Université Paris Saclay (Orsay)

Master 1 M.F.

Année 2021-2022

# M.A.O. Calcul formel

Notes de cours de Joël Riou

Reprises par Stéphane Fischler (cours)

et Kevin Destagnol (T.P.)

# TABLE DES MATIÈRES

# BIBLIOGRAPHIE


[1] Alin Bostan, Frédéric Chyzak, Marc Giusti, Romain Lebreton, Grégoire Lecerf, Bruno Salvy et Eric Schost. Algorithmes Efficaces en Calcul Formel. 686 pages. Imprimé par CreateSpace. Palaiseau : Frédéric Chyzak (auto-édit.), sept. 2017. Disponible en version électronique : https ://hal.archives-ouvertes.fr/AECF/ *Les chapitres de ce livre correspondant aux sujets traités ce semestre constituent une excellente référence ; mais ce livre contient aussi beaucoup d'autres choses qui dépassent largement le niveau de ce cours. Je vous conseille donc de n'imprimer que les chapitres dont vous avez besoin.*

[2] Thomas Cormen, Charles Leiserson, Ronald Rivest et Clifford Stein, Introduction à l'algorithmique, Dunod.

[3] Michel Coste, Annette Paugam et Ronan Quarez, Codes correcteurs, disponible sur le web à l'adresse `http://agreg-maths.univ-rennes1.fr/documentation/docs/codes.pdf`.

[4] Michel Demazure, Cours d'algèbre, Cassini, 1997.

[5] Joachim von zur Gather et Jürgen Gerhard, Modern Computer Algebra, Cambridge.

[6] Maurice Mignotte, Mathématiques pour le calcul formel, PUF.

[7] Bernadette Perrin-Riou, Algèbre, arithmétique et Maple, Cassini.

[8] Oliver Pretzel, Error-correcting codes and finite fields, Clarendon Press, 1992.

[9] Scott Vanstone et Paul van Oorschot, An introduction to error correcting codes with applications, Kluwer.

**PROGRAMME DE L'AGREGATION EXTERNE**

**OPTION CALCUL FORMEL**

*Session 2022*

(a) Représentation et manipulation des entiers, des flottants, des polynômes, des éléments de $\mathbf{Z}/n\mathbf{Z}$. Calcul effectif : addition, multiplication, division. Coût de ces opérations(*).

(b) Algorithmes algébriques élémentaires.

Exponentiation rapide, algorithme d'EUCLIDE étendu. Coût de ces algorithmes(*). Test de primalité de FERMAT.

Application au chiffrement RSA.

(c) Corps finis

Représentation des éléments d'un corps fini.

Calcul effectif : addition, multiplication, inversion. Coût de ces opérations(*).

(d) Matrices à coefficients dans un corps.

Résolution de systèmes linéaires, méthode du pivot de GAUSS. Calcul effectif du rang, du déterminant. Coût de ces calculs(*).

(e) Codes correcteurs linéaires.

Distance de HAMMING, distance minimale d'un code linéaire.

Codes de répétition, codes de HAMMING binaires.

(f) Polynômes à une indéterminée.

Évaluation (schéma de HORNER), interpolation. Coût de ces opérations(*).

Localisation des racines dans $\mathbf{R}$ ou $\mathbf{C}$ : majoration en fonction des coefficients.

(g) Polynômes à plusieurs indéterminées.

Tout polynôme symétrique s'exprime en fonction des polynômes symétriques élémentaires.

Résultant, élimination. Applications : résolution de systèmes polynomiaux, calcul de l'intersection de courbes algébriques planes, passage d'une paramétrisation à une équation implicite. Calcul effectif du résultant, coût de ce calcul(*).

(*) Concernant les différentes mentions du coût d'un algorithme, aucune formalisation d'un modèle de calcul n'est exigée.

# CHAPITRE 1

# ORDRES DE GRANDEUR, COÛT

Buts :

- comparer le comportement asymptotique de suites ;
- fournir un outil de mesure du coût des algorithmes.

Les lettres $u$, $v$, $w$, etc. désignent des suites réelles, autrement dit des applications $\mathbf{N} \to \mathbf{R}$.

## 1.1. Définitions

***Définition 1.1***. —

- si $l \in \mathbf{R}$, on dit que $u$ tend vers $l$ (noté $u \to l$, ou $\lim_{n\to+\infty} u_n = l$) si pour tout $\varepsilon > 0$, il existe un entier $N_\varepsilon$ tel que pour tout $n \geq N_\varepsilon$, $|u_n - l| \leq \varepsilon$ ;
- $u$ est un grand O de $v$ (noté $u = O(v)$), sous-entendu quand la variable $n$ tend vers $+\infty$, s'il existe $k > 0$ tel qu'à partir d'un certain rang, on ait $|u_n| \leq k|v_n|$ ;
- $u$ est un petit o de $v$ (noté $u = o(v)$), si pour tout $\varepsilon > 0$, il existe un entier $N_\varepsilon$ tel que pour tout $n \geq N_\varepsilon$, on ait $|u_n| \leq \varepsilon|v_n|$.
- $u$ est équivalent à $v$ (noté $u \sim v$), si $u - v = o(v)$.

Si $v$ ne s'annule pas (à partir d'un certain rang...), on a les critères suivants :

- $u = O(v)$ si et seulement si la suite $u/v$ est bornée [1] ;
- $u = o(v)$ si et seulement si la suite $u/v$ tend vers 0
- $u \sim v$ si et seulement si la suite $u/v$ tend vers 1.

## 1.2. Quelques règles de calcul

- si $u = o(v)$ ou $u \sim v$, alors $u = O(v)$.
- $u \sim u$ ;
- $u \sim v$ équivaut à $v \sim u$.
- si $u \sim v$ et $v \sim w$, alors $u \sim w$.
- si $u = O(w)$ et $v = O(w)$, alors $u + v = O(w)$.
- si $u = O(v)$ et $v = O(w)$, alors $u = O(w)$.
- si $u = O(v)$ et $u' = O(v')$, alors $uu' = O(vv')$.
- si $u \sim v$, alors $w = O(u)$ équivaut à $w = O(v)$ ;
- si $u \sim v$, alors $w = o(u)$ équivaut à $w = o(v)$.

1. Autrement dit, on peut supprimer « à partir d'un certain rang » dans la définition. Cependant, considérer l'assertion $n^2 = O(n(n-1))$.

## 1.3. Exemples

### 1.3.1. Polynômes. —

***Proposition 1.2.*** — *Soit $P \in \mathbf{R}[X] = a_d X^d + \cdots + a_0$ un polynôme de degré $d$ (ainsi $a_d \neq 0$). Alors $P(n) \sim a_d n^d$. En particulier, $P(n) = O(n^d)$.*

***Proposition 1.3.*** — *Soit $P \in \mathbf{R}[X]$. Il existe un unique polynôme $Q \in \mathbf{R}[X]$ tel que $Q(-1) = 0$ et que $Q(X+1) - Q(X) = P(X+1)$. Ainsi, pour tout $n \in \mathbf{N}$, on a $Q(n) = \sum_{k=0}^{n} P(k)$. Si $P$ est non nul, on a $\deg Q = \deg P + 1$.*

***Corollaire 1.4.*** — *Pour tout entier $k$, $\sum_{i=0}^{n} i^k = O(n^{k+1})$.*

***Corollaire 1.5.*** — *Soit $k \in \mathbf{N}$ [2]. On suppose que $u_n = O(n^k)$. Alors,*

$$\sum_{i=0}^{n} u_i = O(n^{k+1}) \ .$$

**1.3.2. Suites récurrentes.** — Soit $u$ une suite à termes strictement positifs.

- s'il existe une constante $\lambda > 0$ telle que $u_{n+1} \leq \lambda u_n$ (autrement dit $u_{n+1} = O(u_n)$), alors $u_n = O(\lambda^n)$.
- s'il existe $k \in \mathbf{N}$ tel que $u_{n+1} \leq u_n + O(n^k)$, alors $u_n = O(n^{k+1})$.
- si $u_n \leq u_{\lceil \frac{n}{2} \rceil} + O(1)$, alors $u_n = O(\log_2 n)$.
- si $u_n \leq 2u_{\lceil \frac{n}{2} \rceil} + O(n)$, alors $u_n = O(n \log_2 n)$.

*Démonstration.* — La première assertion est claire ; la deuxième découle par télescopage du corollaire 1.5.

*Cas où $u_n \leq u_{\lceil \frac{n}{2} \rceil} + O(1)$.*

Il existe $\lambda > 0$ tel que pour tout $n$, $u_n \leq u_{\lceil \frac{n}{2} \rceil} + \lambda$. Par récurrence sur $d \in \mathbf{N}$, on montre que pour tout $2^d + 1 \leq n \leq 2^{d+1}$, on a $u_n \leq u_1 + \lambda(d+1)$. Ainsi, pour tout $n \geq 2$, on a $u_n \leq u_1 + \lambda(1 + \log_2 n)$, ce qui permet de conclure que $u_n = O(\log_2 n)$.

