Algèbre : le grand combat

Grégory Berhuy

Table des matières

Avant-propos

Petit plaidoyer totalement partial pour l'étude de l'algèbre.–Nous voudrions commencer cet ouvrage par un petit plaidoyer en faveur de cette discipline mal-aimée des étudiants qu'est l'algèbre. Les étudiants se plaignent souvent de sa difficulté à cause de son haut niveau d'abstraction. Bien souvent, les objets manipulés ne leur paraissent pas assez concrets, et l'utilité de l'étude des structures algébriques semblant absconses leur échappe complètement. Pourquoi donc s'embêter à classifier les groupes finis d'ordre donné ou chercher toutes leurs représentations irréductibles, pourquoi s'acharner à savoir si tel groupe est résoluble ou si tel anneau est factoriel, pourquoi calculer la réduite de Jordan d'un endomorphisme et autres joyeusetés ? Pourquoi un tel besoin quasi-irrépressible chez le mathématicien de généraliser à l'extrême, sinon pour torturer les étudiants par pur plaisir sadique [1] ?

Avant de tenter de répondre à ces questions, essayons de définir ce qu'est l'algèbre. La définition la plus juste est sans doute la suivante : c'est l'étude des structures algébriques. Ici, on entend par structure algébrique un ensemble muni d'une ou plusieurs lois de composition internes ou externes, éventuellement muni d'un ordre ou une topologie, le tout satisfaisant un certain nombre d'axiomes. Ainsi, les groupes, les anneaux, les corps, les espaces vectoriels ou les représentations sont des structures algébriques. Une fois les structures définies, on définit souvent la notion de sous-structure, mais surtout les flèches entre deux telles structures, appelées « morphismes », qui sont « les applications compatibles avec les lois internes et externes ». Ce sont d'ailleurs les morphismes qui constituent la part importante la théorie, plus que les objets eux-mêmes, puisque ce sont eux qui établissent les relations entre les objets, à travers la notion de structure quotient, par exemple. Une notion cruciale découlant directement de la notion de morphisme est la notion d'isomorphisme. Elle permet d'identifier des objets a priori différents mais structurellement identiques, et jouissant alors de propriétés similaires à tout point de vue. De manière un peu vague, deux structures isomorphes sont en fait deux façons de voir le même objet, mais dont on aurait étiqueté les éléments différemment. Par exemple, tous les espaces vectoriels de dimension n sur un même corps sont isomorphes, et une propriété vraie pour l'un est aussi vraie pour tous les autres, même si leurs éléments sont de natures différentes. D'ailleurs, après le choix d'une base, tout espace vectoriel de dimension n se comporte comme K^n : l'étiquetage des éléments se fait via les coordonnées dans la base choisie, mais le comportement des éléments vis-à-vis des opérations est exactement le même que celui des vecteurs de K^n muni de ses opérations usuelles. Un des problèmes fondamentaux en algèbre est alors de classifier les objets à isomorphismes près. Pour résumer, ce qui compte, ce n'est pas tant les objets, mais les flèches entre les objets [2].

1. D'aucuns diront que la réponse est contenue dans la question.
2. Ce point de vue a d'ailleurs été poussé à l'extrême et avec succès à travers la théorie des catégories.

Passons maintenant aux difficultés d'appréhender cette théorie. En ce qui concerne le problème du manque de palpabilité des objets, cela provient en grande partie d'un manque de pratique de l'abstraction, les étudiants étant confrontés aux objets abstraits et aux joies de la démonstration de théorèmes formels extrêmement tardivement. Autrement dit, pour maîtriser une nouvelle notion/discipline, il faut s'exercer. Ce principe frappé au coin du bon sens s'applique bien évidemment dans la vie de tous les jours (pratique d'un sport, d'un instrument, d'une langue, de la danse,etc.). Il n'y a aucune raison pour les mathématiques y échappe. Ceci dit, la difficulté avec l'algèbre, et de manière plus générale avec les mathématiques, est que les définitions et théorèmes exposés en cours ou dans les ouvrages sont le fruit d'un long cheminement intellectuel, qui s'est déroulé parfois sur plusieurs dizaines d'années. Par conséquent, on ignore souvent le problème concret qui a motivé l'introduction de telle ou telle notion à l'origine. Par exemple, l'algèbre linéaire trouve essentiellement sa source dans la géométrie. C'est Descartes qui fût le premier à utiliser des coordonnées pour résoudre des problèmes géométriques comme la détermination de l'intersection de deux droites, en termes d'équation linéaire, établissant dès lors un pont entre deux branches mathématiques jusqu'alors séparées : l'algèbre et la géométrie. Il faut attendre le dix-neuvième siècle et Gauss pour qu'une étude approfondie des systèmes linéaires soit faite. Camille Jordan résout également définitivement le problème de la réduction d'endomorphisme. Il faut noter que la notion d'espace vectoriel n'a toujours pas été dégagée à ce moment là ! Une première tentative de formalisation de cette notion a été effectuée par Grassmann en 1844, dans son traité « La théorie de l'extension linéaire ». Ce traité, passé relativement inaperçu, contient déjà l'essentiel de la théorie telle que nous la pratiquons aujourd'hui. Enfin, en 1888, Peano, en se basant sur les travaux de Grassmann, axiomatise définitivement la théorie. On pourrait constater la même chose avec d'autres branches de l'algèbre. La théorie des groupes trouve ainsi son origine première dans le problème de résolubilité des équations polynômiales par radicaux et les travaux de Galois. La notion de groupe est également utilisée par Klein pour étudier les nouvelles géométries ayant émergé (géométrie hyperbolique, géométrie projective). Kummer introduit également un groupe abélien dans ses tentatives de démonstration du grand théorème de Fermat. Néanmoins, il faudra attendre 1882 pour qu'émerge la première définition axiomatique d'un groupe abstrait. Notons au passage que l'étude des anneaux prend également sa source dans les travaux de Dedekind sur le théorème de Fermat. C'est finalement Noether qui établira les fondements de la théorie des anneaux commutatifs unitaires.

Vu la diversité des contextes dans lesquels toutes ces notions sont apparues historiquement, on comprend mieux pourquoi la mise en place une théorie générale est indispensable. D'ailleurs, établir une théorie unificatrice possède plusieurs avantages : éclairer sous un nouveau jour les cas particuliers qui ont motivé son élaboration, obtenir des résultats valables pour tous les objets de la théorie considérée (ce qui évite de redémontrer la même chose dans des dizaines de cas particuliers), mettre à disposition de nouveaux outils plus puissants pouvant être utiles pour démontrer de nouveaux théorèmes, créer des passerelles entre diverses branches des mathématiques, etc. Le prix à payer est souvent une montée dans l'abstraction, mais c'est un mal nécessaire pour pouvoir utiliser une artillerie efficace. Une image qui fonctionne plutôt bien est celle du randonneur complètement perdu : afin de retrouver plus facilement son chemin, il va chercher à prendre de la hauteur et à atteindre un point de vue plus élevé, pour avoir accès à un panorama complet des alentours, avant de redescendre pour revenir sur le bon chemin. Un des exemples les plus frappants est celui du grand théorème de Fermat : pour arriver finalement à une démonstration, il a fallu étudier les structures algébriques d'objets dont on aurait eu peine à imaginer

l'existence au temps de Fermat. C'est ainsi que la théorie des anneaux, la théorie des nombres et de nouveaux pans de la géométrie algébrique se sont développés. Comme dans d'autres situations en mathématiques, le fait d'intégrer le problème de Fermat dans un cadre plus général et apparemment beaucoup plus difficile a permis de grandes avancées, parce que l'on dispose alors de tout un outillage développé pour ce cadre.

Essayons finalement de répondre à la question que beaucoup d'étudiants se posent légitimement : à quoi ça sert ? Mentionnons tout de suite la principale force de l'algèbre, à savoir son ubiquité. Il est en effet intéressant de souligner que les diverses branches de l'algèbre ne sont pas compartimentées. Bien au contraire, il existe des passerelles entre elles et vers d'autres disciplines, ce qui permet une variété d'applications dans divers domaines.

La théorie des groupes illustre parfaitement notre propos. La notion de groupe apparaît naturellement dès qu'il s'agit d'étudier les symétries d'un objet. Elle intervient ainsi en géométrie puisque beaucoup d'exemples de groupes sont de nature géométrique, mais aussi en chimie (simplification des calculs de structure électronique, classification des états électroniques ou vibratoires d'une molécule ou d'un solide), en physique théorique (théorie de la relativité, théorie de l'unification) via la théorie des représentations. La théorie des groupes est aussi intimement liée à la théorie des corps via la théorie de Galois. Cette dernière théorie permet par exemple de donner une jolie caractérisation de la constructibilité d'un point à la règle et au compas, ou de la résolubilité d'une équation polynomiale par radicaux en termes de théorie des groupes. Une autre application de la théorie de Galois moins connue des étudiants (car plus pointue), mais omniprésente dans la recherche mathématique, est la résolution des problèmes dits « de descente ». De manière très floue, un problème de descente est un problème du type suivant : on considère un ensemble d'objets mathématiques et une relation d'équivalence sur ces objets. On se donne deux objets définis sur un corps K, et on suppose que les deux objets sont équivalents lorsque l'on les considère comme des objets définis sur un corps L contenant K. Les deux objets sont-il équivalents sur K ? Par exemple, considérons le problème suivant : soient $M, M' \in \mathrm{M}_n(K)$ deux matrices. On suppose qu'il existe une matrice $Q \in \mathrm{GL}_n(L)$ telle que $M' = QMQ^{-1}$, où L est un corps contenant K. Existe-t-il une matrice $P \in \mathrm{GL}_n(K)$ telle que $M' = PMP^{-1}$? Il est bien connu que la réponse est positive, et la démonstration la plus classique ne nécessite pas de théorie de Galois. On peut alors légitimement se poser la question de savoir si c'est encore le cas lorsque l'on remplace le groupe linéaire par le groupe spécial linéaire ou le groupe orthogonal. La réponse est cette fois négative, et la théorie de Galois permet de caractériser précisément les matrices M' pour lesquelles cela devient vrai, à M fixé. Enfin, dans un registre plus ludique, la théorie des groupes permet de résoudre certains casse-têtes célèbres comme le Rubik's cube ou le jeu du taquin.

De manière plus générale, la théorie des groupes se révèle un outil indispensable dans l'étude des structures mathématiques, quelles qu'elles soient. En effet, pour comprendre un objet mathématique, il est toujours utile d'étudier son groupe d'automorphismes. Il est aussi naturel d'essayer de faire agir un groupe sur cet objet et d'étudier les orbites et stabilisateurs associés. On obtient alors à la fois de plus amples renseignements sur l'objet lui-même et sur le groupe qui agit. Enfin, lorsque l'on s'attaque au problème de classification d'objets mathématiques à isomorphisme près, problème qui s'avère extrêmement ardu en général, il est toujours utile d'associer un groupe à chaque objet, de sorte que deux objets isomorphes aient des groupes associés isomorphes. Cette approche est souvent fructueuse, surtout lorsque

le groupe associé est abélien, puisque les groupes abéliens sont des objets particulièrement bien compris et dans lesquels il est facile de calculer. Ce principe est illustré de manière extrêmement frappante en topologie. À tout espace topologique X et tout point $x \in X$, on peut associer un groupe abélien $H_k(X, X \setminus \{x\})$ pour tout entier $k \geq 0$, de sorte que, si $h : X \xrightarrow{\sim} Y$ est un homéomorphisme, on a un isomorphisme de groupes

$$h_* : H_k(X, X \setminus \{x\}) \xrightarrow{\sim} H_k(Y, Y \setminus \{h(x)\}).$$

Ces groupes peuvent être calculés dans certains cas, grâce à des théorèmes généraux. Par exemple, on peut démontrer que $H_k(\mathbb{R}^n, \mathbb{R}^n \setminus \{x\})$ est trivial si $k \neq n$, et est isomorphe à $\mathbb{Z}$ si $k = n$, et ceci pour tout $x \in \mathbb{R}^n$. On en déduit aisément le résultat non trivial suivant, dont il serait difficilement concevable (voire impossible) d'avoir une démonstration à la main : pour tous entiers $n, m \geq 1$, les espaces topologiques $\mathbb{R}^n$ et $\mathbb{R}^m$ sont homéomorphes si, et seulement si, $n = m$.

L'algèbre linéaire quant à elle s'invite aussi bien dans les autres branches de l'algèbre qu'en analyse (équations différentielles ou aux dérivées partielles, espaces fonctionnels), en mécanique, en théorie des graphes, en théorie du codage (codes linéaires auto-correcteurs), dans l'étude des systèmes dynamiques discrets, en probabilités (chaînes de Markov), en informatique,etc. C'est aussi un excellent outil pour étudier les groupes, grâce à la théorie des représentations linéaires. Par exemple, la table des caractères permet de déterminer tous ses groupes distingués ou de décider de la simplicité d'un groupe, ou de démontrer des théorèmes de structure. D'autre part, comme déjà mentionné plus haut, la théorie des représentations a des applications en chimie et en physique théorique.

Enfin, la théorie des anneaux possède un pouvoir d'ubiquité similaire à la théorie des groupes ou de l'algèbre linéaire. Cette théorie trouve son origine en théorie des nombres, avec les travaux de Dedekind visant à démontrer le théorème de Fermat, le point de départ étant que si $x, y, z \in \mathbb{Z}$ et p est un nombre premier impair, l'égalité $x^p + y^p = z^p$, peut se retraduire dans l'anneau $\mathbb{Z}[\zeta_p] = \{P(\zeta_p) \mid P \in \mathbb{Z}[X]\}$ sous la forme

$$(x + y)(x + \zeta_p y) \cdots (x + \zeta_p^{p-1} y) = z^p,$$

où $\zeta_p = e^{\frac{2i\pi}{p}}$. La notion d'anneau intervient également de manière déterminante en géométrie algébrique, mais aussi en logique formelle (algèbres de Heyting, de Boole, algèbres cylindriques, etc.), en théorie des représentations linéaires (si V est une représentation irréductible d'un groupe G, l'anneau des endomorphismes de V est un anneau à division), en algèbre linéaire (théorie des polynômes annulateurs). Récemment, la théorie des anneaux à division a trouvé des applications en théorie des codes Wifi, et a servi à construire des codes performants pour la transmission de données sans fil (téléphone portable, box Wifi). Enfin, la théorie des anneaux a permis d'étudier globalement les formes quadratiques sur un corps K, en introduisant l'anneau de Witt $W(K)$, ce qui a insufflé un nouvel essor à la théorie des formes quadratiques et a contribué à la découverte de nombreux résultats.

De manière générale, l'utilisation des structures algébriques permettent souvent de démontrer des résultats de manière simple et élégante, une fois que l'on a les bons outils (comme la résolution d'un problème de constructibilité à la règle et au compas, qui est purement géométrique, grâce à la théorie des corps et des groupes, ou comme la résolution d'une question topologique difficile à coup de groupes d'homologie). On pourrait continuer indéfiniment la liste des applications des notions présentées dans cet ouvrage. Nous espérons que ce panorama réduit suffira à convaincre le lecteur du bien-fondé de l'étude des structures algébriques et de la beauté de cette théorie.

Prérequis.– Le texte qui suit est un cours d'algèbre s'adressant à des lecteurs ayant acquis les notions mathématiques de base de première et deuxième années de Licence. Nous supposerons donc que le lecteur maîtrise quelques notions simples de théorie des ensembles, l'arithmétique dans $\mathbb{Z}$, et surtout l'algèbre linéaire et matricielle de base. Le lecteur ne se sentant pas au point sur tout ou partie de ces notions est vivement encouragé à d'abord consolider ses connaissances sur les points précédents avant de s'attaquer à ce cours.

Quelques conseils pour travailler son cours.– Les conseils qui suivent sont évidemment valables pour n'importe quel cours, qu'ils soit suivi en amphi ou qu'il soit étudié en autodidacte à l'aide d'un livre. Chaque étudiant est bien entendu différent, et il n'y a pas de recette universelle pour maîtriser un cours de mathématiques. Néanmoins, il y a quand même quelques conseils frappés au coin du bon sens [3] qui peuvent faciliter l'apprentissage.

(1) S'aménager des conditions de travail propices (que ce soit seul ou en groupe) : couper télé, radio, téléphone portable, internet, afin de pouvoir mieux se concentrer.

(2) Apprendre les définitions et les théorèmes principaux à la virgule près (chaque mot et l'ordre de chaque mot ayant son importance dans un énoncé mathématique). Connaître les exemples classiques des objets que l'on a défini.

(3) Savoir faire les exercices d'application directe du cours.

(4) S'attaquer aux exercices moins évidents, quitte à sécher dessus un bon moment. On apprend les maths en les faisant (et on éprouve généralement une grande satisfaction personnelle lorsqu'on trouve finalement). Il est absolument inefficace d'abandonner au bout de dix minutes. Pour progresser, il faut être patient et tenace.

(5) Une fois ceci fait, et seulement à ce moment là, on peut reprendre chaque énoncé important du cours et essayer de répondre (dans la mesure du possible) aux questions suivantes : pourquoi a-t-on fait telle ou telle hypothèse ? Où intervient-elle dans la démonstration ? Le résultat est-il encore vrai si l'on affaiblit les hypothèses, ou si l'on en enlève ? Si ce n'est pas le cas, peut-on produire un contre-exemple ?

C'est très bien pour comprendre les choses en profondeur.

(6) Si possible, essayer de comprendre les étapes essentielles et l'articulation d'une démonstration d'un théorème important. Savoir refaire une démonstration courte.

Il faut prendre conscience que l'on ne peut pas tout comprendre à fond sur le moment, et accepter que l'on puisse ne pas maîtriser une démonstration tout de suite. Il faut parfois un an de recul sur une notion pour vraiment la digérer.

(7) Souvent, lorsqu'on lit une démonstration du cours ou la solution d'un exercice, on a le sentiment d'avoir bien compris. Il faut savoir que neuf fois sur dix, il n'en est rien. Pour voir si c'est vraiment le cas, nous conseillons au lecteur d'essayer de ré-expliquer la dite démonstration ou solution à des collègues. C'est un test qui ne trompe pas : s'il y avait des points obscurs ou doutes, cela se ressentira lors de la tentative d'explication. Comme on dit souvent : ce qui se conçoit bien s'énonce clairement.

Comment aborder un exercice.– Enfonçons le clou : on apprend les maths en les faisant, et en se trompant, et certainement pas en lisant linéairement un ouvrage,

3. Tout du moins selon l'auteur...

une page Web ou une solution rédigée d'exercice. Voici quelques principes de base pour éviter de rester bloqué devant un exercice.

(1) Lire l'énoncé en entier pour identifier le but de l'exercice [4].

(2) Identifier de quoi ça parle, et lister les théorèmes à disposition sur le sujet.

(3) Repérer les hypothèses, ainsi que les mots du style « en déduire ». Cela donne une bonne indication sur la manière de procéder.

(4) Lorsque que l'on doit faire une démonstration, ne pas hésiter à écrire les hypothèses, et la conclusion, pour visualiser d'où on part et où on veut arriver. Cela aide souvent pour savoir comment démarrer.

(5) Lorsque on rédige une démonstration, il faut savoir justifier chaque étape. Si ce n'est pas le cas, c'est qu'il y a un problème...

(6) Bien faire apparaître où on utilise les hypothèses de l'énoncé dans la rédaction d'une démonstration. Si une hypothèse n'est pas utilisée, il y a deux possibilités : soit l'énoncé est mal fait (ce qui peut arriver), soit il y a une ou plusieurs erreurs dans la démonstration en question (ce qui est le cas le plus probable).

(7) Se forcer à bien rédiger et à écrire des phrases mathématiquement complètes et grammaticalement correcte **même au brouillon.** Il est extrêmement facile de prendre de mauvaises habitudes de rédaction.

Parmi les principes de bonne rédaction, on peut citer : ne pas utiliser d'abréviation, ne pas mélanger symboles mathématiques et français, n'écrire un signe « $\iff$ » que si l'on est certain d'avoir une équivalence logique, dire où on utilise les hypothèses, vérifier que l'on n'utilise pas la conclusion souhaitée pour la démontrer, etc.

(8) Faire des exercices dont on ne dispose pas du corrigé, pour ne pas être tenté d'aller y jeter un coup d'œil [5].

(9) Comme déjà évoqué précédemment, essayer d'expliquer sa proposition de solution d'un exercice à des collègues permet de déceler des erreurs ou des points mal rédigés.

4. Dit comme ça, le conseil a l'air complètement idiot, mais c'est en fait vital pour ne pas perdre de vue le but final !

5. Cela ne sert à rien, et ce n'est pas rentable.

Première partie

Groupes : il faut agir ! !

Chapitre I

Théorie des ensembles, structures algébriques

I.1. Applications

Dans ce paragraphe, E et F seront deux ensembles non vides, et $f : E \longrightarrow F$ désignera une application.

Définition I.1.1. On dit que $f : E \longrightarrow F$ est *injective* si pour tous $x, x' \in E$, on a

$$f(x) = f(x') \Longrightarrow x = x'.$$

Autrement dit, $f : E \longrightarrow F$ est injective si pour tout $y \in F$, l'équation

$$f(x) = y, \ x \in E$$

possède *au plus* une solution.

On dit que $f : E \longrightarrow F$ est *surjective* si pour tout $y \in F$, l'équation

$$f(x) = y, \ x \in E$$

possède *au moins* une solution.

On dit que $f : E \longrightarrow F$ est *bijective* si elle est injective et surjective. Autrement dit, $f : E \longrightarrow F$ est bijective si pour tout $y \in F$, l'équation

$$f(x) = y, \ x \in E$$

possède *exactement une* solution. Dans ce cas, il existe une unique application [1] $g : F \longrightarrow E$ vérifiant $f \circ g = \mathrm{Id}_F$ et $g \circ f = \mathrm{Id}_E$.

On la note f^{-1}.

Définition I.1.2. Si $E' \subset E$, *l'image (directe)* de E' par f est l'ensemble

$$\begin{array}{rcl} f(E') & = & \{f(x') \mid x' \in E'\} \\ & = & \{y \in F \mid \text{ il existe } x' \in E' \text{ tel que } y = f(x')\}. \end{array}$$

L'image de f est l'ensemble $f(E)$. On le note aussi $\mathrm{Im}(f)$.

Si $F' \subset F$, *l'image réciproque* de F' par f est l'ensemble

$$f^{-1}(F') = \{x \in E \mid f(x) \in F'\}.$$

Attention ! La notation est trompeuse, car f n'est pas nécessairement bijective. On trouve parfois la notation $f^{\langle -1 \rangle}(F')$ pour éviter la confusion.

Remarques I.1.3.

(1) On a $\mathrm{Im}(f) = F$ si, et seulement si, f est surjective.

1. Pour tout $y \in F$, $g(y)$ est simplement l'unique solution $x \in E$ de l'équation $f(x) = y$.

(2) Si $f : E \longrightarrow F$ est une application, l'application

$$\begin{aligned} E &\longrightarrow \mathrm{Im}(f) \\ x &\longmapsto f(x) \end{aligned}$$

est surjective. En particulier, si une application $f : E \longrightarrow F$ est injective, elle induit une bijection de E sur son image.

I.2. Relations d'équivalence

La notion de relation d'équivalence sur un ensemble permet de mettre en relation des éléments ayant une propriété commune.

Cela permet alors regrouper ces éléments par « paquets », définissant ainsi la notion de classe d'équivalence, puis de construire un nouvel ensemble en « identifiant » les éléments d'un même paquet à un seul et même élément. On aboutit alors à la notion d'ensemble quotient.

Définition I.2.1. Soit E un ensemble. Une *relation* sur E est une partie R de $E \times E$. On note $x \sim_R y$ si $(x, y) \in R$.

Une *relation d'équivalence sur* E est une relation R satisfaisant les propriétés suivantes :

(1) réflexivité : pour tout $x \in E$, $x \sim_R x$;
(2) symétrie : pour tous $x, y \in E$, si $x \sim_R y$, alors $y \sim_R x$;
(3) transitivité : pour tous $x, y, z \in E$, si $x \sim_R y$ et $y \sim_R z$, alors $x \sim_R z$.

On confondra souvent R et $\sim_R$, et on parlera plutôt de la relation d'équivalence $\sim_R$ ou $\sim$.

Si $x \in E$, *la classe d'équivalence de* x est l'ensemble

$$\overline{x} = \{y \in E \mid y \sim_R x\}.$$

C'est le sous-ensemble des éléments en relation avec x. En particulier, $x \in \overline{x}$ pour tout $x \in E$.

Remarque I.2.2. Si E est vide, alors R est vide. Il n'y a alors qu'une classe d'équivalence, qui est la classe vide. Si E est non vide, alors R est non vide (puisque R contient (x, x) pour tout $x \in E$). Dans ce cas, chaque classe d'équivalence est non vide.

Exemples I.2.3.

(1) La relation d'égalité est une relation d'équivalence sur tout ensemble E, associée à la partie $R = \{(x, x) \mid x \in E.\}$ De plus, pour tout $x \in E$, on a

$$\overline{x} = \{x\}.$$

(2) La relation « avoir la même parité » est une relation d'équivalence sur l'ensemble $\mathbb{Z}$. Il y a deux classes d'équivalence : l'ensemble P des entiers pairs, et l'ensemble I des entiers impairs.

Définition I.2.4. L'ensemble $E/_\sim$ dont les éléments sont les diverses classes d'équivalence de E pour la relation $\sim$ est appelé *ensemble quotient* de E par $\sim$.

Attention ! Les éléments de $E/_\sim$ sont des sous-ensembles de E.

Exemple I.2.5. Si $\mathbb{Z}$ est muni de la relation d'équivalence « avoir même parité », alors l'ensemble quotient a deux éléments, qui sont P et I.

Proposition I.2.6. *Soit E un ensemble non vide, que l'on suppose muni d'une relation d'équivalence $\sim$. Alors :*

(1) *pour tous $x, y \in E$, on a $x \sim y \iff \overline{x} = \overline{y}$;*
(2) *deux classes d'équivalence sont soit disjointes, soit égales ;*
(3) *les classes d'équivalence forment une partition de E. Autrement dit, les classes d'équivalence sont non vides, deux à deux disjointes et de réunion E.*

Démonstration. Si $\overline{x} = \overline{y}$, alors $x \in \overline{x} = \overline{y}$, et donc $x \sim y$. Inversement, supposons que $x \sim y$. Soit $z \in \overline{x}$. Alors $z \sim x$, et puisque $x \sim y$, on a $z \sim y$. Ainsi $z \in \overline{y}$, et donc $\overline{x} \subset \overline{y}$. Mais puisque $x \sim y$, on a $y \sim x$ et le même raisonnement montre que $\overline{y} \subset \overline{x}$, et donc $\overline{x} = \overline{y}$.

Montrons maintenant le deuxième point. Soient $\overline{x}$ et $\overline{y}$ deux classes d'équivalence et supposons qu'elles ne soient pas disjointes. Autrement dit, il existe $z \in E$ tel que $z \in \overline{x}$ et $z \in \overline{y}$. Par définition, on a $x \sim z$ et $y \sim z$. On a alors $x \sim z$ et $z \sim y$, et par suite $x \sim y$. Le premier point implique alors que $\overline{x} = \overline{y}$.

Enfin, pour montrer le dernier point, il reste à montrer que les classes d'équivalence sont non vides, de réunion E, le point (2) montrant que ces classes sont deux à deux disjointes. Mais c'est clair puisque l'on a $x \in \overline{x}$ pour tout $x \in E$. □

Corollaire I.2.7. *Soit $\omega \in E/_\sim$ une classe d'équivalence pour la relation d'équivalence $\sim$. Alors, pour tout $x \in \omega$, on a $\omega = \overline{x}$.*

Démonstration. Par définition d'une classe d'équivalence, on a $\omega = \overline{y}$ pour un certain $y \in E$. Si $x \in \omega = \overline{y}$, alors $x \sim y$, et donc $\overline{x} = \overline{y} = \omega$ par la proposition **??** (1). □

Définition I.2.8. Soit $\omega \in E/_\sim$ une classe d'équivalence pour la relation d'équivalence $\sim$. Un élément $x \in E$ tel que $\omega = \overline{x}$ est appelé un *représentant* de ω. Par le corollaire précédent, $x \in E$ est un représentant de ω si, et seulement si, $x \in \omega$.

Exemple I.2.9. Si $E = \mathbb{Z}$ est muni de la relation d'équivalence « avoir même parité », 0 et 2 sont des représentants de l'ensemble P des entiers pairs.

Définition I.2.10. Soit E un ensemble non vide muni d'une relation d'équivalence $\sim$. L'application surjective

$$\begin{aligned} \pi \colon E &\longrightarrow E/_\sim \\ x &\longmapsto \overline{x} \end{aligned}$$

est appelée la *projection canonique.*

On veut maintenant caractériser l'ensemble des applications

$$\varphi : E/_\sim \longrightarrow F,$$

où F est un ensemble. Remarquons tout d'abord que si $\varphi : E/_\sim \longrightarrow F$ est une application, alors l'application $\varphi \circ \pi : E \longrightarrow F$ est constante sur les classes d'équivalence.

En effet, si $x \sim y$, alors $\overline{x} = \overline{y}$ et donc

$$(\varphi \circ \pi)(x) = \varphi(\pi(x)) = \varphi(\overline{x}) = \varphi(\overline{y}) = \varphi(\pi(y)) = (\varphi \circ \pi)(y).$$

Réciproquement, on a le résultat suivant.

Proposition I.2.11. *Soit E un ensemble non vide, muni d'une relation d'équivalence $\sim$, soit F un ensemble non vide, et soit $f : E \longrightarrow F$ une application. On suppose que f est constante sur les classes d'équivalence, c'est-à-dire*

$$\textit{pour tous } x, y \in E, x \sim y \Longrightarrow f(x) = f(y).$$

Alors, il existe une unique application $\overline{f} : E/\sim \longrightarrow F$ *telle que*

$$f = \overline{f} \circ \pi.$$

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\pi"'] & F \\ E/\sim \arrow[ur, "\overline{f}"'] & \end{tikzcd}$$

Elle est définie par

$$\overline{f}(\omega) = f(x) \quad \textit{pour tout } \omega \in E/_\sim,$$

où $x \in E$ *est un représentant arbitraire de* ω.

En particulier, on a

$$\overline{f}(\overline{x}) = f(x) \quad \textit{pour tout } x \in E.$$

En outre, pour tout ensemble non vide F*, il existe une correspondance bijective entre l'ensemble des applications* $\varphi : E/_\sim \longrightarrow F$ *et l'ensemble des applications* $f : E \longrightarrow F$ *qui sont constantes sur les classes d'équivalence. Cette bijection est donnée par*

$$\varphi \longmapsto \varphi \circ \pi$$

$$\overline{f} \longleftarrow\!\shortmid f$$

Démonstration. Gardons les notations de l'énoncé. Supposons qu'une telle application $\overline{f}$ existe, et soit $\omega \in E/_\sim$. Si $x \in E$ est un représentant de ω, on a $\omega = \overline{x} = \pi(x)$, et ainsi

$$\overline{f}(\omega) = \overline{f}(\pi(x)) = (\overline{f} \circ \pi)(x) = f(x).$$

Cela montre que si $\overline{f}$ existe, elle est définie par

$$\overline{f}(\omega) = f(x) \text{ pour tout } \omega \in E/_\sim,$$

où $x \in E$ est un représentant arbitraire de ω. En particulier, une telle application est alors unique.

Le problème est de vérifier que réciproquement cette formule définit bien une application. Autrement dit, il faut vérifier que $\overline{f}(\omega)$ ne dépend pas du représentant de la classe d'équivalence choisi. Soient $x, y \in E$ deux représentants de ω. On a donc $x, y \in \omega$. On doit montrer alors que $f(x) = f(y)$. Mais c'est exactement l'hypothèse faite sur f. Ceci démontre l'existence de l'application $\overline{f}$. De plus, pour tout $x \in E$, x est un représentant de $\overline{x}$, et on a donc

$$\overline{f}(\overline{x}) = f(x) \quad \text{pour tout } x \in E.$$

Pour achever la démonstration du théorème, il suffit de vérifier que les deux flèches

$$\varphi \longmapsto \varphi \circ \pi$$

$$\overline{f} \longleftarrow\!\shortmid f$$

sont inverses l'une de l'autre.

Soit $\varphi : E/\sim \longrightarrow F$ une application. Alors, l'application $\overline{\varphi \circ \pi}$ est l'unique application $\psi : E/\sim \longrightarrow F$ telle que $\psi \circ \pi = \varphi \circ \pi$. Or, φ est une telle application. On en déduit alors que $\overline{\varphi \circ \pi} = \varphi$. Inversement, si $f : E \longrightarrow F$ est constante sur les classes d'équivalence, on a $\overline{f} \circ \pi = f$ par définition. Ceci achève la démonstration. $\square$

Remarque I.2.12. L'hypothèse sur f est indispensable !

Reprenons l'exemple de la relation d'équivalence « avoir la même parité » sur $\mathbb{Z}$, et soit $f = \mathrm{Id}_{\mathbb{Z}}$. Alors, f n'est pas constante sur les classes d'équivalence, et l'application $\overline{f}$ précédente n'existe pas. En effet, dans le cas contraire, on aurait $\overline{f}(P) = f(0) = 0$, mais aussi $\overline{f}(P) = f(2) = 2$, car les entiers 0 et 2 sont deux représentants de la classe d'équivalence formée par les entiers pairs. On obtiendrait alors l'égalité $0 = 2 \in \mathbb{Z}$. Ainsi, $\overline{f}$ n'est pas définie et les lignes précédentes n'ont aucun sens mathématique.

I.3. Lois internes, structures algébriques

Les structures sont les armes des mathématiciens[2].

Dans tout ce qui suit, E est un ensemble non vide.

Définition I.3.1. Une *loi interne* $*$ sur E est une application

$$\begin{aligned} E \times E &\longrightarrow E \\ (x, y) &\longmapsto x * y. \end{aligned}$$

On dit que la loi $*$:

(1) est *associative* si pour tous $x, y, z \in E$, on a $(x * y) * z = x * (y * z)$;
(2) est *commutative* si pour tous $x, y \in E$, on a $x * y = y * x$.

Si $\perp$ est une autre loi interne sur E, on dit que la loi $*$ est *distributive* par rapport à $\perp$ si, pour tous $x, y, z \in E$, on a

$$(x \perp y) * z = (x * z) \perp (y * z) \quad \text{et} \quad x * (y \perp z) = (x * y) \perp (x * z).$$

Exemple I.3.2. L'addition et le produit sont des lois associatives sur $\mathbb{N}$, ainsi que sur $\mathbb{Z}$. De plus, le produit est distributif par rapport à l'addition.

Définition I.3.3. Un *élément neutre* pour la loi $*$ est un élément $e \in E$ tel que, pour tout $x \in E$, on a $x * e = e * x = x$.

Exemple I.3.4. L'addition sur $\mathbb{N}$ possède un élément neutre, à savoir 0, et le produit sur $\mathbb{N}$ possède un élément neutre, à savoir 1.

Lemme I.3.5. *Soit E un ensemble non vide, muni d'une loi interne $*$. Si la loi $*$ possède un élément neutre, celui-ci est unique.*

Démonstration. Supposons que $*$ possède deux éléments neutres $e_1, e_2 \in E$. Puisque e_1 est un élément neutre, on a

$$e_1 * e_2 = e_2.$$

Mais, e_2 étant aussi un élément neutre, on a également

$$e_1 * e_2 = e_1.$$

On en déduit que $e_1 = e_2$, d'où le résultat. □

Définition I.3.6. Soit E un ensemble non vide, muni d'une loi $*$, et possédant un élément neutre e (nécessairement unique d'après le lemme précédent).

Soit $x \in E$. Un *symétrique* de x est un élément $x' \in E$ tel que

$$x * x' = x' * x = e.$$

2. Nicolas Bourbaki

Exemple I.3.7. Tout $m \in \mathbb{Z}$ possède un symétrique dans $\mathbb{Z}$ pour la loi d'addition, à savoir $-m$. En revanche, le seul entier naturel possédant un symétrique pour l'addition dans $\mathbb{N}$ est 0.

Lemme I.3.8. *Soit E un ensemble non vide, muni d'une loi interne $*$. On suppose que $*$ est associative et possède un élément neutre e. Si $x \in E$ possède un symétrique, celui-ci est unique.*

Démonstration. Supposons que $x \in E$ possède deux symétriques $x'_1, x'_2 \in E$. Alors, on a

$$x'_1 = x'_1 * e = x'_1 * (x * x'_2) = (x'_1 * x) * x_2 = e * x'_2 = x'_2,$$

d'où le résultat. □

On peut maintenant définir la notion de groupe.

Définition I.3.9. Un *groupe* est un couple $(G, *)$, où G est un ensemble non vide, et $* : G \times G \longrightarrow G$ est une loi interne associative, possédant un élément neutre $1_G \in G$, et telle que tout élément $x \in G$ possède un symétrique.

D'après les résultats du paragraphe précédent, le neutre est unique, et un symétrique de $x \in G$ est unique. On le note x^{-1}, et on l'appelle *l'inverse* de x.

Un groupe $(G, *)$ est dit *abélien* (ou commutatif) si la loi $*$ est commutative.

Les groupes seront étudiés de manière approfondie dans la suite. Aussi, nous nous bornerons pour l'instant à donner deux exemples.

Exemples I.3.10.

(1) L'ensemble $\mathbb{Z}$ des entiers relatifs, muni de son addition, est un groupe abélien, de neutre 0.
(2) L'ensemble $\mathbb{R}^\times = \mathbb{R} \setminus \{0\}$, muni du produit usuel, est un groupe abélien, de neutre 1.

Remarques I.3.11.

(1) Dans les énoncés, on omettra de préciser la loi $*$, et on parlera du groupe G, plutôt que du groupe $(G, *)$.
D'ailleurs, si $x, y \in G$, on écrira xy plutôt que $x * y$.
(2) Lorsque la loi de G est commutative, on la note souvent (mais pas toujours) « $+$ ». Dans ce cas, le neutre est noté 0_G, et l'inverse de x est plutôt noté $-x$, et appelé « opposé de x ».

Maintenant que l'on a défini les groupes, on va ajouter une loi supplémentaire, et définir la notion d'anneau. Encore une fois, nous nous limiterons pour l'instant au strict minimum.

Définition I.3.12. Un *anneau* est un triplet $(A, +, \cdot)$, où A est un ensemble non vide, $+ : A \times A \longrightarrow A$ et $\cdot : A \times A \longrightarrow A$ sont deux lois internes, vérifiant les propriétés suivantes :

(1) le couple $(A, +)$ est un groupe abélien ;
(2) la loi « $\cdot$ » est associative, et possède un élément neutre [3] ;

3. Selon certains auteurs, un anneau ne possède pas nécessairement un tel élément neutre, et appellent « anneau unitaire » ce que nous appelons un anneau. Néanmoins, notre convention est la plus répandue actuellement.

(3) la loi « $\cdot$ » est distributive par rapport à « $+$ ».

Les éléments neutres pour les lois « $+$ » et « $\cdot$ » seront notés 0_A et 1_A respectivement.

On dit que l'anneau $(A, +, \cdot)$ est *commutatif* si « $\cdot$ » est commutative.

Exemples I.3.13.

(1) L'ensemble $\mathbb{Z}$ des entiers relatifs, muni de son addition et de son produit, est un anneau commutatif.
(2) L'ensemble $\mathrm{M}_n(\mathbb{R})$ des matrices carrées de taille n à coefficients réels, muni de l'addition et du produit de matrices, est un anneau, non commutatif si $n \geq 2$.
(3) (TD) Soit $n \geq 0$ un entier. On rappelle que deux entiers $a, b \in \mathbb{Z}$ sont *congrus modulo* n si $a - b \in \mathbb{Z}$ est un multiple de n, i.e. s'il existe un entier $k \in \mathbb{Z}$ tel que $a = b + kn$.

On le note $a \equiv b\ [n]$. La relation de congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$. L'ensemble quotient est noté $\mathbb{Z}/n\mathbb{Z}$.

On montre que les deux lois internes

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a} + \overline{b} = \overline{a+b} \end{aligned} \quad \text{et} \quad \begin{aligned} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a}\overline{b} = \overline{ab} \end{aligned}$$

sont bien définies, et munissent $\mathbb{Z}/n\mathbb{Z}$ d'une structure d'anneau commutatif. Si de plus $n \geq 1$, cet anneau possède n éléments, à savoir

$$\overline{0}, \overline{1}, \ldots, \overline{n-1}.$$

Nous aurons aussi besoin rapidement de la notion de corps.

Définition I.3.14. Un *corps* est un ensemble K, muni de deux lois internes

$$\begin{aligned} K \times K &\longrightarrow K \\ (x, y) &\longmapsto x + y \end{aligned} \quad \text{et} \quad \begin{aligned} K \times K &\longrightarrow K \\ (x, y) &\longmapsto xy, \end{aligned}$$

appelées respectivement addition et multiplication, vérifiant les propriétés suivantes :

(1) le triplet $(K, +, \cdot)$ est un anneau commutatif ;
(2) tout élément non nul de K possède un symétrique par rapport à la loi « $\cdot$ » ;
(3) on a $0_K \neq 1_K$.

Le symétrique de $x \in K$ par rapport à l'addition est noté $-x$, et appelé *l'opposé* de x. Le symétrique de $x \in K \setminus \{0_K\}$ par rapport à la multiplication est noté x^{-1}, et appelé *l'inverse* de x.

On notera $K^\times = K \setminus \{0_K\}$.

Remarque I.3.15. Remarquons qu'avec cette définition, un corps est en particulier un anneau commutatif. Ce n'est pas toujours le cas dans les ouvrages. Nous préférons néanmoins réserver le terme « corps » pour désigner ce que d'autres auteurs appelleraient « corps commutatif », et plutôt parler d'anneau à division, ou de corps gauche, lorsque le groupe $(K \setminus \{0_K\}, \cdot)$ n'est pas abélien [4].

Exemples I.3.16.

(1) Les ensembles $\mathbb{Q}, \mathbb{R}, \mathbb{C}$, munis de leurs lois usuelles, sont des corps.
(2) Pour tout nombre premier p, l'anneau $\mathbb{Z}/p\mathbb{Z}$ est un corps, que l'on note $\mathbb{F}_p$. On en verra la démonstration dans un chapitre ultérieur.

4. La convention utilisée ici l'est d'ailleurs aussi chez nos amis anglosaxons.

Chapitre II

Propriétés élémentaires des groupes

Dans ce chapitre, nous introduisons les définitions et résultats élémentaires de la théorie des groupes.

II.1. Généralités

Bien que nous ayons déjà introduit cette notion dans le chapitre **??** (cf. définition **??**), rappelons la définition d'un groupe.

Définition II.1.1. Un *groupe* est un couple $(G, *)$, où G est un ensemble non vide et $* : G \times G \longrightarrow G$ est une loi interne vérifiant les propriétés suivantes :

(1) la loi $*$ est associative : pour tous $x, y, z \in G$, on a $(x * y) * z = x * (y * z)$;

(2) la loi $*$ possède un élément neutre : il existe un élément $e \in G$ tel que l'on ait $x * e = e * x = x$ pour tout $x \in G$;

(3) tout élément possède un symétrique : pour tout $x \in G$, il existe $x' \in G$ tel que $x * x' = x' * x = e$.

D'après les lemmes **??** et **??**, l'élément neutre est unique, et un symétrique de $x \in G$ est unique. On le note x^{-1}, et on l'appelle *l'inverse* de x.

Un groupe $(G, *)$ est dit *abélien* (ou commutatif) si l'on a de plus

$$x * y = y * x \text{ pour tous } x, y \in G.$$

Passons aux exemples de groupes.

Exemples II.1.2.

(1) Un singleton $\{e\}$ peut être muni trivialement d'une structure de groupe en posant $e * e = e$.

(2) Les couples $(\mathbb{Z}, +), (\mathbb{R}, +), (\mathbb{R}^\times, \cdot), (\mathbb{C}, +), (\mathbb{C}^\times, \cdot)$ sont des groupes abéliens.

(3) Le couple $(\mathrm{GL}_n(\mathbb{C}), \cdot)$ est un groupe, non abélien si $n \geq 2$.

(4) Soit E un K-espace vectoriel, où K est un corps. Alors, l'ensemble des applications linéaires bijectives $f : E \longrightarrow E$ est un groupe, pour la composition des applications, noté $\mathrm{GL}(E)$.

(5) Si E est un K-espace vectoriel, $(E, +)$ est un groupe abélien, comme par exemple $(\mathrm{M}_n(\mathbb{C}), +)$ et $(\mathbb{C}^n, +)$.

(6) Le couple $(\mathbb{Z}, \cdot)$ n'est pas un groupe : par exemple, 2 n'a pas d'inverse.

(7) Soit E un ensemble. L'ensemble $\mathfrak{S}(E)$ des bijections de E dans lui-même, muni de la composition, est un groupe, appelé *groupe des permutations* de E ou *groupe symétrique* sur E. Il est non abélien dès que $|E| \geq 3$. Si $E = [\![1, n]\!]$, il est noté $\mathfrak{S}_n$.

(8) Si $(G,*)$ est un groupe et X est un ensemble, on note G^X l'ensemble des fonctions $f : X \longrightarrow G$. Pour toutes fonctions $f, g \in G^X$, on définit la fonction $f * g : X \longrightarrow G$ par

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

Alors, $(G^X, *)$ est un groupe (Quels sont le neutre et l'inverse d'un élément ?).

(9) Si $(G,*)$ est un groupe, l'ensemble des suites $(g_n)_{n\geq 0}$ d'éléments de G est un groupe. C'est un cas particulier du cas précédent, puisqu'une suite n'est rien d'autre qu'une fonction $f : \mathbb{N} \longrightarrow G$.

(10) Si $(G,*)$ et $(G',\bullet)$ sont deux groupes, l'ensemble $G \times G'$, muni de la loi

$$\begin{aligned}(G \times G') \times (G \times G') &\longrightarrow G \times G' \\ ((g_1, g_1'), (g_2, g_2')) &\longmapsto (g_1 * g_2, g_1' \bullet g_2')\end{aligned}$$

est un groupe. Le groupe obtenu s'appelle le *produit direct (externe)* de G et G' .

Par exemple, $(\mathbb{R}^2, +)$ est le produit direct $(\mathbb{R}, +) \times (\mathbb{R}, +)$.

Nous laissons le soin au lecteur de détailler les exemples.

Notations.

(1) Dans les énoncés, on omettra de préciser la loi $*$, et on parlera du groupe G, plutôt que du groupe $(G,*)$.

(2) Le produit de deux éléments x et y de G est en général noté xy, et le neutre est noté 1_G ou 1 s'il n'y a pas d'ambigüité.

Attention néanmoins, cette loi n'est pas forcément la multiplication au sens usuel ; c'est juste une notation commode lorsque l'on travaille avec une structure de groupe non spécifique. Lorsque G est abélien, on utilise aussi la notation additive « + ». Dans ce cas précis, le neutre est noté 0_G ou 0, et l'inverse de x est noté $-x$.

(3) Si $x_1, \ldots, x_n \in G$, le produit $x_1 \cdots x_n$, noté aussi $\prod_{i=1}^{n} x_i$, est défini par récurrence de la manière suivante :

(1) si $n = 0$, $x_1 \cdots x_n = 1_G$;

(2) si $n \geq 1$, on pose $x_1 \cdots x_n = (x_1 \cdots x_{n-1})x_n$.

Remarquons que par associativité, pour tous $k, n \geq 0$, on a

$$(x_1 \cdots x_k)(x_{k+1} \cdots x_n) = x_1 \cdots x_n.$$

(4) Si $n \in \mathbb{Z}$, on pose $x^n = x \cdots x$ si $n \geq 0$ et $x^n = x^{-1} \cdots x^{-1}$ si $n < 0$. En notation additive, on le note $n{\cdot}x$.

On a alors

$$x^{n+m} = x^n x^m \quad \text{et} \quad (x^n)^m = x^{nm} \quad \text{pour tous } n, m \in \mathbb{Z}.$$

Ces deux égalités sont faciles mais pénibles à démontrer. Leur vérification est lâchement laissée au lecteur.

Attention ! En général, $(xy)^n \neq x^n y^n$ si G n'est pas abélien.

Lemme II.1.3. *Soit G un groupe. Alors :*

(1) *la loi de groupe est simplifiable : pour tous $x, y, z \in G$, on a*

$$xz = yz \Longrightarrow x = y, \quad \textit{et} \quad zx = zy \Longrightarrow x = y$$

(2) *pour tous $x, x_1, \ldots, x_n \in G$, on a*

$$(x^{-1})^{-1} = x \quad \textit{et} \quad (x_1 \cdots x_n)^{-1} = x_n^{-1} \cdots x_1^{-1}$$

(3) *pour tout $a \in G$, les translations à gauche et à droite*

$$\begin{array}{rcl} \ell_a \colon G & \longrightarrow & G \\ x & \longmapsto & ax \end{array} \quad \textit{et} \quad \begin{array}{rcl} r_a \colon G & \longrightarrow & G \\ x & \longmapsto & xa \end{array}$$

sont bijectives.

Démonstration.

(1) Il suffit de multiplier par l'inverse et d'utiliser l'associativité.

(2) Puisque $x^{-1}x = xx^{-1} = 1_G$, x est l'unique inverse de x^{-1}, d'où le premier point. Pour le second, il suffit de traiter le cas $n = 2$ et de conclure par récurrence. On a

$$x_1x_2(x_2^{-1}x_1^{-1}) = x_1(x_2x_2^{-1})x_1^{-1} = x_1 1_G x_1^{-1} = x_1x_1^{-1} = 1_G,$$

et de même $(x_2^{-1}x_1^{-1})x_1x_2 = 1_G$.

(3) On vérifie que les applications inverses sont $\ell_{a^{-1}}$ et $r_{a^{-1}}$ respectivement.

□

Notation. Si E est un ensemble fini, on notera $|E|$ son nombre d'éléments.

Définition II.1.4. Si G est un groupe fini, son nombre d'éléments $|G|$ est appelé *l'ordre* de G.

Exemples II.1.5.

(1) Si E est un ensemble à n éléments, $\mathfrak{S}(E)$ est un groupe fini d'ordre $n!$.

(2) Si $n \geq 1$, le groupe abélien $\mathbb{Z}/n\mathbb{Z}$ est un groupe d'ordre n.

Maintenant que l'on a défini les objets que l'on veut étudier, il faut définir les « flèches » entre les objets, appelées morphismes. Évidemment, de tels morphismes n'ont d'intérêt que s'ils « respectent » la structure de groupe.

Définition II.1.6. Soient G, G' deux groupes.

Une application $f : G \longrightarrow G'$ est un *morphisme de groupes* si

$$f(xy) = f(x)f(y) \text{ pour tous } x, y \in G.$$

Attention ! La loi de groupe à droite est la loi de groupe dans G'. Il se peut très bien que G soit noté multiplicativement, et G' additivement, ou l'inverse.

Si $f : G_1 \longrightarrow G_2$ et $g : G_2 \longrightarrow G_3$ sont des morphismes de groupes, il en est de même de $g \circ f : G_1 \longrightarrow G_3$.

Exemples II.1.7. Les applications suivantes sont des morphismes de groupes :

(1) $\begin{array}{rcl} h_x \colon \mathbb{Z} & \longrightarrow & G \\ m & \longmapsto & x^m \end{array}$, où G est un groupe quelconque et $x \in G$;

(2) $\begin{array}{rcl} f \colon \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ n & \longmapsto & 2n \end{array}$;

(3) $\begin{array}{rcl} f : \mathbb{R}^\times & \longrightarrow & \mathbb{R}^\times \\ x & \longmapsto & x^3 \end{array}$;

(4) $\exp : \mathbb{R} \longrightarrow \mathbb{R}^\times$;

(5) $\ln :]0, +\infty[\longrightarrow \mathbb{R}$;

(6) $\det : \mathrm{GL}_2(\mathbb{C}) \longrightarrow \mathbb{C}^\times$.

En revanche, $\det : \mathrm{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ n'est **pas** un morphisme de groupes ;

(7) plus généralement, si K est un corps et E est un K-espace vectoriel de dimension finie, l'application $\det : \mathrm{GL}(E) \longrightarrow K^\times$ est un morphisme de groupes multiplicatifs ;

Encore une fois, on laisse au lecteur le soin de détailler les exemples.

La remarque suivante est très pratique et sera utilisée sans référence ultérieure.

Remarque II.1.8. Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors :

(1) on a $f(1_G) = 1_{G'}$;

(2) pour tout $x \in G$, on a $f(x^{-1}) = f(x)^{-1}$.

En effet, on a
$$f(1_G) = f(1_G \cdot 1_G) = f(1_G)f(1_G),$$
et on conclut par simplification dans G'. On a alors
$$f(x)f(x^{-1}) = f(xx^{-1}) = f(1_G) = 1_{G'},$$
et de même $f(x^{-1})f(x) = 1_{G'}$, donc $f(x^{-1})$ est l'inverse de $f(x)$.

Définition II.1.9. Un *isomorphisme* de groupes est un morphisme de groupes bijectif.

Deux groupes G et G' sont *isomorphes* s'il existe au moins un isomorphisme de groupes $f : G \longrightarrow G'$. On le note $G \simeq G'$.

Un *automorphisme* d'un groupe G est un isomorphisme de groupes de G sur lui-même. L'ensemble des automorphismes de G est noté $\mathrm{Aut}(G)$.

Remarques II.1.10.

(1) Pour tout groupe G, l'identité $\mathrm{Id}_G : G \longrightarrow G$ est un isomorphisme. On a donc $G \simeq G$ pour tout groupe G.

(2) Si $f : G \longrightarrow G'$ est un isomorphisme, alors $f^{-1} : G' \longrightarrow G$ est aussi un morphisme de groupes, et est donc un isomorphisme de groupes.

En effet, soient $x', y' \in G'$. On a alors
$$f(f^{-1}(x')f^{-1}(y')) = f(f^{-1}(x'))f(f^{-1}(y')) = x'y',$$
et donc $f^{-1}(x'y') = f^{-1}(x')f^{-1}(y')$ par définition de f^{-1}.

En particulier, si $G \simeq G'$, alors $G' \simeq G$.

(3) Si $f_1 : G_1 \longrightarrow G_2$ et $f_2 : G_2 \longrightarrow G_3$ sont des isomorphismes de groupes, il en est de même de $f_2 \circ f_1 : G_1 \longrightarrow G_3$.

Par conséquent, si $G_1 \simeq G_2$ et $G_2 \simeq G_3$, alors $G_1 \simeq G_3$.

La vérification est laissée au lecteur.

(4) Les trois points précédents entraînent aisément que l'ensemble $\mathrm{Aut}(G)$ est un groupe pour la composition des applications.

Exemples II.1.11. Les applications suivantes sont des isomorphismes de groupes :

(1) $\begin{array}{rcl} f\colon \mathbb{R}^\times & \longrightarrow & \mathbb{R}^\times \\ x & \longmapsto & x^3 \end{array}$;

(2) $\exp : \mathbb{R} \longrightarrow]0, +\infty[$;

(3) $\ln :]0, +\infty[\longrightarrow \mathbb{R}$;

(4) si E est un K-espace vectoriel de dimension n, le choix d'une base $\mathbf{e}$ induit deux isomorphismes de groupes

$$\begin{aligned} \mathrm{End}(E) &\xrightarrow{\sim} \mathrm{M}_n(K) \\ u &\longmapsto \mathrm{Mat}(u; \mathbf{e}). \\ \mathrm{GL}(E) &\xrightarrow{\sim} \mathrm{GL}_n(K) \\ u &\longmapsto \mathrm{Mat}(u; \mathbf{e}). \end{aligned}$$

Le lemme suivant donne toute une famille d'exemples d'automorphismes de groupes.

Lemme II.1.12. *Soit G un groupe, et soit $g \in G$. L'application*

$$\begin{array}{rcl} \mathrm{Int}_g\colon G & \longrightarrow & G \\ x & \longmapsto & gxg^{-1} \end{array}$$

est un automorphisme de G, appelé la conjugaison par g.

De plus, l'application

$$\begin{array}{rcl} \mathrm{Int}\colon G & \longrightarrow & \mathrm{Aut}(G) \\ g & \longmapsto & \mathrm{Int}_g \end{array}$$

est un morphisme de groupes.

Démonstration. Pour tous $x, y \in G$, on a

$$\mathrm{Int}_g(xy) = gxyg^{-1} = gxg^{-1}gyg^{-1} = \mathrm{Int}_g(x)\mathrm{Int}_g(y).$$

Ainsi, Int_g est un morphisme de groupes. De plus, pour tout $x \in G$, et pour tous $g, g' \in G$, on a

$$\mathrm{Int}_{gg'}(x) = gg'x(gg')^{-1} = gg'xg'^{-1}g^{-1} = g\mathrm{Int}_{g'}(x)g^{-1} = \mathrm{Int}_g(\mathrm{Int}_{g'}(x)).$$

Par conséquent, on a

$$\mathrm{Int}_g \circ \mathrm{Int}_{g'} = \mathrm{Int}_{gg'}.$$

On a en particulier

$$\mathrm{Int}_g \circ \mathrm{Int}_{g^{-1}} = \mathrm{Int}_{g^{-1}} \circ \mathrm{Int}_g = \mathrm{Int}_{1_G} = \mathrm{Id}_G,$$

et donc Int_g est bijectif, d'inverse $\mathrm{Int}_{g^{-1}}$. Le dernier point est clair, au vu des calculs précédents. □

Définition II.1.13. Un *automorphisme intérieur* d'un groupe G est un automorphisme de la forme $\mathrm{Int}_g, g \in G$. L'ensemble $\mathrm{Int}(G)$ des automorphismes intérieurs est un groupe pour la composition des applications.

Les automorphismes intérieurs jouent un rôle important en théorie des groupes (cf. la notion de sous-groupe distingué dans le paragraphe suivant). Nous finissons ce paragraphe en introduisant quelques notations et définitions.

Notation. Si P est une partie de G, pour tout $g \in G$, on note gPg^{-1} l'image de P par Int_g. Autrement dit, on a

$$gPg^{-1} = \{gxg^{-1} \mid x \in P\}.$$

Puisque Int_g est un automorphisme, gPg^{-1} est fini si, et seulement si, P est fini, et dans ce cas, on a

$$|gPg^{-1}| = |P|,$$

et ceci pour tout $g \in G$.

Définition II.1.14. Soit G un groupe, et soient deux parties P, P' de G. On dit que P et P' sont *conjuguées* dans G s'il existe $g \in G$ tel que $P' = gPg^{-1}$.

Deux éléments $x, x' \in G$ sont dits *conjugués* dans G s'il existe $g \in G$ (non nécessairement unique) tel que $x' = gxg^{-1}$. La *classe de conjugaison* de x dans G est l'ensemble $\mathrm{Conj}_G(x)$ des éléments de G conjugués à x, c'est-à-dire l'ensemble

$$\mathrm{Conj}_G(x) = \{gxg^{-1} \mid g \in G\}.$$

Le lecteur vérifiera sans peine que la relation « être conjugués dans G » est une relation d'équivalence sur G, dont les classes d'équivalences sont les classes de conjugaison d'éléments de G.

Remarque II.1.15. Déterminer toutes les classes de conjugaison d'un groupe est une question importante en théorie des groupes, mais difficile à résoudre. Par exemple, si $G = \mathrm{GL}_n(K)$, cela revient à savoir décider si deux matrices données M et M' sont semblables ou non.

Il y a quand même quelques groupes où ce problème peut être résolu facilement. C'est par exemple le cas du groupe symétrique, comme nous le verrons dans un chapitre ultérieur (cf. corollaire **??**).

II.2. Sous-groupes

Introduisons maintenant la notion de sous-groupe.

Définition II.2.1. Soit G un groupe, de neutre 1_G. Un *sous-groupe* de G est un sous-ensemble H de G vérifiant les propriétés suivantes :

(1) on a $1_G \in H$;
(2) pour tous $x, y \in H$, on a $xy \in H$;
(3) pour tout $x \in H$, on a $x^{-1} \in H$.

Clairement, $\{1_G\}$ et G sont des sous-groupes de G, pour tout groupe G. Un sous-groupe H sera dit *trivial* si $H = \{1_G\}$, *strict* si $H \neq G$ et *propre* s'il est distinct de $\{1_G\}$ et de G.

Exemples II.2.2.

(1) L'ensemble $\mathbb{Z}$ est un sous-groupe de $(\mathbb{C}, +)$.
(2) L'ensemble $\{\pm 1\}$ est un sous-groupe de $(\mathbb{R}^\times, \cdot)$.

L'intérêt de cette notion est donné par le lemme suivant.

Lemme II.2.3. *Soit G un groupe, et soit H un sous-groupe de G. Alors, la loi interne de G se restreint en une loi interne sur H, qui confère à H une structure de groupe, dont le neutre est celui de G.*

Démonstration. Le fait que la loi de G se restreigne en une loi interne sur l'ensemble H provient de l'axiome (2) de la définition de sous-groupe. Puisque G est un groupe, sa loi est associative. Sa restriction à H l'est donc aussi. De plus, 1_G étant un élément neutre pour la loi de G, c'est aussi un élément neutre pour sa restriction à H, car $1_G \in H$ par hypothèse sur H. Enfin, si $x \in H$, son symétrique x^{-1} est dans H. Mais alors, x^{-1} est un inverse de x pour la loi de H, puisque $xx^{-1} = x^{-1}x = 1_G = 1_H$. $\square$

Pour montrer qu'un ensemble muni d'une loi est un groupe, il est ainsi plus rapide de montrer que c'est un sous-groupe d'un groupe connu, l'intérêt étant d'éviter de vérifier l'associativité. De plus, le lemme suivant permet encore de raccourcir (parfois) les calculs.

Lemme II.2.4. *Soit G un groupe, et soit H une partie de G. Alors, H est un sous-groupe de G si, et seulement si, H est non vide, et pour tous $x, y \in H$, on a $xy^{-1} \in H$.*

Démonstration. Si H vérifie les conditions du lemme, puisque H est non vide, il contient au moins un élément $x_0 \in H$, et contient donc $x_0 x_0^{-1} = 1_G$. Pour tout $x \in H$, on a alors $x^{-1} = 1_G x^{-1} \in H$. Enfin, si $x, y \in H$, alors d'après le point précédent, $y^{-1} \in H$, et donc $xy = x(y^{-1})^{-1} \in H$. Le sens direct étant clair, ceci achève la démonstration. □

Exemples II.2.5.

(1) Il se peut fort bien qu'un groupe G n'ait aucun sous-groupe propre, comme $G = \{\pm 1\}$, par exemple.
(2) L'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-groupe du groupe des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ (pour l'addition des fonctions).
(3) L'ensemble des matrices triangulaires supérieures (resp. diagonales) est un sous-groupe de $(\mathrm{M}_n(\mathbb{C}), +)$.
(4) L'ensemble des matrices triangulaires supérieures inversibles est un sous-groupe de $(\mathrm{GL}_n(\mathbb{C}), \cdot)$.
(5) Si H est un sous-groupe de G et N est un sous-groupe de H, alors N est un sous-groupe de G.
(6) Si H_i est un sous-groupe de $G_i, i = 1, 2$, alors $H_1 \times H_2$ est un sous-groupe de $G_1 \times G_2$.
(7) La réunion de deux sous-groupes n'est pas un sous-groupe.
    Par exemple, $H_1 = \{(0, y) \mid y \in \mathbb{R}\}$ et $H_2 = \{(x, 0) \mid x \in \mathbb{R}\}$ sont des sous-groupes de $\mathbb{R}^2$, mais pas $H_1 \cup H_2$.

Remarque II.2.6. Il peut être ardu de décrire tous les sous-groupes d'un groupe G donné. Néanmoins, les cas de $\mathbb{Z}$ et de $\mathbb{Z}/n\mathbb{Z}$ sont bien connus, comme on le verra plus loin.

Le lemme suivant est utile pour construire de nouveaux sous-groupes.

Lemme II.2.7. *Soit $f : G \longrightarrow G'$ un morphisme de groupes.*

(1) *Si H est un sous-groupe de G, alors $f(H)$ est un sous-groupe de G'.*
(2) *Si H' est un sous-groupe de G', alors $f^{-1}(H')$ est un sous-groupe de G.*

Démonstration. Soit H un sous-groupe de G, et soit

$$f(H) = \{f(h) \mid h \in H\}.$$

Puisque H est non vide, $f(H)$ aussi. Pour tout $x, y \in H$, on a

$$f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}).$$

Comme H est un sous-groupe, on a $xy^{-1} \in H$, et donc $f(x)f(y)^{-1} \in f(H)$. Ainsi, $f(H)$ est un sous-groupe de G'.

