Structures algébriques fondamentales

Définitions, constructions et propriétés universelles

Patrick Massot

28 mars 2023

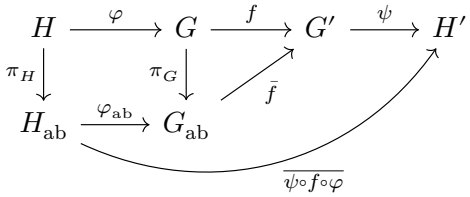
Introduction

Des éléments aux flèches

Le cours d'algèbre des deux premières années d'études en mathématique s'inspire des opérations utilisées en arithmétique, en géométrie et en analyse pour extraire des cadres abstraits pertinents tels que la notion de groupe, d'anneau ou d'espace vectoriel. Dans ce processus, un point crucial est d'oublier la structure interne des éléments des ensembles étudiés. Par exemple la notion d'espace vectoriel s'applique aussi bien aux points de la géométrie euclidienne d'un plan muni d'un repère et aux fonctions de $\mathbb{R}$ dans $\mathbb{R}$ car on peut oublier la complexité individuelle de chaque fonction lorsqu'on manipule abstraitement des sommes de fonctions, des sous-espaces vectoriels de fonctions etc.

Bien sûr ce processus n'interdit pas de retourner à la structure interne des éléments lors de l'étude d'exemples. Le niveau abstrait est une possibilité supplémentaire. Il permet d'abord de diminuer le nombre d'énoncés à démontrer, car les théorèmes abstraits s'appliquent à tous les exemples concrets. Il permet aussi de penser plus efficacement, en étant délivré des détails concrets qui ne font qu'obscurcir la situation, ne serait-ce qu'au niveau des notations. Un autre élément important de cette étude est la notion d'application compatible avec une structure algébrique, par exemple la notion de morphisme de groupe ou d'application linéaire.

Un des buts de ce cours est d'aller plus loin dans cette direction consistant à se détacher des éléments individuels formant un groupe ou un espace vectoriel par exemple, pour se focaliser sur les sous-objets (sous-groupe, sous-espaces) et surtout sur les applications entre ces objets (morphismes de groupes, applications linéaires). L'aspect le plus visible de cette orientation est l'omniprésence des diagrammes d'applications dans les pages qui suivent. Le diagramme suivant est extrait de l'énoncé du corollaire 3.4.6.



Chaque sommet de ce graphe est un groupe et chaque flèche représente un morphisme de groupes. Les éléments de ces groupes ne jouent aucun rôle, toute la discussion tourne autour des morphismes. Là encore, ce mode de pensée est une possibilité qui s'ajoute à l'existant sans interdire quoi que ce soit. La suite naturelle de ce développement est la théorie des catégories, qui n'est pas du tout abordée dans ce cours mais nécessite d'avoir déjà entraîné son esprit à penser ainsi.

Constructions libres et quotients

Les paragraphes précédents se concentraient sur la première partie du sous-titre du cours : les définitions. Le second point est celui des constructions. Loin de se limiter aux exemples ayant inspirés les définitions, l'algèbre introduit de très nombreuses constructions d'ensembles munis de structures algébriques à partir d'autres ensembles munis (ou pas) d'autres structure algébriques. Un premier principe de construction est le passage aux sous-objets (sous-groupes, sous-espaces vectoriels etc.), particulièrement les sous-objets engendrés par une partie. Mais cela ne créé pas vraiment de nouveauté. Il y a deux grands principes de construction véritable, qui travaillent dans des directions opposées : les objets libres et les quotients. Dans les deux cas, la construction s'accompagne d'une porte d'entrée et d'une porte de sortie. La porte d'entrée explique comment fabriquer des éléments de l'objet algébrique construit. En général il n'y a pas tellement de mystère, il s'agit d'utiliser les ingrédients fournis et les opérations algébriques promises par la nouvelle structure. Le point clef est la porte de sortie, appelée propriété universelle, qui explique comment construire des applications qui partent de l'objet construit.

Les objets libres généralisent l'idée de base d'un espace vectoriel. De ce point de vue, les deux caractéristiques principales d'une base e d'un $\mathbb{K}$-espace vectoriel E sont : les éléments e_i de la base et leurs combinaisons linéaires sont dans E (c'est la porte d'entrée) et, pour tout $\mathbb{K}$-espace vectoriel F et toute famille de vecteurs v_i de F, il existe une unique application linéaire de E dans F qui envoie chaque e_i sur le v_i correspondant (c'est la propriété universelle). Dans ce cours, cette idée est adaptée à toutes les structures algébriques étudiées mais en ne supposant pas que l'objet ambiant E est déjà construit. Partant de la collection d'éléments destiné à servir de base, on créé un objet libre accueillant ces éléments et une propriété universelle.

L'exemple le plus familier d'une construction d'objet libre est l'algèbre $\mathbb{K}[X]$ des polynômes à coefficients dans un corps $\mathbb{K}$. On veut construire une $\mathbb{K}$-algèbre commutative (c'est à dire un ensemble muni d'une structure de $\mathbb{K}$-espace vectoriel et d'une structure d'anneau commutatif qui cohabitent en bonne entente) qui contient un élément distingué X et tout ce qu'il faut pour faire une algèbre commutative sans rien imposer d'autre que la présence de X et les axiomes d'algèbre commutative. En particulier $\mathbb{K}[X]$ doit contenir en plus de X des éléments 0 et 1 comme tout anneau, et les résultats de toutes les opérations d'addition, de multiplication et de multiplication par un scalaire, comme par exemple $X+1$ ou $(X+1)X$, mais on n'impose pas d'idée préconçue sur la construction de 0, 1 ou des opérations. La précision « sans rien imposer d'autre » est vague. Sa version précise est exactement la propriété universelle de la construction : pour toute $\mathbb{K}$-algèbre B et tout élément b_0 de B, il existe un unique morphisme de $\mathbb{K}$-algèbre de $\mathbb{K}[X]$ dans B qui envoie X sur b_0. Si on avait imposé plus que la présence de X et les axiomes, il faudrait mettre des restrictions sur l'élément b_0 sur lequel on désire envoyer X. La propriété universelle joue aussi son rôle général d'expliquer comment sortir de $\mathbb{K}[X]$. Par exemple si $B = \mathrm{End}(E)$, l'algèbre des endomorphismes d'un $\mathbb{K}$-espace vectoriel E alors la propriété universelle de $\mathbb{K}[X]$ construit, pour chaque endomorphisme u, l'application qui envoie un polynôme P sur l'endomorphisme $P(u)$. Remarquons enfin qu'en général on ne peut pas tricher en construisant l'algèbre $\mathbb{K}[X]$ comme sous-algèbre de l'algèbre des fonctions de $\mathbb{K}$ dans $\mathbb{K}$. Par exemple pour $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$, le polynôme X^2+X est non nul mais la fonction polynomiale correspondante est identiquement nulle puisqu'elle s'annule

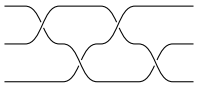
sur les deux seuls éléments de $\mathbb{K}$.

La deuxième grande idée de construction est celle de quotient. Alors que les objets libres sont plus gros que les ingrédients qu'ils ont ingérés, les quotients sont des objets plus petits obtenus en collant entre eux certains éléments des ingrédients. L'exemple le plus connu, qui donne son nom à la procédure, est celui des nombres rationnels. L'ensemble $\mathbb{Q}$ est obtenu à partir de l'ensemble $\mathbb{Z} \times \mathbb{N}^*$ des fractions en collant (p, q) et (p', q') lorsque $pq' = p'q$. Dans le monde des ensembles sans structure algébrique, la construction d'un quotient ne nécessite que deux ingrédients : un ensemble X et une relation d'équivalence sur X. Le résultat est un ensemble Y muni d'une application $\pi : X \to Y$. La porte d'entrée est simple : tous les éléments de Y proviennent de X, c'est à dire que π est surjective. La propriété universelle explique comment sortir de Y à partir d'une fonction f définie sur X vérifiant une condition de compatibilité avec la relation d'équivalence : $\forall x\, x',\ x \sim x' \Rightarrow f(x) = f(x')$. En particulier si X est muni d'opérations algébriques on peut espérer construire ainsi des opérations algébriques sur Y. C'est par exemple ce que l'on fait pour construire une structure de corps sur $\mathbb{Q}$ ou une structure d'anneau sur $\mathbb{Z}/n\mathbb{Z}$. Chaque structure algébrique possède ses propres contraintes pour construire des quotients héritant d'une structure algébrique.

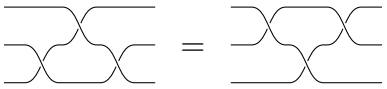
La combinaison des deux idées de construction, objets libres et quotients, est universelle : tout objet est quotient d'un objet libre. Cette observation est réconfortante mais assez tautologique. La démonstration consiste à partir d'un objet E, considérer l'immense objet librement engendré par l'ensemble de tous les éléments de E, oubliant ainsi toutes les relations entre ces éléments, puis quotienter pour recoller tous les éléments qui doivent l'être et retrouver ainsi E. Le cas intéressant est donc bien de partir d'un objet peu structuré, construire un objet libre plus structuré mais sans relation intéressante et enfin quotienter pour obtenir un objet structuré intéressant.

Prenons un exemple familier en analyse. Partant du corps $\mathbb{R}$ des nombres réels, on veut créer un corps contenant $\mathbb{R}$ et un élément i vérifiant $i^2 = -1$. En particulier ce corps doit être une $\mathbb{R}$-algèbre. On part d'un ensemble à un élément noté X et on construit la $\mathbb{R}$-algèbre librement engendré par X. Comme expliqué précédemment, il s'agit tout simplement de $\mathbb{R}[X]$. Dans cette $\mathbb{R}$-algèbre, X ne vérifie aucune relation hormis celles imposées par la définition d'un anneau. Ensuite on quotiente pour imposer la relation $X^2 = -1$. On note $\mathbb{C}$ le quotient obtenu et i l'image de X dans le quotient.

Prenons maintenant un exemple en topologie et théorie des groupes. Soit n un entier strictement positif. On appelle tresse à n brins un diagramme de la forme suivante (ici $n = 3$)



modulo l'opération de déformation à extrémités fixées. Par exemple



Le produit de deux tresses à n brins est simplement obtenu en juxtaposant les deux dessins. On obtient ainsi un groupe (l'inverse d'une tresse est obtenue en prenant son image dans un miroir vertical). Il existe différentes façons de rendre cette « définition »

rigoureuse. La voie algébrique est de commencer par remarquer que la brique de base est un croisement élémentaire. On note σ_i le croisement du i-ème et du $(i+1)$-ème brin (c'est juste un symbole). On forme le groupe libre engendré par ces $n-1$ symboles puis on impose par quotient les relations $\sigma_i\sigma_j = \sigma_j\sigma_i$ lorsque $|i-j| \geqslant 2$ (les croisements sans brin commun commutent) et $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$ pour tout i (c'est la relation correspondant au dessin précédent). On obtient ainsi un groupe B_n parfaitement bien défini une fois compris le chapitre 3. Bien sûr il faut réfléchir un peu pour voir que les relations imposées suffisent à encoder toute l'intuition géométrique de départ. La porte d'entrée est donnée par les symboles σ_i devenus éléments du groupe. La propriété universelle affirme que, pour tout groupe G et toute collection d'éléments $g_1, \dots, g_n$ vérifiant les relations imposées aux σ_i ci-dessus, il existe un unique morphisme de groupes de B_n dans G qui envoie chaque σ_i sur le g_i correspondant. Ainsi, on peut construire un morphisme vers le groupes de permutations $\mathfrak{S}_n$ qui ne retient que le point d'arrivée de chaque brin (par exemple la tresse dessinée ci-dessus correspond à une permutation cyclique). Pour cela il suffit de vérifier que les transposition $\tau_{i,i+1}$ vérifient les relations imposées pour obtenir un morphisme qui envoie σ_i sur $\tau_{i,i+1}$.

Encapsulation et fondements

Dans tout ce programme de construction, il est crucial de veiller à respecter l'encapsulation : on ne retourne pas voir les détails d'une construction terminée. La porte d'entrée (et ses éventuelles propriétés) et la propriété universelle doivent rester les seuls accès. Sans cela la complexité des constructions s'empile et devient ingérable. Prenons l'exemple des polynômes. On construit $\mathbb{K}[X_1, \dots, X_n]$ comme étant l'ensemble des fonctions à support fini sur l'ensemble des fonctions sur $\{1, \dots, n\}$ à valeurs dans $\mathbb{N}$ à valeurs dans $\mathbb{K}$. Oui, la phrase précédente est pénible à lire et non, il n'y a pas de problème de copier-coller ayant mal tourné. Et ce niveau d'offuscation est atteint sans même se demander comment a été construit le corps $\mathbb{K}$ ou même le monoïde $\mathbb{N}$ des entiers naturels.

Ce dernier point permet de mentionner en passant que, en plus du problème de la complexité par empilement il y a un risque de s'approcher trop près des fondements et de devoir s'y confronter. Comme presque toute la pratique mathématique, ce cours est indépendant des fondements, tout en prétendant très vaguement reposer sur les deux étages de fondements fournis par la logique du premier ordre avec égalité et la théorie des ensembles, disons de Zermelo-Fraenkel+choix (ZFC). Si cette prétention était sérieuse, il faudrait construire $\mathbb{N}$. Or ZFC est une théorie complètement amorphe, il n'y a qu'un seul type d'objets, les ensembles, et qu'une relation, l'appartenance. L'encodage le plus fréquent des entiers naturels consiste en l'ensemble $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \dots\}$ (et inutile de dire que les points de suspension ne forment pas une méthode de construction encouragée donc il y a du travail à faire pour construire $\mathbb{N}$ dans ZFC). Pire, les fonctions ne sont pas une notion primitive dans ZFC. Il faut passer par l'idée de graphe d'une fonction : une fonction de A dans B est une partie f de $A \times B$ telle que $\forall a \in A, \exists! b \in B, (a,b) \in f$. Hélas la notion de produit de deux ensembles n'est pas non plus une notion primitive. Par exemple on peut définir $A \times B$ comme l'ensemble des ensembles de la forme $\{\{a\}, \{a,b\}\}$ pour $a \in A$ et $b \in B$. Munis de ces préliminaires (et d'un corps $\mathbb{K}$…), on peut retourner achever de rendre impossible à lire la définition de $\mathbb{K}[X_1, \dots, X_n]$.

Il faut noter qu'on dispose maintenant de fondements alternatifs bien plus sympathiques que ZFC, comme les diverses variantes de la théorie des types dépendants mais en creusant suffisamment on fini toujours par se perdre. Bref, dans ce cours on suppose que $\mathbb{N}$ est déjà connu mais on construit $\mathbb{Z}$ et $\mathbb{Q}$ dans le chapitre 4 (l'apparition très tardive de la construction de $\mathbb{Z}$ est une simple commodité pédagogique qui n'introduit pas problème de cohérence, on peut lire le début de la section 4.4 entre les chapitres 2 et 3).

Théorèmes de structure et de classification

L'essentiel de ce cours élémentaire est consacré aux constructions. Une fois compris l'état d'esprit général il n'y a pas de difficulté majeure, et le mot théorème n'apparait presque jamais, il n'y a que de nombreux lemmes et quelques propositions. Mais l'algèbre est aussi faite de problèmes difficiles où il faut faire la route dans l'autre sens et comprendre un objet inconnu. Les théorèmes correspondants sont les théorèmes de structure (pas vraiment dans le même sens que dans « structure algébrique »…) ou, dans les cas les plus favorables, les théorèmes de classification. Par exemple en algèbre linéaire, le théorème de la base incomplète affirme que tout espace vectoriel est libre (autrement dit il peut être construit comme objet libre). C'est un théorème de structure très fort. Le théorème de la dimension affirme en plus que deux espaces vectoriels sur un même corps sont isomorphes si et seulement si leurs bases ont même cardinal, c'est un théorème de classification. La théorie de la réduction des endomorphismes quant à elle étudie la structure des $\mathbb{K}[X]$-algèbres (ce point de vue sera abordé en M1).

Dans cette direction, les résultats principaux du cours se trouvent dans le chapitre 3 et ses exercices avec les théorèmes de Sylow qui permettent en particulier de classifier les groupes finis de petit cardinal, dans le chapitre 5 avec un théorème de structure grossier pour les modules sur un anneau principal et la classification complète des groupes abéliens de type fini, et dans le chapitre 7 avec un théorème de structure pour le corps des nombres constructibles à la règle et au compas et la classification des corps finis.

1 Quotients et relations d'équivalences

Une des procédures les plus fondamentales des mathématiques consiste à créer de nouveaux ensembles en identifiant certains éléments d'un ensemble donné. On dit que le nouvel ensemble est un *quotient* de l'ensemble de départ. Par exemple l'ensemble des nombres rationnels est obtenu à partir de l'ensemble des fractions, c'est à dire les paires constituées d'un entier relatif et d'un entier strictement positif, en identifiant (p, q) et (p', q') dès que $pq' = p'q$. Dans cette procédure, il y a des conditions à vérifier pour assurer que ces identifications se comportent bien puis des conditions à vérifier pour construire des fonctions entre ensembles construits par identification. Par exemple il n'est pas du tout évident que l'addition des fractions, qui envoie $((p_1, q_1), (p_2, q_2))$ sur $(p_1q_2 + p_2q_1, q_1q_2)$ se comporte bien lorsque qu'on passe à l'ensemble des nombres rationnels. Ce chapitre est dédié à l'étude générale de cette procédure. Son impact va bien au-delà du cours d'algèbre. Par exemple on peut construire l'ensemble des nombres réels comme quotient d'un ensemble de suites de nombres rationnels, et les espaces L^p de Lebesgue sont construits comme quotients de certains espaces des fonctions.

La définition général de quotient Y d'un ensemble X est d'une simplicité presque suspecte, on demande simplement un moyen de transformer les éléments de X en éléments de Y et que tous les éléments de Y proviennent de X. Le travail consiste ensuite à comprendre quelles sont les identifications autorisées, ce sera l'objet de la définition 1.0.2 et surtout du théorème 1.0.3.

Définition 1.0.1. *Un* quotient (Y, π) *d'un ensemble* X *est un ensemble* Y *muni d'une application surjective* $\pi : X \to Y$.

La synecdoque remplaçant la paire (Y, π) par Y ou par π est courante lorsque le contexte ne laisse pas de place à l'ambigüité. Lorsque Y et π sont clairs, l'image $\pi(x)$ d'un élément x est souvent notée $\bar{x}$ ou $[x]$ ou même… x. Lorsque le contexte identifie clairement Y comme quotient de X, on appelle souvent π la *projection canonique* de X sur Y.

Définition 1.0.2. *Une* relation d'équivalence *est une relation* R *qui est*

- réflexive *:* $\forall x$, xRx
- transitive *:* $\forall x\, y\, z$, xRy *et* $yRz \Rightarrow xRz$
- symétrique *:* $\forall x\, y$, $xRy \Leftrightarrow yRx$.

En pratique on note souvent $\sim$ les relations d'équivalence.

Théorème 1.0.3. *Soit* X *un ensemble.*

- *Pour tout ensemble* Y *et toute fonction* $f : X \to Y$*, la relation sur* X *définie par* xRx' *si* $f(x) = f(x')$ *est une relation d'équivalence, dite* associée à f.
- *Réciproquement, pour toute relation d'équivalence* R *sur* X*, il existe un quotient* (Y, π) *de* X *tel que, pour tous* x *et* x'*,* xRx' *si et seulement si* $\pi(x) = \pi(x')$.

Les deux parties sont légèrement dissymétriques mais la symétrie est restaurée dès que l'on remarque que, dans le premier point, on ne change pas R en remplaçant Y par l'image de f. On verra plus loin que le quotient (Y, π) promis par la deuxième partie est essentiellement unique. Par ailleurs la démonstration construit un quotient particulier : l'ensemble des classes d'équivalences de R, qu'on note souvent X/R et qu'on appelle souvent abusivement *le* quotient de X par R.

Démonstration. Les vérifications du premier point sont très faciles et laissées en exercice.

Pour le deuxième point, on se donne une relation d'équivalence R sur X. On note Y l'image dans l'ensemble $\mathcal{P}(X)$ de l'application $\pi : x \mapsto \{x' \mid xRx'\}$. L'application π est surjective de X dans Y par définition de Y. Soit x et x' dans X. Supposons $\pi(x) = \pi(x')$. Par réflexivité de R, $x'Rx'$, donc $x' \in \pi(x')$ par définition de π puis $x' \in \pi(x)$ par notre hypothèse d'égalité. Par définition de π on obtient alors xRx'. Réciproquement, supposons xRx'. On veut montrer que $\pi(x) = \pi(x')$. Comme R est symétrique, il suffit de montrer que, pour tout x_1 et x_2, $x_1Rx_2 \Rightarrow \pi(x_2) \subset \pi(x_1)$. Soit x_1 et x_2 tels que x_1Rx_2. Soit $x_3 \in \pi(x_2)$. On a x_2Rx_3 par définition de π. La transitivité de R donne alors x_1Rx_3, c'est à dire $x_3 \in \pi(x_1)$. □

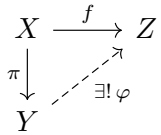
Exemple 1.0.4. Sur $X = \mathbb{Z} \times \mathbb{N}^*$, on considère la relation $(p, q) \sim (p', q')$ si $pq' = p'q$. On vérifie sans peine qu'il s'agit d'une relation d'équivalence. Le quotient $X/\sim$ est appelé ensemble des nombres rationnels et noté $\mathbb{Q}$.

Pour tout ensemble X on peut considérer la relation triviale pour laquelle tout élément n'est équivalent qu'à lui-même. La construction de la démonstration précédente fournit comme quotient $X/\sim = \{\{x\} \ ; \ x \in X\}$ muni de $X \to X/\sim$ qui envoie x sur $\{x\}$. Mais bien sûr X lui-même, muni de l'application Id_X, est aussi un quotient de X pour cette relation d'équivalence.

Théorème 1.0.5 (Propriété universelle des quotients)**.** *Soit $\pi : X \to Y$ un quotient de relation associée $\sim$. Soit $f : X \to Z$. Si*

$$\forall x\ x',\ x \sim x' \Rightarrow f(x) = f(x')$$

alors il existe une unique fonction $\varphi : Y \to Z$ telle que $f = \varphi \circ \pi$. Autrement dit, on a le diagramme commutatif



L'application φ est surjective si et seulement si f l'est. Elle est injective si et seulement si $\forall x\ x', x \sim x' \Leftrightarrow f(x) = f(x')$.
Réciproquement, si φ existe alors $\forall x\ x', x \sim x' \Rightarrow f(x) = f(x')$.

Dans le théorème, le terme « diagramme commutatif » signifie qu'à chaque fois qu'on peut aller d'un ensemble à un autre en suivant les flèches, les composées des applications correspondantes sont égales. Lorsque la condition du théorème est vérifiée, on dit que f est *compatible avec la relation d'équivalence* où encore que f *descend au quotient* en φ, ou induit φ au quotient.

Démonstration. Supposons que f est compatible avec $\sim$. Montrons l'existence de φ. Comme π est surjective, l'axiome du choix fournit $\sigma : Y \to X$ telle que $\pi \circ \sigma = \mathrm{Id}$. On pose $\varphi = f \circ \sigma$. Montrons que $\varphi \circ \pi = f$. Soit x dans X. On a $\varphi \circ \pi(x) = (f \circ \sigma) \circ \pi(x) = f(\sigma(\pi(x))$. Or $\pi(\sigma(\pi(x)) = \pi \circ \sigma(\pi(x)) = \pi(x)$ donc $\sigma(\pi(x)) \sim x$. L'hypothèse sur f assure alors que $f(\sigma(\pi(x)) = f(x)$.

Montrons maintenant l'unicité. Soit φ et φ' telles que que $\varphi \circ \pi = \varphi' \circ \pi = f$. Soit y dans Y. Comme π est surjective, on obtient x tel que $y = \pi(x)$. On a alors $\varphi(y) = \varphi(\pi(x)) = \varphi'(\pi(x)) = \varphi'(y)$.

Supposons φ surjective et montrons que f l'est. Soit z dans Z. Par surjectivité de φ, on obtient y tel que $\varphi(y) = z$. Par surjectivité de π on obtient x tel que $\pi(x) = y$ et on a $f(x) = \varphi(\pi(x)) = \varphi(y) = z$ donc $z \in \operatorname{im} f$. Réciproquement supposons que f est surjective. Soit z dans Z. Par hypothèse on obtient x tel que $f(x) = z$. On a alors $\varphi(\pi(x)) = f(x) = z$ donc $z \in \operatorname{im} \varphi$.

On a déjà supposé $\forall x\ x', x \sim x' \Rightarrow f(x) = f(x')$. On a :

$$\begin{aligned}
\varphi \text{ injective} &\Leftrightarrow \forall y\ y', \varphi(y) = \varphi(y') \Rightarrow y = y' \\
&\Leftrightarrow \forall x\ x', \varphi(\pi(x)) = \varphi(\pi(x')) \Rightarrow \pi(x) = \pi(x') \text{ car } \pi \text{ surj} \\
&\Leftrightarrow \forall x\ x', f(x) = f(x') \Rightarrow \pi(x) = \pi(x') \text{ car } f = \varphi \circ \pi \\
&\Leftrightarrow \forall x\ x', f(x) = f(x') \Rightarrow x \sim x'
\end{aligned}$$

Réciproquement, supposons que f s'écrive sous la forme $\varphi \circ \pi$. Soit x et x' tels que $x \sim x'$. On a alors $\pi(x) = \pi(x')$ donc $f(x) = \varphi(\pi(x)) = \varphi(\pi(x')) = f(x')$. □

Remarque 1.0.6. Dans la démonstration d'existence, l'application σ est loin d'être unique en général, mais la partie unicité du théorème montre que cela n'a pas d'importance. Cette démonstration d'existence est souvent racontée de la façon imagée suivante. Pour chaque y dans Y, on veut définir $\varphi(y)$. Par surjectivité de π, il existe x tel que $y = \pi(x)$. On « pose $\varphi(y) = f(x)$ et on vérifie que le résultat ne dépend pas de x parmi les préimages de y ». La phrase entre guillemets ne veut rien dire mais cela reste un bon moyen de se former une intuition. Il y a de nombreuses variantes encore plus poétiques de cette explication, comme par exemple « on pose $\varphi([x]) = f(x)$ et on vérifie que φ est bien définie ».

On peut aussi noter que, selon les fondements choisis, l'axiome du choix peut être évité dans la démonstration ci-dessus. Par exemple en théorie des ensembles de Zermelo-Fraenkel, tout graphe défini une fonction (sans utiliser d'axiome supplémentaire) et on peut utiliser la surjectivité de π et l'hypothèse sur f pour montrer que

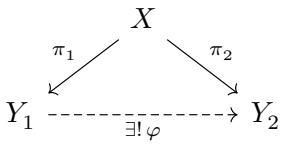
$$\{(y, z) \mid \forall x \in X, \pi(x) = y \Rightarrow f(x) = z\}$$

est un graphe. Mais la preuve donnée plus haut me semble bien plus intuitive et on ne s'occupe pas de fondements dans ce cours.

Exemple 1.0.7. Soit T un réel strictement positif. On considère la relation d'équivalence sur $\mathbb{R}$ définie par $t \sim t'$ si $t' - t \in T\mathbb{Z}$. Une fonction définie sur $\mathbb{R}$ descend au quotient si et seulement si elle est T-périodique.

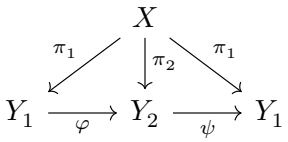
La propriété universelle implique en particulier que le quotient associé à une relation d'équivalence est essentiellement unique, au sens précis de l'énoncé suivant.

Corollaire 1.0.8. *Si $\pi_1 : X \to Y_1$ et $\pi_2 : X \to Y_2$ sont deux quotients associés à la même relation d'équivalence $\sim$ alors il existe une unique bijection $\varphi : Y_1 \to Y_2$ telle que $\pi_2 = \varphi \circ \pi_1$.*



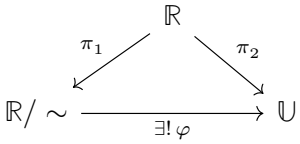
Démonstration. Puisque la fonction π_2 est associée à $\sim$, elle est en particulier compatible avec $\sim$ donc descend au quotient en une unique application $\varphi : Y_1 \to Y_2$ telle que $\pi_2 = \varphi \circ \pi_1$. Il reste à montrer que φ est une bijection. De même π_1 descend au quotient en $\psi : Y_2 \to Y_1$ telle que $\pi_1 = \psi \circ \pi_2$.

On contemple donc le diagramme suivant :



Comme les deux petits triangles commutent, tout le diagramme commute. En effet $(\psi \circ \varphi) \circ \pi_1 = \psi \circ (\varphi \circ \pi_1) = \psi \circ \pi_2 = \pi_1$. Or l'application Id_{Y_1} vérifie aussi cette relation : $\mathrm{Id}_{Y_1} \circ \pi_1 = \pi_1$. Par unicité dans la propriété universelle de (Y, π_1) appliquée à π_1, on obtient donc $\psi \circ \varphi = \mathrm{Id}_{Y_1}$. On montre de même que $\varphi \circ \psi = \mathrm{Id}_{Y_2}$. □

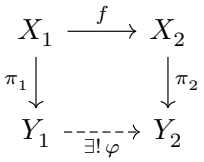
Exemple 1.0.9. Soit T un réel strictement positif. L'application de $\mathbb{R}$ dans $\mathbb{U}$ (le groupe des nombres complexes de module 1) définit par $t \to \exp(2i\pi t/T)$ est un quotient. La relation d'équivalence associée est $t \sim t'$ si $t - t' \in T\mathbb{Z}$. On notera que $\mathbb{U}$ n'est pas égal à $\mathbb{R}/\sim$ l'ensemble des classes d'équivalence pour $\sim$, mais le corollaire 1.0.8 assure qu'il existe une unique bijection φ qui fait commuter le diagramme suivant :



On a vu plus haut que les fonctions qui descendent au quotient par $\sim$ sont les fonctions T-périodiques. Comme $\mathbb{U}$ est un quotient de $\mathbb{R}$ par $\sim$, cet exemple explique précisement en quel sens les fonctions périodiques sur $\mathbb{R}$ peuvent être vues comme des fonctions définies sur un cercle.

Le corollaire suivant de la propriété universelle des quotients explique comment définir une application entre deux quotients.

Corollaire 1.0.10. *Soit $\pi_1 : X_1 \to Y_1$ et $\pi_2 : X_2 \to Y_2$ deux quotients et $f : X_1 \to X_2$. Si $\forall x_1\ x_1', x_1 \sim x_1' \Rightarrow f(x_1) \sim f(x_1')$ alors il existe une unique application $\varphi : Y_1 \to Y_2$ telle que $\pi_2 \circ f = \varphi \circ \pi_1$.*



Réciproquement si φ existe alors $\forall x_1\ x_1', x_1 \sim x_1' \Rightarrow f(x_1) \sim f(x_1')$.

Démonstration. Il suffit d'appliquer le théorème à $\pi_2 \circ f$ car, pour tous x_1 et x_1' dans X_1, $f(x_1) \sim f(x_1') \Leftrightarrow \pi_2 \circ f(x_1) = \pi_2 \circ f(x_1')$. □

Remarque 1.0.11. Dans l'énoncé précédent, le même symbole $\sim$ est utilisé pour les relations d'équivalence sur X_1 et X_2. On pourrait utiliser $\sim_1$ et $\sim_2$ mais en pratique il n'y a aucune ambigüité. On sait que f est une fonction de X_1 dans X_2. Dans l'expression $\forall x_1\ x_1', x_1 \sim x_1' \Rightarrow f(x_1) \sim f(x_1')$, le fait qu'on applique f à x_1 et x_2 force x_1 et x_2 à vivre dans X_1 donc la quantification universelle porte sur les éléments de X_1, le premier $\sim$ est celui de X_1 et le second est celui de X_2. Dans les chapitres qui suivent, ce type d'inférence sera aussi utilisé constamment pour interpréter tous les symboles de lois de composition internes.

Exemple 1.0.12. L'application $(p, q) \mapsto (-p, q)$ de $\mathbb{Z} \times \mathbb{N}^*$ dans lui-même descend en application de $\mathbb{Q}$ dans lui-même.

Un autre cas courant d'utilisation de la propriété universelle consiste à définir une application d'un produit de quotients vers un quotient (par exemple une loi de composition interne sur un quotient).

Corollaire 1.0.13. *Soit $\pi_1 : X_1 \to Y_1$, $\pi_2 : X_2 \to Y_2$ et $\pi_3 : X_3 \to Y_3$ trois quotients et $f : X_1 \times X_2 \to X_3$. Si*

$$\forall x_1\ x_1'\ x_2\ x_2',\ [x_1 \sim x_1' \text{ et } x_2 \sim x_2'] \Rightarrow f(x_1, x_2) \sim f(x_1', x_2')$$

alors il existe une unique application $\varphi : Y_1 \times Y_2 \to Y_3$ telle que $\pi_3 \circ f = \varphi \circ (\pi_1 \times \pi_2)$.

$$\begin{array}{ccc} X_1 \times X_2 & \xrightarrow{f} & X_3 \\ {\scriptstyle \pi_1 \times \pi_2}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ Y_1 \times Y_2 & \xrightarrow[\exists ! \varphi]{} & Y_3 \end{array}$$

Démonstration. Comme π_1 and π_2 sont surjectives, $\pi_1 \times \pi_2$ l'est aussi donc $\pi_1 \times \pi_2 : X_1 \times X_2 \to Y_1 \times Y_2$ est un quotient de $X_1 \times X_2$ et la relation d'équivalence associée est $(x_1, x_2) \sim (x_1', x_2')$ si $x_1 \sim x_1'$ et $x_2 \sim x_2'$. On est donc ramené au corollaire 1.0.10. □

Exemple 1.0.14. La loi de composition interne sur $\mathbb{Z} \times \mathbb{N}^*$ qui envoie $((p_1, q_1), (p_2, q_2))$ sur $(p_1q_2 + p_2q_1, q_1q_2)$ descend au quotient en loi de composition interne sur $\mathbb{Q}$ (la vérification de la condition de l'énoncé est laissée en exercice). On appelle addition des nombres rationnels la loi ainsi définie.

Remarque 1.0.15. Dans la démonstration précédente, on notera l'intérêt de ne pas se limiter à la construction explicite d'un quotient comme ensemble de classes d'équivalences (comme dans la démonstration du théorème 1.0.3). En effet, si on suppose que Y_1 (resp. Y_2) est l'ensemble des classes d'équivalences de X_1 (resp. X_2) alors $Y_1 \times Y_2$ n'est *pas* l'ensemble des classes d'équivalences de $X_1 \times X_2$. Par exemple si $X_1 = \{a\}$ et $X_2 = \{b\}$ (nécessairement munis de la relation d'équivalence triviale), on a $Y_1 \times Y_2 = \{(\{a\}, \{b\})\}$ tandis que $(X_1 \times X_2)/\sim = \{\{(a, b)\}\}$.

On a vu dans les théorèmes 1.0.3 et 1.0.8 que les quotients et les relations d'équivalence décrivent essentiellement le même objet mathématique. Cet objet a une troisième incarnation qui est parfois utile, la notion de partition d'un ensemble.

Définition 1.0.16. *Une partition d'un ensemble X est une collection de parties non-vides de X dont la réunion est X et qui sont deux à deux disjointes.*

Lemme 1.0.17. *Pour tout quotient $\pi : X \to Y$, l'ensemble $\{\pi^{-1}(y) \ ; \ y \in Y\}$ est une partition de X. Réciproquement, si $\{X_i \ ; \ i \in I\}$ est une partition de X alors l'application π de X dans I qui à x associe l'unique i tel que $x \in X_i$ est un quotient.*

Démonstration. Soit $\pi : X \to Y$ un quotient de X. Comme tout élément de X a une image par π, X est la réunion des $\pi^{-1}(y)$. Comme π est surjective, tous les $\pi^{-1}(y)$ sont non-vides. Comme chaque élément de x n'a qu'une image par π, ils sont deux à deux disjoints.

Pour la réciproque, on observe d'une part que π est bien définie en utilisant que les X_i recouvrent X et sont disjoints et d'autre part qu'elle est surjective car chaque X_i est non-vide. □

On a donc trois points de vue essentiellement équivalents. Par ordre décroissant d'importance générale, il s'agit de la notion de quotient, la notion de relation d'équivalence et la notion de partition. La prééminence de la notion de quotient peut choquer car le quotient n'est défini que modulo un unique isomorphisme mais il faut simplement s'habituer au fait que, dans bien des contextes, il s'agit d'une unicité plus naturelle que l'égalité.

On termine par un lemme technique qui sera parfois utile dans la suite.

Lemme 1.0.18. *Soit $\pi : X \to Y$ un quotient.*

- $\forall B \in \mathcal{P}(Y), \pi(\pi^{-1}(B)) = B$
- $\pi^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ *est injective*

Démonstration. Soit B une partie de Y. On a $\pi(\pi^{-1}(B)) \subset B$ sans utiliser la surjectivité de π. Montrons l'autre inclusion. Soit b dans B. La surjectivité de π fournit $x \in X$ tel que $\pi(x) = b$. Comme $b \in B$, $x \in \pi^{-1}(B)$ donc $\pi(x) \in \pi(\pi^{-1}(B))$, c-à-d $b \in \pi(\pi^{-1}(B))$. Soit B et B' des parties de Y telles que $\pi^{-1}(B) = \pi^{-1}(B')$. En appliquant π et le point précédent, on obtient $B = B'$. □

2 Monoïdes

Définition 2.0.1. *Un* monoïde *est un ensemble M muni d'un loi de composition interne, disons notée multiplicativement, et d'un élément e tels que :*

- *la multiplication est associative : $\forall x\ y\ z, x(yz) = (xy)z$*
- *e est neutre à gauche : $\forall x, ex = x$*
- *e est neutre à droite : $\forall x, xe = x$*

On dit qu'un monoïde est commutatif *si $\forall x\, y, xy = yx$.*

On utilise aussi le symbole d'addition pour la loi de composition interne de certains monoïdes commutatifs (mais jamais si la loi n'est pas commutative). L'élément neutre e est souvent noté 1 ou 1_M en notation multiplicative ou bien 0 ou 0_M en notation additive.

Lorsque le contexte est clair, on utilise presque toujours la synecdoque « Soit M un monoïde » plutôt que « Soit $(M, \times, 1)$ un monoïde ».

La structure de monoïde est trop pauvre pour fournir à elle seule une théorie intéressante. On ne l'introduit que pour regrouper des arguments communs à la théorie des groupes et des anneaux ainsi que pour des constructions intermédiaires. Cependant on peut déjà donner des exemples naturels de monoïdes qui ne sont pas des groupes.

Exemple 2.0.2. $(\mathbb{N}, +, 0)$, $(\mathbb{N}, \times, 1)$ et $(\mathbb{Z}, \times, 1)$ sont des monoïdes commutatif. Pour tout ensemble X, l'ensemble des fonctions de X dans X, muni de la composition et de l'identité, est un monoïde qui n'est commutatif que si X a au plus un élément.

Définition 2.0.3. *Soit M un monoïde et x un élément de M. On dit que x est* simplifiable à gauche *si $y \mapsto xy$ est injective de M dans M. On dit qu'il est* simplifiable à droite *si $y \mapsto yx$ est injective. On dit que x est* inversible *s'il existe x' tel que $xx' = x'x = 1$. On dit aussi que x est une* unité *de M. L'ensemble des unités de M est noté $M^\times$.*

Lemme 2.0.4. *Soit M un monoïde. Si $x \in M$ est inversible alors il n'existe qu'un seul élément x' tel que $xx' = 1$ ou $x'x = 1$. On l'appelle l'inverse de x et on le note x^{-1} (ou $-x$ en notation additive). On a $xx^{-1} = x^{-1}x = 1$. Dans ce cas x est simplifiable à gauche et à droite.*

Démonstration. Soit x un élément inversible de M. L'hypothèse fournit x' tel $xx' = x'x = 1$. Si $xx'' = 1$ alors on calcule $x'' = 1x'' = (x'x)x'' = x'(xx'') = x'1 = x'$. De même si $x''x = 1$ on peut calculer $x'' = x''1 = x''(x'x) = (x''x)x' = 1x' = x'$. On peut maintenant utiliser la notation x^{-1}. L'élément x est simplifiable à gauche car la multiplication à gauche par x^{-1} est inverse de la multiplication à gauche par x, et de même pour la droite. □

Exemple 2.0.5. L'élément neutre d'un monoïde est toujours inversible : il est son propre inverse. Dans $(\mathbb{N}, \times)$, le neutre 1 est le seul inversible. Dans $(\mathbb{Q}, \times)$ tous les éléments sauf

zéro sont inversibles. Les inversibles dans les fonctions d'un ensemble dans lui-même sont les bijections.

L'exemple précédent montre qu'il ne faut pas confondre les notations $\mathbb{N}^\times$ et $\mathbb{N}^*$, même si elles coïncident dans le cas de $\mathbb{Q}$, ou plus généralement dans le cas des corps.

Définition 2.0.6. *Un* morphisme de monoïdes *entre M et N est une application $f : M \to N$ telle que :*

- $f(1) = 1$
- $\forall x\, y, f(xy) = f(x)f(y)$.

Lemme 2.0.7. *Soit $f : M \to N$ un morphisme de monoïdes.*

- $\forall x \in M^\times, f(x) \in N^\times$ *et* $f(x)^{-1} = f(x^{-1})$.
- *Si f est bijectif alors f^{-1} est automatiquement un morphisme de monoïdes. On dit alors que f est un isomorphisme entre M et N.*

De plus une composée de morphismes de monoïdes est un morphisme de monoïdes.

Démonstration. Soit x dans $M^\times$. On a $f(x^{-1})f(x) = f(x^{-1}x) = f(1_M) = 1_N$. On montre de même que $f(x)f(x^{-1}) = 1_N$. Ainsi $f(x)$ est inversible, d'inverse $f(x^{-1})$.

Supposons f bijectif. Soit x et y dans N. On calcule

$$\begin{aligned} f^{-1}(xy) &= f^{-1}\Big(f(f^{-1}(x))f(f^{-1}(y))\Big) \\ &= f^{-1}\Big(f(f^{-1}(x)f^{-1}(y))\Big) \\ &= f^{-1}(x)f^{-1}(y). \end{aligned}$$

□

Soit $f : M \to N$ et $g : N \to P$ des morphismes des monoïdes. Soit x et y dans M. On a $g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y))$ et $g(f(1)) = g(1) = 1$ donc $g \circ f$ est un morphisme de monoïdes.

Définition 2.0.8. *Un* quotient *d'un monoïde M est un monoïde N muni d'un morphisme de monoïdes $\pi : M \to N$ surjectif.*

Lemme 2.0.9. *Soit M un monoïde et $\pi : M \to N$ un quotient d'ensembles. Soit $\sim$ la relation d'équivalence associée à π. Il existe une structure de monoïde sur N telle que π soit un morphisme de monoïde si et seulement si la loi de composition interne de M est compatible avec la relation d'équivalence produit sur $M \times M$. La structure de monoïde sur N est alors unique. De plus si M est commutatif alors N l'est aussi.*

Pour tout quotient de monoïde $\pi : M \to N$ et tout morphisme de monoïdes $\varphi : M \to P$ qui est compatible avec $\sim$, la fonction $\bar{\varphi} : N \to P$ induite par φ est automatiquement un morphisme de monoïdes.

Démonstration. Supposons qu'on a une loi de composition interne sur N. L'application π est un morphisme si et seulement si le diagramme suivant commute (les flèches horizontales représentent les loi de composition) :

$$\begin{array}{ccc} M \times M & \longrightarrow & M \\ \downarrow{\scriptstyle \pi\times\pi} & & \downarrow{\scriptstyle \pi} \\ N \times N & \longrightarrow & N \end{array}$$

Ainsi la structure de monoïde cherchée ne peut exister que si la loi de M descend au quotient. Le corollaire 1.0.13 donne le critère d'existence annoncé et assure l'unicité de la loi descendue.

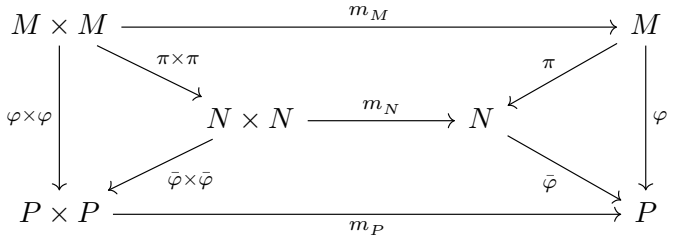
On suppose maintenant cette condition de compatibilité. On pose $1_N = \pi(1_M)$. Il s'agit bien d'un élément neutre dans N car tout élément de N est de la forme $\pi(m)$ et $1_N\pi(m) = \pi(1_M)\pi(m) = \pi(1_M m) = \pi(m)$ et de même pour $\pi(m)1_N$. Dans la suite on notera simplement 1 les neutres de M et de N.

Pour l'associativité, la surjectivité de π montre qu'il suffit de calculer, pour tout x, y et z dans M,

$$(\pi(x)\pi(y))\pi(z) = \pi(xy)\pi(z) = \pi((xy)z) = \pi(x(yz)) = \pi(x)\pi(yz) = \pi(x)(\pi(y)\pi(z)).$$

Supposons un instant que M est commutatif et montrons que N l'est aussi. La surjectivité de π montre qu'il suffit de calculer, pour tous x et y dans M, $\pi(x)\pi(y) = \pi(xy) = \pi(yx) = \pi(y)\pi(x)$.

Supposons maintenant que $\varphi\colon M \to P$ est un morphisme de monoïdes compatible avec $\sim$, de sorte qu'il descend en fonction $\bar{\varphi}\colon N \to P$. On contemple le diagramme suivant où m_M, m_P et m_N sont les multiplications de M, P et N respectivement.



Comme φ est un morphisme, le grand rectangle commute. Par hypothèse sur φ, les deux petits triangles commutent. Comme π est un morphisme, le trapèze du haut commute. Le but est de montrer que le trapèze du bas commute : $m_P \circ (\bar{\varphi} \times \bar{\varphi}) = \bar{\varphi} \circ m_N$. Comme $\pi \times \pi$ est surjective, il suffit de vérifier cette égalité précomposée par $\pi \times \pi$. On calcule en regardant le diagramme et en tenant compte des commutations déjà établies.

$$\begin{aligned} m_P \circ (\bar{\varphi} \times \bar{\varphi}) \circ (\pi \times \pi) &= m_P \circ (\varphi \times \varphi) \\ &= \varphi \circ m_M \\ &= \bar{\varphi} \circ \pi \circ m_M \\ &= \bar{\varphi} \circ m_N \circ (\pi \times \pi). \end{aligned}$$

De plus $\bar{\varphi}(1) = \bar{\varphi}(\pi(1)) = \varphi(1) = 1$. □

Exemple 2.0.10. L'ensemble $\mathbb{N} \times \mathbb{N}$, munie de l'addition composante par composante est un monoïde (noté additivement), de neutre $(0, 0)$. La relation d'équivalence définie par $(a, b) \sim (a', b')$ si $a + b' = a' + b$ est compatible avec l'addition. Le lemme précédent assure que le quotient $(\mathbb{N} \times \mathbb{N})/\sim$, noté $\mathbb{Z}$, est muni d'une unique structure de monoïde additif telle que la projection canonique π est un morphisme. L'application $j\colon \mathbb{N} \to \mathbb{Z}$ qui envoie n sur $\pi(n, 0)$ est un morphisme de monoïdes (comme composée de morphismes) appelé inclusion de $\mathbb{N}$ dans $\mathbb{Z}$. On vérifie facilement que tous les éléments de $\mathbb{Z}$ sont inversibles (toujours pour l'addition). Cet exemple sera considérablement généralisé dans le chapitre 4.

Définition 2.0.11. *Un* sous-monoïde *d'un monoïde M est une partie N de M telle que :*

- $1 \in N$
- *N est stable par multiplication :* $\forall x\, y \in M, x \in N$ *et* $y \in N \Rightarrow xy \in N$

On note alors $N \leqslant M$.

Exemple 2.0.12. Pour tout monoïde M, $\{1\}$ et M sont des sous-monoïdes de M. Dans $(\mathbb{Z}, \times)$, $\mathbb{N}$ est un sous-monoïde.

La définition ci-dessus est écrite pour assurer que la structure de monoïde sur M se restreint en structure de monoïde sur N.

Lemme 2.0.13. *Soit M et M' des monoïdes, $N \subset M$, $N' \subset M'$ et $f : M \to M'$ un morphisme.*

- *Si $N \leqslant M$ alors $f(N) \leqslant M'$. En particulier* $\mathrm{im}(f) \leqslant M'$.
- *Si $N' \leqslant M'$ alors $f^{-1}(N') \leqslant M$.*
- *L'intersection d'une famille de sous-monoïdes de M est un sous-monoïde de M.*

Démonstration. Soit N un sous-monoïde de M. Comme $1 \in N$, on obtient $f(1) \in f(N)$, c'est à dire $1 \in f(N)$. Soit x et y dans $f(N)$. On obtient x_0 et y_0 tels que $x = f(x_0)$ et $y = f(y_0)$. On a alors $xy = f(x_0)f(y_0) = f(x_0y_0)$ donc $xy \in f(N)$.

Soit N' un sous-monoïde de M'. Comme $1 \in N'$ et que $f(1) = 1$, $1 \in f^{-1}(N')$. Soit x et y dans $f^{-1}(N')$. Ainsi $f(x)$ et $f(y)$ sont dans N' donc $f(x)f(y)$ aussi, c'est à dire $f(xy) \in N'$.

Soit $\mathcal{H}$ une famille de sous-monoïde de M et K l'intersection des éléments de $\mathcal{H}$. Tous les éléments de $\mathcal{H}$ contiennent 1 donc leur intersection K aussi. Soit x et y dans K. Pour tout $N \in \mathcal{H}$, x et y sont dans N donc xy aussi. Ainsi xy est dans K. □

Remarque 2.0.14 (Intermède vacuiste). Le dernier point du lemme 2.0.13 autorise l'intersection de la famille vide de sous-monoïdes de M. De quel sous-monoïde s'agit-il ? On pourrait imaginer que ce soit une question de convention, mais alors on pourrait légitimement douter de la validité de la démonstration dans ce cas. En fait il n'y a aucun problème de définition ici. Tous les ensembles intervenant dans cette démonstration sont des parties de M et l'opération d'intersection est parfaitement définie de $\mathcal{P}(\mathcal{P}(M))$ dans $\mathcal{P}(M)$ par la règle

$$\forall x \in M, \left(x \in \bigcap_{N \in \mathcal{H}} N \Leftrightarrow \forall N \in \mathcal{H}, x \in N \right)$$

qui stipule sans aucune ambigüité que $\bigcap_{N \in \varnothing} N = M$. On peut aussi vérifier que cette intersection vérifie bien la propriété universelle attendue d'un opérateur de borne inférieure : $\forall K \in \mathcal{P}(M), (K \subset \bigcap_{N \in \mathcal{H}} N \Leftrightarrow \forall N \in \mathcal{H}, K \subset N)$.

Définition 2.0.15. *Le sous-monoïde engendré par une partie S d'un monoïde M est l'intersection de tous les sous-monoïdes de M contenant S :*

$$\langle S \rangle = \bigcap_{N \leqslant M,\, S \subset N} N.$$

Si $\langle S \rangle = M$, on dit que S engendre *M, ou bien que S est une* partie génératrice *de M.*

Le lemme suivant rassemble les propriétés formelles des sous-monoïdes engendrés.

Lemme 2.0.16. *Soit S une partie d'un monoïde M.*

- *$\langle S\rangle$ est un sous-monoïde de M qui contient S.*
- *Pour tout $N \leqslant M$, $\langle S\rangle \leqslant N \Leftrightarrow S \subset N$ (ainsi $\langle S\rangle$ est le plus petit sous-monoïde de M qui contient S). Cette propriété universelle caractérise $\langle S\rangle$.*
- *L'application $\langle\cdot\rangle$ est croissante : $\forall S\ S', S \subset S' \Rightarrow \langle S\rangle \subset \langle S'\rangle$.*
- *Pour tout morphisme $f\colon M \to M'$, $\langle f(S)\rangle = f(\langle S\rangle)$.*

Démonstration. Le premier point découle directement du lemme 2.0.13 et le second de la propriété universelle de l'intersection (cf. remarque 2.0.14). Montrons que la propriété universelle caractérise $\langle S\rangle$. Supposons qu'un sous-monoïde K vérifie aussi cette propriété. Comme K contient S, la propriété universelle de $\langle S\rangle$ assure que $\langle S\rangle \subset K$. De même, comme $\langle S\rangle$ contient S, la propriété de K assure que $K \subset \langle S\rangle$.

Les troisième et quatrième points découlent de façons complètement formelle des deux premiers (bien sûr on peut aussi donner une démonstration à partir de la définition, mais les raisonnements suivants sont des exemples d'énoncés bien plus généraux portant sur les applications entre ensembles ordonnés et peuvent être copiés sans aucune altération lors de l'étude des sous-groupes, sous-anneaux, idéaux ou sous-algèbres engendrés par une partie).

Pour le troisième point, soit S et S' des parties de M avec $S \subset S'$. Le premier point donne $S' \subset \langle S'\rangle$ donc $S \subset \langle S'\rangle$ par transitivité puis $\langle S\rangle \subset \langle S'\rangle$ par le second point.

Montrons le dernier point. Il suffit de montrer que $f(\langle S\rangle)$ vérifie la propriété universelle de $\langle f(S)\rangle$. Tout d'abord il s'agit bien d'un sous-monoïde d'après le lemme 2.0.13. Soit N' un sous-monoïde de M'. Le même lemme 2.0.13 assure que $f^{-1}(N')$ est un sous-monoïde de N. On a donc :

$$\begin{aligned} f(S) \subset N' &\Leftrightarrow S \subset f^{-1}(N') \\ &\Leftrightarrow \langle S\rangle \subset f^{-1}(N') \text{ d'après la prop univ de } \langle S\rangle \\ &\Leftrightarrow f(\langle S\rangle) \subset N' \end{aligned}$$

Ce qui est bien la propriété universelle de $\langle f(S)\rangle$. □

Lorsque tout le reste a échoué, on peut aussi utiliser la description explicite du lemme suivant.