*Cas où $u_n \leq 2u_{\lceil \frac{n}{2} \rceil} + O(n)$.*

Il existe $\lambda > 0$ tel que pour tout $n$, $u_n \leq 2u_{\lceil \frac{n}{2} \rceil} + \lambda n$. On choisit un nombre réel $M$ supérieur à $u_1$ et à $\lambda$. Par récurrence sur $d \in \mathbf{N}$, on montre pour tout $2^d + 1 \leq n \leq 2^{d+1}$, on a $u_n \leq M a_d$ où la suite $a_d$ est définie par $a_0 = 4$ et la relation de récurrence $a_d = 2a_{d-1} + 2^{d+1}$. On montre que $a_d = (2+d) \cdot 2^{d+1}$ [3]. Par suite, comme $d \leq \log_2 n$, on obtient que $u_n \leq 2M \cdot (2 + \log_2 n) \cdot n$ pour tout $n \geq 1$, d'où $u_n = O(n \log_2 n)$.

□

## 1.4. Coûts de quelques algorithmes

**1.4.1. Exponentiation rapide.** — Dans un anneau $A$, on veut calculer $a^n$, où $a \in A$ et $n \in \mathbf{N}$.

L'algorihme trivial consiste à faire une boucle :

```
def expo_lente(a,n):
    x = 1
    for i in range(n):
        x*=a
    return x
```

2. Le résultat vaut aussi si on remplace l'exposant entier $k$ par un nombre réel $\alpha > -1$.

3. Pour retrouver ce résultat, considérer la suite $b_d = \frac{a_d}{2^d}$ et montrer qu'elle vérifie la relation de récurrence $b_d = b_{d-1} + 2$.

Si on considère la multiplication comme une opération élémentaire, le coût (en temps) de cet algorithme est $O(n)$. Le coût en espace est $O(1)$.

L'algorithme d'exponentiation rapide s'appuie sur l'égalité $a^{2n} = a^n \cdot a^n$. Pour calculer $a^n$ avec $n \geq 1$, on distingue deux cas :

- si $n$ est pair, on détermine l'entier $m$ tel que $n = 2m$, on calcule $b = a^m$ et on renvoie $b^2$ ;
- si $n$ est impair, on détermine l'entier $m$ tel que $n = 2m + 1$, on calcule $b = a^m$ et on renvoie $ab^2$.

Ainsi décrit, il est naturel d'implémenter cet algorithme par une fonction récursive :

```
def exponentiation_rapide_rec(a,n):
    if n == 0:
        return 1
    if n == 1:
        return a
    y = expo(a,n//2);
    if n%2 == 0:
        return y*y
    else:
        return a*y*y
```

Le nombre d'appels récursifs à la fonction `exponentiation_rapide_rec` est en $O(\log_2 n)$. Le coût en temps est donc aussi en $O(\log_2 n)$.

Les appels récursifs à cette fonction s'organisent en un arbre de profondeur $O(\log_2 n)$. Le coût en mémoire n'est pas constant. Il peut être intéressant de récrire la fonction sous forme d'une boucle de façon à conserver un coût en $O(\log_2 n)$ tout en ayant un coût en espace en $O(1)$ :

```
def expo_rapide_iteratif(a,n):
    r=1;
    N=n;
    A=a;
    while(N>0):
        if N%2 == 1:
            r=r*A
        A=A*A;
        N=N//2;
    return r
```

Le calcul termine puisqu'à chaque passage dans la boucle, $N$ décroît strictement ; plus précisément, le nombre de passages dans la boucle est en $O(\log_2 n)$. Pour montrer que cette fonction fait bien ce qu'elle doit, considérer l'invariant de boucle « $a^n = r \cdot A^N$ ». Le coût en espace est constant et le coût en temps est en $O(\log_2 n)$.

***Remarque 1.6***. — On a supposé que le coût d'une multiplication était constant. Si on travaille dans un anneau $\mathbf{Z}/2^n\mathbf{Z}$, c'est le cas et les opérations sur les entiers réalisées par les processeurs rentrent essentiellement dans ce cadre.

***Remarque 1.7***. — L'exponentiation rapide est un des ingrédients essentiels du chiffrement RSA.

**1.4.2. Une application aux groupes finis.** — Dans ce paragraphe on donne, comme application de l'exponentiation rapide, un algorithme efficace pour déterminer si un élément d'un groupe fini l'engendre.

Soit $G$ un groupe fini de cardinal $n$, noté multiplicativement. On suppose connue la factorisation de $n$ en produit de nombres premiers. On dispose alors de l'algorithme suivant pour tester si un élément $g \in G$ est un générateur de $G$ (c'est-à-dire si il est d'ordre $n$).

**Algorithme.** Pour tout nombre premier $p$ qui divise $n$, calculer $g^{n/p}$. Si il existe un tel $p$ pour lequel $g^{n/p} = 1$ (en notant 1 l'élément neutre de $G$), alors $g$ n'est pas générateur. Sinon $g$ est générateur.

**Correction.** Si on trouve $p$ tel que $g^{n/p} = 1$ alors l'ordre de $g$ dans $G$ divise $n/p$, donc ce n'est pas $n$. Réciproquement, si l'ordre de $g$ (noté $d$) est un diviseur strict de $n$, alors il existe $p$ premier qui divise $n/d$. Alors $d$ divise $n/p$ donc $g^{n/p} = 1$.

**Complexité.** Pour chaque nombre premier $p$, le calcul de $g^{n/p}$ se fait en $O(\log(n/p)) = O(\log n)$ opérations dans $G$, par exponentiation rapide. Notons $t$ le nombre de facteurs premiers (deux à deux distincts) de $n$ ; alors le coût de cet algorithme est en $O(t \log n)$ opérations dans $G$. En pratique $t$ est connu puisqu'on suppose connue la factorisation de $n$. On peut toutefois chercher des bornes a priori, dont la dépendance en $n$ soit plus simple à visualiser. Pour cela notons $n = p_1^{e_1} \dots p_t^{e_t}$ avec $p_1, \dots, p_t$ premiers deux à deux distincts et $e_1, \dots, e_t \geq 1$. Alors on a $n \geq p_1 \dots p_t \geq 2^t$ donc $t \leq \frac{\log n}{\log 2}$ d'où un coût en $O((\log n)^2)$ opérations dans $G$. On peut en fait montrer, en utilisant le théorème des nombres premiers, que $t \leq \frac{\log n}{\log \log n}(1 + o(1))$ quand $n \to +\infty$ (et cette majoration est optimale), ce qui raffine légèrement l'estimation du coût : $O\Big(\frac{(\log n)^2}{\log \log n}\Big)$ opérations dans $G$.

**1.4.3. Multiplication rapide.** — Soit $n$ un entier naturel. Un entier naturel à $n$ bits est un entier compris entre 0 et $2^n - 1$ : on représente un entier par le $n$-uplet $(b_{n-1}, \dots, b_0)$ d'entiers valant 0 ou 1 tel que cet entier soit $\sum_{i=0}^{n-1} b_i 2^i$. On cherche à calculer le produit de deux entiers à $n$ bits (le résultat a au plus $2n$ bits).

***Théorème 1.8* (Karatsuba).** — *On peut multiplier deux entiers à $n$ bits en $O(n^{\log_2 3})$* [4] *opérations élémentaires sur les bits.*

Soit $m$ un entier. Des entiers à $2m$ bits $x$ et $y$ peuvent être représentés sous la forme $x = a2^m + b$ et $y = c2^m + d$ où $a, b, c, d$ sont des entiers naturels à au plus $m$ bits. On a

$$xy = ac \cdot 2^{2m} + (ad + bc) \cdot 2^m + bd \, .$$

Un algorithme possible pour faire la multiplication consiste à calculer les quatre produits $ac$, $ad$, $bc$ et $bd$ d'entiers à $m$ chiffres, puis faire les opérations faciles (décalage et additions) pour obtenir le résultat $xy$.

Notons $c_k$ le coût en opérations par bits de la multiplication de deux entiers à $2^k$ bits. En appliquant récursivement la méthode ci-dessus, on obtient une relation de récurrence $c_{k+1} = 4c_k + O(2^k)$. Le quotient $d_k = \frac{c_k}{4^k}$ vérifie une relation $d_{k+1} = d_k + O(\frac{1}{2^k})$ dont on déduit que la suite $d_k$ converge, ainsi $c_k = O(4^k)$. Finalement, on peut en conclure que l'on peut faire le produit d'entiers à $n$ bits en $O(n^2)$ opérations.

L'algorithme de Karatsuba s'appuie sur le fait que l'on peut déterminer $ac$, $ad + bc$ et $bd$ en faisant quelques additions et surtout trois multiplications au lieu de quatre. En effet, on peut calculer $N = (a + b) \cdot (c + d)$, $ac$ et $bd$, et en déduire $ad + bc$ puisqu'on a $ad + bc = N - ac - bd$. Notons $c'_k$ le coût de la multiplication de deux entiers à $2^k$ bits par cette méthode. On a cette fois-ci $c_{k+1} = 3c_k + O(2^k)$, dont on déduit $c_k = O(3^k)$. Si $n$ est un entier naturel non nul, on peut considérer la plus petite puissance de 2 $2^k$ qui lui est supérieure, on a alors $k \leq \log_2 n + 1$. Comme $3^{\log_2 n} = n^{\log_2 3}$, on peut multiplier deux entiers à $n$ bits en $O(n^{\log_2 3})$ (et donc $O(n^{1.585})$).

