_1 x_1 + \cdots + u_n x_n}$$

où $u_1, \ldots, u_n$ sont des constantes complexes, on est évidemment ramené à résoudre l'équation

$$\sum a_{i_1 \ldots i_n} u_1^{i_1} \ldots u_n^{i_n} = 0.$$

i.e. à étudier l'hypersurface algébrique de $\mathbf{C}^n$ définie par cette équation.]

7. Trouver les pôles et les points d'indétermination des fractions rationnelles suivantes :

$$\frac{X}{Y}\ ; \qquad \frac{X - Y}{XY}\ ; \qquad \frac{(X^2 - 1)(Y^2 - 1)}{X^2 + Y^2 - 1}\ ; \qquad \frac{X + Y + Z}{X - Y}\ ; \qquad \frac{X - Z}{Y - Z}$$

(on prendra **C** pour corps de base).

8. Soit K un corps commutatif; on considère (§§ 27, 28, *Exercice* **11**) l'anneau K[[X]] des séries formelles à une variable à coefficients dans K. Comme c'est un anneau d'intégrité on peut former son corps des fractions, qu'on note K((X)). Montrer que tout élément de celui-ci s'écrit d'une façon et d'une seule sous la forme du produit d'une puissance (éventuellement négative) de X par une série formelle dont le terme constant n'est pas nul, i.e. sous la forme d'une « série »

$$(*) \qquad \sum_{n=-\infty}^{n=+\infty} a_n X^n$$

à coefficients a_n dans K, avec la condition que les entiers n *négatifs* tels que $a_n \neq 0$ soient en nombre fini.
On notera que, comme K[X] est un sous-anneau de K[[X]], le corps K((X)) contient un

sous-corps isomorphe à $K(X)$, en sorte que toute fraction rationnelle à une variable à coefficients dans K peut se représenter par une série de la forme (*). Trouver les séries formelles (*) représentant les fractions rationnelles suivantes :

$$\frac{1}{X - X^3}; \qquad \frac{X^3 + X + 1}{X^4 - X^2}.$$

Montrer que la série (*) représentant une fraction rationnelle f ne comporte aucune puissance négative de X lorsque f est définie en $x = 0$, et réciproquement (*).
Montrer que $K[[X]]$ est un anneau de valuation (§ 8, *Exercice* 6) du corps $K((X))$.

¶ **9.** Soient A un anneau commutatif et S une partie de A; on suppose que S contient 1 mais ne contient pas 0, et que l'on a $xy \in S$ quels que soient $x \in S$, $y \in S$ (si A est un anneau d'intégrité on peut prendre par exemple pour S l'ensemble des éléments non nuls de A; dans le cas général, un exemple important s'obtient en prenant pour S l'ensemble des $x \in A$ qui n'appartiennent pas à un idéal premier donné de A).

a) Soit F l'ensemble des couples (x, s) avec $x \in A$, $s \in S$; étant donnés deux éléments $y' = (x', s')$ et $y'' = (x'', s'')$ de F, on désigne par $R\{y', y''\}$ la relation

$$\text{il existe un } s \in S \text{ tel que } \quad s(x's'' - x''s') = 0.$$

Montrer que R est une relation d'équivalence sur F. Que se passe-t-il lorsque A est intègre et qu'on prend pour S l'ensemble des éléments non nuls de A?

b) Soient A_S l'ensemble quotient F/R et θ l'application canonique de F sur F/R. Montrer qu'il existe sur A_S une et une seule structure d'anneau commutatif telle que l'on ait les formules

$$\theta(x, s) + \theta(y, t) = \theta(xt + ys, st)$$
$$\theta(x, s).\theta(y, t) = \theta(xy, st).$$

c) Montrer que l'application j de A dans A_S donnée par

$$j(x) = \theta(x, 1)$$

est un homomorphisme d'anneaux, que $j(s)$ est inversible dans A_S pour tout $s \in S$, et que tout élément de A_S est quotient d'un élément $j(x)$, $x \in A$, par un élément $j(s)$, $s \in S$. Quel est le noyau de l'homomorphisme j? A quelle condition j est-il injectif?

d) Soit f un homomorphisme de A dans un anneau commutatif K. Pour que $f(s)$ soit inversible dans K quel que soit $s \in S$, il faut et il suffit que f soit composé de l'homomorphisme j de la question précédente et d'un homomorphisme de l'anneau A_S dans l'anneau K.

e) Soient A un anneau d'intégrité commutatif, K son corps des fractions, et $\mathfrak{p}$ un idéal premier de A (i.e. tel que $\mathfrak{p} \neq A$ et que la relation $xy \in \mathfrak{p}$ implique $x \in \mathfrak{p}$ ou $y \in \mathfrak{p}$); on prend pour S le complémentaire de $\mathfrak{p}$ dans A. Montrer que l'anneau A_S est isomorphe au sous-anneau $A_{\mathfrak{p}}$ de K formé des fractions qui peuvent se mettre sous la forme x/y avec $x, y \in A$ et $y \notin \mathfrak{p}$.

(*) L'opération qui, étant donnés deux polynômes $f, g \in K[X]$, consiste à écrire sous la forme (*) la fraction rationnelle f/g est connue dans les ouvrages anciens sous le nom de *division de f par g suivant les puissances croissantes de* X; elle consiste aussi, pour chaque entier $r \geqslant 0$ (et si le terme constant de g n'est pas nul, cas auquel on peut toujours se ramener trivialement), à trouver un polynôme q de degré $\leqslant r$, tel que $f(X) - q(X)g(X)$ soit multiple de X^{r+1}. Le polynôme q s'obtient en supprimant les termes de degré $> r$ de la série formelle (*) qui représente f/g.
Dans la pratique, on doit effectuer ces opérations lorsqu'on veut développer en série entière ou de Laurent le quotient de deux polynômes ou séries entières, ou en trouver des développements limités.

f) Les hypothèses restant celles de *e*), on associe à chaque idéal $\mathfrak{a}$ de l'anneau $A_\mathfrak{p}$, distinct de $A_\mathfrak{p}$, l'idéal $A \cap \mathfrak{a}$ de l'anneau A. Montrer qu'on obtient de cette façon une bijection de l'ensemble des idéaux de $A_\mathfrak{p}$ distincts de $A_\mathfrak{p}$ sur l'ensemble des idéaux de A contenus dans $\mathfrak{p}$. Quelle est l'application réciproque?

g) Soient L un corps commutatif et $a_1, \ldots, a_n$ des éléments de L. On prend

$$A = L[X_1, \ldots, X_n]$$

et pour $\mathfrak{p}$ l'ensemble des polynômes $f \in A$ tels que $f(a_1, \ldots, a_n) = 0$. Montrer que $A_\mathfrak{p}$ est l'ensemble des fractions rationnelles $f(X_1, \ldots, X_n)$, à coefficients dans L, qui sont définies au point $(a_1, \ldots, a_n)$ de K^n.

h) Étendre les résultats de la question *f*) au cas d'un anneau de fractions A_S quelconque.

10. Soient A un anneau d'intégrité commutatif et K son corps des fractions. Montrer que le corps des fractions de l'anneau de polynômes A[X] est canoniquement isomorphe au corps de fractions rationnelles K(X).

¶ 11. Soient A un anneau d'intégrité commutatif, M un A-module, et K le corps des fractions de A. On se propose de montrer que, *si M est sans torsion* (§ 10, *Exercice* 11 ; les résultats de cet Exercice ne seront pas utiliser ici), *on peut plonger M dans un espace vectoriel sur* K (exemple trivial : A^n se plonge dans K^n).
On ne fait aucune hypothèse sur M jusqu'à nouvel ordre.

a) Soit F l'ensemble des couples (m, s) avec $m \in M$, $s \in A$ et $s \neq 0$. Étant donnés deux éléments $x' = (m', s')$ et $x'' = (m'', s'')$ de F, on note $R\{x', x''\}$ la relation

$$\text{il existe un } s \in A \text{ tel que } s(s'm'' - s''m') = 0 \text{ et } s \neq 0.$$

Montrer que R est une relation d'équivalence sur l'ensemble F.

b) Soit V l'ensemble quotient F/R ; on note θ l'application canonique de F sur V. Montrer qu'on peut définir la somme de deux éléments de V de telle sorte que l'on ait

$$\theta(m', s') + \theta(m'', s'') = \theta(s''m' + s'm'', s's'')$$

quels que soient (m', s'), $(m'', s'') \in F$, et que V, muni de cette loi de composition, est un groupe commutatif.

c) Montrer qu'il existe une application $(\lambda, x) \to \lambda x$ de $K \times V$ dans V qui vérifie la condition suivante : si $\lambda = u/s$ et si $x = \theta(m, t)$ (avec $u, s, t \in A$, $m \in M$, et s, t non nuls), on a

$$\lambda x = \theta(um, st).$$

Montrer que le groupe commutatif V, muni de cette application, est un espace vectoriel sur K.

d) On définit une application « canonique » j de M dans V par

$$j(m) = \theta(m, 1) ;$$

montrer que c'est un homomorphisme de A-modules (NB : comme V est un espace vectoriel sur K, on peut *a fortiori* regarder V comme un A-module), dont le noyau est le sous-module de torsion de M (i.e. l'ensemble des m tels que l'on ait $sm = 0$ pour au moins un $s \in A$ *non nul*). En déduire que, si M est sans torsion, j est un isomorphisme de M sur un sous-module de V. Exemple (en prenant $A = \mathbf{Z}$) : tout groupe commutatif sans torsion se plonge dans un espace vectoriel rationnel.

e) On suppose M sans torsion et de type fini. Montrer que V est de dimension finie sur K. Soit $n = \dim(V)$ (on dit que n est le rang de M) ; montrer qu'il existe deux bases $(a_i)_{1 \leqslant i \leqslant n}$ et

$(b_i)_{1 \leq i \leq n}$ de V telles que, si l'on désigne par P et Q les sous-A-modules (isomorphes à A^n) de V engendrés par les a_i et b_i respectivement, on ait $P \subset M \subset Q$.

f) Déduire de là et du Théorème 3 du § 18 le résultat suivant : *si* A *est un anneau principal, tout* A*-module* M *sans torsion et de type fini est isomorphe à* A^n *où* n *est le rang de* M. Traduction lorsque $A = \mathbf{Z}$?

g) Soit M un module de type fini sur un anneau principal A. Soit T le sous-module de torsion de M. Montrer que M/T est libre de type fini. En déduire (à l'aide de l'*Exercice* 8 du § 17) que M est isomorphe au produit direct de T et d'un A-module libre de type fini. (Ce résultat ramène l'étude des modules de type fini à celle des modules *de torsion* de type fini, qui sera faite au § 31, *Exercices* 8, 9, 10.)

h) Soient M un module libre de type fini sur un anneau principal A et M′ un sous-module de M. Montrer que les propriétés suivantes sont équivalentes : *i*) M′ est facteur direct dans M *ii*) le module quotient M/M′ est sans torsion *iii*) quels que soient $a \in A$ et $x \in M$, la relation $ax \in M'$ implique $a = 0$ ou $x \in M'$. Retrouver à partir de là le résultat du § 18, *Exercice* 2.

i) Soit M′ un sous-groupe de $\mathbf{Z}^n$ défini par un système d'équations linéaires et homogènes à coefficients entiers. Montrer que toute base de M′ fait partie d'une base de $\mathbf{Z}^n$.

12. On trouve, dans un manuel d'Algèbre destiné aux élèves des Lycées et Collèges, la phrase suivante : « Sous réserve de ne pas donner aux variables des valeurs qui annulent le numérateur ou le dénominateur, l'ensemble des fractions rationnelles muni des lois d'addition et de multiplication présente une structure de corps. » Que pensez-vous de cet énoncé?

§ 30

1. Soit K un corps commutatif de caractéristique 0 (par exemple $K = \mathbf{C}$). Montrer que l'équation

$$\frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + 1 = 0$$

n'a aucune racine multiple dans K.

2. Soit K un corps commutatif. Trouver un polynôme $f \in K[X]$ de degré 7 tel que 1 soit racine multiple d'ordre 4 au moins de $f(X) + 1$, et -1 racine multiple d'ordre 4 au moins de $f(X) - 1$. Généraliser en remplaçant les entiers 4 et 7 par n et $2n - 1$.

3. Chacun des polynômes suivants admet 1 pour racine; déterminer son ordre de multiplicité :

$$X^{2n} - nX^{n+1} + nX^{n-1} - 1; \qquad X^{2n+1} - (2n+1)X^{n+1} + (2n+1)X^n - 1;$$
$$X^{2n} - n^2X^{n+1} + 2(n^2-1)X^n - n^2X^{n-1} + 1.$$

4. Soit f un polynôme à une variable à coefficients dans un corps commutatif K. On suppose $f' = 0$. Montrer que f est constant si K est de caractéristique 0, et que f est un polynôme en X^p si K est de caractéristique $p \neq 0$. Réciproque?

5. Soit K un corps commutatif de caractéristique 0. On se propose de trouver toutes les applications

$$t \mapsto U(t)$$

de K dans l'anneau $M_n(K)$ qui vérifient

$$U(x+y) = U(x)U(y) \quad \text{quels que soient } x, y \in K,$$
$$U(0) = 1_n$$

et qui sont de plus polynomiales, i.e. de la forme

$$U(t) = A_0 + A_1 t + \cdots + A_r t^r + \cdots$$

où les matrices $A_r \in M_n(K)$ sont presque toutes nulles.

a) Montrer que la dérivée $U'(t)$ de la fonction polynomiale $U(t)$ vérifie

$$U'(t) = A_1 U(t).$$

b) Montrer que la matrice $A_1 = N$ est nilpotente et que

$$U(t) = \sum_{r \geqslant 0} N^r \frac{t^r}{r!} = \exp(tN)$$

(cf. § 8, *Exercice* 2).

c) Montrer inversement que, pour toute matrice nilpotente N, l'application

$$t \mapsto \exp(tN)$$

satisfait aux conditions requises.

d) Trouver la matrice N dans le cas de la fonction $U(t)$ du § 12, *Exercice* 11, et *vérifier* dans ce cas qu'on a bien $U(t) = \exp(tN)$.

6. Soit K un corps commutatif de caractéristique 0. Montrer qu'il n'existe aucun polynôme $f \in K[X]$ non nul tel que l'on ait

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

quels que soient $x, y \in K$. Même question pour la relation

$$f(xy) = f(x) + f(y).$$

Quels sont les polynômes tels que

$$f(x + y) = f(x) + f(y)?$$

¶ 7. Soit K un corps de caractéristique $p \neq 0$.

a) Montrer qu'on a

$$(x + y)^p = x^p + y^p$$

quels que soient $x, y \in K$. En déduire plus généralement que

$$(x + y)^q = x^q + y^q$$

si q est une puissance de p.

b) Montrer que l'application $x \to x^p$ est un isomorphisme de K sur un sous-corps de K (que l'on note K^p). Montrer que $K^p = K$ si K est fini.

c) Quels sont les polynômes $f \in K[X]$ vérifiant

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

quels que soient $x, y \in K$?

¶ 8. Soit K un corps de caractéristique $p \neq 0$. Montrer que, pour $n \in \mathbf{Z}$ et $x \in K$, l'élément $nx \in K$ ne dépend que de x et de la classe de n modulo p. En déduire qu'on peut considérer K comme un espace vectoriel sur le corps $\mathbf{Z}/p\mathbf{Z}$.
Montrer que *le nombre d'éléments d'un corps fini de caractéristique p est une puissance de p.*

9. Soit p un nombre premier. Montrer que le coefficient binômial $\binom{p^n}{r}$ est divisible par p pour tout $n \geqslant 1$ et tout r tel que $1 \leqslant r \leqslant p^n - 1$. (Utiliser l'*Exercice* 7 pour le corps $K = \mathbf{Z}/p\mathbf{Z}$.) Démonstration élémentaire?

10. Soit $f(X_1, \ldots, X_n)$ un polynôme à n variables à coefficients dans un anneau commutatif K. On suppose f homogène de degré r. Montrer que

$$X_1 f'_1(X_1, \ldots, X_n) + \ldots + X_n f'_n(X_1, \ldots, X_n) = r.f(X_1, \ldots, X_n)$$

où f'_i est la dérivée partielle de f par rapport à X_i. Cette relation (connue sous le nom d'identité d'Euler) caractérise-t-elle les polynômes homogènes de degré r?

11. Soit

$$(*) \qquad a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

une équation algébrique à coefficients a_i entiers rationnels; on suppose dans ce qui suit les a_i premiers entre eux (cas auquel on peut évidemment toujours se ramener en divisant les a_i par leur pgcd).
Soit $x = p/q$ une racine rationnelle de l'équation $(*)$; on suppose p et q premiers entre eux. Montrer que p divise a_0 et que q divise a_n.
Application : trouver les racines rationnelles des équations suivantes :

$$6x^4 - 11x^3 - x^2 - 4 = 0$$
$$2x^3 + 12x^2 + 13x + 15 = 0$$
$$6x^6 + 11x^4 - x^3 + 5x - 6 = 0$$
$$x^6 + 3x^5 + 4x^4 + 3x^3 - 15x^2 - 16x + 20 = 0$$
$$2x^6 + x^5 - 9x^4 - 6x^3 - 5x^2 - 7x + 6 = 0.$$

12. Soit K un anneau commutatif. On considère l'anneau $L = K[\varepsilon]$ engendré par K et un élément ε tel que

$$\varepsilon^2 = 0$$

(faire $n = 1$ dans l'*Exercice* 23 du § 28). Pour que l'application

$$x \mapsto x + D(x)\varepsilon$$

de K dans L soit un homomorphisme, il faut et il suffit que D doit une dérivation de l'anneau K.

13. Dans quels corps a-t-on l'identité

$$x^4 - x^2 + 1 = (x^2 - 5x + 1)(x^2 + 5x + 1)?$$

14. Soit K un corps de caractéristique $p \neq 0$. Si un $x \in K$ vérifie $x^n = 1$ pour un entier n, il existe un entier r *non divisible par* p tel que

$$x^r = 1.$$

15. Soit K un anneau commutatif. On définit (par récurrence sur l'entier $n \geqslant 0$) les **opérateurs différentiels d'ordre n au plus sur** K comme suit : ce sont des applications D de K dans K, vérifiant

$$D(x + y) = D(x) + D(y) \quad \text{quels que soient } x, y \in K,$$

et possédant en outre la propriété suivante : si $n = 0$, il existe un $a \in K$ tel que

$$D(x) = ax \quad \text{pour tout } x \in K;$$

si $n \geqslant 1$, il existe, pour tout $x \in K$, un opérateur différentiel D_x d'ordre $n - 1$ au plus sur K tel que l'on ait

$$D(xy) = x.D(y) + D_x(y) \quad \text{pour tout } y \in K.$$

a) Déterminer tous les opérateurs différentiels d'ordre 1 au plus sur K.

b) Montrer que si D′ et D″ sont des opérateurs différentiels d'ordres r et s au plus, l'application composée

$$D'' \circ D'$$

est un opérateur différentiel d'ordre $r + s$ au plus, et le crochet de Jacobi

$$D'' \circ D' - D' \circ D''$$

un opérateur différentiel d'ordre $r + s - 1$ au plus.

c) Soit D un opérateur différentiel d'ordre n au plus. On considère une famille $(x_i)_{i \in I}$ d'éléments de K, avec Card (I) $= n + 1$. Pour toute partie F de I, on pose

$$x_F = \prod_{i \in F} x_i \quad \text{et } x_\varnothing = 1.$$

Montrer qu'on a l'identité

$$\sum_{F \subset I} (-1)^{\mathrm{card}(F)} x_F \, D(x_{I-F}) = 0.$$

Montrer réciproquement que toute application D de K dans K, possédant cette propriété et telle que $D(x + y) = D(x) + D(y)$, est un opérateur différentiel d'ordre n au plus dans K.

d) Déduire de là une formule pour calculer les dérivées partielles d'ordre p d'un produit de $n + 1$ polynômes, par récurrence sur n. Cas $p = 1$?

e) On prend

$$K = k[X_1, \ldots, X_n]$$

où k est un anneau commutatif. Construire tous les opérateurs différentiels dans K qui sont nuls sur k.

§ 31

1. Soient $x_1, \ldots, x_n$ des éléments non nuls d'un anneau principal K et d un de leurs pgcd; on choisit $u_1, \ldots, u_n \in K$ tels que $u_1x_1 + \cdots + u_nx_n = d$. Montrer que $u_1, \ldots, u_n$ sont premiers entre eux.

¶ 2. Soient M un module libre de type fini sur un anneau principal K et a un élément non nul de M. Montrer que les cinq propriétés suivantes sont équivalentes : a fait partie d'une base de M; il existe une forme linéaire f sur M telle que $f(a) = 1$; les coordonnées de a par rapport à *une* base de M sont premières entre elles; les coordonnées de a par rapport à *toute* base de M sont premières entre elles; si $a = ux$ pour un $u \in K$ et un $x \neq 0$ dans M, alors u est inversible; si $ux = va$ avec $u, v \in K$ et $x \in M$ non nul alors v est multiple de u. (On utilisera l'*Exercice* 2 du § 18 et l'*Exercice* 11, *h*), du § 29). Un vecteur $a \in M$ satisfaisant aux conditions précédentes est dit **primitif.**

3. Soit M un module libre de type fini sur un anneau principal K. Montrer que tout $x \in M$ est multiple d'au moins un vecteur primitif de M. Faire le calcul en prenant $K = \mathbf{Z}$, $M = \mathbf{Z}^4$ et $x = (126, 210, 168, 504)$.

4. Soient $a_1, \ldots, a_n$ des éléments d'un anneau principal K; pour qu'il existe une matrice

$$U \in GL(n, K)$$

dont la première ligne (resp. colonne) soit précisément $a_1, \ldots, a_n$, il faut et il suffit que $a_1, \ldots, a_n$ soient premiers entre eux. On peut alors choisir U de telle sorte que dét(U) = 1, i.e. supposer

$$U \in SL(n, K).$$

5. Construire une matrice $U \in SL(3, \mathbf{Z})$ dont la première colonne soit 2, 3, 4.

6. Construire une matrice $U \in SL(3, \mathbf{Z})$ dont la seconde colonne soit 2, 3, 4.

¶ 7. Soit M un module libre de type fini sur un anneau principal K.

a) Montrer que, si a est un élément non nul de M, le pgcd des coordonnées de a par rapport à une base de M est indépendant du choix de celle-ci. Quelle est l'interprétation « géométrique » de ce résultat (cf. *Exercice* 3) ?

b) Soit M' un sous-module non nul de M. On choisit une base de M et on considère l'idéal de K engendré par toutes les coordonnées de tous les éléments de M'. Montrer que cet idéal

est indépendant du choix de la base. Montrer qu'il est engendré par les coordonnées d'un ensemble quelconque de générateurs de M'. [Cet idéal, ou l'un quelconque de ses générateurs, est appelé le *premier facteur invariant* de M' dans M; voir l'*Exercice* suivant.]

¶¶ **8.** Soient M un module libre de type fini sur un anneau principal K et M' un sous-module non nul de M; on note n et r les rangs (nombres d'éléments d'une base) de M et M'. On se propose de démontrer le résultat que voici : *il existe une base* $a_1, \ldots, a_n$ *de* M *et des éléments* $d_1, \ldots, d_r$ *de* K *tels que les vecteurs* $d_1a_1, \ldots, d_ra_r$ *forment une base de* M' *et que* d_i *divise* d_{i+1} *pour* $1 \leqslant i \leqslant r-1$.

a) Montrer que, pour toute forme linéaire f sur M, l'ensemble $f(M') \subset K$ est un idéal de K.

b) Montrer qu'il existe une forme linéaire f_1 sur M telle que, pour toute forme linéaire f sur M, la relation

$$f_1(M') \subset f(M') \quad \text{implique} f_1(M') = f(M').$$

Montrer qu'on a alors

$$f_1(M) = K.$$

c) On choisit f_1 satisfaisant à *b*); on pose

$$f_1(M') = (d_1)$$

et on choisit un vecteur $u_1 \in M'$ tel que

$$f_1(u_1) = d_1.$$

Montrer qu'on a

$$f(u_1) \in (d_1)$$

pour toute forme linéaire f sur M (en posant $f(u_1) = d$, montrer qu'il existe une combinaison linéaire g de f et f_1 telle que $g(u_1)$ soit un pgcd de d et d_1).

d) Déduire de (*c*) qu'on a

$$u_1 = d_1e_1$$

pour un vecteur $e_1 \in M$, tel que $f_1(e_1) = 1$.

e) Montrer que M est somme directe du sous-module engendré par e_1 et de $\mathrm{Ker}(f_1)$; et que M' est somme directe du sous-module engendré par u_1 et de $M' \cap \mathrm{Ker}(f_1)$. Montrer que $f(M') \subset f_1(M')$ pour toute f.

f) Achever la démonstration par récurrence sur n.

¶¶ **9.** Soit A une matrice à n lignes et p colonnes à coefficients dans un anneau principal K. Montrer à l'aide de l'*Exercice* 8 qu'il existe des matrices

$$U \in \mathrm{GL}(n, K) \qquad \text{et} \qquad V \in \mathrm{GL}(p, K)$$

telles que l'on ait

$$UAV = \begin{pmatrix} d_1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_r & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

où $d_1, \ldots, d_r$ sont des éléments non nuls de K tels que chacun divise le suivant. Les éléments

$d_1, \ldots, d_r$ sont appelés les **facteurs invariants** de la matrice A; on verra (*Exercice* 11) que les idéaux $(d_1), \ldots, (d_r)$ sont entièrement déterminés par A.

¶ **10.** Montrer que *tout groupe commutatif de type fini est isomorphe au produit direct d'un groupe* $\mathbf{Z}^q$ *et de groupes cycliques finis* $\mathbf{Z}/d_1\mathbf{Z}, \ldots, \mathbf{Z}/d_r\mathbf{Z}$ *où chaque* d_i *divise* d_{i+1}. (Observer qu'un **K-module de** type fini, où K est un anneau arbitraire, est isomorphe à un quotient M/M′ où M est libre de type fini — et appliquer *l'Exercice* 8). Comment ce résultat se généralise-t-il à un anneau principal quelconque?

¶ **11.** On reprend les hypothèses et notations de l'*Exercice* 8, dont on utilise les résultats.

a) Soient $j_1, \ldots, j_h$ des entiers tels que

$$1 \leqslant j_1 < j_2 < \ldots < j_h \leqslant r;$$

montrer que $d_1 \ldots d_h$ divise $d_{j_1} \ldots d_{j_h}$.

b) Soit h tel que $1 \leqslant h \leqslant r$; montrer que, si f est une forme h-linéaire alternée sur M, le produit $d_1 \ldots d_h$ divise $f(x_1, \ldots, x_h)$ quels que soient $x_1, \ldots, x_h \in \mathrm{M}'$; montrer qu'on peut en outre choisir f et $x_1, \ldots, x_h \in \mathrm{M}'$ de telle sorte que

$$d_1 \ldots d_h = f(x_1, \ldots, x_h);$$

en déduire que $d_1 \ldots d_h$ est un pgcd des éléments de K de la forme $f(x_1, \ldots, x_h)$ et en conclure que *les idéaux* (d_i) *sont entièrement déterminés par le module* M *et le sous-module* M′ (i.e. ne dépendent pas du choix des bases construites dans l'*Exercice* 8).

c) Soient $(a_i)_{1 \leqslant i \leqslant n}$ une base quelconque de M et $(b_j)_{1 \leqslant j \leqslant p}$ un système quelconque de générateurs de M′; on note A la matrice (à n lignes et p colonnes) formée avec les coordonnées des b_j par rapport à la base (a_i) de M.
Montrer que, pour $1 \leqslant h \leqslant r$, *l'élément* $d_1 \ldots d_h$ *est un pgcd des mineurs d'ordre* h *de la matrice* A.

d) En déduire que les facteurs $d_1, \ldots, d_r$ de *l'Exercice* 9 se calculent de même.

e) Soient A et B deux matrices à n lignes et p colonnes à coefficients dans un anneau principal K. Pour que A et B soient **équivalentes** (i.e. pour qu'il existe des matrices U et V *inversibles* telles que B = UAV) il faut et il suffit que A et B aient le même rang et les mêmes facteurs invariants.