Si maintenant H' est un sous-groupe de G', soit

$$f^{-1}(H') = \{x \in G \mid f(x) \in H'\}.$$

Puisque $f(1_G) = 1_{G'} \in H'$, on a $1_G \in f^{-1}(H')$. En particulier, $f^{-1}(H')$ n'est pas vide. Soient maintenant $x, y \in f^{-1}(H')$. Alors,

$$f(xy^{-1}) = f(x)f(y)^{-1} \in H',$$

car $f(x), f(y) \in H'$ et H' est un sous-groupe. Ainsi, $xy^{-1} \in f^{-1}(H')$, d'où le résultat. □

Définition II.2.8. Soit $f : G \longrightarrow G'$ un morphisme de groupes. Le *noyau* de f est le sous-groupe de G, noté $\mathrm{Ker}(f)$, défini par

$$\mathrm{Ker}(f) = f^{-1}(\{1_{G'}\}) = \{x \in G \mid f(x) = 1_{G'}\}.$$

L'image de f est le sous-groupe de G', noté $\mathrm{Im}(f)$, défini par

$$\mathrm{Im}(f) = f(G) = \{f(g) \mid g \in G\}.$$

Le lemme suivant donne une propriété très utile des morphismes de groupes.

Lemme II.2.9. *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors, f est injective si, et seulement si,* $\mathrm{Ker}(f) = \{1_G\}$.

Démonstration. Supposons f injective, et soit $x \in \mathrm{Ker}(f)$. On a alors

$$f(x) = 1_{G'} = f(1_G),$$

et par injectivité, on a $x = 1_G$. On a donc $\mathrm{Ker}(f) = \{1_G\}$. Inversement, supposons que $\mathrm{Ker}(f) = \{1_G\}$, et soient $x, y \in G$ tels que $f(x) = f(y)$. Alors, on a $f(x)f(y)^{-1} = f(xy^{-1}) = 1_{G'}$, et donc $xy^{-1} \in \mathrm{Ker}(f)$. Par hypothèse, $xy^{-1} = 1_G$, et donc $x = y$. Ainsi, f est injective. □

Nous allons maintenant définir une classe importante de sous-groupes.

Définition II.2.10. Soit G un groupe, et soit H un sous-groupe. On dit que H est *distingué dans* G si l'on a

$$ghg^{-1} \in H \quad \text{pour tout } h \in H, \text{ et tout } g \in G.$$

Autrement dit, un sous-groupe H est distingué dans G s'il est stable par tout automorphisme intérieur de G. On dit aussi que H est un sous-groupe distingué de G. Nous utiliserons indifféremment les deux expressions.

On le note $H \triangleleft G$.

Exemples II.2.11.

(1) Si G est abélien, tout sous-groupe de G est distingué dans G.

(2) Les sous-groupes $\{1_G\}$ et G sont distingués dans G.

Définition II.2.12. Soit G un groupe. Le *centre* d'un groupe est l'ensemble

$$Z(G) = \{z \in G \mid zg = gz \text{ pour tout } g \in G\}.$$

C'est un sous-groupe abélien de G, distingué dans G.

Remarque II.2.13. On a $Z(G) = G$ si, et seulement si, G est abélien.

Le lemme suivant fournit toute une famille d'exemples de sous-groupes distingués.

Lemme II.2.14. *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors,* $\mathrm{Ker}(f)$ *est un sous-groupe distingué de G.*

Démonstration. On sait déjà que c'est un sous-groupe de G. Soit $g \in G$, et soit $x \in \mathrm{Ker}(f)$. On a

$$f(gxg^{-1}) = f(g)f(x)f(g)^{-1} = f(g)1_{G'}f(g)^{-1} = 1_{G'},$$

et donc $gxg^{-1} \in \mathrm{Ker}(f)$, ce qu'il fallait vérifier. □

Remarque II.2.15. D'après ce lemme, pour démontrer qu'une partie H d'un groupe G est un sous-groupe distingué de G, on peut essayer de l'identifier au noyau d'un morphisme de groupes $f : G \longrightarrow G'$.

Exemple II.2.16. Soit E un K-espace vectoriel de dimension finie. L'ensemble $\mathrm{SL}(E)$ des automorphismes de l'espace vectoriel E et de déterminant 1 est un sous-groupe distingué de $\mathrm{GL}(E)$, puisque c'est le noyau du déterminant, qui est un morphisme de groupes.

Une notion centrale en théorie des groupes est celle de groupe simple. Les groupes simples sont, en quelque sorte, les briques élémentaires dont « tout groupe fini est constitué ».

Définition II.2.17. On dit qu'un groupe G est *simple* si $G \neq \{1_G\}$, et si G n'a pas de sous-groupe propre qui soit distingué dans G.

Autrement dit, G est simple si $G \neq \{1_G\}$, et si $\{1_G\}$ et G sont les seuls sous-groupes distingués de G.

Exemples II.2.18.

(1) Le groupe $\mathrm{GL}(E)$ n'est pas simple, puisque $\mathrm{SL}(E)$ est un sous-groupe propre distingué dans $\mathrm{GL}(E)$ (sauf cas exceptionnel).
(2) Le groupe $\{\pm 1\}$ est simple, comme le lecteur le vérifiera immédiatement.
(3) Un groupe non abélien dont le centre est non trivial n'est pas simple.
(4) On peut démontrer que les groupes abéliens simples sont les groupes $\mathbb{Z}/p\mathbb{Z}$, où p est premier, à isomorphisme près.

Nous verrons dans des chapitres ultérieurs des familles de groupes simples.

Remarque II.2.19. La classification des groupes finis simples à isomorphisme près est connue. Son obtention a été un travail de longue haleine, et la dernière version de la démonstration de cette classification nécessite environ cinq mille pages.

Nous allons maintenant étudier l'ordre des sous-groupes d'un groupe G. On commence par associer deux relations d'équivalence à un sous-groupe quelconque H de G.

Lemme II.2.20. *Soit G un groupe, et soit H un sous-groupe. Pour tout $a \in G$, on note*

$$aH = \{ah \mid h \in H\} \quad \textit{et} \quad Ha = \{ha \mid h \in H\}.$$

(1) *La relation $\sim_g$ sur G définie par*

$$x \sim_g y \textit{ si } x^{-1}y \in H$$

est une relation d'équivalence. De plus, on a

$$\overline{x} = xH \textit{ pour tout } x \in G.$$

(2) *La relation* $\sim_d$ *sur* G *définie par*

$$x \sim_d y \text{ si } yx^{-1} \in H$$

est une relation d'équivalence. De plus, on a

$$\overline{x} = Hx \text{ pour tout } x \in G.$$

Enfin, pour tous $x, y \in G$, *on a*

$$x \sim_g y \iff x^{-1} \sim_d y^{-1}.$$

Démonstration. On démontre le premier point, la démonstration du second point étant similaire. Pour tout $x \in G$, on a $x \sim_g x$ puisque $x^{-1}x = 1_G \in H$. Soient $x, y, z \in G$. Si $x^{-1}y \in H$, alors $(x^{-1}y)^{-1} = y^{-1}x \in H$. Autrement dit, si $x \sim_g y$, alors on a aussi $y \sim_g x$. Enfin, si $x^{-1}y, y^{-1}z \in H$, on a aussi $(x^{-1}y)(y^{-1}z) = x^{-1}z \in H$. Ainsi, si $x \sim_g y$ et $y \sim_g z$, alors $x \sim_g z$. L'égalité $\overline{x} = xH$ est claire.

Enfin, pour tous $x, y \in H$, on a

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

ce qui démontre la dernière partie. □

Définition II.2.21. Soit G un groupe, et soit H un sous-groupe. Un ensemble de la forme $xH, x \in G$ est appelé une *classe à gauche modulo* H. L'ensemble des classes à gauche modulo H est noté G/H.

Un ensemble de la forme $Hx, x \in G$ est appelé une *classe à droite modulo* H. L'ensemble des classes à droite modulo H est noté $H\backslash G$.

Lorsque G est abélien, classes à gauche et classes à droite coïncident. De plus, si G est noté additivement, la classe à gauche ou à droite de x est l'ensemble

$$x + H = \{x + h \mid h \in H\}.$$

Remarque II.2.22. Les classes à gauche (resp. les classes à droite) de G modulo H étant les classes d'équivalence pour une certaine relation d'équivalence sur G, on a immédiatement les propriétés suivantes :

(1) les classes à gauche (resp. à droite) de G modulo H forment une partition de G ;

(2) pour tous $x, y \in G$, on a

$$y \in xH \iff xH = yH \iff x^{-1}y \in H,$$

ainsi que

$$y \in Hx \iff Hx = Hy \iff yx^{-1} \in H.$$

En particulier, pour tout $x \in G$, on a les équivalences

$$xH = H \iff Hx = H \iff x \in H.$$

La dernière partie du lemme **??** implique facilement que l'application

$$\begin{aligned} G/H &\longrightarrow H\backslash G \\ xH &\longmapsto Hx^{-1} \end{aligned}$$

est bien définie et bijective.

En particulier, le nombre de classes à gauche modulo H est fini si, et seulement si, le nombre de classes à droite modulo H est fini. Dans ce cas, ces deux nombres sont égaux.

Définition II.2.23. Soit G un groupe, et soit H un sous-groupe. Le nombre de classes à gauche modulo H, lorsqu'il est fini, est appelé *l'indice de H dans G*, et est noté $[G : H]$. C'est aussi le nombre de classes à droite modulo H.

Si G/H est infini, on pose $[G : H] = +\infty$.

Remarque II.2.24. Lorsque G est fini, alors $[G : H]$ est fini pour tout sous-groupe H de G, puisqu'il y a au plus $|G|$ classes d'équivalence distinctes.

Exemple II.2.25. Soit $G = \mathbb{Z}$, et soit $H = n\mathbb{Z}, n \geq 0$. On vérifie aisément que H est un sous-groupe de G. Alors, la relation d'équivalence associée à H n'est rien d'autre que la relation de congruence modulo n, et la classe à gauche de $x \in G$ modulo H est simplement sa classe de congruence modulo n, i.e. $x + n\mathbb{Z} = \{x + kn \mid k \in \mathbb{Z}\}$. L'ensemble quotient est donc [1] $\mathbb{Z}/n\mathbb{Z}$.

Si $n > 0$, on a donc $[\mathbb{Z} : n\mathbb{Z}] = |\mathbb{Z}/n\mathbb{Z}| = n$. En revanche, on a $[\mathbb{Z} : 0\mathbb{Z}] = +\infty$, puisque la relation de congruence modulo 0 est simplement la relation d'égalité, et qu'il y a donc une infinité de classes d'équivalence.

Les notions de classe à gauche et classe à droite permettent de donner une autre caractérisation des sous-groupes distingués d'un groupe G.

Lemme II.2.26. *Soit G un groupe, et soit H un sous-groupe. Les propriétés suivantes sont équivalentes :*

(1) *le sous-groupe H est distingué dans G ;*
(2) *pour tout $x \in G$, $xH = Hx$;*
(3) *pour tout $x \in G$, $xHx^{-1} = H$, où $xHx^{-1} = \{xhx^{-1} \mid h \in H\}$.*

Démonstration.

(1) $\Longrightarrow$ (2). Supposons que H soit distingué dans G, et soit $x \in G$. Pour tout $h \in H$, on a $xhx^{-1} \in H$. Il existe donc $h' \in H$ tel que $xh = h'x$. Ainsi, $xh \in Hx$, ce qui montre que $xH \subset Hx$. Mais, on a aussi $x^{-1}hx = x^{-1}h(x^{-1})^{-1} \in H$, et donc il existe $h'' \in H$ tel que $hx = xh'' \in xH$. Ainsi, $Hx \subset xH$, d'où finalement $xH = Hx$.

(2) $\Longrightarrow$ (3). Soit $x \in G$. Alors, $xH = Hx$ par hypothèse. En utilisant les définitions, on obtient facilement les égalités

$$xHx^{-1} = (xH)x^{-1} = (Hx)x^{-1} = H(xx^{-1}) = H1_G = H,$$

d'où le résultat voulu.

(3) $\Longrightarrow$ (1). Évident. □

À cette occasion, citons un classique de théorie des groupes.

Proposition II.2.27 ([TD)). *Soit G un groupe. Alors, tout sous-groupe d'indice 2 est distingué dans G.*

On passe maintenant au théorème de Lagrange.

Théorème II.2.28 (Lagrange). *Soit G un groupe* **fini**, *et soit H un sous-groupe. Alors, on a*

$$|G| = [G : H] \cdot |H|.$$

Autrement dit,

$$|G/H| = [G : H] = \frac{|G|}{|H|}.$$

En particulier, $|H| \mid |G|$.

1. Le lecteur restera pantois devant la cohérence des notations.

Démonstration. Soit $r = [G : H]$. Les classes à gauche $x_1H, \ldots, x_rH$ formant une partition de G, on a donc

$$G = x_1H \cup \cdots \cup x_rH,$$

et l'union est disjointe. On a donc

$$|G| = |x_1H| + \cdots + |x_rH|.$$

Mais, pour tout $a \in G$, l'application

$$\begin{aligned} H &\longrightarrow aH \\ h &\longmapsto ah \end{aligned}$$

est bijective. En effet, elle est surjective par définition de aH, et injective par puisque la loi de groupe est simplifiable. En particulier, on a

$$|aH| = |H| \text{ pour tout } a \in G.$$

Ainsi, on obtient

$$|G| = r|H| = [G : H]\cdot|H|,$$

ainsi que $|H| \mid |G|$. □

Attention ! La réciproque du théorème de Lagrange est fausse.

On continue ce paragraphe par une propriété de multiplicativité de l'indice.

Proposition II.2.29. *Soit G un groupe, soit H un sous-groupe de G, et soit K un sous-groupe de H.*

Alors, $[G : K]$ est fini si, et seulement si, $[G : H]$ et $[H : K]$ sont finis. Dans ce cas,on a

$$[G : K] = [G : H][H : K].$$

Démonstration. Écrivons

$$G/H = \{g_iH \mid i \in I\},$$

où les classes à gauche $g_iH, i \in I$, sont deux à deux distinctes. De même, écrivons

$$H/K = \{h_jK \mid j \in J\},$$

où les classes à gauche $h_jK, j \in J$, sont deux à deux distinctes.On va montrer que l'on a

$$G/K = \{g_ih_jK \mid (i, j) \in I \times J\}.$$

Soit $g \in G$. Alors, il existe $i \in I$ tel que $g \in g_iH$. On a donc $g = g_ih$ pour un certain $h \in H$. Mais alors, il existe $j \in J$ tel que $h \in h_jK$. Par conséquent, il existe un indice $k \in K$ tel que $h = h_jk$, et donc $g = g_ih_jk$. On en déduit que $g \sim g_ih_j$, et donc que $gK = g_ih_jK$. Ainsi, les classes

$$g_ih_jK, (i, j) \in I \times J$$

sont les diverses classes à gauche modulo K. Il reste à voir qu'elles sont distinctes.

Supposons que l'on ait $g_ih_jK = g_{i'}h_{j'}K$ pour $(i, j), (i', j') \in I \times J$. Alors, il existe $k \in K$ tel que $g_ih_j = g_{i'}h_{j'}k$. On a donc

$$g_i = g_{i'}h_{j'}kh_j^{-1}.$$

Comme $K \subset H$, on obtient $g_i \in g_{i'}H$, et donc $g_iH = g_{i'}H$. Par choix de I, on obtient $i = i'$. On a alors $h_j = h_{j'}k$, et donc $h_jK = h_{j'}K$. Par choix de J, on obtient $j = j'$. Ceci montre que les classes $g_ih_jK, (i, j) \in I \times J$ sont toutes distinctes.

En particulier, G/K est fini si, et seulement si $I \times J$ est fini, c'est-à-dire si, et seulement si, I et J sont finis. Dans ce cas, $[G : H]$ et $[H : K]$ sont donc finis et on a

$$[G : K] = \mathrm{card}(I \times J) = \mathrm{card}(I)\mathrm{card}(J) = [G : H][H : K].$$

Ceci achève la démonstration. □

II.3. Groupes quotients

Soit G un groupe, et soit H un sous-groupe. Rappelons que nous avons défini une relation d'équivalence sur G de la manière suivante. Si $x, y \in G$, alors $x \sim y$ si $x^{-1}y \in H$. Les classes d'équivalence pour cette relation sont les classes à gauche modulo H. Autrement dit, pour tout $x \in G$, on a

$$\overline{x} = xH = \{xh \mid h \in H\}.$$

On note alors G/H l'ensemble des classes d'équivalence. Les éléments de l'ensemble G/H sont donc des **parties** de G. Ce ne sont **pas** des éléments de G.

On a alors une application surjective

$$\begin{aligned} \pi \colon G &\longrightarrow G/H \\ x &\longmapsto \overline{x}. \end{aligned}$$

On voudrait définir une structure de groupe sur G/H de manière à ce que la projection canonique $\pi : G \longrightarrow G/H$ soit un morphisme de groupes. La seule façon pour cela serait de poser

$$\overline{x}\ \overline{y} = \overline{xy} \ \text{ pour tous } x, y \in G,$$

ceci traduisant exactement ce que l'on veut. Malheureusement, rien ne dit que cette loi soit bien définie.

En effet, la difficulté cachée est que le résultat ne doit pas dépendre des éléments choisis pour représenter les classes d'équivalence de x et de y. Autrement dit, on doit vérifier que si $\overline{x}_1 = \overline{x}_2$ et $\overline{y}_1 = \overline{y}_2$, alors $\overline{x_1 y_1} = \overline{x_2 y_2}$. Ce n'est pas toujours le cas, comme le montre l'exemple suivant.

Exemple II.3.1. Soit $G = \mathfrak{S}_3$. Si $i, j \in [\![1, 3]\!]$ et $i \neq j$, on note (provisoirement) τ_{ij} l'application qui échange i et j et fixe le troisième élément. On note aussi σ l'application définie par

$$\sigma(1) = 2, \sigma(2) = 3,\ \sigma(3) = 1.$$

On vérifie que $\sigma = \tau_{13} \circ \tau_{12}$.

Soit $H = \{\mathrm{Id}, \tau_{12}\}$. On vérifie que c'est un sous-groupe de G. Notons que les deux éléments de H fixent 3. Soient

$$x_1 = \tau_{13}, x_2 = \tau_{13}\tau_{12} = \sigma, y_1 = y_2 = \tau_{23}.$$

Alors, $x_1 \sim x_2, y_1 \sim y_2$, et donc $\overline{x}_1 = \overline{x}_2$ et $\overline{y}_1 = \overline{y}_2$. En revanche, on vérifie que $x_1 y_1$ ne fixe pas 3, et que $x_2 y_2 = \tau_{12}$. Ainsi,

$$x_1 y_1 \notin H \ \text{ et } \ x_2 y_2 \in H,$$

et donc $\overline{x_1 y_1} \neq \overline{x_2 y_2}$.

Examinons le problème d'un peu plus près. Supposons que la loi interne

$$\begin{aligned} G/H \times G/H &\longrightarrow G/H \\ (\overline{x}, \overline{y}) &\longmapsto \overline{xy} \end{aligned}$$

soit bien définie. Dans ce cas, on a

$$\overline{g} = \overline{1_G g} = \overline{1}_G\ \overline{g} \ \text{ pour tout } g \in G.$$

Mais, pour tout $h \in H$, on a $\overline{1}_G = \overline{h}$ par définition de l'ensemble G/H, et ainsi, on obtient

$$\overline{g} = \overline{h}\ \overline{g} = \overline{hg} \ \text{ pour tout } g \in G, \text{ et tout } h \in H.$$

Par définition de la relation de congruence modulo H, cela revient à

$$g^{-1}hg \in H \ \text{ pour tout } g \in G, \text{ et tout } h \in H.$$

Puisque le passage à l'inverse induit une bijection de G sur lui-même, en remplaçant g par g^{-1} dans l'énoncé précédent, on constate que cela équivaut finalement à

$$ghg^{-1} \in H \ \text{ pour tout } g \in G, \text{ et tout } h \in H.$$

Autrement dit, H est nécessairement un sous-groupe distingué de G.

Ceci explique d'ailleurs l'exemple précédent, puisque le sous-groupe H choisi n'est pas distingué dans $\mathfrak{S}_3$.

La proposition suivante montre que, réciproquement, si H est un sous-groupe distingué de G, on peut munir G/H d'une structure de groupe ayant les propriétés voulues.

Proposition II.3.2. *Soit H un sous-groupe distingué de G. Alors, la loi interne*

$$\begin{aligned} G/H \times G/H &\longrightarrow G/H \\ (\overline{x}, \overline{y}) &\longmapsto \overline{xy} \end{aligned}$$

est bien définie, de neutre $\overline{1}_G$, et induit sur G/H une structure de groupe. De plus, l'application

$$\begin{aligned} \pi \colon G &\longrightarrow G/H \\ x &\longmapsto \overline{x} \end{aligned}$$

est un morphisme de groupes.

Démonstration. Vérifions que si $\overline{x}_1 = \overline{x}_2$ et $\overline{y}_1 = \overline{y}_2$, alors $\overline{x_1 y_1} = \overline{x_2 y_2}$. Par définition de la relation d'équivalence, il existe $h, h' \in H$ tels que

$$x_2 = x_1 h, y_2 = y_1 h'.$$

On a alors

$$x_2 y_2 = x_1 h y_1 h' = x_1 y_1 (y_1^{-1} h y_1) h'.$$

Comme H est un sous-groupe distingué de G, on a $y_1^{-1} h y_1 \in H$, et par conséquent $(y_1^{-1} h y_1) h' \in H$. Ainsi, $\overline{x_1 y_1} = \overline{x_2 y_2}$, et l'application

$$\begin{aligned} G/H \times G/H &\longrightarrow G/H \\ (\overline{x}, \overline{y}) &\longmapsto \overline{xy} \end{aligned}$$

est bien définie.

Vérifions maintenant que G/H, muni de cette loi interne, est bien un groupe. Pour tout $\overline{x} \in G/H$, on a

$$\overline{1}_G \overline{x} = \overline{1_G x} = \overline{x},$$

et de même $\overline{x}\overline{1}_G = \overline{x}$. Donc $\overline{1}_G$ est un élément neutre pour cette loi. De plus, la loi est associative. En effet, pour tous $\overline{x}, \overline{y}, \overline{z} \in G/H$, on a

$$(\overline{x}\ \overline{y})\overline{z} = \overline{xy}\ \overline{z} = \overline{(xy)z} = \overline{x(yz)} = \overline{x}\ \overline{yz} = \overline{x}(\overline{y}\ \overline{z}).$$

Enfin, on vérifie facilement que tout $\overline{x} \in G/H$ admet pour inverse la classe à gauche $\overline{x^{-1}}$.

Le fait que π soit un morphisme de groupes provient de la définition de la loi de groupe sur G/H. $\square$

Définition II.3.3. Soit G un groupe, et soit H un sous-groupe distingué dans G. Le groupe G/H est appelé le *groupe quotient* de G par H.

Remarques II.3.4.

(1) La chose la plus importante à retenir de cette définition, outre la définition de la loi de groupe, est la suivante : pour tout $x \in G$, on a

$$\overline{x} = \overline{1}_G \iff x \in H.$$

(2) Si $[G : H]$ est fini, G/H est d'ordre $[G : H]$ par définition. En particulier, si G est fini, par le théorème de Lagrange, on a

$$|G/H| = \frac{|G|}{|H|}.$$

(3) Si G est abélien, tout sous-groupe H est distingué dans G, et on peut toujours former le groupe quotient G/H.

(4) La projection canonique $\pi : G \longrightarrow G/H$ est surjective, de noyau H. Ainsi, tout sous-groupe distingué de G est le noyau d'un certain morphisme de groupes.

Quel est l'intérêt d'un groupe quotient ? En fait, quotienter G par un sous-groupe H distingué dans G revient à imposer des relations entre les éléments du groupe G qui n'existent pas forcément. En un sens, « on force les éléments de H à devenir triviaux ».

Exemples II.3.5.

(1) On a $G/G = \{\overline{1}_G\}$. En effet, tous les éléments sont équivalents dans ce cas. En fait, le groupe G/H est trivial si, et seulement si, $H = G$.

(2) On a $G \simeq G/\{1_G\}$, l'isomorphisme étant donné par la projection canonique. En effet, on a $\overline{x} = \{x\}$ pour tout $x \in G$. La projection canonique est donc bien un isomorphisme. Attention, on a un isomorphisme, mais **pas** une égalité ! !

(3) Soit $n \geq 1$. Le groupe $\mathbb{Z}/n\mathbb{Z}$ n'est rien d'autre que le quotient de $\mathbb{Z}$ par le sous-groupe $n\mathbb{Z}$.

On continue ce paragraphe en étudiant les sous-groupes de G/H.

Proposition II.3.6. *Soit G un groupe et soit H un sous-groupe distingué de G. Alors, la projection canonique $\pi : G \longrightarrow G/H$ induit une correspondance bijective entre l'ensemble des sous-groupes H' de G contenant H et l'ensemble des sous-groupes de G/H.*

Plus précisément, si H' est un sous-groupe de G contenant H, alors

$$H'/H = \pi(H') = \{\overline{h'} \mid h' \in H'\}$$

est un sous-groupe de G/H.

Réciproquement, si S est un sous-groupe de G/H, alors $\pi^{-1}(S)$ est un sous-groupe de G contenant H, et ces deux constructions sont inverses l'une de l'autre.

De plus, cette correspondance établit une bijection entre l'ensemble des sous-groupes distingués de G contenant H et l'ensemble des sous-groupes distingués de G/H.

Démonstration. Puisque π est un morphisme de groupes, $\pi(H')$ et $\pi^{-1}(S)$ sont des sous-groupes de G/H et G respectivement par le lemme **??**. De plus, pour tout $h \in H$, on a $\pi(h) = \overline{h} = \overline{1}_G \in S$, et donc $H \subset \pi^{-1}(S)$. Il reste à vérifier que les deux constructions sont inverses l'une de l'autre. Soit H' un sous-groupe de G contenant H, et soit $S = \pi(H') = H'/H$. Alors, on a

$$\pi^{-1}(S) = \{x \in G \mid \overline{x} \in H'/H\}.$$

Il reste à constater que $\{x \in G \mid \overline{x} \in H'/H\} = H'$.

L'inclusion « $\supset$ » est claire. Si maintenant $\overline{x} \in H'/H$, il existe $y \in H'$ tel que $\overline{x} = \overline{y}$. On a donc $x = yh, h \in H$. Or, $H \subset H'$, et ainsi $x \in H'$. On a donc bien $\pi^{-1}(S) = H'$.

Inversement, soit S un sous-groupe de G/H, et soit $H' = \pi^{-1}(S)$. On a donc $\pi(H') \subset S$ par définition de H'. Soit $\overline{x} \in S$. Alors, $\overline{x} = \pi(x) \in S$, et donc $x \in H'$. Ainsi, on a $\overline{x} \in \pi(H')$, et donc $S = \pi(H')$.

Si H' est un sous-groupe distingué de G contenant H, alors H'/H est distingué dans G/H. En effet, pour tout $\overline{x} \in G/H$ et tout $h' \in H'$, on a

$$\overline{x}\overline{h'}\overline{x}^{-1} = \overline{xh'x^{-1}} \in H'/H,$$

car H' est distingué dans G. Enfin, soit S un sous-groupe distingué de G/H. Alors, pour tout $x \in G$ et tout $h \in \pi^{-1}(S)$, on a

$$\pi(xhx^{-1}) = \pi(x)\pi(h)\pi(x)^{-1} \in S,$$

car $\pi(h) \in S$ et S est distingué dans G/H. Ainsi, $\pi^{-1}(S)$ est distingué dans G. Ceci achève la démonstration. □

Nous allons maintenant nous intéresser aux morphismes de G/H dans un groupe G'.

Théorème II.3.7 (de factorisation). *Soit G un groupe, et soit H un sous-groupe distingué de G. Soit $f : G \longrightarrow G'$ un morphisme de groupes tel que $H \subset \mathrm{Ker}(f)$. Alors, il existe un unique morphisme de groupes*

$$\overline{f} : G/H \longrightarrow G'$$

tel que $f = \overline{f} \circ \pi$.

$$\begin{tikzcd} G \arrow[r, "f"] \arrow[d, "\pi"'] & G' \\ G/H \arrow[ur, "\overline{f}"'] & \end{tikzcd} \qquad f = \overline{f} \circ \pi$$

Ce morphisme de groupes est défini par

$$\overline{f}(\overline{x}) = f(x) \text{ pour tout } \overline{x} \in G/H.$$

De plus, pour tout groupe G', il y a une correspondance bijective entre les ensembles $\mathrm{Hom}(G/H, G')$ et $\{f \in \mathrm{Hom}(G, G') \mid H \subset \mathrm{Ker}(f)\}$.

La correspondance est donnée par

$$\begin{array}{l} \varphi \longmapsto \varphi \circ \pi \\ \overline{f} \longleftarrow\!\shortmid f \end{array}$$

Démonstration. Remarquons tout d'abord qu'un morphisme $f : G \longrightarrow G'$ est constant sur les classes d'équivalences (pour la relation associée à H) si, et seulement si, $H \subset \mathrm{Ker}(f)$.

En effet, si $H \subset \mathrm{Ker}(f)$, alors pour tout $x \in G$ et $h \in H$, on a

$$f(xh) = f(x)f(h) = f(x).$$

Réciproquement, si f est constante sur chaque classe d'équivalence, elle est constante sur H. Mais H contient 1 et on a donc

$$f(h) = f(1_G) = 1_{G'} \text{ pour tout } h \in H,$$

c'est-à-dire $H \subset \mathrm{Ker}(f)$.

La proposition **??** fournit alors l'existence et l'unicité d'une application $\overline{f}$ vérifiant les propriétés voulues. De plus, pour tous $x_1, x_2 \in G$, on a

$$\overline{f}(\overline{x}_1\overline{x}_2) = \overline{f}(\overline{x_1x_2}) = f(x_1x_2) = f(x_1)f(x_2) = \overline{f}(\overline{x}_1)\overline{f}(\overline{x}_2),$$

et $\overline{f}$ est donc un morphisme de groupes.

Démontrons la seconde partie. Par la proposition **??**, on a une bijection entre l'ensemble des applications de G/H dans G' et l'ensemble des applications $G \longrightarrow G'$ constantes sur les classes d'équivalences. Elle est donnée par

$$\begin{array}{l} \varphi \longmapsto \varphi \circ \pi \\ \overline{f} \longleftarrow\!\shortmid f \end{array}$$

Si maintenant φ est un morphisme de groupes, alors $\varphi \circ \pi$ est un morphisme dont le noyau contient H. De plus, si f est un morphisme de groupes, alors $\overline{f}$ est un morphisme de groupes. Ainsi, la correspondance se restreint pour définir une bijection entre les ensembles voulus. Ceci achève la démonstration. □

Remarquons en particulier que tout morphisme de groupes $f : G \longrightarrow G'$ induit un morphisme de groupes

$$\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G',$$

puisque $\mathrm{Ker}(f)$ est un sous-groupe distingué de G qui contient... $\mathrm{Ker}(f)$. Vu la définition du groupe quotient, on peut s'attendre à ce que cette application soit injective, puisque l'on « tue » tous les éléments du noyau. C'est en effet le cas.

Lemme II.3.8. *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors, le morphisme de groupes*

$$\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G'$$

est injectif.

Démonstration. Soit $\overline{x} \in G/\mathrm{Ker}(f)$. On a $\overline{f}(\overline{x}) = 1_{G'}$ si, et seulement si, on a $f(x) = 1_{G'}$, c'est-à-dire si, et seulement si, $x \in \mathrm{Ker}(f)$, ce qui revient encore à dire que $\overline{x} = \overline{1}_G \in G/\mathrm{Ker}(f)$. Ainsi, $\mathrm{Ker}(\overline{f})$ est réduit à l'élément neutre et $\overline{f}$ est injective. □

On obtient alors le résultat fondamental suivant.

Théorème II.3.9 (Premier théorème d'isomorphisme). *Soient G, G' deux groupes, et soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors, le morphisme de groupes*

$$\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G'$$

induit un isomorphisme de groupes

$$G/\mathrm{Ker}(f) \simeq \mathrm{Im}(f).$$

Démonstration. D'après le lemme précédent, $\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G'$ est un morphisme injectif, et induit donc un isomorphisme de $G/\mathrm{Ker}(f)$ sur son image. Or, on a

$$\mathrm{Im}(\overline{f}) = \left\{\overline{f}(\overline{x}) \mid \overline{x} \in G/\mathrm{Ker}(f)\right\} = \left\{f(x) \mid x \in G\right\} = \mathrm{Im}(f),$$

d'où le résultat. □

Le théorème **??** nous permet en pratique d'identifier un groupe quotient G/H à un groupe connu de la manière suivante : on essaye d'identifier H au noyau d'un morphisme de groupes $f : G \longrightarrow G'$ bien choisi, et on calcule l'image de f (en pratique, on essaye de trouver un f qui soit surjectif).

Exemples II.3.10.

(1) Identifions le quotient $\mathbb{C}^\times/\mathbb{U}$, où $\mathbb{U} = \{z \in \mathbb{C}^\times \mid |z| = 1\}$, et considérons l'application
$$\begin{aligned} f\colon \mathbb{C}^\times &\longrightarrow \mathbb{R}^{+\times} \\ z &\longmapsto |z|. \end{aligned}$$
C'est un morphisme de groupes surjectif, de noyau $\mathbb{U}$, et le premier théorème d'isomorphisme nous donne
$$\mathbb{C}^\times/\mathbb{U} \simeq \mathbb{R}^{+\times}.$$

(2) Identifions $\mathbb{R}/2\pi\mathbb{Z}$. On considère le morphisme de groupes surjectif
$$\begin{aligned} f\colon \mathbb{R} &\longrightarrow \mathbb{U} \\ \theta &\longmapsto e^{i\theta}. \end{aligned}$$
Son noyau étant égal à $2\pi\mathbb{Z}$, on obtient
$$\mathbb{R}/2\pi\mathbb{Z} \simeq \mathbb{U}.$$

(3) Identifions $\mathbb{R}^\times/\mathbb{R}^{\times 2}$, où $\mathbb{R}^{\times 2}$ est l'ensemble des carrés non nuls de $\mathbb{R}^\times$. On considère le morphisme de groupes surjectif
$$\begin{aligned} f\colon \mathbb{R}^\times &\longrightarrow \{\pm 1\} \\ x &\longmapsto \frac{x}{|x|}. \end{aligned}$$
On a $\mathrm{Ker}(f) =]0, +\infty[= \mathbb{R}^{\times 2}$, et on obtient donc
$$\mathbb{R}^\times/\mathbb{R}^{\times 2} \simeq \{\pm 1\}.$$

Nous allons maintenant démontrer deux autres théorèmes d'isomorphisme.

Théorème II.3.11 (Deuxième théorème d'isomorphisme). *Soit G un groupe, et soient H et K deux sous-groupes. On suppose que H est distingué dans G. Alors, on a les propriétés suivantes :*

(1) *les ensembles*
$$HK = \{hk \mid h \in H, k \in K\} \quad \textit{et} \quad KH = \{kh \mid h \in H, k \in K\}$$
sont des sous-groupes de G. De plus, $HK = KH$.

(2) *le sous-groupe H est distingué dans HK, $H \cap K$ est distingué dans K, et on a un isomorphisme de groupes*
$$HK/H \simeq K/(H \cap K).$$

Démonstration.

(1) Montrons que HK et KH sont des sous-groupes de G.

Puisque H et K sont des sous-groupes de G, ils contiennent 1_G, et donc HK contient $1_G 1_G = 1_G$. Si $h_1k_1, h_2k_2 \in HK$, on a
$$(h_1k_1)(h_2k_2) = (h_1(k_1h_2k_1^{-1}))(k_1k_2) \in HK,$$
puisque H est distingué dans G, et H et K sont des sous-groupes. Enfin, si $hk \in HK$, on a également
$$(hk)^{-1} = k^{-1}h^{-1} = (k^{-1}h^{-1}k)k^{-1} \in HK,$$

pour les mêmes raisons que précédemment. Ainsi, HK est bien un sous-groupe de G.

Il reste à voir que $HK = KH$, ce qui montrera toutes les assertions manquantes. Comme H, K et HK sont des sous-groupes de G, on a

$$HK = (HK)^{-1} = K^{-1}H^{-1} = KH,$$

où pour toute partie P de G, P^{-1} désigne l'ensemble des inverses des éléments de P. Ceci achève la démonstration de (1).

(2) Pour tout $h' \in H$, et tout $hk \in HK$, on a

$$(hk)h'(hk)^{-1} = h(kh'k^{-1})h^{-1} \in H,$$

et donc H est distingué dans HK. On peut ainsi former le groupe quotient HK/H. Considérons le morphisme de groupes

$$\begin{aligned} f\colon K &\longrightarrow HK/H \\ k &\longmapsto \overline{k}. \end{aligned}$$

Pour tout $hk \in HK$, on a

$$\overline{hk} = \overline{h}\,\overline{k} = \overline{1}\,\overline{k} = \overline{k} = f(k),$$

et ainsi f est surjective. D'autre part, pour tout $k \in K$, on a

$$\overline{k} = \overline{1} \iff k \in H.$$

Autrement dit, $\mathrm{Ker}(f) = H \cap K$. Ainsi, $H \cap K$ est distingué dans K par le lemme **??**, et par le premier théorème d'isomorphisme, on obtient

$$K/(H \cap K) \simeq HK/H,$$

d'où le résultat. □

Théorème II.3.12 (Troisième théorème d'isomorphisme). *Soit G un groupe, et soient H et K deux sous-groupes. On suppose que H et K sont distingués dans G, et que K est un sous-groupe de H. Alors, K est distingué dans H, H/K est distingué dans G/K, et on a un isomorphisme de groupes*

$$(G/K)/(H/K) \simeq G/H.$$

Démonstration. Clairement, K est distingué dans H. Soit $\pi : G \longrightarrow G/H$ la projection canonique. Comme $K \subset H = \mathrm{Ker}(\pi)$, le théorème de factorisation montre que π induit un morphisme de groupes

$$\begin{aligned} \tilde{\pi}\colon G/K &\longrightarrow G/H \\ \tilde{x} &\longmapsto \overline{x}, \end{aligned}$$

où $\tilde{x}$ et $\overline{x}$ désignent respectivement les classes à gauche de x modulo K et modulo H.

Le morphisme $\tilde{\pi}$ est évidemment surjectif. De plus, pour tout $x \in G$, on a

$$\overline{x} = \overline{1} \iff x \in H \iff \tilde{x} \in H/K.$$

Autrement dit, $\mathrm{Ker}(\tilde{\pi}) = H/K$. Ainsi, H/K est distingué dans G/K, et par le premier d'isomorphisme, on a

$$(G/K)/(H/K) \simeq G/H,$$

ce qui achève la démonstration. □

II.4. Sous-groupes engendrés par une partie

Nous avons vu que la notion de sous-groupe ne se comporte pas bien vis-à-vis de la réunion. En revanche, tout se passe bien par intersection.

LEMME II.4.1. *Soit G un groupe, et soit $(H_i)_{i\in I}$ une famille non vide de sous-groupes de G. Alors, $\bigcap_{i\in I} H_i$ est un sous-groupe de G.*

Démonstration. Posons $H = \bigcap_{i\in I} H_i$. Puisque $1_G \in H_i$ pour tout $i \in I$, alors $1_G \in H$. En particulier, H est non vide. Si $x, y \in H$, alors $x, y \in H_i$ pour tout $i \in I$. On a alors $xy^{-1} \in H_i$ pour tout $i \in I$ puisque H_i est un sous-groupe de G, et donc $xy^{-1} \in H$. □

DÉFINITION II.4.2. Soit G un groupe, et soit P une partie de G (qui peut être éventuellement vide). Le *sous-groupe de G engendré par P*, noté $\langle P\rangle$, est l'intersection de tous les sous-groupes de G contenant P.

Autrement dit, $\langle P\rangle$ est le plus petit sous-groupe de G contenant P (au sens de l'inclusion).

Cette définition a bien un sens, puisque d'une part on intersecte sur une famille non vide (elle contient au moins G), et le lemme précédent nous dit que l'on obtient bien un sous-groupe.

REMARQUES II.4.3. Soit G un groupe et soient P, Q deux parties de G. Alors :

(1) on a $P \subset \langle P\rangle$;
(2) soit H un sous-groupe de G. Si $P \subset H$, alors $\langle P\rangle \subset H$;
(3) on a $\langle P\rangle = P$ si, et seulement si, P est un sous-groupe de G ;
(4) si $P \subset Q$, alors $\langle P\rangle \subset \langle Q\rangle$;
(5) on a $\langle\emptyset\rangle = \{1_G\}$.

On a une description précise de $\langle P\rangle$.

THÉORÈME II.4.4. *Soit G un groupe, et soit P une partie de G. On note*

$$P^{-1} = \{x^{-1} \mid x \in P\}.$$

Alors, on a

$$\langle P\rangle = \{x_1 \cdots x_n \mid n \geq 0, x_i \in P \cup P^{-1}\}.$$

Démonstration. Soit $H = \{x_1 \cdots x_n \mid n \geq 0, x_i \in P \cup P^{-1}\}$.

Montrons tout d'abord que $H \subset \langle P\rangle$. Si $x \in P \subset \langle P\rangle$, alors $x^{-1} \in \langle P\rangle$, car $\langle P\rangle$ est un sous-groupe de G. Donc $P \cup P^{-1} \subset \langle P\rangle$, et puisque $\langle P\rangle$ est un sous-groupe, on a $x_1 \cdots x_n \in \langle P\rangle$ pour tous $x_1, \ldots, x_n \in P \cup P^{-1}$ (même si $n = 0$, puisque le neutre de G est dans tout sous-groupe). Il s'ensuit que $H \subset \langle P\rangle$.

Pour montrer que $\langle P\rangle \subset H$, il suffit de montrer que H est un sous-groupe de G contenant P, puisque $\langle P\rangle$ est le plus petit-sous-groupe de G vérifiant cette propriété. En prenant $n = 1$ et $x_1 \in P$, on voit que $P \subset H$. En prenant cette fois $n = 0$, on obtient $1_G \in H$. En particulier, H est non vide. D'autre part, si $x \in P \cup P^{-1}$, on a $x^{-1} \in P \cup P^{-1}$. Par conséquent, si $x = x_1 \cdots x_n, y = y_1 \cdots y_m$ sont des éléments de H, on en déduit que

$$xy^{-1} = x_1 \cdots x_n y_m^{-1} \cdots y_1^{-1} \in H.$$

Ainsi, H est un sous-groupe de G contenant P, et on a fini. □

Notation. Si $P = \{a_1, \dots, a_r\}$, le sous-groupe engendré par P est simplement noté $\langle a_1, \dots, a_r \rangle$.

Exemple II.4.5. Soit G un groupe, et soit $x \in G$. Alors, on a

$$\langle x \rangle = \{x^n \mid n \in \mathbb{Z}\}.$$

En effet, le théorème précédent nous dit que

$$\langle x \rangle = \{x^{\varepsilon_1} \cdots x^{\varepsilon_n} \mid n \geq 0, \varepsilon_i = \pm 1\}.$$

En prenant les ε_i tous égaux à 1 ou -1, on obtient l'inclusion « $\supset$ ». D'autre part, on a

$$x^{\varepsilon_1} \cdots x^{\varepsilon_n} = x^{\varepsilon_1 + \cdots + \varepsilon_n}.$$

Comme $\varepsilon_1 + \cdots + \varepsilon_n \in \mathbb{Z}$, on a l'inclusion manquante.

En particulier, $\langle x \rangle$ est toujours un groupe abélien, puisque

$$x^n x^m = x^{n+m} = x^{m+n} = x^m x^n.$$

Remarque II.4.6. Si G est **abélien**, et si $P = \{a_1, \dots, a_r\}$, on a une description un peu plus simple de $\langle P \rangle$. En effet, on a

$$\langle a_1, \dots, a_r \rangle = \{a_1^{m_1} \cdots a_r^{m_r} \mid m_1, \dots, m_r \in \mathbb{Z}\}.$$

En version additive, cela donne

$$\langle a_1, \dots, a_r \rangle = \{m_1 \cdot a_1 + \cdots + m_r \cdot a_r \mid m_1, \dots, m_r \in \mathbb{Z}\}.$$

Définition II.4.7. Soit G un groupe, et soit P une partie de G. On dit que P *engendre* G si $G = \langle P \rangle$. On dit aussi que P est une partie génératrice, ou un système de générateurs de G. Par exemple, P engendre $\langle P \rangle$.

Exemples II.4.8. (1) (TD) Si $n \geq 3$, le groupe D_n des isométries d'un n-gone régulier est un groupe d'ordre $2n$. Si O est le centre du n-gone, et si S en est un sommet fixé, D_n est engendré par une rotation σ d'angle $\dfrac{2\pi}{n}$ et de centre O, et par une réflexion orthogonale d'axe (OS).

(2) Soit $f : G \longrightarrow G'$ un morphisme de groupes. Pour toute partie P de G, on a $f(\langle P \rangle) = \langle f(P) \rangle$ (exercice laissé au lecteur).

Remarque II.4.9. La notion de partie génératrice est très utile pour simplifier les calculs.

Par exemple, supposons que H et H' soient deux groupes, et que l'on veuille montrer que $H \subset H'$. Si l'on dispose d'une partie génératrice P de H, alors il faut et il suffit de montrer que $P \subset H'$.

De même, supposons que H soit un sous-groupe de G, et que H soit engendré par une partie P. Alors, pour montrer que H est un sous-groupe distingué de G, il faut et il suffit de montrer que pour tout $g \in G$, et pour tout $x \in P$, on a $gxg^{-1} \in H$.

En effet, d'après l'exemple **??** (2), l'ensemble gHg^{-1} est un sous-groupe de G engendré par $gPg^{-1} = \{gxg^{-1} \mid x \in P\}$. Si la condition précédente est vérifiée, on a alors $\langle gPg^{-1} \rangle = gHg^{-1} \subset H$, et H est distingué dans G.

Ces astuces de calcul seront utilisées dans la suite sans référence ultérieure.

Nous allons maintenant faire une pause, et illustrer les notions précédentes sur quelques exemples concrets.

II.5. Exemple du groupe $\mathbb{Z}$

On commence par s'intéresser au groupe $\mathbb{Z}$. Il est clairement engendré par 1, i.e. $\mathbb{Z} = \langle 1 \rangle$. Plus généralement, tout sous-groupe de $\mathbb{Z}$ est engendré par un élément.

PROPOSITION II.5.1. *Les sous-groupes de $\mathbb{Z}$ sont tous de la forme*

$$n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}, n \in \mathbb{Z}.$$

Plus précisément, si H est un sous-groupe de $\mathbb{Z}$ non nul, alors on a $H = n\mathbb{Z}$, où n est le plus petit élément strictement positif de H.

Démonstration. Supposons tout d'abord $H = n\mathbb{Z}$. Alors, H est non vide, et puisque la différence de deux multiples de n est un multiple de n, on en déduit que H est un sous-groupe de $\mathbb{Z}$. Réciproquement, soit H un sous-groupe de $\mathbb{Z}$. Si $H = \{0\}$, le résultat est trivialement vrai. Supposons H non nul. Alors, l'ensemble $\{m \in H \mid m > 0\}$ est une partie non vide de $\mathbb{N}$, et admet donc un plus petit élément $n > 0$. Soit $m \in \mathbb{Z}$. Puisque $n \in H$ et que H est un sous-groupe, alors $nm = n + \cdots + n$ ou $(-n) + \cdots + (-n) \in H$, selon que m est positif ou négatif. On a donc $n\mathbb{Z} \subset H$. Réciproquement, soit $a \in H$. On peut écrire

$$a = qn + r, q, r \in \mathbb{Z}, \quad \text{avec } r \in [\![0, n-1]\!].$$

Puisque $a, qn \in H$, on a $r = a - qn \in H$. Si l'on avait $r > 0$, cela contredirait le choix de n. Donc $r = 0$ et $a = qn \in n\mathbb{Z}$. Cela montre que $H \subset n\mathbb{Z}$ et donc $H = n\mathbb{Z}$. □

D'après ce qui précède, on devrait alors pouvoir décrire l'intersection de sous-groupes et le sous-groupe engendré par une partie de $\mathbb{Z}$ en fonction d'un seul élément.

PROPOSITION II.5.2. *Soient $a, b \in \mathbb{Z}$, et soient $d, m \in \mathbb{Z}$ tels que*

$$\langle a, b \rangle = a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z} \quad \textit{et} \quad a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}.$$

Alors, d est un pgcd de a et b, et m est un ppcm de a et b.

Démonstration. L'égalité $\langle a, b \rangle = a\mathbb{Z} + b\mathbb{Z}$ découle de la remarque **??**. Soit $d \in \mathbb{Z}$ tel que $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. Puisque $a \in \langle a, b \rangle = d\mathbb{Z}$, a est un multiple de d. Autrement dit, $d \mid a$. De même, $d \mid b$. Si maintenant $c \in \mathbb{Z}$ vérifie $c \mid a$ et $c \mid b$, alors c divise tout élément de $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. En particulier, $c \mid d$. On en déduit que d est un pgcd de a et b.

Montrons la deuxième partie de la proposition. Soit $m \in \mathbb{Z}$ tel que

$$a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}.$$

On a $m \in m\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$, donc $m \in a\mathbb{Z}$ et $m \in b\mathbb{Z}$. Par conséquent, $a \mid m$ et $b \mid m$. Si maintenant $c \in \mathbb{Z}$ vérifie $a \mid c$ et $b \mid c$, alors $c \in a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$, et donc $m \mid c$. On en déduit que m est un ppcm de a et b. □

Intéressons-nous maintenant à des groupes issus de l'algèbre linéaire et de la géométrie.

II.6. Exemple du groupe linéaire et du groupe spécial linéaire

Dans ce qui suit, K est un corps et E est un K-espace vectoriel de dimension $n \geq 1$. On note $\mathscr{L}(E)$ l'ensemble des endomorphismes de E. On commence par définir la notion de transvection.

DÉFINITION II.6.1. Soit H un hyperplan de E, et soit D une droite de E telle que $D \subset H$ (ce qui impose $\dim_K(E) \geq 2$). On dit que $u \in \mathscr{L}(E)$ est **une** *transvection* d'hyperplan H et de droite D si les conditions suivantes sont vérifiées :

(1) $u \neq \mathrm{Id}_E$;

(2) $u(x) = x$ pour tout $x \in H$;

(3) $u(x) - x \in D$ pour tout $x \in E$.

On dit que $u \in \mathscr{L}(E)$ est une transvection s'il existe un hyperplan H et une droite D de E incluse dans H tels que u soit une transvection d'hyperplan H et de droite D.

On commence par caractériser les transvections matriciellement.

Proposition II.6.2. *Soit $u \in \mathscr{L}(E)$. Les propriétés suivantes sont équivalentes :*

(1) *u est une transvection ;*

(2) *il existe une base de E dans laquelle la matrice représentative de u est de la forme*

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix},$$

où les coefficients non diagonaux n'apparaissant pas sont nuls ;

(3) *il existe une base de E dans laquelle la matrice représentative de u est de la forme*

$$T_{i,j}(\lambda) = \begin{array}{c} \\ i \rightarrow \end{array} \begin{array}{c} \begin{array}{cccc} & & & j \\ & & & \downarrow \end{array} \\ \begin{pmatrix} 1 & & & \\ & \ddots & & \lambda \\ & & \ddots & \\ & & & 1 \end{pmatrix} \end{array}$$

où $i \neq j$ et $\lambda \in K^\times$.

Démonstration.

(1) $\Longrightarrow$ (2). Puisque $u \neq \mathrm{Id}_E$, il existe $e_n \in E$ tel que $u(e_n) \neq e_n$. En particulier, $e_n \notin H$. Soit $e_{n-1} = u(e_n) - e_n$. C'est donc un élément non nul de D, donc une base de D, que l'on peut compléter en une base $(e_1, \ldots, e_{n-1})$ de H. Comme $e_n \notin H$, la famille $\mathbf{e} = (e_1, \ldots, e_n)$ est une base de H.

On a alors $u(e_i) = e_i$ pour tout $i = 1, \ldots, n-1$ et $u(e_n) = e_{n-1} + e_n$, d'où le résultat voulu.

(2) $\Longrightarrow$ (3). Évident.

(3) $\Longrightarrow$ (1). Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E telle que

$$\mathrm{Mat}(u; \mathbf{e}) = T_{i,j}(\lambda).$$

Posons $H = \mathrm{Vect}_K(e_k, k \neq j)$ et $D = Ke_i$. Alors H est un hyperplan, et D est une droite contenue dans H (puisque $i \neq j$). Des calculs élémentaires montrent alors que u est une transvection d'hyperplan H et de droite D.

Ceci achève la démonstration.

□

Le point (3) nous invite à poser la définition suivante.

Définition II.6.3. Soit $n \geq 1$. Une *matrice de transvection* de $\mathrm{M}_n(K)$ est une matrice de la forme

$$T_{i,j}(\lambda) = \begin{array}{c} \\ i \rightarrow \end{array} \begin{array}{c} \begin{array}{cccc} & & & j \\ & & & \downarrow \end{array} \\ \begin{pmatrix} 1 & & & \\ & \ddots & & \lambda \\ & & \ddots & \\ & & & 1 \end{pmatrix} \end{array}$$

où $i \neq j$ et $\lambda \in K^\times$.

Remarquons que $T_{i,j}(\lambda) \in \mathrm{SL}_n(K)$ et que $T_{i,j}(\lambda)^{-1} = T_{i,j}(-\lambda)$.

Enfin, si $M \in \mathrm{M}_n(K)$, désignons par $L_1, \ldots, L_n$ les lignes de M et par $C_1, \ldots, C_n$ les colonnes de M. Alors l'opération

$$M \longrightarrow T_{i,j}(\lambda)M$$

revient à faire l'opération élémentaire sur les lignes

$$L_i \longleftarrow L_i + \lambda L_j,$$

tandis que l'opération

$$M \longrightarrow MT_{i,j}(\lambda)$$

revient à faire l'opération élémentaire sur les colonnes

$$C_j \longleftarrow C_j + \lambda C_i.$$

On continue en définissant les dilatations.

Définition II.6.4. Soit H un hyperplan de E et D une droite de E tels que $E = H \oplus D$, et soit $\lambda \in K \setminus \{0, 1\}$. La *dilatation* d'hyperplan H, de direction D et de rapport λ est l'unique endomorphisme $u \in \mathscr{L}(E)$ tel que :

(1) on a $u(x) = x$ pour tout $x \in H$;
(2) on a $u(x) = \lambda x$ pour tout $x \in D$.

Autrement dit, on a

$$u(x_H + x_D) = x_H + \lambda x_D,$$

pour tout $x_H \in H$ et tout $x_D \in D$.

On dit que $u \in \mathscr{L}(E)$ est une dilatation s'il existe un hyperplan H et une droite D supplémentaires dans E, et un scalaire $\lambda \in K \setminus \{0, 1\}$ tels que u soit la dilatation d'hyperplan H, de direction D et de rapport λ.

Le résultat suivant caractérise les dilatations. Sa démonstration est un exercice facile laissé au lecteur.

Proposition II.6.5. *Soit E un K-espace vectoriel de dimension $n \geq 1$. Pour tout endomorphisme $u \in \mathscr{L}(E)$, les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme u est une dilatation ;*
(2) *il existe une base de E dans laquelle la matrice représentative de u est de la forme*

$$D(\lambda) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{pmatrix},$$

où $\lambda \in K \setminus \{0,1\}$.

DÉFINITION II.6.6. Soit $n \geq 1$. Une *matrice de dilatation* de $\mathrm{M}_n(K)$ est une matrice de la forme

$$D(\lambda) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{pmatrix},$$

où $\lambda \in K \setminus \{0,1\}$.

Remarquons que $D(\lambda) \in \mathrm{GL}_n(K)$ et que $D(\lambda)^{-1} = D(\lambda^{-1})$.

Enfin, si $M \in \mathrm{M}_n(K)$, l'opération

$$M \longrightarrow D(\lambda)M$$

revient à faire l'opération élémentaire sur les lignes

$$L_n \longleftarrow \lambda L_n,$$

tandis que l'opération

$$M \longrightarrow MD(\lambda)$$

revient à faire l'opération élémentaire sur les colonnes

$$C_n \longleftarrow \lambda C_n.$$

On peut maintenant énoncer le résultat suivant.

PROPOSITION II.6.7. *Soit K un corps, et soit $n \geq 1$.*

(1) *Toute matrice $M \in \mathrm{SL}_n(K)$ est produit de matrices de transvection ;*

(2) *toute matrice $M \in \mathrm{GL}_n(K)$ est produit de matrices de transvection et d'au plus une matrice de dilatation.*

En particulier, $\mathrm{SL}_n(K)$ est engendré par les matrices de transvection, et $\mathrm{GL}_n(K)$ est engendré par les matrices de transvection et les matrices de dilatations.

Autrement dit, si E est un K-espace vectoriel non nul de dimension finie, on a :

(1′) *tout $u \in \mathrm{SL}(E)$ est la composée de transvections ;*

(2′) *tout $u \in \mathrm{GL}(E)$ est la composée de transvections et d'au plus une dilatation.*

En particulier, $\mathrm{SL}(E)$ est engendré par les transvections, et $\mathrm{GL}(E)$ est engendré par les transvections et les dilatations.

Démonstration. Les propriétés (1′) et (2′) sont des conséquences directes de (1) et (2). Il suffit donc de démontrer ces dernières.

Soit $M \in \mathrm{SL}_n(K)$. Nous allons montrer qu'il existe des matrices de transvection $T_1, \ldots, T_r, T'_1, \ldots, T'_s$ telles que

$$T_1 \cdots T_r M T'_1 \cdots T'_s = I_n,$$

ce qui suffira à démontrer la première assertion, en utilisant le fait que l'inverse d'une matrice de transvection est encore une matrice de transvection. Cela revient à dire que l'on peut passer de M à I_n en faisant des opérations du type

$$C_j \longleftarrow C_j + \lambda C_i \ , \ L_i \longleftarrow L_i + \lambda L_j, \ i \neq j.$$

On procède par récurrence sur n.

Si $n = 1$, il n'y a rien à faire. Supposons que le résultat soit vrai pour toute matrice de $\mathrm{SL}_{n-1}(K)$, et soit $M = (a_{ij}) \in \mathrm{SL}_n(K)$. Puisque M est en particulier inversible, L_1 est non nulle, donc il existe j tel que $a_{1j} \neq 0$. Si $j \neq 1$ et si $a_{11} = 0$, on effectue

$$C_1 \longleftarrow C_1 + C_j$$

et on peut donc supposer $a_{11} \neq 0$. Si maintenant $a_{21} = 0$, on effectue

$$L_2 \longleftarrow L_2 + L_1$$

et on peut supposer que l'on a aussi $a_{21} \neq 0$. On fait ensuite l'opération

$$L_1 \longleftarrow L_1 + \frac{1 - a_{11}}{a_{21}} L_2$$

pour obtenir une matrice telle que $a_{11} = 1$. En faisant alors des opérations du type

$$C_j \longleftarrow C_j + \lambda C_1 \ , \ L_j \longleftarrow L_j + \lambda L_1,$$

on obtient une matrice diagonale par blocs de la forme

$$M'' = \begin{pmatrix} 1 & 0 \\ 0 & M' \end{pmatrix}.$$

Faire des opérations élémentaires ne changent pas le déterminant, donc cette matrice est aussi de déterminant 1. En développant par rapport à la première colonne, on obtient $\det(M') = 1$, c'est-à-dire $M' \in \mathrm{SL}_{n-1}(K)$. Par hypothèse de récurrence, on peut faire des opérations sur les lignes et les colonnes de M' du type précédent pour la ramener à I_{n-1}. Il existe donc des matrices de transvection $P_1, \ldots, P_k, P'_1, \ldots, P'_\ell \in \mathrm{SL}_{n-1}(K)$ telles que

$$P_1 \cdots P_k M' P'_1 \cdots P'_\ell = I_{n-1}.$$

Posons alors

$$T_i = \begin{pmatrix} 1 & 0 \\ 0 & P_i \end{pmatrix}, \ T'_j = \begin{pmatrix} 1 & 0 \\ 0 & P'_j \end{pmatrix}.$$

Ce sont des matrices de transvections, et le calcul par blocs montre alors que $T_1 \cdots T_k M'' T'_1 \cdots T'_\ell = I_n$. Comme M'' est obtenue en multipliant M à gauche et à droite par des matrices de transvections, on a le résultat voulu. Ceci achève alors la récurrence. Pour montrer le dernier point, il suffit de remarquer que si $M \in \mathrm{GL}_n(K) \setminus \mathrm{SL}_n(K)$, alors $MD(\det(M))^{-1} \in \mathrm{SL}_n(K)$, et d'appliquer le premier point. □

On achève ces considérations sur les groupes $\mathrm{GL}(E)$ et $\mathrm{SL}(E)$ en calculant leur centre. Le lemme suivant est bien connu.

Lemme II.6.8. *Soit $u \in \mathscr{L}(E)$. On suppose que u stabilise toute droite de E. Alors, u est une homothétie.*

Démonstration. Si u stabilise toute droite de E, alors pour tout $x \in E$, il existe $\lambda_x \in K$ tel que $u(x) = \lambda_x x$. Soient $x, y \in E$ non nuls. Si x et y sont linéairement indépendants sur K, la relation $u(x + y) = u(x) + u(y)$ implique que

$$(\lambda_{x+y} - \lambda_x)x + (\lambda_{x+y} - \lambda_y)y = 0,$$

d'où $\lambda_x = \lambda_{x+y} = \lambda_y$. Si x et y sont liés, par exemple si $y = \mu x, \mu \in K^\times$, la relation $u(y) = \mu u(x)$ montre que $\mu\lambda_y = \mu\lambda_x$, d'où $\lambda_y = \lambda_x$.

Finalement, pour tous $x, y \in E$ non nuls, on a $\lambda_x = \lambda_y$. Si on note λ cette valeur commune, on en déduit que $u(x) = \lambda x$ pour tout $x \in E$ non nul. Comme $u(0) = 0 = \lambda \cdot 0$, on a bien $u = \lambda \mathrm{Id}_E$. □

Proposition II.6.9. *Un endomorphisme $u \in \mathrm{GL}(E)$ commutant à toute transvection est une homothétie.*

En particulier, on a $Z(\mathrm{GL}(E)) = K^\times \mathrm{Id}_E$, ainsi que

$$Z(\mathrm{SL}(E)) = Z(\mathrm{GL}(E)) \cap \mathrm{SL}(E) = \{\lambda \mathrm{Id}_E \mid \lambda \in K \quad \textit{tel que } \lambda^n = 1\}.$$

Démonstration. Les résultats étant clairs si $n = 1$, supposons que $n \geq 2$.

Soit D une droite de E, et soit $e_{n-1} \in D \setminus \{0\}$, que l'on complète en une base $\mathbf{e} = (e_1, \ldots, e_n)$ de E. Soit $v \in \mathscr{L}(E)$ l'unique application linéaire telle que

$$v(e_i) = e_i \text{ pour tout } i = 1, \ldots, n-1 \quad \text{et } v(e_n) = e_{n-1} + e_n.$$

En utilisant la définition, on voit que v est une transvection de droite D et d'hyperplan $H = \mathrm{Vect}_K(e_1, \ldots, e_{n-1})$.

On a alors $\mathrm{Im}(v - \mathrm{Id}_E) = Ke_{n-1} = D$, comme on peut le constater aisément en utilisant la définition de v. Mais on a aussi

$$\mathrm{Im}(u \circ v \circ u^{-1} - \mathrm{Id}_E) = \mathrm{Im}(u \circ (v - \mathrm{Id}_E) \circ u^{-1}) = u(\mathrm{Im}(v - \mathrm{Id}_E)) = u(D).$$

Comme $u \circ v \circ u^{-1} = v$ par hypothèse, on aussi $\mathrm{Im}(uvu^{-1} - \mathrm{Id}_E) = D$. Ainsi, pour toute droite D de E, $u(D) = D$. Autrement dit, u laisse stable toute droite de E, et u est une homothétie par le lemme **??**.

Comme un élément de $Z(\mathrm{GL}(E))$ commute en particulier à toute transvection, on obtient $Z(\mathrm{GL}(E)) \subset K^\times \mathrm{Id}_E$. Puisque toute homothétie commute évidemment à tout endomorphisme, on a bien

$$Z(\mathrm{GL}(E)) = K^\times \mathrm{Id}_E.$$

Comme les transvections sont des éléments de $\mathrm{SL}(E)$, un élément dans le centre de $\mathrm{SL}(E)$ commute à toute transvection. C'est donc une homothétie par le point précédent. On obtient alors le résultat en remarquant que $\det(\lambda \mathrm{Id}_E) = \lambda^n$. □

II.7. Ordre d'un élément

Nous allons maintenant définir l'ordre d'un élément d'un groupe.

