Cours de mathématiques
Partie III – Algèbre 1
MP2I

Alain TROESCH

Version du:

18 février 2023

Table des matières

19 Structures algébriques

THÉORÈME. Soit une équation donnée, dont a, b, c, ... sont les m racines. Il y aura toujours un groupe de permutations des lettres a, b, c, ... qui jouira de la propriété suivante :

1. *Que toute fonction des racines, invariable par les substitutions de ce groupe, soit rationnellement connue ;*
2. *Réciproquement, que toute fonction des racines, déterminable rationnellement, soit invariable par les substitutions.*

[...] Nous appellerons groupe de l'équation le groupe en question.

(Évariste Galois)

Note Historique 19.0.1

Il est fréquent de trouver des propriétés communes dans des situations qui au départ semblent totalement sans rapport. Une des grandes découvertes (et réussites) des mathématiques du 19e siècle a été de parvenir à unifier ces problèmes en apparence distincts, en faisant ressortir de ces différents problèmes des structures ensemblistes et opératoires ayant des propriétés similaires.
C'est Évariste Galois le premier à mettre en avant ces études de structure à l'occasion de ses travaux visant à étudier la résolubilité des équations polynomiales par radicaux. Il y parle de groupes de permutations des solutions d'une équation, et est amené à étudier des propriétés de certains sous-ensembles de ces groupes de permutations. C'est lui qui introduit la terminologie de « groupe », même si la formalisation précise de cette notion est beaucoup plus tardive.
Le groupe des permutations d'un ensemble avait déjà été étudié auparavant par Lagrange (mais sans en faire ressortir cette structure bien particulière de groupe). Il a notamment établi à cette occasion un résultat important, généralisé plus tard pour tout groupe sous le nom de « théorème de Lagrange ».

La notion de structure algébrique repose de façon essentielle sur la notion de loi de composition (c'est-à-dire d'opération définie sur un ensemble, comme l'addition ou la multiplication) et sur les différentes propriétés que ces lois de composition peuvent vérifier. Nous commençons donc notre étude par l'examen de ces propriétés, après avoir défini de façon précise ce qu'est une loi de composition.

I Lois de composition

I.1 Définitions

Dans ce qui suit, E est un ensemble quelconque.

Définition 19.1.1 (Lois de composition)

On distingue deux types de lois de compositions (opérations), suivant que la loi décrit une opération entre deux éléments de l'ensemble E, ou entre un élément de E et un élément d'un ensemble externe Ω, appelé domaine d'opérateur.

- Une *loi de composition interne* est une application de $\varphi : E \times E$ dans E, souvent notée de façon opérationnelle plutôt que fonctionnelle (par exemple $x + y$ au lieu de $\varphi(x, y)$ pour désigner une addition).
- Une *loi de composition externe à gauche* sur E, d'ensemble d'opérateurs Ω, est une application de $\Omega \times E$ dans E, également notée de façon opérationnelle le plus souvent (par exemple $\lambda \cdot x$ au lieu de $\varphi(\lambda, x)$).
- De même, une *loi de composition externe à droite* sur E d'ensemble d'opérateurs Ω est une application $E \times \Omega \longrightarrow E$.

Exemples 19.1.2

- Les lois $+$ et $\times$ sont des lois de composition internes sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ ou $\mathbb{C}$.
- La loi $+$ est une loi de composition interne sur $\mathbb{R}^n$ ou $\mathbb{C}^n$.
- $(\lambda, X) \mapsto \lambda X$ (multiplication d'un vecteur par un scalaire) est une loi de composition externe sur $\mathbb{R}^n$ (ou $\mathbb{C}^n$), d'ensemble d'opérateurs $\mathbb{R}$ (ou $\mathbb{C}$).
- De même pour la multiplication des polynômes par des scalaires.
- La composition $\circ$ définit une loi de composition interne sur E^E.
- Le produit scalaire sur $\mathbb{R}^n$ n'est pas une loi de composition (interne ou externe), car le résultat de l'opération n'est pas un élément de $\mathbb{R}^n$.

I.2 Propriétés d'une loi de composition

Soit E un ensemble, muni d'une loi de composition interne que nous noterons $\star$. Nous étudions ici quelques propriétés pouvant être vérifiées par la loi $\star$.

Définition 19.1.3 (Associativité, commutativité)

- On dit que $\star$ est *associative* ssi : $\forall (x, y, z) \in E^3, \quad (x \star y) \star z = x \star (y \star z)$
- On dit que $\star$ est *commutative* ssi : $\forall (x, y) \in E^2, \quad x \star y = y \star x$.

Ainsi, lorsque E est muni d'une loi associative, on peut effectuer les opérations dans l'ordre que l'on veut, à condition de respecter la position respective des éléments les uns par rapport aux autres. Si la loi est commutative, on peut échanger la position respective des éléments (mais pas nécessairement faire les opérations dans l'ordre qu'on veut si la loi n'est pas associative). Pour énoncer cette propriété d'associativité généralisée, on commence par définir ce qu'est un parenthésage admissible.

Définition 19.1.4 (Parenthésage admissible)

Un parenthésage admissible d'une expression formée de produits $\star$ d'éléments $x_1, \ldots, x_n$ de E est un parenthésage qui permet de regrouper 2 par 2 des éléments $x_1, \ldots, x_n$, ou des termes calculés à partir de ceux-ci par un parenthésage plus fin. De façon plus rigoureuse, on définit cette notion par induction structurelle :

- (initialisation) les expressions x constitués d'un unique élément sont munis d'un parenthésage admissible ;
- si A_1 et A_2 sont deux expressions en $x_1, \ldots, x_k$ et $x_{k+1}, \ldots, x_n$ munis d'un parenthésage admissible, alors $(A_1 \star A_2)$ est muni d'un parenthésage admissible.

Remarque 19.1.5

Le parenthésage le plus externe n'est pas complètement utile, et ne sert qu'à continuer la construction si d'autres termes doivent s'ajouter à l'expression. Ainsi, dans une expression munie d'un parenthésage admissible, on omet souvent le jeu de parenthèses externes. Par exemple, l'expression $(x_1 \star x_2)$ est munie d'un parenthésage admissible, mais on écrira plutôt $x_1 \star x_2$. De même, on écrira $x_1 \star (x_2 \star x_3)$ plutôt que $(x_1 \star (x_2 \star x_3))$.

Exemples 19.1.6

En utilisant la convention de la remarque précédente, lesquelles des expressions ci-dessous sont munies d'un parenthésage admissible ?

- $(x_1 \star x_2) \star (x_3 \star x_4) \star x_5$
- $(x_1 \star x_2) \star ((x_3 \star x_4) \star x_5)$
- $(x_1 \star x_2) \star x_3) \star ((x_4 \star x_5)$

Théorème 19.1.7 (Associativité généralisée)

Soit $\star$ une loi associative sur E, et $x_1, \dots, x_n$ des éléments de E. Alors la valeur de $x_1 \star x_2 \star \cdots \star x_n$ ne dépend pas du parenthésage admissible choisi sur cette expression (donc de l'ordre dans lequel on effectue ces opérations).

◁ **Éléments de preuve.**

Ce résultat qui paraît évident intuitivement n'est pas si évident que cela à démontrer. Une démonstration consiste à montrer par récurrence forte sur n (en se servant de la structure inductive), toute expression convenable parenthésée est égale à l'expression munie du parenthésage croissant, dans lequel les opérations sont faits dans l'ordre de lecture. La démonstration consiste alors à décomposer une expressions en deux, écrire l'expression de droite avec un parenthésage croissant en utilisant l'hypothèse de récurrence, utiliser l'associativité pour isoler x_n puis réutiliser l'hypothèse de récurrence sur l'expression de gauche constituée maintenant de $n-1$ termes. ▷

Notation 19.1.8 (Suppression des parenthèses)

Lorsque $\star$ est associative, nous nous permettons d'omettre le parenthésage, en notant $x \star y \star z$ au lieu de $(x \star y) \star z$ ou $x \star (y \star z)$, la propriété d'associativité levant toute ambiguité sur l'interprétation de cette expression. Plus généralement, d'après la propriété d'associativité généralisée, on peut omettre le parenthésage dans des opérations portant sur un nombre quelconque de termes.

On peut aussi donner une propriété de commutativité généralisée, lorsqu'on a à la fois l'associativité et la commutativité. Dans le cas d'une structure commutative non associative, la description est plus délicate.

Théorème 19.1.9 (Commutativité généralisée)

Soit $\star$ une loi commutative et associative sur E, et $x_1, \dots, x_n$ des éléments de E. Alors, pour tout $\sigma \in \mathfrak{S}_n$,

$$x_1 \star x_2 \star \cdots \star x_n = x_{\sigma(1)} \star x_{\sigma(2)} \star \cdots \star x_{\sigma(n)}.$$

◁ **Éléments de preuve.**

Ici encore, le théorème semble assez évident. Mais une démonstration rigoureuse nécessite un petit effort de reflexion, et l'utilisation de quelques propriétés des permutations. On peut montrer que toute permutation s'écrit comme composée de permutations très simples consistant simplement à

échanger 2 termes consécutifs, en laissant les autres fixes. C'est ce qu'on fait par exemple lorsqu'on effectue un tri à bulles, ou un tri par insertion, si on remonte les éléments au fur et à mesure. En admettant ce résultat, on passe donc de l'expression de gauche à l'expression de droite en faisant une succession d'échanges de deux termes consécutifs. Par associativité généralisée, on peut trouver un parenthésage associé qui regroupe ces deux termes, et donc l'échange de ces deux termes ne change pas la valeur de l'expression (par commutativité).

En attendant de disposer de ce résultat, on peut montrer à la main que l'échange de deux termes quelconques non nécessairement consécutifs ne change pas la valeur de l'expression, ce qui permet d'échanger x_n et $x_{\varphi(n)}$ dans l'expression de droite (si $n \neq \varphi(n)$), puis on termine par récurrence, en considérant l'expression formée des $n-1$ premiers termes. ▷

Exemples 19.1.10 (Lois commutatives, associatives)

1. Les lois $+$ et $\times$ définies sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ et $\mathbb{C}$ sont associatives et commutatives.
2. Le produit matriciel définit une loi associative sur $\mathcal{M}_n(\mathbb{R})$ (ensemble des matrices carrées d'ordre n), mais pas commutative.
3. La composition définit une loi associative sur E^E mais pas commutative.
4. La soustraction dans $\mathbb{Z}$ est non associative et non commutative.
5. La loi définie sur $\mathbb{R}$ par $(a,b) \mapsto (a+b)^2$ est commutative mais non associative.

Avertissement 19.1.11

Attention à toujours bien indiquer le parenthésage lorsque la loi n'est pas associative, ou lorsque plusieurs lois sont en jeu sans qu'il n'ait été défini de façon explicite de relation de priorité sur les opérations.

Convention 19.1.12 (Commutativité d'une loi d'addition, usage)

Nous réserverons la notation additive (signe opératoire $+$) pour des lois de composition commutatives. Cela n'empêche pas en revanche de considérer des lois commutative notées différemment (par exemple multiplicativement).

Convention 19.1.13 (Omission du signe d'opération)

Il est fréquent d'omettre certains signes d'opérations (généralement les multiplications), si l'usage qui est fait de cette suppression est suffisamment clair et ne provoque pas d'ambiguïté.

Ainsi, vous avez déjà l'habitude d'écrire ab au lieu de $a \times b$ ou $a \cdot b$. Cet usage, courant dans $\mathbb{R}$ ou $\mathbb{C}$, est aussi fréquent pour les opéraions matricielles, à la fois pour le produit interne que le produit externe (multiplication d'une matrice par un scalaire). De façon peut-être plus troublante, il est fréquent d'omettre le $\circ$ de la composition, en particulier lorsqu'on compose des applications linéaires (il n'y a alors pas d'ambiguïté sur le sens de ce produit, les éléments de l'espace d'arrivée ne pouvant en général pas se multiplier entre eux)

Notation 19.1.14 (Itération d'une loi)

Soit $\star$ une loi associative sur E, n un élément de $\mathbb{N}^*$ et x un élément de E. On note $x^{\star n}$, ou plus simplement x^n lorsqu'il n'y a pas d'ambiguïté (lorsqu'il n'y a qu'une loi en jeu par exemple), l'itéation de la loi $\star$, c'est à dire :

$$x^{\star n} = x \star x \star \cdots \star x,$$

le nombre de termes x étant égal à n. Pour une définition plus rigoureuse, par récurrence, $x^{\star 1} = x$, et pour tout $n \in \mathbb{N}^*$, $x^{\star(n+1)} = x^{\star n} \star x$.

Si E admet un élément neutre e pour la loi $\star$ (voir ci-dessous), on note par convention $x^{\star 0} = e$. Remarquez qu'alors, la définition par récurrence est aussi valable pour passer de l'exposant 0 à l'exposant 1.
Dans le cas où plusieurs lois sont en jeu, la notation x^n peut prêter à confusion. En général, dans les structures faisant intervenir deux lois dont une commutative, on utilise la multiplication $\times$ (loi multiplicative) et l'addition $+$ (loi additive). On distingue les itérations des lois sans introduire de lourdeur d'écriture en utilisant une notation particulière pour l'itération de l'addition, calquée sur ce qu'il se passe dans $\mathbb{R}$:

Notation 19.1.15 (Itération d'une loi additive)
Si E est muni d'une loi notée additivement $+$, on note $n \cdot x$ au lieu de x^{+n} l'itération de la loi $+$.

Attention au fait que généralement, n n'étant pas élément de E, la notation $\cdot$ est à distinguer d'une éventuelle multiplication dans E (cela définit une loi externe à opérateurs dans $\mathbb{N}$). Si $\mathbb{N} \subset E$, la loi externe $\cdot$ peut coïncider avec le produit, si E est muni d'une structure suffisamment riche. C'est ce qui se produit dans la plupart des structures qui contiendront $\mathbb{N}$ que nous aurons l'occasion de considérer.
Nous voyons maintenant des propriétés liées à l'existence de certains éléments particuliers de E.

Définition 19.1.16 (Élément neutre)
Soit e un élément de E. On dit que e est un *élément neutre* pour la loi $\star$ si pour tout $x \in E$, $e \star x = x = x \star e$
On trouve aussi la notion de neutre à gauche ou à droite si une seule de ces deux égalités est satisfaite.

Pour une loi commutative, e est neutre ssi e est neutre à droite ssi e est neutre à gauche.

Exemple 19.1.17 (Éléments neutres)

1. 0 est élément neutre pour $+$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$. C'est le seul élément neutre pour $+$.
2. 1 est élément neutre pour $\times$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$. C'est le seul élément neutre pour $\times$.
3. I_n est élément neutre pour $\times$ sur $\mathcal{M}_n(\mathbb{R})$, 0_n est élément neutre pour $+$ sur $\mathcal{M}_n(\mathbb{R})$.
4. id_E est élément neutre pour $\circ$ sur E^E.
5. Sur un ensemble E de cardinal supérieur ou égal à 2, la loi $(x, y) \mapsto y$ admet plusieurs neutres à gauche (tout $x \in E$ est neutre à gauche). En revanche, il n'y a pas de neutre à droite.

Une loi ne peut pas admettre plusieurs éléments neutres, comme le montre la propriété suivante.

Proposition 19.1.18 (Unicité du neutre)
L'élément neutre, s'il existe, est unique.

◁ **Éléments de preuve.**
Considérer $e_1 \star e_2$ ▷

Notation 19.1.19 (0_E, 1_E)

- On note généralement 0_E (ou 0 s'il n'y a pas de risque d'ambiguité) le neutre (s'il existe) d'une loi notée additivement $+$.

- On note généralement 1_E (ou 1 s'il n'y a pas de risque d'ambiguité) le neutre (s'il existe) d'une loi notée multiplicativement $\times$.

Définition 19.1.20 (Élément symétrique)

Supposons que E admet un élément neutre e pour la loi $\star$. Soit $x \in E$.

- On dit que y est un symétrique à gauche de x pour la loi $\star$ si $y \star x = e$.
- On dit que y est un symétrique à droite de x pour la loi $\star$ si $x \star y = e$.
- On dit que y est un symétrique de x pour la loi $\star$ si et seulement si y est un symétrique à droite et à gauche de x.
- On dit que x est symétrisable (*resp.* symétrisable à gauche, *resp.* symétrisable à droite) si x admet au moins un symétrique (*resp.* un symétrique à gauche, *resp.* un symétrique à droite).

Terminologie 19.1.21 (Opposé, inverse)

- Dans le cas d'une loi notée additivement, on parle plutôt d'opposé, et en cas d'unicité, on note $-x$ l'opposé de x.
- Dans le cas d'une loi notée multiplicativement, on parle plutôt d'inversibilité (inversibilité à droite, à gauche), et en cas d'unicité, on note x^{-1} l'inverse de x.

Dans une situation plus générale, on trouve souvent la notation x^s pour désigner le symétrique.

Proposition 19.1.22 (Unicité du symétrique)

Si $\star$ est associative, alors, en cas d'existence, le symétrique est unique.

◁ **Éléments de preuve.**

Si y et z sont deux symétriques de x, considérer $y \star x \star z$. ▷

Exemples 19.1.23

1. Dans $\mathbb{N}$ seul 0 admet un opposé pour $+$.
2. Dans $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, tout élément admet un opposé pour $+$.
3. Dans $\mathbb{N}$ seul 1 admet un inverse, dans $\mathbb{Z}$, seuls 1 et -1 admettent un inverse. Dans $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ tous les éléments non nuls admettent un inverse.
4. Dans E^E muni de $\circ$, sous réserve de l'axiome du choix, les éléments symétrisables à gauche sont les injections, les éléments symétrisables à droite sont les surjections, les éléments symétrisables sont les bijections. Une injection non surjective admet plusieurs symétriques à gauche ; une surjection non injective admet plusieurs symétriques à droite.

Proposition 19.1.24 (Symétrique de $x \star y$)

Supposons $\star$ associative. Soit $(x, y) \in E^2$. Si x et y sont symétrisables, de symétriques x^s et y^s, alors $x \star y$ est symétrisable de symétrique $y^s \star x^s$. Notez l'inversion !

◁ **Éléments de preuve.**

Le vérifier ! ▷

Traduisons pour une loi multiplicative : si x et y sont inversibles, d'inverses x^{-1} et y^{-1}, alors xy aussi, et

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

Dans le cas d'une loi additive commutative, on obtient

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

ce qu'on note plus simplement $-x - y$, comme dans $\mathbb{R}$.

Définition 19.1.25 (Élément absobant)

Soit $x \in E$.

- On dit que x est un élément absorbant à gauche pour $\star$ ssi : $\forall y \in E, \quad x \star y = x$.
- On dit que x est absorbant à droite pour $\star$ ssi : $\forall y \in E$, $y \star x = x$.
- On dit que x est absorbant s'il est à la fois absorbant à gauche et à droite.

Exemples 19.1.26

1. 0 est absorbant pour $\times$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
2. Pour la loi $(x, y) \mapsto y$, tout élément y de E est absorbant à droite. Il n'y a pas d'élément absorbant à gauche si E est de cardinal au moins 2.
3. $+\infty$ est absorbant dans $(\mathbb{R} \cup \{+\infty\}, +)$.

Définition 19.1.27 (Élément régulier ou simplifiable)

- Un élément x est dit régulier (ou simplifiable) à gauche ssi :

$$\forall (y, z) \in E^2, \ x \star y = x \star z \Longrightarrow y = z.$$

- Un élément x est dit régulier (ou simplifiable) à droite ssi :

$$\forall (y, z) \in E^2, \ y \star x = z \star x \Longrightarrow y = z.$$

- Un élément x est dit régulier (ou simplifiable) s'il est à la fois régulier à gauche et à droite.

Proposition 19.1.28 (Régularité des éléments symétrisables)

Supposons que E soit muni d'une loi $\star$ associative.

- *Soit x un élément admettant un symétrique à gauche. Alors x est régulier à gauche.*
- *Soit x un élément admettant un symétrique à droite. Alors x est régulier à droite.*
- *Soit x un élément admettant un symétrique. Alors x est régulier.*

◁ **Éléments de preuve.**

Pour simplifier, multiplier par le symétrique ! ▷

Ainsi, le fait de pouvoir simplifier une égalité par un réel ou complexe non nul x ne vient pas tant de la non nullité que de l'inversibilité de x. Par exemple, la non nullité n'est pas un critère suffisant de régularité dans $\mathcal{M}_n(\mathbb{R})$: il est nécessaire d'avoir l'inversibilité de la matrice que l'on veut simplifier.
Il convient toutefois de noter que la condition d'inversibilité, si elle est suffisante, n'est en général pas nécessaire.

Exemple 19.1.29

Donnez des exemples de structures algébriques simples dans lesquelles certains éléments sont réguliers sans être inversibles.

I.3 Ensembles munis de plusieurs lois

Soit E un ensemble muni de deux lois de composition $\star$ et $\diamond$.

Définition 19.1.30 (Distributivité)

- On dit que la loi $\star$ est distributive à gauche sur $\diamond$ ssi : $\forall (x,y,z) \in E^3,\ \ x\star(y\diamond z) = (x\star y)\diamond(x\star z)$.
- On dit que la loi $\star$ est distributive à droite sur $\diamond$ ssi : $\forall (x,y,z) \in E^3,\ \ (y\diamond z)\star x = (y\star x)\diamond(z\star x)$.
- On dit que la loi $\star$ est distributive sur $\diamond$ ssi elle est distributive à droite et à gauche.

Exemples 19.1.31

1. La loi $\times$ est distributive sur $+$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$, $\mathcal{M}_n(\mathbb{R})$...
2. La loi $\cap$ est distributive sur $\cup$ sur $\mathcal{P}(X)$. Inversement, la loi $\cup$ est distributive sur $\cap$.
3. Que peut-on dire de la loi $\cap$ par rapport à elle-même ? De la loi $\cup$?

Remarque 19.1.32

La première relation $x \star (y \diamond z) = (x \star y) \diamond (x \star z)$ a également un sens lorsque $\star$ est une loi externe. Ainsi, dans $\mathbb{R}^n$, muni de l'addition en loi interne, et de la multiplication des scalaires en loi externe, on a $\lambda(X+Y) = \lambda X + \lambda Y$: la loi externe est distributive sur la loi interne.

Théorème 19.1.33 (Distributivité généralisée)

Soit E muni de deux lois $\times$ et $+$ associatives, et $+$ commutative. On suppose que $\times$ est distributive sur $+$. On a alors, pour tout $n \in \mathbb{N}^*$ et tous ensembles finis $J_1, \dots, J_n$ non vides, les $x_{i,j}$ étant des éléments de E, on a :

$$\prod_{i=1}^{n} \sum_{j \in J_i} x_{i,j} = \sum_{(j_1,\dots,j_n) \in J_1 \times \cdots \times J_n} \prod_{i=1}^{n} x_{i,j_i}.$$

La loi $\times$ n'étant pas supposée commutative, les produits $\prod$ sont à comprendre dans l'ordre croissant des indices.

◁ **Éléments de preuve.**

Le montrer pour $n = 2$, puis récurrence sur n. La démonstration est exactement la même que celle faite dans le chapitre sur les sommes dans le cas des nombres réels. ▷

Exemple 19.1.34

Comprendre la formule ci-dessus pour l'expression $(x_1 + x_2) \times (y_1 + y_2 + y_3) \times (z_1 + z_3)$.

Définition 19.1.35 (associativité externe)

Soit E un ensemble muni d'une loi de composition externe $\diamond$ sur $\mathbb{K}$, lui même muni d'une loi de composition interne $\star$. On dit que les lois $\star$ et $\diamond$ vérifient une propriété d'associativité externe si pour tout $(\lambda, \mu) \in \mathbb{K}^2$ et $x \in E$

$$(\lambda \star \mu) \diamond x = \lambda \diamond (\mu \diamond x).$$

Cette propriété est par exemple satisfaite sur $\mathbb{R}^n$ pour la multiplication par un scalaire : $(\lambda\mu)X = \lambda(\mu X)$. Plus généralement, un espace vectoriel vérifiera cette propriété d'associativité externe.
De la même manière, si E est muni d'une loi interne $\times$ et d'une loi externe $\diamond$ à opérateurs dans $\mathbb{K}$, on dispose d'une autre propriété de distributivité externe s'exprimant par la relation :

$$\lambda \diamond (x \times y) = (\lambda \diamond x) \times y.$$

Dans cette situation, qu'on retrouve notamment dans la définition des $\mathbb{K}$-algèbres, on pourra éventuellement avoir en plus l'égalité avec $x \times (\lambda \diamond y)$.
C'est une propriété qu'on retrouve par exemple pour l'ensemble des matrices carrées d'ordre n, muni de la loi interne $\times$ (produit matriciel) et de la loi externe correspondant au produit d'une matrice par un scalaire.

I.4 Stabilité

Définition 19.1.36

Soit E un ensemble muni d'une loi $\star$ et $F \subset E$ un sous-ensemble de E. On dit que F est stable par $\star$, si la restriction de la loi de E à $F \times F$ peut se corestreindre à F, autrement dit si :

$$\forall (x, y) \in F^2,\ x \star y \in F.$$

Dans ce cas, la loi de E se restreint en une loi $\star_F : F \times F \to F$, appelée *loi induite sur F par $\star$*.

II Structures

II.1 Généralités

Définition 19.2.1

- Une structure de « truc » est la donnée d'un certain nombre d'axiomes (définissant ce qu'on appelle un « truc ») portant sur un ensemble fini de lois de composition (internes et/ou externes).
- On dit qu'un ensemble E est muni d'une structure de truc ssi E est muni d'un nombre fini de lois de composition vérifiant les axiomes de structure de truc.

Exemples 19.2.2

1. Une structure de magma se définit comme la donnée d'une loi de composition, et un ensemble vide d'axiomes. Ainsi, tout ensemble E muni d'une loi de composition est muni d'une structure de magma.
2. Une structure de monoïde se définit comme la donnée d'une loi de composition, et de deux axiomes : l'associativité de la loi et l'existence d'un élément neutre. Par exemple $(\mathbb{N}, +)$ est muni d'une structure de monoïde (on dit plus simplement que $(\mathbb{N}, +)$ est un monoïde). L'ensemble des mots sur un alphabet $\mathcal{A}$, muni de l'opération de concaténation est aussi un monoïde (appelé monoïde libre sur l'alphabet $\mathcal{A}$). Contrairement à $\mathbb{N}$, le monoïde libre n'est pas commutatif.
3. Ainsi, la structure de monoïde est plus riche que celle de magma : tout monoïde est aussi un magma ; un monoïde peut être défini comme un magma dont la loi est associative et possède un élément neutre.
4. Une structure de groupe est une structure de monoïde à laquelle on rajoute l'axiome d'existence de symétriques. Par exemple $(\mathbb{Z}, +)$ est un groupe, mais pas $(\mathbb{N}, +)$.

Définition 19.2.3 (Stucture induite)

Soit E un ensemble muni d'une structure de truc, et F un sous-ensemble de E. Si F est stable pour chacune des lois de E, l'ensemble F muni des lois induites sur F par les lois de E est muni d'une structure appelé structure induite sur F par la structure de E.

Avertissement 19.2.4

En général, F ne peut pas être muni d'une structure de truc, mais seulement d'une structure moins riche, certains des axiomes de la structure de truc pouvant ne pas être préservée par restriction.

Exemple 19.2.5

$(\mathbb{N}, +)$ est la structure induite sur $\mathbb{N}$ par la structure de groupe additif de $(\mathbb{Z}, +)$. En revanche, $(\mathbb{N}, +)$ n'est pas un groupe. On a perdu l'existence des opposés par restriction.

Définition 19.2.6 (Sous-truc)

Soit E un ensemble muni d'une structure de truc et F un sous-ensemble de E. On dit que F est un sous-truc de E si F est stable par les lois de E, si F contient les neutres imposés de E, et si les lois induites sur F par les lois de E vérifient les axiomes de la structure de truc.

Nous verrons comment traduire de façon effective cette notion dans le cas de sous-groupes et sous-anneaux.

Remarque 19.2.7 (Restriction des propriétés universelles)

Toutes les propriétés universelles (quantifiées par $\forall$) passent bien aux structures induites. Ainsi, la commutativité, l'associativité, la distributivité, la régularité passent aux structures induites.

II.2 Morphismes

Lorsqu'on dispose d'une structure de truc, on est souvent amené à considérer des applications entre ensembles munis de la structure de truc. Cependant seules nous intéressent les applications compatibles dans un certain sens avec la structure de truc : les autres ne sont pas pertinentes dans le contexte (si on a à s'en servir, c'est qu'on sort de la structure de truc, et que la structure de truc n'est plus le contexte adapté).

Définition 19.2.8 (Homomorphisme)

Soit E et F deux ensembles munis d'une structure de truc, E étant muni des lois de composition interne $(\underset{1}{\star}, \ldots, \underset{n}{\star})$ et F des lois $(\underset{1}{\diamond}, \ldots, \underset{n}{\diamond})$, et des lois de composition externes $(\underset{1}{\top}, \ldots, \underset{m}{\top})$ et $(\underset{1}{\bot}, \ldots, \underset{m}{\bot})$ sur $K_1, \ldots, K_m$ respectivement. On dit qu'une application $f : E \longrightarrow F$ est un homomorphisme de trucs (ou plus simplement un morphisme de trucs) ssi :

- L'application f est compatible avec (ou « respecte ») les lois internes :
$$\forall k \in [\![1, n]\!], \quad \forall (x, y) \in E^2, \quad f(x \underset{k}{\star} y) = f(x) \underset{k}{\diamond} f(y).$$
- L'application f est compatible avec (ou « respecte ») les lois externes :
$$\forall \ell \in [\![1, m]\!], \quad \forall \lambda \in K_\ell, \quad \forall x \in E, \quad f(\lambda \underset{\ell}{\top} x) = \lambda \underset{\ell}{\bot} f(x).$$
- Si l'existence du neutre e_i pour la loi $\underset{i}{\star}$ est imposée dans les axiomes (et donc le neutre e'_i pour la loi $\underset{i}{\diamond}$ existe aussi), f doit être compatible avec le neutre : $f(e_i) = e'_i$.

On peut avoir à rajouter certaines propriétés liées à la structure étudiée. On peut aussi ajouter l'existence d'un homomorphisme nul (ne vérifiant pas la compatibilité avec les neutres non additifs), afin d'obtenir une structure intéressante sur l'ensemble des morphismes.

Ainsi, un homomorphisme entre deux ensembles muni d'une certaine structure est une application « respectant » cette structure.

Pour chaque structure étudiée, nous redéfinirons de façon précise la notion d'homomorphisme associée, si celle-ci est à connaître. Nous donnons une propriété générale, dont la démonstration dans le cadre général nous dispensera des démonstrations au cas par cas.

Proposition 19.2.9 (Composition d'homomorphismes)
Soit $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux morphismes de trucs. Alors $g \circ f$ est un morphisme de trucs.

◁ **Éléments de preuve.**
Vérifier en deux temps le respect par $g \circ f$ de chaque loi interne, chaque loi externe, chaque neutre imposé. ▷

Nous définissons alors :

Terminologie 19.2.10
- Un isomorphisme de trucs est un homomorphisme de truc bijectif.
- Un endomorphisme de truc est un homomorphisme de truc de E dans lui-même (muni des mêmes lois) (il n'y a ici qu'un truc en jeu, ce qui justifie le singulier)
- Un automorphisme de truc est un endomorphisme qui est également un isomorphisme.

Proposition 19.2.11
Si $f : E \longrightarrow F$ est un isomorphisme de trucs, alors f^{-1} est un isomorphisme de trucs.

Ainsi, la réciproque d'un isomorphisme est bijective (ça, ce n'est pas une surprise), et c'est aussi un homomorphisme de trucs (ce qui est moins trivial).

◁ **Éléments de preuve.**
C'est ce dernier point qu'il faut vérifier. Pour une loi interne $\star$, comparer $f(f^{-1}(a) \star f^{-1}(b))$ et $f(f^{-1}(a \star b))$. Même principe pour la loi externe. Le respect des neutres est évident. ▷

II.3 Catégories (HP)

La notion de structure et de morphisme associé permet de définir la notion de catégorie. Grossièrement, une catégorie est la donnée :
- d'une classe d'objets ;
- de flèches entre ces objets ;
- d'une règle de composition entre les flèches.

Par exemple, la catégorie des monoïdes est la catégorie dont les objets sont les monoïdes, les flèches sont les homomorphismes de monoïdes, et la composition des flèches correspond à la composition usuelle des homomorphismes (la composée de deux homomophismes étant encore un homomorphisme, donc une flèche de la catégorie). On définit de même la catégorie des groupes, ou la catégorie des anneaux, ou encore la catégorie des corps.

Cette notion de catégorie nous permet de travailler dans un certain contexte. Se donner une catégorie permet de se concentrer sur un certain type d'objets, et un certain type d'applications, et de les étudier d'un point de vue formel.

La notion de catégorie dépasse largement le cadre de l'étude des structures algébriques, car si les structures algébriques fournissent des catégories, de nombreuses catégories sont issues d'autres contextes, comme :

- la catégorie des ensembles, les morphismes étant toutes les applications ;
- la catégorie des ensembles ordonnés, les morphismes étant les applications croissantes
- la catégorie des espaces topologiques, les morphismes étant les applications continues
- ou encore, la catégorie des catégories, les morphismes étant les foncteurs de C dans D, associant à chaque objet de C un objet de D, et à chaque flèche de C une flèche de D, tout en respectant un certain nombre de règles de compatibilité.
- ou encore, des catégories de foncteurs entre deux catégories, les objets étant cette fois des foncteurs, et les flèches étant des « transformations naturelles » entre foncteurs...

III Groupes

III.1 Axiomatique de la structure groupes

Définition 19.3.1 (Groupe)

Soit G un ensemble. On dit que G est muni d'une structure de groupe si G est muni d'une loi de composition $\star$ telle que :

- $\star$ est associative ;
- il existe un élément neutre e pour la loi $\star$;
- tout élément x admet un symétrique x^s.

En vertu de résultats antérieurs, on peut énoncer :

Proposition 19.3.2 (Unicité du neutre et des symétriques)

Soit $(G, \star)$ un groupe. Alors :

- *G admet un unique élément neutre pour $\star$*
- *Pour tout $x \in G$, il existe un unique symétrique x^s de x.*

◁ **Éléments de preuve.**

Provient de résultats déjà vus ▷

Corollaire 19.3.3 (régularité des éléments d'un groupe)

Tous les éléments d'un groupe sont réguliers pour la loi du groupe.

◁ **Éléments de preuve.**

De même. ▷

Définition 19.3.4 (Groupe abélien ou commutatif)

On dit qu'un groupe $(G, \star)$ est abélien (ou commutatif) si la loi de G est commutative.

Notation 19.3.5 (Notation additive, notation multiplicative)

La loi d'un groupe est le plus souvent notée additivement (signe $+$) ou multiplicativement (signe $\times$, parfois remplacé par un point, voire omis, comme dans $\mathbb{R}$). La notation additive est réservée au cas de groupes abéliens. Nous avons alors les notations suivantes pour désigner des itérées de la loi de composition sur un même élément x :

- loi multiplicative : $x \times \cdots \times x$ (avec n occurrences) est noté x^n ;
  le neutre est noté 1 ;
  par convention, $x^0 = 1$;
- loi additive : $x + \cdots + x$ (avec n occurrences) est noté $n \cdot x$ ou nx ;
  le neutre est noté 0 ;
  par convention $0x = 0$.

Une définition plus rigoureuse par récurrence pourrait être donnée pour ces itérées.

Notation 19.3.6 (Simplifications d'écriture pour la notation additive)

Soit $(G, +)$ un groupe commutatif. Comme mentionné plus haut, l'opposé d'un élément x est noté $-x$. On note alors $x - y$ au lieu de $x + (-y)$. On a alors les règles suivantes :

- $\forall (x, y, z) \in G^3, \quad x - (y + z) = x - y - z$
- $\forall (x, y, z) \in G^3, \quad x - (y - z) = x - y + z$.

En vertu des définitions générales données dans le paragraphe précédent, nous donnons la définition suivante :

Définition 19.3.7 (Homomorphisme de groupes)

Soit $(G, \star)$ et $(H, \diamond)$ deux groupes.

- On dit qu'une application $f : G \longrightarrow H$ est un homomorphisme de groupes (ou plus simplement morphisme de groupes) si pour tout $(x, y) \in G$, $f(x \star y) = f(x) \diamond f(y)$.
  On note $\mathrm{Hom}(G, H)$ l'ensemble des homomorphismes de G dans H.
- Si $(G, \star) = (H, \diamond)$, on dit que f est un endomorphisme de $(G, \star)$.
- Un homomorphisme bijectif est appelé isomorphisme ; en vertu de ce qui précède, la réciproque d'un isomorphisme est un isomorphisme.
- Un endomorphisme bijectif est appelé automorphisme ; en vertu de ce qui précède, la réciproque d'un automorphisme est un automorphisme.
  On note $\mathrm{Aut}(G)$ l'ensemble des automophismes de G.