(NB. — On exprime souvent ce résultat en introduisant, au lieu des facteurs invariants de A, ses **diviseurs élémentaires**

$$e_1 = d_1, e_2 = d_2/d_1, \ldots, e_r = d_r/d_{r-1})$$

¶ **12.** Soient K un anneau principal et

$$a_j = (\alpha_{1j}, \ldots, \alpha_{nj}) \quad (1 \leqslant j \leqslant p)$$

des éléments de K^n. Pour qu'ils fassent partie d'une base de K^n, il faut et il suffit que les mineurs d'ordre p de la matrice

$$\begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1p} \\ \ldots\ldots & & \ldots\ldots \\ \alpha_{n1} & \ldots & \alpha_{np} \end{pmatrix}$$

formés avec les composantes des vecteurs donnés soient *premiers entre eux* (et en particulier non tous nuls).

¶¶ **13.** Soient K un anneau principal et n et p deux entiers tels que $1 \leqslant p \leqslant n$. Soit A une matrice à n lignes et p colonnes à coefficients dans K. Pour qu'on puisse compléter A en une matrice carrée d'ordre n et *inversible* sur l'anneau K, il faut et il suffit que le pgcd des mineurs d'ordre p de A soit égal à 1.

14. Trouver toutes les matrices à coefficients entiers rationnels, de la forme

$$\begin{pmatrix} 1 & 4 & * \\ 2 & 5 & * \\ 1 & 6 & * \end{pmatrix},$$

et de déterminant 1.

¶ **15.** Soit A une matrice à coefficients dans un anneau commutatif K; on appelle opération élémentaire sur A une opération consistant soit à permuter deux lignes (resp. colonnes) de A, soit à ajouter à une ligne (resp. colonne) une combinaison linéaire des autres lignes (resp. colonnes), soit à multiplier une ligne (resp. colonne) par un élément inversible de K.

a) Montrer que toute matrice déduite de A par une succession d'opérations élémentaires est équivalente à A (i.e. de la forme UAV avec U, V inversibles sur K).

b) On suppose $K = \mathbf{Z}$, et $A \neq 0$. Soit d_1 le plus petit entier strictement positif possédant la propriété suivante : il existe une matrice qui se déduit de A par une succession d'opérations élémentaires, et dont d_1 est un coefficient. Montrer qu'il existe alors une matrice de la forme

$$\begin{pmatrix} d_1 & 0 \\ 0 & A_1 \end{pmatrix}$$

(où A_1 possède une ligne et une colonne de moins que A) qui se déduit de A par une succession d'opérations élémentaires, et de plus que tous les coefficients de A_1 sont multiples de d_1.

c) Déduire de là, pour $K = \mathbf{Z}$, une nouvelle démonstration du résultat de l'*Exercice* 9 (donc aussi de l'*Exercice* 8), et une méthode pratique pour réduire une matrice à coefficient dans $\mathbf{Z}$ à la forme canonique de l'*Exercice* 9.

d) Appliquer cette méthode aux matrices suivantes :

$$\begin{pmatrix} 0 & 2 & 4 & -1 \\ 6 & 12 & 14 & 5 \\ 0 & 4 & 14 & -1 \\ 10 & 6 & -4 & 11 \end{pmatrix}, \quad \begin{pmatrix} 0 & 6 & -9 & -3 \\ 12 & 24 & 9 & 9 \\ 30 & 42 & 45 & 27 \\ 66 & 78 & 81 & 63 \end{pmatrix},$$

$$\begin{pmatrix} 17 & -28 & 45 & 11 & 39 \\ 24 & -37 & 61 & 13 & 50 \\ 25 & -7 & 32 & -18 & -11 \\ 31 & 12 & 19 & -43 & -55 \\ 42 & 13 & 29 & -55 & -68 \end{pmatrix}.$$

16. Soit A une matrice à coefficients dans un anneau commutatif K quelconque, et soit B une matrice équivalente à A (i.e. de la forme UAV avec U et V inversibles sur l'anneau K). Montrer que, pour tout entier p inférieur au nombre de lignes et au nombre de colonnes de A, l'idéal de K engendré par les mineurs d'ordre p de A est égal à celui qui est engendré par les mineurs d'ordre p de B.

¶ **17.** Soit A une matrice carrée d'ordre n à coefficients dans un anneau principal K. Montrer qu'il existe une matrice $U \in GL(n, K)$ telle que UA soit triangulaire (utiliser les *Exercices* 1 et 4 et raisonner par récurrence sur n). Interprétation géométrique?

¶ **18.** On considère un système d'équations linéaires

$$\begin{cases} a_{11}x_1 + \cdots + a_{1p}x_p = b_1 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \cdots + a_{np}x_p = b_n \end{cases}$$

à coefficients, seconds membres et inconnues dans un anneau principal K. Montrer que, pour que ce système possède au moins une solution, il faut et il suffit d'une part que les deux matrices

$$\mathrm{A} = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \cdots\cdots & & \cdots\cdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix}, \qquad \mathrm{B} = \begin{pmatrix} a_{11} & \cdots & a_{1p} & b_1 \\ \cdots\cdots & & \cdots\cdots & \cdots\cdots \\ a_{n1} & \cdots & a_{np} & b_n \end{pmatrix}$$

aient le même rang r, d'autre part qu'un pgcd des mineurs d'ordre r de A soit égal à un pgcd des mineurs d'ordre r de B.

¶ **19.** Soient K un anneau principal, f une forme p-linéaire sur le module K^n et soit d un pgcd des coefficients de f par rapport à la base canonique; montrer qu'il existe $x_1, \ldots, x_p \in \mathrm{K}^n$ tels que

$$f(x_1, \ldots, x_p) = d.$$

20. Démontrer que tout nombre rationnel peut s'écrire, d'une façon et d'une seule, sous la forme d'une somme finie de fractions de la forme

$$a/p^n$$

avec p premier, $n \geqslant 1$, et $1 \leqslant a \leqslant p - 1$.
Effectuer cette décomposition pour les nombres

$$\frac{1887}{5400}, \qquad \frac{122}{1323}.$$

¶ **21.** Soit K un anneau d'intégrité commutatif. On dit qu'un $p \in \mathrm{K}$ est **irréductible** s'il n'est pas inversible et s'il n'a pas d'autres diviseurs dans K que ceux qui sont évidents (à savoir les éléments inversibles de K, et les up où $u \in \mathrm{K}$ est inversible).
On dit que K est **factoriel** s'il possède les deux propriétés suivantes :

(UFD 1) : Tout élément non inversible de K est produit d'un nombre fini d'éléments irréductibles.

(UFD 2) : Si $p_1 \ldots p_r = q_1 \ldots q_s$ où les p_i et les q_j sont des éléments irréductibles de K, alors on a $r = s$, et on peut changer l'ordre des q_j de telle sorte que l'on ait

$$\mathrm{K}p_i = \mathrm{K}q_i \quad \text{pour} \quad 1 \leqslant i \leqslant r$$

(ou, ce qui revient au même, $q_i = u_i p_i$ avec u_i inversible).
Ces propriétés expriment que *tout élément de* K *s'écrit, d'une façon essentiellement unique, sous la forme d'un produit d'éléments irréductibles de* K.

a) Montrer que tout anneau principal est factoriel (la réciproque est fausse; cf. § 32, *Exercice* 31).

b) Montrer que, dans la définition d'un anneau factoriel, on peut remplacer la condition (UFD 2) par

(UFD 3) : Si un élément irréductible p de K divise un produit xy, il divise l'un au moins des facteurs de ce produit.

c) Montrer qu'on peut aussi remplacer (UFD 2) par

(UFD 4) Pour tout élément irréductible p de K, l'idéal $\mathrm{K}p$ est premier.

d) Soit K un anneau factoriel. Montrer que, quels que soient $x, y \in K$ il existe un $d \in K$ tel que les diviseurs communs à x et y soient exactement les diviseurs de d (on dit que d est un pgcd de x et y), et que d est unique modulo la possibilité de le multiplier par un élément inversible de K.

e) Soient K un anneau factoriel, L son corps des fractions, et $\mathfrak{p}$ l'idéal premier de K engendré par un élément irréductible p de K. Montrer que l'anneau local $K_{\mathfrak{p}}$ [§ 8, *Exercice* 7, (*g*)] est l'anneau d'une valuation discrète (§ 8, *Exercice* 6) de L.

f) Si un anneau d'intégrité commutatif est à la fois factoriel et de Dedekind, il est principal.

g) Soit K un anneau d'intégrité commutatif noethérien. Montrer que tout élément non inversible de K est produit d'éléments irréductibles — autrement dit que K vérifie (UFD 1). [Mais un anneau noethérien n'est pas nécessairement factoriel, i.e. la décomposition en éléments irréductibles peut ne pas être unique : prendre un anneau de Dedekind non principal; ce dernier phénomène se produit notamment pour l'anneau des entiers d'un corps de nombres algébriques, et a longtemps bloqué les progrès dans l'étude de ces anneaux — jusqu'à Dedekind, qui reconnut le premier que la notion importante, dans ce cas, était celle d'*idéal* premier et non d'*élément* irréductible, contrairement à ce qu'indiquait une analogie trompeuse avec les entiers rationnels.]

§ 32

1. Trouver le quotient et le reste de la division de

$$2X^4 - 3X^3 + 4X^2 - 5X + 6 \quad \text{par} \quad X^2 - 3X + 1$$
$$4X^3 + X^2 \quad \text{par} \quad X + 1 + i$$
$$X^4 - 2X^3 + 4X^2 - 6X + 8 \quad \text{par} \quad X - 1.$$

2. Calculer le pgcd des polynômes suivants :

a) $X^6 - 7X^4 + 8X^3 - 7X + 7$ et $3X^5 - 7X^3 + 3X^2 - 7$;

b) $X^5 + X^4 - X^3 - 3X^2 - 3X - 1$ et $X^4 - 2X^3 - X^2 - 2X - 1$;

c) $X^3 + X^5 - X^4 - 2X^3 - X^2 + X + 1$, $X^5 + X^3 - X^2 - 1$ et $X^4 - 2X^3 - X + 2$.

3. Pour chacun des couples de polynômes p, q indiqués ci-dessous, trouver des polynômes u et v tels que $up + vq$ soit un pgcd de p et q :

a) $X^5 + 3X^4 + X^3 + X^2 + 3X + 1$ et $X^4 + 2X^3 + X + 2$;

b) $3X^5 + 5X^4 - 16X^3 - 6X^2 - 5X - 6$ et $3X^4 - 4X^3 - X^2 - X - 6$;

c) $X^5 + 5X^4 + 9X^3 + 7X^2 + 5X + 3$ et $X^4 + 2X^3 + 2X^2 + X + 1$;

d) X^4 et $(1 - X)^4$

4. Trouver un polynôme de degré aussi petit que possible 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 $2X^2 - 3$.

5. Montrer que le pgcd des polynômes

$$X^m - 1 \quad \text{et} \quad X^n - 1$$

est $X^d - 1$, où d est le pgcd des entiers m et n.

6. Soient p et q deux polynômes à une indéterminée. Si le polynôme $p(X^3) + Xq(X^3)$ est divisible par $X^2 + X + 1$, alors $p(1) = q(1) = 0$.

7. Soit

$$f(X) = \frac{p(X)}{q(X)}.$$

une fraction rationnelle à une variable à coefficients dans un corps K, et soit a une racine *simple* de son dénominateur q, de sorte que la contribution du pôle a dans la décomposition de f en éléments simples est

$$\frac{A}{X-a}$$

pour un certain $A \in K$. Montrer qu'on a

$$A = \frac{p(a)}{q'(a)}$$

[Écrire

$$\frac{p(X)}{q(X)} = \frac{A}{X-a} + \frac{r(X)}{s(X)}$$

avec $r(a) \neq 0$, $s(a) \neq 0$, réduire au même dénominateur, dériver, et faire $X = a$ dans le résultat obtenu. Le processus de « passage à la limite » qu'on utilise en général dans les cours d'Analyse pour démontrer ce résultat ne s'étend pas à un corps quelconque].

8. Décomposer en éléments simples (sur C, puis sur R) les fractions rationnelles suivantes :

$$\frac{X^2+1}{X(X^2-1)}, \quad \frac{2}{(X-1)(X-2)(X-3)}, \quad \frac{X^5-X^3-X^2}{X^2-1}, \quad \frac{4X^3}{(X^2+1)^3},$$

$$\frac{X^6-X^2+1}{(X-1)^3}, \quad \frac{3X^2+3}{X^3-3X-2}, \quad \frac{X^5}{(X^4-1)^2}.$$

Le lecteur qui estimerait ces exemples insuffisants pourra facilement en construire autant qu'il en désire : la méthode consiste à choisir au hasard deux polynômes (en s'arrangeant tout de même pour que les racines du dénominateur soient évidentes, ou en tous cas calculables, si l'on désire avoir des résultats explicites).

¶ 9. Soient L un corps commutatif, K un sous-corps de L, et x un élément de L algébrique sur K. Soit I l'ensemble des polynômes $f \in K[X]$ tels que $f(x) = 0$; montrer que c'est un idéal de K[X]. En déduire qu'il existe un et un seul polynôme

$$(*) \qquad f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

à coefficients dans K qui vérifie $f(x) = 0$ et tel que tout polynôme $g \in K[X]$ vérifiant $g(x) = 0$ soit un multiple de f. Montrer que

$$(**) \qquad x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

est l'équation (à coefficients dans K) de plus petit degré possible vérifiée par x. On dit que $(*)$ est le **polynôme minimal** et $(**)$ l'**équation minimale** de x sur K; son degré n s'appelle **le degré de x sur** K. Montrer que, considéré comme espace vectoriel sur K, le sous-corps K[x] est de dimension n et admet pour base les éléments

$$1, x, \ldots, x^{n-1}.$$

Autrement dit, on a

$$[K[x] : K] = n$$

avec les notations du § 26, *Exercice* 5.

On prend $L = \mathbf{C}$, $K = \mathbf{Q}$ dans ce qui précède. Trouver les équations minimales des éléments suivants de L :

$$\sqrt{2} + \sqrt{3}; \qquad \sqrt[3]{2} + \sqrt{5}$$

¶ **10.** Soit K un corps commutatif.

a) Soient L un sur-corps de K et x un élément de L algébrique sur K. Montrer que le polynôme minimal de x sur K est irréductible (sur K).

b) Soit f un polynôme irréductible à une variable à coefficients dans K, et de coefficient dominant égal à 1. Soit x une racine de f dans une extension de K. Montrer que f est le polynôme minimal de x sur K.

c) Le polynôme f étant comme dans la question précédente, soient x et y deux racines de f dans un sur-corps L de K. Montrer qu'il existe un et un seul isomorphisme j du corps $K(x)$ sur le corps $K(y)$ vérifiant

$$j(x) = y, \qquad j(a) = a \text{ pour tout } a \in K.$$

d) On suppose de plus que K est le corps des fractions d'un anneau A. Avec les notations de la question *c*), montrer que si x est entier sur A (§ 26, *Exercice* 6) il en est de même de y.

¶¶ *e*) Soient A un anneau d'intégrité commutatif, K son corps des fractions et L un sur-corps de K. Soit $x \in L$ entier sur A; montrer que les coefficients du polynôme minimal f de x sur K sont entiers sur A (plonger L dans un corps algébriquement clos, observer que toutes les racines de f sont des entiers sur A et appliquer le § 33, n° 6). En déduire que f est à coefficients dans A si A est intégralement clos (i.e. si tout élément de K entier sur A est dans A).

¶¶ *f*) On suppose dans ce qui précède que L est une extension algébrique de degré fini de K (§ 26, *Exercice* 4). Montrer que, si $x \in L$ est entier sur A, et si A est intégralement clos, on a

$$\mathrm{Tr}_{L/K}(x) \in A, \qquad \mathrm{N}_{L/K}(x) \in A$$

(utiliser l'*Exercice* 5 du § 26).

¶¶ **11.** Soient L un corps commutatif, K un sous-corps de L, et x un élément de L algébrique sur K. On dit que x est **séparable** sur K s'il est racine *simple* de son polynôme minimal f sur K.

a) Pour que x soit séparable sur K, il faut et il suffit que x soit racine simple d'au moins une équation algébrique à coefficients dans K.

b) On a $f' = 0$ si x n'est pas séparable sur K, et réciproquement.

c) Si K est de caractéristique 0, tout x algébrique sur K est séparable sur K, et toutes les racines de tout polynôme irréductible à coefficients dans K sont simples.

d) Si K est de caractéristique $p \neq 0$, pour tout $x \in L$ algébrique sur K il existe un entier $n \geqslant 0$ tel que

$$x^{p^n}$$

soit séparable sur K.

e) Soit L une extension algébrique de degré fini de K (§ 26, *Exercice* 4). Pour que L soit séparable sur K (§ 26, *Exercice* 4, (*h*)) il faut et il suffit que tout $x \in L$ soit séparable sur K.

¶ **12.** Si un polynôme $f \in \mathbf{Z}[X]$ non constant n'est pas irréductible dans l'anneau $\mathbf{Q}[X]$, alors on peut le décomposer de façon non triviale en produit de polynômes à coefficients *entiers* rationnels (utiliser le lemme de Gauss, § 27, *Exercice* 13).

13. Parmi les polynômes

$$X^3 + X + 1, \qquad X^4 + X + 1, \qquad X^4 + X^2 + 1, \qquad X^3 + 7X + 7, \qquad X^6 + 3X + 2,$$

quels sont ceux qui sont irréductibles sur $\mathbf{Q}$?

¶ **14.** Soit $f(X) = a_0 + a_1X + \cdots + a_nX^n$ un polynôme à coefficients dans **Z**. On suppose qu'un certain nombre premier p divise $a_0, \ldots, a_{n-1}$, ne divise pas a_n, et de plus que a_0 n'est pas divisible par p^2. Montrer que f est irréductible sur **Q** (**critère d'irréductibilité d'Eisenstein**; utiliser l'*Exercice* 12).

¶ **15.** (Diviseurs élémentaires d'une matrice à coefficients polynomiaux). Soient k un corps commutatif et $K = k[X]$ l'anneau des polynômes à une indéterminée à coefficients dans k.

a) Montrer que $GL(n, K)$ est l'ensemble des matrices $U \in M_n(K)$ dont le déterminant est un élément non nul du corps k (le déterminant d'une telle matrice est donc « constant »).

b) On utilise dans ce qui suit, pour les matrices à coefficients dans K, la notion d'opération élémentaire du § 31, *Exercice* 15. Soit A une matrice rectangulaire non nulle à coefficients dans K et soit $a_{ij}(X)$ le coefficient situé à l'intersection de la i^e colonne et de la j^e ligne de A. Montrer qu'on peut, à l'aide d'un nombre fini d'opérations élémentaires, remplacer les coefficients situés sur la i^e colonne ou la j^e ligne de A par les restes de leur division par $a_{ij}(X)$.

c) Soit $d_1(X)$ un polynôme non nul, de coefficient dominant égal à 1, et de plus petit degré possible parmi tous les coefficients non nuls de toutes les matrices déduites de A par une succession d'opérations élémentaires. Montrer qu'on peut déduire de A, par une succession d'opérations élémentaires, une matrice de la forme

$$\begin{pmatrix} d_1 & 0 \\ 0 & A_1 \end{pmatrix}$$

où A_1 a une ligne et une colonne de moins que A, et pour coefficients des polynômes tous divisibles par d_1. Montrer que d_1 est un pgcd des coefficients non nuls de A (et est par suite entièrement déterminé par la connaissance de A).

d) Montrer qu'on peut déduire de A, par une succession d'opérations élémentaires, une matrice de la forme

$$\begin{pmatrix} d_1(X) & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & d_2(X) & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & d_r(X) & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & 0 \end{pmatrix}$$

où les d_i sont des polynômes non nuls tels que chacun divise le suivant. Montrer, à l'aide de l'*Exercice* 16 du § 31, que pour tout entier i inférieur au nombre de lignes et au nombre de colonnes de A, le polynôme $d_1(X) \ldots d_i(X)$ (où l'on pose $d_i = 0$ pour $i \geqslant r + 1$) est un pgcd des mineurs d'ordre i de A. En déduire que l'entier r est égal au rang de A, et que pour $1 \leqslant i \leqslant r$ les polynômes $d_i(X)$ sont entièrement déterminés par A si on impose à leurs coefficients dominants d'être égaux à 1.

e) On appelle $d_1, \ldots, d_r$ les **facteurs invariants** de la matrice A, et **diviseurs élémentaires** de A les quotients $d_i(X)/d_{i+1}(X)$. Enfin, on dit que deux matrices A et B à coefficients dans $K = k[X]$, et ayant toutes deux p lignes et q colonnes, sont **équivalentes** s'il existe des matrices

$$U \in GL(p, K) \qquad \text{et} \qquad V \in GL(q, K)$$

telles que $B = UAV$. Montrer que, pour que A et B soient équivalentes, il faut et il suffit que A et B aient même rang et mêmes facteurs invariants, et qu'on peut alors passer de A à B par une succession d'opérations élémentaires.

¶¶ *f*) Soient M un $\mathbb{K}$-module isomorphe à $\mathbb{K}^n$ et M' un sous-module de M. Montrer qu'il existe une base $(a_1, \ldots, a_n)$ de M, et des polynômes $d_1, \ldots, d_p \in \mathbb{K}$ tels que d_i divise d_{i+1} et que M'

soit engendré par $d_1a_1, \ldots, d_pa_p$ (on n'interdit pas à certains des d_i d'être nuls). [Appliquer la question *d*) à la matrice, par rapport à une base de M, d'un endomorphisme de M ayant M' pour image].

¶ *g*) Soit E un K-module de type fini (mais non nécessairement libre). Montrer que E est isomorphe au produit direct d'un module de la forme K^s et de modules de la forme $K/d_1K, \ldots, K/d_rK$ où $d_1, \ldots, d_r$ sont des polynômes non nuls tels que d_i divise d_{i+1} pour $1 \leqslant i \leqslant r-1$. [Choisir un homomorphisme f de K^p *sur* E et appliquer la question précédente à Ker (f)]. Montrer que les entiers r et s, et les polynômes d_i (dont on supposera qu'ils ont 1 pour coefficient dominant), sont entièrement déterminés par E et les conditions qu'on leur a imposées. [On dit que $d_1, \ldots, d_r$ sont les facteurs invariants du K-module E; l'entier s est le rang de E au sens du § 29, *Exercice* 11, *e*)]. Montrer que deux K-modules de type fini sont isomorphes si et seulement si leurs rangs et leurs facteurs invariants sont égaux.

¶ *h*) Déduire les résultats précédents du § 31, *Exercices* 8, 9, 10 et 11 et du fait que l'anneau K est principal.
[Les résultats de cet *Exercice*, qui constituent l'analogue pour les anneaux de polynômes à une variable sur un corps de la théorie des diviseurs élémentaires des matrices à coefficients dans Z (§ 31, *Exercice* 17), ont des applications importantes, notamment à la théorie des systèmes d'équations différentielles linéaires d'ordre quelconque à coefficients constants; on trouvera d'excellents exposés de ces résultats dans certains des ouvrages cités dans la Bibliographie (notamment dans Albert, Gelfand, Schreier-Sperner); mais la véritable explication de ces résultats est évidemment la théorie des modules de type fini sur un anneau principal.]
Dans les *Exercices* 16 à 21 qui suivent (*), on demande de réduire la matrice donnée à la forme canonique de l'*Exercice* 15, *d*), et d'en calculer les diviseurs élémentaires (le corps de base est C).

16. $\begin{pmatrix} X & 1 \\ 0 & X \end{pmatrix}$

17. $\begin{pmatrix} X^2-1 & X+1 \\ X+1 & X^2+2X+1 \end{pmatrix}$

18. $\begin{pmatrix} 1-X & X^2 & X \\ X & X & -X \\ 1+X^2 & X^2 & -X^2 \end{pmatrix}$

19. $\begin{pmatrix} X & 1 & 0 & 0 \\ 0 & X & 1 & 0 \\ 0 & 0 & X & 1 \\ 0 & 0 & 0 & X \end{pmatrix}$

20. $\begin{pmatrix} X & -1 & 0 & 0 & 0 \\ 0 & X & -1 & 0 & 0 \\ 0 & 0 & X & -1 & 0 \\ 0 & 0 & 0 & X & -1 \\ 1 & 2 & 3 & 4 & 5+X \end{pmatrix}$

21. $\begin{pmatrix} X & 1 & 1 & \ldots & 1 \\ 0 & X & 1 & \ldots & 1 \\ 0 & 0 & X & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & X \end{pmatrix}$ (n lignes et colonnes)

Dans les *Exercices* 22 et 23 qui suivent, on demande de trouver des matrices U et V à coefficients polynomiaux, de déterminant constant non nul, telles que, A désignant la matrice donnée, UAV soit mise sous la forme canonique de l'*Exercice* 15.

22. $\begin{pmatrix} X^4+4X^3+4X^2+X+2, & X^3+4X^2+4X \\ X^4+5X^3+8X^2+5X+2, & X^3+5X^2+8X+4 \end{pmatrix}$

23. $\begin{pmatrix} X^4+3X^3-5X^2+X+1, & 2X^4+3X^3-5X^2+X-1, & 2X^4+2X^3-4X^2 \\ X^4-X^3+1, & 2X^4-X^3-X^2, & 2X^3-2X^2 \\ X^4+2X^3-4X^2+X+1, & 2X^4+2X^3-4X^2+X-1, & 2X^4+X^3-3X^2 \end{pmatrix}$

(*) Les *Exercices* 16 à 26 sont extraits du recueil de Proskurjakov, où le lecteur trouvera de nombreux autres énoncés semblables.

24. Vérifier que les deux matrices suivantes, à coefficients dans $\mathbf{C}[X]$, sont équivalentes :

$$\begin{pmatrix} X^3+6X^2+6X+5, & X^3+4X^2+4X+3 \\ X^3+3X^2+3X+2, & X^3+2X^2+2X+1 \\ 2X^3+3X^2+3X+1, & 2X^3+2X^2+2X \end{pmatrix} = A$$

$$\begin{pmatrix} X^3+X^2+X, & 2X^3+X^2+X-1 \\ 3X^3+2X^2+2X-1, & 6X^3+2X^2+2X-4 \\ X^3-X^2-X-2, & 2X^3-X^2-X-3 \end{pmatrix} = B$$

(on calculera des matrices U, V telles que $B = UAV$).

25. Calculer les facteurs invariants de la matrice

$$\begin{pmatrix} X^3+X^2-X+3, & X^3-X^2+X, & 2X^3+X^2-X+4, & X^3+X^2-X+2 \\ X^3+3X^2-3X+6, & X^3-3X^2+3X-2, & 2X^3+3X^2-3X+7, & X^3+3X^2-3X+4 \\ X^3+2X^2-2X+4, & X^3-2X^2+2X-1, & 2X^3+2X^2-2X+5, & X^3+2X^2-2X+3 \\ 2X^3+X^2-X+5, & 2X^3-X^2+X+1, & 4X^3+X^2-X+7, & 2X^3+X^2-X+3 \end{pmatrix}$$

26. Calculer les diviseurs élémentaires de la matrice

$$\begin{pmatrix} X^4+1, & X^7-X^4+X^3-1, & X^4-4X^3+4X-5 \\ 2X^4+3, & 2X^7-2X^4+4X^3-2, & 3X^4-10X^3+X^2+10X-14 \\ X^4+2, & X^7-X^4+2X^3-2, & 2X^4-6X^3+X^2+6X-9 \end{pmatrix}$$

en prenant pour anneau de base soit $\mathbf{Q}[X]$, soit $\mathbf{R}[X]$, soit $\mathbf{C}[X]$.

