Première partie

Des nombres

Les nombres sont au cœur des opérations de codage, de compression et de chiffrement. Cette partie se concentre sur le concept de nombre, les ensemble de nombre, les calculs et les opérations que l'on peut concevoir, la représentation des nombres sur un ordinateur et enfin les structures algébriques qui les architecturent.

L'objectif est de construire des corps finis. Ces corps sont indispensables au chiffrement et au codage contemporain. L'intérêt d'en disposer est principalement qu'on peut utiliser leurs éléments pour encoder les symboles d'un alphabet fini. Au-delà de l'encodage des symboles, il est possible de faire des calculs avec ces éléments dans ces corps : addition, soustraction, multiplication, inversion. Ces opérations linéaires et non linéaires peuvent être effectuées rapidement par les ordinateurs et garantissent la confusion et la non-linéarité des codes.

Les notions exposées au début de cette partie sont fondamentales et probablement que vous les maîtrisez, sans vous être interrogés davantage sur les justifications de ces notions. Elles sont peu nombreuses mais profondes, puissantes et on peut en faire la liste :

1. notion d'ensemble et de cardinal,
2. addition et multiplications sur les ensembles usuels de nombres,
3. fonction partie entière,
4. division euclidienne,
5. algorithmes d'Euclide du PGCD et algorithme étendu,
6. nombres premiers et factorisation.

Les algorithmes engendrés par les notions mathématiques exposées sont détaillés en pseudo-langage. Les algorithmes d'Euclide, même s'ils sont très connus, sont fondamentaux et couramment utilisés encore dans le domaine des codes. C'est pourquoi il est important de les comprendre et de savoir les programmer.

Dans un deuxième temps, cette partie qe concentre sur les corps finis en décrivant les structures algébriques nécessaires à leur construction puis l'arithmétique dans ces corps.

Chapitre 1

Ensembles, nombres et applications

1.1 Des cailloux à compter

Au commencement, il y a les entiers naturels, ces entiers qu'on manipule tous les jours, sans s'en rendre compte : $\mathbb{N} = \{0, 1, 2, 3, \dots\}$. Il a fallu longtemps pour arriver à les abstraire comme nous le faisons aujourd'hui. Pour nos ancêtres, les entiers n'étaient souvent que des petits cailloux[1] que l'on manipulait pour compter des moutons, des mesures de blé ou des personnes.

C'est la théorie des ensembles qui nous a permis d'abstraire ces nombres et de les manipuler comme des ensembles. Cette théorie est née à la fin du XIX^{e} siècle de l'audace de Cantor, l'audace d'avoir osé compter le nombre d'éléments d'un ensemble, même quand celui-ci était infini [**milinowski_uber_1874**].

▪ **Définition 1 — Ensemble.** Un ensemble est une collection de choses qu'on appelle éléments. L'ensemble vide est noté $\emptyset$.

1.2 Cardinal d'un ensemble

▪ **Définition 2 — Cardinal d'un ensemble fini.** Le cardinal d'un ensemble fini est son nombre d'éléments.

▪ **Exemple 1** Soit l'ensemble $A = \{\heartsuit, \blacklozenge, \spadesuit, \clubsuit, \bigstar\}$. On a $\text{card}(A) = 5$.
On a évidemment également $\text{card}(\emptyset) = 0$.

S'il est facile de dire que le cardinal d'un ensemble fini est le nombre de ses éléments, la définition de cardinal d'un ensemble infini est plus délicate. D'ailleurs, plutôt que de le définir directement, on le décrit plutôt [**devoldere_cardinal_2000**] en se donnant les moyens de :

1. dire si deux ensembles E et F ont le même cardinal (cf. définition 3). Il existe alors une bijection[2] entre les deux : on dit qu'ils sont équipotents.
2. comparer les cardinaux de deux ensembles E et F. L'un est contenu dans l'autre s'il existe une injection de l'un dans l'autre.

1. C'est le mot caillou en latin *calculus* qui a donné le mot calcul en français.
2. Il est intéressant de noter que Galiléé déjà avait eu l'idée de faire un appariement bijectif.

On peut définir également une relation d'ordre (cf. définition 65) sur les cardinaux et ainsi tous les comparer.

▪ **Définition 3 — Cardinal d'un ensemble.** Si deux ensembles peuvent être mis en bijection, on dit qu'ils ont le même cardinal, qu'ils sont équipotents.

1.3 Peut-on construire l'ensemble $\mathbb{N}$?

Les entiers jouent un rôle essentiel dans le cadre de la théorie de l'information en général et davantage encore dans le domaine de la cryptographie. La construction de l'ensemble $\mathbb{N}$ peut être établie rigoureusement et simplement dans le cadre de la théorie des ensembles en utilisant l'ensemble vide [3] et trois axiomes : l'axiome de la paire, l'axiome de la réunion et l'axiome de l'infini. C'est la méthode de construction des ensembles dite de Von Neumann.

Pour construire l'ensemble des entiers naturels $\mathbb{N}$ on peut :

1. choisir de noter 0 l'ensemble vide : $0 = \emptyset$,
2. définir une fonction s *successeur de* en posant pour tout ensemble a : $s(a) = a \cup \{a\}$. Le successeur est donc obtenu en ajoutant à l'ensemble l'ensemble de départ comme élément. Si $\text{card}(a) = n$, alors on voit immédiatement que $\text{card}(s(a)) = n + 1$

Le successeur de 0 s'écrit $s(0) = 0 \cup \{0\} = \emptyset \cup \{\emptyset\} = \{0\}$, ensemble que l'on peut noter 1 et dont le cardinal vaut 1. On remarque aussi que $s(1) = 1 \cup \{1\} = \{0\} \cup \{\{0\}\} = \{\emptyset, \{\emptyset\}\} = \{0, 1\}$, ensemble que l'on peut noter 2, dont le cardinal vaut bien 2. Ainsi de suite, $n = \{0, 1, \cdots, n-1\}$, récursivement. Selon cette approche, on peut donc définir chaque nombre entier comme un ensemble.

Afin de garantir l'existence de l'ensemble $\mathbb{N}$, on a besoin de l'axiome de l'infini qui garantit qu'il existe un ensemble contenant 0 fermé pour l'opération successeur, c'est à dire que tout successeur d'un élément de I appartient à I. Grâce à cet axiome, l'ensemble des entiers naturels $\mathbb{N}$ est un ensemble infini d'ensembles dont les cardinaux valent les nombres entiers.

1.4 Ensembles usuels

Dans ce document, on supposera donc que l'on sait construire les ensembles :

1. $\mathbb{N} = \{0, 1, 2, 3 \dots\}$ les entiers naturels,
2. $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3 \dots\}$ les entiers relatifs,
3. $\mathbb{Q}$ l'ensemble des nombres rationnels,
4. $\mathbb{R}$ l'ensemble des nombres réels.

On supposera également que $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ et que les opérations d'addition et de multiplication sont possibles sur tous ces ensembles. Enfin, on suppose pour l'instant que l'on peut comparer les nombres réels entre eux.

Ces ensembles sont représentés sur la figure 1.1. L'ensemble $\mathbb{A}$ est l'ensemble des solutions des équations polynômiales à coefficients rationnels, c'est à dire l'ensemble des racines des polynômes de $\mathbb{Q}[X]$, par exemple $\sqrt{2}$ qui est racine de $X^2 - 2$. L'ensemble $\mathbb{D}$ est l'ensemble des nombres décimaux (cf. définition 4).

3. dont le cardinal vaut 0.

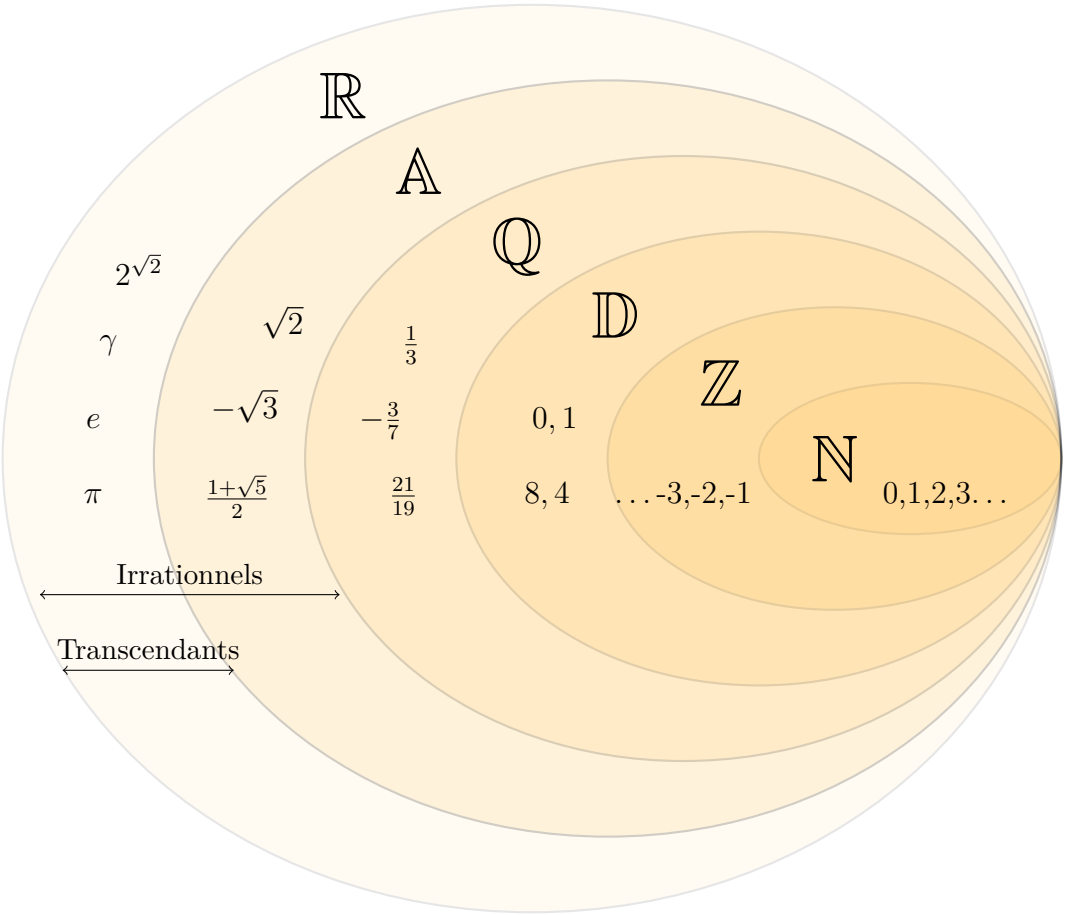


FIGURE 1.1 – Ensembles de nombres : naturels $\mathbb{N}$, relatifs $\mathbb{Z}$, décimaux $\mathbb{D}$, rationnels $\mathbb{Q}$, algébriques $\mathbb{A}$ et réels $\mathbb{R}$.

1.5 Nombres décimaux et dyadiques

■ **Définition 4 — Nombre décimal.** Un nombre décimal est un rationnel qui peut s'écrire sous la forme

$$\frac{a}{10^n}, a \in \mathbb{Z}, n \in \mathbb{N} \tag{1.1}$$

On note $\mathbb{D} = \{\frac{a}{10^n}, a \in \mathbb{Z}, n \in \mathbb{N}\}$ l'ensemble des nombres décimaux.

■ **Exemple 2 — $8,4$ est un nombre décimal.** En effet, on peut l'écrire $\frac{84}{10}$.

■ **Définition 5 — Nombre dyadique.** Un nombre dyadique est un rationnel qui peut s'écrire sous la forme

$$\frac{a}{2^n}, a \in \mathbb{Z}, n \in \mathbb{N} \tag{1.2}$$

On note $\mathcal{D} = \{\frac{a}{2^n}, a \in \mathbb{Z}, n \in \mathbb{N}\}$ l'ensemble des nombres dyadiques.

■ **Exemple 3 — $6,25$ est un nombre dyadique.** En effet, on observe que $6,25_{10} = 110,01_2$. Son développement binaire est fini. On peut l'écrire $\frac{25}{2^2}$.

Proposition 1 — Caractérisation des dyadiques et des décimaux. Un nombre est décimal (resp. dyadyque) si son développement en base 10 (resp. 2) est fini.

■ **Exemple 4 — $1/3$ n'est pas un nombre décimal.** En effet, son développement en base dix n'est pas fini, il se répète. On peut l'écrire $\frac{1}{3} = 0,3333333\ldots$. C'est un nombre rationnel.

■ **Exemple 5 — $8,4$ n'est pas un nombre dyadique.** En effet, on observe que $8,4_{10} = 1000,011001100110011_2\ldots$. Son développement binaire n'est pas fini.

Théorème 1 — Les décimaux ne sont pas les dyadiques. On a $\mathcal{D} \subsetneq \mathbb{D}$. Ce qui signifie qu'il existe des dyadiques qui ne sont pas des décimaux et vice versa.

On a par contre $\mathcal{D} \subset \mathbb{Q}$ et $\mathbb{D} \subset \mathbb{Q}$ (cf. figure 1.1).

■ **Exemple 6 — Décimaux mais pas dyadiques.** $0,1$, $0,2$ et $0,3$ sont des nombres décimaux mais pas dyadiques.

1.6 Applications

■ **Définition 6 — Application.** Soit E et F deux ensembles. Une application de E vers F est un procédé qui associe à chaque élément de E un unique élément de F. E est l'ensemble de départ et F l'ensemble d'arrivée.

On note généralement une application de E vers F de la manière suivante :

$$\begin{aligned} f : E &\longrightarrow F \\ x &\longmapsto y \end{aligned}$$

■ **Définition 7 — Image.** Soit f un application de E vers F. L'élément $f(x) \in F$ est nommée image de x par l'application f.

■ **Définition 8 — Antécédent.** Lorsque $y \in F$ est l'image de $x \in E$, c'est à dire $y = f(x)$, on dit que x est l'antécédent de y.

■ **Définition 9 — Injectivité.** Soit f un application de E vers F. On dit que f est injective si chaque élément de F admet au plus un antécédent par f.

Formulé autrement, une telle application ne peut pas avoir deux fois la même image.

Proposition 2 — Formuler mathématiquement l'injectivité. Une application $f : E \longrightarrow F$ est injective si et seulement si :

$$\forall x \in E, \forall x' \in E, f(x) = f(x') \Longrightarrow x = x' \tag{1.3}$$

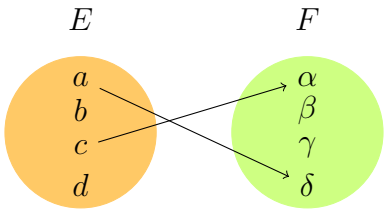


FIGURE 1.2 – Illustration d'une application injective

■ **Définition 10 — Surjectivité.** Soit f un application de E vers F. On dit que f est surjective si chaque élément de F admet au moins un antécédent par f.

Formulé autrement, tous les éléments de F sont atteints par f.

$$\forall y \in F, \exists x \in E, y = f(x) \tag{1.4}$$

■ **Définition 11 — Bijectivité.** Soit f un application de E vers F. On dit que f est bijective si chaque élément de F admet exactement un seul antécédent par f.

Formulé autrement, chaque élément de F est atteint une seule fois par f.

$$\forall y \in F, \exists! x \in E, y = f(x) \tag{1.5}$$

Proposition 3 — Bijectivité. Un application f est bijective si et seulement si elle est injective et surjective.

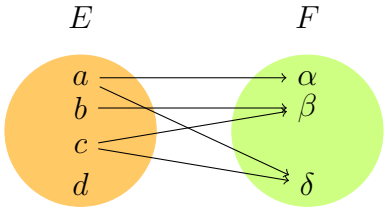


FIGURE 1.3 – Illustration d'une application surjective

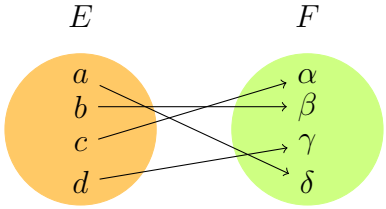


FIGURE 1.4 – Illustration d'une application bijective

1.7 Partie entière

On définit la partie entière d'un nombre réel comme suit :

■ **Définition 12 — Partie entière.** La partie entière d'un nombre réel α est le plus grand entier plus petit ou égal à ce nombre. On le note $\lfloor \alpha \rfloor$.

$$\begin{aligned} \lfloor\ \rfloor : \mathbb{R} &\longrightarrow \mathbb{Z} \\ \alpha &\longmapsto \max\{n \in \mathbb{Z}, n \leqslant \alpha\} \end{aligned}$$

Autrement dit, la partie entière de α est le seul entier $n \in \mathbb{Z}$ tel que $n \leqslant \alpha < n+1$.

Ce plus grand entier existe, car l'ensemble $\{\beta \in \mathbb{Z}, \beta \leqslant \alpha\}$ possède une borne supérieure. La figure 1.5 représente la fonction partie entière.

■ **Exemple 7** Quelques parties entières simples :

$$\lfloor 101,3 \rfloor = 101 \tag{1.6}$$
$$\lfloor 1,2 \rfloor = 1 \tag{1.7}$$
$$\lfloor 0,7 \rfloor = 0 \tag{1.8}$$
$$\lfloor -0,7 \rfloor = -1 \tag{1.9}$$
$$\lfloor -12,08 \rfloor = -13 \tag{1.10}$$

V **Vocabulary 1 — floor & ceil** ⇜⇝ En anglais, partie entière se dit *floor* et signifie plancher. Il existe également la fonction *ceil*, plafond en anglais, qui désigne le plus petit entier immédiatement supérieur ou égal à un nombre réel. On la note $\lceil a \rceil$. Ces deux formulations sont clairement plus concrètes que les appellations en français.

R Il ne faut pas confondre la fonction partie entière avec un arrondi ni avec une troncature à l'unité.

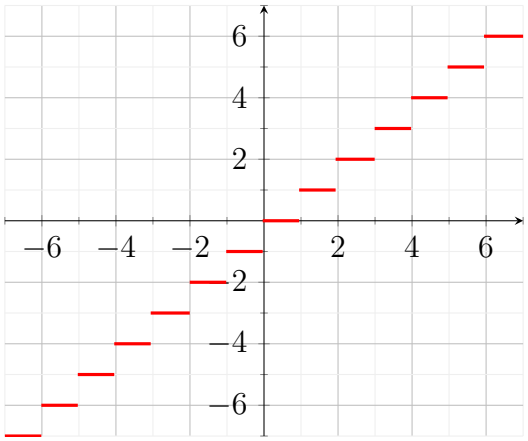


FIGURE 1.5 – Fonction partie entière

R Partie entière n'est pas injective. En effet, un élément de $\mathbb{Z}$ possède plusieurs antécédents dans $\mathbb{R}$ et même une infinité !

Chapitre 2

Diviseurs et multiples

2.1 Divisibilité

■ **Définition 13 — Divisibilité.** On dit que a divise n s'il existe un entier b tel que : $n = ab$.
$\forall a, n \in \mathbb{Z}, a \mid n \Leftrightarrow \exists b \in \mathbb{Z} : n = ab$

Si a *divise* n, on peut dire également que n est un *multiple* de a ou que n est *divisible* par a. Si ce n'est pas le cas, on note $a \nmid n$.

■ **Exemple 8** $21|63$ car $63 = 3 \times 21$. De même, $-21|63$ car $63 = -3 \times -21$.

Théorème 2 — Règles de divisibilité. Soit a, b, c, d et e des éléments de $\mathbb{Z}$.

1. $a \mid b$ et $b \mid c \Rightarrow a \mid c$.
2. $a \mid b \Rightarrow \forall c, ac \mid bc$.
3. $c \mid a$ et $c \mid b \Rightarrow \forall (d, e) \in \mathbb{Z}^2, c \mid ad + be$.
4. $a \mid b$ et $b \neq 0 \Rightarrow \mid a \mid \leqslant \mid b \mid$.
5. $a \mid b$ et $b \mid a \Rightarrow \mid a \mid = \mid b \mid$.

Démonstration. Pour démontrer ces règles, on revient à la définition 13.

1. si $a \mid b$ et $b|c$ alors il existe $k \in \mathbb{Z}$ et $p \in \mathbb{Z}$ tels que $b = ka$ et $c = pb$. On en déduit que $c = pka$ ce qui signifie que $a|c$.
2. si $a \mid b$ alors il exist $k \in \mathbb{Z}$ tel que $b = ka$. On peut multiplier les côtés de cette égalité par c et on obtient $(bc) = k(ac)$, ce qui signifie que $ac \mid bc$.
3. Si $c \mid a$ et $c \mid b$, alors il existe $(k, p) \in \mathbb{Z}^2$ tels que $a = kc$ et $b = pc$. Soient deux entiers relatifs d et e. Alors, $ad = kdc$ et $be = pec$ et $ad + be = (kd + pe)c$. Ce qui signifie que $c \mid ad + be$.
4. et ainsi de suite... Un bon exercice consiste à démontrer les deux derniers points.

■

2.2 Divisibilité et programmation

2.2.1 Opérateur modulo % et divisibilité

En informatique, l'opérateur qui permet de tester la divisibilité est l'opérateur modulo souvent noté %. Si le résultat le l'opération a%b est nulle, alors on a trouvé un facteur du nombre a : on peut dire que b divise a.

R Le résultat de l'opérateur modulo sur a et b, a%b, peut ne pas être égal au reste de la division euclidienne. Cela peut être le cas lorsque le diviseur est négatif. Dans ce cas, la plupart des langages choisissent la convention de prendre le reste du même signe que le diviseur. Le reste peut alors être négatif ne garantissant plus l'unicité de la division euclidienne (cf. 6).

2.2.2 Opérateur division (partie) entière //

L'opérateur `//` permet de réaliser une division de nombre entiers. Si a et b sont des entiers, le résultat de l'opération $a//b$ est un entier et peut être exprimé mathématiquement ainsi : $\lfloor \frac{a}{b} \rfloor$. D'un point de vue logiciel, en langage Python, on a : a//b = math.floor(a/b).

2.3 Division euclidienne

La division euclidienne est un des outils mathématiques parmi les plus accessibles et en même temps, comme on le verra dans les parties suivantes, d'une puissance phénoménale. L'intérêt émerge d'une situation concrète de la vie de tout les jours : comment répartir au mieux (le plus équitablement possible) 23 bonbons entre 5 enfants ?

Le principe de la division euclidienne était déjà connu dans l'antiquité par les babyloniens, les égyptiens, les chinois et bien sûr les grecs. Euclide l'explique par soustractions successives dans les Éléments [**euclides_euclids_2008**] en -300 avant notre ère.

Intuitivement, cette notion correspond à l'action de partager un ensemble en parts équitables (quotient) tout en laissant de côté ce que l'on ne peut pas distribuer équitablement (reste) (cf. figure 2.1 et 2.2).

FIGURE 2.1 – Illustration géométrique de la division euclidienne : le partage en parts égales avec reste éventuel.

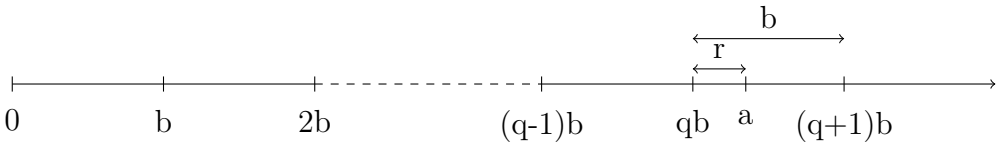


FIGURE 2.2 – Illustration géométrique de la division euclidienne.

Théorème 3 — Division euclidienne . Soient $a \in \mathbb{Z}$ et $b \in \mathbb{N}^\star$. Alors il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que les deux critères suivants sont vérifiés :

$$\begin{cases} a = bq + r \\ 0 \leqslant r < b \end{cases}$$

Démonstration. 1. Existence : a et b étant donné, on pose $q = \lfloor \frac{a}{b} \rfloor$. Par définition (cf. 12), on a : $0 \leqslant \frac{a}{b} - \lfloor \frac{a}{b} \rfloor < 1$. En multipliant par b, on obtient : $0 \leqslant a - b \times \lfloor \frac{a}{b} \rfloor < b$. En choisissant donc $q = \lfloor \frac{a}{b} \rfloor$ et $r = a - b \times \lfloor \frac{a}{b} \rfloor$, on a bien :

$$\begin{cases} a = bq + r \\ 0 \leqslant r < b \end{cases}$$

2. Unicité : supposons que l'on ait deux couples (q, r) et (q', r') appartenant à $\mathbb{Z} \times \mathbb{N}$: $a = bq + r = bq' + r'$ avec $0 \leqslant r < b$ et $0 \leqslant r' < b$. Cela peut également s'écrire : $b(q' - q) = r - r'$. Or, on a l'encadrement $-b < r - r' < b$. On en conclut que $-b < b(q' - q) < b$ et donc que $-1 < q' - q < 1$. Mais q et q' sont des entiers d'après nos hypothèses de départ. Donc, on en déduit de $q' - q = 0$. Il s'en suit que $q = q'$ et que $r = r'$. Il s'agit donc bien du même couple.

■

R Dans le cas où l'on se limite à des entiers naturels, on peut faire la démonstration de la division euclidienne par récurrence sur a. On initialise pour $a = 0$ en vérifiant que $q = r = 0$ conviennent. L'hypothèse de récurrence ($a = bq + r$ et $0 \leqslant r < b$) nous conduit à étudier $a + 1 = bq + r + 1$ et à montrer que l'on peut trouver un couple qui convient dans ce cas pour la division euclidienne de $a + 1$ par b. Si $r + 1 < b$ alors $q' = q$ et $r' = r + 1$ conviennent. Sinon, $q' = q + 1$ et $r' = 0$ conviennent.

R On peut étendre la définition de la division euclidienne aux entiers relatifs : on peut choisir $(a, b) \in \mathbb{Z} \times \mathbb{Z}^\star$. Dans ce cas, on a un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que les deux critères suivants sont vérifiés :

$$\begin{cases} a = bq + r \\ 0 \leqslant r <\mid b \mid \end{cases}$$

Par exemple : $23 = (-5) \times (-4) + 3$.

R On peut également chercher à étendre la définition à un reste négatif et avoir alors $0 \leqslant\mid r \mid <\mid b \mid$. Par exemple : $-23 = (-5) \times (4) - 3$ et on a bien $3 < 5$. **Néanmoins, dans ce cas, on n'a plus l'unicité** car $-23 = (-5) \times 5 + 2$ et $2 < 5$.

R C'est pourquoi, en programmation informatique, selon la convention choisie pour l'extension de la division euclidienne avec restes négatifs, les résultats des opérations divisions entières et modulos peuvent différer pour des opérations dans $\mathbb{Z}$. Généralement, le reste est choisi du même signe que le diviseur.

Chapitre 3

Numération et représentation

3.1 Formalisation de la numération de position

■ **Définition 14 — Numération de position.** La numération de position est un principe de notation selon lequel la signification d'un chiffre dépend de sa position dans le nombre. Dans un tel système, chaque chiffre se voit affecter un «poids» dans un nombre, poids qui est un facteur multiplicatif et qui dépend de la position du chiffre dans ce nombre.

■ **Définition 15 — Base d'un système de numération de position.** Soit g un entier naturel fixé supérieur à 1. g est la base d'un système de numération de position, si, lors de l'écriture d'un nombre, une unité de chaque ordre vaut g unités de l'ordre précédent.

Même si le principe énoncé dans la définition 14 peut apparaître trivial, l'écriture des nombres n'a pas toujours suivi ce principe[1]. Aujourd'hui, nous utilisons la numération de position quelle que soit la base et les conventions prises sont la plupart du temps les suivantes :

- les positions des chiffres se comptent de droite à gauche,
- le chiffre le plus à droite représente les unités et possède l'indice 0,
- le nombre formé par 10 représente exactement la valeur de la base quelle que soit cette base, i.e. la base d'un système de numération se note toujours 10 dans cette base,
- si on ajoute un zéro à droite d'un nombre, cela revient à multiplier ce nombre par sa base.