Le coût en espace est en $O(n)$. En effet, notons $M_k$ le coût en espace de la multiplication des entiers à $2^k$ bits. On obtient

$$M_{k+1} = O(2^k) + M_k \, .$$

En effet, pour faire le calcul pour des entiers à $2^{k+1}$ bits, on a besoin de stocker quelques entiers à au plus $2^{k+2}$ bits pour les résultats des calculs intermédiaires, ceci donne une contribution en $O(2^k)$. Puis, on doit faire trois appels récursifs pour faire des multiplications sur des entiers à $2^k$ chiffres. Le calcul de la première multiplication nécessite un surcoût de $M_k$, la deuxième et la troisième multiplication aussi, mais on peut libérer la mémoire utilisée par la première avant de lancer la deuxième, etc. Le coût en mémoire des trois multiplications de chiffres à $2^k$ bits est donc $M_k$, et pas $3M_k$. Ainsi, $M_k$ est un grand O de la somme partielle de la série des $2^k$, d'où $M_k = O(2^k)$. On en déduit que l'algorithme de Karatsuba permet de faire la multiplication des entiers à $n$ bits avec un coût en espace de $O(n)$.

4. $\log_2 3 \simeq 1.58496\dots$

***Remarque 1.9***. — Il est en fait possible de multiplier deux entiers à $n$ bits en $O(n \ln n \ln\ln n)$ opérations élémentaires grâce à l'algorithme de Schönhage-Strassen (1971), fondé sur la transformation de Fourier rapide (voir chapitre 9). Un résultat de Harvey et van der Hoeven paru en 2021 permet même de passer à $O(n \ln n)$ opérations.

***Théorème 1.10*** **(Karatsuba)**. — *En suivant la même méthode, on peut multiplier deux polynômes de degré $\leq n$ à coefficients dans un anneau commutatif $A$ en $O(n^{\log_2 3})$ opérations arithmétiques dans $A$.*

Ici encore, l'algorithme naïf utilise $O(n^2)$ opérations arithmétiques, et la transformation de Fourier rapide permet de passer à $O(n \ln n \ln\ln n)$ (et même à $O(n \ln n)$ si $A$ est un corps fini). De façon générale, la multiplication des polynômes ressemble beaucoup à celle des entiers, en plus facile car il n'y a pas besoin de gérer les retenues.

**1.4.4. Suite de Fibonacci.** — La suite de Fibonacci $F_n$ est définie par les égalités $F_0 = 0$, $F_1 = 1$ et la relation de récurrence $F_{n+2} = F_n + F_{n+1}$.

***Proposition 1.11***. — *Pour tout entier naturel $n$, on a*

$$F_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}} \ .$$

En particulier, $F_n \sim \frac{1}{\sqrt{5}}\phi^n$ où $\phi = \frac{1+\sqrt{5}}{2}$ est le nombre d'or.

Plusieurs méthodes de calcul des entiers $F_n$ sont possibles. En voici une stupide :

```
def fibonacci(n):
    if n<=1:
        return n
    return fibonacci(n-2)+fibonacci(n-1)
```

Si on considère l'addition entière comme une opération élémentaire, le coût en temps est en $O(\phi^n)$, ce qui est extrêmement mauvais.

Une autre méthode consiste à faire une boucle pour calculer le couple $(F_k, F_{k+1})$ pour $k$ de 0 à $n-1$ :

```
def fibonacci2(n):
    a = 0;b=1
    for i in range(n):
        c=a
        a=b
        b=a+c
    return a
```

Le coût en temps de `fibonacci2` est en $O(n)$.

On peut faire mieux. Rappelons que si on a une suite définie par une équation de récurrence linéaire de la forme $u_{n+2} = au_n + bu_{n+1}$, on peut introduire la matrice

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

La relation de récurrence se récrit :

$$\begin{pmatrix} u_{n+1} \\ u_{n+2} \end{pmatrix} = M \cdot \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix}$$

Par conséquent,

$$\begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} = M^n \cdot \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \ ,$$

ainsi, le calcul de $u_n$ n'est pas plus compliqué que celui de $M^n$. L'algorithme d'exponentiation rapide s'applique à l'exponentiation de matrices. On obtient ainsi un algorithme ayant un coût en $O(\log_2 n)$. Une implémentation possible en Sage est la suivante :

```
def fibonacci3(n):
    m = matrix(2,2,[0,1,1,1])**n
    return m[0,1]
```

# CHAPITRE 2

# ARITHMÉTIQUE, ALGORITHME D'EUCLIDE

Dans tout ce cours, les anneaux sont supposés commutatifs et unitaires. On note $A^\times$ le groupe multiplicatif formé par les éléments inversibles d'un anneau $A$.

## 2.1. Rappels d'arithmétique

### 2.1.1. Anneaux, idéaux, pgcd, ppcm. —

***Définition 2.1***. — Soit $A$ un anneau. Un idéal $I$ de $A$ est un sous-groupe de $(A, +)$ tel que pour tous $a \in A$ et $x \in I$, $ax \in I$.

***Remarque 2.2***. — Si $I$ est un idéal de $A$, le groupe abélien $A/I$ est naturellement muni d'une structure d'anneau telle que le morphisme de projection $A \to A/I$ soit un morphisme d'anneaux. Tout idéal de $A/I$ s'écrit de manière unique sous la forme $J/I$ où $J$ est un idéal de $A$ contenant $I$.

***Définition 2.3***. — Si $a_1, \ldots, a_n$ sont des éléments de $A$, l'idéal engendré par ces éléments, noté $(a_1, \ldots, a_n)$, est l'ensemble des combinaisons linéaires (à coefficients dans A) de ces éléments. Un idéal pouvant s'écrire sous la forme $(a)$ pour un certain $a \in A$ est un idéal principal.

***Définition 2.4***. — Soit $A$ un anneau. Soit $a \in A$, soit $b \in A$. On dit que $a$ divise $b$ (et on le note $a|b$), ou encore que $b$ est multiple de $a$, s'il existe $x \in A$ tel que $b = ax$, c'est-à-dire que $b \in (a)$, ou que $(b) \subset (a)$, ou encore que la classe $\overline{b}$ de $b$ dans $A/(a)$ est nulle.

***Définition 2.5***. — Soit $A$ un anneau. On dit que $A$ est intègre si $A$ est non nul (*i.e.* $1_A \neq 0_A$) et si pour tous $(x, y) \in A^2$ tels que $xy = 0$, on a $x = 0$ ou $y = 0$. Un idéal $\mathfrak{p}$ de $A$ est premier si l'anneau quotient $A/\mathfrak{p}$ est intègre. Un idéal $\mathfrak{m}$ de $A$ est maximal si et seulement si $\mathfrak{m} \neq A$ et que $\mathfrak{m}$ est maximal pour cette propriété ; ceci équivaut à demander que $A/\mathfrak{m}$ soit un corps.

***Proposition 2.6***. — *Soit $A$ un anneau intègre. Soit $a \in A$, soit $b \in A$ ; alors $(a) = (b)$ si, et seulement si, il existe un élément $u \in A^\times$ tel que $b = ua$.*

Ainsi, dans un anneau intègre, se donner un idéal principal revient à se donner une classe d'équivalence d'éléments de $A$ pour la relation d'équivalence qui identifie $a$ et $b$ s'il existe un inversible $u$ tel que $b = ua$.

***Définition 2.7***. — Soit $A$ un anneau intègre. Soit $a \in A$, soit $b \in A$. Le pgcd de $a$ et de $b$ est un élément $y$ (parfois noté $a \wedge b$) de $A$ divisant $a$ et $b$ et tel que tout élément divisant $a$ et $b$ divise aussi $y$. S'il existe, le pgcd est défini à un facteur inversible près. Le ppcm de $a$ et de $b$ est un élément $y$ (parfois noté $a \vee b$) de $A$ multiple de $a$ et de $b$ tel que tout multiple de $a$ et de $b$ soit aussi un multiple de $y$ [1]. S'il existe, le ppcm est aussi défini à facteur inversible près.

1. On peut étendre la définition du pgcd et du ppcm donnée ici pour deux éléments à une famille d'éléments de $A$.

***Proposition 2.8***. — *Soit $A$ un anneau intègre. Soit $a \in A$, soit $b \in A$. Le ppcm de $a$ et $b$ existe si et seulement si l'idéal $(a) \cap (b)$ est principal ; le ppcm est un générateur de cet idéal.*

***Proposition 2.9***. — *Soit $A$ un anneau intègre. Soit $a \in A$, soit $b \in A$. Si l'idéal $(a, b)$ est principal, alors le pgcd de $a$ et $b$ existe et est un générateur de $(a, b)$.*

Soit $x \in A$. Si $x$ divise $a$ et $b$, alors $x$ divise tous les éléments de l'idéal $(a, b)$. Par conséquent, $x$ divise $a$ et $b$ équivaut à $(a, b) \subset (x)$. Supposons que $y \in A$ soit tel que $(y) = (a, b)$. D'après ce qui précède, $y$ divise $a$ et $b$ et tout élément $x$ qui divise $a$ et $b$ est tel que $(y) \subset (x)$, donc $x$ divise $y$, ce qui montre que $y$ est le pgcd de $a$ et de $b$.

***Définition 2.10***. — Soit $A$ un anneau intègre. Soit $x$ un élément non nul et non inversible de $A$. On dit que $x$ est irréductible si dès lors que l'on a $x = ab$ avec $(a, b) \in A^2$, alors $a$ ou $b$ est inversible. On dit que $x$ est réductible s'il n'est pas irréductible, à savoir que l'on peut écrire $x$ comme produit de deux éléments non inversibles.