¶ **27.** On se propose de démontrer que *si* K *est un anneau commutatif noethérien, l'anneau de polynômes* K[X] *est noethérien.* On désigne par I un idéal de K[X].

a) Pour tout entier $n \geqslant 0$, soit $J_n \subset K$ l'ensemble formé de 0 et des $a \in K$ vérifiant la condition suivante : il existe un polynôme $f \in I$, de degré n, dont le coefficient dominant est égal à a. Montrer que les J_n forment une suite croissante d'idéaux de K. En conclure qu'on a

$$J_r = J_{r+1} = \cdots$$

pour un certain entier r.

b) Pour tout entier i tel que $0 \leqslant i \leqslant r$, on choisit dans I des polynômes f_{ij} $(1 \leqslant j \leqslant n_i)$ en nombre fini, de degré i, dont les coefficients dominants a_{ij} engendrent l'idéal J_i. Montrer que, pour tout $f \in I$, il existe des polynômes $q_{ij} \in K[X]$ tels que l'on ait

$$f = \sum_{\substack{1 \leqslant j \leqslant n_i \\ 0 \leqslant i \leqslant r}} q_{ij} f_{ij} + g \qquad \text{avec} \qquad d^0(g) < d^0(f).$$

c) En déduire, par récurrence sur le degré de f, que les $n_0 + \cdots + n_r$ polynômes f_{ij} engendrent l'idéal I.

d) Déduire du résultat précédent que si un anneau commutatif L contient un sous-anneau noethérien K et des éléments $x_1, \ldots, x_n$ en nombre fini tels que $L = K[x_1, \ldots, x_n]$, alors L est noethérien. (Observer que L est un quotient d'un anneau de polynômes à coefficients dans K).

¶ **28.** Soit K un corps commutatif infini. On rappelle qu'une partie V de K^n est appelée une *variété algébrique* s'il existe un nombre *fini* de polynômes $p_1, \ldots, p_s \in K[X_1, \ldots, X_n]$ tels que V soit l'ensemble des $x \in K^n$ où l'on a $p_1(x) = \cdots = p_s(x) = 0$; et qu'une partie A de K^n est appelée un *ouvert de Zariski* si l'ensemble complémentaire $K^n - A$ est une variété algébrique

(§§ 27, 28, *Exercice* 1). En utilisant le fait que l'anneau $K[X_1, \ldots, X_n]$ est noethérien, démontrer les propriétés suivantes :

a) L'intersection d'une famille (finie ou infinie) de variétés algébriques dans K^n est une variété algébrique dans K^n. Toute réunion (finie ou non) d'ouverts de Zariski est un ouvert de Zariski.

b) Toute suite décroissante de variétés algébriques dans K^n est stationnaire. Toute suite croissante d'ouverts de Zariski est stationnaire.

Montrer en outre qu'on a

c) La réunion d'une famille *finie* de variétés algébriques dans K^n est encore une variété algébrique dans K^n. L'intersection d'une famille *finie* d'ouverts de Zariski est encore un ouvert de Zariski.

d) L'intersection de deux ouverts de Zariski non vides est non vide. Si U et V sont deux variétés algébriques dans K^n, telles que $U \neq K^n$ et $V \neq K^n$, alors on a $U \cup V \neq K^n$.

(On aura intérêt, pour chaque variété algébrique V dans K^n, à introduire l'idéal $I(V) \subset K[X_1, \ldots, X_n]$ formé des polynômes qui sont nuls en tout $x \in V$, et à interpréter en termes d'idéaux les opérations qu'on demande d'effectuer sur les variétés algébriques).

¶¶ **29.** Soit K un anneau commutatif noethérien. Montrer que l'anneau de séries formelle $K[[X]]$ (§§ 27, 28, *Exercice* 11) est noethérien. (Étant donné un idéal I de $K[[X]]$, considérer pour tout $n \geqslant 0$ l'idéal J_n de K formé des coefficients du terme en X^n dans les $f \in I$ qui ne comportent aucun terme de degré $\leqslant n-1$).

¶¶¶ 30. Soient V un espace vectoriel de dimension finie sur un corps commutatif K de caractéristique 0, et G un groupe fini, d'ordre r, d'automorphismes de V. On désigne par A l'anneau des fonctions polynomiales sur V (§§ 27, 28, *Exercice* 17; ou bien § 28, n° 2 dans le cas où $V = K^n$, auquel on peut évidemment se ramener).

Étant donnés un $s \in G$ et une fonction polynomiale f sur V, on définit une nouvelle application f_s de V dans K par

$$f_s(x) = f(s^{-1}(x)) \quad \text{pour tout } x \in V.$$

On dit que f est un invariant du groupe G si $f_s = f$ pour tout $s \in G$. On note $I \subset A$ l'ensemble de ces invariants, qui est un sous-anneau de A.

a) Montrer qu'on a $f_s \in A$ pour toute $f \in A$ et tout $s \in G$, et que I est un sous-anneau de A.

b) Pour toute fonction polynomiale $f \in A$, on définit la fonction polynomiale

$$f^\natural = \frac{1}{r} \sum_{s \in G} f_s;$$

montrer que $f^\natural$ est un invariant de G. Montrer qu'on a les relations

$$(f+g)^\natural = f^\natural + g^\natural \quad \text{quels que soient } f, g \in A$$
$$f^\natural = f \quad \text{si et seulement si } f \in I$$
$$(fg)^\natural = f^\natural g \quad \text{quels que soient } f \in A \text{ et } g \in I.$$

c) Montrer que si f est un invariant de G, il en est de même de toutes les composantes homogènes de f (*Exercice* 17, §§ 27, 28).

d) Soit J l'idéal de l'anneau A engendré par I. En tenant compte de la question *c*) et du fait que A est noethérien (*Exercices* 14 et 15), montrer qu'il existe dans I des polynômes homogènes

$$f_1, \ldots, f_p$$

le nombre fini qui engendrent J. On pose dans ce qui suit $q_i = d^0(f_i)$.

e) Pour tout $f \in I$ homogène de degré q, montrer qu'il existe des $u_i \in A$ homogènes de degrés $q - q_i$ (on prendra $u_i = 0$ si $q - q_i < 0$) tels que $f = \sum f_i u_i$. Montrer qu'on peut même prendre les u_i dans I (Ecrire que $f = f^*$).

f) En raisonnant par récurrence sur le degré de f, déduire de là que *tout* $f \in I$ *est un polynôme en les* f_i, *à coefficients dans* K, autrement dit que les invariants du groupe G forment un anneau engendré sur K par un nombre *fini* d'éléments (*théorème des invariants* de Hilbert).

¶¶ **31.** (La résolution de cet *Exercice* suppose acquis les résultats de l'*Exercice* 21 du § 31). On se propose de démontrer que *si* A *est un anneau factoriel, l'anneau* A[X] *est factoriel.*

a) Étant donnés des polynômes $f, g \in A[X]$, soit p un élément irréductible de A qui divise tous les coefficients de fg; montrer que p divise tous les coefficients de f, ou bien tous ceux de g. (Raisonner comme dans l'*Exercice* 13 du § 27).

b) On dit qu'un polynôme $f \in A[X]$ est primitif si le pgcd de ses coefficients est 1. Montrer que si f et g sont primitifs, il en est de même de fg.

c) Pour tout $f \in A[X]$ non nul, on note $c(f)$ un pgcd de ses coefficients. Montrer que

$$c(fg) = c(f)c(g)$$

(lemme de Gauss pour les anneaux factoriels).

d) Soit K le corps des fractions de A. Montrer que si un $f \in A[X]$ n'est pas irréductible dans K[X], il n'est pas non plus irréductible dans A[X].

e) Déduire de là et de la question *a*) que les éléments irréductibles de l'anneau A[X] sont les éléments irréductibles de l'anneau factoriel A, et les polynômes non constants qui sont primitifs, et irréductibles dans l'anneau K[X].

f) Déduire de là et des résultats du § 32, n° 3 (qu'on appliquera à K[X]) que tout élément de A[X] s'écrit, d'une façon essentiellement unique, sous la forme d'un produit d'éléments irréductibles de A[X], et par suite que A[X] est factoriel comme annoncé.

g) Montrer que, si A est un anneau factoriel (par exemple si A est un corps, ou bien si $A = \mathbf{Z}$), l'anneau $A[X_1, \ldots, X_n]$ est factoriel.
En particulier, *tout polynôme* (à n variables) *à coefficients dans un corps* K *se décompose, d'une façon essentiellement unique, en un produit de polynômes irréductibles* à coefficients dans K (un polynôme f à coefficients dans K étant dit irréductible s'il est non constant, et si chacun de ses diviseurs est constant, ou est proportionnel à f).

h) Montrer que l'anneau Z[X] (qui est factoriel d'après ce qui précède) *n'est pas* principal (examiner l'idéal engendré par 2 et X). Même question pour K[X, Y] où K est un corps.

i) Montrer que, pour tout corps commutatif K, le polynôme $Y^2 - X^3$ est irréductible dans l'anneau K[X, Y]. (On écrira K[X, Y] = A[Y] où A = K[X] et on appliquera la question *d*) ci-dessus).

j) Montrer que toute fraction rationnelle f à n variables, à coefficients dans un corps K, peut se mettre sous la forme $f = p/q$ où p et q sont des polynômes à n variables, à coefficients dans K, et premiers entre eux (i.e. n'ayant aucun diviseur commun en dehors des constantes); et que, de plus, p et q sont uniques à des facteurs constants près. Montrer que les points d'indétermination de f sont les éléments de K^n où p et q s'annulent simultanément, et que les pôles de f sont les $x \in K^n$ où l'on a $p(x) \neq 0$ et $q(x) = 0$.

k) Soient p et q deux polynômes non constants à $n \geqslant 2$ indéterminées et à coefficients dans un corps commutatif K. On suppose p et q premiers entre eux; s'ensuit-il qu'il existe des polynômes u et v à n indéterminées, à coefficients dans K, tels que

$$up + vq = 1 ?$$

[L'interprétation géométrique du fait que l'anneau $C[X_1, \ldots, X_n]$, par exemple, est factoriel est la suivante. Soit $f \in C[X_1, \ldots, X_n]$ non constant et soit V l'hypersurface de C^n définie

par l'équation $f(x) = 0$. Soit

$$f(X) = \prod_{i=1}^{i=s} p_i(X)^{r_i}$$

la décomposition de f en produit de facteurs irréductibles, et soit V_i l'hypersurface $p_i(x) = 0$. Alors V_i est irréductible (i.e. ne peut pas se représenter de façon non triviale comme réunion de deux autres variétés algébriques), on a

$$V = V_1 \cup \ldots \cup V_s,$$

et cette décomposition de V en hypersurfaces irréductibles est unique à l'ordre près.
On peut encore se placer au point de vue suivant. Soit V une variété algébrique dans $\mathbf{C}^n$ et soit $\mathfrak{a}$ l'idéal de $\mathbf{C}[X_1, \ldots, X_n]$ formé des polynômes f tels que $f(x) = 0$ pour tout $x \in V$. Comme $\mathbf{C}[X_1, \ldots, X_n]$ est noethérien, l'*Exercice* 9, *b*), du § 18 montre que $\mathfrak{a}$ est intersection finie d'idéaux primaires de $\mathbf{C}[X_1, \ldots, X_n]$ (et même d'idéaux *premiers* : appliquer l'*Exercice* 11 du § 18 en remarquant que l'idéal $\mathfrak{a}$ est identique à son radical, attendu que la relation

$$f(x)^q = 0 \text{ sur } V \quad \text{implique} \quad f(x) = 0 \text{ sur } V$$

pour des raisons triviales); écrivons donc

$$\mathfrak{a} = \mathfrak{p}_1 \cap \ldots \cap \mathfrak{p}_s$$

où les $\mathfrak{p}_i$ sont les idéaux premiers minimaux de $\mathfrak{a}$, et soit V_i la variété algébrique de $\mathbf{C}^n$ formée des x tels que

$$f(x) = 0 \quad \text{pour tout } f \in \mathfrak{p}_i$$

(V_i est définie par un nombre fini d'équations si l'on veut : prendre des générateurs de $\mathfrak{p}_i$). Comme $\mathfrak{p}_i$ est premier, chaque V_i est irréductible, et l'on a

$$V = V_1 \cup \ldots \cup V_s;$$

on dit que les V_i sont les **composantes irréductibles** de V. Ceci dit, le fait que l'anneau $\mathbf{C}[X_1, \ldots, X_n]$ soit factoriel montre que si V est une hypersurface (i.e. peut être définie par une seule équation) il en est de même de ses composantes irréductibles; ou encore : si une variété *irréductible* W est contenue dans une hypersurface V, il existe une *hypersurface* irréductible W' telle que $W \subset W' \subset V$, résultat « évident » géométriquement...
Comme autre exemple important d'anneau factoriel, citons (Weierstrass) l'anneau des séries entières *convergentes* (i.e. à domaine de convergence non réduit à 0) à n variables complexes; cet anneau intervient dans l'étude « locale » des « variétés analytiques » dans $\mathbf{C}^n$ (parties de $\mathbf{C}^n$ définies par des équations dont les premiers membres sont des fonctions holomorphes). Cet anneau est aussi noethérien].

¶¶ 33. Étendre le critère d'irréductibilité d'Eisenstein (*Exercice* 11) aux anneaux factoriels.

¶ 1. Soit G_n l'ensemble des nombres complexes z tels que $z^n = 1$.

a) Montrer que G_n est un sous-groupe d'ordre n du groupe multiplicatif $\mathbf{C}^*$ des nombres complexes non nuls.

b) Soit

$$z = \cos(2k\pi/n) + i.\sin(2k\pi/n)$$

un élément de G_n; pour que z soit un générateur du groupe G_n (i.e. pour que toute racine n^e de l'unité soit une puissance de z) il faut et il suffit que k soit premier à n (on dit alors que z est une **racine primitive** n^e de l'unité).

c) Sans supposer k et n premiers entre eux, montrer que l'ordre de z dans le groupe G_n (i.e. le plus petit entier $d \geqslant 1$ tel que $z^d = 1$) est $n/\mathrm{pgcd}(k, n)$, et qu'alors z est racine primitive d^e de l'unité.

d) Soit $\varphi(n)$ le nombre des racines primitives n^e de l'unité. Montrer que $\varphi(n)$ est le nombre d'entiers k tels que $1 \leqslant k \leqslant n$ qui sont premiers à n, et que c'est aussi le nombre des éléments inversibles de l'anneau $\mathbf{Z}/n\mathbf{Z}$. Montrer que

$$n = \sum_{d|n} \varphi(d),$$

la somme étant étendue à tous les diviseurs d et n (la notation $d|n$ signifie que d divise n). Montrer que

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_r}\right)$$

où $p_1, \ldots, p_r$ sont les divers diviseurs premiers de n.

e) Classer d'après leur ordre les racines n^e de l'unité pour $n = 2, 3, 4, 6, 8, 12, 16, 20, 24$. Calculer les parties réelles et imaginaires de toutes les racines 24^e de l'unité.

¶ 2. Soient K un corps commutatif et n un entier tel que l'équation $x^n = 1$ possède n racines dans K. Montrer que le sous-groupe d'ordre n du groupe multiplicatif K^* formé par les racines de cette équation est *cyclique* (utiliser l'*Exercice* 20 du § 7).

En déduire que si K est un corps fini à q éléments le groupe multiplicatif K^* est cyclique (considérer l'équation

$$x^{q-1} = 1$$

dans K).

En particulier, pour tout nombre premier p, le groupe multiplicatif $(\mathbf{Z}/p\mathbf{Z})^*$ est cyclique.

¶ 3. Soit K un corps commutatif fini à q éléments. Soient $a_1, \ldots, a_{q-1}$ les éléments non nuls de K. Montrer que si X désigne une indéterminée sur K on a

$$(X - a_1) \ldots (X - a_{q-1}) = X^{q-1} - 1$$

(utiliser le Théorème 1 du § 33). En déduire que

$$a_1 \ldots a_{q-1} = -1$$

(on utilisera, pour déterminer le signe, le fait que q est une puissance de la caractéristique p de K : § 30, *Exercice* 8).
En prenant $K = \mathbf{Z}/p\mathbf{Z}$, déduire de là le théorème de Wilson, à savoir que

$$(p-1)! \equiv -1 \pmod{p}$$

pour tout nombre premier p.

¶ 4. Soient p un nombre premier et r un entier ; on dit qu'un entier n premier à p est une puissance r^e **modulo** p (si $r = 2$, on dit un **reste quadratique** modulo p) s'il existe des entiers x tels que l'on ait

$$x^r \equiv n \pmod{p}.$$

En utilisant le fait que le groupe multiplicatif $(\mathbf{Z}/p\mathbf{Z})^*$ est cyclique (*Exercice* 2) montrer que tout x premier à p est une puissance r^e modulo p si r est premier à $p-1$. Si r divise $p-1$ (exemple : $r = 2$ et p impair, cas le plus important), pour qu'un entier n premier à p soit puissance r^e modulo p il faut et il suffit que

$$n^{\frac{p-1}{r}} \equiv 1 \pmod{p}.$$

Les classes modulo p des puissances r^e mod p sont alors en nombre égal à $\frac{p-1}{r}$ (par exemple si p est impair il y a $\frac{p-1}{2}$ restes quadratiques modulo p).
On prend $p = 31$. Pour chaque diviseur r de $p - 1 = 30$, trouver les puissances r^e modulo p.

¶ 5. (Cet Exercice repose sur l'*Exercice* 1). Pour tout entier $n \geqslant 1$, on appelle **polynôme cyclotomique** d'indice n le polynôme

$$\Phi_n(X) = (X - \xi_1) \ldots (X - \xi_h)$$

dont les racines sont les $h = \varphi(n)$ racines *primitives* n^e de l'unité dans le corps $\mathbf{C}$; ce polynôme est, en apparence, à coefficients dans $\mathbf{C}$, mais on va montrer qu'en fait il est à coefficients *entiers rationnels*. On convient de poser $\Phi_1(X) = X - 1$.

a) Montrer que

$$\Phi_p(X) = 1 + X + X^2 + \cdots + X^{p-1}$$

si p est premier.

b) Vérifier que

$$\Phi_{12}(X) = X^4 - X^2 + 1.$$

c) Montrer que, pour tout entier $n \geqslant 1$, on a

$$X^n - 1 = \prod_{d|n} \Phi_d(X)$$

où le produit figurant au second membre est étendu à tous les diviseurs d de n (y compris 1 et n). (Utiliser la décomposition du polynôme $X^n - 1$ en produit de facteurs du premier degré).

d) En utilisant la relation

$$\Phi_n(X) = \frac{X^n - 1}{\prod_{\substack{d \mid n \\ d \neq n}} \Phi_d(X)}$$

et en raisonnant par récurrence sur n, montrer que Φ_n est à coefficients entiers rationnels. (Ces résultats s'étendent à tout corps algébriquement clos K, pourvu qu'on se limite aux entiers n qui ne sont pas divisibles par la caractéristique p de K, restriction qui n'en est d'ailleurs pas une en vertu du § 30, *Exercice* 14).

¶¶ **6.** Montrer qu'il existe, sur l'ensemble des entiers $n \geqslant 1$, une et une seule fonction μ (**fonction de Möbius**) à valeurs entières, vérifiant la relation

$$\sum_{d \mid n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n > 1 \end{cases}$$

(la somme est étendue aux diviseurs d de n tels que $1 \leqslant d \leqslant n$).
Montrer qu'on a

$$\mu(1) = 1$$
$$\mu(p) = -1 \quad \text{si } p \text{ est premier}$$
$$\mu(p^r) = 0 \quad \text{si } p \text{ est premier et si } r \geqslant 2.$$

Montrer qu'on a

$$(*) \qquad \mu(mn) = \mu(m)\,\mu(n) \quad \text{si } m \text{ et } n \text{ sont premiers entre eux}$$

(on observera que tout diviseur de mn, lorsque m et n sont premiers entre eux, s'écrit d'une façon et d'une seule comme produit d'un diviseur de m et d'un diviseur de n; on raisonnera alors par récurrence en supposant (*) déjà établi pour les couples m', n' tels que $m'n' < mn$).
Déduire des résultats précédents que l'on a

$$\mu(n) = \begin{cases} 1 & \text{si } n = 1 \\ (-1)^r & \text{si } n \text{ est produit de } r \text{ facteurs premiers distincts} \\ 0 & \text{si } n \text{ est divisible par le carré d'un nombre premier.} \end{cases}$$

Calculer $\mu(n)$ pour $1 \leqslant n \leqslant 100$.

¶¶ **7.** Soit f une fonction sur l'ensemble des entiers $n \geqslant 1$, et à valeurs dans un groupe additif A. On définit une nouvelle fonction g en posant

$$g(n) = \sum_{d \mid n} f(d)$$

où la somme est étendue aux diviseurs d de n tels que $1 \leqslant d \leqslant n$. Montrer qu'on a inversement

$$f(n) = \sum_{d \mid n} g(d)\,\mu\left(\frac{n}{d}\right)$$

où μ est la fonction de Möbius de l'*Exercice* précédent. (On utilisera exclusivement la propriété ayant servi à définir μ).

Comment modifier les formules précédentes si A est un groupe commutatif écrit multiplicativement ?

¶¶ 8. (Cet Exercice repose sur les *Exercices* 5, 6 et 7). Montrer que les polynômes cyclotomiques sont donnés par la relation

$$\Phi_n(X) = \prod_{d|n} (X^d - 1)^{\mu\left(\frac{n}{d}\right)}.$$

Calculer les polynômes Φ_n pour $n = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15$, et montrer que

$$\begin{aligned}\Phi_{105}(X) = {} & X^{48} + X^{47} + X^{46} - X^{43} - X^{42} - 2X^{41} - X^{40} - X^{39} + X^{35} + X^{36} + X^{34} + X^{33} \\ & + X^{32} + X^{31} - X^{28} - X^{26} - X^{24} - X^{22} - X^{20} + X^{17} + X^{16} + X^{15} + X^{14} + X^{13} \\ & + X^{12} - X^9 - X^8 - 2X^7 - X^6 - X^5 + X^2 + X + 1,\end{aligned}$$

en supposant que Faddeev et Sominskii aient raison, ce que l'auteur n'a pas vérifié...

9. Décomposer en éléments simples complexes la fraction rationnelle

$$\frac{f(X)}{X^n - 1}$$

où f est un polynôme quelconque à coefficients complexes.

¶ 10. Soient $z_1, \ldots, z_n$ les racines n^e de l'unité dans $\mathbf{C}$. Montrer qu'on a

$$z_1^h + \ldots + z_n^h = \begin{cases} n & \text{si } h \equiv 0 \pmod{n} \\ 0 & \text{si } h \not\equiv 0 \pmod{n} \end{cases}$$

(multiplier le premier membre par z_1^h).

11. On désigne par $z_1, \ldots, z_n$ les racines n^e de l'unité dans $\mathbf{C}$. Démontrer les relations suivantes :

$$\prod_{k=1}^{k=n} (a + bz_k) = a^n + (-1)^{n-1} b^n$$

$$\prod_{k=1}^{k=n} (z_k^2 - 2z_k \cos\theta + 1) = 2(1 - \cos n\theta)$$

$$\prod_{k=1}^{k=n} \frac{(t + z_k)^n - 1}{t} = \prod_{k=1}^{k=n-1} [t^n - (z_k - 1)^n].$$

12. Soient u, v, w les trois racines (distinctes ou non) d'une équation

$$ax^3 + bx^2 + cx + d = 0$$

de degré 3 à coefficients complexes. En utilisant les résultats du § 33, n° 6, *Exemple* 4, calculer à l'aide de a, b, c, d les expressions

$$u^3 + v^3 + w^3, \qquad u^6 + v^6 + w^6,$$

¶ **13.** Soient $X_1, \ldots, X_n$ des indéterminées sur un anneau commutatif K. On appelle **fonctions symétriques élémentaires** de $X_1, \ldots, X_n$ les polynômes

$$s_1 = X_1 + X_2 + \cdots + X_n = \sum X_i$$

$$s_2 = X_1X_2 + \cdots + X_{n-1}X_n = \sum_{1 \leqslant i < j \leqslant n} X_iX_j$$

$$s_3 = X_1X_2X_3 + \cdots = \sum_{1 \leqslant i < j < k \leqslant n} X_iX_jX_k$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$s_n = X_1X_2 \ldots X_n$$

(ces expressions interviennent, cf. n° 6 du § 33, dans le calcul des coefficients d'une équation algébrique en fonction de ses racines). D'autre part, on dit qu'un polynôme $f \in K[X_1, \ldots, X_n]$ est **symétrique** si l'on a

$$f(X_{s(1)}, \ldots, X_{s(n)}) = f(X_1, \ldots, X_n)$$

pour toute permutation s des entiers $1, \ldots, n$.

a) Démontrer que *tout polynôme symétrique* $f(X_1, \ldots, X_n)$ *est un polynôme en* $s_1, \ldots, s_n$ *à coefficients dans* K. [On pourra procéder par récurrence sur l'entier $n + d^\circ(f)$. Observer d'abord que $f(X_1, \ldots, X_{n-1}, 0)$ est symétrique en $X_1, \ldots, X_{n-1}$, donc est un polynôme en les fonctions symétriques élémentaires de $X_1, \ldots, X_{n-1}$, lesquelles s'obtiennent en remplaçant X_n par 0 dans les fonctions symétriques élémentaires de $X_1, \ldots, X_n$. En déduire qu'il existe un polynôme $p(s_1, \ldots, s_{n-1})$, de degré au plus égal au degré de f, tel que le polynôme

$$g(X_1, \ldots, X_n) = f(X_1, \ldots, X_n) - p(s_1, \ldots, s_{n-1})$$

s'annule pour $X_n = 0$; en déduire, compte-tenu de la symétrie de g, que

$$g(X_1, \ldots, X_n) = X_1 \ldots X_n h(X_1, \ldots, X_n)$$

avec h symétrique et $d^\circ(h) < d^\circ(f)$].

b) Démontrer que si un polynôme $p \in K[X_1, \ldots, X_n]$ vérifie $p(s_1, \ldots, s_n) = 0$, alors $p = 0$ (raisonner par récurence sur n; prendre p de plus petit degré possible par rapport à s_n et faire $X_n = 0$ dans le résultat).

c) En conclure que *l'expression d'un polynôme symétrique* f *à l'aide de* $s_1, \ldots, s_n$ *est unique.*

d) On suppose f homogène de degré total k en $X_1, \ldots, X_n$, et on pose

$$f(X_1, \ldots, X_n) = p(s_1, \ldots, s_n);$$

montrer que les seuls monômes

$$s_1^{r_1} \ldots s_n^{r_n}$$

figurant effectivement dans p sont ceux pour lesquels on a

$$r_1 + 2r_2 + \cdots + nr_n = k.$$

e) Calculer à l'aide des fonctions symétriques élémentaires les polynômes suivants :

$$X_1^2X_2 + X_1X_2^2 + X_1^2X_3 + X_1X_3^2 + X_2^2X_3 + X_2X_3^2 \quad (n = 3)$$

$$(2X_1 - X_2 - X_3)(2X_2 - X_1 - X_3)(2X_3 - X_1 - X_2) \quad (n = 3)$$

$$(X_1X_2 + X_3X_4)(X_1X_3 + X_2X_4)(X_1X_4 + X_2X_3) \quad (n = 4)$$

$$\sum_{i \neq j} X_i^2X_j; \quad \sum_{i \neq j \neq k} X_i^2X_jX_k; \quad \sum_{i \neq j \neq k} X_iX_jX_k; \quad \sum_{s \in \mathfrak{S}_n} (a_1X_{s(1)} + a_2X_{s(2)} + \cdots + a_nX_{s(n)})^k;$$

$$\sum_{\substack{i \neq k \\ j \neq i,\, j \neq k}} (X_i + X_k - X_j)^n \quad (n \text{ quelconque}).$$

¶ **14.** Les notations étant celles de l'*Exercice* précédent, on considère les **sommes de Newton**

$$\sigma_k(X_1, \ldots, X_n) = X_1^k + X_2^k + \cdots + X_n^k \quad (k = 0, 1, \ldots)$$

Montrer qu'on peut les exprimer en fonction de $s_1, \ldots, s_n$ à l'aide des formules suivantes :

$$\sigma_k - s_1\sigma_{k-1} + s_2\sigma_{k-2} - \cdots + (-1)^{k-1}s_{k-1}\sigma_1 + (-1)^k s_k\sigma_0 = 0 \quad \text{pour } k < n$$
$$\sigma_k - s_1\sigma_{k-1} + \cdots + (-1)^n s_n\sigma_{k-n} = 0 \quad \text{pour } k \geqslant n.$$

Calculer complètement, à l'aide des fonctions symétriques élémentaires, les sommes de Newton σ_k pour $0 \leqslant k \leqslant 6$.