Dans un nombre écrit avec le système de numération de position, un chiffre c en position p ne vaut pas la même chose qu'à la position $p-1$. Si g est la base de ce système, le chiffre c en position p vaut g fois plus que s'il était placé en position $p-1$.

La division euclidienne (cf. théorème 3) est une décomposition unique d'un nombre. C'est pourquoi, on montre dans la section suivante qu'un entier naturel a peut s'écrire d'une manière unique dans un système de numération de position en base g à l'aide d'un nombre minimal de chiffres n. Les coefficients $a_i \in \{0, \ldots, g-1\}$ sont des entiers naturels strictement inférieurs à g (cf. théorème 4). On a alors :

1. Le chiffres romains par exemple.

$$a = a_{n-1}g^{n-1} + \cdots + a_2 g^2 + a_1 g + a_0 = \sum_{k=0}^{n-1} a_k g^k. \tag{3.1}$$

La numération de position revient à représenter le nombre en écrivant seulement les coefficients de ce polynôme, en omettant la plupart du temps la base et en notant tous les coefficients nuls ou non, de manière à ce que leur place soit définie sans ambiguïté. On désigne alors un nombre a, qui possède n chiffres, par un n-uplet en séparant ou non les éléments du n-uplet :

$$a = (a_{n-1}, \ldots, a_1, a_0) = (a_{n-1} \ldots a_1 a_0)_g = a_{n-1} \ldots a_1 a_0 \tag{3.2}$$

■ **Exemple 9** $2021_{10} = 2 \times 10^3 + 0 \times 10^2 + 2 \times 10^1 + 1 \times 10^0 = 2021$

■ **Exemple 10** $2021_3 = 2 \times 3^3 + 0 \times 3^2 + 2 \times 3^1 + 1 \times 3^0 = 61_{10}$ et ne se prononce pas «deux mille vingt et un»...

■ **Exemple 11** $2021_{64} = 2 \times 64^3 + 0 \times 64^2 + 2 \times 64^1 + 1 \times 64^0 = 524417_{10}$

3.2 Écriture d'un entier dans base quelconque

Théorème 4 — Décomposition d'un entier en base g . Soit $g \in \mathbb{N} \setminus \{0, 1\}$. Pour tout $a \in \mathbb{N}$, il existe $n \in \mathbb{N}$ et un n-uplet unique de chiffres $(a_0, a_1, \ldots a_{n-1}) \in [\![0, g-1]\!]^n$ tels que :

$$a = \sum_{k=0}^{n-1} a_k g^k. \tag{3.3}$$

De plus, si $a \in \mathbb{N}^\star$, on peut calculer le nombre de chiffres nécessaires pour représenter un nombre dans la base g :

$$n \leqslant \lfloor \log_g a \rfloor + 1 \tag{3.4}$$

Démonstration. 1. Unicité : on suppose qu'un développement tel que 3.3 existe. Alors, on peut écrire, en regroupant les puissances non nulles de g :

$$a = gA_1 + a_0$$

avec

$$A_1 = a_{n-1}g^{n-2} + \cdots + a_1$$

Ainsi, a_0 peut être vu comme le reste de la division euclidienne de a par g. Celle-ci étant unique, le coefficients a_0 et A_1 sont bien déterminés de manière unique. En considérant les termes $A_k = a_{n-k}g^{n-k-1} + \cdots + a_k$, on trouve de même que a_k et A_{k+1} sont les restes et quotients de la division euclidienne de A_k par a. Par une récurrence immédiate, on montre ainsi que les a_k sont uniques.

2. Existence : on procède en construisant la solution, c'est à dire les coefficients entiers. On les choisit comme suit :

$$c_k = \lfloor \frac{a}{g^k} \rfloor - g \lfloor \frac{a}{g^{k+1}} \rfloor$$

D'après la définition 12 de la partie entière,

$$\frac{a}{g^k} - 1 < \lfloor \frac{a}{g^k} \rfloor \leqslant \frac{a}{g^k}$$

et

$$-\frac{a}{g^{k+1}} \leqslant -\lfloor \frac{a}{g^{k+1}} \rfloor < -\frac{a}{g^{k+1}} + 1$$

En multipliant la dernière inéquation par g et en l'additionnant première, on obtient que :

$$-1 < c_k < g$$

Comme c_k est un entier, il appartient donc à l'ensemble $\{0, \ldots, g-1\}$. Ces coefficients sont nuls à partir d'un certain rang. En effet, si $k \geqslant \lfloor \log_g a \rfloor + 1$, alors $k > \lfloor \log_g a \rfloor$ et

$$\frac{a}{g^k} < \frac{a}{g^{\lfloor \log_g a \rfloor}} = \frac{a}{a} = 1$$

.

Ceci signifie que $\lfloor \frac{a}{g^k} \rfloor = 0$ quelque soit $k > \lfloor \log_g a \rfloor$.

Enfin, en notant $m = 1 + \lfloor \log_g a \rfloor$ on obtient par télescopage :

$$\begin{aligned}
\sum_{k=0}^{m-1} c_k g^k =& \lfloor \frac{a}{g^0} \rfloor - g\lfloor \frac{a}{g^1} \rfloor + (\lfloor \frac{a}{g^1} \rfloor - g\lfloor \frac{a}{g^2} \rfloor)g \\
&+ (\lfloor \frac{a}{g^2} \rfloor - g\lfloor \frac{a}{g^3} \rfloor)g^2 + (\lfloor \frac{a}{g^3} \rfloor - g\lfloor \frac{a}{g^4} \rfloor)g^3 \\
&+ \cdots \\
&+ (\lfloor \frac{a}{g^{m-2}} \rfloor - g\lfloor \frac{a}{g^{m-1}} \rfloor)g^{m-2} + (\lfloor \frac{a}{g^{m-1}} \rfloor - g\lfloor \frac{a}{g^m} \rfloor)g^{m-1} \\
=& \lfloor \frac{a}{g^0} \rfloor - \lfloor \frac{a}{g^m} \rfloor g^m \\
=& a
\end{aligned}$$

car $m > \lfloor \log_g a \rfloor$ et $\lfloor \frac{a}{g^m} \rfloor = 0$. Les coefficients c_k conviennent donc bien pour les a_k du théorème.

■

R On peut définir un système de numération unaire, c'est à dire avec les seuls symboles 0 et 1, mais ce n'est pas l'objet ici. C'est pourquoi on a restreint $g \in \mathbb{N} \setminus \{0, 1\}$.

Proposition 4 Si $a \in \mathbb{N}$ s'écrit $a_{n-1} \ldots a_0$ dans la base g, alors $g^{n-1} \leqslant a < g^n$.

Démonstration. On le démontre par récurrence. Pour $n = 1$, $a = a_0 < g$ d'après le théorème 4. Supposons la proposition vraie pour les nombres à n chiffres et prenons un nombre c à $n+1$ chiffres. Alors, on peut écrire $c = c_n g^n + d$, c'est à dire le nième chiffre multiplié par son poids dans la base plus un nombre $d = d_{n-1} \ldots d_0$ à n chiffres en base g. On applique l'hypothèse de récurrence au nombre d. Avec l'inégalité de droite, on obtient :

$$c < c_n g^n + g^n = g^n(c_n + 1) \leqslant g^{n+1}$$

puisque $c_n \in \{0, \ldots, g-1\}$.

Avec l'inégalité de gauche, on obtient :

$$c \geqslant c_n g^n + g^{n-1} \geqslant c_n g^n \geqslant g^n$$

Donc,

$$g^n \leqslant c < g^{n+1}$$

Donc la proposition est vraie pour tout entier de n chiffres. ■

■ **Exemple 12 — en base 2.** $1111_2 < 2^4$

■ **Exemple 13 — en base 10.** $999 < 10^3$

Théorème 5 — Moins on a de chiffres, plus on est petit . Soient deux entiers écrits dans une même base g et dont le nombre de chiffres est différent, alors le plus petit est celui dont l'écriture possède le moins de chiffres.

Démonstration. Soient a et b deux entiers écrits respectivement $a_{n-1} \dots a_0$ et $b_{m-1} \dots b_0$ dans la base g et que $a \neq b$. Supposons que $n < m$ alors $n \leqslant m - 1$ et puisque $1 < g$, la proposition 4 implique que $a < g^n \leqslant g^{m-1} \leqslant b$. Donc $a < b$. ■

3.3 Changement de base

Pour convertir en base 10 un nombre entier quelconque, il suffit d'appliquer la formule 3.1 comme dans l'exemple 10.

Pour faire l'opération inverse, c'est à dire convertir un nombre entier en base 10 vers une base quelconque, il faut remarquer que si $a = (a_{n-1} \dots a_1 a_0)_g$ alors le quotient de la division euclidienne de a par g est égal à $q = (a_{n-1} \dots a_1)_g$ et le reste à $r = a_0$ puisque $a = gq + r$ avec $0 \leqslant r \leqslant g - 1$.

■ **Exemple 14 — Écrire 61_{10} en base 3.**

$$61_{10} = 3 \times 20 + \mathbf{1} \tag{3.5}$$
$$20 = 3 \times 6 + \mathbf{2} \tag{3.6}$$
$$6 = 3 \times 2 + \mathbf{0} \tag{3.7}$$
$$2 = 3 \times 0 + \mathbf{2} \tag{3.8}$$

C'est pourquoi, $61_{10} = 2021_3 = 2 \times 3^3 + 0 \times 3^2 + 2 \times 3^1 + 1 \times 3^0$

R En langage Python, on peut directement écrire du binaire en préfixant le nombre par **0b** (0b00011) ou de l'hexadécimal en préfixant par **0x** (0xF4E). Les fonctions bin, oct et hex permettent de convertir directement des nombres en binaire, octal et hexédacimal. Inversement, l'instruction int('2021', 3) permet de convertir de la base 3 en décimal (on trouve 61).

3.4 Représentation des nombres entiers

3.4.1 Encoder un entier naturel

La base 2 n'utilise que les chiffres 0 et 1, ce qui correspond également à la capacité de stocker une information selon deux états physiquement différents de la matière, chose bien maîtrisée en électronique. Sur p bits, on peut coder un nombre entier naturel $0 \leqslant a < 2^p$. C'est à dire que l'on peut utiliser la plage de nombre allant de 0 à 2^{p-1} exactement.

■ **Exemple 15** Sur 8 bits, le plus grand entier naturel que l'on peut inscrire est $2^8 - 1 = 255$.

3.4.2 Encoder un entier relatif

Si l'on cherche à représenter un entier relatif et qu'on adopte une approche naïve, on peut imaginer utiliser un bit pour désigner le signe de ce nombre. Par exemple, 0 pour le signe positif et 1 pour le signe négatif. S'en suivrait alors la représentation binaire de la valeur absolue du nombre $\mid a \mid$. Cette méthode possède néanmoins de nombreux inconvénients :

1. le nombre 0 pourrait être représenté par deux symboles différents, positif ou négatif, ce qui n'est pas souhaitable. Il est toujours préférable d'avoir une unicité lors de la représentation d'un objet.
2. l'algorithme de l'addition ne s'appliquerait qu'à des entiers de même signe et on devrait donc utiliser une autre algorithme, et donc un autre circuit électronique, pour les entiers relatifs. Or il est intéressant de pouvoir appliquer toujours le même algorithme quelle que soit la donnée, c'est à dire d'utiliser les mêmes circuits électroniques.

On utilise donc un encodage particulier pour représenter les entiers négatifs, encodage dit **complément à deux**.

■ **Définition 16 — Complément à deux.** Pour représenter les entiers relatifs en mémoire sur n bits, on encode le nombre **négatif** a par le nombre positif $2^n + a$. Le premier chiffre est donc toujours 1 si le nombre est négatif. Cela limite la représentation des nombres à la plage $-2^{n-1} \leqslant a \leqslant 2^{n-1} - 1$.

■ **Exemple 16 — Encodage sur 8 bits en complément à 2.** Sur 8 bits, on peut donc normalement inscrire les entiers relatifs allant de $-128 \rightarrow 11111111$ à $127 \rightarrow 01111111$. Pour écrire le nombre -67, on calcule $2^8 - 67 = 189$ et on écrit alors 10111101.

L'encodage complément à deux permet de donner une représentation **unique** sur n bits à tout nombre entier relatif appartenant à $[\![-2^{n-1}, 2^{n-1} - 1]\!]$. Ces encodages sont dits **signé** ou **signed** en anglais. Ils sont déclinés pour des nombres de bits allant de 8 à 64 sur la plupart des architectures (cf. tableau 3.1). Ils permettent également de réaliser des opérations d'addition, de soustraction, de multiplication et de division facilement.

■ **Exemple 17 — Addition en complément à deux.** Calculons $113 + (-91)$ comme le ferait un ordinateur. Ces deux nombres sont encodables sur 8 bits, car compris dans l'intervalle $[\![-2^7, 2^7 - 1]\!] = [\![-128, 127]\!]$.

$113 = (01110001)_2$ et $-91 = (10100101)_2$. On additionne de manière classique en

Nombre de bits	Intervalle accessible (signé)	Intervalle accessible (non signé)
8	$[\![-128, 127]\!]$	$[\![0, 255]\!]$
16	$[\![-32768, 32767]\!]$	$[\![0, 65535]\!]$
32	$[\![-2147483648, 2147483647]\!]$	$[\![0, 4294967295]\!]$
n	$[\![-2^{n-1}, 2^{n-1} - 1]\!]$	$[\![0, 2^n - 1]\!]$

TABLE 3.1 – Plage d'entiers accessibles en fonction du nombre de bits de la représentation signée.

tronquant le résultat à 8 bits et on obtient : $133 + (-91) = (00010110)_2 = 22_{10}$

■ **Exemple 18 — Opposé en complément à deux.** Si $a > 0$, alors l'opposé de a sera encodé $2^n - a$. Si $a < 0$, alors l'opposé de a sera encodé $2^n + a$.

■ **Définition 17 — Dépassement de capacité.** Lorsque le résultat d'une opération sort de l'intervalle de représentation possible, c'est à dire $a \otimes b \notin [\![-2^{n-1}, 2^{n-1} - 1]\!]$, alors le résultat n'est plus valide et on dit qu'on a dépassé les capacités de stockage du système.

■ **Exemple 19 — Addition et dépassement.** Si on utilise un encodage signé de ces entiers sur 8 bits pour effectuer $72+59$ alors il advient un dépassement de capacité. Le résultat obtenu sur 8 bits est -125 et n'est pas valide. Les processeurs savent détecter ces situations.

3.4.3 Calculs à virgule fixe

La représentation des nombres décrite dans le paragraphe précédent peut servir à encoder des nombres rationnels dont la précision est limitée, des décimaux ou des dyadiques par exemple. On appelle cette représentation à virgule fixe.

V **Vocabulary 2 — Fixed-Point Arithmetics** ⇝ En anglais, ce mode de calcul s'énonce Fixed-Point Arithmetics, étant donné l'utilisation du point à la place de la virgule par les anglo-saxons.

■ **Définition 18 — Représentation à virgule fixe .** Il s'agit de représenter un nombre par un entier divisé par un facteur d'échelle choisi, c'est à dire $\frac{a}{b^n}, a \in \mathbb{Z} b \in \mathbb{N}, n \in \mathbb{N}$.

On affecte un certain nombre de bits à la **partie entière** et à la **partie fractionnaire** de ce nombre.

La notation standard est la suivante : `fixed<m,n>` est le type d'un nombre à virgule fixe codé sur m bits pour la partie entière et sur n bits pour la partie fractionnaire. Le facteur d'échelle vaut donc 2^n.

La base de numération est deux et le facteur d'échelle est une puissance de deux. On peut donc représenter à virgule fixe un sous-ensemble des nombres dyadiques $\mathcal{D}$ (cf. définition 5).

■ **Exemple 20 — 7,5 est dyadique .** En base dix, le nombre décimal (cf. définition 4)

$7,5$ peut être représenté à virgule fixe par le nombre entier 75 et le facteur d'échelle 10, car $75 = \frac{75}{10}$.

Pour coder $7,5$ en binaire, on observe que $7,5 = 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} = 4 + 2 + 1 + 0,5 = 7,5$. C'est un nombre dyadique (cf. définition 5). Pour trouver la partie fractionnaire, on procède comme pour la partie entière, par divisions successives : la partie entière du résultat contient le bit recherché. Dans notre cas : $\lfloor 0,5/2^{-1} \rfloor = \lfloor 0,5 \times 2 \rfloor = 1$. Tous les autres bits sont nuls.

On peut alors choisir de représenter $7,5$ avec une type `fixed<3,1>` et l'écrire : $111,1_2$. En machine, il peut être stocké sur 4 bits $1111_2 = 15_{10}$ et le facteur d'échelle vaut 2. On vérifie que $15/2 = 7,5$.

■ **Exemple 21 — $3,14$ est décimal mais pas dyadique...** π est un nombre transcendant. Mais la valeur approchée $3,14$ en base dix est un nombre décimal. Il peut être représenté à virgule fixe par le nombre entier 314 et le facteur d'échelle 100, car $3,14 = \frac{314}{100}$.

$3,14$ n'est pas un nombre dyadique. Pour coder exactement la valeur $0,14$ en binaire, il faudrait une infinité de bits [a] : $0,14_{10} = 0,00\overline{100011110101111000011}_2$. Cependant, on remarque que $3,14 \simeq 1\times2^1+1\times2^0+1\times2^{-3}+1\times2^{-6} = 3+0,125+0,015625 = 3,140625$. Il s'agit là d'une valeur approchée.

On peut alors choisir de représenter $3,14$ en valeur approchée avec un type `fixed<2,6>` et l'écrire : $11,001001_2$. Le facteur d'échelle vaut 2^6. Il sera donc codé en machine 11001001. On vérifie que $11001001_2/1000000_2 = 201/64 = 3,140625$

a. Les chiffres surmontés d'une barre horizontale constituent le motif répété à l'infini.

R Tout comme en base 10, certains nombres ne peuvent pas être codés en base 2 par un nombre fini de bits. C'est le cas par exemple de $0,1_{10} = 0,0\overline{0011001100110011001}1_2$.

Une telle représentation des nombres permet d'utiliser les circuits dédiés à l'arithmétique des nombres entiers pour effectuer des calculs sur des nombres codés en virgule fixe [2]. On n'a donc pas besoin de modifier les architectures des processeurs qui sont capables de calculer sur les entiers pour calculer en arithmétique à virgule fixe.

R Les nombres rationnels ont une grande importance dans la suite de cet ouvrage. On peut montrer qu'un nombre rationnel est un nombre :

- à représentation décimale répétitive. Par exemple, $1/3 = 0,333\ldots = 0.\overline{3}$ ou $1/7 = 0,142857142857\ldots = 0,\overline{142857}$,
- ou à représentation décimale terminale, lorsque la répétition est 0. Par exemple, $1/2 = 0,5000\ldots = 0,4999\ldots = 0,5$.

Il existe, en base 10, deux représentations équivalentes décimales terminales : soit avec des 0 soit avec des 9 (cf. proposition 5). C'est pourquoi, il faut être attentif à l'interprétation des résultats lorsque ces différentes représentations interviennent.

Proposition 5 — $1 = 0,9999\ldots$.

2. On code souvent une valeur approchée du nombre décimal ou dyadique.

Démonstration. On utilise le résultat de la somme des termes d'une suite géométrique.

$$
\begin{aligned}
0,999\ldots &= \frac{9}{10} + \frac{9}{100} + \frac{9}{1000} + \ldots && (3.9)\\
&= \frac{9}{10} + \frac{9}{10^2} + \frac{9}{10^3} + \ldots && (3.10)\\
&= 9\left(\frac{1}{10} + \frac{1}{10^2} + \frac{1}{10^3} + \ldots\right) && (3.11)\\
&= 9\left(-1 + \sum_{k=0}^{+\infty} \frac{1}{10^k}\right) && (3.12)\\
&= 9\left(-1 + \frac{1}{1-\frac{1}{10}}\right) && (3.13)\\
&= 9\left(-1 + \frac{10}{9}\right) && (3.14)\\
&= 9.\frac{1}{9} = 1 && (3.15)
\end{aligned}
$$

■

Arrondis et erreurs

Il nous faut pouvoir estimer précisément les erreurs que l'on commet lorsqu'on choisit d'encoder en machine un nombre selon un format donné. Les définitions d'un arrondi et des erreurs absolues et relatives vont nous y aider.

■ **Définition 19 — Arrondir un nombre en base 10.** La plupart du temps, on choisit d'arrondir un nombre réel a à 10^{-n} de la manière suivante :

$$arrondi(a,n) = sgn(a)\frac{\lfloor \mid a \times 10^n \mid +0,5 \rfloor}{10^n} \tag{3.16}$$

■ **Exemple 22 — Arrondir à 10^{-2}.** $5,456$ sera arrondi à 10^{-2} par la formule 3.16 à $5,46$.

R Plusieurs modes d'arrondi sont définis pour le binaire dans la norme IEEE 754. Ils sont décrits à la section 3.5.3.

■ **Définition 20 — Erreur d'approximation absolue.** Soit ν un nombre réel et $\tilde{\nu}$ sa valeur approchée.

$$\varepsilon_a = \tilde{\nu} - \nu \tag{3.17}$$

■ **Définition 21 — Erreur d'approximation relative.** Soit ν un nombre réel et $\tilde{\nu}$ sa valeur approchée.

$$\varepsilon_r = \frac{\tilde{\nu} - \nu}{\mid \nu \mid} \tag{3.18}$$

Erreurs d'approximation

L'inconvénient de l'arithmétique à virgule fixe est principalement le lien entre la précision obtenue et l'ordre de grandeur des entiers représentés.

Proposition 6 — Majorant de l'erreur absolue. L'erreur absolue d'un nombre ν et sa valeur approchée $\tilde{\nu}$ encodé par un `fixed<m,n>` est majorée par une constante. On a :

$$\mid \varepsilon_a \mid = \mid \tilde{\nu} - \nu \mid < 2^{-n-1} \tag{3.19}$$

Démonstration. Soit a une nombre codé en `fixed<m,n>`. Le nombre b directement supérieur que l'on peut coder est obtenu en ajoutant 2^{-n}, c'est à dire en ajoutant 1 au bit de poids faible. $b - a = 2^{-n}$: cette différence est la plus petite que l'on puisse coder. Au pire, le nombre réel ν à encoder se situe au milieu de l'intervalle $[a, b]$ et donc l'erreur commise sera la moitié de la largeur de cet intervalle : $\frac{1}{2}(b-a) = 2^{-n-1}$. Cet valeur est un majorant de l'erreur absolue. ■

■ **Exemple 23 — Erreur absolue, ordre de grandeur et virgule fixe.** Imaginons que l'on utilise des entiers codés à virgule fixe par un type `fixed<6,3>` non signé. La plage des valeurs que l'on peut atteindre va de 0 à $63,875_{10} = 111111,111_2$. On peut tout d'abord noter une limitation : on ne pourra coder des nombres supérieurs à $63,875$.

De plus, entre chaque nombre ainsi représenté, l'incrément minimal est de $1/2^3 = 0,125$. L'erreur absolue est majorée par la constante 2^{-n-1}, dans notre cas $0,0625$.

Proposition 7 — Majorant de l'erreur relative. L'erreur relative entre un nombre réel ν et sa valeur approchée $\tilde{\nu}$ encodée par un `fixed<m,n>` est majorée par un terme dépendant de ν. On a :

$$\mid \varepsilon_r \mid = \frac{\mid \tilde{\nu} - \nu \mid}{\mid \nu \mid} \leqslant \frac{2^{-n-1}}{\mid \nu \mid} \tag{3.20}$$

Démonstration. Ce résultat est une conséquence de la proposition précédente. ■

■ **Exemple 24 — Erreur relative en `fixed<5,3>` .** L'erreur relative pour des nombres codés en `fixed<5,3>` non signé est beaucoup plus faible pour les valeurs élevées du nombre.

Sur l'intervalle $[1, 1,125]$, tout nombre supérieur ou égal à $1,0625$ va être arrondi à $1,125$, tout nombre strictement inférieur à 1. L'erreur relative en arrondissant à 1 est majorée par celle commise pour le nombre $1,0625$. Elle vaut $(1 - 1,0625)/1,0625 \simeq -0,059$. Celle commise en arrondissant à $1,125$ est majorée par celle commise pour $1,0625$. Elle vaut $(1,125 - 1.0625)/1,0625 \simeq 0,059$.

Ce résultat est conforme à la proposition précédente et peut s'écrire :

$$\max \mid \epsilon_r(a \in [1, 1,125]) \mid < 0,059 \simeq \frac{2^{-4}}{1,125}$$

En procédant de même avec l'intervalle $[63, 63,125]$, on trouve que :

$$\max \mid \epsilon_r(a \in [63, 63,125]) \mid < 0,001 \simeq \frac{2^{-4}}{63}$$

Cela signifie que l'erreur relative d'arrondi commise sur un nombre a compris entre $[63, 63,125]$ peut être jusqu'à soixante trois fois plus faible que pour un nombre a appartenant à l'intervalle $[1, 1,125]$.

Même si l'arithmétique à virgule fixe est très utilisée dans le calcul embarqué, elle nécessite un savoir faire particulier, car il faut gérer les dépassements correctement. En

outre, une erreur relative variable sur la plage de données manipulée par un algorithme peut être rédhibitoire pour certaines applications. Heureusement, les nombres flottants permettent de dépasser ces limites.

3.5 Représentation à virgule flottante

3.5.1 Cas général

La représentation d'un nombre à virgule flottante ne fixe pas un nombre exact de bits alloués à la partie fractionnaire. Au contraire, la position de la virgule dépend d'un exposant. IEEE 754 est la norme utilisée pour cette représentation par la plupart des systèmes.

■ **Définition 22 — Notation scientifique.** Exprimer un nombre a selon la notation scientifique c'est l'écrire sous la forme :

$$a = \pm m \times 10^e \tag{3.21}$$

où $m \in [1, 10[\subset \mathbb{D}$ est un nombre décimal nommé *mantisse* et $e \in \mathbb{Z}$ l'exposant.

■ **Définition 23 — Représentation normalisé IEEE 754 d'un nombre binaire.** Pour représenter un nombre binaire en utilisant norme IEEE 754, il est nécessaire d'écrire le nombre à représenter sous la forme : $\pm 1, M.2^e$.

- On a appelle M la pseudo-mantisse. Le 1 au début du nombre n'est pas codé en machine, il est implicite. On l'appelle le bit caché.
- $\pm$, le signe de la mantisse est codée par s qui vaut 0 ou 1.
- e est l'exposant qui va être codé par un exposant biaisé E. Le biais dépend du format choisi, simple ou double précision : c est codé sur $n_E = 8$ bits ou $n_E = 11$ bits. En simple précision, il vaut $2^{n_E-1} - 1 = 2^7 - 1 = 127$ En double précision, $2^{n_E-1} - 1 = 2^{10} - 1 = 1023$.

FIGURE 3.1 – Schématisation du format IEEE 754.

■ **Définition 24 — Nombre IEEE 754 normalisé.** Un nombre IEEE 754 est normalisé lorsque le bit caché de la mantisse est 1 et lorsque l'exposant biaisé appartient à $[\![1, 2^{n_E} - 2]\!]$, où n_E est le nombre de bits qui code l'exposant biaisé E.

Si le biais vaut b, cela signifie qu'un nombre normalisé correspond à un exposant $e \in [\![1-b, 2^{n_E}-2-b]\!]$. En simple précision, $b = 127$ et cela se traduit par $e \in [\![-126, 127]\!]$. En double précision, $b = 1023$ et cela se traduit par $e \in [\![-1022, 1023]\!]$.

La norme IEEE 754 définit plusieurs formats qui garantissent des précisions différentes. Dans tous les cas, s est codé sur un seul bit.

Simple précision - 32 bits M est codée sur 23 bits, E sur 8 bits.

Double précision - 64 bits M est codée sur 52 bits, E sur 11 bits.