### 2.1.2. Restes chinois. —

***Définition 2.11***. — Soit $A$ un anneau. On dit de deux éléments $a$ et $b$ qu'ils sont étrangers si $(a, b) = A$.

***Remarque 2.12***. — Dans un anneau $A$, $a$ et $b$ sont étrangers si et seulement si la classe de $a$ dans $A/(b)$ est inversible.

***Proposition 2.13***. — *Soit $A$ un anneau. Soit $(a, b) \in A^2$. On suppose que $a$ et $b$ sont étrangers. Alors, $(a) \cap (b) = (ab)$, et le morphisme évident d'anneaux*

$$\varphi \colon A/(ab) \to A/(a) \times A/(b)$$

*est un isomorphisme.*

On choisit une relation $au + bv = 1$. On a évidemment $(ab) \subset (a) \cap (b)$. Soit $x \in (a) \cap (b)$. On peut écrire $x = ax'$, $x = bx''$, d'où $x = (au + bv)x = abux'' + abvx' \in (ab)$. Le morphisme d'anneaux considéré est donc injectif.

Considérons l'égalité $au + bv = 1$. On en déduit les congruences $au \equiv 0[a]$ et $au \equiv 1[b]$. Par conséquent, $\varphi([au]) = (0, 1)$. De même, $\varphi([bv]) = (1, 0)$. Si $x$ et $y$ sont deux éléments de $A$, il vient $\varphi([xbv + yau]) = ([x], [y])$, ce qui montre que $\varphi$ est surjectif.

***Théorème 2.14***. — *Soit $A$ un anneau. Soit $n \geq 1$. Soit $(a_1, \ldots, a_n)$ un $n$-uplet d'éléments deux-à-deux étrangers. Le morphisme d'anneaux*

$$\varphi \colon A/(a_1 \ldots a_n) \to A/(a_1) \times \cdots \times A/(a_n)$$

*est un isomorphisme.*

On procède par récurrence sur $n$. Si $n \geq 3$, on peut appliquer le théorème au cran $n - 1$ pour les éléments $a_1, \ldots, a_{n-1}$ et au cran 2 pour le couple $a_1 \ldots a_{n-1}, a_n$.

***Lemme 2.15*** **(Gauß)**. — *Soit $A$ un anneau. On se donne des éléments $(a, b, c)$ de $A$ tels que $a|bc$ et que $a$ et $b$ soient étrangers. Alors, $a|c$.*

Comme $a$ et $b$ sont étrangers, l'image $\overline{b}$ de $b$ dans $A/(a)$ est inversible. Par hypothèse, $\overline{b}\overline{c} = 0$ dans $A/(a)$. On en déduit que $\overline{c} = 0$ dans $A/(a)$, ce qui signifie que $a$ divise $c$.

### 2.1.3. Anneaux principaux. —

***Définition 2.16***. — Un anneau principal est un anneau intègre dont tout idéal est principal.

Dans un anneau principal, les pgcd et ppcm existent toujours.

***Proposition 2.17***. — *L'anneau des entiers* $\mathbf{Z}$ *est principal. Pour tout corps* $k$*, l'anneau des polynômes* $k[X]$ *est principal.*

Ceci sera une conséquence de la proposition 2.22 ci-dessous.

***Théorème 2.18*** **(Bézout)**. — *Soit* $A$ *un anneau principal. Soit* $a \in A$*, soit* $b \in A$*. Il existe* $(u, v) \in A^2$ *tels que* $au + bv = a \wedge b$*.*

***Définition 2.19***. — Soit $A$ un anneau principal. On dit deux deux éléments $a$ et $b$ qu'ils sont premiers entre eux si $a \wedge b = 1$. D'après le théorème de Bézout, ceci équivaut à dire qu'ils sont étrangers.

***Proposition 2.20***. — *Soit* $A$ *un anneau principal. Soit* $x$ *un élément non nul de* $A$*. Alors* $x$ *est irréductible si et seulement s'il engendre un idéal premier. Par ailleurs, tout idéal premier non nul de* $A$ *est maximal.*

Supposons $(x)$ premier. Montrons que $x$ est irréductible. Donnons-nous une écriture $x = ab$ avec $(a, b) \in A^2$. En particulier, $ab \in (x)$, donc $a$ ou $b$ appartient à $(x)$. On veut montrer que $a$ ou $b$ est inversible. Par symétrie des rôles, on peut supposer que $a \in (x)$, il existe $c \in A$ tel que $a = cx$, d'où $x = bcx$. Comme $x$ est non nul, il vient $bc = 1$, par conséquent $b$ est inversible [2].

Supposons maintenant que $x$ est irréductible. Montrons que $(x)$ est un idéal maximal (donc en particulier un idéal premier). Supposons donc $(x) \subset I$ avec $I \neq A$. On peut écrire $I = (y)$ avec $y \in A$. Donc, $y|x$, il existe $z \in A$ tel que $x = yz$. Comme $x$ est irréductible et que $y$ est non inversible, $z$ est inversible, d'où $I = (y) = (x)$, l'idéal $(x)$ est maximal.

### 2.1.4. Anneaux euclidiens. —

***Définition 2.21***. — Un anneau euclidien est un anneau intègre $A$ muni d'une fonction (appelée « jauge ») $v\colon A - \{0\} \to \mathbf{N}$ (on la prolonge à $A$ en posant $v(0) = -\infty$) telle que pour tout $(a, b) \in A^2$ tel que $b \neq 0$, il existe $(q, r) \in A^2$ tel que $a = bq + r$ et $v(r) < v(b)$ [3].

L'anneau des entiers est euclidien : prendre pour $v$ la valeur absolue. L'anneau des polynômes $k[X]$ sur un corps $k$ l'est aussi : prendre pour $v$ le degré des polynômes.

La proposition 2.17 ci-dessus est un cas particulier du résultat général :

***Proposition 2.22***. — *Un anneau euclidien est principal.*

Soit $I$ un idéal d'un anneau euclidien $(A, v)$. Montrons que $I$ est principal. On peut supposer $I$ non nul. L'ensemble des $v(x)$ pour $x \in I - \{0\}$ est un sous-ensemble non vide de $\mathbf{N}$, il admet un plus petit élément $v(b)$ avec $b \in I$. L'idéal $I$ est engendré par $b$. En effet, soit $x \in I$, on peut faire la division euclidienne de $x$ par $b$, ce qui fournit des éléments $q$ et $r$ tels que $x = bq + r$ avec $v(r) < v(b)$. Comme $r \in I$, la propriété de minimalité de $v(b)$ interdit que $r \neq 0$. Donc $r = 0$, d'où $x = bq$, ce qui achève de montrer l'inclusion $I \subset (b)$ ; l'autre inclusion étant évidente, on a démontré la proposition.

***Proposition 2.23***. — *Dans* $k[X]$*, le coût en opérations arithmétiques (additions, multiplications, divisions dans* $k$*) de la division euclidienne d'un polynôme* $A$ *par un polynôme* $B$ *est en* $O((\deg A + 1) \cdot (\deg B + 1))$*, soit* $O(\deg A \cdot \deg B)$ *si les polynômes sont non constants. Plus précisément, si* $\deg A > \deg B > 0$*, le coût est un* $O(\deg B \cdot (\deg A - \deg B))$*.*

2. On notera que ce sens d'implication vaut dans un anneau intègre quelconque.
3. Déterminer un tel couple $(q, r)$ s'appelle « faire la division euclidienne de $a$ par $b$ ».

Notons $a = \deg A$ et $b = \deg B$, et supposons pour simplifier que $a > b$. Si on note $Q$ le quotient de la division euclidienne, on a : $Q = q_{a-b}X^{a-b} + \dots$. L'algorithme consiste à déterminer successivement $q_{a-b}, q_{a-b-1}, \dots, q_0$ de la façon suivante.

On définit une suite de polynômes $R_i$ par : $R_0 = A$. Pour $i$ allant de 0 jusqu'à $a - b$, on note $q_{a-b-i}$ le quotient du coefficient de $X^{a-i}$ dans $R_i$ par celui de $X^b$ dans $B$. On calcule la différence $R_{i+1} := R_i - q_{a-b-i}X^{a-b-i}B$. À la sortie de la boucle, les coefficients de $Q$ ont tous été déterminés et le reste $R$ de la division euclidienne est $R_{a-b+1}$.

En effet, on dispose d'un invariant de boucle suivant : au début de l'étape $i$, « $A = \left(\sum_{k=0}^{i-1} q_{a-b-k}X^{a-b-k}\right) \cdot B + R_i$ et $\deg R_i \leq a - i$ ».

À chaque étape de la boucle, on fait $O(b)$ opérations. Comme on passe $a - b + 1$ fois dans la boucle, on obtient un coût en $O(b(a - b))$ dès lors que $a > b$.

## 2.2. Algorithme d'Euclide

### 2.2.1. Calcul du pgcd. —

***Proposition 2.24.*** — *Soit $A$ un anneau euclidien, de jauge $v$. Soit $(a, b) \in A^2$. On suppose que $b \neq 0$. On effectue la division euclidienne de $a$ par $b$ : on détermine $(q, r) \in A^2$ tels que $a = bq + r$ et $v(r) < v(b)$. Alors, $a \wedge b = b \wedge r$.*

Il suffit de montrer l'égalité des idéaux $(a, b) = (a - bq, b)$.