¶ **15.** Soit

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

une équation algébrique à coefficients dans un corps commutatif K. Soient $x_1, \ldots, x_n$ ses racines dans une extension algébriquement close de K (on pourra prendre C en supposant que K est un sous-corps de C, mais bien entendu cette hypothèse ne simplifie rien !). Soit f un polynôme symétrique à n indéterminées, à coefficients dans K. Montrer qu'il existe un polynôme p à n indéterminées, à coefficients dans K, tel que l'on ait

$$f(x_1, \ldots, x_n) = p(a_{n-1}, \ldots, a_0)$$

et que p ne dépend que de f (utiliser l'*Exercice* 13). Applications :

a) Calculer la somme des puissances 5^e des racines de l'équation

$$x^6 - 4x^5 + 3x^3 - 4x^2 + x + 1 = 0.$$

b) Calculer la somme

$$\sum x_i^2 x_j^2 x_k x_h$$

où $x_1, \ldots, x_5$ désignent les racines de l'équation

$$x^5 - 4x^3 + x^2 + 3x + 1 = 0.$$

c) On désigne par x_1, x_2, x_3 les racines de l'équation

$$x^3 + ax^2 + bx + c = 0;$$

former les équations dont les racines sont les quantités suivantes :

i) $x_1 + x_2, x_2 + x_3, x_3 + x_1$;

ii) $x_1^2 - x_2x_3, x_2^2 - x_3x_1, x_3^2 - x_1x_2$

iii) $(x_1 + jx_2 + j^2x_3)^3, (x_1 + j^2x_2 + jx_3)^3$ où $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$.

¶ **16.** Soit

$$(*) \qquad x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

une équation algébrique à coefficients dans un corps commutatif K. On désigne par $x_1, \ldots, x_n$ ses racines (distinctes ou non) dans une extension algébriquement close de K. On appelle **discriminant** de l'équation donnée l'expression

$$D = \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j)^2.$$

a) Montrer qu'il existe un polynôme p à n indéterminées, à coefficients dans **K**, et indépendant de l'équation (*), tel que l'on ait

$$D = p(a_{n-1}, \ldots, a_0).$$

b) Calculer le discrimimant d'une équation de degré 2, 3 ou 4.

c) Pour que l'équation (*) possède au moins une racine double, il faut et il suffit que son discriminant soit nul.

d) Déterminer les valeurs de λ pour lesquelles les équations suivantes possèdent au moins une racine double :

$$x^3 - 3x + \lambda = 0; \qquad x^3 - 8x^2 + (13 - \lambda)x - 6 - 2\lambda = 0;$$
$$x^4 - 4x^3 + (2 - \lambda)x^2 + 2x - 2 = 0.$$

¶¶ *e*) Montrer que le discriminant de l'équation

$$x^n + px + q = 0$$

est égal à

$$(-1)^{\frac{n(n-1)}{2}} n^n q^{n-1} + (-1)^{\frac{(n-1)(n-2)}{2}} (n-1)^{n-1} p^n.$$

f) On désigne par f le polynôme figurant au premier membre de l'équation (*). Montrer que le discriminant D de l'équation (*) est encore donné par la formule

$$(-1)^{\frac{n(n-1)}{2}} D = \prod_{1 \leqslant i \leqslant n} f'(x_i).$$

g) Montrer que le discriminant de l'équation

$$\frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + 1 = 0$$

est égal à

$$(-1)^{\frac{n(n-1)}{2}}.$$

¶¶ *h*) On considère (*Exercice* 5) l'équation cyclotomique

$$\Phi_n(x) = 0.$$

Montrer que son discriminant est égal à

$$(-1)^{\frac{\varphi(n)}{2}} \frac{n^{\varphi(n)}}{\prod_{p|n} p^{\frac{\varphi(n)}{p-1}}}$$

(utiliser la question (*f*) et l'*Exercice* 8).

¶ 20. Soient K un corps commutatif et f un polynôme à une indéterminée, à coefficients dans **K**. On se propose de prouver qu'il existe un corps **L**, extension de K (i.e. dont K est un sous-corps), dans lequel f admet au moins une racine (ce résultat est le premier pas dans la démonstration du théorème de Steinitz).

a) Montrer qu'on peut se ramener au cas où f est irréductible sur K.

b) Dans l'anneau de polynôme $A = K[X]$, on considère l'idéal $I = (f)$ engendré par f et on forme l'anneau quotient $L = A/I$ (§ 8, *Exercice* 7). Montrer que c'est un corps.

c) Soit j l'application de K dans L qui, à chaque $c \in K$, associe la classe mod I du polynôme constant c. Montrer que c'est un isomorphisme de K sur un sous-corps de L (dans ce qui suit, on convient d'identifier chaque $c \in K$ à son image $j(c)$ dans L).

d) Soit $z \in L$ l'image du polynôme X par l'application canonique de K[X] sur L. Montrer que z est racine de f, et que $L = K[z]$. (Ce qui démontre le résultat annoncé).

e) On prend $f(X) = X^2 - d$ où $d \in K$ n'est pas un carré dans K. Montrer que les constructions précédentes se réduisent alors à celles du § 9. On vérifiera en particulier que le corps C est le quotient de l'anneau R[X] par l'idéal engendré par le polynôme $X^2 + 1$ (cette méthode de construction de C est due à Cauchy).

f) On prend $f(X) = X^3 + pX + q$, supposé irréductible sur K. Donner du corps L correspondant une description analogue à celle qu'on a donnée au § 9 pour les anneaux $K[\sqrt{d}]$.

g) Que se passe-t-il, dans les constructions précédentes, si le polynôme f n'est pas irréductible?

¶¶ **21.** Soient $f_1, \ldots, f_n$ des polynômes non constants à une indéterminée à coefficients dans un corps commutatif K. Montrer qu'il existe dans l'anneau $K[X_1, \ldots, X_n]$ un idéal maximal qui contient $f_1(X_1), \ldots, f_n(X_n)$ (utiliser l'*Exercice* 15 du § 27). En raisonnant comme dans l'*Exercice* précédent, en déduire l'existence d'une extension algébrique L de K dans laquelle chaque f_i possède au moins une racine.
(La démonstration complète du théorème de Steinitz est une extension directe du raisonnement de cet *Exercice*; on introduit un anneau de polynômes à une infinité de variables, comportant autant (sic) de variables qu'il y a de polynômes irréductibles à coefficients dans K et de coefficient dominant égal à 1, puis on prend le quotient de cet anneau par un idéal maximal bien choisi)

¶¶ **22.** Les notations restant celles de l'*Exercice* précédent, montrer que tout idéal premier de l'anneau $K[X_1, \ldots, X_n]$ contenant $f_1(X_1), \ldots, f_n(X_n)$ est maximal (cf. § 26, *Exercice* 3).

¶ **23.** Soit K un corps commutatif. On dit qu'une extension L de K (i.e. un corps commutatif admettant K pour sous-corps) est **algébrique** si tout $x \in L$ est algébrique sur K. Montrer que, pour que K soit algébriquement clos, il faut et il suffit qu'on ait $L = K$ pour toute extension algébrique de K.

24. Un corps algébriquement clos possède toujours une infinité d'éléments.

¶ **25.** (Démonstration du théorème de d'Alembert-Gauss). Cet *Exercice* suppose connues les propriétés des fonctions continues dans le plan (en particulier et tout spécialement le fait qu'une fonction continue positive sur un ensemble compact y atteint effectivement son minimum). On désigne par

$$f(z) = a_0 + \cdots + a_n z^n$$

un polynôme non constant à coefficients complexes; on suppose $a_n \neq 0$.

a) Montrer que le rapport

$$f(z)/a_n z^n$$

tend vers 1 quand $|z|$ augmente indéfiniment, i.e. que pour tout $\varepsilon > 0$ il existe $r > 0$ tel que

$$|z| > r \qquad \text{implique} \qquad \left| 1 - \frac{f(z)}{a_n z^n} \right| < \varepsilon.$$

b) Soit

$$m = \inf_{z \in \mathbb{C}} |f(z)|;$$

montrer qu'il existe un nombre $r' > 0$ tel que

$$|z| \geqslant r' \qquad \text{implique} \qquad |f(z)| \geqslant m + 1.$$

En appliquant le théorème du minimum à la fonction continue $|f(z)|$ sur l'ensemble compact $|z| \leqslant r'$, montrer qu'il existe un $z_0 \in \mathbf{C}$ tel que

$$|f(z_0)| = m.$$

[Si le théorème de d'Alembert est vrai, il est clair que $m = 0$; pour montrer que le théorème en question est vrai, il est donc nécessaire, et bien entendu suffisant, de montrer que $m = 0$. C'est le but de la question suivante.]

c) On suppose $m \neq 0$; en remplaçant z par $z - z_0$ et f par $f/f(z_0)$ on se ramène au cas où l'on a

$$f(0) = 1, \qquad |f(z)| \geqslant 1 \qquad \text{pour tout} \quad z \in \mathbf{C}.$$

Soit

$$f(z) = 1 + b_q z^q + b_{q+1} z^{q+1} + \cdots + b_n z^n \qquad \text{avec} \qquad b_q \neq 0;$$

montrer qu'il existe un nombre $M > 0$ tel que

$$|z| \leqslant 1 \qquad \text{implique} \qquad |f(z) - 1 - b_q z^q| \leqslant M.|z|^{q+1};$$

déduire de là qu'on a $|f(z)| < 1$ (contradiction avec l'hypothèse, ce qui achèvera la démonstration) pourvu que z soit choisi de telle sorte qu'on ait

$$|z| \leqslant 1, \qquad |z| < |b_q|/M, \qquad \operatorname{Arg}(b_q) + q.\operatorname{Arg}(z) = \pi.$$

¶¶ **26.** Soient E un corps commutatif algébriquement clos, L un corps commutatif quelconque, et σ un isomorphisme de L sur un sous-corps L' de E. On considère une extension M de L et on suppose $M = L[z]$ où z est algébrique sur L. Soit

$$f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

le polynôme minimal de z sur L; on note

$$f^\sigma(X) = X^n + \sigma(a_{n-1})X^{n-1} + \cdots + \sigma(a_0)$$

le polynôme, à coefficients dans L', qui s'en déduit par σ, et on considère une racine z' de f^σ dans E; soit $M' = L'[z']$. Montrer qu'il existe un et un seul isomorphisme σ' de M sur M' qui coïncide avec σ sur L, et applique z sur z'.

Déduire de là le résultat suivant : soient E une extension algébriquement close d'un corps commutatif K, et L une extension de degré fini de K. Alors il existe un isomorphisme j de L sur un sous-corps de E, tel que $j(x) = x$ pour tout $x \in K$. (Raisonner par récurence sur le degré de L sur K. On peut en fait démontrer que le résultat subsiste pour toute extension *algébrique*, de degré fini ou non, de K).

¶¶ **27.** Soient K un corps commutatif, E une extension algébriquement close de K, et L une extension de degré fini de K; on suppose L *séparable sur* K (§ 26, *Exercice* 4, *h*)) et on pose $n = [L : K]$. Montrer que le nombre d'isomorphismes j de L dans E, tels que $j(x) = x$ pour tout $x \in K$, est exactement n (raisonner comme dans l'*Exercice* 26 en utilisant l'*Exercice* 11 du § 32).

On désigne par $j_1, \ldots, j_n$ les isomorphismes en question. Pour $k \neq h$, on note L_{kh} l'ensemble des $z \in L$ tels que $j_k(z) = j_h(z)$; montrer que c'est un sous-corps de L contenant K, et distinct de L.

On suppose K *infini*; montrer que la réunion des L_{kh} n'est pas L tout entier et qu'il existe un $z \in L$ tel que les n éléments $j_k(z)$ soient deux à deux distincts. En déduire que si L est une extension séparable de degré fini d'un corps K infini, il existe un $z \in L$ tel que $L = K[z]$ (**théorème de l'élément primitif**, démontré d'abord par Dedekind pour les corps de nombres algébriques, i.e. pour $K = \mathbf{Q}$; en fait, il est encore valable pour K fini vu l'*Exercice* 2 ci-dessus).

¶ 28. On considère l'extension

$$L = \mathbf{Q}[\sqrt{3},\ \sqrt[3]{2}]$$

de $\mathbf{Q}$. Construire un nombre algébrique z tel que $L = \mathbf{Q}[z]$.

¶¶ 29. Soient K un corps commutatif, L une extension séparable et de degré fini n de K, E une extension algébriquement close de K, et $j_1, \ldots, j_n$ les n isomorphismes de L dans E tels que $j_k(x) = x$ pour tout $x \in K$ (*Exercice* 27). On se propose de montrer que

$$\mathrm{Tr}_{L/K}(z) = j_1(z) + \cdots + j_n(z)$$
$$\mathrm{N}_{L/K}(z) = j_1(z) \ldots j_n(z)$$

pour tout $z \in L$.

a) Soit $(a_i)_{1 \leq i \leq n}$ une base de L sur K. Montrer que la matrice

$$A = (j_k(a_h))_{1 \leq k,\, h \leq n}$$

(à coefficients dans E) est inversible (utiliser l'*Exercice* 16 des §§ 10, 11 ainsi que la caractérisation des systèmes de Cramer).

b) Pour tout $z \in L$ on pose

$$z a_i = \sum_j \xi_{ij} a_j$$

avec des $\xi_{ij} \in K$; on introduit les matrices

$$U_z = (\xi_{ij})_{1 \leq i,\, j \leq n}$$

et

$$D_z = \begin{pmatrix} j_1(z) & 0 & \ldots & 0 \\ 0 & j_2(z) & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & j_n(z) \end{pmatrix};$$

montrer qu'on a la relation

$$U_z = A D_z A^{-1}.$$

c) Achever la démonstration en observant que $\mathrm{Tr}_{L/K}(z) = \mathrm{Tr}(U_z)$ et que $\mathrm{N}_{L/K}(z) = \det(U_z)$ (cf. § 26, *Exercice* 4).

[Le lecteur notera que ce raisonnement montre aussi que les valeurs propres dans E de l'endomorphisme

$$u_z : x \to zx$$

de L, où l'on regarde L comme un espace vectoriel de dimension n sur K, sont précisément les $j_k(z)$, $1 \leq k \leq n$].

30. Démontrer que le polynôme

$$X^{n_1} + X^{n_2} + \ldots + X^{n_k}$$

où

$$n_r \equiv r - 1 \pmod{k},$$

est divisible par le polynôme

$$1 + X + X^2 + \cdots + X^{q-1}.$$

¶ **31.** Pour quelles valeurs de r le polynôme $\Phi_n(X^r)$ est-il divisible par le polynôme $\Phi_n(X)$?

32. Soit f un polynôme à une indéterminée à coefficients dans un corps commutatif. Si $f(X^n)$ est divisible par $X - 1$, alors $f(X^n)$ est divisible par $X^n - 1$.

¶¶ **33.** (Démonstration du Nullstellensatz de Hilbert).

a) Soient L un corps commutatif et A un sous-anneau de L; on suppose qu'il existe des éléments $y_1, \ldots, y_n$ de L, en nombre fini, tels que

$$L = A[y_1, \ldots, y_n],$$

et de plus que chaque y_j vérifie une relation algébrique non triviale à coefficients dans A. Montrer qu'il existe un élément $b \neq 0$ de A tel que

$$K = A[b^{-1}]$$

(choisir b de telle sorte que L soit un $A[b^{-1}]$-module de type fini et appliquer l'*Exercice* 24 du § 19).

b) Montrer que tout idéal premier non nul de l'anneau A contient b.

c) On suppose qu'il existe un sous-corps K de L tel que A soit le sous-anneau de L engendré par K et par un nombre fini d'éléments de L algébriquement indépendants sur K. Montrer qu'alors $A = K$ et que L est une extension algébrique de degré fini de K (observer que, dans un anneau de polynômes sur un corps, l'intersection des idéaux premiers non nuls se réduit à 0).

d) Soient K un corps commutatif et L une extension de K; on suppose qu'il existe un nombre fini d'éléments $z_1, \ldots, z_r$ de L tels que

$$L = K[z_1, \ldots, z_r];$$

montrer qu'alors L est une extension algébrique de degré fini de K, et qu'en particulier $L = K$ si K est algébriquement clos [extraire de la famille $z_1, \ldots, z_r$ des éléments algébriquement indépendants en nombre aussi grand que possible et appliquer *c*) à l'anneau A engendré par K et ces éléments].

e) Soient K un corps commutatif et $\mathfrak{m}$ un idéal maximal de l'anneau de polynômes $K[X_1, \ldots, X_r]$; montrer que l'anneau quotient $L = K[X_1, \ldots, X_r]/\mathfrak{m}$ est un corps, extension algébrique de degré fini de K [appliquer la question *d*), et l'*Exercice* 7 du § 8]. En déduire le Nullstellensatz [Voir une autre démonstration au § 35, *Exercice* 51].

¶¶ **34.** Soit p un nombre premier et soit $k = \mathbf{Z}/p\mathbf{Z}$ le corps des entiers modulo p. Si

$$f(X) = f(X_1, \ldots, X_n)$$

est un polynôme en n variables, à coefficients dans k, on note $S(f)$ la somme des valeurs de f, autrement dit on pose

$$S(f) = \sum_{x_i \in k} f(x_1, \ldots, x_n).$$

a) On suppose que f est un monôme $X_1^{m_1} \ldots X_n^{m_n}$. Montrer que l'on a alors $S(f) = 0$,

sauf si tous les m_i sont divisibles par $p-1$ et $\geqslant 1$, auquel cas on a $S(f) = (-1)^n$. (Se ramener au cas d'une variable.)

b) Utiliser *a*) pour prouver que $S(f) = 0$ si $\deg(f) < n(p-1)$.

c) Soit $\varphi(X) = \varphi(X_1, \ldots, X_n)$ un polynôme à coefficients dans k. On pose

$$f(X) = 1 - \varphi(X)^{p-1}.$$

Montrer que l'on a

$$f(x) = 1 \quad \text{si} \quad \varphi(x) = 0, \qquad x \in k^n$$
$$f(x) = 0 \quad \text{si} \quad \varphi(x) \neq 0, \qquad x \in k^n.$$

En déduire que le nombre $N(\varphi)$ de zéros de φ dans k^n vérifie la congruence

$$N(\varphi) \equiv S(f) \pmod{p}.$$

d) On suppose que $\deg(\varphi) < n$. Déduire de *b*) et *c*) que l'on a $N(\varphi) \equiv 0 \pmod{p}$. En particulier, si φ est sans terme constant, φ a au moins un zéro distinct de $(0, \ldots, 0)$ (*théorème de Chevalley*).

e) Étendre ce qui précède au cas d'un nombre fini d'équations $\varphi_\alpha(x) = 0$, avec $\Sigma \deg \varphi_\alpha < n$. (Prendre pour f le produit des $1 - \varphi_\alpha^{p-1}$.)

§ 34

Pour chacune des matrices figurant dans les *Exercices* **1** à **14** ci-dessous, répondre aux questions suivantes : *a*) Calculer les valeurs propres (on prendra **C** pour corps de base). *b*) Pour chaque valeur propre, calculer les vecteurs propres correspondants dans $\mathbf{C}^n$ (on identifie chaque matrice carrée d'ordre n à coefficients complexes à un endomorphisme de $\mathbf{C}^n$). *c*) Trouver, s'il y a lieu, une base de $\mathbf{C}^n$ formée de vecteurs propres. *d*) Si la matrice considérée est diagonalisable sur **C**, déterminer le plus petit sous-corps de **C** sur lequel elle est diagonalisable. *e*) Si la matrice considérée n'est pas diagonalisable sur **C**, trouver une base de $\mathbf{C}^n$ par rapport à laquelle l'endomorphisme correspondant de $\mathbf{C}^n$ possède une matrice triangulaire.

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

2. $\begin{pmatrix} 7 & -12 & 6 \\ 10 & -19 & 10 \\ 12 & -24 & 13 \end{pmatrix}$

3. $\begin{pmatrix} 4 & -5 & 7 \\ 1 & -4 & 9 \\ -4 & 0 & 5 \end{pmatrix}$

4. $\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$

5. $\begin{pmatrix} 9 & -6 & -2 \\ 18 & -12 & -3 \\ 18 & -9 & -6 \end{pmatrix}$

6. $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$

7. $\begin{pmatrix} 4 & 6 & -15 \\ 3 & 4 & -12 \\ 2 & 3 & -8 \end{pmatrix}$

8. $\begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$

9. $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$

10. $\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$

11. $\begin{pmatrix} 0 & 0 & 2 & 3 \\ 0 & 0 & -2 & -3 \\ 2 & -2 & 0 & -1 \\ 3 & -3 & -1 & -3 \end{pmatrix}$

12. $\begin{pmatrix} 3 & 2 & 1 & -1 \\ 2 & 2 & 1 & -1 \\ 1 & 1 & 1 & 0 \\ -1 & -1 & 0 & 0 \end{pmatrix}$

13. $\begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 2 & 0 & 0 & \dots & 0 \\ 1 & 2 & 3 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 2 & 3 & 4 & \dots & n \end{pmatrix}$

14. $\begin{pmatrix} 0 & c & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & c & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & c \\ c & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$ (n lignes et colonnes)

montrer que

$$p(A) = f(A) + g(A), \qquad q(A) = f(A)\,g(A) = g(A)\,f(A), \qquad r(A) = f(g(A)).$$

(On s'efforcera d'éviter *tout* calcul en utilisant de façon appropriée les résultats du § 28).

b) Montrer que

$$f(UAU^{-1}) = U.f(A).U^{-1}$$

si $A \in M_n(K)$ et $U \in GL(n, K)$.

c) On suppose A triangulaire et on désigne par $t_1, \ldots, t_n$ ses termes diagonaux. Montrer que $f(A)$ est triangulaire, et que ses termes diagonaux sont $f(t_1), \ldots, f(t_n)$.

d) On suppose que K est un corps. Soient $t_1, \ldots, t_n$ les valeurs propres de $A \in M_n(K)$ (prises dans une extension algébriquement close de K, et chaque valeur propre étant répétée autant de fois que sa multiplicité dans l'équation caractéristique de A; le lecteur pourra supposer $K = C$ s'il ne s'intéresse pas au cas général).

Montrer que les valeurs propres de $f(A)$ sont $f(t_1), \ldots, f(t_n)$, et que

$$\text{dét}\,(f(A)) = f(t_1) \ldots f(t_n), \qquad \text{Tr}\,(f(A)) = f(t_1) + \cdots + f(t_n).$$

19. Soit

$$f(X) = a_1 + a_2X + \cdots + a_nX^{n-1}$$

un polynôme à coefficients complexes. Montrer que

$$\begin{vmatrix} a_1 & a_2 & a_3 & \ldots & a_n \\ a_n & a_1 & a_2 & \ldots & a_{n-1} \\ a_{n-1} & a_n & a_1 & \ldots & a_{n-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ a_2 & a_3 & a_4 & \ldots & a_1 \end{vmatrix} = f(z_1) \ldots f(z_n)$$

où $z_1, \ldots, z_n$ sont les racines n^{e} de l'unité (**déterminants circulants**; utiliser l'*Exercice* précédent).

Appliquer ce résultat au calcul des déterminants

$$\begin{vmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{vmatrix}, \qquad \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{vmatrix}.$$

20. Soit s une permutation des entiers $1, 2, \ldots, n$. On considère l'endomorphisme u_s de C^n donné par

$$u_s(e_i) = e_{s(i)} \quad (1 \leqslant i \leqslant n)$$

où $(e_i)_{1 \leqslant i \leqslant n}$ est la base canonique de C^n. Utiliser la décomposition de s en cycles (§ 7, *Exercice* 24) pour calculer les valeurs propres de u_s, et montrer que u_s est diagonalisable.

On remplace dans ce qui précède C par un corps commutatif K algébriquement clos et de caractéristique $p \neq 0$; on prend $n = p$ et pour s une permutation circulaire. Montrer que u_s n'est pas diagonalisable.

21. Soient V un espace vectoriel de dimension finie sur un corps commutatif K et F un ensemble d'endomorphismes de V; **on dit que F est trigonalisable** (ou encore que **la réduction simultanée à la forme triangulaire** est possible pour F) s'il existe une base de V par rapport à laquelle la matrice de *tout* $u \in F$ soit triangulaire.

Soit V′ un sous-espace vectoriel de V stable par F, i.e. tel que l'on ait

$$u(V') \subset V' \quad \text{pour tout } u \in F$$

(au lieu de stable on dit aussi **invariant**); tout $u \in F$ induit donc un endomorphisme u' de V′ et un endomorphisme $\bar{u}$ de l'espace vectoriel quotient V/V′ (§ 12, *Exercice*). Soient F′ l'ensemble des u' et $\bar{F}$ l'ensemble des $\bar{u}$.

On suppose F′ trigonalisable dans V′, et $\bar{F}$ trigonalisable dans V/V′. Montrer que F est trigonalisable dans V (on construira une base dans V en imitant la démonstration du Théorème 1 du § 18).

Déduire de là une variante de la démonstration du Théorème 3 du § 34.

¶ **22.** Soient V un espace vectoriel de dimension finie sur un corps commutatif K algébriquement clos (on peut supposer $K = \mathbf{C}$, la démonstration est la même...) et F un ensemble d'endomorphismes de V *commutant deux à deux*. On se propose de démontrer que F est trigonalisable (*Exercice* précédent).

a) Pour tout $u \in F$ et toute valeur propre $\lambda \in K$ de u, soit $V_u(\lambda)$ le sous-espace des $x \in V$ tels que $u(x) = \lambda x$. Montrer que $V_u(\lambda)$ est stable par F.

b) Déduire de là que les $u \in F$ ont un vecteur propre commun dans V [utiliser la question *a*) pour raisonner par récurence sur dim (V)].

c) Terminer la démonstration en utilisant l'*Exercice* précédent.

d) On suppose en outre chaque $u \in F$ diagonalisable. Montrer qu'il existe une base de V par rapport à laquelle la matrice de *tout* $u \in F$ est diagonale (« réduction simultanée à la forme diagonale » pour des endomorphismes diagonalisables qui commutent deux à deux).

e) Montrer qu'au lieu de supposer K algébriquement clos il suffit de supposer que tout $u \in F$ est trigonalisable [ou, pour la question *d*), diagonalisable].

¶ **23.** Soient V un espace vectoriel de dimension finie sur un corps K et F un ensemble d'endomorphismes de V. On dit que F est **irréductible** si les seuls sous-espaces vectoriels de V stables par F sont $\{0\}$ et V.

a) Montrer que, si un endomorphisme f de V commute à tout $u \in F$, les sous-espaces vectoriels Ker (f) et Im(f), ainsi que les sous-espaces propres de f, sont stables par F.

b) On suppose F irréductible et K algébriquement clos. Démontrer que les seuls endomorphises de V qui commutent à tout $u \in F$ sont les homothéties (lemme de Schur).

c) On suppose toujours F irréductible mais on ne fait plus d'hypothèse sur K. Montrer que les endomorphismes de V qui commutent à tout $u \in F$ forment un sous-corps (éventuellement non commutatif) de l'anneau des endomorphismes de V.

d) On prend $K = \mathbf{R}$ et $V = \mathbf{R}^4$. Choisir F de telle sorte que le sous-corps de la question précédente soit le corps des quaternions du § 15, *Exercice* 11.