■ **Exemple 25 — Représenter $39,125_{10}$ par un nombre flottant IEEE 754.** Tout d'abord, il est nécessaire de convertir le nombre en base 2, comme expliqué dans l'exemple 21. On obtient $39,125_{10} = 100111,001_2$. Une fois ce résultat obtenu, il est nécessaire d'écrire ce nombre sous la forme $\pm 1,M.2^e$. Cela donne, en binaire, $1,00111001.2^5$.

Dans cette notation, l'exposant n'est pas biaisé. Pour le représenter en machine, il faut donc le translater et la translation dépend du format choisi, simple ou double précision. Au format simple précision, E est codé sur 8 bits et on translate de 127. Donc on a $E = 127 + 5 = 132_{10} = 10000100_2$.

Le nombre étant positif, le bit de signe s vaut 0.

La représentation IEEE 754 simple précision de $39,125_{10}$ est donc 01000010000111001000000000000 comme le montre la figure 3.2.

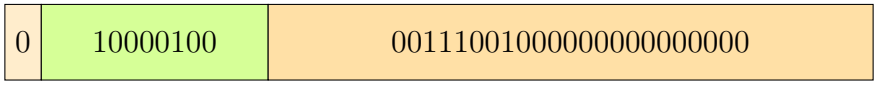


FIGURE 3.2 – Représentation IEEE 754 simple précision de $39,125_{10}$.

R L'intérêt de placer l'exposant avant la mantisse et de le biaisé est de faciliter la comparaison de deux nombres flottants. Si les exposants sont différents, il suffit de comparer ces valeurs positives pour connaître le nombre le plus grand.

R Pourquoi biaiser l'exposant ? Il semble que ce soit pour faire en sorte que l'opération d'inversion du plus petit et du plus grand normalisé ne produise ni dépassement ni sous-dépassement (cf. définitions 3.5.2).

En simple précision, le plus petit normalisé est 2^{-126}. L'inverse de ce nombre vaut 2^{126} : c'est un nombre normalisé. Le plus grand normalisé est $(2 - 2^{-23}).2^{127}$. Son inverse arrondi vaut 2^{-128} : ce nombre est dénormalisé mais il ne s'agit pas d'un sous-dépassement.

Un autre biais ne permettrait pas d'obtenir ce résultat.

3.5.2 Cas particuliers

■ **Définition 25 — Nombre IEEE 754 dénormalisé.** Si n_E est le nombre de bits qui code l'exposant biaisé E, un nombre IEEE 754 est dénormalisé lorsque :

$E = 0$ **:** Alors l'exposant e vaut **par convention** $-2^{n_E-1} + 2$ et le nombre représenté $0,M.2^{-2^{n_E-1}+2}$ (noter le 0 au début).

$E = 2^{n_E} - 1$ **:** la plus grande valeur possible de l'exposant biaisé. Le nombre représente alors des valeurs spéciales comme l'infini ou NaN.

R En simple précision, un nombre dénormalisé correspond à un exposant e non biaisé de -126 commençant par le bit 0 ou de 127.

R La convention prise pour la valeur de e dans le cas où $E = 0$ permet de limiter la distance entre le plus grand des dénormalisés et le plus petit des normalisé. Dans de cas de la simple précision, $0.11111111111111111111111.2^{-126}$ et $1.00000000000000000000000.2^{-126}$.

Il existe trois cas de représentation dénormalisée IEEE 754 binaire : zéro, l'infini et

NaN.

Représentation de zéro

Dans le cas où l'exposant biaisé et la mantisse sont tous les deux nuls, on considère alors que le nombre codé est zéro.

R Il y a donc deux zéros possibles dans la norme IEEE 754, un positif et un négatif, car le bit de signe peut valoir 1 ou 0. Cela est utile notamment pour évaluer les signes de résultats infinis ou dans le cas d'un dépassement de capacité négative (arithmetic underflow).

Représentation de l'infini

Dans le cas où l'exposant vaut $2^{n_E} - 1$ et où la mantisse est nulle, on considère que le nombre représente un infini. Selon le bit de signe, celui-ci peut être considéré comme infiniment négatif ou infiniment positif.

Ceci n'est pas un nombre (Not a Number, NaN)

Dans le cas où l'exposant vaut $2^{n_E} - 1$ et où la mantisse n'est pas nulle, on considère que le nombre représenté n'est pas un nombre.

NaN est une représentation nécessaire aux calculs car les résultats des opérations mathématiques ne sont pas toujours déterminés ou valides.

■ **Définition 26 — NaN silencieux - Quiet NaN.** Un NaN silencieux est émis lorsque le résultat de l'opération est indéterminé. Il est propagé dans le calcul.

■ **Définition 27 — NaN signalé - Signalling NaN.** Un NaN signalé est émis lorsque l'opération n'est pas valide. Il produit immédiatement une exception.

■ **Exemple 26 — QNaN et SNaN.** La division par zéro résulte en un SNaN.
L'opération $\pm\infty / \pm\infty$ résulte en un QNaN.

Dépassements

On peut distinguer plusieurs cas de dépassements liés à la norme IEEE 754. En simple précision,

1. les nombres négatifs inférieurs à $-(2 - 2^{-23}).2^{127}$ constituent un dépassement par valeurs négatives,
2. les nombres négatifs supérieurs à -2^{-149} constituent un sous-dépassement par valeurs négatives,
3. les nombres positifs inférieurs à 2^{-149} constituent un sous-dépassement par valeurs positives,
4. les nombres positifs supérieurs à $(2 - 2^{-23}).2^{127}$ constituent un dépassement par valeurs positives.

V **Vocabulary 3 — Overflow, underflow** ⇜⇝ En anglais on désigne par *overflow* les dépassements et *underflow* les sous-dépassements. On peut ainsi désigner un sous-dépassement par valeurs négatives par *negative underflow*.

R Les sous-dépassements n'engendrent qu'une perte de précision dont l'ordre de grandeur est très faible. Les dépassements sont plus problématiques car ils sont la manifestation d'une incapacité à représenter un nombre en flottant.

Valeurs accessibles aux flottants

Le tableau 3.2 précise les valeurs extrêmes accessibles à la norme IEEE 754. Le tableau 3.3 donne un aperçu de la plage de valeurs accessibles via des flottants en simple et double précision.

Précision	Normalisé	Dénormalisé
simple	$\pm 2^{-126}$ à $(2-2^{-23}).2^{127}$	$\pm 2^{-149}$ à $(1-2^{-23}).2^{-126}$
double	$\pm 2^{-1022}$ à $(2-2^{-52}).2^{1023}$	$\pm 2^{-1074}$ à $(1-2^{-52}).2^{-1022}$

TABLE 3.2 – Plus petit (dé)normalisé et plus grand (dé)normalisé.

Précision	Binaire	Décimal
simple	$[-(2-2^{-23}).2^{127}, (2-2^{-23}).2^{127}]$	$\sim [-10^{38,53}, 10^{38,53}]$
double	$[-(2-2^{-52}).2^{1023}, (2-2^{-52}).2^{1023}]$	$\sim [-10^{308,25}, 10^{308,25}]$

TABLE 3.3 – Plage de valeurs accessibles aux flottants.

Opérations spéciales

La norme IEEE 754 définit le résultat des opérations spéciales qui comportent des valeurs dénormalisées. Elles sont décrites dans le tableau 3.4 et servent à faire en sorte que les résultats classiques sur les limites des fonctions réelles soient respectés par les flottants.

Opération	Résultat
$n/\pm\infty$	0
$n \times \pm\infty$	$\pm\infty$
$\pm nz/0, nz \neq 0$	$\pm\infty$
$\pm 0/\pm 0$	NaN
$\infty + \infty$	∞
$\infty - \infty$	NaN
$\pm\infty \times \pm\infty$	$\pm\infty$
$\pm\infty/\pm\infty$	NaN
$\pm\infty \times 0$	NaN
NaN==NaN	Faux

TABLE 3.4 – Opérations spéciales sur les flottants.

Synthèse de l'écriture des flottants

Le tableau 3.5 résume l'écriture normée IEEE754 des flottants et associe à chaque représentation une valeur ou un sens particulier. La figure 3.3 montre les nombres flottants IEEE754 codables sur l'axe des réels.

Signe	**Exposant**	**Mantisse**	**Sens / valeur**
0	0	0	0
0	0	$M \neq 0$	Nombre positif dénormalisé
0	$0 < E < 2^{n_E} - 1$	M	Nombre positif normalisé
0	$2^{n_E} - 1$	0	$+\infty$
0	$2^{n_E} - 1$	$0 < M < 2^{n_M - 1} - 1$	SNaN
0	$2^{n_E} - 1$	$M \geqslant 2^{n_M - 1}$	QNaN
1	0	0	-0
1	0	$M \neq 0$	Nombre négatif dénormalisé
1	$0 < E < 2^{n_E} - 1$	M	Nombre négatif normalisé
1	$2^{n_E} - 1$	0	$-\infty$

TABLE 3.5 – Synthèse de l'écriture des flottants.

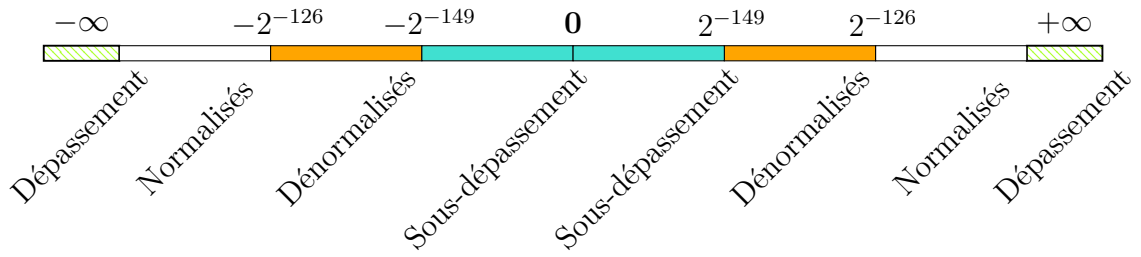


FIGURE 3.3 – Illustration de la répartition des nombres flottants IEEE754 en simple précision sur la droite des réels.

3.5.3 Erreur d'approximation

Erreur absolue

Si on cherche à majorer l'erreur absolue que l'on commet lors de la représentation binaire IEEE 754, il suffit de considérer un incrément de 1 sur le bit de poids faible par rapport à la valeur considérée : celui-ci dépend à la fois du nombre de bits avec lequel est codée la mantisse et de l'exposant.

On trouve alors :

$$\epsilon_a < 2^{-n_M} \times 2^e \tag{3.22}$$

si e est l'exposant non biaisé et n_M le nombre de bits qui code la pseudo-mantisse.

Pour l'exemple 25, l'erreur absolue est donc majorée : $\epsilon_a < 2^{-23} \times 2^5 = 2^{-18} \simeq 3,18.10^{-6}$.

Dans le cas du nombre :

$$500000,875_{10} = 1111010000100100000,111_2 = 1,11101000010010000011 1_2 \times 2^{18}$$

On a : $\epsilon_a < 2^{-23} \times 2^{18} = 2^{-5} \simeq 3,125.10^{-2}$.

L'erreur absolue varie donc selon la plage de valeurs considérée.

Erreur relative

L'intérêt de la représentation à virgule flottante est qu'elle garantit une erreur relative constante quelque soit l'ordre de grandeur des nombres représenté. Avec la norme IEEE 754, on peut représenter 2^{n_M} nombres entre chaque puissance de 2.

L'erreur relative entre chaque nombre flottant en simple précision est donc majorée par l'inverse de ce nombre. On a :

$$\epsilon_r < 2^{-23} \tag{3.23}$$

et ce, quelque soit l'exposant.

R $\epsilon_r < 2^{-23} \simeq \frac{10^{-6}}{8}$. Cela nous garantit 6 chiffres significatifs en base 10.

En double précision, $\epsilon_r < 2^{-52} \simeq \frac{10^{-15}}{4,5}$. IEEE754 nous garantit donc 15 chiffres significatifs en base 10.

Modes d'arrondi

La norme IEEE 754 définit plusieurs modes pour arrondir les résultats des calculs si ceux-ci ne sont pas exacts. Le résultat correct se trouve la plupart du temps entre deux valeurs représentable par la norme. Il faut en choisir un. On utilise alors une des méthodes suivantes :

arrondir au plus proche on choisit de prendre le résultat représentable le plus proche. Si le résultat correct est exactement au milieu de l'intervalle des représentables, on choisit le résultat dont la mantisse se termine par 0 et on parle d'arrondi pair. C'est le mode par défaut. On note qu'il est différent de celui adopté classiquement pour le calcul sur les décimaux (cf. définition 19).

arrondir au dessus on choisit le résultat représentable le plus grand, éventuellement infini ou zéro.

arrondir au dessous on choisit le résultat représentable le plus petit, éventuellement infini ou zéro.

arrondir en troncant on choisit le résultat représentable le plus proche de zéro dans tous les cas.

Le mode par défaut est arrondir au plus proche.

3.5.4 Calcul avec les flottants

■ **Exemple 27 — Addition en simple précision.** Considérons l'opération $(0,1+0,2)-0,3$ en simple précision.

Tout d'abord, il faut remarquer que ni $0,1$ ni $0,2$ ni $0,3$ ne sont des nombres dyadiques (cf définition 5). Cela signifie que leur représentation en machine ne peut être qu'une approximation. Le résultat d'une opération avec ces opérandes est donc

toujours une valeur approchée...

La valeur $0,1_{10}$ en binaire fait apparaître le premier 1 à la quatrième décimale. Donc l'exposant sera -4. De plus, $0,1_{10}$ fait l'objet d'un arrondi au plus prêt en binaire. En l'occurrence, c'est la valeur supérieure qui est plus proche car le milieu de l'intervalle $[0,1]$ est 0.1.

Pour aboutir à ce résultat, on utilise les trois bits GRS après le dernier bit de poids faible. **On nomme ces trois bits GRS pour Guard, Round, Sticky.** Dans ce cas, ils valent 110 et $110 > 100$. Par conséquent, en mode par défaut, on arrondit à la valeur supérieure et on obtient :

$$1.10011001100110011001100\ldots \qquad GRS \tag{3.24}$$
$$1.10011001100110011001100.2^{-4} \qquad 110 \tag{3.25}$$
$$0,1_{10} \mapsto 1.10011001100110011001101.2^{-4} \tag{3.26}$$

Il est de même pour la valeur $0,2_{10}$, mais avec un exposant valant -3 :

$$0,2_{10} \mapsto 1.10011001100110011001101.2^{-3} \tag{3.27}$$

L'addition de $0,1+0,2$ nécessite tout d'abord la mise à l'échelle de $0,1$ par rapport à $0,2$: pour les additionner, on les représente avec le même exposant en l'occurrence -3. On ne conserve que les trois GRS des bits de $0,1$ qui sont rejetés à droite du nouveau bit de poids faible.

$$1.10011001100110011001101.2^{-4} \qquad GRS \tag{3.28}$$
$$\mapsto 0.11001100110011001100110.2^{-3} \qquad 100 \tag{3.29}$$
$$\tag{3.30}$$

On peut alors procéder à l'addition :

$$GRS \tag{3.31}$$
$$0.11001100110011001100110.2^{-3} \qquad 100 \tag{3.32}$$
$$+1.10011001100110011001101.2^{-3} \qquad 000 \tag{3.33}$$
$$---------------------- \tag{3.34}$$
$$10.01100110011001100110011.2^{-3} \qquad 100 \tag{3.35}$$
$$\mapsto 1.00110011001100110011001.2^{-2} \qquad 110 \tag{3.36}$$
$$\mapsto 1.00110011001100110011010.2^{-2} \tag{3.37}$$

Le passage de 3.35 à 3.36 se justifie par le maintien de la norme : on choisit l'exposant de la notation IEEE754 d'après le premier bit à 1. Les bits décalés vers la droite se retrouve positionnés sur les bits GRS.

On effectue ensuite l'opération d'arrondi 3.37 en tenant compte des bits GRS. Le résultat obtenu se termine par 01|110 (en notant $b_1b_0|GRS$). On essaie donc de savoir comment se positionne cette valeur par rapport au milieu de l'intervalle de représentation possible qui est $[01,10]$. Ce milieu vaut 1.100. Le résultat obtenu 1.110 étant supérieur à cette valeur, on arrondit au plus près au nombre strictement supérieur 10. D'où le résultat.

La valeur 0.3_{10} est codée par :

$$0,3_{10} \mapsto 1.00110011001100110011010.2^{-2}. \tag{3.38}$$

En simple précision, on a donc bien $(0,1+0,2)-0,3=0$.

Cependant, les limites de précision des flottants évoquées précédemment engendrent des calculs parfois déroutants et entachés d'erreur. L'associativité de l'addition peut ne plus être vérifiée, la distributivité de la multiplication par rapport à l'addition non plus.

■ **Exemple 28 — Erreur d'arrondis en double précision.** Considérons de nouveau l'opération $(0,1+0,2)-0,3$ mais cette fois-ci **en double précision**. Le résultat de cette opération devrait être 0, mais il n'en est rien et on trouve $5,551115123125783.10^{-17}$.

L'erreut commise provient de la représentation de $0,3_{10}$ en double précision :

$$0,3_{10} \mapsto 1.0011001100110011001100110011001100110011001100110011.2^{-2} \tag{3.39}$$

Pour $0,3_{10}$, les trois derniers bits valent 011 et les bits GRS 001. L'intervalle considéré pour l'arrondi est $[11,100]$, son milieu vaut 11.100 et $11.001 < 11.100$. La valeur choisie pour les trois derniers bits est donc la valeur inférieure de l'intervalle, soit 011 : l'arrondi ne modifie pas la valeur codée en machine. Cette représentation se note $0,29999999999999999$ sur 17 décimales.

Par ailleurs, l'addition de $0,1+0,2$ aboutit à $0,30000000000000004$ en représentant 17 décimales. La soustraction de la représentation de $0,3$ en double précision aboutit donc à une erreur d'arrondi. **En double précision**, on a $(0,1+0,2)-0,3 = 5,551115123125783.10^{-17}$. La différence entre les deux valeurs porte sur les trois derniers bits.

Bien entendu, les chiffres significatifs du résultat, c'est à dire les 15 premiers, sont corrects. Il faut donc juste être vigilant sur l'interprétation des résultats. Un programme en C++ est donnée en annexe **??** pour tester ces limites.

■ **Exemple 29 — Perte de l'associativité de l'addition.** Considérons l'addition suivante :

$$2^{24}+1+1 \tag{3.40}$$

Supposons que le calcul soit mené de la gauche vers la droite, c'est à dire $(2^{24}+1)+1$. Comme l'écart entre deux nombres codables en IEEE 754 est de 2 lorsque l'exposant non biaisé e vaut 24, on obtient que $(2^{24}+1)+1 = 2^{24}+1 = 2^{24}$. Or, si on commence le calcul par la droite, $2^{24}+(1+1) = 2^{24}+2$, car le nombre 2 est représentable avec l'exposant 24. L'addition n'apparaît donc plus associative.

■ **Définition 28 — Mécanisme d'absorption.** Le mécanisme d'absorption apparaît lorsqu'on additionne deux valeurs dont l'écart relatif est très important : cela engendre l'absorption de la plus petite valeur par la plus grande.

■ **Exemple 30 — Multiplication non distributive par rapport à l'addition.** Considérons l'opération

$$100 \times (0,1+0,2)$$

Si le calcul est effectué en commençant par l'addition, le résultat est faux et vaut $30,000000000000004$. Par contre, si on développe le calcul et qu'on l'effectue après développement, on trouve exactement $100 \times 0,1 + 100 \times 0,2 = 30,0$.

Lorsqu'on utilise les flottants, il faut donc éviter d'additionner des calculs sur des nombres dont l'écart relatif est très important. De même, on évitera de soustraire deux valeurs très proches au risque de perdre énormément de chiffres significatifs dans le résultat.

De plus, il faut se poser la question de la nécessité d'utiliser une simple ou une double précision en fonction de l'application.

R L'impact du changement de précision des flottants sur la complexité mémoire est conséquent car il double systématiquement l'espace mémoire nécessaire au stockage des données. Par ailleurs, il a également un impact important sur la complexité temporelle, car les unités de calculs sur le flottants (FPU) sont spécifiques, plus complexes et moins nombreuses que les unités arithmétiques sur les entiers (ALU).La puissance nécessaire calcul nécessaire pour changer de précision n'est donc pas à négliger non plus et dépend fortement de l'architecture du processeur.

R Certains processeurs ne disposent même pas d'unités arithmétiques pour les flottants. Dans ce cas, il faut soit utiliser l'arithmétique fixée soit émuler le calcul des flottants. La première solution, l'arithmétique fixée, est la plus souhaitable pour minimiser l'impact sur la complexité temporelle. Dans le domaine des systèmes embarqués elle est souvent à privilégier.

R Les nombres flottants forment un sous-ensemble des nombres rationnels, tout comme les nombres décimaux. Mais les décimaux et les flottants ne sont pas les mêmes ensembles.

Deuxième partie

Anneaux et corps finis

Chapitre 4

Vers les anneaux euclidiens

La construction et la représentation des nombres ayant été évoquées dans les chapitres précédents, nous pouvons maintenant décrire les caractéristiques du calcul sur les entiers. De ces caractéristiques, les mathématiciens ont dégagé des structures algébriques qui permettent d'identifier d'autres objets mathématiques dont le comportement est similaire. Ces structures se nomment groupes, anneaux ou corps et interviennent dans les calculs liés au codage, au chiffrement et à la compression.

Commençons donc par faire le tour des caractéristiques du calcul sur les entiers.

4.1 Plus grand diviseur commun

4.1.1 Définition et existence

Les entiers possèdent des propriétés remarquables parmi lesquelles on trouve les multiples et les diviseurs communs. Ces éléments sont liés à la décomposition en facteurs premiers qui sera exposée par la suite.

■ **Définition 29 — Diviseur commun.** Soit a et b deux entiers. Un diviseur commun de a et de b est un entier qui divise a et b.

Étudier les diviseurs d'un nombre, c'est le décomposer en éléments plus simples.

Théorème 6 — Existence du PGCD. Parmi tous les diviseurs communs de deux entiers a et b non nuls, il y en a **un** qui est le plus grand. Ce dernier est nommé plus grand commun diviseur de a et de b. On le note $\mathrm{PGCD}(a, b)$.

Démonstration. Soit $a \in \mathbb{N}^\star$. D'après le théorème 2, tous les diviseurs de a sont bornés par $|\ a\ |$. On peut tenir le même raisonnement pour ceux de b. Donc, parmi les diviseurs de a et de b, il y en a donc un plus grand. ■

Une interprétation géométrique du PGCD est donnée sur la figure 4.1. La question qu'on pourrait se poser est la suivante : comment paver un rectangle avec des carrés **identiques les plus grands possibles** ? Le PGCD donne le côté du carré qu'il faut utiliser !

Puisque le PGCD de deux entiers existe, se pose maintenant la question de le calculer.

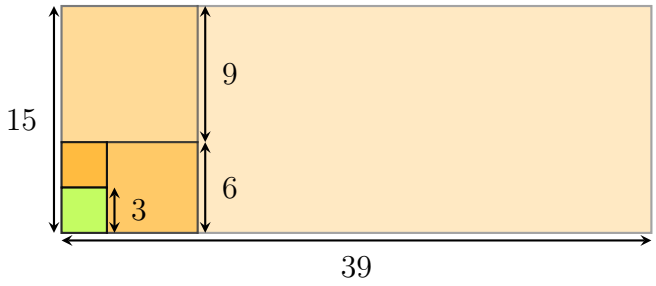


FIGURE 4.1 – Illustration du concept de PGCD en géométrie plane. Le PGCD de 39 et 15 est 3. On peut paver un rectangle de 39 par 15 par 65 carrés de 3 par 3. On peut paver avec un carré de côté plus petit (1 par exemple), mais pas par un carré de côté plus grand.

4.1.2 Algorithme d'Euclide et PGCD

Théorème 7 — . Soit a et b deux entiers.

1. Si $b = 0$, alors $\text{PGCD}(a, b) = a$.
2. Si $b \neq 0$, alors $\text{PGCD}(a, b) = \text{PGCD}(b, a \mod b)$.

Démonstration. Démonstration de l'égalité de l'ensemble $\mathcal{D}_{ab}$ des diviseurs de a et de b et de l'ensemble $\mathcal{D}_{br}$ des diviseurs de b et de r par double inclusion.

$\mathcal{D}_{ab} \subset \mathcal{D}_{br}$ **:** La division euclidienne étant unique comme nous l'avons montré au théorème 3, il existe un entier q tel que $a = qb + r$. Ce qui peut s'écrire : $a - qb = r$. Si γ est un diviseur de a et de b, alors on peut écrire : $a - bq = \gamma a' + \gamma b' q = \gamma(a' - b'q) = r$. On a donc montrer qu'un diviseur de a et de b est un diviseur de r.

$\mathcal{D}_{br} \subset \mathcal{D}_{ab}$ **:** De même, si η est un diviseur de b et de r, alors on a : $a = bq + r = \eta(b'q + r')$, ce qui signifie que η est un diviseur de a.

Donc, $\mathcal{D}_{ab} = \mathcal{D}_{br}$. Ceci est vrai, y compris pour le plus grand des diviseurs de a et de b. ■

De la conclusion du théorème 7, Euclide déduit l'algorithme 1, algorithme fondateur. On en présente une version optimisée dans le sens où l'algorithme d'Euclide procède par soustractions successives, non pas avec l'opérateur modulo.

Algorithme 1 Algorithme d'Euclide (optimisé)

```
1: Fonction PGCD(a, b)                          ▷ On suppose que (a, b) ∈ ℤ, b ⩽ a.
2:     a ←| a |
3:     b ←| b |
4:     r ← a mod b
5:     tant que r > 0 faire                     ▷ On connaît la réponse si r est nul.
6:         a ← b
7:         b ← r
8:         r ← a mod b
9:     retourner b                              ▷ Le pgcd est b
```

Peut-on prouver que cet algorithme se termine, est correct et trouver sa complexité ?

Terminaison de l'algorithme d'Euclide

Dans le but de prouver la terminaison de l'algorithme 1, on introduit une suite auxiliaire.

■ **Définition 30 — Suite des restes de la division euclidienne.** Soient a et b des entiers. On définit la suite des restes de la division euclidienne comme suit :

$$r_0 = |\ a\ | \tag{4.1}$$
$$r_1 = |\ b\ | \tag{4.2}$$
$$q_k = \lfloor r_{k-1}/r_k \rfloor, 1 \leqslant k \leqslant n \tag{4.3}$$

Alors on a :

$$r_{k-1} = q_k r_k + r_{k+1} \tag{4.4}$$
$$r_{k+1} = r_{k-1} \bmod r_k \tag{4.5}$$

On observe qu'un élément de la suite des restes est calculé à chaque tour de boucle (cf. algorithme 1 ligne 8).

À l'entrée de la boucle, $a = r_0$, $b = r_1$ et $r = r_2$. À chaque tour de boucle, les variables valent respectivement $a = r_k$, $b = r_{k+1}$ et $r = r_{k+2}$.

Proposition 8 — Stricte décroissance de $(r_n)_{n\in\mathbb{N}}$. La suite des restes de la division euclidienne est positive, strictement décroissante et minorée par zéro.

Démonstration. D'après le théorème 3, le reste r de la division euclidienne de a et de b est tel que : $0 \leqslant r < b$. Donc, la suite est minorée par zéro. Cette borne est atteinte lorsque a est un multiple de b. C'est une suite positive car elle est initialisée à des valeurs positives. Elle est strictement décroissante car $r < b$. ■

Pour prouver la terminaison, on choisit donc le variant de boucle r. Celui-ci est positif, **strictement** décroissant et minoré par zéro. La condition d'arrêt est donc atteinte. Le programme se termine.

Correction

D'après la proposition 8, il existe un indice n qui correspond au dernier reste non nul de la suite. Il nous reste à nous assurer que r_n est le PGCD de a et de b. Ceci découle directement du point 2 du théorème 7 qui fait de r_n un invariant.