Définition II.7.1. Soit G un groupe. On dit que $x \in G$ est *d'ordre fini* si le sous-groupe $\langle x \rangle$ est fini. Dans ce cas, *l'ordre* de x est par définition l'ordre de $\langle x \rangle$. On le note $o(x)$.

En particulier, $o(x)$ est bien défini pour tout $x \in G$ lorsque G est un groupe fini.

Cette définition n'est a priori pas très commode à manipuler. Le résultat suivant nous fournit un moyen pratique de calculer l'ordre d'un élément.

Théorème II.7.2. *Soit G un groupe, et soit $x \in G$. Alors, x est d'ordre fini si, et seulement si, il existe $m > 0$ tel que $x^m = 1_G$. Dans ce cas, l'ordre de x est le plus petit entier $n > 0$ vérifiant $x^n = 1_G$ et on a*

$$\langle x \rangle = \{1_G, x, \ldots, x^{n-1}\}.$$

De plus, pour tout $m \in \mathbb{Z}$, on a

$$x^m = 1_G \iff o(x) | m.$$

Enfin, si G est un groupe fini, on a $o(x) | \, | G|$.

Démonstration. Supposons que x soit d'ordre fini, c'est-à-dire que $\langle x \rangle$ soit fini. Alors, le morphisme de groupes

$$\begin{aligned} h_x \colon \mathbb{Z} &\longrightarrow G \\ m &\longmapsto x^m \end{aligned}$$

ne peut être injectif, car sinon h_x induirait un isomorphisme de $\mathbb{Z}$ sur son image $\langle x \rangle$, et $\langle x \rangle$ serait alors infini. Ainsi, $\mathrm{Ker}(h_x)$ est non nul, et donc contient un élément $m > 0$. On a alors $x^m = 1_G$.

Réciproquement, supposons qu'un tel entier $m > 0$ existe, c'est-à-dire que le noyau de h_x n'est pas réduit à 0. Puisque $\mathrm{Ker}(h_x)$ est un sous-groupe de $\mathbb{Z}$, on a $\mathrm{Ker}(h_x) = n\mathbb{Z}$, où n est le plus petit élément strictement positif de $\mathrm{Ker}(h_x)$, d'après la proposition **??**. Ainsi pour tout $m \in \mathbb{Z}$, on a

$$x^m = 1_G \iff n \mid m.$$

Montrons alors que $\langle x \rangle = \{1_G, x, \ldots, x^{n-1}\}$. L'inclusion « $\supset$ » est claire.

Soit $m \in \mathbb{Z}$, et écrivons $m = qn + r, q, r \in \mathbb{Z}$, avec $r \in [\![0, n-1]\!]$. Par conséquent,

$$x^m = x^{qn+r} = (x^n)^q x^r = x^r,$$

ce qui montre l'inclusion « $\subset$ ».

Pour finir, il suffit de montrer que les éléments $1_G, x, \ldots, x^{n-1}$ sont deux à deux distincts. Supposons qu'il existe $i, j \in [\![0, n-1]\!]$ tels que $x^i = x^j$. Alors, on a $x^{i-j} = 1_G$. Autrement dit, $i - j \in \mathrm{Ker}(h_x) = n\mathbb{Z}$, i.e. $i - j$ est un multiple de n. Comme $-n < i - j < n$, on a nécessairement $i - j = 0$, soit $i = j$, d'où le résultat voulu par contraposition.

Le fait que $o(x)$ divise l'ordre du groupe G est une application directe du théorème de Lagrange au sous-groupe $\langle x \rangle$. □

Remarque II.7.3. La réciproque du dernier point est fausse.

Par exemple, le groupe $G = \{(1,1), (1,-1), (-1,1), (-1,-1)\} \subset \mathbb{R}^\times \times \mathbb{R}^\times$ est d'ordre 4, mais n'a pas d'élément d'ordre 4.

Corollaire II.7.4. *Soit G un groupe fini d'ordre $n \geq 1$. Pour tout $x \in G$, on a $x^n = 1_G$.*

Démonstration. Par le théorème **??**, on a $o(x) \mid n$ et par le même théorème, on a alors $x^n = 1_G$. □

Corollaire II.7.5. *Soit G un groupe, et soit $x \in G$ un élément d'ordre fini. Alors, pour tout $d \geq 1$, x^d est d'ordre fini, et on a*

$$o(x^d) = \frac{o(x)}{\mathrm{pgcd}(d, o(x))}.$$

En particulier :

(1) *si $d \mid o(x)$, on a $o(x^d) = \dfrac{o(x)}{d}$;*

(2) *si d est premier avec $o(x)$, on a $o(x^d) = o(x)$.*

Démonstration. Soit $x \in G$, et soit $d \geq 1$. Posons $c = \mathrm{pgcd}(d, o(x))$, et écrivons $d = rc$, $o(x) = sc$, où r et s sont premiers entre eux. Soit $m \in \mathbb{Z}$. On a

$$(x^d)^m = 1_G \iff x^{dm} = 1_G \iff o(x) \mid dm \iff s \mid rm \iff s \mid m,$$

la dernière équivalence étant une application du lemme de Gauss.

Ainsi, x^d est d'ordre fini, puisqu'en particulier $(x^d)^s = 1_G$, et la plus petite solution > 0 de $(x^d)^m = 1_G$ est s. On a donc

$$o(x^d) = s = \frac{o(x)}{\mathrm{pgcd}(d, o(x))}.$$

□

Proposition II.7.6. *Soit G un groupe, et soient $x, y \in G$ d'ordre fini tels que $xy = yx$. Alors, xy est d'ordre fini. De plus, on a les propriétés suivantes :*

(1) *supposons que $\langle x\rangle \cap \langle y\rangle = \{1_G\}$. Alors, on a $o(xy) = \mathrm{ppcm}(o(x), o(y))$;*
(2) *si $o(x)$ et $o(y)$ sont premiers entre eux, on a $o(xy) = o(x)o(y)$.*

Démonstration. Soit $m = \mathrm{ppcm}(o(x), o(y))$. On peut donc écrire

$$m = ro(x) = so(y),\ r, s \geq 1.$$

Puisque $xy = yx$, on a

$$(xy)^m = x^m y^m = (x^{o(x)})^r (y^{o(y)})^s = 1_G.$$

Ainsi, xy est d'ordre fini (divisant m).

Supposons que $\langle x\rangle \cap \langle y\rangle = \{1_G\}$. On doit montrer que $o(xy) = m$. On sait déjà que $o(xy)|m$. Soit $\ell = o(xy)$. On a $(xy)^\ell = x^\ell y^\ell = 1_G$. Ainsi, $x^\ell = y^{-\ell}$, et donc $x^\ell \in \langle x\rangle \cap \langle y\rangle$. Comme $\langle x\rangle \cap \langle y\rangle = \{1_G\}$, on obtient

$$x^\ell = 1_G = y^{-\ell}.$$

On en déduit $o(x)|\ell$ et $o(y)|\ell$. Ainsi, $m|\ell$ par définition du ppcm. On a donc l'égalité souhaitée.

Montrons le dernier point. Comme $\langle x\rangle \cap \langle y\rangle$ est un sous-groupe de $\langle x\rangle$ et de $\langle y\rangle$, son ordre divise à la fois $o(x)$ et $o(y)$. Puisqu'ils sont premiers entre eux, on en déduit que $\langle x\rangle \cap \langle y\rangle$ est d'ordre 1, et donc réduit à 1_G. Par le point précédent, on obtient

$$o(xy) = \mathrm{ppcm}(o(x), o(y)) = o(x)o(y),$$

la dernière égalité provenant du fait que $o(x)$ et $o(y)$ sont premiers entre eux. □

II.8. Groupes monogènes, groupes cycliques

On continue ce chapitre en introduisant la notion de groupe monogène et de groupe cyclique.

Définition II.8.1. On dit que G est *monogène* s'il peut être engendré par un élément, et *cyclique* s'il est monogène fini.

Exemples II.8.2.

(1) Le groupe $\mathbb{Z}$ est monogène, non cyclique.
(2) Pour tout $n \geq 1$, le groupe $\mathbb{Z}/n\mathbb{Z}$ est cyclique d'ordre n.

Le théorème suivant élucide la structure des groupes monogènes.

Théorème II.8.3.

(1) *Tout groupe monogène infini est isomorphe à $\mathbb{Z}$.*
(2) *Tout groupe cyclique d'ordre n est isomorphe à $\mathbb{Z}/n\mathbb{Z}$. En particulier, deux groupes cycliques sont isomorphes si, et seulement si, ils ont même ordre.*

Démonstration. Soit $x \in G$ un générateur de G. On a donc $G = \langle x \rangle$. Considérons le morphisme de groupes

$$\begin{aligned} h_x \colon \mathbb{Z} &\longrightarrow G \\ m &\longmapsto x^m. \end{aligned}$$

Puisque x engendre G, h_x est surjectif.

Si G est infini, alors x n'est pas d'ordre fini. Dans ce cas, $\mathrm{Ker}(h_x) = 0$. Sinon, $\mathrm{Ker}(h_x)$ contiendrait un élément strictement positif, et par le théorème **??**, x serait d'ordre fini. Ainsi, h_x est injectif, donc un isomorphisme. Par conséquent, on a $G \simeq \mathbb{Z}$ dans ce cas.

Si G est cyclique d'ordre n, alors $o(x) = n$ par définition, et on a $\mathrm{Ker}(h_x) = n\mathbb{Z}$ par le théorème **??**. Puisque h_x est surjective, le premier théorème d'isomorphisme donne un isomorphisme de groupes $\mathbb{Z}/n\mathbb{Z} \simeq G$. Ceci achève la démonstration. □

Corollaire II.8.4. *Soit p un nombre premier, et soit G un groupe d'ordre p. Alors, G est cyclique. En particulier, $G \simeq \mathbb{Z}/p\mathbb{Z}$.*

Démonstration. Puisque G est d'ordre p, il contient un élément $x \neq 1_G$. En particulier, $o(x) > 1$. Mais par le théorème de Lagrange, $o(x)$ divise p, et donc nécessairement $o(x) = p$. On a donc $\langle x \rangle = G$, et G est cyclique. Il est donc isomorphe à $\mathbb{Z}/p\mathbb{Z}$ d'après le théorème précédent. □

Nous finissons ce paragraphe en nous intéressant aux sous-groupes d'un groupe cyclique.

Théorème II.8.5 ((TD)). *Soit G un groupe cyclique d'ordre n. Pour tout diviseur positif d de n, il existe un unique sous-groupe H_d d'ordre d, et ce sous-groupe est cyclique.*

De plus, si x_0 est un générateur de G, on a les égalités

$$H_d = \langle x_0^{\frac{n}{d}} \rangle = \{x \in G \mid x^d = 1_G\}.$$

II.9. Actions de groupe

Nous allons maintenant définir une notion qui est le cœur même de la théorie des groupes, et qui établit un lien naturel entre théorie des groupes et géométrie.

Définition II.9.1. Soit E un ensemble, et soit G un groupe. Une *action* (ou une *opération*) *à gauche* de G sur E est une application

$$\begin{aligned} G \times E &\longrightarrow E \\ (g, x) &\longmapsto g{\cdot}x \end{aligned}$$

vérifiant les propriétés suivantes :

(1) pour tout $x \in E$, on a $1_G{\cdot}x = x$;
(2) pour tous $g, g' \in G$, et tout $x \in E$, on a $g{\cdot}(g'{\cdot}x) = gg'{\cdot}x$.

Dans ce cas, on dit que G *opère* sur E ou *agit* sur E.

Remarque II.9.2. Supposons que G agisse (à gauche) sur E. Alors, pour tout $g \in G$, et tout $x \in E$, on a

$$g{\cdot}(g^{-1}{\cdot}x) = g^{-1}{\cdot}(g{\cdot}x) = x.$$

En particulier, l'application

$$\begin{aligned} \sigma_g \colon E &\longrightarrow E \\ x &\longmapsto g{\cdot}x \end{aligned}$$

est bijective, d'inverse

$$\begin{aligned} \sigma_{g^{-1}} \colon E &\longrightarrow E \\ x &\longmapsto g^{-1}{\cdot}x. \end{aligned}$$

Autrement dit, on a $\sigma_g \in \mathfrak{S}(E)$. L'application

$$\begin{aligned} \varphi \colon G &\longrightarrow \mathfrak{S}(E) \\ g &\longmapsto \sigma_g \end{aligned}$$

est alors un morphisme de groupes, puisque pour tous $g, g' \in G$, et pour tout $x \in E$, on a

$$\varphi(gg')(x) = gg'{\cdot}x = g{\cdot}(g'{\cdot}x) = g{\cdot}(\varphi(g')(x)) = (\varphi(g) \circ \varphi(g'))(x).$$

Inversement, si $\psi : G \longrightarrow \mathfrak{S}(E)$ est un morphisme de groupes, on vérifie facilement que

$$\begin{aligned} G \times E &\longrightarrow E \\ (g, x) &\longmapsto g{\cdot}x = \psi(g)(x) \end{aligned}$$

est une action de G sur E.

On montre aisément que les deux constructions sont inverses l'une de l'autre. Autrement dit, il y a une correspondance bijective entre les actions de G sur E et les morphismes de groupes de G dans $\mathfrak{S}(E)$.

Dans la suite, l'ensemble E sera implicitement supposé non vide.

Définition II.9.3. Soit G un groupe opérant sur E. On dit que G agit *fidèlement* sur E si pour tout $g \in G$, on a

$$g{\cdot}x = x \text{ pour tout } x \in E \Longrightarrow g = 1_G.$$

Autrement dit, G agit fidèlement sur E si le morphisme $\varphi : G \longrightarrow \mathfrak{S}(E)$ associé est injectif.

On dit que le groupe G agit *transitivement* sur E si pour tous $x, x' \in E$, il existe $g \in G$ tel que $x' = g{\cdot}x$.

Remarque II.9.4. Soit G un groupe, et soit H un sous-groupe de G. Toute action d'un groupe G sur un ensemble E induit par restriction une action de H sur E.

Exemples II.9.5.

(1) Un groupe G agit sur lui-même par translation à gauche, et cette action est fidèle. En effet, pour $g, x \in G$, posons

$$g{\cdot}x = gx.$$

Alors, on a $1_G{\cdot}x = 1_G x = x$, et pour tous $g, g', x \in G$, on a

$$g{\cdot}(g'{\cdot}x) = g{\cdot}(g'x) = g(g'x) = (gg')x = gg'{\cdot}x.$$

Si de plus on a $g{\cdot}x = gx = x$ pour tout $x \in G$, alors on a $g = 1_G$ en prenant $x = 1_G$.

Cette action de G est également transitive. En effet, pour tous $x, y \in G$, on a $g{\cdot}x = y$, avec $g = yx^{-1}$.

(2) Un groupe G agit sur un sous-groupe H distingué dans G, par conjugaison. Autrement dit, l'application

$$\begin{aligned} G \times H &\longrightarrow H \\ (g, x) &\longmapsto g{\cdot}x = gxg^{-1} \end{aligned}$$

est une action de G sur H.

En particulier, G agit sur lui-même par conjugaison.

Cette action n'est pas fidèle en général. Par exemple, si G est abélien, alors $g{\cdot}x = gxg^{-1} = x$ pour tout $g \in G$, et tout $x \in H$.

(3) Pour tout ensemble E, le groupe $\mathfrak{S}(E)$ agit sur E par

$$\begin{aligned} \mathfrak{S}(E) \times E &\longrightarrow E \\ (\sigma, x) &\longmapsto \sigma{\cdot}x = \sigma(x). \end{aligned}$$

Cette action est fidèle et transitive.

(4) Soit $G = \mathrm{GL}_n(\mathbb{C})$ et soit $E = \mathbb{C}^n \setminus \{0\}$. Alors, G agit sur E par

$$\begin{aligned} G \times E &\longrightarrow E \\ (M, v) &\longmapsto Mv, \end{aligned}$$

et cette action est fidèle et transitive.

(5) Soit H un sous-groupe quelconque de G, et soit $E = G/H$ l'ensemble des classes à gauche modulo H. Alors, l'application

$$\begin{aligned} G \times E &\longrightarrow E \\ (g, \overline{g'}) &\longmapsto g{\cdot}\overline{g'} = \overline{gg'} \end{aligned}$$

est bien définie, et donne une action transitive de G sur E.

Vérifions que l'application est bien définie. Si $g', g'' \in G$ sont deux éléments de G tels que $\overline{g''} = \overline{g'}$, alors il existe $h \in H$ tel que $g'' = g'h$, et on a $gg'' = gg'h$. Ainsi, $\overline{gg''} = \overline{gg'}$, ce qu'il fallait montrer. Le fait que l'on obtienne une action transitive de G sur E est alors clair.

(6) Soit G un groupe agissant sur E, et soit $g \in G$. Alors, le groupe $\mathbb{Z}$ agit sur E par

$$\begin{aligned} \mathbb{Z} \times E &\longrightarrow E \\ (m, x) &\longmapsto m * x = g^m{\cdot}x. \end{aligned}$$

Cette action n'est en général ni fidèle, ni transitive.

(7) Soit $n \geq 1$ un entier, et soit E un ensemble. Alors, $\mathfrak{S}_n$ agit sur E^n par

$$\begin{aligned} \mathfrak{S}_n \times E^n &\longrightarrow E^n \\ (\sigma, (x_1, \ldots, x_n)) &\longmapsto (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)}). \end{aligned}$$

Le lecteur consciencieux vérifiera que prendre l'inverse σ^{-1} est bien nécessaire pour avoir une action de groupe.

Nous laissons le soin au lecteur de détailler ces exemples, à titre d'exercice.

L'intérêt de faire agir un groupe G sur un ensemble est multiple. Cela peut permettre de comprendre la structure de G en étudiant le morphisme de groupes $\varphi : G \longrightarrow \mathfrak{S}(E)$ associé. Par exemple, le noyau de φ est le sous-groupe des éléments $g \in G$ qui agissent trivialement sur E. D'autre part, beaucoup de groupes sont de nature géométrique, et définis comme groupes de transformations préservant globalement une figure géométrique (un bon exemple est donné par le Rubik's cube ou autres puzzles apparentés). Les actions de groupes permettent alors de mieux visualiser certains sous-groupes en les interprétant comme le stabilisateur d'un élément de E.

Définition II.9.6. L'ensemble

$$\mathrm{Stab}_G(x) = \{g \in G \mid g{\cdot}x = x\}$$

est appelé le *stabilisateur* de x (sous l'action de G). On vérifie facilement que c'est un sous-groupe de G.

Un autre intérêt majeur d'avoir une action de groupe est qu'elle permet de définir naturellement une relation d'équivalence sur E, comme on va le voir maintenant.

LEMME II.9.7. *Soit G opérant sur un ensemble E. La relation sur E définie par*

$$x \sim y \text{ s'il existe } g \in G \text{ tel que } y = g{\cdot}x$$

est une relation d'équivalence.

Démonstration. Pour tout $x \in E$, on a $x = 1_G{\cdot}x$, et donc $x \sim x$. Si maintenant $y = g{\cdot}x$ pour un certain $g \in G$, alors

$$g^{-1}{\cdot}y = g^{-1}{\cdot}(g{\cdot}x) = x.$$

Ainsi, on a $y \sim x$ dès que $x \sim y$. Enfin, si $y = g{\cdot}x$ et $z = g'{\cdot}y$, alors

$$z = g'{\cdot}(g{\cdot}x) = g'g{\cdot}x,$$

et la relation est donc transitive. □

DÉFINITION II.9.8. Soit G opérant sur un ensemble E, et soit $x \in E$. L'ensemble

$$\mathcal{O}_x = \{g{\cdot}x \mid g \in G\} \subset E$$

est appelé *l'orbite* de x sous l'action de G, ou *G-orbite* de x. C'est la classe d'équivalence de x pour la relation d'équivalence sur E induite par l'action de G sur E.

On dit que $x \in E$ est un *point fixe* sous l'action de G si l'on a $g{\cdot}x = x$ pour tout $g \in G$. L'ensemble des points fixes est noté E^G.

EXEMPLE II.9.9. Si G agit sur lui-même par conjugaison, les orbites correspondantes sont les classes de conjugaison de G.

REMARQUES II.9.10.

(1) Soit G un groupe agissant sur E, et soit $x \in E$. Alors, les propriétés suivantes sont équivalentes :
   - (i) x est un point fixe sous l'action de G ;
   - (ii) $\mathcal{O}_x = \{x\}$;
   - (iii) $|\mathcal{O}_x| = 1$;
   - (iv) $\mathrm{Stab}_G(x) = G$.

(2) Soit G un groupe agissant sur E. Alors, les propriétés suivantes sont équivalentes :
   - (i) G agit transitivement sur E ;
   - (ii) pour tout $x \in E$, $\mathcal{O}_x = E$;
   - (iii) il existe un $x_0 \in E$ tel que $\mathcal{O}_{x_0} = E$;
   - (iv) E n'admet qu'une seule G-orbite, à savoir E lui-même.

Ceci est laissé en exercice au lecteur.

Le lemme suivant découle immédiatement de la définition d'une orbite et des propriétés des relations d'équivalences. Il sera constamment utilisé dans la suite, sans référence ultérieure.

LEMME II.9.11 (Équation aux classes). *Soit G un groupe agissant sur E. Alors, les différentes orbites de E sous l'action de G forment une partition de E. Si de plus E est fini, le cardinal de E est la somme des cardinaux des différentes orbites. Autrement dit, si Ω est l'ensemble des orbites de E sous l'action de G, on a*

$$|E| = \sum_{\omega \in \Omega} |\omega|.$$

Nous allons maintenant faire le lien entre orbites et stabilisateurs.

Proposition II.9.12. *Soit G un groupe opérant sur E. Pour tout $x \in E$, l'application*

$$\begin{aligned} f\colon G &\longrightarrow \mathcal{O}_x \\ g &\longmapsto g{\cdot}x \end{aligned}$$

induit par passage au quotient une bijection

$$\begin{aligned} \overline{f}\colon G/\mathrm{Stab}_G(x) &\xrightarrow{\sim} \mathcal{O}_x \\ \overline{g} &\longmapsto g{\cdot}x. \end{aligned}$$

En particulier, $\mathcal{O}_x$ est un ensemble fini si, et seulement si, $[G : \mathrm{Stab}_G(x)]$ est fini.

Enfin, si G est fini , pour tout $x \in G$, l'ensemble $\mathcal{O}_x$ est fini et son cardinal divise l'ordre de G. On a alors

$$|\mathcal{O}_x| = \frac{|G|}{|\mathrm{Stab}_G(x)|}.$$

Démonstration. L'application

$$\begin{aligned} f\colon G &\longrightarrow \mathcal{O}_x \\ g &\longmapsto g{\cdot}x \end{aligned}$$

est constante sur les classes à gauche modulo $\mathrm{Stab}_G(x)$.

En effet, si $s \in \mathrm{Stab}_G(x)$ et $g \in G$, on a alors

$$f(gs) = gs{\cdot}x = g{\cdot}(s{\cdot}x) = g{\cdot}x = f(g).$$

Comme l'ensemble des classes à gauche $G/\mathrm{Stab}_G(x)$ est l'ensemble des classes d'équivalence pour la relation associée au sous-groupe $\mathrm{Stab}_G(x)$, la proposition **??** entraîne que f induit une application

$$\begin{aligned} \overline{f}\colon G/\mathrm{Stab}_G(x) &\longrightarrow \mathcal{O}_x \\ \overline{g} &\longmapsto g{\cdot}x. \end{aligned}$$

Par définition de l'orbite de x, l'application $\overline{f}$ est surjective. Montrons qu'elle est aussi injective. Soient $\overline{g}, \overline{g'} \in G/\mathrm{Stab}_G(x)$ tels que $g{\cdot}x = g'{\cdot}x$. On a alors

$$x = g^{-1}{\cdot}(g{\cdot}x) = g^{-1}{\cdot}(g'{\cdot}x) = g^{-1}g'{\cdot}x.$$

Ainsi, $g^{-1}g' \in \mathrm{Stab}_G(x)$ et donc $\overline{g} = \overline{g'}$. L'application $\overline{f}$ est donc bijective. Le reste de la proposition provient du théorème de Lagrange. □

On continue ce paragraphe en étudiant brièvement les orbites de E sous l'action d'un groupe monogène.

Définition II.9.13. Soit G un groupe agissant sur un ensemble E.

Pour tout $g \in G$, et pour tout $x \in E$, la *g-orbite*de x est l'orbite de x sous l'action du sous-groupe $\langle g \rangle$.

Lemme II.9.14. *Soit G un groupe agissant sur un ensemble E. Soit $g \in G$, et soit ω la g-orbite d'un élément $x \in E$. Alors, ω est finie si, et seulement si, il existe $k \in \mathbb{Z}$ non nul tel que $g^k{\cdot}x = x$. Dans ce cas, on a*

$$\omega = \{x, g{\cdot}x, \ldots, g^{p-1}{\cdot}x\},$$

où p est le plus petit entier strictement positif tel que $g^p{\cdot}x = x$. De plus, on a $|\omega| = p$.

Démonstration. Remarquons que par définition, on a

$$\omega = \{g^m \cdot x \mid m \in \mathbb{Z}\}.$$

Ce lemme peut se démontrer de la même façon que la première partie du théorème **??**. Néanmoins, nous allons proposer ici une démonstration utilisant les actions de groupes.

Remarquons que l'application

$$\begin{aligned} \mathbb{Z} \times E &\longrightarrow E \\ (m, x') &\longmapsto g^m \cdot x' \end{aligned}$$

définit une action de $\mathbb{Z}$ sur E, et que l'orbite de x pour cette action est exactement ω. Par conséquent, d'après la proposition **??**, ω est un ensemble fini si, et seulement si, $\mathrm{Stab}_{\mathbb{Z}}(x)$ est d'indice fini dans $\mathbb{Z}$. Or, l'ensemble $\mathrm{Stab}_{\mathbb{Z}}(x)$ est un sous-groupe de $\mathbb{Z}$. Par la proposition **??**, il est donc de la forme $p\mathbb{Z}, p \geq 0$, et si $p > 0$, alors p est le plus petit élément strictement positif de $\mathrm{Stab}_{\mathbb{Z}}(x)$. D'après l'exemple **??**, $p\mathbb{Z}$ est d'indice fini si, et seulement si, $p > 0$, i.e. si, et seulement si, $\mathrm{Stab}_{\mathbb{Z}}(x)$ est non trivial. Cela revient à dire qu'il existe $k \in \mathbb{Z}$ non nul tel que $g^k \cdot x = x$. Dans ce cas, on a

$$|\omega| = [\mathbb{Z} : \mathrm{Stab}_{\mathbb{Z}}(x)] = [\mathbb{Z} : p\mathbb{Z}] = p.$$

La bijection entre $\mathbb{Z}/p\mathbb{Z}$ et ω est alors donnée par

$$\begin{aligned} \mathbb{Z}/p\mathbb{Z} &\longrightarrow \omega \\ \overline{m} &\longmapsto g^m \cdot x, \end{aligned}$$

et comme les éléments distincts de $\mathbb{Z}/p\mathbb{Z}$ sont les classes $\overline{0}, \ldots, \overline{p-1}$, on a la conclusion. $\square$

Chapitre III

Groupe symétrique, groupe alterné

On va maintenant illustrer les notions introduites dans les chapitres précédents sur le groupe symétrique.

III.1. Préliminaires

Définition III.1.1. Soit E un ensemble non vide. L'ensemble des bijections de E sur lui-même est un groupe pour la composition des applications, appelé *groupe des permutations de* E ou *groupe symétrique sur* E, et est noté $\mathfrak{S}(E)$. Un élément de $\mathfrak{S}(E)$ est appelé une *permutation*.

Le groupe des permutations de l'ensemble $[\![1, n]\!]$ est noté $\mathfrak{S}_n$. Si $\sigma \in \mathfrak{S}_n$, on la note (provisoirement)

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

Par exemple, $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}$ est la permutation $\sigma \in \mathfrak{S}_4$ définie par

$$\sigma(1) = 3, \sigma(2) = 2, \sigma(3) = 1, \sigma(4) = 4.$$

Lemme III.1.2. *Soient E, E' deux ensembles non vides. Si E et E' sont en bijection, alors $\mathfrak{S}(E) \simeq \mathfrak{S}(E')$.*

Démonstration. Soit $f : E \longrightarrow E'$ une bijection de E sur E'. On vérifie facilement que l'application

$$\begin{aligned} u \colon \mathfrak{S}(E) &\longrightarrow \mathfrak{S}(E') \\ \sigma &\longmapsto f \circ \sigma \circ f^{-1} \end{aligned}$$

est un isomorphisme de groupes. Les détails sont laissés au lecteur. □

Exemple III.1.3. Soit E un ensemble à n éléments, et soient $a_1, \ldots, a_n$ les éléments de E. L'application

$$\begin{aligned} f \colon [\![1, n]\!] &\longrightarrow E \\ i &\longmapsto a_i \end{aligned}$$

est alors bijective, et induit donc un isomorphisme de groupes

$$\begin{aligned} u \colon \mathfrak{S}_n &\xrightarrow{\sim} \mathfrak{S}(E) \\ \sigma &\longmapsto f \circ \sigma \circ f^{-1}. \end{aligned}$$

Si $\sigma \in \mathfrak{S}_n$, alors on voit facilement que

$$u(\sigma)(a_i) = a_{\sigma(i)} \text{ pour } i \in [\![1, n]\!].$$

Le but de ce chapitre est d'étudier $\mathfrak{S}(E)$ lorsque E est un ensemble fini ayant au moins deux éléments (sinon $\mathfrak{S}(E)$ est le groupe trivial). Rappelons que si E a n éléments, alors $\mathfrak{S}(E)$ est d'ordre $n!$.

Au vu du lemme précédent, on peut se restreindre à l'étude du groupe $\mathfrak{S}_n$, ce que nous ferons parfois.

Définition III.1.4. Soit E un ensemble non vide. Le *support* d'une permutation $\sigma \in \mathfrak{S}(E)$ est l'ensemble

$$\mathrm{Supp}(\sigma) = \{a \in E \mid \sigma(a) \neq a\}.$$

Autrement dit, $\mathrm{Supp}(\sigma) = E \setminus \mathrm{Fix}(\sigma)$, où $\mathrm{Fix}(\sigma)$ est l'ensemble des points fixes de σ.

Le support de σ est vide si, et seulement si, σ est l'identité.

Le lemme suivant résume les propriétés du support.

Lemme III.1.5. *Soit E un ensemble non vide.*

(1) *Pour tout $\sigma \in \mathfrak{S}(E)$, l'ensemble* $\mathrm{Supp}(\sigma)$ *est stable par σ. On a même*

$$\sigma(\mathrm{Supp}(\sigma)) = \mathrm{Supp}(\sigma).$$

(2) *Pour tout $\sigma \in \mathfrak{S}(E)$, on a* $\mathrm{Supp}(\sigma) = \mathrm{Supp}(\sigma^{-1})$.

(3) *Pour tout $\sigma \in \mathfrak{S}(E)$, et pour tout $m \in \mathbb{Z}$, on a*

$$\mathrm{Supp}(\sigma^m) \subset \mathrm{Supp}(\sigma).$$

(4) *Pour tous $\sigma, \sigma' \in \mathfrak{S}(E)$, on a*

$$\mathrm{Supp}(\sigma\sigma') \subset \mathrm{Supp}(\sigma) \cup \mathrm{Supp}(\sigma').$$

Si de plus σ et σ' sont à supports disjoints, alors

$$\mathrm{Supp}(\sigma\sigma') = \mathrm{Supp}(\sigma) \cup \mathrm{Supp}(\sigma').$$

Démonstration.

(1) Pour tout $a \in E$, on a $\sigma(\sigma(a)) = \sigma(a) \iff \sigma(a) = a$. On en déduit le résultat voulu par passage au complémentaire.

(2) Pour tout $a \in E$, on a

$$\sigma(a) = a \iff a = \sigma^{-1}(a).$$

On conclut comme précédemment.

(3) C'est clair par passage au complémentaire, puisque tout élément fixé par σ est fixé aussi par σ^{-1}, et donc par σ^m (quelque soit le signe de m).

(4) Soient $\sigma, \sigma' \in \mathfrak{S}(E)$. Si $a \in E$ est fixé par σ et σ', il est fixé par $\sigma\sigma'$, d'où la première inclusion par passage au complémentaire.

Supposons maintenant que σ et σ' soient à supports disjoints.

Soit $a \in \mathrm{Supp}(\sigma) \cup \mathrm{Supp}(\sigma')$.

Si $a \in \mathrm{Supp}(\sigma)$, on a $\sigma(a) \neq a$. Par hypothèse, a est alors fixé par σ', et donc

$$\sigma\sigma'(a) = \sigma(a) \neq a.$$

Ainsi, $a \in \mathrm{Supp}(\sigma\sigma')$.

Si $a \in \mathrm{Supp}(\sigma')$, alors $\sigma'(a) \neq a$. De plus, $\sigma'(a) \in \mathrm{Supp}(\sigma')$ par (1), et $\sigma'(a)$ est alors fixé par σ par hypothèse. Ainsi, on a

$$\sigma(\sigma'(a)) = \sigma'(a) \neq a,$$

ce qui prouve que $a \in \mathrm{Supp}(\sigma\sigma')$, et établit l'égalité voulue. □

Le groupe $\mathfrak{S}(E)$ n'est pas abélien pour dès que E possède au moins trois éléments. En revanche, on a le résultat suivant.

Lemme III.1.6. *Deux permutations à supports disjoints commutent.*

Démonstration. Soient $\sigma, \sigma' \in \mathfrak{S}(E)$ deux permutations à supports disjoints, et soit $a \in E$. Si a est fixé par σ et par σ', alors on a

$$\sigma(\sigma'(a)) = \sigma(a) = a = \sigma'(a) = \sigma'(\sigma(a)).$$

Si $a \in \mathrm{Supp}(\sigma)$, alors $\sigma(a) \in \mathrm{Supp}(\sigma)$ d'après le lemme **??** (1). Par hypothèse, $\sigma(a)$ n'est pas dans le support de σ', et donc $\sigma'(\sigma(a)) = \sigma(a)$. Mais, puisque $a \in \mathrm{Supp}(\sigma)$, a n'est pas dans le support de σ', et donc on a

$$\sigma'(\sigma(a)) = \sigma(a) = \sigma(\sigma'(a)).$$

On raisonne de la même manière lorsque $a \in \mathrm{Supp}(\sigma')$. □

III.2. Décomposition en produit de cycles

Dans tout ce qui suit, E est un ensemble fini à n éléments, où $n \geq 2$.

Soit $\sigma \in \mathfrak{S}(E)$. On rappelle que l'action naturelle de $\mathfrak{S}(E)$ sur E se restreint en une action du sous-groupe $\langle \sigma \rangle$ sur E, et qu'une orbite de E sous cette action s'appelle une *σ-orbite* de E. Si $a \in E$, on notera $\mathrm{Orb}_\sigma(a)$ la σ-orbite correspondante. Par définition, on a donc

$$\mathrm{Orb}_\sigma(a) = \{\sigma^m(a) \mid m \in \mathbb{Z}\}.$$

Il est clair que $\mathrm{Orb}_\sigma(a)$ est réduite à un singleton si, et seulement si, a est un point fixe de σ. Les différentes σ-orbites de E forment alors une partition de E, et la réunion des σ-orbites non réduites à un singleton est égale au support de σ. Puisque $\mathfrak{S}(E)$ est fini, tout sous-groupe est d'indice fini, et le lemme **??** nous donne immédiatement le résultat suivant.

Lemme III.2.1. *Soit $\sigma \in \mathfrak{S}(E)$ une permutation, soit ω une σ-orbite à p éléments, et soit $a \in \omega$. Alors, p est le plus petit entier strictement positif tel que $\sigma^p(a) = a$ et on a*

$$\omega = \{a, \sigma(a), \ldots, \sigma^{p-1}(a)\}.$$

On définit maintenant la notion de cycle.

Définition III.2.2. On dit qu'une permutation $\sigma \in \mathfrak{S}(E)$ est un *cycle* s'il n'existe qu'une seule σ-orbite non réduite à un singleton (nécessairement égale à son support). Si $p \geq 2$ est le nombre d'éléments de $\mathrm{Supp}(\sigma)$, on dit que σ est un p-cycle.

Un 2-cycle est appelé une *transposition.* Autrement dit, une transposition de $\mathfrak{S}(E)$ a $n-2$ points fixes et échange deux éléments.

Exemple III.2.3. Soit $p \geq 2$, et soient $a_1, \ldots, a_p \in E$ p éléments distincts. La permutation $\sigma \in \mathfrak{S}(E)$ définie par

$$\sigma(a_1) = a_2, \ldots, \sigma(a_{p-1}) = a_p, \sigma(a_p) = a_1, \sigma(a) = a, a \neq a_i,\ i \in [\![1, p]\!]$$

est un p-cycle de support $\{a_1, \ldots, a_p\}$.

Pour le voir, il faut montrer qu'il y a une unique σ-orbite non réduite à un élément. Soit $a \in E$. Si $a \neq a_i, i \in [\![1, p]\!]$, alors a est un point fixe sous σ, et sa σ-orbite est réduite à un élément. Pour conclure, il suffit de voir que les éléments $a_1, \ldots, a_p \in E$ forment une seule σ-orbite. Mais ces éléments sont tous dans la σ-orbite de a_1 puisque l'on a

$$\sigma^{i-1}(a_1) = a_i, \quad \text{pour tout } i \in [\![1, p]\!],$$

d'où le résultat.

Notation. Le p-cycle de l'exemple précédent sera noté $(a_1 \cdots a_p)$.

Exemple III.2.4. Soit $E = [\![1, 4]\!]$. Alors, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$ est un 3-cycle. C'est en fait le 3-cycle $(1\ 3\ 2)$. Remarquons que l'on a aussi

$$\sigma = (1\ 3\ 2) = (2\ 1\ 3) = (3\ 2\ 1).$$

En général, on décide de démarrer par le plus petit élément dans le support. De même, $\tau = (2\ 4) = (4\ 2)$ est une transposition, définie par

$$\tau(1) = 1, \tau(2) = 4, \tau(3) = 3, \tau(4) = 2.$$

On va maintenant montrer que toute permutation de $\mathfrak{S}(E)$ se décompose en produit de cycles à supports disjoints. L'idée est que pour connaître une permutation, il suffit de connaître son action sur chaque orbite non réduite à un élément. La première étape est d'associer à chaque orbite ω un cycle de support ω.

Soit $\sigma \in \mathfrak{S}(E)$. Soit Ω^* l'ensemble des σ-orbites non réduites à un élément. Si $\omega \in \Omega^*$, on pose

$$\sigma_\omega(a) = \sigma(a) \text{ si } a \in \omega \quad \text{et} \quad \sigma_\omega(a) = a \text{ si } a \notin \omega.$$

Lemme III.2.5. *Soit $\sigma \in \mathfrak{S}(E)$, et soit $\omega \in \Omega^*$ une orbite non réduite à un élément. Alors, σ_ω est un cycle de support ω, et pour tout $a \in \omega$, on a*

$$\sigma_\omega = (a\ \sigma(a) \cdots \sigma^{p-1}(a)),$$

où p est le nombre d'éléments de ω.

De plus, tout p-cycle de $\mathfrak{S}(E)$ est de cette forme.

Démonstration. Soit $\omega \in \Omega^*$ une σ-orbite de cardinal $p > 1$. Soit $a \in \omega$. Par le lemme **??**, p est le plus petit entier > 0 tel que $\sigma^p(a) = a$, et on a

$$\omega = \{a, \sigma(a), \ldots, \sigma^{p-1}(a)\}.$$

On va montrer l'égalité

$$\sigma_\omega = (a\ \sigma(a) \cdots \sigma^{p-1}(a)).$$

Soit $a' \in E$. Si $a' \in E \setminus \{a, \sigma(a), \ldots, \sigma^{p-1}(a)\}$, alors $a' \notin \omega$ et ainsi $\sigma_\omega(a') = a'$ par définition. De plus, on a

$$\sigma_\omega(\sigma^i(a)) = \sigma^{i+1}(a) \quad \text{pour tout } i \in [\![0, p-1]\!]$$

puisque $\sigma^i(a) \in \omega$. Puisque $\sigma^p(a) = a$, cela montre l'égalité voulue.

D'après l'exemple **??**, le support de σ_ω est alors $\{a, \sigma(a), \ldots, \sigma^{p-1}(a)\}$, c'est-à-dire ω (on peut aussi raisonner de manière directe).

Si maintenant $\sigma \in \mathfrak{S}(E)$ est un p-cycle, posons $\omega = \mathrm{Supp}(\sigma)$. Montrons que $\sigma_\omega = \sigma$. Si $a \in \omega$, on a $\sigma_\omega(a) = \sigma(a)$ par définition de σ_ω, et si $a \notin \omega$, on a $\sigma_\omega(a) = a = \sigma(a)$, la dernière égalité provenant du fait que $E \setminus \omega$ est l'ensemble des points fixes de σ. □

Remarque III.2.6. Le lemme précédent montre que tout p-cycle de $\mathfrak{S}(E)$ est de la forme $(a_1 \cdots a_p)$, où $a_1, \ldots, a_p$ sont p éléments distincts de E.

On a alors le théorème suivant.

Théorème III.2.7. *Soit $\sigma \in \mathfrak{S}(E)$. Alors, σ se décompose en produit de cycles à supports disjoints, et cette décomposition est unique à l'ordre des facteurs près. Cette décomposition est donnée par*

$$\sigma = \prod_{\omega \in \Omega^*} \sigma_\omega,$$

où Ω^* est l'ensemble des σ-orbites non réduites à un singleton.

Démonstration. Soit $\sigma \in \mathfrak{S}(E)$. Montrons l'égalité

$$\sigma = \prod_{\omega \in \Omega^*} \sigma_\omega.$$

Remarquons tout d'abord que les σ_ω sont des cycles dont les supports sont disjoints deux à deux. En effet, $\mathrm{Supp}(\sigma_\omega) = \omega$, et deux σ-orbites distinctes sont disjointes.

Soit $a \in E$. Si a est un point fixe de σ, alors $a \notin \omega$ pour tout $\omega \in \Omega^*$ (sinon ω serait réduite à un élément). Ainsi, $\sigma_\omega(a) = a$ pour tout $\omega \in \Omega^*$ et donc

$$\prod_{\omega \in \Omega^*} \sigma_\omega(a) = a = \sigma(a).$$

Si a est dans le support de σ, alors $\mathrm{Orb}_\sigma(a) \in \Omega^*$ et donc il existe $\omega_0 \in \Omega^*$ tel que $\mathrm{Orb}_\sigma(a) = \omega_0$. On a donc $a \in \omega_0$, et $a \notin \omega$ pour tout $\omega \in \Omega^* \setminus \{\omega_0\}$, car deux σ-orbites distinctes sont disjointes. Par conséquent, $\sigma_\omega(a) = a$ pour tout $\omega \neq \omega_0$. Comme les supports des cycles σ_ω sont disjoints, ils commutent. Ainsi, on a

$$\prod_{\omega \in \Omega^*} \sigma_\omega(a) = \sigma_{\omega_0}\big(\prod_{\omega \in \Omega^* \setminus \{\omega_0\}} \sigma_\omega(a)\big) = \sigma_{\omega_0}(a) = \sigma(a),$$

la dernière égalité provenant de la définition de σ_{ω_0}.

Montrons maintenant l'unicité de la décomposition. Supposons que l'on ait une décomposition

$$\sigma = \sigma_1 \cdots \sigma_r,$$

où les σ_i sont des cycles à supports disjoints. On note E_i le support du cycle σ_i. On va montrer que E_i est une σ-orbite non réduite à un élément.

Soit $a \in E_i$. Comme les supports des σ_i sont disjoints, on a $\sigma_j(a) = a$ dès que $j \neq i$. De plus, puisque les σ_j commutent, on a

$$\sigma(a) = \prod_j \sigma_j(a) = \sigma_i \prod_{j \neq i} \sigma_j(a) = \sigma_i(a).$$

On a donc $\sigma_i(a) = \sigma(a)$ si $a \in E_i$. Puisque $E_i = \mathrm{Supp}(\sigma_i)$ est stable par σ_i par le lemme **??** (1), pour tout $a \in E_i$, on a donc $\sigma(a) = \sigma_i(a) \in E_i$. Par récurrence, on a $\sigma_i^k(a) = \sigma^k(a), k \geq 0$.

Comme σ_i est un cycle de support E_i, alors E_i forme une σ_i-orbite non réduite à un élément. Le lemme **??** appliqué à σ_i montre alors que l'on a $E_i = \{a, \sigma_i(a), \ldots, \sigma_i^{p_i-1}(a)\}$, où p_i est le cardinal de E_i. Par ce qui précède, on obtient donc

$$E_i = \{a, \sigma(a), \ldots, \sigma^{p_i-1}(a)\},$$

ce qui montre que les éléments de E_i forment une σ-orbite (à savoir celle de a) non réduite à un élément.

Remarquons alors que, puisque le support de σ est la réunion des E_i, on obtient bien toutes les orbites non réduites à un élément. Ceci prouve l'unicité de la décomposition, et achève la démonstration. □

Remarque III.2.8. Les deux résultats précédents fournissent l'égalité suivante : soit $\sigma \in \mathfrak{S}(E)$, et soient $\omega_1, \ldots, \omega_r$ les différentes σ-orbites non réduites à un élément. Soit $a_i \in \omega_i$, et soit p_i le cardinal de ω_i.

Alors, on a

$$\sigma = (a_1\ \sigma(a_1) \cdots \sigma^{p_1-1}(a_1)) \cdots (a_r\ \sigma(a_r) \cdots \sigma^{p_r-1}(a_r)).$$

Tout ceci donne une méthode systématique pour trouver la décomposition de σ. Pour fixer les idées, supposons que $E = [\![1, n]\!]$. Si $\sigma \in \mathfrak{S}_n$, on regarde l'image de 1 par les puissances successives de σ jusqu'à reboucler sur 1. Ensuite, on prend un autre entier qui n'est pas encore atteint et on recommence le procédé, jusqu'à épuiser tous les entiers de 1 à n.

Exemple III.2.9. Soit $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 2 & 4 & 1 & 3 \end{pmatrix} \in \mathfrak{S}_6$.

Alors, on a

$$\sigma(1) = 5 \quad \text{et} \quad \sigma(5) = 1.$$

On obtient un premier cycle de longueur 2, à savoir $(1\ 5)$.

On choisit maintenant un entier distinct de 1 et 5, disons 2. On a

$$\sigma(2) = 6,\ \sigma(6) = 3, \quad \text{et} \quad \sigma(3) = 2,$$

et on obtient le cycle $(2\ 6\ 3)$.

Le seul entier non encore utilisé est 4. Or, on a $\sigma(4) = 4$. Ainsi, 4 n'est pas dans le support de σ. On a donc

$$\sigma = (1\ 5)(2\ 6\ 3).$$

Le théorème **??** entraîne immédiatement le résultat suivant.

Corollaire III.2.10. *Soit E un ensemble fini. Le groupe $\mathfrak{S}(E)$ est engendré par les cycles.*

Au vu de ce résultat, on donne souvent un élément de $\mathfrak{S}(E)$ sous forme de produits de cycles, pas nécessairement à supports disjoints. Il faut vite s'habituer à calculer l'image par σ à partir d'une telle décomposition et à en déduire une décomposition en produits de cycles à supports disjoints.

Exemple III.2.11. Soit $\sigma \in \mathfrak{S}_8$, donné par

$$\sigma = (1\ 2\ 3\ 5)(3\ 7)(7\ 4\ 8).$$

Attention ! On compose des applications, donc on calcule à partir de la droite.

$$\begin{array}{ccccccc}
 & (7\ 4\ 8) & & (3\ 7) & & (1\ 2\ 3\ 5) & \\
\underline{1} & \longrightarrow & 1 & \longrightarrow & 1 & \longrightarrow & \underline{2}, \\
\underline{2} & \longrightarrow & 2 & \longrightarrow & 2 & \longrightarrow & \underline{3}, \\
\underline{3} & \longrightarrow & 3 & \longrightarrow & 7 & \longrightarrow & \underline{7}, \\
\underline{7} & \longrightarrow & 4 & \longrightarrow & 4 & \longrightarrow & \underline{4}, \\
\underline{4} & \longrightarrow & 8 & \longrightarrow & 8 & \longrightarrow & \underline{8}, \\
\underline{8} & \longrightarrow & 7 & \longrightarrow & 3 & \longrightarrow & \underline{5}, \\
\underline{5} & \longrightarrow & 5 & \longrightarrow & 5 & \longrightarrow & \underline{1}.
\end{array}$$

On obtient un premier cycle $(1\ 2\ 3\ 7\ 4\ 8\ 5)$. Il reste donc à regarder l'image de 6, qui est fixe. On a donc $\sigma = (1\ 2\ 3\ 7\ 4\ 8\ 5)$.

Corollaire III.2.12. *Le groupe $\mathfrak{S}(E)$ est engendré par les transpositions.*

Démonstration. Puisque toute permutation est produit de cycles, il suffit de montrer qu'un cycle est un produit de transpositions. Or, un simple calcul montre que l'on a

$$(a_1 \cdots a_p) = (a_1\ a_2)(a_2\ a_3) \cdots (a_{p-2}\ a_{p-1})(a_{p-1}\ a_p),$$

d'où le résultat. □

On continue ce paragraphe par un résultat qui calcule l'ordre d'une permutation quelconque. On commence par démontrer deux lemmes.

LEMME III.2.13. *Un p-cycle est d'ordre p.*

Démonstration. Soit $\sigma = (a_1 \cdots a_p)$. Pour $j \in [\![1, p-1]\!]$, on a

$$\sigma^j(a_1) = a_{j+1} \neq a_1.$$

Donc $\sigma, \sigma^2, \ldots, \sigma^{p-1}$ ne sont pas égaux à Id_E. En revanche, pour $k \in [\![1, p]\!]$, on a

$$\sigma^p(a_k) = \sigma^p(\sigma^{k-1}(a_1)) = \sigma^{k-1}(\sigma^p(a_1)) = \sigma^{k-1}(a_1) = a_k,$$

d'où $\sigma^p = \mathrm{Id}_E$. Ainsi, $o(\sigma) = p$ par le théorème **??**. □

LEMME III.2.14. *Soient $\sigma_1, \ldots, \sigma_r \in \mathfrak{S}(E)$ des permutations à supports deux à deux disjoints. Alors, on a*

$$o(\sigma_1 \cdots \sigma_r) = \mathrm{ppcm}(o(\sigma_1), \ldots, o(\sigma_r)).$$

Démonstration. Procédons par récurrence sur r. Si $r = 1$, le résultat est vrai. Montrons-le également pour $r = 2$. Soient $\sigma_1, \sigma_2 \in \mathfrak{S}(E)$ deux permutations à supports disjoints. Par le lemme **??**, σ_1 et σ_2 commutent. D'autre part, on a

$$\langle \sigma_1 \rangle \cap \langle \sigma_2 \rangle = \{\mathrm{Id}_E\}.$$

En effet, soit $\sigma \in \langle \sigma_1 \rangle \cap \langle \sigma_2 \rangle$. En utilisant le lemme **??** (3), on obtient que

$$\mathrm{Supp}(\sigma) \subset \mathrm{Supp}(\sigma_1) \cap \mathrm{Supp}(\sigma_2).$$

L'hypothèse implique alors que σ est de support vide, c'est-à-dire $\sigma = \mathrm{Id}_E$. D'après la proposition **??**, on en déduit l'égalité

$$o(\sigma_1 \sigma_2) = \mathrm{ppcm}(o(\sigma_1), o(\sigma_2)).$$

Supposons maintenant que l'égalité soit vraie au rang $r \geq 1$, et montrons-le au rang $r+1$. Si $\sigma_1, \ldots, \sigma_{r+1} \in \mathfrak{S}(E)$ sont à supports deux à deux disjoints, alors $\sigma_1 \cdots \sigma_r$ et σ_{r+1} sont aussi à supports disjoints.

En effet, une utilisation répétée du lemme **??** (4) montre que l'on a [1]

$$\mathrm{Supp}(\sigma_1 \cdots \sigma_r) \subset \mathrm{Supp}(\sigma_1) \cup \cdots \cup \mathrm{Supp}(\sigma_r).$$

L'hypothèse implique que l'ensemble de droite n'intersecte pas $\mathrm{Supp}(\sigma_{r+1})$, d'où le fait annoncé. Mais alors, le cas $r = 2$ nous donne l'égalité

$$o(\sigma_1 \cdots \sigma_{r+1}) = o((\sigma_1 \cdots \sigma_r)\sigma_{r+1}) = \mathrm{ppcm}(o(\sigma_1 \cdots \sigma_r), o(\sigma_{r+1})).$$

Il suffit alors d'appliquer l'hypothèse de récurrence et les propriétés d'associativité du ppcm pour conclure. □

Les deux lemmes précédents entraînent alors immédiatement le résultat suivant.

THÉORÈME III.2.15. *L'ordre d'une permutation est le ppcm des longueurs des cycles à supports disjoints qui la composent.*

Ainsi, pour calculer l'ordre de $\sigma \in \mathfrak{S}(E)$, il suffit donc de la décomposer en produit de cycles à supports disjoints, et d'appliquer le théorème précédent, ce qui est beaucoup plus rapide que de chercher le plus petit entier $p \geq 1$ tel que $\sigma^p = \mathrm{Id}_E$.

On termine ce paragraphe en déterminant les classes de conjugaison du groupe $\mathfrak{S}(E)$.

1. Il y a même égalité, mais nous n'en aurons pas besoin.

Lemme III.2.16. *Soit $\tau \in \mathfrak{S}(E)$. Pour tout p-cycle $(a_1 \cdots a_p)$, on a*

$$\tau(a_1 \cdots a_p)\tau^{-1} = (\tau(a_1) \cdots \tau(a_p)).$$

En particulier, le conjugué d'un cycle est encore un cycle de même longueur.

Démonstration. Soit $\sigma = \tau(a_1 \cdots a_p)\tau^{-1}$. Alors, on a

$$\sigma(\tau(a_i)) = \tau(a_1 \cdots a_p)(a_i) = \tau(a_{i+1}) \text{ si } i \in [\![1, p-1]\!],$$

et

$$\sigma(\tau(a_p)) = \tau(a_1 \cdots a_p)(a_p) = \tau(a_1) \text{ si } i = p.$$

Soit maintenant $a \in E \backslash \{\tau(a_1), \ldots, \tau(a_p)\}$. Puisque τ est bijective, donc injective, on en déduit que $\tau^{-1}(a) \neq a_i$ pour tout $i \in [\![1, p]\!]$. En particulier, le p-cycle $(a_1 \cdots a_p)$ fixe $\tau^{-1}(a)$, et on a donc

$$\sigma(a) = \tau(\tau^{-1}(a)) = a.$$

Ceci achève la démonstration. □

Théorème III.2.17. *Deux permutations sont conjuguées dans $\mathfrak{S}(E)$ si, et seulement si, les listes (avec répétition) des longueurs des cycles à supports disjoints qui les composent sont les mêmes à l'ordre près.*

Démonstration. Soit $\sigma = \sigma_1\sigma_2 \cdots \sigma_r$, où les σ_i sont des cycles à supports disjoints. Soit $\sigma' \in \mathfrak{S}(E)$. Supposons que σ et σ' soient conjuguées. Alors, il existe $\tau \in \mathfrak{S}(E)$ tel que $\sigma' = \tau\sigma\tau^{-1}$. On a

$$\tau\sigma\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1}) \cdots (\tau\sigma_r\tau^{-1}).$$

D'après le lemme précédent, $\sigma'_i = \tau\sigma_i\tau^{-1}$ est un cycle de même longueur que σ_i. De plus, toujours par le même lemme, le support de σ' est l'image du support de σ_i par τ. Puisque τ est bijective et que les supports des σ_i sont disjoints, il en est de même pour les supports des σ'_i. Ainsi, $\sigma' = \sigma'_1 \cdots \sigma'_r$ est la décomposition de σ' en produits de cycles à supports disjoints, et les listes des longueurs des cycles à supports disjoints qui composent σ et σ' sont les mêmes.

Réciproquement, soient $\sigma, \sigma' \in \mathfrak{S}(E)$. Considérons alors leurs décompositions en produits de cycles à supports disjoints

$$\sigma = \sigma_1 \cdots \sigma_r, \sigma' = \sigma'_1 \cdots \sigma'_r,$$

et supposons que pour tout $i \in [\![1, r]\!]$, les cycles σ_i et σ'_i soient de même longueur, quitte à renuméroter.

Écrivons

$$\sigma_i = (a_1^{(i)} \cdots a_{p_i}^{(i)}),\ \sigma'_i = (b_1^{(i)} \cdots b_{p_i}^{(i)}).$$

Puisque

$$\mathrm{Supp}(\sigma) = \bigcup_{i=1}^{r} \mathrm{Supp}(\sigma_i) \text{ et } \mathrm{Supp}(\sigma') = \bigcup_{i=1}^{r} \mathrm{Supp}(\sigma'_i)$$

par le lemme **??** (4), et que ces deux unions sont disjointes, l'hypothèse entraîne que $\mathrm{Supp}(\sigma)$ et $\mathrm{Supp}(\sigma')$ ont même cardinal. Il en est donc de même des ensembles $X = E \setminus \mathrm{Supp}(\sigma)$ et $X' = E \setminus \mathrm{Supp}(\sigma')$. Choisissons une bijection $f : X \longrightarrow X'$ arbitraire, et soit $\tau : E \longrightarrow E$ définie par

$$\tau(a_j^{(i)}) = b_j^{(i)} \text{ et } \tau(a) = f(a) \text{ si } a \neq a_j^{(i)}.$$

Remarquons que cela définit un élément de $\mathfrak{S}(E)$ par construction. Par définition de τ, et par le lemme précédent, on a

$$\tau\sigma_i\tau^{-1} = \sigma'_i \text{ pour tout } i \in [\![1, r]\!].$$

On a alors

$$\tau\sigma\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1})\cdots(\tau\sigma_r\tau^{-1}) = \sigma'_1\cdots\sigma'_r = \sigma'.$$

Ceci achève la démonstration. □

Exemple III.2.18. Les éléments $\sigma = (1\ 3\ 5)(2\ 4)$ et $\sigma' = (1\ 2\ 3)(6\ 7)$ sont conjugués dans $\mathfrak{S}_7$. Pour trouver un élément τ qui convient, on applique la méthode de la démonstration. On a

$$\begin{array}{rcl} \sigma & = & (1\ 3\ 5)(2\ 4) \\ \sigma' & = & (1\ 2\ 3)(6\ 7) \end{array}$$

Les point fixes de σ sont 6 et 7, et ceux de σ' sont 4 et 5. Ainsi, on peut prendre

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 6 & 2 & 7 & 3 & 4 & 5 \end{pmatrix} = (2\ 6\ 4\ 7\ 5\ 3),$$

car τ induit bien une bijection de $\{6,7\}$ (l'ensemble des points fixes de σ) sur $\{4,5\}$ (l'ensemble des points fixes de σ').

Si l'on préfère, pour faciliter les calculs, on peut rajouter les points fixes dans l'écriture en étendant la notation $(a_1\cdots a_p)$ à $p = 1$. Attention néanmoins, (a) n'est pas un cycle ! ! C'est en fait Id_E. Ainsi, on pourra écrire

$$\begin{array}{rcl} \sigma & = & (1\ 3\ 5)(2\ 4)(6)(7) \\ \sigma' & = & (1\ 2\ 3)(6\ 7)(4)(5) \end{array},$$

ce qui permet de « lire » plus facilement comment doit agir τ sur les points fixes.

Nous allons maintenant donner un résultat qui permet de lister toutes les classes de conjugaison de $\mathfrak{S}(E)$. Donnons tout d'abord quelques définitions.

Définition III.2.19. Soit $n \geq 1$ un entier. Une *partition* de n est une suite d'entiers $\mathbf{p} = (p_k)_{k\geq 1}$ décroissante, nulle à partir d'un certain rang, et telle que

$$\sum_{k\geq 1} p_k = n.$$

Par abus de notation, on confondra souvent $\mathbf{p}$ avec la suite finie décroissante de ses éléments non nuls.

Notation. On note $\mathscr{P}(n)$ l'ensemble des partitions de l'entier n.

Exemple III.2.20. Déterminons toutes les partitions de 4. Puisque

$$4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1,$$

on obtient cinq partitions à savoir

$$(4), (3,1), (2,2), (2,1,1) \ \text{ et } \ (1,1,1,1).$$

Définition III.2.21. Si $\sigma \in \mathfrak{S}(E)$, le *type* de σ est la partition de n dont les éléments non nuls sont les cardinaux des diverses σ-orbites, rangés par ordre décroissant.

On la note $\mathbf{p}_\sigma$.

Exemple III.2.22. Si $\sigma = (1\ 3)(4\ 7\ 6\ 5) \in \mathfrak{S}_9$, on a $\mathbf{p}_\sigma = (4,2,1,1,1)$.

Remarque III.2.23. Connaître les longueurs des p-cycles à supports disjoints composant σ revient à connaître les cardinaux des σ-orbites non réduites à un singleton. Or, le nombre de σ-orbites réduites à un élément, c'est le nombre de points fixes, c'est-à-dire le nombre d'éléments qui ne sont pas dans le support. Mais, le nombre d'éléments du support est la somme des cardinaux des σ-orbites non réduites à un singleton. Bref, connaître les longueurs des p-cycles à supports disjoints composant

σ équivaut à connaître les cardinaux de toutes les σ-orbites, i.e. connaître le type de σ.

Le théorème **??** se retraduit donc de la manière suivante : deux permutations de $\mathfrak{S}(E)$ sont conjuguées si, et seulement si, elles ont même type.

Soit $E = \{a_1, \dots, a_n\}$ une énumération des éléments de E.

Si $\mathbf{p} = (p_1, \dots, p_r)$ est une partition de n, on note $\sigma_{\mathbf{p}} \in \mathfrak{S}(E)$ la permutation

$$\sigma_{\mathbf{p}} = (a_1 \cdots a_{p_1})(a_{p_1+1} \cdots a_{p_1+p_2}) \cdots (a_{p_1+\cdots+p_{r-1}+1} \cdots a_n).$$

Par convention, si $a \in E$, la notation (a) désignera Id_E. Par construction, la permutation $\sigma_{\mathbf{p}}$ est alors de type $\mathbf{p}$.

Si $\sigma \in \mathfrak{S}(E)$, on note $\mathrm{Conj}_{\mathfrak{S}(E)}(\sigma)$ la classe de conjugaison de σ. On a alors le corollaire suivant.

Corollaire III.2.24. *Soit E un ensemble à n éléments, et soit $\mathscr{C}_E$ l'ensemble des classes de conjugaison de $\mathfrak{S}(E)$. Alors, l'application*

$$\begin{aligned} \Psi \colon \mathscr{P}(n) &\longrightarrow \mathscr{C}_E \\ \mathbf{p} &\longmapsto \mathrm{Conj}_{\mathfrak{S}(E)}(\sigma_{\mathbf{p}}) \end{aligned}$$

est une bijection de l'ensemble des partitions de n sur l'ensemble des classes de conjugaison de $\mathfrak{S}(E)$.

Autrement dit, toute permutation de $\mathfrak{S}(E)$ est conjuguée à une unique permutation de la forme $\sigma_{\mathbf{p}}$, où $\mathbf{p}$ est une partition de n.

Démonstration. Montrons que Ψ est injective. Soient $\mathbf{p}$ et $\mathbf{p}'$ deux partitions telles que

$$\Psi(\mathbf{p}) = \Psi(\mathbf{p}').$$

Cela signifie donc que les permutations $\sigma_{\mathbf{p}}$ et $\sigma_{\mathbf{p}'}$ sont conjuguées. Or, deux permutations conjuguées sont de même type. Comme $\sigma_{\mathbf{p}}$ et $\sigma_{\mathbf{p}'}$ sont de type $\mathbf{p}$ et $\mathbf{p}'$ respectivement, on a $\mathbf{p} = \mathbf{p}'$, d'où l'injectivité.

Montrons maintenant que Ψ est surjective. Soit $\sigma \in \mathfrak{S}(E)$, et soit $\mathbf{p}_\sigma$ son type. Alors, $\sigma_{\mathbf{p}_\sigma}$ est de type $\mathbf{p}_\sigma$. Ainsi, $\sigma_{\mathbf{p}_\sigma}$ et σ sont de même type, donc conjuguées par la remarque **??**. On a alors

$$\mathrm{Conj}_{\mathfrak{S}(E)}(\sigma) = \mathrm{Conj}_{\mathfrak{S}(E)}(\sigma_{\mathbf{p}_\sigma}) = \Psi(\mathbf{p}_\sigma),$$

d'où la surjectivité. Ceci achève la démonstration. □

Pour trouver toutes les classes de conjugaison de $\mathfrak{S}(E)$, il suffit donc de décomposer n en somme d'entiers strictement positifs décroissants de toutes les façons possibles.