Le respect du neutre n'a pas été imposé. Et pour cause : il est ici automatiquement vérifié :

Proposition 19.3.8 (Image du neutre par un morphisme)

Soit $f : G \to H$ un morphisme de groupes. Alors $f(e_G) = e_H$.

◁ **Éléments de preuve.**
Considérer $f(e_G \star e_G)$ et utiliser la régularité dans H. ▷

Proposition 19.3.9 (Image par un morphisme d'un inverse)

Soit G, H deux groupes (notés multiplicativement), et f un morphisme de G dans H. Alors $f(x^{-1}) = f(x)^{-1}$. On adaptera aisément cette propriété au cas ou l'un (ou les deux) groupe(s) est (sont) en notation additive.

◁ **Éléments de preuve.**
Considérer $f(x)f(x^{-1})$. ▷

Proposition 19.3.10 (Structure de $(\mathrm{Aut}(G), \circ)$)
Soit G un groupe. Alors, $(\mathrm{Aut}(G), \circ)$ est un groupe.

◁ **Éléments de preuve.**
Utiliser les résultats relatifs aux composées de morphismes et à la réciproque d'un isomophisme. ▷

III.2 Exemples importants

Exemples 19.3.11

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ sont des groupes abéliens notés additivement.
2. $(\mathbb{Q}^*, \times)$, $(\mathbb{R}^*, \times)$, $(\mathbb{C}^*, \times)$, $(\mathbb{Q}_+^*, \times)$, $(\mathbb{R}_+^*, \times)$ sont des groupes abéliens notés multiplicativement.
3. $(\mathbb{N}, +)$, $(\mathbb{Q}, \times)$, $(\mathbb{Z}^*, \times)$, « $(\mathbb{R}_-^*, \times)$ » sont-ils des groupes ?
4. $(\mathbb{U}, \times)$ et $(\mathbb{U}_n, \times)$ sont des groupes.
5. Pour $n \geqslant 2$, $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe.
6. $((\mathbb{Z}/n\mathbb{Z}) \setminus \{0\}, \times)$ est-il en général un groupe ?
7. Étant donné X un ensemble, $(\mathfrak{S}_X, \circ)$, l'ensemble des permutations de X est un groupe pour la loi définie par la composition.
8. $\exp : (\mathbb{R}, +) \longrightarrow (\mathbb{R}_+^*, \times)$ est un homomorphisme de groupes. C'est même un isomorphisme.
9. Sa réciproque est donc aussi un isomorphisme de groupes : $\ln : (\mathbb{R}_+^*, \times) \longrightarrow (\mathbb{R}, +)$.
10. L'application $x \mapsto \mathrm{e}^{\mathrm{i}\,x}$ est un morphisme de groupes de $(\mathbb{R}, +)$ à $(\mathbb{U}, \times)$.
11. L'application $z \mapsto \mathrm{e}^z$ est un morphisme de groupes surjectif (mais non injectif) de $(\mathbb{C}, +)$ sur $(\mathbb{C}^*, \times)$.
12. Soit $n \geqslant 2$ et pour tout $k \in [\![0, n-1]\!]$, $\omega_k = \mathrm{e}^{\mathrm{i}\,\frac{2k\pi}{n}}$. Alors, étant donné $k \in [\![0, n]\!]$:
$$\begin{array}{rccc} f : & (\mathbb{Z}, +) & \longrightarrow & (\mathbb{U}_n, \times) \\ & \ell & \mapsto & \omega_k^\ell \end{array}$$
est un homomorphisme de groupe. Il est surjectif si $k = 1$, et plus généralement si k est premier avec n (d'après le théorème de Bézout). On dit dans ce cas que ω_k est une racine primitive de l'unité (car elle engendre multiplicativement l'ensemble de toutes les racines de l'unité).
13. Puisque $\omega_k^n = 1$, le morphisme précédent « passe au quotient » et définit un homomorphisme de groupe :
$$\begin{array}{rccc} f : & (\mathbb{Z}/n\mathbb{Z}, +) & \longrightarrow & (\mathbb{U}_n, \times) \\ & \ell & \mapsto & \omega_k^\ell \end{array}$$
Cet homomorphisme est un isomophisme si $k = 1$, et plus généralement si k est premier avec n.

III.3 Sous-groupes

Toujours en suivant les définitions plus générales, nous donnons la définition suivante :

Définition 19.3.12 (Sous-groupe)
Soit $(G, \star)$ un groupe. Un sous-ensemble H de G est appelé *sous-groupe de G* si H est stable pour la loi de G et si la loi induite définit sur H une structure de groupe.

Remarquez qu'on n'a pas donné l'appartenance du neutre à H dans la définition, celle-ci étant automatique en vertu de :

Proposition 19.3.13 (Appartenance de l'élément neutre à H)

Soit H un sous-groupe de G. Alors l'élément neutre e de G est dans H et est l'élément neutre du groupe H.

◁ **Éléments de preuve.**

Considérer $e_G \cdot h$ et $e_H \cdot h$, pour $h \in H$. ▷

Dans la pratique, pour vérifier que H est un sous-groupe de G on utilise le résultat suivant, ou sa version compactée :

Théorème 19.3.14 (Caractérisation des sous-groupes)

Un sous-ensemble H d'un groupe $(G, \star)$ (de neutre e_G) est un sous-groupe de G si et seulement si :

(i) H *est non vide,*

(ii) H *est stable pour* $\star$ *:* $\forall (x, y) \in H, \quad x \star y \in H$,

(iii) H *est stable par prise de symétrique :* $\forall x \in H, \quad x^s \in H$.

◁ **Éléments de preuve.**

CN : assez évidente, se servir du fait que le neutre est le même pour le troisième point.

CS : provient du fait que les propriétés universelles (donc l'associativité) se conservent par restriction. Le point (ii) nous assure de la bonne définition de la loi. ▷

On peut rassembler les deux dernières propriétés en une seule vérification :

Théorème 19.3.15 (Caractérisation des sous-groupes, version condensée)

Un sous-ensemble H d'un groupe $(G, \star)$ (de neutre e_G) est un sous-groupe de G si et seulement si :

(i) H *est non vide,*

(ii) $\forall (x, y) \in H^2, \quad x \star y^s \in H$,

◁ **Éléments de preuve.**

Comparer au théorème précédent : un sens est évident. Il suffit de montrer qu'avec les 3 points de ce théorème, on peut séparer la stabilité par somme et par inversion. Commencer par justifier que $e = e_G \in H$, puis considérer $x = e$ pour la stabilité par inverse, puis $y' = y^s$ pour la stabilité par produit. ▷

Proposition 19.3.16

Dans les deux théorèmes précédents, on peut remplacer le point (i) par :

(i') $e_G \in H$.

◁ **Éléments de preuve.**

Si H est un sous-groupe, il contient nécessairement e_G. ▷

On traduit cette dernière propriété dans les deux cas les plus fréquents :

- pour un sous-groupe d'un groupe additif, la vérification de stabilité à faire est donc :

$$\forall (x, y) \in H^2, \ x - y \in H;$$

- pour un sous-groupe d'un groupe multiplicatif, la vérification de stabilité à faire est donc :

$$\forall (x,y) \in H^2, \quad xy^{-1} \in H.$$

Dans certaines situations, il est plus commode de dissocier l'étude de la stabilité par produit et de la stabilité par inversion.

De façon évidente, étant donné G un groupe, d'élément neutre e, $\{e\}$ et G sont des sous-groupes de G.

Définition 19.3.17 (Sous-groupe propre)
Un sous-groupe propre de G est un sous-groupe de G distinct de $\{e\}$ et de G.

Proposition 19.3.18 (Intersection de sous-groupes)
Soit G un groupe, et $(H_i)_{i\in I}$ une famille de sous-groupes de G. Alors $\bigcap_{i\in I} H_i$ est un sous-groupe de G.

◁ **Éléments de preuve.**
Sans difficulté à l'aide de l'une ou l'autre des caractérisations des sous-groupes. ▷

Remarque 19.3.19
Une union de sous-groupes est-elle un sous-groupe ?

Proposition 19.3.20 (Image directe et réciproque de sous-groupes par un homomorphisme)
Soit G et H deux groupes, et soit $f \in \mathrm{Hom}(G,H)$ un morphisme de groupes. Alors l'image par f de tout sous-groupe de G est un sous-groupe de H. L'image réciproque par f de tout sous-groupe de H est un sous-groupe de G.

◁ **Éléments de preuve.**
Vérification facile par caractérisation. ▷

Définition 19.3.21 (Noyau)
Soit G et H deux groupes et $f \in \mathrm{Hom}(G,H)$ un morphisme de groupes. Le noyau de f est le sous-groupe de G défini par :

$$\mathrm{Ker}(f) = f^{-1}(\{e_H\}) = \{y \in G \mid f(y) = e_H\}.$$

◁ **Éléments de preuve.**
Cas particulier de la proposition précédente. ▷

Une propriété importante du noyau est qu'il mesure le défaut d'injectivité :

Théorème 19.3.22 (Caractérisation de l'injectivité)
Soit $f \in \mathrm{Hom}(G,H)$. Le morphisme f est injectif si et seulement si $\mathrm{Ker}(f) = \{e_G\}$.

◁ **Éléments de preuve.**
Remarquer que (en notation multiplicative), $f(x) = f(y)$ équivaut à $f(xy^{-1}) = e_H$. ▷

Plus généralement, on obtient la description des images réciproques des singletons :

Proposition 19.3.23

Soit G un groupe multiplicatif. Soit f un morphisme de $\mathrm{Hom}(G,H)$ et $y \in H$. Soit $x \in f^{-1}(\{y\})$. On a alors

$$f^{-1}(\{y\}) = x \times \mathrm{Ker}(f) = \{x \times z, z \in \mathrm{Ker}(f)\} = \mathrm{Ker}(f) \times x.$$

Ces ensembles sont les fibres *de f.*

La notion de fibre est assez limpide, surtout si on la traduit dans un contexte additif : il s'agit alors d'ensembles $x + \mathrm{Ker}(f)$, autrement dit de translatés du noyau.

Exemple 19.3.24

On a déjà croisé des fibres de morphismes de groupes (c'était même des « applications linéaires »). À quelle(s) occasion(s) ?

III.4 Sous-groupes engendrés par une partie, sous-groupes monogènes

Définition 19.3.25 (Sous-groupe engendré par une partie)

Soit $(G,\times)$ un groupe, et X une partie de G. Le sous-groupe $\langle X\rangle$ de G engendré par X est le plus petit sous-groupe de G contenant X. On trouve aussi souvent la notation $\mathrm{Gr}(X)$.

Proposition 19.3.26 (Description par le haut du sous-groupe engendré par une partie)

Soit X une partie d'une groupe G. Alors :

$$\langle X\rangle = \bigcap_{X\subset H} H,$$

l'intersection étant prise sur tous les sous-groupes H de G contenant X.

◁ **Éléments de preuve.**
L'intersection est bien définie elle est constituée d'au moins un terme), elle est un sous-groupe, et est évidemment minimale. ▷

Proposition 19.3.27 (Description par le bas du sous-groupe engendré par une partie)

Soit X une partie d'une groupe G. Alors $\langle X\rangle$ est l'ensemble des éléments pouvant s'écrire sous la forme $x_1\cdots x_n$, $n\in\mathbb{N}$, où les x_i vérifient soit $x_i\in X$, soit $x_i^{-1}\in X$. Le produit vide est par convention égal au neutre e de G.

◁ **Éléments de preuve.**
Justifier que l'ensemble de ces éléments est forcément inclus dans $\langle X\rangle$, et que c'est une sous-groupe contenant X. ▷

Définition 19.3.28 (Sous-groupe monogène)

1. Soit $X=\{x\}$ un singleton d'un groupe G. Alors $\langle X\rangle$ est appelé sous-groupe monogène engendré par x :
2. Concrètement, le sous-groupe monogène engendré par x est :

$$\langle x\rangle = \{x^n,\ n\in\mathbb{Z}\} \qquad \text{en notation multiplicative}$$

En notation additive, c'est donc l'ensemble des nx, pour $n\in\mathbb{Z}$

3. Un sous-groupe H est dit monogène s'il existe x tel que H soit le sous-groupe monogène engendré par x.
4. Un groupe est dit monogène s'il est un sous-groupe monogène de lui-même.
5. Un groupe est dit cyclique s'il est monogène et fini.

Proposition 19.3.29 (Commutativité d'un groupe monogène)
Un groupe monogène est abélien

◁ **Éléments de preuve.**
C'est une question d'associativité généralisée. ▷

Remarque 19.3.30
À quelle condition nécessaire et suffisante sur $X \subset G$ le sous-groupe $\langle X \rangle$ est-il abélien ?

III.5 Sous-groupes de $\mathbb{Z}$ et $\mathbb{R}$

Théorème 19.3.31 (Sous-groupes de $\mathbb{Z}$)
Les sous-groupes de $\mathbb{Z}$ sont exactement les $n\mathbb{Z}$, $n \in \mathbb{N}$.

◁ **Éléments de preuve.**
Considérer n l'élément minimal de $G \cap \mathbb{N}^*$. Justifier que $n\mathbb{Z} \subset G$, et si l'inclusion est stricte, contredire la minimalité de n en effectuant une division euclidienne. ▷

Les sous-groupes de $\mathbb{Z}$ sont donc tous des groupes monogènes (additifs) .
Nous avons déjà eu l'occasion de prouver en exercice le résultat suivant, pour l'employer dans des situations particulières :

Proposition 19.3.32 (Sous-groupes additifs de $\mathbb{R}$, HP)
Les sous-groupes de $(\mathbb{R}, +)$ sont soit égaux à $a\mathbb{Z}$, $a \in \mathbb{R}_+$, soit denses dans $\mathbb{R}$.

◁ **Éléments de preuve.**
Un peu le même principe pour commencer : soit $a = \inf(G \cap \mathbb{R}_+^*)$.
- si $a > 0$, justifier que $a \in G$ (sinon il existe des éléments de G aussi proches qu'on veut les uns des autres), et terminer comme pour les sous-groupes de $\mathbb{Z}$
- Sinon, il existe des éléments de G aussi petits qu'on veut à partir desquels on peut faire un balayage de $\mathbb{R}$, pour venir en insérer entre deux réels donnés x et y. C'est le même principe que la démonstration de la densité de $\mathbb{Q}$.

▷

Exemple 19.3.33
Ainsi que nous l'avons évoqué dans un chapitre antérieur, l'ensemble des périodes d'une fonction $f : \mathbb{R} \to \mathbb{R}$ est soit de la forme $T\mathbb{Z}$ soit dense dans $\mathbb{R}$.

III.6 Congruences modulo un sous-groupe

Soit G un groupe (multiplicatif) et H un sous-groupe de G.

Définition 19.3.34 (Classes à droite et à gauche modulo H, Spé)

- Les classes à droite modulo H sont les ensembles Ha, $a \in G$.
- Les classes à gauche modulo H sont les ensembles aH, $a \in G$.

Remarque 19.3.35

Les classes à gauche et à droite ne sont en général pas des sous-groupes. Donner une CNS sur a pour que aH soit un sous-groupe de G.

Proposition 19.3.36 (Les classes modulo H sont des classes d'équivalence)

(i) *x et y sont dans la même classe à gauche modulo H ssi $xy^{-1} \in H$*

(ii) *La relation $\equiv_g ...[H]$ définie sur G par*

$$x \equiv_g y \ \ [H] \iff xy^{-1} \in H$$

est une relation d'équivalence, appelée relation de congruence à gauche modulo H.

(iii) *Ainsi, les classes à gauche sont les classes d'équivalence de la relation $\equiv_g ...[H]$*

(iv) *De même, les classes à droite sont les classes d'équivalence de la relation $\equiv_d ...[H]$ définie sur G par*

$$x \equiv_d y \ \ [H] \iff x^{-1}y \in H.$$

Proposition 19.3.37 (Partition des classes à gauche, à droite)

L'ensemble $\{aH, a \in G\}$ des classes à gauche modulo H est une partition de G. De même pour l'ensemble $\{Ha, a \in G\}$ des classes à droite modulo H.

On en déduit le théorème de Lagrange, qui est l'un des résultats élémentaires les plus importantes sur les cardinaux des groupes finis.

Définition 19.3.38 (Ordre d'un groupe fini)

Soit G un groupe fini. L'ordre de G est par définition son cardinal.

Lemme 19.3.39 (Cardinal des classes de congruence)

Soit G un groupe fini, et H un sous-groupe de G. Pour tout $a \in G$, $|aH| = |Ha| = |H|$.

◁ **Éléments de preuve.**

$h \mapsto ah$ est une bijection de H sur aH. ▷

Théorème 19.3.40 (Lagrange, Spé)

Soit G un groupe fini, et H un sous-groupe de G. Alors l'ordre de H divise l'ordre de G.

◁ **Éléments de preuve.**

Les parts de la partition sont toutes de même taille. Ainsi, si k est leur nombre, $|G| = k \cdot |H|$. ▷

Proposition 19.3.41 (Passage au quotient de la loi dans le cas abélien)

- *Si G est un groupe abélien, et H un sous-groupe de G, alors $\equiv_d$ et $\equiv_g$ modulo H sont égales.*
- *Cette relation, plus simplement notée $\equiv ...[H]$, est une congruence pour la loi de G*
- *La loi induite correspond au produit des classes éléments par éléments :*

$$(ab)H = (aH) \cdot (bH) = \{x \cdot y,\ x \in aH,\ y \in bH\}$$

- *La loi induite sur l'ensemble quotient munit celui-ci d'une structure de groupe abélien.*

◁ **Éléments de preuve.**

- Le premier point est immédiat
- le deuxième se vérifie facilement en considérant $(xy)(x'y')^{-1}$ et en utilisant la commutativité pour réordonner les choses.
- Par double inclusion et commutativité, c'est assez immédiat.
- Le point précédent et l'associativité de G permettent de faire toutes les manipulations requises sur les groupes pour obtenir ce résultat. On pourra remarquer que $H \cdot H = H$ (au sens du produit élément par élément pour simplifier).

▷

Définition 19.3.42 (Groupe quotient d'un groupe abélien)

Sous les hypothèses précédentes, on note G/H l'ensemble quotient, muni de sa structure de groupe. Le groupe G/H est appelé groupe quotient de G par H

On voit maintenant comment généraliser cette construction à un groupe G quelconque, avec certaines conditions sur le sous-groupe G. Ce qui fait bien marcher les vérifications précédentes, c'est le fait d'une part que $aH = Ha$ et d'autre part qu'on a une description explicite élément par élément.
On remarquera que la relation $aH = Ha$ imposée pour tout $a \in G$ signifie que les relations de congruence à gauche et à droite sont identiques (puisqu'elles ont les mêmes classes), ce qui nous fait retomber sur le premier point de la propriété précédente.

Proposition/Définition 19.3.43 (Sous-groupe distingué, HP)

Soit G un groupe, et H un sous-groupe de G. Les deux propriétés suivantes sont équivalentes :

(i) $\forall a \in G,\ aH = Ha$

(ii) $\forall a \in G, \forall h \in H,\ aha^{-1} \in H$.

Un sous-groupe H vérifiant cette propriété est appelé sous-groupe distingué (ou sous-groupe normal) de G.

◁ **Éléments de preuve.**

(i) implique (ii) de façon immédiate. Réciproquement, (ii) signifie que tout ah peut s'écrire $h'a$ pour un certain h'. cela donne une inclusion, puis une deuxième par symétrie. ▷

Théorème 19.3.44 (Structure de groupe du quotient, HP)

Soit H un sous-groupe distingué de G. Alors :

- *les relations de congruence à gauche et à droite sont égales. On note $\equiv ...[H]$ cette relation.*
- *$\equiv ...[H]$ est une congruence pour la loi de G. Celle-ci passe au quotient.*
- *La loi sur le quotient se traduit par le produit élément par élément :*

$$(ab)H = (aH)(bH) = \{x \cdot y,\ x \in aH,\ y \in bH\}.$$

- *La loi induite sur le quotient munit celui-ci d'une structure de groupe.*

◁ **Éléments de preuve.**

- Elles ont les mêmes classes
- Mêmes vérifications que dans le cas abélien, mais en remmplaçant la propriété de commutativité par le fait qu'on peut presque commuter un élément par un élément de H, quitte à remplacer l'élément de H par un autre. Plus précisément, ah peut se réécrire $h'a$, pour un $h' \in H$.
- Même principe
- Les vérifications se font de même que dans le cas abélien, puisque basées sur le point précédent. Seule la commutativité ne passe pas bien entendu !

▷

Définition 19.3.45 (Groupe quotient par un sous-groupe distingué)

Soit G un groupe et H un sous-groupe distingué de G. On note G/H le groupe défini sur l'ensemble quotient par la loi induite, telle que décrite dans le théorème précédent.

Remarque 19.3.46

Lorsque G est un groupe abélien, tout sous-groupe H de G est distingué. La première construction donnée est un cas particulier de la situation générale.

On donne maintenant un certain nombre de propriétés utiles pour définir des morphismes sur des groupes quotients, et pour construire des isomorphismes (qui permettent de comparer des structures de groupes)

Lemme 19.3.47 (Passage au quotient d'un morphisme de groupe)

Soit $f : G \to K$ un morphisme de groupes et H un sous-groupe distingué de G. Alors f passe au quotient et définit

$$\tilde{f} : G/H \longrightarrow K$$

telle que $f = \tilde{f} \circ \pi$ si et seulement si $H \subset \mathrm{Ker}(f)$.

◁ **Éléments de preuve.**

Vérifier que pour un morphisme de groupe, cela équivaut au fait que f est constante sur chaque classe d'équivalence. ▷

Théorème 19.3.48 (Premier théorème d'isomorphisme, HP)

Soit $f : G \to H$ un morphisme de groupes. Alors $\mathrm{Ker}(f)$ est un sous-groupe distingué de G, et f passe au quotient, définissant un morphisme de groupes $\tilde{f} : G/\mathrm{Ker}(f) \to H$. Le morphisme $\tilde{f}$ est alors injectif. Sa corestriction à son image est donc un isomorphisme.

◁ **Éléments de preuve.**

Le passage au quotient provient du lemme précédent. La relation de morphisme passe alors aussi au quotient. Le noyau est clairement la classe de e. ▷

Encore une fois, cela traduit le fait que $\mathrm{Ker}(f)$ regroupe tout le défaut d'injectivité. Si on tue le défaut d'injectivité en le considérant comme un unique élément 0, on gagne l'injectivité.

III.7 Les groupes $\mathbb{Z}/n\mathbb{Z}$, groupes cycliques

Définition 19.3.49

Soit $n \in \mathbb{N}^*$. Le groupe $\mathbb{Z}/n\mathbb{Z}$ est le groupe quotient de $\mathbb{Z}$ par son sous-groupe (distingué) $n\mathbb{Z}$

Concrètement, les éléments de $\mathbb{Z}/n\mathbb{Z}$ sont les classes $n\mathbb{Z}+k$, pour $k \in \mathbb{Z}$. On note $\overline{k}$ la classe de k, c'est à dire l'élément $n\mathbb{Z}+k$ de $\mathbb{Z}/n\mathbb{Z}$. Il s'agit donc de l'ensemble des entiers congrus à k modulo n, considérés via le quotient comme un unique et même élément. Deux valeurs k et k' définissent la même classe si et seulement s'ils sont congrus l'un à l'autre modulo n.

Remarque 19.3.50

$\mathbb{Z}/2\mathbb{Z}$ synthétise la parité des entiers. La loi définie sur ce groupe résume les propriétés de parité des sommes.

Proposition 19.3.51

$(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe cyclique.

◁ **Éléments de preuve.**

On dispose d'un générateur évident (et d'autres qui le sont peut-être un peu moins ; pouvez-vous caractériser les générateurs ?). Et c'est un groupe fini. ▷

Ces groupes jouent un rôle important pour l'étude des groupes abéliens finis. On peut en effet montrer que tout groupe abélien fini est produit cartésien de groupes de ce type (théorème de structure).

Définition 19.3.52 (Ordre d'un élément d'un groupe, Spé)

L'ordre d'un élément d'un groupe (multiplicatif) G, dont le neutre est noté e, est

$$\mathrm{ord}(x) = \min\{n \in \mathbb{N}^* \mid x^n = e\}$$

Cet ordre peut être $+\infty$ par convention si l'ensemble ci-dessus est vide.

Proposition 19.3.53 (Résolution de $x^n = 1$, Spé)

Soit x un élément d'un groupe G de neutre e. L'ensemble $A = \{n \in \mathbb{Z} \mid x^n = e\}$ est un sous-groupe de $\mathbb{Z}$, donc de la forme $a\mathbb{Z}$. De plus, x est d'ordre fini si et seulement si $a \neq 0$, et dans ce cas, $\mathrm{ord}(x) = a$. En particulier, si x est d'ordre fini, $x^n = e$ si et seulement si $\mathrm{ord}(x) \mid n$.

◁ **Éléments de preuve.**

Description des sous-groupes de $\mathbb{Z}$ et paraphrase de la définition de l'ordre. ▷

Proposition 19.3.54 (Description des groupes monogènes)

Soit $G = \langle x \rangle$ un groupe monogène. Alors :

- *si $\mathrm{ord}(x) = +\infty$, G est isomorphe à $\mathbb{Z}$;*
- *si $\mathrm{ord}(x) = n$, $n \in \mathbb{N}^*$, alors G est fini, et G est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.*

◁ **Éléments de preuve.**

Considérer $\mathbb{Z} \to G$, $n \mapsto x^n$, et lui appliquer le premier théorème d'isomorphisme. On peut aussi faire les vérifications « à la main ». ▷

En particulier, si G est un groupe monogène d'ordre n, tout générateur de G est d' ordre n.

Exemple 19.3.55

Soit $n \geqslant 2$. On a $\mathbb{U}_n \simeq \mathbb{Z}/n\mathbb{Z}$. En fait, $\mathbb{Z}/n\mathbb{Z}$ est le groupe monogène additif d'ordre n de référence, alors que $\mathbb{U}_n$ est le groupe monogène multiplicatif d'ordre n de référence.

Théorème 19.3.56 (encore Lagrange, Spé)

Soit x un élément d'un groupe fini G. Alors l'ordre de x divise l'ordre de G.

◁ **Éléments de preuve.**

- L'ordre de x est égal à l'ordre d'un sous-groupe de G. On est ramené à la version précédente du théorème de Lagrange.
- Dans le cas où G est abélien, on peut donner une démonstration élémentaire de ce théorème n'utilisant pas les classes de congruence modulo un sous-groupe :
  Simplifier $\prod_{g \in G} (xg)$, en remarquant que $g \mapsto xg$ est une bijection.

▷

Cette preuve (dans le cas abélien) ressemble à une autre (d'un résultat arithmétique classique), que certains d'entre vous ont peut-être déjà vue. Laquelle ? Ce n'est pas anodin, le résultat arithmétique en question étant en fait un cas particulier du théorème de Lagrange.

IV Anneaux et corps

IV.1 Axiomatiques de la structure d'anneau

Définition 19.4.1 (Anneau)

Soit A un ensemble, muni de deux lois de composition internes (généralement notées $+$ et $\times$). On dit que $(A, +, \times)$ (ou plus simplement A) est un anneau si :

(i) $(A, +)$ est un groupe abélien ;

(ii) $(A, \times)$ est un monoïde (autrement dit $\times$ est associative et il existe un élément neutre 1 pour $\times$) ;

(iii) $\times$ est distributive sur $+$.

Remarque 19.4.2

Certains ouvrages (notamment anciens) n'imposent pas l'existence de l'élément neutre 1 pour le produit et parlent alors d'*anneau unifère* ou *unitaire* pour ce que nous appelons ici simplement un *anneau*. La convention que nous adoptons concernant l'existence d'un élément neutre est celle généralement adoptée actuellement, et nous suivons en cela le programme officiel de la classe de MPSI.

Exemples 19.4.3

1. $\{0\}$ muni des opérations triviales est un anneau ; ici le neutre pour le produit est 0.
2. $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ munis des opérations usuelles sont des anneaux.

3. Pour tout $n \in \mathbb{N}^*$, $\mathbb{Z}/n\mathbb{Z}$ est un anneau. La structure circulaire de ces anneaux explique la terminologie.
4. L'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels est un anneau. De même pour $\mathbb{Z}[X]$, $\mathbb{Q}[X]$ ou $\mathbb{C}[X]$.
5. $\mathbb{N}$ n'est pas un anneau.
6. L'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées est un anneau.
7. L'ensemble $(\mathcal{P}(E), \triangle, \cap)$ est un anneau (anneau de Boole).
8. $(\mathbb{R}^{\mathbb{R}}, +, \circ)$ est-il un anneau ?

Lemme 19.4.4 (Les trous noirs existent aussi en mathématiques)
Soit $(A, +, \times)$ un anneau. Alors 0 est absorbant.

◁ **Éléments de preuve.**
Simplifier $(0 + 0) \times x$. ▷

Proposition 19.4.5
Si A est un anneau ayant au moins deux éléments, alors $1 \neq 0$.

◁ **Éléments de preuve.**
En effet 1 n'est pas absorbant. ▷

Proposition 19.4.6 (Produit d'opposés)
Soit x et y dans A. Alors
- $(-x)y = -(xy)$
- $x(-y) = -(xy)$
- $(-x)(-y) = xy$.

◁ **Éléments de preuve.**
Considérer $(x + (-x))y$. ▷

Définition 19.4.7 (Anneau commutatif)
On dit qu'un anneau $(A, +, \times)$ est commutatif si et seulement si la loi $\times$ est commutative.

Les exemples donnés ci-dessus sont des exemples d'anneaux commutatifs, à l'exception d'un exemple. Lequel ?
Enfin, conformément à la définition générale, nous donnons :

Définition 19.4.8 (Homomorphisme d'anneaux)
Soit A et B deux anneaux. Un homomorphisme d'anneaux de A à B est une application $f : A \longrightarrow B$, soit égale à la fonction nulle, soit vérifiant :

$$\forall (x, y) \in A^2, \quad f(x + y) = f(x) + f(y), \qquad f(xy) = f(x)f(y) \qquad \text{et} \qquad f(1_A) = 1_B.$$

Ainsi, un homomorphisme d'anneaux (à part le morphisme nul un peu particulier) est à la fois un homomophisme du groupe $(A, +)$ et du monoïde $(A, \times)$.

IV.2 Sous-anneaux

Conformément à la définition générale, nous avons :

Définition 19.4.9 (Sous-anneau)

Soit $(A, +, \times)$ un anneau. Un sous-ensemble $B \subset A$ est un sous-anneau de A si et seulement si B est stable pour les lois $+$ et $\times$, si $1_A \in B$, et si les lois induites sur B définissent sur B une structure d'anneau (le neutre multiplicatif étant alors nécessairement 1_A).

Remarquez qu'encore une fois, on ne dit rien de l'appartenance de 0_A à B, celle-ci étant ici aussi automatique (puisque $(B, +)$ est un sous-groupe de $(A, +)$). En revanche, l'appartenance de 1_A à B n'est pas automatique, comme le montre l'exemple de $B = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix},\ \lambda \in \mathbb{R} \right\}$, sous-ensemble de $\mathcal{M}_2(\mathbb{R})$, stable pour les lois $+$ et $\times$. Ce n'est pas un sous-anneau au sens que nous en avons donné puisque $I_2 \notin B$. En revanche, les restrictions de $\times$ et $+$ définissent tout de même une structure d'anneau sur B, le neutre étant alors $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Proposition 19.4.10 (Caractérisation des sous-anneaux)

Un sous-ensemble B d'un anneau A est un sous-anneau de A si et seulement si :

(i) $1_A \in B$.

(ii) pour tout $(x, y) \in B$, $x - y \in B$

(iii) pour tout $(x, y) \in B$, $xy \in B$

◁ **Éléments de preuve.**

Comparer à la caractérisation des sous-groupes. Par ailleurs, les propriétés universelles sont conservées par restriction, on ne le dira jamais assez ! ▷

Exemples 19.4.11

1. $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$ qui est un sous-anneau de $\mathbb{R}$ qui est un sous-anneau de $\mathbb{C}$.
2. $\mathbb{Z}/n\mathbb{Z}$ n'a d'autre sous-anneau que lui-même.

Proposition 19.4.12

Soit A un anneau, et $(A_i)_{i \in I}$ une famille de sous-anneaux de A. Alors $\bigcap_{i \in I} A_i$ est un sous-anneau de A.

Proposition 19.4.13 (Image par un homomorphisme)

Soit $f : A \to B$ un homomorphisme d'anneaux.

1. *Soit A' un sous-anneau de A. Alors $f(A')$ est un sous-anneau de B*
2. *Soit B' un sous-anneau de B. Alors $f^{-1}(B')$ est un sous-anneau de A.*

◁ **Éléments de preuve.**

Vérifications sans difficulté avec la caractérisation précédente. ▷

IV.3 Calculs dans un anneau

Du fait de l'existence d'une addition et d'une multiplication dans un anneau et dans un corps, et des règles d'associativité et de commutativité, tous les calculs que l'on a l'habitude de faire dans $\mathbb{Z}$, $\mathbb{R}$ ou $\mathbb{C}$ peuvent se généraliser à un anneau ou un corps quelconque. Il faut toutefois faire attention que dans un anneau, contrairement à ce qu'il se passe dans $\mathbb{R}$ ou $\mathbb{C}$, tous les éléments ne sont pas inversibles, et que par ailleurs, les calculs nécessitant de permuter l'ordre de certains facteurs multiplicatifs ne peuvent pas être effectués en toute généralité dans un anneau non commutatif. Ainsi pour des calculs dans un anneau **considérer l'analogie avec $\mathbb{Z}$ (plutôt que $\mathbb{R}$), et se méfier** :

- des inversions intempestives
- des simplifications abusives
- des problèmes de commutativité, qui nécessitent parfois l'introduction d'hypothèses supplémentaires, à vérifier scrupuleusement.

L'analogie avec $\mathbb{Z}$ n'est pas toujours suffisante, puisqu'il peut se produire des situations bien particulières, n'ayant pas lieu dans $\mathbb{Z}$, comme par exemple l'existence de « diviseurs de zéro » (voir un peu plus loin)
Nous rappelons qu'on peut définir dans un anneau A le produit nx où $n \in \mathbb{Z}$, et $x \in A$ de la manière suivante :

- Si $n = 0$, $nx = 0$
- Si $n > 0$, $nx = x + \cdots + x$ (n facteurs)
- Si $n < 0$, $nx = -|n|x$.

Ces quantités sont compatibles avec le produit dans l'anneau :

Lemme 19.4.14 (Compatibilité du produit avec l'itération de +)
Soit x et y deux éléments d'un anneau A et n, m dans $\mathbb{Z}$. Alors $(mx)(ny) = (mn)(xy)$

◁ **Éléments de preuve.**
Le cas $m = 0$ ou $n = 0$ est évident. Pour $m, n > 0$, faire une récurrence sur m à n fixé. L'initialisation nécessite elle-même une récurrence sur m. Pour les valeurs négatives se servir d'un lemme précédent sur les produits d'opposés. ▷

Nous pouvons définir de même a^n, pour tout $n \in \mathbb{N}$, et même pour tout $n \in \mathbb{Z}$ si a est inversible.
Nous voyons, outre les règles usuelles découlant des règles d'associativité et de distributivité, deux résultats déjà évoqués dans le cas de $\mathbb{R}$ ou $\mathbb{C}$, et que nous voyons plus généralement dans le cadre d'anneaux, mais qui nécessitent une hypothèse de commutativité.

Théorème 19.4.15 (Factorisation de $a^n - b^n$, Bernoulli)
Soit a et b deux éléments d'un anneau A tels que $ab = ba$. Alors pour tout $n \in \mathbb{N}^*$

$$a^n - b^n = (a - b)\sum_{k=0}^{n-1} a^{n-1-k}b^k.$$

◁ **Éléments de preuve.**
Même démonstration que dans $\mathbb{R}$ ou $\mathbb{C}$, por télescopage de la seconde somme (après distribution) ▷

Corollaire 19.4.16 (Factorisation de $1 - a^n$)
Pour tout élément a d'un anneau A,

$$1 - a^n = (1 - a)\sum_{k=0}^{n-1} a^k.$$

Si $1-a$ est inversible (condition plus forte que $a \neq 1$), on peut alors écrire :

$$(1-a)^{-1}(1-a^n) = \sum_{k=0}^{n-1} a^k$$

En revanche, évitez d'écrire cela sous forme de fraction lorsqu'on n'est pas dans une structure commutative, et attention à placer l'inverse du bon côté (même si, pour l'expression considérée, ce ne serait pas gênant car les facteurs considérés commutent, même si globalement l'anneau n'est pas commutatif ; mais autant prendre dès maintenant de bonnes habitudes)

Théorème 19.4.17 (Formule du binôme)

Soit a et b deux éléments d'un anneau tels que $ab = ba$. Alors, pour tout $n \in \mathbb{N}$,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

◁ **Éléments de preuve.**

Même démonstration que dans $\mathbb{R}$ ou $\mathbb{C}$, par récurrence, ou par un argument combinatoire (on remarquera qu'on utilise exclusivement l'associativité, la distributivité généralisée, et la commutativité des deux lois). En quoi utilise-t-on la commutativité de $\times$? ▷

Attention en revanche au cas où on n'a pas commutativité de a et b : il convient de bien distinguer les deux facteurs ab et ba appraissant dans le développement de $(a+b)(a+b)$ (par exemple, pour $n=2$) :

$$(a+b)(a+b) = a^2 + ab + ba + b^2 \neq a^2 + 2ab + b^2,$$

si $ab \neq ba$. Cette situation peut se produire notamment dans le cadre du produit matriciel. Il faut être toujours bien vigilant à vérifier l'hypothèse de commutativité $ab = ba$.