L'algorithme d'Euclide permet de déterminer le pgcd de deux éléments $a$ et $b$ dans un anneau euclidien $A$. On distingue plusieurs cas :

(i) si $b$ est nul, on renvoie $a$ ;
(ii) sinon, on calcule le reste $r$ de la division euclidienne de $a$ par $b$, on remplace $(a, b)$ par $(b, r)$ et on recommence.

Si les éléments $a$ et $b$ vérifient $v(b) < v(a)$, on va conserver cette propriété lors des itérations suivantes. La valeur $v(a)$ va décroître strictement à chaque itération, ce qui assure que l'algorithme termine et qu'il renvoie le bon résultat d'après la proposition ci-dessus.

Si on a $v(b) \geq v(a)$, on aura $v(b) < v(a)$ lors de l'itération suivante, et le raisonnement précédent s'applique. Ainsi, l'algorithme d'Euclide termine toujours et renvoie un résultat correct.

**2.2.2. Détermination de coefficients de Bézout.** — Notons $A$ et $B$ les éléments dont on veut déterminer le pgcd ainsi que des coefficients de Bézout $u$ et $v$ tels que $Au + Bv = A \wedge B$. Pour déterminer ces coefficients de Bézout, il suffit de garder en mémoire une représentation de $a$ et de $b$ sous la forme d'une combinaison linéaire de $A$ et de $B$. En effet, à la terminaison de l'algorithme, le pgcd renvoyé est $a$, si on sait écrire $a$ comme combinaison linéaire de $A$ et de $B$, on peut renvoyer les coefficients de Bézout. On parle alors d'algorithme d'Euclide étendu.

On peut stocker les coefficients des combinaisons linéaires sont la forme d'une matrice $M$ telle que

$$\begin{pmatrix} a \\ b \end{pmatrix} = M \cdot \begin{pmatrix} A \\ B \end{pmatrix} .$$

```
def euclide_etendu(A,B):
    a=A
    b=B
    M=matrix(2,2,[1,0,0,1]);
    while b<>0:
        D=a//b
        r=a%b
        a=b
        b=r
        M=matrix(2,2,[0,1,1,-D])*M
    return (a,M[0,0],M[0,1])
```

On peut montrer que l'algorithme d'Euclide étendu fournit une relation de Bézout minimale, au sens où $|u| < \frac{B}{2A\wedge B}$ et $|v| < \frac{A}{2A\wedge B}$ en supposant $A, B \in \mathbb{N}^*$. La même remarque s'applique au cas polynomial : on obtient $\deg u < \deg B$ et $\deg v < \deg A$.

**2.2.3. Coût en temps.** — On va estimer le nombre d'étapes utilisées par l'algorithme d'Euclide pour déterminer un pgcd. On rappelle que la suite de Fibonacci est définie par $F_0 = 0$, $F_1 = 1$ et $F_{n+2} = F_n + F_{n+1}$ pour tout $n \geq 0$.

***Théorème 2.25* (Lamé).** — *Soit $(a, b)$ un couple d'entiers naturels tel que $a > b$. Soit $n$ un entier naturel non nul. Si l'algorithme d'Euclide pour le calcul du pgcd de $(a, b)$ nécessite au moins $n$ étapes, alors $a \geq F_{n+2}$ et $b \geq F_{n+1}$. En outre, l'algorithme d'Euclide fait le calcul du pgcd de $F_{n+2}$ et de $F_{n+1}$ en exactement $n$ étapes.*

On procède par récurrence sur $n$. Dire que le calcul nécessite zéro étape signifie que $b = 0$ (et $a \geq 1$), l'affirmation est donc fausse pour $n = 0$. Si $n = 1$, on a bien évidemment $b \geq 1$ et donc $a \geq 2$. Supposons que $n \geq 2$ et que le résultat est connu pour $n - 1$. On applique la première étape de l'algorithme d'Euclide à $(a, b)$, on obtient ainsi un nouveau couple $(b, r)$ (avec $r$ le reste de la division euclidienne de $a$ par $b$), et le calcul de l'algorithme d'Euclide pour $(b, r)$ nécessite $n - 1$ étapes. L'hypothèse de récurrence montre que $b \geq F_{n+1}$ et $r \geq F_n$. On a $a = bq + r$, où $q$ est le quotient de la division euclidienne de $a$ par $b$. Bien entendu, $q \geq 1$, d'où $a \geq b + r \geq F_{n+1} + F_n = F_{n+2}$. La « réciproque » est claire.

***Corollaire 2.26.*** — *Le coût en temps de l'algorithme d'Euclide (étendu ou non) sur des entiers $(a, b)$ est en $O(\log_\phi \min(a, b))$.*

La vitesse de calcul d'une relation de Bezout entre deux entiers premiers entre eux est utilisée dans le chiffrement RSA pour calculer la clef privée.

***Proposition 2.27.*** — *Le coût en nombre d'étapes de l'algorithme d'Euclide (étendu ou non) sur des polynômes $A, B \in k[X]$ non constants est en $O(\min(\deg A, \deg B))$ ; son coût en opérations arithmétiques (additions, multiplications, divisions dans $k$) est en $O(\deg A \cdot \deg B)$.*

Pour simplifier, on suppose que $\deg A > \deg B$.

Notons $P_0 = A, P_1 = B, P_2, \ldots, P_e, P_{e+1}$ la suite de polynômes définie par le fait que $P_{i+2}$ soit le reste de la division euclidienne de $P_i$ par $P_{i+1}$. On définit $e$ par $P_e \neq 0$ et $P_{e+1} = 0$. Notons $d_i = \deg P_i$, de sorte que l'on ait $d_0 > d_1 > \cdots > d_e$. D'après le calcul du coût d'une division euclidienne, il existe une constante $\lambda > 0$ tel que le calcul de $P_{i+2}$ connaissant $P_i$ et $P_{i+1}$ ait un coût $\leq \lambda d_{i+1} \cdot (d_i - d_{i+1}) \leq \lambda \deg B \cdot (d_i - d_{i+1})$.

En faisant la somme, on obtient que le pgcd utilise un coût $\leq \lambda \deg B \cdot (\deg A - \deg P_e) \leq \lambda \deg A \cdot \deg B$.

Pour les coefficients de Bézout, il faut une analyse plus précise. Notons $Q_i$ le quotient de la division euclidienne de $P_i$ par $P_{i+1}$ (de sorte que $P_{i+2} = P_i - Q_i P_{i+1}$). Notons $(U_i, V_i)$ des polynômes définis par $(U_0, V_0) = (1, 0)$, $(U_1, V_1) = (0, 1)$, puis $(U_{i+2}, V_{i+2}) = (U_i, V_i) - Q_i \cdot (U_{i+1}, V_{i+1})$. Ils vérifient $P_i = AU_i + BV_i$.

On montre facilement que pour tout $i$, on a $\deg U_{i+2} = \deg Q_1 + \deg Q_2 + \cdots + \deg Q_i = d_1 - d_{i+1}$ et $\deg V_{i+2} = \deg Q_0 + \cdots + \deg Q_i = d_0 - d_{i+1}$. Ceci se vérifie par récurrence, en commençant par observer d'une part que $U_2 = 1$ et $V_2 = -Q_0$ et d'autre part que $U_3 = -Q_1$ et $V_3 = 1 + Q_0 Q_1$. Ensuite, dans la différence $U_{i+2} = U_i - Q_i U_{i+1}$, le terme $Q_i U_{i+1}$ est de degré strictement supérieur à celui de $U_i$, de sorte que $\deg U_{i+2} = \deg U_{i+1} + \deg Q_i$. Un raisonnement semblable vaut pour les $V_i$.

On s'intéresse ainsi au calcul de $U_e$ et de $V_e$, vu que $P_e = U_e A + V_e B$. Les résultats précédents montrent que les polynômes $U_0, U_1, \ldots, U_e$ ont des degrés $\leq \deg B - \deg P_{e-1} < \deg B$. De même, les polynômes $V_0, V_1, \ldots, V_e$ ont des degrés $< \deg A$. Le coût du calcul de $U_{i+2}$ en fonction de $U_i, U_{i+1}, Q_i$ est donc en $O((d_i - d_{i+1}) \cdot \deg B)$. En sommant ce coût pour les différents indice $i$, on obtient un coût en $O(\deg A \cdot \deg B)$. De même, $V_2$ se calcule certainement en $O(\deg A)$, puis pour $i \geq 1$, le coût du calcul de $V_{i+2}$ en fonction de $V_i, V_{i+1}, Q_i$ est en $O(\deg A \cdot (d_i - d_{i+1}))$, ce qui donne au total, un coût aussi en $O(\deg A \cdot \deg B)$.

# CHAPITRE 3

# PIVOT DE GAUSS ET APPLICATIONS

*Ce chapitre est extrait d'un poly de Daniel Perrin.*

## 3.1. Matrices échelonnées, éventuellement réduites

Soit $M$ une matrice rectangulaire (pas nécessairement carrée) à coefficients dans un corps $k$. Rappelons les trois sortes d'opérations élémentaires qu'on peut effectuer sur les lignes de $M$ :

($E_1$) Échanger les lignes $L_i$ et $L_j$. Cette opération est notée $L_i \leftrightarrow L_j$.

($E_2$) Multiplier une ligne $L_i$ par un scalaire $a \neq 0$, opération notée $L_i \leftarrow aL_i$.

($E_3$) Ajouter à une ligne $L_i$ un multiple $aL_j$ d'une autre ligne $L_j$ : $L_i \leftarrow L_i + aL_j$.