Complexité

La complexité de l'algotithme d'Euclide 1 n'est pas triviale à mesurer. Pour y parvenir, nous allons nous appuyer sur la suite des restes et la comparer à la suite de Fibonacci. On note également que n, l'indice du dernier reste non nul donne directement une mesure de la complexité de la boucle.

Proposition 9 — Quotients de la suite des restes. Soit n l'indice de la suite des restes correspondant au dernier reste non nul. Alors on a :

$$\forall k \in [\![1, n-1]\!],\ q_k \geqslant 1 \tag{4.6}$$
$$k = n,\ q_n \geqslant 2 \tag{4.7}$$

Démonstration. D'après la proposition 8, $r_{k-1} > r_k > r_{k+1}$. De plus, q_k est un entier strictement positif d'après l'équation 4.3. Donc $\forall k \in [\![1, n-1]\!], q_k \geqslant 1$.

Par ailleurs, si q_n valait 1, alors on aurait $r_n = r_{n-1}$, ce qui n'est pas possible car la suite est strictement décroissante. C'est pourquoi, $q_n \geqslant 2$. ■

Proposition 10 — Des restes et des éléments de la suite de Fibonacci. Soit n l'indice du dernier reste non nul de la suite des restes de la division euclidienne.

Soit f_i le i$^{\text{e}}$ terme de la suite de Fibonacci définie par $f_{i+1} = f_i + f_{i-1}$, $f_0 = 1$ et $f_1 = 1$. Alors on a :

$$\forall k \in [\![0, n]\!], r_k \geqslant f_{n-k}. \tag{4.8}$$

Démonstration. D'après la proposition 9, on a $r_{n-1} = q_n r_n \geqslant 2 = f_2$, car r_n est non nul. Comme la suite des restes est décroissante, $r_{n-1} \geqslant 2r_n$ et donc $r_n \geqslant f_2/2 = f_1$. En réitérant n fois ce raisonnement en faisant décroitre n, c'est à dire en remontant la suite des restes, on trouve que $r_0 \geqslant f_n$ ainsi que tous les résultats. ■

Supposons qu'il y a n étapes lors de l'algorithme d'Euclide. Alors, on a

$$b = r_1 \geqslant f_{n-1} \tag{4.9}$$

Or, la suite de Fibonacci est une suite récurrente linéaire d'ordre deux. On connaît donc sa forme explicite.

$$f_n = \frac{1}{\sqrt{5}} \left(\phi^n - \left(-\frac{1}{\phi} \right)^n \right) \tag{4.10}$$

avec le nombre d'or $\phi = \frac{1+\sqrt{5}}{2}$. De plus, on peut montrer que

$$f_n \simeq \phi^n \tag{4.11}$$

et donc, au rang $n-1$, on a :

$$\log(f_{n-1}) \simeq (n-1)\log(\phi) \tag{4.12}$$

En utilisant l'équation 4.9, on en conclut que :

$$n \geqslant 1 + \frac{\log(b)}{\log(\phi)} \tag{4.13}$$

La complexité de l'algorithme d'Euclide est donc $O(1 + 1,44 \log_2(b))$, où $\log_2(b)$ est le nombre bits nécessaires pour coder b. Elle est logarithmique en fonction de la taille du codage l'entier et donc très efficace, ce qui est très important pour les opérations de chiffrement ou de codage. Cette conclusion est également le théorème de Lamé.

Théorème 8 — Théorème de Lamé. Le nombre de divisions euclidiennes nécessaires pour calculer PGCD(a, b) par l'algorithme d'Euclide est inférieur ou égal à 5 fois le nombre de chiffres de b en base 10.

4.1.3 Combinaison linéaire d'entiers

Proposition 11 — Les restes sont des combinaisons linéaires. Soit a et b des entiers. Les restes de la division euclidienne de a par b sont des combinaisons linéaires de a et de b. Mathématiquement :

$$\forall k \in [\![0, n]\!], r_k = u_k a + v_k b \tag{4.14}$$

si n est l'indice du dernier reste non nul.

Démonstration. On procède par récurrence sur la propriété $\mathcal{P}_k : r_k = u_k a + v_k b$.

Initialisation : on peut prouver que $\mathcal{P}_0$ et $\mathcal{P}_1$ sont vraies en trouvant u_0, v_0, u_1 et v_1. On peut poser : $r_0 = 1 \times a + 0 \times b$ et $r_1 = 0 \times a + 1 \times b$ ce qui initialise la démonstration.

Hérédité. Supposons que $\mathcal{P}_k$ soit vraie pour un certain rang k et $k-1$, $k > 1$. La suite des restes s'écrit : $r_{k+1} = r_{k-1} - q_k r_k$. En s'appuyant sur l'hypothèse de récurrence, on peut écrire :

$$r_{k+1} = u_{k-1}a + v_{k-1}b - q_k(u_k a + v_k b) \tag{4.15}$$
$$= (u_{k-1} - q_k u_k)a + (v_{k-1} - q_k v_k)b \tag{4.16}$$

Donc, $\mathcal{P}_{k+1}$ est vraie. Comme la propriété a été initialisée, $\mathcal{P}_k$ est vraie pour tout $k \in [\![0, n]\!]$. ■

Théorème 9 — Relation de Bezout. Toute combinaison linéaire de a et de b est un multiple du PGCD de a et de b.Mathématiquement :

$$a\mathbb{Z} + b\mathbb{Z} = \mathrm{PGCD}(a, b)\mathbb{Z} \tag{4.17}$$

Formulé autrement, $\mathrm{PGCD}(a, b)$ est une combinaison linéaire de a et de b.

En particulier,

$$\exists (u, v) \in \mathbb{Z}^2, au + bv = \mathrm{PGCD}(a, b) \tag{4.18}$$

Démonstration. Il suffit d'appliquer la proposition 11 pour l'indice n, l'indice du dernier reste non nul. r_n est une combinaison linéaire de a et de b. On a donc : $\exists (u, v) \in \mathbb{Z}, r_n = \mathrm{PGCD}(a, b) = au + bv$ ■

4.1.4 Algorithme d'Euclide étendu

Les coefficients de la relation de Bezout peuvent être facilement trouvés grâce à la propriété 11.

Théorème 10 — Coefficients de Bezout. Soient a et b des entiers. On définit les suites suivantes :

$$\forall k \in [\![0, n]\!], r_k = r_{k-2} - q_{k-1} r_{k-1}, r_0 = a, r_1 = b \tag{4.19}$$
$$\forall k \in [\![0, n]\!], x_{k+1} = q_k x_k + x_{k-1}, x_0 = 1, x_1 = 0 \tag{4.20}$$
$$\forall k \in [\![0, n]\!], y_{k+1} = q_k y_k + y_{k-1}, y_0 = 0, y_1 = 1 \tag{4.21}$$

Alors on a :

$$r_n = \mathrm{PGCD}(a, b) = (-1)^n x_n a + (-1)^{n+1} y_n b \tag{4.22}$$

Démonstration. On procède par récurrence comme dans la proposition 11. ■

Pour calculer les coefficients de Bezout, il suffit donc de modifier l'algorithme d'Euclide en ajoutant le calcul des x_k et y_k.

Algorithme 2 Algorithme d'Euclide étendu, calcul des coefficients de Bezout

```
1:  Fonction PGCD(a, b)                          ▷ a et b sont des entiers.
2:      a ←| a |
3:      b ←| b |
4:      x_0, y_1 ← 1
5:      y_0, x_1 ← 0
6:      sign ← 1
7:      tant que b > 0 faire                     ▷ On connaît la réponse si b est nul.
8:          r ← a mod b
9:          q ← a//b                             ▷ Division entière
10:         a ← b
11:         b ← r
12:         t_x ← x_1
13:         t_y ← y_1
14:         x_1 ← x_1 × q + x_0
15:         y_1 ← y_1 × q + y_0
16:         x_0 ← t_x
17:         y_0 ← t_y
18:         sign ← −sign
19:     x_0 ← sign × x_0
20:     y_0 ← −sign × y_0
21:     retourner a, x_0, y_0                    ▷ Le pgcd est a
```

R L'algorithme 2 est particulièrement utile pour trouver des combinaisons linéaires d'entiers et donc résoudre des équations. Sa complexité est la même que celle de l'algorithme d'Euclide.

R L'algorithme 2 sert particulièrement à calculer des inverses en arithmétique modulaire, notamment dans le corps finis (cf. section 5.13).

■ **Exemple 31 — Calcul de l'inverse de 3 modulo 4.** On cherche $3^{-1} \mod 4$, c'est à dire $a \in \mathbb{Z}/4\mathbb{Z}$ tel que : $3a \equiv 1 \mod 4$. On peut réécrire cette équation : $\exists k \in \mathbb{Z}/4\mathbb{Z}, 3a + 4k = 1$.

Grâce à l'algorithme d'Euclide étendu, on peut trouver k. On vérifie que : $3 \times 3 + 4 \times (-2) = 1$ ce qui signifie de 3 est l'inverse de 3 modulo 4 !

4.2 Facteurs premiers

4.2.1 Nombres premiers, une infinité d'atomes

■ **Définition 31 — Nombre premier.** Un entier p est un nombre premier si ses diviseurs positifs sont exactement 1 et p.

R 1 n'est donc pas un nombre premier car il ne possède qu'un seul diviseur, lui-même. Le plus petit des nombres premiers est donc 2.

L'ensemble des nombres premiers est souvent noté $\mathbb{P}$.

Théorème 11 — Premiers atomes. Tout nombre entier n possède au moins un diviseur premier.

Démonstration. Démonstration par disjonction des cas et par l'absurde.

Si n est premier, $n \mid n$, ce qui prouve le théorème.

Si n n'est pas premier, alors parmi les diviseurs de n, considérons le plus petit p. Ce diviseur p est nécessairement premier sinon il existerait un diviseur b tel que $b \mid p$ donc $b \mid n$ et alors $1 < b < p < a$. Ce qui contredit l'hypothèse que p est le plus petit des diviseurs de n. ■

Théorème 12 — Une infinité d'atomes. Il existe une infinité de nombres premiers.

Démonstration. Supposons qu'il n'existe pas une infinité de nombres premiers. L'ensemble des nombres premiers serait alors un ensemble fini que l'on pourrait noter $E = \{p_1, p_2, \ldots, p_n\}$. Considérons le nombre $N = p_1 \times p_2 \times \cdots p_n + 1$. Ce nombre ne fait pas partie de l'ensemble fini E, il n'est donc pas premier. N possède donc un diviseur premier que l'on note q. q est différent de tous les éléments de E sinon on aurait $q \mid (N - p_1 \times p_2 \times \cdots \times p_n)$, c'est à dire $q \mid 1$, ce qui est impossible. Donc $q \in E$ et $q \notin E$, ce qui est absurde. ■

4.2.2 Nombres premiers et PGCD

Théorème 13 — Théorème de Gauss-Euclide. Soit p un nombre premier et a et b des entiers. Alors :

$$p \mid ab \longrightarrow p \mid a \text{ ou } p \mid b \tag{4.23}$$

Si p divise le produit ab, alors p divise au moins l'un des deux facteurs a ou b.

Démonstration. Démonstration par l'absurde.

Supposons que $p \mid ab$ et $p \nmid a$. Comme p est premier, $\text{PGCD}(a, p) = 1$. D'après le théorème 9, il existe des entiers u et v tels que $1 = au + pv$. En multipliant par b, on obtient : $b = aub + pvb$. Or, $p \mid aub$ et $p \mid pvb$. Donc $p \mid b$.

En suivant le même raisonnement, on peut montrer que si $p \mid ab$ et $p \nmid b$ alors $p \mid a$. ■

■ **Définition 32 — Premiers entre eux.** Soit a et b deux entiers. a et b sont premiers entre eux si leur plus grand commun diviseur est 1, i.e. $\text{PGCD}(a, b) = 1$.

Théorème 14 — Théorème de Bezout. Soit deux entiers a et b. a et b sont premiers entre eux si et seulement s'il existe deux entiers u et v tels que $au + bv = 1$.

Démonstration. Ce théorème est une conséquence de la définition précédente et du théorème 9. ■

4.2.3 Décomposition en nombres premiers

Proposition 12 — Diviseur d'un produit de premiers. Soit p un nombre premier qui divise un produit de nombres premiers $\prod_{i=1}^{n} q_i$. Alors p est égal à un des facteurs q_i.

Démonstration. Démonstration par récurrence.

Initialisation. Pour $n = 1$, $p \mid q_1$. Comme q_1 est premier, ses seuls diviseurs sont 1 et q_1. On a nécessairement $p = q_1$.

Hérédité. Supposons la propriété vraie au rang n et considérons maintenant que p divise $\prod_{i=1}^{n+1} q_i$. Alors $p \mid q_{n+1} \prod_{i=1}^{n} q_i$. d'après l'hypothèse de récurrence, $\exists q_i, i \in [\![1, n]\!], p = q_i$.

On peut donc conclure que la proposition est vraie. ■

Théorème 15 — Théorème fondamental de l'arithmétique. Tout entier $n > 1$ peut s'écrire comme un produit de nombres premiers. Ce produit, en faisant abstraction des permutations, est uniquement déterminé.

Démonstration. Démonstration de l'existence par récurrence forte.

Initialisation. 2 est premier.

Hérédité. Supposons l'existence d'un produit de nombres premiers pour chaque nombre plus petit que n. Si n est premier, alors $n \mid n$. Si n n'est pas premier, alors, d'après le théorème 11, il existe un nombre premier p et un entier k tel que $n = pk$. D'après l'hypothèse de récurrence, k est un produit de nombres premiers, car $k < n$. Donc n est également un produit de nombres premiers.

Démonstration de l'unicité par l'absurde.

Supposons qu'il existe deux décompositions **distinctes** en nombres premiers d'un entier n telles que $n = \prod_{i=1}^{k} p_i = \prod_{i=1}^{j} q_i$ avec $j > 1$ et $k > 1$. Nécessairement, $p_1 \mid \prod_{i=1}^{j} q_i$ et d'après le théorème 13 il existe donc un nombre premier q_s tel que $p_1 = q_s$. Or, on peut tenir ce raisonnement pour tous les p_i, ce qui est absurde car nous avons supposé les deux décompositions distinctes. ■

R Le problème de la décomposition d'un entier en un produit de nombre premiers s'appelle également la factorisation en nombres premiers. On ne connaît pas d'algorithme vraiment efficace pour résoudre ce problème. Cette observation est le fondement de la sécurité de nombreux systèmes dont le chiffrement RSA.

R Les nombres premiers sont donc les atomes des nombres en général : ils les composent et permettent de les construire.

4.2.4 Indicatrice d'Euler

Les concepts développés dans le cadre du chiffrement de l'information nécessitent souvent l'introduction d'une fonction qualifiant les nombres inférieurs à un nombre et premiers avec celui-ci.

■ **Définition 33 — Fonction indicatrice d'Euler.** On appelle fonction indicatrice d'Euler

la fonction définie par :

$$\begin{aligned}\varphi : \mathbb{N}^\star &\longrightarrow \mathbb{N}^\star \\ n &\longmapsto \text{card}(\{m \in \mathbb{N}^\star, m \leqslant n \text{ et } \text{PGCD}(n,m) = 1\})\end{aligned}$$

■ **Exemple 32 —** $\varphi(12) = 4$**.** En effet, parmi les nombres inférieurs à 12, seuls 1, 5, 7, et 11 sont premiers avec n.

Le calcul de l'indicatrice d'Euler s'appuie sur le théorème fondamental de l'arithmétique. Celui-ci nous dit qu'on peut écrire $n = \prod_{i=1}^{r} p_i^{k_i}$, les nombres p_i étant premiers. On a alors :

$$\varphi(n) = \prod_{i=1}^{r}(p_i - 1)p_i^{k_i - 1} = n\prod_{i=1}^{r}(1 - \frac{1}{p_i}) \tag{4.24}$$

On ne connaît pas non plus d'algorithmes efficaces pour le calcul de $\varphi(n)$.

Théorème 16 — Nombre premier et indicatrice d'Euler. Si p est un nombre premier alors $\varphi(p) = p - 1$

Proposition 13 — Somme des diviseurs d'un entier et indicatrice d'Euler.

$$\sum_{d|m, d>0} \varphi(d) = m \tag{4.25}$$

4.3 Exemples d'applications simples

■ **Exemple 33 — Simplification de fraction.** Le PGCD peut être utilisé de manière programmatique pour automatiser la simplification des fractions.

■ **Définition 34 — Équation diophantienne.** Une équation diophantienne est une équation polynomiale dont les solutions recherchées sont des nombres entiers ou rationnels. Les coefficients du polynôme sont à coefficients rationnels.

■ **Exemple 34 — Équations diophantiennes.** Il existe de nombreux types d'équations diophantiennes. En voici quelques unes :

$$ax + by = c, \quad a, b, c \in \mathbb{Q}, x, y \in \mathbb{Z} \tag{4.26}$$
$$x^2 + y^2 = p, \quad x, y, p \in \mathbb{Z} \tag{4.27}$$
$$x^n + y^n = z^n, \quad n, x, y, z \in \mathbb{Z} \tag{4.28}$$
$$y^2 = x^3 + ax + b, \quad x, y, a, b \in \mathbb{F}_p, p \in \mathbb{P} \tag{4.29}$$
$$\tag{4.30}$$

R Il n'y a pas de méthode générale pour résoudre les équations diophantiennes. Les progrès réalisés dans la recherche de solutions pour un type d'équation diophantienne nécessitent la plupart du temps le développement des mathématiques nouvelles. Ces équations sont très importantes dans le domaine de la cryptographie.

Théorème 17 — Solutions d'une équation diophantienne linéaire à deux inconnues. $\forall a, b, n \in \mathbb{Z}$, l'équation $\mathcal{E} : ax + by = n$ est possède des solutions entières si et seulement si $\text{PGCD}(a, b) \mid n$.

Démonstration. Supposons d'abord que $\mathcal{E}$ possède des solutions entières (x, y). Posons $d = \text{PGCD}(a, b)$. Alors, en notant $a = da'$ et $b = db'$, on peut récrire l'équation $\mathcal{E}$: $d(a'x + b'y) = n$. Comme $\mathcal{E}$ possède des solutions entières (x, y) cela implique que $\exists k \in \mathbb{Z}, n = kd$, ce qui signifie que d divise n.

Inversement, supposons que $d \mid n$. On peut noter $n = dn', n' \in \mathbb{Z}$ et réécrire $\mathcal{E}$: $a'x + b'y = n'$. Or, d'après le théorème de Bezout 14, on est capable de trouver une solution entière à $a'u + b'v = 1$, car a' et b' sont premiers entre eux. Une telle solution (u_0, v_0) multipliée par n', c'est à dire $(n'u_0, n'v_0)$ sera bien une solution de $\mathcal{E}$.

■

M **Méthode 1 — Comment résoudre les équations du type $\mathcal{E} : ax + by = c$?** Pour résoudre dans $\mathbb{Z}$ l'équation

$$\mathcal{E} : ax + by = c$$

où $(a, b, c) \in \mathbb{Z}^3$ il faut :

1. déterminer $d = \text{PGCD}(a, b)$,
2. **si $d \nmid c$, $\mathcal{E}$ n'a pas de solution dans** $\mathbb{Z}^2$ d'après le théorème 17,
3. sinon simplifier l'équation par d et $\mathcal{E} \Leftrightarrow a'x + b'y = c'$ avec $\text{PGCD}(a', b') = 1$,
4. déterminer une solution particulière (x_0, y_0) de $\mathcal{E}$ à l'aide d'une égalité de Bezout triviale ou en utilisant l'**algorithme d'Euclide étendu**.
5. remarquer que si (x, y) et (x_0, y_0) sont solutions de $\mathcal{E}$, alors en effectuant la différence des deux équations $a'x + b'y = c'$ et $a'x_0 + b'y_0 = c'$, on obtient l'égalité $a'(x - x_0) = -b'(y - y_0)$.
6. achever la résolution avec le théorème de Gauss 13 en remarquant que cette égalité signifie que $b' \mid x - x_0$. C'est à dire qu'il existe des entiers $k \in \mathbb{Z}$ tels que : $(x - x_0) = kb'$ et donc $x = x_0 + kb'$. En reportant ce résultat dans l'équation précédente, on trouve : $b'(y - y_0) = -a'kb'$ et donc $y = y_0 - ka'$. Soit $k \in \mathbb{Z}$, une solution de l'équation $\mathcal{E}$ est donc :

$$x = x_0 + kb' \tag{4.31}$$
$$y = y_0 - ka' \tag{4.32}$$
$$\tag{4.33}$$

Il en existe une infinité dans $\mathbb{Z}$.

■ **Exemple 35 — Équation diophantienne du premier degré.** Une puce électronique sous test possède une architecture inconnue. On sait simplement qu'elle ne contient que deux types de circuits, les a et les b. On connaît les courants consommés par a et b, qui valent respectivement 31 mA et 28 mA. On est également capable de mesurer le courant total consommé par la puce qui est de 1460 mA. De combien de circuits a et b est constituée cette puce électronique ?

L'équation à résoudre qui correspond à cette situation est $31x + 28y = 1460$, x et

y étant des nombres entiers représentant les quantités de circuits a et b. On remarque que 31 et 28 sont premiers entre eux. L'algorithme d'Euclide étendu permet de trouver une solution particulière à $31x + 28y = 1$, $x_0 = -9$ et $y_0 = 10$. De cette solution particulière, on en déduit une solution générale :

$$x = -13140 + 28k \tag{4.34}$$
$$y = 14600 - 31k \tag{4.35}$$

Maintenant, on cherche des valeurs de k pour lesquelles x et y sont positifs, car les nombres de circuit de type a ou b sont strictement positifs. Finalement, grâce à un petit script Python, on trouve que $x = 20$ et $y = 30$ conviennent. Ces solutions sont obtenues pour $k = -470$.

Méthode 2 — Comment résoudre un système d'équations PGCD et PPCM Soit le système d'équations suivant à résoudre dans $\mathbb{Z}^2$:

$$\mathcal{S} : \begin{cases} PGCD(x, y) &= d \\ PPCM(x, y) &= m \end{cases} \tag{4.36}$$

- Si $d \nmid m$, alors $\mathcal{S}$ n'a pas de solution dans $\mathbb{Z}^2$.
- Sinon, effectuer le changement de variable $x = dx'$ et $y = dy'$.
- On a $\text{PGCD}(x', y') = 1$, donc $\mathcal{S}$ est équivalent à $x'y' = m'$.
- On achève la résolution en trouvant les couples premiers entre eux qui vérifient l'équation factorisée.

Méthode 3 — Comment résoudre une équation diophantienne par factorisation Soit l'équation $\mathcal{E} : ax^2 + by^2 + cxy + dx + ey + f = n$. Après avoir vérifié que $\text{PGCD}(a, b, c, d, e, f) \mid n$, on simplifie éventuellement l'équation et on la transforme en une écriture factorisée de type $\mathcal{E}_f : XY = K$.

On cherche ensuite tous les couples (X, Y) solutions dans $\mathbb{Z}^2$ et on en déduit l'ensemble $\mathcal{S} = \{(x_i, y_i)\}$ des solutions de l'équation $\mathcal{E}$ dans $\mathbb{Z}^2$.

■ **Exemple 36 — Application au déchiffrement du chiffre affine.** Le chiffrement affine (cf. par exemple le chiffre de César **??**) simple correspond à

$$\mathcal{E} : \lambda \to (a\lambda + b) \mod s$$

avec $(a, b) \in \mathbb{Z}$ connus et $s = 26$. Déchiffrer ce code revient à chercher λ connaissant e tel que :

$$\mathcal{D} : (a\lambda + b) \mod s = e$$

Ce système est équivalent à :

$$\mathcal{D} : a\lambda - ks = e - b$$

qui est une équation diophantienne à deux inconnues (k et λ) et de degré 1. Il suffit donc de résoudre cette équation diophantienne à l'aide de la méthode 1.

$$\mathcal{D} : a'\lambda - ks' = e'$$

avec $\text{PGCD}(a, s) = d, e' = (e-b)/d, a' = a/d, s' = s/d$. On vérifie que $\text{PGCD}(a, s)|e-b$ et on utilisera l'algorithme d'Euclide étendu.

4.4 En résumé

À la lecture attentive des sections précédentes, on se rend compte que la possibilité de définir une division euclidienne sur un ensemble est riche de conséquences :

1. il existe un plus grand commun diviseur (cf. théorème 6),
2. la relation de Bezout est vérifiée (cf. théorème 9),
3. le théorème de Gauss-Euclide est valide (cf. théorème 13),
4. l'ensemble est factoriel, c'est à dire qu'on peut décomposer les éléments en facteurs de nombres premiers (cf. théorème 15),
5. et comme on le verra dans la chapitre suivant, on peut y développer une arithmétique modulaire.

Les mathématiciens ont réalisé les conséquences de la division euclidienne et ont créé des structures algébriques dont les éléments se comportent comme des nombres, même s'ils n'en sont pas, et cela s'est avéré très pratique ! Ces structures sont les anneaux euclidiens.

Chapitre 5

Corps finis

Pour concevoir les anneaux euclidiens et les corps finis, il est nécessaire de définir les structures de groupe, d'anneau et de corps. Elles n'ont été définies que tardivement après de longs travaux de recherche tout au long du XIX^{e} siècle. Il ne faut pas trop chercher de lien entre les noms de ces structures (groupes, anneaux et corps) et le sens qu'on accorde généralement à ces mots : les mathématiciens inventent des concepts abstraits qui ne correspondent pas nécessairement à des réalités tangibles mais ont quand même besoin de les nommer !

5.1 Structures algébriques

5.1.1 Groupe

■ **Définition 35 — Loi de composition interne.** On appelle loi de composition interne toute application :

$$\begin{aligned} \star : E \times E &\longrightarrow E \\ (x, y) &\longmapsto \star(x, y) = x \star y \end{aligned}$$

■ **Définition 36 — Magma.** Tout couple, où E est un ensemble et $\star$ une loi de composition interne est un magma.

■ **Définition 37 — Stabilité.** Une partie A d'un ensemble E est dite stable pour la loi de composition interne $\star$ si : $\forall (x, y) \in A^2, x \star y \in A$

■ **Définition 38 — Morphisme.** Soit $(E, \star)$ et $(F, \bullet)$ deux magmas. Alors l'application $f : E \longrightarrow F$ telle que $\forall (x, y) \in E^2, f(x \star y) = f(x) \bullet f(y)$ est un morphisme de $(E, \star)$ vers $(F, \bullet)$

■ **Définition 39 — Élément neutre.** Un élément e est dit neutre pour la loi de composition interne $\star$ si $\forall x \in E, x \star e = e \star x = x$.

■ **Définition 40 — Associativité.** On dit d'une loi de composition interne $\star$ sur un

ensemble E qu'elle est associative si et seulement si

$$\forall (x, y, z) \in E^3, (x \star y) \star z = x \star (y \star z) \tag{5.1}$$

■ **Définition 41 — Commutativité.** On dit d'une loi de composition interne $\star$ sur un ensemble E qu'elle est commutative si et seulement si

$$\forall (x, y) \in E^2, x \star y = y \star x \tag{5.2}$$

■ **Définition 42 — Monoïde.** Un ensemble E muni d'une loi de composition interne associative et d'un élément neutre e est nommée monoïde $(E, \star)$.

R Par la suite on notera l'élément neutre e. Pour l'opération d'addition classique, l'élément neutre vaut 0. Pour l'opération de multiplication classique, l'élément neutre vaut 1.

■ **Définition 43 — Symétrique.** Un élément x' d'un ensemble E muni d'une loi de composition interne $\star$ est dit symétrique de $x \in E$ si $x \star x' = e$.

■ **Définition 44 — Groupe.** Soit G un ensemble muni d'une loi de composition interne $\star$. G est un groupe si et seulement si :

1. la loi $\star$ est associative,
2. la loi $\star$ possède un élément neutre,
3. tout élément de G possède un symétrique.

Si, de plus, la loi $\star$ est commutative, alors le groupe est dit abélien ou commutatif.