¶ **24.** Soient V un espace vectoriel de dimension finie n sur un corps commutatif K de caractéristique zéro, et G un groupe *fini* d'automorphismes de V; on note r l'ordre de G.

a) Soit f un endomorphisme de V; montrer que l'endomorphisme

$$f^{\natural} = \frac{1}{r} \sum_{s \in G} s \circ f \circ s^{-1}$$

commute à tout $s \in G$, et qu'on a $f^{\natural} = f$ si et seulement si f commute aux éléments de G. Montrer qu'on a

$$(f \circ g)^{\natural} = f^{\natural} \circ g$$

si g commute aux éléments de G.

b) Soit W un sous-espace vectoriel de V invariant par G, i.e. (*Exercice* 21) tel que $s(W) \subset W$ pour tout $s \in G$ (on montrera en passant qu'on a du reste $s(W) = W$ pour tout $s \in G$). On choisit (§ 17, Corollaire du Théorème 2 combiné avec le fait que W admet un supplémentaire dans V) un endomorphisme p de V tel que

$$p^2 = p, \qquad p(V) = W.$$

Montrer que

$$\mathrm{Im}(p^q) = W.$$

c) En considérant, dans la question *b*), le noyau de p^q, démontrer le théorème suivant : tout sous-espace de V *invariant par* G admet dans V un supplémentaire *invariant par* G.

d) Soient V un espace vectoriel de dimension finie sur un corps algébriquement clos de caractéristique 0 (par exemple $\mathbf{C}$) et G un groupe commutatif fini d'automorphismes de V. Montrer qu'il existe une base de V par rapport à laquelle la matrice de *tout* $s \in G$ est diagonale; si de plus G est d'ordre n, les coefficients diagonaux de ces matrices sont des racines n^e de l'unité. [On utilisera la question *b*) de l'*Exercice* 22].

e) Soit X une matrice carrée à coefficients dans un corps algébriquement clos de caractéristique 0; si

$$X^n = 1$$

pour un entier $n \geqslant 1$, alors X est diagonalisable. Montrer à l'aide d'un exemple que ce résultat ne s'étend pas aux corps de caractéristique $p \neq 0$.

f) Montrer que le résultat de la question *c*) est encore valable en caractéristique $p \neq 0$ pourvu que l'ordre r du groupe G ne soit pas multiple de p. Même résultat pour la question *d*).

¶ 25. Soit V un espace vectoriel de dimension finie $n + 1$ sur un corps commutatif K algébriquement clos et de caractéristique 0. On considère trois endomorphismes u, v et h de V satisfaisant aux formules de commutation suivantes :

$$[h, u] = 2u, \qquad [h, v] = -2v, \qquad [u, v] = h$$

où l'on pose d'une façon générale $[f, g] = f \circ g - g \circ f$. On suppose enfin l'ensemble $\{u, v, h\}$ *irréductible*, i.e. que les seuls sous-espaces vectoriels de V stables à la fois par u, v et h sont $\{0\}$ et V.

a) Soient $x \in V$ et $\lambda \in K$ tels que $h(x) = \lambda x$. Montrer que le vecteur $y = u(x)$ vérifie $h(y) = (\lambda + 2)y$, et que le vecteur $z = v(x)$ vérifie $h(z) = (\lambda - 2)z$.

b) Montrer qu'il existe un vecteur $x \neq 0$ et un scalaire $\lambda \in K$ tels que l'on ait

$$h(x) = \lambda x, \qquad u(x) = 0.$$

c) Le vecteur x satisfaisant à la question *b*), on pose

$$x_k = v^k(x)/k! \quad (k \geqslant 0).$$

Démontrer les relations

$$\begin{aligned} h(x_k) &= (\lambda - 2k)x_k, \\ u(x_k) &= (\lambda - k + 1)x_{k-1}, \\ v(x_k) &= (k+1)x_{k+1}. \end{aligned}$$

d) En tenant compte de l'hypothèse d'irréductibilité faite au début, déduire de là que $\lambda = n$ et que les $n + 1$ vecteurs $x_0, x_1, \ldots, x_n$ forment une base de V. Quelles sont les matrices de u, v et h par rapport à cette base? Réciproque? Cas $n = 2$ ou 3?

e) On prend pour V l'espace vectoriel formé des polynômes de degré n au plus, à une indéterminée et à coefficients dans K. Montrer que la situation décrite dans les questions précé-

dentes est effectivement réalisée si l'on définit u, v et h comme suit : u transforme chaque polynôme $f(X)$ en le polynôme $nXf(X) - X^2f'(X)$, v transforme chaque polynôme $f(X)$ en le polynôme $f'(X)$, et h transforme $f(X)$ en $-nf(X) + 2Xf'(X)$; on désigne naturellement par f' le polynôme dérivé de f.

¶ **26.** Soient V un espace vectoriel de dimension finie sur un corps K algébriquement clos de caractéristique 0 (par exemple $K = \mathbb{C}$), et u, v, w trois endomorphismes de V. On suppose

$$[u, w] = [v, w] = 0, \qquad [u, v] = w.$$

Montrer qu'il existe une base de V par rapport à laquelle les matrices de u, v et w sont triangulaires. Déterminer toutes les solutions u, v, w du problème lorsque V est de dimension 3.

¶¶ **27.** Soit V un espace vectoriel de dimension finie sur un corps commutatif K. Un ensemble F d'endomorphismes de V est appelé une **algèbre de Lie** (d'endomorphismes de V) si F est un espace vectoriel (i.e. si l'on a $\alpha u + \beta v \in F$ quels que soient $u, v \in F$ et $\alpha, \beta \in K$), et si de plus on a

$$u \circ v - v \circ u \in F \quad \text{quels que soient } u, v \in F.$$

(Exemple : prendre les combinaisons linéaires de u, v, h dans l'*Exercice* 25, ou de u, v et w dans l'*Exercice* 26).

On dit qu'une algèbre de Lie F d'endomorphismes de V est **résoluble** s'il existe une suite croissante

$$(*) \qquad \{0\} = F_0 \subset F_1 \subset \ldots \subset F_n = F$$

de sous-espaces vectoriels de F tels que l'on ait

$$(**) \qquad u \circ v - v \circ u \in F_{i-1} \quad \text{quels que soient } u, v \in F_i$$

pour tout i tel que $1 \leqslant i \leqslant n$. On se propose de démontrer que *si* K *est algébriquement clos et de caractéristique* 0 (par exemple si $K = \mathbb{C}$), *et si* F *est résoluble, il existe une base de* V *par rapport à laquelle la matrice de tout* $u \in F$ *est triangulaire* [Théorème de Lie, qui généralise le résultat *c*) de l'*Exercice* 22 ainsi que l'*Exercice* 26 comme on le voit facilement].

a) Démontrer le théorème dans le cas où $n = 1$ dans la suite $(*)$.

b) Montrer que le terme F_{n-1} de $(*)$ est une algèbre de Lie résoluble.

c) On suppose trouvé un vecteur $x \neq 0$ dans V tel que l'on ait une relation de la forme

$$(***) \qquad u(x) = \lambda(u).x \quad \text{pour tout } u \in F_{n-1}$$

(autrement dit, x est un vecteur propre commun aux $u \in F_{n-1}$). On prend un $v \in F_n$, et on pose $y = v(x)$. Montrer qu'on a

$$u(y) = \mu(u).y \quad \text{pour tout } u \in F_{n-1},$$

où $\mu(u)$ est un scalaire qu'on calculera. En remplaçant v par ξv dans le résultat obtenu (où ξ est un élément arbitraire de K), en conclure que

$$\mu(u) = \lambda(u) \quad \text{pour tout } u \in F_{n-1}.$$

d) On note $V(\lambda)$ le sous-espace de V formé des $x \in V$ vérifiant $(***)$. Montrer qu'il est stable par tout $v \in F_n$, et que les restrictions à $V(\lambda)$ de deux éléments quelconques de F_n commutent

(utiliser l'*Exercice* 8 des §§ 12, 13, 14). En déduire que les $u \in F$ ont au moins un vecteur propre commun dans $V(\lambda)$.

e) Terminer la démonstration en raisonnant par récurrence sur la dimension de V (utiliser l'*Exercice* 21).

f) Où intervient l'hypothèse que le corps K est de caractéristique 0 ?

g) Montrer (avec les hypothèses indiquées sur K) que le théorème de Lie *caractérise* les algèbres de Lie *résolubles*.

¶ **28.** Pour qu'une matrice carrée X à coefficients dans un corps K algébriquement clos soit nilpotente, il faut et il suffit que toutes ses valeurs propres dans K soient nulles.
Montrer que, si K est de caractéristique 0 (par exemple si $K = C$) on peut remplacer cette condition par les relations

$$\mathrm{Tr}(X) = \mathrm{Tr}(X^2) = \cdots = \mathrm{Tr}(X^n) = 0,$$

où n est l'ordre de X.
Pour qu'une matrice U, à coefficients dans K, soit unipotente (i.e. pour que $1 - U$ soit nilpotente), il faut et il suffit que la seule valeur propre de U soit 1. Quel est le polynôme caractéristique de U ?
Montrer que, si K est de caractéristique $p \neq 0$, on peut remplacer cette condition par la suivante : il existe un entier $n \geqslant 0$ tel que

$$U^{p^n} = 1.$$

29. Soit A une matrice carrée inversible à coefficients dans un corps algébriquement clos. Montrer que les valeurs propres de l'inverse de A sont les inverses des valeurs propres de A, avec les mêmes multiplicités.

¶ **30.** Soit A une matrice carrée d'ordre n à coefficients dans un corps commutatif K. Soit f une fraction rationnelle à une indéterminée à coefficients dans K ; on dit que A est **substituable** dans f s'il existe des polynômes p et q tels que l'on ait

$$f = p/q \qquad \text{et} \qquad \det(q(A)) \neq 0.$$

Montrer qu'alors la matrice

$$f(A) = p(A) \cdot q(A)^{-1}$$

est indépendante du choix de p et q (pourvu que p et q satisfassent aux conditions énoncées). On suppose K algébriquement clos, et on note $\lambda_1, \ldots, \lambda_n$ les valeurs propres de A (comptées avec leurs ordres de multiplicité). Montrer que, pour que A soit substituable dans f, il faut et il suffit que f soit définie en chaque λ_i; les valeurs propres de $f(A)$ sont alors $f(\lambda_1), \ldots, f(\lambda_n)$. (Utiliser l'*Exercice* 18).

31. Soient V un espace vectoriel de dimension finie sur un corps commutatif K, u un endomorphisme de V, et W un sous-espace vectoriel de V stable par u. Montrer que si u est diagonalisable (resp. trigonalisable) il en est de même de la restriction de u à W.

¶ **32.** Soient u et v deux endomorphismes d'un espace vectoriel V de dimension finie sur un corps commutatif. On suppose que u et v sont diagonalisables et commutent. Montrer que $v \circ u$ est diagonalisable.

¶¶ **33.** Soit K un corps commutatif. Étant donnée une matrice $U \in M_n(K)$, on considère l'application $f_U : M_n(K) \to M_n(K)$ donnée par

$$f_U(X) = UX - XU = [U, X] \qquad \text{pour tout } X \in M_n(K),$$

et on considère f_U comme un endomorphisme de l'espace vectoriel $M_n(K)$. Montrer que, pour que f_U soit diagonalisable, il faut et il suffit que U le soit.

34. Soient K un corps commutatif et n un entier. Montrer que, comme espace vectoriel sur K, l'anneau $M_n(K)$ admet une base formée de matrices X possédant la propriété suivante : pour toute matrice *diagonale* $H \in M_n(K)$, on a $[H, X] = \alpha(H).X$ où $\alpha(H)$ est un scalaire dépendant de H, et que l'on calculera.

¶ 35. Soit V un espace vectoriel de dimension finie sur un corps commutatif K; on désigne par $T^p_q(V)$ l'espace vectoriel des tenseurs p fois covariants et q fois contravariants sur V (§ 21, *Exemple* 6). Soit u un *auto*morphisme de V; on considère l'automorphisme $T^p_q(u)$ de $T^p_q(V)$ défini au § 21, *Exercice* 1. Montrer que si u est diagonalisable, il en est de même de $T^p_q(u)$. Montrer de même que, si un endomorphisme u de V est diagonalisable, l'endomorphisme $D^p_q(u)$ de $T^p_q(u)$ défini au § 21, *Exercice* 1, est diagonalisable. Calculer, dans chacun de ces deux cas, les valeurs propres de l'endomorphisme considéré dans $T^p_q(V)$ en fonction de celles de u.

¶ 36. Les notations restant celles de l'*Exercice* précédent, on choisit une base (a_i) de V et on désigne par G le groupe des automorphismes de V dont la matrice par rapport à la base choisie est triangulaire (resp. diagonale). Construire une base de l'espace $T^p_q(V)$ par rapport à laquelle la matrice de $T^p_q(u)$ est triangulaire (resp. diagonale) pour tout $u \in G$.

¶ 37. Soient V un espace vectoriel de dimension finie sur un corps commutatif K, et $S_r(V)$ l'espace vectoriel formé par les fonctions polynomiales homogènes et de degré r sur V (§§ 27, 28, *Exercice* 17); on associe à chaque automorphisme u de V l'automorphisme u_r de $S_r(V)$ donné par $u_r(f) = f \circ u$.
Montrer que si u est diagonalisable (resp. trigonalisable) il en est de même de u_r; calculer les valeurs propres de u_r en fonction de celles de u.
En supposant u diagonalisable, calculer $\mathrm{Tr}(u_3)$ en fonction des coefficients du polynôme caractéristique de u. Le résultat obtenu s'étend-il à tout automorphisme u de V? Existe-t-il des formules analogues pour calculer $\mathrm{Tr}(u_r)$ quel que soit r?

38. Démontrer le Théorème 4 du § 34 sans utiliser le fait que les valeurs propres sont les racines d'une équation algébrique (écrire une relation linéaire non triviale entre $x_1, \ldots, x_n$, lui appliquer u, et en déduire une relation linéaire non triviale entre $n-1$ des vecteurs $x_1, \ldots, x_n$).

39. Démontrer le Théorème 4 du § 34 en utilisant un déterminant de Vandermonde (§ 24, *Exercice* 15) (écrire une relation linéaire non triviale entre $x_1, \ldots, x_n$ et lui appliquer successivement u, u^2, $\ldots$, u^{n-1}).

¶ 40. Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ une matrice carrée d'ordre n à coefficients dans un anneau K. Étant données deux parties H et K de l'ensemble $\{1, 2, \ldots, n\}$ on désigne par $A_{H,K}$ la matrice formée avec les termes a_{ij} de A pour lesquels on a $i \in H$ et $j \in K$.
Soit

$$(-1)^n f_A(X) = X^n - \tau_1(A)X^{n-1} + \tau_2(A)X^{n-2} - \ldots$$

(§ 34, n° 3, formule (9)). Démontrer que les coefficients $\tau_p(A)$ de ce polynôme sont donnés par la formule

$$\tau_p(A) = \Sigma \det(A_{H,H})$$

où la somme est étendue à toutes les parties H de $\{1, 2, \ldots, n\}$ telles que $\mathrm{Card}(H) = p$.

¶¶ **41.** Soient L un anneau commutatif et A un sous-anneau de L; un $x \in L$ est dit **entier sur** A s'il vérifie une équation de la forme

$$x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

avec $a_{n-1}, \ldots, a_0 \in A$ (la condition que le coefficient de x^n est égal à 1 est essentielle si A n'est pas un corps). Un nombre complexe est appelé un **entier algébrique** s'il est entier sur le sous-anneau **Z** de **C**.

a) Dans ce qui précède on suppose que L est de type fini comme A-module. Soit $(m_i)_{1 \leqslant i \leqslant r}$ un système fini de générateurs du A-module L. Montrer que pour tout $x \in L$ il existe des $a_{ij} \in A$ tels que

$$xm_i = \sum_{j=1}^{j=r} a_{ij} m_j \quad (1 \leqslant i \leqslant r).$$

On pose

$$\begin{vmatrix} a_{11} - x & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} - x & \cdots & a_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} - x \end{vmatrix} = d;$$

montrer que $dm_i = 0$ pour $1 \leqslant i \leqslant r$; en déduire que $d = 0$, puis que tout $x \in L$ est entier sur A.

b) L'anneau L n'étant plus supposé de type fini sur A, soient $x_1, \ldots, x_q$ des éléments de L entiers sur A. Montrer que le sous-anneau $A[x_1, \ldots, x_q]$ est un A-module de type fini.

c) Montrer que *l'ensemble* B *des* $x \in L$ *entiers sur* A *est un sous-anneau de* L. (On l'appelle la **clôture intégrale** de A dans L). Exemple : les entiers algébriques forment un sous-anneau de **C**, résultat dû à Dedekind (ainsi que le raisonnement ci-dessus).

d) Soient C un anneau commutatif, B un sous-anneau de C, et A un sous-anneau de B. On suppose tout $x \in C$ entier sur B, et tout $x \in B$ entier sur A. Montrer que tout $x \in C$ est entier sur A.

¶ **42.** Les valeurs propres d'une matrice carrée à coefficients entiers rationnels sont des entiers algébriques. Réciproquement, tout entier algébrique est valeur propre d'une matrice carrée à coefficients dans **Z**.

43. Tout nombre rationnel qui est un entier algébrique est un entier rationnel.

¶ **44.** On considère une extension quadratique

$$L = \mathbf{Q}[\sqrt{d}]$$

du corps des nombres rationnels; on suppose que d est un entier rationnel qui n'est divisible par le carré d'aucun nombre premier (on montrera en passant qu'on peut toujours ramener une extension quadratique de **Q** à être de ce type).

a) Pour qu'un élément $z = x + y\sqrt{d}$ de L soit entier sur **Z**, il faut et il suffit que

$$2x \in \mathbf{Z} \qquad \text{et} \qquad x^2 - y^2 d \in \mathbf{Z}$$

(observer que si z est un entier algébrique il en est de même de $\bar{z} = x - y\sqrt{d}$).

b) Soit B l'anneau des $z \in L$ entiers sur **Z**. Montrer que, comme groupe additif, L admet une

base formée des deux éléments

$$1,\ \sqrt{d} \qquad \text{si } d \equiv 2 \text{ ou } 3 \pmod 4$$
$$1,\ \frac{1+\sqrt{d}}{2} \qquad \text{si } d \equiv 1 \pmod 4.$$

45. Montrer que tout nombre algébrique est le quotient d'un entier algébrique par un entier rationnel non nul.

¶¶ **46.** On dit qu'un anneau d'intégrité commutatif A est **intégralement clos** si tout élément du corps des fractions K de A qui est entier sur A appartient à A; exemple : l'anneau $\mathbf{Z}$ (*Exercice* 43).

a) On suppose que A est un anneau de valuation (i.e. qu'on a $x \in A$ ou $x^{-1} \in A$ pour tout $x \in K$, cf. § 8, *Exercice* 6). Montrer que A est intégralement clos.

b) Si A est intersection d'anneaux de valuation de son corps des fractions K, alors A est intégralement clos [NB — On peut démontrer la réciproque].

c) Tout anneau factoriel (§ 31, *Exercice* 21) est intégralement clos, de même que tout anneau de Dedekind (§§ 10, 11, *Exercice* 14 et § 18, *Exercice* 7).

d) Si A est intégralement clos et si $\mathfrak{p}$ est un idéal premier de A, l'anneau local $A_{\mathfrak{p}}$ [§ 29, *Exercice* 9, *e*) : $A_{\mathfrak{p}}$ est l'ensemble des $x \in K$ qui peuvent s'écrire sous la forme u/v avec $u, v \in A$ et $v \notin \mathfrak{p}$] est intégralement clos.

e) Soient A un anneau d'intégrité commutatif, K son corps des fractions et L une extension de K; alors la clôture intégrale de A dans L est un anneau intégralement clos.

¶¶ **47.** Soient A un anneau intégralement clos et K son corps des fractions.

a) Soient E une extension algébriquement close de K et x un élément de E entier sur A (donc algébrique sur K). Soit f le polynôme minimal (§ 32, *Exercices* 9 et 10) de x sur K. Montrer que toutes les racines de f dans E sont des entiers sur A. En conclure que les coefficients de f appartiennent à A. (Pour vérifier qu'un élément x algébrique sur K est entier sur A, il suffit donc d'examiner son équation minimale sur K).

b) Soit L une extension de degré fini de K. Montrer qu'on a

$$\mathrm{Tr}_{L/K}(x) \in A, \qquad \mathrm{N}_{L/K}(x) \in A$$

pour tout $x \in L$ entier sur A (*Exercices* 4 et 5 du § 26).

c) On suppose, dans la question *b*), que L est extension *séparable* de K (§ 26, *Exercice* 4; on rappelle que cette condition est toujours vérifiée en caractéristique nulle). Soient $(u_i)_{1 \leqslant i \leqslant n}$ une base de L formée d'éléments entiers sur A (on montrera qu'il existe de telles bases) et $(v_i)_{1 \leqslant i \leqslant n}$ la base complémentaire (§ 26, *Exercice* 4); soit B la clôture intégrale de A dans L. Montrer que les composantes par rapport à la base (v_i) de tout $x \in B$ sont dans A. En déduire que *le* A*-module* B *est de type fini si* A *est noethérien, et isomorphe à* A^n *si* A *est principal.*

¶¶ **48.** Soient L un corps de nombres algébriques (§ 26, *Exercice* 4) et B l'anneau des $x \in L$ entiers sur $\mathbf{Z}$ (on dit habituellement que B est l'**anneau des entiers de** L).

a) Montrer que le groupe additif B admet une base à n éléments, où $n = [L : \mathbf{Q}]$ (autrement dit, qu'il existe une base de L sur $\mathbf{Q}$ qui est en même temps une base du $\mathbf{Z}$-module B).

b) Montrer que, pour tout idéal I de B, la relation $I \neq \{0\}$ implique $I \cap \mathbf{Z} \neq \{0\}$ (prendre un $x \in I$ et examiner le premier coefficient non nul de son équation minimale sur $\mathbf{Q}$). En déduire que l'anneau quotient B/I est fini pour tout idéal non nul I de B.

c) Montrer que tout idéal premier $\mathfrak{p}$ non nul de B est maximal, que $B/\mathfrak{p}$ est un corps fini, et que $\mathfrak{p} \cap \mathbf{Z} = p\mathbf{Z}$ où p est la caractéristique de $B/\mathfrak{p}$.

[Ces résultats classiques sont dus à Dedekind, et ceux de l'*Exercice* 47 en sont des généralisations faciles. Le fait que les corps B/$\mathfrak{p}$ soient finis explique en partie l'importance d'une étude générale des corps finis, d'autant plus que *tout* corps fini peut s'obtenir de cette façon. L'un des résultats fondamentaux de Dedekind, que Gauss et Kummer avaient cherché à obtenir avant lui, est que l'anneau des entiers d'un corps de nombres algébriques est un anneau de Dedekind (d'où la terminologie...), autrement dit que *tout idéal de l'anneau* B *s'écrit, d'une façon unique à des permutations près, comme produit d'idéaux premiers.* Les *Exercices* 49 et 50 ont pour but de donner une démonstration de ce fait. On utilisera uniquement le fait que B est un anneau *noethérien* [évident d'après la question *a*) de l'*Exercice* 48], *intégralement clos* [évident d'après la question *e*) de l'*Exercice* 46], *dont tout idéal premier non nul est maximal.*]

49. Soient A un anneau d'intégrité commutatif et K son corps des fractions. On utilise dans ce qui suit la notion d'idéal fractionnaire de A définie au § 10, *Exercice* 14.

a) On dit qu'un idéal fractionnaire I de A est **divisoriel** s'il est l'intersection des idéaux fractionnaires principaux (i.e. de la forme Ax, avec $x \in$ K, $x \neq 0$) qui le contiennent. Montrer que si I et J sont divisoriels il en est de même de (I : J).

Montrer que tout $x \in$ (I : I) est entier sur A si A est noethérien, et en déduire que

$$(\mathrm{I} : \mathrm{I}) = \mathrm{A} \quad \text{si A est noethérien et intégralement clos.}$$

b) On suppose A noethérien. Montrer qu'il existe au moins un idéal premier non nul de A qui est divisoriel (considérer l'ensemble des idéaux divisoriels I tels que I $\subset$ A, I $\neq$ A, et en prendre un élément maximal).

c) On suppose dorénavant que A est un anneau *local* (*), *noethérien, intégralement clos, et que le seul idéal premier non nul de* A *est l'unique idéal maximal* $\mathfrak{p}$ *de* A; un anneau de valuation discrète (§ 8, *Exercice* 6) vérifie ces conditions; on se propose d'établir la réciproque.

Montrer que $\mathfrak{p}$ est divisoriel, et en déduire que

$$(\mathrm{A} : \mathfrak{p}) \neq \mathrm{A}.$$

Montrer que, pour tout $x \in (\mathrm{A} : \mathfrak{p})$, on a

$$x\mathfrak{p} = \mathfrak{p} \qquad \text{ou} \qquad x\mathfrak{p} = \mathrm{A};$$

en déduire que

$$(\mathrm{A} : \mathfrak{p}) . \mathfrak{p} = \mathrm{A}$$

et par suite que l'idéal $\mathfrak{p}$ est *inversible*.

d) Montrer qu'on a $\mathfrak{p} \neq \mathfrak{p}^2$ et $\mathfrak{p} = \mathrm{A}x$ pour tout $x \in \mathfrak{p}$ n'appartenant pas à $\mathfrak{p}^2$.

e) Montrer que pour tout idéal $\mathfrak{a}$ de l'anneau A il existe un entier n tel que $\mathfrak{a}$ soit contenu dans $\mathfrak{p}^n$ mais non dans $\mathfrak{p}^{n+1}$. En utilisant le fait que $\mathfrak{p}$ est inversible, montrer que $\mathfrak{a} = \mathfrak{p}^n$ et en déduire que l'anneau A est principal.

f) Démontrer que A est l'anneau d'une valuation discrète.

50. Soit A un anneau d'intégrité commutatif, de corps des fractions K. On suppose A noethérien et intégralement clos, et que tout idéal premier non nul de A est maximal; on se propose de montrer que A est un anneau de Dedekind, i.e. que tout idéal fractionnaire de A est inversible.