Lemme 2.0.17. *Soit S une partie d'un monoïde M. Les éléments du sous-monoïde engendré par S sont tous les produits (finis) d'éléments de S :*

$$\langle S\rangle = \Big\{ s_1 \cdots s_k \ ;\ k \in \mathbb{N}, s\colon \{1, \dots, k\} \to S \Big\}.$$

où le cas $k = 0$ correspond au produit vide, c'est à dire au neutre de M.

Démonstration. L'ensemble N du membre de droite contient 1 par définition et il est stable par multiplication. C'est donc un sous-monoïde qui contient S. De plus tout sous-monoïde contenant S contient nécessairement N car il doit contenir 1 et S et être stable par multiplication. Donc N vérifie la propriété universelle de $\langle S\rangle$ donc $N = \langle S\rangle$ d'après le lemme précédent. □

3 Groupes

3.1 Définitions, morphismes et sous-objets

Définition 3.1.1. *Un* groupe *est un monoïde dans lequel tous les éléments sont inversibles. Lorsque la loi de composition est commutative, on dit que le groupe est* abélien.

Un morphisme *entre deux groupes G et G' est une fonction $f : G \to G'$ telle que, pour tous x et y dans G, $f(xy) = f(x)f(y)$ (on montrera plus bas que c'est automatiquement un morphisme de monoïdes).*

L'exemple le plus fondamental de groupe est l'ensemble $\mathfrak{S}(E)$ des permutations d'un ensemble E. Plus généralement, l'ensemble des unités d'un monoïde forme toujours un groupe.

Si on déplie toutes les définitions intervenant dans la définition de groupe, on voit qu'un groupe est un ensemble G muni d'une loi de composition interne qu'on notera multiplicativement pour cette discussion, d'un élément $e \in G$ et d'une fonction $x \mapsto x^{-1}$ de G dans lui même tel que :

- la multiplication est associative : $\forall x\ y\ z, x(yz) = (xy)z$
- e est neutre à gauche : $\forall x, ex = x$
- e est neutre à droite : $\forall x, xe = x$
- $^{-1}$ est une inversion à gauche : $\forall x, x^{-1}x = e$.
- $^{-1}$ est une inversion à droite : $\forall x, xx^{-1} = e$.

Il existe de nombreuses variantes équivalentes de la liste ci-dessus. La variante choisie est très symétrique et donne d'emblée toute la structure, ce qui permet d'éviter les empilements de quantificateurs qui arrivent quand e et $^{-1}$ ne sont pas introduits a priori et que les axiomes au-delà de l'associativité se transforment en

$$\exists e, \left[(\forall x, ex = x \text{ et } xe = x) \text{ et } (\forall x, \exists x', x'x = e \text{ et } xx' = e)\right].$$

Cette variante peut sembler plus symétrique car elle ne fixe pas de choix de e et $^{-1}$ mais c'est une illusion comme le montre le lemme suivant qu'on pourra parfois utiliser pour raccourcir certaines vérifications (mais ce genre de raccourci n'est jamais indispensable et le lemme suivant est surtout une curiosité dont la lecture n'est vraiment pas indispensable).

Lemme 3.1.2. *Soit G un ensemble muni d'une loi de composition interne associative. Si*

$$\exists e,\ \forall x,\ (ex = x \text{ et } \exists y, yx = e)$$

alors il existe un unique e et une unique fonction $^{-1}$ tels que G muni de sa loi de composition, de e et de $^{-1}$ est un groupe.

Démonstration. Fixons e promis par la condition. Montrons d'abord que

$$\forall x,\ \exists y,\ yx = e \text{ et } xy = e.$$

Soit x dans G. Par hypothèse sur e appliquée à x on obtient y tel que $yx = e$. Par hypothèse sur e appliquée à y on obtient z tel que $zy = e$. On calcule alors, en utilisant aussi que $\forall w, ew = w$ et l'associativité,

$$xy = (ex)y = ((zy)x)y = z(yx)y = zey = zy = e.$$

Montrons maintenant $\forall w, we = w$. Soit w. On obtient par hypothèse un y tel que $yw = e$ et le paragraphe précédent assure qu'on a aussi $wy = e$. On calcule alors $we = w(yw) = (wy)w = ew = w$.

On en déduit immédiatement l'unicité de e. En effet si e' convient aussi, la propriété de e' donne $e'e = e$ et le paragraphe précédent donne $e'e = e'$.

Montrons que, $\forall x, \exists! y, xy = e$ et $yx = e$. Soit x dans G. On a déjà l'existence de y. Supposons que y' convient aussi. On a alors $xy = xy'$ qu'on multiplie à gauche à par y pour obtenir $y = y'$. On obtient ainsi une unique fonction d'inversion. □

Au vu de ce lemme, on peut être tenté de radiner sur les axiomes dans la définition de groupe. Les lecteurs et lectrices tentés par cette mesquinerie sont invités à s'auto-punir en essayant de montrer que, pour tout ensemble G non vide muni d'une loi de composition interne et d'une fonction ${}^{-1} : G \to G$, il existe $e \in G$ tel que $(G, \cdot, {}^{-1}, e)$ est un groupe si et seulement si

$$\forall u\ x\ y\ z, \quad x\Big(y\Big(((zz^{-1})(uy)^{-1})x\Big)\Big)^{-1} = u.$$

Le lemme suivant liste quelques propriétés fondamentales des morphismes de groupes. On pourrait argumenter que les deux premiers items pourraient faire partie de la définition plutôt que de radiner mais cela permet de souligner le contraste avec les propriétés des morphismes multiplicatifs en général.

Lemme 3.1.3. *Soit $f : G \to G'$ un morphisme de groupes.*

- $f(1_G) = 1_{G'}$ *(f est donc un morphisme de monoïdes)*
- $\forall x, f(x^{-1}) = f(x)^{-1}$.
- *f est injectif si et seulement si* $\ker(f) = \{1_G\}$ *où, par définition,* $\ker(f) = f^{-1}(\{1_{G'}\})$.
- *Si f est bijectif alors f^{-1} est automatiquement un morphisme de groupes. On dit alors que f est un isomorphisme entre G et G′.*

Démonstration. Pour le premier point, on commence par noter que $1_G 1_G = 1_G$ donc on obtient $f(1_G)f(1_G) = f(1_G) = f(1_G)1_{G'}$. Comme $f(1_G)$ est inversible (car tous les éléments de G' le sont), il est simplifiable à gauche d'après le lemme 2.0.4 donc $f(1_G) = 1_{G'}$. Ainsi f est un morphisme de monoïdes donc le lemme 2.0.7 et le premier point règlent les deuxième et quatrième points.

Montrons le troisième point. Soit x et y dans G. On a

$$\begin{aligned} f(x) = f(y) &\Leftrightarrow f(x)f(y)^{-1} = 1_{G'} \\ &\Leftrightarrow f(xy^{-1}) = 1_{G'} \\ &\Leftrightarrow xy^{-1} \in \ker f \end{aligned}$$

donc f est injectif si et seulement si $\forall x\ y, xy^{-1} \in \ker f \Leftrightarrow x = y$. Ainsi si $\ker f = \{1_G\}$ alors f est injective. Réciproquement si f est injective alors $\forall x\ y, xy^{-1} \in \ker f \Leftrightarrow x = y$ et on peut spécialiser cet énoncé à $y = 1_G$ pour obtenir $\forall x, x \in \ker f \Leftrightarrow x = 1_G$, c'est à dire $\ker f = \{1_G\}$. □

Remarque 3.1.4. On peut bien sûr définir le noyau d'un morphisme de monoïde et on obtient ainsi un sous-monoïde. Mais il s'agit d'une notion peu intéressante car l'analogue du dernier point du lemme ci-dessus n'est pas vrai. Considérons par exemple une fonction f idempotente d'un ensemble X dans lui-même qui n'est pas l'identité, disons la projection de $\mathbb{Z}^2$ sur le premier facteur. Le morphisme de $\mathbb{N}$ dans X^X qui envoie n sur f^n a un « noyau » réduit au neutre 0 mais il n'est pas injectif car tous les entiers strictement positifs sont envoyés sur f.

On note $\mathrm{Aut}(G)$ l'ensemble des isomorphismes d'un groupe G dans lui-même. Il s'agit d'une groupe pour la structure induite par $\mathrm{Aut}(G) \subset \mathfrak{S}(G)$ (on reverra dans un instant la notion de sous-groupe).

Exemple 3.1.5. Soit G un groupe.

- Pour tout g dans G, on définit $c_g : G \to G$ comme $h \mapsto ghg^{-1}$. Il s'agit un automorphisme de G appelé la *conjugaison* par g.
- L'application $g \mapsto c_g$ est un morphisme de G dans $\mathrm{Aut}(G)$.
- Pour tout g dans G, on définit $L_g : G \to G$ comme $h \mapsto gh$. Il s'agit d'une bijection de G appelée *translation à gauche* par g, mais ce n'est pas un morphisme de groupe, sauf si $g = 1$. Par contre $g \mapsto L_g$ est un morphisme injectif de G dans $\mathfrak{S}(G)$.

Définition 3.1.6. *Un* sous-groupe *d'une groupe G est une partie H de G telle que :*

- $1 \in H$
- $\forall x\ y \in G, x \in H$ *et* $y \in H \Rightarrow xy \in H$
- $\forall x \in G, x \in H \Rightarrow x^{-1} \in H$.

Autrement dit H est un sous-monoïde de G qui est stable par inversion. On note alors $H \leqslant G$.

Exemple 3.1.7. Soit G un groupe.

- $\{1\}$ et G sont des sous-groupes de G.
- $\mathrm{Aut}(G)$ est un sous-groupe de $\mathfrak{S}(G)$.

La définition ci-dessus est écrite pour assurer que la structure de groupe sur G se restreint en structure de groupe sur H. Le premier point du lemme suivant donne un critère qui est parfois plus commode à vérifier.

Lemme 3.1.8. *Soit G et G' des groupes, $H \subset G$, $H' \subset G'$ et $f : G \to G'$ un morphisme.*

- $H \leqslant G \Leftrightarrow (H \neq \varnothing$ *et* $\forall x\ y,\ x \in H$ *et* $y \in H \Rightarrow xy^{-1} \in H)$
- *Si* $H \leqslant G$ *alors* $f(H) \leqslant G'$. *En particulier* $\mathrm{im}(f) \leqslant G'$.
- *Si* $H' \leqslant G'$ *alors* $f^{-1}(H') \leqslant G$. *En particulier* $\ker(f) \leqslant G$.
- *L'intersection d'une famille de sous-groupes de* G *est un sous-groupe de* G.

Démonstration. Dans le premier point, l'implication de la gauche vers la droite est claire. Supposons maintenant que H n'est pas vide et que pour tous x et y dans H, xy^{-1} est dans H. La première hypothèse fournit $x_0 \in H$. La deuxième assure alors que $x_0 x_0^{-1} \in H$, c'est à dire $1 \in H$. Soit x dans H. Comme $1 \in H$, on obtient $1x^{-1} \in H$ donc $x^{-1} \in H$. Soit x et y dans H. Comme y^{-1} est aussi dans H, on obtient $x(y^{-1})^{-1} \in H$ donc $xy \in H$.

Pour les autres points, il suffit d'appliquer le lemme 2.0.13 concernant les monoïdes en vérifiant en plus la stabilité par inversion.

Soit H un sous-groupe de G et $f(x) \in f(H)$. Comme $f(x)^{-1} = f(x^{-1})$, $f(x^{-1}) \in f(H)$.

Soit H' un sous-groupe de G' et $x \in f^{-1}(H')$. Comme $f(x^{-1}) = f(x)^{-1}$ et que H' est stable par inversion, $f(x^{-1})$ est dans H'.

Soit $\mathcal{H}$ une famille de sous-groupe de G et K l'intersection des éléments de $\mathcal{H}$. Soit $x \in K$. Pour tout $H \in \mathcal{H}$, x^{-1} est dans H, donc x^{-1} est dans K. □

Le second point du lemme ci-dessus et le dernier point de l'exemple 3.1.5 expliquent en quel sens les groupes de permutations sont les exemples fondamentaux.

Corollaire 3.1.9. *Tout groupe est isomorphe à un sous-groupe d'un groupe de permutations, via les translations à gauche qui envoient* G *injectivement dans* $\mathfrak{S}(G)$.

Définition 3.1.10. *Le sous-groupe engendré par une partie* S *d'un groupe* G *est l'intersection de tous les sous-groupes de* G *contenant* S *:*

$$\langle S \rangle = \bigcap_{H \leqslant G, S \subset H} H.$$

Si $\langle S \rangle = G$, *on dit que* S engendre G, *ou bien que* S *est une* partie génératrice *de* G. *On dit que* G *est* cyclique *s'il est engendré par un singleton.*

Remarque 3.1.11. La terminologie « groupe cyclique » est parfois réservée aux groupes finis. On utilise alors la terminologie « groupe monogène » pour les groupes infinis engendré par un singleton. Par ailleurs la formulation « engendré par un singleton » est un peu pédante, en pratique on dit plutôt « engendré par un élément », même si c'est techniquement moins correct.

Le lemme suivant rassemble les propriétés formelles des sous-groupes engendrés.

Lemme 3.1.12. *Soit* S *une partie d'un groupe* G.

- $\langle S \rangle$ *est un sous-groupe de* G *qui contient* S.
- *Pour tout* $H \leqslant G$, $\langle S \rangle \leqslant H \Leftrightarrow S \subset H$ *(ainsi* $\langle S \rangle$ *est le plus petit sous-groupe de* G *qui contient* S*). Cette propriété universelle caractérise* $\langle S \rangle$.
- *L'application* $\langle \cdot \rangle$ *est croissante :* $\forall S\ S', S \subset S' \Rightarrow \langle S \rangle \subset \langle S' \rangle$.
- *Pour tout morphisme* $f : G \to G'$, $\langle f(S) \rangle = f(\langle S \rangle)$.

Démonstration. C'est exactement la même démonstration que pour le lemme 2.0.16 car cette dernière a été rédigé de façon suffisamment abstraite. □

Comme dans le cas des sous-monoïdes, on a aussi une description explicite.

Lemme 3.1.13. *Soit S une partie d'un groupe G. Les éléments du sous-groupe engendré par S sont tous les produits (finis) d'éléments de S et d'inverses d'éléments de S :*

$$\langle S\rangle = \left\{ s_1^{\sigma_1}\cdots s_k^{\sigma_k} \ ;\ k\in\mathbb{N}, s\colon\{1,\dots,k\}\to S, \sigma\colon\{1,\dots,k\}\to\{-1,1\}\right\}.$$

où le cas $k=0$ correspond au produit vide, c'est à dire au neutre de G.

Démonstration. L'ensemble H du membre de droite contient 1 par définition et il est stable par multiplication et inversion. C'est donc un sous-groupe qui contient S. De plus tout sous-groupe contenant S contient nécessairement H car il doit contenir 1 et S et être stable par produit et inversion. Donc H vérifie la propriété universelle de $\langle S\rangle$ donc $H=\langle S\rangle$ d'après le lemme précédent. □

Définition 3.1.14. *L'ordre d'un groupe est son cardinal. L'ordre d'un élément d'un groupe est l'ordre du sous-groupe qu'il engendre.*

Exemple 3.1.15. Dans $(\mathbb{Z}/4\mathbb{Z},+)$ l'élément 1 est d'ordre 4, ce qui est aussi l'ordre de $\mathbb{Z}/4\mathbb{Z}$, car 1 engendre $\mathbb{Z}/4\mathbb{Z}$. L'ordre de 2 est 2 car le sous-groupe engendré par 2 est $\{0,2\}$.

Exemple 3.1.16. En plus de la notion de sous-groupe, l'autre façon facile de créer des nouveaux groupes à partir d'anciens est de prendre des produits. Si $(G_i)_{i\in I}$ est une famille de groupes alors la multiplication et l'inversion composante par composante définissent une structure de groupe sur $\prod_i G_i$. Le groupe obtenu est appelé le *groupe produit* des G_i. Chaque projection $p_j\colon\prod_i G_i\to G_j$ est un morphisme de groupe par construction. Le produit est caractérisé, à unique isomorphisme près, par la propriété universelle suivante : pour tout groupe H et toute collection de morphismes de groupes $\varphi_i : H\to G_i$, il existe un unique morphisme $\varphi : H\to\prod_i G_i$ tel que $p_j\circ\varphi=\varphi_j$ pour tout j.

3.2 Actions de groupes

Définition 3.2.1. *Une* action *(à gauche) d'un groupe G sur un ensemble X est un morphisme de G dans $\mathfrak{S}(X)$.*

Soit ρ une action de G sur X. Il est parfois commode de penser en terme de la fonction décurryfiée $\bar\rho\colon G\times X\to X$ qui envoie (g,x) sur $\rho(g)(x)$ (par exemple si G et X sont munis d'une structure topologique, on peut parler de continuité de $\bar\rho$ sans avoir à discuter de quelle topologie on munit $\mathfrak{S}(X)$). Mais ce point de vue rend beaucoup moins élégant l'écriture de la définition, il faut demander, pour tous g_1, g_2 et x, $\bar\rho(g_1g_2,x)=\bar\rho(g_1,\bar\rho(g_1,x))$ plutôt que d'écrire simplement $\rho(g_1g_2)=\rho(g_1)\rho(g_2)$.

Lorsque le contexte est clair, on omet souvent le symbole désignant le morphisme et les parenthèses indiquant l'application de fonction, on écrit simplement gx plutôt que $\rho(g)(x)$. La définition prend alors l'allure d'une règle d'associativité : pour tous g_1, g_2 et

x, $(g_1 g_2)x = g_1(g_2 x)$. On écrit aussi $G \curvearrowright X$ pour dire « G agit sur X » ou $\rho : G \curvearrowright X$ pour introduire une action ρ de G sur X.

Soit G un groupe, on note G^{op} le groupe obtenu en munissant l'ensemble G de la loi de composition interne $\star$ définie par $g \star h = hg$ pour tous g et h dans G. On vérifie sans peine qu'il s'agit d'un groupe, avec le même neutre et la même fonction d'inversion. On l'appelle *groupe opposé* de G. Une *action à droite* de G sur un ensemble X est une action à gauche de G^{op} sur X. Soit ρ une telle action, on a alors, pour tous g_1, g_2 et x, $\rho(g_1 g_2)x = \rho(g_2 \star g_1)x = \rho(g_2)\rho(g_1)x$. On voit donc qu'il faut écrire l'élément agissant à droite pour retrouver une règle d'associativité : $x(g_1 g_2) = (xg_1)g_2$. On utilise la notation $X \curvearrowleft G$ pour dire que G agit à droite sur X. On notera que l'inversion est un morphisme de G dans G^{op} et que $(G^{op})^{op} = G$ donc on peut toujours convertir une action à gauche en action à droite et vice-versa en précomposant par l'inversion. Dans toute la suite, quand on écrira « action » sans précision, il s'agira toujours d'une action à gauche. Tous les énoncés généraux sur les actions de groupes sont valables, *mutatis mutandis*, pour les actions à droite, la démonstration consistant à appliquer l'énoncé au groupe opposé.

Exemple 3.2.2.

- Pour tout ensemble X, $\mathfrak{S}(X)$ et ses sous-groupes agissent sur X. Par exemple, si V est un espace vectoriel, $\mathrm{GL}(V)$ agit sur V.
- Tout groupe G agit sur lui-même par conjugaison et par translation à gauche (cf. exemple 3.1.5). Il agit aussi à droite sur lui-même par *translation à droite* : $g \mapsto (R_g : h \mapsto hg)$.
- Le groupe des matrices inversibles de taille n à coefficients dans un corps $\mathbb{K}$ agit sur $\mathbb{K}^n$ par multiplication.
- Le groupe $\mathfrak{S}_n$ des permutations de $\{1, \dots, n\}$ agit à droite sur l'ensemble des matrices $\mathcal{M}_n(\mathbb{K})$ par permutations des lignes : $(A\sigma)_{i,j} = A_{\sigma(i),j}$. On peut aussi permuter les colonnes.
- Pour toute action d'un groupe G sur un ensemble X, G agit sur l'ensemble $\mathcal{P}(X)$ des parties de X : gA est l'image directe de A par l'action de g sur X.
- Pour toute action d'un groupe G sur un ensemble X et pour tout ensemble Y, G agit à droite sur Y^X, l'ensemble des fonctions de X dans Y, par pré-composition : $\forall g\ f, fg = (x \mapsto f(gx))$. Le groupe G agit aussi à gauche sur X^Y, par post-composition. On notera que les deux exemples précédents peuvent être vus comme des cas particuliers de celui-ci.

Définition 3.2.3. *Soit G un groupe agissant sur un ensemble X et x un élément de X.*

- *L'*orbite *de x sous l'action de G est $\mathcal{O}_x = \{gx; g \in G\}$, parfois noté simplement $G \cdot x$ ou Gx (quand le groupe n'est pas clair d'après le contexte) ou $\omega(x)$.*
- *Le* stabilisateur *de x est $G_x := \{g \in G \mid gx = x\}$, parfois noté $\mathrm{Stab}_G(x)$.*
- *Le* fixateur *d'un élément g de G est $\mathrm{Fix}_g := \{x \in X \mid gx = x\}$, parfois noté X^g.*

La diversité des notations dans la définition précédente est un peu désagréable mais elle provient directement de l'ubiquité de ces notions. En écriture manuscrite on évitera d'utiliser simultanément G_x et Gx.

Exemple 3.2.4. Pour l'action standard de $\mathrm{SO}_2(\mathbb{R})$ sur $\mathbb{R}^2$, l'orbite de l'origine est réduite à l'origine tandis que celle de tout autre point est un cercle autour de l'origine. C'est l'origine de la terminologie orbite. Le stabilisateur de l'origine est tout $\mathrm{SO}_2(\mathbb{R})$ tandis que les autres points ont tous un stabilisateur trivial (ie. réduit au neutre de $\mathrm{SO}_2(\mathbb{R})$, la matrice I_2). Tous les éléments de $\mathrm{SO}_2(\mathbb{R})$ ont un fixateur réduit à l'origine, sauf I_2 dont le fixateur est tout $\mathbb{R}^2$.

Soit G un groupe et $H \leqslant G$. Les orbites de l'action par translation à droite de H sur G sont appelées *classes à gauche* suivant H. Les orbites de l'action par translation à gauche de H sur G sont appelées *classes à droite* suivant H.

Pour toute action d'un groupe G sur un ensemble X, le stabilisateur d'une partie A pour l'action induite sur $\mathcal{P}(X)$ est l'ensemble des g tels que $g(A) = A$ (on dit que A est invariant sous l'action de g). Il ne faut pas confondre cette condition avec la condition plus forte demandant que A soit fixé point par point : $\forall a \in A, ga = a$. Attention, on dit parfois que A est stable par g si $g(A) \subset A$ mais il s'agit d'une condition plus faible que l'invariance (par exemple $\mathbb{R}_+$ est stable par la translation $x \mapsto x+1$ mais pas invariant). Le lemme suivant assure que cette subtilité n'apparait pas quand une partie est stable par *tous* les éléments d'un groupe.

Lemme 3.2.5. *Soit $G \curvearrowright X$ une action de groupe. Une partie A de X est stable sous l'action de tous les éléments de G si et seulement si elle est invariante sous l'action de tous les éléments de G.*

Démonstration. Si A est invariante alors elle est stable car $\forall g, gA = A \Rightarrow gA \subset A$. Réciproquement supposons $\forall g, gA \subset A$. Soit g dans G. On a $gA \subset A$ mais aussi $g^{-1}A \subset A$ donc $gg^{-1}A \subset gA$, c'est à dire $A \subset gA$. Ainsi $gA = A$ par double inclusion. □

Lemme 3.2.6. *Le stabilisateur d'un point est un sous-groupe du groupe qui agit.*

Démonstration. Soit x un élément de X. Le stabilisateur G_x contient 1 car 1 agit comme Id_X. Soit g et h dans G_x. On a $hx = x$ donc $x = h^{-1}x$ donc $h^{-1} \in G_x$. De plus $ghx = gx = x$ donc $gh \in G_x$. On remarque que le critère du lemme 3.1.8 n'aide pas ici car il n'y a pas de moyen plus simple de vérifier que G_x est non vide que de vérifier qu'il contient 1 ni de façon de montrer que $gh^{-1} \in G_x$ sans montrer que $h^{-1} \in G_x$. □

Lemme 3.2.7. *Si $y = gx$, les stabilisateurs G_x et G_y sont conjugués par g : $G_y = c_g(G_x)$. En particulier ils ont même cardinal.*

Démonstration. Soit g' dans G. On a

$$\begin{aligned} g' \in G_y &\Leftrightarrow g'y = y \\ &\Leftrightarrow g'gx = gx \\ &\Leftrightarrow g^{-1}g'gx = x \\ &\Leftrightarrow c_{g^{-1}}(g') \in G_x \\ &\Leftrightarrow g' \in c_{g^{-1}}^{-1}(G_x) \\ &\Leftrightarrow g' \in c_g(G_x) \end{aligned}$$

□

Définition 3.2.8. *Soit $\rho : G \to \mathfrak{S}(X)$ une action de groupe. On dit que cette action est :*

- libre *si* $\forall g \neq 1, \forall x, gx \neq x$*, autrement dit,* $\forall x, G_x = \{1\}$ *ou encore* $\forall g \neq 1, \mathrm{Fix}_g = \varnothing$ *;*
- fidèle *si* ρ *est injective, autrement dit,* $\bigcap_x G_x = \{1\}$ *ou encore* $\forall g \neq 1, \mathrm{Fix}_g \neq X$ *;*
- transitive *s'il existe* x_0 *tel que* $\mathcal{O}_{x_0} = X$*. De façon équivalente, l'action est transitive si* $\forall x, \mathcal{O}_x = X$*.*

Exemple 3.2.9. Toute action libre sur un ensemble non vide est fidèle.

L'action d'un groupe sur lui-même par translation à gauche (ou à droite) est libre et transitive.

L'action de $\mathbb{Z}$ sur $\mathbb{Z} \times \mathbb{Z}$ par translation de la première composante est libre mais pas transitive.

L'action standard de $\mathrm{SO}_2(\mathbb{R})$ sur $\mathbb{R}^2$ est fidèle mais pas libre ni transitive.

L'action de $\mathbb{Z}^2$ sur $\mathbb{Z}$ définie par $\forall (k,l)\ m, (k,l)m = k + m$ est transitive mais pas fidèle (et donc a fortiori pas libre).

L'observation suivante est très facile mais fondamentale.

Proposition 3.2.10. *Soit* G *un groupe agissant sur un ensemble* X*. La relation sur* X *définie par* $x \sim x'$ *s'il existe* g *tel que* $gx = x'$ *est une relation d'équivalence. On note* $G\backslash X$ *le quotient (ou* X/G *dans le cas d'une action à droite).*

Démonstration. La relation est réflexive car, pour tout x, $1x = x$. Elle est symétrique car, pour tous x et y tels que $x \sim y$, on obtient g tel que $gx = y$ et on alors $y = g^{-1}x$. Elle transitive car pour tous x, y et z tels que $x \sim y$ et $y \sim z$, on obtient g et g' tels que $x = gy$ et $y = g'z$ et on a alors $x = gg'z$. □

On notera qu'il est possible de définir la notion d'action pour des structures algébriques plus faibles que les groupes comme les monoïdes, mais la proposition ci-dessus utilise toute la structure de groupe. Par exemple le corollaire suivant serait faux pour les actions de monoïdes (avec le même contre-exemple que dans la remarque 3.1.4, vu comme action de $\mathbb{N}$ sur $\mathbb{Z}$).

Corollaire 3.2.11. *Les orbites d'une action de groupe forment une partition. De plus, pour tous* x *et* y*,* $\mathcal{O}_x = \mathcal{O}_y \Leftrightarrow \exists g, y = gx$*.*

Démonstration. Les orbites sont les classes d'équivalence de la relation associée à l'action. □

Proposition 3.2.12. *Soit* H *un sous-groupe d'un groupe* G*. Pour tout section* $\sigma : G/H \to G$*, c'est à dire tout fonction telle que* $\pi \circ \sigma = \mathrm{Id}$*, l'application de* $G/H \times H$ *dans* G *qui envoie* (y, h) *sur* $\sigma(y)h$ *est une bijection.*

Démonstration. Montrons l'injectivité. Soit (y, h) et (z, k) tels que $\sigma(y)h = \sigma(z)k$. On a $\sigma(y) = \sigma(z)kh^{-1}$ et kh^{-1} est dans H donc $\sigma(y) \sim \sigma(z)$ donc $\pi(\sigma(y)) = \pi(\sigma(z))$, c'est à dire $y = z$ puisque $\pi \circ \sigma = \mathrm{Id}$. L'égalité $\sigma(y)h = \sigma(z)k$ se réécrit donc $\sigma(y)h = \sigma(y)k$ et on en déduit $h = k$. Ainsi $(y, h) = (z, k)$.

Montrons maintenant la surjectivité. Soit g dans G. On pose $y = \pi(g)$. Comme $\pi(g) = \pi(\sigma(y))$, on a $g \sim \sigma(y)$ donc on obtient $h \in H$ tel que $g = \sigma(y)h$.

Notons qu'on peut aussi expliciter l'inverse de cette application, il s'agit de $g \mapsto (\pi(g), \sigma(\pi(g))^{-1}g)$, mais cela ne rend pas la démonstration plus claire. □

Corollaire 3.2.13 (Théorème de Lagrange)**.** *Pour tout sous-groupe H d'un groupe G, $\sharp G = \sharp(G/H)\sharp H$. En particulier le cardinal de H divise celui de G. En particulier, pour tout $x \in G$, $o(x) \mid \sharp G$ et donc $x^{\sharp G} = 1$.*

Démonstration. Il suffit d'appliquer la proposition précédente à une section de $G \to G/H$ fournie par l'axiome du choix (bien sûr dans le cas où G est fini cet axiome n'est pas nécessaire ici). □

Définition 3.2.14. L'indice *d'un sous-groupe $H \leqslant G$ est le cardinal de G/H, noté $(G : H)$. On a donc $\sharp G = (G : H)\sharp H$ d'après le théorème de Lagrange.*

Lemme 3.2.15. *Soit G un groupe agissant sur un ensemble X. Pour tout $A \subset X$, on a*

$$\pi^{-1}(\pi(A)) = \bigcup_{g \in G} gA.$$

Démonstration. Il suffit de dérouler les définitions. Les trois premières lignes du calcul suivant ne sont pas spécifiques au cas des actions de groupe.

$$\begin{aligned}
\pi^{-1}(\pi(A)) &= \{x | \pi(x) \in \pi(A)\} \\
&= \{x \mid \exists a \in A, \pi(a) = \pi(x)\} \\
&= \{x \mid \exists a \in A, a \sim x\} \\
&= \{x \mid \exists a \in A, \exists g \in G, x = ga\} \\
&= \{x \mid \exists g \in G, x \in gA\} \\
&= \bigcup_{g \in G} gA
\end{aligned}$$

□

Définition 3.2.16. *Soit G un groupe, X et Y des ensembles. Étant donnée une action de G sur X, on dit qu'une fonction $f : X \to Y$ est* G-invariante *si, pour tous g et x, $f(gx) = f(x)$. Dans ce cas f descend en fonction de $G\backslash X$ dans Y.*

Si Y aussi est muni d'une action de G, on dit que f est G-équivariante *si, pour tous g et x, $f(gx) = gf(x)$. Dans ce cas f descend en application de $G\backslash X$ dans $G\backslash Y$.*

Remarque 3.2.17. On notera qu'une fonction G-invariante n'est rien d'autre qu'une fonction qui est G-équivariante pour l'action triviale de G sur Y (celle qui envoie tous les éléments de G sur Id_Y).

Exemple 3.2.18. Soit $\mathbb{K}$ un corps et E un $\mathbb{K}$-espace vectoriel. On note $\mathbb{P}(E)$ l'espace projectif sur E, c'est à dire l'ensemble des droites vectorielles de E. On considère l'action du groupe multiplicatif $\mathbb{K}^\times = \mathbb{K}\backslash\{0\}$ sur $E\backslash\{0\}$ par multiplication scalaire. L'application $f : E \setminus \{0\} \to \mathbb{P}(E)$ qui envoie v sur $\mathbb{K}v$ est $\mathbb{K}^\times$-invariante et induit une bijection de $\mathbb{K}^\times\backslash(E \setminus \{0\})$ sur $\mathbb{P}(E)$.

Toute application linéaire inversible $\varphi \in \mathrm{GL}(E)$ se restreint en permutation $\mathbb{K}^\times$-équivariante de $E \setminus \{0\}$ qui induit une permutation au quotient. Vu dans l'identification précédente, il s'agit de la permutation de $\mathbb{P}(E)$ qui envoie toute droite sur son image par φ.

Le théorème fondamental suivant relie l'action de G sur lui-même par translation à droite et son action sur X.

Théorème 3.2.19. *Soit G un groupe agissant sur un ensemble X. Pour tout x dans X, on considère l'action de G_x sur G par translation à droite. L'application de G dans X qui envoie g sur gx descend au quotient en bijection $G/G_x \to \mathcal{O}_x$. En particulier $\sharp(G/G_x) = \sharp\mathcal{O}_x$.*

Démonstration. On a vu dans le théorème 1.0.5 qu'il suffit de vérifier que, $\forall g\ g', g \sim g' \Leftrightarrow gx = gx'$ et que $g \mapsto gx$ est surjective de G dans $\mathcal{O}_x$. La surjectivité provient directement de la définition de $\mathcal{O}_x$. Montrons la première condition. Soit g et g' dans G. On a :

$$\begin{aligned} g \sim g' &\Leftrightarrow \exists h \in G_x, gh = g' \\ &\Leftrightarrow g^{-1}g' \in G_x \\ &\Leftrightarrow g^{-1}g'x = x \\ &\Leftrightarrow g'x = gx \end{aligned}$$

□

On notera que la conséquence en termes de cardinaux est un énoncé bien plus faible, particulièrement lorsque ces cardinaux sont infinis.

Corollaire 3.2.20 (Équation aux classes)**.** *Soit G un groupe fini agissant sur un ensemble fini X. Soit $\sigma : G\backslash X \to X$ une section de la projection $\pi : X \to G\backslash X$ (c-à-d une application telle que $\pi \circ \sigma = \mathrm{Id}_{G\backslash X}$). On a l'*équation aux classes *:*

$$\sharp X = \sum_{y \in G\backslash X} \frac{\sharp G}{\sharp G_{\sigma(y)}}$$

et chaque terme de la somme est un entier. En particulier, si G agit librement sur X alors $\sharp X = \sharp(G\backslash X)\sharp G$.

Démonstration. Le lemme 1.0.17, qui lie quotients et partitions, et la description de la relation d'équivalence associée à une action assure que les $\mathcal{O}_{\sigma(y)}$ forment une partition de X donc $\sharp X = \sum_y \sharp\mathcal{O}_{\sigma(y)}$ puis le théorème 3.2.19 donne la formule annoncée. □

Remarque 3.2.21. Lorsque G est fini et qu'on considère l'action (libre) d'un sous-groupe par translation, on retrouve le théorème de Lagrange.

Le théorème suivant est une réciproque partielle du théorème de Lagrange. Il en existe de nombreuses démonstrations. Nous utiliserons l'équation aux classes et le lemme arithmétique suivant.

Lemme 3.2.22. *Soit p un nombre premier et m un entier naturel. Si m n'est pas divisible par p alors le coefficient binomial $\binom{p^k m}{p^k}$ ne l'est pas non plus.*

Démonstration. On commence par donner une démonstration sophistiquée puis on expliquera une démonstration artisanale. Pour tout P et Q dans $\mathbb{Z}/p\mathbb{Z}[X]$ on a $(P+Q)^p = P^p + Q^p$. Cette observation classique dans le cas de $\mathbb{Z}/p\mathbb{Z}$ sera généralisée pour inclure notamment le cas de $\mathbb{Z}/p\mathbb{Z}[X]$ dans le lemme 7.4.8. On peut donc calculer $(1+X)^{p^k m} = (1+X^{p^k})^m$ donc le coefficient de X^{p^k} dans ce polynôme est m. Par ailleurs la formule du binôme de Newton assure que ce coefficient est aussi $\binom{p^k m}{p^k}$. Ainsi on a $\binom{p^k m}{p^k} = m$

dans $\mathbb{Z}/p\mathbb{Z}$. Par hypothèse l'image de m dans $\mathbb{Z}/p\mathbb{Z}$ n'est pas nulle donc celle de $\binom{p^k m}{p^k}$ non plus.

Donnons maintenant une démonstration élémentaire qui reste dans $\mathbb{Z}$. Pour tous entiers naturels n et q avec $q \leqslant n$, on a

$$\binom{n}{q} = \frac{\Pi_{j=1}^{q}(n-q+j)}{\Pi_{j=1}^{q} j}.$$

Pour $n = p^k m$ et $q = p^k$, on obtient

$$\left(\prod_{j=1}^{p^k} j\right)\binom{p^k m}{p^k} = \prod_{j=1}^{p^k}(p^k m - p^k + j)$$

Pour tout $j \leqslant p^k$, on écrit $j = p^{v_j} m_j$ où $p \nmid m_j$ et $v_j < k$ sauf pour $j = p^k$ où $v_j = k$ et $m_j = 1$. Dans tous les cas $v_j \leqslant k$. On a alors $p^k m - p^k + j = p^{v_j}(p^{k-v_j}(m-1) + m_j)$. On pose $n_j = p^{k-v_j}(m-1) + m_j$ et on observe que p ne divise pas n_j. En effet, si $j < p^k$ on a $v_j < k$ donc $p | p^{k-v_j}(m-1)$ et p ne divise pas m_j tandis que dans le cas $j = p^k$, $n_j = m$ et c'est notre hypothèse sur m.

Ainsi la grande équation ci-dessus devient

$$\left(\prod_{j=1}^{p^k} m_j\right)\binom{p^k m}{p^k} = \prod_{j=1}^{p^k} n_j$$

où p ne divise aucun des n_j donc, comme p est premier, p ne divise pas non plus le coefficient binomial. □

Théorème 3.2.23 (Premier théorème de Sylow)**.** *Soit G un groupe fini et p un nombre premier. On écrit $\sharp G = p^k m$ avec $p \nmid m$. Il existe un sous-groupe $H \leqslant G$ de cardinal p^k. On dit que H est un p-Sylow de G.*

Démonstration. On note X l'ensemble des parties de G de cardinal p^k. Cet ensemble est de cardinal $\binom{p^k m}{p^k}$. On considère l'action de G sur $\mathcal{P}(G)$ induite par l'action par translation à gauche de G sur lui-même, comme dans l'exemple 3.2.2. Comme les bijections préservent le cardinal, cette action se restreint en action de G sur X. Soit $\sigma : G\backslash X \to X$ une section du quotient. L'équation aux classes montre que $\sharp X$ est la somme des entiers $\sharp G/\sharp G_{\sigma(y)}$ pour y parcourant le quotient. D'après le lemme 3.2.22, le cardinal de X n'est pas divisible par p. Donc, pour au moins un y, $\sharp G/\sharp G_{\sigma(y)}$ n'est pas divisible par p. Posons $A = \sigma(y)$ pour un tel y et $d = \sharp G/\sharp G_{\sigma(y)}$. On a donc $p^k m = \sharp G = d\sharp G_A$ et $p \nmid d$. Comme p est premier, on en déduit $p^k | \sharp G_A$. Or G_A agit sur A et, comme l'action de G sur X est libre, l'action de G_A sur A l'est aussi. Donc $\sharp G_A | \sharp A$ par le cas particulier de l'équation aux classes. Or $\sharp A = p^k$ par définition de X. Donc $\sharp G_A | p^k$ et finalement $\sharp G_A = p^k$ par antisymétrie de la relation de divisibilité sur $\mathbb{N}$. Ainsi G_A est un p-Sylow de G. On notera le coup de théâtre : le p-Sylow n'est pas trouvé comme élément de X mais comme stabilisateur d'un tel élément. □

Théorème 3.2.24 (Formule de Burnside)**.** *Soit G un groupe fini agissant sur un ensemble fini X. On a la* formule de Burnside *:*

$$\sharp G\backslash X = \frac{1}{\sharp G}\sum_{g \in G} \sharp \operatorname{Fix}_g .$$

Démonstration. On procède par double décompte de l'ensemble $A = \{(g, x) \in G \times X \mid gx = x\}$. En comptant par « tranches verticales », c'est à dire à g fixé, on obtient

$$\sharp A = \sum_{g \in G} \sharp\{x \mid gx = x\} = \sum_{g \in G} \sharp \operatorname{Fix}_g .$$

Pour le décompte par tranches horizontales, on va en plus regrouper les tranches par orbites sous l'action de G. On calcule en utilisant le théorème 3.2.19 :

$$\begin{aligned}
\sharp A &= \sum_{x \in X} \sharp\{g \mid gx = x\} \\
&= \sum_{x \in X} \sharp G_x \\
&= \sum_{y \in G\backslash X} \sum_{x \in \pi^{-1}(y)} \sharp G_x \\
&= \sum_{y \in G\backslash X} \sum_{x \in \pi^{-1}(y)} \frac{\sharp G}{\sharp \pi^{-1}(y)} && \text{par le théorème 3.2.19} \\
&= \sum_{y \in G\backslash X} \sharp \pi^{-1}(y) \frac{\sharp G}{\sharp \pi^{-1}(y)} \\
&= \sharp (G\backslash X) \sharp G
\end{aligned}$$

et on conclut en comparant les deux décomptes. □

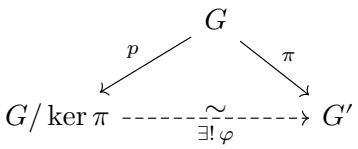
La formule de Burnside a de nombreuses applications amusantes en combinatoire, il y aura des exemples en TD.

3.3 Quotients de groupes et groupes quotients

Définition 3.3.1. *Un* quotient *d'un groupe G est un groupe G' muni d'un morphisme π surjectif de G dans G'.*

Exemple 3.3.2. Le groupe $\mathbb{U}$ des nombres complexes de module 1 est un quotient du groupe des nombres réels, via $t \mapsto \exp it$.

Lemme 3.3.3. *Soit $\pi : G \to G'$ un quotient de groupe. Soit $p : G \to G/\ker\pi$ le quotient associé à l'action par translation à droite du sous-groupe $\ker\pi$ sur G. Il existe une unique bijection $\varphi : G/\ker\pi \to G'$ telle que $\pi = \varphi \circ p$.*



Démonstration. D'après le corollaire 1.0.8, il suffit de montrer que p et π induisent la même relation d'équivalence sur G. Comme π est un morphisme, la relation induite par π peut se réécrire en $x \sim x'$ si $x^{-1}x' \in \ker\pi$ (c'est le même calcul que dans la démonstration du critère d'injectivité du lemme 3.1.3). On peut encore récrire cela comme $\exists g \in \ker\pi, x' = xg$, ce qui est bien la relation associée à l'action par translation à droite par $\ker\pi$, et donc à p. □

Dans le lemme précédent, il n'y a pas de structure de groupe sur $G/\ker\pi$ mais on peut transporter la structure de G' via φ, en définissant le produit comme $(x,y) \mapsto \varphi^{-1}(\varphi(x)\varphi(y))$, le neutre comme $\varphi^{-1}(1)$ et l'inversion comme $x \mapsto \varphi^{-1}(\varphi(x)^{-1})$. On obtient ainsi une structure de groupe telle que p est un morphisme. Il est donc naturel de se demander si, pour tout sous-groupe $H \leqslant G$, il existe une structure de groupe sur G/H qui fasse de $p : G \to G/H$ un morphisme de groupe. Il s'agit d'une idée trop optimiste en général, il faut imposer une condition sur H.

Définition 3.3.4. *On dit qu'un sous-groupe H d'un groupe G est* distingué *s'il est stable sous l'action par conjugaison de tous les éléments de G : $\forall g \in G, c_g(H) \subset H$. On note alors $H \lhd G$.*

Vu le lemme 3.2.5, $H \lhd G$ si et seulement si il est invariant sous l'action par conjugaison de tous les éléments de G : $\forall g \in G, c_g(H) = H$. Autrement dit, $H \lhd G$ si et seulement si $\forall g \in G, gH = Hg$ (pour les actions par translation à gauche et à droite de G sur $\mathcal{P}(G)$).

Exemple 3.3.5. Pour tout groupe G, $\{1\} \lhd G$ et $G \lhd G$. Si G est abélien alors, pour tout sous-groupe $H \leqslant G$, $H \lhd G$.

Théorème 3.3.6. *Soit G un groupe et H un sous-groupe de G. Il existe une structure de groupe sur G/H qui fasse de la projection un morphisme de groupes si et seulement si $H \lhd G$. Cette structure est alors unique et H est le noyau de la projection.*

Démonstration. Le lemme 2.0.9 assure que G/H possède une structure de monoïde qui fasse de $\pi : G \to G/H$ un morphisme de monoïdes si et seulement si la loi de composition interne sur G est compatible avec la relation d'équivalence produit sur G, et que cette structure de monoïde est alors unique. On obtient donc le critère :

$$\forall x_1\ x_2\ y_1\ y_2,\ x_1^{-1}x_2 \in H \text{ et } y_1^{-1}y_2 \in H \Rightarrow (x_1y_1)^{-1}(x_2y_2) \in H \qquad (\star)$$

Supposons la condition ($\star$). Soit g dans G et h dans H. On spécialise ($\star$) à $x_1 = 1$, $x_2 = h$, $y_1 = y_2 = g^{-1}$ et on obtient $1^{-1}h \in H$ et $gg^{-1} \in H \Rightarrow (1g^{-1})^{-1}(hg^{-1}) \in H$. Or on a supposé $h \in H$ et H est un sous-groupe donc contient 1. On obtient donc $ghg^{-1} \in H$. Ainsi H est distingué.

Réciproquement supposons $H \lhd G$ et montrons la condition ($\star$). Soit x_1, x_2, y_1 et y_2 tels que $x_1^{-1}x_2 \in H$ et $y_1^{-1}y_2 \in H$. On a :

$$(x_1y_1)^{-1}(x_2y_2) = y_1^{-1}x_1^{-1}x_2y_2 = \underbrace{y_1^{-1}\underbrace{(x_1^{-1}x_2)}_{\in H}y_1}_{\in c_{y_1^{-1}}(H)=H}\underbrace{(y_1^{-1}y_2)}_{\in H} \in H$$

Donc la loi de composition descend au quotient.

Il reste à voir qu'on obtient une structure de groupe et pas seulement une structure de monoïde. Comme π est un morphisme de monoïde, elle envoie les inversibles de G sur des inversibles de G/H d'après le lemme 2.0.7. Or tous les éléments de G sont inversibles et π est surjective donc tous les éléments de G/H sont inversibles. Remarquons qu'il est instructif de vérifier directement que la fonction d'inversion sur G descend au quotient. □

Lemme 3.3.7. *Soit G et G' deux groupes, et $\varphi : G \to G'$ un morphisme.*

- *Pour tout $H' \lhd G'$, $\varphi^{-1}(H') \lhd G$. En particulier $\ker\varphi \lhd G$.*
- *Si φ est surjectif et $H \lhd G$ alors $\varphi(H) \lhd G'$.*

Démonstration. La clef est que, pour tout $g \in G$, $\varphi \circ c_g = c_{\varphi(g)} \circ \varphi$. Soit $H' \lhd G'$. Le lemme 3.1.8 assure que $\varphi^{-1}(H')$ est un sous-groupe de G. Soit $g \in G$. On a

$$\begin{aligned} \varphi(c_g(\varphi^{-1}(H'))) &= c_{\varphi(g)}(\varphi(\varphi^{-1}(H'))) && \text{par la formule clef} \\ &\subset c_{\varphi(g)}(H') \\ &\subset H' && \text{car } H' \lhd G' \end{aligned}$$

donc $c_g(\varphi^{-1}(H')) \subset \varphi^{-1}(H')$.

Supposons maintenant que φ est surjective et $H \lhd G$. Le lemme 3.1.8 assure que $\varphi(H)$ est un sous-groupe de G'. Soit g' dans G'. Par surjectivité de φ, on obtient $g \in G$ tel que $\varphi(g) = g'$. On calcule

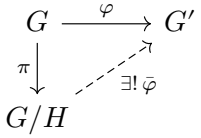
$$c_{g'}(\varphi(H)) = c_{\varphi(g)}(\varphi(H)) = \varphi(c_g(H)) = \varphi(H)$$

en utilisant la formule clef et l'hypothèse $H \lhd G$. □

Corollaire 3.3.8. *Une partie H d'un groupe G est un sous-groupe distingué si et seulement si il existe un groupe G' et un morphisme $\varphi : G \to G'$ tel que $H = \ker\varphi$.*

Démonstration. D'après le théorème 3.3.6, si $H \lhd G$ alors $G' = G/H$ convient, avec comme morphisme la projection canonique. La réciproque est donnée par la première partie du lemme 3.3.7. □

Théorème 3.3.9 (Propriété universelle des groupes quotients)**.** *Soit H un sous-groupe distingué d'un groupe G. Pour tout morphisme $\varphi : G \to G'$ tel que $H \subset \ker\varphi$, il existe un unique morphisme $\bar{\varphi} : G/H \to G'$ tel que $\varphi = \bar{\varphi} \circ \pi$.*



On a alors

- $\ker\bar{\varphi} = \pi(\ker\varphi)$,
- $\bar{\varphi}$ *injectif* $\Leftrightarrow H = \ker\varphi$
- $\bar{\varphi}$ *surjectif* $\Leftrightarrow \varphi$ *surjectif* .

Réciproquement, si $\bar{\varphi}$ existe alors $H \subset \ker\varphi$.

Démonstration. Soit $\varphi : G \to G'$ tel que $H \subset \ker\varphi$. D'après la propriété universelle des quotients d'ensembles (théorème 1.0.5), l'unicité est acquise et pour l'existence il suffit de vérifier que, pour tout x et y dans G, $x \sim y \Rightarrow \varphi(x) = \varphi(y)$. Soit x et y tels que $x \sim y$. On a $x^{-1}y \in H$ et $H \subset \ker\varphi$ on a $x^{-1}y \in \ker\varphi$ donc $\varphi(x) = \varphi(y)$. On obtient

ainsi l'existence de $\bar{\varphi}$. Le lemme 2.0.9 assure qu'il s'agit d'un morphisme. On étudie maintenant le noyau de $\bar{\varphi}$. On calcule :

$$\begin{aligned}\pi^{-1}\ker\bar{\varphi} &= \pi^{-1}\bar{\varphi}^{-1}(\{1\})\\ &= (\bar{\varphi}\circ\pi)^{-1}(\{1\})\\ &= \varphi^{-1}(\{1\})\\ &= \ker\varphi,\end{aligned}$$

et on obtient la conclusion annoncée en appliquant π et le lemme 1.0.18 qui assure que $\pi\circ\pi^{-1} = \mathrm{Id}_{\mathcal{P}(G/H)}$.

En particulier $\bar{\varphi}$ est injectif si et seulement si $\pi(\ker\varphi) = \{1\}$. Or on sait déjà que $\pi^{-1}(\{1\}) = H \subset \ker\varphi$ donc $\{1\} \subset \pi(\ker\varphi)$ (en utilisant encore le lemme 1.0.18). Ainsi

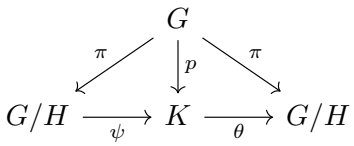
$$\begin{aligned}\bar{\varphi}\text{ injective} &\Leftrightarrow \pi(\ker\varphi)\subset\{1\}\\ &\Leftrightarrow \ker\varphi\subset\pi^{-1}(\{1\})\\ &\Leftrightarrow \ker\varphi\subset H\\ &\Leftrightarrow \ker\varphi = H\end{aligned}$$

où la dernière équivalence provient encore de l'hypothèse $H \subset \ker\varphi$.

La condition de surjectivité provient directement du théorème 1.0.5, il n'y a rien de spécifique à la théorie des groupes ici.

Le fait que la condition annoncée est nécessaire à l'existence de $\bar{\varphi}$ provient aussi de ce théorème et du premier paragraphe de cette démonstration. □

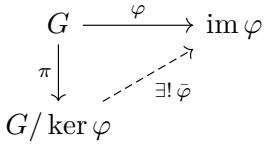
Remarque 3.3.10. La propriété universelle de G/H caractérise G/H à unique isomorphisme près, au sens suivant. Supposons que K soit un groupe muni d'un morphisme $p\colon G\to K$ tel que $H\subset\ker p$ et, pour tout morphisme $\varphi\colon G\to G'$ tel que $H\subset\ker\varphi$, il existe un unique $\bar{\varphi}$ tel que $\varphi = \bar{\varphi}\circ p$. Alors il existe un unique isomorphisme $\psi\colon G/H\to K$ tel que $\psi\circ\pi = p$. En effet, la propriété universelle de G/H appliquée au morphisme p fournit un morphisme $\psi\colon G/H\to K$ tel que $\psi\circ\pi = p$. Ensuite la propriété universelle de K appliquée au morphisme π fournit $\theta\colon K\to G/H$ tel que $\theta\circ p = \pi$. On a donc le diagramme :



où les deux petits triangles commutent donc le grand aussi. L'unicité dans la propriété universelle de G/H appliquée au morphisme π assure que $\theta\circ\psi = \mathrm{Id}$. On montre de même que $\psi\circ\theta = \mathrm{Id}$ et donc ψ est bien un isomorphisme. Cette remarque explique en quel sens les détails de la construction de G/H ne sont pas très important, la propriété universelle contient l'essentiel. On note que cette démonstration est exactement calquée sur celle du corollaire 1.0.8.

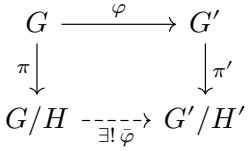
Le corollaire suivant est immédiat et extrêmement utile. On l'appelle souvent le premier théorème d'isomorphisme quand il est associé à deux autres énoncés moins cruciaux qui seront vus en TD.