Lorsqu'on effectue une telle opération, la matrice obtenue est dite *équivalente par lignes* à celle de départ. Plus généralement, deux matrices sont dites *équivalentes par lignes* si on peut passer de l'une à l'autre par une suite d'opérations élémentaires.

On va voir comment se ramener, à l'aide d'opérations élémentaires, à une matrice d'une forme particulière : échelonnée-ligne, voire échelonnée-ligne réduite. Les définitions sont les suivantes ; on omet le mot "ligne" pour alléger.

***Définition 3.1***. — On appelle *matrice échelonnée* toute matrice rectangulaire qui satisfait aux deux conditions suivantes :
1) Si il y a des lignes nulles (c'est-à-dire formées seulement de zéros), alors elles sont tout en bas (en-dessous de toutes les autres).
2) Le premier élément non nul de chaque ligne est toujours plus à droite que le premier élément non nul de la ligne précédente.

***Définition 3.2***. — On appelle *matrice échelonnée réduite* toute matrice rectangulaire qui satisfait aux quatre conditions suivantes :
1) Si il y a des lignes nulles (c'est-à-dire formées seulement de zéros), alors elles sont tout en bas (en-dessous de toutes les autres).
2) Le premier élément non nul de chaque ligne est toujours plus à droite que le premier élément non nul de la ligne précédente.
3) Sur chaque ligne, le premier élément non nul vaut 1 (sauf bien sûr si cette ligne n'est constituée que de zéros).
4) Dans la colonne du premier élément non nul de chaque ligne, il n'y a que des 0 (hormis cet élément, qui vaut 1 d'après ce qui précède).

Bien sûr les conditions 1) et 2) sont celles qui définissent une matrice échelonnée ; toute matrice échelonnée réduite est donc échelonnée (ce qui justifie la terminologie).

L'intérêt de cette définition provient du théorème suivant :

***Théorème 3.3***. — *Toute matrice rectangulaire est équivalente par lignes à une et une seule matrice échelonnée réduite. De plus cette matrice échelonnée réduite s'obtient en appliquant à la matrice de départ la méthode du pivot de Gauss.*

### 3.2. Algorithme du pivot de Gauss

Expliquons maintenant comment appliquer à une matrice la méthode du pivot de Gauss. On part d'une matrice $M$ dont les coefficients sont les $a_{ij} \in k$. On suppose que $M$ n'est pas la matrice nulle (sinon elle est déjà échelonnée réduite, et il n'y a rien à faire).

L'algorithme est alors le suivant.

**Étape 1** Localiser la première colonne de $M$ dans laquelle figure au moins un coefficient non nul ; disons que c'est la $j$-ième.

**Étape 2** Échanger éventuellement deux lignes de telle sorte que le coefficient $a_{1j}$ soit $\neq 0$. (En effet, par définition de $j$ à l'étape précédente on sait que l'un au moins des coefficients de la $j$-ième colonne est non nul)

**Étape 3** Si $a_{1j} \neq 1$, diviser la ligne $L_1$ par $a_{1j}$ pour se ramener au cas où $a_{1j} = 1$ (i.e. appliquer l'opération élémentaire $L_1 \leftarrow \dfrac{1}{a_{1j}} L_1$).

**Étape 4** Utiliser le coefficient $a_{1j}$ (qui vaut maintenant 1) comme pivot pour rendre nul le $j$-ième coefficient de toutes les lignes $L_i$, $i \geq 2$. Pour cela on applique les opérations élémentaires $L_i \leftarrow L_i - a_{ij} L_1$ pour tout $i \geq 2$.

**Étape 5** Répéter les étapes 1, 2, 3, 4 avec la matrice obtenue à partir de $M$ en supprimant la première ligne, jusqu'à ce que cette nouvelle matrice soit la matrice nulle.

**Étape 6** En partant du bas, considérer successivement toutes les lignes non nulles. Pour chaque ligne non nulle, le premier coefficient non nul est un 1 ; disons que ce 1 est sur la ligne $i$ et la colonne $j$. Alors pour chaque $i'$ compris entre 1 et $i-1$, on applique l'opération élémentaire $L_{i'} \leftarrow L_{i'} - a_{i'j} L_i$. Cela permet de mettre un zéro dans chacune des cases situées au-dessus du coefficient $a_{ij} = 1$. Puis on passe à la ligne située juste au-dessus, et on répète ces opérations jusqu'à la première ligne.

A la fin de l'étape 6, la matrice obtenue est échelonnée réduite (il suffit de s'arrêter à l'étape 5 si on veut simplement trouver une matrice échelonnée équivalente par lignes à la matrice de départ). Le théorème énoncé ci-dessus affirme que toute matrice est équivalente par lignes à une et une seule matrice échelonnée réduite. En revanche, une matrice donnée est (en général) équivalente par lignes à de nombreuses matrices échelonnées, que l'on peut trouver en appliquant différentes variantes de l'algorithme du pivot de Gauss.

### 3.3. Application à la résolution des systèmes linéaires

Considérons un système linéaire de $m$ équations à $p$ inconnues :

$$(S) \qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p &= b_2 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad\quad \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mp}x_p &= b_m \end{cases}$$

Les $a_{ij}$ et les $b_i$, avec $1 \leq i \leq m$ et $1 \leq j \leq p$, sont des éléments d'un corps $k$ ; il s'agit respectivement des coefficients et des seconds membres. La *matrice des coefficients* du système $(S)$

est $A = [a_{ij}]$ ; la matrice colonne $B$ du second membre est $[b_i]$. On réunit ces deux matrices pour définir la *matrice augmentée* du système $(S)$, qui est la matrice suivante :

$$M = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1p} & b_1 \\ a_{21} & a_{22} & \dots & a_{2p} & b_2 \\ \vdots & \vdots & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mp} & b_m \end{array}\right] \in \mathbf{M}_{m,p+1}(k).$$

La barre verticale qui sépare la dernière colonne n'a pas vraiment de signification mathématique ; elle indique simplement que la dernière colonne s'interprète comme celle du second membre, alors que les autres correspondent aux coefficients du système.

On remarque que pour passer du système $(S)$ à sa matrice augmentée, on omet simplement d'écrire les inconnues $x_1, \dots, x_p$, les signes $+$, et on remplace le symbole $=$ par une barre verticale. On peut bien sûr procéder dans l'autre sens pour passer de la matrice augmentée au système.

On peut effectuer des opérations élémentaires sur une matrice augmentée, comme sur n'importe quelle matrice rectangulaire : on ne se soucie pas du trait vertical qui sépare la dernière colonne des précédentes. En appliquant l'algorithme du pivot de Gauss à $M$, deux cas peuvent se présenter :

- Si, au cours de l'algorithme, on s'aperçoit que $M$ est équivalente par lignes à une matrice augmentée dont une ligne est formée seulement de zéros, sauf le dernier coefficient qui est non nul, alors le système $(S)$ n'a aucune solution.
- Dans le cas contraire, c'est-à-dire si dans les matrices obtenues à partir de la fin de l'étape 5 le coefficient à droite de la barre est nul dans toutes les lignes pour lesquelles tous les coefficients à gauche de la barre sont nuls, alors il y a au moins une solution. Il peut y avoir exactement une solution, ou bien une infinité de solutions. Précisément, une fois qu'on a supprimé les lignes de zéros qui figurent en bas de la matrice échelonnée réduite donnée par l'algorithme à la fin de l'étape 6, le nombre de lignes qui restent est le rang du système, noté $r$. Seuls deux cas sont possibles, en fonction du nombre de colonnes de la matrice des coefficients (qui est le nombre d'inconnues, noté $p$) :
  - Soit $p = r$ : alors le système $(S)$ a exactement une solution. Pour la trouver, on écrit le système correspondant à la matrice échelonnée réduite (dans laquelle on a supprimé les lignes de zéros éventuelles) : chaque équation donne directement la valeur d'une des inconnues.
  - Soit $p > r$. Alors il y a une infinité de solutions, et les inconnues $x_1, \dots, x_p$ se répartissent en deux groupes : les inconnues $x_j$ pour lesquelles $j$ est l'indice de colonne du premier élément non nul d'une ligne sont dites *principales*, et les autres *secondaires*. Les inconnues secondaires peuvent prendre des valeurs arbitraires ; une fois que ces valeurs sont fixées, chaque inconnue principale est déterminée de manière unique en écrivant le système correspondant à la matrice échelonnée réduite. Il y a $r$ inconnues principales, et $p - r$ secondaires ; les solutions s'expriment donc en fonction de $p - r$ paramètres qui peuvent prendre des valeurs arbitraires.

## 3.4. Complexité et autres applications

Soit $A$ une matrice à coefficients dans un corps $k$. Pour simplifier, on suppose $A$ carrée de taille $n$. Alors le coût en nombre d'opérations arithmétiques dans $k$ (additions, multiplications, divisions) de la mise sous forme échelonnée réduite de $A$ est $O(n^3)$. Les problèmes suivants, qui s'en déduisent facilement, ont donc aussi un coût en $O(n^3)$ :

- Etant donné $B \in M_{n,1}(k)$, la résolution du système linéaire $AX = B$.
- La détermination d'une base du noyau de $A$.
- La détermination d'un système d'équations linéaires de l'image de $A$.
- Le calcul du rang de $A$.
- Le calcul du déterminant de $A$ : une fois sous forme échelonnée réduite le déterminant se calcule trivialement, et à chaque étape de l'algorithme on peut suivre l'évolution du déterminant.