(*) On appelle *anneau local* tout anneau commutatif A tel que l'ensemble des éléments non inversibles de A soit un idéal $\mathfrak{p}$ de A. Cet idéal est alors l'unique idéal maximal de A, et si A est intègre le sous-anneau $\mathrm{A}_{\mathfrak{p}}$ du corps des fractions de A est égal à A lui-même. Inversement, si A est un anneau d'intégrité, et si $\mathfrak{p}$ est un idéal premier de A, l'anneau $\mathrm{A}_{\mathfrak{p}}$ est un anneau local. Les anneaux locaux servent principalement à étudier les propriétés d'une variété algébrique « au voisinage » d'un point donné, ce qui explique la terminologie adoptée pour les désigner.

a) Soit $\mathfrak{p}$ un idéal premier non nul de A; montrer, à l'aide de l'*Exercice* précédent, que l'anneau local $A_{\mathfrak{p}}$ est l'anneau d'une valuation discrète de K. Soit $v_{\mathfrak{p}}$ cette valuation, choisie de telle sorte que

$$v_{\mathfrak{p}}(K^*) = \mathbf{Z}.$$

Montrer que les éléments de A sont caractérisés par le fait qu'on a

$$v_{\mathfrak{p}}(x) \geqslant 0$$

pour tout idéal premier non nul $\mathfrak{p}$ de A.

b) Montrer que, pour tout $x \in K$ non nul, les $\mathfrak{p}$ tels que $v_{\mathfrak{p}}(x) \neq 0$ sont en nombre fini (se ramener au cas où $x \in A$ et appliquer l'*Exercice* 6 du § 18 à l'idéal Ax de A). Montrer qu'étant donnés des idéaux premiers $\mathfrak{p}, \mathfrak{p}_1, \ldots, \mathfrak{p}_r$ de A, deux à deux distincts et en nombre fini, et des entiers $n_1, \ldots, n_r$, il existe un $x \in A$ vérifiant les relations

$$v_{\mathfrak{p}}(x) = 0, \qquad v_{\mathfrak{p}_i}(x) \geqslant n_i \quad \text{pour} \quad 1 \leqslant i \leqslant r.$$

c) Montrer que, pour tout idéal premier non nul $\mathfrak{p}$ de A, il existe un $x \in K$ tel que

$$v_{\mathfrak{p}}(x) = -1, \qquad v_{\mathfrak{q}}(x) \geqslant 0 \quad \text{pour tout} \quad \mathfrak{q} \neq \mathfrak{p}$$

(choisir un x_0 vérifiant $v_{\mathfrak{p}}(x_0) = -1$, poser $x = x_0 y$, et déterminer y à l'aide de la question précédente).

d) Montrer que tout idéal premier non nul de A est inversible, puis que A est un anneau de Dedekind (cf. § 18, *Exercice* 7).

e) Soient A un anneau de Dedekind, L une extension séparable de degré fini du corps des fractions de A, et B la clôture intégrale de A dans L. Montrer que B est un anneau de Dedekind. [Ce procédé, appliqué à $A = k[K]$ où k est un corps commutatif, conduit à des exemples d'anneaux de Dedekind tout à fait différents de ceux de la théorie des entiers algébriques].

f) Montrer que l'anneau $\mathbf{Z}[\sqrt{-5}]$ est un anneau de Dedekind, et que l'idéal engendré dans cet anneau par 3 et $1 + 2\sqrt{-5}$ est premier et non principal.

g) Soit k un corps commutatif algébriquement clos. On pose $A = k[X]$ (où X est une indéterminée sur k), $K = k(X)$, et

$$L = K\left[\sqrt{X^3 + pX + q}\right]$$

où p et q sont des éléments donnés de k (de sorte que L est une extension quadratique de K, correspondant à la « courbe du troisième degré » d'équation

$$y^2 = x^3 + px + q).$$

Trouver la clôture intégrale B de A dans L. Étant donné un $c \in k$, soit $\mathfrak{p}$ l'idéal premier (et maximal) de A formé des $f \in A$ tels que $f(c) = 0$; dans quel cas l'idéal $\mathfrak{p}B$ engendré par $\mathfrak{p}$ dans B est-il premier? S'il n'est pas premier, comment se décompose-t-il en produit de facteurs premiers?

¶¶¶ **51.** (Autre démonstration du Nullstellensatz de Hilbert, qui utilise l'*Exercice* 41).
Soit K un corps commutatif infini et soit

$$L = K[x_1, \ldots, x_n]$$

un anneau d'intégrité commutatif, contenant K et engendré par K et un nombre fini d'éléments $x_1, \ldots, x_n$.

a) On suppose qu'il existe une relation algébrique

$$f(x_1, \ldots, x_n) = 0$$

entre les x_i, où f est un polynôme non nul à n indéterminées et à coefficients dans K, de degré total r. Soit f_r la partie homogène de degré total r de f. Étant donnés des indéterminées $Z_1, \ldots, Z_{n-1}$, Y sur K et des éléments $c_1, \ldots, c_{n-1}$ de K, on pose

$$f(Z_1 + c_1 Y, \ldots, Z_{n-1} + c_{n-1} Y, Y) = \sum_{0 \leqslant k \leqslant r} p_k(Z_1, \ldots, Z_{n-1})\, Y^k;$$

montrer que le polynôme p_r est donné par

$$p_r(Z_1, \ldots, Z_{n-1}) = f_r(c_1, \ldots, c_{n-1}, 1)$$

et est donc constant. Montrer qu'il existe $c_1, \ldots, c_{n-1} \in K$ tels que

$$f_r(c_1, \ldots, c_{n-1}, 1) \neq 0$$

(utiliser l'homogénéité de f_r et le Théorème 1 du § 28). Les $c_i \in K$ étant ainsi choisis, on pose

$$z_i = x_i + c_i x_n \quad (1 \leqslant i \leqslant n-1);$$

montrer que $L = K[z_1, \ldots, z_{n-1}, x_n]$ et que x_n est entier sur le sous-anneau $K[z_1, \ldots, z_{n-1}]$ de L.

b) Déduire de là le résultat suivant (« lemme de normalisation » d'Emmy Noether; il est encore valable si K est fini, mais la démonstration est alors notablement plus difficile) : si $L = K[x_1, \ldots, x_n]$ est un anneau d'intégrité à engendrement fini sur le corps K, et si les $x \in L$ ne sont pas tous algébriques sur K, il existe $d \leqslant n$ éléments $z_1, \ldots, z_d$ de L possédant les propriétés suivantes : (*i*) les z_j sont des combinaisons linéaires des x_i à coefficients dans K (*ii*) $z_1, \ldots, z_d$ sont algébriquement indépendants sur K (*iii*) chaque élément de L est *entier* sur le sous-anneau $K[z_1, \ldots, z_d]$ de L.

c) Montrer que, si les $x \in L$ ne sont pas tous algébriques sur K, l'anneau L ne peut pas être un corps (écrire que l'inverse de z_d dans L est entier sur le sous-anneau $K[z_1, \ldots, z_d]$ et en déduire une contradiction). Autrement dit : si un anneau L à engendrement fini sur un corps commutatif K est un corps, alors L est extension algébrique (de degré fini nécessairement) de K, et en particulier L = K si K est algébriquement clos (d'où le Nullstellensatz : § 33, *Exercice* 33, (*e*)).

§ 35

Mettre sous la forme de Jordan les matrices suivantes (on calculera dans chaque cas le changement de base permettant de se ramener à la forme de Jordan) :

1. $\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$ **2.** $\begin{pmatrix} 12 & -6 & -2 \\ 18 & -9 & -3 \\ 18 & -9 & -3 \end{pmatrix}$ **3.** $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$

4. $\begin{pmatrix} 1 & -3 & 0 & 3 \\ -2 & -6 & 0 & 13 \\ 0 & -3 & 1 & 3 \\ -1 & -4 & 0 & 8 \end{pmatrix}$ **5.** $\begin{pmatrix} 3 & -1 & 1 & -7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$ **6.** $\begin{pmatrix} 0 & 0 & 0 & \dots & n \\ \dots & \dots & \dots & \dots & \dots \\ 0 & n & n-1 & \dots & 2 \\ n & n-1 & n-2 & \dots & 1 \end{pmatrix}$

¶ **7.** Montrer que si

$$A = \begin{pmatrix} a & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & a & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & a & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & a \end{pmatrix}$$

est une matrice de Jordan d'ordre n et si $f(X)$ est un polynôme à une variable, alors

$$f(A) = \begin{pmatrix} f(a) & f_1(a) & f_2(a) & \dots & f_{n-1}(a) \\ 0 & f(a) & f_1(a) & \dots & f_{n-2}(a) \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & f(a) \end{pmatrix}$$

où l'on pose

$$f_k(a) = f^{(k)}(a)/k\,!$$

(ceci suppose le corps de base de caractéristique 0 ; que se passe-t-il en caractéristique p non nulle ?)

¶ **8.** Soit A une matrice carrée à coefficients dans un corps commutatif K. Montrer, sans utiliser le théorème de Hamilton-Cayley, qu'il existe des polynômes non constants $f \in K[X]$ tels que $f(A) = 0$. Montrer que ce sont les multiples de celui d'entre eux qui possède le plus petit degré possible, et que celui-ci est unique si on impose à son coefficient dominant d'être égal à 1. On dit alors que c'est le **polynôme minimal de A sur K** ; il divise le polynôme caractéristique de A, et est donc de degré au plus égal à l'ordre de A.

Montrer que si A est la matrice de Jordan de l'*Exercice* précédent, le polynôme minimal de A est

$$(X - a)^n.$$

Montrer que si

$$A = \begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix},$$

le polynôme minimal de A est le ppcm des polynômes minimaux de A' et A''.
Montrer que deux matrices semblables A et PAP^{-1} ont le même polynôme minimal.
En supposant K algébriquement clos et en utilisant le théorème de Jordan, déduire des résultats précédents le calcul du polynôme minimal d'une matrice carrée quelconque.
Appliquer la méthode aux matrices des *Exercices* 1 à 6 ci-dessus.

9. Soient L un corps commutatif, K un sous-corps de L, et A une matrice carrée à coefficients dans K. Le polynôme minimal de A sur K est-il égal au polynôme minimal de A sur L?

10. Soient k un corps commutatif, V un espace vectoriel de dimension finie sur k, et u un endomorphisme de V. On considère l'anneau de polynôme $K = k[X]$, le K-module V[X] défini dans l'*Exercice* 19 des §§ 27, 28, et enfin le K-module V_u de l'*Exercice* 20 des §§ 27, 28; étant donnés un $x \in V$ et un $f \in K$ on notera

$$f.x = f(u)(x)$$

le produit de x par f dans le module V_u; on note d'autre part $\tilde{u}$ l'endomorphisme du K-module V[X] donné par

$$\tilde{u}(m_0 + m_1X + \cdots) = u(m_0) + u(m_1)X + \cdots$$

quels que soient les $m_i \in V$ presque tous nuls.

a) On considère l'application

$$\theta : V[X] \to V_u$$

donnée par

$$\theta(m_0 + m_1X + m_2X^2 + \cdots) = m_0 + u(m_1) + u^2(m_2) + \cdots ;$$

montrer que c'est un homomorphisme de K-modules, et que θ est surjective.

b) Montrer que le noyau de θ est égal à l'image de l'endomorphisme

$$\tilde{u} - X.j$$

de V[X] (où j désigne l'application identique de V[X] dans lui-même; $X.j$ est donc l'homothétie de rapport X dans ce $k[X]$-module).

c) Soit $A = (a_{ij})_{1 \leq i,j \leq n}$ la matrice de u par rapport à une base de V sur k; on considère la matrice

$$A - X.1_n = \begin{pmatrix} a_{11} - X & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} - X & \cdots & a_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} - X \end{pmatrix}$$

à coefficients dans l'anneau K; montrer, à l'aide de l'*Exercice* 15 du § 32, qu'il existe des matrices $P, Q \in GL_n(K)$ telles que

$$P(A - X.1_n)Q = \begin{pmatrix} d_1(X) & 0 & \cdots & 0 \\ 0 & d_2(X) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n(X) \end{pmatrix}$$

où $d_1, \ldots, d_n$ sont des polynômes non nuls tels que chacun divise le suivant (on fera attention au fait qu'en général les coefficients de P et Q dépendent effectivement de X). Montrer que pour tout i, le polynôme $d_1 \ldots d_i$ est un pgcd des mineurs d'ordre i de $A - X.1_n$. Dans ce qui suit on suppose le coefficient dominant de chaque d_i égal à 1.

d) A l'aide de la question *b*) et de l'*Exercice* 15 du § 32, montrer que le K-module V_u est isomorphe au produit direct des modules quotients K/d_iK. On suppose

$$d_1 = \cdots = d_s = 1$$

et d_{s+1} non constant; on pose

$$d_i(X) = X^{n_i} - a_{i,n_i-1}X^{n_i-1} - \cdots - a_{i,0}$$

pour $s+1 \leqslant i \leqslant n$; enfin on considère les matrices

$$A_i = \begin{pmatrix} 0 & 0 & 0 & \ldots & a_{i,0} \\ 1 & 0 & 0 & \ldots & a_{i,1} \\ 0 & 1 & 0 & \ldots & a_{i,2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & a_{i,n_i-1} \end{pmatrix};$$

montrer qu'il existe une base de V sur k telle que la matrice de u par rapport à cette base soit

$$\begin{pmatrix} A_{s+1} & 0 & \ldots & 0 \\ 0 & A_{s+2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & A_n \end{pmatrix}.$$

Montrer que d_s est le polynôme minimal de u.

e) On dit que $d_1(X), \ldots, d_n(X)$ sont les **invariants de similitude** de u (ou de la matrice A de u par rapport à une base quelconque de V sur k). Montrer que, pour que deux endomorphismes u et v de V soient semblables (i.e. pour qu'il existe un automorphisme w de V tel que $v = w \circ u \circ w^{-1}$) il faut et il suffit qu'ils aient les mêmes invariants de similitude. [Noter que si u et v sont semblables, et si A et B sont leurs matrices par rapport à une base quelconque de V, alors les matrices $A - X.1_n$ et $B - X.1_n$ sont équivalentes sur l'anneau $K = k[X]$, et appliquer le § 32, *Exercice* 15, *e*), ou bien utiliser la fin de la question *c*) ci-dessus].

Ou encore : soient A et B deux matrices carrées d'ordre n à coefficients dans un corps commutatif arbitraire k; pour qu'il existe une matrice $U \in GL(n, k)$ telle que

$$B = UAU^{-1},$$

il faut et il suffit que, pour $1 \leqslant i \leqslant n$, le pgcd des mineurs d'ordre i de la matrice $A - X.1_n$ soit égal au pgcd des mineurs d'ordre i de la matrice $B - X.1_n$.

(Corollaire immédiat : *toute matrice* $A \in M_n(k)$ *est semblable à sa transposée* ${}^tA$).

¶¶ **11.** Montrer que les matrices

$$\begin{pmatrix} 3 & 2 & -5 \\ 2 & 6 & -10 \\ 1 & 2 & -3 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 6 & 20 & -34 \\ 6 & 32 & -51 \\ 4 & 20 & -32 \end{pmatrix}$$

sont semblables en calculant leurs invariants de similitude. Même question pour

$$\begin{pmatrix} 4 & 10 & -19 & 4 \\ 1 & 6 & -8 & 3 \\ 1 & 4 & -6 & 2 \\ 0 & -1 & 1 & 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 41 & -4 & -26 & -7 \\ 14 & -13 & -91 & -18 \\ 40 & -4 & -25 & -8 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

¶¶ **12.** Soient L un corps commutatif, K un sous-corps de L, et A, B deux matrices carrées d'ordre n à coefficients dans K. On suppose A et B semblables comme matrices à coefficients dans L; montrer que A et B sont aussi semblables comme matrices à coefficients dans K (utiliser l'*Exercice* 10). Rapport avec l'*Exercice* 22 du § 27?

¶¶ **13.** Soit A une matrice carrée d'ordre n à coefficients dans un corps commutatif K; on suppose que le polynôme caractéristique de A a toutes ses racines dans K (i.e. que A est trigonalisable sur K). Montrer qu'il existe un tableau diagonal de matrices de Jordan sur K dont les invariants de similitude sont égaux à ceux de A. Déduire de là et de la question *e*) de l'*Exercice* 10 une nouvelle démonstration du théorème de Jordan.

¶ **14.** Soit A une matrice carrée *inversible* d'ordre n à coefficients dans un corps algébriquement clos K. Montrer qu'il existe dans $\mathrm{GL}(n, \mathbf{K})$ une matrice *diagonalisable* D et une matrice *unipotente* U telles que l'on ait

$$A = DU = UD,$$

et que de plus D et U sont uniques (se ramener à une matrice de Jordan). On dit que D et U sont les **composantes semi-simple et unipotente** de A [on peut démontrer que, si K est par exemple de caractéristique 0, ou fini, mais non nécessairement algébriquement clos, alors D et U, calculées dans une extension algébriquement close de K, sont encore à coefficients dans K; si par exemple $\mathbf{K} = \mathbf{R}$, il est clair que si $D, U \in M_n(\mathbf{C})$ conviennent il en est encore même des matrices imaginaires conjuguées, et vu l'unicité de D et U on voit bien qu'on a en fait $D, U \in M_n(\mathbf{R})$ dans ce cas. Naturellement, si K n'est pas algébriquement clos, D est semi-simple mais non nécessairement diagonalisable sur K.]

¶ **15.** Soit $X \in M_n(K)$ où K est algébriquement clos. Montrer qu'il existe une matrice diagonalisable D et une matrice nilpotente N telles que

$$X = D + N, \qquad D.N = N.D,$$

et que ces matrices sont entièrement déterminées par ces conditions. On prend $\mathbf{K} = \mathbf{C}$ et X à coefficients réels; montrer qu'il en est de même de D et N. (En fait, comme dans le cas précédent, on peut montrer que si X est à coefficients dans un sous-corps de caractéristique 0, ou fini, il en est de même de D et N.)

¶ **16.** On désigne par E l'ensemble de toutes les applications (*)

$$f : \mathbf{N} \to \mathbf{C}$$

qui vérifient la relation

$$(*) \qquad f(n + p) = a_{p-1} f(n + p - 1) + \cdots + a_0 f(n) \quad \text{pour tout } n \in \mathbf{N},$$

où $a_0, \ldots, a_{p-1}$ sont des nombres complexes donnés.

a) Montrer qu'il existe une et une seule $f \in E$ pour laquelle les nombres $f(0), \ldots, f(p-1)$ ont des valeurs données (on ne demande pas de calculer f explicitement). En déduire que E est un sous-espace vectoriel de dimension p de l'espace de toutes les applications de N dans C.

b) Pour $0 \leqslant i \leqslant p-1$, on désigne par e_i l'unique élément de E tel que l'on ait

$$e_i(j) = \begin{cases} 0 & \text{si } j \neq i \\ 1 & \text{si } j = i \end{cases} \quad \text{pour } 0 \leqslant j \leqslant p-1.$$

(*) Ces « applications » ne sont autres que les « suites » de nombres complexes, mais il est plus commode ici de les regarder comme des fonctions.

Montrer que $e_0, \ldots, e_{p-1}$ forment une base de E.

c) Montrer qu'il existe un et un seul endomorphisme u de E tel que, pour toute $f \in E$, la fonction $g = u(f)$ soit donnée par

$$g(n) = f(n+1).$$

Calculer la matrice de u par rapport à la base $e_0, \ldots, e_{p-1}$ de E, et montrer que le polynôme caractéristique de u est, au signe près,

$$X^p - a_{p-1}X^{p-1} - \cdots - a_0.$$

d) Montrer que pour tout entier $r \geqslant 0$ et tout $\lambda \in \mathbf{C}$, le sous-espace

$$\mathrm{Ker}[(u-\lambda)^r]$$

de E est formé des $f \in E$ qui sont de la forme

$$f(n) = \lambda^n g(n)$$

où la fonction g est polynomiale de degré $r-1$ au plus.

e) Soient $\lambda_1, \ldots, \lambda_h$ les diverses racines de l'équation

$$\lambda^p - a_{p-1}\lambda^{p-1} - \cdots - a_0 = 0,$$

et $r_1, \ldots, r_h$ leurs ordres de multiplicité. Montrer que, pour qu'une application f de N dans C soit dans E, i.e. vérifie la relation de récurence (*), il faut et il suffit qu'il existe des polynômes

$$g_1, \ldots, g_h \in \mathbf{C}[X],$$

vérifiant

$$d^\circ(g_1) < r_1, \ \ldots, \ d^\circ(g_h) < r_h,$$

et tels que l'on ait

$$f(n) = g_1(n)\lambda_1^n + \cdots + g_h(n)\lambda_h^n \quad \text{pour tout } n \in \mathbf{N};$$

s'il en est ainsi, les polynômes $g_1, \ldots, g_h$ sont entièrement déterminés par f.

f) Trouver toutes les suites $(u_n)_{n \geqslant 0}$ de nombres complexes telles que l'on ait

$$u_{n+5} = u_{n+4} + 5u_{n+3} - u_{n+2} - 8u_{n+1} - 4u_n$$

pour tout $n \geqslant 0$.

¶ **17.** Soit $A = (a_{ij})_{1 \leqslant i,\, j \leqslant p}$ une matrice carrée d'ordre p à coefficients complexes. Trouver toutes les applications

$$f = (f_1, \ldots, f_p) : \mathbf{N} \to \mathbf{C}^p$$

qui vérifient

$$f_i(n+1) = \sum_{j=1}^{j=p} a_{ij} f_j(n)$$

pour $1 \leqslant i \leqslant p$ et tout $n \in \mathbf{N}$ (Mettre A sous la forme de Jordan).

Utiliser les résultats obtenus pour retrouver ceux de l'*Exercice* précédent, en associant à toute solution de la relation (*) la fonction

$$(f(n), f(n+1), \ldots, f(n+p-1))$$

à valeurs dans $\mathbf{C}^p$.

¶ **18.** Trouver toutes les applications (f_1, f_2, f_3, f_4) de $\mathbf{N}$ dans $\mathbf{C}^4$ vérifiant les relations suivantes :

$$\begin{aligned} f_1(n+1) &= -5f_1(n) - 3f_2(n) - 2f_3(n) + 4f_4(n) \\ f_2(n+1) &= 2f_1(n) + f_3(n) - f_4(n) \\ f_3(n+1) &= 10f_1(n) + 7f_2(n) + 4f_3(n) - 9f_4(n) \\ f_4(n+1) &= 2f_1(n) + f_3(n) \end{aligned}$$

(utiliser l'*Exercice* précédent).

¶ **19.** Soit K un corps algébriquement clos et de caractéristique 0; dans cet *Exercice* (*), on considère des séries formelles à une indéterminée à coefficients dans K (§§ 27, 28, *Exercice* 11).

a) Étant donnée une série formelle

$$x = \sum_{n\in\mathbf{N}} f(n)\mathrm{T}^n/n!$$

en une indéterminée T, à coefficients dans K (de sorte que f est une application de l'ensemble N des entiers naturels dans K), on appelle **dérivée** de x la série formelle

$$x' = \sum_{n\in\mathbf{N}} f(n+1)\mathrm{T}^n/n!.$$

Montrer que l'application $x \to x'$ est une dérivation de l'anneau K[[T]]. Dans ce qui suit, on notera

$$x'' = (x')', \quad x''' = (x'')', \quad \ldots, \quad x^{(p)} = (x^{(p-1)})', \quad \ldots$$

les dérivées successives de x.

b) Étant données des constantes $a_0, \ldots, a_{p-1} \in \mathrm{K}$, montrer que la recherche des séries formelles

$$x = \sum_{n\in\mathbf{N}} f(n)\mathrm{T}^n/n!$$

vérifiant l'équation différentielle linéaire et homogène à coefficients constants

$$x^{(p)} = a_{p-1}x^{(p-1)} + \cdots + a_0 x \qquad (**)$$

revient à la résolution de l'équation (*) de l'*Exercice* 16.

c) Pour tout $\lambda \in \mathrm{K}$, on considère la série formelle

$$\exp(\lambda\mathrm{T}) = \sum_{n\in\mathbf{N}} \lambda^n \mathrm{T}^n/n!$$

(*) Le but de l'*Exercice* 19 et des suivants est de montrer au lecteur la liaison existant entre la théorie de la réduction des matrices et celle des systèmes d'équations différentielles. Il va de soi que, vu son importance, le sujet mériterait de beaucoup plus amples développements — mais ceux-ci appartiennent plus à un cours d'Analyse qu'à un cours d'Algèbre. L'intervention de séries formelles dans la théorie est conforme aux meilleures traditions, puisque la méthode de Cauchy-Kowalewska pour établir l'existence de solutions pour des systèmes d'équations différentielles à coefficients analytiques consiste d'abord à construire des séries entières qui vérifient « formellement » les équations données (autrement dit, à se placer, comme nous le faisons ici, dans le cadre des séries formelles, convergentes ou non), puis à démontrer, à l'aide de majorations de leurs coefficients, que ces séries formelles convergent. Dans le cas des systèmes étudiés ici, les démonstrations de convergence (lorsque $\mathrm{K} = \mathbf{C}$ bien entendu) sont triviales vu la forme particulièrement simple des séries obtenues.

(dont la définition est évidemment inspirée du développement en série entière de la fonction exponentielle classique). Étant donnée une série formelle

$$x = \Sigma f(n)\mathrm{T}^n/n!,$$

montrer que les propriétés suivantes sont équivalentes : (*i*) on a $f(n) = g(n)\lambda^n$ où g est une fonction polynomiale sur **N**, à coefficients dans K, et de degré r; (*ii*) la série formelle x est produit de la série $\exp(\lambda\mathrm{T})$ par un polynôme de degré r en T, à coefficients dans K. (Il pourra être utile d'utiliser l'*Exercice* 8 des §§ 27, 28).

d) Soient $\lambda_1, \ldots, \lambda_h$ les racines dans K de l'équation

$$\lambda^p = a_{p-1}\lambda^{p-1} + \cdots + a_0$$

et $r_1, \ldots, r_h$ leurs ordres de multiplicité. Montrer que la solution générale de l'équation différentielle (**) est

$$x = g_1(\mathrm{T})\exp(\lambda_1\mathrm{T}) + \cdots + g_h(\mathrm{T})\exp(\lambda_h\mathrm{T})$$

où chaque g_i est un polynôme de degré $r_i - 1$ au plus à coefficients dans K.

e) On suppose $\mathrm{K} = \mathbf{C}$. Que resterait-il à faire pour déduire des résultats précédents la théorie classique des équations différentielles linéaires et homogènes à coefficients constants?

f) On cherche maintenant p séries formelles

$$x_i = \Sigma f_i(n)\mathrm{T}^n/n!$$

vérifiant le système

$$x_i' = \sum_{j=1}^{j=p} a_{ij}x_j$$

où les a_{ij} sont des éléments donnés de K. Montrer que la résolution de ce problème revient à celle de l'*Exercice* 17, et interpréter les résultats de l'*Exercice* 17 dans le langage de la théorie des systèmes d'équations différentielles.

Dans les *Exercices* suivants, on demande, en utilisant l'*Exercice* 19, *f*), de résoudre les systèmes différentiels donnés :

20.
$$\begin{aligned} x' &= 5x - 3y + 2z \\ y' &= 6x - 4y + 4z \\ z' &= 4x - 4y + 5z \end{aligned}$$

21.
$$\begin{aligned} x' &= 7x - 12y + 6z \\ y' &= 10x - 19y + 10z \\ z' &= 12x - 24y + 13z \end{aligned}$$

22.
$$\begin{aligned} x' &= x - 3y + 3z \\ y' &= -2x - 6y + 13z \\ z' &= -x - 4y + 8z \end{aligned}$$

23.
$$\begin{aligned} x' &= 3x - y \\ y' &= x + y \\ z' &= 3x \quad + 5z - 3u \\ u' &= 4x - y + 3z - u \end{aligned}$$

24. Intégrer l'équation différentielle

$$x^{(5)} - x^{(4)} - 5x^{(3)} + x'' + 8x' + 4x = 0.$$

25. Utiliser l'*Exercice* 16 pour établir l'identité

$$\sum_{p=1}^{p=n} p^2 a^p = \frac{a(a+1)}{(1-a)^3} + \frac{(a-7)\,n - (2a^2 - 5a + 1)\,n^2}{2(1-a)^3}\,a^{n+1}$$

(on suppose $a \neq 1$).

26. Soient K un corps commutatif et n un entier positif. On désigne par V l'espace vectoriel (sur K) formé des polynômes à une indéterminée, à coefficients dans K, de degré au plus égal à n. Quelle est la dimension de V sur K? On désigne par D l'application de V dans V qui transforme chaque polynôme en le polynôme dérivé. Montrer que D est un endomorphisme nilpotent de V, et trouver une base de V sur K par rapport à laquelle la matrice de D ait la forme de Jordan.

§ 36

(Les *Exercices* **1** à **21** sont relatifs à des propriétés valables sur un corps de base K quelconque, à ceci près que le lecteur devra parfois exclure les corps de caractéristique 2, ce que nous n'avons pas mentionné explicitement dans les énoncés en question. Les *Exercices* **22** à **51** supposent par contre que le corps de base est **R** ou **C**, ce qu'on a indiqué dans les énoncés. Il va de soi par ailleurs que les énoncés valables sur un corps de base quelconque, notamment ceux des *Exercices* **1**, **2**, **9** à **21**, sont tout aussi utiles lorsque le corps de base est **R** ou **C**.)