Corollaire 3.3.11 (Premier théorème d'isomorphisme)**.** *Tout morphisme de groupe $\varphi : G \to G'$ induit un unique isomorphisme $\bar{\varphi} : G/\ker\varphi \to \operatorname{im}\varphi$.*



□

Corollaire 3.3.12. *Soit H et H' des sous-groupes distingués dans des groupes G et G'. Soit $\varphi : G \to G'$ un morphisme. Si $\varphi(H) \subset H'$ alors φ descend en un unique morphisme de G/H dans G'/H'.*



On a alors

- $\ker\bar{\varphi} = \pi(\varphi^{-1}(H'))$,
- $\bar{\varphi}$ *injectif* $\Leftrightarrow \varphi^{-1}(H') \subset H \Leftrightarrow \varphi^{-1}(H') = H$
- $\bar{\varphi}$ *surjectif* $\Leftrightarrow \forall x' \in G', \exists x \in G, \exists h' \in H', x' = \varphi(x)h'$.
  En particulier φ surjectif $\Rightarrow \bar{\varphi}$ surjectif.

Démonstration. On applique le théorème 3.3.9 à $\pi' \circ \varphi$, en notant que $\ker(\pi' \circ \varphi) = \varphi^{-1}(\ker\pi') = \varphi^{-1}(H')$ donc la condition à vérifier est $H \subset \varphi^{-1}(H')$, ce qui équivaut à $\varphi(H) \subset H'$.

Supposons cette condition vérifiée, de sorte qu'on obtient un unique $\bar{\varphi}$. Le théorème assure que $\ker\bar{\varphi} = \pi(\ker(\pi' \circ \varphi))$, c'est à dire $\pi(\varphi^{-1}(H'))$ par le calcul ci-dessus, et il assure que $\bar{\varphi}$ est injectif si et seulement si $\varphi^{-1}(H') = H$. Puisqu'on a déjà supposé $H \subset \varphi^{-1}(H')$, cette dernière condition est équivalente à $\varphi^{-1}(H') \subset H$.

Le théorème assure aussi la première équivalence ci-dessous :

$$\begin{aligned}
\bar{\varphi} \text{ surjective} &\Leftrightarrow \pi' \circ \varphi \text{ surjective} \\
&\Leftrightarrow \forall z \in G'/H', \exists g \in G, \pi'(\varphi(g)) = z \\
&\Leftrightarrow \forall g' \in G', \exists g \in G, \pi'(\varphi(g)) = \pi'(g') \\
&\Leftrightarrow \forall g' \in G', \exists g \in G, \varphi(g) \sim g' \\
&\Leftrightarrow \forall g' \in G', \exists g \in G, \exists h' \in H', g' = \varphi(g)h
\end{aligned}$$

□

Théorème 3.3.13 (Théorème de correspondance des sous-groupes)**.** *Soit G un groupe et $N \lhd G$. On note π la projection de G sur G/N. L'ensemble des sous-groupes de G qui contiennent N est en bijection croissante avec l'ensemble des sous-groupes de G/N par l'application $U \mapsto \pi(U)$, de réciproque $H \mapsto \pi^{-1}(H)$. De plus $U \lhd G \Leftrightarrow \pi(U) \lhd G/N$.*

Remarque 3.3.14. Dans le contexte du théorème précédent, soit U un sous-groupe de G. Si U contient N alors N peut aussi être vu comme sous-groupe de U, et il est clair que N

est aussi distingué dans U. Ainsi on peut former le groupe quotient U/N et $\pi|_U$ induit un isomorphisme de U/N sur $\pi(U)$ d'après le corollaire 3.3.11. En fait la construction des quotients comme ensembles de classes d'équivalence assure une égalité d'ensembles $\pi(U) = U/N$. Bien que cela brise la barrière d'abstraction des groupes quotients, il est donc souvent commode d'écrire la correspondance du théorème précédent sous la forme $U \mapsto U/N$.

Démonstration. Le lemme 3.1.8 assure que, pour tout $U \leqslant G$, $\pi(U) \leqslant G/N$ donc l'application est bien définie dans le sens direct. Soit $H \leqslant G/N$. La préimage $\pi^{-1}(H)$ est une sous-groupe de G d'après le lemme 3.1.8. De plus $\{1\} \subset H$ donc $\pi^{-1}(\{1\}) \subset \pi^{-1}(H)$, c'est à dire $N \subset \pi^{-1}(H)$ et l'application réciproque est bien définie aussi.

Il reste à calculer leurs composées. Pour tout $H \leqslant G/N$, on a $\pi(\pi^{-1}(H)) = H$ d'après le lemme 1.0.18.

Soit $U \leqslant G$ tel que $N \subset U$. On calcule

$$\begin{aligned} \pi^{-1}(\pi(U)) &= \bigcup_{n \in N} Un \qquad \text{d'après le lemme 3.2.15} \\ &= U \qquad \text{car } N \subset U \text{ donc } \forall n, Un = U. \end{aligned}$$

Montrons que U est distingué si et seulement si $\pi(U)$ l'est. Si U est distingué, le lemme 3.3.7 montre que $\pi(U)$ est distingué car π est surjective.

Réciproquement supposons que $\pi(U)$ est distingué. En plus du résultat $\forall U \leqslant G, N \subset U \Rightarrow \pi^{-1}(\pi(U)) = U$ établi ci-dessus, le premier point clef est que, pour tout g dans G, $c_{\pi(g)} \circ \pi = \pi \circ c_g$. Le second est que, comme N est distingué, pour tout g dans G, $c_g(U)$ contient N. Soit g dans G. On calcule :

$$\begin{aligned} c_g(U) &= \pi^{-1}(\pi(c_g(U))) \qquad \text{car } N \subset c_g(U) \\ &= \pi^{-1}(c_{\pi(g)}(\pi(U))) \\ &= \pi^{-1}(\pi(U)) \qquad \text{car } \pi(U) \lhd G/N \\ &= U \end{aligned}$$

□

Lemme 3.3.15. *Une intersection de sous-groupes distingués d'un groupe G est un sous-groupe distingué.*

Démonstration. Le lemme 3.1.8 assure déjà qu'une telle intersection est un sous-groupe. La stabilité par conjugaison est claire car l'image directe d'une intersection par une bijection est l'intersection des images directes. □

Définition 3.3.16. *Le sous-groupe distingué engendré par une partie S d'un groupe G est l'intersection de tous les sous-groupes distingués contenant S. C'est un sous-groupe distingué d'après le lemme précédent, on le note $\langle\langle S \rangle\rangle$.*

Lemme 3.3.17. *Soit S une partie d'une groupe G.*

- *$\langle\langle S \rangle\rangle$ est un sous-groupe distingué de G qui contient S.*
- *$\langle S \rangle \subset \langle\langle S \rangle\rangle$.*
- *Pour tout $H \lhd G$, $\langle\langle S \rangle\rangle \leqslant H \Leftrightarrow S \subset H$ (ainsi $\langle\langle S \rangle\rangle$ est le plus petit sous-groupe distingué de G qui contient S). Cette propriété universelle caractérise $\langle\langle S \rangle\rangle$.*

- *L'application* $\langle\!\langle \cdot \rangle\!\rangle$ *est croissante :* $\forall S\ S', S \subset S' \Rightarrow \langle\!\langle S \rangle\!\rangle \subset \langle\!\langle S' \rangle\!\rangle$.
- *Pour tout morphisme* $f : G \to G'$ *surjectif,* $\langle\!\langle f(S) \rangle\!\rangle = f(\langle\!\langle S \rangle\!\rangle)$.
- $\langle\!\langle S \rangle\!\rangle$ *est l'ensemble des produits (finis) de conjugués des éléments de* S *et de leurs inverses.*

Démonstration. C'est exactement la même démonstration que pour les lemmes 3.1.12 et 3.1.13. Le seul point qui n'est pas un analogue est le fait que $\langle S \rangle \subset \langle\!\langle S \rangle\!\rangle$ mais cela découle directement de la propriété universelle de $\langle S \rangle$ puisque que $S \subset \langle\!\langle S \rangle\!\rangle$ et $\langle\!\langle S \rangle\!\rangle$ est un sous-groupe. Dans l'avant-dernier point, la surjectivité de f est utilisée pour assurer que $f(\langle\!\langle S \rangle\!\rangle)$ est un sous-groupe distingué. □

3.4 Abélianisation

Définition 3.4.1. *Soit* G *un groupe. Le* commutateur *de deux éléments* x *et* y *de* G *est* $[x, y] = xyx^{-1}y^{-1}$. *Le* sous-groupe dérivé *de* G *est le sous-groupe engendré par les commutateurs d'éléments de* G. *On le note* $D(G)$ *ou* $[G, G]$.

La terminologie « commutateur » provient du fait que x et y commutent si et seulement si $[x, y] = 1$.

Lemme 3.4.2. *Soit* G *un groupe et* $f : G \to G'$ *un morphisme de groupes.*

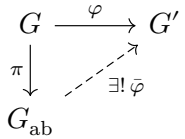
- *pour tous* x *et* y *dans* G, $f([x, y]) = [f(x), f(y)]$.
- $f(D(G)) \subset D(G')$.
- $D(G) \lhd G$

Démonstration. Le premier point est un calcul direct. Soit S l'ensemble des commutateurs d'éléments de G et S' celui de G'. Par définition, $D(G) = \langle S \rangle$ et $D(G') = \langle S' \rangle$. On a donc $f(D(G)) = f(\langle S \rangle) = \langle f(S) \rangle$ par le lemme 3.1.12. Or $f(S) \subset S'$ d'après le premier point donc $\langle f(S) \rangle \subset D(G')$, par un autre point du lemme 3.1.12.

Montrons maintenant le troisième point. Soit g dans G. Le premier point appliqué à la conjugaison c_g montre que $c_g(S) \subset S$. Le lemme 3.1.12 assure alors $c_g(D(G)) \subset D(G)$. □

Définition 3.4.3. *L'*abélianisé *d'un groupe* G *est le groupe quotient* $G_{\text{ab}} := G/D(G)$.

Proposition 3.4.4. *L'abélianisé d'un groupe* G *est un groupe abélien. Pour tout morphisme* $\varphi : G \to G'$ *à valeur dans un groupe abélien, il existe un unique morphisme* $\bar{\varphi} : G_{\text{ab}} \to G'$ *tel que* $\varphi = \bar{\varphi} \circ \pi$ *où* π *est la projection de* G *sur* G_{ab}.



Démonstration. Pour montrer que G_{ab} est abélien, il suffit de montrer que, pour tous x et y dans G, $\pi(x)$ et $\pi(y)$ commutent. Or $[\pi(x), \pi(y)] = \pi([x, y]) = 1$ d'après le lemme 3.4.2.

Soit $\varphi : G \to G'$ un morphisme à valeurs dans un groupe abéliens. On veut descendre φ au quotient G_{ab}. D'après la propriété universelle des groupes quotients (théorème 3.3.9),

il suffit de vérifier que $D(G) \subset \ker \varphi$. Comme $\ker \varphi$ est un sous-groupe, il suffit de vérifier que les commutateurs sont dans $\ker \varphi$. Soit x et y dans G. On a $\varphi([x,y]) = [\varphi(x), \varphi(y)] = 1$ où la première égalité provient encore du lemme 3.4.2 et la deuxième utilise l'hypothèse que G' est abélien. □

Cette proposition explique en quel sens G_{ab} est le plus grand quotient abélien de G. Comme d'habitude, cette propriété caractérise G_{ab} à unique isomorphisme près (comparer avec le corollaire 1.0.8 et la remarque 3.3.10). Ce qui est nouveau ici est que la construction de G_{ab} ne dépend d'aucune donnée auxiliaire (par exemple relation d'équivalence ou un sous-groupe distingué qui ne serait pas livré de série avec le groupe de départ). Le corollaire suivant montre que cette construction est *fonctorielle*, on peut aussi l'appliquer aux morphismes, de façons compatible avec la construction sur les groupes et la composition des morphismes.

Corollaire 3.4.5. *Pour tout morphisme de groupes $f : G \to G'$, il existe un unique morphisme $f_{\mathrm{ab}} : G_{\mathrm{ab}} \to {G'}_{\mathrm{ab}}$ qui fait commuter le diagramme suivant :*

$$\begin{array}{ccc} G & \xrightarrow{f} & G' \\ \downarrow & & \downarrow \\ G_{\mathrm{ab}} & \underset{\exists!\, f_{\mathrm{ab}}}{\dashrightarrow} & {G'}_{\mathrm{ab}} \end{array}$$

De plus $(\mathrm{Id}_G)_{\mathrm{ab}} = \mathrm{Id}_{G_{\mathrm{ab}}}$ et, pour tout morphisme $g : G' \to G''$, on a $(g \circ f)_{\mathrm{ab}} = g_{\mathrm{ab}} \circ f_{\mathrm{ab}}$.

Démonstration. Pour obtenir f_{ab}, on applique la propriété universelle de G_{ab} au morphisme $\pi \circ f : G \to {G'}_{\mathrm{ab}}$ (en utilisant que ${G'}_{\mathrm{ab}}$ est abélien).

Comme $\mathrm{Id}_{G_{\mathrm{ab}}} \circ \pi = \pi \circ \mathrm{Id}_G$, l'unicité dans la propriété universelle assure que $(\mathrm{Id}_G)_{\mathrm{ab}} = \mathrm{Id}_{G_{\mathrm{ab}}}$. Pour la formule de composition, on contemple le diagramme suivant

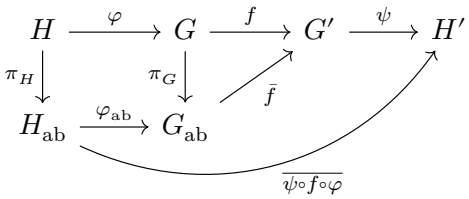
$$\begin{array}{ccccc} G & \xrightarrow{f} & G' & \xrightarrow{g} & G'' \\ \downarrow & & \downarrow & & \downarrow \\ G_{\mathrm{ab}} & \xrightarrow[f_{\mathrm{ab}}]{} & {G'}_{\mathrm{ab}} & \xrightarrow[g_{\mathrm{ab}}]{} & {G''}_{\mathrm{ab}} \end{array}$$

Comme les deux carrés commutent, le grand rectangle commute. L'unicité dans la propriété universelle appliquée à $g \circ f$ assure alors que $(g \circ f)_{\mathrm{ab}} = g_{\mathrm{ab}} \circ f_{\mathrm{ab}}$. □

Il existe un autre aspect de la compatibilité entre l'abélianisé et les morphismes. Pour tout groupe G et tout groupe abélien G', la proposition 3.4.4 fournit une bijection entre les ensembles de morphismes $\mathrm{Hom}(G, G')$ et $\mathrm{Hom}(G_{\mathrm{ab}}, G')$, qui envoie φ sur $\bar{\varphi}$ dans un sens et précompose par la projection dans l'autre. Cette famille de bijections paramétrée par G et G' se comporte bien vis à vis de la composition au départ ou à l'arrivée, au sens du lemme suivant.

Lemme 3.4.6 (Naturalité de l'adjonction entre abélianisation et inclusion des groupes abéliens)**.** *Soit $\varphi : H \to G$ un morphisme de groupes et soit $\psi : G' \to H'$ un morphismes de groupes abéliens. Pour tout $f : G \to G'$,*

$$\overline{\psi \circ f \circ \varphi} = \psi \circ \bar{f} \circ \varphi_{\mathrm{ab}}.$$



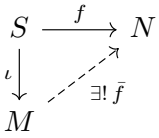
Démonstration. Par unicité dans la propriété universelle de H_{ab} appliquée à $\psi \circ f \circ \varphi$, il suffit de montrer la commutativité du diagramme de l'énoncé, c'est à dire de calculer, en regardant le diagramme, $(\psi \circ \bar{f} \circ \varphi_{\mathrm{ab}}) \circ \pi_H = \psi \circ \bar{f} \circ \pi_G \circ \varphi = \psi \circ f \circ \varphi$. □

3.5 Monoïdes libres

Cette section présente notre premier exemple de construction d'objet libre. Cet exemple est de loin le plus simple, il sert d'échauffement mais aussi de construction intermédiaire en vue de la section suivante.

En lisant la définition suivante, il faut avoir en tête la caractérisation suivante de la notion de base en algèbre linéaire : une famille de vecteurs $e : I \to E$ d'un $\mathbb{K}$-ev E est une base de E si et seulement si, pour tout $\mathbb{K}$-ev F et toute fonction $f : I \to F$, il existe une unique application $\mathbb{K}$-linéaire φ telle que $\forall i \in I, \varphi(e_i) = f_i$.

Définition 3.5.1. *Soit S un ensemble. Un monoïde libre sur S est un monoïde M muni d'une fonction $\iota : S \to M$ qui vérifie la propriété universelle suivante : pour tout monoïde N et toute fonction f de S dans N, il existe un unique morphisme $\bar{f} : M \to N$ tel que $f = \bar{f} \circ \iota$.*



On dit qu'un monoïde M est libre s'il existe un ensemble S et une fonction $\iota : S \to M$ tels que (M, ι) est un monoïde libre sur S.

On peut lire la définition ci-dessus comme : un monoïde libre sur S est un monoïde muni d'une base indexée par S, et un monoïde est libre s'il admet une base (ce qui, contrairement au cas de l'algèbre linéaire, n'est pas du tout automatique).

Exemple 3.5.2. Soit S un ensemble à un élément, noté a. Le monoïde $(\mathbb{N}, +, 0)$ muni de $a \mapsto 1$ est un monoïde libre sur S. Pour tout $f : S \to M$ et $n \in \mathbb{N}$, on a $\bar{f}(n) = f(a)^n$.

Définition 3.5.3. *Étant donné un ensemble S, un* mot *sur S est une suite finie $m = (s_1, \dots, s_n)$ d'éléments de S. On autorise le cas $n = 0$ correspondant à une suite vide notée* 1. *La* longueur *$|m|$ d'un tel mot est l'entier naturel n. L'ensemble des mots sur S est noté $M(S)$. On le munit de l'opération de concaténation qui envoie $((s_1, \dots, s_n), (t_1, \dots, t_m))$ sur $(s_1, \dots, s_n, t_1, \dots, t_m)$ de longueur $n + m$. On note i_S l'application de S dans $M(S)$ qui envoie s sur (s).*

En pratique l'application ι_S est le plus souvent implicite et on note donc $s_1 \cdots s_n$ le mot $(s_1, \dots, s_n)$.

Proposition 3.5.4. *Pour tout ensemble S, l'ensemble des mots $M(S)$ muni de la concaténation et du mot vide est un monoïde. Muni de $\iota_S : S \hookrightarrow M(S)$, il s'agit d'un monoïde libre sur S.*

Comme d'habitude, la propriété universelle des monoïdes libres assure que $M(S)$ est l'unique monoïde libre sur S modulo unique isomorphisme, on l'appelle souvent *le* monoïde libre sur S.

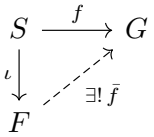
Démonstration. La vérification des axiomes de monoïdes est immédiate, il s'agit donc de démontrer la propriété universelle qui définit les monoïdes libres. Soit N un monoïde et $f : S \to N$ une application. L'unicité de $\bar{f}$ est claire car, pour tout mot $s_1 \cdots s_n$,

$$\bar{f}(s_1 \cdots s_n) = \bar{f}(s_1) \cdots \bar{f}(s_n) = f(s_1) \cdots f(s_n).$$

Il s'agit donc de montrer que l'application $\bar{f}$ définie par la formule ci-dessus est bien un morphisme de monoïde. Les deux vérifications sont immédiates. □

3.6 Groupes libres

Définition 3.6.1. *Soit S un ensemble. Un groupe libre sur S est un groupe F muni d'une application $\iota : S \to F$ qui vérifie la propriété universelle suivante : pour tout groupe G et toute fonction f de S dans G, il existe un unique morphisme de groupes $\bar{f} : F \to G$ tel que $f = \bar{f} \circ \iota$.*



Lorsque ι est l'inclusion d'une partie S de G, on dit aussi que G est librement engendré par S.

Modulo les notations, il s'agit exactement de la même définition que la définition 3.5.1 dans laquelle on a remplacé le mot monoïde par le mot groupe.

Exemple 3.6.2. Le groupe trivial est libre sur l'ensemble vide. Le groupe $(\mathbb{Z}, +)$ est libre sur tout ensemble à un élément, l'unique élément a étant envoyé sur 1. Pour tout $f : S \to G$ et $n \in \mathbb{Z}$, on a $\bar{f}(n) = f(a)^n$. De façon équivalente, on peut dire que $\mathbb{Z}$ est librement engendré par $\{1\}$.

Exemple 3.6.3. On peut montrer par un argument de ping-pong que le sous-groupe de matrices inversibles à coefficients entiers engendré par $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$ et $B = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$ est librement engendré par $\{A, B\}$.

L'existence d'un groupe libre sur n'importe quel ensemble est nettement moins évident que dans le cas des monoïdes. La construction a une saveur très informatique, avec l'apparition de notions de mots comme dans la section précédente mais aussi celle de réécriture. On verra dans la section suivante que cette construction débouche rapidement sur des problèmes algorithmiques difficiles (c'est la remarque 3.7.7).

Pour décrire l'idée, supposons que S a exactement deux éléments a et b. Partant de l'idée des mots de la section précédente, on ajoute à l'alphabet $\{a, b\}$ deux lettres notées

a^{-1} et b^{-1}. Puis on ajoute les règles de réécriture $a^{-1}a = aa^{-1} = b^{-1}b = bb^{-1} = 1$ (où 1 est toujours le mot vide). En ajoutant aucune autre règle de réécriture que celles imposées par les axiomes de groupe, on espère obtenir un objet libre. Il y a deux grandes façons d'implémenter ces règles de réécritures. On peut prendre un quotient de l'ensemble des mots ou considérer le sous-ensemble des mots réduits. On utilisera ici la deuxième option.

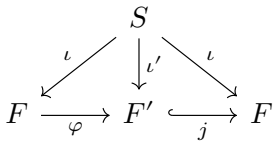
Pour tout ensemble S on note $S^{\pm 1}$ l'ensemble $S \times \{-1, 1\}$, on identifie chaque $s \in S$ avec $(s, 1) \in S^{\pm 1}$ et on définit une fonction inversion de $S^{\pm 1}$ dans lui-même par $(s, \varepsilon) \mapsto (s, -\varepsilon)$. On étend l'opérateur d'inversion à l'ensemble des mots sur $S^{\pm 1}$ en envoyant le mot vide sur lui-même et tout $m_1 \cdots m_n$ sur $m_n^{-1} \cdots m_1^{-1}$. On dit qu'un mot m admet une réduction en position i si $i < |m|$ et $m_{i+1} = m_i^{-1}$. L'opération de réduction élémentaire associée envoie m sur le mot de longueur $|m| - 2$ obtenu en retirant m_i et m_{i+1} de m. Un mot m' est une réduction d'un mot m s'il existe une suite, nécessairement finie, de réductions élémentaires menant de m à m'. Un mot est réduit s'il n'admet aucune réduction. Par exemple le mot vide est réduit et, pour $S = \{a, b\}$, le mot $aaba^{-1}$ est réduit tandis que $ab^{-1}b$ ne l'est pas. On note $F(S) \subset M(S^{\pm 1})$ l'ensemble des mots réduits sur $S^{\pm 1}$.

Soit G un groupe G et $\iota : S \to G$ une fonction. On « étend » ι à $S^{\pm 1}$ en envoyant (s, ε) sur $\iota(s)^{\varepsilon}$. La propriété universelle de $M(S^{\pm 1})$ étend cette fonction en morphisme de monoïdes $\bar{\iota} : M(S^{\pm 1}) \to G$. Les éléments de l'image de $\bar{\iota}$ sont appelés mots en les éléments de $\iota(S)^{\pm 1}$ dans G.

La proposition suivante est une caractérisation des groupes libres analogue au fait qu'une famille de vecteurs constitue une base si et seulement si tout vecteur s'écrit de façon unique comme combinaison linéaire des éléments de la famille. Le cas des groupes est compliqué par l'absence de commutativité.

Proposition 3.6.4. *Soit S un ensemble, F un groupe et $\iota : S \to F$ une fonction. La paire (F, ι) est un groupe libre sur S si et seulement si $\bar{\iota}$ restreinte à $F(S)$ est une bijection. Autrement dit, tout élément de F s'écrit de façon unique comme mot réduit en les éléments de $\iota(S)^{\pm 1}$.*

Démonstration. Supposons que (F, ι) est un groupe libre sur S. On commence par démontrer que $\iota(S)$ engendre F. On pose $F' = \langle \iota(S) \rangle$, on note j l'inclusion de F' dans F et on note ι' l'application ι vue comme allant de S dans F'. Par définition, $\iota = j \circ \iota'$. La propriété universelle de (F, ι) appliquée à ι' fournit un morphisme $\varphi : F \to F'$ tel que $\iota' = \varphi \circ \iota$. On a donc le diagramme commutatif suivant.



L'unicité dans la propriété universelle de (F, ι) appliquée à ι assure que $j \circ \varphi = \mathrm{Id}$. En particulier j est surjective donc $F' = F$.

Ainsi $F = \langle \iota(S) \rangle$ et le lemme 3.1.13 assure que tout élément de F est dans l'image $\bar{\iota}(M(S^{\pm 1}))$. Après réductions, on obtient une écriture comme mot réduit. Il reste à montrer l'unicité de cette écriture. On considère la fonction f de S dans $\mathfrak{S}(F(S))$ qui envoie

s sur

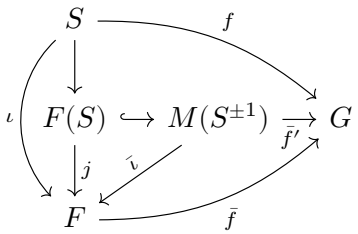
$$m_1 \cdots m_n \mapsto \begin{cases} s m_1 \cdots m_n \text{ si } m_1 \neq s^{-1} \\ m_2 \cdots m_n \text{ si } m_1 = s^{-1} \end{cases}$$

il s'agit bien d'une bijection de $F(S)$ car elle admet pour inverse

$$m_1 \cdots m_n \mapsto \begin{cases} s^{-1} m_1 \cdots m_n \text{ si } m_1 \neq s \\ m_2 \cdots m_n \text{ si } m_1 = s \end{cases}$$

En particulier, pour tout $s \in S$, $f(s)$ envoie le mot vide sur s et $f(s)^{-1}$ l'envoie sur s^{-1}. La propriété universelle de (F, ι) étend f en morphisme $\bar{f} : F \to \mathfrak{S}(F(S))$. Soit $x = \iota(s_1)^{\varepsilon_1} \cdots \iota(s_n)^{\varepsilon_n}$ un mot réduit. Comme $\bar{f}$ est un morphisme et $\bar{f} \circ \iota = f$, $\bar{f}(x) = f(s_1)^{\varepsilon_1} \cdots f(s_n)^{\varepsilon_n}$. Ainsi $\bar{f}(x)$ envoie le mot vide sur $s_1^{\varepsilon_1} \cdots s_n^{\varepsilon_n}$ et on retrouve l'écriture en mot réduit de x qui est donc unique.

Réciproquement supposons que la restriction j de $\bar{\iota}$ à $F(S)$ est une bijection. Soit G un groupe et f une fonction de S dans G. On étend f en $f' : S^{\pm 1}$ en envoyant s^{-1} sur $f(s)^{-1}$. On pose $\bar{f} = \bar{f}' \circ j^{-1}$, où $\bar{f}' : M(S^{\pm 1}) \to G$ provient de la propriété universelle de $M(S^{\pm 1})$.



Cette fonction de F dans G vérifie $\bar{f} \circ \iota = f$ par construction. Il reste à montrer qu'il s'agit d'un morphisme. Ce n'est pas évident car $F(S)$ n'est pas un sous-monoïde de $M(S^{\pm 1})$: la concaténation de deux mots réduits n'est pas forcément réduite. Soit x et y dans F. Soient $m = j^{-1}(x)$ et $p = j^{-1}(y)$ les mots en $S^{\pm 1}$ qui représentent x et y respectivement.

Montrons que $\bar{f}(xy) = \bar{f}(x)\bar{f}(y)$ par récurrence sur $|m|$. Si $|m| = 0$ alors $x = 1$ et l'égalité est claire. Supposons le résultat connu pour les mots de longueur n et supposons $|m| = n + 1$. On écrit $m = m's$ pour $s \in S^{\pm 1}$ et $|m'| = n$. Si s n'est pas l'inverse de la première lettre de p alors mp est réduit et l'égalité visée est évidente. Supposons donc que $p = s^{-1}p'$ pour un mot réduit p'. On pose $x' = j(m')$ et $y' = j(p')$. On a $xy = x'y'$ et on calcule

$$\begin{aligned} \bar{f}(xy) &= \bar{f}(x'y') \\ &= \bar{f}(x')\bar{f}(y') \text{ par hypothèse de récurrence} \\ &= \bar{f}(x')f'(s)f'(s)^{-1}\bar{f}(y') \\ &= \bar{f}(x)\bar{f}(y). \end{aligned}$$

Notons qu'on n'affirme pas que $m'p'$ est réduit (c'est faux en général) et que la simplification de la première ligne a lieu dans F tandis que la complexification de la troisième ligne a lieu dans G. □

Le lemme fondamental qui fait fonctionner toute la construction des groupes libres est le résultat suivant.

Lemme 3.6.5. *Soit S un ensemble. Pour tout mot $m \in M(S^{\pm 1})$, il existe un unique mot réduit obtenu par réduction de m.*

L'unicité n'est pas claire car m peut admettre plusieurs réductions et la provenance des lettres survivantes après réduction n'est pas unique Par exemple, si $m = ss^{-1}s$, on peut réduire en éliminant les deux premières lettres, la troisième survivant ou bien on peut réduire en éliminant les deux dernières lettres, la première survivant. Dans les deux cas on trouve comme mot réduit s.

Démonstration. L'existence est claire car chaque réduction diminue strictement la longueur de m et que le mot vide est réduit. Montrons l'unicité par récurrence forte sur $|m|$. Supposons le résultat connu pour tous les mots de longueur strictement plus petite que $|m|$. Soit $m, p^1, \dots, p^n$ et $m, q^1, \dots, q^k$ deux suites de réductions élémentaires aboutissant à des mots p^n et q^k réduits (ici les exposants ne sont pas des puissances mais une simple numérotation mise en exposant pour éviter une confusion avec l'énumération des lettres de chaque mot). Il a deux cas à considérer. Si les réductions de m à p^1 et q^1 se recouvrent alors $p^1 = q^1$ et on conclut directement par hypothèse de récurrence. En effet la situation est soit le cas trivial où c'est la même réduction soit $m = m'xx^{-1}xm'$ pour un certain $x \in S^{\pm 1}$ où une des réduction élimine xx^{-1} et l'autre élimine $x^{-1}x$. L'autre cas est celui de réductions disjointes. On a alors $m = m^1xx^{-1}m^2yy^{-1}m^3$ pour des mots m^1, m^2 et m^3 (éventuellement vides), et x et y dans $S^{\pm 1}$ et p^1 et q^1 sont obtenus en réduisant une des deux paires. Dans ce cas p^1 et q^1 admettent la réduction élémentaire commune $r^1 = m^1m^2m^3$. On utilise ensuite l'existence pour réduire r^1 en r^l. Par hypothèse de récurrence appliquée à p^1 et r^1, $p^n = r^l$. Par hypothèse de récurrence appliquée à q^1 et r^1, $q^k = r^l$. Ainsi $p^n = q^k$, ce qu'on voulait démontrer. □

Le lemme ci-dessus permet de définir une loi de composition interne sur $F(S)$: la concaténation suivie de la réduction. Le mot vide est clairement neutre pour cette opération. Par ailleurs l'inversion des mots préserve $F(S)$ et on a une application $\iota_S : S \to F(S)$ qui envoie s sur le mot s.

Théorème 3.6.6. *Pour tout ensemble S, l'ensemble $F(S)$ des mots réduits sur $S^{\pm 1}$ munis de la structure décrite ci-dessus forme un groupe libre sur S.*

Démonstration. Le seul axiome de groupe qui n'est pas évident est l'associativité, mais elle découle facilement du lemme 3.6.5. En effet pour tous mots réduits x, y et z, les produits $(xy)z$ et $x(yz)$ sont tous deux des réductions de la concaténation de x, y et z. Le fait que $(F(S), \iota_S)$ est libre sur S découle immédiatement de la proposition 3.6.4. □

On appelle souvent $F(S)$ *le* groupe libre sur S. L'analogue de ce léger abus de langage pour les $\mathbb{R}$-espaces vectoriels serait d'appeler $\mathbb{R}^n$ *le* $\mathbb{R}$-espace vectoriel muni d'une base indexée par $\{1, \dots, n\}$.

Corollaire 3.6.7. *Tout groupe est quotient d'un groupe libre. Si une partie S engendre un groupe G alors G est un quotient de $F(S)$.*

Démonstration. La première partie découle de la seconde puisque tout groupe G est engendré par l'ensemble de tous ses éléments. Montrons la seconde partie. Soit G un groupe et S une partie génératrice de G. La propriété universelle de $F(S)$ appliquée à

l'inclusion $i : S \hookrightarrow G$ fournit un morphisme de groupe $\pi : F(S) \to G$. On a $\pi(F(S)) = \pi(\langle \iota_S(S) \rangle) = \langle \pi(\iota_S(S)) \rangle = \langle \iota(S) \rangle = G$ donc π est bien surjective. □

Comme dans le cas de l'abélianisation, la propriété universelle qui définit les groupes libres assure la fonctorialité de la construction ci-dessus, au sens de la proposition suivante :

Proposition 3.6.8. *Soit S et S' des ensembles. Pour toute fonction $f : S \to S'$, il existe une unique fonction $F(f) : F(S) \to F(S')$ telle que $F(f) \circ \iota_S = \iota_{S'} \circ f$.*

$$\begin{array}{ccc} S & \xrightarrow{f} & S' \\ {\scriptstyle \iota_S}\downarrow & & \downarrow{\scriptstyle \iota_{S'}} \\ F(S) & \underset{\exists!\, F(f)}{\dashrightarrow} & F(S') \end{array}$$

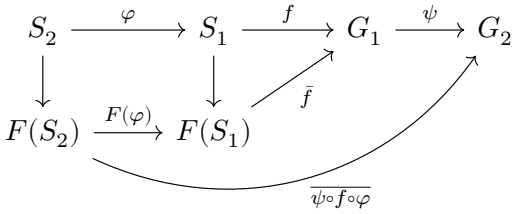
De plus $F(\mathrm{Id}) = \mathrm{Id}$ et cette opération est compatible avec la composition : $F(g \circ f) = F(g) \circ F(f)$.

Pour tout groupe G, la bijection $\mathrm{Hom}(S, G) \simeq \mathrm{Hom}(F(S), G)$ est naturelle : pour toute fonction $\varphi : S_2 \to S_1$ et tout morphisme $\psi : G_1 \to G_2$, on a pour tout fonction $f : S_1 \to G_1$, $\psi \circ \overline{f} \circ F(\varphi) = \overline{\psi \circ f \circ \varphi}$.

Démonstration. C'est exactement la même démonstration que dans le cas de l'abélianisation. On définit $F(f)$ comme $\overline{\iota_{S'} \circ f}$. Comme $\mathrm{Id}_{F(S)} \circ \iota_S = \iota_S \circ \mathrm{Id}_S$, l'unicité dans la propriété universelle assure $F(\mathrm{Id}) = \mathrm{Id}$. La formule de composition découle de la commutativité de :

$$\begin{array}{ccccc} S_1 & \xrightarrow{f} & S_2 & \xrightarrow{g} & S_3 \\ \downarrow & & \downarrow & & \downarrow \\ F(S_1) & \xrightarrow[F(f)]{} & F(S_2) & \xrightarrow[F(g)]{} & F(S_3) \end{array}$$

et de l'unicité dans la construction de $F(g \circ f)$. Pour la naturalité on contemple le diagramme suivant :



et la formule découle de l'unicité dans la propriété universelle de $F(S_2)$ appliquée à $\psi \circ f \circ \varphi$. □

La démonstration ci-dessus suit exactement celle des corollaires 3.4.5 et 3.4.6. Ce fait est étonnant car l'abélianisation d'un morphisme de groupes transforme un morphisme en morphisme entre des groupes *plus petits* tandis que la transformation d'une fonction f en morphisme de groupes libres est une *extension* à des objets bien plus gros (par exemple $F(S)$ est infini dès que S n'est pas vide). Ces deux opérations peuvent se traiter

exactement de la même façon car les propriétés universelles correspondantes ont exactement la même forme abstraite. Bien sûr on peut aussi expliciter $F(f)$ et démontrer à la main la formule de composition, mais on perd alors le parallèle formel (et la possibilité de trouver l'énoncé abstrait qui englobe les deux cas).

Théorème 3.6.9. *Soit S et S' deux ensembles finis. Les groupes $F(S)$ et $F(S')$ sont isomorphes si et seulement si S et S' ont le même cardinal.*

Démonstration. Supposons d'abord que S et S' ont même cardinal. On obtient alors une bijection $f : S \to S'$ puis, par la proposition précédente, des morphismes $F(f) : F(S) \to F(S')$ et $F(f^{-1}) : F(S) \to F(S')$ tels que $F(f) \circ F(f^{-1}) = F(f \circ f^{-1}) = F(\mathrm{Id}) = \mathrm{Id}$ et de même $F(f^{-1}) \circ F(f)$ donc $F(f)$ est un isomorphisme.

Réciproquement, supposons qu'il existe un isomorphisme $\varphi : F(S) \to F(S')$. La précomposition par φ fournit une bijection entre les ensembles de morphismes de groupes $\mathrm{Hom}(F(S'), \mathbb{Z}/2\mathbb{Z})$ et $\mathrm{Hom}(F(S), \mathbb{Z}/2\mathbb{Z})$. Or les propriétés universelles de $F(S')$ et $F(S)$ respectivement mettent en bijection ces ensembles avec les ensembles de fonctions de S' dans $\mathbb{Z}/2\mathbb{Z}$ et de S dans $\mathbb{Z}/2\mathbb{Z}$ qui sont de cardinaux $2^{\sharp S'}$ et $2^{\sharp S}$ respectivement. Ainsi $2^{\sharp S'} = 2^{\sharp S}$ et, comme S et S' sont finis, on conclut que $\sharp S' = \sharp S$. □

Le théorème précédent reste vrai si S et S' sont infinis mais la démonstration est plus compliquée. La démonstration précédente fonctionne jusqu'à l'égalité $2^{\sharp S'} = 2^{\sharp S}$ mais le passage à $\sharp S = \sharp S'$ est indépendant des axiomes usuels de la théorie des ensembles (même en y ajoutant l'hypothèse du continu, il faut ajouter une hypothèse du continu généralisée). Pour contourner ce problème on peut se ramener au cas des groupes abéliens libres (discuté dans le chapitre 5) puis au théorème de la dimension (à condition de disposer de celui-ci en dimension infinie).

Au vu de ce théorème, on pourrait imaginer que si $F(S)$ s'injecte comme sous-groupe de $F(S')$ alors $\sharp S \leqslant \sharp S'$ (la réciproque est claire). Mais ce n'est pas du tout ce qui se passe.

Proposition 3.6.10. *Soit S un ensemble à deux éléments. Il existe un morphisme injectif de $F(\mathbb{N})$ dans $F(S)$. En particulier $F(S)$ contient des sous-groupes isomorphes à $F(\{1, \dots n\})$ pour tout n.*

Démonstration. On note a et b les deux éléments de S. On définit $f : \mathbb{N} \to F(S)$ par $n \mapsto c_{b^n}(a) \in F(S)$. Montrons que le morphisme $\varphi : F(\mathbb{N}) \to F(S)$ correspondant à f est injectif. Soit $x = n_1^{\varepsilon_1} \cdots n_N^{\varepsilon_N}$ un élément de $F(\mathbb{N})$ écrit comme mot réduit non vide, avec chaque n_1 dans $\mathbb{N}$. On veut montrer $\varphi(x) \neq 1$. Montrons par récurrence sur N qu'on peut écrire $\varphi(x)$ sous la forme $b^{m_1}a^{j_1} \cdots b^{m_M}a^{j_M}b^{-n_N}$ où $M \geqslant 1$, les m_i et les j_i sont dans $\mathbb{Z}$ et non nuls, et j_M est du signe de ε_N. Ce résultat est plus fort que $\varphi(x) \neq 1$. Pour $N = 1$ le résultat est clair. Supposons-le jusqu'à N. On a alors par hypothèse de récurrence

$$\varphi(x) = b^{m_1}a^{j_1} \cdots b^{m_M}a^{j_M}b^{-n_N}b^{n_{N+1}}a^{\varepsilon_{N+1}}b^{-n_{N+1}}$$

avec j_M du signe de ε_N. Le mot de départ était réduit donc il y a deux possibilités. Si $n_{N+1} \neq n_N$ on a :

$$\varphi(x) = b^{m_1}a^{j_1} \cdots b^{m_M}a^{j_M}b^{n_{N+1}-n_N}a^{\varepsilon_{N+1}}b^{-n_{N+1}}$$

qui est bien de la forme annoncée car $n_{N+1} - n_N \neq 0$. Sinon, $n_{N+1} = n_N$ mais $\varepsilon_{N+1} = \varepsilon_N$. Dans ce cas

$$\varphi(x) = b^{m_1} a^{j_1} \cdots b^{m_M} a^{j_M + \varepsilon_{N+1}} b^{-n_{N+1}}$$

et j_M est du même signe que ε_N donc du même signe que ε_{N+1}. Ainsi $j_M + \varepsilon_{N+1}$ n'est pas nul et est du même signe que ε_{N+1}. □

Remarque 3.6.11. On peut montrer que tout sous-groupe d'un groupe libre est libre, c'est le théorème de Nielsen-Schreier. Toutes les démonstrations éclairantes de ce théorème font intervenir de la géométrie.

3.7 Présentations de groupes

Définition 3.7.1. *Soit S un ensemble et R une partie de $F(S)$. Le groupe défini par l'ensemble de générateurs S et l'ensemble de relations R est le quotient $F(S)/\langle\langle R \rangle\rangle$. On le note $\langle S \mid R \rangle$.*

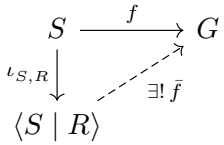
Lorsque S et R sont finis et explicite on utilise cette notation en listant les éléments de S et de R sans les entourer d'accolades. Lorsqu'un groupe G est déjà connu par ailleurs et muni d'une famille génératrice S, on dit aussi que $\langle S \mid R \rangle$ est une *présentation* de G si $\langle\langle R \rangle\rangle$ est le noyau de l'application de $F(S)$ dans G induite par l'inclusion de S dans G (comme dans la démonstration du corollaire 3.6.7). Par extension on appelle aussi présentation d'un groupe G toute paire (S, R) telle que G est isomorphe à $\langle S \mid R \rangle$. On dit qu'un groupe G est de *type fini* s'il admet une famille génératrice finie et de *présentation finie* s'il admet une présentation (S, R) où S et R sont tous deux finis.

On notera que S s'envoie dans $\langle S \mid R \rangle$ en composant ι_S et la projection de $F(S)$ sur $\langle S \mid R \rangle$. En pratique la composée $\iota_{S,R}$ de ces deux applications est presque toujours implicite quand on l'applique à un élément : on utilise le même symbole pour un élément de S et son image.

Exemple 3.7.2. Le groupe $\mathbb{Z}/n\mathbb{Z}$ admet pour présentation $\langle x \mid x^n \rangle$. En effet $\mathbb{Z}$ est isomorphe à $F(\{x\})$ par $m \mapsto x^m$ et $n\mathbb{Z}$ est envoyé sur le sous-groupe (distingué) engendré par x^n.

Proposition 3.7.3 (Propriété universelle des présentations de groupes)**.** *Soit S un ensemble et R une partie de $F(S)$. L'application $\iota_{S,R} \colon S \to \langle S \mid R \rangle$ vérifie que, pour tout $r = \iota_S(s_1)^{\varepsilon_1} \cdots \iota_S(s_n)^{\varepsilon_n}$ dans R, $\iota_{S,R}(s_1)^{\varepsilon_1} \cdots \iota_{S,R}(s_n)^{\varepsilon_n} = 1$ (on dit que $\iota_{S,R}$ est compatible avec R). De plus l'image de $\iota_{S,R}$ engendre $\langle S \mid R \rangle$.*

Soit G un groupe et f une application de S dans G. Si f est compatible avec R alors il existe un unique morphisme $\bar{f} \colon \langle S \mid R \rangle \to G$ qui fait commuter



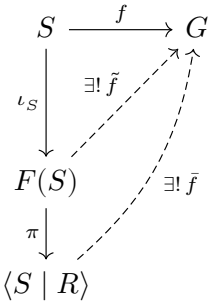
Le morphisme $\bar{f}$ est surjectif si et seulement si $f(S)$ engendre G.

Démonstration. Soit $r = s_1^{\varepsilon_1} \cdots s_n^{\varepsilon_n}$ dans R. On a

$$\begin{aligned}\iota_{S,R}(s_1)^{\varepsilon_1} \cdots \iota_{S,R}(s_n)^{\varepsilon_n} &= \pi(\iota_S(s_1))^{\varepsilon_1} \cdots \pi(\iota_S(s_n)^{\varepsilon_n} \\ &= \pi(r) \\ &= 1.\end{aligned}$$

On sait déjà que $\iota_S(S)$ engendre $F(S)$. Comme la projection de $F(S)$ sur $\langle S \mid R\rangle$ est surjective, $\iota_{S,R}(S)$ engendre $\langle S \mid R\rangle$.

Soit G un groupe et $f\colon S \to G$ une application compatible avec R. On a le diagramme suivant :



L'existence $\tilde{f}$ faisant commuter le triangle du haut découle directement de la propriété universelle de $F(S)$ appliquée à f, il n'y a aucune condition à vérifier. L'existence de $\bar{f}$ faisant commuter le triangle du bas provient de la propriété universelle des groupes quotients, il suffit de vérifier que $R \subset \ker \tilde{f}$, ce qui est précisément la condition de compatibilité imposée.

Il reste à expliquer comment les unicités des deux propriétés universelles invoquées se composent pour donner l'unicité voulue ici. Supposons que $\hat{f} : \langle S \mid R\rangle \to G$ convient aussi. On a alors $\hat{f} \circ (\pi \circ \iota_S) = f$ donc $(\hat{f} \circ \pi) \circ \iota_S = f$ et l'unicité dans la propriété universelle de $F(S)$ assure que $\hat{f} \circ \pi = \tilde{f}$. L'unicité dans la propriété universelle du quotient assure ensuite que $\hat{f} = \bar{f}$.

Enfin on a $\bar{f}(\langle S \mid R\rangle) = \bar{f}(\langle \iota_{S,R}(S)\rangle) = \langle \bar{f}(\iota_{S,R}(S))\rangle = \langle f(S)\rangle$, ce qui démontre le critère de surjectivité. □

Exemple 3.7.4. Pour tout groupe G, l'application de $\mathrm{Hom}(\mathbb{Z}/n\mathbb{Z}, G)$ dans $\{x \in G \mid x^n = 1\}$ qui envoie φ sur $\varphi(1)$ est une bijection. Un morphisme φ est surjectif si et seulement si $\varphi(1)$ engendre G et il est injectif si et seulement si $\varphi(1)$ est d'ordre n.

Montrons que $G := \langle a, b \mid a^2, b^2, (ab)^3\rangle$ est isomorphe à $\mathfrak{S}_3$. On considère l'application $f\colon \{a, b\} \to \mathfrak{S}_3$ qui envoie a sur la transposition (12) et b sur la transposition (23). Cette application vérifie la condition de propriété universelle car $f(a)^2$, $f(b)^2$ et $(f(a)f(b))^3$ valent tous 1. De plus $\{f(a), f(b)\}$ engendre $\mathfrak{S}_3$ donc $\bar{f}$ est surjective. Il suffit maintenant de montrer que $\sharp G \leqslant \sharp\mathfrak{S}_3$. Puisque a et b engendrent G, tout élément de G s'écrit comme $\prod_{i=1}^{n} a^{k_i} b^{l_i}$ avec n un entier naturel, chaque k_i et l_i dans $\mathbb{Z}$ et non nul sauf éventuellement k_1 ou l_n. En utilisant que $a^2 = b^2 = 1$, on peut supposer que chaque k_i et l_i vaut un sauf éventuellement k_1 ou l_n. Enfin la relation $(ab)^3 = 1$ assure que $bab = a^{-1}b^{-1}a^{-1}$, c'est à dire $bab = aba$ en utilisant encore $a^2 = b^2 = 1$. Donc on peut supposer qu'il n'y a qu'une seul b dans l'écriture des éléments de G. Ainsi $G \subset \{1, a, b, ab, ba, aba\}$ et $\sharp G \leqslant 6 = \sharp\mathfrak{S}_3$.

On a donc trouvé une présentation de $\mathfrak{S}_3$ par les générateurs $\tau_1 := (12)$ et $\tau_2 = (23)$ et les relations τ_1^2, τ_2^2 et $(\tau_1\tau_2)^3$. Ainsi pour tout groupe H, les images de τ_1 et τ_2 fournissent une bijection de $\mathrm{Hom}(\mathfrak{S}_3, H)$ vers $\{(x, y) \in H \mid x^2 = y^2 = (xy)^3 = 1\}$.

Corollaire 3.7.5. *Soit S et S' des ensembles, R une partie de $F(S)$ et R' une partie de $F(S')$. Pour tout fonction $f : S \to F(S')$, si $\bar{f}(R) \subset \langle\langle R' \rangle\rangle$ alors il existe un unique morphisme de groupe $\bar{f} : \langle S \mid R\rangle \to \langle S' \mid R'\rangle$ qui fait commuter le digramme*

$$\begin{array}{ccc} S & \xrightarrow{\quad f \quad} & F(S') \\ {\scriptstyle \iota_{S,R}}\downarrow & & \downarrow{\scriptstyle \pi} \\ \langle S \mid R\rangle & \xrightarrow[\exists!\, \bar{f}]{} & \langle S' \mid R'\rangle \end{array}$$

Démonstration. On applique la propriété universelle de $\langle S \mid R\rangle$ à $\pi \circ f$, en utilisant que π envoie $\langle\langle R' \rangle\rangle$ sur $\{1\}$. □

Exemple 3.7.6. On peut utiliser les présentations de groupes pour construire le *coproduit* (ou « produit libre ») $\ast_i G_i$ d'une famille de groupes G_i. Il s'agit d'un groupe muni de morphismes (injectifs) $\varphi_j : G_j \to \ast_i G_i$ vérifiant la propriété universelle duale de celle du produit : pour tout groupe H et toute collection de morphismes de groupes $\psi_i : G_i \to H$, il existe un unique morphisme $\Psi : \ast_i G_i \to H$ tel que $\Psi \circ \varphi_j = \psi_j$ pour tout j. Pour chaque i, on note $\pi_i : F(G_i) \to G_i$ le morphisme induit par l'identité et $\theta_i : F(G_i) \to F(\bigsqcup_j G_j)$ le morphisme induit par l'inclusion. On pose

$$\ast_i G_i = \left\langle \bigsqcup_i G_i \;\middle|\; \bigcup_i \theta_i(\ker \pi_i) \right\rangle .$$

La démonstration de la propriété universelle est laissée en exercice.

Remarque 3.7.7. La notion de présentation de groupe semble être un façon très concrète de manipuler les éléments d'un groupe, mais elle cache un redoutable problème algorithmique, connus sous le nom de « problème du mot ». On peut donner des exemples explicites de présentations de groupes pour lesquelles on peut démontrer qu'il n'existe aucun algorithme permettant de décider si deux mots en les générateurs et leurs inverses représentent le même élément du groupe (ou, de façon équivalente, si un mot représente l'élément neutre).

4 Anneaux et corps

4.1 Définitions, morphismes et sous-objets

Définition 4.1.1. *Un* anneau *(unitaire) est un ensemble A muni de deux lois de composition internes, notées additivement et multiplicativement, de deux éléments 0 et 1 et d'une fonction $-$ de A dans lui-même tels que :*

- $(A, +, 0, -)$ *est un groupe commutatif*
- $(A, \times, 1)$ *est un monoïde*
- 0 *est absorbant à gauche et à droite pour la multiplication :* $\forall a, 0a = a0 = 0$.
- *la multiplication est distributive à gauche et à droite sur l'addition :* $\forall abc, a(b+c) = ab + ac$ *et* $(b + c)a = ba + ca$.

Exemple 4.1.2. Munis de leurs lois et neutres usuels, les ensembles suivants sont des anneaux : $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{Z}/n\mathbb{Z}$. Si A un anneau, M un groupe abélien, $\mathbb{K}$ est un corps et E un $\mathbb{K}$-espace vectoriel, les ensembles $\mathrm{End}(M)$, $\mathbb{K}[X]$, $\mathcal{M}_n(A)$, $\mathrm{End}(E)$ ainsi que l'ensemble des fonctions d'un ensemble X fixé dans A sont des anneaux pour leurs lois et neutre usuels.

Si un ensemble A ne contient qu'un élément, noté 0, alors en utilisant cet élément aussi comme 1, l'unique loi de composition interne de A comme addition et multiplication ainsi que l'unique fonction de A dans lui-même comme négation forment une structure d'anneau. On dit que A est *trivial*.

Comme dans le cas des groupes, on peut radiner sur les axiomes dans la définition précédente. Sans tomber dans l'excès, on peut noter les observations suivantes.

Lemme 4.1.3. *Dans la définition d'un anneau, la commutativité de l'addition et l'absorptivité de zéro pour la multiplication sont redondantes avec les autres conditions, on peut les omettre (simultanément).*

Démonstration. Soit a et b dans un anneau A. La distributivité à gauche puis à droite et la neutralité de 1 donnent $(1 + 1)(a + b) = (1 + 1)a + (1 + 1)b = 1a + 1a + 1b + 1b = a + a + b + b$. En appliquant d'abord la distributivité à droite puis à gauche on calcule $(1 + 1)(a + b) = 1(a + b) + 1(a + b) = 1a + 1b + 1a + 1b = a + b + a + b$. Ainsi $a + a + b + b = a + b + a + b$. Comme $(A, +)$ est un groupe, on peut simplifier à gauche par a et à droite par b pour obtenir $a + b = b + a$.