- Le calcul de l'inverse de $A$, si $A$ est inversible.

### 3.5. Déterminant sur Z et $k[X]$

Les méthodes précédentes fonctionnent bien sur un corps. Lorsque les coefficients des matrices sont dans $\mathbf{Z}$ ou dans $k[X]$ (avec $k$ un corps), l'algorithme du pivot de Gauss nécessite de passer dans $\mathbf{Q}$ ou $k(X)$, ce qui rend moins efficaces les calculs. Dans ce paragraphe on esquisse une méthode pour éviter l'apparition de dénominateurs dans le calcul d'un déterminant.

Soit $M \in \mathbf{M}_n(\mathbf{Z})$. Si tous les coefficients de $M$ sont compris entre $-h$ et $h$ alors on a $|\det M| \leq H$ en posant $H = n!h^n$. On choisit des nombres premiers $p_1, \ldots, p_r$ deux à deux distincts tels que $p_1 \cdots p_r > 2H + 1$. Pour tout $i$ compris entre 1 et $r$, on calcule la classe modulo $p_i$ de $\det M$ : c'est le déterminant de la matrice obtenue en réduisant modulo $p_i$ les coefficients de $M$, et on le détermine en $O(n^3)$ opérations élémentaires dans $\mathbf{Z}/p_i\mathbf{Z}$. Le théorème chinois montre alors que $\det M$ est l'unique entier compris entre $-H$ et $H$ vérifiant ces congruences. On détermine $\det M$ d'autant plus rapidement que les nombres $p_i$ peuvent être choisis à l'avance (en fonction d'un majorant des valeurs absolues des coefficients de $M$), donc une partie de la mise en œuvre du théorème chinois n'a pas besoin d'être répétée avec chaque nouvelle matrice $M$.

La méthode analogue pour calculer le déterminant de $M \in \mathbf{M}_n(k[X])$ consiste à évaluer $\det M$ en des points $x_1, \ldots, x_r$ de $k$ bien choisis ; on calcule $(\det M)(x_i) = \det(M(x_i))$ en $O(n^3)$ opérations élémentaires dans $k$. Ensuite, il ne reste qu'à interpoler en implémentant l'algorithme d'interpolation de Lagrange ; évidemment il faut choisir $r > \deg \det M$.

# CHAPITRE 4

# CORPS FINIS

## 4.1. Premières observations

***Proposition 4.1***. — *Soit $n$ un entier naturel. L'anneau $\mathbf{Z}/n\mathbf{Z}$ est un corps si et seulement si $n$ est un nombre premier.*

Ceci découle de la proposition 2.20 appliquée à l'anneau principal $\mathbf{Z}$.

***Définition 4.2***. — Soit $p$ un nombre premier. On note $\mathbf{F}_p$ le corps $\mathbf{Z}/p\mathbf{Z}$.

***Proposition 4.3***. — *Soit $\mathbf{F}$ un corps fini. Il existe un unique nombre premier $p$ tel que $p \cdot 1_{\mathbf{F}} = 0$, c'est la caractéristique de $\mathbf{F}$. On peut considérer $\mathbf{F}$ comme un $\mathbf{F}_p$-espace vectoriel. Le cardinal $q$ de $\mathbf{F}$ est $p^d$ où $d = \dim_{\mathbf{F}_p} \mathbf{F}$.*

***Définition 4.4***. — Si on dispose d'un corps fini à $q$ éléments, il est permis de le noter $\mathbf{F}_q$.

***Proposition 4.5***. — *Soit $\mathbf{F}_q$ un corps fini à $q$ éléments. Pour tout $x \in \mathbf{F}_q$, on a $x^q = x$.*

Le résultat est vrai pour $x = 0$. Pour $x \neq 0$, il signifie que $x^{q-1} = 1$, ce qui résulte du théorème de Lagrange appliqué au groupe multiplicatif $\mathbf{F}_q^\times$.

***Proposition 4.6***. — *Soit $p$ un nombre premier. Soit $A$ un anneau dans lequel $p$ est nul. Si $(a, b) \in A^2$, on a $(a + b)^p = a^p + b^p$.*

On utilise la formule du binôme et on utilise le fait que $p$ divise $\begin{pmatrix} p \\ i \end{pmatrix}$ pour $1 \leq i \leq p - 1$.

***Définition 4.7***. — Soit $p$ un nombre premier. Pour tout anneau $A$ dans lequel $p$ est nul, on appelle Frobenius le morphisme d'anneaux $\varphi\colon A \to A$ qui à $a$ associe $a^p$.

***Proposition 4.8***. — *Soit $\mathbf{F}$ un corps fini à $q$ éléments. Le groupe multiplicatif $\mathbf{F}^\times$ est cyclique.*

Si $q = 2$, le résultat est trivial. D'après le théorème de structure des groupes abéliens finis, il existe d'uniques entiers non nuls $(a_1, \ldots, a_d)$ avec $a_d | a_{d-1} | \ldots | a_2 | a_1$ avec $a_d \geq 2$ tel que $\mathbf{F}^\times$ soit isomorphe à $\mathbf{Z}/a_1\mathbf{Z} \times \ldots \mathbf{Z}/a_d\mathbf{Z}$. Pour tout $x \in \mathbf{F}^\times$, on a $x^{a_1} = 1$, donc $\mathbf{F}^\times$ a au plus $a_1$ éléments, d'où $a_1 \ldots a_d \leq a_1$, ce qui montre que $d = 1$, et par conséquent $\mathbf{F}^\times \simeq \mathbf{Z}/(q-1)\mathbf{Z}$.

## 4.2. Terminologie générale

### 4.2.1. Extensions, degrés. —

***Définition 4.9***. — Une extension de corps $K/k$ est la donnée de deux corps $K$ et $k$ et d'un morphisme (forcément injectif) $\varphi\colon k \to K$ (parfois appelé « plongement »). Bien souvent, on identifie $k$ à un sous-corps de $K$. Si $K/k$ et $L/k$ sont deux extensions, un $k$-plongement de $K$ dans $L$ est un plongement $K \to L$ qui vaille l'identité sur $k$. Deux extensions $K/k$ et $L/k$ d'un même corps $k$ sont dites isomorphes s'il existe un isomorphisme $\sigma\colon K \to L$ « fixant » les éléments de

$k$. Une extension $K/k$ est finie si $K$ est un $k$-vectoriel de dimension finie. Un $k$-sous-corps d'une extension $L$ de $k$ est un corps $K$ tel que $k \subset K \subset L$.

Par exemple, pour tout corps fini $\mathbf{F}$ de caractéristique $p$, on a une extension finie $\mathbf{F}/\mathbf{F}_p$.

***Définition 4.10***. — Soit $K/k$ une extension finie. On note $[K : k]$ la dimension de $K$ comme $k$-vectoriel, c'est le degré de l'extension $K/k$.

***Proposition 4.11***. — *Si $K/k$ et $L/K$ sont des extensions finies, alors l'extension composée $L/k$ est une extension finie, et $[L : k] = [L : K] \cdot [K : k]$.*

Si $a_1, \ldots, a_n$ est une $k$-base de $K$ et $b_1, \ldots, b_m$ une $K$-base de $L$, alors les $a_i b_j$ forment une $k$-base de $L$.

***Proposition 4.12***. — *Soit $k$ un corps. Soit $A$ une $k$-algèbre de dimension finie comme $k$-espace vectoriel. Si $A$ est intègre, alors $A$ est un corps.*

La multiplication par un élément non nul de $A$ définit un endomorphisme $k$-linéaire injectif de $A$. Il est donc bijectif. En particulier, tout élément non nul de $A$ est inversible.

### 4.2.2. Corps de rupture. —

***Proposition 4.13***. — *Soit $k$ un corps. Soit $P \in k[X]$ un polynôme non constant. L'anneau $k[X]/(P)$ est un corps si et seulement si $P$ est irréductible.*

Ceci découle de la proposition 2.20 appliquée à l'anneau principal $k[X]$.

***Définition 4.14***. — Soit $k$ un corps. Le corps de rupture d'un polynôme irréductible $P \in k[X]$ est une extension de $k$ munie d'un isomorphisme avec l'extension $k[X]/(P)$ de $k$. Il s'agit d'une extension finie de degré $\deg P$.

Comme le montre la proposition suivante, la donnée d'un corps de rupture de $P$ est équivalente à celle d'une extension $K$ de $k$ et d'un élément $x \in K$ tels que $P(x) = 0$ et $K = k(x)$.

***Proposition 4.15***. — *Soit $k$ un corps. Soit $P \in k[X]$ irréductible. Soit $K/k$ le corps de rupture de $P$. Soit $L/k$ une extension. Se donner un $k$-plongement $K \to L$ revient à se donner un élément $x$ de $L$ tel que $P(x) = 0$.*

On peut supposer que $K = k[X]/(P)$. On considère l'application $\varphi$ qui à un $k$-morphisme $f\colon K \to L$ associe $f([X])$. Pour tout tel $f$, on a évidemment $f([\sum_{i=0}^n a_i X^i]) = \sum_{i=0}^n a_i f(X)^i$ pour tous $a_0, \ldots, a_n$ éléments de $k$. L'application $\varphi$ est injective : si $f([X]) = f'([X])$, alors d'après la formule précédente, $f$ et $f'$ coïncident sur tous les éléments de $k[X]$. L'application $\varphi$ est surjective. En effet, soit $x \in L$ tel que $P(x) = 0$. On considère l'application $g\colon k[X] \to L$ qui à $\sum_i a_i X^i$ associe $\sum_i a_i x^i$. Il s'agit évidemment d'un morphisme d'anneaux. On a évidemment $g(P) = 0$. Par suite, $g$ s'annule sur l'idéal engendré par $P$ ; par conséquent, $g$ induit un morphisme d'anneaux $\overline{g}\colon k[X]/(P) \to L$ envoyant bien $X$ sur $x$.