Exemple III.2.25. D'après l'exemple **??**, $\mathfrak{S}_4$ possède 5 classes de conjugaison distinctes, représentées par

$$(1\ 2\ 3\ 4), (1\ 2\ 3), (1\ 2)(3\ 4), (1\ 2), \mathrm{Id}.$$

III.3. Systèmes de générateurs

Dans ce paragraphe, on donne des systèmes de générateurs variés de $\mathfrak{S}_n$.

Corollaire III.3.1. *Le groupe $\mathfrak{S}_n$ est engendré par chacune des familles suivantes :*

(1) *les transpositions* $(1\ i), i \in [\![2, n]\!]$;
(2) *les transpositions* $(i\ i+1), i \in [\![1, n-1]\!]$.

Démonstration. Puisque toute permutation est produit de transpositions, il reste à vérifier que toute transposition est produit de transpositions de la forme $(1\ i)$ (resp. de la forme $(i\ i+1)$). Or, pour tout $i < j$, on a

$$(i\ j) = (1\ i)(1\ j)(1\ i).$$

L'autre point est à peine plus compliqué.

En utilisant le lemme **??**, on voit que l'on a

$$(i\ i+1 \cdots j-1)(j-1\ j)(i\ i+1 \cdots j-1)^{-1} = (i\ j).$$

Mais comme $(i\ i+1 \cdots j-1) = (i\ i+1) \cdots (j-2\ j-1)$, on obtient le résultat. □

Proposition III.3.2. *Le groupe $\mathfrak{S}_n$ est engendré par $(1\ 2)$ et $(1\ 2 \cdots n)$.*

Démonstration. Pour $i \in [\![1, n-1]\!]$, la permutation $(1\ 2 \cdots n)^{i-1}$ envoie 1 sur i et 2 sur $i+1$. Par le lemme **??**, on a

$$(i\ i+1) = (1\ 2 \cdots n)^{i-1}(1\ 2)(1\ 2 \cdots n)^{1-i}.$$

Ainsi, $(i\ i+1)$ est dans le sous-groupe engendré par $(1\ 2)$ et $(1\ 2 \cdots n)$. On utilise alors la dernière partie du corollaire précédent pour conclure. □

Remarque III.3.3. Ce résultat devient faux en général si l'on remplace $(1\ 2)$ et $(1\ 2 \cdots n)$ par une transposition et un n-cycle arbitraires. Le lecteur pourra chercher un contre-exemple lorsque $n = 4$.

III.4. Signature, groupe alterné

Soit E un ensemble fini à n éléments, $n \geq 2$. Nous allons ici définir la signature d'une permutation de E ainsi que le groupe alterné sur E, pour un ensemble fini E quelconque, et présenter un ou deux systèmes de générateurs de ce groupe.

Théorème III.4.1. *Soit E un ensemble fini, avec $|E| \geq 2$. Alors, il existe un unique morphisme $\varepsilon_E : \mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$ non trivial. Si $\sigma \in \mathfrak{S}(E)$ s'écrit comme produit de s transpositions, alors on a*

$$\varepsilon_E(\sigma) = (-1)^s.$$

Démonstration. Soit $\varepsilon : \mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$ un morphisme de groupes non trivial. Remarquons que ε est constant sur les classes de conjugaison de $\mathfrak{S}(E)$. En effet, si $\sigma' = \tau\sigma\tau^{-1}$, on a

$$\varepsilon(\sigma') = \varepsilon(\tau)\varepsilon(\sigma)\varepsilon(\tau)^{-1} = \varepsilon(\sigma)\varepsilon(\tau)\varepsilon(\tau)^{-1} = \varepsilon(\sigma),$$

puisque $\mathbb{C}^\times$ est abélien. En particulier, toutes les transpositions ont même image par ε par le théorème **??**.

Si $\tau \in \mathfrak{S}(E)$ est une transposition, alors $\tau^2 = \mathrm{Id}_E$, et par conséquent

$$\varepsilon(\tau)^2 = \varepsilon(\tau^2) = \varepsilon(\mathrm{Id}_E) = 1.$$

Ainsi, $\varepsilon(\tau) = \pm 1$. Si l'on avait $\varepsilon(\tau) = 1$ pour toute transposition τ de $\mathfrak{S}(E)$, alors le morphisme ε serait trivial, puisque toute permutation de $\mathfrak{S}(E)$ est produit de transpositions par le corollaire **??**. On a donc $\varepsilon(\tau) = -1$ pour toute transposition τ, et donc $\varepsilon(\sigma) = (-1)^s$ si σ est produit de s transpositions. Ainsi, si un tel morphisme ε existe, il est unique.

Montrons l'existence d'un morphisme non trivial $\mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$. Supposons avoir montré l'existence d'un tel morphisme $\varphi : \mathfrak{S}_n \longrightarrow \mathbb{C}^\times$. Choisissons une numérotation $a_1, \ldots, a_n$ des éléments de E et soit $u : \mathfrak{S}_n \xrightarrow{\sim} S(E)$ l'isomorphisme correspondant. Alors $\varphi \circ u^{-1} : S(E) \longrightarrow \{\pm 1\}$ est un morphisme de groupes. De plus, il est non

trivial. En effet, si $\sigma \in \mathfrak{S}_n$ vérifie $\varphi(\sigma) \neq 1$, alors $(\varphi \circ u^{-1})(u(\sigma)) = \varphi(\sigma) \neq 1$. Il reste donc à construire φ. On dénote par X l'ensemble des paires d'éléments $\{i, j\}$ de l'ensemble $\{1, \ldots, n\}$. Soit $\sigma \in \mathfrak{S}_n$. Remarquons que si $1 \leq i, j \leq n, i \neq j$, alors on a

$$\frac{\sigma(i) - \sigma(j)}{i - j} = \frac{\sigma(j) - \sigma(i)}{j - i}.$$

Ainsi, le rationnel $\dfrac{\sigma(i) - \sigma(j)}{i - j}$ ne dépend que de la paire $\{i, j\}$. On peut donc poser

$$\varphi(\sigma) = \prod_{\{i,j\} \in X} \frac{\sigma(i) - \sigma(j)}{i - j} \in \mathbb{C}^\times.$$

Montrons que $\varphi : \mathfrak{S}_n \longrightarrow \mathbb{C}^\times$ est un morphisme de groupes non trivial.

Soit $\sigma, \tau \in \mathfrak{S}_n$. On a

$$\begin{aligned} \varphi(\sigma\tau) &= \prod_{\{i,j\} \in X} \frac{\sigma\tau(i) - \sigma\tau(j)}{i - j} \\ &= \prod_{\{i,j\} \in X} \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)} \prod_{\{i,j\} \in X} \frac{\tau(i) - \tau(j)}{i - j}. \end{aligned}$$

Comme τ est bijective, l'application

$$f_\tau : X \longrightarrow X, \{i, j\} \longmapsto \{\tau(i), \tau(j)\}$$

est bijective, d'inverse $f_{\tau^{-1}}$, et donc on a

$$\varphi(\sigma\tau) = \prod_{\{i,j\} \in X} \frac{\sigma(i) - \sigma(j)}{i - j} \prod_{\{i,j\} \in X} \frac{\tau(i) - \tau(j)}{i - j} = \varphi(\sigma)\varphi(\tau).$$

Calculons maintenant $\varphi((12))$. Soit $\sigma = (1\ 2)$. Si $\{i, j\} \in X$, on peut toujours supposer que $i < j$. On a donc

$$\varphi(\sigma) = \prod_{i=1}^{n} \prod_{j>i} \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Si $i \geq 3$, la paire $\{i, j\}$ ne contient ni 1, ni 2, et on a

$$\prod_{j>i} \frac{\sigma(i) - \sigma(j)}{i - j} = 1.$$

Supposons $i = 2$. Alors $j > 2$ et donc $\sigma(j) = j$. On a alors

$$\prod_{j>2} \frac{\sigma(2) - \sigma(j)}{2 - j} = \prod_{j>2} \frac{1 - j}{2 - j}.$$

Supposons enfin $i = 1$. Alors $j \geq 2$. Si $j = 2$, on a $\sigma(j) = 1$, et si $j > 2$, on a $\sigma(j) = j$. Ainsi, on a

$$\prod_{j>1} \frac{\sigma(1) - \sigma(j)}{1 - j} = \frac{2 - 1}{1 - 2} \cdot \prod_{j>2} \frac{2 - j}{1 - j} = -\prod_{j>2} \frac{2 - j}{1 - j}.$$

On a ainsi $\varphi((12)) = -1$, et donc $\varphi : \mathfrak{S}_n \longrightarrow \mathbb{C}^\times$ est un morphisme de groupes non trivial. Ceci achève la démonstration. □

Remarque III.4.2. Le théorème donne aussi implicitement le résultat suivant : la parité du nombre de transpositions nécessaires pour écrire une permutation de $\mathfrak{S}(E)$ ne dépend pas de la décomposition choisie.

Définition III.4.3. Si $\sigma \in \mathfrak{S}(E)$, l'élément $\varepsilon_E(\sigma) \in \{\pm 1\}$ est appelé la *signature* de σ.

En général, le morphisme de signature est simplement noté ε.

Exemple III.4.4. La signature d'un p-cycle $\sigma = (a_1 \cdots a_p)$ est égale à $(-1)^{p-1}$. En effet, on a

$$(a_1 \cdots a_p) = (a_1\ a_2)(a_2\ a_3) \cdots (a_{p-2}\ a_{p-1})(a_{p-1}\ a_p),$$

et donc σ est produit de $p-1$ transpositions.

Proposition III.4.5. *Soit E un ensemble à n éléments, et soit $\sigma \in \mathfrak{S}(E)$. Alors,*

$$\varepsilon(\sigma) = (-1)^{n-N_\sigma},$$

où N_σ est le nombre de σ-orbites.

Démonstration. Soit $\sigma \in \mathfrak{S}(E)$, et soit $\sigma = \sigma_1 \cdots \sigma_r$ sa décomposition en produit de cycles à supports disjoints. Si on note p_i la longueur du cycle σ_i, alors $p_1 + \cdots + p_r$ est le nombre d'éléments de E qui ne sont pas fixes. Il y a donc $n - (p_1 + \cdots + p_r)$ points fixes, i.e. $n - (p_1 + \cdots + p_r)$ σ-orbites réduites à un élément. Il y a donc $N_\sigma = n - (p_1 + \cdots + p_r) + r$ σ-orbites distinctes, puisque les σ-orbites non réduites à un élément sont les supports des cycles σ_i. D'autre part, on a d'après l'exemple précédent

$$\varepsilon(\sigma) = \varepsilon(\sigma_1) \cdots \varepsilon(\sigma_r) = (-1)^{p_1-1} \cdots (-1)^{p_r-1} = (-1)^{p_1+\cdots+p_r-r},$$

et donc $\varepsilon(\sigma) = (-1)^{n-N_\sigma}$. Ceci achève la démonstration. □

Nous pouvons maintenant définir le groupe alterné.

Définition III.4.6. Le *groupe alterné*, noté $\mathfrak{A}(E)$, est l'ensemble des permutations de $\mathfrak{S}(E)$ de signature 1. C'est un sous-groupe distingué de $\mathfrak{S}(E)$ (puisque c'est le noyau d'un morphisme de groupes ; cf. le lemme **??**). En fait, $\mathfrak{A}(E)$ est un sous-groupe d'indice 2 de $\mathfrak{S}(E)$.

En effet, si le théorème de factorisation appliqué au morphisme de signature donne immédiatement un isomorphisme

$$\mathfrak{S}(E)/\mathfrak{A}(E) \simeq \{\pm 1\}.$$

Si E a n éléments ($n \geq 2$) , on a donc

$$|\mathfrak{A}(E)| = \frac{n!}{2}.$$

Si $E = [\![1, n]\!]$, on le note $\mathfrak{A}_n$.

Remarque III.4.7. Si $n = 2$, $\mathfrak{A}(E)$ est trivial, et si $n = 3$, $\mathfrak{A}(E)$ est cyclique, engendré par un 3-cycle.

Le résultat suivant, bien qu'il ne nous serve pas pour la suite, est classique.

Proposition III.4.8. *Soit E un ensemble ayant au moins deux éléments. Alors, le groupe alterné $\mathfrak{A}(E)$ est l'unique sous-groupe d'indice 2 de $\mathfrak{S}(E)$.*

Démonstration. Soit H un sous-groupe d'indice 2 de $\mathfrak{S}(E)$. On sait qu'il est nécessairement distingué (cf. proposition **??**). Le groupe quotient G/H est donc d'ordre 2, donc isomorphe à $\{\pm 1\}$. Fixons un isomorphisme $u : G/H \xrightarrow{\sim} \{\pm 1\}$, et soit $f : \mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$ défini par

$$f(\sigma) = u(\overline{\sigma}) \ \text{ pour tout } \sigma \in \mathfrak{S}(E).$$

Alors, on a

$$\sigma \in \mathrm{Ker}(f) \iff u(\overline{\sigma}) = 1 \iff \overline{\sigma} = \overline{\mathrm{Id}}_E \iff \sigma \in H.$$

Ainsi, $\mathrm{Ker}(f) = H$. En particulier, f est non trivial, puisque $H \neq \mathfrak{S}(E)$. Par le théorème **??**, f est donc le morphisme de signature. En particulier, $H = \mathrm{Ker}(f) = \mathrm{Ker}(\varepsilon_E) = \mathfrak{A}(E)$. □

Proposition III.4.9. *Soit E un ensemble à n éléments. Si $n \geq 3$, le groupe alterné $\mathfrak{A}(E)$ est engendré par chacune des familles suivantes :*

(1) *les produits de deux transpositions (non nécessairement à supports disjoints) ;*

(2) *les 3-cycles.*

Démonstration. D'après le corollaire **??**, $\mathfrak{S}(E)$ est engendré par les transpositions. Pour tout $\sigma \in \mathfrak{S}(E)$, on peut donc écrire

$$\sigma = \tau_1 \cdots \tau_r,$$

où chaque τ_i est une transposition. Or, on a $\varepsilon(\sigma) = (-1)^r$, et donc $\sigma \in \mathfrak{A}(E)$ si, et seulement si, r est pair. Autrement dit, les éléments de $\mathfrak{A}(E)$ sont les produits d'un nombre pair de transpositions. En particulier, les produits de deux transpositions engendrent $\mathfrak{A}(E)$.

Pour montrer le deuxième point, il suffit de montrer qu'un produit de deux transpositions est un produit de 3-cycles. Soient $\tau_1, \tau_2 \in \mathfrak{S}(E)$ deux transpositions. Si elles ont même support, alors $\tau_1 = \tau_2$. Ainsi, $\tau_1\tau_2 = \mathrm{Id}_E$ et il n'y a rien à démontrer. Si leurs supports ont exactement un élément en commun, alors on a $\tau_1 = (a\ b), \tau_2 = (b\ c)$ (avec $a, b, c \in E$ distincts), et alors $\tau_1\tau_2 = (a\ b\ c)$. Enfin, si les supports de τ_1 et τ_2 sont disjoints, on peut écrire $\tau_1 = (a\ b), \tau_2 = (c\ d)$ avec $a, b, c, d \in E$ distincts, et on a

$$\tau_1\tau_2 = (a\ b)(c\ d) = (a\ c\ b)(a\ c\ d).$$

Ceci achève la démonstration. □

III.5. Structure des groupes symétrique et alterné

Dans ce paragraphe, on démontre que $\mathfrak{A}(E)$ est un groupe simple dès que l'on a $|E| \neq 2, 4$. Ce résultat dû à Galois est historiquement très important, car il permet de démontrer l'impossibilité de résoudre une équation polynomiale de degré $n \geq 5$ par radicaux en général.

On commence par un lemme.

Lemme III.5.1. *Soit E un ensemble à n éléments. Si $n \geq 5$, les 3-cycles sont conjugués dans $\mathfrak{A}(E)$.*

Démonstration. Soient $\sigma = (a_1\ a_2\ a_3)$ et $\sigma' = (b_1\ b_2\ b_3)$ deux 3-cycles. D'après le théorème **??**, il existe $\tau \in \mathfrak{S}(E)$ tel que

$$\tau\sigma\tau^{-1} = \sigma'.$$

Si $\tau \in \mathfrak{A}(E)$, c'est fini. Supposons que $\varepsilon(\tau) = -1$. Puisque $n \geq 5$, on peut trouver $c \neq d \in E$ distincts de b_1, b_2, b_3. Posons alors $\tau' = (c\ d)\tau$. Par construction, $\tau' \in \mathfrak{A}(E)$ puisque

$$\varepsilon(\tau') = \varepsilon(\tau)\varepsilon((c\ d)) = (-1)(-1) = 1.$$

Mais alors, on a

$$\tau'\sigma\tau'^{-1} = (c\ d)(b_1\ b_2\ b_3)(c\ d)^{-1} = (b_1\ b_2\ b_3)(c\ d)(c\ d)^{-1} = \sigma',$$

car $(b_1\ b_2\ b_3)$ et $(c\ d)$ sont à supports disjoints, donc commutent. On a donc le résultat annoncé.

□

Remarque III.5.2. Le lecteur se convaincra lui-même que ce résultat est faux si $n = 3, 4$.

On peut maintenant démontrer le résultat principal de ce paragraphe.

Théorème III.5.3. *Soit E un ensemble à n éléments ($n \geq 3$). Alors, le groupe alterné $\mathfrak{A}(E)$ est simple si, et seulement si, $n \neq 4$.*

Démonstration. Le cas $n = 3$ est facile puisque dans ce cas, $\mathfrak{A}(E)$ est cyclique d'ordre 3. On vérifie aisément que $\mathfrak{A}(E)$ est simple dans ce cas.

Si $n = 4$, notons $E = \{a, b, c, d\}$. Un simple calcul montre que

$$H = \{\mathrm{Id}_E, (a\ b)(c\ d), (a\ c)(b\ d), (a\ d)(b\ c)\}$$

est un sous-groupe de $\mathfrak{A}(E)$. Il est bien entendu non trivial, et il est distingué dans $\mathfrak{S}(E)$ puisque le conjugué d'une permutation de type $(2, 2)$ est une permutation de type $(2, 2)$. Il est donc en particulier distingué dans $\mathfrak{A}(E)$. Ainsi, $\mathfrak{A}(E)$ n'est pas simple.

On suppose jusqu'à la fin que $n \geq 5$. Faisons d'abord quelques remarques préliminaires. Soit H un sous-groupe distingué de $\mathfrak{A}(E)$, non réduit à l'identité. On doit montrer que $H = \mathfrak{A}(E)$.

(1) Si H contient un 3-cycle, alors $H = \mathfrak{A}(E)$.

En effet, puisque H contient un 3-cycle, il les contient tous puisque H est distingué dans $\mathfrak{A}(E)$ et que tous les 3-cycles sont conjugués par le lemme précédent. Le groupe $\mathfrak{A}(E)$ étant engendré par les 3-cycles par la proposition **??**, on obtient $H = \mathfrak{A}(E)$.

(2) Si $\sigma \in H$, $\tau \in \mathfrak{A}(E)$, alors $\sigma\tau\sigma^{-1}\tau^{-1} \in H$.

En effet, comme H est un sous-groupe, on a $\sigma^{-1} \in H$, et comme il est distingué dans $\mathfrak{A}(E)$, on a $\tau\sigma^{-1}\tau^{-1} \in H$. On a alors aussi $\sigma\tau\sigma^{-1}\tau^{-1} \in H$.

(3) Si H contient un produit de deux transpositions à supports disjoints, alors H contient un 3-cycle. En particulier, $H = \mathfrak{A}(E)$.

En effet, supposons que l'on ait $\sigma = (a\ b)(c\ d) \in H$, où $a, b, c, d \in E$ sont distincts. Puisque $n \geq 5$, on peut choisir un élément $e \in E \setminus \{a, b, c, d\}$. Posons $\tau = (a\ b\ e)$, et remarquons que l'on a $\sigma^{-1} = (c\ d)(a\ b)$. On a alors

$$\tau\sigma^{-1}\tau^{-1} = (\tau(c\ d)\tau^{-1})(\tau(a\ b)\tau^{-1}) = (\tau(c)\ \tau(d))(\tau(a)\ \tau(b)) = (c\ d)(b\ e).$$

Ainsi,

$$\sigma\tau\sigma^{-1}\tau^{-1} = (a\ b)(c\ d)(c\ d)(b\ e) = (a\ b)(b\ e) = (a\ b\ e).$$

Par (2), $(a\ b\ e) \in H$, et par (1), on obtient $H = \mathfrak{A}(E)$.

On tire de tout ceci qu'il suffit de montrer que H contient un 3-cycle ou un produit de deux transpositions à supports disjoints, ce que l'on va faire maintenant. Soit $\sigma \in H \setminus \{\mathrm{Id}_E\}$.

Supposons en premier lieu que tous les cycles à supports disjoints composant σ soient des 3-cycles. Si σ est un 3-cycle, il n'y a rien à faire. Sinon, il y en a au moins deux, disons $(a\ b\ c)$ et $(d\ e\ f)$. Soit $\tau = (a\ d)(b\ e)$. Alors, la permutation $\sigma' = \sigma\tau\sigma^{-1}\tau^{-1}$ est dans H par (2). On a de plus

$$\sigma\tau\sigma^{-1} = (\sigma(a)\ \sigma(d))(\sigma(b)\ \sigma(e)) = (b\ e)(c\ f),$$

et alors

$$\sigma' = (b\ e)(c\ f)(b\ e)(a\ d) = (c\ f)(a\ d) \in H.$$

Supposons maintenant que la décomposition en cycles à supports disjoints de σ ne contienne pas uniquement des cycles d'ordre 3.

Admettons provisoirement l'existence d'une partie F de E à trois éléments telle que $F \cup \sigma(F)$ contienne exactement quatre éléments (en particulier, on a $\sigma(F) \neq F$), et soit τ un cycle de support F. C'est donc un 3-cycle, et ainsi $\tau \in \mathfrak{A}(E)$. On a de plus

$$\mathrm{Supp}(\sigma\tau\sigma^{-1}) = \sigma(F) \neq F = \mathrm{Supp}(\tau).$$

En particulier, $\sigma\tau\sigma^{-1} \neq \tau$ et donc $\sigma' = \sigma\tau\sigma^{-1}\tau^{-1} \neq \mathrm{Id}_E$. Par (2), on a alors $\sigma' \in H$. De plus, par le lemme **??** (3), on a

$$\mathrm{Supp}(\sigma') \subset \mathrm{Supp}(\sigma\tau\sigma^{-1}) \cup \mathrm{Supp}(\tau^{-1}) = \sigma(F) \cup F.$$

Comme $\sigma(F) \cup F$ a quatre éléments et que $\sigma' \neq \mathrm{Id}_E$, σ' est soit un 3-cycle, soit un produit de deux transpositions à supports disjoints car $\sigma' \in \mathfrak{A}(E)$ (en effet, σ' ne peut être une transposition ou un 4-cycle, car de telles permutations sont de signature -1).

Il reste donc à vérifier l'existence d'une telle partie F. Si au moins un des cycles qui composent σ est de longueur ≥ 4, disons $(a\ b\ c\ d \cdots)$, il suffit de poser $F = \{a, b, c\}$, puisqu'alors

$$F \cup \sigma(F) = \{a, b, c, d\}.$$

Supposons maintenant que tous les cycles qui composent σ soient de longueur 2 ou 3. Par hypothèse, il existe au moins un cycle de longueur 2 dans la décomposition de σ, donc il en existe au moins deux, disons $(a\ b)$ et $(c\ d)$. Dans le cas contraire, la signature de σ serait en effet égale à -1 (la signature d'un 2-cycle est -1 et celle d'un 3-cycle est 1), ce qui est impossible puisque $\sigma \in H \subset \mathfrak{A}(E)$. On prend alors $F = \{a, b, c\}$.

Ceci montre le résultat voulu. □

Remarque III.5.4. On peut démontrer que $\mathfrak{A}_5$ est l'unique groupe simple d'ordre 60, à isomorphisme près.

On va maintenant déduire du théorème **??** la liste des sous-groupes distingués de $\mathfrak{S}(E)$. On commence par un lemme.

Lemme III.5.5. *Soit E un ensemble à n éléments. Si $n \geq 3$, le centre du groupe symétrique $\mathfrak{S}(E)$ est trivial.*

Démonstration. Soit $\sigma \neq \mathrm{Id}_E$, et soit $a \in \mathrm{Supp}(\sigma)$ (le support de σ n'est pas vide puisque $\sigma \neq \mathrm{Id}_E$). On a donc $b = \sigma(a) \neq a$. Puisque $n \geq 3$, on peut choisir $c \in E \setminus \{a, b\}$. Soit $\tau = (b\ c)$. On a alors

$$\sigma\tau(a) = \sigma(a) = b \quad \text{et} \quad \tau\sigma(a) = \tau(b) = c.$$

Ainsi, σ et τ ne commutent pas, et donc $\sigma \notin Z(\mathfrak{S}(E))$. Par conséquent, on a $Z(\mathfrak{S}(E)) = \{\mathrm{Id}_E\}$. □

Remarque III.5.6. Le résultat est faux pour $n = 2$, puisque $\mathfrak{S}_2$ est commutatif.

Corollaire III.5.7. *Soit E un ensemble à n éléments ($n \geq 2$). Si $n \neq 4$, les sous-groupes distingués de $\mathfrak{S}(E)$ sont*

$$\{\mathrm{Id}_E\}, \mathfrak{A}(E) \quad \textit{et} \quad \mathfrak{S}(E).$$

Démonstration. Si $n = 2$, le résultat est clair puisque $\mathfrak{A}(E)$ est réduit à l'identité. On peut donc supposer que $n \geq 3$ et $n \neq 4$. Soit H un sous-groupe distingué de $\mathfrak{S}(E)$. Alors, $H \cap \mathfrak{A}(E)$ est un sous-groupe distingué de $\mathfrak{A}(E)$, et donc $H \cap \mathfrak{A}(E) = \{\mathrm{Id}_E\}$ ou $\mathfrak{A}(E)$ par le théorème **??**.

Si $H \cap \mathfrak{A}(E) = \mathfrak{A}(E)$, alors $\mathfrak{A}(E) \subset H$. Si $H \subset \mathfrak{A}(E)$, alors $H = \mathfrak{A}(E)$. Si H contient une permutation σ_0 de signature -1, il contient alors deux classes à gauche modulo $\mathfrak{A}(E)$ distinctes, à savoir $\mathfrak{A}(E)$ et $\sigma_0\mathfrak{A}(E)$, c'est-à-dire qu'il contient $\mathfrak{S}(E)$. Dans ce cas $H = \mathfrak{S}(E)$.

Si $H \cap \mathfrak{A}(E) = \{\mathrm{Id}_E\}$, alors la signature induit par restriction un morphisme injectif $H \hookrightarrow \{\pm 1\}$ (car la signature d'une permutation vaut ± 1), de sorte que $|H| \leq 2$.

Supposons que $|H| = 2$, et soit $\tau \in H$ l'unique élément non trivial de H. Soit $\sigma \in \mathfrak{S}(E)$. Comme $\sigma\tau\sigma^{-1} \in H$ et $\sigma\tau\sigma^{-1} \neq \mathrm{Id}_E$ (puisque $\tau \neq \mathrm{Id}_E$), on a $\sigma\tau\sigma^{-1} = \tau$. On en déduit que $\tau \in Z(\mathfrak{S}(E))$. Comme $n \geq 3$, le centre de $\mathfrak{S}(E)$ est trivial d'après le lemme précédent, et donc $\tau = \mathrm{Id}_E$, d'où une contradiction. Ainsi, $|H| = 1$, et donc $H = \{\mathrm{Id}_E\}$ dans ce cas. Ceci achève la démonstration. $\square$

Remarque III.5.8. Le résultat est faux si $n = 4$. En effet, si $E = \{a, b, c, d\}$, l'ensemble

$$H = \{\mathrm{Id}_E, (a\ b)(c\ d), (a\ c)(b\ d), (a\ d)(b\ c)\}$$

est un sous-groupe distingué de $\mathfrak{S}(E)$, distinct de $\{\mathrm{Id}_E\}, \mathfrak{A}(E)$ et $\mathfrak{S}(E)$.

On peut d'ailleurs montrer que les sous-groupes distingués de $\mathfrak{S}(E)$ sont dans ce cas

$$\{\mathrm{Id}_E\}, H, \mathfrak{A}(E) \quad \text{et} \quad \mathfrak{S}(E).$$

Chapitre IV

Produits directs et semi-directs

IV.1. Préliminaires

Si G un groupe, et si $H_1, \ldots, H_r$ sont des sous-groupes de G, on pose

$$H_1 \cdots H_r = \{h_1 \cdots h_r \mid h_i \in H_i\}.$$

Ce n'est pas un groupe en général.

On note $\langle H_1, \ldots, H_r \rangle$ le sous-groupe de G engendré par $H_1, \ldots, H_r$ (c'est-à-dire par le sous-ensemble $H_1 \cup \ldots \cup H_r$). On a clairement

$$H_1 \cdots H_r \subset \langle H_1, \ldots, H_r \rangle,$$

mais l'inclusion est stricte en général.

Le lemme suivant généralise la première partie du théorème **??**.

Lemme IV.1.1. *Soit G un groupe, et soient $H_1, \ldots, H_r$ r sous-groupes de G, avec $r \geq 2$. On suppose que $H_1, \ldots, H_{r-1}$ sont distingués dans G. Alors, l'ensemble $H_1 \cdots H_r$ est un sous-groupe de G, et on a*

$$\langle H_1, \ldots, H_r \rangle = H_1 \cdots H_r.$$

Démonstration. Pour vérifier que $H_1 \cdots H_r$ est bien un sous-groupe de G, on procède par récurrence sur r. Le cas $r = 2$ n'est rien d'autre que la première partie du théorème **??**. Supposons maintenant que le résultat soit vrai pour $r \geq 2$, et considérons $r+1$ sous-groupes $H_1, \ldots, H_{r+1}$ dont les r premiers sont distingués dans G. L'hypothèse de récurrence montre que $H_2 \cdots H_{r+1}$ est un sous-groupe de G. Comme H_1 est un sous-groupe distingué de G, le cas $r = 2$ montre que l'ensemble

$$H_1(H_2 \cdots H_{r+1}) = H_1 \cdots H_{r+1}$$

est un sous-groupe de G. Ceci achève la récurrence.

Mais alors, l'ensemble $H_1 \cdots H_r$ est un sous-groupe de G, qui contient clairement $H_1, \ldots, H_r$. Ainsi, $\langle H_1, \ldots, H_r \rangle \subset H_1 \cdots H_r$. L'autre inclusion étant évidente, ceci achève la démonstration. □

On finit par un lemme qui nous sera utile par la suite.

Lemme IV.1.2. *Soit G un groupe, et soient H, K deux sous-groupes finis de G. Alors, HK est un ensemble fini et on a*

$$|HK| = \frac{|H| \cdot |K|}{|H \cap K|}.$$

Démonstration. Considérons l'application

$$\begin{aligned} f \colon H \times K &\longrightarrow HK \\ (h, k) &\longmapsto hk. \end{aligned}$$

Puisque f est surjective, on a

$$|HK| \leq |H \times K| < +\infty,$$

et donc HK est fini. De plus, l'ensemble $H \times K$ est l'union disjointe des ensembles $f^{-1}(\{hk\})$, lorsque hk parcourt HK. Or, pour tout $h \in H$ et tout $k \in K$, on a

$$f^{-1}(\{hk\}) = \{(hg, g^{-1}k) \mid g \in H \cap K\}.$$

En effet, l'inclusion « $\supset$ » est évidente. Si maintenant $(h', k') \in H \times K$ vérifie $h'k' = hk$, on a $h^{-1}h' = kk'^{-1}$. Mais alors, $g = h^{-1}h' \in H \cap K$, et on a

$$hg = h(h^{-1}h') = h' \quad \text{et} \quad g^{-1}k = (kk'^{-1})^{-1}k = k'.$$

De plus, pour tous $g_1, g_2 \in H \cap K$, on a

$$(hg_1, g_1^{-1}k) = (hg_2, g_2^{-1}k) \Longrightarrow g_1 = g_2,$$

comme on peut le voir par simplification dans G. Ainsi, pour tout $hk \in HK$, on a $|f^{-1}(hk)| = |H \cap K|$. On obtient donc

$$|H \times K| = |HK||H \cap K|,$$

d'où la conclusion. □

IV.2. Produits directs

Le but de ce paragraphe est de définir la notion de produit direct interne et externe de groupes, qui peut se comparer à la somme directe interne et externe d'espaces vectoriels. On commence par introduire le produit direct externe.

Définition IV.2.1. Soient $G_1, \ldots, G_r$ des groupes.

Le groupe $G_1 \times \cdots \times G_r$ (muni de la loi de groupe évidente) est appelé le *produit direct externe* des groupes $G_1, \ldots, G_r$.

Lemme IV.2.2. *Soit $G = G_1 \times \cdots \times G_r, r \geq 2$. Pour $i \in [\![1, r]\!]$, on pose*

$$H_i = \{(1_{G_1}, \ldots, 1_{G_{i-1}}, g_i, 1_{G_{i+1}} \ldots, 1_{G_r}) \mid g_i \in G_i\}.$$

Alors, on a les propriétés suivantes :

(1) *les ensembles $H_1, \ldots, H_r$ sont des sous-groupes distingués de G ;*
(2) *on a $h_i h_j = h_j h_i$ pour tout $h_i \in H_i$, et tout $h_j \in H_j$, dès que $i \neq j$;*
(3) *tout élément $g \in G$ s'écrit de manière unique sous la forme*

$$g = h_1 \cdots h_r, \ h_i \in H_i.$$

Démonstration. Exercice laissé au lecteur. □

Ce lemme motive la définition suivante.

Définition IV.2.3. Soit G un groupe, et soient $H_1, \ldots, H_r$ r sous-groupes de G (avec $r \geq 2$). On dit que G est le *produit direct interne* de $H_1, \ldots, H_r$ si les deux conditions suivantes sont réalisées :

(1) on a $h_i h_j = h_j h_i$ pour tout $h_i \in H_i$, et tout $h_j \in H_j$, dès que $i \neq j$;
(2) tout élément $g \in G$ s'écrit de manière unique sous la forme

$$g = h_1 \cdots h_r, \ h_i \in H_i.$$

On le note $G = H_1 \odot \cdots \odot H_r$.

Remarque IV.2.4. Soit G un groupe, et soient $H_1, \dots, H_r$ des sous-groupes de G. Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, l'application

$$\begin{aligned} H_1 \times \cdots \times H_r &\longrightarrow G \\ (h_1, \dots, h_r) &\longmapsto h_1 \cdots h_r. \end{aligned}$$

est un isomorphisme. Ceci est laissé en exercice au lecteur.

Nous allons maintenant donner des conditions nécessaires et suffisantes pour qu'un groupe G soit un produit direct interne de sous-groupes.

Proposition IV.2.5. *Soit G un groupe, et soient $H_1, \dots, H_r$ r sous-groupes de G ($r \geq 2$). Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *on a $h_i h_j = h_j h_i$ pour tout $h_i \in H_i$, et tout $h_j \in H_j$, dès que $i \neq j$;*
(2) *on a $G = \langle H_1, \dots, H_r \rangle$;*
(3) *pour tout $i \in [\![1, r-1]\!]$, on a $H_1 \cdots H_i \cap H_{i+1} = \{1_G\}$.*

Démonstration. Supposons que $G = H_1 \odot \cdots \odot H_r$. Alors, (1) est vérifiée par définition. De plus, on a $G = H_1 \cdots H_r \subset \langle H_1, \dots, H_r \rangle$, et on a donc montré (2), puisque l'autre inclusion est triviale. Montrons enfin que (3) est réalisée. Soit $i \in [\![1, r-1]\!]$, et soit $g \in H_1 \cdots H_i \cap H_{i+1}$. Alors, il existe des éléments $h_1 \in H_1, \dots, h_i \in H_i$ tels que

$$g = h_1 \cdots h_i.$$

Comme on a aussi $g \in H_{i+1}$, l'unicité de la décomposition d'un élément entraîne que tous les h_i sont égaux à 1_G, et donc que $g = 1_G$.

Inversement, supposons les conditions $(1) - (3)$ vérifiées. Il reste à montrer que tout élément $g \in G$ se décompose de manière unique sous la forme

$$g = h_1 \cdots h_r, \ h_i \in H_i.$$

Remarquons tout d'abord que $H_1, \dots, H_r$ sont distingués dans G. En effet, soit $h_i \in H_i$ pour tout $i \in [\![1, r]\!]$. Comme on a $G = \langle H_1, \dots, H_r \rangle$, il suffit de vérifier que l'on a $h_j h_i h_j^{-1}$ pour tout $h_j \in H_j$ et tout entier $j \in [\![1, r-1]\!]$. Si $j = i$, cela provient du fait que H_i est un sous-groupe de G. Si $j \neq i$, cela provient de l'égalité $h_j h_i h_j^{-1} = h_i$.

Le lemme **??** implique alors que $G = H_1 \cdots H_r$, c'est-à-dire l'existence de la décomposition recherchée. Pour montrer l'unicité, supposons que l'on ait une égalité de la forme

$$h_1 \cdots h_r = h'_1 \cdots h'_r, \text{ avec } h_i, h'_i \in H_i.$$

En tenant compte de la condition (1), cela se récrit

$$h_r (h'_r)^{-1} = h_1^{-1} h'_1 \cdots h_{r-1}^{-1} h'_{r-1}.$$

La condition (3) appliquée à $i = r - 1$ montre alors que $h_r (h'_r)^{-1} = 1_G$ et que $h_1^{-1} h'_1 \cdots h_{r-1}^{-1} h'_{r-1} = 1_G$. On a donc $h_r = h'_r$, et par simplification on obtient

$$h_1 \cdots h_{r-1} = h'_1 \cdots h'_{r-1}.$$

Une récurrence facile montre alors l'égalité des décompositions. Ceci achève la démonstration. □

Théorème IV.2.6. *Soit G un groupe, et soient $H_1, \dots, H_r$ r sous-groupes de G (avec $r \geq 2$). Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *le sous-groupe* H_i *est distingué dans* G *pour tout* $i \in [\![1, r]\!]$;

(2) *on a* $G = \langle H_1, \dots, H_r \rangle$;

(3) *pour tout* $i \in [\![1, r-1]\!]$, *on a* $H_1 \cdots H_i \cap H_{i+1} = \{1_G\}$.

Démonstration. Si $G = H_1 \odot \cdots \odot H_r$, la proposition **??** montre que les conditions (2) et (3) sont vérifiées. De plus, on a vu au cours de la démonstration de cette proposition que la commutativité des éléments de H_i et H_j couplée à l'égalité $G = \langle H_1, \dots, H_r \rangle$ impliquait que $H_1, \dots, H_r$ sont des sous-groupes distingués de G.

Inversement, supposons les trois conditions du théorème vérifiées. D'après la proposition **??**, il suffit de montrer que l'on a

$$h_i h_j = h_j h_i \text{ pour tout } h_i \in H_i, \text{ et tout } h_j \in H_j, \text{ dès que } i \neq j$$

pour pouvoir conclure. Soient $i \neq j$, $h_i \in H_i, h_j \in H_j$. On peut toujours supposer $i < j$. Posons $g = h_i h_j h_i^{-1} h_j^{-1}$. Le fait que H_i et H_j soient distingués dans G entraîne que $g \in H_i \cap H_j$. Or,

$$H_i \cap H_j \subset H_1 \dots H_{j-1} \cap H_j = \{1_G\}.$$

Donc $g = 1_G$, et ainsi $h_i h_j = h_j h_i$, d'où le résultat souhaité. □

Remarque IV.2.7. Au vu de la condition (1) sur les H_i, on pourrait remplacer (2) par la condition

(2′) on a $G = H_1 \cdots H_r$.

Néanmoins, il est plus facile a priori de vérifier (2) que (2′).

Dans le cas des groupes finis, on peut remplacer la condition (2) par une condition plus simple à vérifier.

Corollaire IV.2.8. *Soit* G *un groupe fini, et soient* $H_1, \dots, H_r$ r *sous-groupes de* G *(*$r \geq 2$*). Alors,* $G = H_1 \odot \cdots \odot H_r$ *si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *le sous-groupe* H_i *est distingué dans* G *pour tout* $i \in [\![1, r]\!]$;

(2) *on a* $|G| = |H_1| \cdots |H_r|$;

(3) *pour tout* $i \in [\![1, r-1]\!]$, *on a* $H_1 \cdots H_i \cap H_{i+1} = \{1_G\}$.

Démonstration. Si $G = H_1 \odot \cdots \odot H_r$, alors $G \simeq H_1 \times \cdots \times H_r$ par la remarque **??**, d'où la condition (2). Les deux autres conditions proviennent des résultats précédents.

Pour montrer la réciproque, supposons tout d'abord que $r = 2$. On sait déjà que les sous-groupes H_1 et H_2 sont distingués dans G, d'intersection triviale. D'après le théorème **??**, il reste à démontrer que $G = \langle H_1, H_2 \rangle$ pour pouvoir conclure. Comme H_1 est distingué dans G, cela revient à vérifier l'égalité $G = H_1 H_2$ par le lemme **??**. Or, les conditions (2) et (3) impliquent que $|G| = |H_1 H_2|$ d'après le lemme **??**. Comme $H_1 H_2 \subset G$, on en déduit $G = H_1 H_2$, d'où le résultat.

Supposons maintenant que $r \geq 3$, et posons $H = \langle H_1, \dots, H_{r-1} \rangle$. Comme les sous-groupes $H_1, \dots, H_{r-1}$ sont distingués dans G, ils sont distingués dans H. Puisque $H_1 \cdots H_i \cap H_{i+1} = \{1\}$ pour tout $i \in [\![1, r-2]\!]$, le théorème **??** montre que l'on a l'égalité

$$H = H_1 \odot \cdots \odot H_{r-1}.$$

En particulier, on a $H = H_1 \cdots H_{r-1}$, ainsi que

$$H \simeq H_1 \times \cdots \times H_{r-1}.$$

Par conséquent, $|H| = |H_1| \cdots |H_{r-1}|$. La condition (2) se récrit alors

$$|G| = |H| \cdot |H_r|.$$

De plus, la condition (3) appliquée à $i = r - 1$ montre que $H \cap H_r = \{1_G\}$. Enfin, le fait que $H_1, \ldots, H_r$ soient distingués dans G impliquent aisément que H et H_r sont distingués dans G. Le cas $r = 2$ montre alors que l'on a

$$G = H \odot H_r = (H_1 \odot \cdots \odot H_{r-1}) \odot H_r = H_1 \odot \cdots \odot H_r,$$

la dernière égalité découlant facilement des définitions. Ceci achève la démonstration. □

Corollaire IV.2.9. *Soit G un groupe fini, et soient H et K deux sous-groupes de G. Alors, on a $G = H \odot K$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *les sous-groupes H et K sont distingués dans G (ou alternativement : on a $hk = hk$ pour tout $h \in H$, et tout $k \in K$) ;*
(2) *on a $|G| = |H| \cdot |K|$;*
(3) *on a $H \cap K = \{1_G\}$.*

IV.3. Produits semi-directs

Soit G un groupe, et soient H et K deux sous-groupes de G. Reprenons l'étude de HK. On a vu dans le premier paragraphe que si H est distingué dans G, alors $\langle H, K \rangle = HK$. En particulier, HK est un sous-groupe de G. On devrait pouvoir ainsi écrire le produit de deux éléments $hk, h'k' \in HK$ sous la forme $h''k''$. C'est possible en utilisant la distinction de H. En effet, on a

$$(hk)(h'k') = \big(h(kh'k^{-1})\big)kk' \in HK.$$

Bien sûr, un élément de HK peut avoir plusieurs écritures différentes. L'écriture sera unique si, et seulement si, $H \cap K = \{1_G\}$.

En effet, si l'écriture est unique, alors pour tout $g \in H \cap K$, on a

$$g1_G = 1_G g \in HK,$$

et donc $g = 1_G$. Inversement, si $H \cap K = \{1_G\}$, alors pour tous $h, h' \in H$, et tous $k, k' \in K$, on a

$$hk = h'k' \Longrightarrow (h')^{-1}h = k'k^{-1},$$

et donc $k'k^{-1} \in H \cap K$. Ainsi, $k'k^{-1} = 1_G$, soit $k = k'$, et on déduit ensuite que $h = h'$.

En particulier, si $H \cap K = \{1_G\}$, on a une bijection

$$\begin{aligned} H \times K &\longrightarrow HK \\ (h, k) &\longmapsto hk, \end{aligned}$$

mais cette bijection n'est **pas** un morphisme de groupes en général. En fait, le groupe HK est isomorphe à $H \times K$, mais muni de la loi de groupe

$$(h, k) * (h', k') = (h(kh'k^{-1}), kk').$$

Remarquons maintenant que l'application

$$\begin{aligned} \theta \colon K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \theta_k = \mathrm{Int}(k)_{|H} \end{aligned}$$

est un morphisme de groupes et que l'on a

$$(h,k)*(h',k') = (h\theta_k(h'), kk').$$

Plus généralement, on a le résultat suivant.

PROPOSITION IV.3.1. *Soient H et K deux groupes, et soit*

$$\begin{aligned} \rho\colon K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \rho_k \end{aligned}$$

un morphisme de groupes. Alors, la loi interne

$$\begin{aligned} (H\times K)\times(H\times K) &\longrightarrow H\times K \\ ((h,k),(h',k')) &\longmapsto (h,k)*(h',k') = (h\rho_k(h'), kk') \end{aligned}$$

est une loi de groupe, de neutre $(1_H, 1_K)$.

Démonstration. Montrons que $(1_H, 1_K)$ est un élément neutre. On a

$$(h,k)*(1_H,1_K) = (h\rho_k(1_H), k) = (h,k) \text{ pour tout } h\in H, k\in K,$$

car ρ_k est un élément de $\mathrm{Aut}(H)$, donc en particulier un morphisme de groupes. On a aussi

$$(1_H,1_K)*(h',k') = (\rho_{1_K}(h'), k') = (h',k') \text{ pour tout } h'\in H, \text{ et tout } k'\in K,$$

car ρ étant un morphisme de groupes, il envoie 1_K sur Id_H. Donc $(1_H, 1_K)$ est un élément neutre pour la loi interne.

Soit $(h,k)\in H\times K$, et montrons que $(\rho_k^{-1}(h^{-1}), k^{-1})$ est un inverse de (h,k). On a

$$(h,k)*(\rho_k^{-1}(h^{-1}), k^{-1}) = (h\rho_k(\rho_k^{-1}(h^{-1})), 1_K) = (hh^{-1}, 1_K) = (1_H, 1_K).$$

On a aussi

$$(\rho_k^{-1}(h^{-1}), k^{-1})*(h,k) = (\rho_k^{-1}(h^{-1})\rho_{k^{-1}}(h), 1_K).$$

Comme ρ est un morphisme de groupes, on a $\rho_{k^{-1}} = \rho_k^{-1}$, et donc on obtient

$$\begin{aligned} (\rho_k^{-1}(h^{-1}), k^{-1})(h,k) &= (\rho_{k^{-1}}(h^{-1})\rho_{k^{-1}}(h), 1_K) \\ &= (\rho_{k^{-1}}(h^{-1}h), 1_K) \\ &= (\rho_{k^{-1}}(1_H), 1_K) = (1_H, 1_K), \end{aligned}$$

les deux dernières égalités provenant du fait que $\rho_{k^{-1}}$ est un morphisme de groupes. Il reste à vérifier l'associativité. On a d'une part

$$\begin{aligned} \big((h,k)*(h',k')\big)*(h'',k'') &= (h\rho_k(h'), kk')*(h'',k'') \\ &= (h\rho_k(h')\rho_{kk'}(h''), (kk')k''), \end{aligned}$$

et d'autre part

$$\begin{aligned} (h,k)*\big((h',k')*(h'',k'')\big) &= (h,k)*(h'\rho_{k'}(h''), k'k'') \\ &= (h\rho_k(h'\rho_{k'}(h'')), k(k'k'')). \end{aligned}$$

Puisque K est un groupe, on a bien sûr $(kk')k'' = k(k'k'')$. Maintenant, ρ étant un morphisme de groupes, on a $\rho_{kk'} = \rho_k\circ\rho_{k'}$, et donc

$$h\rho_k(h')\rho_{kk'}(h'') = h\rho_k(h')\rho_k(\rho_{k'}(h'')).$$

Comme ρ_k est un morphisme de groupes, on a alors

$$h\rho_k(h')\rho_{kk'}(h'') = h\rho_k(h'\rho_{k'}(h'')).$$

Ceci achève la démonstration. □

Définition IV.3.2. Soient H et K deux groupes, et soit

$$\rho : K \longrightarrow \mathrm{Aut}(H)$$

un morphisme de groupes. L'ensemble $H \times K$ muni de la loi de groupe précédente est appelé *produit semi-direct externe* de H et K par ρ, et est noté $H \times_\rho K$.

Remarque IV.3.3. On vérifie aisément que le sous-groupe $H' = \{(h, 1_K) \mid h \in H\}$ est distingué dans $H \times_\rho K$, et isomorphe à H.

Lemme IV.3.4. *Soient H et K deux groupes, et soit*

$$\begin{aligned} \rho\colon K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \rho_k \end{aligned}$$

un morphisme de groupes. Les propriétés suivantes sont équivalentes :

(1) *on a $H \times_\rho K = H \times K$;*

(2) *le morphisme $\rho : K \longrightarrow \mathrm{Aut}(H)$ est trivial ;*

(3) *le sous-groupe $K' = \{(1_H, k) \mid k \in K\}$ est distingué dans $H \times_\rho K$.*

Démonstration.

(3) $\Longrightarrow$ (2). Soient $h \in H, k \in K$. On a

$$(h, 1_K) * (1_H, k) * (h, 1_K)^{-1} = (h, k) * (h, 1_K)^{-1} = (h, k) * (h^{-1}, 1_K),$$

soit encore

$$(h, 1_K) * (1_H, k) * (h, 1_K)^{-1} = (h\rho_k(h^{-1}), k).$$

Puisque K' est distingué dans $H \times_\rho K$, on en déduit que

$$h\rho_k(h^{-1}) = h\rho_k(h)^{-1} = 1_H \quad \text{pour tout } h \in H, \text{ et tout } k \in K.$$

Autrement dit, on a $\rho_k(h) = h$ pour tout $h \in H$, et tout $k \in K$, i.e. ρ est trivial.

(2) $\Longrightarrow$ (1). Évident.

(1) $\Longrightarrow$ (3). Découle des propriétés du produit direct. □

Remarque IV.3.5. Ce lemme devient faux is on remplace l'égalité par un isomorphisme dans le point (1).

EN effet, soit $\rho : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathfrak{S}_3)$ l'unique morphisme tel que $\rho_{\bar{1}} = \mathrm{Int}_{(12)}$. Alors, ρ est non trivial, mais on peut montrer que l'on a un isomorphisme

$$\mathfrak{S}_3 \times_\rho \mathbb{Z}/2\mathbb{Z} \simeq \mathfrak{S}_3 \times \mathbb{Z}/2\mathbb{Z}.$$

On laisse le lecteur démontrer cette affirmation à titre d'exercice.

On s'intéresse maintenant au produit semi-direct interne.

Définition IV.3.6. Soit G un groupe, et soient H et K deux sous-groupes de G. On dit que G est le *produit semi-direct interne* de H et K si les conditions suivantes sont satisfaites :

(1) le sous-groupe H est distingué dans G ;

(2) tout élément $g \in G$ s'écrit de manière unique $g = hk, h \in H, k \in K$;

On le note $G = H \rtimes K$.

Remarque IV.3.7. Soit G un groupe, et soient H et K deux sous-groupes, avec H distingué dans G. Considérons le morphisme de groupes

$$\begin{aligned} \theta\colon K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \mathrm{Int}(k)_{|H}. \end{aligned}$$

Alors, $G = H \rtimes K$ si, et seulement si, l'application

$$\begin{aligned} H \times_\theta K &\longrightarrow G \\ (h,k) &\longmapsto hk \end{aligned}$$

est un isomorphisme de groupes.

En particulier, un produit semi-direct interne est isomorphe à un certain produit semi-direct externe.

Exemples IV.3.8. Voici quelques exemples de produits semi-directs internes :

(1) $\mathfrak{S}(E) = \mathfrak{A}(E) \rtimes \langle \tau \rangle$, où τ est une transposition fixée ;
(2) $\mathrm{GL}_n(K) = \mathrm{SL}_n(K) \rtimes D$, où D est le sous-groupe de $\mathrm{GL}_n(K)$ formés des matrices de dilatation et de la matrice identité ;
(3) soit $\sigma \in D_n$ une rotation d'angle $\dfrac{2\pi}{n}$ et $\tau \in D_n$ une symétrie orthogonale dont l'axe passe par le centre du n-gone et un des sommets. Posons $H = \langle \sigma \rangle$ et $K = \langle \tau \rangle$. Alors, $D_n = H \rtimes K$.
On a dans ce cas $D_n \simeq H \times_\rho K$, où $\rho_1 = \mathrm{Id}_H$ et ρ_τ est le passage à l'inverse.

Les vérifications sont laissées au lecteur.

Proposition IV.3.9. *Soit G un groupe, et soient H et K deux sous-groupes. Alors, $G = H \rtimes K$ si, et seulement si, les conditions suivantes sont vérifiées :*

(1) *le sous-groupe H est distingué dans G ;*
(2) *on a $G = \langle H, K \rangle$ (ou, de manière équivalente, $G = HK$) ;*
(3) *on a $H \cap K = \{1_G\}$.*

Démonstration. Supposons que $G = H \rtimes K$. Alors, par définition H est distingué dans G. De plus, comme tout élément de G est produit d'un élément de H et d'un élément de K, on a $G = HK$ et donc $G = \langle H, K \rangle$. On a déjà constaté que l'unicité d'écriture dans HK implique (3).

Inversement, les conditions (1) et (2) impliquent les égalités

$$G = HK = \langle H, K \rangle$$

par le lemme **??**. La condition $H \cap K = \{1_G\}$ implique l'unicité de l'écriture, comme on l'a déjà remarqué. □

Remarque IV.3.10. Si $G = H \rtimes K$, alors $G = H \odot K$ si, et seulement si, K est distingué dans G.

En effet, le sens direct est clair. Réciproquement, si K est un sous-groupe distingué de G, on utilise la proposition précédente et le théorème **??** pour conclure.

On continue en donnant des conditions nécessaires et suffisantes pour avoir un produit semi-direct interne dans le cas des groupes finis.

Corollaire IV.3.11. *Soit G un groupe fini, et soient H et K deux sous-groupes. Alors, $G = H \rtimes K$ si, et seulement si, les conditions suivantes sont vérifiées :*

(1) *le sous-groupe H est distingué dans G ;*

(2) *on a* $|G| = |H|\cdot|K|$ *;*

(3) *on a* $H \cap K = \{1_G\}$.

Démonstration. Similaire à celle du corollaire **??**. □

Remarque IV.3.12. Comme on l'a déjà vu un produit semi-direct interne est isomorphe à un certain produit semi-direct externe. Réciproquement, un produit semi-direct externe $H \times_\rho K$ est isomorphe au produit semi-direct interne $H' \rtimes K'$, où on a posé $H' = H \times \{1_K\}$ et $K' = \{1_H\} \times K$ (**Exercice**). On peut donc légitimement se demander l'intérêt d'introduire deux notions différentes. Ceci est également vrai pour les produits directs (i.e. le cas où ρ est trivial).

En fait, l'intérêt de la version externe pour les produits directs ou semi-directs est le même que pour les sommes directes d'espaces vectoriels : celui de pouvoir construire de nouveaux groupes à partir de groupes abstraits ne vivant pas dans un même groupe. La version interne est là pour décortiquer la structure d'un groupe donné.

Chapitre V

Applications des groupes opérants à la structure des groupes finis

V.1. Premières applications

Nous allons maintenant donner quelques applications de la théorie des groupes opérants. D'autres seront proposées dans les exercices et dans les chapitres suivants.

On commence par un résultat dû à Cayley.

THÉORÈME V.1.1 (Cayley). *Soit G un groupe fini d'ordre n. Alors, G est isomorphe à un sous-groupe de $\mathfrak{S}_n$.*

Démonstration. Pour le voir, faisons opérer G sur lui-même par translation à gauche. Cette action étant fidèle (cf. exemple **??** (1)), le morphisme de groupes correspondant

$$\varphi : G \longrightarrow \mathfrak{S}(G)$$

est injectif. Puisque G est d'ordre n, on a un isomorphisme [1]

$$\mathfrak{S}(G) \simeq \mathfrak{S}_n$$

et en composant par φ, on obtient un morphisme injectif $G \hookrightarrow \mathfrak{S}_n$, qui identifie donc G à son image dans $\mathfrak{S}_n$. □

Avant de donner d'autres applications, nous allons détailler un exemple d'action de groupe.

Soit p un nombre premier fixé, et soit $n \geq 1$ un entier. Considérons G un groupe d'ordre n (il en existe au moins un, par exemple $G = \mathbb{Z}/n\mathbb{Z}$), et soit

$$E = \{x = (g_1, \ldots, g_p) \in G^p \mid g_1 \cdots g_p = 1_G\}.$$

C'est un ensemble fini, de cardinal n^{p-1}. En effet, l'ensemble E se récrit

$$E = \{x = (g_1, \ldots, g_{p-1}, (g_1 \cdots g_{p-1})^{-1}) \mid g_1, \ldots, g_{p-1} \in G\}.$$

Rappelons que $\mathfrak{S}_p$ agit sur G^p par

$$\begin{aligned} \mathfrak{S}_p \times G^p &\longrightarrow G^p \\ (\tau, (g_1, \ldots, g_p)) &\longmapsto (g_{\tau^{-1}(1)}, \ldots, g_{\tau^{-1}(p)}). \end{aligned}$$

Soit $\sigma = (1\ p\ p-1 \cdots 2) \in \mathfrak{S}_p$, et soit $H = \langle \sigma \rangle$. Rappelons que l'on a

$$H = \{\mathrm{Id}, \sigma, \ldots, \sigma^{p-1}\},$$

puisque σ est d'ordre p. Nous allons vérifier que l'action de $\mathfrak{S}_p$ sur G^p se restreint en une action de H sur E.

Notons tout d'abord que si $(g_1, g_2 \ldots, g_p) \in E$, alors

$$(g_{\sigma^{-1}(1)}, \ldots, g_{\sigma^{-1}(p)}) = (g_2, \ldots, g_p, g_1) \in E.$$

1. Le lecteur pointilleux pourra, soit établir cet isomorphisme lui-même, soit se référer à l'exemple **??** du chapitre suivant.

En effet, si $g_1 g_2 \cdots g_p = 1_G$, alors

$$g_1^{-1}(g_1 g_2 \cdots g_p) g_1 = g_2 \cdots g_p g_1 = g_1^{-1} g_1 = 1_G.$$

Par récurrence, on obtient que si $(g_1, g_2 \ldots, g_p) \in E$, alors

$$(g_{\sigma^{-m}(1)}, \ldots, g_{\sigma^{-m}(p)}) \in E \text{ pour tout } m \in [\![0, p-1]\!],$$

ce qu'il fallait vérifier.

Étudions les orbites de E sous l'action de H d'un peu plus près. D'après la proposition **??**, le cardinal d'une orbite divise l'ordre de H, qui est p. Comme p est premier, on en déduit que le cardinal d'une orbite est 1 ou p. Décrivons maintenant les points fixes. Clairement, $x \in E^H$ si, et seulement si, toutes ses coordonnées sont égales.

Pour tout $x \in E$, on a donc $|\mathcal{O}_x| = 1$ si, et seulement si, on a $x = (g, \ldots, g)$, où $g \in G$ vérifie $g^p = 1_G$. Remarquons que

$$g^p = 1_G \iff o(g) \mid p \iff o(g) = 1 \text{ ou } p.$$

Ainsi, $|\mathcal{O}_x| = 1$ si, et seulement si, $x = (g, \ldots, g)$, avec $g = 1_G$ ou $o(g) = p$. Notons qu'il y a au moins une orbite réduite à un élément, à savoir l'orbite de $(1_G, \ldots, 1_G)$.

Nous allons maintenant montrer l'équivalence des propriétés suivantes :

(1) on a $p \nmid n$;

(2) le groupe G ne possède pas d'élément d'ordre p ;

(3) il n'y a qu'une seule orbite réduite à un élément. Dans ce cas, on a $n^{p-1} \equiv 1\ [p]$.

(1) $\Longrightarrow$ (2). Supposons que $p \nmid n$. Alors, G ne possède pas d'éléments d'ordre p. En effet, si un tel élément existait, son ordre diviserait l'ordre de G, qui est n, ce qui contredirait le choix de p.

(2) $\Longrightarrow$ (3). Si G ne possède pas d'éléments d'ordre p, il n'y a donc qu'une seule orbite réduite à un élément d'après l'analyse précédente. Mais alors, toutes les autres orbites ont p éléments. Comme les orbites distinctes forment une partition de E, le cardinal de E est la somme du cardinal des orbites [2]. En réduisant modulo p, on obtient $n^{p-1} \equiv 1\ [p]$.

(3) $\Longrightarrow$ (1). Supposons qu'il y ait une unique orbite réduite à un élément. On a alors $n^{p-1} \equiv 1\ [p]$. En particulier, on a $p \nmid n$.

Les trois propriétés sont donc bien équivalentes.

En appliquant (1) $\Longrightarrow$ (3), on retrouve le petit théorème de Fermat.

Théorème V.1.2 (Fermat). *Soit $n \geq 1$ un entier, et soit p un nombre premier. Si $p \nmid n$, alors $n^{p-1} \equiv 1\ [p]$.*

La contraposée de l'implication (2) $\Longrightarrow$ (1) nous donne aussi un premier résultat non trivial sur la structure des groupes finis.

Théorème V.1.3 (Cauchy). *Soit G un groupe fini, et soit p un nombre premier. Si $p \mid |G|$, alors G possède au moins un élément d'ordre p.*

Continuons avec une autre application des actions de groupe. Le résultat suivant généralise le fait que tout sous-groupe d'indice 2 dans un groupe G est distingué dans G.

2. Plus rapidement, c'est simplement l'équation aux classes.

Théorème V.1.4 (Frobenius). *Soit G un groupe fini, et soit H un sous-groupe de G. Si tout diviseur premier de $|H|$ est $\geq [G : H]$, alors H est distingué dans G. En particulier, si p est le plus petit diviseur premier de $|G|$, alors tout sous-groupe d'indice p est distingué dans G.*

Démonstration. Pour appliquer la théorie des groupes opérants, nous allons retraduire la conclusion voulue en termes d'actions de groupe. On peut supposer que $H \neq G$, le cas $H = G$ étant trivial.

Soit $E = G/H$ l'ensemble des classes à gauche de G modulo H. Soit $h \in H$. Alors, on a les équivalences suivantes :

$$\begin{aligned} \text{pour tout } g \in G,\ ghg^{-1} \in H \quad &\Longleftrightarrow \quad \text{pour tout } g \in G,\ g^{-1}hg \in H \\ &\Longleftrightarrow \quad \text{pour tout } g \in G,\ hg \in gH \\ &\Longleftrightarrow \quad \text{pour tout } g \in G,\ \overline{hg} = \overline{g} \text{ dans } E. \end{aligned}$$

L'action de G sur E de l'exemple **??** (5) induit une action de H sur E par restriction, donnée par

$$\begin{aligned} H \times E &\longrightarrow E \\ (h, \overline{g}) &\longmapsto \overline{hg}. \end{aligned}$$

Au vu des équivalences précédentes, H est distingué dans G si, et seulement si, tout élément de E est un point fixe sous l'action de H. Supposons donc qu'il existe $\overline{g} \in E$ qui ne soit pas un point fixe, et soit $\mathcal{O}_{\overline{g}}$ l'orbite correspondante. Alors, $\mathcal{O}_{\overline{g}}$ a au moins deux éléments.