On se concentre maintenant sur l'absorptivité. On calcule $a0 = a(0 + 0) = a0 + a0$ et on simplifie par $a0$ pour obtenir $a0 = 0$. On montre de même que $0a = 0$. □

Ainsi la commutativité de l'addition dans un anneau n'est pas négociable. En revanche on ne demande pas que la multiplication soit commutative et cela n'a rien d'automatique. L'astuce de la commutativité automatique ne fonctionne pas non plus dans la variante affaiblie suivante (qui ne servira que très épisodiquement, dans la proposition 4.3.26 et la section 5.3).

Définition 4.1.4. *Un* semi-anneau *est un ensemble A muni de deux lois de composition internes $+$ et $\times$ et de deux éléments 0 et 1 tels que $(A, +, 0)$ est un monoïde commutatif, $(A, \times, 1)$ est un monoïde, la multiplication est distributive sur l'addition à gauche et à droite et zéro est absorbant pour la multiplication à gauche et à droite. Ainsi c'est un anneau dans lequel on ne demande pas l'existence d'opposés.*

Exemple 4.1.5. $(\mathbb{N}, +, \times, 0, 1)$ est un semi-anneau.

Remarque 4.1.6. Dans un anneau A, $0 = 1$ si et seulement si A est trivial.

Démonstration. Supposons que $0 = 1$. Soit a dans A. On calcule $a = 1a = 0a = 0$. La réciproque est claire par définition. □

Définition 4.1.7. *Soit A un anneau.*

- *On dit que A est* commutatif *si sa multiplication l'est.*
- *Le groupe des* unités *de A, noté $A^\times$, est le groupe des unités du monoïde $(A, \times)$, c'est à dire l'ensemble des éléments de A qui sont inversibles pour la multiplication.*
- *On dit que A est un* corps gauche *s'il n'est pas trivial et si $A^\times = A \setminus \{0\}$.*
- *On dit que A est un* corps *si A est un corps gauche commutatif.*
- *On dit qu'un élément x de A est un* diviseur de zéro *s'il existe $y \neq 0$ tel que $xy = 0$ ou $yx = 0$.*
- *On dit que A est* intègre *s'il est commutatif, non trivial et ne possède aucun diviseur de zéro non nul. Autrement dit, A est commutatif, non trivial et, pour tous x et y dans A, $xy = 0 \Rightarrow x = 0$ ou $y = 0$.*

Exemple 4.1.8. Pour tout entier n, l'anneau $\mathbb{Z}/n\mathbb{Z}$ a pour unités les images des entiers qui sont premiers à n. Il est intègre si et seulement si n est premier. Dans ce cas il s'agit d'un corps.

On ne confondra pas la notation $A^\times$ avec la notation A^* qui désigne parfois $A \setminus \{0\}$ (mais qu'on n'utilisera pas dans ce cours). Les deux ne coïncident que dans le cas de l'anneau trivial et celui des corps gauche.

Définition 4.1.9. *Un* morphisme *entre deux anneaux A et B est une application $f : A \to B$ telle que :*

- *f est un morphisme de groupes de $(A, +)$ dans $(B, +)$,*
- *f est un morphisme de monoïdes de $(A, \times)$ dans $(B, \times)$.*

Il est important de noter que la condition d'unitarité dans le deuxième point n'est pas automatique. Par exemple la fonction nulle entre deux anneaux est additive et multiplicative mais n'est unitaire que si l'anneau but est trivial.

Lemme 4.1.10. *Si un morphisme d'anneau est bijectif alors sa réciproque est un morphisme d'anneau. On dit alors que ce morphisme est un isomorphisme d'anneaux.*

Démonstration. Soit $f : A \to B$ un morphisme d'anneaux bijectif. Le lemme 3.1.3 assure que la réciproque f^{-1} est un morphisme de groupes et le lemme 2.0.7 assure que c'est un morphisme de monoïdes. □

Lemme 4.1.11. *Pour tout anneau A, il existe un unique morphisme d'anneaux de $\mathbb{Z}$ dans A.*

Démonstration. Montrons l'unicité. Soit φ un morphisme de $\mathbb{Z}$ dans A et n dans $\mathbb{Z}$. $\varphi(n) = \varphi(n1) = n\varphi(1)$ car φ est un morphisme de groupes. De plus $\varphi(1) = 1$ donc $\varphi(n) = n1$ (on remarquera que c'est l'unicité dans la propriété universelle de $\mathbb{Z}$ comme groupe libre).

Pour l'existence, il suffit de montrer que la formule ci-dessus définit bien un morphisme d'anneau. Les vérifications sont immédiates. □

Définition 4.1.12. *Soit $(A_i)_{i\in\mathcal{J}}$ une famille d'anneaux. Le* produit *des A_i est l'ensemble $P = \prod_i A_i$ équipé de sa structure de groupe additif produit et de monoïde multiplicatif produit.*

Remarque 4.1.13. Les projections sur les facteurs d'un produit d'anneaux sont des morphismes d'anneaux. Le produit d'une famille d'anneaux vérifie une propriété universelle analogue à celle des produits de groupes : pour construire un morphisme d'anneaux à valeurs dans un produit il suffit de donner des morphismes à valeurs dans chacun des facteurs. Comme dans le cas des groupes, la construction de coproduits (vérifiant la propriété universelle duale) est nettement plus compliquée.

Définition 4.1.14. *Un* sous-anneau *de A est un sous-groupe de $(A, +)$ qui est aussi un sous-monoïde de $(A, \times)$.*

Comme dans le cas des groupes, la condition ci-dessus assure qu'un sous-anneau hérite d'une structure d'anneau. Contrairement au cas des groupes, il s'agit essentiellement du seul intérêt de cette définition car cette condition n'est pas suffisante pour quotienter.

Lemme 4.1.15. *Soit A et A' des anneaux $f : A \to A'$ un morphisme.*

- *L'image par f d'un sous-anneau de A est un sous-anneau de A'. En particulier* $\mathrm{im}(f)$ *est un sous-anneau.*
- *La préimage par f d'un sous-anneau de A' est un sous-anneau de A. Il faut prendre garde au fait que $\{0\}$ n'est pas sous-anneau de A' (sauf si A' est trivial) donc ce point n'affirme rien sur* $\ker(f)$. *La bonne notion concernant* $\ker(f)$ *sera discutée dans la section suivante.*
- *L'intersection d'une famille de sous-anneaux de A est un sous-anneau de A.*

Démonstration. Tout découle directement des lemmes 2.0.13 et 3.1.8 concernant les sous-monoïdes et les sous-groupes. □

Remarque 4.1.16. Le lemme précédent permet de construire le *sous-anneau engendré par une partie* d'un anneau exactement comme dans le cas des monoïdes et des groupes, avec les mêmes propriétés.

4.2 Anneaux quotients

Définition 4.2.1. *Un* quotient *d'un anneau A est un anneau B muni d'un morphisme $\pi : A \to B$ surjectif.*

On sait déjà que le noyau d'un morphisme d'anneau $\varphi : A \to B$ est un sous-groupe (distingué mais cette précision est inutile pour cause de commutativité). On observe aussi que, pour tous x et y dans A, si $x \in \ker \varphi$ alors $xy \in \ker \varphi$, car $\varphi(xy) = \varphi(x)\varphi(y) = 0\varphi(y) = 0$ et de même $yx \in \ker \varphi$. Cela motive la définition suivante.

Définition 4.2.2. *Un* idéal *d'un anneau A est un sous-groupe I de A qui vérifie :*

$$\forall x\, y \in A, x \in I \Rightarrow xy \in I \text{ et } yx \in I.$$

On note $I \lhd A$.

Exemple 4.2.3. Dans tout anneau A, $\{0\}$ et A sont des idéaux. Si A est commutatif alors, pour tout $x \in A$, l'ensemble xA des multiples de x est un idéal de A. En particulier, pour tout $n \in \mathbb{Z}$, $n\mathbb{Z} \lhd \mathbb{Z}$.

On peut aussi définir la notion d'idéal à gauche ou à droite en n'imposant qu'une seule des deux conclusions dans la définition. On appelle alors idéaux bilatères les idéaux définis ci-dessus, mais nous n'aurons aucun usage de ces notions dans ce cours.

Remarque 4.2.4. Un idéal n'est presque jamais un sous-anneau. En effet un sous-anneau contient 1 donc un idéal qui est un sous-anneau contient tout A.

Un anneau quotient est en particulier un groupe quotient donc le théorème 3.3.6 assure que, en tant que groupe, un anneau quotient est nécessairement isomorphe à A/A_0 pour un sous-groupe A_0 de A. La question restante est la définition éventuelle d'un produit sur ce quotient.

Proposition 4.2.5. *Soit A un anneau.*

- *Pour tout morphisme d'anneaux $\varphi : A \to B$, $\ker \varphi \lhd A$.*
- *Soit A_0 un sous-groupe de A. Il existe une structure d'anneau qui étend la structure de groupe de A/A_0 et telle que la projection soit un morphisme d'anneaux si et seulement si A_0 est un idéal de A.*

Démonstration. On a déjà démontré le premier point pour motiver la définition d'idéal.

Soit A_0 un sous-groupe de A. Supposons qu'il existe une structure d'anneau sur A/A_0 qui étend sa structure de groupe et telle que la projection $\pi : A \to A/A_0$ soit un morphisme d'anneau. Le noyau π en tant que morphisme d'anneau est son noyau en tant que morphisme de groupe, c'est à dire A_0. Le premier point assure donc que A_0 est un idéal de A.

Réciproquement, supposons que A_0 est un idéal de A. Un produit sur A/A_0 rend multiplicative la projection $\pi : A \to A/A_0$ si et seulement si on peut compléter le diagramme

$$\begin{array}{ccc} A \times A & \xrightarrow{\quad m \quad} & A \\ {\scriptstyle \pi\times\pi}\downarrow & & \downarrow{\scriptstyle \pi} \\ A/A_0 \times A/A_0 & \xrightarrow[\bar{m}]{} & A/A_0 \end{array}$$

Le corollaire 1.0.13 et la définition de la relation d'équivalence associée à A_0 assurent qu'il suffit de vérifier :

$$\forall x_1\, x_2\, y_1\, y_2,\ x_2 - x_1 \in A_0 \text{ et } y_2 - y_1 \in A_0 \Rightarrow x_2y_2 - x_1y_1 \in A_0 \qquad (\star)$$

Soit x_1, x_2, y_1 et y_2 tels que $x_2 - x_1 \in A_0$ et $y_2 - y_1 \in A_0$. On a :

$$x_2y_2 - x_1y_2 = \underbrace{\underbrace{(x_2 - x_1)}_{\in A_0} y_2}_{\in A_0} + \underbrace{x_1 \underbrace{(y_2 - y_1)}_{\in A_0}}_{\in A_0} \in A_0$$

donc la condition $\star$ est vérifiée. Le lemme 2.0.9 assure que la multiplication obtenue sur A/A_0 est associative et admet pour neutre $\pi(1)$.

Le fait que 0 est absorbant et la distributivité découlent très facilement de la surjectivité de π et du fait que π est additive et multiplicative. □

Exemple 4.2.6. Dans $\mathbb{Z}$, le sous-groupe $n\mathbb{Z}$ est un idéal. L'anneau quotient correspondant est bien $\mathbb{Z}/n\mathbb{Z}$, noté aussi $\mathbb{Z}/n$ (ou même parfois $\mathbb{Z}_n$ mais cela créé un grave conflit de notations avec l'anneau des entiers n-adiques).

Dans $\mathbb{R}[X]$ le sous-groupe des multiples de $X^2 + 1$ est un idéal. Le quotient est appelé ensemble des nombres complexes et noté $\mathbb{C}$.

Théorème 4.2.7 (Propriété universelle des anneaux quotients)**.** *Soit A et B des anneaux, $\varphi\colon A \to B$ un morphisme et $I \lhd A$.*

- *Si $I \subset \ker\varphi$ alors il existe un unique morphisme d'anneaux qui fait commuter*

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ {\scriptstyle\pi}\downarrow & \nearrow_{\exists!\,\bar\varphi} & \\ A/I & & \end{array}$$

  On a alors $\ker\bar\varphi = \pi(\ker\varphi)$. Réciproquement, si $\bar\varphi$ existe alors $I \subset \ker\varphi$.

- *Le morphisme φ induit un isomorphisme de $A/\ker\varphi$ sur $\operatorname{im}\varphi$.*

- *Pour tout $J \lhd B$, si $\varphi(I) \subset J$ alors il existe un unique morphisme $\bar\varphi$ tel que*

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ A/I & \xrightarrow[\exists!\,\bar\varphi]{} & B/J \end{array}$$

  De plus $\bar\varphi$ injectif $\Leftrightarrow \varphi^{-1}(J) \subset I \Leftrightarrow \varphi^{-1}(J) = I$. Réciproquement, si $\bar\varphi$ existe alors $\varphi(I) \subset J$.

Démonstration. Pour le premier point, le théorème 3.3.9 assure déjà l'unicité et l'existence d'un morphisme de groupes $\bar\varphi$ et la description de son noyau. Le lemme 2.0.9 assure que $\bar\varphi$ est automatiquement un morphisme de monoïdes multiplicatifs.

Pour le deuxième point, on sait déjà par le corollaire 3.3.11 qu'on obtient un isomorphisme de groupes, et le lemme 4.1.10 assure que c'est un isomorphisme d'anneaux. Pour le troisième point, on applique le premier point à $\pi \circ \varphi$ (la discussion de l'injectivité provient directement du cas des groupes). □

Comme dans le cas des sous-groupes distingués, les idéaux sont stables par préimage et par image par un morphisme surjectif.

Lemme 4.2.8. *Soit A et B des anneaux et $\varphi\colon A \to B$ un morphisme.*

- *La préimage d'un idéal de B par φ est un idéal de A (en particulier on retrouve que* $\ker\varphi$ *est un idéal).*
- *Si φ est surjective alors l'image d'un idéal de A par φ est un idéal de B.*

Démonstration. Le lemme 3.1.8 assure que la préimage d'un idéal J de B par φ est un sous-groupe de A. Soit $x \in \varphi^{-1}(J)$ et $a \in A$. On a $\varphi(ax) = \varphi(a)\varphi(x)$ et $\varphi(x)$ est dans J donc $\varphi(ax)$ aussi. De même $\varphi(xa) = \varphi(x)\varphi(a)$ est dans $\varphi^{-1}(J)$.

Supposons maintenant que φ est surjective. Le même lemme assure que l'image d'un idéal I de A par φ est sous-groupe de B. Soit $x = \varphi(a)$ dans $\varphi(I)$ et b dans B. Par surjectivité de φ on obtient a' dans A tel que $b = \varphi(a')$ et on a $bx = \varphi(a')\varphi(a) = \varphi(a'a)$ et $a'a$ est dans I donc bx est dans $\varphi(I)$. De même $xb = \varphi(aa')$ est dans $\varphi(I)$. □

En particulier on peut comprendre les idéaux des anneaux quotients de façon très analogue à la description des sous-groupes des groupes quotients.

Lemme 4.2.9. *Soit A un anneau et $I \lhd A$. On note π la projection de A sur A/I. L'ensemble des idéaux de A qui contiennent I est en bijection croissante avec l'ensemble des idéaux de A/I par l'application $\Phi : J \mapsto \pi(J)$, de réciproque $K \mapsto \pi^{-1}(K)$.*

Démonstration. Le théorème 3.3.13 assure déjà que Φ est une bijection entre les sous-groupes de A qui contiennent I et les sous-groupes de A/I. De plus le lemme 4.2.8 assure que cette bijection envoie les idéaux sur les idéaux (dans les deux sens). □

On veut maintenant comprendre dans quel cas un anneau quotient est intègre et dans quel cas c'est un corps. La définition suivante introduit les conditions correspondantes sur un idéal. L'origine de la terminologie « idéal premier » sera expliquée dans la section suivante.

Définition 4.2.10. *Soit A un anneau* commutatif *et $I \lhd A$.*

- *On dit que I est* premier *si $I \neq A$ et,*

$$\forall a\, b,\ ab \in I \Rightarrow a \in I \text{ ou } b \in I.$$

  Autrement dit, I est premier si son complémentaire est un sous-monoïde multiplicatif (car $I \neq A \Leftrightarrow 1 \in I^c$).
- *On dit que I est* maximal *si $I \neq A$ et*

$$\forall J \lhd A,\ J \neq A \text{ et } I \subset J \Rightarrow J = I.$$

  Autrement dit, I est maximal s'il est maximal pour l'inclusion parmi les idéaux propres de A.

Pour lier ces définitions au propriétés des anneaux quotient, on utilisera l'observation suivante (qui est aussi utile ailleurs).

Lemme 4.2.11. *Un anneau commutatif A est un corps si et seulement si il possède exactement deux idéaux : 0 et A.*

Démonstration. On commence par noter que, dans les idéaux de A, $0 \neq A$ si et seulement si A n'est pas trivial.

Supposons que A est un corps. Comme A n'est pas trivial, il possède au moins deux idéaux : 0 et A. Soit $I \lhd A$ différent de 0. Soit x un élément non nul de I. Par hypothèse x est inversible donc $1 = x^{-1}x$ est dans I donc $I = A$.

Réciproquement, supposons que A a exactement deux idéaux, 0 et A. En particulier A n'est pas trivial. Soit x dans A non nul. L'idéal xA n'est pas nul car il contient x donc c'est A. En particulier $1 \in xA$ donc il existe x' tel que $xx' = 1$ et x est inversible. □

Voici maintenant le lien promis entre les propriétés de I et celles de A/I.

Lemme 4.2.12. *Soit A un anneau commutatif et $I \lhd A$.*

- *I est premier $\Leftrightarrow$ A/I est intègre*
- *I est maximal $\Leftrightarrow$ A/I est un corps*

En particulier que les idéaux maximaux sont premiers (ce qu'on peut aussi vérifier directement).

Démonstration. On commence par observer que $I = A \Leftrightarrow A/I = 0$ donc les conditions de non-trivialité intervenant dans les définition d'idéal premier et d'anneau intègre sont équivalentes. Comme $\pi : A \to A/I$ est un morphisme surjectif et $\ker \pi = I$, on a

$$\begin{aligned} A/I \text{ intègre} &\Leftrightarrow \forall a\, b \in A,\ \pi(a)\pi(b) = 0 \Rightarrow \pi(a) = 0 \text{ ou } \pi(b) = 0 \\ &\Leftrightarrow \forall a\, b \in A,\ \pi(ab) = 0 \Rightarrow \pi(a) = 0 \text{ ou } \pi(b) = 0 \\ &\Leftrightarrow \forall a\, b \in A,\ ab \in I \Rightarrow a \in I \text{ ou } b \in I. \end{aligned}$$

Pour le second point, on observe encore que $I \neq A$ si et seulement A/I n'est pas trivial. Par ailleurs le lemme 4.2.11 assure que A/I est un corps si et seulement il possède exactement deux idéaux et le lemme 4.2.9 décrit les idéaux de A/I donc on a :

$$\begin{aligned} A/I \text{ corps} &\Leftrightarrow \forall J \lhd A/I,\ J = 0 \text{ ou } J = A/I \\ &\Leftrightarrow \forall K \lhd A,\ I \subset K \Rightarrow \pi(K) = 0 \text{ ou } \pi(K) = A/I \\ &\Leftrightarrow \forall K \lhd A,\ I \subset K \Rightarrow K = I \text{ ou } K = A \end{aligned}$$

□

Soit A un anneau commutatif. Si A n'est pas intègre, il n'y a aucun espoir de l'injecter dans un corps (ni même dans un anneau intègre). Par contre on peut essayer de lui trouver un quotient qui est un corps (par exemple $\mathbb{Z}/3\mathbb{Z}$ est un quotient de $\mathbb{Z}/6\mathbb{Z}$). Vu le lemme 4.2.12 l'existence d'un tel quotient est équivalente à l'existence d'un idéal maximal dans A.

On admet l'énoncé suivant qui est un théorème ou un axiome selon les fondements choisis.

Lemme 4.2.13 (Lemme de Zorn)**.** *Soit $(X, \leqslant)$ un ensemble (partiellement) ordonné. Si toute partie totalement ordonnée de X admet un majorant alors X admet un élément maximal.*

On rappelle qu'un majorant d'une partie A d'un ensemble X ordonné est un élément de X qui est plus grand que tous les éléments de A, il n'est pas nécessairement dans A. Un élément x_0 de X est maximal s'il n'y a pas d'élément de X strictement plus grand que x_0, il s'agit d'une condition plus faible que de demander que x_0 soit plus grand que tous les éléments de X.

Proposition 4.2.14 (Théorème de Krull)**.** *Soit A un anneau commutatif. Tout idéal $I \lhd A$ propre (c'est à dire que $I \neq A$) est contenu dans un idéal maximal. En particulier si A est non trivial alors il possède un idéal maximal et donc un quotient qui est un corps.*

Démonstration. La seconde partie découle du lemme 4.2.12 et de la première partie car dans un anneau non trivial l'idéal nul est propre.

Soit $I \lhd A$ propre. Par définition, un idéal maximal contenant I est un élément maximal de l'ensemble

$$\mathcal{J} = \{J \lhd A \mid I \subset J \text{ et } J \neq A\}.$$

Il suffit donc de montrer que cet ensemble vérifie l'hypothèse du lemme de Zorn. Soit $(J_l)_{l\in L}$ une partie totalement ordonnée de $\mathcal{J}$ (pour l'inclusion). On pose $J = \bigcup_{l\in L} J_l$. Il s'agit d'une partie de A qui contient tous les J_l. Le point clef est que J est dans $\mathcal{J}$.

Soit x dans J et a dans A. Par définition de J on obtient $l \in L$ tel que $x \in J_l$. Comme J_l est un idéal, ax est dans J_l donc dans J.

Soit x et y dans J. Soit l et l' dans L tels que $x \in J_l$ et $y \in J_{l'}$. Comme l'ensemble des J_l est totalement ordonné, $J_l \subset J_{l'}$ ou $J_{l'} \subset J_l$. Dans le deux cas on trouve un membre de la famille qui contient à la fois x et y donc contient leur somme.

Enfin J ne contient pas 1 car aucun des J_l ne le contient, donc $J \neq A$. □

Le lemme de Zorn est inévitable dans la démonstration ci-dessus : on peut démontrer que le théorème de Krull implique le lemme de Zorn (sous des hypothèses raisonnables sur les autres fondements).

4.3 Opérations sur les idéaux et arithmétique

4.3.1 Divisibilité

On commence par rappeler la notion de divisibilité et les notions associées dans un anneau commutatif quelconque mais nous verrons qu'en général ces notions se comportent mal et qu'il existe des notions correspondantes au niveau des idéaux qui se comportent mieux.

Dans presque toute cette section, A sera un anneau commutatif *intègre*. L'intégrité n'intervient pas dans la définition de la divisibilité mais elle est nécessaire pour obtenir les propriétés attendues.

Définition 4.3.1. *Soit a et b deux éléments d'un anneau commutatif A. On dit que a* divise *b, et on note $a \mid b$ s'il existe c dans A tel que $b = ac$. On dit aussi que a est un* diviseur *de b et que b est un* multiple *de a.*

Remarque 4.3.2. Soit a et b deux éléments d'un anneau commutatif A.

- $0 \mid b \Leftrightarrow b = 0$

- $a \mid 1 \Leftrightarrow a \in A^\times$

L'intégrité intervient dès le lemme suivant.

Lemme 4.3.3. *Soit a et b deux éléments d'un anneau commutatif intègre A.*

$$(a \mid b \text{ et } b \mid a) \Leftrightarrow \exists u \in A^\times, a = ub$$

Lorsque ces conditions sont vérifiées, on dit que a et b sont associés.

Démonstration. Supposons que $a \mid b$ et $b \mid a$. On obtient ainsi c et d tels que $b = ac$ et $a = bd$. On en déduit $b = bcd$, donc $b(1 - cd) = 0$ et, par intégrité de A, $b = 0$ ou $1 - cd = 0$. Si $b = 0$ alors la condition $b \mid a$ assure que $a = 0$ et on peut choisir $u = 1$. Sinon on obtient $cd = 1$ donc d est une unité convenable.

Réciproquement, supposons que $a = ub$ pour $u \in A^\times$. On a directement $b \mid a$ et aussi $a \mid b$ car $b = u^{-1}a$. □

Remarque 4.3.4. Le lemme précédent montre que la relation de divisibilité est très rarement une relation d'ordre. En général ce n'est qu'une relation de préordre (c'est à dire une relation réflexive et transitive). Cela n'empêche pas de parler de maximum ou de minimum d'un ensemble dans la définition suivante. C'est l'occasion de mentionner qu'on peut retrouver la notion de divisibilité dans $\mathbb{N}$ si on autorise les semi-anneaux dans la définition (et dans $\mathbb{N}$ on a bien une relation d'ordre car 1 est la seule unité). Comme exemple d'anneau dans lesquels la relation de divisibilité est une relation d'ordre, on peut mentionner $(\mathbb{Z}/2\mathbb{Z})^n$ ou $\mathbb{Z}/2\mathbb{Z}[X]$.

Définition 4.3.5. *Soit A un anneau commutatif. Soit a et b deux éléments de A.*

- *On dit que a et b sont* premiers entre eux *si tous leurs diviseurs communs sont inversibles : $\forall d, [d \mid a \text{ et } d \mid b] \Rightarrow d \in A^\times$.*
- *On dit qu'un élément de d de A est un* pgcd *(plus grand commun diviseur) de a et de b s'il est un diviseur de a et b et si tout diviseur de a et b divise d (autrement dit d est un maximum de l'ensemble des diviseurs communs à a et b, pour la relation de divisibilité). En particulier a et b sont premier entre eux si et seulement si 1 est un pgcd de a et de b.*
- *On dit qu'un élément de m de A est un* ppcm *(plus petit commun multiple) de a et de b s'il est un multiple non nul de a et b et si tout multiple de a et b est multiple de m (autrement dit m est un minimum de l'ensemble des multiples non nuls communs à a et b, pour la relation de divisibilité).*

Lemme 4.3.6. *Soit A un anneau commutatif intègre et soit a et b des éléments de A. Tous les pgcd de a et b sont associés. Tous les ppcm de a et b sont associés.*

Démonstration. Supposons que d et d' sont des pgcd de a et de b. Comme d est un diviseur commun à a et b et que d' est un pgcd de a et b, on apprend que $d \mid d'$. De même on obtient $d' \mid d$ et, par le lemme 4.3.3, d et d' sont associés. La démonstration dans le cas des ppcm est complètement analogue. □

Jusqu'ici tout se passe comme dans $\mathbb{Z}$ ou $\mathbb{K}[X]$ (en supposant qu'on travaille dans un anneau commutatif intègre). Mais les choses tournent mal dès qu'on veut généraliser le

lemme de Bézout qui affirme que deux éléments a et b dans $\mathbb{Z}$ ou $\mathbb{K}[X]$ sont premiers entre eux si et seulement si il existe u et v tel que $ua + vb = 1$. Par exemple dans l'anneau $\mathbb{Z}[X]$, les éléments 2 et X sont premiers entre eux (il suffit de faire la liste complète des diviseurs de 2 et de X) mais il n'existe pas de polynômes U et V tels que $2U + XV = 1$. En effet on peut étudier les solutions de cette équation dans $\mathbb{Q}[X]^2$ par la méthode habituelle et voir qu'aucune d'entre elles n'est dans $\mathbb{Z}[X]^2$. En restant dans le monde des corps mais en ajoutant des indéterminées, on peut aussi utiliser l'exemple de X et Y dans $\mathbb{K}[X, Y]$ (les polynômes à plusieurs indéterminées seront étudiés dans un chapitre ultérieur). La solution à ce problème consiste à travailler au niveau des idéaux en utilisant les constructions des sections suivantes.

4.3.2 Idéaux engendrés et idéaux principaux

Dans cette section, sauf mention explicite du contraire, on retourne au cas général des anneaux non nécessairement commutatifs. On commence par construire l'idéal engendré par une partie. Comme d'habitude, l'énoncé crucial est le suivant.

Lemme 4.3.7. *Une intersection d'idéaux est un idéal.*

Démonstration. Soit $\mathcal{J}$ une famille d'idéaux d'un anneau A. On sait déjà que l'intersection I des éléments de $\mathcal{J}$ est un sous-groupe de A. Soit $x \in I$ et $a \in A$. Pour tout $J \in \mathcal{J}$, x est dans J donc ax aussi. Ainsi ax est dans $\mathcal{J}$. De même xa est dans $\mathcal{J}$. □

On en déduit comme d'habitude une notion d'idéal engendré par une partie.

Définition 4.3.8. *L'idéal engendré par une partie S d'un anneau A est l'intersection de tous les idéaux de A contenant S :*

$$(S) = \bigcap_{I \lhd A, S \subset I} I.$$

Le lemme suivant rassemble les propriétés formelles des idéaux engendrés.

Lemme 4.3.9. *Soit S une partie d'un anneau A.*

- *(S) est un idéal de A qui contient S.*
- *Pour tout $I \lhd A$, $(S) \subset I \Leftrightarrow S \subset I$ (ainsi (S) est le plus petit idéal de A qui contient S). Cette propriété universelle caractérise (S).*
- *L'application $(\cdot)$ est croissante : $\forall S\ S', S \subset S' \Rightarrow (S) \subset (S')$.*
- *Pour tout morphisme $f : A \to A'$ surjectif, $(f(S)) = f((S))$.*

Démonstration. C'est exactement la même démonstration que pour le lemme 2.0.16 car cette dernière a été rédigée de façon suffisamment abstraite. Dans le dernier point, l'hypothèse de surjectivité sert à assurer que $f((S))$ est un idéal. □

Lemme 4.3.10. *Soit A un anneau et S un partie de A. L'idéal engendré par S est l'ensemble des sommes d'éléments de la forme axb avec x dans S, a et b dans A (en incluant la somme vide qui donne 0). Si A est commutatif, il suffit de sommer des éléments de la forme ax avec a dans A et x dans S.*

Démonstration. L'ensemble décrit vérifie bien la propriété universelle du lemme précédent. □

Cette construction permet en particulier de transformer les éléments d'un anneau en idéaux.

Définition 4.3.11. *Un idéal est* principal *s'il est engendré par un singleton. Un anneau est* principal *s'il est intègre et si tous ses idéaux sont principaux.*

Comme dans le cas des groupes cycliques, on dira plutôt qu'un idéal principal est engendré par « un élément » et on notera $I = (x)$ plutôt que $I = (\{x\})$. En pratique dans ce cours on ne parlera d'idéaux principaux que dans un anneau commutatif. Dans cas le lemme 4.3.10 assure que $(x) = xA$, l'ensemble des produits de x et d'un élément de A. On retrouve le cas familier des idéaux $n\mathbb{Z}$ de $\mathbb{Z}$.

Exemple 4.3.12. Dans un anneau A, l'idéal (1) engendré par 1 est A tout entier. On notera donc 1 l'idéal constitué de tout A (on verra plus loin qu'il s'agit de l'élément neutre d'une opération de multiplication des idéaux). On rappelle que cet idéal est le seul qui contienne 1 (le 1 de A).

Il est bien connu que $\mathbb{Z}$ et $\mathbb{K}[X]$ sont des anneaux principaux. La définition suivante permet de faire une seule démonstration pour ces deux cas (et guère plus que cela, malgré son aspect très général).

Définition 4.3.13. *Un anneau* euclidien *est un anneau commutatif intègre A tel qu'il existe une fonction $v\colon A \setminus \{0\} \to \mathbb{N}$ telle que :*

$$\forall a\, b \in A, b \neq 0 \Rightarrow \exists q\, r \in A, a = bq + r \text{ et } r = 0 \text{ ou } v(r) < v(b).$$

Une telle fonction v est appelée stathme pour A et une égalité de la forme $a = bq + r$ avec $r = 0$ ou $v(r) < v(b)$ est appelée division euclidienne de a par b. On ne demande aucune unicité de la division euclidienne.

Exemple 4.3.14. L'anneau $\mathbb{Z}$ est euclidien, avec la valeur absolue comme stathme. Pour tout corps $\mathbb{K}$, l'anneau $\mathbb{K}[X]$ est euclidien, avec le degré comme stathme. Les autres exemples d'anneaux euclidiens sont très sporadiques.

Lemme 4.3.15. *Tout anneau euclidien est principal.*

Démonstration. Soit A un anneau euclidien et v un stathme pour A. Par définition des anneaux euclidiens, A est intègre. Montrons que tous ses idéaux sont principaux. Soit I un idéal de A. Si $I = 0$ alors I est principal. Sinon on fixe $b \in I \setminus \{0\}$ qui minimise v. Montrons que $I = (b)$. Soit a dans I et $a = bq + r$ une division euclidienne de a par b. Supposons pas l'absurde $r \neq 0$. On a alors $v(r) < v(b)$. Or $r = a - bq$ donc r est dans I et la minimalité de $v(b)$ est contredite. Ainsi $a = bq$ et donc $a \in (b)$. □

Pour un anneau commutatif A qui n'est pas un corps, il n'y a aucune raison que $A[X]$ soit euclidien. Par exemple $\mathbb{Z}[X]$ n'est pas principal. Cependant on a une division euclidienne par les polynômes dont le coefficient dominant est inversible. Cela suffit déjà à rendre des services, comme dans le lemme suivant.

Lemme 4.3.16. *Soit A un anneau commutatif et $P \in A[X]$ à coefficient dominant inversible. La restriction de $\pi : A[X] \to A[X]/(P)$ à l'ensemble des polynômes de degré strictement plus petit que celui de P est une bijection.*

Démonstration. Soit $x \in A[X]/(P)$. On veut montrer qu'il existe un unique polynôme R tel que $\deg(R) < \deg(P)$ et $\pi(R) = x$. Soit $S \in A[X]$ tel que $\pi(S) = x$. Puisque le coefficient dominant de P est inversible, on a une division euclidienne $S = PQ + R$ avec $R = 0$ ou $\deg(R) < \deg(Q)$. Comme $x = \pi(S) = \pi(PQ) + \pi(R) = \pi(R)$, R convient. Montrons maintenant l'unicité. Supposons que R et R' conviennent. On a alors $\deg(R - R') < \deg(P)$ et $\pi(R - R') = 0$ donc $P \mid R - R'$. On obtient ainsi Q tel que $R - R' = PQ$. Comme le coefficient dominant de P est inversible, soit $Q = 0$ soit $\deg(PQ) \geq \deg(P)$. On obtient donc bien $Q = 0$ et donc $R = R'$. □

Pour conclure ces considérations d'arithmétiques des polynômes, on rappelle le lien entre racines des polynômes et divisibilité.

Lemme 4.3.17. *Soit A un anneau commutatif et $P \in A[X]$. Un élément a de A est racine de P si et seulement si $(X - a)$ divise P. Si A est intègre et P n'est pas nul alors P admet au plus* $\deg(P)$ *racines.*

Démonstration. Comme $X - a$ est unitaire, on a une division euclidienne $P = (X - a)Q + R$ avec $R = 0$ ou $\deg(R) < \deg(X - a)$ donc $\deg(R) = 0$. En évaluant en a on obtient $P(a) = R(a)$ donc $P(a) = 0 \Leftrightarrow R = 0 \Leftrightarrow (X - a) \mid P$.

Supposons maintenant que A est intègre. Cette hypothèse assure que, pour tous polynômes Q et R non nuls, $\deg(QR) = \deg(Q) + \deg(R)$ (car le coefficient dominant de QR est le produit des coefficients dominants de Q et de R). En particulier les polynômes $X - a$ sont premiers entre eux deux à deux. Soit P un polynôme non nul. Supposons que P possède des racines a_i pour $1 \leqslant i \leqslant N$. Comme les polynômes $(X - a_i)$ sont premiers entre eux deux à deux, $\prod_i (X - a_i) \mid P$. On obtient ainsi $Q \in A[X]$ tel que $P = Q \prod_i (X - a_i)$. Comme A est intègre et que P est non nul, $\deg(P) = \deg(Q) + N$ et donc $N \leqslant \deg(P)$. Remarque : la section suivante rendra plus confortable cet genre d'argument en plongeant tout anneau intègre dans un corps. □

4.3.3 Idéaux et divisibilité

La notion d'idéal principal permet déjà de relier les idéaux à l'arithmétique via l'observation suivante.

Lemme 4.3.18. *Soit A un anneau commutatif. Pour tous les éléments a et b de A, on a les équivalences $a \mid b \Leftrightarrow b \in (a) \Leftrightarrow (b) \subset (a)$.*

Démonstration. La première équivalence provient de la description concrète de l'idéal engendré (lemme 4.3.10). La deuxième provient de la propriété universelle de l'idéal engendré (lemme 4.3.9). □

Pour comprendre à quoi correspondent les pgcd du côté des idéaux, on introduit la construction suivante.

Lemme 4.3.19. *Soit I et J deux idéaux d'un anneau A. L'ensemble $I + J$ des sommes d'un élément de I et d'un élément de J est un idéal appelé* somme *de I et J. C'est l'idéal engendré par $I \cup J$. L'idéal nul est neutre pour cette opération.*

Démonstration. Pour vérifier que $I+J$ est un sous-groupe, on utilise le lemme 3.1.8. L'ensemble $I+J$ n'est pas vide car il contient zéro. Soit $i+j$ et $i'+j'$ des éléments de $I+J$. On a $(i'+j')-(i+j)=(i'-i)+(j'-i)$ qui est dans $I+J$. Ainsi $I+J$ est bien un sous-groupe de A. Soit $i+j\in I+J$ et a dans A. On a $a(i+j)=ai+aj$ qui est bien dans $I+J$ car I et J sont des idéaux. De même $(i+j)a$ est dans $I+J$. Ainsi $I+J$ est bien un idéal de A.

Il contient $I\cup J$ car I et J contiennent 0. Donc $I+J$ contient $(I\cup J)$. Montrons l'autre inclusion. Soit x dans $(I\cup J)$. Le lemme 4.3.10 fournit n et des familles a, b et y telles que $x=\sum_{i=1}^{n} a_i y_i b_i$ avec $y_i\in I\cup J$ pour tout i. En mettant d'un côté la somme des $a_i y_i b_i$ tels que y_i est dans I et de l'autre celle de ceux pour lesquels y_i est dans J, on voit que x est dans $I+J$. □

Le lemme suivant relie l'addition des idéaux avec les pgcd.

Lemme 4.3.20. *Soit a, b, c, et d des éléments d'un anneau commutatif A.*

- $[d \mid a \text{ et } d \mid b] \Leftrightarrow (a)+(b)\subset(d)$
- *Si $(d)=(a)+(b)$ alors d est un pgcd de a et de b. En particulier, si $(a)+(b)=1$ alors a et b sont premiers entre eux.*
- *Si A est principal alors a et b ont un pgcd, et d est un pgcd de a et de b si et seulement si $(d)=(a)+(b)$. En particulier, sous cette hypothèse que A est principal, a et b sont premiers entre eux si et seulement si il existe u et v dans A tels que $au+bv=1$ (ce résultat est souvent appelé théorème de Bézout).*

Démonstration. Pour le premier point, on utilise la propriété universelle de la somme d'idéaux et le fait que (d) est un idéal pour obtenir

$$\begin{aligned}(a)+(b)\subset(d) &\Leftrightarrow (a)\cup(b)\subset(d)\\ &\Leftrightarrow (a)\subset(d) \text{ et } (b)\subset(d)\\ &\Leftrightarrow d \mid a \text{ et } d\mid b\end{aligned}$$

Supposons maintenant que $(a)+(b)=(d)$. En particulier $(a)+(b)\subset(d)$ donc le point précédent assure que d est un diviseur commun à a et b. Soit d' un autre diviseur commun à a et b. Par l'autre implication du point précédent, $(a)+(b)\subset(d')$, c'est à dire $(d)\subset(d')$ donc $d'\mid d$.

Supposons que A est principal. Comme tous les idéaux de A sont principaux, on obtient un générateur d de l'idéal $(a)+(b)$. Par le point précédent, d est un pgcd de a et b. La réciproque est directement fournie par le point précédent.

Toujours sous l'hypothèse que A est principal, supposons que a et b sont premiers entre eux. On a alors $(a)+(b)=1$ par le point précédent. Donc $(a)+(b)$ contient 1, c'est à dire qu'il existe u et v tels que $au+bv=1$. Réciproquement si un tel couple (u,v) existe alors $1\in(a)+(b)$ donc $(a)+(b)=1$ et a et b sont premiers entre aux par le troisième point (sans utiliser que A est principal). □

L'énoncé précédent suggère une définition de coprimalité pour les idéaux.

Définition 4.3.21. *On dit que deux idéaux I et J d'un anneau A sont* premiers entre eux *si $I+J=1$.*

Le lemme suivant relie les opérations sur les idéaux avec les ppcm (mais il est nettement moins utile que la version concernant les pgcd).

Lemme 4.3.22. *Soit a, b, c et m des éléments d'un anneau commutatif A.*

- $(a \mid m \text{ et } b \mid m) \Leftrightarrow (m) \subset (a) \cap (b)$
- *Si $(m) = (a) \cap (b)$ alors m est un ppcm de a et de b.*
- *Si A est principal alors a et b ont un ppcm, et m est un ppcm de a et de b si et seulement si $(m) = (a) \cap (b)$.*

Démonstration. Le premier point est clair car $(a \mid m \text{ et } b \mid m) \Leftrightarrow (m) \subset (a)$ et $(m) \subset (b)$. Les points suivants en découlent comme dans le cas du pgcd. □

4.3.4 Opérations sur les idéaux et théorème chinois

Notre objectif suivant est de généraliser le théorème des restes chinois. La version élémentaire de ce théorème fait intervenir des produits d'éléments premiers entre eux. Nous venons de voir dans la définition 4.3.21 ce que devient la condition de coprimalité, il nous reste à généraliser les produits.

L'analogue du lemme 4.3.19 pour les produits est faux. Si a et a' sont dans I et b et b' sont dans J, il n'y a aucune raison que $ab + a'b'$ puisse s'écrire sous la forme $a''b''$ avec a'' dans I et b'' dans J. On utilise donc la définition un peu piégeuse suivante.

Définition 4.3.23. *Soit I et J des idéaux d'un anneau A. L'idéal* produit IJ *est l'idéal* engendré *par les produits d'un élément de I et d'un élément de J.*

Remarque 4.3.24. L'ensemble des produit d'un élément de I et d'un élément de J est déjà absorbant pour la multiplication donc les éléments de IJ s'écrivent simplement comme sommes de produits d'un élément de I et d'un élément de J.

Lemme 4.3.25. *Si A est commutatif, l'application de A dans l'ensemble de ses idéaux qui envoie x sur l'idéal principal (x) est multiplicative : pour tout x et y dans A, $(xy) = (x)(y)$.*

Démonstration. Soit x et y dans A. Comme $xy \in (x)(y)$, $(xy) \subset (x)(y)$. Pour l'autre inclusion, on utilise que A est commutatif donc $(x) = \{ax \ ; \ a \in A\}$, $(y) = \{bx \ ; \ b \in A\}$ donc $(x)(y)$ est formé des sommes $\sum_i a_i x b_i y$. Or une telle somme peut se réécrire $(\sum_i a_i b_i)xy$ donc appartient à (xy). □

Dans le cas d'un anneau commutatif, le lemme ci-dessus explique pourquoi Dedekind a utilisé le mot idéal : d'un point de vue multiplicatif, l'application $x \mapsto (x)$ permet de voir les idéaux de A comme des éléments supplémentaires de A (et le mot « imaginaire » était déjà pris dans un contexte analogue). Dans cette image il faut se méfier un peu du fait que $x \mapsto (x)$ n'est pas injectif. La motivation de Dedekind était de retrouver dans certains cas une unique décomposition en produit de facteurs premiers (comme dans $\mathbb{Z}$).

Il faut tout de fois prendre garde à ne pas pousser cette analogie en direction de l'addition. Nous avons vu plus haut que $(x) + (y)$ n'a rien à voir avec $(x + y)$. Par contre l'addition et la multiplication des idéaux interagissent comme on l'imagine. La proposition suivante détaille cela et d'autre propriétés utiles de ces opérations.

Proposition 4.3.26. *Soit A un anneau.*

- *L'ensemble de idéaux de A, muni des opérations d'addition et de multiplication, de l'idéal nul et de l'idéal $A = (1)$ est un semi-anneau. On notera donc 1 l'idéal A.*
- *Pour tout morphisme d'anneaux $\varphi : A \to B$, l'application induite des idéaux de A dans les idéaux de B envoie 0 sur 0 et est compatible avec l'addition et la multiplication (par contre elle n'envoie 1 sur 1 que si φ est surjective).*
- *Pour tous idéaux I et J, $(I \cap J)^2 \subset IJ \subset I \cap J$.*
- *Si A est commutatif alors, pour tous idéaux I et J, $IJ = JI$ et $I + J = 1 \Rightarrow IJ = I \cap J$.*
- *Soit $I \lhd A$. On note π la projection de A sur A/I. La bijection $\Phi : J \mapsto \pi(J)$ des idéaux contenant I vers les idéaux de A/I est compatible avec l'addition, envoie I sur 0 et 1 sur 1. Si J et K sont des idéaux tels que $I \subset JK$ alors $\Phi(JK) = \Phi(J)\Phi(K)$.*

Remarque 4.3.27. Il est utile d'avoir en tête un exemple de produit d'idéaux qui n'est pas égal à leur intersection. Vu le critère donné, on essaie $I = 2\mathbb{Z}$ et $J = 4\mathbb{Z}$ dans $A = \mathbb{Z}$ et on obtient $IJ = 8\mathbb{Z}$ tandis que $I \cap J = 4\mathbb{Z}$ puisque $J \subset I$.

Démonstration. Soit I, J et K des idéaux de A. On a $I + J = (I \cup J) = (J \cup I) = J + I$ en utilisant le lemme précédent. De même l'associativité de l'addition des idéaux découle de celle de la réunion mais il faut être un peu plus prudent. Soit I, J et K des idéaux de A. On veut montrer que $(I + J) + K = I + (J + K)$. Vu l'associativité de la réunion, il suffit de montrer que $I + (J + K) = \langle I \cup (J \cup K)\rangle$ et $(I + J) + K = \langle (I \cup J) \cup K\rangle$. Soit P un idéal de A.

$$\begin{aligned} I \cup (J \cup K) \subset P &\Leftrightarrow I \subset P \text{ et } J \cup K \subset P \\ &\Leftrightarrow I \subset P \text{ et } J + K \subset P \\ &\Leftrightarrow I \cup (J + K) \subset P \\ &\Leftrightarrow I + (J + K) \subset P \end{aligned}$$

Donc l'idéal $I + (J + K)$ vérifie la propriété universelle qui caractérise $\langle I \cup (J \cup K)\rangle$. Le cas de $(I + J) + K$ fonctionne exactement de la même façon.

De plus $I + 0 = (I \cup 0) = (I) = I$.

Notons $\star$ l'ensemble des produits d'éléments de I et de J, de sorte que $IJ = (I \star J)$. L'opération $\star$ est associative car la multiplication dans A est associative. Elle est commutative si A est commutatif. On a $I1 = (I \star A) = (I) = I$ et de même $1I = I$. L'associativité de la multiplication découle de celle de $\star$ comme l'associativité de la somme découle de celle de la réunion.

Montrons la distributivité. Par la remarque ci-dessus, les éléments de $I(J+K)$ s'écrivent comme $\sum_\lambda i_\lambda(j_\lambda + k_\lambda)$ avec, pour tout λ, $i_\lambda \in I$, $j_\lambda \in J$ et $k_\lambda \in K$. En développant distribuant les multiplication et en utilisant la commutativité de l'addition on voit que ces éléments sont dans $IJ + IK$. Montrons l'inclusion réciproque. Comme $J \subset J + K$, on obtient $I \star J \subset I \star (J + K)$ donc $IJ \subset I(J + K)$. De même $IK \subset I(J + K)$ donc $IJ \cup IK \subset I(J + K)$ donc $IJ + IK \subset I(J + K)$ et finalement $I(J + K) = IJ + IK$. On montre de même que $(J + K)I = JI + KI$.

Pour le second point, soit $\varphi : A \to B$ un morphisme d'anneaux. Le fait que φ envoie l'idéal nul sur l'idéal nul est clair. Montrons l'additivité. Soit I et J des idéaux dans A. On a $\varphi(I+J) = \varphi((I \cup J)) = (\varphi(I \cup J)) = (\varphi(I) \cup \varphi(J)) = \varphi(I) + \varphi(J)$. Comme φ est multiplicative, $\varphi(I \star J) = \varphi(I) \star \varphi(J)$ et on en déduit que $\varphi(IJ) = \varphi(I)\varphi(J)$ comme pour l'addition.

Pour le troisième point, il est clair que $I \star J \subset I \cap J$ donc $IJ \subset I \cap J$. Par ailleurs les éléments de $(I \cap J)^2$ s'écrivent comme sommes de $a_\lambda b_\lambda$ avec $a_\lambda \in I \cap J$ donc $a_\lambda \in I$ et $b_\lambda \in I \cap J$ donc $b_\lambda \in J$ donc ces éléments sont dans IJ.

Montrons le quatrième point. On suppose A commutatif. L'opération $\star$ est alors commutative donc le produit des idéaux aussi. Soit I et J des idéaux de A. Supposons $I+J = 1$. On sait déjà que $IJ \subset I \cap J$. On a $I \cap J = (I \cap J)(I+J) = (I \cap J)I + (I \cap J)J \subset JI + IJ = IJ$ (en utilisant la commutativité dans la dernière égalité).

Enfin fixons un idéal I de A, notons π la projection de A sur A/I et $\Phi : J \mapsto \pi(J)$. Le fait que Φ envoie I sur 0 est clair, $\Phi(1) = 1$ car π est surjectif, et le reste provient directement des points précédents (la condition $I \subset JK$ ne sert qu'à assurer que $\Phi(JK)$, $\Phi(J)$ et $\Phi(K)$ sont tous trois bien définis). □

Remarque 4.3.28. Dans l'avant-dernier point de la proposition précédente, la condition $I + J = 1$ est suffisante mais pas nécessaire. Par exemple, dans $A = \mathbb{Z}/6\mathbb{Z}$, on considère $I = J = (2)$. On a $IJ = (4) = (2) = I \cap J$ car les multiples de 4 sont, dans l'ordre, 0, 4, 2, 0, 4 et 2. Mais $I + J = I = (2) \neq 1$.

Muni de toutes ces propriétés, il est facile d'étendre le théorème des restes Chinois comme promis (sans même une hypothèse de commutativité). La démonstration utilisera le lemme suivant.

Lemme 4.3.29. *Soit I, I_1, …, I_n des idéaux d'un anneau A. Si I est premier avec chacun des I_i alors il est premier avec leur intersection.*

Démonstration. On procède par récurrence sur $n \geqslant 1$. Le cas $n = 1$ est tautologique. Supposons le théorème démontré jusqu'à n. Soit I, I_1, …, I_{n+1} des idéaux d'un anneau A tels que I est premier avec chacun des I_i. On pose $J = I_1 \cap \cdots \cap I_n$, de sorte que l'intersection des I_i est $J \cap I_{n+1}$. Par hypothèse de récurrence, $I + J = 1$. Par ailleurs on a supposé $I + I_{n+1} = 1$. On a donc, en utilisant les propriétés garanties par la proposition 4.3.26,

$$1 = I + J = I + J(I + I_{n+1}) = (1 + J)I + JI_{n+1} \subset I + I_{n+1} \cap J$$

donc $I + I_{n+1} \cap J = 1$ (car $1 = A$). □

Théorème 4.3.30 (Théorème des restes chinois)**.** *Soit A un anneau et I_1, …, I_n des idéaux de A. On note π_i la projection de A sur A/I_i. Si les idéaux I_i sont premiers entre eux deux à deux alors le morphisme $\pi_1 \times \cdots \pi_n : A \to \prod_i A/I_i$ induit un isomorphisme entre $A/\bigcap_i I_i$ et $\prod_i A/I_i$. Si de plus A est commutatif alors $\bigcap_i I_i = \prod_i I_i$.*

Démonstration. On applique le second point du théorème 4.2.7. On a $\ker(\pi_1 \times \cdots \pi_n) = \bigcap \ker \pi_i = \bigcap I_i$ donc tout le travail consiste à montrer que $\pi_1 \times \cdots \pi_n$ est surjectif. Soit x_1, …x_n dans A. On veut x dans A tel que $\forall i, \pi_i(x) = \pi_i(x_i)$.

Construisons une famille d'éléments $e_i \in A$ tels que, pour tout i, $\pi_1(e_i) = 1$ et, pour tout $j \neq i$, $\pi_j(e_i) = 0$. Soit i. Le lemme 4.3.29 assure que I_i est premier avec l'intersection

J_i de tous les autres I_j. On obtient donc des éléments $u_i \in I_i$ et $e_i \in J_i$ tels que $u_i + e_i = 1$. On a alors $\pi_i(e_i) = \pi_i(1 - u_i) = \pi_i(1) - \pi_i(u_i) = 1$ car $u_i \in I_i$ tandis que, pour tout $j \neq i$, $\pi_j(e_i) = 0$ car $e_i \in J_i$ et $J_i \subset I_j$.

On retourne maintenant au problème de départ et on pose $x = \sum_i x_i e_i$ qui convient d'après les propriétés des e_i.

Supposons maintenant que A est commutatif. Montrons par récurrence sur n que $\bigcap_i I_i = \prod_i I_i$. Pour $n = 1$ il n'y a rien à démontrer. Supposons le résultat connu jusqu'à n et considérons une famille de $n + 1$ idéaux premiers entre eux deux à deux. D'après le lemme 4.3.29, I_{n+1} est premier avec $\bigcap_{i=1}^{n} I_i$. Le dernier point de la proposition 4.3.26 assure alors que $\bigcap_{i=1}^{n+1} I_i = \left(\bigcap_{i=1}^{n} I_i\right) I_{n+1}$. On conclut par l'hypothèse de récurrence qui garantit que $\bigcap_{i=1}^{n} I_i = \prod_{i=1}^{n} I_i$. □

Dans le cas des anneaux principaux, on retrouve une généralisation directe du cas de $\mathbb{Z}$ et $\mathbb{K}[X]$.