IV.4 Éléments inversibles

Un anneau n'étant pas nécessairement commutatif, il convient de distinguer la notion d'inversibilité à droite, inversibilité à gauche. Un inverse est alors à la fois un inverse à gauche et à droite, et en cas d'existence, il est unique, comme nous l'avons montré dans une situation générale. De plus, dans le cas d'un anneau, l'ensemble des éléments inversibles possède une structure particulière.

Théorème 19.4.18 (Groupe des inversibles d'un anneau)

Soit A un anneau. Alors l'ensemble des éléments inversibles de A, généralement noté $A^\times$ ou $U(A)$, est stable pour la loi $\times$, et la loi induite munit $A^\times$ d'une structure de groupe multiplicatif.

◁ **Éléments de preuve.**

C'est plus généralement le groupe des inversibles d'un monoïde (on ne considère pas la structure additive ici). Vérifier les différents points de la définition. Ici, on ne peut pas le voir comme un sous-groupe de quelque chose. ▷

Exemples 19.4.19

1. $\mathbb{R}^\times = \mathbb{R}^*$, $\mathbb{Q}^\times = \mathbb{Q}^*$, $\mathbb{C}^\times = \mathbb{C}^*$.
2. $(\mathcal{M}_n(\mathbb{R}))^\times = \mathrm{GL}_n(\mathbb{R})$, ensemble des matrices inversibles, appelé *groupe linéaire*
3. L'ensemble des inversibles de $\mathbb{Z}$: $\mathbb{Z}^\times = U(\mathbb{Z}) = \{-1, 1\}$.

4. $(\mathbb{Z}/4\mathbb{Z})^\times = \{1, 3\}$
5. $(\mathbb{Z}/p\mathbb{Z})^\times = (\mathbb{Z}/p\mathbb{Z}) \setminus \{0\}$, si p est premier. On peut montrer que ce groupe multiplicatif est isomorphe au groupe additif $\mathbb{Z}/(p-1)\mathbb{Z}$.
6. Que dire plus généralement de $(\mathbb{Z}/n\mathbb{Z})^\times$?

L'étude de $\mathbb{Z}/4\mathbb{Z}$ amène un résultat peu commun pour qui n'a pas l'habitude de travailler dans des structures algébriques abstraites : $2 \times 2 = 0$. Autrement dit, on a deux éléments a et b non nuls, et vérifiant $ab = 0$. La vieille règle, bien pratique pour résoudre des équations, qui nous dit que si $ab = 0$, alors $a = 0$ ou $b = 0$, ne s'applique donc pas dans ce contexte. Comme elle est bien pratique tout de même, nous allons établir un contexte dans lequel elle est vraie, en définissant une propriété adéquate des anneaux nous permettant de l'utiliser.

Définition 19.4.20 (Diviseurs de zéro, HP)

Soit a un élément d'un anneau A. On dit que a est un diviseur de 0 à gauche si $a \neq 0$ et s'il existe $b \in A$, $b \neq 0$, tel que $ab = 0$. On définit de façon symétrique les diviseurs de zéro à droite.

La notion de diviseur de 0 caractérise en fait la non régularité :

Proposition 19.4.21

Un élément a non nul d'un anneau est régulier à gauche (resp. à droite) si et seulement s'il n'est pas diviseur de 0 à gauche (resp. à droite). L'élément a est régulier s'il est régulier à gauche et à droite.

◁ **Éléments de preuve.**

- Pour simplifier $ax = ay$, passer tout du même côté, et factoriser par a
- Réciproquement, si $ax = ay$ avec $x \neq y$, exprimer une relation de divisibilité de 0.

▷

Corollaire 19.4.22

Un diviseur de 0 n'est pas inversible.

◁ **Éléments de preuve.**

Les éléments inversibles (resp. inversibles à gauche, resp. inversibles à droite) sont réguliers (resp. réguliers à gauche, resp. réguliers à droite). ▷

Définition 19.4.23 (anneau intègre, HP)

Un anneau intègre A est un anneau commutatif non réduit à $\{0\}$ et sans diviseur de 0.

Remarque 19.4.24 (Acceptation plus large)

Dans certains ouvrages, le caractère commutatif n'est pas imposé. Dans ce cas, A ne doit avoir ni diviseur de 0 à gauche ni diviseur de 0 à droite.

En particulier, dans un anneau intègre, toutes les simplifications par des éléments non nuls sont possibles, puisque le seul élément non régulier est 0.

Exemples 19.4.25

1. $\mathbb{Z}$ est intègre, $\mathbb{R}[X]$ est intègre, tout corps (défini plus loin) est intègre.
2. $\mathcal{M}_n(\mathbb{R})$ n'est pas intègre lorsque $n \geqslant 2$.
3. À quelle condition sur n, $\mathbb{Z}/n\mathbb{Z}$ est-il intègre ?

IV.5 Corps

Un corps est un anneau vérifiant une condition supplémentaire :

Définition 19.4.26 (Corps)

Soit K un ensemble muni de deux lois $+$ et $\times$. On dit que $(K, +, \times)$ (ou plus simplement K) est un corps si K est un anneau commutatif tel que $(K^*, \times)$ soit un groupe, où $K^* = K \setminus \{0\}$.

Ainsi K est un corps si et seulement si c'est un anneau commutatif non réduit à $\{0\}$, et tel que tout élément non nul soit inversible. En particulier, $\{0\}$ n'est en général pas considéré comme un corps. En effet, la définition impose que les deux éléments 0 et 1 soient distincts.

Remarque 19.4.27

- Conformément au programme, nous adoptons la convention stipulant que tout corps doit être commutatif. Là encore, les ouvrages anciens n'imposent pas cette condition. Il est d'usage actuellement d'appeler *corps gauche* un ensemble muni d'une structure vérifiant tous les axiomes de la structure de corps, à l'exception de la commutativité de la multiplication. On rencontre aussi parfois la terminologie *anneau à divisions*, traduction de la terminologie anglaise *division ring*.
- Dans le cas des corps finis, les deux notions coïncident, d'après le théorème de Wedderburn, stipulant que « tout corps fini est commutatif », ce qui, avec notre terminologie, se réexprime : « tout corps gauche fini est un corps. ».

Exemples 19.4.28

1. $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ sont des corps.
2. $\mathbb{Z}$ n'est pas un corps.
3. En général $\mathbb{Z}/n\mathbb{Z}$ n'est pas un corps. Par exemple, 2 n'est pas inversible dans $\mathbb{Z}/4\mathbb{Z}$.

L'exemple suivant, du fait de son importance, est donné en théorème. La démonstration doit être retenue.

Théorème 19.4.29 (Le corps $\mathbb{F}_p$)

L'anneau $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ est un corps si et seulement si p est premier. Ce corps est en général noté $\mathbb{F}_p$.

◁ **Éléments de preuve.**

Exprimer une relation de Bézout entre x et p. ▷

La notation utilisée s'explique par la terminologie anglaise (*field*) pour un corps. Par exemple, $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ est un corps à 2 éléments. C'est le plus petit corps possible, puisqu'un corps contient 0, et par définition, n'est pas réduit à $\{0\}$.

Remarque 19.4.30

On peut montrer que tout corps fini a un cardinal égal à p^n, où p est un nombre premier et n un entier. On peut également montrer que pour de telles données, il existe (à isomorphisme près) un unique corps à $q = p^n$ éléments, qu'on note $\mathbb{F}_q$. Lorsque $n = 1$, on retrouve les corps $\mathbb{F}_p$ du point précédent.

Définition 19.4.31 (Sous-corps)

Soit $L \subset K$ un sous-ensemble d'un corps K. On dit que L est un sous-corps de K si L est stable par $+$ et $\times$, $1_K \in L$, et si les lois induites sur L par celles de K le munissent d'une structure de corps.

Remarque 19.4.32

On pourrait remplacer l'hypothèse $1_K \in L$ par le fait que L contient un élément $x \neq 0_K$.

Proposition 19.4.33 (Caractérisation des sous-corps)

$L \subset K$ est un sous-corps de K si et seulement si :

- $1_K \in L$
- *pour tout* $(x, y) \in L$, $x - y \in L$
- *pour tout* $(x, y) \in L$ *tel que* $y \neq 0$, $xy^{-1} \in L$.

◁ **Éléments de preuve.**

Combiner caractérisation des groupes (multiplicatifs) et caractérisation des anneaux. ▷

Exemples 19.4.34

$\mathbb{Q}$ est un sous-corps de $\mathbb{R}$, $\mathbb{R}$ est un sous-corps de $\mathbb{C}$. Entre $\mathbb{Q}$ et $\mathbb{R}$, il existe un grand nombre de corps intermédiaires (corps de nombres), par exemple $\mathbb{Q}[\sqrt{2}]$, plus petit sous-corps de $\mathbb{R}$ contenant les rationnels et $\sqrt{2}$. Plus explicitement,

$$\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2},\ (a, b) \in \mathbb{Q}\}$$

(voir exercices pour une justification).

Définition 19.4.35 (Homomorphisme de corps)

Soit K et L deux corps. Un homomorphisme de corps $f : K \longrightarrow L$ est un homomorphisme des anneaux sous-jacents.

Proposition 19.4.36 (Image par un homomorphisme)

Soit $f : K \to L$ un homomorphisme de corps.

1. *Soit K' un sous-corps de K. Alors $f(K')$ est un sous-corps de L*
2. *Soit L' un sous-corps de L. Alors $f^{-1}(L')$ est un sous-corps de K.*

◁ **Éléments de preuve.**

Toujours pareil. On peut repartir de l'énoncé similaire pour les anneaux, pour ne pas avoir à tout refaire. ▷

Proposition 19.4.37 (Injectivité des homomorphismes de corps, HP)

Un homomorphisme de corps est toujours injectif.

◁ **Éléments de preuve.**

Si x est inversible, $f(x)$ est inversible (ce fait est vérifié pour tout anneau). Comment exprimez-vous $f(x)^{-1}$? Quelle conséquence sur le noyau (additif) ? ▷

Définition 19.4.38 (Caractéristique d'un corps)

Soit K un corps, d'élément neutre $1_K \neq 0_K$. Soit $H = \{n \cdot 1_K, n \in \mathbb{Z}\}$ le sous-groupe monogène de $(K, +)$ engendré par 1_K.

- Si H est infini, on dit que K est de caractéristique nulle.
- Si H est fini, de cardinal $p \in \mathbb{N}$, on dit que K est de caractéristique p.

Proposition 19.4.39

Soit K un corps de caractéristique finie p. Alors, pour tout $x \in K$, $px = 0$.

◁ **Éléments de preuve.**

Utiliser la relation $(nx)y = n(xy)$ en choisissant convenablement n, x et y. ▷

Théorème 19.4.40 (Primalité de la caractéristique d'un corps)

Soit K un corps de caractéristique non nulle. Alors sa caractéristique p est un nombre premier.

◁ **Éléments de preuve.**

Si $p = ab$, $(a \times 1)(b \times 1) = 0$. Dans quelle mesure est-ce possible ? ▷

Remarque 19.4.41

- Un corps fini est toujours de caractéristique non nulle, donc première.
- Il existe des corps infinis de caractéristique p (par exemple le corps des fractions rationnelles à coefficients dans $\mathbb{F}_p$)
- On peut définir de la même manière la caractéristique d'un anneau. Il s'agit de $+\infty$, ou d'un entier strictement positif, qui peut cette fois ne pas être premier.

Proposition 19.4.42

Si $\mathbb{K}$ est un corps de caractéristique nulle, et $(n, x) \in \mathbb{Z} \times \mathbb{K}$, alors $n \cdot x = 0$ si et seulement si $n = 0$ ou $x = 0$.

IV.6 Idéaux d'un anneau (Spé)

La notion de sous-anneau est souvent trop restrictive, et on est souvent amené à considérer une structure moins riche :

Définition 19.4.43 (Idéal d'un anneau commutatif, Spé)

Soit A un anneau commutatif, et I un sous-ensemble de A. On dit que I est un idéal si et seulement si I est un sous-groupe du groupe $(A, +)$ et si pour tout $a \in I$ et tout $\lambda \in A$, $\lambda a \in I$.

Ainsi, I est un sous-groupe de $(A,+)$, stable par multiplication par un élément de A.
On peut aussi définir la notion d'idéal dans un anneau non commutatif. Il faut alors distinguer la notion d'idéal à droite, à gauche, ou bilatère, suivant qu'on impose la stabilité par multiplication à droite, à gauche ou des deux côtés, par un élément quelconque de A.
Nous n'étudierons pas les idéaux cette année, mais nous illuminerons parfois quelques résultats à l'éclat de cette notion, notamment en arithmétique. Nous donnons tout de même quelques exemples importants :

Exemples 19.4.44

1. Pour tout $n \in \mathbb{N}$, $n\mathbb{Z}$ est un idéal de $\mathbb{Z}$. Réciproquement, tout idéal de $\mathbb{Z}$ est de cette forme.
2. L'ensemble des polynômes de $\mathbb{R}[X]$ s'annulant en 0 est un idéal de $\mathbb{R}[X]$. Comment généraliseriez-vous ce résultat ?
3. L'ensemble des polynômes $\{XP(X,Y)+YQ(X,Y),\ \ (P,Q) \in \mathbb{R}[X,Y]\}$ est un idéal de l'anneau $\mathbb{R}[X,Y]$ des polynômes à deux indéterminées.
4. Que peut-on dire des idéaux d'un corps ?

Dans les deux premiers exemples, on constate que l'idéal considéré est de la forme $\{\lambda a, \lambda \in A\}$, donc engendré par un unique élément a, par multiplication par les éléments λ de A. Un idéal vérifiant cette propriété est appelé *idéal principal* :

Définition 19.4.45 (Idéal principal)

Un idéal principal est un idéal engendré par un unique élément, c'est à dire de la forme

$$I = aA = \{ay, y \in A\},$$

pour un certain $a \in A$. On note souvent $(a) = aA$.

Tout idéal n'est pas principal, comme le montre le troisième exemple.

Définition 19.4.46 (Anneau principal)

Un anneau intègre dont tous les idéaux sont principaux est appelé *anneau principal*.

Théorème 19.4.47

$\mathbb{Z}$ est un anneau principal.

◁ **Éléments de preuve.**

Les idéaux de $\mathbb{Z}$ sont en particulier des sous-groupes, dont on connaît la description ! ▷

Cette définition, qui peut paraître anodine, est à la base d'une généralisation possible de la notion de pgcd et de ppcm à des anneaux autres que $\mathbb{Z}$. Cette propriété, aussi vérifiée pour $\mathbb{R}[X]$ ou $\mathbb{C}[X]$, permet par exemple de généraliser l'arithmétique connue de $\mathbb{Z}$ au cas des polynômes.
Pour terminer, nous faisons la remarque que la notion d'idéal est la notion adaptée pour définir des quotients dans la catégorie des anneaux commutatifs.

Théorème 19.4.48 (Anneau quotient, HP)

Soit A un anneau commutatif et I un idéal de A. Le quotient de groupes A/I peut être muni d'une multiplication telle que pour tout $(a,b) \in A^2$, $\overline{ab} = \overline{a}\,\overline{b}$. De plus, A/I est ainsi muni d'une structure d'anneaux.

20 Calcul matriciel

La Matrice est universelle. Elle est omniprésente. Elle est avec nous ici, en ce moment même.

(Matrix – Lana et Andy Wachowski)

On a donc autant d'équations linéaires qu'il n'y a d'inconnues à trouver ; les valeurs de ces inconnues seront obtenues par l'élimination ordinaire.
Voyons maintenant, si cette élimination est toujours possible, ou si la solution peut quelquefois devenir indéterminée ou même impossible

Carl Friedrich Gauss (traduction Edmond Dubois)

Et les shadoks pivotaient, pivotaient, pivotaient...

(Libre adaptation de l'oeuvre de Jacques Rouxel)

Voici venir les temps où vibrant sur sa tige
Chaque fleur s'évapore ainsi qu'un encensoir ;
Les sons et les parfums tournent dans l'air du soir ;
Valse mélancolique et langoureux vertige !

(Charles Baudelaire ; le troisième vers est aussi le titre d'un prélude de Debussy)

Introduction

L'objectif de ce chapitre est d'introduire le calcul matriciel, à savoir les opérations matricielles (somme et produit), ainsi, que les diverses manipulations et techniques usuelles permettant d'obtenir des informations intéressantes sur les matrices (échelonnement, calcul de l'inverse, résolution de systèmes linéaires définis par une matrice, et par ces méthodes, dans un chapitre ultérieur, calcul du rang d'une matrice). Pour ces dernières méthodes, une technique sera centrale : il s'agit de la technique de l'élimination de Gauss-Jordan, aussi connue sous le nom de pivot de Gauss. Cela explique un certain nombre des citations débutant ce chapitre.

I Opérations matricielles

I.1 L'ensemble des matrices de type (n, p)

Dans ce qui suit, $\mathbb{K}$ désigne un corps quelconque, par exemple $\mathbb{R}$, $\mathbb{C}$ ou $\mathbb{Q}$. Beaucoup de constructions, de méthodes et de propriétés se généralisent au cas de coefficients dans un anneau A, dès lors qu'elles ne font pas intervenir d'inverses.

Définition 20.1.1 (Matrice)

Une matrice de taille $n \times p$ (n lignes et p colonnes) à coefficients dans $\mathbb{K}$ est la donnée d'une famille $A = (a_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ d'éléments de $\mathbb{K}$. On utilise la représentation planaire suivante :

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}.$$

Définition 20.1.2 (Ensemble des matrices)

L'ensemble des matrices de taille $n \times p$ (on dit aussi *de type* (n, p)) à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{n,p}(\mathbb{K})$.

Définition 20.1.3 (Matrices carrées)

Si $n = p$, on dit que la matrice est *carrée*, et on note simplement $\mathcal{M}_n(\mathbb{K})$ au lieu de $\mathcal{M}_{n,n}(\mathbb{K})$ l'ensemble des matrices carrées de taille n. Par ailleurs on parlera souvent plutôt de matrice carrée *d'ordre* n plutôt que de matrice de taille n ou $n\times n$. Cette notion d'ordre est à distinguer de la notion d'ordre des éléments d'un groupe ; cela n'a aucun rapport.

Une matrice est souvent représentée sous forme d'un tableau, explicite dans le cas d'une matrice déterminée de petite taille, ou avec des « ... » dans le cas de matrice explicites ou non, de taille variable. Par exemple :

$$M_1 = \begin{pmatrix} 1 & 4 & 2 \\ 5 & 2 & 2 \\ 7 & 6 & 1 \end{pmatrix} \qquad M_2 = (i+j-1)_{1\leqslant i,j\leqslant n} = \begin{pmatrix} 1 & \cdots & n \\ \vdots & & \vdots \\ n & \cdots & 2n-1 \end{pmatrix}$$

$$M_3 = (a_{i,j})_{1\leqslant i,j\leqslant n} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}.$$

Dans une telle représentation, le premier indice est toujours l'indice de ligne, et le second l'indice de colonne.

Terminologie 20.1.4 (Matrices colonnes, matrices lignes)

- Une matrice $X \in \mathcal{M}_{n,1}(\mathbb{K})$ est appelée matrice colonne, et est représentée par : $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
- Une matrice $X \in \mathcal{M}_{1,n}(\mathbb{K})$ est appelée matrice ligne, et est représentée par : $X = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}$
- On peut identifier $\mathbb{K}^n$ indifféremment à $\mathcal{M}_{n,1}(\mathbb{K})$ ou $\mathcal{M}_{1,n}(\mathbb{K})$, suivant la situation.
- $\mathcal{M}_{1,1}(\mathbb{K})$ peut être identifié à $\mathbb{K}$ (matrice constituée d'un unique coefficient)

I.2 Combinaisons linéaires de matrices

Pour une matrice $A \in \mathcal{M}_{n,p}$, et $(i, j) \in [\![1, n]\!] \times [\![1, p]\!]$, on désignera par $[A]_{i,j}$ le coefficient de A en position (i, j) (notation non standard)

Définition 20.1.5 (Somme de matrices)

Soit $A, B \in \mathcal{M}_{n,p}(KK)$ deux matrices *de même format*. On définit la somme de A et B coefficient par coefficient par

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,p]\!], \quad [A+B]_{i,j} = [A]_{i,j} + [B]_{i,j}.$$

Ainsi, il s'agit d'une matrice de même format que A et B.

On peut réécrire cette définition de la façon suivante :

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} + \begin{pmatrix} b_{1,1} & \cdots & b_{1,p} \\ \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,p} \end{pmatrix} = \begin{pmatrix} a_{1,1}+b_{1,1} & \cdots & a_{1,p}+b_{1,p} \\ \vdots & & \vdots \\ a_{n,1}+b_{n,1} & \cdots & a_{n,p}+b_{n,p} \end{pmatrix}$$

Proposition 20.1.6 (Associativité et commutativité de la somme)

La somme matricielle définie sur $\mathcal{M}_{n,p}(\mathbb{K})$ est associative et commutative.

◁ **Éléments de preuve.**

Cela provient directement de l'associativité et de la commutativité de la somme dans $\mathbb{K}$, appliquées à chaque coefficient. ▷

Proposition 20.1.7 (Neutre et opposés)

Soit $0_{\mathcal{M}_{n,p}(\mathbb{K})}$ la matrice constituée uniquement de 0.

(i) $0_{\mathcal{M}_{n,p}(\mathbb{K})}$ est neutre pour l'addition : pour tout $A \in \mathcal{M}_{n,p}(\mathbb{K})$, $A + 0_{\mathcal{M}_{n,p}(\mathbb{K})} = A$.

(ii) Toute matrice A admet un opposé $-A$ tel que $A + (-A) = 0_{\mathcal{M}_{n,p}(\mathbb{K})}$. De façon explicite :

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,p]\!] \quad [-A]_{i,j} = -[A]_{i,j}.$$

Ainsi, $-A = (-1) \cdot A$, avec la définition ci-dessous du produit par un scalaire.

◁ **Éléments de preuve.**

Ce sont des vérifications évidentes coefficient par coefficient. ▷

Définition 20.1.8 (Produit externe par un scalaire de $\mathbb{K}$)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $\lambda \in \mathbb{K}$. On définit le produit externe λA coefficient par coefficient par :

$$\forall (i,j) \in [\![1,n]\!] \times [\![1,p]\!], \quad [\lambda A]_{i,j} = \lambda [A]_{i,j}$$

On peut réécrire cette définition de la façon suivante :

$$\lambda \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} = \begin{pmatrix} \lambda a_{1,1} & \cdots & \lambda a_{1,p} \\ \vdots & & \vdots \\ \lambda a_{n,1} & \cdots & \lambda a_{n,p} \end{pmatrix}$$

Proposition 20.1.9 (Première associativité externe)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $(\lambda, \mu) \in \mathbb{K}^2$. Alors $(\lambda\mu)A = \lambda(\mu A)$.

◁ **Éléments de preuve.**

C'est une conséquence directe de la définition coefficient par coefficient, et de l'associativité du produit dans $\mathbb{K}$. ▷

Les deux définitions précédentes nous permettent donc de définir des combinaisons linéaires $\lambda A + \mu B$ de deux matrices de même format (λ et μ sont des éléments de $\mathbb{K}$).
Nous verrons dans un chapitre ultérieur que ces opérations munissent $\mathcal{M}_n(\mathbb{K})$ d'une structure de $\mathbb{K}$-espace vectoriel.

Définition 20.1.10 (Matrice élémentaire, ou matrice de la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$)

Soit $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$. On définit la *matrice élémentaire* $E_{i,j} \in \mathcal{M}_{n,p}(\mathbb{K})$ par :

$$\forall (k,\ell) \in [\![1,n]\!] \times [\![1,p]\!], \quad [E_{i,j}]_{k,\ell} = \delta_{(i,j),(k,\ell)} = \delta_{i,j}\delta_{k,\ell}.$$

Il s'agit donc de la matrice constituée d'un 1 en position (i,j) et de 0 partout ailleurs.

On prendra garde au fait que la notation ne fait pas référence au format (n,p). Si on est amené à considérer des matrices élémentaires de différents formats, il faut introduire des notations distinctes (par exemple avec des primes).
L'importance de ces matrices réside dans la propriété suivante :

Proposition/Définition 20.1.11 (Base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$)

La famille $(E_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$. En d'autres termes, pour tout matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$, il existe d'uniques scalaires $\lambda_{i,j}$, $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$, tels que

$$A = \sum_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} \lambda_{i,j} E_{i,j}.$$

◁ **Éléments de preuve.**

Il suffit de remarquer que la combinaisons linéaire des $E_{i,j}$ place les $\lambda_{i,j}$ en des positions distinctes de la matrice. Le résultat découle alors immédiatement du fait que 2 matrices sont égales ssi elles ont les mêmes coefficients en toutes les positions. ▷

I.3 Produit matriciel

Nous donnons ci-dessous la définition coefficient par coefficient du produit matriciel.

Définition 20.1.12 (Produit matriciel)

Soit $n,p,q \in \mathbb{N}^*$. On définit le produit matriciel :

$$\times : \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K}) \longrightarrow \mathcal{M}_{n,q}(\mathbb{K})$$

pour $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$ par :

$$\forall (i,k) \in [\![1,n]\!] \times [\![1,q]\!] \quad [AB]_{i,k} = \sum_{j=1}^{p} [A]_{i,j}[B]_{j,k} = \left\langle L_i(A)^\top, C_j(B) \right\rangle,$$

où $L_i(A)$ et $C_j(B)$ sont repectivement la i-ième ligne de A et la j-ième colonne de B, et pour une matrice ligne L, $L^\top$ (aussi souvent notée ${}^tL$) représente la matrice colonne obtenue en redressant L (transposée). Le produit scalaire considéré est le produit scalaire canonique de $\mathbb{K}^p$, les éléments de $\mathbb{K}^p$ étant identifiés aux matrices colonnes.

Avertissement 20.1.13

Attention aux formats ! Le nombre de colonnes de A doit être égal au nombre de lignes de B. Le produit AB a alors autant de lignes que A et autant de colonnes que B.

Remarque 20.1.14

La description un peu compliquée de ce produit matriciel est motivée par le fait que les applications linéaires entre deux espaces de dimension finie peuvent se décrire par des matrices (après avoir choisi des bases). Le produit matriciel correspond dans ce cadre à la matrice associée à la composée de deux applications linéaires. Cette remarque justifie aussi la condition sur les formats (nombre de ligne de A égal au nombre de colonnes de B) : c'est la condition de compatibilité des domaines pour pouvoir considérer la composée.

Note Historique 20.1.15

C'est bien cette correspondance entre applications linéaires et matrices qui a motivé historiquement la définition du produit matriciel, sous une forme légèrement détournée. En effet, Gauss s'est intéressé à l'expression de changements de variables dans des formes quadratiques à n variables, c'est-à-dire des changements de base. Ces changements de base peuvent se décrire par des matrices carrées qui correspondent à des applications linéaires (envoyant une base sur l'autre). Gauss s'est ensuite intéressé à la description des coefficients de la matrice qui traduit deux changements de variables successifs (ce qui revient à composer les applications linéaires décrites ci-dessus pour chacun des deux changements de variable). Il s'est rendu compte qu'elle s'exprimait facilement en fonction des matrices de chacun des deux changements de variable : c'est la naissance du produit matriciel.

Proposition 20.1.16 (Associativité)

Soit $n, p, q, r \in \mathbb{N}^*$ *et* $(A, B, C) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K}) \times \mathcal{M}_{q,r}(\mathbb{K})$. *Alors*

$$(AB)C = A(BC).$$

◁ **Éléments de preuve.**

C'est essentiellement une interversion de deux signes $\sum$. ▷

Proposition 20.1.17 (Deuxième associativité externe)

Soit $n, p, q \in \mathbb{N}^*$ *et* $\lambda \in \mathbb{K}$. *Soit* $(A, B) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K})$. *Alors :*

$$(\lambda A)B = \lambda(AB) \qquad \text{et} \qquad A(\lambda B) = \lambda AB.$$

◁ **Éléments de preuve.**

Vérification immédiate coefficient par coefficient : on se ramène à l'associativité dans $\mathbb{K}$. ▷

Proposition 20.1.18 (Distributivités)

Soit $n, p, q \in \mathbb{N}^*$,

1. $\forall (A, B, C) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K})$, $(A + B)C = AC + BC$
2. $\forall (A, B, C) \in \mathcal{M}_{n,p}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K}) \times \mathcal{M}_{p,q}(\mathbb{K})$, $A(B + C) = AB + AC$.

◁ **Éléments de preuve.**
Vérification immédiate coefficient par coefficient, c'est juste regrouper deux sommes et utiliser la distributivité dans $\mathbb{K}$. ▷

La combinaison des deux propriétés ci-dessus montre que sous réserve de compatibilité des formats,

$$A(\lambda B + \mu C) = \lambda AB + \mu AC,$$

et de même pour la deuxième distributivité. Ces relations traduisent la bilinéarité du produit matriciel.

Avertissement 20.1.19
Le produit matriciel n'est pas commutatif, même lorsque les tailles sont compatibles pour effectuer les opérations dans les deux sens (par exemple pour des matrices carrées).

Exemple 20.1.20
Comparer $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Dans de nombreuses situations, une description plus globale du produit matriciel permet de mieux visualiser les calculs et de moins s'embrouiller sur la position des éléments au sein des matrices en jeu. Le résultat suivant, même si en théorie il donne exactement le même nombre d'opérations à faire, permet souvent poue cette raison d'effectuer des produits matriciels plus rapidement, en particulier lorsque la matrice de droite (ici B) a beaucoup de 0.

Proposition 20.1.21 (Description explicite du produit matriciel, par colonnes)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{i,j}) \in \mathcal{M}_{p,q}(\mathbb{K})$. Notons $C_1(A), \dots, C_p(A)$ les colonnes de la matrice A. Alors la j-ième colonne $C_j(AB)$ du produit AB est :

$$C_j(AB) = \sum_{i=1}^{p} b_{i,j} C_i(A).$$

Ainsi, la j-ième colonne du produit AB est obtenu en faisant la combinaison linéaire des colonnes de A par les coefficients de la j-ième colonne de B :

$$\left(\begin{array}{c|c|c} & & \\ C_1(A) & \dots & C_p(A) \\ & & \end{array}\right) \times \left(\begin{array}{c|c|c} \cdots & b_{1,j} & \cdots \\ & \vdots & \\ \cdots & b_{p,j} & \cdots \end{array}\right) = \left(\begin{array}{c|c|c} & & \\ \cdots & b_{1,j}C_1(A) + \cdots + b_{p,j}C_p(A) & \cdots \\ & & \end{array}\right)$$

◁ **Éléments de preuve.**
Vérifier qu'on retrouve bien la même chose, coefficient par coefficient ▷

La description initiale du produit matriciel présentant une certaine dualité entre les lignes de A et les colonnes de B et inversement, on obtient une description duale par combinaison des lignes de B cette fois. Cette description est notamment très efficace lorsque la matrice A possède beaucoup de 0.

Proposition 20.1.22 (Expression du produit à l'aide des lignes)
Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}$, dont les lignes sont $L_1(B), \dots, L_p(B)$. Soit $i \in [\![1, n]\!]$. La i-ième ligne de AB et

$$L_i(AB) = \sum_{j=1}^{p} a_{i,j} L_j(B).$$

Ainsi :

$$\left(\begin{array}{ccc} \vdots & & \vdots \\ \hline a_{i,1} & \cdots & a_{i,p} \\ \hline \vdots & & \vdots \end{array}\right)\left(\begin{array}{c} L_1(B) \\ \hline \vdots \\ \hline L_p(B) \end{array}\right) = \left(\begin{array}{c} \vdots \\ \hline a_{i,1}L_1(B)+\cdots+a_{i,p}L_p(B) \\ \hline \vdots \end{array}\right).$$

Exemple 20.1.23

1. Calculer $\begin{pmatrix} 2 & 6 & 12 \\ 2 & -7 & 2 \\ 8 & 5 & 11 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
2. Quel est l'effet de la multiplication à droite par $C_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}$?
3. Effet de la multiplication à gauche par C_n ?
4. Mêmes questions avec la matrice « de Jordan » $J_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$
5. Calculer pour tout $k \in \mathbb{N}$, C_n^k et J_n^k. Que constatez-vous ?

Définition 20.1.24 (Matrice identité)

La matrice identité I_n est la matrice carrée de format (n,n) dont tous les coefficients sont nuls sauf les coefficients diagonaux égaux à 1 :

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \qquad \text{soit:} \qquad \forall (i,j) \in [\![1,n]\!]^2,\ [I_n]_{(i,j)} = \delta_{i,j}.$$

Proposition 20.1.25

La matrice I_n est neutre pour le produit matriciel. Plus précisément :

- *Pour toute matrice $M \in \mathcal{M}_{m,n}(\mathbb{K})$, $MI_n = M$ $(m \in \mathbb{N}^*)$;*
- *Pour toute matrice $N \in \mathcal{M}_{n,p}(\mathbb{K})$, $I_nN = N$ $(p \in \mathbb{N}^*)$.*

◁ **Éléments de preuve.**

Vérification facile par description sur les lignes ou colonnes. ▷

Dans le même ordre d'idée, on peut exprimer le produit de deux matrices élémentaires :

Proposition 20.1.26 (Produit des éléments de la base canonique)

Soit $(E_{i,j})$ la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$, $(E'_{j,k})$ la base canonique de $\mathcal{M}_{p,q}(\mathbb{K})$ et $(E''_{i,k})$ la base

canonique de $\mathcal{M}_{n,q}(\mathbb{K})$. Soit $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$, et $(k,\ell) \in [\![1,p]\!] \times [\![1,q]\!]$. Alors :

$$E_{i,j} \times E'_{k,\ell} = \delta_{j,k} E''_{i,\ell} = \begin{cases} E''_{i,\ell} & \text{si } j = k \\ 0 & \text{sinon.} \end{cases}$$

◁ **Éléments de preuve.**

De même, c'est assez facile par description du produit par les lignes ou colonnes ▷

I.4 Transposition

On décrit dans ce paragraphe une autre construction sur les matrices, définissant une application linéaire sur l'ensemble des matrices.

Définition 20.1.27 (Transposée d'une matrice)

Soit $A = (a_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} \in \mathcal{M}_{n,p}(A)$. Alors la matrice transposée de A, notée ${}^tA$ ou $A^\top$, est la matrice de $\mathcal{M}_{p,n}\mathbb{K}$ définie par :

$${}^tA = A^\top = (a_{j,i})_{(i,j)\in[\![1,p]\!]\times[\![1,n]\!]}.$$

Ainsi, si $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}$, alors ${}^tA = A^\top = \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & & \vdots \\ a_{1,p} & \cdots & a_{n,p} \end{pmatrix}$

Le programme impose plutôt la notation $A^\top$, dans un souci d'universalité. Nous essayerons de nous y conformer, mais il est possible que je laisse traîner ici ou là un ${}^tA$.

Exemple 20.1.28

Transposée de $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$?

Proposition 20.1.29 (Linéarité de la transposition)

L'application $A \mapsto A^\top$ de $\mathcal{M}_{n,p}(\mathbb{K})$ dans $\mathcal{M}_{p,n}(\mathbb{K})$ est une application linéaire :

$$\forall (A,B) \in \mathcal{M}_{n,p}(\mathbb{K}), \forall (\lambda,\mu) \in \mathbb{K}, \quad (\lambda A + \mu B)^\top = \lambda A^\top + \mu B^\top$$

Proposition 20.1.30 (Transposition d'un produit)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors

$$(AB)^\top = B^\top A^\top.$$

◁ **Éléments de preuve.**

Vérification directe coefficient par coefficient. ▷

II Matrices carrées

II.1 L'algèbre $\mathcal{M}_n(\mathbb{K})$

Le terme « algèbre » se réfère à une structure dans laquelle on peut additionner et multiplier les objets entre eux, ainsi que les multiplier par un scalaire, avec un certain nombre de propriétés (associativités internes et externes, commutativité de l'addition, distributivité, compatibilité de la multiplication par le neutre des scalaires, neutre et existence des opposés). Cette structure sera étudiée plus généralement plus tard. Ici, le point important qui fait qu'on dispose d'une structure d'algèbre sur $\mathcal{M}_n(\mathbb{K})$ est le fait que le produit de deux matrices carrées d'ordre n est toujours bien défini, et est lui-même une matrice d'ordre n.
En particulier en oubliant la multiplication par un scalaire, on obtient une structure d'anneau :

Proposition 20.2.1 (Structure de $\mathcal{M}_n(\mathbb{K})$)
L'ensemble $\mathcal{M}_n(\mathbb{K})$ muni de l'addition et de la multiplication des matrices est un anneau non commutatif.