■ **Exemple 37 — Groupes usuels.** $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, +), (\mathbb{C}, +), (\mathbb{Q}^\star, \times), (\mathbb{R}^\star, \times), (\mathbb{C}^\star, \times)$ sont des groupes abéliens. $(\mathcal{GL}_n(\mathbb{K}), \times)$, c'est à dire l'ensemble des matrices carrées d'ordre n inversibles à coefficients dans $\mathbb{K}$ muni de la multiplication matricielle, est un groupe non commutatif.

■ **Définition 45 — Ordre d'un groupe.** L'ordre d'un groupe est le cardinal de son ensemble sous-jacent.

M **Méthode 4 — Comment montrer que $(G, \star)$ est un groupe ?** Pour montrer que $(G, \star)$ est un groupe, on peut procéder de deux manières :

1. Montrer que $(G, \star)$ est un sous-groupe d'un groupe de référence $(\hat{G}, \star)$ montrant que :
   - $\hat{G}$ est bien un groupe connu pour $\star$,
   - $G \subset \hat{G}$,
   - en utilisant la caractérisation des sous-groupes.
2. Montrer que en revenant à la définition (loi de composition interne associative, élément neutre, tout élément est inversible).

■ **Définition 46 — Sous-groupe.** Soit $(G, \star)$ un groupe et $H \subset G$. H est un sous-groupe de G si :

1. H est stable par $\star$,
2. et $(H, \star)$ est un groupe.

Théorème 18 — Caractérisation d'un sous-groupe. Soit $(G, \star)$ un groupe et $H \subset G$. H est un sous-groupe de G si et seulement si :

1. H n'est pas vide,
2. $\forall (x, y) \in H^2, x \star y' \in H$.

où y' est le symétrique de y dans G.

M **Méthode 5 — Comment montrer que $(H, \star)$ est un sous-groupe de $(G, \star)$?** On peut procéder de la manière suivante, en utilisant la caractérisation des sous-groupes.

1. vérifier que $H \subset G$,
2. vérifier que H est non vide, en pratique on montre que $e_G \in H$, car l'élément neutre appartient toujours au groupe !
3. vérifier la stabilité pour la composition interne et le passage au symétrique : $\forall (x, y) \in H^2, x \star y' \in H$

On peut décomposer la vérification de la stabilité pour la loi de composition interne et pour le symétrique : $\forall (x, y) \in H^2, x \star y \in H$ et $\forall x \in H, x' \in H$

5.1.2 Anneau

■ **Définition 47 — Distributivité.** Soit un ensemble E muni de deux lois de composition interne $\star$ et $\bullet$. On dit que $\bullet$ est distributive par rapport à $\star$ lorsque $\forall (x, y, z) \in E^3$:

$$x \bullet (y \star z) = (x \bullet y) \star (x \bullet z) \tag{5.3}$$
$$(x \star y) \bullet z = (x \bullet z) \star (y \bullet z) \tag{5.4}$$

La première équation correspond à la distributivité à gauche et la seconde à droite.

R C'est typiquement le cas avec $\mathbb{R}$ et les opérations $\times$ et $+$. La multiplication est distributive par rapport à l'addition : $x(y + z) = xy + xz$ et $(x + y)z = xz + yz$.

■ **Définition 48 — Anneau.** Soit A une ensemble muni de deux lois de composition interne $+$ et $\times$. Alors on dit que $(A, +, \times)$ est un anneau si :

1. $(A, +)$ est un groupe,
2. $\times$ est associative,
3. $\times$ est distributive par rapport à $+$,
4. $\times$ possède un élément neutre noté 1_A,

Si, de plus, $\times$ est commutative, alors l'anneau est dit commutatif.

■ **Exemple 38 — Anneaux usuels.** $(\mathbb{Z}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des anneaux commutatifs. Si A est un anneau alors $(\mathcal{M}_n(A), +, \times)$, c'est à dire l'ensemble des matrices carrées d'ordre n à coefficients dans A muni de l'addition et de la multiplication matricielles, est un anneau. De même $A[X]$ est un anneau.

R $(\mathbb{N}, +, \times)$ n'est pas un anneau.

■ **Définition 49 — Élément inversible.** Soit $(A, +, \times)$ un anneau et $a \in A$. On dit que a est inversible s'il possède un symétrique pour $\times$.

■ **Définition 50 — Diviseur de zéro.** Soit $(A, +, \times)$ un anneau, $a \in A$ et $a \neq 0_A$, où 0_A est l'élément neutre de la loi $+$ de l'anneau A. On dit que a est un diviseur de zéro s'il existe un élément $b \neq 0_A$ de A tel que $a \times b = 0_A$.

■ **Définition 51 — Anneau intègre.** Soit $(A, +, \times)$ un anneau. On dit que A est intègre s'il ne possède pas de diviseurs de zéro.

■ **Exemple 39 — $(\mathbb{Z}, +, \times)$ est un anneau intègre..** On peut raisonner par l'absurde pour le démontrer. Soit a et b des éléments de $\mathbb{Z}$ tels que $ab = 0$. Supposons que $a \neq b \neq 0$. Alors on a $ab = b+b+\cdots+b = 0$ ou encore $b = 0-(a-1)b = -(a-1)b$. Cela implique que $-(a-1) = 1$, c'est à dire $a = -1+1 = 0$ ce qui contredit notre hypothèse.

■ **Définition 52 — Caractéristique d'un anneau.** La caractéristique d'un anneau A est le plus petit entier $n > 0$ tel que $n \times 1_A = 1_A + 1_A + \cdots + 1_A = 0$.

Proposition 14 — Éléments inversibles d'un anneau. Soit $(A, +, \times)$ un anneau. On note A^* l'ensemble des éléments inversibles (pour $\times$) de A. $(A^*, \times)$ est un groupe.

Théorème 19 — Formule du binôme de Newton. Soit $(A, +, \times)$ un anneau et a et b deux éléments de A qui commutent, c'est à dire que $a \times b = b \times a$. Alors, pour tout entier naturel n on a :

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \times b^{n-k} \tag{5.5}$$

Démonstration. Démonstration par récurrence sur n qui utilise la formule de Pascal et un changement de variable. ■

R Il faut noter la puissance de ce résultat qui ne nécessite qu'un anneau et deux éléments qui commutent...

5.1.3 Corps

■ **Définition 53 — Corps.** Soit $\mathbb{K}$ un ensemble muni de deux lois de compositions internes $+$ et $\times$. $(\mathbb{K}, +, \times)$ est un corps si :

- $(\mathbb{K}, +, \times)$ est un anneau commutatif non réduit à $\{0_\mathbb{K}\}$,
- $\mathbb{K}^* = \mathbb{K} \setminus \{0_\mathbb{K}\}$, c'est à dire que tout élément non nul est inversible.

■ **Exemple 40 — Corps usuels.** $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des corps.

R $(\mathbb{Z}, +, \times)$ n'est pas un corps.

R Un corps est un anneau intègre.

5.2 Relation d'équivalence et congruences

■ **Définition 54 — Relation binaire sur un ensemble.** Une relation binaire $\mathcal{R}$ sur un ensemble E est une propriété vraie pour certains couples (x, y) de E et fausse pour les autres. Elle peut être considérée comme un sous-ensemble du produit cartésien $E \times E$.
Si deux éléments x et y sont en relation selon $\mathcal{R}$, on le note $x\mathcal{R}y$.

■ **Définition 55 — Réflexivité.** Soit $\mathcal{R}$ une relation binaire sur E. On dit qu'elle est réflexive si $\forall x \in E, x\mathcal{R}x$.

■ **Définition 56 — Symétrie.** Soit $\mathcal{R}$ une relation binaire sur E. On dit qu'elle est symétrique si $\forall (x, y) \in E^2, x\mathcal{R}y \implies y\mathcal{R}x$.

■ **Définition 57 — Transitivité.** Soit $\mathcal{R}$ une relation binaire sur E. On dit qu'elle est transitive si $\forall (x, y, z) \in E^3, x\mathcal{R}y \text{ et } y\mathcal{R}z \implies x\mathcal{R}z$.

■ **Définition 58 — Antisymétrie.** Soit $\mathcal{R}$ une relation binaire sur E. On dit qu'elle est antisymétrique si $\forall (x, y) \in E^2, x\mathcal{R}y \text{ et } y\mathcal{R}x \implies x = y$.

■ **Définition 59 — Divisibilité.** Soit un anneau $(A, +, \times)$ et a et b deux éléments de A. On dit que b divise a que l'on note $b \mid a$ si et seulement s'il existe un élément q de A tel que $a = b \times q$. Cette définition est une généralisation de 13.

■ **Exemple 41 — La divisibilité est une relation antisymétrique.** $\forall (a, b) \in Z^2, a \mid b \text{ et } b \mid a \implies a = b$

■ **Définition 60 — Relation d'équivalence.** On appelle relation d'équivalence toute relation binaire sur un ensemble réflexive, symétrique et transitive.

Exemple de relation d'équivalence

On peut définir une relation nommée association dans un anneau et montrer que cette relation est une relation d'équivalence.

■ **Définition 61 — Éléments associés.** Soit un anneau $(A, +, \times)$ et a et b deux éléments de A. On dit que a et b sont associés si et seulement si $a \mid b$ et $b \mid a$.

Proposition 15 — L'association d'éléments est une relation d'équivalence.

Démonstration. On vérifie :

- la réflexivité : on a $a = 1 \times a = a \times 1$. Donc $a \mid a$, a est donc associé à a.

- la symétrie : si a est associé à b, $a \mid b$ et $b \mid a$, ce qui, lu dans l'ordre inverse affirme que b est associé à a.
- la transitivité : si a est associé à b et b associé à c, alors comme $a \mid b$ et $b \mid c$, on a $a \mid c$. Comme $c \mid b$ et $b \mid a$ on a $c \mid a$. Donc a est associé à c.

■

Classes des restes modulo

On peut regrouper les éléments équivalent dans des classes.

■ **Définition 62 — Classe d'équivalence.** Soit $\mathcal{R}$ une relation d'équivalence sur E. On note $Cl(x)$ le sous-ensemble de E constitué des éléments en relation avec x. $Cl(x) = \{y \in E, y\mathcal{R}x\}$

R Si $y \in Cl(x)$ alors $Cl(y) = Cl(x)$. Tout élément d'une classe d'équivalence la détermine.

■ **Définition 63 — Congruence modulo un entier sur $\mathbb{Z}$.** On dit que deux entiers a et b sont congrus modulo un entier naturel n si leur différence est un multiple de n. On note :

$$a \equiv b \mod n \Leftrightarrow \exists k \in \mathbb{Z}, x = y + kn \tag{5.6}$$

Proposition 16 — La congruence modulo un entier est une relation d'équivalence. .

Démonstration. Considérons les congruences modulo un entier m. On vérifie :

- la réflexivité : on a $a = a + 0m$. Donc $a = a \mod m$, a est donc congru à a modulo m.
- la symétrie : si a est congru à b, $\exists k \in \mathbb{Z}, a = b + km$ et donc $b = a - km$, ce qui signifie que b est congru à a modulo m.
- la transitivité : si a est congru à b et b congru à c, alors $\exists k \in \mathbb{Z}, a = b + km$ et $\exists q \in \mathbb{Z}, b = c + qm$. Donc, $a = c + (k + q)m$. Donc a est congru à c modulo m.

■

La relation de congruence est fondamentale pour la création des anneaux quotient et des corps finis. Elle définit une nouvelle norme de calcul dont les règles sont données par le théorème 20. Ce théorème est important car il permet de définir les opérations d'addition et de multiplication sur l'anneau des restes et des corps finis.

Théorème 20 — Règles de calcul des congruences. Les opérations addition, multiplication et puissance sont compatibles avec la relation de congruence. Soit $n \in \mathbb{N}$ et $a, b, c, d \in \mathbb{N}$ tels que

1. $a \equiv b \mod n$
2. $c \equiv d \mod n$

alors on a :

1. $\forall k \in \mathbb{N}, a^k \equiv b^k \mod n$
2. $a + c \equiv b + d \mod n$
3. $ac \equiv bd \mod n$

De plus, si $n = n_1 n_2$ et $\text{PGCD}(n_1, n_2) = 1$, alors $x \equiv a \mod n_1 n_2 \Leftrightarrow x \equiv a \mod n_1$ et $x \equiv a \mod n_2$

■ **Définition 64 — Classe des restes modulo m.** La classe d'équivalence de $a \in \mathbb{Z}$ est constituée de tous les entiers obtenus à partir de a en lui ajoutant un multiple entier de m. C'est pourquoi on écrit :

$$\{b : b \equiv a \mod m\} = a + m\mathbb{Z} \tag{5.7}$$

On note souvent cette classe $\overline{a}$

■ **Exemple 42 — Classes des restes modulo 4.** La classe des restes de 1 modulo 4 est l'ensemble $\{1, 1 \pm 4, 1 \pm 2 \times 4, \ldots\} = \{1, -3, 5, -7, 9, -11, 13, \ldots\} = 1 + 4\mathbb{Z}$.

En ce qui concerne les valeurs négatives, on se rend vite compte qu'elles sont englobées dans les classes représentées par des valeurs positives.

Par exemple, $-3 \in \overline{1}$, $-2 \in= \overline{2}$ et $-1 \in= \overline{3}$. C'est pourquoi on représente souvent uniquement les classes des restes par les valeurs positives.

On peut remarquer également que $\overline{-a} = -\overline{a}$.

5.3 Relations d'ordre

■ **Définition 65 — Relation d'ordre.** Une relation d'ordre est une relation binaire réflexive antisymétrique et transitive.

■ **Définition 66 — Ordre partiel, ordre total.** Soit E un ensemble muni d'une relation d'ordre $\prec$. On dit que l'ordre défini par $\prec$ est total si $\forall (x, y) \in E^2, x \prec y \; ou \; y \prec x$. Sinon, on dit que l'ordre est partiel.

■ **Exemple 43 — Ordre total sur $\mathbb{R}$.** $\leqslant$ définit un ordre total sur $\mathbb{R}$.

■ **Exemple 44 — Ordre partiel de la divisibilité sur les entiers naturels.** $\mid$ définit un ordre partiel sur l'ensemble des entiers naturels $\mathbb{N}$. C'est une relation d'ordre car :

$$\forall n \in \mathbb{N}, n \mid n$$

$$\forall (a, b, c) \in \mathbb{N}^3, a \mid b \; et \; b \mid c \Longrightarrow a \mid c$$

et

$$\forall (a, b) \in \mathbb{N}^2, a \mid b \; et \; b \mid a \Longrightarrow a = b$$

Cependant, $3 \nmid 7$ par exemple. Tous les naturels ne sont pas ordonnés par la relation de divisibilité. Il s'agit donc d'un ordre partiel.

5.4 Anneaux des restes

5.4.1 Restes des entiers

■ **Définition 67 — Ensemble des classes des restes modulo m.** L'ensemble des classes des restes modulo m est noté $\mathbb{Z}/m\mathbb{Z}$. On l'appelle également l'ensemble des restes modulo m.

Théorème 21 — $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ est un anneau commutatif appelé l'anneau des restes ou anneau quotient.

Démonstration. On utilise le théorème 20 des règles de calcul sur les congruences modulo m pour définir les opérations d'addition et de multiplication sur $\mathbb{Z}/m\mathbb{Z}$.

Soit $\overline{a} = a + m\mathbb{Z}$, $\overline{b} = b + m\mathbb{Z}$ et $\overline{c} = c + m\mathbb{Z}$ trois éléments de $\mathbb{Z}/m\mathbb{Z}$.

On montre que $(\mathbb{Z}/m\mathbb{Z}, +)$ est un groupe commutatif :

- la loi est bien associative car $\overline{a}+(\overline{b}+\overline{c}) = a+b+c+m\mathbb{Z} = (a+b)+c+m\mathbb{Z} = (\overline{a}+\overline{b})+\overline{c}$.
- Par ailleurs, l'élément neutre de $+$ est le $\overline{0} = m\mathbb{Z}$ car $\overline{a} + \overline{0} = a + m\mathbb{Z} + m\mathbb{Z} = a + m\mathbb{Z} = \overline{a}$.
- Tout élément de $(\mathbb{Z}/m\mathbb{Z}, +)$ possède un symétrique : $\overline{a} + \overline{s} = \overline{0} \iff a + s + mZ = 0 + m\mathbb{Z} \iff s = -a \iff \overline{s} = -\overline{a}$.
- Enfin, on note que $\overline{a} + \overline{b} = a + b + m\mathbb{Z} = b + a + m\mathbb{Z} = \overline{b} + \overline{a}$. $(\mathbb{Z}/m\mathbb{Z}, +)$ est donc un bien groupe commutatif.

On vérifie, toujours à l'aide du théorème 20 que $\times$ est associative, commutative, distributive par rapport à $+$, possède un éléments neutre $\overline{1}$ car $a \times 1 \mod m = a \mod m$.

Donc, d'après la définition 48 $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ est un anneau commutatif. ■

Théorème 22 — $\mathrm{card}(\mathbb{Z}/m\mathbb{Z}) = m$.

Démonstration. Cela découle de la division euclidienne : pour que deux nombres soient congrus modulo m, il faut et il suffit que les restes des divisions euclidiennes de ces deux nombres par m soient les mêmes. Or, il y a au plus m restes possibles de 0 à $m - 1$. Il y a donc au maximum m classes d'équivalence. ■

■ **Exemple 45 — Addition et multiplication dans $(\mathbb{Z}/4\mathbb{Z}, +, \times)$.** Les tableaux 5.1 et 5.2 détaillent les tables d'addition et de multiplication dans l'anneau $(\mathbb{Z}/4\mathbb{Z}, +, \times)$. On notera par exemple que $\overline{3} + \overline{3} = \overline{2}$ ainsi que la présence de diviseurs de zéro tels que $\overline{2}$.

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

TABLE 5.1 – Table d'addition dans $(\mathbb{Z}/4\mathbb{Z}, +, \times)$.

Théorème 23 — Théorème des restes chinois. Soit N un élément de $\mathbb{Z}$ tel que :

1. $N = n_1 \times n_2 \times \cdots \times n_k$,
2. les n_i sont tous premiers entre eux.

$\times$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

TABLE 5.2 – Table de multiplication dans $(\mathbb{Z}/4\mathbb{Z}, +, \times)$.

Soit $a_1, a_2, \cdots a_k$ des entiers.

Alors le système d'équations :

$$\begin{aligned} x_1 &= a_1 \mod n_1 \\ x_2 &= a_2 \mod n_2 \\ &\vdots \\ x_k &= a_k \mod n_k \end{aligned}$$

possède une solution unique modulo N. De plus, toute paire de solutions est congruente modulo N, c'est à dire $x_i = x_j \mod N$.

Ce théorème peut également se formuler d'une manière plus abstraite en disant que les anneaux $(\mathbb{Z}/N\mathbb{Z}, +, \times$ et $(\mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \times \cdots \times \mathbb{Z}/n_k\mathbb{Z}, +, \times)$ sont isomorphes.

Ce théorème permet de résoudre des congruences simultanées. Pour trouver les solutions, il suffit de résoudre chaque équation séparément à l'aide de l'algorithme d'Euclide étendu (cf. algorithme 2). Un crible partant des n_i les plus grands permet d'être également très efficace.

■ **Définition 68 — Isomorphisme.** Un isomorphisme f est un morphisme bijectif (cf. définitions 38 et11). Le morphisme réciproque est noté f^{-1}.

Deux ensembles sont dits isomorphes s'il existe un isomorphisme de l'un vers l'autre. L'intérêt des isomorphismes est qu'ils préservent les propriétés structurelles algébriques des ensembles de départ et d'arrivée.

■ **Exemple 46 — Applications du théorème des restes chinois.** Historiquement, il a permis de calculer rapidement le nombre d'éléments d'ensemble difficiles à dénombrer en les regroupant de manière différente : par exemple le nombre de soldats d'une armée sur un champ de bataille. Au niveau des calendriers, il a permis de prédire les phases de la lune à un moment donné de l'année.

Dans le domaine du chiffrement, le système RSA (SSL, HTTPS) utilise ce théorème qui permet d'accélérer l'exponentiation modulaire dans les phases de déchiffrement et de signature.

Dans le domaine du radar, il permet de résoudre les ambiguïtés liées aux répétitions des radars à impulsions (Pulse Repetition Frequency).

Dans le domaine du traitement du signal, il existe un algorithme de transformée de Fourier rapide utilisant la décomposition en facteurs premiers pour accélérer le traitement.

Enfin, ce théorème possède une importance capitale en mathématiques et en informatique. Il est utilisé pour créer des séquences de nombres par Gödel dans ses démonstrations de l'incomplétude.

5.4.2 Groupe multiplicatif des restes

Théorème 24 — $\overline{a}$ est inversible dans $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ si et seulement si $\text{PGCD}(a, m) = 1$.

Démonstration. Soit $\overline{a}$ un élément de $(\mathbb{Z}/m\mathbb{Z}, +, \times)$. $\overline{a}$ est inversible s'il existe $\overline{x}$ de $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ tel que $ax \equiv 1 \mod m$. On peut également écrire ceci $ax + km = 1$. $\overline{a}$ est inversible dans $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ si l'équation diophantienne précédente possède une solution. Or, on sait d'après le théorème 17 que c'est le cas seulement si $\text{PGCD}(a, m) = 1$. ■

■ **Exemple 47 — Inverse de $\overline{3}$ dans $(\mathbb{Z}/4\mathbb{Z}, +, \times)$.** On a $\text{PGCD}(3, 4) = 1$, donc $\overline{3}$ est inversible dans $(\mathbb{Z}/4\mathbb{Z}, +, \times)$. D'après le tableau 5.8, l'inverse de $\overline{3}$ est $\overline{3}$ car $\overline{3} \times \overline{3} = \overline{1}$.

Théorème 25 — Les éléments inversibles de $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ forment le groupe commutatif $(\mathbb{Z}/m\mathbb{Z}, \times)^*$. On l'appelle le groupe multiplicatif des restes modulo m. La notation usuelle procède par l'ajout de * au groupe considéré.

Démonstration. On utilise la propriété 24 pour sélectionner les éléments et les définitions d'un groupe et d'un anneau. ■

R D'une manière générale, on note E^* l'ensemble des éléments inversibles d'un ensemble E. Par exemple, $(\mathbb{R}, \times)^* = (\mathbb{R} \setminus \{0\}, \times)$ est l'ensemble des éléments inversibles de $(\mathbb{R}, \times)$. L'ensemble des inversibles lui-même est le plus souvent noté $\mathbb{R}^\star$.

■ **Exemple 48 — Groupe multiplicatif $(\mathbb{Z}/4\mathbb{Z}, \times)^*$.** L'ordre de ce groupe est le cardinal de son ensemble, c'est à dire le nombre d'éléments de l'ensemble. En observant la table de multiplication 5.2, on en déduit que l'ordre de $(\mathbb{Z}/4\mathbb{Z}, \times)^*$ vaut 2 car seuls $\overline{1}$ et $\overline{3}$ sont inversibles.

Théorème 26 — L'ordre du groupe $(\mathbb{Z}/m\mathbb{Z}, \times)^*$ vaut $\varphi(m)$.

Démonstration. On s'appuie sur la définition de la fonction indicatrice d'Euler 33. ■

■ **Exemple 49 — Ordre du groupe $(\mathbb{Z}/12\mathbb{Z}, \times)^*$.** L'ordre de $(\mathbb{Z}/12\mathbb{Z}, \times)^*$ vaut 4 d'après le tableau 5.3.

m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\varphi(m)$	1	1	2	2	4	2	6	4	6	4	10	4	12	6	8

TABLE 5.3 – Valeurs de la fonction indicatrice d'Euler.

■ **Définition 69 — Ordre d'un élément d'un groupe multiplicatif.** Soit g un élément d'un groupe G multiplicatif dont l'élément neutre est noté 1_G. S'il existe un entier positif e tel que $g^e = 1_G$, alors le plus petit de ces entiers e est appelé l'ordre de l'élément g. On

dit alors que g est d'ordre fini.

■ **Exemple 50 — Ordre de $\overline{2}$ dans $(\mathbb{Z}/13\mathbb{Z},\times)^*$.** L'ordre de $\overline{2}$ dans $(\mathbb{Z}/13\mathbb{Z},\times)^*$ peut être visualisé sur le tableau 5.4. Il vaut 12.

k	0	1	2	3	4	5	6	7	8	9	10	11	**12**	13	14
2^k	1	2	4	8	3	6	12	11	9	5	10	7	**1**	2	4

TABLE 5.4 – Détermination de l'ordre de $\overline{2}$ dans $(\mathbb{Z}/13\mathbb{Z},\times)^*$

■ **Définition 70 — Sous-groupe engendré par un élément.** Soit g un élément d'un groupe G. On note $\langle g\rangle = \{g^k, k\in\mathbb{Z}\}$ le sous-groupe des éléments de G généré par g.

Si l'ordre de g est fini et vaut e alors $\langle g\rangle = \{g^k, 0\leqslant k<e\}$. On a de plus $\forall x\in\mathbb{Z}, g^x = g^{x \mod e}$.

■ **Exemple 51 — $\langle g\rangle$ dans $(\mathbb{Z}/13\mathbb{Z},\times)^*$.** D'après le tableau 5.4, on a $\langle 2\rangle = (\mathbb{Z}/13\mathbb{Z},\times)^*$.

Par contre, on a $\langle 4\rangle = \{k+13\mathbb{Z}, k=1,3,4,9,10,12\}$ et $4^6 \mod 13 = 1$.

■ **Définition 71 — Groupe cyclique.** Soit G un groupe engendré par un de ses éléments g. On a $G=\langle g\rangle$. Alors G est dit cyclique et g est un générateur.

■ **Exemple 52 — $(\mathbb{Z}/13\mathbb{Z},\times)^*$ est cyclique.** Comme l'ordre de 2 est 12 dans $(\mathbb{Z}/13\mathbb{Z},\times)^*$, $(\mathbb{Z}/13\mathbb{Z},\times)^*$ est un groupe cyclique.

Théorème 27 — Petit théorème de Fermat. Soit $a\in\mathbb{Z}$ et $p\in\mathbb{P}$ ou bien $\text{PGCD}(a,p)=1$. Alors

$$a^{p-1}\equiv 1 \mod p \tag{5.8}$$

Euler a généralisé ce théorème. Pour tout entier naturel $n\geqslant 2$ on a :

$$a^{\varphi(n)}\equiv 1 \mod n \tag{5.9}$$

Démonstration. Démonstration par récurrence sur a en utilisant la formule du binôme de Newton $(1+a)^p$ et en remarquant que tous les coefficients binomiaux à l'exception du premier et du dernier sont des multiples de p car p est premier.

Soit $a\in\mathbb{Z}$, on désigne par $\mathcal{P}(a)$ la propriété :

$$\forall p\in\mathbb{P}, \mathcal{P}(a) : a^{p-1}\equiv 1 \mod p \tag{5.10}$$

Initialisation : pour $a=0$, on a bien $\forall p\in\mathbb{P}, 0^{p-1}=1=1 \mod p$

Hérédité : supposons que la propriété $\mathcal{P}(a)$ est vraie et calculons $(a+1)^p$.

$$(a+1)^p = \sum_{k=0}^{p}\binom{p}{k}a^k \tag{5.11}$$

$$= \binom{p}{0}a^0 + \binom{p}{1}a^1 + \binom{p}{2}a^2 \ldots + \binom{p}{p-1}a^{p-1} + \binom{p}{p}a^p \tag{5.12}$$

Or, pour tout i appartenant à l'intervalle entier $[\![1, p-1]\!]$, on a :

$$\begin{aligned}\binom{p}{i} =& \frac{p!}{i!(p-i)!} \\ =& \frac{p \times (p-1) \ldots \times (p-i) \times \ldots 2 \times 1}{i!(p-i) \times \ldots \times 2 \times 1} \\ =& \frac{p \times (p-1) \ldots \times (p-i+1)}{i!} \\ =& p \times \frac{(p-1) \ldots \times (p-i+1)}{i!} \\ \equiv& 0 \mod p\end{aligned}$$

On a donc :

$$\begin{aligned}(a+1)^p =& \binom{p}{0} a^0 + \binom{p}{p} a^p && (5.13)\\ =& 1 + a^p && (5.14)\\ =& 1 + a \times a^{p-1} && (5.15)\end{aligned}$$

Or, d'après l'hypothèse de récurrence, on a $a^{p-1} \equiv 1 \mod p$. D'où :

$$(a+1)^p \mod p \equiv 1 + a \mod p \tag{5.16}$$

et donc, en simplifiant par $a+1$:

$$(a+1)^{p-1} \mod p \equiv 1 \mod p \tag{5.17}$$

$\mathcal{P}(a+1)$ est vraie. Donc $\mathcal{P}(a)$ est vraie pour tout $a \in \mathbb{Z}$.