Puisque E est réunion disjointe des orbites distinctes, on a en particulier

$$|\mathcal{O}_{\overline{g}}| \leq |E| = [G : H].$$

Soit ℓ un diviseur premier de $|\mathcal{O}_{\overline{g}}|$ (qui existe, car $|\mathcal{O}_{\overline{g}}| \geq 2$). Alors, d'après la proposition **??**, on a

$$\ell \mid |\mathcal{O}_{\overline{g}}| \mid |H|.$$

Par hypothèse, on a alors

$$|\mathcal{O}_{\overline{g}}| \geq \ell \geq [G : H].$$

On a donc

$$|\mathcal{O}_{\overline{g}}| = [G : H] = |E|,$$

et en particulier $\mathcal{O}_{\overline{g}}$ est l'unique orbite de E sous l'action de G. On a donc

$$E = G/H = \mathcal{O}_{\overline{g}} = \mathcal{O}_{\overline{1}_G} = \{\overline{1}_G\},$$

d'où la contradiction $G = H$. Ainsi, tous les points de E sont fixes sous l'action de H, et H est distingué dans G. Le dernier point est une application directe du point précédent. □

Nous finissons ce paragraphe en donnant un résultat général sur l'existence de points fixes sous une action donnée. L'existence de tels points fixes n'est en général pas assurée. Par exemple, il suffit de considérer l'action d'un groupe G sur lui-même par translation à gauche.

En revanche, nous allons voir que la situation devient nettement plus sympathique pour une certaine classe de groupes finis.

Définition V.1.5. Soit p un nombre premier. Un *p-groupe* est un groupe fini dont l'ordre est une puissance de p.

On a alors le résultat suivant.

Théorème V.1.6. *Soit G un p-groupe agissant sur un ensemble fini E. Alors, on a*

$$|E| \equiv |E^G| \ [p].$$

Si de plus $p \nmid |E|$, alors $E^G \neq \emptyset$.

Démonstration. On peut toujours supposer G non réduit à 1_G, sinon le résultat est clair (tout élément de E est fixé par 1_G). Posons $E' = E \setminus E^G$. Nous allons vérifier que E' est stable sous l'action de G. Cela revient à montrer que, pour tout $g \in G$ et tout $x \in E$, si $g{\cdot}x \in E^G$, alors $x \in E^G$.

Soient $g \in G$ et $x \in E$. Si $g{\cdot}x \in E^G$, on a en particulier

$$(gg'g^{-1}){\cdot}(g{\cdot}x) = g{\cdot}x \text{ pour tout } g' \in G,$$

c'est-à-dire

$$gg'{\cdot}x = g{\cdot}x \text{ pour tout } g' \in G.$$

En faisant agir g^{-1} des deux côtés de cette égalité, on en déduit

$$g'{\cdot}x = x \text{ pour tout } g' \in G.$$

Ainsi, $x \in E^G$. L'ensemble E' est alors stable sous l'action de G. En particulier, cette action se restreint en une action sur E'.

Pour tout $x' \in E'$, on a $|\mathcal{O}_{x'}| \mid |G|$ par la proposition **??**. De plus, l'orbite $\mathcal{O}_{x'}$ n'est pas réduite à un élément, car x' n'est pas un point fixe. Par conséquent, $|\mathcal{O}_{x'}|$ est une puissance non triviale de p (puisque $|G|$ est une puissance de p). Comme l'ensemble des orbites distinctes forment une partition de E', $|E'|$ est la somme du cardinal de ces orbites, et donc on obtient

$$|E'| = |E| - |E^G| \equiv 0 \ [p],$$

soit encore

$$|E| \equiv |E^G| \ [p].$$

Supposons maintenant que $p \nmid |E|$. Si E^G était vide, on aurait $|E| \equiv 0 \ [p]$ d'après la congruence précédente, d'où une contradiction. □

On en déduit alors le résultat fort utile suivant.

Proposition V.1.7. *Soit G un p-groupe non trivial. Alors, son centre $Z(G)$ est non trivial.*

Démonstration. On fait agir G sur lui-même par conjugaison. Alors, $Z(G)$ est exactement l'ensemble des points fixes pour cette action. D'après le théorème **??**, on a

$$|G| \equiv |Z(G)| \ [p].$$

Si l'on avait $Z(G) = \{1_G\}$, on obtiendrait

$$|G| \equiv 1 \ [p].$$

Mais, par hypothèse $p \mid |G|$, et on obtient une contradiction. □

De tout ceci, on peut déduire le théorème suivant.

Théorème V.1.8. *Soit G un p-groupe d'ordre p^n. Alors il existe une suite de sous-groupes distingués de G*

$$\{1_G\} = G_0 \subset G_1 \subset \cdots \subset G_{n-1} \subset G_n = G$$

telle que G_i soit d'ordre p^i pour tout $0 \leq i \leq n$.

En particulier, G_i/G_{i-1} est cyclique pour tout $i \in [\![1, n]\!]$.

Démonstration. Soit G un groupe d'ordre p^n, avec p premier, et $n \geq 0$. Le cas $n = 0$ étant évident, on peut supposer que $n \geq 1$.

On procède par récurrence sur n. Soit (H_n) la propriété :

(H_n) Le théorème est vrai pour tout groupe d'ordre p^n.

Si $n = 1$, G est cyclique d'ordre p, et le résultat est clair. Ainsi, (H_1) est vraie. Supposons que (H_n) soit vraie pour un entier $n \geq 1$, et montrons (H_{n+1}). Soit G un groupe d'ordre p^{n+1}. Par la proposition **??**, $Z(G)$ est non trivial. En particulier, son ordre est divisible par p, et le lemme de Cauchy assure l'existence d'un élément $x \in Z(G)$ d'ordre p. Alors, $\langle x \rangle$ est cyclique d'ordre p, et distingué dans G car $x \in Z(G)$. Mais, alors, le groupe $G/\langle x \rangle$ est d'ordre p^n. Par hypothèse de récurrence, il existe une suite de sous-groupes

$$G_0/\langle x \rangle \subset \cdots \subset G_n/\langle x \rangle$$

de $G/\langle x \rangle$ telle que $G_i/\langle x \rangle$ est un sous-groupe distingué de $G/\langle x \rangle$ d'ordre p^i, et ceci pour tout $i \in [\![0, n]\!]$. Puisque $G_i/\langle x \rangle$ est distingué dans $G/\langle x \rangle$, G_i est distingué dans G par la proposition **??**. De plus, par le théorème de Lagrange, on a

$$|G_i| = |G_i/\langle x \rangle| \cdot |\langle x \rangle| = p^{i+1}.$$

Ainsi, la suite de sous-groupes

$$\langle x \rangle, G_0, \ldots, G_n$$

vérifie les conditions demandées. Ceci achève la récurrence.

Le dernier point vient du fait que, pour tout $i \in [\![1, n]\!]$, G_i/G_{i-1} est d'ordre p premier, donc cyclique. □

On finit ce paragraphe en énonçant quelques exercices.

Exercice : Soit G un groupe fini, soit p le plus petit diviseur premier de $|G|$, et soit H un sous-groupe distingué d'ordre p. Montrer que $H \subset Z(G)$(*Indication :* Faire opérer G sur H par conjugaison).

Exercice : Soit G un groupe infini. On suppose que G admet un sous-groupe H, distinct de G et d'indice fini. Montrer que G n'est pas simple (*Indication :* Faire opérer G sur G/H par translation).

Exercice : Soit G un groupe fini. On fait opérer G sur lui-même par translation à gauche. Soit $g \in G$, et soit $\sigma_g \in S(G)$ la permutation correspondante.

(1) Montrer que $\varepsilon(\sigma_g) = (-1)^{\frac{(o(g)-1)|G|}{o(g)}}$.

(2) Soit G un groupe fini avec un 2-Sylow cyclique. En utilisant la question précédente, construire un morphisme non trivial

$$\rho : G \longrightarrow \{\pm 1\}.$$

En déduire que G n'est pas simple dès que $|G| \geq 3$.

V.2. Théorème de Sylow

Que devient un p-Sylow lorsqu'il est devenu fou ? Un Sylow à grain.

Soit G un groupe fini. Comme nous l'avons vu auparavant, l'ordre de tout sous-groupe H divise l'ordre du groupe, mais étant donné un diviseur d de l'ordre de G, il n'existe pas nécessairement de sous-groupe de G d'ordre d. Grâce aux actions de groupes, nous allons voir que cela devient vrai lorsque d est une puissance d'un nombre premier.

Définition V.2.1. Soit G un groupe fini, et soit p un nombre premier.

Un *p-sous-groupe* de G est un sous-groupe de G qui est un p-groupe.

Écrivons $|G| = p^m q$, où $p \nmid q$ et $m \geq 0$. Un *p-sous-groupe de Sylow* de G, ou p-Sylow est un p-sous-groupe de G d'ordre p^m. Autrement dit, un sous-groupe S est un p-Sylow du groupe G si, et seulement si, S est un p-groupe et $p \nmid [G : S]$.

Évidemment, rien ne garantit qu'un tel p-Sylow existe. Nous allons maintenant donner un exemple de groupes où un p-Sylow existe bel et bien.

Exemple V.2.2. Soit $G = \mathrm{GL}_n(\mathbb{F}_p)$. C'est un groupe fini d'ordre

$$|G| = (p^n - 1)(p^n - p) \cdots (p^n - p^{n-1}).$$

En effet, pour qu'une matrice soit inversible, il faut et il suffit que ces colonnes soient linéairement indépendantes. On a donc $p^n - 1$ choix pour la première colonne (on ne peut pas prendre une colonne nulle). La deuxième colonne ne doit pas être multiple de la première. On a donc $p^n - p$ choix. La troisième ne doit pas être combinaison linéaire des deux premières, on a donc $p^n - p^2$ choix, etc.

On a donc

$$|G| = (p^n - 1) \cdots (p^n - p^{n-1}) = p^{1+2+\cdots+(n-1)} q = p^{n(n-1)/2} q, \ p \nmid q.$$

Soit S le sous-groupe des matrices triangulaires supérieures avec des $\overline{1}$ sur la diagonale. Autrement dit,

$$S = \{(a_{ij}) \in G \mid a_{ij} = \overline{0} \text{ si } i > j, a_{ii} = \overline{1}\}.$$

Alors, S est un sous-groupe de G d'ordre $p^{n(n-1)/2}$, donc un p-Sylow.

Nous allons maintenant montrer l'existence d'un p-Sylow pour tout groupe G. On rappelle que si H un sous-groupe, et si $g \in G$, alors

$$gHg^{-1} = \{ghg^{-1} \mid g \in G\}.$$

C'est un sous-groupe de G, puisque c'est l'image de H par $\mathrm{Int}(g)$. De plus, si H est fini, alors il en est de même de gHg^{-1}, et on a

$$|gHg^{-1}| = |H|,$$

puisque Int_g est un automorphisme.

Si H et H' deux sous-groupes, on rappelle que H et H' sont *conjugués* s'il existe $g \in G$ tel que $H' = gHg^{-1}$. On dit aussi que H' est un conjugué de H (cf. définition **??**).

Remarque V.2.3. Soit G un groupe fini, et soit p un nombre premier. Alors, tout conjugué d'un p-Sylow de G est encore un p-Sylow de G (cela provient des considérations précédentes).

Le point-clé pour établir l'existence d'un p-Sylow est le lemme suivant.

Lemme V.2.4. *Soit G un groupe fini avec $|G| = n = p^m q, m \geq 0, p \nmid q$, et soit H un sous-groupe de G. Supposons que G admette un p-Sylow S. Alors, il existe $g \in G$ tel que $gSg^{-1} \cap H$ soit un p-Sylow de H.*

Démonstration. Soit $X = G/S$. Alors, G opère par translation sur G/S par l'exemple **??** (5). Cela induit alors une action de H sur G/S par restriction. Si $\overline{g} \in G/S$, on a alors

$$\mathrm{Stab}_H(\overline{g}) = \{h \in H \mid \overline{g} = \overline{hg}\} = \{h \in H \mid g^{-1}(hg) \in S\} = gSg^{-1} \cap H.$$

Son ordre est donc une puissance de p. En effet, gSg^{-1} est un p-Sylow de G par la remarque **??**. Comme $gSg^{-1} \cap H$ est un sous-groupe de gSg^{-1}, c'est aussi un p-groupe (puisque l'ordre d'un sous-groupe divise l'ordre du groupe). Pour montrer le lemme, il suffit donc de vérifier qu'il existe $g \in G$ tel que $[H : gSg^{-1} \cap H]$ soit premier à p. D'après la proposition **??**, on a

$$[H : gSg^{-1} \cap H] = \mathrm{card}(\mathcal{O}_{\overline{g}}).$$

Si toutes les orbites de G/S sous l'action de G avaient un cardinal divisible par p, alors $\mathrm{card}(G/S)$ serait aussi divisible par p, car $|G/S|$ est la somme des cardinaux des orbites distinctes. Or, S étant un p-Sylow du groupe G, on a $p \nmid [G : S]$. Ainsi, on obtient une contradiction, et on a fini. □

Nous allons maintenant montrer le résultat suivant.

Théorème V.2.5 (Sylow). *Soit G un groupe fini, et soit p un nombre premier. Écrivons $|G| = p^m q, m \geq 0, p \nmid q$. Alors :*

(1) *il existe des p-Sylow de G, et tout p-sous-groupe de G est contenu dans un p-Sylow ;*

(2) *le conjugué d'un p-Sylow de G est un p-Sylow de G, et tous les p-Sylow de G sont conjugués. En particulier, si S est un p-Sylow de G, alors S est distingué dans G si, et seulement si, S est l'unique p-Sylow de G ;*

(3) *si N_p est le nombre de p-Sylow de G, on a $N_p \equiv 1\ [p]$ et $N_p \mid q$.*

Démonstration. Commençons par démontrer l'existence d'un p-Sylow.

D'après le théorème de Cayley, le groupe G s'identifie à un sous-groupe de $\mathfrak{S}_n$, où $n = |G|$. Maintenant, $\mathfrak{S}_n$ s'identifie lui-même à un sous-groupe de $\mathrm{GL}_n(\mathbb{F}_p)$ de la manière suivante. Si $\sigma \in \mathfrak{S}_n$, on pose $M_\sigma = (m_{ij})$, où

$$m_{\sigma(j)j} = \overline{1} \ \text{ et } \ m_{ij} = \overline{0} \text{ si } i \neq \sigma(j).$$

On vérifie alors aisément que l'application

$$\begin{aligned} \mathfrak{S}_n &\longrightarrow \mathrm{GL}_n(\mathbb{F}_p) \\ \sigma &\longmapsto M_\sigma \end{aligned}$$

est un morphisme de groupes injectif. Ainsi, G est isomorphe à un sous-groupe G' de $\mathrm{GL}_n(\mathbb{F}_p)$. D'après le lemme précédent et l'exemple **??**, G' admet un p-Sylow, et donc il en est de même de G.

Soit maintenant H un p-sous-groupe de G et soit S un p-Sylow de G. Par le lemme **??**, il existe $g \in G$ tel que $gSg^{-1} \cap H$ soit un p-Sylow de H. Mais comme H est un p-groupe, on a donc $gSg^{-1} \cap H = H$. Ainsi $H \subset gSg^{-1}$, qui est un p-Sylow de G d'après la remarque **??**. Si maintenant H est un autre p-Sylow de G, on a alors $H = gSg^{-1}$ (puisque H et gSg^{-1} ont alors même ordre). Si de plus S est le seul p-Sylow de G, on a donc $gSg^{-1} = S$ pour tout $g \in G$, et S est donc distingué dans G. Réciproquement, si S est un p-Sylow de G qui est distingué dans G, alors on a $gSg^{-1} \subset S$. Comme ces deux groupes ont même ordre, on obtient donc $gSg^{-1} = S$. Comme tous les p-Sylow sont conjugués, on obtient que S est l'unique p-Sylow de G. Ceci établit les points (1) et (2).

Montrons (3). Soit E l'ensemble des p-Sylow de G. D'après le point (2), l'action de G sur E par conjugaison est transitive. Autrement dit, si S est un p-Sylow de G, on a $\mathcal{O}_S = E$. Comme le cardinal d'une orbite divise l'ordre de G par la proposition

??, on a $|E| = N_p \mid n$. Remarquons maintenant que S opère aussi sur E, et que l'on a

$$N_p \equiv |E^S| \ [p],$$

en vertu du théorème **??**. Nous allons montrer que $E^S = \{S\}$ (S est bien entendu fixe sous l'action de S, puisque $sSs^{-1} = S$ pour tout $s \in S$). Soit T un p-Sylow, et supposons que $T \in E^S$. On a donc

$$sTs^{-1} = T \quad \text{pour tout } s \in S.$$

Soit G' le sous-groupe de G engendré par S et T. On a $S \subset G'$ et $T \subset G'$, et ce sont deux p-Sylow de G'. En effet, puisque G' est un sous-groupe de G, son ordre divise celui de G. On a donc

$$|G'| = p^\ell q', \ell \in [\![0, m]\!], p \nmid q'.$$

Comme G' contient S, qui est d'ordre p^m, on a donc $\ell = m$, et les p-Sylow de G' sont donc d'ordre p^m. Puisque $gTg^{-1} \subset T$ pour tout $g \in S \cup T$, on a $xTx^{-1} \subset T$ pour tout $x \in G'$. Ainsi, T est un p-Sylow de G' qui est distingué dans G'. C'est donc le seul, d'où $T = S$. Ainsi, $\mathrm{card}(E^S) = 1$, et donc

$$N_p \equiv 1 \ [p].$$

Puisque $N_p \mid p^m q$ et que N_p est donc premier à p (du fait de la congruence précédente), on en déduit que $N_p \mid q$. Ceci achève la démonstration. □

Voici deux exemples d'application de ce théorème.

Exemple V.2.6. Un groupe d'ordre 63 n'est pas simple.

En effet, soit G un tel groupe. Puisque $63 = 3^2 \cdot 7$, on a $N_7 \mid 9$, et donc nécessairement $N_7 = 1, 3$ ou 9. Mais on a aussi $N_7 \equiv 1$ [7], donc $N_7 = 1$. Ainsi, G n'admet qu'un 7-Sylow, qui est alors distingué dans G. Ainsi, G n'est pas simple.

Exemple V.2.7. Un groupe d'ordre 33 est cyclique.

En effet, si G est d'ordre $33 = 3 \cdot 11$, on a donc $N_3 \mid 11$ et donc $N_3 = 1$ ou 11. Mais on a aussi $N_3 \equiv 1$ [3], donc $N_3 = 1$. Ainsi, un 3-Sylow H_1 est distingué dans G.

De même, on a $N_{11} \mid 3$ et $N_{11} \equiv 1$ [11], et donc $N_{11} = 1$. Par conséquent, un 11-Sylow H_2 est distingué dans G.

Comme 3 et 11 sont premiers entre eux, on obtient $H_1 \cap H_2 = \{1_G\}$. De plus, on a $|G| = |H_1| \cdot |H_2|$. Par le corollaire **??**, on a $G = H_1 \odot H_2$. Comme 3 et 11 sont des nombres premiers, les groupes H_1 et H_2 sont cycliques. Soit h_1 un générateur de H_1, et soit h_2 un générateur de H_2. Alors, on a $o(h_1) = 3$ et $o(h_2) = 11$. Puisque $G = H_1 \odot H_2$, les éléments de H_1 commutent aux éléments de H_2. On a alors $o(h_1 h_2) = 33$ d'après la proposition **??**.

Le groupe G étant d'ordre 33, on en déduit que G est cyclique, engendré par $h_1 h_2$.

Deuxième partie

Anneaux : battons-nous pour nos idéaux !

Chapitre VI

Propriétés élémentaires des anneaux et des corps

VI.1. Généralités

Rappelons la définition d'anneau, déjà introduite dans le chapitre **??** (cf. définition **??**).

Définition VI.1.1. Un *anneau* est un triplet $(A, +, \cdot)$, où A est un ensemble non vide, $+ : A \times A \longrightarrow A$ et $\cdot : A \times A \longrightarrow A$ sont deux lois internes, vérifiant les propriétés suivantes :

(1) le couple $(A, +)$ est un groupe abélien ;
(2) la loi « $\cdot$ » est associative, et possède un élément neutre ;
(3) la loi « $\cdot$ » est distributive par rapport à « $+$ ».

Les éléments neutres pour les lois « $+$ » et « $\cdot$ » seront notés 0_A et 1_A respectivement.

On dit que l'anneau $(A, +, \cdot)$ est *commutatif* si « $\cdot$ » est commutative.

Attention ! $(A, \cdot)$ n'est pas un groupe multiplicatif (penser à $A = \mathbb{Z}$) !

Remarque VI.1.2. Le neutre 1_A pour la multiplication est unique d'après le lemme **??**.

Exemples VI.1.3.

(1) Les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ munis de leurs lois usuelles sont des anneaux commutatifs.
(2) Pour tout $n \geq 0$, l'ensemble $\mathbb{Z}/n\mathbb{Z}$, muni des lois
$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a+b} \end{aligned} \quad \text{et} \quad \begin{aligned} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{ab} \end{aligned}$$
est un anneau commutatif.
(3) Soit A un anneau, et soit $n \geq 1$ un entier. Soit $\mathrm{M}_n(A)$ l'ensemble des matrices $n \times n$ à coefficients dans A. Si $M = (a_{ij})$, et $N = (b_{ij})$, on pose
$$M + N = (a_{ij} + b_{ij}) \quad \text{et} \quad MN = \Big(\sum_{k=1}^{n} a_{ik} b_{kj}\Big).$$
On obtient un anneau non commutatif si $n \geq 2$ (même si A l'est).
(4) Si $A_1, \ldots, A_n$ sont des anneaux, alors $A_1 \times \cdots \times A_n$ est un anneau pour les lois
$$\begin{aligned} (a_1, \ldots, a_n) + (a'_1, \ldots, a'_n) &= (a_1 + a'_1, \ldots, a_n + a'_n) \\ (a_1, \ldots, a_n)(a'_1, \ldots, a'_n) &= (a_1 a'_1, \ldots, a_n a'_n), \end{aligned}$$
appelé *produit direct* des anneaux $A_1, \ldots, A_n$.

Les éléments neutres pour l'addition et la multiplication sont respectivement $(0_{A_1}, \ldots, 0_{A_n})$ et $(1_{A_1}, \ldots, 1_{A_n})$.

(5) Si X est un ensemble et A est un anneau, alors l'ensemble A^X des fonctions $f : X \to A$ est un anneau pour les lois suivantes : pour toutes fonctions $f, g \in A^X$, on définit les fonctions $f + g$ et fg par

$$(f+g)(x) = f(x) + g(x), \quad \text{et} \quad (fg)(x) = f(x)g(x) \text{ pour tout } x \in X.$$

(6) Soit $d \in \mathbb{Z} \setminus \{0\}$ un entier sans facteurs carrés. On définit $\sqrt{d}$ comme étant l'unique racine carrée positive de d dans $\mathbb{R}$ si $d > 0$, et comme égale à $i\sqrt{-d}$ si $d < 0$.

Soient $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Q}[\sqrt{d}]$ les sous-ensembles de $\mathbb{C}$ définis par :

$$\mathbb{Z}[\sqrt{d}] = \big\{a + b\sqrt{d} \in \mathbb{C} \mid a, b \in \mathbb{Z}\big\},$$
$$\mathbb{Q}[\sqrt{d}] = \big\{a + b\sqrt{d} \in \mathbb{C} \mid a, b \in \mathbb{Q}\big\}.$$

Alors, $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Q}[\sqrt{d}]$ sont des anneaux commutatifs pour les lois de l'anneau $\mathbb{C}$.

De plus, si $d \equiv 1[4]$, il en est de même de

$$\mathbb{Z}\big[\frac{-1+\sqrt{d}}{2}\big] = \big\{a + b\frac{-1+\sqrt{d}}{2} \mid a, b \in \mathbb{Z}\big\}.$$

La vérification de ces faits est laissée en exercice.

Attention ! La dernière partie est fausse si $d \not\equiv 1[4]$. On laisse le soin au lecteur de trouver un contre-exemple.

Par exemple, les ensembles

$$\mathbb{Z}[i], \mathbb{Z}[\sqrt{2}], \mathbb{Z}[i\sqrt{2}], \mathbb{Z}[j], \mathbb{Q}[i], \mathbb{Q}[\sqrt{2}], \mathbb{Q}[i\sqrt{2}], \mathbb{Q}[j]$$

sont des anneaux commutatifs pour les lois de $\mathbb{C}$ (où $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$).

Le résultat suivant montre que les résultats familiers comme les règles de signes sont valables dans n'importe quel anneau.

Proposition VI.1.4. *Soit A un anneau. Pour tous $a, b \in A$, on a*

(1) $a\ 0_A = 0_A = 0_A\ a$;
(2) $a(-b) = -(ab) = (-a)b$;
(3) $(-1_A)a = -a$;
(4) $-(-a) = a$;
(5) $(-a)(-b) = ab$;
(6) $(-1_A)(-1_A) = 1_A$.

De plus, si A est commutatif, pour tous $a, b \in A$, et pour tout entier $n \geq 0$, on a

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot a^k b^{n-k}.$$

Démonstration. Par exemple, pour (1), puisque $0_A = 0_A + 0_A$, on

$$a\ 0_A = a(0_A + 0_A) = a\ 0_A + a\ 0_A.$$

On a donc $0_A = a\ 0_A + (-(a\ 0_A)) = a\ 0_A + a\ 0_A + (-(a\ 0_A)) = a\ 0_A$. La démonstration des autres identités est laissée au lecteur. □

Cette proposition implique le fait suivant : soit A un anneau pour lequel on a $0_A = 1_A$. Dans ce cas, A possède exactement un élément, puisque pour tout $a \in A$, on a

$$a = a\ 1_A = a\ 0_A = 0_A.$$

Définition VI.1.5. Un anneau A dans lequel $1_A = 0_A$ est appelé *anneau trivial*.

Comme pour le cas des groupes, il peut être pénible de montrer qu'un ensemble est un anneau. Heureusement, on a la notion de sous-anneau, qui joue le même rôle que celui de sous-groupe.

Définition VI.1.6. Soit A un anneau. Un *sous-anneau* de A est un sous-ensemble B de A vérifiant les conditions suivantes :

(1) on a $1_A \in B$;
(2) pour tous $x, y \in B$, $x + y, xy, -x \in B$.

Dans ce cas, l'axiome (2) montre que les lois internes de A induisent par restriction deux applications $B \times B \longrightarrow B$. De plus, on a $0_A \in B$. En effet, on a $1_A \in B$ par l'axiome (1), et donc $0_A = 1_A - 1_A \in B$ par l'axiome (2). Il est alors immédiat que B muni de ces deux opérations est un anneau, de neutres $0_B = 0_A$ et $1_B = 1_A$.

Autrement dit, un sous-anneau de A est un anneau pour les lois induites par celles de A.

Comme pour le cas des groupes, pour montrer qu'un ensemble muni de certaines opérations est un anneau, il est plus rapide de montrer que c'est un sous-anneau d'un anneau connu.

Exemples VI.1.7.

(1) L'ensemble $\mathbb{Z}[\sqrt{d}]$ est un sous-anneau de $\mathbb{C}$.

De même, si $d \equiv 1\ [4]$, $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ est un sous-anneau de $\mathbb{C}$.

(2) L'anneau $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$.
(3) L'ensemble $\mathrm{M}_n(\mathbb{Z})$ est un sous-anneau de $\mathrm{M}_n(\mathbb{Q})$.
(4) Plus généralement, si B est un sous-anneau de A, alors $\mathrm{M}_n(B)$ est un sous-anneau de $\mathrm{M}_n(A)$.
(5) L'ensemble $C^0([0,1], \mathbb{R})$ des fonctions continues $f : [0,1] \longrightarrow \mathbb{R}$ est un sous-anneau de l'anneau des fonctions de $[0,1]$ dans $\mathbb{R}$.
(6) L'ensemble $Z(A) = \{z \in A \mid az = za \text{ pour tout } a \in A\}$ est un sous-anneau de A, appelé le *centre* de A. [1]

La démonstration du résultat suivant est laissée en exercice.

Proposition VI.1.8. *Soit A un anneau. L'intersection d'une famille quelconque de sous-anneaux de A est encore un sous-anneau de A.*

Maintenant que nous avons défini les objets avec lesquels nous allons travailler, il reste à définir les morphismes entre les objets.

Définition VI.1.9. Un *morphisme d'anneaux* $f : A \longrightarrow B$ est une application satisfaisant :

1. Attention ! C'est bien le produit de l'anneau qui est utilisé ici. Ainsi, $Z(A)$ n'est pas défini comme le centre du groupe $(A, +)$, ce qui n'aurait d'ailleurs aucun intérêt puisque, $(A, +)$ étant abélien, on aurait alors $Z(A) = A$.

(1) $f(1_A) = 1_B$;
(2) pour tous $a, a' \in A$, $f(a + a') = f(a) + f(a')$;
(3) pour tous $a, a' \in A$, $f(aa') = f(a)f(a')$.

On note $\mathrm{Hom}(A, B)$ l'ensemble des morphismes d'anneaux $f : A \longrightarrow B$. Puisque $(B, +)$ est abélien, $\mathrm{Hom}(A, B)$ est un groupe abélien. On le notera aussi parfois $\mathrm{Hom}_{\mathrm{Ann}}(A, B)$ s'il y a risque de confusion avec l'ensemble des morphismes de groupes de A vers B.

Un *isomorphisme d'anneaux* $f : A \longrightarrow B$ est un morphisme d'anneaux bijectif. Dans ce cas, on vérifie comme dans la remarque **??** (2) que l'application inverse est aussi un isomorphisme d'anneaux. On dit que deux anneaux A et B sont isomorphes s'il existe un isomorphisme de A sur B. On le note $A \simeq B$.

Remarques VI.1.10.

(1) Un morphisme d'anneaux $f : A \longrightarrow B$ est un morphisme de groupes additifs. En particulier, on a toujours $f(0_A) = 0_B$.
(2) Si $f : A \longrightarrow B$ est un morphisme d'anneaux, alors, pour tout sous-anneau A' de A, $f(A')$ est un sous-anneau de B. En particulier, $\mathrm{Im}(f)$ est un sous-anneau de B.

    En effet, $f(A')$ est un sous-groupe de B par le lemme **??** (1). On a donc $0_B \in f(A')$, et $f(A')$ est stable par addition et par opposé. De plus, puisque A' est un sous-anneau de A, on a $1_A \in A'$, et par conséquent, $1_B = f(1_A) \in f(A')$. Enfin, pour tous $a'_1, a'_2 \in A'$, on a
    $$f(a'_1)f(a'_2) = f(a'_1a'_2) \in f(A'),$$
    car $a'_1a'_2 \in A'$. Ainsi, $f(A')$ est stable par multiplication et $f(A')$ est un sous-anneau de B.

Les morphismes d'anneaux de $\mathbb{Z}$ vers un anneau A sont bien connus, et il n'y en a pas beaucoup !

Lemme VI.1.11. *Soit A un anneau. Alors, il existe un unique morphisme d'anneaux $\Theta_A : \mathbb{Z} \longrightarrow A$. Il est défini par*
$$\Theta_A(m) = m{\cdot}1_A \text{ pour tout } m \in \mathbb{Z}.$$

Démonstration. Si $f : \mathbb{Z} \longrightarrow A$ est un morphisme d'anneaux, on a nécessairement $f(1) = 1_A$, et par suite
$$f(m) = f(1 + \cdots + 1) = f(1) + \cdots + f(1) = m{\cdot}1_A \quad \text{pour tout } m \geq 0.$$
Si $m \leq 0$, on a alors
$$f(m) = f(-(-m)) = -f(-m) = -(-m{\cdot}1_A) = m{\cdot}1_A,$$
par la proposition **??**. Ainsi, on a nécessairement $f = \Theta_A$. Il reste à vérifier que Θ_A est bien un morphisme d'anneaux, ce qui est un exercice longuet, mais facile, lâchement laissé au lecteur. □

Le noyau de cet unique morphisme va maintenant nous permettre de définir un entier très important attaché à un anneau A, à savoir, sa caractéristique.

Définition VI.1.12. Soit A un anneau. Le morphisme d'anneaux
$$\begin{aligned} \Theta_A \colon \mathbb{Z} &\longrightarrow A \\ m &\longmapsto m{\cdot}1_A. \end{aligned}$$

est en particulier un morphisme de groupes abéliens. Son noyau est ainsi un sous-groupe de $\mathbb{Z}$, donc de la forme $c\mathbb{Z}$ d'après la proposition **??**.

La *caractéristique* de A est l'unique entier $c \geq 0$ défini par

$$\mathrm{Ker}(\Theta_A) = c\mathbb{Z}.$$

On le note $\mathrm{car}(A)$. Autrement dit, $\mathrm{car}(A) > 0$ si, et seulement si, l'équation

$$m{\cdot}1_A = 0_A,\ m \in \mathbb{Z}$$

possède au moins une solution non nulle. Dans ce cas, $\mathrm{car}(A)$ est le plus petit entier $c > 0$ tel que $c{\cdot}1_A = 0_A$. De plus, pour tout $m \in \mathbb{Z}$, on a

$$m{\cdot}1_A = 0_A \Longrightarrow \mathrm{car}(A) \mid m.$$

Remarque VI.1.13. La caractéristique de A est non nulle si, et seulement si, 1_A est d'ordre fini dans $(A, +)$. Dans ce cas, on a $\mathrm{car}(A) = o(1_A)$.

Ceci provient simplement des définitions.

Exemple VI.1.14. Les anneaux $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sont de caractéristique 0, tandis que $\mathrm{car}(\mathbb{Z}/n\mathbb{Z}) = n$ pour tout $n \geq 1$.

Remarque VI.1.15. Si B est un sous-anneau de A, alors A et B ont même caractéristique.

Cela provient de la définition, et du fait que $1_B = 1_A$.

On a alors la propriété importante suivante.

Proposition VI.1.16. *Soit A un anneau. Alors, A contient un sous-anneau isomorphe à $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$.*

Démonstration. Le théorème de factorisation pour les groupes appliqué au morphisme Θ_A montre que l'on a un morphisme de groupes injectif

$$\overline{\Theta}_A : \mathbb{Z}/\mathrm{car}(A)\mathbb{Z} \hookrightarrow A.$$

On vérifie aisément que c'est aussi un morphisme d'anneaux. □

Remarque VI.1.17. Si $c = \mathrm{car}(A)$, alors, pour tout $m \in \mathbb{Z}$, on a

$$c \mid m \Longrightarrow m{\cdot}a = 0_A \ \text{ pour tout } a \in A.$$

En effet, soit $m \in \mathbb{Z}$ tel que $c \mid m$.

Supposons tout d'abord que $m \geq 0$. Si $a \in A$, on a

$$m{\cdot}a = a + \cdots + a = a(1_A + \cdots + 1_A) = a(m{\cdot}1_A) = a0_A = 0_A.$$

Si $m \leq 0$, on a alors

$$m{\cdot}a = (-a) + \cdots + (-a) = (-m){\cdot}(-a) = 0_A,$$

d'où l'assertion.

On en déduit aisément que, si $m{\cdot}1_A = 0_A$ a une solution non nulle dans $\mathbb{Z}$, alors $\mathrm{car}(A)$ est le plus petit entier $c > 0$ tel que $c{\cdot}a = 0_A$ pour tout $a \in A$.

Nous terminons ce paragraphe par un lemme calculatoire très utile.

Lemme VI.1.18. *Soit A un anneau commutatif de caractéristique p, où p est un nombre premier. Alors, pour tout $m \geq 0$, on a*

$$(a+b)^{p^m} = a^{p^m} + b^{p^m},\ (ab)^{p^m} = a^{p^m} b^{p^m} \quad \textit{pour tous } a, b \in A.$$

En particulier, l'application

$$\begin{aligned} A &\longrightarrow A \\ a &\longmapsto a^{p^m} \end{aligned}$$

est un morphisme d'anneaux.

Démonstration. Le seul point non évident est la première égalité. Il suffit de traiter le cas $m = 1$, le cas général s'en déduisant par une récurrence évidente. Puisque A est commutatif, pour tous $a, b \in A$, on a

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} \cdot a^k b^{p-k}.$$

Or, il est bien connu que pour tout $k \in [\![1, p-1]\!], p \mid \binom{p}{k}$. On applique alors la remarque précédente pour conclure. □

VI.2. Diviseurs de zéro, éléments inversibles

L'existence d'une loi produit dans un anneau permet de définir quelques familles d'éléments remarquables. On commence par les éléments inversibles.

Définition VI.2.1. Soit A un anneau. On dit que $a \in A$ est *inversible* (ou une *unité*) s'il possède un symétrique pour la loi produit, i.e. s'il existe $b \in A$ tel que $ab = 1_A = ba$.

Dans ce cas, ce symétrique est unique d'après le lemme **??**. Il est noté a^{-1} et est appelé *inverse* de a.

L'ensemble des éléments inversibles de A est noté $A^\times$.

Proposition VI.2.2. *Soit A un anneau. Alors, $A^\times$ est un groupe pour la multiplication. De plus, pour tous $a, b \in A^\times$, on a*

$$(a^{-1})^{-1} = a, \qquad (ab)^{-1} = b^{-1}a^{-1}.$$

Démonstration. Remarquons que $1_A \in A^\times$, puisque $1_A\, 1_A = 1_A$. Ainsi, $A^\times$ est non vide. D'autre part, si a, $b \in A^\times$, alors

$$(ab)b^{-1}a^{-1} = a(bb^{-1})a^{-1} = a\ 1_A a^{-1} = aa^{-1} = 1_A.$$

De même, $b^{-1}a^{-1}(ab) = 1_A$. Par conséquent, ab est bien inversible, d'inverse $b^{-1}a^{-1}$, et la loi de multiplication sur A se restreint donc en une loi interne sur $A^\times$. Puisque A est un anneau, cette loi est associative, de neutre 1_A. Enfin, si $a \in A^\times$, alors $a^{-1} \in A$ est aussi inversible, d'inverse a, puisque

$$a^{-1}a = aa^{-1} = 1_A.$$

Ainsi, si $a \in A^\times$, on a $a^{-1} \in A^\times$ et $(a^{-1})^{-1} = a$. De plus, a^{-1} est clairement un inverse de a dans $A^\times$. Cela achève la démonstration. □

Exemples VI.2.3.

(1) Dans un anneau A non trivial, 0_A n'est jamais inversible. En particulier, si A est non trivial, on a $A^\times \subset A \setminus \{0_A\}$.

D'autre part, 1_A et -1_A sont toujours inversibles dans n'importe quel anneau A, donc $\{\pm 1_A\} \subset A^\times$. Attention néanmoins, ces deux éléments ne sont pas nécessairement distincts (penser par exemple à $A = \mathbb{Z}/2\mathbb{Z}$).

(2) On a $2 \in \mathbb{Q}^\times$, mais $2 \notin \mathbb{Z}^\times$. En effet, puisque $\mathbb{Z} \subset \mathbb{Q}$, si 2 a un inverse dans $\mathbb{Z}$, il a un inverse dans $\mathbb{Q}$, qui est nécessairement $\frac{1}{2}$. En revanche, on a $\frac{1}{2} \notin \mathbb{Z}$. En fait, on a $\mathbb{Z}^\times = \{\pm 1\}$.

(3) On a $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$.

Remarque VI.2.4. Tout morphisme d'anneaux $f : A \longrightarrow B$ induit par restriction un morphisme de groupes $f : A^\times \longrightarrow B^\times$.

En effet, si $a \in A^\times$, on a

$$f(a)f(a^{-1}) = f(aa^{-1}) = f(1_A) = 1_B,$$

et de même, on a

$$f(a^{-1})f(a) = f(a^{-1}a) = f(1_A) = 1_B.$$

Ainsi, $f(a) \in B^\times$ pour tout $a \in A^\times$. Par conséquent, f restreint en une application $f : A^\times \longrightarrow B^\times$. C'est un morphisme de groupes puisque un morphisme d'anneaux respecte les lois de multiplication.

La proposition **??** montre que les règles de calculs dans un anneau quelconque sont similaires aux règles de calculs bien connues sur les entiers ou les réels. Ceci est vrai jusqu'à un certain point. Par exemple, dans $\mathbb{Z}$, le produit de deux facteurs est nul si, et seulement si, un des deux facteurs est nul.

Ceci devient faux dans un anneau quelconque.

Exemples VI.2.5.

(1) Dans $\mathbb{Z}/6\mathbb{Z}$, on a $\overline{2} \neq \overline{0}$ et $\overline{3} \neq \overline{0}$, mais $\overline{2}\cdot\overline{3} = \overline{6} = \overline{0}$.

(2) Dans $\mathrm{M}_2(\mathbb{R})$, on a

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ceci motive la définition suivante.

Définition VI.2.6. Soit A un anneau. On dit que A possède des *diviseurs de zéro* s'il existe $a, b \in A \setminus \{0_A\}$ tels que $ab = 0_A$. Dans ce cas, a est appelé un *diviseur de zéro à gauche* et b est appelé un *diviseur de zéro à droite*.

Si A est commutatif, les notions de diviseur de zéro à gauche et à droite coïncident, et on parlera simplement de diviseur de zéro.

Exemples VI.2.7.

(1) Si $A = \mathbb{C} \times \mathbb{C}$, alors $(1, 0)$ et $(0, 1)$ sont des diviseurs de zéro.

(2) Si $A = \mathbb{Z}/6\mathbb{Z}$, alors $\overline{2}$ et $\overline{3}$ sont des diviseurs de zéro.

(3) Si $A = \mathrm{M}_2(\mathbb{R})$, alors

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

est un diviseur de zéro à gauche et à droite, puisque $M^2 = 0_{\mathrm{M}_2(\mathbb{R})}$.

Les anneaux ne possédant pas de diviseurs de zéro sont évidemment parmi les plus agréables, puisque les choses redeviennent plus naturelles du point de vue calculatoire. Ceux-ci méritent une attention particulière et font l'objet du paragraphe suivant.

VI.3. Anneaux intègres, corps

Définition VI.3.1. Un anneau A est dit *intègre* s'il est non trivial, commutatif et ne possède pas de diviseurs de zéros. Autrement dit, A est intègre s'il est non trivial, commutatif, et pour tous $a, b \in A$, on a

$$ab = 0_A \Longrightarrow a = 0_A \quad \text{ou} \quad b = 0_A.$$

Remarque VI.3.2. Un sous-anneau d'un anneau intègre est un anneau intègre. La démonstration de ce fait est laissée en exercice.

Exemples VI.3.3.

(1) Les anneaux $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont intègres.

(2) Dans l'anneau $\mathbb{Z}/4\mathbb{Z}$, on a $\overline{2} \neq \overline{0}$, puisque 4 ne divise pas 2, mais on a par ailleurs $\overline{2}\ \overline{2} = \overline{4} = \overline{0}$, donc $\overline{2}$ est un diviseur de zéro dans $\mathbb{Z}/4\mathbb{Z}$. Ainsi, $\mathbb{Z}/4\mathbb{Z}$ n'est pas intègre.

(3) Si A est non trivial, l'anneau $A \times A$ n'est pas intègre, puisque

$$(1_A, 0_A)\ (0_A, 1_A) = (0_A, 0_A).$$

(4) Les anneaux $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Z}[\frac{-1+\sqrt{d}}{2}]$ (lorsque $d \equiv 1\ [4]$) sont intègres.

Revenons maintenant aux éléments inversibles d'un anneau. Comme on l'a déjà constaté, tout élément non nul d'un anneau n'est pas nécessairement inversible. Lorsque c'est le cas, on peut donc « diviser » par n'importe quel élément non nul, d'où la définition suivante.

Définition VI.3.4. Un anneau A est un *anneau à division* s'il est non trivial et si tout élément non nul est inversible.

Un *corps* est un anneau à division qui est commutatif. Autrement dit, un corps est un anneau non trivial, commutatif, tel que tout élément non nul est inversible.

Remarque VI.3.5. Le lecteur vérifiera que cette définition est en accord avec celle donné dans le chapitre **??** (cf. définition **??**).

Exemples VI.3.6.

(1) L'anneau $\mathbb{Z}$ n'est pas un corps.

(2) Les anneaux $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps.

(3) Pour tout $d \in \mathbb{Z} \setminus \{0\}$ sans facteurs carrés, les anneaux $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Z}\big[\frac{-1+\sqrt{d}}{2}\big]$ (si $d \equiv 1[4]$) ne sont pas des corps.

(4) Pour tout $d \in \mathbb{Z} \setminus \{0\}$ sans facteurs carrés, l'anneau $\mathbb{Q}[\sqrt{d}]$ est un corps.

Exercice : Soit $\mathbb{H} \subset \mathrm{M}_2(\mathbb{C})$ l'ensemble défini par

$$\mathbb{H} = \{M \in M_2(\mathbb{C}) | M = \begin{pmatrix} z_1 & -z_2 \\ \overline{z}_2 & \overline{z}_1 \end{pmatrix}, z_1, z_2 \in \mathbb{C}\}.$$

(1) Montrer que $\mathbb{H}$ est un sous-anneau non trivial de $\mathrm{M}_2(\mathbb{C})$.

(2) Montrer que $\mathbb{H}$ n'est pas commutatif.

(3) Montrer que $\mathbb{H}$ est un anneau à division (*Indication :* Si $M \in \mathbb{H}$ a un inverse dans $\mathbb{H}$, il a un inverse dans $\mathrm{M}_2(\mathbb{C})$. Quel est-il ? Vérifier alors que cet inverse dans $\mathrm{M}_2(\mathbb{C})$ appartient à $\mathbb{H}$).

PROPOSITION VI.3.7. *Un diviseur de zéro à gauche ou à droite n'est pas inversible. En particulier, un anneau à division ne contient pas de diviseurs de zéro, et tout corps est un anneau intègre.*

Démonstration. Soient $a, b \in A \setminus \{0_A\}$ tels que $ab = 0_A$. Si a était inversible, on aurait

$$b = 1_A\, b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0_A = 0_A,$$

d'où une contradiction, et de même pour b. Le reste est clair. □

Attention ! Un anneau intègre n'est pas nécessairement un corps (penser à l'anneau $\mathbb{Z}$).

Examinons maintenant le cas de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

PROPOSITION VI.3.8. *Soit $n \geq 1$ un entier. Les propriétés suivantes sont équivalentes :*

(1) *$\mathbb{Z}/n\mathbb{Z}$ est intègre*
(2) *$\mathbb{Z}/n\mathbb{Z}$ est un corps*
(3) *n est premier.*

Démonstration. Si $n = 1$, alors n n'est pas premier. De plus, $\mathbb{Z}/1\mathbb{Z} = \{\overline{0}\}$. On a donc $\overline{0} = \overline{1}$, et donc $\mathbb{Z}/n\mathbb{Z}$ est trivial, et n'est pas un corps.

Si p est un nombre premier, alors l'anneau $\mathbb{Z}/p\mathbb{Z}$ est non trivial, puisqu'il a $p \geq 2$ éléments. Soit $\overline{a} \in \mathbb{Z}/p\mathbb{Z}$ non nul. Puisque $\overline{a} \neq \overline{0}$, a n'est pas un multiple de p. Alors, a et p sont premiers entre eux. Par le théorème de Bézout, il existe donc $u, v \in \mathbb{Z}$ tels que

$$ua + vp = 1.$$

Ceci revient à dire que $ua \equiv 1\ [p]$, ou encore $\overline{ua} = \overline{1}$. Mais alors,

$$\overline{a}\cdot\overline{u} = \overline{u}\cdot\overline{a} = \overline{ua} = \overline{1}.$$

Ainsi, $\overline{a}$ possède un symétrique pour la loi produit, à savoir $\overline{u}$. Ceci démontre que $\mathbb{Z}/p\mathbb{Z}$ est un corps, et donc un anneau intègre par la proposition précédente.

Si $n \geq 2$ n'est pas premier, il existe $n_1, n_2 \in [\![1, n-1]\!]$ tels que $n = n_1 n_2$. Remarquons que les classes $\overline{n}_1$ et $\overline{n}_2$ sont non nulles, puisque $n_1, n_2 \in [\![1, n-1]\!]$. En revanche, on a

$$\overline{n}_1\overline{n}_2 = \overline{n_1 n_2} = \overline{n} = \overline{0}.$$

Ainsi, $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre, et ceci achève la démonstration. □

Notation. Si p est premier, le corps $\mathbb{Z}/p\mathbb{Z}$ est noté $\mathbb{F}_p$.

COROLLAIRE VI.3.9. *La caractéristique d'un anneau intègre est soit nulle, soit un nombre premier.*

Démonstration. Soit A un anneau intègre. Par la proposition **??**, A contient un sous-anneau isomorphe à $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$, qui est donc lui-même intègre. Or, $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$ est intègre si, et seulement si, $\mathrm{car}(A) = 0$ (car cet anneau est isomorphe à $\mathbb{Z}$ dans ce cas) ou si $\mathrm{car}(A)$ est un nombre premier, par la proposition **??**. Ceci achève la démonstration. □

Attention ! La réciproque est fausse. En effet, $\mathbb{C}\times\mathbb{C}$ est un anneau de caractéristique 0, mais n'est pas intègre.

On a bien entendu la notion de sous-anneau à division et de sous-corps.

Définition VI.3.10. Soit A un anneau à division (resp. un corps). Un *sous-anneau à division* (resp. un *sous-corps*) est un sous-anneau B de A tel que, pour tout $x \in B \setminus \{0_A\}$, on a $x^{-1} \in B$.

Autrement dit, c'est un sous-ensemble contenant $0_A, 1_A$, et stable par addition, multiplication, passage à l'opposé, et passage à l'inverse.

Dans ce cas, B est un anneau à division (resp. un corps) pour les lois induites par celles de A.

Exemple VI.3.11. L'anneau $\mathbb{Q}$ est un sous-corps de $\mathbb{R}$, qui est lui-même un sous-corps de $\mathbb{C}$.

VI.4. Exemple de l'anneau $\mathbb{Z}/n\mathbb{Z}$

Nous allons maintenant donner des résultats qui élucident la structure de l'anneau $\mathbb{Z}/n\mathbb{Z}$. Les démonstrations seront données en TD.

Théorème VI.4.1. *Soient $m_1, \dots, m_r \geq 0$ des entiers premiers deux à deux. Aors on a un isomorphisme d'anneaux*

$$\mathbb{Z}/m_1 \cdots m_r\mathbb{Z} \simeq \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_r\mathbb{Z}.$$

En particulier, si $n \geq 2$, et si $n = p_1^{n_1} \cdots p_r^{n_r}$ est la décomposition de n en facteurs premiers, on a un isomorphisme d'anneaux

$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{n_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{n_r}\mathbb{Z},$$

et un isomorphisme de groupes

$$(\mathbb{Z}/n\mathbb{Z})^\times \simeq (\mathbb{Z}/p_1^{n_1}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_r^{n_r}\mathbb{Z})^\times.$$

Lemme VI.4.2. *Soit $n \geq 1$. Alors, $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$ est inversible si et seulement a et n sont premiers entre eux.*

Nous allons maintenant appliquer ceci au calcul de la fonction d'Euler.

Définition VI.4.3. **La fonction d'Euler** est la fonction

$$\varphi : \mathbb{N}^* \longrightarrow \mathbb{N}, n \longmapsto \varphi(n) = |(\mathbb{Z}/n\mathbb{Z})^\times|.$$

Proposition VI.4.4. *Soient $n, m \geq 1$ deux entiers premiers entre eux. Alors on a*

$$\varphi(nm) = \varphi(n)\varphi(m).$$

De plus, pour tout nombre premier p et tout entier $r \geq 1$, on a

$$\varphi(p^r) = p^{r-1}(p-1).$$

En particulier, si $n = p_1^{n_1} \cdots p_r^{n_r}$ est la décomposition en facteurs premiers de n, on a

$$\varphi(n) = \prod_{i=1}^{r} p_i^{n_r - 1}(p_i - 1) = n \prod_{i=1}^{r} (1 - \frac{1}{p_i}).$$

Théorème VI.4.5. *On a les isomorphismes de groupes suivants :*

(1) $(\mathbb{Z}/2^r\mathbb{Z})^\times \simeq \mathbb{Z}/2^{r-1}\mathbb{Z}$ *si* $r = 1, 2$.
(2) $(\mathbb{Z}/2^r\mathbb{Z})^\times \simeq \mathbb{Z}/2^{r-2}\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ *si* $r \geq 3$.
(3) $(\mathbb{Z}/p^r\mathbb{Z})^\times \simeq \mathbb{Z}/p^{r-1}(p-1)\mathbb{Z}$ *si* $p > 2$ *est premier et* $r \geq 1$.

En particulier, $(\mathbb{Z}/p\mathbb{Z})^\times$ est un groupe cyclique d'ordre $p - 1$ pour tout premier p.

VI.5. Idéaux

Nous allons maintenant nous attarder un peu sur une notion cruciale en théorie des anneaux, et qui jouera le même rôle que les sous-groupes distingués d'un groupe en théorie des groupes. D'après le lemme **??**, le noyau d'un morphisme de groupes $\varphi : G \longrightarrow G'$ est un sous-groupe distingué de G. D'autre part, tout sous-groupe H distingué dans G est le noyau de la projection canonique $G \longrightarrow G/H$. Ainsi, les sous-groupes distingués de G sont exactement les sous-groupes apparaissant comme noyau d'un morphisme de groupes de source G.

Regardons maintenant les propriétés du noyau d'un morphisme d'anneaux. Outre le fait que le noyau d'un morphisme d'anneaux $f : A \longrightarrow B$ est un sous-groupe de $(A, +)$, il vérifie également la propriété suivante : pour tout $a \in A$, et pour tout $x \in \mathrm{Ker}(f)$, on a $ax \in \mathrm{Ker}(f)$.

Cela motive la notion d'idéal d'un anneau. On verra plus loin que cette notion vient naturellement lorsque l'on cherche à construire les anneaux quotients, tout comme les sous-groupes distingués d'un groupe interviennent de façon naturelle dans la construction des groupes quotients.

Dans la suite, tous les anneaux seront commutatifs.

Définition VI.5.1. Soit A un anneau commutatif.

Un sous-ensemble $\mathfrak{a} \subset A$ est un *idéal* si :

(1) $\mathfrak{a}$ est un sous-groupe additif de $(A, +)$;

(2) pour tout $a \in A$ et tout $x \in \mathfrak{a}$, on a $ax \in \mathfrak{a}$.

Exemples VI.5.2.

(1) Les ensembles $\{0_A\}$ et A sont des idéaux s de A.

(2) Pour tout $x \in A$, l'ensemble $Ax = \{ax \mid a \in A\}$ est un idéal de A.

(3) Tout idéal de $\mathbb{Z}$ est de la forme $n\mathbb{Z}(= \mathbb{Z}n), n \in \mathbb{Z}$ (on peut même choisir $n \geq 0$, et dans ce cas n est unique). En effet, un idéal étant un sous-groupe additif de $\mathbb{Z}$, il est nécessairement de cette forme par la proposition **??**, et on vérifie aisément que c'est bien un idéal de $\mathbb{Z}$.

Le lemme suivant facilite les vérifications pour montrer qu'un sous-ensemble $\mathfrak{a}$ est un idéal.

Lemme VI.5.3. *Soit $\mathfrak{a}$ un sous-ensemble de A. Alors, $\mathfrak{a}$ est un idéal si, et seulement si, :*

(1) *$\mathfrak{a}$ est non vide ;*

(2) *pour tous $x, y \in \mathfrak{a}$ et tout $a \in A$, on a $x + ay \in \mathfrak{a}$.*

Démonstration. Le sens direct est clair. Supposons maintenant que $\mathfrak{a}$ satisfasse aux conditions ci-dessus. Pour tous $x, y \in A$. Alors, on a

$$x + (-1_A)y = x - y \in \mathfrak{a},$$

et donc $\mathfrak{a}$ est un sous-groupe additif de A (puisque $\mathfrak{a}$ est aussi non vide). En particulier, il contient 0_A, et donc pour tout $x \in \mathfrak{a}$ et tout $a \in A$, on a

$$ax = 0_A + ax \in \mathfrak{a}.$$

Ainsi, $\mathfrak{a}$ est un idéal. □

Le lemme suivant est l'analogue du lemme **??**.

LEMME VI.5.4. *Soit $f : A \longrightarrow B$ un morphisme d'anneaux, et soit $\mathfrak{b}$ un idéal de B. Alors, $f^{-1}(\mathfrak{b})$ est un idéal de A. En particulier,* $\mathrm{Ker}(f)$ *est un idéal de A.*

Démonstration. On sait déjà que $f^{-1}(\mathfrak{b})$ est un sous-groupe de A par le lemme **??**. Soit $a \in A$ et soit $x \in f^{-1}(\mathfrak{b})$. Alors, on a $f(x) \in \mathfrak{b}$, et donc

$$f(ax) = f(a)f(x) \in \mathfrak{b},$$

car $\mathfrak{b}$ est un idéal de B. Ainsi, $ax \in f^{-1}(\mathfrak{b})$ et $f^{-1}(\mathfrak{b})$ est donc un idéal de A. Le dernier point est immédiat puisque $\{0_B\}$ est un idéal de B. □

Attention ! Si $\mathfrak{a}$ est un idéal de A, $f(\mathfrak{a})$ n'est pas nécessairement un idéal de B. On laisse le soin au lecteur de trouver un contre-exemple. Cela devient vrai néanmoins si f est surjectif.

LEMME VI.5.5. *L'intersection d'une famille d'idéaux de A est un idéal de A.*

Démonstration. Soit $(\mathfrak{a}_i)_{i\in I}$ une famille d'idéaux de A. On sait déjà que c'est un sous-groupe additif de A par le lemme **??**. Si maintenant $x \in \mathfrak{a}_i$ pour tout $i \in I$ et $a \in A$, on a $ax \in \mathfrak{a}_i$ pour tout $i \in I$, car $\mathfrak{a}_i$ est un idéal, et donc ax appartient à l'intersection des $\mathfrak{a}_i$. □

DÉFINITION VI.5.6. Soit A un anneau, et soit S une partie de A (qui peut être éventuellement vide). *L'idéal de A engendré par S*, noté (S), est l'intersection de tous les idéaux de A contenant S. Autrement dit, (S) est le plus petit idéal de A contenant S (au sens de l'inclusion).

L'idéal engendré jouit des mêmes propriétés que le sous-groupe engendré. Par exemple, $(S) = S$ si, et seulement si, S est un idéal de A.

Si $S = \{s_1, \dots, s_m\}$, on note $(s_1, \dots, s_m)$ plutôt que $(\{s_1, \dots, s_m\})$, et de même pour les autres cas.

De même que dans le cas des sous-groupes, on a une description précise de l'idéal engendré par une partie.

PROPOSITION VI.5.7. *Soit A un anneau, et soit S une partie de A. Alors, on a*

$$(S) = \{a_1 s_1 + \cdots + a_n s_n \mid n \geq 0, s_i \in S, a_i \in A\}.$$

Démonstration. Montrons tout d'abord l'inclusion « $\supset$ ». Puisque $S \subset (S)$ par définition, pour tous $s_1, \dots, s_n \in S$, et tous $a_1, \dots, a_n \in A$, on a $a_1 s_1, \dots, a_n s_n \in (S)$, car (S) est un idéal. Puisque (S) est aussi stable par addition, on en déduit que $a_1 s_1 + \cdots + a_n s_n \in (S)$. Pour montrer l'autre inclusion, on doit vérifier que l'ensemble

$$\{a_1 s_1 + \cdots + a_n s_n \mid n \geq 0, s_i \in S, a_i \in A\}$$

est un idéal de A qui contient S, ce qui est facile en utilisant le lemme **??**. □

REMARQUE VI.5.8. Pour tous $s_1, \dots, s_m \in A$, on a

$$\begin{aligned}(s_1, \dots, s_m) &= \{a_1 s_1 + \cdots + a_m s_m \mid a_i \in A\} \\ &= As_1 + \cdots + As_m.\end{aligned}$$

En particulier, pour tout $s \in A$, on a $(s) = \{as \mid a \in A\} = As$.

Les idéaux pouvant être engendrés par un seul élément méritent un nom spécifique.

DÉFINITION VI.5.9. Un idéal $\mathfrak{a}$d'un anneau A est dit *principal* s'il peut être engendré par un élément, c'est-à-dire s'il existe $a \in A$ tel que $\mathfrak{a} = (a)$.

Un anneau A est *principal* s'il est intègre, et si tout idéal de A est principal.

Exemple VI.5.10. L'anneau $\mathbb{Z}$ est principal d'après l'exemple **??** (2).

Remarque VI.5.11. Nous verrons au fur et à mesure des exemples d'anneaux principaux ou non principaux.

On continue ce paragraphe par un lemme élémentaire, mais fort utile.

Lemme VI.5.12. *Soit A un anneau, et soit $\mathfrak{a}$ un idéal.*

(1) *On a $\mathfrak{a} = A$ si, et seulement si, $\mathfrak{a}$ contient un élément inversible de A.*
(2) *Pour tout $a \in A$, on a $(a) = A$ si, et seulement si, $a \in A^\times$.*
(3) *Si A est intègre, pour tous $a, b \in A$, on a $(a) = (b)$ si, et seulement si, il existe $u \in A^\times$ tel que $b = au$.*

Démonstration. Si $\mathfrak{a} = A$, alors $\mathfrak{a}$ contient 1_A qui est inversible. Si maintenant $\mathfrak{a}$ contient $u \in A^\times$, alors pour tout $a \in A$, on a

$$a = (au^{-1})u \in \mathfrak{a},$$

car $\mathfrak{a}$ est un idéal, et donc $\mathfrak{a} = A$.

Soit $a \in A$. Si a est inversible, alors $(a) = A$ par le point précédent. Inversement, si $(a) = A$, alors $1_A \in A$, et donc il existe $b \in A$ tel que $ba = 1_A$. Ainsi, $a \in A^\times$ (car $ab = ba = 1_A$).

Supposons enfin que A soit intègre. Si $b = au, u \in A^\times$, on a $(b) \subset (a)$. Comme $u \in A^\times$, on a $a = u^{-1}b$ et donc $(a) \subset (b)$. Ainsi, on a $(a) = (b)$. Réciproquement, si $(a) = (b)$, il existe $u, v \in A$ tels que $b = au$ et $a = bv$, et donc $a = uva$. Si $a = 0_A$, alors $b = 0_A$ et donc $b = a = a1_A$. Si $a \neq 0_A$, puisque A est intègre, on obtient $1_A = uv$ et donc $u \in A^\times$. □

On a alors le corollaire suivant.

Corollaire VI.5.13. *Soit A un anneau commutatif. Alors, A est un corps si, et seulement si, A possède exactement deux idéaux, à savoir (0_A) et A.*

Démonstration. Supposons que A soit un corps. Alors, $0_A \neq 1_A$, et par suite (0_A) et A sont deux idéaux distincts. Si $\mathfrak{a}$ est un idéal non nul, il contient un élément non nul, donc inversible, et est donc égal à A par le point (1) du lemme précédent.

Inversement, supposons que (0_A) et A soient distincts, et soient les seuls idéaux. Alors, on a $0_A \neq 1_A$. Soit $a \in A \setminus \{0\}$. Alors, $(a)_g$ est un idéal non nul, donc égal à A. Il existe donc $b \in A$ tel que $ba = 1_A$. Remarquons que b est non nul, car $0_A \neq 1_A$. Par le même raisonnement, il existe $c \in A$ tel que $cb = 1_A$. On a alors

$$a = 1_A\ a = (cb)a = c(ba) = c\ 1_A = c.$$

Par conséquent, on a aussi $ab = 1_A$, et a est donc inversible. Ainsi, A est un corps. Ceci achève la démonstration. □

Dorénavant, on écrira 0 et 1, plutôt que 0_A et 1_A, afin de ne pas alourdir les notations.

Le prochain paragraphe est consacré à l'étude d'un exemple, à savoir l'anneau des polynômes.

VI.6. Anneaux de polynômes

Dans ce paragraphe, on définit l'anneau des polynômes à coefficients dans un anneau A, et on étudie quelques unes de ses propriétés.

Soit A un anneau, et soit $\mathcal{S}$ le groupe additif des suites $(a_n)_{n\geq 0}$ d'éléments de A à support fini (i.e. dont tous les termes sont nuls à partir d'un certain rang).

Pour tout $n \geq 0$, on note $e_n \in \mathcal{S}$ la suite dont tous les termes sont nuls, sauf celui d'indice n qui vaut 1.

Si $s = (a_n)_{n\geq 0}, s' = (b_n)_{n\geq 0} \in \mathcal{S}$, on pose

$$ss' = \Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)_{n\geq 0} = \Big(\sum_{\substack{i,j\geq 0\\ i+j=n}} a_i b_j\Big)_{n\geq 0}.$$

Proposition VI.6.1. *L'ensemble $\mathcal{S}$ muni de l'addition et de la loi de multiplication précédente est un anneau, de neutre multiplicatif e_0.*

Démonstration. Il reste en fait à montrer les propriétés relatives à la multiplication, puisque l'on sait déjà que $(\mathcal{S}, +)$ est un groupe additif. Commençons par montrer que si $s = (a_n)_{n\geq 0}, s' = (b_n)_{n\geq 0} \in \mathcal{S}$, alors $ss' \in \mathcal{S}$.