1. Soit V un espace vectoriel de dimension finie sur un corps commutatif K; on appelle **forme quadratique** sur V toute fonction polynomiale homogène de degré 2 sur V (§§ 27, 28, *Exercice* 17); i.e. toute application q de V dans K donnée par une relation de la forme

$$q(x) = \sum a_{ij} x_i x_j$$

où les a_{ij} sont des éléments donnés de K, et où les $x_i \in K$ sont les coordonnées du vecteur $x \in V$ par rapport à une base de V.

a) Montrer que si $f(x, y)$ est une forme bilinéaire symétrique sur V, la fonction

$$q(x) = f(x, x)$$

est une forme quadratique sur V (dite **associée** à f).

b) On suppose K de caractéristique $\neq 2$. Montrer que la donnée de q permet de reconstituer f, à l'aide de la formule

$$f(x, y) = \frac{q(x + y) - q(x - y)}{4}.$$

c) Inversement, si q est une forme quadratique sur V, la formule précédente définit une forme bilinéaire symétrique f sur V, et on a $q(x) = f(x, x)$.

d) Pour qu'une base de V soit orthogonale par rapport à f, il faut et il suffit que l'expression de q par rapport à cette base soit de la forme

$$q(x) = c_1 x_1^2 + c_2 x_2^2 + \cdots + c_r x_r^2$$

(on dit alors que q est **réduite à une somme de carrés**).

[Les résultats précédents montrent que l'étude des formes bilinéaires symétriques est équivalente, en caractéristique $\neq 2$, à celle des formes quadratiques. On utilisera souvent le langage des formes quadratiques dans les *Exercices* suivants.]

2. Soit K un corps commutatif de caractéristique $\neq 2$. On considère une forme quadratique

$$q(x) = \sum_{1 \leqslant i,j \leqslant n} a_{ij} x_i x_j$$

sur K^n.

a) On suppose $a_{11} \neq 0$; montrer qu'il existe alors une forme linéaire f_1 sur K^n telle que l'on ait

$$q(x) = a_{11} . f_1(x)^2 + q_1(x)$$

où la forme quadratique q_1 ne dépend plus de la variable x_1 (écrire $q(x)$ comme un trinôme du second degré en x_1, dont les coefficients dépendent de $x_2, \ldots, x_n$, et mettre ce trinôme sous la forme canonique des Lycées et Collèges).

b) On suppose $a_{11} = 0$ mais $a_{12} \neq 0$, et on choisit comme nouvelles coordonnées dans K^n les formes linéaires

$$\begin{aligned} y_1 &= x_1 + x_2 \\ y_2 &= x_1 - x_2 \\ y_i &= x_i \qquad (3 \leqslant i \leqslant n). \end{aligned}$$

Montrer qu'on a

$$q(x) = \sum_{1 \leqslant i,j \leqslant n} b_{ij} y_i y_j$$

avec $b_{11} \neq 0$, et par suite qu'on peut, dans le nouveau système de coordonnées, appliquer la question *a*).

c) Déduire de ce qui précède une méthode pratique pour réduire une forme quadratique à une somme de carrés *Exercice* [1, *d*)], ou pour construire une base orthogonale pour une forme bilinéaire symétrique donnée.

Dans les *Exercices* 3 à 8, on demande de mettre la forme quadratique donnée sous forme d'une somme de carrés en utilisant la méthode indiquée dans l'*Exercice* 2; on prendra $\mathbb{Q}$ pour corps de base, et on indiquera dans chaque cas le changement de coordonnées qui conduit au résultat cherché

3. $x_1^2 + x_2^2 + 3x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3$.

4. $x_1^2 + 5x_2^2 - 4x_3^2 + 2x_1x_2 - 4x_1x_3$.

5. $2x_1^2 + 18x_2^2 + 8x_3^2 - 12x_1x_2 + 8x_1x_3 - 27x_2x_3$.

6. $x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_1x_3 + 2x_1x_4 + 2x_2x_3 + 2x_2x_4 + 2x_3x_4$.

7. $3x_1^2 + 2x_2^2 - x_3^2 - 2x_4^2 + 2x_1x_2 - 4x_2x_3 + 2x_3x_4$.

8. $3x_1^2 - 2x_2^2 + 2x_3^2 + 4x_1x_2 - 3x_1x_3 - x_2x_3$.

9. Soit f une forme bilinéaire sur un espace vectoriel E de dimension finie sur un corps K.

a) Soit

$$A = (f(a_i, a_j))_{1 \leqslant i,j \leqslant n}$$

la matrice de f par rapport à une base (a_i) de E. On identifie chaque vecteur $x \in E$ à la matrice colonne formée avec les composantes de x par rapport à la base en question. Montrer qu'on a

alors

$$f(x, y) = {}^t y . A . x$$

quels que soient $x, y \in E$.

b) Soit B la matrice de f par rapport à une autre base (b_i) de E. On note P la matrice de passage de la base (a_i) à la base (b_i). Montrer que

$$B = {}^t PAP.$$

c) Déduire de là que, pour toute matrice symétrique $S \in M_n(K)$, il existe une matrice diagonale D et une matrice $P \in GL(n, K)$ telles que

$$S = {}^t PDP,$$

et réciproquement; montrer de plus que, si K est algébriquement clos, on peut supposer tous les coefficients de D égaux à 1 ou 0, et que si $K = \mathbf{R}$ on peut supposer qu'ils sont égaux à 0, + 1 ou — 1.

10. Pour toute matrice symétrique S d'ordre n, à coefficients dans un corps algébriquement clos K, il existe une matrice $X \in M_n(K)$ telle que

$$S = {}^t XX,$$

et réciproquement.

11. Soit f une forme hermitienne sur un espace vectoriel E de dimension finie sur un corps commutatif K (muni d'une involution, cf. l'introduction du § 36).

a) Soit $A = (f(a_i, a_j))_{1 \leq i, j \leq n}$ la matrice de f par rapport à une base (a_i) de E. Montrer que, si l'on identifie chaque $x \in E$ à la matrice colonne formée avec les composantes de x par rapport à la base en question, on a

$$f(x, y) = y^* . A . x$$

quels que soient $x, y \in E$.

b) Soit u une endomorphisme de E, dont la matrice par rapport à la base (a_i) est U; montrer que la matrice par rapport à cette base de l'adjoint de u relativement à f (on suppose maintenant f non dégénérée) est

$$A^{-1}U^*A.$$

c) En utilisant l'existence pour toute forme hermitienne d'une base orthogonale, montrer que, pour toute matrice hermitienne $A \in M_n(K)$, il existe une matrice hermitienne diagonale $D \in M_n(K)$ et une matrice inversible $P \in GL(n, K)$ telles que

$$A = P^*DP.$$

12. Pour que le produit de deux matrices hermitiennes soit une matrice hermitienne, il faut et il suffit que les deux matrices données commutent.

13. Soient E un espace vectoriel complexe de dimension finie et f une forme hermitienne définie positive sur E.

a) Montrer que, pour tout sous-espace vectoriel M de E, l'opérateur de projection orthogonale p_M est autoadjoint relativement à f.

b) Inversement, pour tout endomorphisme p de E tel que

$$p = p^* = p^2,$$

il existe un sous-espace vectoriel M tel que $p = p_M$.

c) Soient M et N deux sous-espaces vectoriels de E; soit M′ (resp. N′) le sous-espace formé des $x \in M$ (resp. $x \in N$) orthogonaux à $M \cap N$. Montrer que, pour que p_M et p_N commutent, il faut et il suffit que M′ et N′ soient orthogonaux (la situation généralise alors celle de deux plans perpendiculaires dans l'espace usuel). Si cette condition est réalisée, on a

$$p_{M \cap N} = p_M \circ p_N,$$
$$p_{M+N} = p_M + p_N - p_M \circ p_N.$$

d) Généralisation à un corps de base quelconque? (Considérer des sous-espaces non isotropes).

14. Soit f une forme bilinéaire symétrique sur un espace vectoriel E de dimension finie sur un corps commutatif K. Soient $a_1, \ldots, a_r$ des éléments de E et U le sous-espace vectoriel qu'ils engendrent. Les deux propriétés suivantes sont équivalentes: (*i*) U est non isotrope et les a_i forment une base de U; (*ii*) le déterminant des $f(a_i, a_j)$ n'est pas nul. Corollaire: si f ne possède aucun vecteur isotrope (exemple: $K = \mathbf{R}$ et f définie positive), pour que des vecteurs $a_1, \ldots, a_r$ soient linéairement indépendants il faut et il suffit que le déterminant des $f(a_i, a_j)$ soit non nul.

¶ **15.** Soit

$$S = (a_{ij})_{1 \leq i, j \leq n}, \qquad a_{ij} = a_{ji},$$

une matrice symétrique à coefficients dans un corps commutatif K. On suppose que les mineurs principaux

$$D_p = \begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix} \quad (1 \leq p \leq n)$$

de S ne sont pas nuls. Montrer qu'il existe alors une matrice diagonale

$$D = \begin{pmatrix} c_1 & 0 & 0 & \cdots & 0 \\ 0 & c_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & c_n \end{pmatrix}$$

et une matrice triangulaire de la forme

$$T = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ t_{21} & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ t_{n1} & t_{n2} & t_{n3} & \cdots & t_{n,n-1} & 1 \end{pmatrix},$$

à coefficients dans K, telles que

$$S = TD{}^tT;$$

que D et T sont entièrement déterminées par S; et que les termes diagonaux de D sont donnés par les relations

$$c_1 = D_1, \quad c_2 = D_2/D_1, \quad \ldots, \quad c_n = D_n/D_{n-1}.$$

Corollaire (Jacobi) : soit

$$q(x) = \Sigma a_{ij} x_i x_j \quad (a_{ij} = a_{ji})$$

une forme quadratique telle que les mineurs principaux de la matrice (a_{ij}) ne soient pas nuls; alors on peut réduire q à une somme de carrés à l'aide d'un changement *triangulaire* de coordonnées.

(Pour établir l'existence de T et D on pourra utiliser le § 23, *Exercice* 11, ou bien raisonner par récurrence sur n, en écrivant

$$S = \begin{pmatrix} S_1 & {}^t u \\ u & a_{nn} \end{pmatrix}$$

où S_1 est une matrice carrée d'ordre $n-1$, u une matrice colonne à $n-1$ éléments, et utiliser une décomposition en blocs analogues pour D et T).

¶ **16.** Les notations restant celles de l'*Exercice* précédent, on suppose S de rang r quelconque, et on cherche à mettre S sous la forme

$$S = TD{}^tT$$

avec $c_1 \neq 0, \ldots, c_r \neq 0, c_{r+1} = \cdots = c_n = 0$. Montrer que, pour que le problème admette une solution, il faut et il suffit qu'on ait

$$D_p \neq 0 \quad \text{pour } 1 \leqslant p \leqslant r.$$

17. Soient E un espace vectoriel de dimension finie sur un corps commutatif K, et f une forme bilinéaire symétrique non dégénérée sur E; on note q la forme quadratique $f(x, x)$. On suppose qu'il existe des vecteurs isotropes non nuls pour f. Montrer alors que, pour tout $c \in K$, il existe $x \in E$ tel que

$$q(x) = c.$$

¶ **18.** Soient E un espace vectoriel de dimension finie n sur un corps commutatif K, et f une forme bilinéaire symétrique non dégénérée sur E.

a) Soit H un sous-espace vectoriel de E, non isotrope pour f. Montrer qu'il existe un et un seul automorphisme s_H de f pour lequel on ait

$$s_H(x) = \begin{cases} x & \text{si } x \in H \\ -x & \text{si } x \in H^\perp \end{cases}$$

(**symétrie par rapport à H**).

b) Soient u un automorphisme de f et x un vecteur non isotrope pour f. Montrer que l'un au moins des vecteurs $u(x) - x$, $u(x) + x$ est non isotrope. En déduire qu'il existe dans E un sous-espace H non isotrope tel que l'on ait $s_H(u(x)) = x$. [Prendre pour H soit le sous-espace orthogonal à $u(x) + x$, soit le sous-espace engendré par $u(x) - x$].

c) En déduire, par récurence sur n, que *tout automorphisme de f est produit d'au plus n symétries par rapport à des sous-espaces non isotropes de* E.

¶¶ **19.** Soient E un espace vectoriel de dimension finie sur un corps commutatif K et f une forme bilinéaire symétrique sur E. On dit qu'un sous-espace vectoriel U de E est **totalement isotrope** pour f si $f(x, x) = 0$ pour tout $x \in U$; on dit que U est un sous-espace totalement isotrope **maximal** s'il n'est contenu dans aucun autre sous-espace totalement isotrope pour f. On se propose d'établir que *tous les sous-espaces totalement isotropes maximaux de f ont la même dimension*, qu'on appelle l'**indice** de f (ou de la forme quadratique correspondante).

a) Pour que U soit totalement isotrope pour f, il faut et il suffit que

$$f(x, y) = 0 \quad \text{quels que soient } x, y \in U$$

i.e. que $U \subset U^\perp$ (utiliser l'*Exercice* 1).

b) Soient U et V deux sous-espaces totalement isotropes pour f. Montrer que pour tout, $x \in U \cap V^\perp$, le sous-espace $V + Kx$ est totalement isotrope.

c) Soient U et V deux sous-espaces totalement isotropes pour f; soient M un supplémentaire de $U \cap V$ dans U, et N un supplémentaire de $U \cap V$ dans V. Montrer que

$$U \cap V^{\perp} = (U \cap V) \oplus (M \cap N^{\perp}).$$

Montrer que les éléments de $M \cap N^{\perp}$ s'obtiennent en résolvant un système de $s = \dim(N)$ équations linéaires et homogènes à $r = \dim(M)$ inconnues. En déduire que, si

$$\dim(V) < \dim(U),$$

il existe un $x \in U$ non dans V tel que le sous-espace $V + Kx$ soit totalement isotrope.

d) Montrer que tout sous-espace totalement isotrope est contenu dans au moins un sous-espace totalement isotrope maximal. En déduire, à l'aide de la question *c*), le résultat annoncé.

¶¶ **20.** (Démonstration du théorème de Witt). Soit f une forme bilinéaire symétrique non dégénérée sur un espace vectoriel E de dimension finie sur un corps commutatif K de caractéristique différente de 2. Soient M et N deux sous-espaces de même dimension de E, et u une application linéaire bijective de M sur N. On se propose de montrer que les deux propriétés suivantes sont équivalentes : (*i*) il existe un automorphisme de f qui coïncide avec u sur M; (*ii*) on a $f[u(x), u(y)] = f(x, y)$ quels que soient $x, y \in M$. Comme il est trivial que (*i*) implique (*ii*), on se borne ci-dessous à montrer que (*ii*) implique (*i*).

a) Soient x et y deux éléments de E tels que

$$f(x, x) = f(y, y) \neq 0;$$

montrer qu'il existe un automorphisme de f appliquant x sur y (montrer que $x - y$ et $x + y$ ne sont pas tous les deux isotropes, et prendre la symétrie par rapport à H, où H est soit l'hyperplan orthogonal à $x - y$, soit la droite engendrée par $x + y$).

b) On suppose que M et N ne sont pas totalement isotropes. Montrer à l'aide de la question *a*) qu'on peut supposer $u(x) = x$ pour un $x \in M$ non isotrope. Soit E' l'hyperplan orthogonal à x; montrer que, pour construire un automorphisme de f prolongeant u, il suffit de construire un automorphisme de la restriction f' de f à E' égal à u sur $E' \cap M$. En déduire dans ce cas le théorème de Witt par récurence sur dim(E).

c) On suppose dorénavant M et N totalement isotropes. On choisit un $x \notin M^{\perp}$. Montrer qu'il existe un $y \notin N^{\perp}$ tel que l'on ait

$$f[y, u(z)] = f(x, z) \quad \text{pour tout } z \in M.$$

Montrer qu'on peut en outre supposer

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

(remplacer y par $y + tn$ avec $t \in K$ et $n \in N$ convenablement choisis).

d) Déduire de la question *c*) qu'il existe des sous-espaces non totalement isotropes $M' \supset M$ et $N' \supset N$, ainsi qu'un isomorphisme u' de M' sur N', tels que l'on ait

$$f[u'(x), u'(y)] = f(x, y) \quad \text{quels que soient } x, y \in M'$$
$$u' = u \text{ sur } M,$$

et achever à partir de là la démonstration du théorème de Witt.

¶¶ **21.** Soient E un espace vectoriel de dimension finie n sur un corps commutatif K, f une forme bilinéaire symétrique non dégénérée sur E, et M un sous-espace totalement isotrope de dimension r de E.

a) Montrer qu'il existe dans E des vecteurs non isotropes non orthogonaux à M.

b) Montrer qu'il existe un sous-espace totalement isotrope N tel que

$$E = M^{\perp} \oplus N$$

(prendre le symétrique de M par rapport à la droite engendrée par l'un des vecteurs construits dans la question précédente).

c) Soit $H = M^{\perp} \cap N^{\perp}$; montrer que

$$E = M \oplus H \oplus N,$$

et que H ne contient aucun vecteur isotrope non nul si M est totalement isotrope maximal. On forme une base de E en réunissant des bases de M, H et N; montrer que la matrice de f par rapport à cette base est de la forme

$$S = \begin{pmatrix} 0 & 0 & A \\ 0 & S_1 & 0 \\ {}^tA & 0 & 0 \end{pmatrix}$$

où A est une matrice carrée inversible d'ordre r, et S_1 une matrice symétrique d'ordre $n - 2r$.

d) Trouver à quelles conditions une matrice de la forme

$$\begin{pmatrix} U & 0 & 0 \\ 0 & V & 0 \\ 0 & 0 & W \end{pmatrix}$$

(où U et W sont carrées d'ordre r, et V carrée d'ordre $n - 2r$) représente, par rapport à la base considérée dans E, un automorphisme de f. En déduire que pour tout automorphisme u de l'espace vectoriel M, il existe un automorphisme de f qui se réduit à u sur M. Pouvez-vous déduire ce résultat du théorème de Witt?

22. Soit $q(x)$ une forme quadratique sur un espace vectoriel réel (resp. complexe) E de dimension finie; montrer qu'il existe une base de E par rapport à laquelle l'expression de q est de la forme

$$(*) \qquad x_1^2 + \cdots + x_p^2 - (x_{p+1}^2 + \cdots + x_{p+q}^2)$$

(resp.

$$x_1^2 + \cdots + x_r^2).$$

Trouver une telle base (dans le cas réel et dans le cas complexe) pour les formes quadratiques des *Exercices* 3 à 8.

23. Soient $f_1, \ldots, f_{p+q}$ des formes linéaires sur un espace vectoriel réel U de dimension finie; on suppose que, sur U, la forme quadratique

$$q(x) = f_1(x)^2 + \cdots + f_p(x)^2 - f_{p+1}(x)^2 - \cdots - f_{p+q}(x)^2$$

soit définie positive, i.e. vérifie

$$q(x) > 0 \qquad \text{pour tout} \qquad x \neq 0.$$

Montrer qu'on a alors

$$\dim (U) \leqslant p.$$

(Remarquer que dans le cas contraire il existerait un $x \neq 0$ où $f_1, \ldots, f_p$ seraient toutes nulles).

¶ **24.** Soit $q(x)$ une forme quadratique sur un espace vectoriel réel E de dimension finie. On choisit une base de E par rapport à laquelle q est mise sous la forme (*) de l'*Exercice* 22 ; montrer, à l'aide de l'*Exercice* précédent, que tout sous-espace vectoriel U de E sur lequel $q(x)$ est définie positive est de dimension p au plus. En déduire (**loi d'inertie** des formes quadratiques) que les entiers p et q de l'*Exercice* 22 sont indépendants du choix de la base de E par rapport à laquelle $q(x)$ se met sous la forme (*).
[Le couple (p, q) formé du nombre p de carrés positifs et du nombre q de carrés négatifs dans la formule (*) s'appelle la **signature de la forme quadratique** considérée ou de la forme bilinéaire symétrique correspondante. Le nombre p est la dimension maximum des sous-espaces sur lesquelles la forme donnée est définie positive, et le nombre q la dimension maximum des sous-espaces sur lesquels elle est définie négative].

¶ **25.** Deux formes quadratiques q et q' sur un espace vectoriel E de dimension finie sur un corps commutatif sont dites **équivalentes s'il existe** un automorphisme u de E tel que l'on ait

$$q'(x) = q(u(x)) \qquad \text{pour tout } x \in \mathrm{E}.$$

On suppose que le corps de base soit **R**; montrer que, pour que q et q' soient équivalentes, il faut et il suffit qu'elles aient même signature.

26. Montrer que les deux formes quadratiques suivantes sur $\mathbf{R}^3$ sont équivalentes, et construire un automorphisme de $\mathbf{R}^3$ transformant la première en la seconde :

$$2x^2 + 9y^2 + 3z^2 + 8xy - 4xz - 10yz$$
$$3x^2 + 3y^2 + 6z^2 - 4xy - 4xz + 8yz.$$

Même problème pour les formes

$$5x^2 + 5y^2 + 2z^2 + 8xy + 6xz + 6yz \qquad \text{et} \qquad 4x^2 + y^2 + 9z^2 - 12xz.$$

¶ **27.** On considère sur $\mathbf{R}^n$ la forme bilinéaire symétrique f correspondant à la forme quadratique

$$q(x) = x_1^2 + \cdots + x_p^2 - x_{p+1}^2 - \cdots - x_{p+q}^2$$

et on suppose $p \quad q$. Montrer que le sous-espace engendré par les vecteurs

$$e_1 + e_{p+1},\ e_2 + e_{p+2},\ \ldots,\ e_p + e_{2p},\ e_{p+q+1},\ \ldots,\ e_n$$

est totalement isotrope maximal pour f (*Exercice* 19).
En déduire que si f est une forme bilinéaire symétrique de signature (p, q) sur un espace vectoriel réel de dimension n, l'indice de f (*Exercice* 14) est $r + n - p - q$ où r est le plus petit des deux entiers p et q.
Quelle est la dimension des sous-espaces totalement isotropes maximaux de la forme de Lorentz?

¶ **28.** Soit f une forme bilinéaire symétrique non dégénérée sur un espace vectoriel réel E de dimension finie. Soient M et N deux sous-espaces vectoriels de E, tels que $\dim(\mathrm{M}) = \dim(\mathrm{N})$. Pour qu'il existe un automorphisme de f appliquant M sur N, il faut et il suffit que les restrictions de f à M et N aient la même signature (utiliser le théorème de Witt et les *Exercices* 24 et 25)
On prend pour f la forme de Lorentz et on considère, sur l'ensemble de tous les sous-espaces vectoriels de E, la relation d'équivalence « il existe un automorphisme u de f tel que $u(\mathrm{M}) = \mathrm{N}$ ». Combien d'éléments l'ensemble quotient comporte-t-il?
Même question pour la forme quadratique

$$x^2 + y^2 + z^2 - t^2 - u^2$$

sur $\mathbf{R}^5$.

29. Pour qu'une matrice hermitienne complexe $H \in M_n(\mathbf{C})$ soit positive, il faut et il suffit qu'il existe une matrice $X \in M_n(\mathbf{C})$ telle que

$$H = X^*X;$$

si de plus H est réelle on peut supposer X réelle [utiliser l'*Exercice* 11, *c*), et noter que les termes diagonaux de D sont positifs si H est positive].
En déduire que si

$$H = (a_{ij})_{1 \leqslant i,j \leqslant n}$$

est une matrice hermitienne complexe positive, on a

$$D_p = \begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix} \geqslant 0 \quad \text{pour } 1 \leqslant p \leqslant n$$

et que

$$D_p = 0 \quad \text{implique} \quad D_{p+1} = \cdots = D_n = 0$$

(Montrer d'abord que $D_n \geqslant 0$, puis remplacer H par la matrice dont D_p est le déterminant, et utiliser l'*Exercice* 16). Cas $n = 2$? (Retrouver l'inégalité de Cauchy-Schwarz, et le « signe du trinôme ».)

¶ **30.** Soit

$$H = (h_{ij})_{1 \leqslant i,j \leqslant n}$$

une matrice hermitienne complexe positive. Montrer qu'il existe une matrice triangulaire complexe

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix}, \text{ avec } t_{ii} \text{ réel positif}$$

telle que

$$H = T^*T$$

(utiliser les *Exercices* 15 et 16). Si H est inversible, T est unique.
Déduire de ce qui précède que pour toute matrice hermitienne positive H on a l'inégalité

$$0 \leqslant \det(H) \leqslant h_{11}h_{22}\ldots h_{nn};$$

si H est inversible, on ne peut, de plus, avoir l'égalité que si H est diagonale.

¶ **31.** Soit A une matrice carrée inversible à coefficients complexes. Montrer qu'il existe une matrice unitaire U et une matrice triangulaire $T = (t_{ij})$, avec $t_{ii} > 0$ pour tout i, telles que

$$A = U.T,$$

et que U et T sont uniques (appliquer l'*Exercice* précédent à A^*A). Montrer que U et T sont réelles si A est réelle.

¶ **32.** Soit $A = (a_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice complexe; montrer qu'on a

$$|\det(A)|^2 \leqslant \prod_{j=1}^{j=n} (|a_{1j}|^2 + \cdots + |a_{nj}|^2).$$

¶ **33.** Pour qu'une matrice hermitienne complexe soit définie positive, il faut et il suffit que tous ses mineurs principaux soient strictement positifs (utiliser l'*Exercice* 30 et calculer les t_{ii} en fonction des mineurs principaux de la matrice donnée).

34. Trouver, à l'aide de l'*Exercice* précédent, les valeurs réelles de t pour lesquelles les formes quadratiques suivantes sont définies positives :

$$5x_1^2 + x_2^2 + tx_3^2 + 4x_1x_2 - 2x_1x_3 - 2x_2x_3$$
$$2x_1^2 + x_2^2 + 3x_3^2 + 2tx_1x_2 + 2x_1x_3$$
$$2x_1^2 + 2x_2^2 + x_3^2 + 2tx_1x_2 + 6x_1x_3 + 2x_2x_3$$
$$t(x_1^2 + x_2^2 + x_3^2) + 4x_1x_2 + 6x_1x_3 + 8x_2x_3$$

¶¶ **35.** Montrer que toute matrice hermitienne positive est somme de matrices de la forme

$$\begin{pmatrix} c_1\bar{c}_1 & \dots & c_1\bar{c}_n \\ \dots\dots\dots & & \\ c_n\bar{c}_1 & \dots & c_n\bar{c}_n \end{pmatrix},$$

et réciproquement.
En déduire que si deux matrices complexes $(a_{ij})_{1\leqslant i,\,j\leqslant n}$ et $(b_{ij})_{1\leqslant i,\,j\leqslant n}$ sont hermitiennes positives, il en est de même de la matrice

$$(a_{ij}b_{ij})_{1\leqslant i,\,j\leqslant n}$$

obtenue en multipliant les termes de mêmes indices des deux matrices données.

36. Soit H une matrice hermitienne complexe. Montrer que $1 - iH$ est inversible, que

$$U = (1 + iH)(1 - iH)^{-1} = (i - H)(i + H)^{-1}$$

est unitaire, et que -1 n'est pas valeur propre de U. Inversement, toute matrice unitaire U dont -1 n'est pas valeur propre peut s'obtenir de cette façon **(transformation de Cayley)**.

¶ **37.** Soit H une matrice complexe hermitienne positive. Montrer (en réduisant H à la forme diagonale à l'aide d'une matrice unitaire) qu'il existe une et une seule matrice hermitienne positive H' telle que

$$H = H'^2.$$

(On dit que H' est la **racine carrée positive** de H, et on écrit $H' = H^{\frac{1}{2}}$).

¶ **38.** Soient S et T deux matrices hermitiennes complexes; on suppose S définie positive. Montrer que les valeurs propres de ST sont réelles, et même positives si T est positive (utiliser l'*Exercice* 37).