On en déduit notamment que les procédés calculatoires valables dans un anneau sont valables pour les matrices :

Théorème 20.2.2 (Factorisation de $A^n - B^n$)
Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = BA$. Alors pour tout $n \in \mathbb{N}^*$

$$A^n - B^n = (A - B)\sum_{k=0}^{n-1} A^{n-1-k}B^k.$$

Corollaire 20.2.3 (Factorisation de $I_n - A^n$)
En particulier, pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$,

$$I_n - A^n = (I_n - A)\sum_{k=0}^{n-1} A^k.$$

Théorème 20.2.4 (Formule du binôme)
Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = BA$. Alors, pour tout $n \in \mathbb{N}$,

$$(A + B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}.$$

Exemple 20.2.5

1. Déterminer A^n, pour tout $n \in \mathbb{N}$, où $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$
2. Plus généralement, si J_n est définie comme dans l'exemple 20.1.23, calculer $(aI_n + J_n)^k$, pour tout $k \in \mathbb{N}$.
3. Calculer $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^n$

On en déduit une méthode assez efficace, mais pas toujours réalisable, de calcul des puissances.

Remarque 20.2.6

L'anneau $\mathcal{M}_n(\mathbb{K})$ est-il intègre ? (i.e. vérifie-t-il $AB = 0 \Longrightarrow A = 0$ ou $B = 0$?)

Définition 20.2.7 (Polyôme annulateur, polyôme minimal)

1. Un polynôme annulateur de $A \in \mathcal{M}_n(\mathbb{K})$ est un polynôme $P = \sum_{k=0}^{q} a_k X^k$ à coefficients dans $\mathbb{K}$ tel que
$$P(A) = \sum_{k=0}^{q} a_k A^k = 0.$$
2. Le polynôme minimal est le polynôme annulateur unitaire (donc non nul) de plus petit degré.

Proposition 20.2.8 (Existence et unicité du polynôme minimal, à moitié admis)

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors :

1. *A admet un polynôme annulateur non nul (admis momentanément)*
2. *A admet un unique polynôme minimal.*

◁ **Éléments de preuve.**

Existence découlant de la propriété fondamentale de $\mathbb{N}$. Unicité : s'il y en a deux distincts, obtenir une contradiction en formant la différence. ▷

Exemple 20.2.9

1. Déterminer un polynôme annulateur de $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.
2. Déterminer un polynôme annulateur de C_n
3. Déterminer un polynôme annulateur de J_n. Un polynôme minimal.

Le polynôme annulateur peut être efficace pour la recherche des puissances successives d'une matrice A.

Méthode 20.2.10 (Calcul de A^n à l'aide d'un polynôme annulateur)

- Déterminer un polynôme annulateur P de petit degré de A.
- Chercher le reste R_n de la division euclidienne de X^n par P.
- Évaluer l'égalité de division euclidienne en A, il reste $A^n = R_n(A)$. Ainsi A^n s'exprime comme combinaison linéaire des premières puissances de A.

Exemple 20.2.11

Calculer $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^n$ à l'aide d'un polynôme annulateur.

II.2 Matrices triangulaires et diagonales

La matrice I_n a une particularité notable : elle est nulle, sauf sur sa diagonale. Les matrices vérifiant cette propriété jouent un rôle central, notamment dans la théorie de la diagonalisation. En effet, les applications linéaires associées $X \mapsto AX$ laissent stables les axes définis par les vecteurs de la base, et sont donc faciles à étudier. En particulier les produits de matrices diagonales sont simples à exprimer (et donc aussi les puissances). C'est une des motivations de la théorie de la diagonalisation. Nous présentons d'autres formes de matrices (matrices triangulaires).

Définition 20.2.12 (Matrice diagonale)

Soit D une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que D est une matrice diagonale si tous ses coefficients sont nuls, à l'exception éventuellement de ses coefficients diagonaux.

Ainsi, une matrice diagonale est de la forme :

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

Définition 20.2.13 (Matrice triangulaire)

Soit T une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que T est une matrice triangulaire supérieure (resp. inférieure) si tous ses coefficients situés strictement en-dessous (resp. au-dessus) de sa diagonale sont nuls.

Ainsi, une matrice triangulaire supérieure (resp. inférieure) est de la forme :

$$T = \begin{pmatrix} \bullet & \cdots & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \bullet \end{pmatrix} \quad \left(\text{resp. } T = \begin{pmatrix} \bullet & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \bullet & \cdots & \cdots & \bullet \end{pmatrix}\right)$$

où les $\bullet$ désignent des coefficients quelconques. On définit également les matrices strictement triangulaires supérieures ou inférieures, nulles également sur la diagonale.

Notation 20.2.14

Nous noterons dans ce cours $\mathcal{D}_n(\mathbb{K})$ l'espace des matrices diagonales d'ordre n, $\mathcal{T}_n^+(\mathbb{K})$ l'espace des matrices triangulaires supérieures, $\mathcal{T}_n^-(\mathbb{K})$ l'espace des matrices triangulaires inférieures.

Proposition 20.2.15 (Produit de deux matrices triangulaires ou diagonales)

(i) Le produit de deux matrices triangulaires supérieures est une matrice triangulaire supérieure.

(ii) Le produit de deux matrices triangulaires inférieures est une matrice triangulaire inférieure.

(iii) Le produit de matrices diagonales est une matrice diagonale.

◁ **Éléments de preuve.**

Cas des triangulaires supérieures : la k-ième colonne du produit AB est une CL des k premières colonnes de A. ▷

On peut regarder plus précisément les termes diagonaux de ces produits. On obtient sans peine le complément suivant :

Proposition 20.2.16 (Termes diagonaux du produit de deux matrices triangulaires)

1. *On obtient les termes diagonaux du produit de deux matrices triangulaires par produit des termes diagonaux correspondants des deux matrices initiales :*

$$\begin{pmatrix} a_{1_1} & \cdots & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n} \end{pmatrix} \begin{pmatrix} b_{1,1} & \cdots & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & b_{n,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,1} & \cdots & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{n,n}b_{n,n} \end{pmatrix}$$

2. *En particulier, si l'une des deux matrices est strictement triangulaire supérieure, le produit l'est aussi.*

◁ **Éléments de preuve.**

Revenir à la description coefficient par coefficient. ▷

Vu l'importance de cette règle, on explicite pour les matrices diagonales :

Proposition 20.2.17 (Produit de matrices diagonales)

On a :

$$\begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_n \end{pmatrix} \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix} = \begin{pmatrix} c_1d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_nd_n \end{pmatrix}.$$

En particulier, pour tout $k \in \mathbb{N}$, on a :

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}^k = \begin{pmatrix} d_1^k & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n^k \end{pmatrix}.$$

II.3 Matrices symétriques et antisymétriques

Voici d'autres matrices carrées jouant un rôle important, notamment en algèbre bilinéaire :

Définition 20.2.18 (Matrices symétriques, antisymétriques)

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

(i) On dit que A est *symétrique* si $A = A^\top$.

(ii) On dit que A est *antisymétrique* si $A = -A^\top$.

On note $\mathcal{S}_n(\mathbb{K})$ l'ensemble des matrices symétriques de $\mathcal{M}_n(\mathbb{K})$, et $\mathcal{A}_n(\mathbb{K})$ l'ensemble des matrices antisymétriques.

Théorème 20.2.19 (Supplémentarité de $\mathcal{S}_n(\mathbb{K})$ et $\mathcal{A}_n(\mathbb{K})$)

On suppose que $\mathrm{car}(\mathbb{K}) \neq 2$. *Soit $M \in \mathcal{M}_n(\mathbb{K})$, il existe un unique couple $(S, A) \in \mathcal{S}_n(\mathbb{K}) \times \mathcal{A}_n(\mathbb{K})$ tel que $M = S + A$.*

Remarque 20.2.20

À quoi ressemble la diagonale d'une matrice antisymétrique ?

II.4 Matrices inversibles

Définition 20.2.21 (Matrice inversible)

Une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est dite inversible s'il existe $N \in \mathcal{M}_n(\mathbb{K})$ tel que

$$MN = I_n \qquad \text{et} \qquad NM = I_n.$$

Avertissement 20.2.22

Par définition, une matrice inversible est toujours une matrice carrée.

Proposition 20.2.23 (Caractérisation de l'inversibilité, admis pour le moment)

Une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement s'il existe $N \in \mathcal{M}_n(\mathbb{K})$ tel que $MN = I_n$ OU $NM = I_n$.

◁ **Éléments de preuve.**

On verra que cela résulte de caractérisations de l'injectivité ou la surjectivité d'applications linéaires, issues d'un résultat important appelé formule du rang, reliant la dimesion de l'image et la dimension du noyau d'une application linéaire. ▷

Définition 20.2.24 (Groupe linéaire)

L'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est noté $\mathrm{GL}_n(\mathbb{K})$, et est appelé *n-ième groupe linéaire.*

Proposition 20.2.25 (Structure de $\mathrm{GL}_n(\mathbb{K})$)

$(\mathrm{GL}_n(\mathbb{K}), \times)$ *est un groupe (c'est le groupe des inversibles de l'anneau $\mathcal{M}_n(\mathbb{K})$).*

Conformément aux propriétés générales concernant l'inverse de produits dans un anneau, on a :

Proposition 20.2.26 (Inverse d'un produit)

Soit A et B deux matrices inversibles de $\mathcal{M}_n(\mathbb{K})$. Alors AB est inversible, et son inverse est $B^{-1}A^{-1}$.

◁ **Éléments de preuve.**

Comme dans tout groupe ! N'oubliez pas l'interversion ! ▷

Par ailleurs, l'inversion commute avec la transposition :

Proposition 20.2.27 (Inverse d'une transposée)

Soit A une matrice inversible. Alors $A^\top$ l'est également, et

$$(A^\top)^{-1} = (A^{-1})^\top.$$

◁ **Éléments de preuve.**
Exprimer le produit $A^\top \cdot (A^{-1})^\top$. ▷

Voici un exemple important de famille de matrices inversibles. Important car, comme on va le voir, via un pivot de Gauss, l'étude de l'inversibilité d'une matrice explicite peut toujours se ramener à ce cas.

Proposition 20.2.28 (CNS d'inversibilité pour les matrices triangulaires)
Soit $T \in \mathcal{T}_n^+(\mathbb{K})$ une matrice triangulaire. Alors T est inversible si et seulement si tous ses coefficients diagonaux sont non nuls (donc inversibles), et dans ce cas, T^{-1} est une matrice triangulaire dont les coefficients diagonaux sont les inverses des coefficients diagonaux de T.

◁ **Éléments de preuve.**
On peut s'y prendre à la main, en montrant que le système $TX = 0$ admet le vecteur nul comme unique solution (le système est triangulaire donc facile à résoudre en partant d'en bas). Quelle information liée à la bijectivité cela apporte-t-il ?
La description des coefficients diagonaux découle de l'expression des coefficients diagonaux d'un produit de matrices triangulaires. ▷

En particulier, une matrice diagonale est inversible si et seulement si tous ses coefficients diagonaux sont non nuls, et dans ce cas :

$$D^{-1} = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}^{-1} = \begin{pmatrix} d_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n^{-1} \end{pmatrix}.$$

Dans le cas de matrices 2×2, on obtient facilement une formule explicite de l'inverse, à connaître par coeur vu son utilité :

Théorème 20.2.29 (Inverse des matrices 2×2 par la comatrice)
Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$. Alors M est inversible si et seulement si $ad - bc \neq 0$, et

$$M^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

◁ **Éléments de preuve.**
Le plus simple est ici de faire une vérification. ▷

Définition 20.2.30 (Déterminant d'une matrice 2×2)
La quantité $ad - bc$ est appelée *déterminant* de M, et est noté $\det(M)$ ou $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Remarque 20.2.31
Il existe une notion de déterminant pour des matrices carrées de taille quelconque n. Nous définirons cette notion générale dans le chapitre suivant. La non nullité du déterminant caractérise alors l'inversibilité de la matrice, et comme dans le cas $n = 2$, il existe une formule générale de l'inverse d'une matrice basée sur la notion de comatrice. Dans des situations concrètes, cette formule est cependant assez peu efficace, sauf pour $n = 2$, et éventuellement $n = 3$ (et encore...)

Enfin, voici une méthode efficace pour calculer l'inverse lorsqu'on connaît un polynôme annulateur.

Méthode 20.2.32 (Calcul de l'inverse d'une matrice avec un polynôme annulateur)

Soit P un polynôme annulateur. Si A est inversible, quitte à multiplier plusieurs fois par A^{-1}, il existe alors un polynôme annulateur à coefficient constant non nul, et quitte à diviser par cette constante, on peut supposer que ce coefficient constant est égal à 1. En notant $P(X) = XQ(X) - 1$, on a alors :

$$0 = P(A) = AQ(A) - I \qquad \text{donc:} \qquad -AQ(A) = I.$$

Ainsi, $-Q(A)$ est inverse de A.

Dans des situations numériques, on aura davantage intérêt à utiliser la méthode décrite dans le paragraphe suivant, basé sur la technique du pivot de Gauss.

III Pivot de Gauss et matrices équivalentes par lignes

Nous exposons dans ce paragraphe le principe du pivot de Gauss, dont nous verrons par la suite qu'il permet de calculer explicitement l'inverse de matrices. Cette méthode a bien d'autres applications, comme par exemple la résolution de systèmes linéaires, que nous évoquerons aussi, ou le calcul du rang d'une matrice, qu'on verra dans un chapitre ultérieur, ou encore le calcul des déterminants. Par ailleurs, l'algorithme du pivot s'implémente facilement informatiquement, ce qui permet d'écrire des programmes calculant l'inverse, résolvant des systèmes linéaires, calculant le rang d'une matrice, ou calculant un déterminant.

III.1 Opérations sur les lignes d'une matrice

Étant donnée une matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ dont les lignes sont désignées par $L_1, \dots, L_n$, les opérations admissibles pour le pivot sont les trois opérations suivantes décrites ci-dessous :

Définition 20.3.1 (Opérations admissibles sur les lignes d'une matrice)

- Opération de **permutation** : échange des lignes L_i et L_j de la matrice, codée par $L_i \leftrightarrow L_j$
- Opération de **dilatation** : multiplication d'une ligne L_i par un scalaire (réel ou complexe) *non nul* λ, codée par $L_i \leftarrow \lambda L_i$
- Opération de **transvection** : ajout à une ligne donnée L_i d'une autre ligne L_j éventuellement multipliée par un scalaire λ (le résultat remplaçant la ligne L_i). Cette opération est codée par $L_i \leftarrow L_i + \lambda L_j$.

Remarque 20.3.2

- L'itération de la première opération (permutation) autorise les permutations quelconques des lignes d'une matrice. En effet, nous justifierons plus tard que toute permutation de $\mathcal{S}_n$ peut se décomposer en composition de transpositions, c'est-à-dire en produit de permutations particulières n'effectuant que l'échange de deux valeurs. On peut remarquer que la justification de la correction de certains algorithmes de tri basés sur des transpositions (par exemple le tri par insertion ou le tri à bulles) permet de prouver cette affirmation.
- La combinaison des deux dernières règles amène la règle suivante souvent bien pratique :

$$L_i \leftarrow \lambda L_i + \mu L_j, \text{ si } \lambda \neq 0.$$

  Attention à ce que le coefficient devant la ligne modifiée par cette opération ne soit pas nul !

Avertissement 20.3.3

Les différentes opérations s'effectuent successivement (même si on les note ensemble dans la même étape) : on ne peut pas effectuer des opérations simultanées. Ainsi, si on a dans la même étape deux opérations $L_1 \leftarrow L_1 + L_2$ et $L_2 \leftarrow L_1 + L_2$, cela signifie que la seconde est effectuée avec la ligne L_1 obtenue à l'issue de la première opération, et non avec la ligne L_1 initiale.

Définition 20.3.4 (Équivalence par ligne)

Deux matrices A et B sont équivalentes par lignes si et seulement si B s'obtient de A en effectuant un nombre fini (pouvant être nul) d'opérations sur les lignes successives. On note $A \equiv_L B$.

Proposition 20.3.5

La relation $\equiv_L$ est une relation d'équivalence.

◁ **Éléments de preuve.**

En effet, seule la symétrie mérite qu'on s'y attarde : il suffit de savoir faire les opérations inverses de celles allant de A à B, dans le sens inverse.

- L'opération inverse de $L_i \leftrightarrow L_j$ est $L_i \leftrightarrow L_j$
- L'opération inverse de $L_i \leftarrow \lambda L_i$ (pour $\lambda \neq 0$) est $L_i \leftarrow \frac{1}{\lambda} L_i$
- L'opération inverse de $L_i \leftarrow L_i + \lambda L_j$ est $L_i \leftarrow L_i - \lambda L_j$.

▷

Évidemment, il est aussi possible de définir des opérations sur les colonnes, et de définir la relation $\equiv_C$ d'équivalence par colonnes de la même façon.

III.2 Échelonnement d'une matrice par la méthode du pivot de Gauss

Le but de ce paragraphe est de donner un algorithme permettant de trouver explicitement une matrice échelonnée (en lignes) équivalente par lignes à une matrice donnée.

Définition 20.3.6 (Matrice échelonnée)

Soit m et n deux entiers non nuls, et $M = (a_{i,j})_{\substack{1 \leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$. On dit que M est une matrice échelonnée s'il existe un entier $k \in [\![1, m]\!]$ et une suite croissante $j_1 < j_2 < \cdots < j_k$ d'éléments de $[\![1, n]\!]$ tels que :

(i) $\forall i \in [\![1, k]\!]$, $a_{i,j_i} \neq 0$;

(ii) $\forall i \in [\![1, k]\!]$, $\forall j \in [\![1, j_i - 1]\!]$, $a_{i,j} = 0$;

(iii) $\forall i \in [\![k+1, m]\!]$, $\forall j \in [\![1, n]\!]$, $a_{i,j} = 0$

Autrement dit, les lignes nulles sont regroupées au bas de la matrice (lignes $k+1$ à m), les autres lignes sont classées suivant la position de leur premier élément non nul, ces positions étant deux à deux distinctes. Une matrice échelonnée admet donc la représentation suivante :

$$M = \begin{pmatrix}
0 & \cdots & 0 & a_{1,j_1} & \bullet & & & \cdots & & & \bullet \\
0 & \cdots & & \cdots & 0 & a_{2,j_2} & \bullet & & \cdots & & \bullet \\
\vdots & & & & & & & & & & \vdots \\
0 & & & \cdots & \cdots & & 0 & a_{k,j_k} & \bullet & \cdots & \bullet \\
0 & & & & \cdots & & & \cdots & & & 0 \\
\vdots & & & & & & & & & & \vdots \\
0 & & & & \cdots & & & \cdots & & & 0
\end{pmatrix},$$

les coefficients indiqués d'un • étant quelconques, et les a_{i,j_i} étant non nuls.

Méthode 20.3.7 (Algorithme du pivot de Gauss, ou élimination de Gauss-Jordan)

1. On cherche la première colonne non nulle de la matrice A.
2. Sur cette colonne, on effectue un choix de pivot : n'importe quel coefficient non nul de la colonne convient, mais on a intérêt à choisir un pivot donnant le moins de calculs possible, si on effectue ces calculs à la main. Il y a trois critères pour cela :
   - Le pivot lui-même doit être facile à inverser. L'idéal est un pivot égal à 1.
   - Les autres coefficients de la ligne du pivot doivent être « simples », de préférence des entiers.
   - Plus il y a de zéros sur la ligne contenant le pivot, moins il y aura de calculs !
3. On fait un échange de lignes pour ramener le pivot choisi sur la première ligne.
4. On annule tous les coefficients situées sous le pivot à l'aide d'opérations élémentaires $L_i \leftarrow L_i + \lambda L_1$, ou bien pour éviter d'introduire des fractions, $L_i \leftarrow \alpha L_i + \beta L_1$, avec $\alpha \neq 0$
5. On recommence récursivement en considérant la sous-matrice située strictement en-dessous à droite du pivot.

Remarque 20.3.8

Bien entendu, les critères de choix du pivot sont donnés ici en vue d'un calcul à la main. En vue d'une implémentation sur ordinateur, le choix du pivot doit se faire de sorte à diminuer au maximum les erreurs d'arrondi. Les critères sont, dans cette optique, différents de ceux énoncés ci-dessus. À première approximation, le choix d'un pivot de valeur absolue maximale est à privilégier.

Nous donnons ci-dessous une description purement algorithmique en pseudo-code un peu lâche.

Algorithme 20.1 : Pivot de Gauss

Entrée : A une matrice
Sortie : A' matrice échelonnée équivalente par lignes à A

Initialiser l'indice j de colonne à 1;
Initialiser l'indice i de ligne à 1;
tant que *les indices i et j ne font pas sortir de la matrice A* **faire**
  si *le bas de la colonne j (en-dessous de la ligne i au sens large) est nul* **alors**
    Passer à la colonne suivante ($j \leftarrow j + 1$)
  sinon
    Placer un élément non nul en position (i, j) par une opération $L_i \leftrightarrow L_k$;
    pour $k \leftarrow i + 1$ **à** *dernière ligne* **faire**
      $L_k \leftarrow L_k + \lambda L_i$, où $\lambda = -\dfrac{L_k[j]}{L_i[j]}$
    fin pour
    Passer à la ligne suivante ($i \leftarrow i + 1$);
    Passer à la colonne suivante ($j \leftarrow j + 1$)
  fin si
fin tant que
renvoyer A

Théorème 20.3.9 (Terminaison et correction de l'algorithme du pivot de Gauss)

Cet algorithme se termine, et la matrice A' renvoyée par l'algorithme est échelonnée et équivalente par lignes à A.

◁ **Éléments de preuve.**

- La terminaison s'étudie avec le variant de boucle $n+m-i-j$, entier positif strictement décroissant au fil des itérations de la boucle while.
- La correction s'étudie avec l'invariant de boucle :
  « avant d'entrer dans la boucle, les $i-1$ premières lignes de A sont échelonnées, et les coefficients en position (i', j'), $i' \geqslant i$, $j' < j$ sont tous nuls. »

▷

Exemple 20.3.10

Recherche d'une matrice échelonnée équivalente par lignes à :

$$A = \begin{pmatrix} 1 & -4 & -2 & 3 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Méthode 20.3.11 (Pivot remontant)

- Une fois la matrice échelonnée A' obtenue, on peut continuer en reprenant les pivots en sens inverse, à partir du dernier, et en annulant les coefficients situés au-dessus de ce pivot.
- Clairement, cet algorithme se termine, ne modifie pas l'échelonnement (les pivots ne sont pas modifiés, ni les coefficients nuls qui déterminent l'échelonnement).
- De plus, les opérations faites au-dessus d'un pivot laissent nuls les coefficients déja annulés au dessus d'un pivot situé au-delà.
- Ainsi, à l'issue de cet algorithme, on obtient une matrice A échelonnée, tel que tous les coefficients situés en-dessous ET au-dessus des pivots soient nuls.
- En divisant chacune des lignes contenant un pivot par la valeur de ce pivot, on peut même normaliser la matrice échelonnée obtenue, c'est-à-dire s'arranger pour que tous les pivots soient égaux à 1.

Définition 20.3.12 (Matrice échelonnée réduite)

Une matrice M est dite échelonnée réduite si :

(i) elle est échelonnée,

(ii) ses pivots sont tous égaux à 1

(iii) les coefficients au-dessus des pivots sont tous nuls

Ainsi, la méthode précédente montre que toute matrice est équivalente par lignes à une matrice échelonnée réduite (appelée réduite de Gauss).

Remarque 20.3.13

Toutes les opérations décrites dans l'algorithme du pivot initial, et l'algorithme du pivot remontant ne nécessitent pas l'opération de dilatation, à part la normalisation finale. Ainsi, si le but est de trouver une matrice échelonnée (y compris avec pivot remontant), non normalisée, on peut se dispenser de ce type d'opérations. Cette remarque sera utile lors de l'étude des déterminants.

Note Historique 20.3.14 (Pivot de Gauss)

- Le nom de la méthode du pivot est un hommage aux deux mathématiciens Gauss et Jordan.
- Gauss utilise cette méthode dans ses ouvrages, en l'appelant *élimination ordinaire*, ou, en latin (langue qu'il emploie pour ses publications scientifiques), *eliminatio vulgaris*.

- Gauss et Jordan utilisent cette méthode d'élimination ordinaire notamment dans le cadre de la classification des formes quadratiques.
- Ce n'est que vers 1880 que Frobenius publie plusieurs mémoires faisant un état des lieux de la théorie des matrices, et élucide complètement à l'occasion la théorie des systèmes linéaires à coefficients réels ou complexes.
- Mais la méthode est en fait beaucoup plus ancienne : elle est déjà exposée dans un ouvrage chinois du III^{e} siècle *Jiuzhang suanshu* (Prescriptions de calcul en 9 chapitres) de Liu Hui. Le huitième chapitre est entièrement consacré à la méthode d'élimination par pivot, appelée *fang cheng* (disposition, ou modèle rectangulaire).
- La méthode elle-même est sans doute plus ancienne, puisque Liu Hui en attribue la paternité à Chang Ts'ang, 3 ou 4 siècles plus tôt, auteur d'un ouvrage aujourd'hui disparu.

III.3 Interprétation matricielle des opérations du pivot

Nous voyons dans ce paragraphe comment les opérations sur les lignes (ou colonnes) de A se traduisent par la multiplication de A par certaines matrices codant ces opérations.

Définition 20.3.15 (Matrices de codage des opérations élémentaires)

Soit $n \in \mathbb{N}^*$. On définit les trois familles suivantes de matrices :

(i) Codage des échanges de lignes (matrice de transposition) :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad i \neq j, \quad E(i,j) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & 0 & \ddots & & 1 & & \vdots \\ \vdots & & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & 1 & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ i \\ \\ \\ j \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & & i & & & j & & \end{matrix}$$

(ii) Codage d'une combinaison linéaire (matrice de transvection) : pour $i \neq j$:

$$E_{i,j}(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ i \\ \\ \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & & j & & \end{matrix}$$

(iii) Codage de la multiplication d'une ligne par un scalaire (matrice de dilatation) :

$$E_i(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & & i & & & \end{matrix}$$

Proposition 20.3.16 (Interprétation matricielle des opérations élémentaires sur les lignes)

Soit $M \in \mathcal{M}_{m,n}(\mathbb{K})$. Soit $i \neq j$ dans $[\![1,m]\!]$, et $\lambda \in \mathbb{K}$.

(i) La matrice N obtenue de M par l'opération $L_i \leftrightarrow L_j$, est $N = E(i,j) \cdot M$;

(ii) La matrice N obtenue de M par l'opération $L_i \leftarrow L_i + \lambda L_j$ est $N = E_{i,j}(\lambda) \cdot M$;

(iii) La matrice N obtenue de M par l'opération $L_i \leftarrow \lambda L_i$, $\lambda \neq 0$, est $N = E_i(\lambda) \cdot M$.

◁ **Éléments de preuve.**

Utiliser la description du produit matriciel par les lignes. ▷

Ainsi, partant d'une matrice M à laquelle on applique le pivot de Gauss, on obtient une matrice N, et une matrice P telles que $N = PM$, telle que P soit un produit de matrices de transvection, de dilatation et de permutation.
La validité du pivot de Gauss provient alors du résultat suivant, exprimant que toute opération valide du pivot est réversible :

Proposition 20.3.17 (Inversibilité des matrices de codage des opérations élémentaires)
Pour $i \neq j$, les matrices $E(i,j)$, $E_{i,j}(\lambda)$ sont inversibles, ainsi que $E_i(\lambda)$ lorsque $\lambda \neq 0$.

◁ **Éléments de preuve.**
Pour beaucoup, cela se règle en remarquant qu'elles sont triangulaires. Quant à $E(i,j)$, son inverse n'est pas dur à trouver. ▷

On peut préciser ce résultat en exprimant les inverses de ces matrices :

Proposition 20.3.18 (Inverses des matrices de codage des opérations élémentaires)
- $E(i,j)^{-1} = ...$
- $E_{i,j}(\lambda)^{-1} = ...$
- $E_i(\lambda)^{-1} = ...$

◁ **Éléments de preuve.**
Pensez en terme de réversibilité des opérations sur les lignes.
On peut vérifier ensuite, une fois deviné l'expression de l'inverse. ▷

III.4 Calcul pratique de l'inverse d'une matrice

Proposition 20.3.19 (Caractérisation de l'inversibilité par le pivot)
Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si elle est équivalente par lignes à une matrice échelonnée (donc triangulaire) à coefficients diagonaux non nuls.

Dans ce cas, toute matrice échelonnée obtenue par opérations sur les lignes vérifiera cette propriété.

◁ **Éléments de preuve.**
Si A' est échelonnée équivalente par lignes à A, alors elle s'écrit $A' = PA$, où P est une matrice inversible, produit de toutes les matrices d'opérations élémentaires utilisées dans le pivot. Ainsi, A est inversible ssi A' l'est. On remarque ensuite qu'une réduite de Gauss d'une matrice carrée est toujours triangulaire. On est donc ramené à la caractérisation de l'inversibilité des matrices triangulaires. ▷

Théorème 20.3.20
Toute matrice inversible est équivalente par ligne à I_n. On trouve la succession d'opérations à effectuer en opérant un pivot de Gauss, suivi d'un pivot remontant, suivi d'une normalisation

◁ **Éléments de preuve.**
En effet, le résultat précédent montre que la matrice échelonnée réduite obtenue à l'issue de cet algorithme est nécessairement I_n. ▷

Notant P la matrice obtenue par produit des matrices des opérations élémentaires nous ayant amené de A à I_n, on a donc $PA = I_n$. Ainsi, $P = A^{-1}$.
Or, $P = PI_n$, et est donc obtenu en appliquant à la matrice I_n les mêmes opérations sur ses lignes que celles qui ont permis de transformer A en I_n. On en déduit la méthode suivante :

Méthode 20.3.21 (Calcul de l'inverse d'une matrice (seconde méthode, pivot de Gauss))

- Juxtaposer la matrice A et la matrice I_n (séparées d'une barre verticale)
- Effectuer un pivot sur A, en faisant les mêmes opérations sur la matrice I_n, pour obtenir une matrice échelonnée à la place de A
- La matrice A est inversible si et seulement si la matrice échelonnée obtenue est inversible (c'est-à-dire s'il s'agit d'une matrice triangulaire supérieure à coefficients diagonaux non nuls)
- Dans ce cas, faire un pivot remontant, pour annuler les coefficients au dessus de chaque pivot, et toujours en effectuant les mêmes opérations sur la matrice de droite.
- En normalisant les coefficients diagonaux, on obtient à gauche la matrice identité, et à droite la matrice A^{-1}.

Corollaire 20.3.22 (Un système de générateurs de $\mathbf{GL_n(\mathbb{K})}$)

Toute matrice inversible est produit de matrices d'opérations élémentaires.

◁ **Éléments de preuve.**

On a écrit A^{-1} comme produit de matrices d'opérations élémentaires. Tout inverser, ou alors, partir de A^{-1} au lieu de A. ▷

Corollaire 20.3.23 (Caractérisation de l'équivalence par lignes)

Deux matrices A et B sont équivalentes par lignes si et seulement s'il existe une matrice inversible P telle que $B = PA$

◁ **Éléments de preuve.**

Le sens direct a déjà été évoqué. Le sens réciproque est une conséquence immédiate du corollaire précédent. ▷

IV Résolution d'un système linéaire

L'algorithme du pivot de Gauss permet également de résoudre des systèmes linéaires. Cela est en fait une généralisation de la méthode d'inversion. En effet, un système linéaire peut se réécrire sous la forme $AX = B$, où A est la matrice des coefficients du système X la matrice colonne des inconnues et B la matrice colonne des seconds membres (on l'explicite mieux ci-dessous). Si la matrice A du système est inversible, alors cette équation équivaut à $X = A^{-1}B$. Ainsi, le calcul de l'inverse (par la méthode du pivot par exemple) permet de résoudre le système. On va voir que la méthode du pivot permet en fait de résoudre le système plus généralement, mais que dans la situation générale, il n'y aura pas existence et unicité de la solution.

IV.1 Position du problème, réexpression et structure

Nous nous intéressons à la résolution d'un système de n équations linéaires à p inconnues réelles ou complexes (ou plus généralement dans un corps $\mathbb{K}$). Il s'agit donc de trouver tous les p-uplets $(x_1, \ldots, x_p)$

tels que :

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p &= b_1 \\ a_{2,1}x_1 + \cdots + a_{2,p}x_p &= b_2 \\ \vdots & \vdots \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p &= b_n \end{cases} \tag{20.1}$$

où les $a_{i,j}$ et les b_i sont des éléments de $\mathbb{K}$.
Les x_i sont les *inconnues* du système, les $a_{i,j}$ les *coefficients*, et les b_i contituent le *second membre*.

Proposition 20.4.1

En posant

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \qquad \text{et} \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix},$$

le système précédent équivaut à l'équation matricielle

$$AX = B$$

de l'inconnue vectorielle X.

Définition 20.4.2 (Système de Cramer)

Le système représenté par $AX = B$ est de Cramer si A est inversible (donc carrée). Dans ce cas, le système a une unique solution $X = A^{-1}B$.

Théorème 20.4.3 (Structure de l'ensemble des solutions)

L'ensemble des solutions $\mathcal{S}$ de l'équation $AX = B$, s'il est non vide, est un sous-espace affine de $\mathbb{K}^p$. Ainsi, si X_0 désigne une solution particulière, et si $\mathcal{S}_0$ désigne l'ensemble des solutions de l'équation homogène associée $AX = 0$, alors :

$$\mathcal{S} = \{X_0 + X \mid X \in \mathcal{S}_0\} = X_0 + \mathcal{S}_0,$$

et $\mathcal{S}_0$ contient la solution nulle et est stable par combinaisons linéaires (sous-espace vectoriel de $\mathbb{R}^p$).

◁ **Éléments de preuve.**

C'est en fait toujours le même principe : on peut montrer que X est solution si et seulement $X - X_0$ est solution de l'équation homogène. On peut aussi remarquer qu'il s'agit d'une fibre du morphisme $\varphi : X \mapsto AX$, et se rattacher à la dercription des fibres

La structure de $\mathcal{S}_0 = \mathrm{Ker}(\varphi)$ se vérifie facilement, et sera justifiée plus tard plus généralement pour tout noyau d'une application linéaire. ▷

IV.2 Système échelonné réduit associé

On commence par la remarque suivante :

Proposition 20.4.4

Soit P une matrice inversible (de format compatible au produit). Alors le système $AX = B$ équivaut au système $PAX = PB$.

Ainsi, effectuer des opérations sur les lignes de A ne modifie pas les solutions du système, si on effectue les mêmes opérations sur les lignes du second membre B (i.e. multiplier par la même matrice d'opération). C'est en fait assez évident, puisque cela consiste à faire les opérations correspondantes directement sur les équations du système.
Si P représente la matrice obtenue en effectuant le produit des matrices d'opérations élémentaires permettant de calculer une réduite de Gauss de A, on peut donc se ramener à un système dont la matrice est échelonnée.

Méthode 20.4.5 (Se ramener à un système dont la matrice est échelonnée réduite)
Effectuer un pivot, suivi d'un pivot remontant, suivi d'une normalisation sur la matrice A, en efectuant les mêmes opérations sur B. Pour cela, on aura intérêt à présenter comme pour le calcul d'inverse, en juxtaposant initialement les matrices A et B (séparées d'une barre verticale).

On peut bien sûr s'arrêter avant et se contenter d'un système échelonné non réduit, qui peut parfois se résoudre simplement aussi, en remontant.
Un système échelonné réduit (i.e. associé à une matrice A échelonnée réduite) n'admet pas toujours une solution unique, ni même une solution tout court. Pour trouver une solution particulière, il y aura donc des choix à faire. Ainsi, si la dernière équation fait par exemple intervenir 3 inconnues, en fixer 2 à sa guise impose la troisième. Ainsi, en remontant un système échelonné, on ajoute à chaque nouvelle ligne considérée de nouvelles inconnues n'intervenant pas dans les lignes suivantes. Si on a déjà obtenu des valeurs pour les inconnues des dernières lignes, donner des valeurs quelconques à toutes les nouvelles inconnues sauf une détermine alors la dernière inconnue.
On peut donc remonter le système en donnant des valeurs quelconques à toutes les nouvelles inconnues de chaque ligne, sauf une. L'inconnue qui jouera ce rôle particulier est l'une quelconque des nouvelles inconnues (en remontant), mais un choix s'impose naturellement : celui de la première inconnue intervenant de façon effective dans la ligne. Autrement dit, l'inconnue se plaçant à la position du pivot utilisé sur cette ligne pour obtenir l'échelonnement.