Par hypothèse, il existe $N, N' \geq 0$ tels que

$$a_n = 0 \text{ pour tout } n \geq N \quad \text{et} \quad b_n = 0 \text{ pour tout } n \geq N'.$$

Mais alors pour tout $n \geq N + N'$, on a $\displaystyle\sum_{k=0}^{n} a_k b_{n-k} = 0$, car soit $k \geq N$, et dans ce cas $a_k = 0$, soit $k < N$, mais dans ce cas

$$n - k \geq N + N' - k > N',$$

et $b_{n-k} = 0$. Ainsi, $A[X]$ est aussi stable par produit, d'où le résultat souhaité.

Continuons par l'associativité. Si $s = (a_n)_{n\geq 0}, s' = (b_n)_{n\geq 0}, s'' = (c_n)_{n\geq 0} \in \mathcal{S}$, on a par définition

$$(ss')s'' = \Big(\sum_{\ell=0}^{n}(\sum_{k=0}^{\ell} a_k b_{\ell-k})c_{n-\ell}\Big)_{n\geq 0} = \Big(\sum_{i_1+i_2+i_3=n} (a_{i_1} b_{i_2})c_{i_3}\Big)_{n\geq 0}.$$

De même, on a

$$s(s's'') = \Big(\sum_{i_1+i_2+i_3=n} a_{i_1}(b_{i_2} c_{i_3})\Big)_{n\geq 0}.$$

Comme la multiplication de A est associative, il en est donc de même de celle de $\mathcal{S}$.

La distributivité et la commutativité proviennent aisément de la définition. Enfin, on vérifie que e_0 est un élément neutre pour la multiplication. □

On introduit maintenant de nouvelles notations. Si $a \in A$ et $(a_n) \in \mathcal{S}$, on pose

$$a(a_n)_{n\geq 0} = (aa_n)_{n\geq 0}.$$

C'est tout simplement le produit de $(a, 0, 0, \ldots)$ par $(a_n)_{n\geq 0}$. Un élément de $\mathcal{S}$ s'écrit donc

$$\sum_{n\geq 0} a_n e_n.$$

On vérifie facilement en utilisant la définition que l'on a

$$e_1^n = e_n \text{ pour tout } n \geq 0.$$

Posons $X = e_1$ et notons $A[X]$ l'anneau $\mathcal{S}$. D'après ce qui précède un élément $P \in A[X]$ s'écrit de manière unique

$$S = \sum_{n \geq 0} a_n X^n,$$

où les coefficients a_n sont tous nuls sauf un nombre fini.

De plus, les opérations sur $A[X]$ sont définies par

$$\sum_{n \geq 0} a_n X^n + \sum_{n \geq 0} b_n X^n = \sum_{n \geq 0} (a_n + b_n) X^n$$

et

$$\Big(\sum_{n \geq 0} a_n X^n\Big)\Big(\sum_{n \geq 0} b_n X^n\Big) = \sum_{n \geq 0} \Big(\sum_{k=0}^{n} a_k b_{n-k}\Big) X^n.$$

Par définition, on a $\sum_{n \geq 0} a_n X^n = 0$ si, et seulement si, $a_n = 0$ pour tout entier $n \geq 0$.

Le neutre multiplicatif est donné par X^0, que l'on notera 1, et le neutre additif est donné par $\sum_{n \geq 0} 0\ X^n$, que l'on notera 0.

Définition VI.6.2. Soit A un anneau. L'anneau $A[X]$ est appelé anneau des *polynômes* à coefficients dans A .

Remarque VI.6.3. On montre aisément que $A[X]$ est commutatif si, et seulement si, A l'est.

On vérifie que l'application

$$\begin{aligned} A &\longrightarrow A[X] \\ a &\longmapsto aX^0 \end{aligned}$$

est un morphisme d'anneaux injectif, qui identifie l'anneau A à un sous-anneau de $A[X]$, et même à un sous-anneau de $A[X]$. On identifiera A avec son image dans $A[X]$, et on notera a plutôt que $a\ 1$ ou aX^0.

Les éléments de A, vus comme polynômes, seront appelés des polynômes *constants*.

Remarque VI.6.4. Pour tout $n \geq 0$, on définit l'anneau $A[X_1, \ldots, X_n]$ par récurrence en posant $A[X_1, \ldots, X_n] = A$ pour $n = 0$, et

$$A[X_1, \ldots, X_n] = A[X_1, \ldots, X_{n-1}][X_n]$$

pour tout $n \geq 1$.

Nous allons maintenant illustrer les notions introduites dans ce chapitre dans le cas de l'anneau $A[X]$. On commence par une définition.

Définition VI.6.5. Soit A un anneau, et soit $f \in A[X] \setminus \{0\}$. Puisque $f \neq 0$, on peut écrire

$$f = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_0,$$

avec $a_i \in A$, et $a_d \in A \setminus \{0\}$. L'entier $d \geq 0$ est appelé le *degré* de f, et est noté $\deg(f)$. Par convention, on pose $\deg(0) = -\infty$. On a donc $\deg(f) \geq 0$ si, et seulement si, f est non nul.

Le coefficient a_d est appelé le *coefficient dominant* de f. On dit qu'un polynôme est *unitaire* si son coefficient dominant est égal à 1.

Le lemme suivant est souvent utile dans les calculs.

LEMME VI.6.6. *Soit A un anneau commutatif non trivial, et soit $g \in A[X]$ un polynôme non nul. On suppose que le coefficient dominant de g n'est pas un diviseur de zéro dans A. Alors, pour tout $f \in A[X]$, on a*

$$\deg(fg) = \deg(f) + \deg(g).$$

En particulier, l'application

$$\begin{array}{rcl} r_g \colon A[X] & \longrightarrow & A[X] \\ f & \longmapsto & fg \end{array}$$

est injective.

Démonstration. Si f est nul, l'égalité désirée est triviale. Supposons maintenant que f soit non nul, et écrivons

$$f = a_n X^n + \cdots + a_1 X + a_0,\ g = b_m X^m + \cdots + b_1 X + b_0,\ a_n, b_m \in A \setminus \{0\}.$$

En utilisant la définition, on voit que l'on a

$$fg = a_n b_m X^{n+m} + \text{ des termes en } X^i, i < n + m.$$

Comme b_m n'est pas un diviseur de zéro et $a_n \neq 0$, on a $a_n b_m \neq 0$, ce qui démontre l'égalité

$$\deg(fg) = n + m = \deg(f) + \deg(g).$$

Cela démontre aussi que si $f \neq 0$, on a $fg \neq 0$.

Remarquons maintenant que r_g est un morphisme de groupes additifs. Or, le point précédent montre que $\mathrm{Ker}(r_g) = \{0\}$. Par conséquent, r_g est injective. □

Contrairement au cas de l'anneau des polynômes à coefficients dans un corps, l'anneau $A[X]$ n'est pas nécessairement intègre. Le lemme suivant donne des conditions nécessaires et suffisantes pour cela soit le cas.

LEMME VI.6.7. *Soit A un anneau commutatif non trivial. Alors, les propriétés suivantes sont équivalentes :*

(1) *l'anneau $A[X]$ est intègre ;*

(2) *l'anneau A est intègre ;*

(3) *pour tous $f, g \in A[X]$, on a $\deg(fg) = \deg(f) + \deg(g)$.*

Démonstration.

(1) $\Longrightarrow$ (2). C'est immédiat, car un sous-anneau d'un anneau intègre est intègre.

(2) $\Longrightarrow$ (3). Si g est nul, l'égalité est triviale. Si g est non nul, son coefficient dominant est non nul, et n'est donc pas un diviseur de zéro puisque A est intègre. Le résultat découle alors du lemme **??**.

(3) $\Longrightarrow$ (1). Puisque l'anneau A est commutatif et non trivial, il en est de même de l'anneau $A[X]$.

Soient f et g deux polynômes non nuls. On a donc $\deg(f) \geq 0$ et $\deg(g) \geq 0$. Mais alors, on a

$$\deg(fg) = \deg(f) + \deg(g) \geq 0,$$

ce qui entraîne que fg est non nul, d'où l'intégrité de $A[X]$. □

On continue ce paragraphe en expliquant comment diviser deux polynômes.

THÉORÈME VI.6.8 (Division de polynômes généralisée). *Soit A un anneau commutatif. Soient $f, g \in A[X]$, avec $g \neq 0$. On suppose que le coefficient dominant de g est un élément inversible de A. Alors, il existe deux polynômes $Q, R \in A[X]$, avec $\deg(R) < \deg(g)$ tels que $f = Qg + R$. De plus, Q et R sont uniques.*

Démonstration. Si $\deg(g) > \deg(f)$, on pose $Q = 0$ et $R = f$. On suppose maintenant que $\deg(g) \leq \deg(f)$. On va démontrer l'existence de Q et R par récurrence sur $\deg(f)$. Soit (H_n) la propriété :

(H_n) : Pour tous $f, g \in A[X], g \neq 0$ tels que $\deg(g) \leq \deg(f) \leq n$, et tels que le coefficient dominant de g soit inversible, il existe $Q, R \in A[X]$, tels que $\deg(R) < \deg(g)$ et $f = Qg + R$.

Si f est le polynôme nul, c'est clair, car on prend $Q = R = 0$. Si $\deg(f) = 0$, alors on a $\deg(g) = 0$. On a donc $f = a, g = b, a, b \in A, b \in A^\times$. Dans ce cas, on pose $Q = ab^{-1}$ et $R = 0$. Ainsi, (H_0) est vraie.

Supposons (H_n) vraie pour un certain entier $n \geq 0$, et montrons (H_{n+1}).

Si $\deg(f) \leq n$, alors Q et R existent, car (H_n) est vraie, donc on peut supposer que $\deg(f) = n + 1$. Écrivons

$$f = a_{n+1}X^{n+1} + \cdots + a_1X + a_0,\ g = b_mX^m + \cdots + b_1X + b_0.$$

Par hypothèse, on a $n + 1 \geq m$.

Soit $Q_1 = a_{n+1}b_m^{-1}X^{n+1-m}$ et $R_1 = f - a_{n+1}b_m^{-1}X^{n+1-m}g$. Remarquons que b_m^{-1} a un sens dans A, puisque $b_m \in A^\times$. Puisque $f, g \in A[X]$, on a aussi $Q_1, R_1 \in A[X]$. Par construction, le coefficient en X^{n+1} de R_1 est 0, donc $\deg(R_1) \leq n$.

Si $\deg(R_1) < \deg(g)$, on a fini, car on peut écrire

$$f = a_{n+1}b_m^{-1}X^{n+1-m}g + R_1,$$

et donc poser $R = R_1$ et $Q = a_{n+1}b_m^{-1}X^{n+1-m}$.

Supposons que $\deg(R_1) \geq \deg(g)$. Comme $\deg(g) \leq \deg(R_1) \leq n$, on peut appliquer (H_n), et donc il existe $Q_2, R_2 \in A[X]$ tels que $R_1 = Q_2g + R_2$, avec $\deg(R_2) < \deg(g)$.

Ainsi, on a $f = (a_{n+1}b_m^{-1}X^{n+1-m} + Q_2)g + R_2$, et $\deg(R_2) < \deg(g)$. Maintenant, il suffit de poser $Q = a_{n+1}b_m^{-1}X^{n+1-m} + Q_2$ et $R = R_2$.

Par conséquent, (H_{n+1}) est vraie, ce qui achève la récurrence.

Montrons maintenant l'unicité de Q et R. Supposons que l'on ait

$$f = Q_1g + R_1 = Q_2g + R_2,$$

avec $Q_i, R_i \in A[X]$ et $\deg(R_i) < \deg(g)$. Alors, on a $(Q_1 - Q_2)g = R_2 - R_1$.

Puisque le coefficient dominant de g est inversible, ce n'est pas un diviseur de zéro. Par le lemme **??**, on a donc

$$\deg\big((Q_1 - Q_2)g\big) = \deg(Q_1 - Q_2) + \deg(g).$$

Si $Q_1 - Q_2 \neq 0$, on a donc

$$\deg(R_2 - R_1) = \deg(Q_1 - Q_2) + \deg(g) \geq \deg(g).$$

Néanmoins, il est facile de constater que $\deg(R_2 - R_1) < \deg(g)$, d'où une contradiction. Ainsi, $Q_1 - Q_2 = 0$, puis $R_2 - R_1 = 0$, d'où l'unicité voulue. □

DÉFINITION VI.6.9. Les polynômes Q et R s'appellent respectivement le *quotient* et le *reste* de la division de f par g.

En pratique, la meilleure façon de calculer Q et R est de procéder comme dans le cas où $A = \mathbb{R}, \mathbb{C}$ (en posant la division).

Attention ! Ce résultat n'est pas vrai si le coefficient dominant de g n'est pas inversible. Par exemple, $X, 2X \in \mathbb{Z}[X]$, mais le quotient de X par $2X$ est $\frac{1}{2}$, qui n'est pas dans $\mathbb{Z}[X]$.

Remarque VI.6.10. Si K est un corps, on peut donc diviser n'importe quel polynôme par un polynôme non nul.

Comme dans le cas de l'anneau $\mathbb{Z}$, l'existence d'une division euclidienne dans $K[X]$ a des conséquences intéressantes sur les idéaux.

Lemme VI.6.11. *Soit K un corps. Alors, $K[X]$ est un anneau principal.*

Démonstration. Puisqu'un corps K est un anneau intègre, il en est de même de $K[X]$ par le lemme **??**. Soit $\mathfrak{a}$ un idéal de $K[X]$. Montrons qu'il est principal. Si $\mathfrak{a} = (0)$, il n'y a rien à faire. Supposons donc que $\mathfrak{a} \neq (0)$, et soit f_0 un élément de $\mathfrak{a} \setminus \{0\}$ de degré minimal. Montrons que $\mathfrak{a} = (f_0)$. Puisque $f_0 \in \mathfrak{a}$, on a $(f_0) \subset \mathfrak{a}$. Si maintenant $f \in \mathfrak{a}$, il existe $Q, R \in K[X]$ tels que $f = Qf_0 + R$ avec $\deg(R) < \deg(f_0)$ (cf. remarque **??**). Puisque $f, f_0 \in \mathfrak{a}$, on a $R = f - Qf_0 \in \mathfrak{a}$, car $\mathfrak{a}$ est un idéal. Si R était non nul, cela contredirait la minimalité du degré de f_0. On a donc $R = 0$, soit encore $f = Qf_0 \in (f_0)$. On a donc finalement l'inclusion manquante, ce qui achève la démonstration. □

Remarque VI.6.12. On verra dans un chapitre ultérieur que la description des idéaux de $\mathbb{Z}$ et $K[X]$ s'inscrit dans un contexte plus général, à savoir celui des anneaux euclidiens.

On s'intéresse maintenant aux morphismes d'anneaux de $A[X]$ vers un anneau quelconque B.

Proposition VI.6.13. *Soit $u : A \to B$ un morphisme d'anneaux commutatifs, et soit $b \in B$. Alors, il existe un unique morphisme d'anneaux*

$$\psi_{u,b} : A[X] \to B$$

tel que

$$\psi_{u,b}(a) = u(a) \text{ pour tout } a \in A \quad \text{et} \quad \psi_{u,b}(X) = b.$$

Il est défini par

$$\psi_{u,b}\Big(\sum_{n\geq 0} a_n X^n\Big) = \sum_{n\geq 0} u(a_n) b^n.$$

De plus, pour tout morphisme d'anneaux $\varphi : A[X] \longrightarrow B$, il existe un unique morphisme d'anneaux $u : A \longrightarrow B$ et un unique élément $b \in B$ tel que $\varphi = \psi_{u,b}$.

Démonstration. L'application $\psi_{u,b}$ satisfait clairement les conditions requises. Il reste à vérifier que c'est bien un morphisme d'anneaux. Pour commencer, on a $\psi_{u,b}(1) = u(1) = 1$. De plus,

$$\psi_{u,b}\Big(\sum_{n\geq 0} a_n X^n + \sum_{n\geq 0} b_n X^n\Big) = \psi_{u,b}\Big(\sum_{n\geq 0} (a_n + b_n) X^n\Big) = \sum_{n\geq 0} u(a_n + b_n) b^n.$$

Puisque u est un morphisme d'anneaux, on obtient

$$\begin{aligned}\psi_{u,b}\Big(\sum_{n\geq 0} a_nX^n + \sum_{n\geq 0} b_nX^n\Big) &= \sum_{n\geq 0}(u(a_n)+u(b_n))b^n \\ &= \sum_{n\geq 0} u(a_n)b^n + \sum_{n\geq 0} u(b_n)b^n,\end{aligned}$$

et donc

$$\psi_{u,b}\Big(\sum_{n\geq 0} a_nX^n + \sum_{n\geq 0} b_nX^n\Big) = \psi_{u,b}\Big(\sum_{n\geq 0} a_nX^n\Big) + \psi_{u,b}\Big(\sum_{n\geq 0} b_nX^n\Big).$$

Finalement, on a

$$\begin{aligned}\psi_{u,b}\Big((\sum_{n\geq 0} a_nX^n)(\sum_{n\geq 0} b_nX^n)\Big) &= \psi_{u,b}\Big(\sum_{n\geq 0}(\sum_{j+k=n} a_jb_k)X^n\Big) \\ &= \sum_{n\geq 0} u(\sum_{j+k=n} a_jb_k)b^n.\end{aligned}$$

Puisque u est un morphisme d'anneaux, on a

$$\begin{aligned}\psi_{u,b}\Big((\sum_{n\geq 0} a_nX^n)(\sum_{n\geq 0} b_nX^n)\Big) &= \sum_{n\geq 0}\sum_{j+k=n} u(a_j)u(b_k)b^n \\ &= \sum_{n\geq 0}\sum_{j+k=n} u(a_j)u(b_k)b^jb^k \\ &= \sum_{n\geq 0}\sum_{j+k=n} (u(a_j)b^j)(u(b_k)b^k) \\ &= (\sum_{n\geq 0} u(a_n)b^n)(\sum_{n\geq 0} u(b_n)b^n),\end{aligned}$$

et ainsi

$$\psi_{u,b}\Big((\sum_{n\geq 0} a_nX^n)(\sum_{n\geq 0} b_nX^n)\Big) = \psi_{u,b}\Big(\sum_{n\geq 0} a_nX^n\Big)\psi_{u,b}\Big(\sum_{n\geq 0} b_nX^n\Big).$$

Soit maintenant un autre morphisme d'anneaux $\phi : A[X] \longrightarrow B$ satisfaisant

$$\phi(a) = u(a) \ \text{ pour tout } a \in A, \text{ et } \ \phi(X) = b.$$

On a alors (puisque ϕ est un morphisme d'anneaux)

$$\phi\Big(\sum_{n\geq 0} a_nX^n\Big) = \sum_{n\geq 0}\phi(a_n)(\phi(X))^n = \sum_{n\geq 0} u(a_n)b^n = \psi_{u,b}\Big(\sum_{n\geq 0} a_nX^n\Big),$$

et donc $\phi = \psi_{u,b}$.

Enfin, soit $\varphi : A[X] \longrightarrow B$ un morphisme d'anneaux. Si $\varphi = \psi_{u,b}$, par définition de $\psi_{u,b}$, on a nécessairement $u = \varphi_{|A}$ et $b = \varphi(X)$. Ainsi, si u et b existent, ils sont uniques.

Posons $u = \varphi_{|A}$ et $b = \varphi(X)$, et montrons que $\varphi = \psi_{u,b}$. Puisque φ est un morphisme d'anneaux, il en est de même de $u : A \longrightarrow B$. On vient de voir que l'on a

$$\varphi\Big(\sum_{n\geq 0} a_nX^n\Big) = \sum_{n\geq 0}\varphi(a_n)(\varphi(X))^n.$$

Par définition de u et b, cela s'écrit

$$\varphi\Big(\sum_{n\geq 0} a_nX^n\Big) = \sum_{n\geq 0} u(a_n)b^n = \psi_{u,b}\Big(\sum_{n\geq 0} a_nX^n\Big).$$

Ainsi, $\varphi = \psi_{u,b}$. □

Exemple VI.6.14. Supposons que A soit un sous-anneau de B, et soit $b \in B$. Si l'on applique ce qui précède à l'inclusion $A \subset B$, on obtient un morphisme d'anneaux, appelé *évaluation en* b. Dans ce cas, l'image de $f \in A[X]$ est notée $f(b)$. Si $f = \sum_{n \geq 0} a_n X^n \in A[X]$ et $b \in B$, on a

$$f(b) = \sum_{n \geq 0} a_n b^n.$$

En d'autres termes, pour un $b \in B$ donné, l'application

$$\begin{aligned} A[X] &\longrightarrow B \\ f = \sum_{n \geq 0} a_n X^n &\longmapsto f(b) = \sum_{n \geq 0} a_n b^n \end{aligned}$$

est un morphisme d'anneaux.

Remarque VI.6.15. Si $P \in A[X]$, on peut en particulier définir une fonction polynomiale

$$\begin{aligned} \tilde{P} \colon A &\longrightarrow A \\ a &\longmapsto P(a), \end{aligned}$$

mais cette fonction ne caractérise pas nécessairement le polynôme P. Plus précisément, l'application

$$\begin{aligned} A[X] &\longrightarrow \mathscr{F}(A, A) \\ P &\longmapsto \tilde{P} \end{aligned}$$

n'est pas nécessairement injective. Par exemple, si $A = \mathbb{F}_2$, la fonction polynomiale associée au polynôme $P = X^2 - X$ est identiquement nulle.

L'exemple **??** nous conduit tout naturellement à nous intéresser au noyau du morphisme de spécialisation. Cela conduit à la notion de racine d'un polynôme, notion extrêmement importante qui sera étudiée tout particulièrement dans la partie sur la théorie des corps.

Définition VI.6.16. Soit A un anneau, et soit $f \in A[X]$. On dit que $a \in A$ est une *racine* (ou un *zéro*) de f si $f(a) = 0$.

Lemme VI.6.17. *Soit A un anneau commutatif. Soit $f \in A[X]$ un polynôme, et soit $a \in A$. Alors, a est une racine de f si, et seulement si, il existe un polynôme $g \in A[X]$ tel que $f = (X - a)g$.*

Démonstration. D'après la proposition **??**, il existe $Q, R \in A[X]$ tels que

$$f = (X - a)Q + R,$$

où $\deg(R) < 1$. Autrement dit, R est un polynôme constant, et on a donc

$$f = (X - a)Q + r, r \in A.$$

Puisque A est commutatif, l'évaluation en a

$$\begin{aligned} A[X] &\longrightarrow A \\ P &\longmapsto P(a) \end{aligned}$$

est un morphisme d'anneaux. On en déduit que $f(a) = (a - a)Q(a) + r = r$. Par conséquent, $f(a) = 0$ si, et seulement si, $r = 0$, ce qui entraîne le résultat voulu. $\square$

Il est bien connu que si K est un corps, et si $f \in K[X]$ est un polynôme non nul, alors f possède au plus $\deg(f)$ racines dans K. Ceci se généralise aisément à un anneau intègre (voir le lemme ci-dessous), mais cela devient faux en général, puisque $\overline{2}X \in \mathbb{Z}/4\mathbb{Z}[X]$ a 2 racines.

En fait, on a le résultat suivant.

Lemme VI.6.18. *Soit A un anneau commutatif non trivial. Alors, tout polynôme non nul $f \in A[X]$ possède au plus* $\deg(f)$ *racines dans A si, et seulement si, A est intègre.*

Démonstration. Supposons que A soit intègre. Nous allons démontrer le résultat par récurrence sur le degré d de f. Si $d = 0$, f est un polynôme constant non nul, et n'a aucune racine dans A. Supposons avoir démontré que tout polynôme non nul de degré $d \geq 0$ possède au plus d racines dans A, et soit f un polynôme non nul de degré $d + 1$. Si f n'a pas de racines dans A, on a fini. Supposons donc que f possède au moins une racine $a \in A$. Par le lemme **??**, il existe $g \in A[X]$ tel que $f = (X - a)g$. Puisque A est intègre, on a

$$\deg((X - a)g) = \deg(X - a) + \deg(g) = 1 + \deg(g).$$

Ainsi, $\deg(g) = d$ et par hypothèse de récurrence g a au plus d racines. Pour conclure, il suffit de constater qu'une racine de f est soit égale à a, soit une racine de g. En effet, si $b \in A$ vérifie $f(b) = 0$, alors l'évaluation en b (qui est un morphisme d'anneaux, car A est commutatif) nous donne

$$0 = (b - a)g(b).$$

Par intégrité de A, on en déduit que $b = a$ ou $g(b) = 0$, ce qu'il fallait vérifier. On en déduit donc que f a au plus $d + 1$ racines, ce qui achève la récurrence.

Supposons maintenant que A ne soit pas intègre. Puisque A est commutatif et non trivial, cela implique l'existence de deux éléments $a, b \in A$ non nuls tels que $ab = 0$. Mais alors, le polynôme aX possède au moins deux racines distinctes (qui sont 0 et b). □

Remarque VI.6.19. La commutativité de A est cruciale. En effet, si A n'est pas commutatif, les résultats précédents deviennent faux, même si A n'a pas de diviseurs de zéro. Par exemple, soit $\mathbb{H} \subset \mathrm{M}_2(\mathbb{C})$ l'ensemble défini par

$$\mathbb{H} = \left\{ M \in \mathrm{M}_2(\mathbb{C}) | M = \begin{pmatrix} z_1 & -z_2 \\ \overline{z}_2 & \overline{z}_1 \end{pmatrix}, z_1, z_2 \in \mathbb{C} \right\}.$$

On vérifie sans mal que $\mathbb{H}$ est un anneau à division. En particulier, $\mathbb{H}$ est sans diviseurs de zéros.

Posons

$$e = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, f = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Alors, le polynôme $X^2 + I_2 \in \mathbb{H}[X]$ possède au moins 4 racines dans $\mathbb{H}$, qui sont $e, -e, f, -f$ (en fait, il en admet même une infinité). Ceci s'explique par le fait que la non-commutativité de $\mathbb{H}$ implique que l'évaluation d'un polynôme en un élément de $\mathbb{H}$ n'est plus un morphisme d'anneaux (et donc la démonstration précédente n'est plus valable).

Par exemple, on a $X^2 + I_2 = (X - e)(X + e)$, $f^2 + I_2 = 0$, mais on vérifie sans peine que $(f - e)(f + e) \neq 0$, car $fe \neq ef$.

En particulier, la relation entre factorisation et racines n'a plus lieu. Par exemple, on a $X^2 + I_2 \neq (X - e)(X - f)$.

Pour finir, nous allons nous intéresser aux éléments inversibles de $A[X]$ lorsque A est intègre.

Lemme VI.6.20. *Soit A un anneau intègre. Alors, on a $A[X]^\times = A^\times$.*

Démonstration. Si $a \in A^\times$, alors $a \in A[X]^\times$, puisque les relations

$$aa^{-1} = a^{-1}a = 1 \in A$$

sont également vraies dans $A[X]$. Réciproquement, soit $f \in A[X]^\times$. Alors, il existe $g \in A[X]$ tel gue $fg = gf = 1$. En particulier, f et g sont non nuls. Comme A est intègre, on a

$$\deg(fg) = \deg(f) + \deg(g) = \deg(1) = 0$$

par le lemme **??**. Mais le degré d'un polynôme non nul étant un entier positif ou nul, on en déduit

$$\deg(f) = \deg(g) = 0,$$

donc $f = a$ et $g = b$ pour certains $a, b \in A$. Mais alors, on a $ab = ba = 1$, et donc $f = a \in A^\times$. □

Ce résultat devient faux lorsque A est quelconque.

Exemple VI.6.21. On a $\overline{2}X \neq 0 \in \mathbb{Z}/4\mathbb{Z}[X]$, mais $(\overline{2}X)(\overline{2}X) = 0$. De plus, on vérifie que $(\overline{1} + \overline{2}X)^2 = \overline{1}$ et donc $\overline{1} + \overline{2}X$ est inversible dans $\mathbb{Z}/4\mathbb{Z}[X]$.

Chapitre VII

Arithmétique dans les anneaux principaux

VII.1. Divisibilité, pgcd, ppcm

Définition VII.1.1. Soit A un anneau commutatif, et soient $a, b \in A$. On dit que a *divise* b s'il existe $x \in A$ tel que $b = ax$, ou de manière équivalente, si l'on a $(b) \subset (a)$. On le note $a \mid b$.

On dit aussi que a est un *diviseur* de b, et que b est un *multiple* de a.

Remarques VII.1.2.

(1) Cette définition soulève une petite difficulté lorsque $b = 0$. En effet, tout élément $a \in A$ est un diviseur de 0 au sens précédent (puisque l'on peut toujours écrire $0 = a\ 0$), mais pas nécessairement un diviseur de zéro au sens de la définition **??**.

(2) La notion de divisibilité est relative à un anneau ambiant fixé (cf. les points (2) et (3) de la série d'exemples qui suit). Il est donc important que l'élément x de la définition soit un élément de l'anneau A dans lequel on travaille.

Exemples VII.1.3.

(1) Dans un anneau non trivial, 0 ne divise aucun élément non nul.

(2) Dans $\mathbb{Z}$, l'entier 2 ne divise pas 3. Sinon, on aurait $3 = 2m$ pour $m \in \mathbb{Z}$, ce qui est faux, puisque 3 est impair.

(3) On a $2 \mid 3$ dans $\mathbb{Q}$ puisque $3 = 2 \cdot \dfrac{3}{2}$.

(4) Dans $\mathbb{Z}[i]$, 2 ne divise pas $1+i$. Sinon, on aurait $1+i = 2z, z \in \mathbb{Z}[i]$. En écrivant $z = a + bi, a, b, \in \mathbb{Z}$, on obtient $1 + i = 2a + 2bi$, et en particulier $1 = 2a$ dans $\mathbb{Z}$, ce qui est impossible.

(5) On a $1 + i \mid 2$ dans $\mathbb{Z}[i]$ puisque $2 = (1 + i)(1 - i)$, et $1 - i \in \mathbb{Z}[i]$.

Définition VII.1.4. On dit que deux éléments $a, b \in A$ sont *associés* si $a \mid b$ et $b \mid a$, autrement dit si $(a) = (b)$. Si l'anneau A est intègre, cela revient à dire que $b = au, u \in A^\times$, d'après le lemme **??** (3).

La relation « être associés » est une relation d'équivalence.

Maintenant que l'on a la notion de divisibilité appropriée, on peut introduire la notion d'éléments premiers entre eux.

Définition VII.1.5. On dit que deux éléments a et b de A sont *premiers entre eux* si pour tout $d \in A$, on a

$$d \mid a \ \text{ et } \ d \mid b \Longrightarrow d \in A^\times.$$

Autrement dit, deux éléments sont premiers entre eux si leurs seuls diviseurs communs sont les éléments inversibles.

On dit que deux éléments a et b de A sont *étrangers* s'il existe $u, v \in A$ tels que $ua + vb = 1$.

Exemple VII.1.6. Deux éléments $a, b \in A$ étrangers sont premiers entre eux.

En effet, supposons que a et b soient étrangers. Il existe donc $u, v \in A$ tels que $ua + bv = 1$. Soit $d \in A$ tels que $d \mid a$ et $d \mid b$. Alors, $d \mid (ua + vb)$, c'est-à-dire $d \mid 1$. La définition entraîne alors immédiatement que $d \in A^\times$.

Remarque VII.1.7. On peut démontrer que la réciproque est fausse en général.

Par exemple, 2 et X sont premiers entre eux dans $\mathbb{Z}[X]$, mais ne sont pas étrangers (**exercice**).

On peut aussi généraliser la notion de pgcd et de ppcm.

Définition VII.1.8. Soit A un anneau, et soient $a, b \in A$. On dit que $d \in A$ est un *pgcd* de a et b si d vérifie les propriétés suivantes :

(1) on a $d \mid a$ et $d \mid b$;
(2) pour tout $c \in A$ tel que $c \mid a$ et $c \mid b$, on a $c \mid d$.

On dit que $m \in A$ est un *ppcm* de a et b si m vérifie les propriétés suivantes :

(1) on a $a \mid m$ et $b \mid m$;
(2) pour tout $c \in A$ tel que $a \mid c$ et $b \mid c$, on a $m \mid c$.

Attention ! En général, un pgcd et un ppcm n'existent pas nécessairement. Par exemple, si $A = \mathbb{Z}[i\sqrt{5}]$, on peut montrer que 9 et $3(2 + i\sqrt{5})$ n'ont pas de pgcd, et que 3 et $2 + i\sqrt{5}$ n'ont pas de ppcm.

Remarques VII.1.9.

(1) Si $a = 0$, alors un pgcd de a et b est b. En effet, b divise à la fois 0 et b. Si maintenant $c \in A$ divise 0 et b, il divise b.
    En particulier, si $a = b = 0$, un pgcd de a et b est 0.
(2) Si $(a, b) \neq (0, 0)$, alors un pgcd de a et b (lorsqu'il existe) est non nul. En effet, si par exemple $a \neq 0$, on a $a = da', a' \in A$. Comme $a \neq 0$, on a nécessairement $d \neq 0$.
(3) Si $a = 0$ ou $b = 0$, alors 0 est un ppcm de a et b. En effet, 0 est un multiple commun à a et b. De plus, si c est un multiple commun à a et b, alors $c = 0$. Dans ce cas, c est bien multiple de 0.
(4) Si a et b sont non nuls, un ppcm m de a et b (s'il existe) est non nul dès que A est intègre, puisque m divise ab, qui est non nul par intégrité.

Le pgcd et le ppcm, lorsqu'ils existent, sont essentiellement uniques.

Lemme VII.1.10. *Soit A un anneau. Alors, un pgcd et un ppcm de deux éléments, lorsqu'ils existent, sont uniques à association près. Autrement dit, les idéaux engendrés par un pgcd et un ppcm respectivement sont indépendants du pgcd et ppcm choisis.*

En particulier, si A est intègre, deux pgcd/ppcm sont égaux à multiplication par un élément inversible près.

Démonstration. On le montre pour le pgcd, le cas du ppcm étant similaire. Supposons que d et d' soient deux pgcd de a et b. Alors, puisque d' est un pgcd de a et b et que d est un diviseur commun à a et b, on a $d \mid d'$. De même, on a $d' \mid d$. Autrement dit, d et d' sont associés, ce qui revient à dire que $(d) = (d')$. Lorsque l'anneau A est intègre, cela équivaut à dire qu'il existe $u \in A^\times$ tel que $d = ud'$ (cf. définition **??**). □

Remarque VII.1.11. Soient $a, b \in A$. Alors, a et b sont premiers entre eux si, et seulement si, 1 est un pgcd de a et b. Dans ce cas, tout pgcd de a et b est inversible.

En effet, si a et b sont premiers entre eux, alors tout diviseur commun est inversible. En particulier, la définition montre alors que 1 est un pgcd de a et b, puisque tout élément inversible divise 1. Tout pgcd de a et b est alors inversible par le lemme précédent. Inversement, si 1 est un pgcd de a et b, alors tout diviseur commun divise 1, et est donc inversible. Mais cela veut exactement dire que a et b sont premiers entre eux.

On peut trouver des exemples d'anneaux dans lesquels un pgcd ou un ppcm n'existent pas toujours. Néanmoins, dans les anneaux principaux, tout se passe bien.

La démonstration du résultat suivant est identique à celle de la proposition **??**.

Proposition VII.1.12. *Soit A un anneau principal. Soient $a, b \in A$, et soient $d, m \in A$ tels que*

$$(a) + (b) = (d) \quad \textit{et} \quad (a) \cap (b) = (m).$$

Alors, d est un pgcd de a et b, et m est un ppcm de a et b.

Le résultat suivant est alors une conséquence immédiate de la proposition précédente et de la remarque **??**, et généralise le théorème de Bézout.

Corollaire VII.1.13. *Soit A un anneau principal.*

Deux éléments $a, b \in A$ sont premiers entre eux si, et seulement si, ils sont étrangers. Autrement dit, a et b sont premiers entre eux si, et seulement si, il existe $u, v \in A$ tels que $ua + vb = 1$.

On peut alors démontrer une généralisation du lemme de Gauss à tout anneau principal.

Lemme VII.1.14 (Gauss). *Soit A un anneau principal. Soient $a, b \in A$ deux éléments premiers entre eux. Alors, pour tout $c \in A$, $a \mid bc \Longrightarrow a \mid c$.*

Démonstration. Gardons les notations de l'énoncé. Puisque a et b sont premiers entre eux, par le corollaire précédent, il existe $u, v \in A$ tels que $ua + bv = 1$. On a alors $uac + bcv = c$. Puisque a divise a et bc, il divise alors c. □

L'égalité $(a) + (b) = (d)$ montre que si d est un pgcd de a et b, il existe $u, v \in A$ tels que $ua + vb = d$ (relation de Bézout).

En général, il n'est pas possible calculer un pgcd de deux éléments et encore moins d'expliciter une relation de Bézout. En revanche, tout se passe pour le mieux lorsque l'on dispose dune division euclidienne (comme pour les anneaux $\mathbb{Z}$ et $K[X]$).

VII.2. Anneaux euclidiens

Définition VII.2.1. Un anneau *euclidien* est un couple (A, δ), où A est un anneau intègre et $\delta : A \setminus \{0\} \longrightarrow \mathbb{N}$ est une application telle que, pour tout $a \in A$, et pour tout $b \in A \setminus \{0\}$, il existe $q, r \in A$ tels que

$$a = qb + r, \ r = 0 \ \text{ ou } \ \delta(r) < \delta(b).$$

Attention ! On ne demande pas l'unicité de q et r.

Lorsque δ est spécifiée, on dit simplement que A est euclidien.

Exemples VII.2.2.

(1) L'anneau $\mathbb{Z}$ est euclidien (pour la valeur absolue).
(2) Si K est un corps, K est euclidien pour la fonction constante 1.
(3) Si K est un corps, $K[X]$ est euclidien pour le degré des polynômes, d'après le théorème **??** et la remarque qui suit.

L'intérêt des anneaux euclidiens est donné par le résultat suivant.

PROPOSITION VII.2.3. *Tout anneau euclidien est principal.*

Démonstration. Supposons (A, δ) euclidien. Alors, A est intègre. Soit $\mathfrak{a}$ un idéal de A. Si $\mathfrak{a} = (0)$, il n'y a rien à faire. On peut donc supposer $\mathfrak{a}$ non nul. L'ensemble $\{\delta(a) \mid a \in \mathfrak{a} \setminus \{0\}\}$ est alors une partie non vide de $\mathbb{N}$, et admet donc un plus petit élément. Soit $a_0 \in \mathfrak{a} \setminus \{0\}$ réalisant ce minimum. On va montrer que $\mathfrak{a} = (a_0)$. Puisque $\mathfrak{a}$ est un idéal de A contenant a_0, on a $(a_0) \subset \mathfrak{a}$. Réciproquement, soit $a \in \mathfrak{a}$. Écrivons

$$a = qa_0 + r \text{ avec } r = 0 \text{ ou } \delta(r) < \delta(a_0).$$

Puisque a et a_0 sont des éléments de $\mathfrak{a}$, on a $r = a - qa_0 \in \mathfrak{a}$. Si r était non nul, on aurait alors $r \in \mathfrak{a} \setminus \{0\}$ et $\delta(r) < \delta(a_0)$, ce qui contredirait le choix de a_0. Par conséquent, $r = 0$ et on a $a = qa_0 \in (a_0)$. Ceci montre l'inclusion manquante. $\square$

EXEMPLE VII.2.4. On retrouve le fait que $\mathbb{Z}$ et $K[X]$ (K corps) sont des anneaux principaux. Le lecteur attentif notera d'ailleurs que les démonstrations de ces deux faits sont identiques à la précédente.

On continue en donnant quelques exemples d'anneaux euclidiens.

LEMME VII.2.5. *Si $A = \mathbb{Z}[i], \mathbb{Z}[i\sqrt{2}]$ et $\mathbb{Z}[j]$, alors A est euclidien pour l'application*

$$\begin{aligned} \delta\colon A \setminus \{0\} &\longrightarrow \mathbb{N} \\ z &\longmapsto |z|^2. \end{aligned}$$

De même, si $A = \mathbb{Z}[\sqrt{2}]$, alors A est euclidien pour l'application

$$\begin{aligned} \delta\colon A \setminus \{0\} &\longrightarrow \mathbb{N} \\ a + b\sqrt{2} &\longmapsto |a^2 - 2b^2|. \end{aligned}$$

Démonstration. Tous ces anneaux sont intègres, puisque ce sont des sous-anneaux de $\mathbb{C}$.

Supposons tout d'abord que $A = \mathbb{Z}[i], \mathbb{Z}[i\sqrt{2}]$ ou $\mathbb{Z}[\sqrt{2}]$, et posons

$$\omega = i, i\sqrt{2} \text{ ou } \sqrt{2},$$

selon les cas. Tout élément de A s'écrit donc de manière unique sous la forme

$$z = a + b\omega,\ a, b \in \mathbb{Z}.$$

Posons $B = \mathbb{Q}[\omega]$. Tout élément de B s'écrit de manière unique sous la forme

$$z = u + v\omega,\ u, v \in \mathbb{Q}.$$

On définit alors $z^* = u - v\omega \in B$. Dans tous les cas, pour tout $z \in B$, on a

$$zz^* = u^2 - dv^2 \in \mathbb{Q}, \quad \text{avec } d \in \{-1, -2, 2\},$$

et $zz^* = 0$ si, et seulement si, $z = 0$. On étend l'application δ en une application de B dans $\mathbb{Q}$, encore notée δ par abus, en posant

$$\begin{aligned} \delta\colon B &\longrightarrow \mathbb{Q} \\ a + b\omega &\longmapsto |a^2 - db^2|, \end{aligned}$$

où $d = -1, -2$ ou 2.

On remarque alors que

$$\delta(z) = |zz^*| \text{ pour tout } z \in B.$$

Remarquons aussi que pour tous $z_1, z_2 \in A$, on a $(z_1 z_2)^* = z_1^* z_2^*$. Cela résulte en effet d'un simple calcul laissé au lecteur. En particulier, pour tous $z_1, z_2 \in B$, on a

$$\delta(z_1 z_2) = |z_1 z_2 z_1 z_2^*| = |z_1 z_2||z_1^* z_2^*| = \delta(z_1)\delta(z_2).$$

Notons enfin que, si $z \in A \setminus \{0\}$, $\delta(z)$ est un entier strictement positif.

Supposons maintenant que $A = \mathbb{Z}[j]$, et soit $B = \mathbb{Q}[j]$. Dans ce cas, δ est simplement le carré du module, et s'étend bien entendu à B. De plus, pour tous $u, v \in \mathbb{Q}$, on a

$$\delta(u + vj) = |u + vj|^2 = \left|(u - \frac{v}{2}) + iv\frac{\sqrt{3}}{2}\right|^2 = (u - \frac{v}{2})^2 + \frac{3v^2}{4} = u^2 - uv + v^2.$$

En particulier, si $z \in A$, $\delta(z)$ est un entier positif, qui est nul si, et seulement si, $z = 0$. Notons également que A et B sont stables par conjugaison complexe.

Pour harmoniser les notations avec les autres cas, si $z \in B$, on note $z^* \in B$ son conjugué. On note également $\omega = j$ dans ce cas.

Passons maintenant à la démonstration proprement dite. Soient $z_1, z_2 \in A$, avec $z_2 \neq 0$. Alors, on a

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*}.$$

Comme $z_2 z_2^* \in \mathbb{Z} \setminus \{0\}$ et que $z_1 z_2^* \in A$, on peut donc écrire $\dfrac{z_1}{z_2} = u + v\omega$, avec $u, v \in \mathbb{Q}$. Soient $m, n \in \mathbb{Z}$ deux entiers vérifiant

$$|u - m| \leq \frac{1}{2} \text{ et } |v - n| \leq \frac{1}{2}.$$

Posons $q = m + n\omega$ et $r = z_1 - qz_2 = \left(\dfrac{z_1}{z_2} - q\right)z_2$. On a alors

$$\delta(r) = \delta(((u - m) + (v - n)\omega)z_2) = \delta((u - m) + (v - n)\omega)\delta(z_2).$$

Nous voulons maintenant montrer que $\delta((u - m) + (v - n)\omega)) < 1$.

Si $A = \mathbb{Z}[i], \mathbb{Z}[i\sqrt{2}]$ ou $\mathbb{Z}[\sqrt{2}]$, on a

$$\delta((u - m) + (v - n)\omega) = |(u - m)^2 - d(v - n)^2|,$$

avec $d = -1, -2$ ou 2. Dans les trois cas, on a donc

$$\delta((u - m) + (v - n)\omega) \leq (u - m)^2 + 2(v - n)^2 \leq \frac{1}{4} + \frac{1}{2} < 1.$$

Si $A = \mathbb{Z}[j]$, on a

$$\delta((u - m) + (v - n)\omega) \leq |u - m|^2 + |u - m|^2 + |u - m||v - n| \leq \frac{3}{4} < 1.$$

Ainsi, dans tous les cas, on a $z_1 = qz_2 + r$ avec, soit $r = 0$, soit $\delta(r) < \delta(z_2)$, ce qui démontre que (A, δ) est euclidien. □

Remarque VII.2.6. La démonstration précédente nous donne un moyen de calculer un quotient dans chaque cas. Si $z_1, z_2 \in A$, avec $z_2 \neq 0$, on écrit

$$\frac{z_1}{z_2} = u + vi, u + vj, u + iv\sqrt{2}, u + v\sqrt{2},\ u, v \in \mathbb{Q}$$

selon le cas dans lequel on se trouve. On choisit ensuite $m, n \in \mathbb{Z}$ vérifiant

$$|u - m| \leq \frac{1}{2}, |v - n| \leq \frac{1}{2}.$$

Alors, un quotient q de z_1 par z_2 est respectivement donné par

$$q = m + ni, m + nj, m + in\sqrt{2}, m + n\sqrt{2}.$$

On pose ensuite $r = z_1 - qz_2$.

Remarquons au passage que l'on pourrait avoir plusieurs choix possibles pour m ou n, dans le cas où $u = k + \frac{1}{2}$ ou $v = k + \frac{1}{2}$ avec $k \in \mathbb{Z}$.

Nous allons maintenant généraliser l'algorithme d'Euclide bien connu pour l'anneau $\mathbb{Z}$ à tout anneau euclidien.

Algorithme d'Euclide. Soit (A, δ) un anneau euclidien, soient a et b deux éléments non nuls. Tant que c'est possible (i.e. tant que l'on ne divise pas par 0), on définit des éléments q_i et r_{i+1} comme suit par divisions euclidiennes successives en posant :

$$r_{-1} = a,\ r_0 = b,\ r_{i-1} = q_i r_i + r_{i+1}.$$

En particulier, si $r_i \neq 0$, on a soit $r_{i+1} = 0$, soit $\delta(r_{i+1}) < \delta(r_i)$.

Remarquons que q_i n'est défini que pour $i \geq 0$.

Pour toute valeur $i \geq -1$ pour laquelle $r_i \neq 0$, on peut alors définir des éléments u_i et v_i par récurrence en posant :

$$u_{-1} = 1,\ u_0 = 0,\ u_{i+1} = -q_i u_i + u_{i-1},$$
$$v_{-1} = 0,\ v_0 = 1,\ v_{i+1} = -q_i v_i + v_{i-1}.$$

Proposition VII.2.7. *Il existe un entier $i \geq 1$ tel que $r_i = 0$. Si $n \geq 0$ est le plus grand entier tel que $r_n \neq 0$, on pose*

$$d = r_n, u = u_n, v = v_n.$$

Alors, d est un pgcd de a et b, et on a la relation de Bézout

$$ua + vb = d.$$

Démonstration. Supposons que r_i soit non nul pour tout $i \geq -1$. Par construction, la suite $(\delta(r_i))_{i \geq -1}$ est alors une suite d'entiers naturels strictement décroissante. Une telle suite n'existant pas, on obtient une contradiction. Ainsi, il existe un entier[1] $i \geq 1$ tel que $r_i = 0$. Soit $n \geq 0$ le plus grand entier tel que r_n soit non nul. On a donc $r_{n+1} = 0$, et on dispose de suites finies d'éléments

$$(r_i)_{i \in [\![-1,n+1]\!]}, (q_i)_{i \in [\![0,n+1]\!]}, (u_i)_{i \in [\![-1,n+1]\!]}, (v_i)_{i \in [\![-1,n+1]\!]}.$$

Remarquons que les égalités définissant $(r_i)_{i \in [\![-1,n]\!]}$ se récrivent

$$\begin{pmatrix} r_{i-1} \\ r_i \end{pmatrix} = M_i \begin{pmatrix} r_i \\ r_{i+1} \end{pmatrix},$$

avec $M_i = \begin{pmatrix} q_i & 1 \\ 1 & 0 \end{pmatrix}$, pour tout $i \in [\![0, n]\!]$.

En particulier, la matrice produit $M = M_0 \cdots M_n$ est à coefficients dans l'anneau A. Or, on a

$$\begin{pmatrix} r_{-1} \\ r_0 \end{pmatrix} = M \begin{pmatrix} r_n \\ r_{n+1} \end{pmatrix},$$

1. Rappelons que $r_{-1} = a$ et $r_0 = b$ sont supposés non nuls.

soit

$$\begin{pmatrix} a \\ b \end{pmatrix} = M \begin{pmatrix} d \\ 0 \end{pmatrix}.$$

On en déduit immédiatement que $d \mid a$ et $d \mid b$.

Remarquons que l'on a également

$$\begin{pmatrix} u_{i-1} \\ u_i \end{pmatrix} = M_i \begin{pmatrix} u_i \\ u_{i+1} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} v_{i-1} \\ v_i \end{pmatrix} = M_i \begin{pmatrix} v_i \\ v_{i+1} \end{pmatrix}$$

pour tout $i \in [\![0, n]\!]$.

Ainsi,

$$\begin{pmatrix} u_{-1} \\ u_0 \end{pmatrix} = M \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} v_{-1} \\ v_0 \end{pmatrix} = M \begin{pmatrix} v_n \\ v_{n+1} \end{pmatrix}.$$

Matriciellement, cela se traduit par

$$I_2 = M \begin{pmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{pmatrix}.$$

En particulier, $M \in \mathrm{GL}_2(K_A)$ et on a $M^{-1} = \begin{pmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{pmatrix}$.

L'égalité $\begin{pmatrix} a \\ b \end{pmatrix} = M \begin{pmatrix} d \\ 0 \end{pmatrix}$ nous donne alors

$$\begin{pmatrix} d \\ 0 \end{pmatrix} = M^{-1} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

En particulier, on a $au_n + bv_n = d$, soit $ua + vb = d$ avec les notations du lemme. Mais alors, si $c \in A$ est un diviseur commun à a et b, il divise également $ua + vb$, soit $c \mid d$. Cela démontre que d est un pgcd de a et b, et achève la démonstration. □

Nous décrivons maintenant un moyen pratique de faire fonctionner l'algorithme.

Disposition manuelle. On remarque que l'on a en fait

$$\begin{aligned} r_{-1} = a, r_0 = b, r_{i+1} &= r_{i-1} - q_i r_i, \\ u_{-1} = 1, u_0 = 0, u_{i+1} &= u_{i-1} - q_i u_i, \\ v_{-1} = 0, v_0 = 1, v_{i+1} &= v_{i-1} - q_i v_i. \end{aligned}$$

On commence par remplir un tableau à quatre colonnes (ou seulement deux colonnes si l'on n'est pas intéressés par le calcul d'une relation de Bézout) comme suit :

$-q$	r	u	v
	a	1	0
	b	0	1

Expliquons le procédé sur un exemple. On prend $A = \mathbb{Z}, a = 32, b = 7$. On part donc du tableau suivant :

$-q$	r	u	v
	32	1	0
	7	0	1

Remarquons que q_{-1} n'est pas défini. L'étape suivante est de calculer q_0. On dispose d'un algorithme pour le faire dans les anneaux

$$\mathbb{Z}, K[X], \mathbb{Z}[i], \mathbb{Z}[j], \mathbb{Z}[i\sqrt{2}], \text{ et } \mathbb{Z}[\sqrt{2}],$$

mais cela peut être plus délicat dans le cas d'un anneau euclidien quelconque.

Dans notre exemple, $32 = 4 \times 7 + 4$, et donc $q_0 = 4$. On remplit alors le tableau comme suit :

$-q$	r	u	v
	32	1	0
-4	7	0	1

Pour calculer r_1, u_1, v_1, on multiplie les éléments des colonnes $2, 3$ et 4 de la dernière ligne par l'élément de la colonne 1, et on ajoute aux éléments correspondants dans la ligne du dessus.

Dans notre exemple, $r_1 = -4 \times 7 + 32$, $u_1 = -4 \times 0 + 1$, et $v_1 = -4 \times 1 + 0$. On obtient :

$-q$	r	u	v
	32	1	0
-4	7	0	1
	4	1	-4

Pour calculer les valeurs suivantes, on considère les deux dernières lignes, et on calcule le nouveau quotient en effectuant la division des deux dernières valeurs de la colonne des restes successifs (i.e. la colonne 2).

Dans notre exemple, on doit diviser 7 par 4. On a $7 = 1 \times 4 + 3$, et le nouveau quotient est 1. On obtient :

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4

On répète le procédé pour calculer les valeurs de la ligne suivante. La nouvelle valeur de r_i est $-1 \times 4 + 7$, celle de u_i est $-1 \times 1 + 0$, et celle de v_i est $-1 \times (-4) + 1$.

On obtient donc :

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4
	3	-1	5

Pour trouver le nouveau quotient, on divise 4 par 3 : on a $4 = 1 \times 3 + 1$, et ainsi le nouveau quotient est 1. On a ainsi :

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4
-1	3	-1	5

La nouvelle valeur de r_i est $-1 \times 3 + 4$, celle de u_i est $-1 \times (-1) + 1$, et celle de v_i est $-1 \times 5 + (-4)$. On a donc :

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4
-1	3	-1	5
	1	2	-9

On constate alors que la valeur du reste suivant est 0, puisque l'on doit diviser 3 par 1. On arrête donc ici l'algorithme. Ainsi, un pgcd de 32 et 7 est 1 (la dernière valeur de r_i dans le tableau) et on a $u = 2$ et $v = -9$.

On a bien $32 \times 2 + 7 \times (-9) = 64 - 63 = 1$.

Donnons maintenant un exemple dans $\mathbb{Z}[i]$.

Exemple VII.2.8. Soient $A = \mathbb{Z}[i], a = 12 + 45i$ et $b = 4 + 12i$. On part donc du tableau suivant :

$-q$	r	u	v
	$12 + 45i$	1	0
	$4 + 12i$	0	1

On a ensuite

$$\frac{12 + 45i}{4 + 12i} = \frac{147}{40} + \frac{9}{40}i.$$

Un quotient est donc donné par $q_0 = 4$, et donc on obtient :

$-q$	r	u	v
	$12 + 45i$	1	0
-4	$4 + 12i$	0	1
	$-4 - 3i$	1	-4

On a maintenant

$$\frac{4 + 12i}{-4 - 3i} = -\frac{52}{25} - \frac{36}{25}i,$$

d'où $q_1 = -2 - i$. Tous calculs faits, le tableau se complète comme suit :

$-q$	r	u	v
	$12 + 45i$	1	0
-4	$4 + 12i$	0	1
$2 + i$	$-4 - 3i$	1	-4
	$-1 + 2i$	$2 + i$	$-7 - 4i$

On a ensuite

$$\frac{-4 - 3i}{-1 + 2i} = -\frac{2}{5} + \frac{11}{5}i,$$

soit $q_2 = 2i$. On obtient donc :

$-q$	r	u	v
	$12 + 45i$	1	0
-4	$4 + 12i$	0	1
$2 + i$	$-4 - 3i$	1	-4
$-2i$	$-1 + 2i$	$2 + i$	$-7 - 4i$
	$-i$	$3 - 4i$	$-12 + 14i$

Comme $-i$ est inversible, le prochain reste sera nécessairement nul. Un pgcd de a et b est donc $-i$, et on a bien la relation

$$(3 - 4i)(12 + 45i) + (-12 + 14i)(4 + 12i) = -i.$$

Chapitre VIII

Anneaux quotients, théorème chinois

VIII.1. Anneaux quotients, théorème de factorisation

Rappelons que l'on a muni le groupe additif $\mathbb{Z}/n\mathbb{Z}$ d'une structure d'anneau en posant

$$\overline{a}\ \overline{b} = \overline{ab}.$$

Il est facile de voir alors que l'application

$$\begin{aligned} \pi \colon \mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ a &\longmapsto \overline{a} \end{aligned}$$

est un morphisme d'anneaux.

On se propose maintenant de généraliser cette construction à un anneau commutatif arbitraire. Soit A un anneau commutatif, et soit $\mathfrak{a}$ un sous-groupe additif de A. On veut maintenant munir $A/\mathfrak{a}$ d'une structure d'anneau telle que la projection canonique $\pi : A \longrightarrow A/\mathfrak{a}$ soit un morphisme d'anneaux.

Si une telle structure d'anneau existe, alors pour tous $\overline{a}, \overline{b} \in A/\mathfrak{a}$, on doit avoir

$$\overline{a} + \overline{b} = \pi(a) + \pi(b) = \pi(a+b) = \overline{a+b},$$

ainsi que

$$\overline{a}\ \overline{b} = \pi(a)\pi(b) = \pi(ab) = \overline{ab}.$$

Puisque $\mathfrak{a}$ est un sous-groupe additif du groupe abélien $(A, +)$, $\mathfrak{a}$ est distingué dans A. Ainsi, la loi d'addition précédente est bien définie, et $(A/\mathfrak{a}, +)$ est un groupe abélien. Examinons le cas de la loi produit. Dans le cas où elle est bien définie, pour tout $x \in \mathfrak{a}$ et tout $a \in A$, on doit avoir

$$\overline{ax} = \overline{a}\ \overline{x} = \overline{a}\ \overline{0} = \overline{a0} = \overline{0}.$$

Autrement dit, on a $ax \in \mathfrak{a}$. Ainsi, si la loi produit est bien définie, $\mathfrak{a}$ est nécessairement un idéal de A.

La proposition suivante montre que réciproquement, si $\mathfrak{a}$ est un idéal de A, ces deux lois sont bien définies, et répondent à notre problème.

Proposition VIII.1.1. *Soit A un anneau commutatif, et soit $\mathfrak{a}$ un idéal de A. Les applications*

$$\begin{aligned} A/\mathfrak{a} \times A/\mathfrak{a} &\longrightarrow A/\mathfrak{a} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a+b} \end{aligned}$$

et

$$\begin{aligned} A/\mathfrak{a} \times A/\mathfrak{a} &\longrightarrow A/\mathfrak{a} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{ab} \end{aligned}$$

sont bien définies, et confèrent à $A/\mathfrak{a}$ une structure d'anneau, de neutre additif $\overline{0}$, et de neutre multiplicatif $\overline{1}$.

De plus, l'application

$$\pi : A \longrightarrow A/\mathfrak{a}$$

est un morphisme d'anneaux.

Démonstration. Comme $\mathfrak{a}$ est un sous-groupe de $(A,+)$, donc nécessairement distingué dans A, on sait déjà que l'addition est bien définie et que le couple $(A/\mathfrak{a},+)$ est un groupe abélien, de neutre $\overline{0}$. Montrons que la multiplication est bien définie. Supposons que $a,a',b,b' \in A$ vérifient $\overline{a'}=\overline{a}$ et $\overline{b'}=\overline{b}$. Il existe donc $x,y \in \mathfrak{a}$ tels que $a'=a+x, b'=b+y$, et on a alors

$$a'b' = (a+x)(b+y) = ab+ay+xb+xy.$$

Comme $x,y \in \mathfrak{a}$ et que $\mathfrak{a}$ est un idéal, on a $ay+xb+xy \in \mathfrak{a}$ et donc $\overline{a'b'}=\overline{ab}$, ce qu'il fallait vérifier. Les propriétés de commutativité, associativité, distributivité découlent alors des propriétés correspondantes des lois de A. De plus, $\overline{1}$ est clairement un neutre pour la multiplication. Enfin, le fait que π soit un morphisme d'anneaux provient de la définition même de la multiplication de $A/\mathfrak{a}$. □

Définition VIII.1.2. Soit A un anneau, et soit $\mathfrak{a}$ un idéal. L'anneau $A/\mathfrak{a}$ est appelé *l'anneau quotient* de A par $\mathfrak{a}$. L'application

$$\pi : A \longrightarrow A/\mathfrak{a}$$

est appelée la *projection canonique.*

On a alors le même genre de propriétés que pour les groupes quotients.

Proposition VIII.1.3. *Soit A un anneau et soit $\mathfrak{a}$ un idéal. Alors, la projection canonique $\pi : A \longrightarrow A/\mathfrak{a}$ induit une correspondance bijective entre l'ensemble des idéaux de A contenant $\mathfrak{a}$ et l'ensemble des idéaux de $A/\mathfrak{a}$.*

Plus précisément, si $\mathfrak{a}'$ est un idéal de A contenant $\mathfrak{a}$, alors

$$\mathfrak{a}'/\mathfrak{a} = \pi(\mathfrak{a}') = \{\overline{a'} \mid a' \in \mathfrak{a}'\}$$

est un idéal de $A/\mathfrak{a}$. Réciproquement, si $\mathfrak{b}$ est un idéal de $A/\mathfrak{a}$, alors $\pi^{-1}(\mathfrak{b})$ est un idéal de A contenant $\mathfrak{a}$, et ces deux constructions sont inverses l'une de l'autre.

Démonstration. D'après la proposition **??**, il reste simplement à voir que $\pi^{-1}(\mathfrak{b})$ est un idéal de A pour tout idéal $\mathfrak{b}$ de $A/\mathfrak{a}$, et que $\pi(\mathfrak{a}')$ est un idéal de $A/\mathfrak{b}$ pour tout idéal $\mathfrak{a}'$ de A contenant $\mathfrak{a}$.

Puisque π est un morphisme d'anneaux, $\pi^{-1}(\mathfrak{b})$ est un idéal par le lemme **??**. On sait déjà que $\pi(\mathfrak{a}')$ est un sous-groupe de $A/\mathfrak{a}$ par le lemme **??** (1). Or, pour tout $\overline{a} \in A/\mathfrak{a}$, et tout $x \in \mathfrak{a}'$, on a

$$\overline{a}\pi(x) = \overline{a}\ \overline{x} = \overline{ax} \in \pi(\mathfrak{a}'),$$

car $x \in \mathfrak{a}'$ et $\mathfrak{a}'$ est un idéal de A. Ainsi, $\pi(\mathfrak{a}')$ est un idéal de $A/\mathfrak{a}$. Ceci achève la démonstration. □

Comme dans le cas des groupes, on a alors à notre disposition un théorème de factorisation, et divers théorèmes d'isomorphisme.

Théorème VIII.1.4 (de factorisation). *Soit A un anneau, et soit $\mathfrak{a}$ un idéal. Soit $f : A \longrightarrow B$ un morphisme d'anneaux (commutatifs) tel que $\mathfrak{a} \subset \mathrm{Ker}(f)$. Alors, il existe un unique morphisme d'anneaux*

$$\overline{f} : A/\mathfrak{a} \longrightarrow B$$

tel que $f = \overline{f} \circ \pi$. Ce morphisme d'anneaux est défini par

$$\overline{f}(\overline{x}) = f(x) \quad \textit{pour tout } \overline{x} \in A/\mathfrak{a}.$$

De plus, pour tout anneau B, il y a une bijection entre l'ensemble $\mathrm{Hom}(A/\mathfrak{a}, B)$ *et l'ensemble* $\{f \in \mathrm{Hom}(A, B) \mid \mathfrak{a} \subset \mathrm{Ker}(f)\}$. *Elle est donnée par est donnée par*

$$\begin{array}{ccc} \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longleftarrow\!\shortmid & f. \end{array}$$

Démonstration. On sait déjà qu'il existe un unique morphisme de groupes additifs $\overline{f}$ vérifiant les propriétés requises, par le théorème de factorisation pour les groupes, et qu'il est défini par la formule précédente. Il suffit donc vérifier que $\overline{f}$ est aussi un morphisme d'anneaux. Mais on a

$$\overline{f}(\overline{1}_A) = f(1_A) = 1_B,$$

et pour tous $\overline{x}, \overline{y} \in A/\mathfrak{a}$, on a

$$\overline{f}(\overline{x})\overline{f}(\overline{y}) = f(x)f(y) = f(xy) = \overline{f}(\overline{xy}) = \overline{f}(\overline{x}\ \overline{y}).$$

D'après le théorème **??** la correspondance entre $\mathrm{Hom}(A/\mathfrak{a}, B)$ et l'ensemble A' des morphismes $f : A \longrightarrow B$ s'annulant sur $\mathfrak{a}$ est donnée par

$$\begin{array}{ccc} \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longleftarrow\!\shortmid & f. \end{array}$$

Ceci achève la démonstration. □

Soit $f : A \longrightarrow B$ un morphisme d'anneaux. Puisque $\mathrm{Ker}(f)$ est un idéal de A, on a en particulier un morphisme d'anneaux

$$\overline{f} : A/\mathrm{Ker}(f) \longrightarrow B.$$

Le lemme **??** entraîne alors le résultat suivant.