Corollaire 4.3.31 (Théorème des restes chinois dans un anneau principal)**.** *Soit A un anneau commutatif principal, a_1, …, a_n des éléments de A. On note π_i la projection de A sur $A/(a_i)$. Si les a_i sont premiers entre eux deux à deux alors le morphisme $\pi_1 \times \cdots \pi_n \colon A \to \prod_i A/(a_i)$ induit un isomorphisme entre $A/\left(\prod_i a_i\right)$ et $\prod_i A/(a_i)$.*

Démonstration. Supposons les a_i premiers entre eux deux à deux. Comme A est principal, le dernier point du lemme 4.3.20 assure que les (a_i) sont premiers entre eux deux à deux. Le théorème 4.3.30 assure alors que $\pi_1 \times \cdots \pi_n$ descend en isomorphisme de $A/\prod_i(a_i)$ vers $\prod_i A/(a_i)$. Or $\prod_i(a_i) = \left(\prod_i a_i\right)$ d'après le lemme 4.3.25. □

4.3.5 Éléments premiers et irréductibles

Pour conclure cette partie arithmétique du cours, on discute brièvement ce que devient la notion de nombre premier dans un anneau commutatif général et comment elle s'exprime en terme d'idéaux. Il y a deux généralisations possibles qui coïncident dans le cas des anneaux principaux mais pas en général.

Définition 4.3.32. *On dit qu'un élément a d'un anneau commutatif A est*

- irréductible *si $a \notin A^\times$ et $\forall b\, c,\ a = bc \Rightarrow b \in A^\times$ ou $c \in A^\times$*
- premier *si $a \notin A^\times$, $a \neq 0$ et $\forall b\, c,\ a \mid bc \Rightarrow a \mid b$ ou $a \mid c$*

On notera quand 0 n'est jamais irréductible ou premier (même dans un anneau trivial).

Proposition 4.3.33. *Soit a un élément d'un anneau commutatif intègre A.*

- *Si a est premier alors a est irréductible.*
- *Si $a \neq 0$, l'idéal (a) est premier si et seulement si a est premier (par contre l'idéal (0) est premier tandis que le 0 de A n'est pas un élément premier).*
- *Si l'idéal (a) est maximal et $a \neq 0$ alors a est irréductible.*
- *Si A est principal alors la réciproque du point précédent est vraie. Dans ce cas, sous l'hypothèse $a \neq 0$,*

$$a \text{ premier} \Leftrightarrow a \text{ irréductible} \Leftrightarrow (a) \text{ premier} \Leftrightarrow (a) \text{ maximal.}$$

Dans le troisième point de l'énoncé ci-dessus, le fait de supposer $a \neq 0$ n'est vraiment pas restrictif car, d'après le lemme 4.2.11, l'idéal nul n'est maximal que lorsque A est un corps, et dans cas la définition d'élément irréductible n'a aucun intérêt.

Démonstration. Supposons a premier. En particulier a n'est ni inversible ni nul. Supposons que $a = bc$ pour b et c dans A. En particulier $a \mid bc$ donc par primalité, $a \mid b$ ou $a \mid c$. Disons que $a \mid b$. Soit k dans A tel que $b = ka$. On obtient $b = kbc$ donc $b(kc - 1) = 0$. Par intégrité de A, on en déduit que $b = 0$ ou $kc = 1$. Le premier cas est exclu car $a = bc$ n'est pas nul. Ainsi $kc = 1$ et c est inversible.

Le second point découle directement des définitions. En effet, (a) est premier si et seulement si $(a) \neq 1$ et $\forall x\, y,\ xy \in (a) \Rightarrow x \in (a)$ ou $y \in (a)$, ce qui se traduit par $a \notin A^\times$ et $\forall x\, y,\ a \mid xy \Rightarrow a \mid x$ ou $a \mid y$. En supposant $a \neq 0$ on a donc bien l'équivalence. On remarque que l'idéal nul est premier car A est intègre.

Supposons maintenant que (a) est maximal et $a \neq 0$. En particulier $(a) \neq 1$ donc a n'est pas inversible. Supposons que $a = bc$ pour b et c dans A. En particulier $b \mid a$ donc $(a) \subset (b)$. Par maximalité de (a), $(b) = (a)$ ou $(b) = 1$. Dans le premier cas on obtient une unité u telle que $a = bu$ et donc $bc = bu$. Or $b \neq 0$ car sinon on aurait $a = 0$. Donc par intégrité de A on obtient $c = u$ et c est inversible. Dans le second cas on obtient directement que b est inversible.

Enfin supposons que A est principal et $a \neq 0$. On sait déjà que

$$(a) \text{ maximal } \Rightarrow (a) \text{ premier } \Leftrightarrow a \text{ premier } \Rightarrow a \text{ irréductible}$$

sans utiliser l'hypothèse que A est principal. Il reste à montrer que si a est irréductible alors (a) est maximal. Supposons a irréductible. En particulier a n'est pas inversible donc $(a) \neq 1$. Soit J un idéal contenant (a). Soit b un générateur de J. On a $b \mid a$ car $(a) \subset J$. Soit d tel que $a = bd$. Comme a est irréductible, b ou d est inversible. Si b est inversible alors $J = 1$. Si d est inversible alors $J = (a)$. □

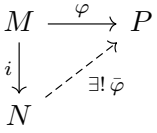
Notons que le lemme précédent assure que dans $\mathbb{Z}$ et dans $\mathbb{K}[X]$, qui sont principaux, les éléments irréductibles et premiers sont les mêmes. Pour des raisons historiques, on utilise systématiquement le terme premier dans le cas de $\mathbb{Z}$ et irréductible dans celui de $\mathbb{K}[X]$…

4.4 Localisation

Dans cette section on veut abstraire le passage de $\mathbb{Z}$ à $\mathbb{Q}$ qui créé des inverses pour les éléments non nuls de $\mathbb{Z}$, ainsi que la construction des nombres décimaux qui ne créé des inverses qu'aux puissances de dix. Dans le contexte très simple des entiers, on peut voir les nombres décimaux comme un sous-anneau du corps $\mathbb{Q}$ mais ce ne sera plus le cas en général si on part d'un anneau qui n'est pas intègre. Il faut donc un cadre général permettant d'inverser seulement certains éléments. On commence par le cas plus simple des monoïdes (qui a d'autres applications, en particulier la construction de $\mathbb{Z}$ à partir de $\mathbb{N}$).

Définition 4.4.1. *Soit M un monoïde commutatif et S une partie de M. Une* localisation *de M par rapport à S est un monoïde N muni d'un morphisme $i : M \to N$ tel que*

$i(S) \subset N^\times$ et qui sont minimaux pour cette propriété, c'est à dire qu'ils satisfont la propriété universelle suivante : pour tout monoïde P et tout morphisme $\varphi\colon M \to P$ tel que $\varphi(S) \subset P^\times$, il existe un unique $\bar{\varphi}\colon N \to X$ tel que $\varphi = \bar{\varphi} \circ i$.



Comme d'habitude, la propriété universelle de la définition assure l'unicité à unique isomorphisme près. Dans la définition de localisation, on ne demande rien à S mais le lemme suivant montre que cette notion ne voit que le sous-monoïde engendré par S.

Lemme 4.4.2. *Si $i : M \to N$ est une localisation par rapport à S alors c'est une localisation par rapport au sous-monoïde $\langle S\rangle$ engendré par S.*

Démonstration. Par hypothèse $i(S) \subset N^\times$. Or $N^\times$ est un sous-monoïde de N donc, d'après le lemme 2.0.16, il contient aussi le sous-monoïde engendré par $i(S)$, c'est à dire $i(\langle S\rangle)$. Soit P un monoïde et $\varphi : M \to P$ un morphisme tel que $\varphi(\langle S\rangle) \subset P^\times$. En particulier $\varphi(S) \subset P^\times$ donc l'hypothèse fournit l'unique $\bar{\varphi}\colon N \to P$ désiré. □

Théorème 4.4.3. *Soit M un monoïde commutatif et S une partie de M. Il existe une localisation $(S^{-1}M, i_S)$ de M par rapport à S. De plus $S^{-1}M$ est commutatif, et tous les éléments de $S^{-1}M$ s'écrivent sous la forme $i_S(a)i_S(s)^{-1}$ pour un certain $a \in M$ et $s \in \langle S\rangle$. Deux éléments a et b de S ont même image si et seulement si il existe $s \in \langle S\rangle$ tel que $sa = sb$. En particulier i_S est injective si et seulement si tous les éléments de $\langle S\rangle$ sont simplifiables.*

Démonstration. On s'inspire de la construction des nombres rationnels mais en prenant garde à l'existence potentielle d'éléments non simplifiables. On définit une relation sur $M \times \langle S\rangle$ par $(a, s) \sim (b, t)$ s'il existe $r \in \langle S\rangle$ tel que $rat = rbs$. On note que $\langle S\rangle$ contient 1 donc il s'agit d'une relation qui contient la relation naïve définie par $at = bs$. En particulier cette relation est réflexive. La symétrie est claire aussi. C'est la transitivité qui nécessite la précaution d'inclure r. Supposons que $(a, s) \sim (b, t)$ et $(b, t) \sim (c, w)$. On obtient q et r dans $\langle S\rangle$ tels que $qat = qbs$ et $rbw = rct$. On a alors $(tqr)aw = (qat)rw = (qbs)rw = (rbw)sq = (rct)sq = (tqr)cs$ et $tqr \in \langle S\rangle$ donc $(a, s) \sim (c, w)$. Avec la relation naïve on aurait seulement pu calculer $taw = tcs$ mais cela ne permet pas de conclure que $aw = cs$ si t n'est pas simplifiable.

On définit $S^{-1}M$ comme le quotient de $M \times \langle S\rangle$ par cette relation d'équivalence. On note a/s l'image d'une paire (a, s) dans ce quotient et on définit $i_S : M \to S^{-1}M$ comme envoyant a sur $a/1$. On définit 1 dans $S^{-1}M$ comme $1/1$.

Pour définir la multiplication sur $S^{-1}M$ on descend $((a, s), (b, t)) \mapsto (ab, st)$ qui est la multiplication sur le monoïde $M \times \langle S\rangle$. La vérification des conditions de compatibilité permettant la descente est directe. Le lemme 2.0.9 assure alors que $S^{-1}M$ est un monoïde. L'image de S est bien formée d'éléments inversibles car $(s/1)(1/s) = s/s = 1$ puisque $1s1 = 1 \cdot 1s$ donc $(s, s) \sim (1, 1)$. L'application i_S est clairement un morphisme de monoïdes.

Soit $\varphi\colon M \to P$ un morphisme de monoïdes tel que $\varphi(S) \subset P^\times$ (et donc $\varphi(\langle S\rangle) \subset P^\times$). L'image de φ est un sous-monoïde commutatif de P car M est commutatif, donc on

commutera librement les éléments de cette image. L'application ψ de $M \times \langle S \rangle$ dans P définie par $(a, s) \mapsto \varphi(a)\varphi(s)^{-1}$ descend au quotient car si $(a, s) \in M \times \langle S \rangle$, $(b, t) \in M \times \langle S \rangle$ et $r \in \langle S \rangle$ vérifient $rat = rbs$ alors $\varphi(r)\varphi(a)\varphi(t) = \varphi(r)\varphi(b)\varphi(s)$ et, comme $\varphi(r)$ est inversible, $\varphi(a)\varphi(t) = \varphi(b)\varphi(s)$ puis, comme $\varphi(s)$ et $\varphi(t)$ sont inversibles et commutent, $\varphi(a)\varphi(s)^{-1} = \varphi(b)\varphi(t)^{-1}$. Le lemme 2.0.9 assure que l'application descendue $\bar{\varphi}$ est un morphisme de monoïdes. Enfin pour tout a dans A, on a $\bar{\varphi}(a) = \bar{\varphi}(a/1) = \varphi(a)\varphi(1)^{-1} = \varphi(a)$ donc $\bar{\varphi}$ est bien une extension de φ. Montrons maintenant l'unicité de $\bar{\varphi}$. Soit φ' une autre extension de φ. Montrons que $\varphi' = \varphi$. Vu l'unicité dans la descente au quotient, il suffit de montrer que φ' descend ψ. Soit $(a, s) \in M \times \langle S \rangle$. On $\varphi'(a/s) = \varphi'(as^{-1}) = \varphi'(a)\varphi'(s)^{-1} = \psi(a/s)$ car φ' est un morphisme de monoïdes qui étend φ.

Vérifions enfin à quelle condition a et b dans M ont la même image dans $S^{-1}M$. $i_S(a) = i_S(b) \Leftrightarrow (a, 1) \sim (b, 1) \Leftrightarrow \exists s \in \langle S \rangle, sa1 = sb1 \Leftrightarrow \exists s \in \langle S \rangle, sa = sb.$ □

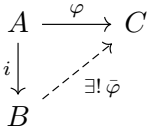
Exemple 4.4.4. La localisation du monoïde $(\mathbb{N}, +)$ par rapport à $\mathbb{N}$ est notée $(\mathbb{Z}, +)$. Comme tous les éléments de $\mathbb{N}$ sont simplifiables, l'application $\iota_{\mathbb{N}} : \mathbb{N} \to \mathbb{Z}$ est injective (ici il faut attentif au fait que la notation est additive, donc on ajoute des opposés aux éléments de $\mathbb{N}$, pas des inverses). Plus généralement tout monoïde commutatif s'envoie dans un groupe avec la propriété universelle que les morphismes à valeur dans un groupe se factorisent. Par contre l'application obtenue n'est pas injective en présence d'éléments non simplifiables.

On peut aussi étendre la multiplication de $\mathbb{N}$ à $\mathbb{Z}$ et obtenir une structure d'anneau. On peut ensuite étendre l'addition de $\mathbb{Z}$ à la localisation du monoïde $(\mathbb{Z}, \times)$ par rapport à $\mathbb{Z} \setminus \{0\}$ et obtenir un nouvel anneau noté $\mathbb{Q}$. Cette construction sera détaillée ci-dessous.

Remarque 4.4.5. L'exemple précédent, qui construit $\mathbb{Z}$ à partir de $\mathbb{N}$, incite naturellement à se demander comme est construit $\mathbb{N}$. Cette question est trop proche des fondements pour faire partie de ce cours : la réponse dépend de façon critique des fondements choisis.

On revient maintenant à notre objectif initial de localiser des anneaux.

Définition 4.4.6. *Soit A un anneau commutatif et S une partie de A. Une* localisation *de A par rapport à S est un anneau B muni d'un morphisme $i : A \to B$ tel que $i(S) \subset B^\times$ et qui sont minimaux pour cette propriété, c'est à dire qu'ils satisfont la propriété universelle suivante : pour tout anneau C et tout morphisme $\varphi : A \to C$ tel que $\varphi(S) \subset C^\times$, il existe un unique $\bar{\varphi} : B \to X$ tel que $\varphi = \bar{\varphi} \circ i$.*



Bien sûr cette définition est complètement parallèle à la définition 4.4.1 mais il n'est pas évident que le monoïde multiplicatif sous-jacent à une localisation d'un anneau A soit une localisation du monoïde $(A, \times)$ car les propriétés universelles demandées sont incomparables (il y a plus de monoïdes que d'anneaux mais la contrainte de morphisme d'anneaux est plus forte). Cependant on verra qu'on peut réutiliser le théorème 4.4.3.

On commence par observer que cette définition on ne demande rien à S mais, comme dans le cas des monoïdes, cette notion ne voit que le sous-monoïde engendré par S.

Lemme 4.4.7. *Si $i : A \to B$ est une localisation par rapport à S alors c'est une localisation par rapport au sous-monoïde (multiplicatif) $\langle S \rangle$ engendré par S.*

Démonstration. On ne peut pas appliquer directement l'analogue démontré pour les localisations de monoïdes car la propriété universelle n'est pas la bonne, mais la même démonstration fonctionne. Par hypothèse $i(S) \subset B^\times$. Or $B^\times$ est un sous-monoïde de B donc il contient aussi le sous-monoïde engendré par $i(S)$, c'est à dire $i(\langle S \rangle)$ d'après le lemme 4.3.9. Soit C un anneau et $\varphi : A \to C$ un morphisme tel que $\varphi(\langle S \rangle) \subset C^\times$. En particulier $\varphi(S) \subset C^\times$ donc l'hypothèse fournit l'unique $\bar{\varphi} : B \to C$ désiré. □

Théorème 4.4.8. *Soit A un anneau commutatif et S une partie de A. Il existe une localisation $(S^{-1}A, i_S)$ de A par rapport à S. De plus $S^{-1}A$ est commutatif, $\ker i_S = \{a | \exists s \in \langle S \rangle, as = 0\}$ et tous les éléments de $S^{-1}A$ s'écrivent sous la forme $i_S(a)i_S(s)^{-1}$ pour un certain $a \in A$ et $s \in \langle S \rangle$.*

Démonstration. On utilise le théorème 4.4.3 pour obtenir une localisation $(S^{-1}A, \iota_S)$ de $(A, \times)$ telle que tout élément de $S^{-1}A$ s'écrit sous la forme $a/s = i_S(a)i_S(s)^{-1}$ pour un certain $a \in A$ et $s \in \langle S \rangle$. Ainsi $S^{-1}A$ est le quotient de $A \times \langle S \rangle$ défini par $(a, s) \mapsto a/s$. On définit $0 \in S^{-1}A = i_S(0) = 0/1$. Pour définir l'addition on descend $((a, s), (b, t)) \mapsto (at + bs, st)$. La vérification de la condition de compatibilité permettant la descente est directe. De même on peut vérifier directement les axiomes de groupes pour l'addition et la distributivité. On sait déjà que i_S est un morphisme de monoïde multiplicatif et la compatibilité avec l'addition est claire. Comme $i_S(0) = 0$, la description du noyau de i_S provient directement du théorème 4.4.3.

Soit $\varphi : A \to C$ un morphisme d'anneaux tel que $\varphi(S) \subset C^\times$ (et donc $\varphi(\langle S \rangle) \subset C^\times$). On a déjà un morphisme de monoïdes multiplicatifs $\bar{\varphi} : S^{-1}A \to C$ défini en descendant $(a, s) \mapsto \varphi(a)\varphi(s)^{-1}$. Il ne reste à vérifier que la compatibilité avec l'addition. Soit a/s et b/t dans $S^{-1}A$. On calcule en utilisant que l'image de φ est un sous-anneau commutatif :

$$\begin{aligned}
\bar{\varphi}(a/s + b/t) &= \bar{\varphi}((at + bs)/(st)) \\
&= \varphi(at + bs)\varphi(st)^{-1} \\
&= (\varphi(a)\varphi(t) + \varphi(b)\varphi(s))\varphi(s)^{-1}\varphi(t)^{-1} \\
&= \varphi(a)\varphi(s)^{-1} + \varphi(b)\varphi(t)^{-1} \\
&= \bar{\varphi}(a/s) + \bar{\varphi}(b/t).
\end{aligned}$$

□

Comme d'habitude, la propriété universelle de la définition montre que les localisations sont uniques à unique isomorphisme près. Par abus de langage on appellera souvent $S^{-1}A$ *la* localisation de A par rapport à S. Le corollaire suivant montre ce que le modèle $S^{-1}A$ nous apprend sur les autres localisations, abstraites ou concrètes.

Corollaire 4.4.9. *Soit A un anneau commutatif et S une partie de A.*

- *Pour toute localisation (B, i) de A par rapport à S*
  - *B est commutatif*
  - *tous les éléments de B s'écrivent sous la forme $i(a)i(s)^{-1}$ pour un certain $a \in A$ et $s \in \langle S \rangle$*
  - *$\ker i = \{a | \exists s \in \langle S \rangle, as = 0\}$ et i est injectif si et seulement si S ne contient pas de diviseur de zéro.*

- $B = 0 \Leftrightarrow 0 \in \langle S \rangle$
- *Si A est un sous-anneau d'un corps K et S ne contient pas 0 alors le sous-monoïde $\langle S \rangle$ engendré par S ne contient pas non plus zéro et $B = \{a/s \ ; \ a \in A, s \in \langle S \rangle\}$ est un sous-anneau de K qui, muni de l'inclusion $A \hookrightarrow B$, est une localisation A par rapport à S.*

Démonstration. Pour le premier point, la propriété universelle de (B, i) assure qu'il existe un (unique) isomorphisme $\psi \colon S^{-1}A \to B$ tel que $i = \psi \circ i_S$ donc ces propriétés découlent directement de celles de $S^{-1}A$.

Si $B = 0$ alors $1 \in \ker i$ donc le théorème donne l'existence d'un $s \in \langle S \rangle$ tel que $1s' = 0$ donc $0 \in \langle S \rangle$. Réciproquement si 0 est dans $\langle S \rangle$ alors $i(0) = 0$ est inversible dans B donc $B = 0$.

Le théorème donne directement que i est injectif si et seulement si $\langle S \rangle$ ne contient pas de diviseur de zéro. Il suffit de montrer que S contient un diviseur de zéro si et seulement si $\langle S \rangle$ en contient un. Un des sens est clair, l'autre se démontre facilement par récurrence sur le nombre de facteur nécessaire pour écrire un élément de $\langle S \rangle$ comme produit d'éléments de S.

Supposons maintenant $A \subset K$ et $0 \notin S$. Le sous-monoïde $\langle S \rangle$ est constitué de produits d'éléments de S. Or K est intègre donc aucun de ces produits ne peut être nul. L'inclusion de A dans K est un morphisme d'anneau qui envoie les éléments de S sur des inversibles de K puisque $0 \notin S$ et K est un corps. On obtient donc une application $\varphi \colon S^{-1}A \to K$ qui étend l'inclusion et envoie a/s sur as^{-1}. Cette application est injective car son noyau est trivial et son image est B par définition de B. Ainsi on a un isomorphisme de $S^{-1}A$ vers B compatible avec les inclusions donc B est aussi une localisation de A par rapport à S. □

Exemple 4.4.10. Le nombres décimaux sont une localisation de $\mathbb{Z}$ par rapport à $\{10\}$.

Remarque 4.4.11. Soit S une partie d'un anneau commutatif A. Si $0 \in \langle S \rangle$ alors S contient au moins un diviseur de zéro mais la réciproque est fausse. Par exemple dans $A = \mathbb{Z}/6\mathbb{Z}$, $S = \{2\}$ contient un diviseur de zéro car $2 \times 3 = 0$ alors que 3 n'est pas nul mais $\langle S \rangle = \{1, 2, 4\}$ ne contient pas 0.

Exemple 4.4.12. On considère $A = \mathbb{Z}/6\mathbb{Z}$ et $S = \{2\}$. Comme $6\mathbb{Z} \subset 3\mathbb{Z}$, l'identité de $\mathbb{Z}$ descend en morphisme $i : \mathbb{Z}/6\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$. On peut vérifier que $(\mathbb{Z}/3\mathbb{Z}, i)$ est une localisation de $\mathbb{Z}/6\mathbb{Z}$ par rapport à $\{2\}$: l'image de 2 est bien inversible et la propriété universelle découle facilement de la surjectivité de i. Non seulement i n'est pas injectif, comme promis par le corollaire puisque 2 est diviseur de zéro dans $\mathbb{Z}/6\mathbb{Z}$, mais en plus il est surjectif.

Corollaire 4.4.13. *Soit A un anneau commutatif intègre. La localisation de A par rapport à $A \setminus \{0\}$ est un corps appelé* corps des fractions *de A et noté* $\mathrm{Frac}(A)$. *Le morphisme $i \colon A \to \mathrm{Frac}(A)$ de localisation est injectif.*

Ce corps vérifie la propriété universelle suivante : pour tout corps K et tout morphisme d'anneau $\varphi \colon A \to K$ injectif, il existe un unique morphisme $\bar{\varphi} \colon \mathrm{Frac}(A) \to K$ tel que $\varphi = \bar{\varphi} \circ i$.

Ainsi $\mathrm{Frac}(A)$ *est le plus petit corps dans lequel on peut injecter A (et cela le caractérise à unique isomorphisme près).*

Démonstration. L'intégrité de A assure que $S = A \setminus \{0\}$ est un sous-monoïde de A qui ne contient pas de diviseur de zéro. Ainsi le corollaire 4.4.9 assure que $S^{-1}A$ est non trivial et que i est injectif. Soit a/s un élément non nul de $S^{-1}A$. Comme $a \neq 0$, a est dans S donc on peut calculer $(a/s)(s/a) = 1$ et a/s est inversible.

Soit K un corps et $\varphi : A \to K$ un morphisme injectif. Par injectivité, φ envoie S dans $K \setminus \{0\}$, c'est à dire dans $K^\times$ puisque K est un corps. On conclut donc par la propriété universelle de $S^{-1}A$. □

Exemple 4.4.14. Le corps des fractions de $\mathbb{Z}$ est noté $\mathbb{Q}$. Le corps des fractions de $\mathbb{K}[X]$ est noté $\mathbb{K}(X)$ et appelé corps des fractions rationnelles à coefficients dans $\mathbb{K}$.

Plus généralement, on peut localiser tout anneau commutatif par rapport au sous-monoïde de ses éléments qui ne sont pas des diviseurs de zéro. On obtient ainsi un anneau dans le lequel on a inversé un maximum d'éléments sans perte d'information.

5 Modules

5.1 Définitions, morphismes et sous-objets

Dans ce chapitre on étudie la structure de module qui est la généralisation de la structure d'espace vectoriel obtenue lorsqu'on remplace le corps des scalaires par un anneau. Si A est un anneau général, on ne dit pas « A-espace vectoriel » mais « A-module ». Le début de l'histoire ne nécessite pas de supposer que cet anneau est commutatif mais au fera cette hypothèse systématiquement à partir de la troisième section. Dans cette première section on généralise les notions d'espace vectoriel, de sous-espace vectoriel et d'application linéaire. Il n'y a aucune surprise, même si les définitions sont exprimées de façon plus concise que dans un cours d'introduction à l'algèbre linéaire.

Les différences qui interviennent ultérieurement ont deux sources principales (qui sont en fait liées comme le verra dans les exemples ci-dessous et dans le corollaire 5.3.12) : sur un anneau commutatif général, un module n'a pas nécessairement de base et un sous-module n'a pas nécessairement de supplémentaire. Ces phénomènes ne nécessitent pas un anneau exotique. Le $\mathbb{Z}$-module $\mathbb{Z}/2\mathbb{Z}$ n'admet pas de base tandis que le sous-module $2\mathbb{Z}$ du $\mathbb{Z}$-module $\mathbb{Z}$ n'admet pas de supplémentaire. Le problème de l'absence de supplémentaire est résolu par la construction des modules quotients qui fait l'objet de la section 5.2. Ensuite la section 5.3 étudie les modules libres, c'est-à-dire ceux qui admettent une base. La section 5.4 assemble tous ces éléments pour fournir une classification grossière des modules ayant une famille génératrice finie lorsque l'anneau des coefficients est principal, ainsi qu'une classification complète dans le cas où cet anneau est $\mathbb{Z}$, c'est la classification des groupes abéliens de type fini. Cette section est de loin la plus technique de ce cours, ce qui est attendu puisqu'elle démontre des théorèmes de structure et de classification comme expliqué dans l'introduction générale du cours.

5.1.1 Modules et applications linéaires

Définition 5.1.1. *Soit A un anneau. Un* A-module *est un groupe abélien M muni d'un morphisme d'anneaux de A vers* $\mathrm{End}(M)$.

Remarque 5.1.2. La définition de module est très concise. Il est important de prendre le temps de la relier à la longue définition qui est habituellement donnée pour les espaces vectoriels. Tout d'abord il faut savoir que le morphisme d'anneau intervenant dans la définition est habituellement noté par un symbole invisible. Notons-le provisoirement ρ. Ensuite pour des éléments a de A et x de M l'élément $\rho(a)(x)$ est noté ax et appelé multiplication de x par le scalaire a. La définition implique les propriétés suivantes pour tous a et b dans A et tous x et y dans M :

- $a(x+y) = ax + ay$ car $\rho(a)$ est un élément de $\mathrm{End}(M)$, l'anneau des morphismes de groupes de M dans M
- $(a+b)x = ax + bx$ car ρ est additif

- $a(bx) = (ab)x$ car ρ est multiplicatif
- $1x = x$ car $\rho(1) = 1$ et le 1 de $\mathrm{End}(M)$ est le morphisme identité.

On retrouve bien les axiomes apparaissant dans la définition élémentaire d'espace vectoriel et on voit qu'on peut définir la notion de module par une telle liste. Comme dans la discussion de la définition d'anneau, on peut noter qu'on peut dresser une liste d'axiomes plus courte. En particulier la commutativité de l'addition dans un module découle des autres axiomes. La définition retenue en terme de morphisme de A dans $\mathrm{End}(M)$ dispense de ces discussions.

Exemple 5.1.3. Tout anneau A est un A-module, en utilisant la multiplication comme multiplication scalaire. Si M est un groupe abélien, c'est un $\mathrm{End}(M)$-module, via l'identité de $\mathrm{End}(M)$. Les groupes abéliens sont exactement les $\mathbb{Z}$-modules. Pour tout corps $\mathbb{K}$, les $\mathbb{K}$-espaces vectoriels sont exactement les $\mathbb{K}$-modules. Tout $\mathbb{K}$-espace vectoriel E est aussi un $\mathrm{End}(E)$-module. Pour tout endomorphisme $u \in \mathrm{End}(E)$ on obtient aussi une structure de $K[X]$-module sur E, l'image du scalaire $P \in \mathbb{K}[x]$ dans $\mathrm{End}(V)$ étant $v \mapsto P(u)v$.

En une seule occasion, dans le lemme 5.1.16, on rencontrera la notion de semi-module sur un semi-anneau, qui est obtenu en remplaçant les mots anneau et groupe abélien par semi-anneau et monoïde commutatif respectivement.

Lorsque A n'est pas commutatif, on peut préciser que les modules définis ci-dessus sont des A-modules à gauche et appeler A-modules à droite les modules sur A^{op} (l'anneau obtenu à partir de A en utilisant la multiplication $(a, b) \mapsto ba$).

Définition 5.1.4. *Un* morphisme *de A-modules entre M et M' est un morphisme de groupe $f : M \to M'$ qui est équivariant pour les actions du monoïde $(A, \times)$ sur M et M' :*

$$\forall a \in A, \forall x \in M, f(ax) = af(x).$$

On dit aussi que f est une application linéaire *de M dans M', ou même une application A-linéaire quand il y a un risque d'ambigüité (par exemple tout A-module est aussi un $\mathbb{Z}$-module et une application peut-être $\mathbb{Z}$-linéaire sans être A-linéaire). On note $\mathrm{Hom}_A(M, M')$ l'ensemble des applications A-linéaires de M dans M', on note $\mathrm{End}_A(M) = \mathrm{Hom}_A(M, M)$ et $\mathrm{Aut}_A(M) = \mathrm{End}_A(M)^\times$.*

Exemple 5.1.5. Une application entre groupes abéliens est $\mathbb{Z}$-linéaire si et seulement si c'est un morphisme de groupes. Les applications linéaires de l'algèbre linéaire sont bien des applications linéaires au sens des modules. Les applications A-linéaires de A dans lui-même sont les homothéties, c'est à dire les applications de la forme $x \mapsto ax$ pour un a fixé.

Définition 5.1.6. *Soit A un anneau et $(M_i)_{i \in \mathcal{J}}$ une famille de A-modules. Le* produit *des M_i est le groupe abélien $P = \prod_i M_i$ équipé de la structure de A-module définie composante par composante : $\forall a \in A, \forall m \in P, \forall i \in \mathcal{J}, (am)_i = am_i$. La* somme $\bigoplus_i M_i$ *est le sous-module du produit P constitué des éléments m tels que $\{i \mid m_i \neq 0\}$ est fini. On l'appelle aussi le* coproduit *des M_i.*

Lemme 5.1.7. *Les produits et coproduits de modules vérifient les propriétés universelles analogues à celles de produits et coproduits de groupes. Le produit est muni d'applications*

linéaires vers *ses facteurs et pour construire une application linéaire* vers *un produit il suffit de donner des applications linéaires* vers *les facteurs. Le coproduit est muni d'application linéaire* depuis *ses facteurs et pour construire une application linéaire* depuis *un coproduit il suffit de donner des applications linéaires* depuis *les facteurs.*

Démonstration. Le cas des produits est clair. Voyons comment la condition de finitude qui intervient dans la définition de la somme assure la propriété universelle. Soit $(M_i)_{i\in\mathcal{J}}$ une famille de A-modules. Soit N un A-module et $(\varphi_i : M_i \to N)_{i\in\mathcal{J}}$ une famille d'applications linéaires. Pour tout $m \in \bigoplus_i M_i$, on pose $\varphi(m) = \sum_i \varphi_i(m_i)$. Cette somme n'a un sens que parce que tous les m_i sont nuls sauf un nombre fini. □

5.1.2 Sous-modules

Définition 5.1.8. *Un* sous-module *d'un A-module M est un sous-groupe N de M qui est stable par multiplication scalaire : $\forall a \in A, \forall n \in N, an \in N$. On peut préciser le terme sous-module en sous-A-module en cas d'ambigüité sur l'anneau des scalaires.*

Exemple 5.1.9. Dans un $\mathbb{K}$-espace vectoriel, les sous-$\mathbb{K}$-modules sont exactement les sous-$\mathbb{K}$-espaces vectoriels. Les sous-$\mathbb{Z}$-modules d'un groupe abélien sont exactement ses sous-groupes. Les sous-A-modules de A sont exactement ses idéaux. Le singleton $\{0\}$ et le module entier sont des sous-A-modules.

Lemme 5.1.10. *L'image d'un sous-module par une application linéaire est un sous-module. En particulier l'image d'une application linéaire est un sous-module. La préimage d'un sous-module par une application linéaire est un sous-module. En particulier le noyau d'une application linéaire est un sous-module.*

Une intersection de sous-modules est un sous-module. En particulier on a une notion de sous-module engendré *par une partie, avec toutes les propriétés habituelles pour les sous-objets engendrés.*

Le sous-module engendré par une partie S est l'ensemble des sommes de la forme $\sum_{s\in s} a_s s$ pour une fonction $a : S \to A$ à support fini *(c'est à dire nulle sauf sur un ensemble fini).*

Démonstration. Soit A un anneau commutatif. Soit $\varphi : M \to M'$ une application A-linéaire entre A-modules et N un sous-module de M. Le lemme 3.1.8 assure que $\varphi(N)$ est un sous-groupe de M'. Il reste à voir la stabilité par multiplication scalaire. Soit $n \in N$ et $a \in A$. On $a\varphi(n) = \varphi(an)$ et N est stable donc $an \in N$ puis $a\varphi(n) \in \varphi(N)$.

Soit N' un sous-module de M'. Le même lemme assure que $\varphi^{-1}(N')$ est un sous-groupe de M. Soit $n \in \varphi^{-1}(N')$ et $a \in A$. On a $\varphi(an) = a\varphi(n)$ et N' est stable donc $a\varphi(n) \in N'$ puis $an \in \varphi^{-1}(N')$.

Soit $\mathcal{N}$ une famille de sous-modules de M. Le lemme 3.1.8 assure que $N_0 = \bigcap_{N\in\mathcal{N}} N$ est un sous-groupe de M. Soit $n \in N_0$ et $a \in A$. Pour tout $N \in \mathcal{N}$, $n \in N$ et N est stable donc $an \in N$. On a donc $an \in N_0$.

Montrons maintenant la description du sous-module engendré par une partie S de M. Notons N l'ensemble de l'énoncé. On a $S \subset N$. De plus la stabilité par multiplication scalaire force $as \in \langle S\rangle$ pour tout $s \in S$ et $a \in A$ puis la stabilité par somme force $N \subset \langle S\rangle$. Ainsi il suffit de montrer que N est un sous-module de M, ce qui est clair. □

Remarque 5.1.11. Le premier point du lemme ci-dessus peut paraître choquant quand on pense que les sous-A-modules de A sont ses idéaux et que l'image directe d'un idéal par un morphisme d'anneau n'est un idéal que pour les morphismes surjectifs en général. Mais ces deux résultats parlent de classes différentes d'applications. Les applications A-linéaires de A dans lui-même sont les homothéties et on peut vérifier directement qu'elles envoient les idéaux sur des idéaux. Les applications $\mathbb{Z}$-linéaires de A dans un autre anneau commutatif B sont simplement les morphismes de groupes entre A et B et les sous-$\mathbb{Z}$-modules sont les sous-groupe de A. Dans ce cas on retrouve l'absence de condition de surjectivité du lemme 3.1.8 concernant l'image d'un sous-groupe.

Lemme 5.1.12. *Soit A un anneau et M un A-module. L'ensemble des sous-modules de M est un monoïde commutatif pour l'opération qui envoie (N, N') sur le sous-module $N + N'$ engendré par $N \cup N'$ et appelé* somme *de N et N'.*

Démonstration. Ces propriétés se démontrent exactement comme dans le cas particulier des idéaux qui fait l'objet de la proposition 4.3.26. Montrons l'associativité. Soit N, N' et N'' des sous-modules de M. Vu l'associativité de la réunion, il suffit de montrer que $N + (N' + N'') = \langle N \cup (N' \cup N'')\rangle$ et $(N + N') + N'' = \langle (N \cup N') \cup N''\rangle$. Soit P un sous-module de M.

$$\begin{aligned} N \cup (N' \cup N'') \subset P &\Leftrightarrow N \subset P \text{ et } N' \cup N'' \subset P \\ &\Leftrightarrow N \subset P \text{ et } N' + N'' \subset P \\ &\Leftrightarrow N \cup (N' + N'') \subset P \\ &\Leftrightarrow N + (N' + N'') \subset P \end{aligned}$$

Donc le sous-module $N + (N' + N'')$ vérifie la propriété universelle qui caractérise $\langle N \cup (N' \cup N'')\rangle$. Le cas de $(N + N') + N''$ fonctionne exactement de la même façon.

Le sous-module nul est neutre car $N+0 = \langle N \cup 0\rangle = \langle N\rangle = N$ et de même $0+N = N$.

La commutativité découle directement de celle de la réunion puisque, pour tous sous-modules N et N', $N + N' = \langle N \cup N'\rangle = \langle N' \cup N\rangle = N' + N$. □

Plus généralement, on peut définir de même la somme d'une famille quelconque de sous-modules. Cette notion de somme de sous-modules est liée à la somme de modules de la définition 5.1.6. La somme d'un famille de sous-modules est l'image de leur somme en tant que modules abstraits par l'application induite par les inclusions. Pour la suite il suffira de considérer le cas où cette application est injective, c'est l'objet de la définition suivante.

Définition 5.1.13. *Soit A un anneau, M un A-module et $(M_i)_{i\in\mathcal{J}}$ une famille de sous-modules de M. On dit que les M_i sont en* somme directe *si l'application de $\bigoplus_i M_i$ dans M induite par les inclusions est injective. Lorsqu'elle est de plus surjective, on écrit abusivement $M = \bigoplus_i M_i$. Dans le cas de deux sous-modules M_1 et M_2 tels que $M = M_1 \oplus M_2$, on dit que M_1 et M_2 sont supplémentaires l'un de l'autre.*

Comme en algèbre linéaire sur un corps, les décompositions en somme directes correspondent à la notion de projecteur.

Définition 5.1.14. *Soit M un A-module. Un projecteur sur M est une application $p \in \mathrm{End}_A(M)$ telle que $p \circ p = p$.*

Lemme 5.1.15. *Soit M un A-module et $p \in \mathrm{End}_A(M)$ un projecteur. On a $M = \ker p \oplus \mathrm{im}\, p$. Réciproquement si $M = M_1 \oplus M_2$ pour des sous-modules M_1 et M_2 alors il existe des projecteurs p_1 et p_2 tels que $\ker p_1 = M_2$, $\mathrm{im}\, p_1 = M_1$, $\ker p_2 = M_1$, $\mathrm{im}\, p_2 = M_2$ et $\mathrm{Id}_M = p_1 + p_2$.*

Démonstration. On a $M = \ker p + \mathrm{im}\, p$ car $\forall m, m = (m - p(m)) + p(m)$ où le premier morceau est dans $\ker p$ car p est un projecteur. Montrons que la somme est directe. Soit $x \in \ker p$ et z tels que $x + p(z) = 0$. On a $x = -p(z) = p(-z)$ puis, en appliquant p et en utilisant que x est dans $\ker p$, on obtient $0 = p(x) = p(p(-z)) = p(-z)$. Ainsi $p(-z) = 0$ et donc $x = 0$ et $p(z) = 0$.

Réciproquement on suppose que $M = M_1 \oplus M_2$. La propriété universelle des sommes fournit p_1 et p_2 tels que p_1 coïncide avec l'inclusion sur M_1 et l'application nulle sur M_2 tandis que p_2 coïncide avec l'inclusion sur M_2 et l'application nulle sur M_1. Ces applications conviennent. □

Les opérations sur les sous-modules décrites jusqu'ici existent déjà dans le cas des espaces vectoriels. Passons maintenant à une opération qui n'a d'intérêt que lorsque l'anneau des scalaires n'est pas un corps car les seuls idéaux d'un corps sont 0 et 1 (voir le lemme 4.2.11). Il est donc inutile de chercher ce que le lemme suivant généralise en algèbre linéaire sur un corps.

Lemme 5.1.16. *Soit A un anneau, $I \lhd A$ un idéal, M un A-module. Le sous-module engendré par l'ensemble des im avec $i \in I$ et $m \in M$ est l'ensemble des sommes de tels éléments. On le note IM.*

L'opération $I \mapsto (N \mapsto IN)$ fait du monoïde des sous-modules de M un semi-module sur le semi-anneau des idéaux de A.

Démonstration. Le sous-module engendré par les im contient leurs sommes donc il suffit de montrer que l'ensemble des $\sum_\lambda i_\lambda m_\lambda$ est bien un sous-module. Il s'agit clairement d'un sous-groupe. Pour tout a dans A on a $a \sum_\lambda i_\lambda m_\lambda = \sum_\lambda (a i_\lambda) m_\lambda$ et chaque $a i_\lambda$ est dans I car I est un idéal.

Soit I un idéal de A et N et N' des sous-modules de M. Pour tout i dans I, on note $\mu_i \in \mathrm{End}_A(M)$ la multiplication par i. Montrons que $I(N+N')$ vérifie la propriété universelle de $IN + IN$. Soit P un sous-module de M.

$$\begin{aligned}
I(N+N') \subset P &\Leftrightarrow \forall i \in I, \forall m \in N+N', im \in P \\
&\Leftrightarrow \forall i \in I, \mu_i(N+N') \subset P \\
&\Leftrightarrow \forall i \in I, \mu_i(\langle N \cup N' \rangle) \subset P \\
&\Leftrightarrow \forall i \in I, \langle \mu_i(N) \cup \mu_i(N') \rangle \subset P \\
&\Leftrightarrow \forall i \in I, \mu_i(N) \cup \mu_i(N') \subset P \\
&\Leftrightarrow (\forall i \in I, \mu_i(N) \subset P) \text{ et } (\forall i \in I, \mu_i(N') \subset P) \\
&\Leftrightarrow IN \subset P \text{ et } IN' \subset P \\
&\Leftrightarrow IN \cup IN' \subset P \\
&\Leftrightarrow IN + IN' \subset P
\end{aligned}$$

Ainsi $I(N+N') = IN + IN'$. On montre de façon analogue que, pour tous idéaux I et J dans A et tout sous-module N de M, $I(JN) = (IJ)N$. Le fait que $0N = 0$ est

clair (mais cela ne découle pas immédiatement de l'additivité car les idéaux de A ne forment pas un groupe). Enfin, pour tout sous-module N, $1N = N$ car N est stable par multiplication scalaire donc $IN \subset N$ pour tout I et $1N \supset \mu_1(N) = N$. □

Exemple 5.1.17. Dans le cas où $M = A^r$, le sous-module IM du lemme précédent est simplement I^r. Dans le cas où $A = \mathbb{Z}$, $I = \mathbb{Z}/2\mathbb{Z}$ et $M = \mathbb{Z}/\mathbb{Z}4$, $IM = \{0, 2\}$.

5.2 Modules quotients et suites exactes courtes

5.2.1 Modules quotients

En algèbre linéaire sur un corps, on peut cacher longtemps l'importance des espaces vectoriels quotients en utilisant que tout sous-espace vectoriel admet un supplémentaire. Au prix de quelques contorsions, un tel supplémentaire peut jouer le rôle d'un quotient. Souvent les premiers espaces vectoriels quotients rencontrés explicitement le sont en cours d'intégration où le sous-espace des fonctions nulles presque partout n'a vraiment pas de supplémentaire agréable dans l'espace des fonctions intégrables. Le quotient, noté L^1, est alors difficile à éviter.

Toujours sur un corps, la dimension finie permet aussi d'éviter les quotients avec une perte d'information limitée. Par exemple la proposition 5.2.2 expliquera entre autres comment toute application linéaire $\varphi : E \to F$ induit un isomorphisme entre $E/\ker\varphi$ et $\operatorname{im}\varphi$. On en déduit facilement le corollaire $\dim(E) = \dim(\ker\varphi) + \dim(\operatorname{im}\varphi)$ appelé théorème du rang. Il s'agit d'une information bien plus faible mais déjà utile si les dimensions intervenant sont finies.

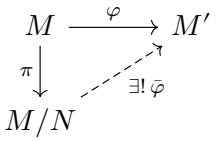
La notion de module quotient est complètement analogue aux cas des groupes et anneaux.

Définition 5.2.1. *Un quotient d'un A-module M est un A module N équipé d'une application A-linéaire surjective $\pi : M \to N$. On dit que c'est un quotient de M par un sous-module $M' \subset M$ si* $\ker\pi = M'$.

Comme une application linéaire est en particulier un morphisme de groupes pour l'addition, un module quotient est en particulier un groupe quotient. La théorie du chapitre 3, particulièrement le lemme 3.3.3 et le théorème 3.3.6, se combine avec le lemme 5.1.10 pour assurer que tout quotient de M est isomorphe comme groupe abélien à M/N pour un sous-module N. La proposition suivante assure réciproquement qu'à tout sous-module N est associé un module quotient, unique à unique isomorphisme près.

Proposition 5.2.2. *Soit A un anneau, M un A-module et $N \subset M$ un sous-module. Il existe une unique structure de A-module sur le groupe quotient M/N qui fasse de $\pi : M \to M/N$ une application A-linéaire.*

Le quotient $\pi : M \to M/N$ *vérifie la propriété universelle suivante. Pour toute application A-linéaire $\varphi : M \to M'$ telle que $N \subset \ker\varphi$, il existe une unique application $\bar{\varphi}$ qui fait commuter*



En particulier φ induit un isomorphisme A-linéaire de $M/\ker\varphi$ dans $\operatorname{im}\varphi$.

Pour toute application A-linéaire $\varphi\colon M \to M'$ et tout sous-module $N' \subset M'$, tel que $\varphi(N) \subset N'$, il existe une unique application A-linéaire $\hat{\varphi} : M/N \to M'/N'$ qui fait commuter

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & M' \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ M/N & \xrightarrow[\exists!\,\hat{\varphi}]{} & M'/N' \end{array}$$

Démonstration. Le théorème 3.3.6 assure déjà l'existence d'une unique structure de groupe sur M/N qui fasse de π un morphisme de groupes. Il s'agit donc de s'occuper de la multiplication scalaire. Soit $a \in A$ et $\mu_a \in \operatorname{End}(M)$ la multiplication par a. On veut compléter le diagramme

$$\begin{array}{ccc} M & \xrightarrow{\mu_a} & M \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ M/N & \xrightarrow[\bar{\mu}_a]{} & M/N \end{array}$$

Le corollaire 3.3.12 assure l'unicité de $\bar{\mu}_a$ et donne la condition nécessaire d'existence $\mu_a(N) \subset N$ qui est bien vérifié car N est un sous-module. Il reste à vérifier que $a \mapsto \bar{\mu}_a$ est un morphisme d'anneaux de A dans $\operatorname{End}(M/N)$. Montrons que, en plus des propriétés de $a \mapsto \mu_a$, cela découle de la commutativité du diagramme ci-dessus, de la surjectivité de π et du fait que π est un morphisme de groupes. Soit a et b dans A. Pour tout m dans M on a $\bar{\mu}_{a+b}(\pi(m)) = \pi(\mu_{a+b}(m)) = \pi(\mu_a(m) + \mu_b(m)) = \pi(\mu_a(m)) + \pi(\mu_a(m)) = \bar{\mu}_a(\pi(m)) + \bar{\mu}_b(\pi(m))$. De même $\bar{\mu}_{ab}(\pi(m)) = \pi(\mu_{ab}(m)) = \pi(\mu_a \circ \mu_b(m)) = \bar{\mu}_a(\pi(\mu_b(m))) = \bar{\mu}_a(\bar{\mu}_b(\pi(m)))$ et $\bar{\mu}_1(\pi(a)) = \pi(\mu_1(a)) = \pi(a)$.

Montrons maintenant la propriété universelle. Soit $\varphi\colon M \to M'$ une application A-linéaire telle que $N \subset \ker\varphi$. Le théorème 3.3.9 assure que φ descend en morphisme de groupe $\bar{\varphi}\colon M/N \to M'$. Il reste à vérifier que $\bar{\varphi}$ est A-équivariant. Soit $m \in M$ et $a \in A$. On a $\bar{\varphi}(a\pi(m)) = \bar{\varphi}(\pi(am)) = \varphi(am) = a\varphi(m) = a\bar{\varphi}(\pi(m))$.

La dernière partie de l'énoncé découle directement de la partie précédente appliquée à $\pi \circ \varphi$. □

Expliquons maintenant en quel sens un éventuel supplémentaire peut jouer le rôle d'un quotient. On verra dans la section suivante que l'existence d'un supplémentaire n'a rien d'automatique.

Lemme 5.2.3. *Soit A un anneau, M un A-module, N un sous-A-module de M et $\pi\colon M \to M/N$ la projection sur le quotient. On suppose que N admet un supplémentaire N'. Alors la restriction de π à N' est un isomorphisme.*

Démonstration. On a $\ker\pi|_{N'} = \ker\pi \cap N' = N \cap N' = \{0\}$ donc $\pi|_{N'}$ est injective. Montrons la surjectivité. Soit $z \in M/N$. Par surjectivité de π, on obtient $x \in M$ tel que $\pi(x) = z$. Soit p et p' les projecteurs associés à la décomposition $M = N \oplus N'$ par le lemme 5.1.15. On a $\pi(x) = \pi(p(x) + p'(x)) = \pi(p(x)) + \pi(p'(x)) = \pi(p'(x))$ car $\operatorname{im} p = N = \ker\pi$. Ainsi $\pi(p'(x)) = z$ et comme $p'(x) \in N'$, on a bien la préimage cherchée. □

Ainsi la construction des modules quotients permet de remplacer l'utilisation d'un supplémentaire. Elle est utile même dans le cas des espaces vectoriels. Au contraire le lemme suivant ne généralise pas un résultat concernant les espaces vectoriels. Il est vraiment spécifique au cas des anneaux car il repose sur la construction du lemme 5.1.16.

Lemme 5.2.4. *Soit A un anneau, $I \lhd A$ un idéal, M un A-module. Il existe une unique structure de A/I-module sur le groupe abélien M/IM qui fasse commuter*

$$\begin{array}{ccc} A & \longrightarrow & \mathrm{End}_{\mathbb{Z}}(M/IM) \\ {\scriptstyle\pi}\downarrow & \nearrow_{\exists!} & \\ A/I & & \end{array}$$

Pour toute application A-linéaire $\varphi : M \to M'$, il existe une unique application A/I-linéaire $\bar{\varphi}$ qui fait commuter

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & M' \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ M/IM & \xrightarrow[\exists!\,\bar{\varphi}]{} & M'/IM' \end{array}$$

Cette construction est fonctorielle : $\overline{\mathrm{Id}_M} = \mathrm{Id}_{M/IM}$ pour tout M et $\overline{\varphi \circ \psi} = \bar{\varphi} \circ \bar{\psi}$. En particulier si φ est un isomorphisme alors $\bar{\varphi}$ aussi.

Démonstration. On a déjà vu dans la proposition 5.2.2 d'où provient la structure de A-module sur M/IM. Il s'agit de voir que le morphisme d'anneau μ de A dans $\mathrm{End}_{\mathbb{Z}}(M/IM)$ descend de façon unique à A/I. D'après la propriété universelle des anneaux quotients, il suffit de vérifier que $I \subset \ker \mu$. Soit $i \in I$ et $x \in M/IM$. Soit $m \in M$ tel que $x = \pi(m)$. On a $\mu_i(x) = \mu_i(\pi(m)) = \pi(\mu_i(m)) = 0$ où la dernière égalité provient de $\mu_i(m) \in IM$.

Pour la deuxième partie, on considère une application A-linéaire $\varphi : M \to M'$. On veut la descendre en application A-linéaire de M/IM dans M'/IM'. Toujours d'après la même proposition 5.2.2, il s'agit de vérifier que $\varphi(IM) \subset IM'$. Puisque IM' est un sous-module, il suffit de vérifier que, pour tout i dans I et m dans M, $\varphi(im) \in IM'$, ce qui est clair par linéarité de φ. La fonctorialité découle de l'unicité comme d'habitude (voir par exemple la démonstration du corollaire 3.4.5). □

Le lemme précédent permet par exemple de fabriquer à partir d'un groupe abélien M un $\mathbb{Z}/p\mathbb{Z}$-espace vectoriel $M/p\mathbb{Z}M$.

5.2.2 Suites exactes courtes

Nous avons vu dans le lemme 5.2.3 que la construction des modules quotients peut être vu comme généralisation de la notion de supplémentaire d'un sous-module. Dans cette section on étudie dans quels cas cette généralisation est essentielle car un supplémentaire n'existe pas.