Définition 20.4.6 (Inconnue principale)
Nous appelons *inconnue principale* du système échelonné $AX = B$ une inconnue x_i se plaçant en tête d'une des lignes du système. Il s'agit donc des inconnues d'indice j_i de la définition 20.3.6

Les autres inconnues sont appelées secondaires.

Remarque 20.4.7
- La définition d'inconnue principale donnée ici est fortement dépendante de l'ordre des variables. Une permutation des variables préservant l'échelonnement (c'est en général possible) donnerait un autre système d'inconnues principales.
- Il existe une notion plus générale d'inconnues principales, se définissant bien à l'aide de déterminants, portant sur des systèmes non nécessairement échelonnés. En gros, il s'agit d'un système d'inconnues tel que le choix quelconque des autres inconnues déterminent de façon unique les inconnues principales. Mais pour un système donné, le choix des inconnues principales n'est pas unique.
- Dans ce cadre plus général, notre définition des inconnues principales n'est qu'un des plusieurs choix possibles pour le système échelonné donné.

IV.3 Résolution d'un système échelonné réduit

Méthode 20.4.8 (recherche d'une solution particulière d'un système échelonné)

Soit A une matrice échelonnée réduite, et $AX = B$ un système associé.

1. S'il existe dans ce système une ligne du type $0 = b_i$, avec b_i non nul, alors le système n'admet pas de solution. On dit que $AX = B$ est un système non compatible.
2. Sinon, donner des valeurs quelconques aux variables non principales détermine entièrement les autres.
3. On trouve par exemple une solution particulière en attribuant à toutes les inconnues secondaires la valeur 0. La i-ième inconnue principale prendra alors la valeur b_i.

Remarque 20.4.9

C'est facilement implémentable.

Exemple 20.4.10

Recherche d'une solution particulière du système $AX = B$ lorsque

$$A = \begin{pmatrix} 1 & -4 & -2 & 3 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Méthode 20.4.11 (Résolution du système homogène associé)

- La matrice A étant échlonnée réduite, le système $AX = 0$ permet d'exprimer chacune des inconnues principales en fonction des inconnues secondaires, en les isolant sur chaque ligne
- On a donc l'ensemble des solutions, paramétré par les inconnues secondaires, choisies quelconques.
- Chaque coordonnée du vecteur X est une fonction linéaire en les inconnues secondaires. Ainsi, en isolant les contributions de chaque paramètre, on peut voir l'ensemble des solutions comme l'ensemble des combinaisons linéaires d'un certain nombre de vecteurs bien déterminés.

Exemple 20.4.12

Terminer la résolution du système de l'exemple 20.4.10.

Proposition 20.4.13 (Caractérisation des systèmes de Cramer)

Un système de n équations à n inconnues est de Cramer si et seulement s'il admet une et une seule solution.

◁ **Éléments de preuve.**

S'il n'est pas de Cramer, soit il est incompatible, soit la matrice échelonnée réduite associée possède strictement moins que n pivots. Une inconnue secondaire pourra être choisie arbitrairement . La réciproque est déjà vue. ▷

IV.4 Retour sur le calcul de l'inverse

Méthode 20.4.14 (Calcul de l'inverse par résolution d'un système)

Soit A une matrice carrée

- Pour un second membre Y arbitraire (i.e. indéterminé), résoudre le système $AX = Y$, et, si le résultat est unique, l'écrire sous la forme $X = BY$.
- Alors $B = A^{-1}$

La justification de l'inversibilité provient de l'unicité de la solution trouvée, et de la caractérisation précédente des systèmes de Cramer. Puisque Y est arbitraire, on a alors

$$\forall Y \in \mathcal{M}_{n,1}(\mathbb{K}), \quad BY = A^{-1}Y.$$

Il est facile de s'assurer alors que $B = A^{-1}$, par exemple par des choix judicieux de Y, permettant d'identifier les matrices colonne par colonne.

Remarque 20.4.15

- Si on résout ce système par la méthode du pivot, c'est exactement la même méthode que la précédente par juxtaposition avec I_n. Mais la présentation est moins commode, les coefficients étant plus mélangés (les coefficients de la j-ième colonne de la matrice de droite correspondent aux coefficients devant la j-ième coordonnée de Y). Il vaut mieux privilégier la présentation matricielle précédente, plus claire dans les alignements.
- En revanche, cette méthode peut être intéressante si on peut s'éloigner un peu de la méthode du pivot, par exemple pour exploiter certaines symétries de la matrice A, qui parfois permettent une résolution plus efficace que par la méthode du pivot.

V Produit matriciel par blocs

Nous voyons dans cette dernière section qu'on peut effectuer un produit matriciel en groupant les termes par blocs rectangulaires, en utilisant les règles usuelles, c'est-à-dire en remplaçant dans les formules les coefficients par des blocs matriciels. Il faut cependant prendre garde au fait que cela n'est possible que si le découpage en blocs des deux matrices A et B fournit des formats compatibles pour les produits intervenant de la sorte.

Théorème 20.5.1 (Produit par blocs)

Soit $A \in \mathcal{M}_{n,p}$ et $B \in \mathcal{M}_{p,m}$ deux matrices, et

$$0 = i_0 < i_1 < i_2 < \cdots < i_{q-1} < i_q = n,$$
$$0 = j_0 < j_1 < j_2 < \cdots < j_{r-1} < j_r = p,$$
$$0 = k_0 < k_1 < k_2 < \cdots < k_{s-1} < k_s = m.$$

Les deux premières suites définissent un découpage par blocs $A = (A_{i,j})_{(i,j)\in[\![1,q]\!]\times[\![1,r]\!]}$ de A et les deux dernières définissent un découpage par blocs $B = (B_{j,k})_{(j,k)\in[\![1,r]\!]\times[\![1,s]\!]}$.
Le produit AB admet alors une représentation par blocs :

$$AB = (C_{i,k})_{(i,k)\in[\![1,q]\!]\times[\![1,s]\!]},$$

où pour tout $(i,k) \in [\![1,q]\!] \times [\![1,s]\!]$, $C_{i,k} = \sum_{j=1}^{r} A_{i,j}B_{j,k}$.

◁ **Éléments de preuve.**

Regarder coefficient par coefficient. C'est un peu technique et désagréable, mais pas difficile en soi.

▷

En particulier, si on a une représentation diagonale par blocs (avec les A_k carrées) :

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k \end{pmatrix}, \qquad \text{alors} \qquad A^n = \begin{pmatrix} A_1^n & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_k^n \end{pmatrix}.$$

Exemple 20.5.2
Calcul des puissances successives de $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

21

Arithmétique des entiers

La mathématique est la reine des sciences, mais la théorie des nombres est la reine des sciences mathématiques.

(Carl-Friedrich Gauss)

Les nombres sont le plus haut degré de la connaissance. Le nombre est la connaissance même.

(Platon)

Les nombres sont l'essence des choses.

(Pythagore)

Décomposer un cube en deux autres cubes, une quatrième puissance, et généralement une puissance quelconque en deux puissances de même nom, au dessus de la seconde puissance, est chose impossible et j'en ai assurément trouvé l'admirable démonstration. La marge trop exiguë ne la contiendrait pas.

(Pierre de Fermat)

Note Historique 21.0.1

- L'arithmétique désigne dans un premier temps l'étude des opérations élémentaires entre entiers (arithmétique élémentaire), et les algorithmes permettant de faire ces opérations (algorithmes de multiplication, division euclidienne...). C'est une des disciplines fondamentales des mathématiques dans le sens où, avec la géométrie et le calcul numérique (algébrique), elle constitue le point de départ de toutes les mathématiques.
- En découle de façon naturelle (et déjà en Grèce antique) l'étude des propriétés de divisibilité, et donc de primalité.
- Plus généralement, l'arithmétique désigne l'étude de problèmes relatifs à des nombres entiers. Ainsi, les problèmes de Diophante, liés à la recherche de solutions entières d'équations relèvent de l'arithmétique. De telles équations sont d'ailleurs appelées *équations diophantiennes*. L'exemple le plus célèbre en est certainement le fameux théorème de Fermat-Wiles stipulant que pour tout $n \geqslant 3$, il n'existe pas de triplet (a, b, c) d'entiers naturels non nuls tels que $a^n + b^n = c^n$. Il est instructif d'ailleurs de noter que l'ouvrage dans lequel Pierre de Fermat a écrit en marge sa fameuse note (citation en début de chapitre) concernant une preuve de ce résultat n'est autre que *Les arithmétiques* de Diophante.
- L'exemple du théorème de Fermat-Wiles justifie la nécessité de sortir du cadre des entiers pour résoudre des problèmes arithmétiques en apparence simples. Ainsi, l'arithmétique a évolué en diverses branches (théorie algébrique des nombres, théorie analytique des nombres, géométrie algébrique...)

- Les concepts essentiels de l'arithmétique ont également été généralisés dans des contextes différents de celui des entiers. C'est une des motivations de l'introduction de la notion d'anneau et d'idéal. Un exemple que vous aurez l'occasion d'étudier prochainement est l'étude de l'arithmétique des polynômes. Mais cela ne s'arrête pas là !

I Divisibilité, nombres premiers

I.1 Notion de divisibilité

Définition 21.1.1 (Divisibilité, diviseur, multiple)

- Soit a et b deux entiers relatifs, $b \neq 0$. On dit que b *divise* a, et on écrit $b \mid a$, si et seulement s'il existe $q \in \mathbb{Z}$ tel que $a = bq$.
- On dit dans ce cas que b est un *diviseur* de a, et que a est un *multiple* de b.

On note $a \mid b$ pour dire que a divise b.
Ainsi, $2 \mid 4$, $-2 \mid 4$, $2 \mid -4$, et $-2 \mid -4$.

Proposition 21.1.2 (Caractérisation de la divisibilité en termes d'idéaux)

Soit a et b deux entiers positifs, $a \neq 0$. Alors $a \mid b$ si et seulement si $b\mathbb{Z} \subset a\mathbb{Z}$.

Définition 21.1.3 (couple d'entiers associés)

On dit que deux entiers a et b sont associés si et seulement si $a \mid b$ et $b \mid a$, c'est-à-dire $a\mathbb{Z} = b\mathbb{Z}$.

Proposition 21.1.4 (Caractérisation des entiers associés)

Les entiers a et b sont associés si et seulement si il existe $\varepsilon \in \{-1, 1\} = \mathbb{Z}^\times = U(\mathbb{Z})$ tel que $a = \varepsilon b$.

◁ **Éléments de preuve.**
Essayer de le démontrer (sous la forme $U(\mathbb{Z})$) en n'utilisant que l'intégrité de $\mathbb{Z}$ (et donc la régularité multiplicative). Cela permet de généraliser à d'autres anneaux. ▷

Remarque 21.1.5

- Ce résultat peut sembler trivial et sans intérêt. Sa version plus générale, pour un anneau intègre A, est plus intéressante, et affirme que les éléments associés diffèrent d'une constante multiplicative appartenant au groupe A^* des inversibles de A.
- Par exemple, dans $\mathbb{K}[X]$, les éléments associés à un polynôme P sont tous les λP, pour $\lambda \in \mathbb{K}^*$.
- Dans $\mathbb{Z}[\mathrm{i}]$ (entiers de Gauss), les éléments associés à un nombre z sont les 4 éléments z, $-z$, $\mathrm{i}\,z$ et $-\mathrm{i}\,z$.
- Les éléments associés de x sont les éléments qui jouissent des mêmes propriétés de divisibilité que x.

Théorème/Définition 21.1.6 (Théorème de la division euclidienne)

Soit $(a, b) \in \mathbb{Z}^2$, $b \neq 0$.

- *Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que :*

$$a = bq + r \qquad \text{et} \qquad 0 \leqslant r < |b|.$$

- *L'entier q est appelé quotient de la division euclidienne de a par b.*
- *L'entier r est appelé reste de la division euclidienne de a par b.*

◁ **Éléments de preuve.**

On pourrait se raccrocher à la division euclidienne réelle, mais c'est maladroit dans le sens où les entiers existent avant les réels et indépendemment d'eux. Pour l'existence, pour a positif, on peut faire une récurrence forte, en initialisant pour tout $a \in [\![0, |b| - 1]\!]$. Une autre récurrence dans l'autre sens pour compléter à $a < 0$ (ou par symétrisation). L'unicité se montre bien à peu près de la même façon que dans le cas réel. ▷

Remarquez que b peut être négatif.
La plupart des propriétés arithmétiques de $\mathbb{Z}$ (pour ne pas dire toutes) découlent de l'existence de cette division euclidienne. On peut définir de façon similaire dans certains anneaux une division euclidienne, la condition sur le reste étant un peu plus dure à exprimer. On parle dans ce cas d'anneau euclidien.

Définition 21.1.7 (Anneau euclidien, HP)

Soit A un anneau. On dit que A est euclidien s'il est intègre, et muni d'un stathme, c'est-à-dire d'une application $v : A \setminus \{0\} \to \mathbb{N}$ telle que :

$$\forall a \in A, \quad \forall b \in A \setminus \{0\}, \quad \exists (q, r)^2 \in A, \quad a = bq + r \qquad \text{et} \qquad (r = 0 \text{ ou } v(r) < v(b))$$

Exemple 21.1.8

- Quel est le stathme pour la division euclidienne dans $\mathbb{Z}$?
- Quel est le stathme pour la division euclidienne dans $\mathbb{C}[X]$?
- On peut montrer que $\mathbb{Z}[\mathrm{i}]$ est euclidien, de stathme $z \mapsto |z|^2$.

Ainsi, $\mathbb{Z}$ et $\mathbb{R}[X]$ sont des anneaux euclidiens. Cette dernière propriété nous permettra d'établir un certain nombre de propriétés arithmétiques pour les polynômes, très similaires à celles qu'on a pour les entiers.

Remarques 21.1.9

- Certains auteurs appellent préstathme la notion de stathme telle que nous l'avons définie. Ils imposent une condition supplémentaire pour les stathmes. La différence n'est pas trop gênante dans la mesure où on peut montrer qu'avec leur terminologie, tout anneau intègre muni d'un préstathme peut aussi être muni d'un stathme.
- Dans la notion générale de division euclidienne définie par stathme, on n'impose pas de propriété d'unicité. Par exemple, dans $\mathbb{Z}[\mathrm{i}]$, on n'a pas de propriété d'unicité. D'ailleurs, dans $\mathbb{Z}$, la définition générale de division euclidienne nous donne deux divisions euclidiennes possibles, la division euclidienne usuelle n'est que l'une des deux divisions possibles (ou on impose en plus la positivité du reste).

Note Historique 21.1.10

La division euclidienne est appelée ainsi par référence à Euclide qui décrit dans ses éléments le procédé algorithmique de soustractions répétées permettant d'obtenir le quotient et le reste. Cependant, on trouve trace de cette notion à des époques antérieures, notamment en Égypte.
C'est Gauss le premier, avec l'étude de $\mathbb{Z}[\mathrm{i}]$, qui remarque que de nombreuses propriétés arithmétiques ne sont pas spécifiques à $\mathbb{Z}$ et découlent de façon plus générale de l'existence d'une division euclidienne dans un anneau. Cette remarque est évidemment à la base de la notion d'anneau euclidien.

I.2 Congruences

Nous rappelons la définition suivante, indissociable de la notion de division euclidienne :

Définition 21.1.11 (Congruences d'entiers)

Soit $n \in \mathbb{N}^*$, et $(a, b) \in \mathbb{Z}^2$. On dit que a et b sont *congrus* modulo n, et on écrit $a \equiv b\ [n]$, si et seulement si n divise $b - a$, ou encore si les divisions euclidiennes de a et b par n ont même reste.

On trouve aussi assez souvent la notation $a \equiv b \mod n$, ou un mélange des 2 : $a \equiv b\ [\text{mod } n]$.
Nous rappelons les résultats suivants, que nous avons déjà eu l'occasion de démontrer.

Théorème 21.1.12

La relation de congruence modulo n est une relation d'équivalence.

Théorème 21.1.13

La relation de congruence modulo n est compatible avec le produit et la somme : soit $(a, a', b, b') \in \mathbb{Z}^4$ tels que $a \equiv a'\ [n]$ et $b \equiv b'\ [n]$. Alors $a + b \equiv a' + b'\ [n]$ et $ab \equiv a'b'\ [n]$

En d'autre terme, c'est une congruence sur les monoïdes $(\mathbb{Z}, +)$ et $(\mathbb{Z}, \times)$, au sens vu dans le chapitre sur les ensembles.
Ces règles sont importantes pour pouvoir mener à bien le calcul modulaire de façon efficace : il permet de faire lors d'une succession d'opérations, des réductions modulo n étape par étape, plutôt que de tout calculer dans $\mathbb{N}$ et de réduire à la fin.

Exemples 21.1.14

Calculer le reste de la division euclidienne de $12 \times 21 \times 28 \times 18 \times 75 \times 23$ par 11.

Cette possibilité de réduire les opérations à chaque étape est également important pour l'implémentation informatique du calcul modulaire, permettant ainsi de travailler avec des entiers plus petit, diminuant de la sorte la complexité des calculs. On peut ainsi, contrairement au cas du calcul dans $\mathbb{Z}$, borner explicitement le temps de calcul des opérations modulo n par un réel dépendant de n mais ne dépendant pas des opérandes.
Nous avons aussi, de façon immédiate :

Proposition 21.1.15

Si n divise m alors pour tout a et b dans $\mathbb{Z}$:

$$a \equiv b\ [m] \Longrightarrow a \equiv b\ [n].$$

Enfin, le résultat suivant est souvent bien utile :

Proposition 21.1.16 (Périodicité des puissances modulo n)

Soit $n > 1$, et $a \in \mathbb{Z}$. Alors la suite $(a^p)_{p \in \mathbb{N}^*}$ est périodique modulo n à partir d'un certain rang. On trouve une période dès lors qu'on trouve deux valeurs distinctes telles que $a^{p_1} \equiv a^{p_2}\ [n]$.

◁ **Éléments de preuve.**

C'est un principe des tiroirs : il y a plus d'exposants distincts que de classes de congruences possibles !

▷

Remarque 21.1.17

À quelle condition nécessaire et suffisante sur a la suite (a^p) est-elle périodique modulo n dès le rang $n = 0$?

Ainsi, le calcul des premières puissances jusqu'à obtenir un résultat déjà obtenu auparavant permet de trouver la période, puis d'en déduire toutes les autres puissances. On peut donc de cette manière calculer a^p modulo n :

Méthode 21.1.18 (Calcul de a^p modulo n)

- On commence par calculer les puissances successives de a modulo n, jusqu'à obtenir deux valeurs $p_1 < p_2$ telles que a^{p_1} et a^{p_2} soient congrus modulo n. On peut remarquer que si $a \wedge n = 1$, une période peut être obtenue par le théorème de Fermat ou d'Euler, mais qu'il ne s'agit pas forcément de la période minimale.
- Si a est inversible modulo n, on peut prendre $p_1 = 0$ (quel résultat permet d'en être sûr ?), la suite est alors périodique dès le rang initial. Mais ce n'est pas toujours le cas lorsque a n'est pas inversible modulo n.
- Soit $T = p_2 - p_1$ la longueur d'une période, et $p \geqslant p_1$. On réduit alors l'exposant modulo T pour trouver l'unique représentant q de p dans $[\![p_1, p_2 - 1]\!]$.
- On a alors $a^p \equiv a^q \mod [n]$.

Exemple 21.1.19

Calculer le reste de la division euclidienne de 1685^{1750} par 42.

I.3 Nombres premiers

Nous les avons déjà rencontrés, évidemment. Nous rappelons :

Définition 21.1.20 (Nombres premiers)

Soit $p \in \mathbb{N}^*$. On dit que p est un nombre premier si p admet exactement 2 diviseurs positifs distincts (à savoir 1 et p lui-même)

Remarquez que l'existence de deux diviseurs distincts exclut d'office 1 de l'ensemble des nombres premiers, puisqu'il n'a qu'un diviseur.

Définition 21.1.21 (Nombres composés)

Soit $n \in \mathbb{N}^*$. On dit que n est un nombre composé si n possède au moins 3 diviseurs positifs distincts, ou en d'autres termes, si n possède un diviseur positif distinct de 1 et de n.

Proposition 21.1.22

Tout nombre composé admet un diviseur strict premier.

◁ **Éléments de preuve.**

Récurrence forte. ▷

Cette proposition est à la base de l'existence de la décomposition primaire.

Théorème 21.1.23 (Combien de nombres premiers ? Euclide)
Il y a une infinité de nombres premiers.

◁ **Éléments de preuve.**
De très (très très) nombreuses démonstrations de ce fait existent. La démonstration d'Euclide est liée à l'étude des diviseurs de $p_1 \cdots p_n$ en supposant par l'absurde que $p_1, \ldots, p_n$ sont tous les nombres premiers. ▷

C'est bien joli tout ça, mais comment faire pour déterminer les nombres premiers (pas trop gros) ? Erathostène, mathématicien, astronome, bibliothécaire en chef d'Alexandrie (excusez du peu), astéroïde et cratère lunaire, répondit à cette question il y a déjà très longtemps, par un procédé d'élimination.

Méthode 21.1.24 (Crible d'Érathostène)
Pour trouver tous les nombres premiers inférieurs ou égaux à n :

1. Écrire tous les nombres entiers de 2 à n.
2. Le plus petit d'eux, à savoir 2, est premier (il n'a pas de diviseur strictement plus petit que lui, autre que 1)
3. Les multiples stricts de 2 ne sont pas premiers, on les barre tous.
4. Parmi les nombres restants (en excluant les nombres premiers précédents, à savoir 2 dans la première étape, et en excluant les nombres barrés), le plus petit est premier (il n'est divisible par aucun nombre premier strictement plus petit que lui et différent de 1, sinon il serait barré). On barre tous ses multiples stricts qui ne peuvent pas être premiers, et on recommence cette étape jusqu'à épuisement de tous les entiers de la liste.

Cet algorithme est très facile à implémenter dans un langage informatique. Il n'est évidemment efficace que pour des petites valeurs de n, mais ne peut pas servir à la recherche de très grands nombres premiers. Notamment, il est à peu près inutilisable pour répondre à la question de savoir si un très grand nombre donné est premier ou non (question cruciale dans certaines situations en rapport avec des cryptages, comme la méthode RSA).

II PGCD et PPCM

II.1 PGCD et PPCM d'un couple d'entiers

Lemme 21.2.1 (Somme de deux groupes abéliens)
Soit H et K deux sous groupes d'un groupe abélien $(G, +)$. Alors $H+K$ est le plus petit groupe contenant $H \cup K$.

◁ **Éléments de preuve.**
Il faut bien comprendre que $H + K$ désigne l'ensemble de toutes les sommes d'un élément de H et d'un élément de K. Justifier que $H + K \subset < H \cup K >$ et que c'est un groupe. ▷

On vérifie facilement la propriété suivante :

Lemme 21.2.2 (intersection et somme de deux idéaux)
Soit I et J deux idéaux d'un anneau commutatif A. Alors $I \cap J$ et $I + J$ sont des idéaux de A.

Proposition/Définition 21.2.3 (PGCD)

Soit a et b deux entiers positifs tels que l'un au moins des entiers a et b est non nul, et $m \in \mathbb{N}^*$. Les propositions suivantes sont équivalentes :

(i) l'entier m est le maximum (pour l'ordre usuel) de $\{d \in \mathbb{N}^* \mid d \text{ divise } a \text{ et } d \text{ divise } b\}$

(ii) l'entier m est le maximum (pour l'ordre de divisibilité) de $\{d \in \mathbb{N}^* \mid d \text{ divise } a \text{ et } d \text{ divise } b\}$.

(iii) $m = \inf_{(\mathbb{N}^*,|)}(a, b)$

(iv) $a\mathbb{Z} + b\mathbb{Z} = m\mathbb{Z}$

Si l'une de ces quatre conditions équivalentes est satisfaite, on dit que m est le *plus grand commun diviseur* de a et b (en abrégé : PGCD), et on le note $a \wedge b$.

◁ **Éléments de preuve.**
Par définition $(ii) \iff (iii)$, et $(ii) \implies (i)$ est facile, ainsi que $(iv) \implies (ii)$. Le point délicat est $(i) \implies (iv)$ pour compléter la boucle. Se rappeler que $\mathbb{Z}$ est principal, ce qui permet d'exprimer $a\mathbb{Z} + b\mathbb{Z}$ sous la forme $n\mathbb{Z}$. Comparer n et m. ▷

Ainsi, le PGCD de a et b est entièrement caractérisé par l'égalité des idéaux (en notant (a) l'idéal engendré par a) :

$$(a) + (b) = (a \wedge b).$$

Remarquez que pour établir ce point partant de la description usuelle (premier point), on se sert du fait que tout idéal de $\mathbb{Z}$ s'écrit $\mathbb{Z}\cdot a$, donc que $\mathbb{Z}$ est principal. Le fait que $\mathbb{Z}$ est principal nous assure également que le plus petit idéal contenant a et b est engendré par un élément. C'est là une façon de définir le pgcd, comme élément générateur de l'idéal engendré par a et b.
Cette définition est valide dans tout anneau principal :

Définition 21.2.4 (PGCD dans un anneau principal, HP)

Soit A un anneau principal et a et b deux éléments de A. Un PGCD d de a et b est un élément défini de façon unique à inversibles près par :

$$(d) = (a) + (b).$$

Dans certains cas, un choix de pgcd s'impose (le pgcd positif dans $\mathbb{Z}$ par exemple). Dans ce cas on peut utiliser une notation non ambiguë ($a \wedge b$ par exemple), et parler du PGCD. Dans les autres cas, on parle d'UN PGCD, et on ne peut utiliser une notation qu'à abus près.
Le PGCD se détermine très facilement algorithmiquement, en se basant sur les lemmes suivants :

Lemme 21.2.5 (Simplification d'un pgcd)

Soit $k \in \mathbb{Z}$. Alors $a \wedge (b + ka) = a \wedge b$.

Lemme 21.2.6 (Étape de l'algorithme d'Euclide)

Soit a et b deux entiers, et r le reste de la division euclidienne de a par b. Alors

$$a \wedge b = b \wedge r.$$

Ainsi, en prenant des restes divisions euclidiennes successives le dernier reste non nul fournira le PGCD. La preuve de la terminaison de l'algorithme, faite en cours d'informatique, repose sur le variant de boucle b, entier positif strictement décroissant, et sa correction provient de l'invariant de boucle $a \wedge b$ (son invariance provenant du lemme).

Algorithme 21.1 : Algorithme d'Euclide pour le calcul du PGCD

Entrée : a, b : entiers naturels
Sortie : $a \wedge b$
tant que $b > 0$ **faire**
  $a,\ b \leftarrow b,\ a \,\%\, b$
fin tant que
renvoyer a

Le PPCM se définit par des propriétés équivalentes similaires, symétriques de celles définissant le PGCD

Proposition/Définition 21.2.7 (PPCM)

Soit a et b deux entiers non nuls, et $M \in \mathbb{N}^*$. Les propositions suivantes sont équivalentes :

(i) l'entier M est le minimum (pour l'ordre usuel) de $\{m \in \mathbb{N}^* \mid a \text{ divise } m \text{ et } b \text{ divise } m\}$

(ii) l'entier M est le minimum (pour l'ordre de divisibilité) de $\{m \in \mathbb{N}^* \mid a \text{ divise } m \text{ et } b \text{ divise } m\}$

(iii) $M = \sup_{(\mathbb{N}^*,|)}(a, b)$

(iv) $M\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$.

Si l'une de ces quatre propositions équivalentes est satisfaite, on dit que M est le *plus petit commun multiple* de a et b (PPCM en abrégé), et on note $M = a \vee b$.

◁ **Éléments de preuve.**
Même principe que pour le PGCD. ▷

Encore une fois, le dernier point peut être pris comme définition dans un anneau principal.

Définition 21.2.8 (PPCM dans un anneau principal, HP)

Soit A un anneau principal et a et b deux éléments non nuls de A. Alors un PPCM de a et b est un élément m tel que

$$(m) = (a) \cap (b).$$

Ce qui est parfois évident sur les idéaux ne l'est pas toujours autant pour les autres descriptions :

Proposition 21.2.9 (Distributivité du produit sur $\wedge$ et $\vee$)

Soit a et b deux entiers naturels, et c un entier naturel non nul.

1. *Si a et b ne sont pas tous les deux nuls, $(a \wedge b) \cdot c = (ac) \wedge (bc)$.*
2. *Si a et b sont non nuls, $(a \vee b) \cdot c = (ac) \vee (bc)$.*

◁ **Éléments de preuve.**
On peut le démontrer avec les propriétés de minimalité/maximalité. Mais c'est plus limpide par des manipulations sur les idéaux. ▷

II.2 Identité de Bézout

L'identité de Bézout est elle aussi une conséquence immédiate de la caractérisation par les idéaux :

Théorème 21.2.10 (identité de Bézout, ou théorème de Bachet-Bézout)

1. *Soit a et b deux entiers dont l'un au moins est non nul. Alors il existe des entiers relatifs x et y tels que $ax + by = a \wedge b$.*

2. *Réciproquement, étant donné un entier $d \in \mathbb{N}^*$, s'il existe des entiers relatifs x et y tels que*

$$d = ax + by,$$

alors $a \wedge b \mid d$.

◁ **Éléments de preuve.**
C'est juste une réexpression du point (iv) de la définition. ▷

Note Historique 21.2.11

- C'est le nom d'Étienne Bézout, mathématicien français du 18e siècle, qui est le plus souvent associé à ce résultat. C'est pourtant à Claude-Gaspard Bachet de Méziriac que l'on doit la première preuve, parue dans son ouvrage *Problèmes plaisans et délectables qui se font par les nombres*, paru en 1624. Sa preuve est celle que nous présentons ci-dessous (par l'algorithme d'Euclide)
- Qu'a fait Bézout alors pour avoir droit à tous ces honneurs ? Il a généralisé le résultat à d'autres situations, notamment au cas des polynômes.
- Il est intéressant de noter que le fameux ouvrage dans lequel Fermat écrivit dans une marge qu'il savait démontrer ce qu'on appelle aujourd'hui le théorème de Fermat-Wiles est en fait une traduction par Bachet de Mériziac de l'*Arithmétique* de Diophante. Le monde est petit...

La démonstration passant par les idéaux peut se généraliser dans un anneau principal. Elle possède l'inconvénient de ne pas être constructive. Il peut être intéressant de trouver explicitement des entiers x et y assurant l'égalité $ax + by = a \wedge b$. L'algorithme de la division euclidienne itéré permet à la fois de déterminer $a \wedge b$, et d'obtenir une identité de Bézout.

Méthode 21.2.12 (Déterminer une relation de Bézout)
C'est un complément apporté par l'algorithme d'Euclide. Écrire successivement des combinaisons de a et b égales aux restes successifs utilisés dans l'algorithme. Pour cela, pour passer d'une identité à la suivante, combiner la précédente avec la relation de division euclidienne.

Cette méthode est valide dans tout anneau euclidien.
Ainsi, en écrivant $r_0 = a$, $r_1 = b$ puis les divisions euclidiennes successives :

$$\begin{cases} r_0 & = r_1 q_2 + r_2 \\ r_1 & = r_2 q_3 + r_3 \\ \vdots & \\ r_{k-2} & = r_{k-1} q_k + r_k \\ r_{k-1} & = r_k q_{k+1} + r_{k+1}, \end{cases}$$

avec $r_2 \neq 0$, $r_3 \neq 0, \ldots, r_k \neq 0$ et $r_{k+1} = 0$, on a $r_k = a \wedge b$. De plus, en posant $x_0 = 1$, $x_1 = 0$, $y_0 = 0$, $y_1 = 1$ et pour tout $i \leqslant [\![3, k]\!]$

$$x_i = x_{i-2} - q_i x_{i-1} \qquad \text{et} \qquad y_i = y_{i-2} - q_i y_{i-1},$$

on obtient pour tout $i \in [\![1, n]\!]$,

$$r_i = a x_i + b y_i,$$

donc en particulier pour $i = k$, on obtient une identité de Bézout :

$$a \wedge b = a x_k + b y_k.$$

On peut donc décrire de façon plus algorithmique :

Algorithme 21.2 : Algorithme d'Euclide étendu

Entrée : a, b : entiers naturels non nuls
Sortie : m, u, v tels que $m = a \wedge b = ua + bv$
$u, v, w, x, r, s \leftarrow 1, 0, 0, 1, a, b$;
tant que $s \neq 0$ **faire**
  $q,\ s,\ r \leftarrow r$ // $s,\ r$ % $s,\ s$;
  $w,\ u \leftarrow u - qw,\ w$;
  $x,\ v \leftarrow v - qx,\ x$
fin tant que
renvoyer *(r,u,v)*

En pratique, pour ne pas s'embrouiller, il vaut mieux vaut écrire les différentes relations obtenues par la division euclidienne, en remplaçant étape par étape les restes obtenus par leur expression obtenue récursivement en fonction de a et b.

Exemple 21.2.13

- Trouver à l'aide de l'algorithme d'Euclide le pgcd de 27 et 33, ainsi qu'une identité de Bézout.
- Comment trouver une autre identité de Bézout ?
- À retenir : on n'a pas unicité de la relation de Bézout !
- Comment trouver toutes les relations de Bézout ?

II.3 PGCD et PPCM d'une famille finie d'entiers

La notion de PGCD et de PPCM de deux entiers peut être généralisée à un plus grand nombre d'entiers :

Proposition/Définition 21.2.14 (PGCD d'un nombre fini d'entiers)

Soit $a_1, \ldots, a_n$ des entiers naturels, non tous nuls, et m un entier naturel. Les propriétés suivantes sont équivalentes :

(i) m est le maximum (au sens de l'ordre usuel) des entiers d qui divisent chacun des a_i, $i \in [\![1, n]\!]$.

(ii) m est le maximum (au sens de la divisibilité) des entiers d qui divisent chacun des a_i, $i \in [\![1, n]\!]$.

(iii) $m = \inf_{(\mathbb{N}^*, |)}(a_1, \ldots, a_n)$

(iv) $m\mathbb{Z} = a_1\mathbb{Z} + a_2\mathbb{Z} + \cdots + a_n\mathbb{Z}$.

Si l'une de ces quatre propositions équivalentes est satisfaite, on dit que m est le PGCD de la famille $(a_1, \ldots, a_n)$ et on note $m = a_1 \wedge a_2 \wedge \cdots \wedge a_n$.

De la même façon :

Proposition/Définition 21.2.15 (PPCM d'un nombre fini d'entiers)

Soit $a_1, \ldots, a_n$ des entiers naturels, non nuls, et m un entier naturel. Les propriétés suivantes sont équivalents :

(i) m est le minimum (au sens de l'ordre usuel) des entiers m multiples de chacun des a_i, $i \in [\![1, n]\!]$.

(ii) m est le minimum (au sens de la divisibilité) des entiers m multiples de chacun des a_i, $i \in [\![1, n]\!]$.

(iii) $m = \sup_{(\mathbb{N}^*, |)}(a_1, \ldots, a_n)$

(iv) $m\mathbb{Z} = a_1\mathbb{Z} \cap a_2\mathbb{Z} \cap \cdots \cap a_n\mathbb{Z}$.

Si l'une de ces quatre propositions équivalentes est satisfaite, on dit que m est le PPCM de la famille $(a_1, \ldots, a_n)$ et on note $m = a_1 \vee a_2 \vee \cdots \vee a_n$.

◁ **Éléments de preuve.**

Par très différent du cas $n = 2$. ▷

La caractérisation par idéaux, ou encore la caractérisation par borne inférieure (et l'associativité des bornes inférieures) nous assure que ces notions correspondent aux PGCD et PPCM itérés :

Proposition 21.2.16 (Associativité du PGCD et du PPCM)

Soit $a_1, \dots, a_n$ des entiers strictement positifs non tous nuls. Alors

$$a_1 \wedge \cdots \wedge a_n = ((a_1 \wedge a_2) \wedge \dots) \wedge a_n.$$

S'ils sont tous non nuls, on a également

$$a_1 \vee \cdots \vee a_n = ((a_1 \vee a_2) \vee \dots) \vee a_n.$$

En particulier, pour $n = 3$, on en déduit :

$$a_1 \wedge (a_2 \wedge a_3) = (a_1 \wedge a_2) \wedge a_3 \qquad \text{et} \qquad a_1 \vee (a_2 \vee a_3) = (a_1 \vee a_2) \vee a_3.$$

On peut étendre le théorème de Bachet-Bézout à cette situation, toujours en utilisant la caractérisation par les idéaux :

Théorème 21.2.17 (Relation de Bézout)

Soit $a_1, \dots, a_n$ des entiers naturels non tous nuls. Alors il existe des entiers relatifs $x_1, \dots, x_n$ tels que

$$a_1 \wedge \cdots \wedge a_n = x_1 a_1 + \cdots + x_n a_n.$$

Réciproquement, s'il existe des entiers $x_1, \dots, x_n$ tels que

$$d = x_1 a_1 + \cdots + x_n a_n,$$

alors d est un multiple de $a_1 \wedge \cdots \wedge a_n$.