On peut également le démontrer d'une manière plus générale en utilisant la propriété des ordres des éléments d'un groupe qui divise toujours l'ordre d'un groupe. ■

Théorème 28 — Soit $n \in \mathbb{N}$. Le nombre de générateur d'un groupe cyclique est $\varphi(n)$.

Théorème 29 — Si $p \in \mathbb{P}$ alors $(\mathbb{Z}/p\mathbb{Z}, \times)^*$ est cyclique. De plus, il possède $\varphi(p-1)$ générateurs d'ordre $p-1$

5.5 Anneaux euclidiens

■ **Définition 72 — Anneau euclidien.** Un anneau $(A, +\times)$ est dit euclidien si et seulement si on peut le munir d'un **stasthme**, c'est à dire d'une application $\sigma : A \setminus \{0\} \longrightarrow \mathbb{N}$ telle que pour tout élément a et b de A on peut trouver q et r tels que $a = bq + r$ et $r = 0$ ou $\sigma(r) < \sigma(b)$.

■ **Exemple 53 — Stasthmes.** On peut recenser quelques stasthmes usuels :

- $(\mathbb{Z}, +, \times)$ muni du stasthme valeur absolue est euclidien.
- Pour tout corps $\mathbb{K}$, $(\mathbb{K}[X], +, \times)$ muni du stasthme degré est euclidien.

■ **Définition 73 — Irréductibilité.** On dit qu'un élément p d'un anneau A est irréductible si et seulement si :

- $p \neq 0_A$,
- p n'est pas inversible,
- p est premier avec tout élément qu'il ne divise pas (cf. définition 32).

On peut également formuler cette définition ainsi : p est irréductible si et seulement si $p = ab \implies a \in A^*$ *ou* $b \in A^*$. C'est à dire que les seuls diviseurs de p sont les inversibles de A.

Exprimée en contraposément, sur un anneau de polynômes $(\mathbb{K}[X], +, \times)$ et sur tout corps, les polynômes factorisables (réductibles) sont ceux qui se décomposent en produit de plusieurs polynômes non constants. Les polynômes de degré 1 sont donc toujours irréductibles.

Un polynôme de degré 2 produit de deux polynômes de degré 1 est factorisable : il possède des racines. Par exemple, $X^2+2X+1 = (X+1)(X+1)$ n'est donc pas irréductible.

Par contraposée, un polynôme de degré 2 ne possédant pas de racines est irréductible. Dans le cas de $(\mathbb{R}[X], +, \times)$, tous les polynômes de degré 2 dont le discriminant est négatif sont irréductibles. Par exemple, $X^2 + X + 1$ est irréductible dans $(\mathbb{R}[X], +, \times)$.

■ **Exemple 54 — Éléments irréductibles.** Dans le cas des anneaux euclidiens (et donc principaux) que l'on manipule couramment, les éléments irréductibles sont équivalents aux éléments premiers. Par exemple :

- Les irréductibles de $\mathbb{Z}$ sont les nombres premiers.
- Les irréductibles de $\mathbb{K}[X]$ sont les $(X - a)$ si $\mathbb{K}$ est algébriquement clos ($\mathbb{C}$ par exemple). On note que $(X - a)$ n'est divisible que par 1 ou lui-même. Il est donc premier.
- Dans le cas de $\mathbb{R}[X]$, les irréductibles sont les polynômes de degré un ou deux dont le discriminant est négatif. Ces polynômes sont premiers également.

Néanmoins, dans les corps finis, les éléments premiers sont irréductibles, mais l'inverse n'est pas vrai (cf. 5.10). Enfin, il faut également remarquer que les irréductibles d'un corps $\mathbb{K}$ sont des éléments de $\mathbb{K}^*$.

Pour tout anneau euclidien :

1. on peut définir la relation d'équivalence divisibilité comme dans la définition 59,
2. il existe la notion de plus grand commun diviseur (cf. théorème 6),
3. la relation de Bezout est vérifiée (cf. théorème 9),
4. le théorème de Gauss-Euclide est valide (cf. théorème 13),
5. l'ensemble est factoriel, c'est à dire qu'on peut décomposer les éléments en facteurs irréductibles (cf. théorème 15),
6. on peut y développer une arithmétique modulaire avec une anneau quotient.

5.6 Corps premiers finis $\mathbb{F}_p$

■ **Définition 74 — Corps fini.** On dit qu'un corps est fini lorsque le cardinal de l'ensemble sous-jacent l'est.

Théorème 30 — $\mathbb{Z}/p\mathbb{Z}$ est un corps si et seulement si $p \in \mathbb{P}$.

Démonstration. Tous les éléments de l'anneau $(\mathbb{Z}/p\mathbb{Z}, +, \times)$ sauf 0 sont inversibles car p est premier (cf. théorème 24). C'est donc un corps. ■

R Comme $\text{card}(\mathbb{Z}/p\mathbb{Z}) = p$ d'après le théorème 22, $\mathbb{Z}/p\mathbb{Z}$ est un **corps fini**.

R Comme $p \in \mathbb{P}$, $\mathbb{Z}/p\mathbb{Z}$ est un **corps premier fini**.

R Comme qui peut le plus peut le moins, un corps est un anneau. Il possède donc une caractéristique. La caractéristique de $\mathbb{Z}/p\mathbb{Z}, p \in \mathbb{P}$ est p. En effet, $\overline{p} = \underbrace{\overline{1} + \ldots + \overline{1}}_{\text{p fois}} = \overline{0}$

■ **Définition 75 — Corps premier fini.** On appelle les corps $\mathbb{Z}/p\mathbb{Z}$ avec $p \in \mathbb{P}$ les corps premiers finis à p éléments. On les note $\mathbb{F}_p$. Ils possèdent p éléments et sont de caractéristique p.

Proposition 17 — Propriétés du corps $\mathbb{Z}/2\mathbb{Z}$. $\mathbb{Z}/2\mathbb{Z}$ est un corps fini à deux éléments, car 2 est un nombre premier. On a : $\overline{0} = \{0, 2, -2, 4, -4, \cdots\}$ l'élément neutre de $+$ et $\overline{1} = \{1, -1, 3, -3, \cdots\}$ l'élément neutre de $\times$.

Les tableaux 5.5 et 5.6 montrent les tables d'addition et de multiplication dans ce anneau. L'addition et la soustraction sont donc une seule et même opération dans $\mathbb{Z}/2\mathbb{Z}$. On peut remplacer l'addition par un **ou exclusif** (xor) et la multiplication par un **et** (and).

On observe que seul $\overline{1}$ est inversible car $\overline{1} \times \overline{1} = \overline{1}$.

$+$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

TABLE 5.5 – Table d'addition dans $\mathbb{Z}/2\mathbb{Z}$.

$\times$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$

TABLE 5.6 – Table de multiplication dans $\mathbb{Z}/2\mathbb{Z}$.

Proposition 18 — $\mathbb{Z}/2\mathbb{Z}$ est le corps fini de plus petit cardinal. Il est noté $\mathbb{F}_2$.

Démonstration. Pour que ce soit un corps, il faut nécessairement un élément neutre pour chaque loi de composition interne $+$ et $\times$. Donc, on ne peut pas un corps fini dont le cardinal est inférieur à 2. Or $\text{card}(\mathbb{F}_2) = 2$. ■

■ **Exemple 55 — $\mathbb{F}_2$, le corps premier fini des électroniciens et informaticiens.** Les éléments de $\mathbb{F}_2$ sont 0 et 1. L'addition et la soustraction dans $\mathbb{F}_2$ sont réalisées par un ou exclusif (cf. tableau 5.5). La multiplication est celle décrite sur le tableau 5.6.

C'est le plus petit des corps finis. Un seul élément est inversible : 1.

5.7 Anneaux et polynômes

5.7.1 Généralités

■ **Définition 76 — Indéterminée X.** Dans le but de définir des polynômes dont la variable n'est pas un nombre, on définit X comme une indéterminée : cette variable peut représenter un nombre, une matrice ou une fonction.

■ **Définition 77 — Polynôme.** Soient $a_0, a_1, a_2, \cdots, a_n$ des éléments d'un corps $\mathbb{K}$. On dit que P est un polynôme à coefficient dans $\mathbb{K}$ lorsque :

$$P(X) = a_0 + a_1 X + \cdots + a_n X^n = \sum_{k=0}^{n} a_k X^k \tag{5.18}$$

Cette écriture d'un polynôme est dite canonique. X, l'indéterminée, est un polynôme. Les a_i sont appelés coefficients.

■ **Définition 78 — Ensemble des polynômes.** On note $\mathbb{K}[X]$ l'ensemble des polynômes d'interminée X à coefficients dans $\mathbb{K}$.

■ **Définition 79 — Polynôme nul.** Le polynôme dont tous les coefficients sont nuls est noté $O_{\mathbb{K}[X]}$.

■ **Définition 80 — Degré d'un polynôme.** On appelle degré d'un polynôme $P(X) = \sum_{k=0}^{n} a_k X^k$ de $\mathbb{K}[X]$ et on note $\deg(P)$ le stasthme définit par :

- n si $a_n \neq 0$, c'est à dire l'indice du coefficient non nul le plus élevé, autrement appelé coefficient dominant car lié à la puissance de X la plus élevée.
- $-\infty$ si $P(X) = 0_{\mathbb{K}[X]}$.

Théorème 31 — Deux polynômes sont égaux si et seulement s'ils ont même degré et si tous leurs coefficients sont égaux..

■ **Définition 81 — Opérations sur le polynômes.** On peut définir les opérations suivantes sur les polynômes de $\mathbb{K}[X]$:

- multiplication par un scalaire de $\mathbb{K}$: $\forall \lambda \in \mathbb{K}, \lambda P = \sum_{k=0}^{n} \lambda a_k X^k$
- addition de deux polynômes : Soit P et Q deux polynômes. Alors on définit $P + Q = \sum_{k=0}^{n} \lambda a_k X^k + \sum_{k=0}^{m} \lambda b_k X^k$ en regroupant les termes de même degré k. On obtient ainsi une forme canonique.

- multiplication de deux polynômes : Soit P et Q deux polynômes. Alors on définit $PQ = \sum_{k=0}^{n} \lambda a_k X^k \times \sum_{k=0}^{m} \lambda b_k X^k$ que l'on développe selon les règles classiques de calcul. On rassemble ensuite les termes pour trouver une forme canonique.

Proposition 19 — Opération sur les degrés. Soit P et Q deux éléments de $\mathbb{K}[X]$. Alors on a :

$$\deg(P+Q) \leqslant \max(\deg(P), \deg(Q)) \tag{5.19}$$
$$\deg(PQ) = \deg(P) + \deg(Q) \tag{5.20}$$

■ **Définition 82 — Racine d'un polynôme.** On dit que $a \in \mathbb{K}$ est racine d'un polynôme $P \in \mathbb{K}[X]$ si $P(a) = 0_{\mathbb{K}[X]}$

■ **Définition 83 — Polynôme monique ou unitaire.** Un polynôme monique possède un coefficient dominant égal à 1. Par conséquent, il est de la forme :

$$X^n + a_{n-1}X^{n-1} + \ldots + a_2X^2 + a_1X + a_0 \tag{5.21}$$

Un polynôme non monique possède les mêmes racines que son monique associé, c'est à dire le polynôme dont on a divisé les coefficients par a_n (qui devient donc monique). En effet, multiplier un polynôme par un scalaire ne modifie pas les racines d'un polynôme.

Si l'on cherche les racines d'un polynôme, il suffit donc de considérer les racines du polynôme monique associé.

■ **Exemple 56 — Racines de $2X^2 + 4X - 6$.** Les racines de $2X^2 + 4X - 6$ sont les mêmes que celles de $X^2 + 2X - 3$. En effet, $2X^2 + 4X - 6 = 2(X-1)(X-3)$ et $X^2 + 2X - 3 = (X-1)(X-3)$.

■ **Définition 84 — Divisibilité de deux polynômes.** Soit A et B deux éléments de $\mathbb{K}[X]$. On dit que B divise A et on note $B \mid A$ s'il existe un polynôme Q tel que $A = BQ$.

Théorème 32 — Division euclidienne de deux polynômes. Soit A et B deux éléments de $\mathbb{K}[X]$ et $B \neq O_{\mathbb{K}[X]}$. Alors il n'existe qu'un seul couple (Q, R) de polynômes tels que :

$$A = BQ + R \tag{5.22}$$
$$et\ \deg(R) < \deg(B) \tag{5.23}$$

Démonstration. Comme dans l'anneau des entiers, on peut procéder par récurrence sur le degré de A pour prouver l'existence. Pour l'unicité, on procède également par l'absurde. ■

M **Méthode 6 — Comment diviser deux polynômes dans $\mathbb{R}[X]$** Soient les polynômes $A(X) = X^4 - 5X^3 + 6X - 2$ et $B(X) = X^2 - 2$, deux polynômes de $\mathbb{R}[X]$. On cherche donc deux polynômes Q et R tels que $A = BQ + R$ et $\deg(R) < \deg(B)$.

$$
\begin{array}{rl|l}
(A) & X^4 - 5X^3 \quad + 6X - 2 & X^2 - 2 \quad (B) \\
& -(X^4 \quad - 2X^2) & \\
\hline
& -5X^3 + 2X^2 + 6X - 2 & X^2 - 5X + 2 \quad (Q) \\
& -(-5X^3 \quad + 10X) & \\
\hline
& 2X^2 - 4X - 2 & \\
& -(2X^2 \quad - 4) & \\
\hline
& (R) \quad - 4X + 2 &
\end{array}
$$

On a donc $Q(X) = X^2 - 5X + 2$ et $R(X) = -4X + 2$. On vérifie qu'on a bien $\deg(R) = 1 < \deg(B) = 2$

Proposition 20 — B divise A si et seulement si le reste de la division euclidienne de A par B est nul.

Théorème 33 — $(\mathbb{K}[X], +, \times)$ muni du stasthme défini par le degré d'un polynôme est un anneau euclidien.

Démonstration. Cela découle de la définition de la division euclidienne pour les polynômes. ■

5.7.2 Anneau $\mathbb{F}_2[X]$

On peut également utiliser un corps premier pour créer un anneau des polynômes. Par exemple, construire $\mathbb{F}_2[X]$ à partir de $\mathbb{F}_2$. Mais à quelle arithmétique aboutit-on alors ?

R La notation $\mathbb{R}_n[X]$ signifie l'ensemble de polynômes de degrés n à coefficients dans $\mathbb{R}$. Il faut remarquer qu'avec la notation $\mathbb{F}_p[X]$ il n'en est pas ainsi : il s'agit bien de l'ensemble des polynômes de **degré quelconque** dont les coefficients sont à valeur dans $\mathbb{F}_p$. Dans ce cas, p est le nombre d'éléments du corps fini, pas le degré maximal des polynômes.

M **Méthode 7 — Comment additionner et multiplier deux polynômes dans $\mathbb{F}_2[X]$?**
Tout d'abord, il faut remarquer que les seuls éléments de $\mathbb{F}_2$ sont 0 et 1. Ainsi, les coefficients des polynômes ne peuvent être que 0 ou 1. En effet, on a : $\overline{0} = \{0, 2, -2, 4, -4, \cdots\}$ et $\overline{1} = \{1, -1, 3, -3, \cdots\}$. Ces éléments seront maintenant notés 0 et 1.

Les tableaux 5.5 et 5.6 montrent les tables d'addition et de multiplication dans l'anneau $\mathbb{F}_2[X]$. Tout comme dans le corps $\mathbb{F}_2$, l'addition et la soustraction sont donc une seule et même opération, le ou exclusif.

$$\forall P \in \mathbb{F}_2[X], P + P = 0 \tag{5.24}$$
$$\forall P, Q \in \mathbb{F}_2[X], P + Q = P - Q \tag{5.25}$$

On procède de même pour la multiplication, c'est à dire en multipliant termes à termes classiquement puis en appliquant l'arithmétique de $\mathbb{F}_2$.

■ **Exemple 57 — Calculs dans $\mathbb{F}_2[X]$.** Soit $X^2 + X$ dans $\mathbb{F}_2[X]$. Ce polynôme a pour racine 1 et 0. On peut donc l'écrire $X(X-1) = X(X+1)$.

Soit $X^2 + 1$ dans $\mathbb{F}_2[X]$. Ce polynôme a pour racine 1, car dans $\mathbb{F}_2$, $1 + 1 = 0$. On peut donc l'écrire $X^2 + 1 = (X-1)(X-1) = X^2 - 2X + 1$, car $2X = 0$.

Soit $P(X) = X^2 + X + 1$ dans $\mathbb{F}_2[X]$. Ce polynôme n'a pas racines dans $\mathbb{F}_2$, car $P(0) = 1$ et $P(1) = 1$. Il est irréductible dans $\mathbb{F}_2[X]$.

M **Méthode 8 — Comment diviser deux polynômes dans $\mathbb{F}_2[X]$** Soient les polynômes $A(X) = X^4 + X^3 + X$ et $B(X) = X^2 + 1$, deux polynômes de $\mathbb{F}_2[X]$. On cherche donc deux polynômes Q et R tels que $A = BQ + R$ et $\deg(R) < \deg(B)$.

$$
\begin{array}{rrcll}
(A) & X^4 + X^3 \quad + X & | & X^2 + 1 & (B) \\
 & -X^4 \quad - X^2 & | & \multicolumn{2}{l}{\text{---------------}} \\
\cline{1-2}
 & -X^3 - X^2 + X & | & X^2 + X + 1 & (Q) \\
 & -X^3 \quad - X & | & & \\
\cline{2-2}
 & -X^2 & | & & \\
 & -X^2 + 1 & | & & \\
\cline{2-2}
 & (R) \quad 1 & | & &
\end{array}
$$

On a donc $Q(X) = X^2 + X + 1$ et $R(X) = 1$. On vérifie qu'on a bien $\deg(R) = 0 < \deg(B) = 2$. Par ailleurs, on peut vérifier en développant que :

$$
\begin{aligned}
BQ(X) + R(X) &= (X^2+1)(X^2+X+1) + 1 \\
&= X^4 + X^3 + X^2 + X^2 + X + 1 + 1 \\
&= X^4 + X^3 + X \\
&= A(X)
\end{aligned}
$$

■ **Exemple 58 — Polynôme réductible dans $\mathbb{F}_2[X]$.** Le polynôme $X^3 + X^2 + X + 1$ est réductible dans $\mathbb{F}_2[X]$. En effet, 1 est racine de ce polynôme et on peut donc le factoriser ainsi : $X3 + X^2 + X + 1 = (X+1)(X^2+1)$.

■ **Exemple 59 — Polynôme irréductible dans $\mathbb{F}_2[X]$.** Il n'existe qu'un seul polynôme de degré 2 irréductible dans $\mathbb{F}_2[X]$.

Les polynômes de degré 2 sont : X^2, X^2+1 et X^2+X+1. Les deux premiers sont factorisables. Les facteurs possibles sont X si 0 est racine et $1 + X$ si 1 est racine.

La factorisation de X^2 est $X^2 = X \times X$ et X n'est pas un élément inversible de $\mathbb{F}_2[X]$ (cf. définition 73). Il est donc réductible dans $\mathbb{F}_2[X]$.

De même, $X^2 + 1 = (X+1)^2 = (X+1) \times (X+1)$ et $X + 1$ n'est pas un élément inversible de $\mathbb{F}_2[X]$. Il est donc réductible dans $\mathbb{F}_2[X]$.

Enfin, ni 0 ni 1 ne sont racines de $X^2 + X + 1$. On ne peut donc pas le factoriser et il est donc irréductible dans $\mathbb{F}_2[X]$. Il y a donc bien un seul polynôme irréductible de degré 2 dans $\mathbb{F}_2[X]$ comme le prédit le théorème 43.

Plus loin dans ce cours, il est montré que les racines des polynômes de degré deux

irréductibles sur $\mathbb{F}_2$ sont les éléments de $\mathbb{F}_{2^2} = \mathbb{F}_4$ qui n'appartiennent pas à $\mathbb{F}_2$. Ce sont donc les deux générateurs du groupe multiplicatif $\mathbb{F}_4^*$ à trois éléments.

5.8 Construction des corps finis

L'anneau $(\mathbb{Z}/m\mathbb{Z}, +, \times)$ a été construit grâce à la relation de congruence et à l'anneau $(\mathbb{Z}, +, \times)$. Comme $\mathbb{K}[X]$ est également euclidien, on peut procéder de la même manière et construire un nouvel anneau à partir l'anneau des polynômes $\mathbb{K}[X]$ en choisissant un polynôme irréductible m pour l'opération de modulo. C'est ainsi qu'on peut parler de l'anneau des restes des polynômes que l'on note $\mathbb{K}[X]/m$.

Corollaire 1 — Anneau des restes des polynômes. Une conséquence du théorème 33 est qu'on peut définir un anneau des restes des polynômes ou anneau quotient $(\mathbb{K}[X]/m, +, \times)$ en utilisant un polynôme m irréductible de degré n.

R En utilisant un corps premier, la construction d'un anneau quotient avec un élément irréductible permet de construire de nouveaux corps finis de type $\mathbb{F}_p[X]/m$, différents des corps premiers.

Proposition 21 — L'anneau $(\mathbb{F}_p[X]/m, +, \times)$ est un corps si et seulement si m est irréductible.

V **Vocabulary 4 — Corps de rupture (Splitting Fields)** ⇜⇝ En anglais, les corps $(\mathbb{F}_p[X]/m, +, \times)$ ainsi généré se nomme Splitting Fields, corps de rupture en français.

Théorème 34 — Le cardinal du corps $\mathbb{F}_p[X]/m$ avec $m \in \mathbb{F}_p[X]$ irréductible et $\deg(m) = n$ est p^n .

Démonstration. D'après la relation de congruence et les propriétés de la division euclidienne, $\mathbb{F}_p[X]/m$ est l'ensemble des polynômes à coefficients dans $\mathbb{F}_p$ dont le degré est strictement inférieur à n. Ce sont des polynômes à n coefficients a_i à valeur dans $\mathbb{F}_p$. Pour chaque a_i, on a donc p valeurs possibles. D'où le résultat. ■

Théorème 35 — Isomorphisme par rapport à un polynôme de degré n. Pour tout nombre premier p et tout entier n strictement positif, il existe un corps fini de cardinal $q = p^n$. Il est noté $\mathbb{F}_q$ ou $\mathbb{F}_{p^n}$. Ce corps est unique à un isomorphisme près.

R Cette méthode de construction des corps finis ne dépend donc que du degré du polynôme irréductible choisi dans $\mathbb{F}_p[X]$. Un corps fini $\mathbb{F}_p[X]/m$ avec $\deg(m) = n$ sera donc noté $\mathbb{F}_{p^n}$ d'après le résultat du théorème 45.

V **Vocabulary 5 — Galois Fields - Corps de Galois** $GF(p)$ ⇜⇝ On peut désigner les corps finis $\mathbb{F}_p$ par le terme *corps de Galois*, en hommage à Évariste Galois qui a été à l'origine de leur découverte. Les anglo-saxons les notent parfois $GF(p)$.

Nous allons étudier ces corps en détails dans les sections suivantes et notamment la manière dont on peut conduire les calculs.

5.9 Propriétés fondamentales des corps finis

Théorème 36 — Tout corps fini est commutatif.

Théorème 37 — Le groupe multiplicatif $\mathbb{F}_q^*$ d'un corps fini $\mathbb{F}_q$ est cyclique. Il possède $\varphi(q-1)$ éléments générateurs d'ordre $q-1$.

Théorème 38 — Tout corps fini $\mathbb{F}_{p^n}$ a pour caractéristique un nombre premier p et pour cardinal une puissance n de ce nombre.

■ **Exemple 60 — Caractéristiques de corps finis.** Dans le cas de $\mathbb{F}_3$, comme 3 est premier, on peut assimiler ce corps à $\mathbb{Z}/3\mathbb{Z}$. La caractéristique de ce corps est 3, son cardinal 3^1.

Dans le cas de $\mathbb{F}_4$, comme il s'agit d'un corps fini non premier, il faut l'interpréter comme un anneau des restes par un polynôme de degré 2 dans $\mathbb{F}_2[X]$. Dans ce cas, la caractéristique de $\mathbb{F}_4$ est 2, puisque ses coefficients sont à valeur dans $\mathbb{F}_2$, et son cardinal 2^2.

R Il est important de noter pour la suite que pour tous les corps de caractéristique 2 (c'est le cas des corps $\mathbb{F}_{2^n}$), la soustraction et l'addition sont les mêmes opérations et on peut l'assimiler à un ou exclusif.

Un corps fini est élaboré à partir d'un polynôme irréductible à coefficients dans le corps premier sous-jacent. Il est donc important de pourvoir identifier ces polynômes irréductibles.

R Dans $\mathbb{F}_2[X]$, les seules racines possibles sont 0 ou 1. Donc, si ni 0 ni 1 ne sont racines, un polynôme de $\mathbb{F}_2[X]$ est irréductible.

Les corps finis possèdent des éléments générateurs qui permettent de construire tous les éléments du corps en prenant leurs puissances successives. Ces éléments générateurs sont les racines de polynômes qu'on appelle les polynômes premiers.

■ **Définition 85 — Polynôme primitif.** Un polynôme irréductible de $\mathbb{F}_p[X]$ est primitif si une de ses racines est un élément générateur de $\mathbb{F}_{p^n}$.

Théorème 39 — Degré d'un polynôme primitif. Un polynôme primitif de $\mathbb{F}_p[X]$ dont la racine est un élément générateur de $\mathbb{F}_{p^n}$ est de degré n.

R Les polynômes primitifs sont irréductibles, mais les irréductibles ne sont pas nécessairement primitifs.

■ **Définition 86 — Polynôme minimal de $\alpha \in \mathbb{F}_{p^n}$.** Soit $\alpha \in \mathbb{F}_{p^n}$. Soit $P \in \mathbb{F}_p[X]$ le polynôme monique de plus petit degré pour lequel $P(\alpha) = 0$. On dit que P est un polynôme minimal de $\alpha \in \mathbb{F}_{p^n}$.

Théorème 40 — Un polynôme minimal de $\alpha \in \mathbb{F}_{p^n}$ est irréductible et son degré est inférieur à n.

Théorème 41 — Si P est le polynôme minimal d'un élément de $\mathbb{F}_{p^n}$, alors P divise $X^{p^n} - X$.

Théorème 42 — Le polynôme minimal d'un élément générateur du corps $\mathbb{F}_{p^n}$ est un polynôme primitif. .

R Cela signifie que lorsqu'on cherche un polynôme primitif pour un générateur α de $\mathbb{F}_{p^n}$, il suffit de chercher parmi les moniques de degré n.

5.10 Arithmétique dans un corps fini

Si la convention veut que l'on représente la plupart du temps les éléments d'un corps premier fini $\mathbb{F}_p$ par les symboles $\overline{0}, \overline{1}, \ldots \overline{p-1}$ voire $0, 1, 2 \ldots, p-1$, il n'en est pas de même pour les corps finis en général. La raison en est que les éléments de ces corps finis ne se comportent pas du tout comme nombres que nous manipulons tous les jours.