Définition 5.2.5. *Soit A un anneau commutatif. Une* suite exacte *de A-modules est une suite $(M_i)_{i\in\mathbb{Z}}$ de A-modules munie d'applications A-linéaires $f_i : M_i \to M_{i+1}$ telles que*

$$\forall i, \mathrm{im}\, f_i = \ker f_{i+1}.$$

Une suite exacte est courte *si elle comporte au plus trois termes M_i non nuls et que ces termes sont successifs. On écrit une telle suite sous la forme*

$$0 \longrightarrow M \xrightarrow{i} N \xrightarrow{p} Q \longrightarrow 0.$$

Lemme 5.2.6. *Dans une suite exacte courte*

$$0 \longrightarrow M \xrightarrow{i} N \xrightarrow{p} Q \longrightarrow 0,$$

i est injective et p est un quotient de noyau l'image $i(M)$.

Démonstration. La première flèche est forcément l'application nulle car c'est la seule application A-linéaire partant du module nul. La condition d'exactitude en M est donc $\operatorname{im} 0 = \ker i$ donc $\ker i = 0$. De même la dernière flèche est nulle car c'est la seule application à valeur dans le module nul. La condition d'exactitude en Q est donc $\operatorname{im} p = \ker 0$ donc $\operatorname{im} p = Q$ et p est surjective. Enfin la condition d'exactitude en N est exactement $\ker p = \operatorname{im} i$. □

Exemple 5.2.7. Soit A un anneau commutatif.

- Pour tout sous-module M d'un A-module N, on a la suite exacte
$$0 \longrightarrow M \hookrightarrow N \xrightarrow{\pi} N/M \longrightarrow 0.$$
- Pour tous A-modules M_1 et M_2, on a la suite exacte
$$0 \longrightarrow M_1 \xrightarrow{i_1} M_1 \oplus M_2 \xrightarrow{\mathrm{pr}_2} M_2 \longrightarrow 0$$
où $i_1 : m_1 \mapsto (m_1, 0)$ et $\mathrm{pr}_2 : (m_1, m_2) \mapsto m_2$.

Le premier exemple ci-dessus est, à isomorphisme près, la seule façon de construire une suite exacte courte. Par contre l'aspect symétrique du second exemple est très spécial, en particulier parce que le module du milieu contient une copie du module de droite. Ainsi la suite exacte

$$0 \longrightarrow n\mathbb{Z} \hookrightarrow \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/n\mathbb{Z} \longrightarrow 0$$

ne peut pas être de cette forme spéciale car $\mathbb{Z}$ ne contient pas de copie de $\mathbb{Z}/n\mathbb{Z}$ si $n \neq 0$ ($\mathbb{Z}$ ne contient aucun élément d'ordre n). À retenir : en algèbre linéaire sur un corps, utiliser un supplémentaire plutôt qu'un quotient est une simple faute de goût, sur un anneau général ce peut être une erreur cruciale. Cet exemple explique aussi pourquoi $\mathbb{Z}/n\mathbb{Z}$ est souvent le premier quotient rencontré explicitement : il n'y a pas de supplémentaire de $n\mathbb{Z}$ susceptible de le remplacer.

Définition 5.2.8. *Une suite exacte courte*

$$0 \longrightarrow M \xrightarrow{i} N \xrightarrow{p} Q \longrightarrow 0$$

est scindée *s'il existe des modules M_1 et M_2 et des isomorphismes qui font commuter le diagramme suivant*

$$\begin{array}{ccccccccc}
0 & \longrightarrow & M & \xrightarrow{i} & N & \xrightarrow{p} & Q & \longrightarrow & 0 \\
 & & \downarrow\wr & & \downarrow\wr & & \downarrow\wr & & \\
0 & \longrightarrow & M_1 & \xrightarrow{i_1} & M_1 \oplus M_2 & \xrightarrow{\mathrm{pr}_2} & M_2 & \longrightarrow & 0
\end{array}$$

On peut obtenir une caractérisation commode des suites scindées en utilisant les projecteurs (définition 5.1.14).

Lemme 5.2.9. *Soit*

$$0 \longrightarrow M \xrightarrow{\ i\ } N \xrightarrow{\ p\ } Q \longrightarrow 0$$

une suite exacte courte. Les propriétés suivantes sont équivalentes :

1. *la suite est scindée ;*
2. *il existe une* rétraction *de i, c'est-à-dire une application linéaire $r\colon N \to M$ telle que $r \circ i = \mathrm{Id}_M$;*
3. *il existe une* section *de p, c'est-à-dire une application linéaire $s\colon Q \to N$ telle que $p \circ s = \mathrm{Id}_Q$.*

Dans le lemme ci-dessus, il existe toujours des fonctions s et r telles que $p \circ s = \mathrm{Id}_Q$ et $r \circ i = \mathrm{Id}_M$. Toute la question est de savoir si on peut en trouver qui soient linéaires.

Démonstration. Montrons la suite d'implications (1) ⇒ (2) ⇒ (3) ⇒ (1). Supposons la suite scindée. On a donc un diagramme de la forme

$$\begin{array}{ccccccccc}
0 & \longrightarrow & M & \xrightarrow{\ i\ } & N & \xrightarrow{\ p\ } & Q & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \psi_1} & & \downarrow{\scriptstyle \psi_2} & & \downarrow{\scriptstyle \psi_3} & & \\
0 & \longrightarrow & M_1 & \underset{\mathrm{pr}_1}{\overset{i_1}{\rightleftarrows}} & M_1 \oplus M_2 & \xrightarrow{\ \mathrm{pr}_2\ } & M_2 & \longrightarrow & 0
\end{array}$$

et on peut poser $r = \psi_1^{-1} \circ \mathrm{pr}_1 \circ \psi_2$. La relation $r \circ i = \mathrm{Id}_M$ découle directement de la commutativité du carré de gauche dans le diagramme et de la relation $\mathrm{pr}_1 \circ i_1 = \mathrm{Id}_{M_1}$:

$$\begin{aligned}
r \circ i &= \psi_1^{-1} \circ \mathrm{pr}_1 \circ \psi_2 \circ i \\
&= \psi_1^{-1} \circ \mathrm{pr}_1 \circ i_1 \circ \psi_1 \\
&= \psi_1^{-1} \circ \mathrm{Id}_{M_1} \circ \psi_1 \\
&= \mathrm{Id}_M .
\end{aligned}$$

Supposons maintenant qu'il existe une rétraction r de i et montrons qu'il existe une section de p. Il suffit de montrer que la restriction p' de p à $\mathrm{im}(\mathrm{Id} - i \circ r)$ est un isomorphisme car alors l'inverse de p' sera une section de p. On a $\ker p' = \ker p \cap \mathrm{im}(\mathrm{Id} - i \circ r)$. Comme la suite est exacte, $\ker p = \mathrm{im}\, i$. Ainsi tout élément x de $\ker p \cap \mathrm{im}(\mathrm{Id} - i \circ r)$ s'écrit comme $i(m)$ et $m' - i(r(m'))$ pour des éléments m et m' de M. En appliquant r à l'égalité $i(m) = m' - i(r(m'))$ et en utilisant l'hypothèse $r \circ i = \mathrm{Id}$ on obtient $m = r(m') - r(m')$ donc m puis x sont nuls. Ainsi $\ker p' = 0$. Montrons maintenant que p' est surjective. Soit $q \in Q$. Comme la suite est exacte, p est surjective donc on obtient $n \in N$ tel que $p(n) = q$. On a alors $p(n - i \circ r(n)) = p(n) - p \circ i \circ r(n) = p(n) = q$ car la suite est exacte donc $p \circ i = 0$. Ainsi q est dans l'image de p'.

Enfin supposons que s existe et montrons que la suite est scindée. On a $(s \circ p)^2 = s \circ p \circ s \circ p = s \circ \mathrm{Id} \circ p = s \circ p$. Le lemme 5.1.15 assure alors que $N = \ker(s \circ p) \oplus \mathrm{im}(s \circ p)$. Comme $p \circ s = \mathrm{Id}$, s est injective donc $\ker(s \circ p) = \ker p$. Ce dernier est égal à $\mathrm{im}\, i$ car la suite de départ est exacte. Ainsi $\ker(s \circ p) = \mathrm{im}\, i$. De plus p est surjective donc $\mathrm{im}(s \circ p) = \mathrm{im}\, s$. En rassemblant les morceaux on obtient $N = \mathrm{im}\, i \oplus \mathrm{im}\, s$ et donc un isomorphisme $i \oplus s\colon M \oplus Q \to N$ qui fournit le diagramme

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M & \xrightarrow{i} & N & \xrightarrow{p} & Q & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle (i\oplus s)^{-1}} & & \downarrow{\scriptstyle \mathrm{Id}} & & \\ 0 & \longrightarrow & M & \xrightarrow{i_1} & M\oplus Q & \xrightarrow{\mathrm{pr}_2} & Q & \longrightarrow & 0 \end{array}$$

qui est bien de la forme annoncée.

Remarque : la démonstration ci-dessus économise un maximum d'implications mais on peut aussi montrer directement toutes les autres avec des démonstrations très analogues. □

5.3 Modules libres

5.3.1 Définition, existence et unicité

Les modules libres sont ceux qui admettent une base, ce qui n'est pas automatique lorsque l'anneau des scalaires n'est pas un corps. Dans la définition suivante, la base est l'application ι.

Définition 5.3.1. *Soit S un ensemble et A un anneau commutatif. Un A-*module libre *sur S est un A-module M muni d'une application $\iota : S \to M$ qui vérifie la propriété universelle suivante : pour tout A-module M' et toute fonction f de S dans M', il existe une unique application A-linéaire $\bar{f} : M \to M'$ tel que $f = \bar{f} \circ \iota$.*

$$\begin{array}{ccc} S & \xrightarrow{f} & M' \\ {\scriptstyle \iota}\downarrow & \nearrow_{\exists!\, \bar{f}} & \\ M & & \end{array}$$

Exemple 5.3.2. Soit $\mathbb{K}$ un corps, E un $\mathbb{K}$-espace vectoriel et $e : \{1, \dots, n\} \to E$ une base de E. Alors (E, e) est un module libre sur $\{1, \dots, n\}$ comme on le verra dans le corollaire 5.3.4. Pour tout anneau commutatif A, A^n muni de la base canonique est un module libre sur $\{1, \dots, n\}$ comme on le verra dans l'exemple 5.3.6.

Pour tout anneau, le seul module M libre sur l'ensemble vide est le module nul car sinon l'identité de M la fonction nulle de M dans M sont deux applications A-linéaires distinctes qui font commuter l'unique diagramme pertinent (on rappelle qu'il existe exactement une fonction du vide dans M).

Soit n un entier strictement positif. Montrons que le $\mathbb{Z}$-module $\mathbb{Z}/n\mathbb{Z}$ n'est pas libre. Supposons que $\iota : S \to \mathbb{Z}/n\mathbb{Z}$ est une base de $\mathbb{Z}/n\mathbb{Z}$. L'exemple précédent montre que S n'est pas vide. Soit $s_0 \in S$. Soit f une fonction de S dans $\mathbb{Z}$ qui envoie s_0 sur 1 et soit $\bar{f}$ l'extension promise par la propriété universelle de ι. On a $\bar{f}(n\iota(s_0)) = n\bar{f}(\iota(s_0)) = nf(s_0) = n$. Or $n\iota(s_0) = 0$ donc $\bar{f}(n\iota(s_0)) = 0$ et on a une contradiction. Cet exemple sera généralisé dans le lemme 5.3.19.

On fait maintenant le lien avec la notion de combinaison linéaire.

Proposition 5.3.3. *Soit S un ensemble et (M, ι) un A-module libre sur S. Tout élément de M s'écrit de façon unique comme $\sum_{s\in S} a_s \iota(s)$ pour une fonction $a : S \to A$ à support fini (la somme a donc bien un sens).*

Réciproquement, pour tout un A-module M et tout $\iota : S \to M$, si tout élément de M s'écrit de façon unique comme $\sum_{s\in S} a_s\iota(s)$ pour une fonction $a : S \to A$ à support fini alors (M, ι) est un A-module libre sur S.

Démonstration. Montrons d'abord que $\iota(S)$ engendre M. La projection canonique de M sur $M/\langle\iota(S)\rangle$ et le morphisme nul entre ces modules étendent tous deux la fonction nulle de S dans $M/\langle\iota(S)\rangle$. Par unicité dans la propriété universelle, ces deux morphismes sont égaux. Or la projection est surjective donc $M/\langle\iota(S)\rangle = \{0\}$.

Soit x un élément de M. Le paragraphe précédent et le lemme 5.1.10 assurent l'existence d'une écriture de x comme annoncé. Supposons maintenant que x s'écrive $\sum_{s\in S} a_s\iota(s)$ et $\sum_{s\in S} b_s\iota(s)$ pour deux fonctions a et b à support fini. Pour tout s dans S, on considère la fonction $\delta_s : S \to A$ qui vaut un en s et zéro ailleurs. On note Δ_s son extension à M promise par la propriété universelle. En appliquant Δ_s aux deux sommes on obtient $a_s = b_s$.

Réciproquement, supposons maintenant que (M, ι) vérifie cette condition d'écriture unique. Soit M' un A-module et f une fonction de S dans M'. Montrons d'abord l'unicité de $\bar{f}$ car cela aidera pour l'existence. Soit $x \in M$. Par hypothèse on peut écrire $x = \sum_s a_s\iota(s)$. On calcule :

$$\begin{aligned}\bar{f}(x) &= \bar{f}\left(\sum_s a_s\iota(s)\right)\\ &= \sum_s a_s\bar{f}(\iota(s)) \text{ car } \bar{f} \text{ est } A\text{-linéaire}\\ &= \sum_s a_s f(s) \text{ car } \bar{f}\circ\iota = f\end{aligned}$$

Donc la valeur de $\bar{f}(x)$ est uniquement spécifiée par nos contraintes.

Pour l'existence de $\bar{f}$, on peut utiliser l'unicité de l'écriture comme somme pour définir $\bar{f}$ par la formule ci-dessus. Il reste à montrer qu'on obtient bien une application linéaire. Soit x et y dans M. Comme $(M, +)$ est abélien, on peut écrire $x + y = \sum_s (a_s + b_s)\iota(s)$ et calculer :

$$\begin{aligned}\bar{f}(x+y) &= \bar{f}\left(\sum_s (a_s+b_s)\iota(s)\right)\\ &= \sum_s (a_s+b_s)f(s) \text{ par définition de } \bar{f}\\ &= \sum_s a_s f(s) + \sum_s b_s f(s)\\ &= \bar{f}(x) + \bar{f}(x).\end{aligned}$$

La compatibilité avec la multiplication par un scalaire est encore plus directe. □

Corollaire 5.3.4. *Soit $\mathbb{K}$ un corps et S un ensemble. Un $\mathbb{K}$-espace vectoriel libre sur S est un $\mathbb{K}$ espace vectoriel équipé d'une base indexée par S. En particulier tout $\mathbb{K}$-espace vectoriel est libre.*

Remarque 5.3.5. La proposition précédente reste valable pour les semi-modules libres sur un semi-anneau commutatif. La démonstration ne nécessite aucune modification si ce n'est que nous n'avons pas défini les quotients de monoïdes commutatifs (il y a une soustraction cachée dans la définition des quotients de groupes abéliens). On peut se contenter de la deuxième partie qui est celle qui sert en pratique, ou bien construire ces quotients. Le point clef est que la bonne relation à associer à un sous-monoïde commutatif M' d'un monoïde commutatif M est définie par $x \sim x'$ s'il existe y et y' dans M' tels que $x + y = x' + y'$.

La réciproque dans la proposition précédente fournit nos premiers exemples de modules libres sur un anneau quelconque.

Exemple 5.3.6. Pour tout entier naturel $n > 0$, le module A^n est libre sur $\{1, \dots, n\}$ via la « base canonique » qui envoie chaque i sur le n-uplet e_i dont toutes les composantes sont nulles sauf la i-ème qui vaut un. La propriété universelle de cette base appliquée au module A^p pour un autre entier p fournit une bijection entre les applications A-linéaires de A^n dans A^p et les fonctions de $\{1, \dots, n\}$ dans A^p. Cet ensemble de fonction est lui-même en bijection avec l'ensemble des fonctions de $\{1, \dots, n\} \times \{1, \dots, p\}$ dans A (par décurryfication, comme dans la discussion suivant la définition 3.2.1). On retrouve ainsi la correspondance entre matrices de taille (n, p) à coefficients dans A et applications A-linéaires de A^n dans A^p comme cas particulier de la propriété universelle des modules libres.

L'exemple ci-dessus est limité aux ensembles finis mais il s'agit d'un obstacle psychologique, la démonstration de la proposition suggère la construction générale suivante.

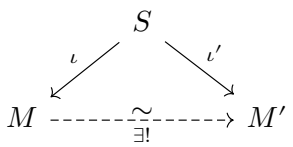
Définition 5.3.7. *Pour tout ensemble S, on note $A[S]$ l'ensemble des fonctions à support fini de S dans A muni de l'addition et la négation ponctuelles et de fonction nulle comme élément neutre. On note δ la fonction de S dans $A[S]$ qui envoie s sur le Dirac en s, c'est à dire la fonction qui envoie s' sur 1 si $s = s'$ et zéro sinon.*

Corollaire 5.3.8. *Pour tout ensemble S, $A[S]$ muni de $\delta : S \to A[S]$ est un A-module libre sur S.*

Démonstration. La vérification des axiomes de module pour $A[S]$ est immédiate à partir des axiomes de module de A. La proposition 5.3.3 assure le reste. □

On appelle souvent $A[S]$ *le* A-module libre sur S ou, de façon plus imagée, « le A-module librement engendré par S ». Comme d'habitude, cet abus est justifié par le fait que la propriété universelle entraîne une caractérisation à unique isomorphisme près :

Lemme 5.3.9. *Soit S un ensemble et A un anneau commutatif. Si (M, ι) et (M', ι') sont deux A-modules libres sur S alors il existe un unique isomorphisme de M vers M' qui fait commuter*



On a aussi l'analogue suivant du corollaire 3.4.5.

Corollaire 5.3.10. *Pour toute application f entre ensembles S et S', il existe une unique application A-linéaire $A[f]\colon A[S] \to A[S']$ telle que*

$$\begin{array}{ccc} S & \xrightarrow{f} & S' \\ {\scriptstyle\delta}\downarrow & & \downarrow{\scriptstyle\delta} \\ A[S] & \xrightarrow[\exists!\, A[f]]{} & A[S'] \end{array}$$

De plus $A[\mathrm{Id}_S] = \mathrm{Id}_{A[S]}$ et, pour toute fonction $g\colon S' \to S''$, $A[g \circ f] = A[g] \circ A[f]$.

Démonstration. Pour obtenir $A[f]$, on applique la propriété universelle de $A[S]$ à l'application $\delta \circ f\colon S \to A[S']$. Comme $\mathrm{Id}_{A[S]} \circ \iota = \iota \circ \mathrm{Id}_S$, l'unicité dans la propriété universelle assure $A[\mathrm{Id}_S] = \mathrm{Id}_{A[S]}$. Pour la formule de composition, on contemple le diagramme suivant

$$\begin{array}{ccccc} S & \xrightarrow{f} & S' & \xrightarrow{g} & S'' \\ \downarrow & & \downarrow & & \downarrow \\ A[S] & \xrightarrow[A[f]]{} & A[S'] & \xrightarrow[A[g]]{} & A[S''] \end{array}$$

Comme les deux carrés commutent, le grand rectangle commute. L'unicité dans la définition de $A[g \circ f]$ assure alors que $A[g \circ f] = A[g] \circ A[f]$. □

Corollaire 5.3.11. *Pour tout ensemble S, il existe un groupe abélien libre sur S et un monoïde commutatif libre sur S.*

Démonstration. Il suffit d'appliquer les résultats de cette section à $A = \mathbb{Z}$ et à $A = \mathbb{N}$ respectivement, en utilisant la remarque 5.3.5 pour ce dernier cas. □

Montrons maintenant que la liberté est liée à la question de l'existence de supplémentaires.

Corollaire 5.3.12. *Soit*

$$0 \longrightarrow M \xrightarrow{i} N \xrightarrow{p} Q \longrightarrow 0$$

une suite exacte courte de A-modules. Si Q est libre alors la suite est scindée.

Démonstration. D'après le lemme 5.2.9, il suffit de construire une section de p.

$$\begin{array}{ccccccccc} & & & & & & S & & \\ & & & & & {\scriptstyle\sigma}\swarrow & \downarrow{\scriptstyle\iota} & & \\ 0 & \longrightarrow & M & \xrightarrow{i} & N & \underset{s}{\overset{p}{\rightleftarrows}} & Q & \longrightarrow & 0 \end{array}$$

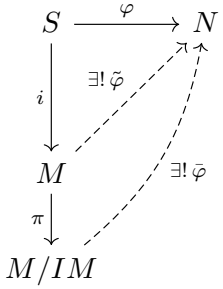
Par hypothèse, il existe un ensemble S et $\iota\colon S \to Q$ tel que (Q, ι) est libre sur S. Comme p est surjectif, on obtient une fonction $\sigma : S \to N$ telle que $p \circ \sigma = \iota$. La propriété universelle de Q donne $s\colon Q \to N$ telle que $\sigma = s \circ \iota$. On a $p \circ s \circ \iota = p \circ \sigma = \iota$ et, comme $i(S)$ engendre Q et que $p \circ s$ est linéaire, $p \circ s = \mathrm{Id}_Q$. □

5.3.2 Rang d'un module libre

On veut maintenant étendre la théorie de la dimension des espaces vectoriels aux modules libres généraux en se ramenant au cas déjà connu. L'ingrédient crucial est la proposition 4.2.14 qui affirme que tout anneau commutatif non trivial admet un quotient qui est un corps, ce qui permettra de transformer les anneaux en corps. Pour faire suivre les modules libres, on utilisera le lemme suivant.

Lemme 5.3.13. *Soit A un anneau commutatif, $I \lhd A$ et S un ensemble. Pour tout A-module (M, i) libre sur S, $(M/IM, \pi \circ i)$ est libre sur A/I.*

Démonstration. Soit N un A/I module et $\varphi : S \to N$ une fonction. Comme N est un A/I-module, c'est aussi un A-module. Par liberté de (M, i), φ s'étend en application A-linéaire $\tilde{\varphi} : M \to N$.



Pour montrer que $\tilde{\varphi}$ descend à M/IM, il suffit de montrer que, pour tout $i \in I$ et $m \in M$, $\tilde{\varphi}(im) = 0$ (car IM est engendré par ces éléments). Or $\tilde{\varphi}$ est A-linéaire donc $\tilde{\varphi}(im) = i\tilde{\varphi}(m)$ qui est nul car l'action scalaire de A sur N passe par A/I. L'unicité s'obtient comme d'habitude en composant les unicités des deux propriétés universelles utilisées (voir par exemple la démonstration de la proposition 3.7.3). □

Voici maintenant l'extension promise de la théorie de la dimension.

Proposition 5.3.14. *Soit A un anneau commutatif non trivial et S et S' deux ensembles. Deux A-modules libres (M, i) et (M', i') sur S et S' respectivement sont isomorphes si et seulement si S et S' ont même cardinal.*

Plus généralement s'il existe une application linéaire injective (resp. surjective) de M dans M' alors $\sharp S \leqslant \sharp S'$ (resp. $\sharp S \geqslant \sharp S'$).

Démonstration. Si $\varphi : S \to S'$ est une bijection alors on obtient un isomorphisme entre M et M' par le corollaire 5.3.10.

Réciproquement supposons que $\varphi : M \to M'$ soit un isomorphisme linéaire. Soit I un idéal maximal de A fournit par la proposition 4.2.14, de sorte que A/I est un corps d'après le lemme 4.2.12. Le lemme 5.1.16 fournit un isomorphisme A/I-linéaire $\bar{\varphi} : M/IM \to M'/IM'$. Or le lemme 5.3.13 assure que M/IM et M'/IM' sont des A/I-modules libres sur S et S'. Le corollaire 5.3.4 assure donc que ce sont des A/I-espaces vectoriels de dimension $\sharp S$ et $\sharp S'$. On conclut par le théorème de la dimension.

De même si φ est injective (resp. surjective) alors $\bar{\varphi}$ est injective (resp. surjective) et on conclut encore par la théorie de la dimension. □

On peut définir la notion de module libre sur un anneau non commutatif mais la proposition précédente n'est plus plus vraie en général. Par exemple on peut montrer que les « matrices » à coefficients dans un anneau commutatif ayant une infinité dénombrable de lignes et de colonnes mais dont chaque colonne n'a qu'un nombre fini d'entrées non nulles forment un anneau A qui vérifie que toutes les puissances finies non nulles de A sont deux à deux isomorphes en tant que A-modules.

Définition 5.3.15. *Le* rang *d'un module libre est le cardinal commun assuré par la proposition précédente. On ne l'appelle pas « dimension » pour prévenir tout optimisme excessif.*

Exemple 5.3.16. Soit A un anneau commutatif intègre. Les sous-A-modules libres de rang 1 dans A sont exactement les idéaux principaux non nuls de A. L'idéal nul est le seul sous-module de rang 0 dans A.

5.3.3 Torsion

Le but de cette section est d'introduire une obstruction à la liberté appelée torsion. On verra dans la section suivante que la torsion est la seule obstruction à la liberté pour les modules ayant une famille génératrice finie lorsque l'anneau des scalaires est principal.

On commence par remarquer que si un anneau commutatif n'est pas intègre alors il n'admet aucun module libre non nul. Dans toute cette section, A est un anneau commutatif et *intègre*.

Définition 5.3.17. *Soit A un anneau commutatif et intègre. Soit M un A-module. La partie de* torsion *de M est*

$$\mathrm{Tor}(M) = \{m \in M \mid \exists a \in A \setminus \{0\}, am = 0\}.$$

Les éléments de $\mathrm{Tor}(M)$ *sont appelés* éléments de torsion *de M. On dit que M est de torsion si* $\mathrm{Tor}(M) = M$ *et sans torsion si* $\mathrm{Tor}(M) = 0$.

En cas de doute sur l'anneau sous-jacent, on utilise la notation $\mathrm{Tor}_A(M)$. Par exemple on peut toujours considérer $\mathrm{Tor}_{\mathbb{Z}}(M)$. L'intégrité ne sert à rien dans la définition ci-dessus mais elle est cruciale dans le lemme suivant.

Lemme 5.3.18. *Soit A un anneau commutatif et intègre. Soit M un A-module.*

- $\mathrm{Tor}(M)$ *est un sous-module de M.*
- $M/\mathrm{Tor}(M)$ *est sans torsion.*

Démonstration. Montrons d'abord que $\mathrm{Tor}(M)$ est un sous-groupe de M en appliquant le lemme 3.1.8. Il contient 0 car A est intègre donc non trivial donc $1 \neq 0$ dans A et $1 \cdot 0 = 0$ dans M. Soit x et y dans $\mathrm{Tor}(M)$ et soit a et b non nuls dans A tels que $ax = by = 0$. On a $ab(x - y) = 0$ et $ab \neq 0$ car A est intègre, donc $x - y$ est dans $\mathrm{Tor}(M)$. Montrons maintenant que $\mathrm{Tor}(M)$ est stable sous l'action scalaire. Soit x dans $\mathrm{Tor}(M)$ et a dans A non nul tel que $ax = 0$. Pour tout b dans A on a $a(bx) = b(ax) = 0$ donc bx est dans $\mathrm{Tor}(M)$.

Montrons maintenant que $M/\mathrm{Tor}(M)$ est sans torsion. Soit x dans M et a dans A non nul tel que $a\pi(x) = 0$. Comme π est A-linéaire, $\pi(ax) = a\pi(x) = 0$, donc ax est dans $\mathrm{Tor}(M)$. On obtient donc b non nul dans A tel que $bax = 0$. Comme A est intègre, ab n'est pas nul donc x est de torsion et donc $\pi(x) = 0$. □

Voyons maintenant que, comme promis, la torsion est une obstruction à la liberté.

Lemme 5.3.19. *Soit M un module sur un anneau commutatif intègre A. Si M est libre alors il est sans torsion.*

Démonstration. Soit S un ensemble et $i : S \to M$ tel que (M, i) est libre sur S. Soit m non nul dans M. On sait par la proposition 5.3.3 que $m = \sum_s a_s i(s)$ pour une fonction $a : S \to A$ à support fini et non nulle. Soit s_0 tel que $a_{s_0} \neq 0$. Soit b dans A tel que $bm = 0$. On a donc $\sum_s ba_s i(s) = 0$. Par unicité de ces décompositions, tous les ba_s sont nuls. En particulier $ba_{s_0} = 0$. Comme A est intègre et $a_{s_0} \neq 0$, on en déduit que b est nul. □

5.4 Modules de type fini

Le but de cette dernière section est de voir ce qu'il reste du théorème de structure des espaces vectoriels. Nous avons déjà vu qu'il est inutile d'espérer l'existence de bases car la torsion est une obstruction. Le mieux qu'on puisse espérer pour un module M est donc de trouver un supplémentaire libre au sous-module de torsion $\mathrm{Tor}(M)$. C'est effectivement ce que donnera le théorème 5.4.9 dans le cas des modules ayant une partie génératrice finie sur un anneau principal. Dans le cas des $\mathbb{Z}$-modules nous donnerons une classification complète.

5.4.1 Cas général

On commence par quelques points qui ne nécessitent aucune hypothèse sur A.

Définition 5.4.1. *Un module M sur un anneau commutatif A est de* type fini *s'il est engendré par une partie finie.*

Lemme 5.4.2. *Un module M sur un anneau commutatif A est de type fini si et seulement si il est quotient d'un A-module libre sur un ensemble fini.*

Démonstration. Supposons M de type fini. Soit $S \subset M$ un ensemble fini qui engendre M. La propriété universelle du module libre $A[S]$ étend l'inclusion en application linéaire dont l'image contient S donc contient M.

Réciproquement si (N, i) est libre sur S fini et $\pi : N \to M$ est surjective alors $\pi(i(S))$ est fini et engendre M. □

Le lemme suivant garantit qu'un module libre de type fini est bien ce qu'on croit et permet de parler indifféremment de module libre de type fini et de module libre de rang fini.

Lemme 5.4.3. *Un module libre est de type fini si et seulement si son rang est fini.*

Démonstration. Soit $(M, i : S \to M)$ un module libre sur un anneau commutatif A. Si M est de rang fini alors il est clairement de type fini. Réciproquement, supposons que $(N, j : T \to N)$ est un module libre sur un ensemble fini T et qu'on a une application linéaire surjective $\pi : N \to M$. La proposition 5.3.14 assure $\sharp S \leqslant \sharp T$ donc S est fini. □

5.4.2 Cas des anneaux principaux

Proposition 5.4.4. *Soit A un anneau commutatif principal et M un A-module libre de type fini. Tout sous-module M' de M est libre de rang inférieur à celui de M.*

Démonstration. On raisonne par récurrence sur le rang r de M. Le cas $r = 0$ correspond au module trivial dont tous les sous-modules sont triviaux donc libres de rang zéro. Supposons le résultat démontré pour tous le modules libres de rang r. Soit M un A-module libre de rang $r + 1$. On considère une suite exacte de A-modules

$$0 \longrightarrow A \xrightarrow{i} M \xrightarrow{p} Q \longrightarrow 0$$

où Q est libre de rang r. Une telle suite existe car, d'après la proposition 5.3.14, tous les A-modules libres de même rang sont isomorphes donc on peut supposer que $M = A^{r+1}$ et choisir $i : a \to (a, 0, \dots, 0)$ et p la projection sur les r dernières coordonnées.

Soit N un sous-module de M. Le sous-module $p(N)$ est libre de range au plus r dans Q par hypothèse de récurrence. Le sous-module $E = i^{-1}(N)$ dans A est un idéal de A. Comme A est principal, E est libre de rang au plus un. On a la suite exacte

$$0 \longrightarrow E \xrightarrow{i|_E} N \xrightarrow{p|_N} p(N) \longrightarrow 0$$

Comme $p(N)$ est libre, le corollaire 5.3.12 assure que cette suite est scindée donc N est isomorphe à $E \oplus p(N)$ qui est libre de rang au plus $r + 1$. □

Remarque 5.4.5. Dans le proposition précédente, l'hypothèse que A est principal est indispensable. D'abord l'intégrité est indispensable car tout diviseur de zéro non nul engendre un sous-module de A qui n'est pas libre. Supposons donc que A est intègre. A est un A-module libre de rang 1 et ses sous-module sont ses idéaux, qui sont libres de rang inférieur à 1 précisément s'ils sont principaux.

Corollaire 5.4.6. *Soit M un module de type fini sur un anneau commutatif principal A. Tout sous-module de M est de type fini.*

Démonstration. Par hypothèse, on obtient un module libre L de rang fini et une application linéaire surjective $\pi : L \to M$. Soit N un sous-module de M. Le sous-module $\pi^{-1}(N)$ est libre de rang fini d'après la proposition 5.4.4. Comme π est surjective, $N = \pi(\pi^{-1}(N))$ donc N est bien l'image d'un module libre de rang fini. □

Remarque 5.4.7. Là encore l'hypothèse que A est principal est indispensable et, comme pour tous les énoncés ayant cette hypothèse dans cette section, on peut trouver des contre-exemples en rang 1, c'est à dire avec $M = A$. On verra dans le chapitre suivant comment construire des anneaux de polynômes avec un ensemble quelconque d'indéterminées, par exemple $\mathbb{Z}[X_1, X_2, \dots]$ avec une infinité dénombrable d'indéterminées. Dans cet anneau, l'idéal engendré par tous les X_i n'est pas de type fini.

On peut maintenant démontrer comme promis que, dans le cas des modules de type finis sur un anneau principal, la torsion est l'unique obstruction à la liberté.

Proposition 5.4.8. *Soit M un module sur un anneau commutatif principal A. Si M est de type fini alors il est libre si et seulement si il est sans torsion.*

Démonstration. Le lemme 5.3.19 assure déjà la première implication, sans hypothèse sur A. Supposons donc que M est sans torsion et montrons qu'il est libre. Par hypothèse on obtient un ensemble fini S et une fonction $\iota : S \to M$ dont l'image engendre M (dans toute cette démonstration, « engendre » est toujours entendu dans le contexte des sous-modules). Soit T une partie maximale de S telle que $\iota(T)$ est libre, c'est à dire que $(\langle \iota(T)\rangle, \iota)$ est libre sur T. Une telle partie existe car $(\langle \iota(\varnothing)\rangle, \iota)$ est libre et S est fini. On pose $N = \langle \iota(T)\rangle$.

Montrons que, pour tout s dans S, il existe a dans $A\setminus\{0\}$ tel que $a\iota(s)$ est dans N. Soit s dans S. Si s est dans T on utilise $a = 1$ (qui est bien non nul car A est principal donc non trivial). Sinon, a provient de la maximalité de T. En effet si aucun a ne convenait alors $\iota(s)$ engendrerait un sous-module libre de rang un qui serait en somme directe avec N.

On fixe un élément a_s comme ci-dessus pour tout s et on pose $a = \prod_s a_s$. Comme A est intègre, a n'est pas nul. Comme M est sans torsion, l'homothétie $\varphi : m \to am$ est une application linéaire injective de M dans M. Montrons que son image est contenue dans N. Comme cette image est un sous-module et que M est engendré par $\iota(S)$, il suffit de montrer que $\varphi(\iota(S))$ est dans N. Soit s dans S. On a

$$\varphi(\iota(s)) = a\iota(s) = \left(\prod_{s'\neq s} a_{s'}\right) \underbrace{a_s\iota(s)}_{\in N} \in N.$$

Ainsi φ est un isomorphisme de M sur un sous-module $\varphi(M)$ de N. Comme N est libre, la proposition 5.4.4 assure que $\varphi(M)$ est libre donc M l'est aussi. □

Tous les énoncés précédents se combinent maintenant pour donner le théorème de classification grossière promis en introduction de cette section.

Théorème 5.4.9. *Soit A un anneau commutatif principal et M un A-module. Si M est de type fini alors*

$$M \simeq (M/\operatorname{Tor}(M)) \oplus \operatorname{Tor}(M).$$

De plus $M/\operatorname{Tor}(M)$ est sans torsion et les deux morceaux sont de type fini.

Démonstration. On a la suite exacte

$$0 \longrightarrow \operatorname{Tor}(M) \hookrightarrow M \xrightarrow{\pi} M/\operatorname{Tor}(M) \longrightarrow 0$$

et $M/\operatorname{Tor}(M)$ est sans torsion d'après le lemme 5.3.18. La proposition 5.4.8 assure donc que $M/\operatorname{Tor}(M)$ est libre. Le corollaire 5.3.12 en déduit que la suite est scindée donc on a bien l'isomorphisme annoncé. Enfin $M/\operatorname{Tor}(M)$ est de type fini car il est quotient de M qui est de type fini et le corollaire 5.4.6 montre que $\operatorname{Tor}(M)$ est de type fini, car c'est un sous-module de M. □

Dans le théorème précédent il faut prendre garde au fait que l'image de $M/\operatorname{Tor}(M)$ dans M n'est pas unique. Il n'y a pas de supplémentaire naturel au sous-module $\operatorname{Tor}(M)$. Le théorème ne fait qu'énoncer l'existence d'un supplémentaire et le fait qu'un tel supplémentaire est libre et de type fini.

Définition 5.4.10. *Le* rang $\operatorname{rg}(M)$ *d'un module M de type fini sur un anneau commutatif principal est le rang de $M/\operatorname{Tor}(M)$ (qui est bien libre de type fini d'après le théorème précédent). On a donc $M/\operatorname{Tor}(M) \simeq A^{\operatorname{rg}(M)}$*

5.4.3 Cas des groupes abéliens

On se concentre maintenant sur le cas des groupes abéliens (les $\mathbb{Z}$-modules). Vu le théorème précédent, la compréhension des groupes abéliens de type finis est complètement ramenée à celle des groupes abéliens finis. Cette question sera complètement réglée par le théorème ci-dessous, dont la démonstration occupe toute la fin de ce chapitre.

Théorème 5.4.11. *Soit G un groupe abélien fini non trivial. Il existe un unique entier $N > 0$ et un unique N-uplet $(d_1, \dots, d_N)$ d'entiers tels que*

- $$G \simeq \prod_{i=1}^{N} \mathbb{Z}/d_i\mathbb{Z}$$
- $\forall i,\ d_i > 1$
- $\forall i < N,\ d_i \mid d_{i+1}$

Les entiers d_i sont appelés les facteurs invariants de G.

Dans toute la fin de ce chapitre, G est un groupe abélien fini (noté additivement). On cherche à comprendre la structure de G par récurrence forte sur le cardinal de G. Le point clef est de démontrer, pour tout G non trivial, l'existence d'un sous-groupe cyclique $H \leqslant G$ non trivial et tel que $0 \to H \to G \to G/H \to 0$ est scindée. On peut alors appliquer l'hypothèse de récurrence à G/H.

Cette stratégie ne fonctionne pas du tout dans le cas des modules de type fini sur un anneau principal quelconque parce que les modules de type fini de torsion n'ont aucune raison d'être de cardinal fini en général. Le théorème de classification s'étend tout de même mais la démonstration est bien plus technique.

Pour identifier un sous-groupe H convenable, on observe l'énoncé du théorème. On remarque que l'entier d_N est le plus petit entier n qui vérifie $\forall x \in G, nx = 0$ (à cause des relations de divisibilité). Cela motive la définition suivante, et les lemmes qui suivront.

Définition 5.4.12. *Soit G un groupe abélien fini. L'*exposant *de G est l'entier*

$$\exp(G) = \min\{n \in \mathbb{N} \setminus \{0\} \mid \forall x \in G, nx = 0\}.$$

Il n'est pas complètement évident que l'exposant d'un groupe fini est fini mais cela découle du lemme suivant. La notation exp n'est pas vraiment standard (il n'y a pas de notation standard). On rappelle que l'ordre d'un élément x est noté $o(x)$.

Lemme 5.4.13. *Soit G un groupe abélien fini.*

- $\exp(G) = \mathrm{ppcm}_{x \in G}\, o(x)$. *En particulier* $\exp(G) \mid \sharp G$.
- $\exp(G) = \max_{x \in G} o(x)$. *En particulier il existe x tel que $o(x) = \exp(G)$.*

Démonstration. Soit x dans G. L'application de $\mathbb{Z}$ dans G qui envoie n sur nx est un morphisme, son noyau est donc un sous-groupe de $\mathbb{Z}$ donc de la forme $m\mathbb{Z}$. Ce morphisme a pour image le sous-groupe engendré par x donc $\mathbb{Z}/m\mathbb{Z} \simeq \langle x \rangle$ et donc $m = o(x)$, par définition de $o(x)$. De plus $nx = 0$ si et seulement si n est dans $o(x)\mathbb{Z}$, c'est à dire que n est un multiple de $o(x)$. Ainsi $\exp(G)$ est bien le plus petit multiple commun à tous les $o(x)$. Or on sait d'après le théorème de Lagrange (corollaire 3.2.13) que tous les $o(x)$ divisent $\sharp G$ donc leur ppcm aussi.

Montrons maintenant le deuxième point. Comme $\exp(G)$ et les $o(x)$ sont positifs, la divisibilité entraîne l'inégalité donc le premier point assure que $o(x) \leqslant \exp(G)$ pour tout x. Montrons qu'il existe x tel que $o(x) = \exp(G)$. Soit $\mathcal{P}$ l'ensemble des nombres premiers qui interviennent dans la décomposition en produit de facteurs premier de $\exp(G)$ (qui est non nul par définition). On a $\exp(G) = \prod_{p\in\mathcal{P}} p^{r_p}$ et le premier point assure que, pour tout $p \in \mathcal{P}$, $r_p = \max_{x\in G} v_p(o(x))$ où $v_p(n)$ désigne le plus grand entier k tel que p^k divise n. Pour chaque p on fixe un x_p tel que $r_p = v_p(o(x_p))$. Ainsi $o(x_p) = p^{r_p}q_p$ pour un entier q_p premier à p. On déduit $o(q_p x_p) = p^{r_p}$. On pose $x = \sum_p q_p x_p$. Comme les p^{r_p} sont premiers entre eux, x est bien d'ordre $\exp(G)$. □

Corollaire 5.4.14. *Un groupe abélien fini G est cyclique si et seulement si $\exp(G) = \sharp G$.*

Le sous-groupe cyclique apparaissant dans la stratégie de démonstration du théorème sera engendré par un élément réalisant l'exposant du groupe. Le lemme suivant permettra de montrer que la suite exacte associée est scindée.

Lemme 5.4.15. *Soit G un groupe abélien fini et x un élément de G tel que $o(x) = \exp(G)$. La suite exacte*

$$0 \longrightarrow \langle x\rangle \xhookrightarrow{i} G \xrightarrow{\pi} G/\langle x\rangle \longrightarrow 0$$

est scindée.

Remarque 5.4.16. Dans le lemme précédent, l'hypothèse $o(x) = \exp(G)$ est cruciale. Par exemple, dans $G = \mathbb{Z}/4\mathbb{Z}$, $x = 2$ est d'ordre 2 qui n'est pas $\exp(G)$, et la suite exacte correspondante n'est pas scindée car G n'est pas isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

Démonstration. D'après le lemme 5.2.9, il suffit de trouver une rétraction de l'inclusion i de $\langle x\rangle$, c'est-à-dire un morphisme de G dans $\langle x\rangle$ qui est l'identité sur $\langle x\rangle$. Plus progressivement, montrons par récurrence que, pour tout sous-groupe H de G contenant $\langle x\rangle$, il existe une rétraction de l'inclusion de $\langle x\rangle$ dans H. Le cas $H = \langle x\rangle$ est clair. Il suffit donc de considérer un sous-groupe H contenant $\langle x\rangle$ et muni d'une rétraction $r : H \to \langle x\rangle$, et un élément y de G qui n'est dans H et de montrer que r s'étend en morphisme r' de $H + \langle y\rangle$ dans $\langle x\rangle$. Un tel morphisme sera automatiquement une rétraction de i puisque que cette propriété ne dépend que de la restriction de r' à $\langle x\rangle$ et que r' étend r.

Le groupe $\langle y\rangle \cap H$ est un sous-groupe du groupe cyclique fini $\langle y\rangle$ donc il est engendré par βy pour certain β divisant $o(y)$. Comme βy est dans H, on peut considérer $r(\beta y)$, qui est de la forme αx pour un certain α.

L'affirmation clef où l'hypothèse $o(x) = \exp(G)$ va servir est que β divise α. Soit β' l'entier tel que $o(y) = \beta'\beta$. On note que β' ne saurait être nul puisqu'on aurait sinon $o(y) = 0$, ce qui contredirait que y n'est pas dans H. On calcule $\beta'\alpha x = \beta' r(\beta y) = r(\beta'\beta y) = r(o(y)y) = r(0) = 0$. Donc $o(x) \mid \beta'\alpha$. Or $o(x) = \exp(G)$ donc $\beta'\beta$, qui vaut $o(b)$, divise $o(x)$. Ainsi $\beta'\beta \mid \beta'\alpha$. Comme β' n'est pas nul, on obtient bien $\beta \mid \alpha$.

Soit n l'entier tel que $\alpha = n\beta$. Le sous-groupe $H + \langle y\rangle$ est un quotient de $H \times \mathbb{Z}$ via le morphisme $\pi : (h, l) \mapsto h + ly$. Soit $\varphi : H \times \mathbb{Z} \to \langle x\rangle$ le morphisme qui envoie (h, l) sur $r(h) + lnx$. Montrons que φ descend au quotient en extension de r. La condition de descente est $\ker\pi \subset \ker\varphi$. Soit (h, l) tel que $h + ly = 0$. En particulier ly est dans H donc l s'écrit sous la forme $k\beta$. On calcule $\varphi(h, l) = r(h) + lnx = r(-k\beta y) + k\beta nx =$

$-k\alpha x + k\alpha x = 0$. On obtient donc bien un morphisme $r' : H + \langle y \rangle \to \langle x \rangle$. Montrons qu'il étend r. Soit h dans H. On a $r'(h) = r'(\pi(h,0)) = \varphi(h,0) = r(h)$. □

Comme le verra plus loin, le lemme précédent est suffisant pour décomposer tout groupe abélien fini comme produit de groupes cycliques. Pour étudier l'unicité dans la décomposition obtenue, on va utiliser le lemme suivant (qui n'est pas spécifique aux groupes abéliens).

Lemme 5.4.17. *Soit G, H et H' des groupes finis. Si $G \times H$ est isomorphe à $G \times H'$ alors H est isomorphe à H'.*

Démonstration. Pour tous groupes finis G_1 et G_2, on note $m(G_1, G_2)$ le nombre de morphismes de G_1 dans G_2 et $i(G_1, G_2)$ le nombre de ceux qui sont injectifs. On a $m(G_1, G_2) \geqslant 1$ puisque l'application constante de valeur 1 est un morphisme. Pour calculer $m(G_1, G_2)$ on regroupe les morphismes par noyau. Le premier théorème d'isomorphisme (corollaire 3.3.11) montre que

$$m(G_1, G_2) = \sum_{N \lhd G_1} i(G_1/N, G_2) \qquad (\star)$$

On revient maintenant au lemme. Supposons que $G \times H$ et $G \times H'$ sont isomorphes. Montrons d'abord que, pour tout groupe fini L, $m(L, H) = m(L, H')$. Soit L un groupe fini. La propriété universelle du produit de groupe (exemple 3.1.16) et l'hypothèse d'isomorphisme assurent que

$$m(L, G)m(L, H) = m(L, G \times H) = m(L, G \times H') = m(L, G)m(L, H').$$

Comme $m(L, G)$ est non nul, on obtient $m(L, H) = m(L, H')$

Montrons que, pour tout groupe fini L, $i(L, H) = i(L, H')$. On raisonne par récurrence forte sur $\sharp L$. Le cas $\sharp L = 0$ est trivial car tout groupe est de cardinal au moins un. Soit L un groupe fini. Supposons la formule établie pour tous les groupes de cardinal strictement inférieur à $\sharp L$. On utilise l'équation $(\star)$ pour (L, H) et (L, H'), en mettant à part les morphismes injectifs, qui correspondant au sous-groupe distingué trivial :

$$m(L, H) = i(L, H) + \sum_{\substack{N \lhd L \\ N \neq 1}} i(L/N, H)$$

et

$$m(L, H') = i(L, H') + \sum_{\substack{N \lhd L \\ N \neq 1}} i(L/N, H').$$

On a vu que les membres de gauche des ces deux équations sont égaux. Chaque terme de la première somme est égal au terme correspondant de la seconde somme par hypothèse de récurrence puisque les sous-groupes N sont non triviaux donc $\sharp L/N < \sharp L$. Ainsi on a bien $i(L, H) = i(L, H')$.

En particulier $i(H, H') = i(H', H') \geqslant 1$ donc il existe un morphisme injectif $\varphi : H \to H'$. Or H et H' sont finis et de même cardinal puisque $\sharp G \sharp H = \sharp(G \times H) = \sharp(G \times H') = \sharp G \sharp H'$ et $\sharp G \geqslant 1$. Donc φ est bijective. C'est donc un isomorphisme d'après le lemme 3.1.3. □

Nous avons maintenant tous les ingrédients pour démontrer le théorème de classification.

Démonstration du théorème 5.4.11. On démontre le théorème par récurrence forte sur $\sharp G$. Supposons que G n'est pas trivial et que le résultat est établi pour tous les groupes abéliens de cardinal strictement inférieur à $\sharp G$.

Le lemme 5.4.13 fournit x_0 tel que $o(x_0) = \exp(G)$. Comme G n'est pas trivial, $\exp(G) > 1$ donc $\sharp(G/\langle x_0\rangle) < \sharp G$. L'hypothèse de récurrence appliquée à $G/\langle x_0\rangle$ fournit donc un entier N' et des entiers d_1, ..., $d_{N'}$ comme dans l'énoncé. Le lemme 5.4.15 assure que la suite exacte associée à ce quotient est scindée donc

$$G \simeq \left(\prod_{i=1}^{N'} \mathbb{Z}/d_i\mathbb{Z}\right) \times \langle x_0\rangle \simeq \left(\prod_{i=1}^{N'} \mathbb{Z}/d_i\mathbb{Z}\right) \times \mathbb{Z}/\exp(G)\mathbb{Z}$$

On pose $N = N' + 1$ et $d_N = \exp(G)$. Il ne reste qu'à expliquer pourquoi $d_{N'} \mid d_N$. Soit y un générateur du facteur $\mathbb{Z}/d_{N'}\mathbb{Z}$. On a $o(y) = d_{N'}$ et $o(y) \mid \exp(G)$ donc $d_{N'} \mid d_N$.

Montrons maintenant l'unicité. Supposons qu'on ait deux telles décompositions

$$G \simeq \prod_{i=1}^{N} \mathbb{Z}/d_i\mathbb{Z} \simeq \prod_{j=1}^{N'} \mathbb{Z}/d'_j\mathbb{Z}.$$

On a alors $d_N = \exp(G) = d'_{N'}$. Le lemme 5.4.17 donne donc un isomorphisme

$$\prod_{i=1}^{N-1} \mathbb{Z}/d_i\mathbb{Z} \simeq \prod_{j=1}^{N'-1} \mathbb{Z}/d'_j\mathbb{Z}.$$

L'unicité dans l'hypothèse de récurrence assure que $N-1 = N'-1$ et que $d_i = d'_i$ pour tout $i < N$. □

Remarque 5.4.18. Une fois connu le théorème de classification, le théorème 4.3.30 des restes Chinois montre qu'il existe aussi une décomposition de la forme

$$G \simeq \prod_{i=1}^{N} \mathbb{Z}/p_i^{\alpha_i}\mathbb{Z}$$

où les p_i sont des nombres premiers (pas distincts en général). On peut montrer que cette décomposition est unique modulo permutation. Il existe un algorithme pour passer de la décomposition en termes de nombres premiers aux facteurs invariants. Le plus simple est de le voir fonctionner sur un exemple. Considérons deux nombres premiers p et q et le groupe

$$G = \mathbb{Z}/p\mathbb{Z} \times (\mathbb{Z}/p^2\mathbb{Z})^2 \times (\mathbb{Z}/p^3\mathbb{Z}) \times \mathbb{Z}/q^4\mathbb{Z} \times \mathbb{Z}/q^6\mathbb{Z}$$

On écrit les puissances de p et q en lignes croissantes avec répétitions, alignées à droite puis on fait les produits de chaque colonne pour obtenir les d_i.

p	p^2	p^2	p^3
		q^4	q^6
p	p^2	p^2q^4	p^3q^6

La décomposition en facteurs invariants de G est

$$G \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p^2q^4\mathbb{Z} \times \mathbb{Z}/p^3q^6\mathbb{Z}.$$

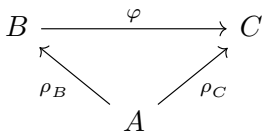
6 Algèbres

6.1 Définitions, morphismes et sous-objets

La dernière structure algébrique étudiée dans ce cours est celle d'algèbre. Étant donné un anneau commutatif A, une A-algèbre est un ensemble muni à la fois d'une structure de A-module et d'une multiplication, avec des conditions de compatibilité. Comme dans le cas des modules, on donne une définition très concise.

Définition 6.1.1. *Soit A un anneau commutatif, une* A-algèbre *(associative unitaire) est un anneau B muni d'un morphisme $\rho\colon A \to B$ dont l'image commute avec tous les éléments de B.*

Un morphisme *de A-algèbres entre (B, ρ_B) et (C, ρ_C) est un morphisme d'anneau $\varphi\colon B \to C$ tel que $\varphi \circ \rho_B = \rho_C$.*



En particulier toute A-algèbre est un A-module, en faisant agir a par multiplication par $\rho(a)$, et une application entre deux A-algèbres est un morphisme de A-algèbre si et seulement si c'est un morphisme d'anneaux qui est A-linéaire.

Remarque 6.1.2. On peut définir des algèbres qui ne sont pas unitaires ou mêmes des algèbres qui ne sont pas associatives mais la définition est moins compacte (on ne peut pas se reposer sur la théorie des anneaux) et ces objets plus généraux n'apparaitront pas du tout dans ce cours.

Exemple 6.1.3. D'après le lemme 4.1.11, tout anneau est, de façon unique, une $\mathbb{Z}$-algèbre.