Méthode 21.2.18

Les coefficients $x_1, \dots, x_n$ peuvent se trouver explicitement, par itération de l'algorithme d'Euclide : on cherche d'abord une relation de Bézout entre $d_1 = a_1 \wedge a_2$, a_1 et a_2, puis entre $d_2 = d_1 \wedge a_3$, d_1 et a_2 ; en substituant à d_1 la première relation trouvée, on obtient une relation de Bézout entre $a_1 \wedge a_2 \wedge a_3$, a_1, a_2 et a_3. On continue alors de la sorte, de proche en proche.

Enfin, toutes les notions introduites dans ce paragraphe peuvent être généralisées à des entiers relatifs quelconques ; le pgcd et le ppcm ne sont alors définis correctement qu'au signe près (c'est le cas général dans un anneau principal, ou le pgcd ne peut être déterminé qu'à un facteur multiplicatif inversible près). Dans le cas de $\mathbb{Z}$, on a un choix privilégié qui consiste à prendre la valeur positive. Le pgcd et les relations de Bézout se trouvent de la même façon, en les cherchant d'abord pour les valeurs absolues, puis en modifiant les signes de façon adéquate.

III Entiers premiers entre eux

III.1 Couple d'entiers premiers entre eux

Définition 21.3.1 (Entiers premiers entre eux)

Soit a et b deux entiers naturels non tous les deux nuls. On dit que a et b sont premiers entre eux si et seulement si $a \wedge b = 1$, donc si a et b n'ont pas d'autre diviseur positif commun que 1.

Cela peut aussi s'exprimer par la relation $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$.

Note Historique 21.3.2

La première apparition de cette notion est dans le Livre VII des *Éléments* d'Euclide.

Proposition 21.3.3 (Caractérisation par les diviseurs premiers)

Deux entiers naturels a et b non tous les deux nuls sont premiers entre eux si et seulement s'ils n'ont aucun diviseur premier en commun.

Proposition 21.3.4 (Simplification des fractions)

- *Soit a et b deux entiers naturels, $b \neq 0$. Alors $\frac{a}{a\wedge b}$ et $\frac{b}{a\wedge b}$ sont premiers entre eux.*
- *En particulier, il est toujours possible d'écrire un rationnel $\frac{a}{b}$ sous forme irréductible $\frac{a'}{b'}$, c'est-à-dire de sorte que $a' \wedge b' = 1$, en simplifiant par $a \wedge b$. En imposant $b' > 0$, cette représentation irréductible est unique.*

◁ **Éléments de preuve.**

C'est une propriété de distributivité du produit sur le pgcd. ▷

On déduit des résultats de la section précédente :

Théorème 21.3.5 (Bézout, ou Bachet-Bézout)

Deux entiers naturels a et b sont premiers entre eux si et seulement s'il existe des entiers relatifs x et y tels que $ax + by = 1$.

◁ **Éléments de preuve.**

Les deux sens d'implication découlent respectivement des 2 points du théorème général. ▷

En particulier :

Corollaire 21.3.6 (Éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$, HP)

1. *Soit $n \in \mathbb{N}^*$, et $k \in [\![0, n-1]\!]$. La classe de k modulo n est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si k et n sont premiers entre eux.*
2. *En particulier, $\mathbb{Z}/p\mathbb{Z}$ et un corps si et seulement si p est premier.*

◁ **Éléments de preuve.**

Écrire une relation de Bézout et la réduire modulo n. Réciproquement, écrire une relation d'inversion, et prendre des représentants dans $\mathbb{Z}$; cela fournit une relation de Bézout. ▷

On déduit de ce dernier point le théorème important suivant, parfois appelé « petit » théorème de Fermat, par opposition à un autre qui a donné tant de fil à retordre à des générations de mathématiciens.

Théorème 21.3.7 (Fermat)

Soit p un nombre premier, et a un entier naturel quelconque. Alors $a^p \equiv a[p]$. Si de plus a n'est pas divisible par p, $a^{p-1} \equiv 1[p]$.

◁ **Éléments de preuve.**

Considérer l'ordre de a dans le groupe multiplicatif de $\mathbb{Z}/p\mathbb{Z}$. ▷

Méthode 21.3.8 (Calcul d'un inverse modulo n)

Soit k premier avec n. Pour calculer l'inverse de k modulo n, c'est-à-dire l'inverse de k dans $\mathbb{Z}/n\mathbb{Z}$, déterminer une relation de Bézout $xk+yn=1$, par l'algorithme d'Euclide. On obtient alors $xk \equiv 1\ [n]$.

On en arrive à l'un des résultats les plus importants, qui est à la base de théorème fondamental de l'arithmétique, qui est l'existence et l'unicité de la décomposition en facteurs premiers.

Lemme 21.3.9 (Lemme ou théorème de Gauss)

Soit a, b et c trois entiers naturels tels que $a \mid bc$ et $a \wedge b = 1$. Alors $a \mid c$.

◁ **Éléments de preuve.**

Multiplier par c une relation de Bézout entre a et b et aviser. ▷

Corollaire 21.3.10 (Lemme d'Euclide)

Soit p un nombre premier tel que p divise ab. Alors p divise a ou p divise b.

◁ **Éléments de preuve.**

Si p ne divise pas a, alors p est premier avec a (car ... ?) ▷

Corollaire 21.3.11

Soit a, b et c des entiers tels que a soit premier avec b et avec c. Alors a est premier avec le produit bc.

◁ **Éléments de preuve.**

Par contraposée, en considérant p diviseur premier commun à a et bc. ▷

Note Historique 21.3.12

Gauss démontre le lemme d'Euclide de façon directe et élémentaire, par un argument de récurrence qu'on peut assimiler à une descente infinie. Il en déduit puis en déduit l'existence et l'unicité de la décomposition primaire, qu'il utilise pour démontrer son propre théorème ci-dessus. Sa démarche est donc totalement différente de celle que nous adoptons dans ce chapitre.

À l'aide d'une relation de Bézout, on obtient également :

Proposition 21.3.13

Si a et b sont premiers entre eux et $a|c$ et $b|c$, alors $ab|c$.

◁ **Éléments de preuve.**
Multiplier une relation de Bézout par c et remarquer que $ab \mid ac$ et $ab \mid bc$. ▷

Corollaire 21.3.14
Si $a \wedge b = 1$, alors

$$\begin{cases} x \equiv y \ [a] \\ x \equiv y \ [b] \end{cases} \implies x \equiv y \ [ab].$$

◁ **Éléments de preuve.**
C'est une réexpression de 21.3.13 avec $c = y - x$. ▷

Corollaire 21.3.15 (PPCM de deux nombres premiers entre eux)
Si a et b sont premiers entre eux, $a \vee b = ab$

◁ **Éléments de preuve.**
La proposition 21.3.13 fournit la propriété de minimalité requise. ▷

On en déduit une relation entre PGCD et PPCM :

Proposition 21.3.16 (relation liant PGCD et PPCM)
Soit a et b deux entiers strictement positifs. Alors $ab = (a \wedge b)(a \vee b)$.

◁ **Éléments de preuve.**
Se ramener au cas de deux entiers premiers entre eux en divisant par $a \wedge b$. ▷

Cette relation sera limpide lorsqu'on aura la description du PGCD et du PPCM en terme de décomposition primaire.

III.2 Famille finie d'entiers premiers entre eux

Enfin, nous définissons deux notions sur un nombre quelconque d'entiers, à bien distinguer l'une de l'autre :

Définition 21.3.17 (Nombres premiers entre eux deux à deux)
Soit $a_1, \ldots, a_n$ des entiers naturels. On dit que $a_1, \ldots, a_n$ sont premiers entre eux deux à deux si deux entiers pris au hasard parmi ces n entiers sont toujours premiers entre eux, c'est-à-dire :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad (i \neq j) \implies a_i \wedge a_j = 1.$$

La notion suivante est moins forte :

Définition 21.3.18 (Nombres premiers entre eux dans leur ensemble)
Soit $a_1, \ldots, a_n$ des entiers naturels. On dit que $(a_1, \ldots, a_n)$ sont premiers entre eux dans leur ensemble si $a_1 \wedge \cdots \wedge a_n = 1$, ou de façon équivalente, s'il existe des entiers $x_1, \ldots, x_n$ tels que

$$x_1 a_1 + \cdots + x_n a_n = 1.$$

Par exemple 10, 12 et 15 sont premiers entre eux dans leur ensemble. Vous remarquerez en revanche que deux quelconques d'entre eux ne sont pas premiers entre eux !
La reciproque, en revanche est vraie :

Proposition 21.3.19
Soit $(a_1, \dots, a_n) \in \mathbb{N}^n$. Si $a_1, \dots, a_n$ sont premiers entre eux deux à deux (il suffit même en fait que deux d'entre eux soient premiers entre eux) alors ils sont premiers entre eux dans leur ensemble.

◁ **Éléments de preuve.**
Évident en utilisant la bonne caractérisation. ▷

III.3 Fonction indicatrice d'Euler

Arrivé à ce stade, nous ne pouvons nous empêcher de parler de la fonction indicatrice d'Euler, d'une importance capitale en arithmétique.

Définition 21.3.20 (Fonction indicatrice d'Euler, ou fonction phi d'Euler, HP)
La fonction φ d'Euler est la fonction qui à tout n de $\mathbb{N}^*$ associe $\varphi(n)$ le nombre d'entiers de $[\![1, n]\!]$ premiers avec n.

En particulier, les résultats précédents amènent :

Proposition 21.3.21 (Cardinal de $(\mathbb{Z}/n\mathbb{Z})^\times$, HP)
Soit $n \in \mathbb{N}^*$. Alors $(\mathbb{Z}/n\mathbb{Z})^\times$ est de cardinal $\varphi(n)$.

◁ **Éléments de preuve.**
On a décrit les inversibles. ▷

Grâce au théorème de Lagrange, on en déduit notamment la généralisation suivante du petit théorème de Fermat

Corollaire 21.3.22 (Théorème d'Euler, HP)
Soit $n \in \mathbb{N}^*$. Alors pour tout $x \in \mathbb{N}^*$ tel que $x \wedge n = 1$, $x^{\varphi(n)} \equiv 1\ [n]$.

On peut montrer (voir exercices ou problèmes) que φ est multiplicative : si $a \wedge b = 1$, $\varphi(ab) = \varphi(a)\varphi(b)$. On peut également calculer facilement $\varphi(p^k)$, pour p premier. On peut en déduire une expression de $\varphi(n)$ pour tout n, à condition de connaître la décomposition primaire de n.

IV Décomposition primaire d'un entier

IV.1 Décomposition primaire

Un théorème incontournable de l'arithmétique est bien sûr :

Théorème 21.4.1 (Décomposition primaire)
Tout entier strictement positif n s'écrit de façon unique sous la forme

$$n = p_1 \times \cdots \times p_k,$$

où $p_1 \leqslant \cdots \leqslant p_k$ sont des nombres premiers, ce produit étant éventuellement vide si $n = 1$.

On montre d'abord le lemme suivant, obtenue par itération du lemme d'Euclide.

Lemme 21.4.2

Soit $n \in \mathbb{N}*$, tel que $n = p_1 \dots p_k$, où les p_i sont premiers. Soit $p \in \mathbb{P}$. Si $p \mid n$, alors il existe $i \in [\![1, k]\!]$ tel que $p = p_i$.

◁ **Éléments de preuve.**
Récurrence sur k en appliquant le lemme d'Euclide pour diminuer le nombre de facteurs. ▷

◁ **Éléments de preuve du théorème 21.4.1.**
Récurrence forte pour l'existence (en distinguant les cas p premier, et p composé). Récurrence forte aussi pour l'unicité en divisant par exemple par le plus petit diviseur premier. ▷

Un anneau dans lequel on a une propriété d'existence et d'unicité (à facteurs multiplicatifs inversibles près, et à l'ordre près des facteurs) d'une décomposition en facteurs irréductibles est appelé *anneau factoriel*. Ainsi, quitte à multiplier par l'élément inversible -1 pour obtenir la décomposition d'un entier relatif, ce résultat se réexprime en disant que $\mathbb{Z}$ est un anneau factoriel. On peut montrer que tout anneau principal est factoriel. Par ailleurs tout anneau euclidien est principal (même démonstration que dans $\mathbb{Z}$). Donc tout anneau euclidien est factoriel. C'est par exemple le cas de $\mathbb{K}[X]$, lorsque $\mathbb{K}$ est un corps (ainsi tout polynôme se décompose de façon unique, à éléments inversibles près, comme produit de polynômes irréductibles). C'est par exemple aussi le cas de l'anneau $\mathbb{Z}[i]$ des entiers de Gauss. La question se pose alors, notamment dans ce dernier cas, de savoir décrire les éléments irréductibles. C'est une question pas complètement triviale dans $\mathbb{Z}[i]$, en rapport avec le théorème des deux carrés (donnant la description des entiers s'écrivant comme somme de deux carrés).

IV.2 Valuations p-adique

Un nombre premier p pouvant apparaître plusieurs fois dans la décomposition de n, nous définissons :

Définition 21.4.3 (Valuation p-adique d'entiers)

Soit n un entier et p un entier premier. On appelle valuation p-adique de l'entier n, et on note $v_p(n)$, le nombre d'occurrences (éventuellement nul) de l'entier p dans la décomposition primaire de n.

Il s'agit donc de l'unique entier v tel que p^v divise n mais pas p^{v+1} (et donc pas les puissances suivantes non plus) :

$$v_p(n) = \max\{v \text{ tel que } p^v \mid n\}.$$

En notant $\mathbb{P}$ l'ensemble des nombres premiers, il vient donc :

Proposition 21.4.4 (Reexpression de la décomposition primaire)

Pour tout $n \in \mathbb{N}^*$

$$n = \prod_{p \in \mathbb{P}} p^{v_p(n)},$$

ce produit ayant un sens, puisque constitué d'un nombre fini de termes non égaux à 1.

Proposition 21.4.5 (Règles sur les valuations)

Soit a et b deux entiers strictement positifs, et p un nombre premier.

1. *On a : $v_p(ab) = v_p(a) + v_p(b)$.*
2. *Si b divise a, on a : $v_p\left(\frac{a}{b}\right) = v_p(a) - v_p(b)$.*

◁ **Éléments de preuve.**
Ce n'est rien de plus que les règles de manipulation des exposants ! ▷

Proposition 21.4.6 (Caractérisation de la divisibilité par les valuations)
Soit a et b des entiers strictement positifs. Alors $a \mid b$ si et seulement si pour tout $p \in \mathbb{P}$, $v_p(a) \leqslant v_p(b)$.

Proposition/Définition 21.4.7 (Valuations de rationnels (HP))
Soit $q \in \mathbb{Q}$ et $p \in \mathbb{P}$. Alors la quantité $v_p(a) - v_p(b)$ est indépendante de la représentation $q = \dfrac{a}{b}$. On définit alors la valuation p-adique de q par :

$$v_p(q) = v_p(a) - v_p(b),$$

où $q = \frac{a}{b}$ est une représentation quelconque de q.

Proposition 21.4.8 (règles sur les valuations de rationnels)
Les propriétés de la proposition 21.4.5 restent valides pour des rationnels.

Les valuations permettent alors de caractériser les entiers parmi les rationnels :

Proposition 21.4.9 (Caractérisation des entiers parmi les rationnels par les valuations)
Soit $q \in \mathbb{Q}$. Alors q est un entier (relatif) si et seulement si pour tout $p \in \mathbb{P}$, $v_p(q) \geqslant 0$.

◁ **Éléments de preuve.**
Sens direct évident. Sens réciproque : poser $q = \frac{a}{b}$, et exprimer les décompositions primaires de a et b. ▷

Le théorème suivant n'est pas explicitement au programme, mais est souvent très utile pour faire de l'arithmétique avec des factorielles :

Proposition 21.4.10 (Formule de Legendre, HP)
Soit p un nombre premier, et n un nombre entier naturel. Alors

$$v_p(n!) = \sum_{k=1}^{+\infty} \left\lfloor \frac{n}{p^k} \right\rfloor,$$

cette somme étant en fait finie (les termes sont tous nuls pour k assez grand)

◁ **Éléments de preuve.**
Compter tous les facteurs multiples de p, puis une deuxième fois (on les a déjà compté une fois) ceux multiples de p^2, puis une troisième fois ceux multiples de p^3 etc. ▷

On peut de la sorte calculer $v_p\left(\binom{n}{k}\right)$. Un cas particulier important (qu'on démontre plus simplement de façon directe) est :

Lemme 21.4.11
Soit p un nombre premier. Alors pour tout $k \in [\![1, p-1]\!]$, $\binom{p}{k} \equiv 0\ [p]$.

◁ **Éléments de preuve.**

Localiser les facteurs p dans la fraction. Il n'y en a pas beaucoup. ▷

De ce lemme, on tire :

Proposition 21.4.12

Soit p un nombre premier, et $\mathbb{K}$ un corps de caractéristique p, et Soit a et b deux éléments de $\mathbb{K}$. Alors $(a+b)^p = a^p + b^p$.

Ce résultat affirme en fait que l'application $x \mapsto x^p$ est un endomorphisme du corps $\mathbb{K}$. Ce morphisme s'appelle *morphisme de Frobenius*.
Dans le cas particulier de $\mathbb{K} = \mathbb{F}_p$, la proposition ci-dessus se réécrit

$$(a+b)^p \equiv a^p + b^p \ [p],$$

et en particulier :

$$(a+1)^p \equiv a^p + 1 \ [p].$$

Cette dernière identité est à la base d'une démonstration possible du petit théorème de Fermat, par récurrence sur a.
Évidemment, cette preuve explique moins bien la raison profonde du résultat que la preuve voyant ce résultat comme un cas particulier du théorème de Lagrange, appliqué au groupe $(\mathbb{Z}/p\mathbb{Z})^*$.

Remarque 21.4.13

Le petit de théorème de Fermat est notamment beaucoup utilisé dans les tests de non primalité (avec un ordinateur !). En effet, pour montrer qu'un entier p n'est pas premier, il suffit de trouver un entier a tel que $a^p \not\equiv a \ [p]$. Ainsi, par exemple, à l'aide d'un ordinateur, on peut trouver facilement, pour $n = \frac{1}{9}(10^{31} - 1)$ (nombre contitué de 31 chiffres 1) que $2^n \not\equiv 2 \ [n]$. Ainsi, n n'est pas premier. Trouver une décomposition de n est une autre paire de manches...

En revanche, déduire de la validité de tests de Fermat qu'un nombre est premier est beaucoup plus délicat, car le petit théorème de Fermat ne caractérise pas les nombres premiers : il existe des nombres composés vérifiant les identités du théorème de Fermat (la seconde identité étant alors donnée pour tout x premier avec n). Ces nombres sont appelés *nombres de Carmichael*.

IV.3 PGCD et PPCM vus sous l'angle de la décomposition primaire

Nous traduisons d'abord la divisibilité en terme de décomposition primaire :

Lemme 21.4.14 (Caractérisation de la divisibilité par les valuations)

Soit a et b deux entiers non nuls. Alors $a|b$ si et seulement si pour tout $p \in \mathbb{P}$, $v_p(a) \leqslant v_p(b)$.

◁ **Éléments de preuve.**

Sens directe par définition de la valuation. Sens réciproque par la décomposition primaire. ▷

Étant donné deux nombres a et b, le pgcd et le ppcm de a et b s'obtiennent facilement à l'aide de leur décomposition primaire. Par exemple :

$$150 = 2 \times 3 \times 5^2 \qquad \text{et} \qquad 180 = 2^2 \times 3^2 \times 5.$$

Ainsi, $150 \wedge 180 = 2 \times 3 \times 5 = 30$ et $150 \vee 180 = 2^2 \times 3^2 \times 5^2 = 900$.
Plus généralement, en utilisant le lemme énoncé ci-dessus, on obtient :

Proposition 21.4.15

Soit a et b deux entiers strictement positifs. Alors, pour tout $p \in \mathbb{P}$,

$$v_p(a \wedge b) = \min(v_p(a), v_p(b)) \qquad \text{et} \qquad v_p(a \vee b) = \max(v_p(a), v_p(b)).$$

◁ **Éléments de preuve.**

C'est la caractérisation de la divisibilité par les valuations, et les propriétés de minimalité et maximalité imposées. ▷

La relation

$$(a \wedge b) \times (a \vee b) = ab$$

devient alors évidente.
Cette description du PGCD et du PPCM peut bien sûr être généralisée au calcul du PGCD et du PPCM d'une famille finie quelconque d'entiers naturels.

V Théorème des restes chinois (HP)

Nous nous intéressons dans cette section à la résolution de systèmes de congruences.

Note Historique 21.5.1

Le nom de théorème des restes chinois provient du fait que le mathématicien chinois Sun Zi du III^{e} siècle répond à la question suivante : « Soit une armée. Si on range les soldats par 3 il en reste 2, si on les range par 5, il en reste 3 et si on les range par 7 il en reste 2. Combien y a-t-il de soldat ? ». La réponse de Sun Zi est : « Multiplie le reste de la division par 3, c'est-à-dire 2, par 70, ajoute-lui le produit du reste de la division par 5, c'est-à-dire 3, avec 21 puis ajoute le produit du reste de la division par 7, c'est-à-dire 2 par 15. Tant que le nombre est plus grand que 105, retire 105. »

V.1 Cas de modulo premiers entre eux

La justification du théorème des restes chinois repose sur le lemme suivant portant sur les groupes (évidemment, c'est un point de vue dont ne disposait pas Sun Zi) :

Théorème 21.5.2 (Produit d'anneaux $\mathbb{Z}/n\mathbb{Z}$ d'ordre premiers entre eux)

1. *Soit a et b deux entiers premiers entre eux. Alors l'anneau $\mathbb{Z}/ab\mathbb{Z}$ est isomorphe à l'anneau produit $\mathbb{Z}/a\mathbb{Z} \times \mathbb{Z}/b\mathbb{Z}$, un isomorphisme explicite étant donné par $\overline{x} \mapsto x(1,1) = (x \mod a, x \mod b)$.*
2. *Plus généralement, si $a_1, \ldots, a_n$ sont deux à deux premiers entre eux, alors*

$$\mathbb{Z}/(a_1 \cdots a_n)\mathbb{Z} \simeq \mathbb{Z}/a_1\mathbb{Z} \times \cdots \times \mathbb{Z}/a_n\mathbb{Z},$$

l'isomorphisme étant donné explicitement par $\overline{x} = (x \mod a_1, \ldots, x \mod a_n)$.

◁ **Éléments de preuve.**

Pour montrer que c'est un isomorphisme, étudier le noyau en se ramenant à des propriétés arithmétiques. ▷

Corollaire 21.5.3 (Unicité modulo $a_1 \cdots a_n$ de la solution d'un système)

Soit $a_1, \dots, a_n$ des entiers deux à deux premiers entre eux, et $b_1, \dots, b_n$ des entiers. Alors le système

$$\begin{cases} x \equiv b_1 \ [a_1] \\ \vdots \\ x \equiv b_n \ [a_n] \end{cases}$$

admet une solution, unique modulo $a_1 \cdots a_n$.

Corollaire 21.5.4 (cas d'un second membre commun)

En particulier, si $b_1 = \cdots = b_n = b$, la seule solution modulo $a_1 \cdots a_n$ est b.

En notant $A = a_1 \cdots a_n$, et pour tout $B = (b_1, \dots, b_n)$, $X(B)$ l'unique élément de $\mathbb{Z}/A\mathbb{Z}$ solution du système ci-dessus, on obtient :

Proposition 21.5.5

L'application $B \mapsto X(B)$ est un morphisme de groupes, de $(\mathbb{Z}^n, +)$ dans $\mathbb{Z}/A\mathbb{Z}$.

◁ **Éléments de preuve.**

Vérifications faciles. ▷

Ainsi, il suffit de déterminer les valeurs de $X(e_i)$ pour les vecteurs $e_i = (0, \dots, 0, 1, 0, \dots, 0)$, le 0 étant en position i. En effet, pour un vecteur $B = (b_1, \dots, b_n)$, on aura alors

$$X(B) = \sum_{i=1}^{n} b_i X(e_i).$$

Pour déterminer $X(e_i)$, on part de la constatation que les résultats précédents permettent de réduire le système au système à deux inconnues :

$$\begin{cases} x \equiv 1 \ [a_i] \\ x \equiv 0 \ [\widehat{a_i}], \end{cases}$$

où nous avons noté $\widehat{a_i} = \prod_{j \neq i} a_j$. Ce système peut être résolu en utilisant une relation de Bézout, du fait que a_i et $\widehat{a_i}$ sont premiers entre eux. On commence donc par déterminer (par l'algorithme d'Euclide étendu) u_i et v_i tels que

$$u_i a_i + v_i \widehat{a_i} = 1.$$

Nous avons alors :

$$v_i \widehat{a_i} \equiv 0 \ [\widehat{a_i}] \qquad \text{et} \qquad v_i \widehat{a_i} \equiv 1 \ [a_i].$$

Ainsi, $X(e_i) = \overline{v\widehat{a_i}}$ (classe dans $\mathbb{Z}/A\mathbb{Z}$).
On énonce :

Théorème 21.5.6 (Théorème des restes chinois)

Si les a_i sont deux à deux premiers entre eux, il existe modulo $a_1 \dots a_n$ une unique solution au système $x \equiv b_i \ [a_i]$, $i \in [\![1, n]\!]$ donné par

$$x \equiv \sum_{i=1}^{n} b_i v_i \widehat{a_i} \ [a_1 \dots a_n],$$

où $\widehat{a_i} = \prod_{j \neq i} a_j$ et v_i est un coefficient d'une relation de Bézout $u_i a_i + v_i \widehat{a_i} = 1$.

◁ **Éléments de preuve.**
Résulte des explications précédentes ▷

V.2 Résolution d'un système quelconque

Les a_i n'étant plus supposés premiers entre eux, en écrivant chaque a_i sous la forme $p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, on remplace chaque ligne du système par un sous-système équivalent :

$$\begin{cases} x \equiv b_i \ [p_1^{\alpha_1}] \\ \vdots \\ x \equiv b_i \ [p_k^{\alpha_k}] \end{cases}$$

Ainsi, toutes les équations du nouveau système obtenu sont réduites modulo une puissance d'un nombre premier. Un même nombre premier peut intervenir dans plusieurs lignes, avec un exposant éventuellement différent. La compatibilité de ces équations est aisée à vérifier (pour chaque paire d'équations modulo p^α et p^β, avec disons $\alpha \leqslant \beta$, on doit obtenir les mêmes équations en réduisant la seconde modulo p^α). Si les équations ne sont pas compatibles, il n'y a pas de solution, sinon, on garde la plus contraignante des équations (à savoir celle faisant intervenir le plus grand exposant). On est alors ramené à un système tel qu'étudié plus haut, auquel on peut appliquer le théorème des restes chinois.

Exemple 21.5.7

Résoudre les deux systèmes suivants :

1. $\begin{cases} x \equiv 3 \ [42] \\ x \equiv 10 \ [49] \end{cases}$
2. $\begin{cases} x \equiv 3 \ [42] \\ x \equiv 9 \ [49] \end{cases}$

22

Polynômes et fractions rationnelles

On a ainsi traité le problème comme s'il s'agissait simplement de déterminer la forme des racines, dont l'existence est admise sans démonstration, manière de raisonner qui est ici entièrement illusoire et en fait une véritable petitio principis.

(C. F. Gauss, à propos des « démonstrations » antérieures du théorème de d'Alembert-Gauss)

Le but de ce chapitre est d'étudier les fonctions polynomiales $x \mapsto a_d x^d + \cdots + a_1 x + a_0$. On s'intéresse notamment aux propriétés arithmétiques (produit, somme, divisibilité...) et aux propriétés analytiques (racines, dérivation...)

On se placera dans un cadre plus formel dans le but notamment de généraliser des constructions *a priori* uniquement valables pour des polynômes à coefficients réels (comme la dérivation) à des polynômes à coefficients dans des anneaux plus généraux.

Seuls les polynômes à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$ sont théoriquement au programme. Nous donnerons les définitions formelles plus généralement pour les polynômes à coefficients dans un anneau. Ce point de vue a une certaine importance, car c'est lui qui permet d'itérer ensuite la construction pour obtenir les polynômes de plusieurs indéterminés, puisque si $\mathbb{A}$ est un anneau, $\mathbb{A}[X]$ hérite de cette structure d'anneau. En revanche, si $\mathbb{K}$ est un corps, on n'a pas de structure de corps sur $\mathbb{K}[X]$, mais uniquement d'anneau.

Pour l'étude des propriétés de l'anneau des polynômes, nous nous limiterons au cas où les coefficients sont dans un corps. On dispose dans ce cas de propriétés plus fortes que dans le cas général des polynômes sur un anneau, en particulier toutes les propriétés permettant de faire de l'arithmétique. On a même parfois besoin de certaines hypothèses supplémentaires (par exemple sur la caractéristique du corps). Ainsi, pour certaines propriétés, nous reviendrons aux exigences du programme ($\mathbb{R}$ ou $\mathbb{C}$), en précisant parfois ce qu'il en est dans les autres situations. Il convient de remarquer que dans ce cas, ces propriétés ne sont en général plus satisfaites pour les polynômes à plusieurs indéterminées, puisque $\mathbb{A}[X]$ n'est pas un corps.

I Polynômes à coefficients dans un anneau commutatif

Soit $\mathbb{A}$ un anneau, qu'on supposera commutatif.

I.1 Polynômes formels

Remarque 22.1.1 (Motivation de la définition)

Une fonction polynomiale réelle est entièrement déterminée par la suite de ses coefficients. Les différentes constructions telles la somme, le produit, la dérivation, peuvent s'écrire uniquement sur les coefficients.

La remarque précédente semble justifier de considérer un polynôme comme une suite de coefficients. Seul un nombre fini de ces coefficients doit être non nul.

Définition 22.1.2 (Polynômes formels)

- Un polynôme formel P à coefficients dans $\mathbb{A}$ est une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{A}$, nulle à partir d'un certain rang.
- Le réel a_k est appelé k-ième coefficient de P, ou coefficient du monôme de degré k de P.
- L'ensemble des polynômes (formels) à coefficients dans $\mathbb{A}$ est noté $\mathbb{A}[X]$.

Exemples 22.1.3

1. $\mathbb{R}[X]$ ou $\mathbb{C}[X]$, polynômes formels à coefficients réels ou complexes ;
2. $\mathbb{Q}[X]$, ensemble des polynômes à coefficients rationnels ;
3. $\mathbb{Z}[X]$, ensemble des polynômes à coefficients entiers ;
4. $(\mathbb{Z}/n\mathbb{Z})[X]$, polynômes à coefficients dans $\mathbb{Z}/n\mathbb{Z}$, dont un cas particulier important est $\mathbb{F}_p[X]$.

I.2 Opérations arithmétiques sur les polynômes

Les polynômes considérés dans cette section sont à coefficients dans un anneau commutatif.

Définition 22.1.4 (Somme de polynômes de $\mathbb{A}[X]$)

La *somme* de deux polynômes $P = (a_n)_{n\in\mathbb{N}}$ et $Q = (b_n)_{n\in\mathbb{N}}$ de $\mathbb{A}[X]$ est la suite $P+Q = (a_n+b_n)_{n\in\mathbb{N}}$, qui est bien nulle à partir d'un certain rang.

Définition 22.1.5 (Produit d'un polynôme de $\mathbb{A}[X]$ par un élément de $\mathbb{A}$)

Le produit de $P = (a_n)_{n\in\mathbb{N}}$ par $\lambda \in \mathbb{A}$ est $\lambda P = (\lambda \cdot a_n)_{n\in\mathbb{N}}$.

Lorsque $\mathbb{A}$ est un corps, on parle de multiplication par un scalaire.

Définition 22.1.6 (Produit de deux polynômes de $\mathbb{A}[X]$)

Soit $P = (a_n)_{n\in\mathbb{N}}$ et $Q = (b_n)_{n\in\mathbb{N}}$ deux polynômes de $\mathbb{A}[X]$. Soit pour tout $n \in \mathbb{N}$, $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Alors $(c_n)_{n\in\mathbb{N}}$ est un polynôme. On définit alors $PQ = (c_n)_{n\in\mathbb{N}}$.
La suite $\mathbf{c} = (c_n)_{n\in\mathbb{N}}$ est appelée *produit de convolution* des suites $\mathbf{a} = (a_n)$ et $\mathbf{b} = (b_n)$, et est parfois notée $\mathbf{c} = \mathbf{a} \star \mathbf{b}$.

Théorème 22.1.7 (Structure d'anneau de $\mathbb{A}[X]$)

La somme et le produit définis ci-dessus munissent $\mathbb{A}[X]$ d'une structure d'anneau commutatif.

◁ **Éléments de preuve.**
Vérifier tous les axiomes de la structure. Notamment l'associativité du produit, qui nécessite quelques manipulations sur les sommes. ▷

Avertissement 22.1.8

Ne généralisez pas trop vite « $\mathbb{A}$ anneau $\implies$ $\mathbb{A}[X]$ anneau » en « $\mathbb{K}$ corps $\implies$ $\mathbb{K}[X]$ corps ». Cette dernière affirmation est fausse ! Ainsi, si $\mathbb{K}$ est un corps, tout ce qu'on peut dire, c'est que $\mathbb{K}[X]$ est un anneau.

Cependant, lorsque $\mathbb{K}$ est un corps on peut munir $\mathbb{K}[X]$ d'une structure plus riche. En effet, la multiplication par un scalaire munit $\mathbb{K}[X]$ d'une structure d'espace vectoriel sur le corps $\mathbb{K}$, compatible d'une certaine manière avec la structure d'anneau (voir chapitre ultérieur). La structure totale (espace vectoriel + anneau) est appelée structure d'algèbre sur le corps $\mathbb{K}$.
Dans le cas où $\mathbb{A}$ n'est pas un corps, on peut adapter la définition des espaces vectoriels, en définissant la notion de *module* sur un anneau $\mathbb{A}$: $\mathbb{A}[X]$ est alors muni d'une structure de module sur $\mathbb{A}$. Associé à sa structure d'anneau, on parle aussi de structure d'algèbre (sur l'anneau $\mathbb{A}$)

Note Historique 22.1.9

Les premiers polynômes apparaissant dans l'histoire des mathématiques sont des polynômes de petits degrés associés à des équations traduisant des problèmes concrets : ainsi trouve-t-on dès l'époque babylonienne des résolutions d'équations polynomiales de degré 2. Ces méthodes sont systématisées par Al Khwarizmi à la fin du premier millénaire. Ainsi, les polynômes sont d'abord introduits de façon fonctionnelle. La notation par exposants pour les puissances apparaît plus tard ; elle est introduite par Nicolas Chuquet au 15^{e} siècle. Auparavant, on répétait l'inconnue autant de fois que le degré, ce qui était une limitation à une étude générale. Cependant, dès le 14^{e} siècle, le point de vue formel apparaît dans les travaux de Ibn al-Banna, qui présente les polynômes sous la forme de suites de coefficients. C'est exactement l'approche que nous venons d'en faire.

I.3 Indéterminée formelle

Par commodité, on adopte une notation plus proche de la notation fonctionnelle qu'on connaît pour les fonctions polynomiales ; cette notation est plus facile à manipuler que la définition formelle par les suites. De ce fait, à partir du moment où nous aurons défini l'indéterminée formelle X (remplaçant la notion de variable pour les fonctions polynomiales), nous n'utiliserons plus la définition formelle des polynômes par les suites.
On rappelle que par définition, l'anneau $\mathbb{A}$ considéré contient un élément neutre pour le produit, noté 1.

Définition 22.1.10 (Indéterminée formelle)

On définit dans $\mathbb{A}[X]$ l'indéterminée formelle X comme étant le polynôme $X = (0, 1, 0, 0, \ldots)$.

Ainsi, X n'est pas une variable (au sens fonctionnel), mais un polynôme bien précis, auquel on donne un nom particulier, et auquel on attribue une notation bien particulière, dont le but est l'analogie avec les fonctions polynomiales.

Avertissement 22.1.11

- En particulier, l'indéterminée formelle X n'étant pas une variable, elle ne doit pas être quantifiée, et ne peut pas être utilisée pour résoudre des équations.
- Un polynôme n'est pas une fonction de l'indéterminée formelle, donc la notation $P(X)$ en lieu et place de P n'est pas de mise. On l'utilise néanmoins dans certaines situations, notamment lorsque plusieurs indéterminées sont en jeu. Cette notation peut être justifiée rigoureusement par la notion de spécialisation qu'on verra un peu plus loin.

Proposition 22.1.12 (Monômes)

Pour tout $n \in \mathbb{N}$, on a $X^n = (\underbrace{0, \ldots, 0}_{n \text{ zéros}}, 1, 0, \ldots)$, le 1 étant donc à l'indice n (en commençant l'indexation à 0).