¶ **39.** Soit A une matrice carrée inversible à coefficients complexes; montrer que la matrice

$$A^*A$$

est hermitienne et définie positive. En considérant sa racine carrée positive, montrer qu'on peut écrire

$$A = U.H$$

avec U unitaire, et H hermitienne définie positive; montrer de plus que le problème n'admet qu'une solution.
Montrer que, si A est réelle, il en est de même de U et de H.

¶¶ **40.** Soit $F \subset M_n(C)$ un ensemble de matrices normales deux à deux permutables. Montrer qu'il existe une matrice unitaire $U \in M_n(C)$ telle que UAU^{-1} soit diagonale pour toute $A \in F$. Interprétation géométrique? (Utiliser l'*Exercice* 22 du § 34).

Dans les *Exercices* **41** à **47** on demande de ramener la matrice symétrique réelle donnée à une matrice diagonale à l'aide d'une matrice orthogonale réelle.

41. $\begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}$ **42.** $\begin{pmatrix} 11 & 8 & 2 \\ 8 & 5 & -10 \\ 2 & -10 & 2 \end{pmatrix}$ **43.** $\begin{pmatrix} 8 & 4 & -1 \\ 4 & -7 & 4 \\ -1 & 4 & 8 \end{pmatrix}$

44. $\begin{pmatrix} 17 & -2 & -2 \\ -2 & 14 & -4 \\ -2 & -4 & 14 \end{pmatrix}$ **45.** $\begin{pmatrix} 0 & 1 & -3 & 0 \\ 1 & 0 & 0 & -3 \\ -3 & 0 & 0 & 1 \\ 0 & -3 & 1 & 0 \end{pmatrix}$

46. $\begin{pmatrix} 5 & -5 & 1 & 3 \\ -5 & 5 & 3 & 1 \\ 1 & 3 & 5 & -5 \\ 3 & 1 & -5 & 5 \end{pmatrix}$ **47.** $\begin{pmatrix} 9 & 0 & 0 & 0 \\ 0 & 5 & 4 & -2 \\ 0 & 4 & 5 & 2 \\ 0 & -2 & 2 & 8 \end{pmatrix}$

¶ **48.** Soient E un espace vectoriel réel de dimension finie, f une forme bilinéaire symétrique définie positive sur E, et u un automorphisme de f. On suppose que les seuls sous-espaces vectoriels de E stables par u sont E et $\{0\}$. Montrer qu'on est alors dans l'une des situations suivantes : (*i*) E est de dimension *un* et u est égal à $+1$ ou à -1 ; (*ii*) E est de dimension *deux*, et il existe une base orthonormale pour f par rapport à laquelle la matrice de u est de la forme

$$\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}$$

où φ est un nombre réel non multiple de π.

¶ **49.** Soient E un espace vectoriel réel de dimension finie, f une forme bilinéaire symétrique définie positive sur E, et u un automorphisme de f. Montrer qu'il existe une base de E, orthonormale relativement à f, et par rapport à laquelle la matrice de u est de la forme

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & \ddots & & \\ & & & & & -1 & \\ & & & & & & \begin{matrix} \cos\varphi_1 & -\sin\varphi_1 \\ \sin\varphi_1 & \cos\varphi_1 \end{matrix} \\ & & & & & & & \ddots \\ & & & & & & & & \begin{matrix} \cos\varphi_h & -\sin\varphi_h \\ \sin\varphi_h & \cos\varphi_h \end{matrix} \end{pmatrix}$$

où les φ_i sont des nombres réels non multiples de π.

50. Pour tout nombre réel t, on pose

$$U(t) = \begin{pmatrix} \cos t & \sin t & 0 \\ -\sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad V(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{pmatrix}.$$

Montrer que, pour toute matrice orthogonale réelle X d'ordre 3 et de déterminant 1 (matrice d'une rotation dans l'espace usuel) il existe des nombres réels φ, ψ, θ tels que

$$X = U(\varphi)V(\theta)U(\psi).$$

(Observer que si l'on a deux bases orthonormales i, j, k et u, v, w, on passe de la première à la seconde en effectuant d'abord une rotation autour de k de façon à amener i dans le plan engendré par u et v, puis une rotation autour de l'intersection des plans ij et uv de façon à amener k sur w, et enfin une rotation autour de w). Les nombres φ, ψ et θ sont appelés les **angles d'Euler** de la matrice X (ou de la rotation correspondante).

qq **51.** On considère la matrice hermitienne

$$S = \begin{pmatrix} 1_p & 0 \\ 0 & -1_q \end{pmatrix}$$

d'ordre $p + q$ et de signature (p, q). Montrer que le groupe des automorphismes de S, i.e. le groupe des matrices $W \in GL(p + q, C)$ telles que

$$W^*SW = S,$$

est l'ensemble des matrices

$$W = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

possédant la propriété suivante : il existe une matrice complexe Z à p colonnes et q lignes, telle que la matrice

$$1_p - Z^*Z$$

soit définie positive, et deux matrices unitaires U et V, d'ordres p et q, telles que l'on ait les relations

$$A = (1 - Z^*Z)^{-\frac{1}{2}}U, \qquad B = Z^*(1 - ZZ^*)^{-\frac{1}{2}}V$$
$$C = Z(1 - Z^*Z)^{-\frac{1}{2}}U, \qquad D = (1 - ZZ^*)^{-\frac{1}{2}}V,$$

et de plus U, V et Z sont entièrement déterminées par W. Si W est réelle, il en est de même de U, V et Z. [NB — On pose

$$H^{-\frac{1}{2}} = \left(H^{\frac{1}{2}}\right)^{-1}$$

lorsque H est définie positive].

q **52.** Soit V un espace vectoriel de dimension paire $n = 2m$ sur un corps commutatif k.

a) Soit f une forme bilinéaire alternée (§ 22) non dégénérée sur V. On choisit deux vecteurs a et b tels que $f(a, b) = 1$ et on considère le plan $P \subset V$ engendré par a et b. Montrer que la restriction de f à P est non dégénérée, et en déduire de V est somme directe de P et du sous-espace V' des vecteurs orthogonaux à P pour f. En déduire que V admet une base par rapport

à laquelle l'expression de $f(x, y)$ en fonction des coordonnées de x et y est

$$f(x, y) = \sum_{i=1}^{i=m} x_i y_{2m-i} \qquad x_i y_{2m-i}$$

En déduire que deux formes bilinéaires alternées non dégénérées sur V peuvent être transformées l'une en l'autre par un automorphisme de V. Existe-t-il des formes bilinéaires alternées non dégénérées sur un espace de dimension impaire ?

b) On choisit une forme bilinéaire alternée non dégénérée f sur V, et on appelle plan non dégénéré de V tout sous-espace P de dimension 2 de V sur lequel f n'est pas identiquement nulle. Soit $S\rho(V)$ le groupe des automorphismes u de V tels que

$$f(u(x), u(y)) = f(x, y)$$

quels que soient $x, y \in V$ (« groupe symplectique »). Montrer que $S\rho(V)$ opère transitivement sur l'ensemble X des plans non dégénérés, et que si $P \in X$ le sous-groupe de $S\rho(V)$ qui laisse fixe P (i.e. l'ensemble des $u \in S\rho(V)$ tels que $u(P) = P$) est isomorphe au produit

$$SL_1(2, k) \times S\rho(W),$$

où W est l'orthogonal de P dans V par rapport à f.

c) On suppose maintenant que k est fini à q éléments. Soit x un élément non nul de V. Montrer que x appartient à q^{n-2} plans non dégénérés. En déduire la formule

$$\text{Card}(X) \quad \begin{matrix} q^n \\ q^2 \end{matrix} \quad \begin{matrix} \\ \\ \end{matrix} \quad \begin{matrix} 1 \\ 1 \end{matrix} \quad q^{n-2}.$$

Soient P un plan non dégénéré et W son orthogonal dans V. Montrer que l'on a

$$\text{Card}(S\rho(V)) = (q^{n-1} - 1) q^{n-1} \text{Card}(S\rho(W)).$$

En déduire, par récurrence sur n, la formule

$$\text{Card}(S\rho(V)) = q^{n(2n+1)} \prod_{i=1}^{i=n} (1 - 1/q^{2i}).$$

¶ **53.** Soit X un ensemble ayant 6 éléments. Soit Y l'ensemble des parties de X ayant 0 ou 2 éléments. On définit une loi de composition symétrique sur Y en posant :

$$\begin{array}{lll} A + \varnothing = A & \text{pour tout} & A \in Y \\ A + A = \varnothing & \text{pour tout} & A \in Y \\ A + B = A \cup B - A \cap B & \text{si} & A \cap B \text{ a un élément} \\ A + B = X - (A \cup B) & \text{si} & A \text{ et } B \text{ sont disjoints et non vides.} \end{array}$$

a) Montrer que cette loi de composition fait de Y un groupe commutatif, d'ordre 16 et d'élément neutre $\varnothing$. Montrer que Y peut être muni (de façon unique) d'une structure d'espace vectoriel sur le corps $k = Z/2Z$ et que sa dimension est égale à 4.

b) Si $A, B \in Y$, on désigne par $f(A, B)$ l'élément de k défini par la congruence

$$f(A, B) \equiv \text{Card}(A \cap B) \pmod 2.$$

Montrer que f est une forme bilinéaire alternée non dégénérée sur Y.

c) Soient A, B, C trois parties de X, disjointes, et ayant chacune deux éléments. Montrer que $\{\varnothing, A, B, C\}$ est un sous-espace totalement isotrope de dimension 2 de Y (relativement à la forme f).

d) Soit S_X le groupe des permutations de X ; c'est un groupe isomorphe au groupe symétrique

S_6. Soit d'autre part $Sp(Y)$ le groupe des automorphismes de Y qui conservent la forme f (ce groupe est souvent noté $Sp_4(F_2)$). Montrer que tout élément de S_X définit un élément de $Sp(Y)$, et que l'homomorphisme $\varepsilon : S_X \to Sp(Y)$ ainsi obtenu est un isomorphisme. (On montrera d'abord que ε est injectif, puis on comparera les ordres de S_X et $Sp(Y)$.)

e) Soit F le k-espace vectoriel des fonctions $f : X = k$ telles que $\sum_{x \in X} f(x) = 0$. Soit $V = F/\{0,1\}$ le quotient de F par le sous-espace de dimension 1 engendré par la fonction constante 1. A tout $A \in Y$, on associe sa fonction caractéristique θ_A (égale à 1 sur A et à 0 sur $X - A$). Montrer que $A \to \theta_A$ définit un *isomorphisme* de l'espace vectoriel Y sur l'espace vectoriel V. Montrer que cet isomorphisme transforme la forme bilinéaire f de *b*) en la forme

$$u(\theta, \theta') = \sum_{x \in X} \theta(x)\ \theta'(x).$$

BIBLIOGRAPHIE

A. A. ALBERT, *Introduction to algebraic Theories*, University of Chicago Press, 1941.

Polynômes, matrices, déterminants, équations linéaires, réduction des matrices (y compris le théorème de Jordan), généralités sur les anneaux et les corps. Exposé élémentaire et classique, très concentré et rigoureux. Exercices.

A.A. ALBERT, *Fundamental concepts of higher algebra*, University of Chicago Press, 1956.

Groupes, anneaux, corps, espaces vectoriels et matrices, extensions algébriques des corps commutatifs, corps finis. Niveau sensiblement plus élevé que l'ouvrage précédent, et destiné à des étudiants déjà avancés. Exercices.

E. ARTIN, *Geometric Algebra*, Interscience Publishers, New York, 1957.

E. ARTIN, *Algèbre Géométrique*, Gauthier-Villars, Paris, 1962.

Espaces vectoriels, dualité, équations linéaires, groupes, corps (exposés très rapides); construction axiomatique de la Géométrie élémentaire (affine et projective); étude géométrique des groupes orthogonal, symplectique et linéaire. A utiliser principalement comme complément au § 36 du présent ouvrage. Exposé de premier ordre commençant à un niveau extrêmement élémentaire, et aboutissant directement, dans le domaine traité, aux questions ouvertes.

G. BIRKHOFF and S. MACLANE, *A Survey of modern Algebra*, Macmillan, New York, 1941.

Couvre à peu près tous les sujets traités dans le présent ouvrage, et expose en outre la théorie des extensions algébriques des corps commutatifs (théorie de Galois, etc...); exposé très accessible et moderne, avec de nombreux exercices en général très simples.

N. BOURBAKI, *Eléments d'histoire des mathématiques*, Hermann, Paris, 1960.

La lecture de ce recueil des « notes historiques » ajoutées par N. Bourbaki aux divers volumes de son Traité n'est évidemment pas indispensable si l'on ne s'intéresse qu'aux aspects pratiques des Mathématiques. Par contre, ceux qui désirent comprendre pourquoi on a été conduit à introduire les notions exposées ici, les plus classiques comme les plus modernes, ne pourront mieux faire que d'étudier ces notes historiques. Lecture assez difficile, l'auteur ne se bornant pas à écrire en français basique.

I. M. GEL'FAND, *Leksii po linĕinoĭ algebre*, Gostekhizdat, Moscou, 1951.

I. M. GEL'FAND, *Lectures on linear Algebra*, Interscience Publishers, New York, 1961.

La majeure partie de ce livre expose, d'une façon en général plus détaillée (notamment en ce qui concerne le théorème de Jordan), les sujets traités dans les §§ 34, 35 et 36; la théorie des déterminants et des systèmes d'équations linéaires est supposée déjà connue du lecteur. Exposé évidemment conçu comme introduction à la théorie des espaces de Hilbert et à celle des groupes classiques (orthogonal, unitaire, linéaire, etc...). On peut

lui reprocher d'une part de ne pas suffisamment séparer les propriétés de nature purement linéaire de celles qui reposent sur l'emploi d'une forme quadratique, et d'autre part de négliger les corps autres que **R** et **C**. Style extrêmement clair.

W. GRAEUB, *Lineare Algebra*, Springer, Berlin, 1958.

Couvre les §§ 10 à 25 et 34 à 36 du présent ouvrage, d'une façon plus détaillée sur certains points (notamment en ce qui concerne les formes multilinéaires alternées). Excellent exposé — et la plus belle typographie du monde. Seconde édition en anglais.

P. R. HALMOS, *Finite-dimensional vector spaces*, Van Nostrand, Princeton, 1958.

Espaces vectoriels, dualité, espaces quotients, notions sur les produits tensoriels et les formes multilinéaires alternées, applications linéaires et matrices, valeurs propres, forme de Jordan, espaces euclidiens et matrices hermitiennes. Ce livre (d'abord publié en 1942) est visiblement conçu comme une introduction à la théorie des espaces de Hilbert (l'étude des formes quadratiques positives occupe un tiers environ du volume), et néglige volontairement d'autres aspects importants (systèmes d'équations linéaires, déterminants, diviseurs élémentaires). Exposé très simple et très géométrique, à peu près accessible au débutant. Exercices théoriques en général très faciles.

H. HASSE, *Höhere Algebra*, 2 vol., Walter de Gruyter (Sammlung Göschen), Berlin, 1951-1957.

H. HASSE, *Higher Algebra*, 2 vol., F. Ungar, New York, 1954.

Le premier volume traite des anneaux, corps, polynômes, groupes, équations linéaires (sur un corps non nécessairement commutatif), et de la théorie des déterminants; le second, de la division des polynômes, des extensions algébriques des corps commutatifs, et de la théorie de Galois. Dû à l'un des fondateurs de l'algèbre «abstraite», cet ouvrage (dont la première édition remonte à 1926-1927) n'est pas dépassé, et reste même nettement plus moderne, sur de nombreux points, que beaucoup de ceux qu'on a publiés depuis, y compris dans les dernières années; style d'une clarté et d'une précision exemplaires. Le seul inconvénient est un excès de concentration, imposé par le très faible volume des livres de la collection Göschen; la traduction américaine n'a pas supprimé cet inconvénient, mais a du moins permis l'emploi d'une typographie normalement aérée.

K. HOFFMAN and R. KUNZE, *Linear Algebra*, Prentice-Hall, Englewood Cliffs, 1961.

L'un des nombreux ouvrages élémentaires parus aux U.S.A. dans les dernières années. La plupart de ces livres sont d'un intérêt mathématique plutôt douteux, mais celui-ci est très supérieur à la moyenne par la rigueur de l'exposé et la correction du style. Destiné initialement aux Undergraduates du M.I.T., et tout à fait accessible au débutant français. Assez nombreux exercices. Couvre les §§ 10 à 24 et 34 à 36.

A. G. KUROŠ, *Kurs vysšeĭ algebry*, Fizmatgiz, Moscou, 1959.

Couvre tout le programme du présent ouvrage (sauf la théorie des ensembles) et va nettement plus loin sur certains points (démonstration du théorème de d'Alembert, racines réelles des équations algébriques à coefficients réels, fonctions symétriques des racines et élimination, structure des groupes commutatifs finis, etc...). Exposé dans l'ensemble assez classique, dans lequel les notions « modernes » (espaces vectoriels, corps, groupes) sont introduites lorsqu'on n'en a pratiquement plus besoin, ce que certains trouveront peut-être agréable.

S. LANG, *Algebra*, Addison-Wesley, Reading, 1965.

Groupes, anneaux, corps, polynômes, modules noethériens; théorie des corps (extensions algébriques et transcendantes, théorie de Galois, valuations); formes bilinéaires, théorème de Jordan, algèbre tensorielle, représentations des groupes finis. Bien que commençant à un niveau très élémentaire, l'exposé de Lang est évidemment à réserver pour ceux qui, après avoir à peu près assimilé le contenu du présent volume, désirent aller plus loin; la lecture de Lang leur permettra d'avancer alors à une vitesse vertigineuse. Nombreux exercices instructifs.

A. LENTIN et J. RIVAUD, *Leçons d'Algèbre moderne*, Vuibert, Paris, 1961.

Couvre une partie substantielle des sujets traités ici; plutôt destiné aux élèves des classes de Mathématiques Spéciales, mais utilisable également par les étudiants en Mathématiques Générales puisque les programmes d'Algèbre de ces deux enseignements sont maintenant très voisins l'un de l'autre. Nombreux exercices.

A. LICHNÉROWICZ, *Algèbre et Analyse linéaires*, Masson, Paris, 19.

Espaces vectoriels, calcul des matrices, équations linéaires, déterminants et formes alternées, valeurs propres, ainsi que plusieurs chapitres d'Analyse qui ne sont pas destinés aux débutants. Manque d'exemples et, plus encore, d'Exercices.

A. I. MAL'CEV, *Foundations of linear Algebra*, W. H. Freeman, San Francisco, 1963.

Matrices et déterminants, espaces vectoriels, théorème de Jordan, espaces euclidiens, formes quadratiques, transformations orthogonales, symplectiques, unitaires, etc..., formes multilinéaires, tenseurs, algèbre extérieure. Exercices. Excellent exposé, que le lecteur débutant ne pourra sans doute pas assimiler instantanément, mais qui lui sera utile durant plusieurs années.

L. MIRSKY, *An Introduction to linear Algebra*, Clarendon Press, Oxford, 1955.

Couvre les §§ 10 à 24 et 34 à 36, tout en étant nettement plus détaillé sur certains points (notamment en ce qui concerne le théorème de Jordan). Exposé très complet et correct, avec quelques détails étranges (par exemple l'introduction des « vector spaces », qui sont des parties de K^n, et des « linear manifold » qui sont les espaces vectoriels du présent ouvrage et de tous les mathématiciens), mais très utile à consulter. Très nombreux et intéressants exercices. Style et typographie prodigieusement britanniques.

G. D. MOSTOW, J. H. SAMPSON and J. P. MEYER, *Fundamental Structures of Algebra*, McGraw Hill, New York, 1963.

Programme et niveau à peu près identiques à celui du présent volume, avec quelques suppléments sur les équations différentielles linéaires et les tenseurs, et un accès peut-être plus facile au début. Exposé très clair, avec beaucoup d'exemples et d'exercices. Mis à part, bien entendu, le présent volume, nous paraît constituer la meilleure introduction d'ensemble à l'algèbre actuellement disponible sur le marché mondial.

O. SCHREIER und E. SPERNER, *Einführung in die analytische Geometrie und Algebra*, 2 vol., Vandenhoeck und Ruprecht, Göttingen, 1955.

O. SCHREIER and E. SPERNER, *Introduction to modern Algebra and Matrix theory*, Chelsea Publishing Cy., New York, 1951.

Espaces vectoriels R^n, équations linéaires, déterminants et formes multilinéaires alternées, formes quadratiques et déplacements, corps, polynômes, théorème de d'Alembert-Gauss, structure des groupes commutatifs de type fini, matrices, valeurs propres, théorème de Jordan. L'édition allemande comporte en outre un long chapitre sur la Géométrie projective, omis de la traduction américaine. L'ouvrage de Schreier et Sperner a été publié pour la première fois en 1931, et c'était, à l'époque, avec le petit livre de Hasse cité plus haut, le premier exposé élémentaire faisant systématiquement usage des méthodes « géométriques », et influencé par les travaux des mathématiciens allemands des années vingt. Ce qu'on a dit de Hasse, à savoir qu'il s'agit d'un ouvrage plus moderne que la plupart de ceux qui l'ont suivi, s'applique également à Schreier et Sperner. Style beaucoup plus agréable que Hasse (on ne cherche pas à économiser le papier). Lecture recommandée. Quelques exercices.

G. E. ŠILOV, *Vvedenie v teoriju linejnykh prostranstv*, Gostekhizdat, Moscou, 1956.

G. E. SHILOV, *An Introduction to the Theory of linear Spaces*, Prentice-Hall, Englewood Cliffs, 1961.

Déterminants, espaces vectoriels, équations linéaires, applications linéaires et matrices

formes bilinéaires et quadratiques, vecteurs propres (sans le théorème de Jordan), classification des surfaces du second degré, notions sur les espaces de Hilbert et les opérateurs complètement continus. Exposé très élémentaire et très clair, orienté, comme ceux de Gel'fand et de Halmos, vers l'Analyse et les espaces de Hilbert beaucoup plus que vers d'autres branches des Mathématiques. Seul inconvénient sérieux (qu'on retrouve du reste dans Gel'fand, Mirsky, et beaucoup d'exposés que nous n'avons pas cités) : l'utilisation de la théorie des déterminants pour établir les propriétés des espaces vectoriels de dimension finie (en particulier le fait que toutes les bases ont le même nombre d'éléments). Mais ce point, facile à rectifier si on le désire, ne concerne qu'une très faible partie de l'ouvrage. Lecture très recommandée au débutant.

B. L. van der WAERDEN, *Moderne Algebra*, 2 vol., Springer, Berlin, 1955.

B. L. van der WAERDEN, *Modern Algebra*, 2 vol., F. Ungar, New York, 1950.

Cet ouvrage célèbre, dont la première édition remonte à 1930, a pendant longtemps été le seul exposé d'ensemble de l'algèbre « moderne ». Style extraordinairement clair et concis. Niveau beaucoup plus élevé que les ouvrages précédents. Destiné uniquement aux étudiants avancés désirant s'engager dans la recherche mathématique.

O. ZARISKI and P. SAMUEL, 2 vol., *Commutative Algebra*, D. Van Nostrand Cy., Princeton, 1958.

Expose tout le matériel nécessaire en Arithmétique « supérieure » (corps de nombres algébriques) et en Géométrie Algébrique : extensions de corps, anneaux noethériens, anneaux de Dedekind, anneaux de valuation, etc... Accessible, comme le précédent, au lecteur ayant assimilé le contenu du présent ouvrage, et désirant se spécialiser.

RECUEILS D'EXERCICES

Certains des ouvrages indiqués ci-dessus contiennent des énoncés d'exercices, principalement Birkhoff-MacLane, Hoffman-Kunze, et Mirsky. Le livre de Hasse est accompagné d'un recueil d'Exercices (avec solutions) dans le style « Algèbre abstraite » et qu'on recommande aux futurs mathématiciens :

H. HASSE und W. KLOBE, *Aufgabensammlung zur Höheren Algebra*, Walter de Gruyter (Sammlung Göschen), Berlin, 1952.

H. HASSE and W. KLOBE, *Exercises to Higher Algebra*, F. Ungar, New York, 1954.

Les deux meilleures sources d'Exercices actuellement disponibles en ce qui concerne les calculs pratiques et effectifs — on y trouvera peu d'énoncés théoriques — sont de loin les suivantes :

D. K. FADEEV i I. S. SOMINSKJI, *Sbornik zadač po vyššej algebre*, Fizmatgiz, Moscou, 1961.

D. K. Faddeev and I. S. Sominskii, *Problems in Higher Algebra*, W. H. Freeman, San Francisco, 1965.

I. V. PROSKURJAKOV, *Sbornik zadač po linejnoj algebre*, Gostekhizdat, Moscou, 1962.

Le premier recueil (980 énoncés, avec les solutions détaillées s'il y a lieu) couvre à peu près tous les sujets traités ici; le second (1753 énoncés, avec les réponses et de très brèves explications lorsqu'il y a lieu) couvre seulement, mais en beaucoup plus grand détail, l'algèbre linéaire. La plupart des Exercices « pratiques » qu'on trouvera dans le présent ouvrage ont été extraits des deux recueils précédents (surtout du second).

En langue française, on trouvera naturellement de nombreux énoncés d'Exercices dans les cours de Mathématiques Spéciales (par exemple dans Lentin et Rivaud). D'autre part, l'ouvrage suivant, destiné aux étudiants de Mathématiques Générales, comporte une centaine d'énoncés d'Algèbre, avec des solutions détaillées :

G. LAFORET, *Algèbre et Analyse. Exercices*, Dunod, Paris, 1961.

INDEX DES NOTATIONS

Pour les notations introduites dans le texte, on renvoie au § et au n° où elles apparaissent pour la première fois. Pour les notations définies dans les Exercices, on indique le § et, entre parenthèses, le n° de l'Exercice où elles sont définies.

INDEX TERMINOLOGIQUE

Pour les termes définis dans le texte, on renvoie au § et au n° où ils apparaissent pour la première fois. Pour les termes définis dans les Exercices, on indique le § et, entre parenthèses, le n° de l'Exercice où ils apparaissent pour la première fois.

IMPRIMÉ EN FRANCE. DURAND, 28-LUISANT, 1963
TROISIÈME TIRAGE JOSEPH FLOCH, MAYENNE, 1969
DÉPOT LÉGAL TROISIÈME TRIMESTRE 1969
NUMÉRO D'ÉDITION : 2358
HERMANN, ÉDITEURS DES SCIENCES ET DES ARTS

Roger Godement

Cours d'algèbre

Cet ouvrage fondamental contient une mine d'exercices sans égale dans les ouvrages similaires, français ou autres. Il constitue un bagage minimal de tout étudiant et de tout enseignant en mathématiques et, à ce titre, reste irremplaçable. Dégagé du langage superficiel et sophistiqué qui sévit ailleurs, il survit et survivra aux modes.

Le raisonnement logique. Les relations d'égalité et d'appartenance. La notion de fonction. Réunions et intersections. Relations d'équivalence. Ensembles finis et nombres entiers. Lois de composition. La notion de groupe. Anneaux et corps. Nombres complexes. Modèles et espaces vectoriels. Relations linéaires dans un module. Applications linéaires. Matrices. Addition des homomorphismes et matrices. Produits de matrices. Matrices inversibles et changements de base. Transposée d'une application linéaire. Sommes de sous-modules. Théorèmes de finitude. La notion de dimension. Systèmes d'équations linéaires. Fonctions multilinéaires. Applications bilinéaires et trilinéaires alternées. Applications multilinéaires alternées. Développement d'un déterminant. Formules de Cramer. Variétés linéaires affines. Relations algébriques. Anneaux de polynômes. Fonctions polynomiales. Corps des fractions d'un anneau d'intégrité. Fractions rationnelles. Dérivation des polynômes et fractions rationnelles. Formule de Taylor. Anneaux principaux. Propriétés de divisibilité des polynômes. Nombre de racines d'une équation algébrique. Vecteurs propres et valeurs propres. Forme canonique d'une matrice. Formes hermitiennes. – Exercices.

ISBN 2 7056 5241 8

9 782705 652418

42 euros

HERMANN ÉDITEURS