Soit $\mathbb{K}$ un corps. L'anneau $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$, muni de l'inclusion usuelle de $\mathbb{K}$, est une $\mathbb{K}$-algèbre commutative. On généralisera cet exemple dans la suite.

Soit A un anneau et M un A-module. L'anneau $\mathrm{End}_A(M)$ des endomorphismes de M muni de $\lambda \mapsto \lambda\,\mathrm{Id}$ est une A-algèbre. Elle n'est commutative que si M est de dimension au plus un. En particulier A lui-même est une A-algèbre (dans ce cas $\rho = \mathrm{Id}$).

Dans les exemples précédents, la structure d'algèbre est claire et en pratique l'application ρ est notée par un symbole invisible. C'est le cas pour la plupart des algèbres.

Définition 6.1.4. *Une* sous-algèbre *d'une A-algèbre B est un sous-anneau de B qui est aussi un sous A-module. Autrement dit, c'est une partie de B qui contient 0 et 1 et est stable par addition, passage à l'opposé, multiplication et multiplication par les images des éléments de A.*

Lemme 6.1.5. *L'image d'une sous-algèbre par un morphisme d'algèbres est une sous-algèbre. La préimage d'une sous-algèbre par un morphisme d'algèbre est une sous-algèbre.*

Une intersection de sous-algèbres est une sous-algèbre. En particulier on a une notion de sous-algèbre engendrée *par une partie, avec toutes les propriétés habituelles pour les sous-objets engendrés.*

Démonstration. Ce lemme découle immédiatement de la combinaison des lemmes 4.1.15 et 5.1.10. □

6.2 Algèbre d'un monoïde

Dans cette section on introduit une algèbre associée à tout anneau commutatif et tout monoïde. Cette construction est importante en théorie des représentations mais pour nous il s'agira simplement d'une étape intermédiaire sur la route des algèbres de polynômes.

Lemme 6.2.1. *Soit A un anneau commutatif et M un monoïde. Il existe une unique structure de A-algèbre sur le module libre $A[M]$ telle que l'application $M \to A[M]$ soit un morphisme de monoïdes. L'application de A dans $A[M]$ est $a \mapsto a1_M$ et la multiplication est donnée par*

$$\left(\sum_m a_m m\right)\left(\sum_{m'} b_{m'} m'\right) = \sum_{m,m'} a_m b_{m'} mm'.$$

Si M est commutatif alors $A[M]$ l'est aussi et la réciproque est vraie dès que A est non trivial.

Cette algèbre est appelé la A-algèbre du monoïde M. *Elle vérifie la propriété universelle suivante. Soit B une A-algèbre et $\varphi\colon M \to (B, \times)$ un morphisme de monoïdes. Il existe un unique morphisme de A-algèbres $\bar{\varphi}\colon A[M] \to B$ qui étende φ.*

Démonstration. On commence par l'unicité. On sait que chaque élément de $A[M]$ s'écrit de façon unique comme somme $\sum_m a_m m$ pour une fonction $a\colon M \to A$ à support fini. Les axiomes d'anneau et de morphisme de monoïde forcent la formule de l'énoncé.

Réciproquement cette formule défini bien une loi de composition interne sur $A[M]$ (la somme du membre de droite est bien finie car $(m, m') \mapsto a_m b_{m'}$ est à support fini). L'inclusion de M dans $A[M]$ est clairement compatible avec cette multiplication et l'élément neutre de M, vu comme élément de $A[M]$, est neutre. On voit aussi que les éléments de A sont envoyés sur des éléments qui commutent avec tous les éléments de $A[M]$.

On peut vérifier directement les axiomes d'anneau sans difficulté (à condition de faire les choses dans un ordre judicieux, en particularité la distributivité avant l'associativité de la multiplication). Mais il est plus facile de passer par la propriété universelle des A-module libre $A[M]$. Soit a dans $A[M]$. L'application de M dans $A[M]$ qui envoie m sur am (défini par la formule ci-dessus) admet une unique extension A-linéaire de $A[M]$ dans $A[M]$ et cette extension est bien donnée par notre formule. Son additivité montre la distributivité à gauche. De même l'additivité de l'extension de $m \mapsto ma$ montre la distributivité à droite. La compatibilité avec la multiplication scalaire de l'extension de

$m \mapsto am$ montre que

$$\begin{aligned} a(bc) &= \sum_{m',m''} (b_{m'}c_{m''})a(m'm'') \\ &= \sum_{m,m',m''} a_m(b_{m'}c_{m''})m(m'm'') \end{aligned}$$

De même en utilisant l'extension de $m \mapsto mc$, on obtient

$$(ab)c = \sum_{m,m',m''} (a_m b_{m'})c_{m''}(mm')m''$$

et on conclut par l'associativité des multiplications dans A et M.

Si M est commutatif, la formule de produit montre que $A[M]$ est commutative. Réciproquement si A est non trivial et $A[M]$ est commutative alors pour tout m et m' dans M on a $mm' = m'm$ dans $A[M]$ et on conclut par injectivité de $M \to A[M]$ (qui provient de l'hypothèse que A n'est pas trivial).

Montrons la propriété universelle. Soit B une A-algèbre. En particulier B est un A-module. Soit $\varphi : M \to B$ un morphisme de monoïdes. La propriété universelle des A-modules libres permet d'étendre φ de façon unique en application A-linéaire $\bar{\varphi} : A[M] \to B$. On a $\bar{\varphi}(1) = \varphi(1) = 1$ donc il ne reste à montrer que la multiplicativité. Soit a et b dans $A[M]$. On calcule en utilisant la linéarité de $\bar{\varphi}$ et le fait qu'elle étend φ :

$$\begin{aligned} \bar{\varphi}(ab) &= \bar{\varphi}\left(\sum_{m,m'} a_m b_{m'} mm'\right) \\ &= \sum_{m,m'} a_m b_{m'} \varphi(mm') \\ &= \sum_{m,m'} a_m b_{m'} \varphi(m)\varphi(m') \\ &= \left(\sum_m a_m \varphi(m)\right)\left(\sum_{m'} b_{m'}\varphi(m')\right) \\ &= \bar{\varphi}(a)\bar{\varphi}(b). \end{aligned}$$

Ainsi $\bar{\varphi}$ est aussi compatible avec la multiplication. □

Exemple 6.2.2. Soit $\mathbb{K}$ un corps. La $\mathbb{K}$-algèbre du monoïde $\mathbb{N}$ est l'algèbre des polynômes à coefficients dans $\mathbb{K}$. L'image d'un entier n dans cette algèbre est notée X^n. Soit E un $\mathbb{K}$-espace vectoriel et u un endomorphisme de E. La $\mathbb{K}$-algèbre du sous-monoïde engendré par u est l'algèbre des polynômes en u, elle est isomorphe au sous-anneau de $\mathrm{End}(E)$ engendré par u et Id. Ces deux exemples seront généralisés dans la section suivante.

6.3 Algèbres de polynômes

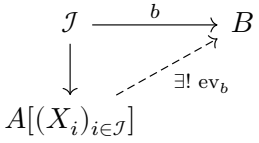
On peut maintenant introduire les algèbres de polynômes qui sont des algèbres commutatives libres et généralisent le cas des polynômes à une indéterminée.

Définition 6.3.1. *Soit A un anneau commutatif et $\mathcal{I}$ un ensemble. L'algèbre des polynômes d'indéterminées indexées par $\mathcal{I}$ à coefficients dans A est la A-algèbre du monoïde commutatif libre sur $\mathcal{I}$. On choisit des symboles, par exemple X_i pour désigner les images des éléments i de $\mathcal{I}$ dans cette algèbre et on note l'algèbre $A[(X_i)_{i\in\mathcal{I}}]$.*

Le cas le plus courant est celui où $\mathcal{I}$ est un ensemble fini mais le cas général n'est pas plus compliqué. Lorsque deux anneaux A et B n'ayant rien à voir entre eux sont en jeu, il est crucial de varier les notations, par exemple pour savoir si X doit être interprété comme un élément de $A[X]$ ou de $B[X]$. Par contre si B est une A-algèbre on conservera souvent la même notation.

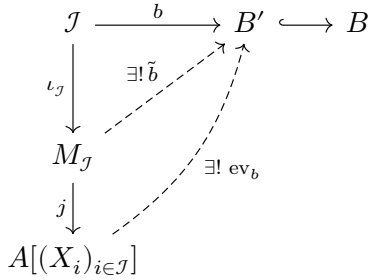
Les propriétés universelles des monoïdes commutatifs libres et des algèbres de monoïdes impliquent bien sûr que deux ensembles isomorphes (c'est à dire de même cardinal) donnent lieu à des algèbres de polynômes isomorphes.

Proposition 6.3.2 (Propriété universelle des algèbres de polynômes)**.** *Soit A un anneau commutatif, B une A-algèbre et $\mathcal{I}$ un ensemble. Soit $b\colon \mathcal{I}\to B$ une fonction dont l'image est constituée d'éléments qui commutent deux à deux. Il existe un unique morphisme de A-algèbres $\mathrm{ev}_b\colon A[(X_i)_{i\in\mathcal{I}}]\to B$ qui fait commuter*



Ce morphisme est appelé évaluation *en b. Pour toute A-algèbre C et tout morphisme de A-algèbres $\varphi\colon B\to C$, $\varphi\circ\mathrm{ev}_b=\mathrm{ev}_{\varphi\circ b}$.*

Démonstration. Soit B' la sous-algèbre de B engendrée par l'image de b. L'hypothèse de commutation assure que B' est commutative. En particulier $(B',\times)$ est un monoïde commutatif.



La propriété universelle du monoïde commutatif libre $M_{\mathcal{I}}$ sur $\mathcal{I}$ assure l'existence d'un unique morphisme de monoïdes $\tilde{b}$ de $M_{\mathcal{I}}$ dans B' qui étend b. La propriété universelle des algèbres de monoïdes appliquée à $\tilde{b}$ fournit alors le morphisme ev_b souhaité (plus exactement on obtient le morphisme souhaité en composant avec l'inclusion de B' dans B).

Comme d'habitude, l'unicité de ev_b découle de l'unicité dans les deux propriétés universelles utilisées mais il faut faire un tout petit peu attention à l'inclusion de B' dans B qui complique très légèrement la discussion. Soit $\psi\colon A[(X_i)_{i\in\mathcal{I}}]\to B$ un morphisme d'algèbre tel que $\psi\circ j\circ\iota_{\mathcal{I}}=b$. Comme $A[(X_i)_{i\in\mathcal{I}}]$ est engendrée (en tant que A-algèbre)

par $j \circ \iota_{\mathcal{J}}(\mathcal{J})$, l'image de ψ est engendrée par $\psi \circ j \circ \iota_{\mathcal{J}}(\mathcal{J}) = b(\mathcal{J})$ donc c'est B'. On peut donc oublier B et voir ψ comme morphisme de $A[(X_i)_{i\in\mathcal{J}}]$ dans B' et dérouler l'unicité.

Montrons la dernière partie de l'énoncé. Soit $\varphi\colon B \to C$ un morphisme de A-algèbres. L'unicité dans la propriété universelle appliquée à $\varphi \circ b$ assure que, pour montrer que $\varphi \circ \mathrm{ev}_b = \mathrm{ev}_{\varphi\circ b}$, il suffit de remarquer que $\varphi \circ \mathrm{ev}_b$ est un morphisme de A-algèbres et que, pour tout i, $\varphi \circ \mathrm{ev}_b(X_i) = \varphi \circ b(i)$. □

Remarque 6.3.3. En utilisant que chaque anneau est, de façon unique, une $\mathbb{Z}$-algèbre, on obtient comme cas particulier de la proposition précédente que $\mathbb{Z}[(X_i)_{i\in\mathcal{J}}]$ est un anneau commutatif libre sur $\mathcal{J}$: pour tout anneau commutatif B et toute fonction $b\colon \mathcal{J} \to B$, il existe un unique morphisme d'anneaux de $\mathbb{Z}[(X_i)_{i\in\mathcal{J}}]$ dans B qui étend b.

Définition 6.3.4. *Dans le cas où $\mathcal{J} = \{1, \dots, n\}$, l'ensemble des fonctions b intervenant dans la propriété universelle ci-dessus est naturellement un sous-ensemble de B^n, en identifiant b au n-uplet de ses images $(b_1, \dots, b_n)$. Pour tout $P \in A[X_1, \dots, X_n]$, on écrit $P(b_1, \dots, b_n)$ plutôt que $\mathrm{ev}_b(P)$.*

Remarque 6.3.5. La propriété universelle des algèbres de polynômes est exprimée en termes de A-algèbres. En général, étant donnée B, le morphisme de A dans B est évident. Mais on peut aussi partir d'un anneau commutatif A, d'un anneau B et d'un morphisme $\psi\colon A \to B$ d'image centrale. Dans ce cas la propriété universelle, disons dans le cas d'un ensemble fini d'indéterminées, se lit : pour toute famille b_1, …, b_n d'éléments de B qui commutent entre eux deux à deux, il existe un unique morphisme $\bar{\psi}\colon A[X_1, \dots, X_n] \to B$ tel que $\bar{\psi}$ étend ψ et envoie chaque X_i sur le b_i correspondant. La propriété d'extension de ψ correspond exactement au fait que $\bar{\psi}$ est compatible avec la structure de A-algèbre fournie par ψ.

Lemme 6.3.6. *Si B est commutative, l'application qui envoie $P \in A[X_1, \dots, X_n]$ sur la fonction $(b_1, \dots, b_n) \mapsto P(b_1, \dots, b_n)$ est un morphisme de A-algèbres. Son image est appelée l'algèbre des fonctions polynomiales sur B^n.*

Démonstration. Les opérations sur l'ensemble de fonctions de B^n dans B sont définie ponctuellement donc il suffit de vérifier la compatibilité avec l'addition et la multiplication pour chaque point $(b_1, \dots, b_n)$. Celles-ci sont immédiatement fournies par la propriété universelle. L'unité de $A[X_1, \dots, X_n]$ est bien envoyée sur 1 et les polynômes constants (qui forment l'image de $A \to A[X_1, \dots, X_n]$) sont bien envoyés sur les fonctions constantes correspondantes. □

Remarque 6.3.7. Le morphisme du lemme n'est pas injectif en général. C'est une de raisons qui imposent le point de vue algébrique sur les polynômes. Par exemple, dans $\mathbb{Z}/2\mathbb{Z}[X]$, le polynôme $X(X+1)$ n'est pas nul mais la fonction polynômiale correspondante s'annule partout. Plus généralement, si A est un anneau fini non trivial, l'application de $A[X]$ dans les fonctions polynomiales sur A n'est pas injective, son noyau est engendré par le produit des $(X - a)$ pour a parcourant A.

Lemme 6.3.8. *Soit A un anneau commutatif et B une A-algèbre commutative. Pour tout ensemble $\mathcal{J}$, il existe un unique morphisme $A[(X_i)_{i\in\mathcal{J}}] \to B[(X_i)_{i\in\mathcal{J}}]$ qui fait commuter*

$$\begin{array}{ccc} A & \longrightarrow & B \\ \downarrow & & \downarrow \\ A[(X_i)_{i\in\mathcal{J}}] & \dashrightarrow & B[(X_i)_{i\in\mathcal{J}}] \end{array}$$

Pour tout anneau commutatif A et tout $n \geqslant 1$, il existe un unique isomorphisme de A-algèbre entre $A[X_1, \dots, X_n][X_{n+1}]$ et $A[X_1, \dots, X_{n+1}]$ qui envoie X_i sur X_i pour tout i. Cet isomorphisme sera noté par un symbole invisible et on écrira $A[X_1, \dots, X_n][X_{n+1}] = A[X_1, \dots, X_{n+1}]$.

Démonstration. Pour la première partie, on applique la propriété universelle de $A[(X_i)_{i\in\mathcal{J}}]$ à l'application naturelle de $\mathcal{J}$ dans $B[(X_i)_{i\in\mathcal{J}}]$. Pour la seconde on voit $A[X_1, \dots, X_{n+1}]$ comme $A[X_1, \dots, X_n]$-algèbre via l'inclusion induite par l'inclusion de $\{1, \dots, n\}$ dans $\{1, \dots, n+1\}$ et on envoie X_{n+1} sur... X_{n+1} (mais l'autre). Bref, ces anneaux sont « égaux ». □

Lemme 6.3.9. *Soit A un anneau commutatif. Si A est intègre alors tous les anneaux de polynômes à coefficients dans A sont intègres.*

Démonstration. Supposons A intègre. Pour tout ensemble $\mathcal{I}$ et tout diviseur de zéro P dans $A[(X_i)_{i\in\mathcal{J}}]$, il existe un ensemble $\mathcal{J} \subset \mathcal{I}$ fini tel que $A[\mathcal{J}]$ contient ce diviseur de zéro et un élément Q non nul tel que $PQ = 0$ (en effet chaque élément de $A[(X_i)_{i\in\mathcal{J}}]$ est contenu dans un tel sous-anneau). Il suffit donc de montrer le lemme pour des anneaux de polynômes n'ayant qu'un nombre fini d'indéterminées.

On raisonne par récurrence sur le nombre d'indéterminées. Le cas de base est notre hypothèse sur A. Supposons par récurrence que $A[X_1, \dots, X_n]$ est intègre. Supposons qu'il existe un diviseur de zéro P non nul dans $A[X_1, \dots, X_{n+1}]$. Soit Q non nul tel que $PQ = 0$. En utilisant l'isomorphisme $A[X_1, \dots, X_{n+1}] = A[X_1, \dots, X_n][X_{n+1}]$, on écrit $P = \sum_{i=0}^{d} P_i X_{n+1}^i$ et $Q = \sum_{j=0}^{d'} Q_j X_{n+1}^j$ où les P_i et Q_j sont dans $A[X_1, \dots, X_n]$, $P_d \neq 0$ et $Q_{d'} \neq 0$ (on utilise ici que P et Q sont non nuls). Le coefficient de $X_{n+1}^{d+d'}$ dans PQ est $P_d Q_{d'}$ donc $P_d Q_{d'} = 0$. Comme $A[X_1, \dots, X_n]$ est intègre par hypothèse de récurrence, $P_d = 0$ ou $Q'_d = 0$, ce qui est contradictoire. □

Remarque 6.3.10 (Intermède logique). Le pas de récurrence dans la démonstration ci-dessus n'est pas une démonstration par l'absurde, c'est une démonstration d'une négation. Une démonstration par l'absurde suppose le contraire de l'objectif, démontre une contradiction et en déduit l'objectif. Il s'agit d'un procédé indirect qui repose sur l'axiome du tiers-exclu qui affirme que tout énoncé mathématique est soit vrai soit faux. La négation d'un énoncé affirme que cet énoncé implique une contradiction. La démonstration directe d'une négation consiste donc, comme toute démonstration d'une implication, à supposer l'énoncé et à démontrer la conclusion, ici une contradiction. La confusion entre ces deux types de raisonnements est étonnamment fréquente (mais assez inoffensive dans un contexte ordinaire).

Lemme 6.3.11. *Soit A un anneau commutatif. Il existe une unique application A-linéaire de $A[X]$ dans lui-même qui étend $X^n \mapsto nX^{n-1}$. On l'appelle la* dérivation *des polynômes et on la note $P \mapsto P'$. Elle vérifie la formule de Leibniz $\forall P\, Q$, $(PQ)' = P'Q + PQ'$.*

Démonstration. L'existence et l'unicité provient de la propriété universelle des modules libres. L'application $(P,Q) \mapsto (PQ)'$ est A-bilinéaire (c'est à dire A-linéaire par rapport à P quand Q est fixé et inversement) car chaque application partielle est composée d'applications linéaires. Il suffit donc de vérifier la formule de Leibniz pour $P = X^k$ et $Q = X^l$ pour tous k et l dans $\mathbb{N}$. C'est un calcul immédiat. □

6.4 Déterminants

On rappelle la définition suivante (qui n'est pas la plus élégante façon de définir le déterminant mais qui a l'avantage de ne nécessiter aucun prérequis).

Définition 6.4.1. *Soit A un anneau commutatif et $n \geqslant 1$ un entier. Le déterminant d'une matrice $M \in \mathcal{M}_n(A)$ est*

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} M_{i\sigma(i)}.$$

La comatrice de M est la matrice $\mathrm{co}(M) \in \mathcal{M}_n(A)$ définie par

$$\mathrm{co}(M) \colon (i,j) \mapsto (-1)^{i+j} \det(M^{i,j})$$

où $M^{i,j}$ est la matrice obtenue en supprimant la i-ème ligne et la j-ème colonne de M.

Un observation évidente mais cruciale pour généraliser aux anneaux commutatifs de nombreux résultats concernant le déterminant est que le déterminant d'une matrice M est l'évaluation en $(M_{11}, \dots, M_{nn})$ du polynôme

$$\mathrm{Det} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \prod_{i=1}^{n} X_{i\sigma(i)} \in \mathbb{Z}[X_{11}, \dots, X_{nn}].$$

Cette observation et ses variantes permettent de faire des calculs dans des anneaux de polynômes à coefficients dans $\mathbb{Z}$ et d'en déduire des résultats sur les matrices. Le premier niveau d'utilisation de cette méthode ne fait que formaliser l'idée de calcul qui n'utilise que la définition d'un anneau (commutatif) et donc fonctionne sur tout anneau. Le lemme ci-dessous fait partie de cette catégorie. Le suivant utilise le second niveau qui permet de remplacer des arguments topologiques.

Lemme 6.4.2. *Soit A un anneau commutatif et $n \geqslant 1$ un entier. Pour toute matrice $M \in \mathcal{M}_n(A)$, $M\,{}^t\mathrm{co}(M) = {}^t\mathrm{co}(M)M = \det(M) I_n$.*

Pour toutes matrices M et N dans $\mathcal{M}_n(A)$, $\det(MN) = \det(M)\det(N)$

Démonstration. Chaque coefficient $(M\,{}^t\mathrm{co}(M))_{i,j}$ est clairement l'évaluation en les $M_{k,l}$ d'un polynôme $P_{i,j}$ de $\mathbb{Z}[X_{11}, \dots, X_{nn}]$. De même ${}^t\mathrm{co}(M)M$ et $\det(M)I_n$ correspondent à des familles de polynômes $Q_{i,j}$ et $R_{i,j}$. Les calculs qui sont habituellement présentés comme une démonstration du résultat pour des matrices à coefficient dans un corps montrent en fait que $P_{i,j} = Q_{i,j} = R_{i,j}$ pour tous i et j. En évaluant ces égalités sur les coefficients d'une matrice M on obtient le résultat général.

Le deuxième résultat se démontre de la même façon mais plus simplement car il n'y a qu'un polynôme par côté de l'égalité. □

Voyons maintenant un véritable exemple d'application de cette idée.

Proposition 6.4.3 (Formule de Sylvester)**.** *Soit A un anneau commutatif et $n \geqslant 1$ un entier. Pour toutes matrices M et N dans $\mathcal{M}_n(A)$,*

$$\det(I_n + MN) = \det(I_n + NM).$$

Démonstration. On présente d'abord la démonstration habituelle pour $A = \mathbb{R}$ ou $\mathbb{C}$. On observe que $M(I_n + NM) = M + MNM = (I_n + MN)M$ donc la multiplicativité du déterminant donne $\det(M)\det(I_n + NM) = \det(I_n + MN)\det(M)$. Si M est inversible, on peut diviser $\det(M)$ et obtenir le résultat. On en déduit le cas général par densité des matrices inversibles et continuité de $(M, N) \mapsto \det(I_n + NM)$ et $(M, N) \mapsto \det(I_n + MN)$.

Cet argument topologique est sans espoir dans un anneau commutatif général (particulièrement s'il n'est pas intègre donc ne se plonge pas dans un corps). Soit $R = \mathbb{Z}[X_{11}, \dots, X_{nn}, Y_{11}, \dots, Y_{nn}]$. Soit M_0 et N_0 les matrices à coefficients dans R définies par X_{ij} et Y_{ij} respectivement. Soit $P_0 = \det(I_n + M_0N_0) \in R$ et $Q_0 = \det(I_n + N_0M_0) \in R$. Le même argument que dans le premier paragraphe montre que $\det(M_0)P = Q\det(M_0)$. Or $\det(M_0) \neq 0$ et R est intègre d'après le lemme 6.3.9, donc $P_0 = Q_0$. Or, pour tout anneau commutatif A et toutes matrices M et N, $\det(I_n + MN) = P_0(M_{11}, \dots, N_{nn})$ et $\det(I_n + NM) = Q_0(M_{11}, \dots, N_{nn})$ donc la formule générale est démontrée. □

7 Extensions d'algèbres et de corps

Dans ce chapitre on étudie les extensions d'algèbres, autrement dit les paires d'algèbres emboîtées. Le cas le plus important est celui des corps. La construction principale de ce chapitre consiste à étendre une algèbre en ajoutant un élément qui sera racine d'un polynôme donné. Il s'agit d'une généralisation du passage de $\mathbb{R}$ à $\mathbb{C}$ qui se fait en « ajoutant » à $\mathbb{R}$ une racine, notée i, du polynôme X^2+1. En répétant ce processus à partir de $\mathbb{Z}/p\mathbb{Z}$ on peut construire tous les corps finis. Plus précisément, le théorème 7.4.9 donne une classification complète de ces corps. L'objectif suivant est d'utiliser la théorie des extensions de corps pour caractériser les nombres constructibles à la règle et au compas, ce qui permet de démontrer de célèbres résultats d'impossibilité de constructions. Enfin la dernière section montre que tout corps a une extension dans laquelle tout polynôme de degré strictement positif admet au moins une racine.

7.1 Éléments entiers et adjonction de racine

Définition 7.1.1. *Soit A un anneau commutatif. Un élément b d'une A-algèbre B est* entier *sur A s'il existe $P \in A[X]$ unitaire tel que $P(b) = 0$.*

Exemple 7.1.2. Comme tous les anneaux, $\mathbb{Q}$ est une $\mathbb{Z}$-algèbre. Montrons qu'un élément de $\mathbb{Q}$ est entier sur $\mathbb{Z}$ si et seulement si il appartient à $\mathbb{Z}$. Soit $x_0 \in \mathbb{Q}$. Si $x_0 \in \mathbb{Z}$ alors $P = X - x_0$ est bien un élément unitaire de $\mathbb{Z}[X]$ qui annule x_0 (ie $P(x_0) = 0$). Réciproquement, supposons que x_0 soit racine d'un polynôme unitaire $P \in \mathbb{Z}[X]$. On écrit $x_0 = a/b$ avec a et b dans $\mathbb{Z}$ premiers entre eux et $b \neq 0$. On écrit $P = X^n + \sum_{i=0}^{n-1} a_i X^i$. En multipliant la relation $P(x_0) = 0$ par b^n, on obtient $a^n + \sum_{i=0}^{n-1} a_i a^i b^{n-i} = 0$. Dans le membre de gauche, b divise la somme donc b divise a^n. Or b et a sont premiers entre eux donc b est une unité de $\mathbb{Z}$, c'est à dire $b = \pm 1$ et x_0 est dans $\mathbb{Z}$.

Dans l'algèbre des endomorphismes d'un espace vectoriel de dimension finie sur un corps $\mathbb{K}$, tous les éléments sont entiers sur $\mathbb{K}$ puisque tout endomorphisme admet un polynôme annulateur unitaire.

Proposition 7.1.3. *Soit A un anneau commutatif et B une A-algèbre. Si B est commutative alors l'ensemble des éléments entiers sur A est une sous-algèbre de B.*

Démonstration. On note d'abord que les éléments de A sont entiers car, pour tout $a \in A$, $X - a$ annule a. Soit x et y des éléments de B entiers sur A. On veut montrer que $x \pm y$ et xy sont entiers sur A. Notons z l'un de ces éléments (l'argument qui suit sera le même dans les trois cas). On écrit sous la forme $X^n - P$ un polynôme annulateur de x unitaire dans $A[X]$, avec $\deg(P) < n$. De même on écrit $X^m - Q$ un polynôme annulateur de y.

On pose $N = nm$. Soit $(v_1, \dots, v_N)$ une énumération de tous les produits $x^i y^j$ avec $0 \leqslant i < n$ et $0 \leqslant j < m$ en commençant par $v_1 = 1$. Pour tout k entre 1 et nm, zv_k s'écrit sous la forme $\sum_{l=1}^{N} c_{kl} v_l$ avec $c_{kl} \in A$ pour tous k et l. En effet la multiplication

de v_k par x ou y fait au pire apparaître un x^n qu'on peut remplacer par $P(x)$ qui a bien la forme attendue ou bien un y^m qu'on peut remplacer par $Q(y)$.

On note $C \in \mathcal{M}_N(A)$ la matrice de coefficients c_{kl}. Par construction le vecteur $v \in B^N$ de coordonnées v_k vérifie $Cv = zv$. D'après le lemme 6.4.2, on a ${}^t\mathrm{co}(C - zI_N)(C-zI_N) = \det(C-zI_N)I_N$ donc ${}^t\mathrm{co}(C - zI_N)(Cv-zv) = \det(C-zI_N)v$ et donc $\det(C-zI_N)v = 0$. Or $v_1 = 1$ donc on en déduit $\det(C-zI_N) = 0$. Ainsi le polynôme $(-1)^N \det(C-XI_N)$, qui est bien unitaire et à coefficients dans A, convient. □

Remarque 7.1.4. La fin de la démonstration précédente nécessite un peu de précautions car le lien entre polynôme caractéristique et valeurs propre est plus subtile lorsque l'anneau des scalaires n'est pas intègre. Sous l'hypothèse que A est intègre, on aurait pu passer dans le corps des fractions de A, conclure directement $\chi_C(z) = 0$ à partir de $Cv = zv$ et $v \neq 0$ puis se souvenir que χ_C est à coefficients dans A.

On notera aussi que la démonstration précédente est complètement effective. Si on connait un polynôme annulateur unitaire pour x et y on peut écrire la matrice C de façon complètement explicite et obtenir un polynôme annulateur explicite pour z en calculant un déterminant. Le résultat est difficile à imaginer même dans des cas très simples, par exemple si $A = \mathbb{Z}$ et $B = \mathbb{R}$, $x = \sqrt{2}$ et $y = \sqrt{3}$ on trouve $v = (1, \sqrt{2}, \sqrt{3}, \sqrt{6})$ et, pour $z = \sqrt{2} + \sqrt{3}$,

$$C = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 3 & 0 & 0 & 1 \\ 0 & 3 & 2 & 0 \end{pmatrix}$$

qui donne $X^4 - 10X^2 + 1$ comme polynôme annulateur de z. Ce résultat n'est pas complètement évident et on ne peut pas faire mieux en terme de degré.

Remarque 7.1.5. Dans la proposition précédente, la commutativité de B est cruciale. Par exemple, si $A = \mathbb{Z}$ et $B = \mathcal{M}_2(\mathbb{Q})$, $a = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ et $b = \left(\begin{smallmatrix} 0 & 1/2 \\ 0 & 0 \end{smallmatrix}\right)$ sont entiers sur $\mathbb{Z}$ (car annulés par X^2) mais $ab = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1/2 \end{smallmatrix}\right)$ ne l'est pas car, pour tout polynôme unitaire $P \in \mathbb{Z}[X]$, $P(ab) = \left(\begin{smallmatrix} 0 & 0 \\ 0 & P(1/2) \end{smallmatrix}\right)$ qui n'est jamais nul d'après l'exemple 7.1.2.

Voyons maintenant comment construire une A-algèbre à partir d'un anneau commutatif A en ajoutant un élément ayant un polynôme annulateur prescrit. Il s'agit donc de généraliser la construction qui ajoute une racine de $X^2 + 1$ à $\mathbb{R}$ pour obtenir $\mathbb{C} = \mathbb{R}[X]/(X^2 + 1)$.

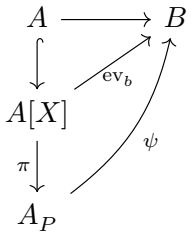
Proposition 7.1.6. *Soit A un anneau commutatif et $P \in A[X]$. L'anneau $A[X]/(P)$, muni du morphisme $j\colon A \to A[X]/(P)$ obtenu en composant $A \hookrightarrow A[X]$ avec la projection canonique $\pi\colon A[X] \to A[X]/(P)$, est une A-algèbre dans laquelle P admet une racine, à savoir $\pi(X)$. Si P est unitaire et de degré strictement positif alors j est injectif.*

Cette algèbre vérifie la propriété universelle suivante : pour toute A-algèbre B et pour toute racine b de P dans B, il existe un unique morphisme de A-algèbre $\psi\colon A[X]/(P) \to B$ qui envoie $\pi(X)$ sur b. Comme d'habitude, cette propriété caractérise $A[X]/(P)$ modulo unique isomorphisme.

Démonstration. Posons $A_P = A[X]/(P)$. Avec les notations de la propriété universelle des algèbres de polynômes (proposition 6.3.2), il s'agit de montrer que l'évaluation $\mathrm{ev}_{\pi(X)} : A[X] \to A_P$ envoie P sur zéro. On remarque que le morphisme d'anneaux

$\pi : A[X] \to A_P$ est aussi un morphisme de A-algèbres, par construction de la structure d'algèbre sur A_P. Or l'unicité dans la propriété universelle des algèbres de polynômes assure que $\mathrm{ev}_{\pi(X)}$ est l'unique morphisme de A-algèbres qui envoie X sur $\pi(X)$. Donc $\mathrm{ev}_{\pi(X)} = \pi$ et $\mathrm{ev}_{\pi(X)}(P) = \pi(P) = 0$. Le critère d'injectivité sur j découle du lemme 4.3.16 qui assure que la projection de $A[X]$ sur A_P est injective en restriction aux polynômes de degré strictement inférieur à celui de P, donc en particulier en restriction à l'image de A dans $A[X]$.

Soit B une A-algèbre et b une racine de P dans B. La propriété universelle des algèbres de polynômes fournit un unique morphisme d'algèbres $\mathrm{ev}_b : A[X] \to B$ tel que $\mathrm{ev}_b(X) = b$. Pour montrer que ev_b descend, de façon unique, en morphisme d'anneaux $\psi : A_P \to B$, il suffit de montrer que $(P) \subset \ker \mathrm{ev}_b$. Pour cela il suffit de montrer que $P \in \ker \mathrm{ev}_b$. Or $\mathrm{ev}_b(P) = 0$ par définition de « b est racine de P ». Dans le diagramme



les deux petits triangles commutent donc le grand aussi, donc ψ est aussi un morphisme de A-algèbres. Comme d'habitude, l'unicité de ψ s'obtient en combinant les unicités dans les deux propriétés universelles invoquées. □

7.2 Algèbres sur un corps

Dans cette section on suppose que l'anneau des scalaires est un corps. Le premier objectif est de généraliser la théorie du polynôme minimal rencontrée en algèbre linéaire. On étudiera ensuite dans quel cas une algèbre obtenue par adjonction d'une racine (au sens de la proposition 7.1.6) à une corps est elle-même un corps, comme c'est le cas quand on adjoint à $\mathbb{R}$ une racine de $X^2 + 1$ pour obtenir $\mathbb{C}$.

Notation 7.2.1. *Soit $\mathbb{K}$ un corps et A une $\mathbb{K}$-algèbre. Pour toute famille $a_1, \dots, a_n$ d'éléments de A qui commutent deux à deux, on note $\mathbb{K}[a_1, \dots, a_n]$ la sous-algèbre de A qu'ils engendrent. Il s'agit de l'image du morphisme d'algèbre $\mathrm{ev}_a : \mathbb{K}[X_1, \dots, X_n] \to A$ (en effet cette image une sous-algèbre qui vérifie clairement la propriété universelle de la sous-algèbre engendrée par les a_i).*

Remarque 7.2.2. La notation ci-dessus peut sembler ambigüe mais en pratique il y a peu de risque de confusion avec une algèbre de polynôme et on sait dans quoi vivent les a_i donc on retrouve le A implicite dans la notation.

Définition 7.2.3. *Soit $\mathbb{K}$ un corps. Un élément a d'une $\mathbb{K}$-algèbre A est* algébrique *sur $\mathbb{K}$ s'il existe $P \in \mathbb{K}[X] \setminus \{0\}$ tel que $P(a) = 0$, autrement dit $\ker \mathrm{ev}_a \neq 0$. On appelle alors* polynôme minimal *de a, et on note μ_a, l'unique générateur unitaire de $\ker \mathrm{ev}_a$. Le degré de μ_a est appelé* degré *de a.*

Un élément qui n'est pas algébrique est dit transcendant.

Remarque 7.2.4. Un élément a d'une $\mathbb{K}$-algèbre A est algébrique sur $\mathbb{K}$ si et seulement si il est entier sur $\mathbb{K}$ au sens de la section précédente. En effet tout polynôme unitaire est non nul et l'existence d'un polynôme annulateur non nul implique celle d'un polynôme annulateur unitaire en divisant par le coefficient dominant (qui est inversible car $\mathbb{K}$ est un corps). Cette duplication de terminologie est un accident historique. En particulier la proposition 7.1.3 s'applique ici : si A est commutative, l'ensemble des éléments de A algébriques sur $\mathbb{K}$ forme une sous-algèbre.

Le lemme suivant décrit la sous-algèbre engendré par un élément en discutant selon que cet élément est algébrique ou transcendant.

Lemme 7.2.5. *Soit $\mathbb{K}$ un corps, A une $\mathbb{K}$-algèbre et a un élément de A. Si a est algébrique sur $\mathbb{K}$ alors son degré n est le minimum des degrés des polynômes annulateurs de a et la famille $(1, a, \dots, a^{n-1})$ est une $\mathbb{K}$-base de $\mathbb{K}[a]$ (on rappelle que toute $\mathbb{K}$-algèbre est en particulier un $\mathbb{K}$-espace vectoriel). En particulier $\dim_{\mathbb{K}}(\mathbb{K}[a]) = \deg(\mu_a)$. L'anneau $\mathbb{K}[a]$ est un corps si et seulement si μ_a est irréductible.*

Si au contraire a est transcendant alors ev_a est un isomorphisme de $\mathbb{K}$-algèbres entre $\mathbb{K}[X]$ et $\mathbb{K}[a]$. En particulier $\mathbb{K}[a]$ est de dimension infinie.

Démonstration. Supposons d'abord que a est algébrique sur $\mathbb{K}$. L'étude des idéaux de $\mathbb{K}[X]$ assure qu'il existe un unique polynôme unitaire qui engendre $\ker \mathrm{ev}_a$ et que son degré est minimal parmi les éléments de $\ker \mathrm{ev}_a$.

La proposition 7.1.6 assure que ev_a induit un morphisme de $\mathbb{K}$-algèbres de $\mathbb{K}[X]/(\mu_a)$ dans A qui envoie $\pi(X)$ sur a. Comme $\ker \mathrm{ev}_a = (\mu_a)$, ce morphisme est injectif. Son image est $\mathrm{im}\, \mathrm{ev}_a = \mathbb{K}[a]$. Le lemme 4.3.16 assure que, dans $\mathbb{K}[X]/(\mu_a)$, la famille 1, $\pi(X)$, …, $\pi(X)^{n-1}$ est libre et engendre donc elle forme bien une base et son image 1, a, …, a^{n-1} dans $\mathbb{K}[a]$ aussi. Enfin le lemme 4.2.12 assure que $\mathbb{K}[a]$ est un corps si et seulement si (μ_a) est un idéal maximal. D'après la proposition 4.3.33, cette condition est équivalente à l'irréductibilité de μ_a.

Dans le cas où a est transcendant, le théorème 4.2.7 assure directement que ev_a est un isomorphisme sur son image $\mathbb{K}[a]$. □

Remarque 7.2.6. L'histoire du polynôme minimal ne se généralise pas bien aux cas des algèbres sur des anneaux qui ne sont pas des corps. D'abord il n'y a aucun raison que l'anneau de polynômes pertinent soit principal. Mais en plus il n'y a même pas unicité du polynôme unitaire de degré minimal annulant un élément entier. Par exemple $\mathbb{Z}/2\mathbb{Z}$ est une $\mathbb{Z}$-algèbre, 1 est entier et il est annulé par $X - k \in \mathbb{Z}[X]$ pour tout entier impair k.

On peut aussi fabriquer des exemples où il existe un polynôme annulateur non unitaire de degré strictement plus petit que tous les polynômes annulateurs unitaires. Par exemple si l'anneau de base est $R = \mathbb{Z}/4\mathbb{Z}$ et si $A = R[Y]$ alors $a = 2Y$ est entier sur R car annulé par X^2 mais il est aussi annulé par $2X$ sans être annulé par aucun polynôme unitaire de degré 1.

Lemme 7.2.7. *Soit $\mathbb{K}$ un corps et A une $\mathbb{K}$-algèbre. Pour tout élément algébrique a dans A, $a \in A^\times \Leftrightarrow \mu_a(0) \neq 0$. De plus, si a est algébrique et inversible alors a^{-1} est dans $\mathbb{K}[a]$.*

Démonstration. Soit a un élément algébrique de A. On écrit $\mu_a = XP + \mu_a(0)$ pour un certain $P \in \mathbb{K}[X]$ de degré $\deg(\mu_a)-1$. Par minimalité du degré de μ_a, $P(a) \neq 0$. Comme $\mu_a(a) = 0$, on obtient $aP(a) = -\mu_a(0)$. Si $\mu_a(0) = 0$ alors, comme $P(a) \neq 0$, a ne peut pas être inversible. Si $\mu_a(0) \neq 0$ alors on a $a[(-\mu_a(0)^{-1}P(a)] = 1$ et $[(-\mu_a(0)^{-1}P(a)]a = 1$ donc a est inversible (notons que ces calculs utilisent que les éléments de $\mathbb{K}$ tels que $\mu_a(0)$ commutent avec tous les éléments de A et que les puissances de a commutent entre elles). Dans ce dernier cas, $a^{-1} = -\mu_a(0)^{-1}P(a)$ est bien dans $\mathbb{K}[a]$. □

7.3 Extensions de corps

Définition 7.3.1. *Une* extension *d'un corps $\mathbb{K}$ est un corps $\mathbb{L}$ muni d'un morphisme d'anneaux $\rho\colon \mathbb{K} \to \mathbb{L}$. Autrement c'est une $\mathbb{K}$-algèbre qui est un corps. On abrège souvent « Soit $\mathbb{L}$ une extension de $\mathbb{K}$ » en « Soit $\mathbb{L}/\mathbb{K}$ », bien qu'il n'y ait aucun rapport avec les quotients.*

Exemple 7.3.2. Le corps des réels est une extension du corps des rationnels. Le corps des nombres complexes est une extension du corps des réels. Le corps $\mathbb{K}(X)$ des fractions rationnelles à coefficients dans un corps $\mathbb{K}$ est une extension de $\mathbb{K}$.

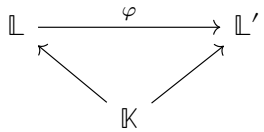
Pour toute extension de corps $\mathbb{L}/\mathbb{K}$ et toute famille $\alpha_1, \dots, \alpha_n$ d'éléments de $\mathbb{L}$, on note $\mathbb{K}(\alpha_1, \dots, \alpha_n)$ le sous-corps de $\mathbb{L}$ engendré par l'image de $\mathbb{K}$ et $\{\alpha_1, \dots, \alpha_n\}$. Il s'agit aussi d'une extension de $\mathbb{K}$.

Remarque 7.3.3. Dans une extension de corps $\rho\colon \mathbb{K} \to \mathbb{L}$, le morphisme ρ est automatiquement injectif. En effet son noyau est un idéal de $\mathbb{K}$ donc 0 ou 1 d'après le lemme 4.2.11 et le cas $\ker\rho = 1$ est exclus car on aurait alors $1_{\mathbb{K}} \in \ker\rho$ tandis que $\rho(1_{\mathbb{K}}) = 1_{\mathbb{L}}$ et $1_{\mathbb{L}} \neq 0_{\mathbb{L}}$ car $\mathbb{L}$ est un corps.

Pour cette raison on écrit parfois qu'une extension $\mathbb{L}$ d'un corps $\mathbb{K}$ est un sur-corps de $\mathbb{K}$, c'est à dire un corps tel que $\mathbb{K} \subset \mathbb{L}$. Ce n'est ni plus ni moins abusif que d'écrire $\mathbb{Q} \subset \mathbb{R}$ ou $\mathbb{R} \subset \mathbb{C}$.

Par ailleurs le fait de ne presque jamais expliciter ρ ne doit pas faire oublier que $\mathbb{K}$ et $\mathbb{L}$ ne déterminent pas uniquement ρ. Par exemple $\mathbb{Q}[X]/(X^2-2)$ est un corps (car X^2-2 est irréductible dans $\mathbb{Q}[X]$) qui se plonge de deux façons dans $\mathbb{R}$ par $a+b\bar{X} \mapsto a+b\sqrt{2}$ et par $a+b\bar{X} \mapsto a-b\sqrt{2}$. Ces deux plongements proviennent de la propriété universelle de l'adjonction de racines (proposition 7.1.6) appliquées aux deux racines de X^2-2 dans $\mathbb{R}$. Cet exemple conduit à la définition suivante.

Définition 7.3.4. *Soit $\mathbb{K}$ un corps. Un isomorphisme d'extensions de $\mathbb{K}$ est un isomorphisme de $\mathbb{K}$-algèbres entre deux extensions de $\mathbb{K}$. Autrement dit il s'agit d'un isomorphisme $\varphi\colon \mathbb{L} \to \mathbb{L}'$ tel que*



commute. On y pense comme l'ensemble des isomorphismes entre $\mathbb{L}$ et $\mathbb{L}'$ qui « sont l'identité sur $\mathbb{K}$ ». L'ensemble des automorphismes d'une extension $\mathbb{L}$ de $\mathbb{K}$ est appelé le groupe de Galois de $\mathbb{L}$ sur $\mathbb{K}$ et noté $\mathrm{Gal}(\mathbb{L}/\mathbb{K})$.

Exemple 7.3.5. La conjugaison complexe est un automorphisme de $\mathbb{C}/\mathbb{R}$ (encore une fois, cette notation signifie « $\mathbb{C}$ vu comme extension de $\mathbb{R}$ », cela n'indique pas un quotient).

La première partie du lemme suivant généralise le fait que les racines complexes d'un polynôme à coefficients réels viennent par paires complexes conjuguées. La deuxième partie est une sorte de réciproque partielle.

Lemme 7.3.6. *Soit $\mathbb{K}$ un corps et $\mathbb{L}$ et $\mathbb{L}'$ deux extensions de $\mathbb{K}$.*

- *Pour tout isomorphisme d'extensions $\varphi\colon \mathbb{L} \to \mathbb{L}'$, pour tout polynôme $P \in \mathbb{K}[X]$ et tout $\alpha \in \mathbb{L}$, $P(\varphi(\alpha)) = \varphi(P(\alpha))$. En particulier $P(\alpha) = 0 \Leftrightarrow P(\varphi(\alpha)) = 0$.*
- *Soit $\alpha \in \mathbb{L}$ et $\beta \in \mathbb{L}'$ des éléments algébriques sur $\mathbb{K}$. Il existe un isomorphisme d'extensions $\varphi\colon \mathbb{K}(\alpha) \to \mathbb{K}(\beta)$ qui envoie α sur β si et seulement si $\mu_\alpha = \mu_\beta$.*

Démonstration. Soit $\varphi : \mathbb{L} \to \mathbb{L}'$ un isomorphisme d'extensions et $\alpha \in \mathbb{L}$. La dernière partie de la proposition 6.3.2 assure que $\mathrm{ev}_{\varphi(\alpha)} = \varphi \circ \mathrm{ev}_\alpha$ donc, pour tout P, $P(\varphi(\alpha)) = \mathrm{ev}_{\varphi(\alpha)}(P) = \varphi \circ \mathrm{ev}_\alpha(P) = \varphi(P(\alpha))$.

Il est logiquement inutile mais néanmoins instructif d'écrire un calcul plus explicite en prétendant que $\mathbb{K}$ est vraiment inclus dans $\mathbb{L}$ et $\mathbb{L}'$ et que $\varphi|_{\mathbb{K}} = \mathrm{Id}$ de sorte que les coefficients a_i de P vérifient $\varphi(a_i) = a_i$:

$$\begin{aligned} P(\varphi(\alpha)) &= \sum_i a_i \varphi(\alpha)^i \\ &= \sum_i \varphi(a_i)\varphi(\alpha)^i \\ &= \sum_i \varphi(a_i \alpha^i) \\ &= \varphi\left(\sum_i a_i \alpha^i\right) \\ &= \varphi(P(\alpha)). \end{aligned}$$

Montrons maintenant le second point. Supposons que φ soit un isomorphisme d'extensions qui envoie α sur β. D'après le premier point, pour tout $P \in \mathbb{K}[X]$, $P(\alpha) = 0 \Leftrightarrow P(\beta) = 0$ donc ev_α et ev_β ont même noyau et donc $\mu_\alpha = \mu_\beta$. Réciproquement supposons que $\mu_\alpha = \mu_\beta$. On a alors les isomorphismes de $\mathbb{K}$-algèbres

$$\mathbb{K}(\alpha) \simeq \mathbb{K}[X]/(\mu_\alpha) = \mathbb{K}[X]/(\mu_\beta) \simeq \mathbb{K}(\beta)$$

dont la composée envoie bien α sur β. □

Exemple 7.3.7. Pour $\mathbb{K} = \mathbb{Q}$ et $\mathbb{L} = \mathbb{L}' = \mathbb{C}$, avec $\alpha = \sqrt[3]{2}$ et $\beta = e^{2i\pi/3}\alpha$, on a $\mu_\alpha = \mu_\beta = X^3 - 2$ donc il existe un isomorphisme de $\mathbb{Q}$-algèbres qui envoie $\mathbb{Q}(\alpha)$ sur $\mathbb{Q}(\beta)$ bien que le premier soit inclus dans $\mathbb{R}$ et pas le second.

Le lemme suivant est immédiat mais très important.

Lemme 7.3.8. *Soit $\mathbb{L}/\mathbb{K}$ une extension de corps. Pour tout automorphisme $\varphi \in \mathrm{Gal}(\mathbb{L}/\mathbb{K})$, l'ensemble $\mathrm{Fix}(\varphi)$ des points fixes de φ est un sous-corps de $\mathbb{L}$ qui contient (l'image de) $\mathbb{K}$.*

Exemple 7.3.9. La conjugaison complexe est un automorphisme de $\mathbb{C}/\mathbb{Q}$ dont l'ensemble des points fixes est $\mathbb{R}$. Dans $\mathbb{Q}[\sqrt{2}]$, l'application $a+b\sqrt{2} \mapsto a-b\sqrt{2}$ est un automorphisme de $\mathbb{Q}[\sqrt{2}]/\mathbb{Q}$ dont l'ensemble des points fixes est $\mathbb{Q}$.

Proposition 7.3.10. *Soit $\mathbb{L}/\mathbb{K}$ une extension de corps. L'ensemble des éléments de $\mathbb{L}$ qui sont algébriques sur $\mathbb{K}$ est un sous-corps de $\mathbb{L}$ (qui contient $\mathbb{K}$).*

Démonstration. On sait déjà que cet ensemble $\mathbb{L}'$ est une sous-algèbre de $\mathbb{L}$ par la proposition 7.1.3. Il reste à voir que $\mathbb{L}' \setminus \{0\}$ est stable par inversion. Cela découle directement du lemme 7.2.7. □

Définition 7.3.11. *Le* degré *$\mathbb{L}/\mathbb{K}$ est la dimension de $\mathbb{L}$ comme $\mathbb{K}$-espace vectoriel. On le note $[\mathbb{L} : \mathbb{K}]$. On dit que $\mathbb{L}/\mathbb{K}$ est :*

- *une* extension finie *si son degré est fini,*
- *une* extension algébrique *si tous les éléments de $\mathbb{L}$ sont algébriques sur $\mathbb{K}$,*
- *une* extension transcendante *si elle n'est pas algébrique.*

Exemple 7.3.12. L'extension $\mathbb{C}/\mathbb{R}$ est finie, de degré 2. Plus généralement, le lemme 7.2.5 assure que pour tout $\mathbb{L}/\mathbb{K}$ et tout $\alpha \in \mathbb{L}$ algébrique, $\mathbb{K}(\alpha)/\mathbb{K}$ est finie, de degré $\deg(\alpha) = \deg(\mu_\alpha)$ (c'est de cet exemple que provient la terminologie « degré d'une extension »).

Le même lemme 7.2.5 assure aussi que toute extension finie est algébrique puisqu'un élément transcendant engendre un sous-corps de dimension infinie. La réciproque de cette observation est fausse : on peut montrer que l'ensemble des réels algébriques sur $\mathbb{Q}$ est de degré infini (il s'agit bien d'un sous-corps de $\mathbb{R}$ contenant $\mathbb{Q}$ d'après la proposition 7.3.10).

Proposition 7.3.13. *Soit $\mathbb{F}/\mathbb{E}$ et $\mathbb{E}/\mathbb{K}$ des extensions de corps. On a*

$$[\mathbb{F} : \mathbb{K}] = [\mathbb{F} : \mathbb{E}]\ [\mathbb{E} : \mathbb{K}].$$

En particulier si $\mathbb{F}/\mathbb{E}$ et $\mathbb{E}/\mathbb{K}$ sont finies alors $\mathbb{F}/\mathbb{K}$ est finie.