Lemme VIII.1.5. *Soit $f : A \longrightarrow B$ un morphisme d'anneaux. Alors, le morphisme d'anneaux*

$$\overline{f} : A/\mathrm{Ker}(f) \longrightarrow B$$

est injectif.

On a alors un théorème d'isomorphisme semblable au théorème **??**.

Théorème VIII.1.6 (Premier théorème d'isomorphisme). *Pour tout morphisme d'anneaux commutatifs $f : A \longrightarrow B$, le morphisme d'anneaux*

$$\overline{f} : A/\mathrm{Ker}(f) \longrightarrow B$$

induit un isomorphisme d'anneaux

$$A/\mathrm{Ker}(f) \simeq \mathrm{Im}(f).$$

Démonstration. D'après le théorème **??**, on sait que $\overline{f}$ est un isomorphisme de groupes, donc une application bijective. Par le théorème de factorisation précédent, on sait aussi que $\overline{f}$ est un morphisme d'anneaux. C'est donc un isomorphisme d'anneaux, d'où le résultat. □

On a également un autre théorème d'isomorphisme, correspondant au troisième théorème d'isomorphisme pour les groupes (le deuxième ne donnant rien au niveau de la structure d'anneau).

Théorème VIII.1.7 (Second théorème d'isomorphisme). *Soit A un anneau, et soient $\mathfrak{a}, \mathfrak{b}$ deux idéaux de A. On suppose que $\mathfrak{a} \subset \mathfrak{b}$. Alors, le groupe quotient $\mathfrak{b}/\mathfrak{a}$ est un idéal de $A/\mathfrak{a}$, et on a un isomorphisme d'anneaux*

$$(A/\mathfrak{a})/(\mathfrak{b}/\mathfrak{a}) \simeq A/\mathfrak{b}.$$

Démonstration. La première partie découle de la proposition **??**. Pour le reste, on procède comme dans la démonstration du théorème **??** (mais en constatant cette fois que l'on travaille avec des morphismes d'anneaux). □

VIII.2. Idéaux premiers, maximaux

Définition VIII.2.1. Soit A un anneau commutatif. Un idéal $\mathfrak{p}$ de A est dit *premier* si $\mathfrak{p} \neq A$, et si pour tous $a, b \in A$, on a

$$ab \in \mathfrak{p} \Longrightarrow a \in \mathfrak{p} \quad \text{ou} \quad b \in \mathfrak{p}.$$

Autrement dit, $\mathfrak{p}$ est premier si $\mathfrak{p} \neq A$, et pour tous éléments $a, b \in A \setminus \mathfrak{p}$, on a $ab \in A \setminus \mathfrak{p}$.

Un idéal $\mathfrak{m}$ est dit *maximal* si $\mathfrak{m} \neq A$ et si pour tout idéal $\mathfrak{a}$ de A, on a

$$\mathfrak{m} \subset \mathfrak{a} \subset A \Longrightarrow \mathfrak{a} = \mathfrak{m} \quad \text{ou} \quad \mathfrak{a} = A.$$

Autrement dit, $\mathfrak{m}$ est maximal si c'est un élément maximal de l'ensemble des idéaux de A qui sont distincts de A.

Exemples VIII.2.2.

(1) L'idéal (0) est premier si, et seulement si, A est intègre.

Cela provient immédiatement de la définition d'un anneau intègre. Les vérifications sont laissées au lecteur.

(2) Si p est premier, $p\mathbb{Z}$ est un idéal premier de $\mathbb{Z}$. Là encore, on laisse le lecteur vérifier ce fait à titre d'exercice.

L'idéal $p\mathbb{Z}$ est aussi maximal.

En effet, il est clair que $1 \notin p\mathbb{Z}$, donc $p\mathbb{Z} \neq \mathbb{Z}$. Supposons qu'un idéal $\mathfrak{a} = n\mathbb{Z}, n \geq 0$ vérifie $p\mathbb{Z} \subset n\mathbb{Z} \subset \mathbb{Z}$. Alors, $n|p$, et donc soit $n = 1$ ou p. Ainsi, soit $n\mathbb{Z} = \mathbb{Z}$, soit $n\mathbb{Z} = p\mathbb{Z}$.

On veut maintenant démontrer l'existence d'idéaux maximaux, existence qui n'est a priori pas garantie. On commence par des considérations générales sur les ensembles ordonnés.

Définition VIII.2.3. Soit E un ensemble, et soit « $\leq$ » une relation sur E. On dit alors que « $\leq$ » est une *relation d'ordre* si l'on a les propriétés suivantes :

(1) réflexivité : pour tout $x \in E$, on a $x \leq x$;

(2) antisymétrie : pour tous $x, y \in E$, $(x \leq y \text{ et } y \leq x) \Longrightarrow x = y$;

(3) transitivité : pour tous $x, y, z \in E$, $(x \leq y \text{ et } y \leq z) \Longrightarrow x \leq z$.

On note alors $x < y$ si $x \leq y$ et $x \neq y$.

On dit que la relation d'ordre « $\leq$ » est *totale*, ou que $(E, \leq)$ est totalement ordonné, si pour tous $x, y \in E$, on a $x \leq y$ ou $y \leq x$.

Exemples VIII.2.4.

(1) Soit E un ensemble. Alors, la relation d'inclusion est une relation d'ordre sur l'ensemble $\mathcal{P}(E)$ des parties de E. Cette relation d'ordre n'est pas totale dès que E possède au moins deux éléments.

(2) La relation d'ordre usuelle sur l'ensemble $\mathbb{N}$ est une relation d'ordre total.

Définition VIII.2.5. Soit $(E, \leq)$ un ensemble ordonné.

On dit que $z \in E$ est *maximal* si pour tout $x \in E$, on a $z \leq x \Longrightarrow x = z$.

Autrement dit, $z \in E$ est maximal s'il n'existe aucun élément $x \in E$ tel que $z < x$.

On dit que $z \in E$ est un *maximum* (ou un *plus grand élément*) si $x \leq z$ pour tout $x \in E$.

Remarques VIII.2.6.

(1) Un maximum, s'il existe, est unique, et c'est aussi le seul élément maximal.

En effet, si $z, z' \in E$ sont deux maximums de E, on a $z' \leq z$ et $z \leq z'$ par définition d'un maximum, et donc $z' = z$, par définition d'une relation d'ordre.

De plus, si $z \in E$ est un maximum, et si $x \in E$ vérifie $z \leq x$, alors puisque $x \leq z$ par définition d'un maximum, on a $x = z$. D'autre part, si $z' \in E$ est maximal, et puisque $z' \leq z$ par définition d'un maximum, on a $z = z'$.

(2) En revanche, un élément maximal n'est pas nécessairement un maximum.

Par exemple, soit E l'ensemble des parties strictes de $X = \{0, 1, 2\}$, ordonné par l'inclusion. Alors, toute partie de X à deux éléments est un élément maximal, mais E ne possède pas d'élément maximum, car il n'existe pas de partie stricte de X qui contienne toutes les autres.

(3) Tout ensemble fini totalement ordonné possède un maximum.

Définition VIII.2.7. Soit $(E, \leq)$ un ensemble ordonné, et soit F une partie non vide de E. On dit que $x \in E$ est un *majorant* de F si l'on a

$$y \leq x \quad \text{pour tout } y \in F.$$

On dit que E est *inductif* si tout sous-ensemble de E totalement ordonné admet un majorant.

On a alors le résultat suivant.

Théorème VIII.2.8. [Lemme de Zorn] *Tout ensemble inductif non vide possède un élément maximal.*

Théorème VIII.2.9 (Krull). *Soit A un anneau commutatif.*

Alors, tout idéal $\mathfrak{a} \neq A$ est contenu dans un idéal maximal. En particulier, tout anneau commutatif non trivial possède au moins un idéal maximal.

Démonstration. Considérons l'ensemble E défini par

$$E = \{\mathfrak{b} \mid \mathfrak{b} \neq A \text{ idéal de } A, \mathfrak{a} \subset \mathfrak{b}\}.$$

Alors, E est non vide (car il contient $\mathfrak{a}$), partiellement ordonné par l'inclusion. Soit $(I, \leq)$ un ensemble totalement ordonné, et soit $F = (\mathfrak{b}_i)_{i \in I}$ une famille d'éléments de E totalement ordonnée indexée par I. Autrement dit, pour tous $i, j \in I$ tels que $i \leq j$, on a $\mathfrak{b}_i \subset \mathfrak{b}_j$. Posons

$$\mathfrak{b} = \bigcup_{i \in I} \mathfrak{b}_i.$$

Alors, $\mathfrak{a} \subset \mathfrak{b}$. De plus, $\mathfrak{b}$ est un idéal de l'anneau A. En effet, si $x, y \in \mathfrak{b}$, il existe des indices $i, j \in I$ tels que $x \in \mathfrak{b}_i$ et $y \in \mathfrak{b}_j$. Soit $k = \max(i, j)$. On a ainsi $i \leq k$ et $j \leq k$. Puisque F est totalement ordonné, on a alors

$$x \in \mathfrak{b}_i \subset \mathfrak{b}_k \quad \text{et} \quad y \in \mathfrak{b}_j \subset \mathfrak{b}_k.$$

Puisque $\mathfrak{b}_k$ est un idéal, alors pour tout $a \in A$, on a $x + ay \in \mathfrak{b}_k \subset \mathfrak{b}$. Donc $\mathfrak{b}$ est un idéal de A. De plus, on a $\mathfrak{b} \neq A$. Sinon, $\mathfrak{b}$ contiendrait 1. Mais alors, il existerait $i \in I$ tel que $1 \in \mathfrak{b}_i$ et donc on aurait $\mathfrak{b}_i = A$, ce qui contredirait le fait que $\mathfrak{b}_i \in E$.

Ainsi $\mathfrak{b} \in E$, et puisque $\mathfrak{b}_i \subset \mathfrak{b}$ pour tout $i \in I$, $\mathfrak{b}$ est un majorant de F. Ainsi, E est inductif, non vide et donc contient un élément maximal $\mathfrak{m}$ par le lemme de Zorn. On a donc $\mathfrak{m} \neq A$ et $\mathfrak{a} \subset \mathfrak{m}$. Montrons que $\mathfrak{m}$ est un idéal maximal. Soit $\mathfrak{b}'$ un idéal de A tel que $\mathfrak{m} \subset \mathfrak{b}' \subset A$. Supposons que $\mathfrak{b}' \neq A$. Alors, on a $\mathfrak{b}' \in E$ et donc $\mathfrak{b}' = \mathfrak{m}$ par maximalité de $\mathfrak{m}$. Ceci achève la démonstration du premier point. Pour montrer le dernier point, il suffit de prendre $\mathfrak{a} = (0)$, qui est distinct de A, car A est non trivial. □

On achève ce paragraphe en donnant une caractérisation des idéaux premiers et maximaux en termes de quotients.

Théorème VIII.2.10. *Soit A un anneau commutatif.*

(1) *Un idéal $\mathfrak{p}$ est premier si, et seulement si, $A/\mathfrak{p}$ est intègre.*
(2) *Un idéal $\mathfrak{m}$ est maximal si, et seulement si, $A/\mathfrak{m}$ est un corps.*

En particulier, tout idéal maximal est premier.

Démonstration.

(1) Supposons que $A/\mathfrak{p}$ soit intègre. En particulier, ce n'est pas l'anneau trivial, et donc $\mathfrak{p} \neq A$. Soient maintenant $a, b \in A$ tels que $ab \in \mathfrak{p}$. Alors, on a

$$\overline{a}\ \overline{b} = \overline{ab} = \overline{0}.$$

Puisque $A/\mathfrak{p}$ est intègre, on a $\overline{a} = \overline{0}$ ou $\overline{b} = \overline{0}$, c'est-à-dire $a \in \mathfrak{p}$ ou $b \in \mathfrak{p}$. Ainsi, $\mathfrak{p}$ est premier.

Inversement, si $\mathfrak{p}$ est premier, alors $\mathfrak{p} \neq A$, donc $A/\mathfrak{p}$ n'est pas trivial. Si maintenant $\overline{a}\ \overline{b} = \overline{0}$ dans $A/\mathfrak{p}$, alors $\overline{ab} = \overline{0}$, et donc $ab \in \mathfrak{p}$. Comme $\mathfrak{p}$ est premier, on obtient $a \in \mathfrak{p}$ ou $b \in \mathfrak{p}$, c'est-à-dire $\overline{a} = \overline{0}$ ou $\overline{b} = \overline{0}$. Ainsi, $A/\mathfrak{p}$ est intègre.

(2) Supposons $A/\mathfrak{m}$ soit un corps. Encore une fois, l'anneau quotient n'est pas trivial, et $\mathfrak{m} \neq A$. Soit $\mathfrak{a}$ un idéal de A tel que $\mathfrak{m} \subset \mathfrak{a} \subset A$, et supposons que $\mathfrak{a} \neq \mathfrak{m}$. On doit montrer que $\mathfrak{a} = A$. Soit $a \in \mathfrak{a} \setminus \mathfrak{m}$. Alors, $\overline{a} \neq \overline{0}$, et donc $\overline{a}$ est inversible, puisque $A/\mathfrak{m}$ est un corps. Ainsi, il existe $\overline{b} \in A/\mathfrak{m}$ tel que $\overline{a}\ \overline{b} = \overline{1}$. On a donc $\overline{1 - ab} = \overline{0}$, et donc $1 - ab \in \mathfrak{m} \subset \mathfrak{a}$. Puisque $a \in \mathfrak{a}$ et $\mathfrak{a}$ est un idéal, $ab \in \mathfrak{a}$, donc $1 = (1 - ab) + ab \in \mathfrak{a}$. Ainsi $\mathfrak{a}$ contient un élément inversible, et donc $\mathfrak{a} = A$ par le lemme **??**. Il s'ensuit que $\mathfrak{m}$ est maximal.

Inversement, si $\mathfrak{m}$ est un idéal maximal, alors $\mathfrak{m} \neq A$, et $A/\mathfrak{m}$ est non trivial. Soit $\overline{a} \in (A/\mathfrak{m}) \setminus \{\overline{0}\}$. Ainsi, $a \notin \mathfrak{m}$. Soit $\mathfrak{a} = aA + \mathfrak{m}$; c'est un idéal de A (en fait, c'est l'idéal engendré par a et les éléments de $\mathfrak{m}$). Alors, $\mathfrak{m} \subset \mathfrak{a}$ et $\mathfrak{m} \neq \mathfrak{a}$, puisque $a \in \mathfrak{a} \setminus \mathfrak{m}$. Par maximalité de $\mathfrak{m}$, on obtient $\mathfrak{a} = A$. En particulier, $\mathfrak{a}$ contient 1, et donc il existe $b \in A$ et $x \in \mathfrak{m}$ tels que $1 = ba + x$. Ainsi, $\overline{1} = \overline{b}\overline{a}$ et $\overline{a}$ est inversible. L'anneau $A/\mathfrak{m}$ est donc bien un corps.

Un corps étant intègre par la proposition **??**, on obtient le dernier point. Ceci achève la démonstration. □

Exemples VIII.2.11.

(1) Les idéaux premiers de $\mathbb{Z}$ sont exactement (0) et les $p\mathbb{Z}$, p premier.
(2) Les idéaux maximaux de $\mathbb{Z}$ sont les $p\mathbb{Z}$, p premier.

VIII.3. Exemple de l'anneau $K[X]$

Dans ce paragraphe, K est un corps.

On commence par rappeler la notion de polynôme irréductible.

Définition VIII.3.1. Un polynôme $P \in K[X]$ est *irréductible* s'il est non constant, et s'il ne peut pas s'écrire comme produit de deux polynômes non constants.

Exemples VIII.3.2.

Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.

Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré et les polynômes de degré 2 de discriminant < 0.

Pour information, on rappelle le théorème suivant.

Théorème VIII.3.3. *Soit K un corps. Alors, tout polynôme $P \in K[X]$ non nul s'écrit de manière unique (à l'ordre des facteurs près) sous la forme*

$$P = c\pi_1^{m_1} \cdots \pi_r^{m_r},$$

où $r \geq 0, c \in K^\times, m_i \geq 1$, et $\pi_1, \ldots, \pi_r$ sont des polynômes irréductibles unitaires deux à deux distincts.

La démonstration du résultat suivant est très similaire à celle de la proposition **??**.

Proposition VIII.3.4. *Soit $P \in K[X]$ un polynôme non nul. Les propriétés suivantes sont équivalentes :*

(1) *$K[X]/(P)$ est intègre*
(2) *$K[X]/(P)$ est un corps*
(3) *P est irréductible.*

Corollaire VIII.3.5. *Soit K un corps. Alors :*

(1) *les idéaux premiers de $K[X]$ sont (0) et les idéaux (π), où π est irréductible (unitaire)*
(2) *les idéaux maximaux de $K[X]$ sont les idéaux (π), où π est irréductible (unitaire).*

L'anneau quotient $K[X]/(P)$ jouit d'une propriété supplémentaire : c'est un K-espace vectoriel pour la loi externe

$$\begin{aligned} K \times K[X]/(P) &\longrightarrow K[X]/(P) \\ (\lambda, \overline{f}) &\longmapsto \lambda \cdot \overline{f} = \overline{\lambda f}. \end{aligned}$$

On peut facilement en calculer une base.

Proposition VIII.3.6. *Soit $P \in K[X]$ un polynôme de degré $d \geq 1$. Alors, toute classe $\overline{f} \in K[X]/(P)$ peut être représentée par un* **unique** *polynôme de degré $\leq d-1$.*

Autrement dit, la famille $(\overline{1}, \overline{X}, \ldots, \overline{X^{d-1}})$ est une K-base de $K[X]/(P)$.

En particulier, $K[X]/(P)$ est de dimension d.

Démonstration. Soit $f \in K[X]$. On peut écrire $f = PQ + R$, où $Q, R \in K[X]$ et R est de degré $\leq d-1$. Mais alors, on a $\overline{f} = \overline{R}$, puisque $PQ \in (P)$. Si maintenant $\overline{f} = \overline{R_1} = \overline{R_2}$, où R_i est de degré $\leq d-1$, alors $\overline{R_1 - R_2} = \overline{0}$, i.e. $R_1 - R_2 \in (P)$. Autrement dit, $R_1 - R_2$ est un multiple de P. Mais $\deg(R_1 - R_2) \leq d-1 < d = \deg(P)$. Ceci n'est donc possible que si $R_1 - R_2 = 0$, et ainsi $R_1 = R_2$.

Le deuxième point de la proposition vient simplement de l'égalité

$$\lambda_0 \cdot \overline{1} + \cdots + \lambda_{d-1} \cdot \overline{X^{d-1}} = \overline{\lambda_0 + \cdots + \lambda_{d-1} X^{d-1}},$$

et du point précédent. Le reste en découle immédiatement. □

La proposition précédente implique en particulier que le morphisme d'anneaux

$$\begin{aligned} K &\longrightarrow K[X]/(P) \\ \lambda &\longmapsto \overline{\lambda} \end{aligned}$$

est injective. Ainsi, $K[X]/(P)$ contient un sous-anneau isomorphe à K.

Si on identifie K à son image dans $K[X]/(P)$, on a alors le phénomène suivant : P possède une racine dans L ! En effet, si $\alpha = \overline{X}$, et si $P = a_d X^d + \cdots + a_1 X + a_0$, on a

$$a_d \alpha^d + \cdots + a_1 \alpha + a_0 = \overline{a_d X^d + \cdots + a_1 X + a_0} = \overline{P} = \overline{0} \in L.$$

En particulier, si l'on part d'un polynôme $P \in K[X]$ **irréductible** de degré $d \geq 1$ (donc qui ne possède pas de racines dans L), l'anneau $L = K[X]/(P)$ possède les propriétés suivantes :

(1) L est un corps qui contient K
(2) L est un K-espace vectoriel de dimension d
(3) P possède une racine dans L.

Exemple VIII.3.7. Prenons $K = \mathbb{R}, P = X^2 + 1$. Le corps L obtenu contient $\mathbb{R}$, est un $\mathbb{R}$-espace vectoriel de dimension 2, et contient un élément α tel que $\alpha^2 + 1 = 0$. On se doute bien que l'on obtient $\mathbb{C}$.

D'ailleurs, c'est un bon exercice que de montrer que le morphisme $\mathbb{R}[X] \longrightarrow \mathbb{C}$ d'évaluation en i induit par passage au quotient un isomorphisme d'anneaux $\mathbb{R}[X]/(X^2 + 1) \simeq \mathbb{C}$ (utiliser par exemple une division euclidienne).

VIII.4. Théorème chinois

Nous allons maintenant démontrer un théorème qui permet parfois de décortiquer un peu la structure d'un anneau quotient. On commence par une définition.

Définition VIII.4.1. Soit A un anneau commutatif. On dit que deux idéaux $\mathfrak{a}$ et $\mathfrak{b}$ sont *comaximaux* si on a $\mathfrak{a} + \mathfrak{b} = A$.

Cela revient à dire qu'il existe $a \in \mathfrak{a}$ et $b \in \mathfrak{b}$ tels que $a + b = 1$.

Exemple VIII.4.2. Deux idéaux $n\mathbb{Z}$ et $m\mathbb{Z}$ de $\mathbb{Z}$ sont comaximaux si, et seulement si, les entiers n et m sont premiers entre eux, puisque d'après la proposition **??**, on a

$$n\mathbb{Z} + m\mathbb{Z} = \mathrm{pgcd}(n, m)\mathbb{Z}.$$

Lemme VIII.4.3. *Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_{n+1}$ des idéaux de A deux à deux comaximaux. Alors, $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$ et $\mathfrak{a}_{n+1}$ sont comaximaux.*

Démonstration. Puisque $\mathfrak{a}_i + \mathfrak{a}_{n+1} = A$ pour tout $i \in [\![1, n]\!]$ par hypothèse, il existe $a_i \in \mathfrak{a}_i$ et $b_1, \ldots, b_n \in \mathfrak{a}_{n+1}$ tels que

$$a_i + b_i = 1 \quad \text{pour tout } i \in [\![1, n]\!].$$

On a donc

$$(a_1 + b_1) \cdots (a_n + b_n) = 1.$$

En développant et en utilisant le fait que $\mathfrak{a}_{n+1}$ est un idéal, on peut donc écrire

$$a_1 \cdots a_n + b = 1,\ b \in \mathfrak{a}_{n+1}.$$

Mais $a_1 \cdots a_n \in \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$, car $a_i \in \mathfrak{a}_i$ et chaque $\mathfrak{a}_i$ est un idéal de A, d'où le résultat. □

Notation. Soit $\mathfrak{a}$ un idéal de A, et soient $x, y \in A$. Si $x - y \in \mathfrak{a}$, on le notera

$$x \equiv y \mod \mathfrak{a}.$$

Si de plus $\mathfrak{a} = (a)$, on le notera aussi

$$x \equiv y\ [a].$$

Proposition VIII.4.4. *Soit A un anneau commutatif, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. Alors, pour tous $x_1, \ldots, x_n \in A$, il existe $x \in A$ tel que*

$$\begin{cases} x & \equiv & x_1 & \mod \mathfrak{a}_1 \\ & \vdots & & \\ x & \equiv & x_n & \mod \mathfrak{a}_n \end{cases}$$

Démonstration. On procède par récurrence sur n. Si $n = 1$, x_1 est une solution du système. Supposons que $n = 2$. Puisque $\mathfrak{a}_1$ et $\mathfrak{a}_2$ sont comaximaux, il existe $a_1 \in \mathfrak{a}_1, a_2 \in \mathfrak{a}_2$ tels que $a_1 + a_2 = 1$. Soient alors $x_1, x_2 \in A$, et posons $x = x_1 a_2 + x_2 a_1$. Alors, on a

$$x = x_1(1 - a_1) + x_2 a_1 = x_1 + a_1(x_2 - x_1),$$

et donc $x \equiv x_1 \mod \mathfrak{a}_1$. De même, on a

$$x = x_1 a_2 + x_2(1 - a_2) = x_2 + a_2(x_1 - x_2),$$

et donc $x \equiv x_2 \mod \mathfrak{a}_2$. Le théorème est donc vrai pour $n = 2$.

Supposons maintenant que le résultat soit vrai pour $n \geq 2$, et montrons le au rang $n + 1$. Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_{n+1}$ des idéaux de A deux à deux comaximaux, et soient $x_1, \ldots, x_{n+1} \in A$. Par hypothèse de récurrence, il existe $y \in A$ tel que

$$y \equiv x_i \mod \mathfrak{a}_i \text{ pour tout } i \in [\![1, n]\!].$$

D'après le lemme précédent, $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$ et $\mathfrak{a}_{n+1}$ sont comaximaux. Par le cas $n = 2$, il existe $x \in A$ tel que

$$x \equiv y \mod \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \text{ et } x \equiv x_{n+1} \mod \mathfrak{a}_{n+1}.$$

Mais alors, on a aussi

$$x \equiv y \equiv x_i \mod \mathfrak{a}_i \text{ pour tout } i \in [\![1, n]\!],$$

et le résultat est vrai au rang $n + 1$. Ceci achève la récurrence et la démonstration. □

On a alors le résultat suivant.

Théorème VIII.4.5 (chinois). *Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. Pour tout $i \in [\![1, n]\!]$, et pour tout $a \in A$, on note $[a]_i$ la classe de a modulo $\mathfrak{a}_i$.*

Alors, le morphisme

$$\begin{aligned} f \colon A &\longrightarrow A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n \\ x &\longmapsto ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$. En particulier, on a un isomorphisme d'anneaux

$$A/\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \simeq A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n.$$

Démonstration. La surjectivité de f est une reformulation de la proposition précédente. Un élément $x \in A$ est dans le noyau de f si, et seulement si, on a $[x]_i = [0]_i$ pour tout $i \in [\![1, n]\!]$, c'est-à-dire si, et seulement si, $x \in \mathfrak{a}_i$ pour tout $i \in [\![1, n]\!]$. Ainsi, $\mathrm{Ker}(f) = \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$. Le dernier point est une simple application du premier théorème d'isomorphisme à f. □

Nous continuons ces considérations en donnant une formulation un peu différente du théorème chinois, mais qui s'avére extrêmement pratique dans certains cas.

On commence par introduire une nouvelle notation.

Notation. Soit $n \geq 0$ un entier. Si $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ sont n idéaux de l'anneau A, on note $\mathfrak{a}_1 \cdots \mathfrak{a}_n$ l'idéal engendré par les éléments de la forme

$$a_1 \cdots a_n, \ a_i \in \mathfrak{a}_i.$$

Lorsque $\mathfrak{a}_1 = \cdots = \mathfrak{a}_n = \mathfrak{a}$, on le note simplement $\mathfrak{a}^n$.

Remarques VIII.4.6.

(1) Lorsque $n = 0$, on obtient l'idéal engendré par le produit vide, à savoir 1, et donc $\mathfrak{a}_1 \cdots \mathfrak{a}_n = A$ si $n = 0$.

(2) le produit d'idéaux est associatif : si $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ sont des idéaux de A, on a

$$(\mathfrak{a}\mathfrak{b})\mathfrak{c} = \mathfrak{a}\mathfrak{b}\mathfrak{c} = \mathfrak{a}(\mathfrak{b}\mathfrak{c}).$$

Il suffit pour s'en convaincre d'écrire explicitement les éléments dans chaque cas.

Exemple VIII.4.7. Si $\mathfrak{a}_i = (\alpha_i)$ pour $i \in [\![1, n]\!]$, on vérifie facilement que

$$\mathfrak{a}_1 \cdots \mathfrak{a}_n = (\alpha_1 \cdots \alpha_n).$$

On a alors le lemme suivant.

Lemme VIII.4.8. *Soit $n \geq 0$, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A. Alors, on a*

$$\mathfrak{a}_1 \cdots \mathfrak{a}_n \subset \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n.$$

Si de plus $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ sont deux à deux comaximaux, alors on a

$$\mathfrak{a}_1 \cdots \mathfrak{a}_n = \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n.$$

Démonstration. Montrons la première partie. Pour cela, il suffit de vérifier que pour tous $a_1 \in \mathfrak{a}_1, \ldots, a_n \in \mathfrak{a}_n$, on a $a_1 \cdots a_n \in \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$, ce qui est clair, car $\mathfrak{a}_i$ est un idéal. Nous allons démontrer la seconde partie par récurrence. Il reste simplement à démontrer l'inclusion non triviale

$$\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \subset \mathfrak{a}_1 \cdots \mathfrak{a}_n.$$

Si $n = 0$ ou 1, il n'y a rien à faire. Montrons-le pour $n = 2$. Les idéaux $\mathfrak{a}_1$ et $\mathfrak{a}_2$ étant comaximaux, il existe $a_1 \in \mathfrak{a}_1$ et $a_2 \in \mathfrak{a}_2$ tels que $a_1 + a_2 = 1$.

Soit maintenant $x \in \mathfrak{a}_1 \cap \mathfrak{a}_2$. On a alors

$$x = x\ 1 = xa_1 + xa_2 = a_1 x + xa_2.$$

Or, $a_1 x, xa_2 \in \mathfrak{a}_1\mathfrak{a}_2$ par définition, puisque $x \in \mathfrak{a}_1 \cap \mathfrak{a}_2$. Mais alors, $x \in \mathfrak{a}_1\mathfrak{a}_2$, et on a fini.

Supposons maintenant avoir montré le résultat pour un produit de n idéaux, avec $n \geq 2$, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_{n+1}$ $n + 1$ idéaux deux à deux comaximaux. D'après le

lemme **??**, les idéaux $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$ et $\mathfrak{a}_{n+1}$ sont comaximaux. En utilisant le cas $n = 2$, puis l'hypothèse de récurrence, on obtient donc

$$\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \cap \mathfrak{a}_{n+1} = (\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n)\mathfrak{a}_{n+1} = (\mathfrak{a}_1 \cdots \mathfrak{a}_n)\mathfrak{a}_{n+1}.$$

On achève alors la récurrence en utilisant la remarque **??** (2). □

On obtient alors une nouvelle formulation du théorème chinois.

Théorème VIII.4.9. *Soit A un anneau, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. Pour tout $i \in [\![1, n]\!]$, et tout $a \in A$, on note $[a]_i$ la classe de a modulo $\mathfrak{a}_i$.*

Alors, le morphisme

$$\begin{aligned} f\colon A &\longrightarrow A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n \\ x &\longmapsto ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $\mathfrak{a}_1 \cdots \mathfrak{a}_n$. En particulier, on a un isomorphisme d'anneaux

$$A/\mathfrak{a}_1 \cdots \mathfrak{a}_n \simeq A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n.$$

Lorsque $\mathfrak{a} = (a)$ et $\mathfrak{b} = (b)$, les idéaux $\mathfrak{a}$ et $\mathfrak{b}$ sont comaximaux si, et seulement si, il existe $u, v \in A$ tels que $ua + vb = 1$, c'est-à-dire s'ils sont étrangers.

Dans le cas où tous les idéaux $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ sont principaux (par exemple lorsque A est un anneau principal), le théorème chinois se retraduit alors comme suit.

Théorème VIII.4.10. *Soit A un anneau commutatif, et soient $a_1, \ldots, a_n \in A$ deux à deux étrangers.*

Alors, le morphisme

$$\begin{aligned} f\colon A &\longrightarrow A/(a_1) \times \cdots \times A/(a_n) \\ x &\longmapsto ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $(a_1 \cdots a_n)$. En particulier, on a un isomorphisme d'anneaux

$$A/(a_1 \cdots a_n) \simeq A/(a_1) \times \cdots \times A/(a_n).$$

Remarque VIII.4.11. Si A est un anneau principal, deux éléments sont étrangers si et seulement s'ils sont premiers entre eux (corollaire **??**). En particulier, si A est principal, le théorème précédent s'applique à des éléments premiers entre eux deux à deux.

Exemples VIII.4.12.

(1) Si n et m sont deux entiers premiers entre eux, on a un isomorphisme d'anneaux
$$\mathbb{Z}/nm\mathbb{Z} \simeq \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

(2) Si $n \geq 2$, et si $n = p_1^{m_1} \cdots p_r^{m_r}$ est la décomposition de n en facteurs premiers, on a un isomorphisme d'anneaux
$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}.$$

(3) Si P et Q sont deux polynômes de $K[X]$ premiers entre eux (où K est un corps), on a un isomorphisme d'anneaux
$$K[X]/(PQ) \simeq K[X]/(P) \times K[X]/(Q).$$

(4) Si $P \in K[X]$ est non constant, et si $P = c\pi_1^{m_1} \cdots \pi_r^{m_r}$ est la décomposition de P en facteurs irréductibles unitaires, on a un isomorphisme d'anneaux
$$K[X]/(P) \simeq K[X]/(P_1^{m_1}) \times \cdots \times K[X]/(P_r^{m_r}).$$

Troisième partie

Algèbre linéaire : réduisons en miettes !

Chapitre IX

Dualité en dimension finie

Soit E un espace vectoriel réel de dimension finie $n \geq 1$, et soit $\langle\ ,\rangle$ un produit scalaire sur E.

On sait que pour tout sous-espace F de E, $F^\perp$ est de dimension $n-k$, et de plus $(F^\perp)^\perp = F$, si bien que l'on obtient une bijection entre les sous-espaces de dimension k de E, et les sous-espaces de dimension $n-k$.

De plus, si $u \in \mathscr{L}(E)$, il existe un unique endomorphisme adjoint $u^* \in \mathscr{L}(E)$ tel que

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \quad \text{pour tous } x, y \in E.$$

De plus, on a $(u^*)^* = u$, ainsi que les égalités $\mathrm{Ker}(u^*) = (\mathrm{im}(u))^\perp$ et $\mathrm{im}(u^*) = (\mathrm{Ker}(u))^\perp$.

Le but de ce chapitre est de généraliser ces résultats dans un contexte plus global.

IX.1. Formes linéaires et bilinéaires, espace dual

Définition IX.1.1. Soit E un K-espace vectoriel.

Une *forme linéaire* sur E est un élément de $\mathscr{L}(E, K)$, autrement dit une application linéaire $\varphi : E \longrightarrow K$.

Le *dual* de E est le K-espace vectoriel $\mathscr{L}(E, K)$. On le note E^*.

Exemples IX.1.2. (1) Pour tout $n \geq 0$, l'application

$$\begin{aligned} \mathbb{R}[X] &\longrightarrow \mathbb{R} \\ P &\longmapsto \frac{P^{(n)}(0)}{n!} \end{aligned}$$

est une forme linéaire.

(2) L'application

$$\begin{aligned} C^0([0,1], \mathbb{R}) &\longrightarrow \mathbb{R} \\ f &\longmapsto \int_0^1 f(t)\mathrm{d}t \end{aligned}$$

est une forme linéaire.

(3) Soit E un K-espace vectoriel, et soit $(e_1, \ldots, e_n)$ une base de E. Alors il existe une unique forme linéaire $e_i^* : E \longrightarrow K$ telle que

$$e_i^*(e_k) = \delta_{ik} \quad \text{pour tout } k \in [\![1, n]\!].$$

Elle est simplement définie par

$$\begin{aligned} e_i^* : \quad E &\longrightarrow K \\ \sum_{k=1}^n x_k e_k &\longmapsto x_i. \end{aligned}$$

On a alors

$$x = \sum_{i=1}^{n} e_i^*(x) e_i \quad \text{pour tout } x \in E,$$

par définition.

Définition IX.1.3. Soit E un K-espace vectoriel, et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Les formes linéaires $e_i^*, i \in [\![1, n]\!]$ sont appelées les *formes coordonnées* par rapport à la base $\mathbf{e}$.

Ces formes coordonnées jouent un rôle important pour l'étude des formes linéaires en dimension finie, au vu de la proposition suivante.

Proposition IX.1.4. *Soit E un K-espace vectoriel de dimension finie $n \geq 0$, et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Alors, la famille $\mathbf{e}^* = (e_1^*, \ldots, e_n^*)$ est une base de E^*.*

De plus, pour tout $x \in E$, et tout $\varphi \in E^*$, on a

$$x = \sum_{i=1}^{n} e_i^*(x) e_i \quad \textit{et} \quad \varphi = \sum_{i=1}^{n} \varphi(e_i) e_i^*.$$

Enfin, $\dim_K(E^*) = \dim_K(E)$.

Démonstration. Montrons que $\mathbf{e}^*$ est libre. Supposons que l'on ait

$$\sum_{i=1}^{n} \lambda_i e_i^* = 0_{E^*}, \ \lambda_i \in K.$$

Soit $k \in [\![1, n]\!]$. En appliquant e_k à cette égalité, par définition des formes coordonnées, on obtient $\lambda_k = 0$, d'où la liberté requise. De plus,

$$\dim_K(E^*) = \dim_K(\mathscr{L}(E, K)) = \dim_K(E) \dim_K(K) = \dim_K(E) = n.$$

Dans ce cas, comme $\mathbf{e}^*$ est libre de cardinal n, c'est une base de E^*.

Soit $\varphi \in E^*$. Pour tout $k \in [\![1, n]\!]$, on a

$$\Big(\sum_{i=1}^{n} \varphi(e_i) e_i^*\Big)(e_k) = \sum_{i=1}^{n} \varphi(e_i) e_i^*(e_k) = \varphi(e_k).$$

Ainsi, φ et $\displaystyle\sum_{i=1}^{n} \varphi(e_i) e_i^*$ coïncident sur une base, et elle sont donc égales (Notons que cela démontre aussi que la famille $\mathbf{e}^*$ est génératrice, sans utiliser le fait que l'on connait la dimension de E^*). □

On rappelle maintenant la définition d'une forme bilinéaire.

Définition IX.1.5. Soient E, F deux K-espaces vectoriels. Une *forme bilinéaire* est une application

$$\begin{aligned} b \colon E \times F &\longrightarrow K \\ (x, y) &\longmapsto b(x, y) \end{aligned}$$

telle que pour tout $x \in E, y \in F$ les applications

$$\begin{aligned} b(x, _) \colon F &\longrightarrow K \\ y &\longmapsto b(x, y) \end{aligned} \quad \text{et} \quad \begin{aligned} b(_, y) \colon E &\longrightarrow K \\ x &\longmapsto b(x, y) \end{aligned}$$

soient linéaires, c'est-à-dire une application linéaire en chaque argument.

Exemples IX.1.6. (1) L'application

$$\begin{aligned} \mathbb{R}[X] \times \mathbb{R}[X] &\longrightarrow \mathbb{R} \\ (P, Q) &\longmapsto \int_0^1 P(t)Q(t)e^{-t^2}\mathrm{d}t \end{aligned}$$

est bilinéaire. C'est même un produit scalaire.

(2) L'application

$$\begin{aligned} \langle\ ,\ \rangle_E \colon E \times E^* &\longrightarrow K \\ (x, \varphi) &\longmapsto \langle x, \varphi\rangle_E = \varphi(x) \end{aligned}$$

est bilinéaire. Elle est appelée *crochet de dualité.*

La démonstration du lemme suivant est facile et laissée en exercice au lecteur.

Lemme IX.1.7. *Soit $b : E \times F \longrightarrow K$ une application bilinéaire. Alors, pour tout $x \in E$, et tout $y \in F$, les applications $b(x,_\,)$ et $b(_\,, y)$ sont des formes linéaires, et les applications*

$$\begin{aligned} {}_*b \colon E &\longrightarrow F^* \\ x &\longmapsto b(x,_\,) \end{aligned} \quad \textit{et} \quad \begin{aligned} b_* \colon F &\longrightarrow E^* \\ y &\longmapsto b(_\,, y) \end{aligned}$$

sont linéaires.

Exemple IX.1.8. Si $F = E^*$ et $b : E \times E^* \longrightarrow K$ est le crochet de dualité, on a $b_* = \mathrm{Id}_{E^*}$ et ${}_*b$ est l'application $\chi : E \longrightarrow E^{**}$ qui à $x \in E$ associe l'évaluation en x, i.e.

$$\chi(x)(\varphi) = \varphi(x) \ \text{ pour tout } \varphi \in E^*.$$

Ici, E^{**} désigne l'espace $(E^*)^*$.

De même que l'on a défini une notion d'orthogonalité pour les produits scalaires ou pour le crochet de dualité, on a une notion d'orthogonalité par rapport à une forme bilinéaire $b : E \times F \longrightarrow K$.

Définition IX.1.9. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire.

Si $x \in E$ et $y \in F$, on dit que x et y sont *orthogonaux* si $b(x, y) = 0$.

On le note parfois $x \perp_b y$ ou $x \perp y$, lorsque il n'y a pas d'ambiguïté sur le choix de b.

Si V un sous-espace de E, on définit *l'orthogonal* de V par

$$V^\perp = \{y \in F \mid b(x, y) = 0 \ \text{ pour tout } x \in V\}.$$

C'est un sous-espace vectoriel de F.

Si W un sous-espace de F, *l'orthogonal* de W est

$${}^\perp W = \{x \in E \mid b(x, y) = 0 \ \text{ pour tout } y \in W\}.$$

C'est un sous-espace vectoriel de E.

Attention ! La formulation « x et y orthogonaux » semble supposer que la relation d'orthogonalité est symétrique. Il n'en est rien, puisqu'a priori les vecteurs x et y ne vivent pas dans le même espace vectoriel. Cela peut être très piégeux lorsque $E = F$. En effet, si $b : E \times F \longrightarrow K$ est une application bilinéaire quelconque, en général, $x \perp_b y$ n'implique **PAS** que $y \perp_b x$.

Exemples IX.1.10.

(1) Si $F = E^*$ et $b = \langle\ ,\ \rangle_E$ est le crochet de dualité, on a
$$V^\perp = \{\varphi \in E^* \mid \varphi_{|V} = 0\}$$
pour tout sous-espace V de E et
$${}^\perp W = \bigcap_{\varphi \in W} \mathrm{Ker}(\varphi)$$
pour tout sous-espace W de E^*.

(2) Lorsque $E = F$ et b est symétrique (par exemple si E est un espace vectoriel réel et b est un produit scalaire), on a $b_* = {}_*b$, et les deux définitions d'orthogonal d'un sous-espace coïncident. On retrouver alors la définition classique d'orthogonalité, et d'orthogonal d'un sous-espace.

Remarque IX.1.11. Si $b : E \times F \longrightarrow K$ est une application bilinéaire, les définitions montrent que l'on a
$$V \subset {}^\perp(V^\perp) \ \text{ et } \ W \subset ({}^\perp W)^\perp$$
pour tout sous-espace V de E et tout sous-espace W de F.

En effet, soit $x \in V$. Pour tout $y \in V^\perp$, on a $b(x, y) = 0$, et par conséquent $x \in {}^\perp(V^\perp)$. L'autre inclusion se montre de même.

Dans ce qui suit, on considère deux K-espaces vectoriels E et F, et une application K-bilinéaire
$$b : E \times F \longrightarrow K.$$

Sauf mention expresse du contraire, E et F seront non nuls.

Les deux notions d'orthogonal permettent d'associer un sous-espace de F à un sous-espace de E et réciproquement. Il est donc naturel de se demander si

les applications
$$\begin{array}{ccc} V & \longmapsto & V^\perp \\ {}^\perp W & \longleftarrow & W \end{array}$$
sont inverses l'une de l'autre, établissant ainsi une correspondance bijective entre les sous-espaces de E et les sous-espaces de F.

Lorsque c'est le cas, on a
$${}^\perp(V^\perp) = V \ \text{ et } \ ({}^\perp W)^\perp = W$$
pour tout sous-espace V de E et tout sous-espace W de F. En particulier, on doit avoir
$${}^\perp(\{0_E\}^\perp) = \{0_E\} \ \text{ et } \ ({}^\perp\{0_F\})^\perp = \{0_F\}.$$

Puisque $b(0_E, y) = 0$ pour tout $y \in F$, on a $\{0_E\}^\perp = F$, et ainsi
$$\begin{array}{rcl} {}^\perp(\{0_E\}^\perp) & = & {}^\perp F \\ & = & \{x \in E \mid b(x, y) = 0 \ \text{ pour tout } y \in F\} \\ & = & \{x \in E \mid b(x, _\,) = 0\} \\ & = & \mathrm{Ker}({}_*b). \end{array}$$

De même, on a
$$({}^\perp\{0_F\})^\perp = \mathrm{Ker}(b_*).$$

Ainsi, une condition nécessaire pour avoir cette correspondance[1] est l'injectivité des deux applications ${}_*b$ et b_*. Ceci conduit à la définition suivante.

1. Via les deux applications décrites ci-dessus.

Définition IX.1.12. On dit qu'une forme bilinéaire $b : E \times F \longrightarrow K$ est *non dégénérée* si les applications ${}_*b : E \longrightarrow F^*$ et $b_* : F \longrightarrow E^*$ sont injectives. Autrement dit, $b : E \times F \longrightarrow K$ est non dégénérée si l'on a :

(1) pour tout $x \in E$, $b(x, y) = 0$ pour tout $y \in F \Longrightarrow x = 0_E$;

(2) pour tout $y \in F$, $b(x, y) = 0$ pour tout $x \in E \Longrightarrow y = 0_F$.

Tout produit scalaire sur un espace vectoriel réel est bien entendu une forme bilinéaire non dégénérée. Voici un autre exemple.

Exemple IX.1.13. Soit E un K-espace vectoriel de dimension finie $n \geq 1$. Alors, le crochet de dualité est une forme bilinéaire non dégénérée.

En effet, comme on l'a déjà remarqué plus haut, l'application $(\langle\ ,\ \rangle_E)_*$ est simplement Id_{E^*}, qui est injective.

Montrons maintenant que $\chi : E \longrightarrow E^{**}$ est injective. Il faut montrer que si $x \in E$ vérifie $\varphi(x) = 0$ pour tout $\varphi \in E$, alors $x = 0$.

Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Écrivons $x = \displaystyle\sum_{i=1}^{n} x_i e_i$. Par hypothèse, on a $e_i^*(x) = 0 = x_i$ pour tout $i \in [\![1, n]\!]$, d'où $x = 0$.

Remarque IX.1.14. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée. Si un des deux espaces E ou F est de dimension finie, alors il en est de même de l'autre. Dans ce cas, on a $\dim_K(E) = \dim_K(F)$, et les deux applications linéaires ${}_*b$ et b_* sont des isomorphismes.

En effet, supposons par exemple que E soit de dimension finie sur K. Alors, E^* est aussi de dimension finie. Puisque $b_* : F \longrightarrow E^*$ est injective, F s'identifie à un sous-espace de E^* et est donc de dimension finie. De plus, on a

$$\dim_K(F) \leq \dim_K(E^*) = \dim_K(E).$$

Comme ${}_*b : E \longrightarrow F^*$ est injective, on a

$$\dim_K(E) \leq \dim_K(F^*) = \dim_K(F).$$

Ainsi, $\dim_K(E) = \dim_K(F)$. Les K-espaces vectoriels E, E^*, F et F^* ont alors même dimension, ce qui implique que ${}_*b$ et b_* sont des isomorphismes (puisqu'elle sont injectives).

On a alors le résultat attendu.

Proposition IX.1.15. *Soient E, F deux K-espaces vectoriels de dimension $n \geq 1$, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée. Alors pour tout sous-espace V de E, et tout sous-espace W de F, on a*

$$\dim_K(V^\perp) = n - \dim_K(V) \ \textit{et}\ \dim_K({}^\perp W) = n - \dim_K(W).$$

De plus, on a ${}^\perp(V^\perp) = V$ et $({}^\perp W)^\perp = W$.

En particulier, pour tout $k \in [\![0, n]\!]$, on a une correspondance bijective entre l'ensemble $\mathcal{V}_k(E)$ des sous-espaces de E de dimension k de E et l'ensemble $mathcalV_{n-k}(F)$ des sous-espaces de F de dimension $n - k$. Elle est donnée par

$$\begin{array}{ccc} \mathcal{V}_k(E) & & \mathcal{V}_{n-k}(F) \\ V & \longmapsto & V^\perp \\ {}^\perp W & \longleftarrow & W. \end{array}$$

Démonstration. Considérons l'application linéaire

$$\begin{aligned} f\colon F &\longrightarrow V^* \\ y &\longmapsto b(_, y)_{|V}. \end{aligned}$$

On a $\mathrm{Ker}(f) = V^\perp$. De plus, f est surjective. En effet, soit $\psi \in V^*$, et soit G un supplémentaire de V dans E. Posons

$$\begin{aligned} \varphi\colon \quad E \quad &\longrightarrow K \\ x = v + g &\longmapsto \psi(v). \end{aligned}$$

Alors, $\varphi \in E^*$ et $\varphi_{|V} = \psi$. Or, b étant non dégénérée, l'application $b_* : F \longrightarrow E^*$ est un isomorphisme. Il existe donc $y \in F$ tel que $b(_, y) = \varphi$. Ainsi $b(_, y)_{|V} = \varphi_{|V} = \psi$, et f est surjective. Le théorème du rang donne alors

$$\dim_K(F) = n = \dim_K(V^\perp) + \dim_K(V^*) = \dim_K(V^\perp) + \dim_K(V).$$

On démontre de même l'autre égalité.

On en déduit facilement que l'on a

$$\dim_K({}^\perp(V^\perp)) = \dim_K(V) \text{ et } \dim_K(({}^\perp W)^\perp) = \dim_K(W).$$

Or, on a les inclusions

$$V \subset {}^\perp(V^\perp) \text{ et } W \subset ({}^\perp W)^\perp.$$

L'égalité des dimensions implique alors les égalités voulues. Le dernier point n'est qu'une reformulation de ce qui précède. □

Lorsque b est un produit scalaire sur un espace vectoriel réel E de dimension finie, on retrouve les propriétés de l'orthogonal d'un sous-espace. Le cas du crochet de dualité donne aussi des choses intéressantes.

Exemple IX.1.16. Soit E un K-espace vectoriel de dimension $n \geq 1$. Soit V un sous-espace de E, et soit W un sous-espace de E^*.

L'égalité $({}^\perp W)^\perp = W$ nous dit que W est exactement l'ensemble des formes linéaires $\varphi \in E^*$ telles que $\bigcap_{\psi \in W} \mathrm{Ker}(\psi) \subset \mathrm{Ker}(\varphi)$.

Pour retraduire ceci de manière plus frappante, remarquons que si $(\varphi_1, \ldots, \varphi_{n-k})$ engendre W, alors

$${}^\perp W = \bigcap_{i=1}^{n-k} \mathrm{Ker}(\varphi_i).$$

On obtient alors le résultat suivant : pour toute forme linéaire $\varphi \in E^*$, on a $\bigcap_{i=1}^{n-k} \mathrm{Ker}(\varphi_i) \subset \mathrm{Ker}(\varphi)$ si et seulement si φ est une combinaison linéaire de $\varphi_1, \ldots, \varphi_{n-k}$.

On peut d'ailleurs démontrer ce résultat même si E n'est pas de dimension finie, mais c'est un peu plus délicat.

L'égalité ${}^\perp(V^\perp) = V$ montre que l'intersection des noyaux de toutes les formes linéaires s'annulant sur V est V (ce n'est pas tout à fait évident, quand on y réfléchit).

D'autre part, si V est de dimension k, et si $(\varphi_1, \ldots, \varphi_{n-k})$ est une base de $V^\perp$, on obtient

$$V = \bigcap_{i=1}^{n-k} \mathrm{Ker}(\varphi_i).$$

Ainsi, tout sous-espace de dimension k est l'intersection des noyaux de $n-k$ formes linéaires indépendantes.

Lorsque $E = K^n$, cela revient à dire que tout sous-espace de dimension k est solution d'un système linéaire de $n-k$ équations (indépendantes) à n inconnues.

Attention ! Le résultat précédent ne fonctionne plus en dimension infinie, comme le montre l'exemple suivant.

Exemple IX.1.17. Soit E l'ensemble des suites réelles $u = (u_n)_{n\geq 1}$ bornées, et soit V le sous-espace des suites réelles nulles à partir d'un certain rang. Pour $u, v \in E$, on pose

$$b(u,v) = \sum_{n\geq 1} \frac{u_n v_n}{n^2}.$$

Alors, b est bilinéaire symétrique, et non dégénérée, car définie positive. En effet, si $u \in E \setminus \{0\}$, et si $k \geq 1$ est un entier tel que $u_k \neq 0$, alors $b(u,u) \geq \dfrac{u_k}{k^2} > 0$.

D'autre part, pour tout $k \geq 1$, la suite $e_k = (\delta_{kn})_{n\geq 1}$ est un élément de V. Si $u \in V^\perp$, on a en particulier

$$b(u, e_n) = \frac{u_n}{n^2} = 0 \quad \text{pour tout } n \geq 1,$$

d'où $V^\perp = \{0\}$. Mais alors, ${}^\perp(V^\perp) = E \neq V$.

IX.2. Bases duales.

Dans ce qui suit, E et F seront des K-espaces vectoriels non nuls de dimension finie, et $b : E \times F \longrightarrow K$ sera une forme bilinéaire non dégénérée. Le but de ce court paragraphe est d'établir une correspondance entre les bases de E et celles de F.

On commence par étudier à quelles conditions b est non dégénérée. Pour cela, on introduit la notion de matrice représentative d'une forme bilinéaire.

Définition IX.2.1. Soient E, F deux espaces vectoriels de dimension finie, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire (pas nécessairement non dégénérée !). Soient $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{f} = (f_1, \ldots, f_m)$ des bases de E et F respectivement. La *matrice représentative* de b dans les bases $\mathbf{e}$ et $\mathbf{f}$ est la matrice

$$\mathrm{Mat}(b; \mathbf{e}, \mathbf{f}) = (b(e_i, f_j))_{i,j}.$$

Lemme IX.2.2. *Soient E, F deux espaces vectoriels de dimension finie, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire.*

Soient $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{f} = (f_1, \ldots, f_m)$ des bases de E et F respectivement, et soit $B = \mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$. Alors, on a

$$b(x,y) = [x]_{\mathbf{e}}^t B [y]_{\mathbf{f}}.$$

Démonstration. Écrivons $x = \displaystyle\sum_{i=1}^n x_i e_i$ et $y = \displaystyle\sum_{j=1}^m y_j f_j$. Par bilinéarité, on a

$$b(x,y) = b\Big(\sum_{i=1}^n x_i e_i, \sum_{j=1}^m y_j f_j\Big) = \sum_{i,j} x_i b(e_i, f_j) y_j = [x]_{\mathbf{e}}^t B [y]_{\mathbf{f}}.$$

Les détails sont laissés au lecteur. □

Proposition IX.2.3. *Soient E, F deux espaces vectoriels de dimension finie, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire.*

Soient $\mathbf{e} = (e_1, \dots, e_n)$ *et* $\mathbf{f} = (f_1, \dots, f_m)$ *des bases de* E *et* F *respectivement. Alors, on a*

$$\mathrm{Mat}(b; \mathbf{e}, \mathbf{f}) = \mathrm{Mat}(b_*; \mathbf{f}, \mathbf{e}^*) = \mathrm{Mat}({}_*b; \mathbf{e}, \mathbf{f}^*)^t.$$

En particulier, la forme bilinéaire b *est non dégénérée si, et seulement si, sa matrice représentative dans des bases arbitraires de* E *et* F *est inversible.*

Démonstration. Pour tout $j \in [\![1, m]\!]$, on a

$$b_*(f_j) = b(_, f_j) = \sum_{i=1}^{n} b(e_i, f_j) e_i^*$$

d'après la proposition **??**. De même pour tout $j \in [\![1, n]\!]$, on a

$${}_*b(e_j) = b(e_j, _\,) = \sum_{i=1}^{m} b(e_j, f_i) f_i^*,$$

d'où la première assertion. Si b est non dégénérée, on a vu que b_* était un isomorphisme. Sa matrice représentative est donc inversible et donc le point précédent montre que $\mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$ est inversible pour toutes bases $\mathbf{e}$ et $\mathbf{f}$. Inversement, supposons que $\mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$ soit inversible. Les égalités précédentes montrent alors que les matrices $\mathrm{Mat}(b_*; \mathbf{f}, \mathbf{e}^*)$ et $\mathrm{Mat}({}_*b; \mathbf{e}, \mathbf{f}^*)$ sont inversibles, et donc que ${}_*b$ et b_* sont des isomorphismes. En particulier, la forme bilinéaire $b : E \times F \longrightarrow K$ est non dégénérée. □

Nous allons maintenant définir la notion de bases duales par rapport à b.

Définition IX.2.4. Soient E et F deux espaces vectoriels de même dimension finie $n \geq 1$ sur K, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire.

Deux bases $\mathbf{e} = (e_1, \dots, e_n)$ et $\mathbf{f} = (f_1, \dots, f_n)$ de E et F respectivement sont dites *duales* par rapport à $b : E \times F \longrightarrow K$ si l'on a

$$b(e_i, f_j) = \delta_{ij} \text{ pour tous } i, j \in [\![1, n]\!].$$

Autrement dit, $\mathbf{e}$ et $\mathbf{f}$ sont duales par rapport à b si, et seulement si, on a

$$\mathrm{Mat}(b; \mathbf{e}, \mathbf{f}) = I_n \in \mathrm{M}_n(K).$$

Remarque IX.2.5. La proposition **??** implique en particulier que s'il existe deux bases $\mathbf{e}$ et $\mathbf{f}$ duales par rapport à b, alors b est nécessairement non dégénérée.

Exemple IX.2.6. Soit E un K-espace vectoriel de dimension finie, et soit $\mathbf{e}$ une base de E. Alors, les bases $\mathbf{e}$ et $\mathbf{e}^*$ sont duales par rapport au crochet de dualité.

Le lemme suivant explique l'intérêt de cette nouvelle notion.

Lemme IX.2.7. *Si* $\mathbf{e}$ *et* $\mathbf{f}$ *sont deux bases duales par rapport à* b*, alors pour tous* $x \in E, y \in F$*, on a*

$$x = b(x, f_1)e_1 + \cdots + b(x, f_n)e_n$$

et

$$y = b(e_1, y)f_1 + \cdots + b(e_n, y)f_n.$$

Démonstration. Écrivons $x = \sum_{i=1}^{n} x_i e_i$. Pour tout $j \in [\![1, n]\!]$, on a

$$b(x, f_j) = \sum_{i=1}^{n} x_i b(e_i, f_j) = x_j.$$

L'autre égalité se montre de la même manière. □

Exemples IX.2.8.

(1) Soit E un K-espace vectoriel de dimension finie, et soit $\mathbf{e}$ une base de E. En appliquant ce résultat au crochet de dualité avec $\mathbf{f} = \mathbf{e}^*$, on retrouve les formules énoncées dans la proposition **??**.

(2) Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien, et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Alors, $\mathbf{e}$ est orthonormée si, et seulement si, les bases $\mathbf{e}$ et $\mathbf{e}$ sont duales par rapport au produit scalaire. Si $\mathbf{e}$ est orthonormée, le lemme précédent redonne alors l'égalité classique

$$x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j \quad \text{pour tout } x \in E.$$

On va maintenant montrer le résultat principal de ce paragraphe.

Théorème IX.2.9. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée, où E et F sont de même dimension finie. Alors, pour toute base $\mathbf{e}$ de E, il existe une unique base $\mathbf{e}^*$ de F telle que les bases $\mathbf{e}$ et $\mathbf{e}^*$ soient duales par rapport à b. De même, pour toute base $\mathbf{f}$ de F, il existe une unique base ${}^*\mathbf{f}$ de E telle que les bases ${}^*\mathbf{f}$ et $\mathbf{f}$ soient duales par rapport à b. De plus, on a*

$${}^*(\mathbf{e}^*) = \mathbf{e} \quad \textit{et} \quad ({}^*\mathbf{f})^* = \mathbf{f}.$$

Autrement dit, on a une correspondance bijective entre l'ensemble $\mathscr{B}(E)$ des bases de E et l'ensemble $\mathscr{B}(F)$ des bases de F, donnée par

$$\begin{array}{ccc} \mathscr{B}(E) & & \mathscr{B}(F) \\ \mathbf{e} & \longmapsto & \mathbf{e}^* \\ {}^*\mathbf{f} & \longleftarrow & \mathbf{f} \end{array}$$

Démonstration. Soit $\mathbf{e}$ une base de E, et soit $\mathbf{f}$ une base quelconque de F. Puisque b est non dégénérée, la matrice $B = \mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$ est inversible par la proposition **??**.

Écrivons $B^{-1} = (m_{ij})$ et pour $j \in [\![1, n]\!]$, posons

$$e_j^* = \sum_{k=1}^{n} m_{kj} f_k.$$

Puisque B^{-1} est inversible, la famille $\mathbf{e}^*$ est une base de F. De plus, pour tous $i, j \in [\![1, n]\!]$, on a

$$b(e_i, e_j^*) = b(e_i, \sum_{k=1}^{n} m_{kj} f_k) = \sum_{k=1}^{n} b(e_i, f_k) m_{kj} = \delta_{ij},$$

la dernière égalité provenant du fait que $BB^{-1} = I_n$. Ainsi, $\mathbf{e}$ et $\mathbf{e}^*$ sont duales.

Supposons que $\mathbf{f^{(1)}}$ et $\mathbf{f^{(2)}}$ soient deux bases de F telles que $\mathbf{e}$ et $\mathbf{f^{(k)}}$ soient duales pour $k = 1, 2$. Alors, pour tous $i, j \in [\![1, n]\!]$, on a

$$b(e_i, f_j^{(1)} - f_j^{(2)}) = b(e_i, f_j^{(1)}) - b(e_i, f_j^{(2)}) = \delta_{ij} - \delta_{ij} = 0.$$

Puisque b est non dégénérée, on en déduit que $f_j^{(1)} - f_j^{(2)} = 0$ pour tout entier $j \in [\![1, n]\!]$ et donc $\mathbf{f^{(1)}} = \mathbf{f^{(2)}}$.

Soit maintenant $\mathbf{f}$ une base de F, et soit $\mathbf{e}$ une base quelconque de E. Puisque b est non dégénérée, la matrice $B = \mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$ est inversible par la proposition **??**.

Écrivons $B^{-1} = (m_{ij})$ et pour $j \in [\![1, n]\!]$, posons cette fois

$$^*f_j = \sum_{k=1}^{n} m_{jk} e_k.$$

Puisque $(B^{-1})^t$ est inversible, la famille $^*\mathbf{f}$ est une base de E. De plus, pour tous $i, j \in [\![1, n]\!]$, on a

$$b(^*f_i, f_j) = b(\sum_{k=1}^{n} m_{ik} e_k, f_j) = \sum_{k=1}^{n} m_{ik} b(e_k, f_j) = \delta_{ij},$$

la dernière égalité provenant du fait que $B^{-1}B = I_n$. Ainsi, $^*\mathbf{f}$ et $\mathbf{f}$ sont duales. L'unicité d'une base duale à $\mathbf{f}$ se démontre comme précédemment.

Montrons le dernier point. Par définition $^*(\mathbf{e}^*)$ est l'unique base de E telle que $^*(\mathbf{e}^*)$ et $\mathbf{e}^*$ soient duales. Mais par définition, $\mathbf{e}$ et $\mathbf{e}^*$ sont duales, d'où l'égalité désirée. La deuxième se démontre de manière analogue. □

Remarques IX.2.10.

(1) La notation $\mathbf{e}^*$ est un peu dangereuse. Si b n'est pas le crochet de dualité, $\mathbf{e}^*$ est une base de F et non pas de E^*, et en particulier n'est **pas** la base des formes coordonnées ! ! !

(2) La démonstration précédente fournit un moyen de calculer explicitement les bases duales.

   (i) Si $\mathbf{e}$ est une base de E, on choisit une base arbitraire $\mathbf{f}$ de F, et on pose $B = \mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$. Alors, pour tout entier $j \in [\![1, n]\!]$, $[e_j^*]_{\mathbf{f}}$ est la j-ième colonne de B^{-1}.

   (ii) Si $\mathbf{f}$ est une base de F, on choisit une base arbitraire $\mathbf{e}$ de E, et on pose $B = \mathrm{Mat}(b; \mathbf{e}, \mathbf{f})$. Alors, pour tout entier $j \in [\![1, n]\!]$, $[^*f_j]_{\mathbf{e}}$ est la j-ième colonne de $(B^{-1})^t$.

Définition IX.2.11. Si $\mathbf{e}$ est une base de E, la base $\mathbf{e}^*$ est appelée la *base duale* de $\mathbf{e}$ (par rapport à b). De même, si $\mathbf{f}$ est une base de F, la base $^*\mathbf{f}$ est appelée la *base duale* de $\mathbf{f}$ (par rapport à b).