Chaque corps fini possède sa propre arithmétique : les tables d'addition et de multiplication ne sont pas les mêmes ! C'est pourquoi il est nécessaire :

- de bien détailler ces mécanismes selon le corps fini utilisé,
- de choisir une représentation pour les éléments de ces corps qui ne prête pas à confusion.

■ **Exemple 61 — $\mathbb{F}_4$ et ses symboles.** $\mathbb{F}_4$ est le plus petit corps fini non premier. $\mathbb{F}_4 = \mathbb{F}_{2^2}$ est de caractéristique 2. Il possède quatre éléments. Comment les noter ?

La caractéristique de ce corps implique que pour tout élément λ, on a $\lambda + \lambda = 0_{\mathbb{F}_4}$, si on note l'élément neutre de l'addition $0_{\mathbb{F}_4}$. Cette remarque est valable pour l'élément neutre de la multiplication : $1_{\mathbb{F}_4} + 1_{\mathbb{F}_4} = 0_{\mathbb{F}_4}$ ce qui nous met déjà dans une situation qui n'est pas courante...

Pour les deux derniers éléments que l'on peut noter α et β, la table de multiplication nous dit que $\alpha\beta = 1_{\mathbb{F}_4}$. Ces éléments ne se comportent donc pas du tout comme 2 et 3. On n'a donc pas intérêt à les noter ainsi si on ne veut pas engendrer de la confusion.

Néanmoins, on conserve souvent la notation $0_{\mathbb{F}_p}$ pour l'élément neutre de l'addition des corps finis.

5.10.1 Tables de $\mathbb{F}_4$

Pour aller plus loin dans cette réflexion, nous avons besoin des tables d'addition et de multiplication dans $\mathbb{F}_4$ (cf. tableaux 5.7 et 5.8). Ces tables découlent essentiellement des raisons suivantes :

- $\mathbb{F}_4$ est un groupe commutatif pour l'addition,
- $\mathbb{F}_4$ est un groupe commutatif pour la multiplication.

Cela implique que :

- le résultat de toute opération dans $\mathbb{F}_4$ appartient à $\mathbb{F}_4$,
- il existe une élément neutre pour l'addition,
- tous les éléments possèdent un symétrique,

- il existe un élément neutre pour la multiplication,
- tous les éléments non nuls sont inversibles.

$+$	0	1	α	β
0	0	1	α	β
1	1	0	β	α
α	α	β	0	1
β	β	α	1	0

TABLE 5.7 – Table d'addition dans $\mathbb{F}_4$.

$\times$	0	1	α	β
0	0	0	0	0
1	0	1	α	β
α	0	α	β	1
β	0	β	1	α

TABLE 5.8 – Table de multiplication dans $\mathbb{F}_4$.

La table d'addition est facile à établir pour l'opérande 0 qui est l'élément neutre de l'addition. Pour $\mathbb{F}_4$, comme la caractéristique du corps est 2, il est également facile de compléter la diagonale de la table d'addition : elle est nulle.

On peut ensuite déduire que $1 + \alpha = \beta$. Si $1 + \alpha$ valait 1 alors on aurait $\alpha = 0$ ce qui n'est pas vrai, le corps ayant quatre éléments distincts. On ne peut pas avoir $1 + \alpha = \alpha$, car 1 n'est pas l'élément neutre de l'addition. On ne peut pas avoir $1 + \alpha = 0$ car $1 = -1$ dans $\mathbb{F}_4$, car de caractéristique 2 et $\alpha \neq -1$, c'est un élément distinct. La seule solution est donc $1 + \alpha = \beta$.

Par élimination, on a ensuite nécessairement $1 + \beta = \alpha$, c'est le seul élément distinct qu'il reste sur la ligne. Finalement, $\alpha + \beta = 1 + \beta + \beta = 1$. La table d'addition 5.7 est complète. On remarque qu'on pourrait se passer du symbole β en le notant $1 + \alpha$ et ne garder donc que trois symboles $0, 1, \alpha$.

Pour la multiplication, les lignes et les colonnes dont une opérande est 0 ou 1 sont triviales. Il reste à étudier les produits α^2 et de $\alpha\beta$.

On ne peut pas avoir $\alpha\beta = \alpha$, sinon β serait l'élément neutre 1. On ne peut pas avoir $\alpha\beta = \beta$, sinon α serait l'élément neutre 1. On ne peut pas avoir $\alpha\beta = 0$ car $\mathbb{F}_4$ est un corps : il ne possède donc pas de diviseurs de zéro et $\alpha \neq 0$ et $\beta \neq 0$ puisque $\mathbb{F}_4$ possède quatre éléments distincts. On a donc nécessairement $\alpha\beta = \beta\alpha = 1$. On en déduit que $\alpha^2 = \beta$ et $\beta^2 = \alpha$, ces éléments sont inverses l'un de l'autre. La table de multiplication est complétée.

5.10.2 Représentation des éléments d'un corps fini

On considère maintenant un corps fini $\mathbb{F}_{p^n}$ de caractéristique p et de le cardinal p^n. On peut considérer ce corps comme l'extension en dimension n du corps $\mathbb{Z}/p\mathbb{Z}$. De ce point de vue, on peut le représenter comme un espace vectoriel de dimension n sur le corps $\mathbb{F}_p$: $\mathbb{F}_{p^n} \mapsto (\mathbb{F}_p, +, .)^n$.

■ **Exemple 62 — $\mathbb{F}_4$ vu comme un espace vectoriel de dimension 2 sur $\mathbb{F}_2$.** Considérons $\mathbb{F}_4 = 0, 1, \alpha, \beta$. Chaque élément de ce corps peut être considéré comme un élément de l'espace vectoriel $(\mathbb{F}_2, +, .)^2$ ou extension quadratique de $\mathbb{F}_2$. On notera alors $\mathbb{F}_4 = \{(0,0), (0,1), (1,0), (1,1)\}$ et on appellera les sont des vecteurs.

L'avantage de cette représentation est que l'on peut faire des calculs facilement sur un ordinateur : l'addition se fait alors élément par élément avec un ou exclusif ce qui est confortable pour l'électronicien ! La multiplication par contre n'est pas élémentaire.

Une autre manière de représenter les éléments d'un corps fini est de choisir un élément générateur et d'exprimer tous les autres d'après une puissance de celui-ci. On sait que ce corps possède $\varphi(3) = 2$ éléments générateurs, car le groupe multiplicatif sous-jacent au corps fini $\mathbb{F}_4$ compte trois éléments. D'après la table de multiplication, on voit que α et β sont ces générateurs.

■ **Exemple 63 — $\mathbb{F}_4$ généré par un élément.** Considérons de nouveau $\mathbb{F}_4 = 0, 1, \alpha, \beta$. α est un élément générateur de $\mathbb{F}_4$ d'après la table de multiplication 5.8. Chaque élément de ce corps peut être considéré comme une puissance de α. On notera alors $\mathbb{F}_4 = \{0, \alpha, \alpha^2, \alpha^3\}$. Cela revient à exprimer tous les éléments du corps à partir d'une racine d'un polynôme primitif.

L'avantage de cette représentation est que la multiplication est facile : il suffit de choisir la puissance égale à l'addition des exposants modulo $p - 1$. L'inconvénient est que l'addition est moins évidente...

Pour représenter correctement les éléments de ces corps, on choisit souvent une représentation mixte, selon les besoins de l'application.

Une dernière représentation est possible : considérer que $\mathbb{F}_4$ est bien le résultat de $\mathbb{F}_2[X]/m$ ou m est un polynôme de degré 2. Les éléments du corps sont alors des polynômes. Ces polynômes peuvent être notés tels quels avec l'indéterminée X ou sous forme binaire ou hexadécimale en ne conservant que les coefficients binaires des polynômes (0 ou 1).

■ **Exemple 64 — $\mathbb{F}_4$ et les polynômes.** $\mathbb{F}_4$ peut être interprété comme le corps $\mathbb{F}_2[X]/(X^2 + X + 1)$.

On vérifie que le polynôme $X^2 + X + 1$ est irréductible dans $\mathbb{F}_2[X]$: ni 0 ni 1 ne sont des racines.

Le polynôme $X^2 + X + 1$ est de degré 2. Conformément au théorème 45, il permet donc de créer $\mathbb{F}_4$.

On peut vérifier également que ce polynôme est primitif : une racine α de ce polynôme dans $\mathbb{F}_4$ vérifie $\alpha^2 + \alpha + 1 = 0$, c'est à dire $\alpha^2 = -\alpha - 1 = \alpha + 1$ car $\mathbb{F}_4$ est de caractéristique 2. On retrouve ainsi l'arithmétique décrite ci-dessus ! De plus, α est un élément générateur de $\mathbb{F}_4$. Donc, $X^2 + X + 1$ est un polynôme primitif de $\mathbb{F}_4$.

Le tableau 5.9 présente une synthèse des différentes représentations possibles pour les éléments de $\mathbb{F}_4$.

■ **Exemple 65 — Polynôme primitif de $\mathbb{F}_4$.** Soit α un générateur de $\mathbb{F}_4$. On cherche les polynômes primitifs de $\mathbb{F}_4$ dont α est racine. Ce sont des polynômes moniques de degré

Espace vectoriel	**Élément réduit**	**Puissance d'un générateur**	Polynôme	Polynôme (Binaire)	Polynôme (décimal)
$(0,0)$	0	0	0	00	0
$(0,1)$	1	α^0	1	01	1
$(1,0)$	α	α^1	X	10	2
$(1,1)$	$1+\alpha$	α^2	$X+1$	11	3

TABLE 5.9 – Représentations possibles des éléments de $\mathbb{F}_4$.

2, car $\mathbb{F}_4 = \mathbb{F}_{2^2}$. Les seuls polynômes qui conviennent sont :

$$X^2 \tag{5.26}$$
$$X^2+1 \tag{5.27}$$
$$X^2+X+1 \tag{5.28}$$

0 est racine du premier, 1 du second. Le seul irréductible dans $\mathbb{F}_2[X]$ est X^2+X+1.

On vérifie, grâce aux tables de $\mathbb{F}_4$, que α et β sont bien racines de ce polynôme : $\alpha^2+\alpha+1 = \alpha+1+\alpha+1 = 0$ et $\beta^2+\beta+1 = \beta+\beta = 0$. C'est donc un polynôme primitif.

Élément	**Polynôme minimal associé**	**Polynôme primitif ?**
0	X	non
1	$1+X$	non
α	X^2+X+1	oui, car α est un générateur de $\mathbb{F}_4$.
$1+\alpha$	X^2+X+1	oui, car $1+\alpha$ est un générateur de $\mathbb{F}_4$.

TABLE 5.10 – Éléments de $\mathbb{F}_4$ et polynômes minimaux associés.

5.11 Opérations dans $\mathbb{F}_{256}$

Le corps $\mathbb{F}_{256}$ possède 2^8 éléments. Ce corps est utilisé par l'algorithme de chiffrement AES si bien qu'on l'appelle parfois le corps de Rinjdael, expression issue des noms des créateurs d'AES. C'est pourquoi il est intéressant d'examiner l'arithmétique dans ce corps.

Le polynôme $m(X) = X^8+X^4+X^3+X+1$ a été choisi par les créateurs d'AES pour créer le corps $\mathbb{F}_{256}$. C'est un polynôme monique de degré 8 irréductible dans $\mathbb{F}_2[X]$ puisque ni 0 ni 1 ne sont racines. Il y a en tout 30 polynômes irréductibles de degré 8 dans $\mathbb{F}_2[X]$ (cf. exemple 70). Il faut noter que ce polynôme n'est pas primitif, certains le sont.

On peut lister tous les éléments de ce corps à partir d'un élément générateur g comme le montre le tableau 5.11. Le polynôme $X+1$ représente un tel élément générateur dont l'ordre est 255. Il permet d'engendrer tous les éléments de $\mathbb{F}_{256}$. Sa représentation en décimal est 3.

Ce corps représente tous les polynômes de degré inférieur à 8 à coefficients dans $\mathbb{F}_2 = \{0,1\}$ et il y en a exactement 256. Un octet $\{b_7,b_6,b_5,b_4,b_3,b_2,b_1,b_0\}$ peut donc représenter un polynôme de ce corps $\sum_{k=0}^{7} b_k X^k$ avec $b_i \in \{0,1\}$. Sur le tableau 5.11, chaque valeur de g^i est un octet qui représente un polynôme. Par exemple, $3^3 = 15_{10} = 00001111_2 =$

$\{0,0,0,0,1,1,1,1\} = \{0x0F\}$ (la dernière expression étant en hexadécimal) représente le polynôme $X^3 + X^2 + X + 1$.

Une conséquence importante du paragraphe précédent est qu'additionner ou multiplier des polynômes dans $\mathbb{F}_{2^8}$ revient à additionner ou multiplier des octets dans $\mathbb{F}_2$.

L'addition dans $\mathbb{F}_{2^8}$ fonctionne comme suit :

$$(X^4 + X^3 + 1) + (X^3 + X^2 + 1) = X^4 + (1+1)X^3 + X^2 + 1 + 1 = X^4 + X^2$$

C'est à dire qu'on effectue une addition de polynômes normale mais les coefficients obéissent aux règles de calcul de $\mathbb{F}_2$. Dans ce corps, $1+1=0$, l'addition est le ou exclusif. Donc $X^p + X^p = 0, \forall p < n$.

On peut effectuer la même addition sans utiliser les polynômes directement mais en les représentant par des octets : $00011001_2 \oplus 00001101_2 = 00010100_2$ si $\oplus$ représente le ou exclusif.

Pour la multiplication, on procède encore une fois comme pour un polynôme normal en gardant en mémoire que les coefficients suivent les règles de $\mathbb{F}_2$, il n'y a donc pas de retenues. Puis, on calcule le reste de la division euclidienne du polynôme obtenu par m. Contrairement à l'addition, cette opération n'est pas simple voire très laborieuse. Mais on a développé des méthodes efficaces pour l'effectuer.

5.12 Multiplication dans $\mathbb{F}_p$ et logarithmes discrets

5.12.1 Calcul direct

La méthode la plus évidente pour multiplier deux éléments de $\mathbb{F}_p$ est de calculer le produit des polynômes associés aux éléments, puis le reste de la division euclidienne de ce produit modulo le polynôme irréductible choisi pour construire le corps.

L'algorithme 3 décrit comment procéder dans $\mathbb{F}_{2^8}$. Il s'agit de la division euclidienne de deux polynômes mais réalisée via des octets.

Algorithme 3 Multiplication dans $\mathbb{F}^*_{256}$

```
1:  Fonction MULTIPLIER_F256(x, y)        ▷ x et y sont des éléments de F_256.
2:      r ← 0
3:      pour i de 0 à 7 faire
4:          si (y&0x01) ≠ 0 alors          ▷ & est l'opérateur et bit à bit.
5:              r ← r ⊕ x                  ▷ ⊕ est l'opérateur ou exclusif bit à bit.
6:          b ← x&0x80                     ▷ Le masque est donné en hexadécimal.
7:          x ← (x << 1)&0xFF              ▷ << 1 décale les bits vers la gauche de 1.
8:          si b alors
9:              x ← x ⊕ 0x1B
10:         y ← (y >> 1)&0xFF              ▷ >> 1 décale les bits vers la droite de 1.
11:     retourner r
```

5.12.2 Passage par les logarithmes discrets

Afin de multiplier efficacement dans $\mathbb{F}_{2^8}$, c'est à dire en évitant le calcul du reste la division euclidienne, on cherche à utiliser les propriétés des générateurs du corps.

Si l'on choisit le polynôme $m(x) = X^8 + X^4 + X^3 + X + 1$ alors on peut remarquer que l'élément $\{03\}$, c'est à dire $X + 1$ est un générateur du corps (cf. tableau 5.11).

■ **Définition 87 — Logarithme discret.** Soit G un groupe cyclique d'ordre n, g un générateur de G et a un élément de G. Alors il existe un entier naturel i tel que $a = g^i$. On appelle cet entier i logarithme discret de a.

Soient a et b deux éléments de $\mathbb{F}_p^*$. Alors il existe une puissance entière i de g telle que $a = g^i$ et une puissance entière j de 3 telle que $b = g^j$. Le produit ab peut donc s'écrire $ab = g^i \times g^j = g^{i+j}$.

On devine qu'on peut utiliser la mémoïsation pour stocker dans une table les correspondances entre les entiers i et les puissances g^i, car il n'y en a que 255. On peut appeler ces tables des tables de logarithmes discrets, d'après la définition 87. Ainsi, pour calculer ab, il suffit de trouver i et j dans la table des éléments puis regarder dans la table des logarithmes la valeur de g^{i+j}. On n'omettra pas l'opération modulo $p - 1$ dans l'addition des deux logarithmes i et j, le groupe multiplicatif possédant $p - 1$ éléments.

■ **Exemple 66 — Calcul de 37×253 dans $\mathbb{F}_{2^8}$.** On dispose des tables 5.11 et 5.12. On trouve que $37 = 3^{185}$ et que $253 = 3^{247}$. Alors $37 \times 253 = 3^{185+247} = 3^{177} = 182$.

■ **Exemple 67 — Interprétation du tableau 5.11.** On choisit de prendre $3^i = 26$. Sur le tableau 5.11, on voit que cela correspond à $i = 8$.

Il faut interpréter ceci comme suit :

$$
\begin{aligned}
3^8 =& (1 + X)^8 && (5.29)\\
=& X^8 + 8X^7 + 28X^6 + 56X^5 + 70X^4 + 56X^3 + 28X^2 + 8X + 1 && (5.30)\\
=& X^8 + 1 && (5.31)\\
=& 1 \times m(X) - X^4 - X^3 - X && (5.32)\\
=& X^4 + X^3 + X && (5.33)\\
=& 11010_2 && (5.34)\\
=& 26_{10} && (5.35)\\
=& \{0x1A\} && (5.36)
\end{aligned}
$$

5.13 Calculs d'inverses multiplicatifs dans $\mathbb{F}_p^*$

On peut, comme les corps finis sont également des anneaux euclidiens, appliquer l'algorithme d'Euclide étendu pour trouver l'inverse d'un élément a de $\mathbb{F}_p^*$ que l'on note a^{-1}. Cet élément existe car a appartient à un corps. C'est une méthode efficace (cf. exemple 31).

On peut également chercher à utiliser les puissances pour calculer a^{-1}. Si g est un générateur de $\mathbb{F}_p^*$, alors il existe une puissance entière i de g telle que $a = g^i$ et donc telle que $a^{-1} = g^{-i} = g^{p-i}$.

Ainsi, on peut déduire un inverse multiplicatif dans $\mathbb{F}_p^*$ directement à partir d'une table de logarithmes discrets.

■ **Exemple 68 — Calcul de l'inverse de** 42 **dans** $\mathbb{F}_{2^8}^*$**.** On trouve dans la table 5.11 que $42 = 3^{166}$. D'après la table 5.12, l'inverse de 42 vaut $42^{-1} = 3^{255-166} = 3^{89} = 152$. On vérifie qu'on a bien $42 \times 152 = 1$ dans $\mathbb{F}_{2^8}^*$.

5.14 Polynômes irréductibles moniques

La construction d'un corps fini $\mathbb{F}_{p^n}$ dépend de la capacité à trouver un polynôme irréductible de degré n dans $\mathbb{F}_p[X]$. Combien peut-il y en avoir ? Pour répondre à cette question, il est nécessaire d'introduire la fonction de Möbius.

■ **Définition 88 — Fonction de Möbius.** La fonction de Möbius

$$\begin{aligned} \mu : \mathbb{N}^\star &\longrightarrow \{-1, 0, 1\} \\ n &\longmapsto \mu(n) \end{aligned}$$

est définie par :

$$\mu(n) = \begin{cases} 1 & \text{si } n \text{ est le produit d'un nombre pair de nombres premiers distincts,} \\ 0 & \text{si } n \text{ est divisible par un carré différent de 1} \\ -1 & \text{si } n \text{ est le produit d'un nombre impair de nombres premiers distincts.} \end{cases} \tag{5.37}$$

Théorème 43 — Le nombre $N_q(m)$ de polynômes moniques irréductibles de degré m de $\mathbb{F}_q$ est :.

$$N_q(m) = \frac{1}{m} \sum_{d|m} \mu(d) q^{\frac{m}{d}} \tag{5.38}$$

■ **Exemple 69 — Polynômes moniques irréductibles de degré 2 dans** $\mathbb{F}_2$**.**

$$\begin{aligned} N_2(2) =& \frac{1}{2} \sum_{d|2} \mu(d) 2^{\frac{2}{d}} && (5.39) \\ =& \frac{1}{2}(\mu(1)2^2 + \mu(2)2) && (5.40) \\ =& \frac{1}{2}(2^2 - 2) && (5.41) \\ =& \frac{4-2}{2} && (5.42) \\ =& 1 && (5.43) \end{aligned}$$

Ce résultat est à rapprocher des polynômes X^2, X^2+1 et X^2+X+1. Les deux premiers sont factorisables. Les facteurs possibles sont X si 0 est racine et $1+X$ si 1 est racine.

La factorisation de X^2 est $X^2 = X \times X$ et X n'est pas un élément inversible de $\mathbb{F}_2[X]$. Il est donc réductible dans $\mathbb{F}_2[X]$.

De même, $X^2+1 = (X+1)^2 = (X+1) \times (X+1)$ et est donc réductible dans $\mathbb{F}_2[X]$.

Enfin, ni 0 ni1 ne sont racines de $X^2 + X + 1$ qui est donc irréductible dans $\mathbb{F}_2[X]$. Il y a donc bien un seul polynôme irréductible de degré 2 dans $\mathbb{F}_2[X]$.

■ **Exemple 70 — Polynômes moniques irréductibles de degré 8 dans $\mathbb{F}_2$.**

$$N_2(8) = \frac{1}{8}\sum_{d|8}\mu(d)2^{\frac{8}{d}} \tag{5.44}$$
$$= \frac{1}{8}(\mu(1)2^8 + \mu(2)2^4 + \mu(4)2^2 + \mu(8)2) \tag{5.45}$$
$$= \frac{1}{8}(2^8 - 2^4) \tag{5.46}$$
$$= \frac{256 - 16}{8} \tag{5.47}$$
$$= 30 \tag{5.48}$$

Théorème 44 — Produit des moniques irréductibles. Soit p un nombre premier et d un entier naturel non nul. Le produit de tous les polynômes moniques irréductibles de $\mathbb{F}_p[X]$ dont le degré divise d est égal à $X^{p^d} - X$.

Considérons une dernière fois le corps fini $\mathbb{F}_4 = \mathbb{F}_{2^2}$. D'après le théorème 5.38, il existe 3 polynômes irréductibles dans $\mathbb{F}_2[X]$, 2 de degré 1 et 1 de degré 2.

Considérons le polynôme $X^4 - X$. 0 et 1 sont des racines. On peut donc le factoriser par $X(X-1)$, qui sont les polynômes irréductibles de degré 1 (et $1 \mid 2$). On peut également le factoriser par le polynôme irréductible de degré 2 car $2 \mid 2$.

$$X^4 - X = X(X-1)Q(X) = X(X-1)(X^2 + X + 1) \tag{5.49}$$

On a trouvé tous les polynômes irréductibles de $\mathbb{F}_4$. Seul $X^2 + X + 1$ est primitif.

À l'issue de ces développements, nous disposons donc de corps finis avec lesquels on peut faire des calculs linéaires et non linéaires :

- sereinement, c'est à dire sans se poser la question de savoir si l'inverse d'un élément existe, si on a le droit de faire l'opération ou si le résultat est nul.
- rapidement, car les additions, multiplications et inversions possèdent des implémentations électroniques rapides.

Ces propriétés sont fondamentales pour les applications qui sont déclinées dans les parties suivantes de ce cours : CRC, codes correcteurs d'erreur et générateurs de séquences pseudo-aléatoires utilisés dans le domaine du chiffrement. Un outil en particulier est commun à toutes ces approches : les registres à décalage à rétroaction linéaire.

5.15 Représentation des polynômes générateurs

Les polynômes générateurs peuvent être représentés de plusieurs manières en héxadécimal.

■ **Définition 89 — Représentation normale.** Le bit de poids fort, c'est à dire le coefficient de X^n si le polynôme est de degré n, n'est pas représenté. La raison est que celui-ci vaut toujours 1.

■ **Définition 90 — Représentation inverse.** Le bit de poids fort, c'est à dire le coefficient de X^n si le polynôme est de degré n, n'est pas représenté et l'expression est lue à l'envers du à partir du bit de poids faible. Le bit de poids fort de cette représentation

est donc le coefficient de X^0.

■ **Définition 91 — Représentation de Koopman.** Le bit de poids faible n'est pas représenté. Le nombre hexadécimal ne représente donc que les coefficients de X^n à X^1.

■ **Exemple 71 — Représentation de $X^8+X^4+X^3+X^2+1$.** La forme normale est $0x1D$. Sa forme inverse est $0xB8$. Sa forme de Koopman est $0x8E$.

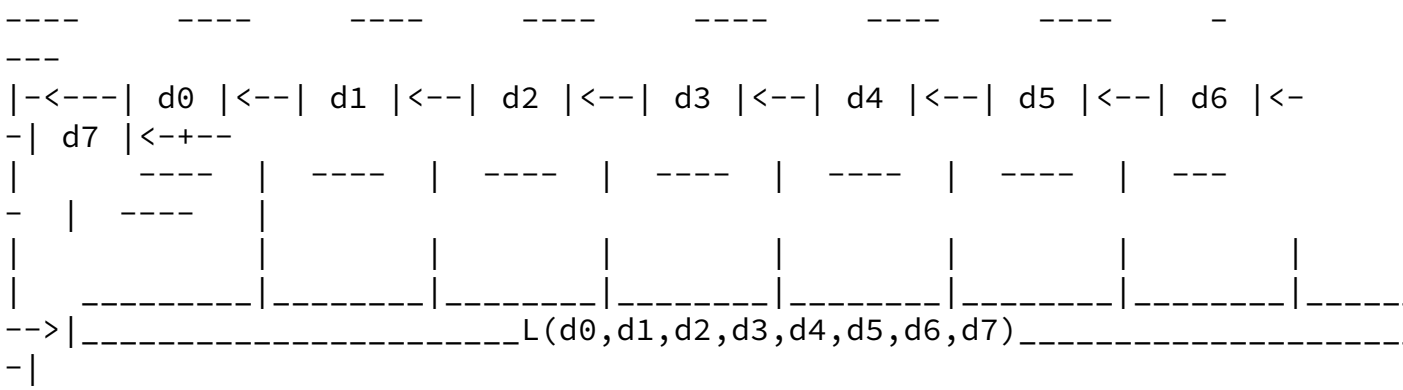
R D'une manière générale, on privilégie la forme normale pour l'étude des CRC et la forme de Koopman pour les générateurs pseudo-aléatoires. Cela n'affecte que notre manipulation des polynômes, pas leur usage dans les calculs.

5.16 Registres à décalage à rétroaction linéaire

Les registres à décalage constituent un outil formidable pour l'électronicien. Ils permettent notamment créer :

- des codes détecteurs d'erreur,
- des générateurs pseudo-aléatoires,
- des compteurs.

Électroniquement, un registre à décalage est composé d'une succession de bascules D. Si on procède à un rebouclage de la sortie sur l'entrée, on dit que le registre est à rétroaction. Cette rétroaction est linéaire, si la fonction L qui calcule l'élément suivant est linéaire.



V **Vocabulary 6 — Linear Feedback Shift Register (LFSR)** ⇜⇝ En anglais, désigne les registres à décalages à rétroaction linéaires par LFSR.

■ **Définition 92 — Rétroaction linéaire.** La rétroaction linéaire est une fonction logique linéaire qui calcule la valeur du bit qui est réinjecté à l'entrée du registre à décalage.

Elle est typiquement donnée sous la forme d'un polynôme mais peut s'exprimer d'une manière purement logique.

5.16.1 Modélisation d'un registre à décalage

Les éléments d'un registre à décalage à rétroaction linéaire peuvent être vus comme des éléments d'un corps premier $\mathbb{F}_p$. Un registre de longueur n constitue alors de manière

équivalente :

- soit l'élément $d_0 d_1 \dots d_{n-1}$ de $\mathbb{F}_{p^n}$,
- soit un n-uplet $(d_0, d_1, \dots, d_{n-1})$ de $(\mathbb{F}_p, \oplus, .)^n$, l'espace vectoriel associé.