Démonstration. Dans cette démonstration, les applications entre corps sont implicites. Soit $(e_i)_{i\in I}$ une $\mathbb{K}$-base de $\mathbb{E}$ et $(f_j)_{j\in J}$ une $\mathbb{E}$-base de $\mathbb{F}$. Montrons que $(e_i f_j)_{(i,j)\in I\times J}$ est une $\mathbb{K}$-base de $\mathbb{F}$. Soit $x \in \mathbb{F}$. On décompose x sur la base f en $\sum_j x_j f_j$ avec chaque x_j dans E. Puis on décompose chaque x_j sur la base e en $\sum_i x_{j,i} e_i$. On a ainsi

$$x = \sum_{i\in I,\, j\in J} x_{j,i} e_i f_j$$

donc la base promise engendre $\mathbb{F}$. Montrons qu'elle est libre. Supposons

$$\sum_{i,j} \lambda_{j,i} e_i f_j = 0.$$

Comme la famille f est libre, on obtient $\forall j, \sum_{i,j} \lambda_{j,i} e_i = 0$. Puis, pour chaque j, on utilise que la famille e est libre pour obtenir $\forall i, \lambda_{j,i} = 0$. □

Corollaire 7.3.14. *Si $\mathbb{L}/\mathbb{K}$ est une extension finie alors, pour tout $\alpha \in \mathbb{L}$, $\deg(\alpha) \mid [\mathbb{L} : \mathbb{K}]$ (on sait déjà que α est algébrique sur $\mathbb{K}$).*

Démonstration. D'après la proposition précédente, $[\mathbb{L} : \mathbb{K}] = [\mathbb{L} : \mathbb{K}(\alpha)][\mathbb{K}(\alpha) : \mathbb{K}]$ et, d'après le lemme 7.2.5, $[\mathbb{K}(\alpha) : \mathbb{K}] = \deg(\alpha)$. □

Corollaire 7.3.15. *Si $\mathbb{F}/\mathbb{E}$ et $\mathbb{E}/\mathbb{K}$ sont algébriques alors $\mathbb{F}/\mathbb{K}$ est algébrique.*

Démonstration. Soit $\alpha \in \mathbb{F}$. Comme $\mathbb{F}/\mathbb{E}$ est algébrique, on obtient $n \in \mathbb{N}$ et $a_0, \dots, a_{n-1} \in \mathbb{E}$ tels que $\alpha^n = \sum_i a_i \alpha^i$. Par hypothèse, chaque a_i est algébrique sur $\mathbb{K}$. On considère la tour d'extensions

$$\mathbb{K}(a_0, \dots, a_{n-1})/\mathbb{K}(a_0, \dots, a_{n-2})/\cdots/\mathbb{K}(a_0)/\mathbb{K}.$$

Comme chaque a_i est algébrique sur $\mathbb{K}$, il est a fortiori algébrique sur $\mathbb{K}(a_0, \dots, a_{i-1})$ donc chacune de ces extensions est finie d'après l'exemple 7.3.12. Par la proposition précédente, on en déduit par récurrence sur n que $\mathbb{K}(a_0, \dots, a_{n-1})/\mathbb{K}$ est finie. Par ailleurs α est algébrique sur $\mathbb{K}(a_0, \dots, a_{n-1})$ donc $\mathbb{K}(a_0, \dots, a_{n-1}, \alpha)$ est finie sur $\mathbb{K}(a_0, \dots, a_{n-1})$ donc sur $\mathbb{K}$. Ainsi α est dans une extension finie de $\mathbb{K}$ donc est algébrique sur $\mathbb{K}$. □

Définition 7.3.16. *Soit $\mathbb{K}$ un corps et $P \in \mathbb{K}[X]$. Un corps de décomposition pour P est une extension $\mathbb{L}/\mathbb{K}$ telle que :*

- *P est* scindé *sur L :*
$$P = a \prod_{i=1}^{\deg(P)} (X - \alpha_i)$$
pour des $\alpha_i \in \mathbb{L}$ (pas nécessairement distincts)
- $\mathbb{L} = \mathbb{K}(\alpha_1, \dots, \alpha_{\deg(P)})$.

Proposition 7.3.17. *Tout polynôme à coefficients dans un corps admet un corps de décomposition.*

Remarque 7.3.18. Dans le résultat précédent, le corps de décomposition est unique modulo isomorphisme d'extensions du corps de base, mais nous n'aurons pas besoin de ce fait.

Démonstration. Montrons par récurrence sur $n \in \mathbb{N}$ que

$$\forall \mathbb{K}, \forall P \in \mathbb{K}[X], \deg(P) \leqslant n \Rightarrow P \text{ admet un corps de décomposition.}$$

On notera que $\mathbb{K}$ n'est pas fixé. Si $n = 0$ alors pour tout $\mathbb{K}$ et tout P, $\mathbb{K}$ convient. Supposons maintenant le résultat prouvé jusqu'à n. Soit $\mathbb{K}$ un corps et $P \in \mathbb{K}[X]$ avec $\deg(P) = n + 1$. Comme $\deg(P) \geqslant 1$, on obtient un facteur irréductible Q de P avec $\deg(Q) \geqslant 1$. Le corps $\mathbb{L}_1 = \mathbb{K}[X]/(Q)$ est une extension de $\mathbb{K}$ dans laquelle Q a une racine α et $\mathbb{L}_1 = \mathbb{K}(\alpha)$. Dans $\mathbb{L}_1[X]$, $(X - \alpha) \mid Q$ donc $(X - \alpha) \mid P$. On obtient ainsi $R \in \mathbb{L}_1[X]$ tel que $P = (X - \alpha)R$ et $\deg(R) = n$. Par hypothèse de récurrence appliquée à $\mathbb{L}_1$ et R, on obtient un corps de décomposition $\mathbb{L}$ pour R. Sur ce corps, R et donc P sont scindés et $\mathbb{L} = \mathbb{L}_1(\alpha_1, \dots, \alpha_n) = \mathbb{K}(\alpha_1, \dots, \alpha_n, \alpha)$ donc $\mathbb{L}$ est un corps de décomposition pour P. □

7.4 Corps finis

Définition 7.4.1. *Soit A un anneau et $\iota : \mathbb{Z} \to A$ l'unique morphisme d'anneau de $\mathbb{Z}$ dans A. La caractéristique de A est l'unique entier naturel* $\mathrm{car}(A)$ *tel que*

$$\ker(\iota : \mathbb{Z} \to A) = \mathrm{car}(A)\mathbb{Z}.$$

Exemple 7.4.2. L'anneau $\mathbb{Z}$ est de caractéristique nulle. L'anneau $\mathbb{Z}/n\mathbb{Z}$ est de caractéristique n. Un anneau est de caractéristique 1 si et seulement si il est trivial.

Remarque 7.4.3. Si un anneau A est intègre alors sa caractéristique est nulle ou un nombre premier. En effet $\iota : \mathbb{Z} \to A$ induit un isomorphisme de $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$ vers un sous-anneau de A. On réserve parfois le mot caractéristique à ce cas des anneaux intègres.

Le même argument montre que si $\mathrm{car}(A) = 0$ alors A est infini. La réciproque est fausse, par exemple $\mathbb{Z}/2\mathbb{Z}(X)$ est un corps infini de caractéristique 2.

Lemme 7.4.4. *Soit $\mathbb{K}$ un corps. Pour toute $\mathbb{K}$-algèbre A non triviale,* $\mathrm{car}(A) = \mathrm{car}(\mathbb{K})$. *En particulier toute extension de $\mathbb{K}$ a la même caractéristique que $\mathbb{K}$.*

Démonstration. Soit $\rho : \mathbb{K} \to A$ le morphisme d'anneau qui donne la structure de $\mathbb{K}$-algèbre de A. On note $\iota_{\mathbb{K}}$ et ι_A les uniques morphismes de $\mathbb{Z}$ dans $\mathbb{K}$ et A respectivement. Par cette unicité, $\iota_A = \rho \circ \iota_{\mathbb{K}}$. Or ρ est nécessairement injectif car $\ker \rho$ est un idéal de $\mathbb{K}$ donc 0 ou 1 et $\rho(1_{\mathbb{K}}) = 1_A$ et on a supposé A non trivial donc $1_A \neq 0_A$. On a donc $\ker \iota_A = \ker(\rho \circ \iota_{\mathbb{K}}) = \ker \iota_{\mathbb{K}}$. □

Lemme 7.4.5. *Un corps est de caractéristique nulle si et seulement si il contient une copie de $\mathbb{Q}$. Cette copie est alors unique, c'est le sous-corps engendré par* 1.

Démonstration. Si $\mathbb{Q}$ s'injecte dans un corps $\mathbb{K}$ alors on a $\mathrm{car}(\mathbb{K}) = \mathrm{car}(\mathbb{Q}) = 0$ par le lemme précédent. Réciproquement, supposons $\mathrm{car}(\mathbb{K}) = 0$. Alors $\iota : \mathbb{Z} \to \mathbb{K}$ est injective. La propriété universelle du corps des fractions (corollaire 4.4.13) assure alors que ι s'étend de façon unique en morphisme (nécessairement injectif) de $\mathbb{Q}$ dans $\mathbb{K}$. Toute autre injection de $\mathbb{Q}$ étend nécessairement ι par unicité de ι. □

Lemme 7.4.6. *Un anneau A est de caractéristique p avec p premier si et seulement si il contient une copie de $\mathbb{Z}/p\mathbb{Z}$. Cette copie est alors unique, c'est le sous-anneau de A engendré par* 1.

Démonstration. Si A contient une copie de $\mathbb{Z}/p\mathbb{Z}$ alors $\mathrm{car}(A) = \mathrm{car}(\mathbb{Z}/p\mathbb{Z}) = p$ d'après le lemme 7.4.4. Réciproquement supposons que $\mathrm{car}(A) = p$. On a vu que $\iota : \mathbb{Z} \to A$ induit alors un morphisme injectif de $\mathbb{Z}/p\mathbb{Z}$ dans A. Son image est le sous-anneau engendré par 1 car $\mathbb{Z}/p\mathbb{Z}$ est engendré par 1 (comme anneau). Il reste à montrer l'unicité. Toute copie de $\mathbb{Z}/p\mathbb{Z}$ dans A contient nécessairement 1 donc le sous-anneau qu'il engendre. Or ce sous-anneau est de cardinal p donc on conclut par inclusion et égalité des cardinaux (finis). □

Notation 7.4.7. *Dans toute la suite de ce chapitre, on notera $\mathbb{F}_p$ le corps $\mathbb{Z}/p\mathbb{Z}$.*

Lemme 7.4.8. *Soit p un nombre premier et A un anneau commutatif de caractéristique p. L'application $F : A \to A$ qui envoie x sur x^p est un morphisme d'anneaux appelé morphisme de Frobenius de A. Si A est intègre alors* $\mathrm{Fix}(F) = \mathbb{F}_p$ *(le membre de droite désignant l'unique copie de $\mathbb{F}_p$ dans A promise par le lemme 7.4.6).*

Démonstration. Le fait que F est un morphisme de monoïdes multiplicatifs est clair. C'est l'addition qui nécessite l'hypothèse de caractéristique. Soit x et y dans A. Comme A est commutatif, on a la formule du binôme de Newton :

$$(x+y)^p = \sum_{k=0}^{p} \binom{p}{k} x^k y^{p-k}.$$

Il s'agit de voir que seuls les termes correspondant à $k=0$ et $k=p$ sont non nuls. Par définition de $\mathrm{car}(A)$, il suffit de montrer que, pour tous les autres k, $p \mid \binom{p}{k}$. Pour $k>0$, on a

$$\binom{p}{k} = \frac{p(p-1)\cdots(p-k-1)}{k!} = \frac{p}{k}\frac{(p-1)\cdots(p-k-1)}{(k-1)!}$$

donc $\binom{p}{k} = \frac{p}{k}\binom{p-1}{k-1}$ et p divise $k\binom{p}{k}$. Si on suppose en plus $k<p$ alors p ne divise pas k et, comme p est premier, p divise $\binom{p}{k}$.

Par définition, les points fixes de F dont les racines du polynôme $X^p - X$. Si A est intègre alors ce polynôme a au plus p racines d'après le lemme 4.3.17. Or on sait, par le petit théorème de Fermat, que tous les éléments de $\mathbb{Z}/p\mathbb{Z}$ sont racines donc on conclut par inclusion et égalité de cardinaux (finis). □

Théorème 7.4.9 (Classification des corps finis)**.** *Pour tout $q \in \mathbb{N}$, il existe un corps fini de cardinal q si et seulement si q est de la forme p^n avec p premier et $n \geqslant 1$. Soit p un nombre premier, $n \geqslant 1$ un entier et $\mathbb{K}$ un corps de cardinal $q = p^n$.*

- $\mathrm{car}(\mathbb{K}) = p$
- *Le groupe multiplicatif $(\mathbb{K}^\times, \times)$ est cyclique*
- *Dans $\mathbb{K}[X]$, $X^q - X = \prod_{a\in\mathbb{K}}(X-a)$. En particulier tous les éléments a de $\mathbb{K}$ vérifient $a^q = a$.*
- *Il existe $\alpha \in \mathbb{K}$ tel que $\mathbb{K} = \mathbb{F}_p[\alpha]$*
- *Les sous-corps de $\mathbb{K}$ sont exactement les $\mathrm{Fix}(F^k)$ pour k divisant n. Il y en a donc $\varphi(n)$. De plus $\sharp\,\mathrm{Fix}(F^k) = p^k$.*
- *Si $\mathbb{L}$ est un corps fini à p^n élément alors il existe un isomorphisme entre $\mathbb{K}$ et $\mathbb{L}$ (mais il n'est pas unique en général).*

Remarque 7.4.10. Le dernier point du théorème affirme qu'un corps fini est déterminé modulo isomorphisme par son cardinal. Malgré l'absence d'unicité dans ce résultat, il est traditionnel de dire et d'écrire « soit $\mathbb{F}_q$ *le* corps à q éléments ».

Le fait que le groupe multiplicatif d'un corps fini soit cyclique est non trivial, même dans le cas d'un cardinal premier. Par exemple 3 engendre $\mathbb{F}_7^\times$ et ses puissances énumèrent $\mathbb{F}_7^\times$ dans l'ordre $(3,2,6,4,5,1)$.

Vu le quatrième point de l'énumération ci-dessus, l'existence d'un corps fini à p^n éléments est équivalente à l'existence d'un polynôme irréductible de degré n dans $\mathbb{F}_p[X]$. Là encore cette existence n'a rien d'évident.

On verra dans la démonstration que les deux points les plus spectaculaires, l'information sur le groupe des inversibles et l'unicité modulo isomorphisme, sont très liés.

Démonstration. Soit $\mathbb{K}$ un corps de cardinal fini q. Vu la remarque 7.4.3, $\mathbb{K}$ est de caractéristique p pour un nombre premier p. Le lemme 7.4.6 assure alors que $\mathbb{K}$ est une $\mathbb{F}_p$-algèbre. Cette algèbre est nécessairement de dimension finie car $\mathbb{K}$ est fini. On note n cette dimension. La classification des espaces vectoriels assure que $\mathbb{K}$ est isomorphe à $\mathbb{F}_p^n$ (en temps que $\mathbb{F}_p$-espace vectoriel) donc il est de cardinal p^n.

Réciproquement, soit q un entier de la forme p^n avec p premier et $n \geqslant 1$. La proposition 7.3.17 fournit un corps de décomposition $\mathbb{L}$ pour $P = X^q - X \in \mathbb{F}_p[X]$. En particulier P est scindé sur $\mathbb{L}$: $P = \prod_{i=1}^{q}(X - \alpha_i)$. Montrons que les α_i sont deux à deux distincts. Supposons qu'il existe α tel que $P = (X-\alpha)^2 Q$ pour un $Q \in \mathbb{L}[X]$. D'après le lemme 6.3.11, on a alors $P' = 2(X-\alpha)Q + (X-\alpha)^2 Q'$ donc $P'(\alpha) = 0$. Or $P' = qX^{q-1} - 1 = -1$ car $p \mid q$ et $\mathrm{car}(\mathbb{L}) = p$ d'après le lemme 7.4.4. On obtient donc $0 = 1$ qui absurde car $\mathbb{L}$ est un corps. Ainsi l'ensemble des racines de P est de cardinal exactement q. De plus cet ensemble est un sous-corps de $\mathbb{L}$ d'après le lemme 7.3.8 car il s'agit de $\mathrm{Fix}(F^n)$ où F est le morphisme de Frobenius de $\mathbb{L}$ qui est un morphisme de corps d'après le lemme 7.4.8. On a bien trouvé un corps de cardinal q. A posteriori on peut même affirmer qu'il s'agit exactement de $\mathbb{L}$ puisque $\mathbb{L}$ est engendré par les α_i.

Dans toute la suite, $\mathbb{K}$ désigne un corps de cardinal $q = p^n$. Montrons que tous les éléments de $\mathbb{K}$ sont racines de $X^q - X$. Le théorème de Lagrange (corollaire 3.2.13) assure que, pour tout $x \in \mathbb{K}^\times$, $x^{\sharp\mathbb{K}^\times} = 1$, c'est-à-dire $x^{q-1} = 1$. Ainsi tous les éléments de $\mathbb{K}^\times$ sont racines de $X^{q-1} - 1$ et donc tous les éléments de $\mathbb{K}$ sont racines de $X(X^{q-1} - 1)$, c'est-à-dire de $X^q - X$. Comme ce polynôme a au plus q racines, on apprend qu'il est scindé à racines simples et on a bien la factorisation annoncée dans $\mathbb{K}[X]$.

Pour montrer que $\mathbb{K}^\times$ est cyclique, on va montrer l'énoncé plus général suivant : pour tout corps $\mathbb{L}$, tout sous-groupe fini de $\mathbb{L}^\times$ est cyclique. Soit $\mathbb{L}$ un corps et G un sous-groupe fini de $\mathbb{L}^\times$. Pour tout $x \in G$, $x^{\exp(G)} = 1$ (cf. définition 5.4.12). Ainsi les éléments de G sont tous des racines de $X^{\exp(G)} - 1$. Or ce polynôme a au plus $\exp(G)$ racines donc $\sharp G \leqslant \exp(G)$. Par ailleurs, comme pour tout groupe fini, $\exp(G) \leqslant \sharp G$. Donc $\sharp G = \exp(G)$ et le corollaire 5.4.14 assure que G est cyclique. Remarque : de nombreuses sources déduisent ce résultat de la classification complète des groupes abéliens finis, mais ce n'est pas raisonnable, le lemme 5.4.13 est une étape importante mais complètement élémentaire et auto-contenue dans cette classification.

Le paragraphe précédent fourni un générateur α de $\mathbb{K}^\times$. On a alors $\mathbb{K} = \mathbb{F}_p(\alpha)$ car $\mathbb{F}_p(\alpha)$ contient à la fois zéro et le sous groupe multiplicatif engendré par α, c'est à dire $\mathbb{K}^\times$.

Montrons maintenant la description des sous-corps de $\mathbb{K}$. Soit $\mathbb{L}$ un sous-corps de $\mathbb{K}$. D'après le lemme 7.4.4, $\mathrm{car}(\mathbb{K}) = \mathrm{car}(\mathbb{L})$ donc $\mathrm{car}(\mathbb{L}) = p$ et on a une tour d'extensions $\mathbb{K}/\mathbb{L}/\mathbb{F}_p$. Posons $k = [\mathbb{L} : \mathbb{F}_p]$. En particulier $\mathbb{L} = p^k$ d'après la première partie de ce théorème. On a aussi vu que tous les éléments de $\mathbb{L}$ sont racines de $X^{p^k} - X$. Comme ce polynôme a au plus p^k racines dans $\mathbb{K}$, on obtient que $\mathbb{L}$ est l'ensemble des racines de $X^{p^k} - X$ dans $\mathbb{K}$, c'est-à-dire l'ensemble des points fixe de F^k où F est le morphisme de Frobenius de $\mathbb{K}$. La proposition 7.3.13 assure que $n = [\mathbb{K} : \mathbb{F}_p] = [\mathbb{K} : \mathbb{L}] \cdot [\mathbb{L} : \mathbb{F}_p]$. Ainsi $k \mid n$. Réciproquement, soit k un diviseur de n. On sait que $\mathrm{Fix}(F^k)$ est un sous-corps de $\mathbb{K}$. Montrons qu'il est de cardinal p^k. Le lemme arithmétique ci-dessous assure que $X^{p^k} - X$ divise $X^{p^n} - X$ dans $\mathbb{Z}[X]$ et donc dans $\mathbb{F}_p[X]$ (car le morphisme d'anneaux de $\mathbb{Z}$ dans $\mathbb{F}_p$ induit un morphisme d'anneaux de $\mathbb{Z}[X]$ dans $F_p[X]$). Comme on a vu que $X^{p^n} - X$ est scindé à racines simples, c'est aussi le cas de son diviseur $X^{p^k} - X$ qui a

donc p^k racines.

Enfin montrons que si $\mathbb{L}$ est un corps à p^n éléments alors $\mathbb{L}$ est isomorphe à $\mathbb{K}$. Soit α tel que $\mathbb{K} = \mathbb{F}_p(\alpha)$. On note μ_α le polynôme minimal de α sur $\mathbb{F}_p$. On a vu que α est racine de $X^q - X$ donc $\mu_\alpha \mid X^q - X$ dans $\mathbb{F}_p[X]$. Or on sait que $X^q - X$ est scindé sur $\mathbb{L}$ donc μ_α aussi. Ainsi μ_α a une racine β dans $\mathbb{L}$. Comme μ_α est irréductible d'après le lemme 7.2.5 et unitaire, il s'agit du polynôme minimal de β. Le lemme 7.3.6 fournit alors un isomorphisme entre $\mathbb{F}_p(\alpha) = \mathbb{K}$ et $\mathbb{F}_p(\beta) \subset \mathbb{L}$. En particulier $\sharp\mathbb{F}_p(\beta) = p^n$ et, comme $\sharp\mathbb{L} = p^n$, on a $\mathbb{F}_p(\beta) = \mathbb{L}$ et l'isomorphisme construit est un isomorphisme de $\mathbb{K}$ vers $\mathbb{L}$. □

La démonstration précédente a laissé de côté le petit lemme arithmétique suivant.

Lemme 7.4.11. *Soit p, n et m des entiers naturels. Si $m \mid n$ alors $X^{p^m} - X \mid X^{p^n} - X$ dans $\mathbb{Z}[X]$.*

Démonstration. Posons $k = p^n$ et $l = p^m$. Supposons $m \mid n$ et posons $r = n/m$. On observe que, pour tout anneau commutatif intègre A, pour tout $a \in A$ et tout $s \in \mathbb{N}^*$, $a - 1 \mid a^s - 1$ car $a^s - 1 = (a-1)(a^{s-1} + \cdots + a + 1)$. En appliquant cette observation à $A = \mathbb{Z}$, $a = l$ et $s = r$, on obtient que $l - 1 \mid l^r - 1$, c'est à dire $l - 1 \mid k - 1$. Posons $d = (k-1)/(l-1)$. On applique l'observation à $A = \mathbb{Z}[X]$, $a = X^{l-1}$ et $s = d$ pour obtenir $X^{l-1} - 1 \mid X^{d(l-1)} - 1$, c'est à dire $X^{l-1} - 1 \mid X^{k-1} - 1$ puis, en multipliant par X, $X^l - X \mid X^k - X$. □

Exemple 7.4.12. Cette classification et sa démonstration sont extrêmement concrètes, elles permettent en principe de tout de calculer dans les corps finis. Voyons par exemple comment construire un corps à $2^3 = 8$ éléments. Le théorème dit qu'on doit chercher un polynôme irréductible de degré 2 dans $\mathbb{F}_2[X]$. En degré trois l'irréductibilité est équivalente à l'absence de racine puisque toute décomposition en produits de facteurs de degrés strictement positifs ferait intervenir un facteur de degré un. On trouve comme polynômes possibles X^3+X^2+1 et X^3+X+1. Explorons les deux possibilités en parallèle. On pose $\mathbb{K} = \mathbb{F}_2[X]/(X^3 + X^2 + 1)$ et on note α l'image de X. On pose $\mathbb{L} = \mathbb{F}_2[X]/(X^3 + X + 1)$ et on note β l'image de X. D'après le lemme 4.3.16, on a :

$$\mathbb{K} = \{0, 1, \alpha, \alpha+1, \alpha^2, \alpha^2+\alpha, \alpha^2+1, \alpha^2+\alpha+1\}$$

et une énumération similaire pour $\mathbb{L}$ avec β à la place de α. De plus on a $\alpha^3+\alpha^2+1 = 0$ et $\beta^3 + \beta + 1 = 0$ d'après la proposition 7.1.6 donc $\alpha^3 = \alpha^2 + 1$ et $\beta^3 = \beta + 1$ (car ces corps sont de caractéristique 2 donc tous les signes moins disparaissent). Ces informations suffisent à décrire entièrement l'addition et la multiplication dans ces corps. Par exemple $(\alpha+1)\alpha^2 = \alpha^3 + \alpha^2 = \alpha^2 + \alpha^2 + 1 = 1$ et on voit en particulier que $\alpha + 1$ et α^2 sont bien inversibles.

Le fait que les groupes multiplicatifs de $\mathbb{K}$ et $\mathbb{L}$ soient cycliques est trivial car ils sont de cardinal $8-1 = 7$ qui est premier donc tous les éléments sauf 1 engendrent le groupe. La démonstration du théorème de classification explique ensuite comment trouver un isomorphisme entre $\mathbb{K}$ et $\mathbb{L}$. Il suffit de choisir un générateur de $\mathbb{K}^\times$ et de trouver un élément de $\mathbb{L}$ ayant le même polynôme minimal. On choisit α comme générateur de $\mathbb{K}^\times$. On a $\mu_\alpha = X^3 + X^2 + 1$ par construction. Ensuite, par recherche exhaustive, on voit que les racines de μ_α dans $\mathbb{L}$ sont $\beta+1$, $\beta^2+\beta+1$ et β^2+1. Prenons la première. On obtient un

unique isomorphisme de corps de $\mathbb{K}$ sur $\mathbb{L}$ qui envoie α sur $\beta+1$. Vérifions que $\alpha+1$ et α^2 sont bien envoyés sur des éléments inverses l'un de l'autre. On a $\varphi(\alpha+1) = \beta+1+1 = \beta$, $\varphi(\alpha^2) = (\beta+1)^2 = \beta^2+1$ et $\beta(\beta^2+1) = \beta^3+\beta = \beta+1+\beta = 1$.

7.5 Nombres constructibles

Dans cette section, on utilise la théorie des extensions de corps pour expliquer la solution de problèmes très célèbres qui ont résisté aux mathématiciens pendant plus de deux mille ans : en utilisant uniquement une règle non graduée et un compas, on ne peut pas construire un cube de volume double d'un cube donné (c'est le problème de la duplication du cube), ou bien couper en trois parts égales un angle de $\pi/3$ (c'est le problème de la trisection de l'angle).

Définition 7.5.1. *Étant donnés deux points distincts dans le plan, on définit par récurrence les points, droites et cercles* constructibles *à la règle et au compas par :*

- *Le deux points donnés sont constructibles.*
- *Si A et B sont constructibles et distincts alors la droite (AB) est constructible et le cercle $\mathcal{C}(A,B)$ de centre A passant par B est constructible.*
- *Tout point d'intersection isolé entre droites ou cercles constructibles est constructible.*

On dit qu'un nombre réel x est constructible si $|x|$ est la distance entre deux points constructibles, pour l'unité de distance fournie par les deux points de départ.

Le but de cette section est de démontrer le théorème suivant et quelques uns de ses corollaires.

Théorème 7.5.2. *L'ensemble $\mathbb{K}$ des nombres réels constructibles est un sous-corps de $\mathbb{R}$ qui est stable par racine carré. Un nombre réel x est constructible si et seulement si il existe une tour d'extensions $\mathbb{R}/K_N/\cdots/K_1/\mathbb{Q}$ telle que $x \in K_N$ et, pour tout $i < N$, $K_{i+1} = K_i(\sqrt{d_i})$ pour un certain $d_i \in K_i \cap \mathbb{R}_+$ tel que $\sqrt{d_i} \notin K_i$.*

On verra dans la démonstration qu'un point est constructible dans un plan muni d'un repère orthonormé à partir de l'origine et du point de coordonnée $(1,0)$ si et seulement si ses deux coordonnées sont des nombres constructibles.

Corollaire 7.5.3. *Tout réel constructible x est algébrique sur $\mathbb{Q}$ et $\deg(x) = 2^n$ pour un entier n. Le nombre $\sqrt[3]{2}$ n'est pas constructible donc on ne peut pas réaliser la duplication du cube à la règle et au compas. Le nombre $\cos(\pi/9)$ n'est pas constructible donc on ne peut pas réaliser la trisection d'un angle de $\pi/3$ à la règle et au compas.*

Démonstration. Soit x un réel constructible. Le théorème fournit une tour d'extensions $\mathbb{R}/K_N/\cdots/K_1/\mathbb{Q}$ telle que $x \in K_N$ et, pour tout $i < N$, $[K_{i+1} : K_i] = 2$. La proposition 7.3.13 assure que $[K_N : \mathbb{Q}] = 2^N$. De plus le corollaire 7.3.14 assure que $\deg(x) \mid [K_N : \mathbb{Q}]$ donc $\deg(x)$ est bien de la forme 2^n.

Ainsi, pour montrer que $x_0 = \sqrt[3]{2}$ n'est pas constructible, il suffit de montrer qu'il n'est pas algébrique de degré une puissance de deux. Montrons qu'il est algébrique de degré 3. On sait que X^3-2 annule x_0 donc x_0 est algébrique et $\deg(x_0) \mid 3$. Ainsi x_0 est

de degré 1 ou 3. La première possibilité signifierait que x_0 est rationnel. Montrons que ce n'est pas le cas. Supposons que $x_0 = p/q$ avec p et q des entiers strictement positifs premiers entre eux. On alors $p^3 = 2q^3$. Donc, en notant $v_2(n)$ la valuation 2-adique d'un entier n, on obtient $3v_2(p) = 3v_2(q) + 1$ puis $3(v_2(p) - v_2(q)) = 1$, ce qui est absurde car 3 n'est pas inversible dans $\mathbb{Z}$.

Montrons de même que $x_1 = \cos(\pi/9)$ est algébrique de degré 3. Pour tout réel θ, on a $\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta)$. Comme $\cos(\pi/3) = 1/2$, on en déduit que $1/2 = 4x_1^3 - 3x_1$ donc x_1 est racine de $8X^3 - 6X - 1$. Ainsi x_1 est algébrique et son degré divise 3 donc il suffit de montrer qu'il est irrationnel. Supposons que $x_0 = p/q$ avec p et q des entiers strictement positifs premiers entre eux. On a alors $8p^3 - 6pq^2 - q^3 = 0$ donc $p(8p^2 - 6q^2) = q^3$ et $p \mid q^3$. Comme p et q sont premiers entre eux, on obtient $p = 1$ et $8 - 6q^2 = q^3$. Ainsi $q^2(q + 6) = 8$. L'unicité de la décomposition en facteurs premiers d'un entier fournit k et l entiers naturels tels que $q = 2^k$, $q + 6 = 2^l$ et $2k + l = 3$. En particulier $2^l - 6 = 2^k \geqslant 1$ donc $l \geqslant 3$. Comme $2k + l = 3$, la seule possibilité est $(k, l) = (0, 3)$. Or $2^0 + 6 \neq 2^3$ donc x_1 n'est pas rationnel. □

Démonstration du théorème 7.5.2. On commence par observer que si une droite Δ et un point P sont construits alors la perpendiculaire Δ' à Δ passant par P est constructible. La figure 7.1 présente cette construction, les numéros indique l'ordre des étapes. La droite Δ contient au moins deux points construits et l'un au moins de ces points n'est pas sur Δ'. Appelons A un tel point. Le cercle $\mathcal{C}(P, A)$ intersecte Δ aussi en $B \neq A$ (car A n'est pas sur Δ'). Le point B est donc constructible, puis les cercles $\mathcal{C}(A, B)$ et $\mathcal{C}(B, A)$ le sont et leur intersection $\{C, D\}$ l'est. On a $(CD) = \Delta'$ qui est donc constructible.

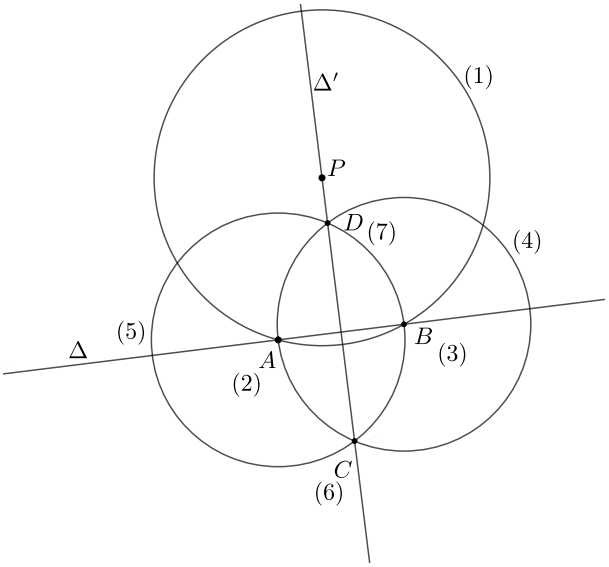


FIGURE 7.1 – Construction d'une perpendiculaire

En appliquant deux fois cette construction de perpendiculaire, on peut aussi construire la parallèle à une droite construite passant par un point construit.

Montrons que cela permet de reporter la distance entre deux points A et B construits sur une droite Δ construite à partir d'un point C construit (figure 7.2). On commence

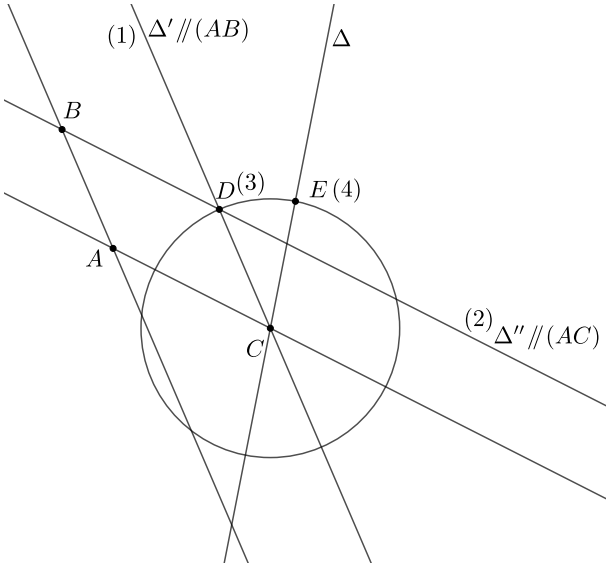


FIGURE 7.2 – Report de longueur

par construire la parallèle Δ' à (AB) passant par C puis la parallèle Δ'' à (AC) passant par B. L'intersection de ces droites est un point D et $ABDC$ est un parallélogramme donc $CD = AB$. Les points d'intersection de $\mathcal{C}(C, D)$ et de Δ réalisent le report de la longueur AB sur le droite Δ en partant de C (d'un côté ou de l'autre de C).

Muni de cette construction de report, on obtient facilement que les nombres constructibles forment un sous-groupe additif de $\mathbb{R}$. On note au passage que la construction de perpendiculaire permet de construire un repère orthonormé dont deux des points sont les points de départ et de montrer qu'un point est constructible si et seulement si ses deux coordonnées dans ce repère sont des nombres constructibles.

Supposons que x et y sont des nombres constructibles et montrons que xy l'est aussi (figure 7.3). Puisque, par définition, un nombre est constructible si et seulement si sa valeur absolue l'est, on peut supposer que x et y sont positifs. Soit A et B des points construits tels que $AB = x$. On construit Δ perpendiculaire à (AB) et passant par A.

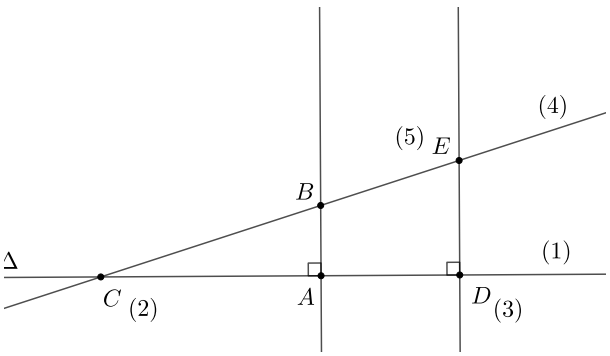


FIGURE 7.3 – Multiplication

Puis on reporte la longueur 1 sur Δ à partir de A pour obtenir un point C tel que

$CA = 1$. À partir de C on reporte la longueur y sur Δ du même côté de C que A (sur la figure $y > 1$ mais cela n'a pas d'importance dans la suite). On obtient ainsi un point D tel que $CD = y$. On construit ensuite la droite (CB) puis la perpendiculaire à Δ passant par D. On note E l'intersection de ces deux droites. Comme (AB) et (DE) sont toutes deux perpendiculaires à Δ, elles sont parallèles entre elles. Le théorème de Thalès assure alors que $DE/AB = DC/AC$, c'est à dire $DE/x = y/1$ donc $DE = xy$ et xy est bien construit. La construction de l'inverse d'un nombre constructible est une variante de cet argument, en utilisant encore le théorème de Thalès.

Montrons maintenant la stabilité par racine carré en utilisant le théorème de Pythagore (figure 7.4). Soit a un nombre constructible positif et $[AB]$ un segment de longueur a. On reporte sur la droite (AB) la longueur 1 pour obtenir un point C tel que $BC = 1$

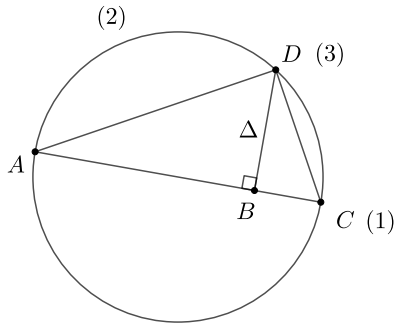


FIGURE 7.4 – Racine carrée

et $B \in [AC]$. On construit ensuite le cercle de diamètre $[AC]$ (pour cela on intersecte $\mathcal{C}(A, C)$ et $\mathcal{C}(C, A)$ pour construire la médiatrice de $[AC]$ puis le milieu I de $[AC]$ et enfin $\mathcal{C}(I, A)$). On construit la perpendiculaire Δ à (AC) passant par B et D un des points d'intersection de Δ et du cercle de diamètre $[AC]$. Par le théorème de Pythagore dans ABD, on obtient $AB^2 + BD^2 = AD^2$. Par le théorème de Pythagore dans CBD, on obtient $CB^2 + BD^2 = CD^2$. Par le théorème de Pythagore dans ACD, qui est rectangle en D car D est sur le cercle de diamètre $[AC]$, on obtient $AD^2 + CD^2 = (AB + BC)^2$. En combinant ces trois équations, on élimine AC et CB pour obtenir $2BD^2 = (AB + BC)^2 - AB^2 - CB^2$, c'est à dire $2BD^2 = (a + 1)^2 - a^2 - 1^2$ donc $BD^2 = a$ et $BD = \sqrt{a}$. Ainsi $\sqrt{a}$ est constructible.

Montrons maintenant la deuxième partie du théorème. La première partie montre que si x est dans une tour d'extensions obtenues par adjonctions de racines carrées alors x est constructible. Réciproquement, supposons que x est constructible. Montrons l'existence de la tour d'extensions par récurrence sur le nombre d'étapes de constructions, en appelant étape de construction l'intersection entre deux droites ou cercles construits.

Comme expliqué précédemment, on peut construire un repère orthonormé tel qu'un point un constructible si et seulement si ses coordonnées le sont. Après zéro étape, seul 1 est constructible. Montrons que chaque nouvelle étape préserve le corps engendré sur $\mathbb{Q}$ par les coordonnées des points construits (appelé « corps courant » dans la suite) ou l'étend en ajoutant une racine carré d'un élément du corps.

Si A et B sont construits et distincts, alors la droite (AB) est d'équation $(x_B - x_A)(y - y_A) = (x - x_A)(y_B - y_A)$ (en effet il s'agit bien de l'équation d'une droite qui contient A et B). Les coefficients de cette droite appartiennent au corps engendré sur $\mathbb{Q}$ par les coordonnées de A et de B. Les coordonnées de l'intersection de deux telles droites sont obtenues en résolvant un système d'équations linéaires. La formule de Cramer montre donc que le corps courant ne change pas (on notera que le déterminant associé n'est pas nul car les droites ne sont pas parallèles puisqu'on ne considère que des intersections isolées).

Si A et B sont construits, alors le cercle $\mathcal{C}(A, B)$ est d'équation

$$(x - x_A)^2 + (y - y_A)^2 = (x_B - x_A)^2 + (y_B - y_A)^2$$

qui est aussi à coefficient dans le corps courant. Si C et D sont construits alors on peut paramétrer la droite (CD) par

$$t \mapsto (x_C + t(x_D - x_C), y_C + t(y_D - y_C)).$$

qui est une fonction linéaire à coefficients dans le corps courant. Plus généralement toute droite admettant une équation à coefficients dans le corps courant admet un tel paramétrage. L'intersection entre $\mathcal{C}(A, B)$ et (CD) se calcule en injectant cette fonction de t dans l'équation du cercle et en cherchant les racines du polynôme de degré 2 en t ainsi obtenu. Ce polynôme est à coefficients dans le corps courant. Ainsi il suffit au pire d'ajouter une racine carrée.

L'intersection de deux cercles s'obtient par résolution d'un système de la forme

$$\begin{cases} (x - x_A)^2 + (y - y_A)^2 & = R_1^2 \\ (x - x_C)^2 + (y - y_C)^2 & = R_2^2 \end{cases}$$

où R_1 et R_2 sont dans le corps courant. La différence entre ces deux équations est linéaire car la partie quadratique est $x^2 + y^2$ dans les deux cas. De plus cette différence est non nulle car on ne considère que les intersections isolées donc les deux cercles sont distincts. Donc nous sommes de nouveau ramenés à un système formé d'une équation de cercle et d'une équation de droite, toutes deux à coefficients dans le corps courant. Notons que cet argument pour faire disparaître $x^2 + y^2$ ne fonctionnerait pas si on voulait intersecter des ellipses générales et de fait il peut y avoir exactement quatre points d'intersection entre deux ellipses donc l'espoir de se ramener à une équation de degré deux serait mince. □

7.6 Clôtures algébriques

Définition 7.6.1. *Un corps $\mathbb{K}$ est* algébriquement clos *si tout polynôme $P \in \mathbb{K}[X]$ de degré strictement positif admet une racine dans $\mathbb{K}$. On dit qu'une extension de corps $\mathbb{L}/\mathbb{K}$ est une* clôture algébrique *de $\mathbb{K}$ si $\mathbb{L}/\mathbb{K}$ est algébrique et $\mathbb{L}$ est algébriquement clos.*

Exemple 7.6.2. Le corps $\mathbb{C}$ est un clôture algébrique de $\mathbb{R}$, c'est le théorème de d'Alembert-Gauß. Cela sera démontré en cours d'analyse complexe, il n'y a pas de démonstration purement algébrique, bien que ce résultat soit souvent appelé le « théorème fondamental de l'algèbre ».

Remarque 7.6.3. Un corps algébriquement clos est nécessairement infini. En effet, si $\mathbb{K} = \{x_1, \dots, x_n\}$ alors $P = \prod_i (X - x_i) + 1$ est de degré strictement positif et n'a aucune racine dans $\mathbb{K}$.

La définition de corps algébriquement clos ne mentionne qu'une racine par polynôme de degré strictement positif, mais le lemme suivant assure que cela suffit à scinder tous les polynômes.

Lemme 7.6.4. *Si un corps $\mathbb{K}$ est algébriquement clos alors tout polynôme P est scindé dans $\mathbb{K}[X]$, c'est à dire qu'on peut écrire P comme produit de polynômes de degrés au plus un.*

Démonstration. Supposons que tout polynôme est scindé. Soit P un polynôme de degré strictement positif. Comme P a au moins un facteur de degré un, il a au moins une racine. La réciproque est une récurrence facile sur le degré en utilisant que si $P(\alpha) = 0$ alors $(X - \alpha) \mid P$. □

On termine ce cours en démontrant que tout corps admet une clôture algébrique, unique à isomorphisme d'extensions près. La démonstration n'est pas très explicite car elle utilise le lemme de Zorn à la fois pour l'existence et pour l'unicité. Mais bien sûr pour des exemples concrets de corps il existe souvent des descriptions plus explicites.

Le lemme suivant sera utile pour démontrer l'unicité.

Lemme 7.6.5. *Soit $\mathbb{L}/\mathbb{K}$ une clôture algébrique d'un corps $\mathbb{K}$. Pour toute extension algébrique $\mathbb{E}/\mathbb{K}$, il existe un morphisme de $\mathbb{K}$-algèbres de $\mathbb{E}$ dans $\mathbb{L}$.*

Démonstration. Il s'agit d'une application du lemme de Zorn. On considère l'ensemble

$$\mathcal{Z} = \{(\mathbb{F}, \varphi) \ ; \ \mathbb{F} \subset \mathbb{E} \text{ sous-extension}, \ \varphi \colon \mathbb{F} \to \mathbb{L} \text{ morphisme de } \mathbb{K}\text{-algèbre}\}$$

muni de la relation d'ordre $(\mathbb{F}, \varphi) \leqslant (\mathbb{F}', \varphi')$ si $\mathbb{F} \subset \mathbb{F}'$ et $\varphi'|_{\mathbb{F}} = \varphi$. Par hypothèse on a des morphismes de corps $\rho \colon \mathbb{K} \to \mathbb{E}$ et $\theta \colon \mathbb{K} \to \mathbb{L}$. L'ensemble $\mathcal{Z}$ n'est donc pas vide car il contient $(\rho(\mathbb{K}), \theta \circ \rho^{-1})$ où ρ^{-1} est bien définie sur $\rho(\mathbb{K})$ car ρ est injective. Soit $((\mathbb{F}_i, \varphi_i))_{i \in I}$ une famille totalement ordonnée d'éléments de $\mathcal{Z}$. On pose $\mathbb{F} = \bigcup_i \mathbb{F}_i$. Comme la famille est totalement ordonnée, il s'agit d'un sous-corps de $\mathbb{L}$. De plus la condition de restriction dans la définition de la relation d'ordre assure que les φ_i se recollent en un morphisme de $\mathbb{K}$-algèbre $\varphi \colon \mathbb{F} \to \mathbb{L}$. La paire $(\mathbb{F}, \varphi)$ est alors un majorant de la famille. Ainsi le lemme de Zorn (lemme 4.2.13) fournit un élément $(\mathbb{F}, \varphi) \in \mathcal{Z}$ maximal.

Montrons que $\mathbb{F} = \mathbb{E}$, de sorte que φ est le morphisme recherché. Comme $\mathbb{F} \subset \mathbb{E}$ par définition de $\mathcal{Z}$, il suffit de montrer que $\mathbb{E} \subset \mathbb{F}$. Soit $\alpha \in \mathbb{E}$. Comme $\mathbb{E}/\mathbb{K}$ est algébrique, α est algébrique sur $\mathbb{K}$ donc sur $\mathbb{F}$. Soit $\mu_\alpha \in \mathbb{F}[X]$ son polynôme minimal sur $\mathbb{F}$. On utilise φ pour munir $\mathbb{L}$ d'une structure de $\mathbb{F}$-algèbre. Comme $\mathbb{L}$ est algébriquement clos, μ_α admet au moins une racine β dans $\mathbb{L}$. Puisque μ_α est irréductible dans $\mathbb{F}[X]$ d'après le lemme 7.2.5, μ_α est aussi le polynôme minimal de β et le lemme 7.3.6 fournit un isomorphisme de $\mathbb{F}$-algèbres $\psi \colon \mathbb{F}(\alpha) \to \mathbb{F}(\beta)$. En particulier ψ envoie $\mathbb{F}(\alpha)$ dans $\mathbb{L}$ en étendant φ (puisque c'est un morphisme de $\mathbb{F}$-algèbres et que φ définit la structure de $\mathbb{F}$-algèbre sur $\mathbb{L}$). Ainsi $(\mathbb{F}(\alpha), \psi)$ est dans $\mathcal{Z}$ et par maximalité de $(\mathbb{F}, \varphi)$, $\mathbb{F}(\alpha) = \mathbb{F}$ et $\psi = \varphi$. En particulier $\alpha \in \mathbb{F}$ et on a montré $\mathbb{E} \subset \mathbb{F}$. □

Théorème 7.6.6. *Tout corps admet une clôture algébrique, unique à isomorphisme d'extensions près (mais un tel isomorphisme n'est pas unique en général).*

Démonstration. Soit $\mathbb{K}$ un corps. On note $A = \mathbb{K}[(X_P)_{P\in\mathbb{K}[X]}]$ la $\mathbb{K}$-algèbre des polynômes d'indéterminées indexées par $\mathbb{K}[X]$. On note I l'idéal de A engendré par les $P(X_P)$ pour $P \in \mathbb{K}[X]$ de degré strictement positif ($P(X_P) = \mathrm{ev}_{X_P}(P)$ a bien un sens puisque $P \in \mathbb{K}[X]$ et que A est une $\mathbb{K}$-algèbre). Concrètement, en écrivant $P = \sum_i c_i X^i$, $P(X_P) = \sum_i c_i X_P^i \in A$.

Montrons que I est strictement inclus dans A. Supposons que 1 est dans I. On obtient alors $N \in \mathbb{N}$, $a_1, \dots, a_N \in A$ et $P_1, \dots, P_N \in \mathbb{K}[X]$ tels que

$$\sum_{i=1}^{N} a_i P_i(X_{P_i}) = 1 \qquad (\star)$$

Comme les P_i sont de degré strictement positif, en utilisant N fois l'existence d'un corps de décomposition garantie par la proposition 7.3.17, on obtient une extension $\mathbb{L}/\mathbb{K}$ telle que chaque P_i admettent une racine α_i dans $\mathbb{L}$. La propriété universelle de A fournit un morphisme de $\mathbb{K}$-algèbres $\varphi\colon A \to \mathbb{L}$ tel que, pour tout i, $\varphi(X_{P_i}) = \alpha_i$ et, pour tout P qui n'est pas parmi les P_i, $\varphi(P) = 0$ (cette dernière condition ne sera pas utile, on la mentionne juste pour spécifier φ entièrement). En appliquant φ à l'égalité ($\star$), on obtient $\sum_i \varphi(a_i)\varphi(P_i(X_{P_i})) = 1$. La dernière partie de la proposition 6.3.2 assure que, pour tout i, $\varphi(P_i(X_{P_i})) = P_i(\varphi(X_{P_i})) = P_i(\alpha_i) = 0$ donc on obtient $0 = 1$ dans $\mathbb{L}$, ce qui est absurde car $\mathbb{L}$ est un corps. Ainsi on a montré que I est un idéal propre.

Le théorème de Krull (proposition 4.2.14) fournit donc un idéal maximal $J \lhd A$ qui contient I et le lemme 4.2.12 assure que $\mathbb{K}_1 = A/J$ est un corps. La composée de $\mathbb{K} \to A \to A/J$ fait de $\mathbb{K}_1$ une extension de $\mathbb{K}$. Par construction de I, chaque $P \in \mathbb{K}[X]$ de degré strictement positif admet une racine dans la $\mathbb{K}$-algèbre A/I (à savoir X_P) donc a fortiori dans l'extension $\mathbb{K}_1$.

Par contre on ne sait pas si les polynômes de degré strictement positif à coefficients dans $\mathbb{K}_1$ ont tous une racine (en fait on peut montrer que c'est vrai, mais c'est plus compliqué que l'argument qui va suivre). On itère donc toute la construction pour obtenir une tour d'extension $\mathbb{K}_1$, $\mathbb{K}_2$, …. On pose $\mathbb{K}_0 = \mathbb{K}$ et on note ρ_i le morphisme d'extension de $\mathbb{K}_i$ dans $\mathbb{K}_{i+1}$ pour tout $i \in \mathbb{N}$ et $\rho_i^j = \rho_{j-1} \circ \cdots \circ \rho_i\colon \mathbb{K}_i \to \mathbb{K}_j$. On note $\mathbb{L}_0$ la « réunion » des $\mathbb{K}_i$, c'est à dire le quotient de $\bigsqcup_i \mathbb{K}_i$ par la relation qui associe $x_i \in \mathbb{K}_i$ et $x_j \in \mathbb{K}_j$ s'il existe $k \geqslant \max(i,j)$ tel que $\rho_i^k(x_i) = \rho_j^k(x_j)$. Construisons une structure de corps sur $\mathbb{L}_0$. Pour $x_i \in \mathbb{K}_i$ et $x_j \in \mathbb{K}_j$, on pose $k = \max(i,j)$ et $x_i \hat{+} x_j = \rho_i^k(x_i) + \rho_j^k(x_j)$. On vérifie que cette loi de composition interne descend à $\mathbb{L}$. On définit de même une multiplication sur $\mathbb{L}_0$ et on vérifie que $\mathbb{L}_0$ est un corps.

Montrons que $\mathbb{L}_0$ est algébriquement clos. Soit P un polynôme de degré strictement positif à coefficients dans $\mathbb{L}_0$. Comme P n'a qu'un nombre fini de coefficient, il existe N tel que ces coefficients proviennent de $\mathbb{K}_N$. Ainsi P a une racine dans $\mathbb{K}_{N+1}$ donc dans $\mathbb{L}_0$.

Enfin on note $\mathbb{L}$ l'ensemble des éléments de $\mathbb{L}_0$ qui sont algébriques sur $\mathbb{K}$. Il s'agit d'un sous-corps de $\mathbb{L}_0$ d'après la proposition 7.3.10 et l'extension $\mathbb{L}/\mathbb{K}$ est algébrique par définition. Ce $\mathbb{L}$ est algébriquement clos car tout polynôme à coefficients dans $\mathbb{L}_0$ de degré strictement positif admet une racine dans $\mathbb{L}_0$ et une telle racine est algébrique sur $\mathbb{L}$ donc sur $\mathbb{K}$ donc appartient à $\mathbb{L}$. Ainsi $\mathbb{L}$ est une clôture algébrique de $\mathbb{K}$.

Montrons maintenant l'unicité modulo isomorphisme. Soit $\mathbb{L}$ et $\mathbb{L}'$ deux clôtures algébriques de $\mathbb{K}$. Le lemme 7.6.5 appliqué à $\mathbb{E} = \mathbb{L}'$ fournit un morphisme de $\mathbb{K}$-algèbres $\rho : \mathbb{L}' \to \mathbb{L}$. On sait que ρ est automatiquement injective donc son image est un sous-corps algébriquement clos de $\mathbb{L}$ qui contient (l'image de) $\mathbb{K}$. Or $\mathbb{L}$ est algébrique sur $\mathbb{K}$ donc sur $\rho(\mathbb{L}')$ donc $\mathbb{L} = \rho(\mathbb{L}')$ et ρ est un isomorphisme d'extensions de $\mathbb{K}$. □

Table des propriétés universelles

Index

Table des matières