Exemples IX.2.12.

(1) Si $\mathbf{e}$ est une base de E, la base duale par rapport au crochet de dualité associée est la base des formes coordonnées $\mathbf{e}^*$, d'où la notation choisie.

(2) Soit $E = \mathbb{R}_n[X]$ l'espace vectoriel réel des polynômes de degré inférieur ou égal à n, et soit $\mathbf{e} = (1, X, \ldots, X^n)$. Pour tout entier $j \in [\![0, n]\!]$, posons

$$\begin{aligned} f_j \colon E &\longrightarrow \mathbb{R} \\ P &\longmapsto \tfrac{P^{(j)}(0)}{j!}. \end{aligned}$$

Alors, $\mathbf{f} = (f_0, \ldots, f_n)$ est la base duale de $\mathbf{e}$ par rapport au crochet de dualité.

(3) Soit $E = \mathbb{R}_n[X]$, et soient $a_0, \ldots, a_n \in \mathbb{R}$ des réels **distincts**. Pour tout entier $j \in\in [\![1, n]\!]$, soit

$$\begin{aligned} f_j \colon E &\longrightarrow \mathbb{R} \\ P &\longmapsto P(a_j). \end{aligned}$$

Alors, $\mathbf{f}$ est une base de E^*, et sa base duale est $\mathbf{e} = (L_0, \ldots, L_n)$, où L_i est défini par

$$L_i = \prod_{j \neq i} \frac{X - a_j}{a_i - a_j}.$$

Les vérifications sont laissées en exercice.

IX.3. Endomorphismes adjoints

THÉORÈME IX.3.1. *Soient E et F des K-espaces vectoriels de dimension n, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée.*

(1) *Pour tout $u \in \mathscr{L}(E)$, il existe un unique $u^* \in \mathscr{L}(F)$ tel que*
$$b(u(x), y) = b(x, u^*(y)) \text{ pour tout } x \in E, y \in F.$$

(2) *Pour tout $\varphi \in \mathscr{L}(F)$, il existe un unique ${}^*\varphi \in \mathscr{L}(E)$ tel que*
$$b(x, \varphi(y)) = b({}^*\varphi(x), y) \text{ pour tout } x \in E, y \in F.$$

De plus, pour tous $u, u_1, u_2 \in \mathscr{L}(E), \varphi, \varphi_1, \varphi_2 \in \mathscr{L}(F)$, et tout $\lambda \in K$, on a
$$(u_1 + \lambda u_2)^* = u_1^* + \lambda u_2^* \quad \text{et } {}^*(\varphi_1 + \lambda\varphi_2) = {}^*\varphi_1 + \lambda {}^*\varphi_2,$$
ainsi que
$${}^*(u^*) = u \quad \text{et } ({}^*\varphi)^* = \varphi.$$
Autrement dit, on a un isomorphisme d'espace vectoriels
$$\begin{array}{ccc} \mathscr{L}(E) & & \mathscr{L}(F) \\ u & \longmapsto & u^*, \end{array}$$
dont l'isomorphisme réciproque est donné par
$$\begin{array}{ccc} \mathscr{L}(E) & & \mathscr{L}(F) \\ {}^*\varphi & \longleftarrow & \varphi. \end{array}$$

Enfin, on a $\mathrm{Id}_E^* = \mathrm{Id}_F, {}^*\mathrm{Id}_F = \mathrm{Id}_E$, et pour tous $,u_1, u_2 \in \mathscr{L}(E)$, et tous $\varphi_1\varphi_2 \in \mathscr{L}(F)$, on a
$$(u_1 \circ u_2)^* = u_2^* \circ u_1^* \quad \text{et } {}^*(\varphi_1 \circ \varphi_2) = {}^*\varphi_2 \circ {}^*\varphi_1.$$

Démonstration. Démontrons (1). Montrons d'abord l'unicité de u^*. Soient $\mathbf{e}, \mathbf{f}$ des bases de E et F respectivement. Soit $B = \mathrm{Mat}(b, \mathbf{e}, \mathbf{f})$. Supposons que $u^* \in \mathscr{L}(F)$ existe, et soient $M = \mathrm{Mat}(u; \mathbf{e})$ et $M^* = \mathrm{Mat}(u^*; \mathbf{f})$. D'après le lemme **??**, l'égalité désirée s'écrit alors matriciellement
$$X^t M^t B Y = X^t B M^* Y \text{ pour tout } X, Y \in K^n.$$
En appliquant cette égalité aux divers vecteurs de base de K^n, on obtient $M^t B = BM^*$ (**Exercice**). Puisque b est non dégénérée, B est inversible, et on a donc $M^* = B^{-1}M^tB$. Ainsi, si u^* existe, il est unique. Pour montrer l'existence, définissons $u^* \in \mathscr{L}(F)$ comme l'unique endomorphisme tel que
$$\mathrm{Mat}(u^*; \mathbf{f}) = B^{-1}M^tB.$$
Alors par construction, u^* vérifie la condition souhaitée (il suffit de remonter les calculs précédents pour le voir).

On montre de même l'existence et l'unicité de ${}^*\varphi$. Si $N = \mathrm{Mat}(\varphi; \mathbf{f})$, alors la matrice ${}^*N$ de ${}^*\varphi$ dans la base $\mathbf{e}$ est donnée par
$${}^*N = (B^t)^{-1}N^tB^t.$$

Montrons maintenant le dernier point. Si $u \in \mathscr{L}(E)$, alors ${}^*(u^*) \in \mathscr{L}(E)$ est l'unique endomorphisme de E vérifiant
$$b({}^*(u^*)(x), y) = b(x, u^*(y)) \text{ pour tout } x \in E, y \in E.$$
Mais u vérifie
$$b(u(x), y) = b(x, u^*(y)) \text{ pour tout } x \in E, y \in E,$$

et par la propriété d'unicité, on en déduit que ${}^*(u^*) = u$. On procède de même pour l'autre égalité.

Les reste est un exercice laissé au lecteur. □

Définition IX.3.2. Si $u \in \mathscr{L}(E)$ et $\varphi \in \mathscr{L}(F)$, les endomorphismes u^* et ${}^*\varphi$ s'appellent les endomorphismes *adjoints* à u et φ respectivement.

Remarque IX.3.3. Si $E = F$ et b est symétrique, les deux notions coïncident et on retrouve la définition usuelle de l'adjoint d'un endomorphisme.

Proposition IX.3.4. *Soient E et F des K-espaces vectoriels de dimension n, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée. Soient $u \in \mathscr{L}(E), \varphi \in \mathscr{L}(F)$. Soient $\mathbf{e}$ et $\mathbf{f}$ des bases de E et F respectivement, et soient*

$$B = \mathrm{Mat}(b; \mathbf{e}, \mathbf{f}), M = \mathrm{Mat}(u; \mathbf{e}), N = \mathrm{Mat}(\varphi; \mathbf{f}).$$

Alors on a

$$\mathrm{Mat}(u^*; \mathbf{f}) = B^{-1}M^tB \text{ et } \mathrm{Mat}({}^*\varphi; \mathbf{e}) = (B^t)^{-1}N^tB^t.$$

En particulier, si $\mathbf{e}$ et $\mathbf{f}$ sont duales par rapport à b, on a

$$\mathrm{Mat}(u^*; \mathbf{f}) = M^t \text{ et } \mathrm{Mat}({}^*\varphi; \mathbf{e}) = N^t.$$

Démonstration. La première partie a déjà été établie au cours de la démonstration du théorème précédent. Pour la seconde, il suffit de remarquer que si $\mathbf{e}$ et $\mathbf{f}$ sont duales par rapport à b, alors $\mathrm{Mat}(b; \mathbf{e}, \mathbf{f}) = I_n$ par définition même. □

Exemple IX.3.5. Soit E un K-espace vectoriel de dimension n, et soit $u \in \mathscr{L}(E)$. Alors l'adjoint $u^* \in \mathscr{L}(E^*)$ de u pour le crochet de dualité est donné par

$$\begin{aligned} E^* &\longrightarrow E^* \\ f &\longmapsto f \circ u. \end{aligned}$$

En effet, $u^* \in \mathscr{L}(E^*)$ est l'unique endomorphisme de E^* satisfaisant

$$f(u(x)) = (u^*(f))(x) \text{ pour tout } x \in E, f \in E^*.$$

Cet endomorphisme s'appelle la *transposée* de u ou *l'endomorphisme transposé* de u .

Soit $\mathbf{e}$ une base de E et soit $\mathbf{e}^*$ la base duale. Posons $M = \mathrm{Mat}(u; \mathbf{e})$. D'après ce qui précède, on a

$$\mathrm{Mat}(u^*; \mathbf{e}^*) = M^t.$$

De plus, on a

$$\mathrm{Mat}({}^*(u^*); \mathbf{e}) = (M^t)^t,$$

et comme ${}^*(u^*) = u$, on obtient $(M^t)^t = M$.

La dualité fournit ainsi une explication conceptuelle à l'égalité matricielle $(M^t)^t = M$ lorsque M est une matrice carrée.

Théorème IX.3.6. *Soient E et F des K-espaces vectoriels de dimension n, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée. Soient $u \in \mathscr{L}(E), \varphi \in \mathscr{L}(F)$. Alors on a*

$$\mathrm{Ker}(u^*) = \mathrm{Im}(u)^\perp, \mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp$$

et de même

$$\mathrm{Ker}({}^*\varphi) = {}^\perp\mathrm{Im}(\varphi), \mathrm{Im}({}^*\varphi) = {}^\perp\mathrm{Ker}(\varphi).$$

En particulier, on a $\mathrm{rg}(u^*) = \mathrm{rg}(u)$ et $\mathrm{rg}({}^*\varphi) = \mathrm{rg}(\varphi)$.

Démonstration. On montre la première série d'égalités. Soit $y \in F$. Alors on a $y \in \mathrm{Ker}(u^*) \iff u^*(y) = 0_F$. Comme b est non dégénérée, on a donc

$$\begin{aligned} y \in \mathrm{Ker}(u^*) \quad &\iff \quad b(x, u^*(y)) = 0 \quad \text{pour tout } x \in E \\ &\iff \quad b(u(x), y) = 0 \quad \text{pour tout } x \in E. \end{aligned}$$

Ceci équivaut à $y \in \mathrm{Im}(u)^\perp$.

Montrons maintenant la deuxième égalité.

D'après la proposition **??**, il suffit de montrer ${}^\perp\mathrm{Im}(u^*) = \mathrm{Ker}(u)$. Soit $x \in E$. Alors on a

$$\begin{aligned} x \in {}^\perp\mathrm{Im}(u^*) \quad &\iff \quad b(x, u^*(y)) = 0 \quad \text{pour tout } y \in F \\ &\iff \quad b(u(x), y) = 0 \quad \text{pour tout } y \in F. \end{aligned}$$

Comme b est non dégénérée, cela équivaut à $u(x) = 0_E$, c'est-à-dire $x \in \mathrm{Ker}(u)$. On a donc

$$\mathrm{rg}(u^*) = \dim_K(\mathrm{Ker}(u)^\perp) = n - \dim_K(\mathrm{Ker}(u)) = \mathrm{rg}(u).$$

On établit de même les autres égalités. □

Exemple IX.3.7. Lorsque l'on traduit ceci matriciellement dans des bases duales, on obtient en particulier que $\mathrm{rg}(M) = \mathrm{rg}(M^t)$ pour toute matrice carrée, ce qui est bien connu. On obtient aussi

$$Y \in \mathrm{Im}(M^t) \iff \text{pour tout } X \in K^n, (MX = 0 \Rightarrow X^tY = 0)$$

et

$$X \in \mathrm{Ker}(M^t) \iff X^tMY = 0 \text{ pour tout } Y \in K^n.$$

Chapitre X

Groupe orthogonal d'un espace euclidien ou hermitien

X.1. Rappels

On commence par faire des rappels sur les espaces euclidiens et hermitiens.

Définition X.1.1. Soit $K = \mathbb{R}$ ou $\mathbb{C}$, et soit E un K-espace vectoriel. Un *produit scalaire* sur E est une application $\langle\,,\rangle : E\times E \longrightarrow K$ telle que, pour tous $x, x', y, y' \in E$ et tout $\lambda \in K$, on a :

(1) $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$ et $\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle$;

(2) $\langle \lambda x, y\rangle = \overline{\lambda}\langle x, y\rangle$ et $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$;

(3) $\langle y, x\rangle = \overline{\langle x, y\rangle}$.
En particulier, $\langle x, x\rangle \in \mathbb{R}$ pour tout $x \in E$;

(4) pour tout $x \in E$, $\langle x, x\rangle \geq 0$;

(5) pour tout $x \in E$, $\langle x, x\rangle = 0 \Longrightarrow x = 0$.

Remarquons que lorsque $K = \mathbb{R}$, les conditions $(1) - (3)$ se résument simplement en disant que l'application $\langle\,,\rangle$ est une forme bilinéaire symétrique, puisque l'on a $\overline{\lambda} = \lambda$ pour tout $\lambda \in \mathbb{R}$.

Deux vecteurs $x, y \in E$ sont dits *orthogonaux* pour le produit scalaire $\langle\,,\rangle$ si $\langle x, y\rangle = 0$.

Deux sous-espaces F et G d'un espace euclidien E sont dit *orthogonaux* si pour tout $x \in F$ et tout $y \in G$, on a $\langle x, y\rangle = 0$.

Une famille de sous-espaces $(F_i)_{i\in I}$ de E est dite en *somme directe orthogonale* si pour tous $i, j \in I, i \neq j$, F_i et F_j sont orthogonaux. Il est facile de voir que ces sous-espaces sont alors en somme directe. On le note

$$\bigoplus_{i\in I}^{\perp} F_i.$$

Une famille de vecteurs $(x_i)_{i\in I}$ de E est dite *orthonormée* si elle vérifie

$$\langle x_i, x_j\rangle = \delta_{ij} \quad \text{pour tous } i, j \in I.$$

Si F est un sous-espace vectoriel de E, on définit *l'orthogonal* $F^{\perp}$ de F par

$$F^{\perp} = \{x \in E \mid \langle x, y\rangle = 0 \text{ pour tout } y \in F\}.$$

C'est un sous-espace vectoriel de E.

Définition X.1.2. Un couple $(E, \langle\,,\rangle)$, où E est un K-espace vectoriel **de dimension finie** et $\langle\,,\rangle$ est un produit scalaire est appelé un espace *euclidien* si $K = \mathbb{R}$ et un espace *hermitien* si $K = \mathbb{C}$.

On dira souvent « soit E un espace euclidien/hermitien » en oubliant de mentionner le produit scalaire.

Notation. Si $M = (a_{ij}) \in \mathrm{M}_{m\times n}(K)$, on note $M^* = (\overline{a}_{ji})$. En particulier, si $K = \mathbb{R}$, on a $M^* = M^t$.

Exemple X.1.3. Soit $n \geq 1$. L'application

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (X, Y) &\longmapsto X^*Y \end{aligned}$$

est un produit scalaire sur $\mathbb{C}^n$. De même, l'application

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{C} \\ (X, Y) &\longmapsto X^tY \end{aligned}$$

est un produit scalaire sur $\mathbb{R}^n$.

Dans les deux cas, on l'appellera le produit scalaire usuel (ou standard).

Le résultat suivant rappelle quelques propriétés des espaces euclidiens et hermitiens.

Proposition X.1.4. *Soit $(E, \langle\,,\rangle)$ un espace euclidien ou hermitien. Alors :*

(1) *toute famille orthonormée de E est libre, et se complète en une base orthonormée. En particulier, tout sous-espace de E admet une base orthonormée (i.e. une base qui soit une famille orthonormée) ;*

(2) *si $\mathbf{e}$ est une base orthonormée, alors pour tous $x, y \in E$, on a*

$$\langle x, y\rangle = [x]_{\mathbf{e}}^*[y]_{\mathbf{e}}$$

(3) *pour tout sous-espace F de E, on a $(F^\perp)^\perp = F$, et $E = F \overset{\perp}{\oplus} F^\perp$.*
En particulier, $\dim(F^\perp) = \dim(E) - \dim(F)$;

(4) *pour tout $u \in \mathscr{L}(E)$, il existe un unique $u^* \in \mathscr{L}(E)$ tel que*

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle \quad \textit{pour tout } y \in E.$$

Si $\mathbf{e}$ est une base **orthonormée** *de E, alors on a*

$$\mathrm{Mat}(u^*; \mathbf{e}) = \mathrm{Mat}(u; \mathbf{e})^*.$$

En particulier, $\det(u^*) = \overline{\det(u)}$.
De plus, $\mathrm{Id}_E^* = \mathrm{Id}_E$, *et pour tous $u_1, u_2 \in \mathscr{L}(E)$, on a*

$$(u_1 \circ u_2)^* = u_2^* \circ u_1^*.$$

Définition X.1.5. Si E est un espace euclidien ou hermitien, et si $u \in \mathscr{L}(E)$, l'endomorphisme $u^* \in \mathscr{L}(E)$ s'appelle l'endomorphisme *adjoint* à u, ou l'adjoint de u.

On dit que $u \in \mathscr{L}(E)$ est :

(1) un endomorphisme *normal* si $u^* \circ u = u \circ u^*$;

(2) une *isométrie* si l'on a[1] $u^* \circ u = \mathrm{Id}_E$;

(3) un endomorphisme *autoadjoint* si $u^* = u$;

(4) un endomorphisme *antiautoadjoint* si $u^* = -u$.

1. Dans ce cas, puisque on est en dimension finie, u est un automorphisme de E, et on a $u^{-1} = u^*$.

On remarque que les isométries, ainsi que les endomorphismes adjoints ou antiautoadjoints sont normaux.

Lorsque $K = \mathbb{R}$, on parle aussi d'endomorphisme symétrique ou antisymétrique, plutôt que d'endomorphisme autoadjoint ou antiautoadjoint.

L'ensemble des isométries a une structure particulièrement agréable.

Définition X.1.6. L'ensemble des isométries de E forme un sous-groupe[2] de $\mathrm{GL}(E)$, que l'on note parfois $\mathrm{Iso}(E)$. C'est exactement l'ensemble des endomorphismes de E qui conservent le produit scalaire.

Lorsque $K = \mathbb{R}$, le groupe $\mathrm{Iso}(E)$ est plutôt noté $\mathrm{O}(E)$, et appelé le *groupe orthogonal* de E .

Lorsque $K = \mathbb{C}$, le groupe $\mathrm{Iso}(E)$ est plutôt noté $\mathrm{U}(E)$, et appelé le *groupe unitaire* de E . Dans ce cas, une isométrie est aussi appelé un endomorphisme *unitaire*.

Nous allons maintenant introduire une définition supplémentaire, afin de pouvoir retraduire ce qui précède en termes matriciels.

Définition X.1.7. On dit qu'une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est *hermitienne* si l'on a $M^* = M$ et *antihermitienne* si l'on a $M^* = -M$.

Une matrice $M \in \mathrm{M}_n(\mathbb{R})$ sera dite *symétrique* si l'on a $M^t = M$ et *antisymétrique* si l'on a $M^t = -M$.

Une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est dite *unitaire* si $M^*M = I_n$.

L'ensemble $\mathrm{U}_n(\mathbb{C})$ des matrices unitaires est un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$.

Une matrice $M \in \mathrm{M}_n(\mathbb{R})$ est dite *orthogonale* si l'on a $M^tM = I_n$.

L'ensemble $\mathrm{O}_n(\mathbb{R})$ des matrices orthogonales est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$.

Remarques X.1.8.

(1) Soit $u \in \mathscr{L}(E)$ d'un endomorphisme d'un espace hermitien.

Si $M \in \mathrm{M}_n(\mathbb{C})$ est la matrice représentative de u dans une base **orthonormée**, alors u est :

(i) normal si, et seulement si, $M^*M = MM^*$;

(ii) une isométrie si, et seulement si, M est unitaire ;

(iii) auto-adjoint si, et seulement si, M est hermitienne ;

(iv) antiautoadjoint si, et seulement si, M est antihermitienne.

(2) Soit $u \in \mathscr{L}(E)$ d'un endomorphisme d'un espace euclidien.

Si $M \in \mathrm{M}_n(\mathbb{R})$ est la matrice représentative de u dans une base **orthonormée**, alors u est :

(i) normal si, et seulement si, $M^tM = MM^t$;

(ii) une isométrie si, et seulement si, M est orthogonale ;

(iii) symétrique si, et seulement si, M est symétrique ;

(iv) antisymétrique si, et seulement si, M est antisymétrique.

2. Cela se démontre facilement en utilisant le dernier point de la proposition précédente.

Définition X.1.9. Une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est dite *unitaire* si $M^*M = I_n$.

L'ensemble $\mathrm{U}_n(\mathbb{C})$ des matrices unitaires est un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$.

Une matrice $M \in \mathrm{M}_n(\mathbb{R})$ est dite *orthogonale* si l'on a $M^tM = I_n$.

L'ensemble $\mathrm{O}_n(\mathbb{R})$ des matrices orthogonales est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$.

Si E est un espace hermitien (resp. euclidien), $u \in \mathscr{L}(E)$ est une isométrie si, et seulement si, sa matrice représentative dans une base orthonormée est unitaire (resp. orthogonale).

Remarquons que si $u \in \mathrm{O}(E)$, alors on a $1 = \det(u^*)\det(u) = \det(u)^2$, et donc $\det(u) = \pm 1$.

Définition X.1.10. Une isométrie $u \in \mathrm{O}(E)$ d'un espace euclidien E est dite *directe* si l'on a $\det(u) = 1$, et *indirecte* si l'on a $\det(u) = -1$. L'ensemble des isométries directes forment un sous-groupe de $\mathrm{O}(E)$, noté $\mathrm{O}^+(E)$ ou encore $\mathrm{SO}(E)$, appelé *groupe spécial orthogonal*. L'ensemble des isométries indirectes est noté parfois $\mathrm{O}^-(E)$. Ce n'est pas un sous-groupe de $\mathrm{O}(E)$.

X.2. Structure du groupe orthogonal

Nous allons maintenant étudier d'un peu plus près la structure de $\mathrm{O}^{\pm}(E)$ lorsque $\dim_{\mathbb{R}}(E) = 2$ ou 3. Pour comprendre ces ensembles, il suffit de décrire les matrices représentatives de leurs éléments dans une base orthonormée.

Notation. Si $n \geq 1$, on pose

$$\mathrm{O}_n^+(\mathbb{R}) = \{M \in \mathrm{O}_n(\mathbb{R}) \mid \det(M) = 1\}$$

et

$$\mathrm{O}_n^-(\mathbb{R}) = \{M \in \mathrm{O}_n(\mathbb{R}) \mid \det(M) = -1\}.$$

Si E est un espace euclidien de dimension n, une isométrie u de E est directe (resp. indirecte) si, et seulement si, sa matrice représentative dans une base orthonormée est dans $\mathrm{O}_n^+(\mathbb{R})$ (resp. $\mathrm{O}_n^-(\mathbb{R})$).

Un calcul direct montre le résultat suivant.

Proposition X.2.1. *On a les égalités suivantes :*

$$(1)\quad \begin{aligned} \mathrm{O}_2^+(\mathbb{R}) &= \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a^2 + b^2 = 1, a, b \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mid \theta \in \mathbb{R} \right\} \end{aligned}$$

$$(2)\quad \begin{aligned} \mathrm{O}_2^-(\mathbb{R}) &= \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \mid a^2 + b^2 = 1, a, b \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \mid \theta \in \mathbb{R} \right\}. \end{aligned}$$

L'interprétation géométrique est la suivante : si E un plan euclidien, les isométries directes de E sont les rotations planes, et les isométries indirectes sont les symétries orthogonales par rapport à une droite.

Remarque X.2.2. Si E est un espace euclidien de dimension 2, et si u est une isométrie indirecte, il existe une base orthonormée de E pour laquelle la matrice de u dans cette base est

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

Ceci se comprend bien géométriquement : les vecteurs parallèles à l'axe de symétrie sont fixes, et les vecteurs orthogonaux à l'axe sont envoyés sur leurs opposés. Il suffit donc de prendre un vecteur unitaire de l'axe de symétrie, et un vecteur unitaire orthogonal à cet axe.

On passe au cas de la dimension 3.

Proposition X.2.3. *Soit E un espace euclidien de dimension 3, et soit u une isométrie de E.*

(1) *Si $u \in \mathrm{O}^+(E)$, alors il existe une base orthonormée de E dans laquelle la matrice représentative de u est de la forme*
$$\begin{pmatrix} 1 & & \\ & \cos(\theta) & -\sin(\theta) \\ & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

(2) *Si $u \in \mathrm{O}^-(E)$, alors il existe une base orthonormée de E dans laquelle la matrice représentative de u est de la forme*
$$\begin{pmatrix} -1 & & \\ & \cos(\theta) & -\sin(\theta) \\ & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Cela sera la conséquence directe d'un résultat plus général qui sera démontré plus loin. L'interprétation géométrique est la suivante : si E est un espace euclidien de dimension 3, les éléments de $\mathrm{O}^+(E)$ sont les rotations de E, et ceux de $\mathrm{O}^-(E)$ sont obtenues en effectuant une rotation dans un plan de $\mathbb{R}^3$, puis une réflexion orthogonale par rapport à ce plan.

On va maintenant montrer qu'il existe des isométries indirectes en toute dimension.

Définition X.2.4. Soit E un espace euclidien, et soit $x \in E \setminus \{0\}$.

La *réflexion orthogonale* d'hyperplan $H = (\mathbb{R}x)^\perp$ est la symétrie orthogonale par rapport à H (et de direction $H^\perp = \mathbb{R}x$), c'est-à-dire l'unique endomorphisme $\tau_x \in \mathscr{L}(E)$ vérifiant :

(1) $(\tau_x)_{|H} = \mathrm{Id}_H$;
(2) $\tau_x(x) = -x$.

Autrement dit, pour tout $\lambda \in K$, et tout $y \in H$, on a
$$\tau_x(\lambda x + y) = -\lambda x + y.$$

Lemme X.2.5. *Pour tout $x \in E \setminus \{0\}$, on a $\tau_x \in \mathrm{O}^-(E)$. De plus, pour tout vecteur $v \in E$, on a*
$$\tau_x(v) = v - 2\frac{\langle x, v\rangle}{\langle x, x\rangle}x.$$
Enfin, $\tau_x^2 = \mathrm{Id}_E$, et $\dim_{\mathbb{R}} \mathrm{Ker}(\tau_x + \mathrm{Id}_E) = 1$.

Démonstration. La restriction du produit scalaire à H étant un produit scalaire, on peut trouver une base $e_2, \ldots, e_n$ de H qui soit orthonormée. La famille $e_1 = \dfrac{x}{\sqrt{\langle x, x\rangle}}, e_2, \ldots, e_n$ est alors une base orthonormée de E, dans laquelle la matrice représentative de τ_x est
$$M = \begin{pmatrix} -1 & 0 \\ 0 & I_{n-1} \end{pmatrix}.$$

Il est alors clair que $M^tM = I_n$ et donc que τ_x est une isométrie indirecte, puisque $\det(M) = -1$. Pour montrer le second point, il suffit de vérifier que l'application

$$E \longrightarrow E, v \longmapsto v - 2\frac{\langle x, v\rangle}{\langle x, x\rangle}x$$

est linéaire, envoie x sur $-x$ et se restreint à l'identité sur $\{x\}^\perp$, ce qui est clair. Le dernier point est évident, puisque $M^2 = I_n$ et $\dim_{\mathbb{R}}(\mathrm{Ker}(M + I_n)) = 1$. □

On continue par un lemme.

Lemme X.2.6. *Soit $u \in \mathrm{O}(E)$, et soit F un sous-espace de E stable par u. Alors $F^\perp$ est aussi stable par u. De plus, on a $u_{|_F} \in \mathrm{O}(F)$ et $u_{|_{F^\perp}} \in \mathrm{O}(F^\perp)$.*

Démonstration. Soit $x \in F^\perp$. Alors pour tout $y \in F$, on a

$$\langle u(x), y\rangle = \langle u(x), u(u^{-1})(y)\rangle = \langle x, u^{-1}(y)\rangle = 0,$$

car F étant stable par u, il l'est aussi par u^{-1}. Ainsi $u(x) \in F^\perp$. Le second point vient du fait que si u conserve le produit scalaire, il en est de même pour $u_{|_F}$ et $u_{|_{F^\perp}}$. □

On peut maintenant énoncer le théorème suivant.

Théorème X.2.7. *Soit E un espace euclidien de dimension n, et soit u une isométrie de E. Alors :*

(1) *u est la composée de $n - \dim_{\mathbb{R}}(\mathrm{Ker}(u - \mathrm{Id}_E))$ réflexions orthogonales ;*

(2) *si u est la composée de p réflexions orthogonales, alors*

$$p \geq n - \dim_{\mathbb{R}}(\mathrm{Ker}(u - \mathrm{Id}_E)).$$

En particulier, toute isométrie est le produit d'au plus n réflexions orthogonales, et $\mathrm{O}(E)$ est engendré par les réflexions orthogonales.

Démonstration. Soit $u \in \mathrm{O}(E)$, et soit $F_u = \mathrm{Ker}(u - \mathrm{Id}_E)$. Supposons que l'on ait $u = \tau_{x_1} \circ \cdots \circ \tau_{x_p}$, et posons $F = \mathrm{Vect}_{\mathbb{R}}(x_1, \ldots, x_p)$. Par définition, on a $\dim_{\mathbb{R}}(F) \leq p$. Si $x \in F^\perp$, alors x est orthogonal à x_i, et ainsi $\tau_{x_i}(x) = x$ pour tout $i \in [\![1, p]\!]$. Il s'ensuit que $u(x) = x$ pour tout $x \in F^\perp$, d'où l'inclusion $F^\perp \subset F_u$. On a ainsi $\dim_{\mathbb{R}}(F_u) \geq \dim_{\mathbb{R}}(F^\perp)$. Or,

$$\dim_{\mathbb{R}}(F^\perp) = n - \dim_{\mathbb{R}}(F) \geq n - p,$$

d'où $p \geq n - \dim_{\mathbb{R}}(F_u)$.

Démontrons maintenant que u est la composée de $n - \dim_{\mathbb{R}}(F_u)$ réflexions orthogonales. D'après ce qui précède, il suffit de démontrer que u peut s'écrire comme la composée d'au plus $n - \dim_{\mathbb{R}}(F_u)$ réflexions orthogonales.

Nous allons procéder par récurrence sur $p_u = n - \dim_{\mathbb{R}}(F_u)$. Plus précisément, soit (H_k) la proposition

(H_k) Toute isométrie $u \in \mathrm{O}(E)$ telle que $p_u \in [\![0, k]\!]$ est la composée d'au plus p_u réflexions orthogonales.

Si $k = 0$, alors $p_u = 0$ et donc $u = \mathrm{Id}_E$. Dans ce cas, u est le produit de 0 réflexions orthogonales. Ainsi, (H_0) est vraie.

Soit $k \in [\![0, n-1]\!]$. Supposons que (H_k) soit vraie, et montrons que (H_{k+1}) est vraie. Il reste à montrer que le résultat attendu est vrai lorsque $u \in \mathrm{O}(E)$ vérifie $p_u = k + 1 \geq 1$.

L'idée est trouver une réflexion orthogonale τ telle que $p_{\tau\circ u} < p_u$. Dans ce cas, $\tau\circ u$ sera un produit d'au plus $p_{\tau\circ u}$ réflexions orthogonales par hypothèse de récurrence. Mais alors, l'isométrie $u = \tau\circ(\tau\circ u)$ sera produit d'au plus $p_{\tau\circ u}+1 \leq p_u$ réflexions, ce qui achèvera la récurrence. En fait, on va trouver une réflexion orthogonale τ telle que $F_u \subsetneq F_{\tau\circ u}$. On aura alors $\dim_{\mathbb{R}}(F_u) < \dim_{\mathbb{R}}(F_{\tau\circ u})$ et donc $p_{\tau\circ u} < p_u$. On doit donc exhiber une réflexion orthogonale vérifiant cette propriété. Puisque le sous-espace $F_u^\perp$ est de dimension $p_u \geq 1$, il existe $x \in F_u^\perp \setminus \{0\}$. On a donc $x \neq u(x)$, puisque l'on a $F_u \cap F_u^\perp = \{0\}$. Posons $y = u(x)$. Comme F_u est un sous-espace vectoriel stable par u, il en est de même de son orthogonal $F_u^\perp$, d'après le lemme **??**. Ainsi, on a $y \in F_u^\perp$, et de plus $y \neq x$, d'où $x-y \in F_u^\perp \setminus \{0\}$. Remarquons enfin que l'on a

$$\langle x-y, x+y\rangle = \langle x,x\rangle - \langle y,y\rangle = \langle x,x\rangle - \langle u(x),u(x)\rangle = 0,$$

car u est une isométrie. Cela illustre simplement le fait que les diagonales d'un losange sont orthogonales. On a alors

$$\tau_{x-y}(y) = -\frac{1}{2}\tau_{x-y}(x-y) + \frac{1}{2}\tau_{x-y}(x+y) = -\frac{1}{2}(y-x) + \frac{1}{2}(x+y) = x.$$

Géométriquement, cela s'explique par le fait qu'une diagonale d'un losange est la bissectrice de l'angle formé par deux côtés adjacents.

On va montrer que $\tau = \tau_{x-y}$ vérifie la propriété souhaitée. Soit $z \in F_u$ (on a donc $u(z) = z$). Comme $x-y \in F_u^\perp$, z est orthogonal à $x-y$ et on obtient alors

$$\tau_{x-y}(u(z)) = \tau_{x-y}(z) = z.$$

Ainsi, $F_u \subset F_{\tau_{x-y}\circ u}$, et l'inclusion est stricte, car $x \notin F_u$ mais

$$\tau_{x-y}(u(x)) = \tau_{x-y}(y) = x,$$

et donc $x \in F_{\tau_{x-y}\circ u}$. Ceci achève la récurrence, ainsi que la démonstration. □

On continue en étudiant les classes de conjugaison des symétries orthogonales.

THÉORÈME X.2.8. *Soit E un espace euclidien, et soient $s_1, s_2 \in \mathrm{O}(E)$ deux symétries orthogonales.*

Alors, il existe $u \in \mathrm{O}^+(E)$ tel que $s_2 = u\circ s_1\circ u^{-1}$ si, et seulement si, les sous-espaces $\mathrm{Ker}(s_1+\mathrm{Id}_E)$ et $\mathrm{Ker}(s_2+\mathrm{Id}_E)$ ont même dimension.

Démonstration. Soient $s_1, s_2 \in \mathrm{O}(E)$. Si $u \in \mathrm{O}^+(E)$ vérifie $s_2 = u\circ s_1\circ u^{-1}$, on a

$$\mathrm{Ker}(s_2+\mathrm{Id}_E) = \mathrm{Ker}(u\circ(s_1+\mathrm{Id}_E)\circ u^{-1}) = u(\mathrm{Ker}(s_1+\mathrm{Id}_E)),$$

et comme u est inversible, les sous-espaces $\mathrm{Ker}(s_1+\mathrm{Id}_E)$ et $\mathrm{Ker}(s_2+\mathrm{Id}_E)$ ont même dimension.

Inversement, supposons que $E_1 = \mathrm{Ker}(s_1+\mathrm{Id}_E)$ et $E_2 = \mathrm{Ker}(s_2+\mathrm{Id}_E)$ aient même dimension $r \geq 0$. Soit $(e_1,\ldots,e_r)$ une base orthonormée de E_1, soit $(f_1,\ldots,f_r)$ une base orthonormée de E_2. Soit $(e_{r+1},\ldots,e_n)$ une base orthonormée de $E_1^\perp$, et soit $(f_{r+1},\ldots,f_n)$ une base orthonormée de $E_2^\perp$. Les familles

$$\mathbf{e} = (e_1,\ldots,e_n) \quad \text{et} \quad \mathbf{f} = (f_1,\ldots,f_n)$$

sont alors deux bases orthonormées de E. Soit $u \in \mathscr{L}(E)$ l'unique endomorphisme de E tel que

$$u(e_i) = f_i \quad \text{pour tout } i \in [\![1,n]\!].$$

Puisque u envoie une base orthonormée sur une base orthonormée, u est une isométrie. Quitte à remplacer un vecteur de $\mathbf{e}$ par son opposé (ce qui multiplie le déterminant de u par -1, comme on peut le voir matriciellement), on peut supposer que $u \in \mathrm{O}^+(E)$.

Pour tout $i \in [\![1, r]\!]$, on a alors

$$(u \circ s_1 \circ u^{-1})(f_i) = (u \circ s_1 \circ u^{-1})(u(e_i)) = u(s_1(e_i)) = u(-e_i) = -f_i = s_2(f_i).$$

D'autre part, puisque s_1 et s_2 sont des symétries orthogonales, on a

$$E_1^\perp = \mathrm{Ker}(s_1 - \mathrm{Id}_E) \quad \text{et} \quad E_2^\perp = \mathrm{Ker}(s_2 - \mathrm{Id}_E).$$

Par conséquent, pour tout $i \in [\![r+1, n]\!]$, on a

$$(u \circ s_1 \circ u^{-1})(f_i) = (u \circ s_1 \circ u^{-1})(u(e_i)) = u(s_1(e_i)) = u(e_i) = f_i = s_2(f_i).$$

Ainsi, $u \circ s_1 \circ u^{-1}$ et s_2 coïncident sur une base de E, et donc $u \circ s_1 \circ u^{-1} = s_2$. Ceci achève la démonstration. □

Corollaire X.2.9. *Soit E un espace euclidien. Alors, toutes les réflexions orthogonales sont conjuguées dans* $\mathrm{O}(E)$.

Démonstration. Une réflexion orthogonale est une symétrie orthogonale s telle que $\mathrm{Ker}(s + \mathrm{Id}_E)$ soit de dimension 1. On applique alors le théorème précédent pour conclure. □

On finit ces considérations sur le groupe orthogonal en calculant son centre.

Lemme X.2.10. *Soit E un espace euclidien. Alors, $Z(\mathrm{O}(E)) = \{\pm \mathrm{Id}_E\}$.*

Démonstration. Avant de passer à la démonstration proprement dite, remarquons que, pour tout $x \in E \setminus \{0\}$, on a

$$u \circ \tau_x \circ u^{-1} = \tau_{u(x)}.$$

En effet, on a

$$(u \circ \tau_x \circ u^{-1})(u(x)) = u(\tau_x(x)) = u(-x) = -u(x).$$

De plus, puisque u et u^{-1} sont des isométries, pour tout $y \in E$, les vecteurs x et y sont orthogonaux si, et seulement si $u(x)$ et $u(y)$ le sont. Autrement dit, on a $\{u(x)\}^\perp = u(\{x\}^\perp)$. Pour tout $y' \in \{u(x)\}^\perp$, il existe donc un vecteur $y \in \{x\}^\perp$ tel que $y' = u(y)$, et on a ainsi

$$(u \circ \tau_x \circ u^{-1})(y') = (u \circ \tau_x \circ u^{-1})(u(y)) = u(\tau_x(y)) = u(y) = y',$$

d'où l'égalité souhaitée.

Revenons à notre problème. Clairement, Id_E et $-\mathrm{Id}_E$ sont des isométries commutant à tout élément de $\mathrm{O}(E)$. Montrons la réciproque. Pour cela, on va montrer le fait plus précis suivant : si $u \in \mathrm{O}(E)$ commute à toute réflexion orthogonale, alors $u = \pm \mathrm{Id}_E$.

Supposons donc que pour tout $x \in E \setminus \{0\}$, on ait $\tau_x \circ u = u \circ \tau_x$. On obtient alors

$$\tau_x = u \circ \tau_x \circ u^{-1} = \tau_{u(x)} \quad \text{pour tout } x \in E \setminus \{0\}.$$

Écrivons $u(x) = \lambda_x x + h$, avec $\lambda_x \in \mathbb{R}$ et $h \in \{x\}^\perp$. On a d'une part

$$\tau_{u(x)}(u(x)) = -u(x) = -\lambda_x x - h,$$

et d'autre part

$$\tau_{u(x)}(u(x)) = \tau_x(u(x)) = \tau_x(\lambda_x + h) = -\lambda_x x + h.$$

On en déduit que $h = 0$, soit $u(x) = \lambda_x x$. Ainsi, u stabilise toute droite de E, et u est donc une homothétie d'après le lemme **??**. Comme u est une isométrie, on obtient $u = \pm \mathrm{Id}_E$.

Puisqu'un élément de $Z(\mathrm{O}(E))$ commute en particulier à toute réflexion orthogonale, on a fini. □

On s'intéresse maintenant au groupe spécial orthogonal.

DÉFINITION X.2.11. Soit E un espace euclidien. On dit que $u \in \mathrm{O}(E)$ est un *renversement* si $u^2 = \mathrm{Id}_E$ et $\mathrm{Ker}(u + \mathrm{Id}_E)$ est de dimension 2.

On vérifie facilement que l'on a $E = \mathrm{Ker}(u - \mathrm{Id}_E) \overset{\perp}{\oplus} \mathrm{Ker}(u + \mathrm{Id}_E)$, et la matrice représentative de u dans une base bien choisie est donc

$$\begin{pmatrix} I_{n-2} & 0 \\ 0 & -I_2 \end{pmatrix},$$

puisque $\mathrm{Ker}(u + \mathrm{Id}_E)$ est de dimension 2. En particulier, un renversement est de déterminant 1, et est donc un élément de $\mathrm{O}^+(E)$. En fait, toute isométrie directe est un produit de renversements.

THÉORÈME X.2.12. *Soit E un espace euclidien de dimension n. Si $n \geq 3$, toute isométrie directe est le produit d'au plus n renversements.*

Démonstration. Soit $u \in \mathrm{O}^+(E)$ une isométrie directe.

Puisque $\det(u) = 1$ et que le déterminant d'une réflexion est -1, le théorème **??** montre que $u = \tau_1 \circ \cdots \circ \tau_{2p}, 2p \leq n$, où chaque τ_i est une réflexion. On peut toujours supposer que $u \neq \mathrm{Id}_E$ (car Id_E est le produit de 0 renversements).

Supposons tout d'abord que $n = 3$. Alors, $u = \tau_1 \circ \tau_2$ est le produit d'exactement 2 réflexions (car $u \neq \mathrm{Id}_E$). Puisque τ_i est une réflexion, sa matrice dans une base orthonormée bien choisie est

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

comme on l'a déja remarqué. Il devient alors clair que $-\tau_i$ est un renversement, et on a fini, puisque $u = (-\tau_1) \circ (-\tau_2)$.

Supposons maintenant que $n > 3$. D'après les considérations précédentes, il suffit de montrer que le produit de deux réflexions s'écrit comme produit de deux renversements. Soient τ_1, τ_2 deux réflexions orthogonales d'hyperplans respectifs H_1, H_2, et posons $u = \tau_1 \circ \tau_2$. Alors, on a $\dim_{\mathbb{R}}(H_1 \cap H_2) = n - 2$ ou $n - 1$. En particulier, $H_1 \cap H_2$ contient dans tous les cas un sous-espace de dimension $n - 3$, disons F. Puisque F est contenu dans H_1 et dans H_2, on a $u_F = \mathrm{Id}_F$, et F est donc stable par u. Par suite, $F^\perp$ est stable par u par le lemme **??**. En recollant une base orthonormée de F et une base orthonormée de $F^\perp$, on obtient une base orthonormée de E dans laquelle la matrice représentative de u est

$$M = \begin{pmatrix} I_{n-3} & \\ & M' \end{pmatrix},$$

où M' est la matrice de $u_{F^\perp}$ dans la base orthonormée de $F^\perp$ choisie. On en déduit alors que $\det(u_{F^\perp}) = 1$, et la relation $M^t M = I_n$ implique facilement que $M'^t M' = I_3$. Ainsi, $u_{|_{F^\perp}} \in \mathrm{O}^+(F^\perp)$. Or, $\dim_{\mathbb{R}}(F^\perp) = 3$, donc par le cas précédent, il existe des renversements σ'_1, σ'_2 de $F^\perp$ tels que

$$u_{F^\perp} = \sigma'_1 \circ \sigma'_2.$$

On définit alors $\sigma_i \in \mathscr{L}(E)$ par

$$\sigma_i(x + y) = x + \sigma'_i(y) \text{ pour tout } x \in F, y \in F^\perp.$$

On vérifie facilement [3] que σ_i est un renversement et que $u = \sigma_1 \circ \sigma_2$. Ceci achève la démonstration. □

Puisqu'un renversement est une symétrie orthogonale s telle que $\mathrm{Ker}(s + \mathrm{Id}_E)$ soit de dimension 2, le théorème **??** fournit immédiatement le résultat suivant.

Corollaire X.2.13. *Soit E un espace euclidien de dimension $n \geq 2$. Alors, tous les renversements sont conjugués dans $\mathrm{O}^+(E)$.*

On finit ce long paragraphe en calculant le centre du groupe spécial orthogonal.

Lemme X.2.14. *Soit E un espace euclidien de dimension $n \geq 1$. Alors, on a*

$$Z(\mathrm{O}^+(E)) = \begin{cases} \mathrm{O}^+(E) & \text{si } n = 1, 2 \\ \{\pm \mathrm{Id}_E\} & \text{si } n \geq 4 \text{ est pair} \\ \{\mathrm{Id}_E\} & \text{si } n \geq 3 \text{ est impair.} \end{cases}$$

Démonstration. Remarquons tout d'abord que Id_E et $-\mathrm{Id}_E$ sont des isométries qui commutent à tout élément de $\mathrm{O}(E)$. De plus, Id_E est bien entendu une isométrie directe, tandis que $\det(-\mathrm{Id}_E) = (-1)^n$. Ainsi, $-\mathrm{Id}_E \in \mathrm{O}^+(E)$ si, et seulement si, n est pair. Dans ce cas, c'est une isométrie directe qui commute à tout élément de $\mathrm{O}^+(E)$. Enfin, si $n = 1, 2$, le groupe $\mathrm{O}^+(E)$ est abélien, donc égal à son centre.

On peut donc supposer que $n \geq 3$ dans la suite. Dans ce cas, on va démontrer que si $u \in \mathrm{O}(E)$ commute à tout renversement, alors $u = \pm \mathrm{Id}_E$. Le résultat voulu en découlera immédiatement.

Soit P un sous-espace de E de dimension 2, et soit $s \in \mathscr{L}(E)$ l'unique endomorphisme de E tel que

$$s(x) = x \quad \text{pour tout } x \in P^\perp \quad \text{et} \quad s(x) = -x \quad \text{pour tout } x \in P.$$

Ainsi, s est une symétrie orthogonale telle que $\mathrm{Ker}(s + \mathrm{Id}_E) = P$. C'est donc un renversement. Par conséquent, on a $s \circ u = u \circ s$ par hypothèse. Autrement dit, $u \circ s \circ u^{-1} = s$. Or, on a

$$\mathrm{Ker}(u \circ s \circ u^{-1} + \mathrm{Id}_E) = \mathrm{Ker}(u \circ (s + \mathrm{Id}_E) \circ u^{-1}) = u(\mathrm{Ker}(s + \mathrm{Id}_E)) = u(P),$$

et on obtient donc en particulier $u(P) = P$.

Ainsi, u stabilise tout plan de E. Montrons alors que u stabilise toute droite de E.

Soit $D = \mathbb{R}x_0$ une droite de E. Puisque $n \geq 3$, il existe $x_1, x_2 \in E$ tels que la famille (x_0, x_1, x_2) soit libre. Les sous-espaces

$$P_1 = \mathrm{Vect}_{\mathbb{R}}(x_0, x_1) \quad \text{et} \quad P_2 = \mathrm{Vect}_{\mathbb{R}}(x_0, x_2)$$

sont deux plans de E tels que $P_1 \cap P_2 = D$. Par conséquent

$$u(D) = u(P_1 \cap P_2) = u(P_1) \cap u(P_2) = P_1 \cap P_2 = D,$$

la deuxième égalité découlant de la bijectivité de u. Ainsi, u stabilise toute droite de E, et u est donc une homothétie d'après le lemme **??**. Comme u est une isométrie, on obtient $u = \pm \mathrm{Id}_E$. □

X.3. Réduction des endomorphismes normaux

On s'intéresse ici à la réduction des endomorphismes normaux d'un espace euclidien ou hermitien. Rappelons que si E est un espace euclidien ou hermitien de dimension $n \geq 1$, $u \in \mathscr{L}(E)$ est normal si $u^*u = uu^*$. Si $\mathbf{e}$ une base orthonormée de E, et si $M = \mathrm{Mat}(u; \mathbf{e})$, ceci est équivalent à $M^*M = MM^*$, où $M^* = \overline{M}^t$.

3. Ceci est laissé en exercice au lecteur.

On commence par quelques résultats intermédiaires.

LEMME X.3.1. *Soit E un espace euclidien ou hermitien, soit $u \in \mathscr{L}(E)$ et soit F un sous-espace vectoriel de E stable par u. Si u est respectivement un endomorphisme normal, autoadjoint, antiautoadjoint ou une isométrie, alors F et $F^\perp$ sont stables par u et u^*, et les endomorphismes $u_F, u_{F^\perp}, u_F^*$ et $u_{F^\perp}^*$ sont respectivement normaux, autoadjoints, antiautoadjoints ou des isométries.*

Démonstration. Choisissons une base orthonormée de F, et une base orthonormée de $F^\perp$. En recollant ces deux bases, on obtient une base orthonormée $\mathbf{e}$ de E. Soit M la matrice représentative de u dans cette base. Comme F est stable par u, on a

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

et la matrice représentative de u^* est

$$M^* = \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}.$$

Si u vérifie une des propriétés de l'énoncé, on a dans tous les cas

$$M^*M = MM^*,$$

ce qui se traduit par l'égalité

$$\begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix}.$$

En particulier, $A^*A = AA^* + BB^*$. En appliquant la trace, et en utilisant le fait que $\mathrm{tr}(MN) = \mathrm{tr}(NM)$ pour toutes matrices carrées M, N, on obtient l'égalité $\mathrm{tr}(BB^*) = 0$. Or, si $B = (b_{ij})$, on a $\mathrm{tr}(BB^*) = \sum_{i,j} |b_{ij}|^2$. Ainsi, on en déduit que $b_{ij} = 0$ pour tous i, j, c'est-à-dire $B = 0$.

Par construction de la base $\mathbf{e}$, on en déduit que F et $F^\perp$ sont stables par u et u^*. Il s'ensuit que A, C, A^* et C^* sont les matrices de $u_F, u_{F^\perp}, u_F^*$ et $u_{F^\perp}^*$ dans les bases orthonormées de F et $F^\perp$ choisies au départ. Si maintenant on a $MM^* = M^*M, M^* = M, M^* = -M$ ou $M^*M = I_n$, il en est de même de A, A^*, C et C^*, d'où le résultat. □

LEMME X.3.2. *Soit E un espace vectoriel réel non nul, et soit $u \in L(E)$. Alors il existe un sous-espace F de E stable par u de dimension* 1 *ou* 2.

DÉMONSTRATION. Écrivons $\chi_u = P_1 \cdots P_r$, où $P_i \in \mathbb{R}[X]$ est irréductible unitaire. Puisque P_i est irréductible, on a $\deg(P_i) = 1$ ou 2. Par Cayley-Hamilton, on a

$$P_1(u) \circ \cdots \circ P_r(u) = 0.$$

En particulier, $P_1(u) \circ \cdots \circ P_r(u)$ n'est pas inversible, et donc un des endomorphismes $P_i(u)$ n'est pas injectif (sinon, tous les $P_i(u)$ seraient inversibles car on est en dimension finie, et par suite $P_1(u) \circ \cdots \circ P_r(u)$ serait inversible).

Soit $x \in \mathrm{Ker} P_i(u), x \neq 0$.

Si P_i est de degré 1, alors $P_i = X - a$. On a alors $P_i(u)(x) = u(x) - ax = 0$, soit $u(x) = ax$, et donc $F = \mathbb{R}x$ est stable par u, de dimension 1.

Si P_i est de degré 2, alors $P_i = X^2 + aX + b$. On a alors $P_i(u)(x) = u^2(x) + au(x) + b = 0$, et donc $u(u(x)) = -a(ux) - b(x)$. On en déduit aisément que $F = \mathrm{Vect}(x, u(x))$ est stable par u, de dimension ≤ 2. □

LEMME X.3.3. *Soit E un plan euclidien*[4]*, et soit $u \in \mathscr{L}(E)$ un endomorphisme normal. Alors, il existe une base orthonormée $\mathbf{e}$ de E dans laquelle la matrice représentative de u dans cette base est, soit diagonale, soit de la forme*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, a, c \in \mathbb{R}, c \neq 0.$$

Démonstration. Soit $\mathbf{e}$ une base orthonormée de E, et soit

$$M = \mathrm{Mat}(u; \mathbf{e}) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Puisque u est normal, on a $M^t M = M M^t$, ce qui est équivalent à

$$a^2 + c^2 = a^2 + b^2,\ ab + cd = ac + bd,\ c^2 + d^2 = b^2 + d^2.$$

On a donc $c = \varepsilon b$, avec $\varepsilon = \pm 1$. On a alors

$$b(a + \varepsilon d) = b(a\varepsilon + d).$$

Soit $\chi_u = \lambda^2 - (a + d)\lambda + ad - bc$.

Supposons d'abord que $\varepsilon = 1$, i.e. $b = c$. Alors, le discriminant de ce polynôme en λ est

$$(a + d)^2 - 4ad + 4b^2 = (a - d)^2 + 4b^2 \geq 0.$$

Ains, χ_u possède au moins une racine réelle. Autrement dit, u possède une valeur propre $\lambda \in \mathbb{R}$. Soit e'_1 un vecteur propre associé,l'on peut supposer unitaire, et soit $F = \mathbb{R}e'_1$. Alors, F est stable par u. Par le lemme **??**, $F^\perp$ est aussi stable par u. Comme $F^\perp$ est de dimension 1, il est engendré par un vecteur non nul e'_2, que l'on peut supposer unitaire. Alors, $\mathbf{e}' = (e'_1, e'_2)$ est une base orthonormée de E, et la matrice $\mathrm{Mat}(u; \mathbf{e}')$ est diagonale (car F et $F^\perp$ sont stables par u et de dimension 1).

Supposons maintenant que $\varepsilon = -1$. Si $b = 0$, alors $c = 0$ et M est diagonale. Si $b \neq 0$, on a $c = -b$ et $-a + d = a - d$, soit $d = a$, ce qui achève la démonstration. □

On alors le résultat suivant.

THÉORÈME X.3.4. *Soit $n \geq 1$ un entier.*

(1) *Soit E un espace hermitien de dimension n, et soit $u \in \mathscr{L}(E)$ un endomorphisme normal. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale*[5]*.*

(2) *Soit E un espace euclidien de dimension n, et soit $u \in \mathscr{L}(E)$ un endomorphisme normal. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est de la forme*

4. C'est-à-dire un espace euclidien de dimension 2.

5. Dans le langage de la réduction des endomorphismes, dont on parlera ultérieurement, on dira que u est diagonalisable dans une base orthonormée.

$$\begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_r & & & & & \\ & & & a_1 & -b_1 & & & \\ & & & b_1 & a_1 & & & \\ & & & & & \ddots & & \\ & & & & & & a_s & -b_s \\ & & & & & & b_s & a_s \end{pmatrix},$$

où $r, s \geq 0$ sont des entiers vérifiant $r + 2s = n, \lambda_i, a_j \in \mathbb{R}, b_j \in \mathbb{R}^\times$, pour tout $i \in [\![1, r]\!], j \in [\![1, s]\!]$.

Démonstration. On le montre par récurrence sur n. Commençons par le cas hermitien.

Si $n = 1$, il n'y a rien à faire. Supposons que le résultat soit vrai pour tout endomorphisme normal d'un espace hermitien de dimension $n - 1$, et soit u un endomorphisme normal d'un espace hermitien E de dimension n.

Par le lemme **??**, il existe au moins une droite F de E stable par u. Par le lemme **??**, $F^\perp$ est stable par u et $u_{F^\perp}$ est normal. Par hypothèse de récurrence, puisque $\dim(F^\perp) = n - 1$, il existe une base orthonormée de $F^\perp$ dans laquelle la matrice représentative de $u_{|_{F^\perp}}$ est diagonale. Pour conclure, il suffit prendre un vecteur unitaire dirigeant F et de lui accoler cette base de $F^\perp$ pour obtenir une base orthonormée de E convenable.

Supposons maintenant que E soit un espace euclidien. Encore une fois, il n'y a rien à faire si $n = 1$, et le cas $n = 2$ est donné par le lemme précédent. Supposons que le résultat soit vrai pour tout endomorphisme normal d'un espace euclidien de dimension $\leq n-1$, pour un certain $n \geq 3$, et soit u un endomorphisme normal d'un espace euclidien E de dimension n.

Par le lemme **??**, il existe un sous-espace F de E de dimension ≤ 2 stable par u. Par le lemme **??**, F et $F^\perp$ sont stables par u et les endomorphismes u_F sont $u_{F^\perp}$ sont normaux. Par les cas $n = 1, 2$, il existe une base orthonormée de F dans laquelle la matrice représentative de u_F a la forme voulue. Puisque $\dim(F^\perp) = n - \dim(F) \leq n - 1$, par hypothèse de récurrence, il existe une base orthonormée de $F^\perp$ dans laquelle la matrice représentative de $u_{F^\perp}$ a la forme voulue. Pour conclure, il suffit de recoller ces deux bases et de changer l'ordre des vecteurs si nécessaire pour obtenir une base orthonormée de E ayant la propriété requise. □

On peut maintenant élucider la structure des endomorphismes autoadjoints, antiautoadjoints et unitaires d'un espace hermitien.

Corollaire X.3.5. *Soit E un espace hermitien.*

(1) *Si $u \in \mathscr{L}(E)$ est autoadjoint, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale à coefficients réels ;*

(2) *si $u \in \mathscr{L}(E)$ est antiautoadjoint, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale à coefficients purement imaginaires ;*

(3) *si $u \in \mathrm{U}(E)$ est unitaire, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale, dont les coefficients diagonaux sont de module 1.*

En particulier :

(i) *pour toute matrice* $M \in \mathrm{M}_n(\mathbb{C})$ *hermitienne, il existe* $P \in \mathrm{U}_n(\mathbb{C})$ *telle que* P^*MP *soit diagonale à coefficients réels ;*

(ii) *pour toute matrice* $M \in \mathrm{M}_n(\mathbb{C})$ *antihermitienne, il existe* $P \in \mathrm{U}_n(\mathbb{C})$ *telle que* P^*MP *soit diagonale à coefficients purement imaginaires ;*

(iii) *pour toute matrice* $M \in \mathrm{U}_n(\mathbb{C})$*, il existe* $P \in \mathrm{U}_n(\mathbb{C})$ *telle que* P^*MP *soit diagonale, dont les coefficients diagonaux sont de module* 1.

Démonstration. Un endomorphisme autoadjoint, antiautoadjoint ou unitaire est normal. On choisit alors une base orthonormée $\mathbf{e}$ comme dans l'énoncé du théorème précédent. La matrice $M = \mathrm{Mat}(u; \mathbf{e})$ est alors diagonale. Si u est autoadjoint, antiautoadjoint ou unitaire, M est hermitienne, antihermitienne ou unitaire. Le résultat provient alors d'un simple calcul. Le dernier point est juste une réinterprétation matricielle de ce qui précède. □

Passons maintenant au cas des espaces euclidiens.

Corollaire X.3.6. *Soit* E *un espace euclidien et soit* $u \in \mathscr{L}(E)$ *un endomorphisme symétrique. Alors, il existe une base orthonormée de* E *dans laquelle la matrice de* u *est diagonale à coefficients réels.*

En particulier, pour toute matrice $M \in \mathrm{M}_n(\mathbb{R})$ *telle que* $M^t = M$*, il existe une matrice* $P \in \mathrm{O}_n(\mathbb{R})$ *telle que* P^tMP *soit diagonale.*

Démonstration. Rappelons qu'un endomorphisme symétrique est normal. On choisit alors une base orthonormée $\mathbf{e}$ comme dans l'énoncé du théorème précédent. Soit $M = \mathrm{Mat}(u; \mathbf{e})$. La matrice M est donc symétrique. Mais alors, avec les notations du théorème, on a $s = 0$ dans ce cas (car les b_i sont non nuls) et M est donc diagonale.

Le dernier point est juste une réinterprétation matricielle du résultat. □

Corollaire X.3.7. *Soit* E *un espace euclidien, et soit* $u \in \mathscr{L}(E)$ *un endomorphisme antisymétrique. Alors, il existe une base orthonormée de* E *dans laquelle la matrice de* u *est de la forme*

$$\begin{pmatrix} 0 & & & & & & & \\ & \ddots & & & & & & \\ & & 0 & & & & & \\ & & & 0 & -b_1 & & & \\ & & & b_1 & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & 0 & -b_s \\ & & & & & & b_s & 0 \end{pmatrix},$$

où $s \geq 0$ *est un entier et* $b_j \in \mathbb{R}^\times$ *pour* $j \in [\![1, s]\!]$.

Démonstration. Puisque $u^* = -u$, la matrice représentative de u dans n'importe quelle base orthonormée est antisymétrique, en particulier dans la base orthonormée fournie par le théorème précédent. Mais une matrice réelle antisymétrique a des coefficients diagonaux nuls. □

Corollaire X.3.8. *Soit* E *un espace euclidien, et soit* $u \in \mathrm{O}(E)$*. Alors, il existe une base orthonormée de* E *dans laquelle la matrice de* u *est de la forme*

$$\begin{pmatrix} I_p & & & & & & \\ & -I_q & & & & & \\ & & \cos(\theta_1) & -\sin(\theta_1) & & & \\ & & \sin(\theta_1) & \cos(\theta_1) & & & \\ & & & & \ddots & & \\ & & & & & \cos(\theta_s) & -\sin(\theta_s) \\ & & & & & \sin(\theta_s) & \cos(\theta_s) \end{pmatrix},$$

où $p, q, s \geq 0$ vérifient $p + q + 2s = \dim_{\mathbb{R}}(E)$, et $\theta_j \in \mathbb{R} \setminus \pi\mathbb{Z}$.

De plus, on a

$$p = \dim_{\mathbb{R}}(\mathrm{Ker}(u - \mathrm{Id}_E)), q = \dim_{\mathbb{R}}(\mathrm{Ker}(u + \mathrm{Id}_E)).$$

En particulier, p, q et s sont uniquement déterminés par u.

Démonstration. Puisque $u^* \circ u = \mathrm{Id}_E$, sa matrice représentative dans n'importe quelle base orthonormée est dans $\mathrm{O}_n(\mathbb{R})$, en particulier dans la base fournie par le théorème précédent. Chaque bloc doit être alors dans $\mathrm{O}_1(\mathbb{R})$ ou $\mathrm{O}_2(\mathbb{R})$. Ainsi, on a $\lambda_i = \pm 1$ pour tout $i \in [\![1, r]\!]$ et on a $a_j^2 + b_j^2 = 1$ pour tout $j \in [\![1, s]\!]$. On a donc le résultat voulu en utilisant la paramétrisation du cercle unité, et en réordonnant les vecteurs de base, ce qui donne toujours une base orthonormée. Une rotation plane différente de $\pm\mathrm{Id}_E$ ne stabilisant aucune droite, un simple calcul montre que l'on a $p = \dim_{\mathbb{R}}(\mathrm{Ker}(u - \mathrm{Id}_E))$ et $q = \dim_{\mathbb{R}}(\mathrm{Ker}(u + \mathrm{Id}_E))$. □

Remarque X.3.9. Remarquons que l'on a $\det(u) = (-1)^q$. En particulier, si $u \in \mathrm{O}^-(E)$, $q \geq 1$. Si de plus $\dim_{\mathbb{R}}(E) = 2$, vu que l'on a $p + q + 2s = 2$, on a $s = 0$ et donc $p = 1$, et on retrouve le résultat de la remarque **??**.

Si $\dim_{\mathbb{R}}(E)$ est impair et $u \in \mathrm{O}^+(E)$, alors q est pair, et donc p est impair, car $p + q + 2s = n$. En particulier, $p \geq 1$. Si de plus E est de dimension 3, on a $(p, q, s) = (1, 0, 1)$ ou $(p, q, s) = (1, 2, 0)$. On retrouve alors la première partie de la proposition **??**, à la formulation près.

Si $u \in \mathrm{O}^-(E)$, alors q est impair, et on a $(p, q, s) = (2, 1, 0)$ ou $(0, 1, 1)$. On retrouve ainsi la seconde partie de la proposition **??**, encore à la formulation près.

Ce dernier résultat s'interprète comme suit : une isométrie d'un espace euclidien se décompose en produit de réflexions orthogonales et de rotations planes.