◁ **Éléments de preuve.**
Récurrence sans difficulté. ▷

Corollaire 22.1.13 (Expression d'un polynôme à l'aide de l'indéterminée formelle)
Soit $P = (a_n)_{n\in\mathbb{N}}$ un polynôme de $\mathbb{A}[X]$. Alors $P = \sum_{k=0}^{+\infty} a_k X^k$, cette somme ayant un sens puisqu'elle est en fait finie, les a_k étant nuls à partir d'un certain rang.

Encore une fois, il faut bien comprendre ce que signifie cette égalité : il s'agit bien d'une somme de polynômes, et non d'éléments de $\mathbb{A}$ (signification de l'indéterminée).
De la définition même, il vient :

Proposition 22.1.14 (principe d'identification)
Soit P et Q deux polynômes de $\mathbb{A}[X]$. Notons $P = \sum_{k=0}^{+\infty} a_k X^k$ et $Q = \sum_{k=0}^{+\infty} b_k X^k$, en étant bien conscient que ces sommes sont en fait finies. Alors $P = Q$ si et seulement si pour tout $k \in \mathbb{N}$, $a_k = b_k$.

Les règles de calcul sur les polynômes de $\mathbb{A}[X]$ résultent alors des règles usuelles de calcul dans un anneau découlant de l'associativité, de la commutativité commutativités, et de la distributivité des lois.

Exemples 22.1.15

1. Calcul de $(2 + 3X + 2X^2)(3X + X^2 + 2X^3)$ dans $\mathbb{F}_5[X]$.
2. Calcul de $(X^2 + X + 2)^7$ dans $\mathbb{F}_7[X]$.

Définition 22.1.16 (Monômes)
Un monôme est un polynôme de la forme aX^n, $a \neq 0$. L'entier n est appelé *degré* du monôme.

Ainsi, tout polynôme est une somme de monômes de degrés deux à deux distincts.
Les polynômes formels peuvent aussi se composer. En effet, étant donné un polynôme P, on peut considérer, pour tout $n \in \mathbb{N}$, le polynôme P^k (en a en particulier $P^0 = 1$ et $P^1 = P$). On définit alors la composée de deux polynômes de la sorte :

Définition 22.1.17 (Composition de polynômes formels)
Soit P et Q deux polynômes de $\mathbb{A}[X]$. On note $Q = \sum_{k=0}^{d} a_k X^k$. On définit alors le polynôme $Q \circ P$ par :

$$Q \circ P = \sum_{k=0}^{d} a_k P^k.$$

I.4 Dérivation

On sait facilement dériver (au sens analytique) une fonction polynomiale à coefficients réels, $x \mapsto x^n$ se dérivant en $x \mapsto nx^{n-1}$. Cette règle de dérivation peut être vue de façon purement formelle, permettant de généraliser la dérivation des polynômes à un anneau quelconque (dans lequel on ne dispose pas des techniques d'analyse, spécifiques à $\mathbb{R}$).

Définition 22.1.18 (Dérivée formelle d'un polynôme)
Soit $P = \sum_{k=0}^{d} a_k X^k$ un polynôme à coefficients dans un anneau commutatif $\mathbb{A}$. Le *polynôme dérivé* est défini par :

$$P' = \sum_{k=1}^{d} k a_k X^{k-1}.$$

Des vérifications élémentaires montrent :

Proposition 22.1.19 (Linéarité de la dérivation)
Soit P, Q deux polynômes de $\mathbb{A}[X]$, et $a \in \mathbb{A}$.

1. $(P+Q)' = P' + Q'$.
2. $(aP)' = aP'$.

La linéarité s'exprime en terme de structures en affirmant que la dérivation est une application linéaire (c'est-à-dire un homomorphisme d'espaces vectoriels) lorsque $\mathbb{A}$ est un corps, ou un homomophisme de $\mathbb{A}$-modules sinon.

Vu que la définition de la dérivation est calquée sur la dérivée analytique des fonctions polynomiales réelles, on a, sans surprise, des règles de dérivation similaires, et notamment :

Proposition 22.1.20 (Dérivée de produits)

1. *Soit P et Q deux polynômes à coefficients dans $\mathbb{A}$. Alors*
$$(PQ)' = P'Q + PQ'.$$
2. *Soit $P_1, \dots, P_n$ des polynômes à coefficients dans $\mathbb{A}$. Alors*
$$(P_1 \cdots P_n)' = \sum_{i=1}^{n} P_1 \cdots P_{i-1} P_i' P_{i+1} \cdots P_n.$$
3. *(Formule de Leibniz) Soit P et Q deux polynômes à coefficients dans $\mathbb{A}$. Alors*
$$(PQ)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}.$$

◁ **Éléments de preuve.**
Les points 2 et 3 se déduisent du premier de la même manière que pour la dérivation analytique. Le point 1 se ramène, par linéarité, au cas trivial où $P = X^k$ et $Q = X^\ell$. ▷

Corollaire 22.1.21 (Dérivée de P^n)
En particulier, étant donné $n \in \mathbb{N}^*$, $(P^n)' = nP'P^{n-1}$.

Avec un petit abus de notation, on pourrait considérer cette égalité également au rang $n = 0$ (le terme P^{-1} n'est pas bien défini, mais le produit par n annule l'ensemble). On généralise le corollaire précédent de la sorte

Proposition 22.1.22 (Dérivée d'une composition)
Soit P et Q dans $\mathbb{A}[X]$. Alors
$$(Q \circ P)' = P' \times (Q' \circ P).$$

◁ **Éléments de preuve.**
Écrire $Q \circ P$ comme somme de monômes en P et appliquer le corollaire précédent. ▷

I.5 Degré et valuation

Définition 22.1.23 (Degré et valuation)

Soit $P = (a_n)_{n\in\mathbb{N}}$ un polynôme à coefficients dans un anneau commutatif $\mathbb{A}$.

1. Le *degré de* P est $\deg(P) = \max\{n \in \mathbb{N} \mid a_n \neq 0\}$.
   Si P est non nul, cet ensemble est non vide, et majoré. Ainsi, $\deg(P) \in \mathbb{N}$.
   Si $P = 0$, par convention, $\deg(P) = -\infty$.
2. La *valuation de* P est $\mathrm{val}(P) = \min\{n \in \mathbb{N} \mid a_n \neq 0\}$.
   Si P est non nul, cet ensemble est non vide, et minoré. Ainsi, $\mathrm{val}(P) \in \mathbb{N}$.
   Si $P = 0$, par convention, $\mathrm{val}(P) = +\infty$.

On trouve aussi parfois la notation $\omega(P)$ pour la valuation.
On utilise souvent la filtration suivante de $\mathbb{A}[X]$ (une filtration de E est une chaîne d'inclusions d'union totale E)

Notation 22.1.24 (Filtration par les degrés)

Soit $\mathbb{A}$ un anneau et $n \in \mathbb{N}$. On note $\mathbb{A}_n[X]$ l'ensemble des polynômes de degré au plus n.

Proposition 22.1.25

On a évidemment $\mathbb{A}_0[X] \subset \mathbb{A}_1[X] \subset \cdots \subset \mathbb{A}_n[X] \subset \cdots$ *et*

$$\mathbb{A}[X] = \bigcup_{n=0}^{+\infty} \mathbb{A}_n[X].$$

Remarque 22.1.26

On a évidemment $\mathbb{A}_0[X] \simeq \mathbb{A}$. On identifie souvent les deux, de sorte à pouvoir considérer que $\mathbb{A} \subset \mathbb{A}[X]$.

Définition 22.1.27 (Monôme dominant, coefficient dominant, polynôme unitaire)

Soit $P = \sum_{k=0}^{d} a_k X^k$ un polynôme de $\mathbb{A}[X]$, de degré d.

1. Le monôme dominant de P est le monôme $a_d X^d$, donc le monôme de plus haut degré de P.
2. Le coefficient dominant de P est l'élément a_d de $\mathbb{A}$, donc le coefficient du monôme dominant.
3. Le polynôme P est dit unitaire si son coefficient dominant vérifie $a_d = 1_{\mathbb{A}}$.

Proposition 22.1.28 (Degré d'une somme, d'un produit, d'une dérivée)

Soit P et Q deux polynômes de $\mathbb{A}[X]$, et $\lambda \in \mathbb{A}$. Alors :

1. $\deg(P+Q) \leqslant \max(\deg(P), \deg(Q))$.
   Si $\deg(P) \neq \deg(Q)$, *alors* $\deg(P+Q) = \max(\deg(P), \deg(Q))$.
2. *Si $\mathbb{A}$ est intègre, et si $\lambda \neq 0$,* $\deg(\lambda P) = \deg(P)$.
3. *Si $\mathbb{A}$ est intègre (en particulier si $\mathbb{A}$ est un corps) et si P et Q sont non nuls,* $\deg(PQ) = \deg(P) + \deg(Q)$.
4. $\deg(P') \leqslant \deg(P) - 1$.

◁ **Éléments de preuve.**
Vérifier avec les règles opératoires définies que les monômes de degré plus grand que le degré voulu sont tous nuls, et, pour avoir l'égalité, que le terme de degré maximal est non nul. C'est pour ce point qu'il faut disposer d'une propriété d'intégrité. ▷

Exemples 22.1.29

1. Trouver dans $(\mathbb{Z}/6\mathbb{Z})[X]$ un exemple contredisant le point 2.
2. Trouver dans $(\mathbb{Z}/6\mathbb{Z})[X]$ un exemple contredisant le point 3.
3. Trouver un exemple d'un polynôme non constant pour lequel l'inégalité du point 4 est stricte.

Corollaire 22.1.30 (Théorème de permanence de l'intégrité)
Si $\mathbb{A}$ est intègre, alors $\mathbb{A}[X]$ est intègre.

◁ **Éléments de preuve.**
Quel est alors le degré d'un produit de deux polynômes non nuls ? ▷

Corollaire 22.1.31 (Intégrité des anneaux usuels de polynômes)
Si $\mathbb{K}$ est un corps, $\mathbb{K}[X]$ est intègre. En particulier, les anneaux $\mathbb{R}[X]$, $\mathbb{C}[X]$, $\mathbb{F}_p[X]$, $\mathbb{Q}[X]$ sont intègres.

L'anneau $\mathbb{Z}$ étant également intègre, on obtient aussi l'intégrité de $\mathbb{Z}[X]$.

Corollaire 22.1.32 (Propriétés de stabilité)

1. *$\mathbb{A}_n[X]$ est un sous-groupe additif de $\mathbb{A}[X]$.*
2. *La dérivation $D : \mathbb{A}[X] \longrightarrow \mathbb{A}[X]$ induit un homomophisme de groupes $D_n : \mathbb{A}_n[X] \longrightarrow \mathbb{A}_{n-1}[X]$ (et même un homomorphisme de $\mathbb{A}$-modules dans le sens où D_n respecte aussi le produit externe)*
3. *Si $\mathbb{K}$ est un corps de caractéristique nulle, $D_n : \mathbb{K}_n[X] \longrightarrow \mathbb{K}_{n-1}[X]$ est une surjection. Autrement dit, tout polynôme de $\mathbb{K}_{n-1}[X]$ est primitivable formellement dans $\mathbb{K}_n[X]$.*

◁ **Éléments de preuve.**
Ce sont des vérifications assez immédiates. Pour le dernier point, il faut pouvoir primitiver. La primitivation d'un monoôme X^k nécessite l'inversibilité de $k+1$. ▷

Remarques 22.1.33

- $\mathbb{A}_n[X]$ est-il un sous-anneau de $\mathbb{A}[X]$?
- Trouver un polynôme de $\mathbb{F}_p[X]$ n'admettant pas de primitive.

On précise un peu dans certains cas la relation entre degré de P et degré de P' : dans les situations que vous connaissez, dériver un polynôme non constant baisse son degré de 1. Cependant, ceci n'est pas vrai en toute généralité. Nous avons besoin pour cela d'hypothèses plus fortes.

Proposition 22.1.34 (Degré d'une dérivée dans $\mathbb{K}[X]$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)
Soit $\mathbb{K}$ un corps de caractéristique nulle (par exemple $\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$), et P un polynôme non constant de $\mathbb{K}[X]$. Alors $\deg(P') = \deg(P) - 1$.

◁ **Éléments de preuve.**

Regarder le monôme dominant. ▷

Remarque 22.1.35

1. Trouver dans $\mathbb{F}_p[X]$ un polynôme non constant pour lequel cette égalité est fausse.
2. Si $\mathbb{K}$ est un corps de caractéristique p, quelle condition donner au degré de P pour avoir $\deg(P') = \deg(P) - 1$?

Corollaire 22.1.36

Soit $\mathbb{K}$ un corps de caractéristique nulle, et soit P et Q deux polynômes de $\mathbb{K}[X]$. Si $P' = Q'$, alors P et Q diffèrent d'une constante additive.

◁ **Éléments de preuve.**

La proposition précédente donne une condition nécessaire pour que $(P - Q)' = 0$ ▷

Exemple 22.1.37

Donner un contre-exemple dans le cas où $\mathbb{K} = \mathbb{F}_p$.

On pourrait établir pour les valuations des règles similaires à celles qu'on a pour les degrés. La notion de valuation n'étant pas explicitement au programme, nous laissons le lecteur intéressé établir ces règles par lui-même.

II Arithmétique dans $\mathbb{K}[X]$

On considère ici des polynômes à coefficients dans un corps $\mathbb{K}$. Vous pouvez considérer $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, mais, sauf mention explicite du contraire, les résultats exposés sont valables dans le cadre plus général d'un corps quelconque.

II.1 Division euclidienne

L'anneau $\mathbb{K}[X]$ est euclidien (c'est-à-dire qu'il y existe une notion de division euclidienne) :

Théorème 22.2.1 (Théorème de la division euclidienne dans $\mathbb{K}[X]$)

Soit $\mathbb{K}$ un corps. Pour tout polynôme A et $B \neq 0$ de $\mathbb{K}[X]$ il existe d'uniques polynômes Q et R tels que :

(i) $A = BQ + R$

(ii) $\deg(R) < \deg(B)$.

Les polynômes Q et R sont appelés respectivement quotient et reste de la division euclidienne de A par B.

◁ **Éléments de preuve.**

Comme dans le cas entier, par récurrence (sur quoi ?) ▷

Méthode 22.2.2 (Algorithme de la division euclidienne)

- On pose la division euclidienne comme la division des entiers, en disposant A à gauche et B à droite, les monômes étant écrits dans l'ordre décroissant des degrés (donc en marquant d'abord les monômes de plus haut degré).
- On trouve le monôme aX^k tel que aX^kB ait même monôme dominant que A, puis on effectue la différence $A_1 = A - aX^kB$, qui a donc un degré strictement plus petit que A.
- On recommence sur A_1, et on en déduit A_2
- On recommence ainsi jusqu'à obtenir A_k de degré strictement plus petit que B. Alors A_k est le reste recherché, et le quotient est la somme des monômes par lesquels on a multiplié B pour obtenir les A_i successifs.

Cet algorithme peut facilement être implémenté dans un langage informatique dans $\mathbb{R}[X]$ ou $\mathbb{C}[X]$; un polynôme est dans ce cas représenté par la liste de ses coefficients (on revient à la définition formelle des polynômes sous forme d'une suite finie).

Exemple 22.2.3

Division euclidienne de $X^6 + 3X^2 + 1$ par $X^2 + X + 1$.

On verra un peu plus loin une méthode basée sur l'étude des racines pour déterminer rapidement le reste d'une division euclidienne par un polynôme de petit degré dont on connaît les racines.

Remarque 22.2.4

L'algorithme de la division euclidienne peut-il être mené sans restriction dans $\mathbb{A}[X]$ lorsque $\mathbb{A}$ est un anneau commutatif quelconque ? Donner une condition sur le polynôme B pour qu'on puisse effectuer dans $\mathbb{A}[X]$ la division euclidienne d'un polynôme A quelconque par B.

II.2 Idéaux de $\mathbb{K}[X]$

Comme pour l'arithmétique de $\mathbb{Z}$, il est commode de raisonner en terme d'idéaux. Le résultat rendant la situation totalement similaire à celle de $\mathbb{Z}$ est le fait que tous les idéaux de $\mathbb{K}[X]$ sont principaux, donc engendrés par un unique polynôme.

Théorème 22.2.5 (Description des idéaux de $\mathbb{K}[X]$)

Soit $\mathbb{K}$ un corps. Alors $\mathbb{K}[X]$ est un anneau principal. De plus, deux polynômes P et Q engendrent le même idéal si et seulement s'il existe un $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

◁ **Éléments de preuve.**

S'aider de la division euclidienne, comme pour montrer que $\mathbb{Z}$ est principal. C'est une propriété général de tout anneau muni d'une division euclidienne (anneau euclidien). ▷

On notera (P) l'idéal engendré par un polynôme P.

Remarques 22.2.6

- $\{XP+YQ,\ P,Q \in \mathbb{R}[X,Y]\}$ est un idéal non principal de $\mathbb{R}[X,Y] = \mathbb{R}[X][Y] = \mathbb{R}[Y][X]$. Ainsi, le résultat précédent rentre en défaut lorsque $\mathbb{K}$ n'est pas un corps (dans l'exemple, on considère les polynômes)à coefficients dans l'anneau $\mathbb{R}[X]$).
- En fait, on peut montrer que $\mathbb{K}[X]$ est principal si et seulement si $\mathbb{K}$ est un corps.
- Nous avons vu en exercice que tout anneau euclidien est principal. Ce n'est donc ici qu'une conséquence de ce résultat plus général.

II.3 Divisibilité

Soit $\mathbb{K}$ un corps. Ainsi, $\mathbb{K}[X]$ est principal. On note (P) l'idéal engendré par un élément P de $\mathbb{K}[X]$, à savoir

$$(P) = \{PQ,\ Q \in \mathbb{K}[X]\}.$$

Toutes les propriétés de cette section se démontrant comme dans le cas entier, nous nous dispensons d'en indiquer des preuves.

Définition 22.2.7 (Divisibilité dans $\mathbb{K}[X]$)

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que B divise A s'il existe un polynôme Q tel que $A = BQ$. Inversement, on dit que A est un multiple de B.

Ainsi, B divise A si et seulement si le reste de la division euclidienne de A par B est nul.
Comme dans $\mathbb{Z}$, on a la caractérisation suivante :

Proposition 22.2.8 (Caractérisation en termes d'idéaux)

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A divise B si et seulement si $B \in (A)$, ou encore si et seulement si $(B) \subset (A)$.

Comme dans le cadre général, on dit que le couple (A, B) est un couple de polynômes associés si A divise B et B divise A. Il vient alors de la description des idéaux de $\mathbb{A}[X]$ que :

Proposition 22.2.9 (Polynômes associés)

Soit $(A, B) \in \mathbb{K}[X]^2$. Alors (A, B) est un couple de polynômes associés si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $A = \lambda B$.

II.4 PGCD et PPCM

Proposition/Définition 22.2.10 (PGCD de deux polynômes)

Soit $\mathbb{K}$ un corps. Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul et $P \in \mathbb{K}[X]$. Les propositions suivantes sont équivalents :

(i) P divise A et B et est de degré maximal pour cette propriété.

(ii) P divise A et B et tout autre diviseur de A et B est aussi un diviseur de P

(iii) $(P) = (A) + (B)$.

Si ces propriétés sont vérifiées on dit que P est un PGCD de A et B.

Il n'y a pas unicité d'un PGCD de A et B. Plus précisément :

Proposition 22.2.11 (Description des PGCD)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul, et P un PGCD de A et B. Alors un polynôme Q est un PGCD de A et B si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

Notation 22.2.12 ($A \wedge B$)

En particulier, si A et B sont deux polynômes de $\mathbb{K}[X]$ dont l'un au moins est non nul, il existe un unique PGCD unitaire de A et B. Ce PGCD unitaire est noté $A \wedge B$.

Comme dans le cas de $\mathbb{Z}$, on déduit du troisième point équivalent de la définition l'existence de relations de Bézout.

Proposition 22.2.13 (Relation de Bézout)

Soit A et B deux polynômes de $\mathbb{K}[X]$ dont l'un au moins est non nul.

1. *Il existe des polynômes U et V tels que $AU + BV = A \wedge B$*
2. *Soit $P \in \mathbb{K}[X]$ tel qu'il existe U et V dans $\mathbb{K}[X]$ tels que $AU + BV = P$. Alors P est un multiple de $A \wedge B$.*

Comme dans $\mathbb{Z}$, on peut déterminer un PGCD et une relation de Bézout par l'algorithme d'Euclide étendu, en utilisant le lemme suivant :

Lemme 22.2.14

Soit A et B deux polynômes tels que $B \neq 0$. Soit Q et R le quotient et le reste de la division de A par B. Alors $A \wedge B = B \wedge R$

Méthode 22.2.15 (Calcul d'un PGCD et d'une relation de Bézout)

La méthode est la même que dans $\mathbb{Z}$, par divisions euclidiennes successives, jusqu'à obtenir un reste nul. Le dernier reste non nul est le PGCD, et en combinant les relations de division obtenues, on trouve de la même façon que dans $\mathbb{Z}$ une relation de Bézout (quitte à diviser par un scalaire, pour obtenir le PGCD unitaire)

Exemple 22.2.16

Trouver les PGCD de $X^3 + 2X^2 + 2X + 1$ et $X^3 + 3X^2 + 4X + 2$, et une relation de Bézout.

Comme pour le cas de $\mathbb{Z}$, la définition du PGCD s'étend au cas du PGCD de n polynômes. On obtient alors :

Proposition 22.2.17 (Propriétés du PGCD)

L'opération $\wedge$ est commutative et associative. Par ailleurs, si C est unitaire, $(A \wedge B)C = AC \wedge BC$.

Évidemment, on peut aussi définir les PPCM :

Proposition/Définition 22.2.18 (PPCM de deux polynômes)

Soit $\mathbb{K}$ un corps. Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$, et $P \in \mathbb{K}[X]$. Les propositions suivantes sont équivalents :

(i) A et B divisent P et P est de degré minimal pour cette propriété.

(ii) A et B divisent P et tout autre multiple de A et B est aussi un multiple de P

(iii) $(P) = (A) \cap (B)$.

Si ces propriétés sont vérifiées on dit que P est un PPCM de A et B.

Proposition 22.2.19 (Description des PPCM)

Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$, et P un PPCM de A et B. Alors un polynôme Q est un PPCM de A et B si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

Notation 22.2.20 ($A \vee B$)

En particulier, si A et B sont deux polynômes non nuls de $\mathbb{K}[X]$, il existe un unique PPCM unitaire de A et B. Ce PPCM unitaire est noté $A \vee B$.

Exemple 22.2.21

$P = (X+1)^2$ et $Q = (X+1)(X-1)$.

II.5 Polynômes premiers entre eux

Définition 22.2.22

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que A et B sont premiers entre eux si $A \wedge B = 1$.

Autrement dit, les seuls diviseurs communs à A et B sont les polynôems constants non nuls.
Plus généralement, on définit comme dans $\mathbb{Z}$ la notion de famille finie de polynômes deux à deux premiers entre eux, ou premiers entre eux dans leur ensemble.
Ici encore, les propriétés valables dans $\mathbb{Z}$ se généralisent :

Théorème 22.2.23 (Théorème de Bézout)

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A et B sont premiers entre eux si et seulement s'il existe deux polynômes U et V tels que $AU + BV = 1$.

Exemple 22.2.24

Soit $\lambda \neq \mu$ dans $\mathbb{K}$. Alors les polynômes $X - \lambda$ et $X - \mu$ sont premiers entre eux.

Lemme 22.2.25 (Lemme de Gauss)

Soit A, B et C trois polynômes de $\mathbb{K}[X]$ tels que A divise BC et A et B soient premiers entre eux. Alors A divise C.

Corollaire 22.2.26

Soit A, B et C trois polynômes tels que A et B divisent C et A et B soient premiers entre eux. Alors AB divise C.

Comme dans $\mathbb{Z}$, on a une relation simple entre PPCM et PGCD, à ceci près que comme ces notions sont définies à constante multiplicative près, il faut faire attention au coefficient dominant :

Proposition 22.2.27 (relation entre PGCD et PPCM)

Soit A et B deux polynômes de coefficients dominants a et b respectivement. Alors

$$ab(A \wedge B)(A \vee B) = AB.$$

II.6 Décomposition en facteurs irréductibles

Définition 22.2.28 (Polynôme irréductible)
Un polynôme non constant P de $\mathbb{K}[X]$ est irréductible si et seulement s'il n'est, à une constante multiplicative non nulle près, divisible que par lui-même et par 1.

Exemples 22.2.29

1. Les polynômes $X - \lambda$ sont irréductibles ($\lambda \in \mathbb{K}$)
2. Dans $\mathbb{R}[X]$, tout polynôme $aX^2 + bX + c$ tel que $\Delta < 0$ est irréductible.
3. Ces polynômes ne sont pas irréductibles dans $\mathbb{C}[X]$.

Lemme 22.2.30 (Existence d'un diviseur irréductible)
Tout polynôme non constant P admet un diviseur irréductible.

Lemme 22.2.31
Soit P un polynôme irréductible de $\mathbb{K}[X]$ et A un polynôme, non multiple de P. Alors A et P sont premiers entre eux.

Le lemme de Gauss fournit facilement la généralisation suivante du lemme d'Euclide :

Lemme 22.2.32 (Euclide)
Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors si P divise AB, P divise A ou P divise B.

De façon équivalente, la contraposée fournit :

Corollaire 22.2.33
Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors, si P ne divise ni A ni B, P ne divise pas AB.

Enfin, voici l'analogue du théorème de la décomposition primaire :

Théorème 22.2.34 (Décomposition en facteurs irréductibles)
Soit P un polynôme non nul de $\mathbb{K}[X]$.

1. *Il existe un élément $\lambda \in \mathbb{K}^*$ et des polynômes irréductibles $P_1, \dots, P_k$ tels que*
$$P = \lambda P_1 \cdots P_k.$$
2. *Cette décomposition est unique, à l'ordre près des facteurs, et à multiplication près de chaque facteur (y compris λ) par un élément non nul de $\mathbb{K}$.*
3. *En particulier, si on impose que les P_i soient unitaires, cette décomposition est unique, à l'ordre près des facteurs.*

Nous verrons un peu plus loin la description complète des polynômes irréductibles de $\mathbb{R}[X]$ et de $\mathbb{C}[X]$. Pour cela, il nous faut étudier d'un peu plus près les propriétés liées aux racines d'un polynôme.

III Racines d'un polynôme

Pour pouvoir définir la notion de racine d'un polynôme, il faut d'abord pouvoir « appliquer » un polynôme à un élément de $\mathbb{A}$, donc transformer un polynôme formel en une fonction polynomiale.

III.1 Spécialisation, évaluation

Le lien entre les polynômes formels de $\mathbb{R}$ et les fonctions polynomiales sur $\mathbb{R}$ est assez clair : étant donné un polynôme formel $P = \sum_{k=0}^{d} a_k X^k$ de $\mathbb{R}[X]$, on peut lui associer la fonction polynomiale

$$\tilde{P} : x \mapsto \sum_{k=0}^{d} a_k x^k.$$

La seule condition pour pouvoir faire cela de façon plus générale dans $\mathbb{A}[X]$ est de pouvoir faire dans $\mathbb{A}$ des produits (donc calculer des puissances) et des sommes. Comme $\mathbb{A}$ est un anneau commutatif, cela ne pose pas de problème particulier, et on peut donc définir :

Définition 22.3.1 (Fonction polynomiale associée à un polynôme)

1. Soit $P \in \mathbb{A}[X]$, donné par $P = \sum_{k=0}^{d} a_k X^k$. La fonction polynomiale $\tilde{P} : \mathbb{A} \longrightarrow \mathbb{A}$ associée à P est la fonction définie par :
$$\forall b \in \mathbb{A}, \quad \tilde{P}(b) = \sum_{k=0}^{d} a_k b^k.$$
On rappelle que par convention, $b^0 = 1_{\mathbb{A}}$.
2. L'ensemble des fonctions polynomiales sur $\mathbb{A}$ est l'ensemble :
$$\mathbb{A}[x] = \{\tilde{P} \mid P \in \mathbb{A}[X]\}.$$
Par définition, on a donc $\mathbb{A}[x] \subset \mathbb{A}^{\mathbb{A}}$.

Définition 22.3.2 (Évaluation d'un polynôme)

Soit P un polynôme de $\mathbb{A}[X]$. L'évaluation de P en $b \in \mathbb{A}$ est l'élément de $\mathbb{A}$ défini par $\tilde{P}(b)$. Pour simplifier les notations, on désigne souvent cette évaluation plus simplement par $P(b)$.

Proposition 22.3.3 (Respect des structures)

Soit $\mathbb{A}$ un anneau commutatif. L'application $\varphi : \mathbb{A}[X] \longrightarrow \mathbb{A}[x]$ définie par $\varphi(P) = \tilde{P}$ est un homomorphisme d'anneaux surjectif.

◁ **Éléments de preuve.**

C'est comme cela qu'on a défini les lois de $\mathbb{A}[X]$, de sorte à copier celles de $\mathbb{A}[x]$! ▷

Intuitivement, il apparaît clair que lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, on peut identifier les polynômes formels à coefficients dans $\mathbb{K}$ et les fonctions polynomiales sur $\mathbb{K}$. C'est ce que nous exprimons dans le théorème suivant :

Théorème 22.3.4 ($\mathbb{K}[X] \simeq \mathbb{K}[x]$ pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $\varphi : \mathbb{K}[X] \longrightarrow \mathbb{K}[x]$ définie par $\varphi(P) = \tilde{P}$. Alors φ est un isomophisme d'anneaux.

◁ **Éléments de preuve.**

On verra une version plus générale et algébrique de ce résultat. Pour le moment, on se contente d'une démonstration analytique, consistant par exemple, dans $\mathbb{A}^{\mathbb{A}}$, à considérer la limite de $x \mapsto P(x)$ (avec $P \in \mathrm{Ker}(\varphi)$), divisé par son monôme dominant. ▷

Remarques 22.3.5

1. En considérant le petit théorème de Fermat, montrer que cette propriété n'est pas vraie pour tous les corps $\mathbb{K}$.
2. On montrera un peu plus loin qu'une condition suffisante pour que cette identification soit vraie est que $\mathbb{K}$ soit un corps infini. C'est le cas en particulier lorsque $\mathbb{K}$ est de caractéristique nulle.

On retiendra donc l'avertissement suivant :

Avertissement 22.3.6

Si $\mathbb{K}$ n'est pas un corps infini (par exemple $\mathbb{K} = \mathbb{F}_p$), deux polynômes distincts P et Q de $\mathbb{K}[X]$ peuvent correspondre à la même application polynomiale. Ainsi, les polynômes sont davantage différenciés dans $\mathbb{K}[X]$ que dans $\mathbb{K}[x]$. On n'a donc pas possibilité en général d'identifier polynômes formels et fonctions polynomiales.

Enfin, dans le cas spécifique de $\mathbb{R}$ (seul cas dans lequel on peut considérer la dérivée au sens analytique), on a également, du fait même des définitions :

Proposition 22.3.7

Pour tout polynôme P de $\mathbb{R}[X]$, $\widetilde{P'} = \widetilde{P}'$.

◁ **Éléments de preuve.**

Encore une fois, c'est ainsi qu'on a défini la dérivée formelle ! ▷

Ainsi, les opérations définies formellement coïncident avec les opérations sur les fonctions polynomiales, y compris la dérivation dans le cas de $\mathbb{R}$.

Il est important de constater que le cadre formel qu'on s'est donné pour définir les polynômes permet d'« appliquer » un polynôme à des éléments qui sortent du cadre initialement fixé. Pour prendre un exemple, étant donné un polynôme $P = \sum_{k=0}^{d} a_k X^k$ de $\mathbb{R}[X]$ et M une matrice carrée à coefficients réels, on peut considérer le polynôme de matrices

$$P(M) = \sum_{k=0}^{k} a_k M^k,$$

où il faut bien prendre garde au fait que M^0 désigne la matrice identité I_n.

Exemple 22.3.8

Soit $P = 2 + 3X + 3X^2$, et $M = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. Calculer $P(M)$.

On peut formaliser ce type de construction. Le bon cadre à se fixer est celui donné par la structure de $\mathbb{A}$-algèbre.

Définition 22.3.9 (Algèbre)

Soit $\mathbb{A}$ un anneau commutatif. On dit que $\mathbb{B}$ est une algèbre (unitaire) sur $\mathbb{A}$ si $\mathbb{B}$ est lui-même un anneau commutatif, muni d'une loi externe à opérateurs dans $\mathbb{A}$, telle que, pour tout λ, μ dans $\mathbb{A}$ et tout x, y dans $\mathbb{B}$:

- $(\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x)$
- $\lambda \cdot (xy) = (\lambda \cdot x)y = x(\lambda \cdot y)$
- $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$
- $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$
- $1_{\mathbb{A}} \cdot x = x$.

En gros, ce qu'il faut en retenir, c'est que dans une $\mathbb{A}$-algèbre $\mathbb{B}$, on peut faire, avec des règles de calcul raisonnablement semblables aux situations usuelles, la somme et le produit d'éléments de $\mathbb{B}$ ainsi que le produit d'un élément de $\mathbb{A}$ par un élément de $\mathbb{B}$. En particulier, étant donné un polynôme $P = \sum_{k=0}^{d} a_k X^k$ de $\mathbb{A}[X]$ et $b \in \mathbb{B}$, l'expression suivante a un sens :

$$P(b) = \sum_{k=0}^{d} a_k b^k.$$

Il convient de bien noter que par convention $b^0 = 1_{\mathbb{B}}$.
On parle de *spécialisation* du polynôme P en $b \in \mathbb{B}$.

Exemples 22.3.10

1. L'ensemble des matrices carrées de taille n, à coefficients réels, est une $\mathbb{R}$-algèbre : la situation décrite plus haut est un cas particulier de cette situation plus générale.
2. On utilisera beaucoup en algèbre linéaire des polynômes d'endomorphismes (applications linéaires d'un espace vectoriel dans lui-même), l'ensemble des endomorphismes d'un espace vectoriel E sur $\mathbb{K}$ étant une $\mathbb{K}$-algèbre pour la somme usuelle et le produit défini par la composition. Ainsi, f^n désigne dans ce cas la composition itérée de f, et f^0 désigne la fonction identité id_E.
3. $\mathbb{K}[X]$ est lui-même une $\mathbb{K}$-algèbre. On peut donc spécialiser P en un autre polynôme Q. On peut constater ici que $P(Q) = P \circ Q$ (ce qui est incorrect dans le cadre de notations fonctionnelles !). En particulier, $P(X) = P \circ X = P$. Ainsi, on peut confondre les notations P et $P(X)$ (encore une fois contrairement aux notations fonctionnelles similaires).

III.2 Racines et multiplicité

Soit $\mathbb{A}$ un anneau commutatif.

Définition 22.3.11 (Racine d'un polynôme)

Soit $P \in \mathbb{A}[X]$ et $a \in \mathbb{A}$. On dit que a est une racine de P si $P(a) = 0$. On note $\mathrm{Rac}(P)$ l'ensemble des racines de P

S'il y a une ambiguïté sur l'anneau des coefficients, on peut noter $\mathrm{Rac}_{\mathbb{A}}(P)$ l'ensemble des racines de P en tant que polynôme à coefficients dans $\mathbb{A}$. Cette ambiguïté peut se produire par exemple dans le cadre de l'étude de polynômes à coefficients réels, lorsqu'on considère les racines complexes de P. L'ensemble des racines complexes de P se notera alors $\mathrm{Rac}_{\mathbb{C}}(P)$, par opposition à $\mathrm{Rac}_{\mathbb{R}}(P)$, ensemble des racines réelles de P.

Théorème 22.3.12 (Caractérisation des racines par la divisibilité)
Soit $\mathbb{K}$ un corps, $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$. Alors r est racine de P si et seulement si $X - r$ divise P, donc s'il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - r)Q$.

◁ **Éléments de preuve.**
Effectuer une division euclidienne. ▷

Remarque 22.3.13
Ce théorème reste valable dans $\mathbb{A}[X]$ pour un anneau commutatif quelconque. Pourquoi ?

Si après factorisation $P = (X - r)Q$, r est encore racine de Q, alors r est « plusieurs fois » racine de P. En comptant le nombre de fois qu'on peut mettre $X - r$ en facteur, on obtient la multiplicité de r :

Définition 22.3.14 (Multiplicité d'une racine)
Soit $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$. On dit que r est racine d'ordre de multiplicité $k \in \mathbb{N}^*$ si et seulement si $(X-r)^k$ divise P et $(X-r)^{k+1}$ ne divise pas P. Autrement dit, il existe $Q \in \mathbb{R}[X]$ tel que $P = (X-r)^k Q$, avec $Q(r) \neq 0$.

Remarque 22.3.15
La multiplicité de la racine r correspond donc à la valuation du facteur $(X - r)$ dans la décomposition primaire de P.

Par convention, on dira que r est racine de multiplicité 0 si r n'est pas racine de P. Une racine de multiplicité 1 est aussi appelée racine simple de P, et une racine de multiplicité 2 est appelée racine double. Lorsque la multiplicité est supérieure ou égale à 2, on parlera de racine multiple.
Cette mise en facteur maximale de $(X - r)^r$ peut être mise en valeur par la formule de Taylor pour les polynômes, nécessitant une hypothèse supplémentaire sur $\mathbb{K}$.