■ **Exemple 72 — Registre à décalage dans $\mathbb{F}_8$.** Le corps premier sous-jacent à ce corps fini est $\mathbb{F}_2$, ce qui signifie que les éléments du registre seront 0 ou 1. La longueur du registre est 3. Le registre représente alors de manière équivalente :

- un élément de $\mathbb{F}_8$, par exemple $X + 1$.
- ou un 3-uplet (d_0, d_1, d_2) de $(\mathbb{F}_2, \oplus, .)^3$, par exemple $(0, 1, 0)$.

Si n est la taille du registre, la fonction logique de rebouclage s'écrit d'une manière générale :

$$L(d_0, d_1, \dots, d_{n-1}) = \sum_{i=0}^{n-1} a_i d_i \tag{5.50}$$

où les a_i sont des coefficients à valeur dans $\mathbb{F}_2$ et les d_i les éléments de $\mathbb{F}_2$ contenus dans le registre.

D'un point de vue arithmétique, le rebouclage, c'est à dire recopier la sortie sur l'entrée, se traduit simplement par l'équation suivante :

$$d_n = \sum_{i=0}^{n-1} a_i d_i \tag{5.51}$$

Cette équation signifie que l'élément suivant est fonction linéaire des n premiers éléments du registre à décalage.

En adoptant une notation polynômiale pour les éléments du corps, on peut réécrire cette équation :

$$\sum_{i=0}^{n-1} a_i X^i = X^n \tag{5.52}$$

$$X^n - \sum_{i=0}^{n-1} a_i X^i = 0 \tag{5.53}$$

$$X^n - a_{n-1} X^{n-1} - \dots - a_1 X - a_0 = 0 \tag{5.54}$$

où les X_i sont les valeurs de $\mathbb{F}_p$ prises par le registre et les a_i les coefficients du polynôme associé à la fonction logique. On reconnaît là un polynôme de degré n appartenant au corps fini $\mathbb{F}_{p^n}$.

■ **Définition 93 — Polynôme de rétroaction.** Soit un registre à décalage de taille n et sa fonction logique de rétroaction L. Le polynôme de $\mathbb{F}_{p^n}$ défini par

$$f(X) = X^n - L(X) = X^n - \sum_{i=0}^{n-1} a_i X_i \tag{5.55}$$

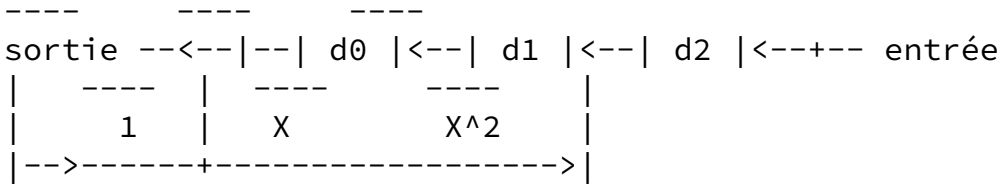
est appelé polynôme de rétroaction.

Posons $L(X) = \sum_{i=0}^{n-1} a_i X_i$. La valeur réinjectée à l'entrée du registre à décalage satisfait l'équation $X^n - L(X) = 0$ et ce, à chaque décalage. Cela signifie que le contenu du

registre $d_0 d_1 \dots d_{n-1}$ qui représente un élément de $\mathbb{F}_{p^n}$ est une racine de f, quelque ce soit le décalage effectué.

On choisit alors astucieusement la fonction L de telle manière que, pour α, un élément générateur de $\mathbb{F}_{p^n}$, f soit minimal. Alors, à chaque décalage, le contenu du registre $d_0 d_1 \dots d_{n-1}$ va successivement prendre pour valeur toutes les puissances de α, c'est à dire toutes les valeurs de $\mathbb{F}_{p^n}$ excepté 0. L'explication réside dans l'équation 5.52 dont α est racine : si le registre vaut $\alpha^i = d_0 d_1 \dots d_{n-1}$, alors au prochain décalage, en appliquant le polynôme minimal, on aura $\sum_{i=0}^{n-1} a_i d_i = d_{i+1}$ et $\alpha^{i+1} = d_1 d_2 \dots d_n$.

■ **Exemple 73 — Registre à décalage dans $\mathbb{F}_8$ et $X^3 + X + 1$.** Le polynôme $f(X) = X^3 + X + 1$ est primitif (et minimal) dans $\mathbb{F}_8$ et une de ses racines, notons la α, est générateur de $\mathbb{F}_8$. Les solutions de $f(X) = 0$ vérifient la relation $X^3 = X + 1$ qui permet de créer un registre à décalage associé.



Selon ce registre, on vérifie toujours $d_{k+1} = d_{k-1} + d_{k-2}$ quelque soit le décalage k. Imaginons qu'on initialise le registre à $(d_0, d_1, d_2) = (0, 0, 1)$, un élément un de $(\mathbb{F}_2, \oplus, .)^3$. Cet élément est nécessairement une puissance de α, puisque α est générateur : on a donc $(d_0, d_1, d_2) = \alpha^i$ pour un certain entier naturel i. Or, à chaque décalage, on calcule une puissance supérieure de α :

$$\alpha^3 = \alpha + 1 \tag{5.56}$$
$$\alpha^4 = \alpha^2 + \alpha \tag{5.57}$$
$$\alpha^5 = \alpha^3 + \alpha^2 = \alpha^2 + \alpha + 1 \tag{5.58}$$
$$\alpha^6 = \alpha^4 + \alpha^3 = \alpha^2 + 1 \tag{5.59}$$
$$\alpha^7 = \alpha^5 + \alpha^4 = 1 \tag{5.60}$$
$$\alpha^8 = \alpha^6 + \alpha^5 = \alpha \tag{5.61}$$
$$\alpha^9 = \alpha^7 + \alpha^6 = \alpha^2 \tag{5.62}$$
$$\alpha^{10} = \alpha + 1 = \alpha^3 \dots \tag{5.63}$$
$$\tag{5.64}$$

Pour décrire les éléments de $\mathbb{F}_8$ on n'a donc besoin que de trois symboles : 1, α et α^2. Il faut mettre ces trois symboles comme des vecteurs directeurs de l'espace vectoriel à trois dimensions associé à $\mathbb{F}_8$. Les d_i ne sont alors que les coordonnées d'un élément dans cet espace, les coefficients devant les vecteurs de la base 1, α et α^2.

On voit qu'à l'initialisation du registre, on avait choisi $(d_0, d_1, d_2) = (0, 0, 1) = \alpha^2$. $(1, 0, 0)$ est 1, $(0, 1, 0)$ est α.

On remarque finalement que la suite composée par le registre est cyclique : elle parcours tous les éléments inversibles du corps fini.

Les schémas des LFSR représentés ci-dessus sont de type Fibonacci. C'est la représentation la plus naturelle d'un LFSR. La valeur suivante du registre n'est qu'un décalage

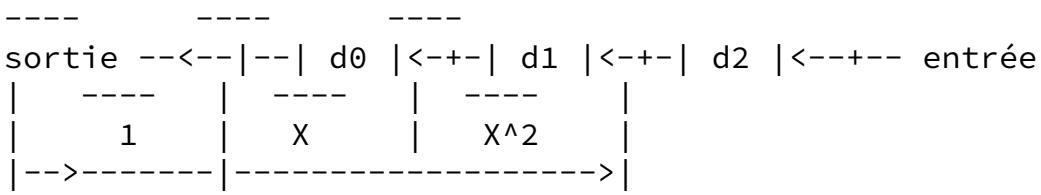
à gauche avec insertion en queue du terme suivant qui résulte d'une fonction linéaire des termes précédents.

Il existe cependant une autre représentation possible équivalente.

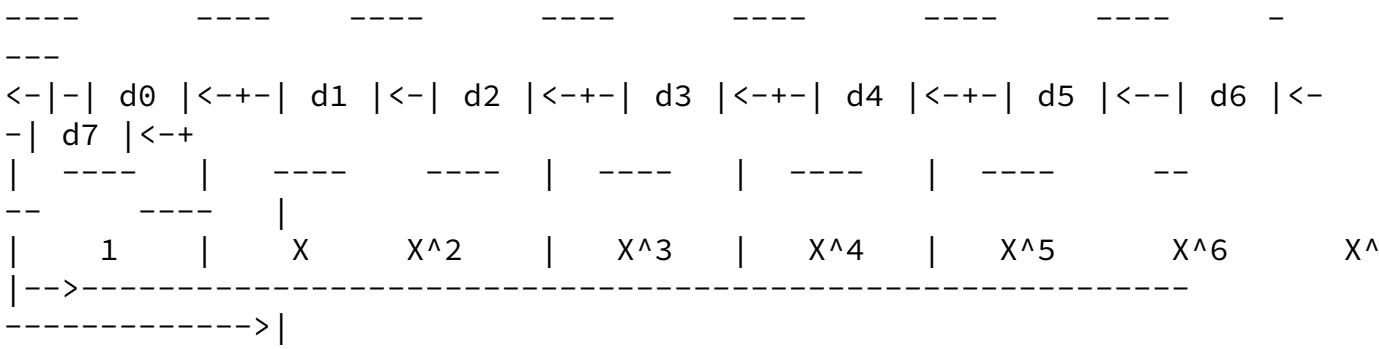
5.16.2 Modélisation sous la forme de Galois

Une autre représentation possible est la suivante : au lieu d'effectuer la rétroaction d'une fonction logique, on insère la fonction logique dans le registre et on reboucle avec la sortie directement. Cette forme donne exactement les mêmes résultats. Elle peut être plus performante, notamment car les opérations logiques peuvent s'effectuer «in place», de manière asynchrone, alors qu'avec Fibonacci il faut d'abord calculer la fonction logique puis décaler.

■ **Exemple 74 — Registre à décalage dans $\mathbb{F}_8$ et X^3+X+1 ou forme normale $0x3$.** On décline l'exemple précédent mais sous la forme d'un LFSR de Galois.



■ **Exemple 75 — LFSR 8 bits avec rétroaction linéaire $0x1D$ en forme normale.** Le polynôme correspondant à ce LFSR est $X^8+X^4+X^3+X^2+1$. On peut le schématiser ainsi sous la forme de Galois :



5.17 Générateurs pseudo-aléatoires

Comme on l'a montré à la section 73, Si le polynôme choisi pour un LFSR est minimal dans $\mathbb{F}_{p^n}$, alors le LFSR balaie toutes les valeurs du corps successivement.

5.17.1 Des différents exigences aléatoires

On peut qualifier de plusieurs manières l'aléatoire produit par un générateur selon l'application visée.

■ **Définition 94 — Aléatoire statistique.** On cherche à ce que la séquence pseudo-aléatoire produite ne puisse pas être distinguée d'une séquence véritablement aléatoire.

On se sert des tests statistiques pour le vérifier.

■ **Définition 95 — Aléatoire cryptographique.** On cherche à ce que la séquence pseudo-aléatoire produite ne puisse pas être distinguée d'une séquence véritablement aléatoire **et** que l'on ne puisse pas prédire la suite de la séquence, même en connaissant l'algorithme de génération.

L'aléatoire cryptographique est donc bien plus exigeant qu'un aléatoire statistique.

5.17.2 Qualité des LFSR en tant que générateur aléatoire

Dans le cas des LFSR, on peut montrer que :

1. La suite de bits générée par un LFSR à séquence maximale (polynôme minimal) possède de bonnes qualités statistiques.
2. La suite des valeurs d'un LFSR (valeurs du registre complet) ne possède pas d'aussi bonnes qualités statistiques. On peut montrer que son spectre ne produit pas une distribution uniforme que l'on attend d'un véritable générateur aléatoire mais privilégie les fréquences basses. Un spectre plat ne peut être obtenu qu'en prenant un seul bit du registre. Si on choisit tous les bits, on note que le 0 est absent. De plus, en y réfléchissant davantage, comme une séquence pseudo-aléatoire de LFSR balaie les éléments d'un corps fini, cette séquence ne se répète pas et elle est donc probablement trop bien ordonnée. Le hasard aussi bégaie parfois...
3. Les LFSR ne sont pas adaptés à la cryptographie en l'état, car pour un LFSR à N bits, il possible, si on dispose de la sortie de 2N bits, de reconstruire le LFSR sans le connaître (algorithme de Berlekamp-Massey).

C'est pourquoi on combine souvent les LFSR avec d'autres fonctions dans le cas de la cryptographie qui permettent notamment d'initialiser le registre d'une manière sophistiquée. Cependant, pour les télécommunication qui n'en ont pas le besoin, par exemple pour étaler un spectre, le multiplexage par code, les LFSR peuvent être adaptés.

5.18 Détection d'erreur via code cyclique

■ **Définition 96 — Contrôle à Redondance Cyclique.** Un CRC est un code par blocs qui consiste à concaténer au message à transmettre un séquence calculée à partir du message et valant le reste de la division euclidienne par un polynôme déterminé. À la réception du message, il suffit de recalculer le CRC à partir du message reçu et de le comparer à celui transmis. Si les résultats diffèrent, alors on a détecté une erreur de transmission.

V **Vocabulary 7 — Cyclic Redundancy Check (CRC)** ⇜⇝

■ **Définition 97 — Distance de Hamming.** La distance de Hamming est une distance mathématique qui permet de quantifier la différence entre deux séquences de symboles.

Soit deux séquences de symboles $a = (a_i)_{i\in[\![0,n-1]\!]}$ et $b = (b_i)_{i\in[\![0,n-1]\!]}$. Alors la distance de Hamming $d(a,b)$ vaut :

$$d(a,b) = \text{card}\,\{i : a_i \neq b_i\} \tag{5.65}$$

■ **Définition 98 — Poids de Hamming.** Le poids de Hamming d'une séquence est sa distance de Hamming par rapport à la séquence constituée du symbole nul.

5.18.1 Principe du CRC et choix du polynôme

On considère que le message à transmettre comporte n bits. Ce message est alors considéré comme les coefficients d'un polynôme de degré n dans $\mathbb{F}_2[X]$ et on le note alors m(X).

On choisit un polynôme $D(X)$ de degré $d \geqslant n$, $D(X) = X^d + a_{d-1}X^{d-1} + \ldots + a_1X + a_0$. On calcule alors le reste du produit $X^d m(X)$

On peut démontrer qu'un tel CRC détecte :

- si $a_0 = 1$, les erreurs sur un bit (c'est à dire de poids de Hamming 1),
- si $f(X)$ possède un facteur qui comporte au moins trois termes, alors le code détecte les erreurs sur deux bits, c'est à dire poids de Hamming 2.
- si $f(X)$ possède $(x + 1)$ comme facteur, alors le code détecte les nombres impairs d'erreurs.
- si $f(X) = p(x).q(x)$ avec p polynôme irréductible de degré d, alors le code détecte toutes les erreurs sur deux bits distantes d'au plus $2^d - 1$ bits consécutifs

C'est pourquoi, il est courant de choisir $f(X) = (X + 1).p(X)$ de degré n dans $\mathbb{F}_{p^n}$ et avec p irréductible. On peut ainsi détecter un nombres impaires d'erreurs et toutes les erreurs sur deux bits jusqu'à une distance de $2^n - 1$.

■ **Exemple 76 — CRC à partir de $f(X) = X^8 + X^2 + X + 1$.** On place dans $\mathbb{F}_{2^8}$ et $f(X) = (X + 1)(X^7 + X^6 + X^5 + X^4 + X^3 + X^2 + 1)$.

On peut vérifier que $X^7 + X^6 + X^5 + X^4 + X^3 + X^2 + 1)$ est irréductible dans $\mathbb{F}_{2^8}$, car il n'est ni facteur de X, ni de $X + 1$, ni de $X^3 + X + 1$ ni de $X^3 + X^2 + 1$ (qui sont les polynômes primitifs de $\mathbb{F}_{2^8}$).

On peut alors détecter toutes les erreurs de poids 1 car $a_0 = 1$ et 2 car p possède plus de trois termes.

R Le résultat d'un CRC tient sur $d - 1$ bits, car le degré du reste de la division euclidienne dans $\mathbb{F}_2[X]$ est plus petit que d, par définition de la division euclidienne.

■ **Exemple 77 — Vérification d'un CRC 8 bits.** Soit le polynôme de degré 8 : $f(X) = X^8 + X^2 + X + 1$. Le degré de f est $d = 8$. Soit un message à transmettre constitué de 8 bits : $m(X) = 11101010$.

Effectuons la division euclidienne le $X^8 m(X)$ par $f(X)$. Seul le reste de la division nous importe, on va donc se concentrer sur cette partie.

Le registre avec lequel on travail est le message lui-même, sa taille est donc de 8 bits. Il faut remarquer que le polynôme est codé sur 9 bits, c'est normal. Le degré du polynôme pourrait être plus grand encore.

```
X^8 m(X)        1110101000000000
f(X)    +       100000111
----------------
11010011
```

```
f(X)    +       100000111
----------------
10100001
f(X)    +        100000111
----------------
01000101
DECALAGE A GAUCHE  10001010
f(X)    +          100000111
----------------
00010011
DECALAGE A GAUCHE      10011000
```

Le résultat de l'opération, et donc le CRC, vaut $0x98$. Afin de vérifier le calcul, il suffit recommencer l'opération en concaténant le CRC au message reçu.

```
X^8 m(X)        1110101010011000
f(X)    +       100000111
----------------
1101001000011000
f(X)    +       100000111
----------------
10100011011000
f(X)    +        100000111
----------------
01000000111000
DECALAGE A GAUCHE  100000111000
f(X)    +          100000111
----------------
00000000000
00000000
```

Le résultat est nul, c'est donc qu'il n'y pas eu d'erreur de transmission.

La performance d'un CRC est un compromis entre :

- le nombre de restes différents possibles,
- le nombre de types d'erreurs détectables.

Dans les deux cas, ces nombres dépendent directement du polynôme choisi. C'est pourquoi les polynômes CRC ont fait l'objet d'études poussées afin de sélectionner les meilleurs pour une application donnée. Le choix d'un polynôme primitif n'est pas forcément le meilleur : on maximise le nombre de restes différents mais pas forcément le nombre de types d'erreurs détectables.

5.18.2 Implémentation via des registres

Les registres à décalages permettent d'implémenter efficacement l'algorithme du CRC : à titre d'exemple, sur une plateforme STM32 F4, le circuit accélérateur de calcul du CRC est 60 fois plus rapide que le logiciel correspondant.

3^i	i	3^i	i	3^i	i	3^i	i	3^i	i	3^i	i	3^i	i	g^i	i
1	255	33	194	65	143	97	110	129	88	161	12	193	178	225	17
2	25	34	29	66	219	98	72	130	168	162	246	194	135	226	146
3	1	35	181	67	189	99	195	131	80	163	111	195	144	227	217
4	50	36	249	68	54	100	163	132	244	164	23	196	97	228	35
5	2	37	185	69	208	101	182	133	234	165	196	197	190	229	32
6	26	38	39	70	206	102	30	134	214	166	73	198	220	230	46
7	198	39	106	71	148	103	66	135	116	167	236	199	252	231	137
8	75	40	77	72	19	104	58	136	79	168	216	200	188	232	180
9	199	41	228	73	92	105	107	137	174	169	67	201	149	233	124
10	27	42	166	74	210	106	40	138	233	170	31	202	207	234	184
11	104	43	114	75	241	107	84	139	213	171	45	203	205	235	38
12	51	44	154	76	64	108	250	140	231	172	164	204	55	236	119
13	238	45	201	77	70	109	133	141	230	173	118	205	63	237	153
14	223	46	9	78	131	110	61	142	173	174	123	206	91	238	227
15	3	47	120	79	56	111	186	143	232	175	183	207	209	239	165
16	100	48	101	80	102	112	43	144	44	176	204	208	83	240	103
17	4	49	47	81	221	113	121	145	215	177	187	209	57	241	74
18	224	50	138	82	253	114	10	146	117	178	62	210	132	242	237
19	14	51	5	83	48	115	21	147	122	179	90	211	60	243	222
20	52	52	33	84	191	116	155	148	235	180	251	212	65	244	197
21	141	53	15	85	6	117	159	149	22	181	96	213	162	245	49
22	129	54	225	86	139	118	94	150	11	182	177	214	109	246	254
23	239	55	36	87	98	119	202	151	245	183	134	215	71	247	24
24	76	56	18	88	179	120	78	152	89	184	59	216	20	248	13
25	113	57	240	89	37	121	212	153	203	185	82	217	42	249	99
26	8	58	130	90	226	122	172	154	95	186	161	218	158	250	140
27	200	59	69	91	152	123	229	155	176	187	108	219	93	251	128
28	248	60	53	92	34	124	243	156	156	188	170	220	86	252	192
29	105	61	147	93	136	125	115	157	169	189	85	221	242	253	247
30	28	62	218	94	145	126	167	158	81	190	41	222	211	254	112
31	193	63	142	95	16	127	87	159	160	191	157	223	171	255	7
32	125	64	150	96	126	128	175	160	127	192	151	224	68		

TABLE 5.11 – Table des éléments de $\mathbb{F}^*_{256}$ avec le polynôme $m(x) = X^8 + X^4 + X^3 + X + 1$ et d'après les puissances i d'un élément générateur $g = 3$, c'est à dire $X + 1$. Les valeurs décimales de 3^i peuvent être transformées en octet et représentent ainsi les polynômes associés à chaque élément de $\mathbb{F}^*_{256}$. On note qu'on a bien $3^{255} = 1$ et que tous les éléments de $\mathbb{F}^*_{256}$ sont listés.

i	3^i	i	3^i	i	3^i	i	3^i	i	3^i	i	3^i	i	3^i	i	g^i
1	3	33	52	65	212	97	196	129	22	161	186	193	31	225	54
2	5	34	92	66	103	98	87	130	58	162	213	194	33	226	90
3	15	35	228	67	169	99	249	131	78	163	100	195	99	227	238
4	17	36	55	68	224	100	16	132	210	164	172	196	165	228	41
5	51	37	89	69	59	101	48	133	109	165	239	197	244	229	123
6	85	38	235	70	77	102	80	134	183	166	42	198	7	230	141
7	255	39	38	71	215	103	240	135	194	167	126	199	9	231	140
8	26	40	106	72	98	104	11	136	93	168	130	200	27	232	143
9	46	41	190	73	166	105	29	137	231	169	157	201	45	233	138
10	114	42	217	74	241	106	39	138	50	170	188	202	119	234	133
11	150	43	112	75	8	107	105	139	86	171	223	203	153	235	148
12	161	44	144	76	24	108	187	140	250	172	122	204	176	236	167
13	248	45	171	77	40	109	214	141	21	173	142	205	203	237	242
14	19	46	230	78	120	110	97	142	63	174	137	206	70	238	13
15	53	47	49	79	136	111	163	143	65	175	128	207	202	239	23
16	95	48	83	80	131	112	254	144	195	176	155	208	69	240	57
17	225	49	245	81	158	113	25	145	94	177	182	209	207	241	75
18	56	50	4	82	185	114	43	146	226	178	193	210	74	242	221
19	72	51	12	83	208	115	125	147	61	179	88	211	222	243	124
20	216	52	20	84	107	116	135	148	71	180	232	212	121	244	132
21	115	53	60	85	189	117	146	149	201	181	35	213	139	245	151
22	149	54	68	86	220	118	173	150	64	182	101	214	134	246	162
23	164	55	204	87	127	119	236	151	192	183	175	215	145	247	253
24	247	56	79	88	129	120	47	152	91	184	234	216	168	248	28
25	2	57	209	89	152	121	113	153	237	185	37	217	227	249	36
26	6	58	104	90	179	122	147	154	44	186	111	218	62	250	108
27	10	59	184	91	206	123	174	155	116	187	177	219	66	251	180
28	30	60	211	92	73	124	233	156	156	188	200	220	198	252	199
29	34	61	110	93	219	125	32	157	191	189	67	221	81	253	82
30	102	62	178	94	118	126	96	158	218	190	197	222	243	254	246
31	170	63	205	95	154	127	160	159	117	191	84	223	14	255	1
32	229	64	76	96	181	128	251	160	159	192	252	224	18		

TABLE 5.12 – Table des logarithmes discrets i de l'élément générateur 3 de $\mathbb{F}_{256}^*$ avec le polynôme $m(x) = X^8 + X^4 + X^3 + X + 1$.

m	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\mu(m)$	1	-1	-1	0	-1	1	-1	0	0	1	-1	0	-1	1	1

TABLE 5.13 – Quelques valeurs de la fonction de Möbius.

Troisième partie

test

Chapitre 6

rfee

Objectifs d'apprentissage :

- ☐ Savoir utiliser les opérateurs.
- ☐ Connaître les opérations de l'algèbre relationnelle.

■ **Définition 99 — Loi de composition interne.** On appelle loi de composition interne toute application :

$$\begin{aligned} \star : E \times E &\longrightarrow E \\ (x, y) &\longmapsto \star(x, y) = x \star y \end{aligned}$$

R En utilisant un corps premier, la construction d'un anneau quotient avec un élément irréductible permet de construire de nouveaux corps finis de type $\mathbb{F}_p[X]/m$, différents des corps premiers.

Proposition 22 — L'anneau $(\mathbb{F}_p[X]/m, +, \times)$ est un corps si et seulement si m est irréductible.

V **Vocabulary 8 — Corps de rupture (Splitting Fields)** ⇜⇝ En anglais, les corps $(\mathbb{F}_p[X]/m, +, \times)$ ainsi généré se nomme Splitting Fields, corps de rupture en français.

Théorème 45 — Le cardinal du corps $\mathbb{F}_p[X]/m$ avec $m \in \mathbb{F}_p[X]$ irréductible et $\deg(m) = n$ est p^n .

Démonstration. D'après la relation de congruence et les propriétés de la division euclidienne, $\mathbb{F}_p[X]/m$ est l'ensemble des polynômes à coefficients dans $\mathbb{F}_p$ dont le degré est strictement inférieur à n. Ce sont des polynômes à n coefficients a_i à valeur dans $\mathbb{F}_p$. Pour chaque a_i, on a donc p valeurs possibles. D'où le résultat. ■

■ **Exemple 78 — Application au déchiffrement du chiffre affine.** Le chiffrement affine (cf. par exemple le chiffre de César **??**) simple correspond à

$$\mathcal{E} : \lambda \to (a\lambda + b) \mod s$$

avec $(a, b) \in \mathbb{Z}$ connus et $s = 26$. Déchiffrer ce code revient à chercher λ connaissant e tel que :

$$\mathcal{D} : (a\lambda + b) \mod s = e$$

Ce système est équivalent à :

$$\mathcal{D} : a\lambda - ks = e - b$$

qui est une équation diophantienne à deux inconnues (k et λ) et de degré 1. Il suffit donc de résoudre cette équation diophantienne à l'aide de la méthode 1.

$$\mathcal{D} : a'\lambda - ks' = e'$$

avec $\text{PGCD}(a, s) = d, e' = (e-b)/d, a' = a/d, s' = s/d$. On vérifie que $\text{PGCD}(a, s)|e-b$ et on utilisera l'algorithme d'Euclide étendu.

M **Méthode 9 — Comment résoudre un système d'équations PGCD et PPCM** Soit le système d'équations suivant à résoudre dans $\mathbb{Z}^2$:

$$\mathcal{S} : \begin{cases} PGCD(x, y) &= d \\ PPCM(x, y) &= m \end{cases} \tag{6.1}$$

- Si $d \nmid m$, alors $\mathcal{S}$ n'a pas de solution dans $\mathbb{Z}^2$.
- Sinon, effectuer le changement de variable $x = dx'$ et $y = dy'$.
- On a $\text{PGCD}(x', y') = 1$, donc $\mathcal{S}$ est équivalent à $x'y' = m'$.
- On achève la résolution en trouvant les couples premiers entre eux qui vérifient l'équation factorisée.