Théorème 22.3.16 (Formule de Taylor pour les polynômes)
Soit $\mathbb{K}$ un corps de caractéristique nulle, P un polynôme de $\mathbb{K}[X]$, de degré d, et $a \in \mathbb{K}$. Alors,

$$P = \sum_{n=0}^{d} \frac{P^{(n)}(a)}{n!}(X - a)^n.$$

◁ **Éléments de preuve.**
Récurrence sur le degré. Utiliser l'HR sur P'. Pourquoi a-t-on besoin de l'hypothèse sur la caractéristique ? ▷

Ainsi, si v est le plus petit indice pour lequel le terme de la somme est non nul, on obtient

$$P = (X - a)^v \sum_{n=v}^{d} \frac{P^{(n)}(a)}{n!}(X - a)^{n-v}.$$

Ainsi, l'ordre de multiplicité de r correspond à la valuation de P après changement d'indéterminée $Y = X - r$.
On en déduit de façon immédiate :

Théorème 22.3.17 (Caractérisation de la multiplicité par les dérivées successives)

Soit $\mathbb{K}$ un corps de caractéristique nulle, $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. Le réel a est racine d'ordre de multiplicité k de P si et seulement si : $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$ et $P^{(k)}(a) \neq 0$.

Ainsi, il faut toujours garder à l'esprit des deux facettes de la multiplicité des racines : la propriété de divisibilité, et la caractérisation par les dérivées.

Corollaire 22.3.18

Soit $\mathbb{K}$ un corps de caractéristique nulle. Soit $P \in \mathbb{K}[X]$ et $r \in \mathbb{R}$. Si r est racine de multiplicité $k > 0$ de P, alors r est racine de multiplicité $k-1$ de P'.

III.3 Majoration du nombre de racines

Le corollaire du lemme de Gauss amène :

Théorème 22.3.19

Soit $\mathbb{K}$ un corps. Soit $P \in \mathbb{K}[X]$, et $r_1, \ldots, r_k$ des racines deux à deux distinctes de P, de multiplicités respectives $\alpha_1, \ldots, \alpha_k$. Alors $(X - r_1)^{\alpha_1} \cdots (X - r_k)^{\alpha_k}$ divise P, et $r_1, \ldots, r_k$ ne sont pas racines du quotient.

Corollaire 22.3.20 (Majoration du nombre de racines)

Soit $P \in \mathbb{K}[X]$ non nul, de degré n. Alors P admet au plus n racines (comptées avec multiplicité). En particulier, tout polynôme non nul a un nombre fini de racines.

Exemple 22.3.21

Soit $\mathbb{K}$ un corps dont $\mathbb{F}_p$ est un sous-corps. Montrer que pour tout $x \in \mathbb{K}$, $x^p = x$ si et seulement si $x \in \mathbb{F}_p$.

On en déduit le résultat très important suivant, qu'on décline sous 3 formes équivalentes.

Théorème 22.3.22 (Rigidité des polynômes)

1. *Soit P un polynôme de $\mathbb{K}[X]$ degré au plus n. Alors, si P admet strictement plus de n racines, $P = 0$.*
2. *Si deux polynômes P et Q de $\mathbb{K}_n[X]$ coïncident en strictement plus de n valeurs distinctes. Alors $P = Q$.*
3. *Soit $n \in \mathbb{N}^*$. Étant donnés $x_1, \ldots, x_n$ des éléments deux à deux distincts de $\mathbb{K}$ et $y_1, \ldots, y_n$ des éléments de $\mathbb{K}$ non nécessairement distincts, il existe au plus un polynôme $P \in \mathbb{K}_{n-1}[X]$ tel que pour tout $i \in [\![1, n]\!]$, $P(x_i) = y_i$. Ainsi, sous réserve d'existence, un polynôme de degré au plus $n-1$ est entièrement déterminé par sa valeur en n points distincts.*

◁ **Éléments de preuve.**

1. C'est la contraposée du corollaire précédent.
2. Appliquer le point 1 à $P - Q$.
3. C'est une paraphrase du point 2.

▷

Un corollaire très utilisé de ce résultat est le suivant :

Corollaire 22.3.23 (Le seul polynôme ayant une infinité de racines)
Soit P un polynôme de $\mathbb{K}[X]$ s'annulant en une infinité de points de $\mathbb{K}$. Alors P est le polynôme nul.

On déduit notamment de cette propriété un résultat annoncé un peu plus haut :

Théorème 22.3.24 (Polynômes formels et fonctions polynomiales)
Soit $\mathbb{K}$ un corps infini. Alors l'application de $\mathbb{K}[X]$ dans $\mathbb{K}[x]$ qui à un polynôme formel associe sa fonction polynomiale est un isomorphisme d'anneaux.

◁ **Éléments de preuve.**
Un élément du noyau a alors une infinité de racines ! ▷

Par ailleurs, le dernier point du théorème ci-dessus affirme l'unicité sous réserve d'existence d'un polynôme de degré au plus $n-1$ prenant des valeurs données en n points fixés. Il n'est pas dur de construire explicitement un tel polynôme, fournissant ainsi l'existence. C'est le but du paragraphe suivant.

III.4 Interpolation de Lagrange

On recherche un polynôme de degré au plus n coïncidant en $n+1$ points distincts avec une fonction f, ou, de façon équivalente, prenant en $n+1$ points distincts $x_0, \dots, x_n$, $n+1$ valeurs (non nécessairement distinctes) imposées $y_0, \dots, y_n$.
Pour ce faire, on commence par le cas où les valeurs imposées sont toutes nulles, sauf une égale à 1. On trouvera le cas général en formant une combinaison linéaire.

Définition 22.3.25 (Polynômes interpolateurs de Lagrange)
Soit $x_0, \dots, x_n$ des réels 2 à 2 distincts et $i \in [\![0, n]\!]$. Le i^{e} polynôme interpolateur de Lagrange associé à la famille $x_0, \dots, x_n$ est

$$L_i = \frac{\displaystyle\prod_{\substack{j=0 \\ j\neq i}}^{n} (X - x_j)}{\displaystyle\prod_{\substack{j=0 \\ j\neq i}}^{n} (x_i - x_j)} = \prod_{\substack{j=0 \\ j\neq i}}^{n} \frac{X - x_j}{x_i - x_j}.$$

Lemme 22.3.26
Le polynôme L_i est l'unique polynôme de degré au plus n tel que pour tout $k \in [\![0, n]\!] \setminus \{i\}$, $L_i(x_k) = 0$, et $L_i(x_i) = 1$.

◁ **Éléments de preuve.**
Le fait qu'il vérifie ces propriétés relève de vérifications élémentaires. L'unicité est une propriété de rigidité. ▷

Théorème 22.3.27 (Polynômes d'interpolation de Lagrange)

Soit $\mathbb{K}$ un corps, $n \in \mathbb{N}^*$, $x_0, \dots, x_n$ des éléments distincts de $\mathbb{K}$, et $y_0, \dots, y_n$ des éléments de $\mathbb{K}$. Alors il existe un et un seul polynôme P de $\mathbb{K}_n[X]$ tel que pour tout $i \in [\![0, n]\!]$, $P(x_i) = y_i$, et ce polynôme est donné explicitement par :

$$P = \sum_{i=0}^{n} y_i L_i.$$

En particulier, si f est une fonction définie sur un intervalle I contenant les x_i, le polynôme d'interpolation de Lagrange de f associé à la famille (x_i) est l'unique polynôme P_f coïncidant avec f sur les x_i, à savoir :

$$P_f = \sum_{i=0}^{n} f(x_i) L_i.$$

◁ **Éléments de preuve.**

De même. ▷

Ces polynômes, appelés polynômes d'interpolation de Lagrange, permettent en particulier d'approcher une fonction réelle f par une fonction polynomiale de degré au plus n coïncidant avec f en $n+1$ points distincts.

Corollaire 22.3.28

Soit P le polynôme d'interpolation de Lagrange associée à la famille $(x_i)_{i \in [\![0,n]\!]}$, et aux valeurs $(y_i)_{i \in [\![0,n]\!]}$. Soit $P_0 = (X - x_0) \dots (X - x_n)$. L'ensemble E des polynômes Q (sans restriction de degré) tels que pour tout $i \in [\![0, n]\!]$, $Q(x_i) = y_i$ est alors décrit par :

$$E = P + (P_0) = \{P + (X - x_0) \dots (X - x_n)R, \quad R \in \mathbb{K}[X].\}.$$

◁ **Éléments de preuve.**

Double inclusion facile. ▷

Remarque 22.3.29

Quelle est la structure algébrique de l'ensemble E du théorème précédent ?

III.5 Polynômes scindés

Définition 22.3.30 (Polynôme scindé)

Soit $\mathbb{K}$ un corps. On dit qu'un polynôme non nul P de $\mathbb{K}[X]$ est scindé s'il possède autant de racines (comptées avec multiplicité) que son degré, autrement dit si son nombre de racines est maximal.

Théorème 22.3.31 (Factorisation d'un polynôme scindé)

1. *Un polynôme est scindé si et seulement s'il peut se factoriser de la façon suivante :*

$$P = \lambda(X - x_1)(X - x_2) \cdots (X - x_n),$$

où λ est un scalaire non nul (égal au coefficient dominant de P), n est le degré de P, et $x_1, \dots, x_n$ sont les racines, non nécessairement distinctes, de P.

2. *Si on renomme $y_1, \dots, y_k$ les racines 2 à 2 distinctes de P, de multiplicités respectives $\alpha_1, \dots, \alpha_k$, cette factorisation se réécrit :*

$$P = \lambda(X - y_1)^{\alpha_1} \cdots (X - y_k)^{\alpha_k},$$

et on a $\alpha_1 + \cdots + \alpha_k = n$.

◁ **Éléments de preuve.**
Il suffit de compter les degrés. ▷

Ainsi, un polynôme est scindé si et seulement si sa décomposition en facteurs irréductibles ne fait intervenir que des polynômes irréductibles de degré 1.
Dans $\mathbb{R}[X]$, certaines techniques d'analyse peuvent aider à étudier cette propriété. Ainsi, le théorème de Rolle permet de montrer facilement que :

Proposition 22.3.32 (HP, mais à savoir redémontrer)
Soit P un polynôme scindé non constant de $\mathbb{R}[X]$, à racines simples. Alors P' est scindé à racines simples, et ses racines séparent celles de P.

◁ **Éléments de preuve.**
Appliquer le théorème de Rolle entre les racines de P, et compter les racines de P'. ▷

Voici une propriété plus générale, constituant un exercice (ou un début d'exercice) classique :

Théorème 22.3.33 (HP, mais à savoir redémontrer)
Soit P un polynôme scindé non constant de $\mathbb{R}[X]$. Alors P' est scindé.

◁ **Éléments de preuve.**
De même, mais comptabiliser également les racines multiples, qui restent racines de P'. Plus précisément, que pouvez-vous dire de la « localisation » des racines de P' par rapport à celles de P. ▷

Une propriété importante des polynômes scindés est la possibilité de trouver facilement des relations entre les coefficients et les racines :

Théorème 22.3.34 (Relations coefficients/racines, ou relations de Viète)
Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme de degré n, scindé, de racines (éventuellement non distinctes, apparaissant dans la liste autant de fois que sa multiplicité) $r_1, \dots, r_n$. Alors pour tout $k \in [\![1, n]\!]$:

$$\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} r_{i_1} r_{i_2} \cdots r_{i_k} = (-1)^k \frac{a_{n-k}}{a_n}.$$

◁ **Éléments de preuve.**
Développer la forme factorisée, par la formule de distributivité généralisée. Identifier les coefficients. ▷

Le terme de gauche de cette expression est appelé polynôme symétrique élémentaire de degré k en les racines et souvent noté $\Sigma_k(r_1, \dots, r_n)$.
On peut montrer que tout expression symétrique en $r_1, \cdots, r_n$ peut s'exprimer comme expression polynomiale (à coefficients dans $\mathbb{K}$ des polynômes symétriques en $r_1, \dots, r_n$, donc comme expression polynomiale

en les coefficients du polynôme. Par une construction itérative, on peut même déterminer cette expression polynomiale. Ainsi, par exemple, calculer la somme des racines cubiques d'un polynôme peut se faire sans calculer explicitement les racines (ce qui bien souvent, est de toute façon impossible), juste en se servant des coefficients.
Cette propriété est notamment très utile pour des propriétés d'algébricité puisqu'elle permet de dire que toute expression symétrique à coefficients rationnels en les racines d'un polynôme lui aussi à coefficients rationnels est rationnelle.

IV Polynômes irréductibles dans $\mathbb{C}[X]$ et $\mathbb{R}[X]$

Cette section étudie spécifiquement les polynômes à coefficients complexes ou réels.

IV.1 Factorisations dans $\mathbb{C}[X]$

Nous avons plus ou moins défini $\mathbb{C}$ comme le corps de rupture du polynôme X^2+1, donc le plus petit corps contenant $\mathbb{R}$ dans lequel ce polynôme admet une racine i. Un théorème essentiel, parfois appelé *théorème fondamental de l'algèbre* (c'est dire son importance) est le théorème suivant, que d'Alembert croyait avoir démontré, que Gauss a démontré par différentes méthodes :

Théorème 22.4.1 (d'Alembert-Gauss, admis)
Tout polynôme non constant de $\mathbb{C}[X]$ admet au moins une racine.

Corollaire 22.4.2
Tout polynôme de $\mathbb{C}[X]$ est scindé, donc admet exactement autant de racines (comptées avec multiplicité) que son degré.

◁ **Éléments de preuve.**
Par récurrence sur le degré. Ou bien sans récurrence, en factorisant par tous les $X-r_i$, et s'il reste au bout un polynôme de degré au moins 2, lui appliquer le théorème de d'Alembert-Gauss. ▷

Corollaire 22.4.3
Dans $\mathbb{C}[X]$, les seuls polynômes irréductibles sont les polynômes de degré 1, c'est-à-dire les polynômes $aX+b$, $a\neq 0$.

◁ **Éléments de preuve.**
Ceux de degré plus grand peuvent être factorisés non trivialement. ▷

Exemple 22.4.4
Quelles sont les racines et leurs multiplicités du polynôme X^n-1 ? Factoriser ce polynôme en facteurs irréductibles dans $\mathbb{C}[X]$.

Tous les polynômes de $\mathbb{C}[X]$ se factorisant en polynômes non constants de degré minimal, on obtient alors une caractérisation simple de la divisibilité :

Théorème 22.4.5 (Caractérisation de la divisibilité dans $\mathbb{C}[X]$)
Soit P et Q deux polynômes de $\mathbb{C}[X]$. Alors P divise Q si et seulement si toute racine de P est aussi racine de Q, et que sa multiplicité dans Q est supérieure ou égale à sa multiplicité dans P.

◁ **Éléments de preuve.**

C'est l'analogue de la caractérisation dans $\mathbb{Z}$ par les valuations. Ici, les $X - r$ jouent le même rôle que les entiers premiers, et les multiplicités correspondent au valuations. ▷

Remarque 22.4.6

Est-ce vrai dans $\mathbb{R}[X]$?

IV.2 Facteurs irréductibles dans $\mathbb{R}[X]$

On commence par caractériser les polynômes à coefficients réels parmi les polynômes à coefficients dans $\mathbb{C}$.

Théorème 22.4.7 (Caractérisation des polynômes à coefficients réels)

Soit $P \in \mathbb{C}[X]$. Les propositions suivantes sont équivalentes :

(i) P est à coefficients réels ;

(ii) $P(\mathbb{R}) \subset \mathbb{R}$

(iii) pour tout $z \in \mathbb{C}$, $P(\overline{z}) = \overline{P(z)}$.

◁ **Éléments de preuve.**

$(i) \implies (iii) \implies (ii)$ est facile. L'implication $(ii) \implies (i)$ provient de la propriété de rigidité, en comparant P et $\overline{P}$ sur $\mathbb{R}$. ▷

Corollaire 22.4.8 (Racines complexes d'un polynôme réel)

Soit P un polynôme à coefficients réels, et r une racine de P dans $\mathbb{C}$. Si $r \notin \mathbb{R}$, alors $\overline{r}$ est aussi racine de P, et elles ont même multiplicité.

◁ **Éléments de preuve.**

Appliquer le théorème précédent aux dérivées successives de P. ▷

Ainsi, les racines non réelles d'un polynôme à coefficients réels peuvent être groupées en paires de racines conjuguées de même multiplicité.
Le théorème de d'Alembert-Gauss amène alors :

Théorème 22.4.9 (Polynômes irréductibles de $\mathbb{R}[X]$)

1. *Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif.*
2. *Ainsi, tout polynôme $P \in \mathbb{R}[X]$ peut être factorisé en produit de polynômes de $\mathbb{R}[X]$ de degré 1, ou de degré 2, de discriminant strictement négatif.*

◁ **Éléments de preuve.**

Si P sans racine réelle est de degré au moins 3, considérer une racine complexe (pourquoi existe-t-elle ?) et son conjugué, et regrouper les facteurs complexes correspondants. ▷

Exemple 22.4.10

Factorisation dans $\mathbb{R}[X]$ de $X^n - 1$.

V Fractions rationnelles

La contruction formelle que nous avons donnée des polynômes nous empêche *a priori* de former des quotients de polynômes (donc des fractions rationnelles), comme nous pouvons le faire pour les fonctions polynomiales. En effet, si $\mathbb{K}$ est un corps, les seuls polynômes inversibles sont les polynômes constants non nuls.

Remarque 22.5.1

1. Quels sont les polynômes inversibles de $\mathbb{A}[X]$ lorsque $\mathbb{A}$ est intègre ?
2. Trouver un polynôme inversible et non constant de $(\mathbb{Z}/4\mathbb{Z})[X]$

Une construction similaire à celle permettant de définir $\mathbb{Q}$ à partir de $\mathbb{Z}$ nous permet cependant de définir formellement des quotients de polynômes.

V.1 Définition des fractions rationnelles formelles

Soit, dans tout ce qui suit, $\mathbb{K}$ un corps. On définit sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$ la relation suivante :

$$(P, Q) \sim (R, S) \iff PS = QR,$$

l'égalité étant donnée dans $\mathbb{K}[X]$.

Proposition 22.5.2

La relation ci-dessus est une relation d'équivalence sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$.

◁ **Éléments de preuve.**

C'est la même que celle qui permet de définir $\mathbb{Q}$ à partir de $\mathbb{Z}$. ▷

Définition 22.5.3 (Fraction rationnelle)

Une fraction rationnelle est une classe d'équivalence de la relation ci-dessus. La classe d'équivalence de (P, Q) sera notée $\dfrac{P}{Q}$. L'ensemble des fractions rationnelles sur le corps $\mathbb{K}$ est noté $\mathbb{K}(X)$.

Ainsi, la relation $PS = QR$ amène assez logiquement l'égalité des fractions rationnelles $\dfrac{P}{Q} = \dfrac{R}{S}$.
Pour définir les lois de composition sur $\mathbb{K}(X)$, on commence par les définir sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$. On définit, pour tout (P_1, Q_1) et (P_2, Q_2) de $\mathbb{K}[X] \times \mathbb{K}[X]^*$:

$$(P_1, Q_1) \times (P_2, Q_2) = (P_1P_2, Q_1Q_2) \qquad \text{et} \qquad (P_1, Q_1) + (P_2, Q_2) = (P_1Q_2 + P_2Q_1, Q_1Q_2).$$

Lemme 22.5.4

1. *Les opérations* $\times$ *et* $+$ *sont associatives et commutatives*
2. *La relation* $\sim$ *est une congruence sur les monoïdes* $(\mathbb{K}[X] \times \mathbb{K}[X]^*, \times)$ *et* $(\mathbb{K}[X] \times \mathbb{K}[X]^*, +)$.

Lemme 22.5.5 (Simplification des fractions)

Pour tous polynômes P, Q *et* R *tels que* Q *et* R *soient non nuls,* $\dfrac{PR}{QR} = \dfrac{P}{Q}$.

◁ **Éléments de preuve.**

Vérifier l'équivalence des deux couples en jeu. ▷

Théorème 22.5.6 (Structure de $\mathbb{K}(X)$)

Les lois $+$ et $\times$ induisent des lois de composition, également notées $+$ et $\times$, sur $\mathbb{K}(X)$. L'ensemble $\mathbb{K}(X)$ muni de ces deux lois est un corps.

◁ **Éléments de preuve.**

Même démonstration que pour $\mathbb{Q}$. ▷

Les lois de composition ainsi définies se réécrivent sans surprise :

$$\frac{P_1}{Q_1}\times\frac{P_2}{Q_2}=\frac{P_1P_2}{Q_1Q_2},\qquad \frac{P_1}{Q_1}+\frac{P_2}{Q_2}=\frac{P_1Q_2+P_2Q_1}{Q_1Q_2}\qquad\text{et}\qquad \frac{P_1}{Q}+\frac{P_2}{Q}=\frac{P_1+P_2}{Q}.$$

Définition 22.5.7 (Inclusion canonique de $\mathbb{K}[X]$ dans $\mathbb{K}(X)$)

L'application $P\mapsto\frac{P}{1}$ de $\mathbb{K}[X]$ dans $\mathbb{K}(X)$ est un homomorphisme injectif d'anneaux. La fraction $\frac{P}{1}$ seront désormais identifiée au polynôme P de $\mathbb{K}[X]$.

En particulier, si $P=AB$, alors $B=\frac{P}{A}$.

Proposition/Définition 22.5.8

Soit $F\in\mathbb{K}(X)$ une fraction rationnelle. Il existe un représentant (P,Q), unique à multiplication près par un scalaire non nul, tel que $P\wedge Q=1$ et $F=\frac{P}{Q}$. On dit que $\frac{P}{Q}$ est la forme irréductible de la fraction rationnelle F.

◁ **Éléments de preuve.**

De même que pour $\mathbb{Q}$. ▷

Proposition/Définition 22.5.9 (Dérivation d'une fraction rationnelle)

Soit $F=\frac{P}{Q}$ une fraction rationnelle. La fraction rationnelle $\frac{P'Q-PQ'}{Q^2}$ ne dépend pas du représentant $\frac{P}{Q}$ de F. On peut alors définir la dérivée formelle de la fraction F par :

$$F'=\frac{P'Q-PQ'}{Q^2}$$

◁ **Éléments de preuve.**

Considérer deux couples équivalents, et vérifier l'équivalence des deux couples dérivés. ▷

On peut remarquer que cette dérivation est compatible avec celle des polynômes lorsque $F=\frac{P}{1}$.
La dérivée formelle des fractions rationnelles vérifie les mêmes propriétés que la dérivée analytique :

Proposition 22.5.10 (Propriétés des dérivées des fractions rationnelles)

Soit F et G deux fractions rationnelles, et λ un scalaire. Alors :

1. $(F+\lambda G)'=F'+\lambda G'$
2. $(FG)'=F'G+FG'$
3. $\left(\frac{1}{G}\right)'=-\frac{G'}{G^2}$
4. $\left(\frac{F}{G}\right)'=\frac{F'G-FG'}{G^2}$
5. $(F\circ G)'=(F'\circ G)\cdot G'$

◁ **Éléments de preuve.**

Vérifications élémentaires. Pour la dernière, traiter d'abord le cas où F est un monôme, puis un polynôme. ▷

V.2 Degré, racines, pôles

Proposition/Définition 22.5.11 (Degré d'une fraction rationnelle)

Soit $F \in \mathbb{K}(X)$. La quantité $\deg(P) - \deg(Q)$ ne dépend pas de la représentation $\frac{P}{Q}$ choisie de la fraction F. On appelle degré de F et on note $\deg(F)$ la quantité

$$\deg(F) = \deg(P) - \deg(Q) \text{ où } F = \frac{P}{Q}$$

Il s'agit d'un entier relatif, ou de $-\infty$ si $P = 0$.

◁ **Éléments de preuve.**

Vérification facile. ▷

Les degrés des fractions rationnelles vérifient des propriétés semblables aux degrés des polynômes :

Proposition 22.5.12 (Propriétés des degrés)

Soit F, G des éléments de $\mathbb{K}(X)$.

- $\deg(F+G) \leqslant \max(\deg(F), \deg(G))$, *avec égalité si* $\deg(F) \neq \deg(G)$.
- $\deg(FG) = \deg(F) + \deg(G)$
- $\deg(F^{-1}) = -\deg(F)$

◁ **Éléments de preuve.**

Pour le premier point, mettre sur le même dénominateur (pourquoi est-ce possible). Le dernier résulte du deuxième, ou peut se montrer directement sur une représentation. ▷

Proposition 22.5.13 (Partie entière)

Soit F une fraction rationnelle de $\mathbb{K}(X)$. Il existe un unique polynôme P de $\mathbb{K}[X]$ et une fraction rationnelle G de $\mathbb{K}(X)$ tels que

$$F = P + G \qquad \textit{et} \qquad \deg(G) < 0.$$

Le polynôme P est appelée partie entière de la fraction rationnelle F.

◁ **Éléments de preuve.**

À quel important théorème arithmétique cela vous fait-il penser ? ▷

Définition 22.5.14 (Racine, pôle, multiplicité)

Soit F une fraction rationnelle de $\mathbb{K}(X)$, écrit sous forme irréductible $F = \frac{P}{Q}$.

1. Une racine de F est une racine de P, sa multiplicité est sa multiplicité en tant que racine de P.
2. Un pôle de F est une racine de Q, sa multiplicité est sa multiplicité en tant que racine de Q.

Remarque 22.5.15

Puisque $\frac{P}{Q}$ est irréductible, r ne peut pas être à la fois racine de P et racine de Q.

Exemple 22.5.16
Racines, pôles et leurs multiplicités, de $\dfrac{(X-2)^2}{X^3(X-1)^4}$?

Définition 22.5.17 (Fonction rationnelle associée)
Soit $F = \frac{P}{Q}$ une fraction rationnelle formelle sous forme irréductible, et $\mathcal{P}$ l'ensemble de ses pôles. La fonction rationnelle associée est $\tilde{F} : \mathbb{K} \setminus \mathcal{P} \longrightarrow \mathbb{K}$ définie par

$$\forall x \in \mathbb{K} \setminus \mathcal{P}, \quad F(x) = \frac{P(x)}{Q(x)}.$$

V.3 Décomposition en éléments simples sur un corps quelconque

On étudie dans ce paragraphe l'existence et l'unicité d'une décomposition d'une fraction en somme de fractions simples (appelés éléments simples), de la forme $\frac{Q}{P^\alpha}$, où P est un polynôme irréductible et $\deg(Q) < \deg(P)$. On se donne un corps $\mathbb{K}$.

Lemme 22.5.18
Soit $F = \frac{A}{B}$ une fraction rationnelle écrite sous forme irréductible ($A \wedge B = 1$), et soit $B = \lambda P_1^{\alpha_1} \cdots P_k^{\alpha_k}$ la décomposition de B en facteurs irréductibles. Alors, il existe des d'uniques polynômes $Q_1, \ldots, Q_k$ et un unique polynôme E tels que

$$F = E + \sum_{i=1}^{k} \frac{Q_i}{P_i^{\alpha_i}},$$

et $\deg(Q_i) < \deg(P_i^{\alpha_i})$

◁ **Éléments de preuve.**

- Existence : Appliquer Bézout à une famille bien choisie, diviser par ce qu'il faut et sortir la partie entière.
- Unicité : Par unicité de la partie entière, on se ramène à justifier que si les Q_i vérifient $\deg(Q_i) < \deg(P_i^{\alpha_i})$ et

$$\sum_{i=1}^{k} \frac{Q_i}{P_i^{\alpha_i}} = 0,$$

alors les Q_i sont tous nuls. Cela peut se montrer par récurrence sur k, en multipliant par tous les dénominateurs, et en utilisant le lemme de Gauss pour montrer que $P_k^{\alpha_k}$ divise Q_k, puis que $Q_k = 0$.

▷

Lemme 22.5.19
Soit $\alpha \in \mathbb{N}^*$, P un polynôme de degré d et Q un polynôme de degré strictement inférieur à $d\alpha$. Alors il existe d'uniques polynômes R_j de degré strictement inférieur à d tels que

$$Q = \sum_{j=1}^{\alpha} P^{\alpha-j} R_j.$$

◁ **Éléments de preuve.**

On trouve R_α par division euclidienne. On peut itérer ensuite le procédé, en redivisant à chaque fois le quotient. On peut bien sûr le rédiger par récurrence sur α. L'unicité provient de l'unicité de la division euclidienne. ▷

Théorème 22.5.20 (Décomposition en éléments simples dans $\mathbb{K}(X)$, DÉS)

Avec les notations du lemme précédent, il existe d'uniques polynômes $R_{i,j}$ et un unique polynôme E tels que $\deg(R_{i,j}) < \deg(P_i)$ *et :*

$$F = E + \sum_{i=1}^{k} \sum_{j=1}^{\alpha_i} \frac{R_{i,j}}{P_i^j}.$$

De plus, le polynôme E est la partie entière de la fraction rationnelle F, donc obtenue en effectuant la division euclidienne de P par Q, où $F = \frac{P}{Q}$

◁ **Éléments de preuve.**

Immédiat avec les deux lemmes précédents. ▷

V.4 Décomposition en éléments simples dans $\mathbb{C}(X)$

Les facteurs irréductibles dans $\mathbb{C}[X]$ étant de degré 1, le théorème de décomposition en éléments simples se réexprime assez facilement dans ce cadre :

Théorème 22.5.21 (Décomposition en éléments simples dans $\mathbb{C}(X)$)

Soit F une fraction rationnelle de $\mathbb{C}(X)$, et $r_1, \dots, r_k$ ses pôles, de multiplicités $\alpha_1, \dots, \alpha_k$. Alors il existe un unique polynôme E et d'uniques complexes $\lambda_{i,j}$ ($1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant \alpha_i$) tels que

$$F = E + \sum_{i=1}^{k} \sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X - r_i)^j}.$$

De plus, le polynôme E est la partie entière de la fraction rationnelle F, donc obtenue en effectuant la division euclidienne de P par Q, où $F = \frac{P}{Q}$

Définition 22.5.22 (Partie polaire)

Avec les notations du théorème précédent, la somme $\displaystyle\sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X - r_i)^j}$ est appelée partie polaire de F relativement au pôle r_i.

Exemples 22.5.23

1. Forme de la décomposition en éléments simples de $\dfrac{X^7}{(X-1)^3(X+1)^4}$.
2. Forme de la décomposition en éléments simples de $\dfrac{1}{(1+X^2)^3}$

Proposition 22.5.24 (cas d'un pôle simple)

Soit r un pôle simple de $F = \frac{P}{Q}$ (sous forme irréductible), et soit $\hat{Q}$ le polynôme $\frac{Q}{X-r}$. Alors le coefficient λ du terme $\dfrac{1}{X-r}$ de la DÉS de F est :

$$\lambda = \frac{P(r)}{\hat{Q}(r)} = \frac{P(r)}{Q'(r)}.$$

◁ **Éléments de preuve.**
La première égalité s'obtient en multipliant la forme *a priori* de la DÉS par $X - r$ et en évaluant en r. Cette dernbière manipulation est justifiée par le fait qu'on travaille avec des fractions rationnelles formelles.
La deuxième égalité s'obtient en dérivant Q donné comme produit de $X - r_i$ et en évaluant en r. ▷

Un cas important de décomposition en éléments simples est le cas de la fraction rationnelle $\frac{P'}{P}$.

Théorème 22.5.25 (Décomposition en éléments simples de $\frac{P'}{P}$)
Soit P un polynôme non nul de $\mathbb{C}[X]$. Soit $r_1, \dots, r_k$ les racines de P, de muliplicités $\alpha_1, \dots, \alpha_k$. Alors les pôles de $\frac{P'}{P}$ sont $r_1, \dots, r_k$ et sont tous des pôles simples. La DÉS de $\frac{P'}{P}$ est :

$$\frac{P'}{P} = \sum_{i=1}^{k} \frac{\alpha_i}{X - r_i}.$$

◁ **Éléments de preuve.**
Écrire P sous forme factorisée et dériver le produit. ▷

Nous avons déjà vu que les racines de la dérivée P' d'un polynôme scindé sont situées entre la racine minimale et la racine maximale de P. De la décomposition de $\frac{P'}{P}$, on déduit une propriété similaire dans $\mathbb{C}[X]$:

V.5 Décomposition en éléments simples dans $\mathbb{R}[X]$

Théorème 22.5.26 (DÉS dans $\mathbb{R}(X)$)
Soit $F = \frac{P}{Q}$ une fraction rationnelle sous forme irréductible, et $Q = Q_1^{\alpha_1} \cdots Q_k^{\alpha_k}$ la décomposition en facteurs irréductibles de Q dans $\mathbb{R}[X]$. Ainsi, les Q_i sont de degré 1 ou 2. Alors il existe un unique polynôme E, et d'uniques polynômes $A_{i,j}$, de degré strictement plus petit que Q_i, tels que

$$F = E + \sum_{i=1}^{k} \sum_{j=1}^{\alpha_i} \frac{A_{i,j}}{Q_i^j}.$$

◁ **Éléments de preuve.**
Cas particulier du théorème général, les facteurs irréductibles étant dans ce cas de degré 1 et 2. ▷

Remarques 22.5.27

1. Si Q_i est de degré 1, la partie correspondante dans la DÉS dans $\mathbb{R}(X)$ est la même que dans $\mathbb{C}(X)$.
2. Si Q_i est de degré 2, il admet deux racines complexes conjuguées r et $\overline{r}$. La partie correspondante dans la DÉS est obtenue en regroupant les parties polaires relatives à r et $\overline{r}$, si ces racines sont simples.
3. On a déjà vu en pratique comment déterminer des DÉS dans des cas simples. On a aussi déjà vu l'intérêt que peuvent avoir ces DÉS, notamment pour le calcul d'intégrales.

VI Primitivation des fractions rationnelles réelles

Nous montrons maintenant comment, connaissant une décomposition en éléments simples, on peut primitiver une fraction rationelle.

Le fait important à retenir est qu'on sait primitiver toutes les fractions rationnelles, à condition de connaître explicitement ses pôles. Ce fait est souvent un phare guidant les naufragés du calcul intégral vers des rivages cléments, grâce à cette idée fixe : « se ramener à l'intégrale d'une fraction rationnelle ».

Méthode 22.6.1 (Primitivation d'une fraction rationnelle F)

1. Trouver la décomposition en éléments simples de F dans $\mathbb{R}(X)$.
2. La partie polynomiale se primitive facilement.
3. Les termes en $\dfrac{1}{(x-a)^\alpha}$ se primitivent en $\dfrac{1}{(1-\alpha)(x-a)^{\alpha-1}}$ si $\alpha \neq 1$, et $\ln|x-a|$ si $\alpha = 1$.
4. Pour les facteurs irréductibles de degré 2, faire une mise sous forme canonique du dénominateur et factoriser par le terme constant restant, ce qui ramène à :
$$\int \frac{cx+d}{((ax+b)^2+1)^\alpha}\,\mathrm{d}x.$$
Le changement de variable $y = ax+b$ nous ramène alors à
$$\int \frac{sy+t}{(y^2+1)^\alpha}\,\mathrm{d}y.$$
Le terme $\dfrac{y}{(y^2+1)^\alpha}$ se primitive en $\dfrac{1}{2}\ln(y^2+1)$ si $\alpha = 1$ et en $\dfrac{1}{2(1-\alpha)(y^2+1)^{\alpha-1}}$ sinon. On est donc ramené à $\displaystyle\int \frac{1}{(y^2+1)^\alpha}\,\mathrm{d}y$.
Avec un peu d'habitude, on peut se débarasser du terme cx avant la mise sous forme canonique, en le faisant partir dans une primitivation du type $\dfrac{u'}{u^\alpha}$.
5. Le calcul de $\displaystyle\int \frac{1}{(y^2+1)^\alpha}\,\mathrm{d}y$ se fait par réduction du degré α par intégrations par partie successives, jusqu'à se ramener au cas $\alpha = 1$, pour lequel la primitivation se fait en $\operatorname{Arctan} y$.
Plus précisément, pour faire cette réduction sur l'exposant α, écrire le numérateur sous la forme $1 = 1 + y^2 - y^2$, séparer la fraction entre les deux termes y^2, ce qui permet de se ramener au calcul de $\displaystyle\int \frac{1}{(y^2+1)^{\alpha-1}}\,\mathrm{d}y - \int \frac{y^2}{(y^2+1)^\alpha}\,\mathrm{d}y$. La première intégrale nous ramène à l'exposant précédent, c'est bien ce qu'on voulait. La seconde s'intègre par partie, en primitivant $y \mapsto \dfrac{y}{(y^2+1)^\alpha}$ et en dérivant $y \mapsto y$. Cela nous ramène également à l'exposant précédent.
6. Une alternative pour le point précédent est de poser le changement de variable $x = \operatorname{Arctan}(y)$, de tout réexprimer comme puissance d'un cosinus, puis de linéariser.