***Proposition 4.16***. — *Soit $K/k$ une extension finie de corps. Soit $x \in K$. Il existe un unique polynôme unitaire irréductible $P \in k[X]$ tel que $P(x) = 0$ : c'est le polynôme minimal de $x$. Le plus petit $k$-sous-corps de $K$ contenant $x$, noté $k(x)$, est le corps de rupture de $P$ : $k(x) \simeq k[X]/(P)$. Le degré de $x$ sur $k$ est le degré de $P$, qui est aussi $[k(x) : x]$.*

La famille $1, x, x^2, x^3, \ldots$ n'est pas une famille libre du $k$-vectoriel $K$. Le noyau $I$ du morphisme canonique de $k$-algèbres $k[X] \to K$ envoyant $X$ sur $x$ est non nul. Notons $P$ le polynôme unitaire engendrant $I$. Le morphisme induit $k[X]/(P) \to K$ est injectif, donc $k[X]/(P)$ est intègre : $(P)$ est un idéal premier, c'est-à-dire que $P$ est irréductible, etc.

***Définition 4.17***. — Soit $K/k$ une extension finie. Soit $(x_1, \ldots, x_n)$ un $n$-uplet d'éléments de $K$. On note $k[x_1, \ldots, x_n]$ (ou $k(x_1, \ldots, x_n)$) le plus petit $k$-sous-corps de $K$ contenant les $x_i$ : il est formé des valeurs des polynômes de $k[X_1, \ldots, X_n]$ évalués en les $x_i$.

***Définition 4.18***. — Soit $K/k$ une extension finie. On dit que $K/k$ admet un élément primitif $x$ s'il existe $x \in K$ tel que $K = k[x]$.

***Remarque 4.19***. — Toute extension $K/k$ de corps finis admet un élément primitif : prendre un générateur du groupe $K^\times$.

***Remarque 4.20***. — Si $P \in k[X]$ est un polynôme irréductible de degré $d$, on peut représenter les éléments du corps de rupture $k[X]/(P)$ (on note $x$ la classe de $X$), sous la forme $a_0 + a_1 x + \cdots + a_{d-1}x^{d-1}$ avec $(a_0, \ldots, a_{d-1}) \in k^d$, autrement dit sous la forme $Q(x)$ où $Q \in k[X]$ est un polynôme de degré au plus $d-1$. Pour exprimer un produit $Q_1(x)Q_2(x)$ sous cette forme, on considère le reste $R$ de la division euclidienne de $Q_1Q_2$ par $P$, et on a $Q_1(x)Q_2(x) = R(x)$ dans le corps de rupture $k[X]/(P)$.

***Proposition 4.21***. — *Soit $P \in k[X]$ un polynôme de degré $d \geq 2$. Alors, $P$ est irréductible si et seulement si $P$ n'a de racine dans aucune extension de $k$ de degré $e \leq \lfloor \frac{d}{2} \rfloor$.*

Si $P$ est réductible, il s'écrit $P = P_1 \ldots P_k$ avec $P_i$ des polynômes irréductibles. On peut supposer que $P_1$ est celui qui a le plus petit degré $e$. Comme $k \geq 2$, on a bien sûr $\deg P_1 = e \leq \frac{d}{2}$. Notons $K$ le corps de rupture de $P_1$ sur $P$. $K$ est une extension de degré $e$ dans laquelle $P_1$, et donc $P$, admettent une racine.

Inversement, soit $K$ une extension de degré $e \leq \frac{d}{2}$ de $k$ dans laquelle $P$ admet une racine $x$. Notons $Q$ le polynôme minimal de $x$, c'est un polynôme irréductible. Comme $P(x) = 0$, $Q$ est bien un diviseur de $P$. Le degré de $Q$ est celui de l'extension $k[x]/k$ qui est inférieur à celui de $K/k$ qui est lui-même inférieur à $\frac{d}{2}$.

***Remarque 4.22***. — Ce critère est très utile pour les corps finis puisque nous verrons que ceux-ci n'admettent à isomorphisme près qu'une seule extension de degré donné.

### 4.2.3. Composées de deux extensions. —

***Proposition 4.23***. — *Soit $K/k$ et $L/k$ deux extensions finies. Alors, il existe une extension $M/k$ munie de deux $k$-plongements $M/K$ et $M/L$, telle que tout élément de $M$ s'écrive $\sum_i x_i y_i$ avec les $x_i$ dans $K$ et les $y_i$ dans $L$ (on note $M = KL$). On dit que $M/k$ est une extension composée de $K/k$ et de $L/k$.*

On commence par le cas où $K/k$ admet un élément primitif $x$. Soit $P$ le polynôme minimal de $x$ sur $k$. On a $K = k[X]/(P)$. On choisit un facteur irréductible $Q$ de $P$ vu comme élément de $L[X]$. On note $M = L[X]/(Q)$. Par construction, $M$ est une $k$-extension de $L$. Par ailleurs, $M$ contient une racine privilégiée de $Q$, donc de $P$, on peut aussi en faire une extension de $K$ en envoyant $X$ sur cette racine. Par conséquent, on a bien $M = KL$.

Dans le cas général, on procède par récurrence sur $[K : k]$. Si $[K : k] = 1$, c'est trivial, sinon, on choisit $x \in K - k$, on note $K' = k[x]$. On applique ce qui précède pour construire une extension composée $M'/k$ de $K'/k$ et de $L/k$. On considère ensuite les extensions $K$ et $M'$ de $K'$, le degré $[K : K']$ est strictement moindre que $[K : k]$, l'hypothèse de récurrence permet de trouver une extension composée $M/K'$ de $K/K'$ et de $M'/K'$.

### 4.2.4. Polynômes scindés, corps de décomposition. —

***Définition 4.24***. — Un polynôme non nul $P \in k[X]$ est dit scindé s'il peut s'écrire $\lambda(X - a_1)\ldots(X - a_n)$ avec $a_1, \ldots, a_n$ et $\lambda$ des éléments de $k$. On dit aussi que $P$ a toutes ses racines dans $k$.

Un diviseur d'un polynôme scindé est encore scindé. Le produit de deux polynômes scindés est scindé.

***Proposition 4.25***. — *Soit $P \in k[X]$ un polynôme non nul. Il existe une extension finie $K/k$ telle que, en tant qu'élément de $K[X]$, $P$ soit scindé.*

On peut supposer que $P$ est unitaire. On procède par récurrence sur $\deg P$. Si $\deg P \leq 1$, $K = k$ convient. Si $P$ n'est pas irréductible, on écrit $P = QR$ avec $Q$ et $R$ non constants (unitaires), et donc de degrés strictement plus petits que celui de $P$. On choisit $L$ une extension finie de $k$ dans laquelle $Q$ est scindé. On considère ensuite $R$ comme élément de $L[X]$, et on choisit une extension finie $K/L$ telle que $R$ soit scindé sur $K$. Les deux polynômes $Q$ et $R$ sont scindés sur $K$, leur produit $P$ l'est donc aussi.

Ainsi, on peut supposer que $P$ est irréductible. On peut considérer un corps de rupture $L$ de $P$. Il existe $a \in L$ tel que $P(a) = 0$. On peut factoriser $P \in L[X]$ sous la forme $P = (X - a)Q$, avec $Q \in L[X]$. Par hypothèse de récurrence, on peut choisir une extension $K$ de $L$ scindant le polynôme $Q$. Le polynôme $P$ est scindé sur $K$.

***Définition 4.26***. — Soit $k$ un corps. Soit $P \in k[X]$, $P$ non constant. Un corps de décomposition de $P$ est une extension $K/k$ telle que $P$ soit scindé sur $K$, et que si on note $x_1, \ldots, x_n$ les racines de $P$ dans $K$, alors $K = k(x_1, \ldots, x_n)$.

***Proposition 4.27***. — *Soit $k$ un corps. Soit $P \in k[X]$, $P$ non constant. Si $K/k$ et $L/k$ sont deux corps de décomposition de $P$, alors $K$ et $L$ sont isomorphes.*

On peut supposer que $K$ et $L$ sont contenus dans une même extension finie $M$ de $k$. Les racines de $P$ dans $M$ sont à la fois dans $K$ et dans $L$. Or, $K$ et $L$ sont précisément les sous-$k$-algèbres engendrées par ces racines, donc $K = L$.

### 4.2.5. Polynômes séparables. —

***Définition 4.28***. — Soit $P \in k[X]$. Si $P = \sum_i a_i X^i$, on pose $P' = \sum_{i \neq 0} i a_i X^{i-1}$.

***Proposition 4.29***. — *Soit $k$ un corps. Soit $P = (X - a_1) \ldots (X - a_n)$ avec $a_i \in k$. Les conditions suivantes sont équivalentes :*

(i) $P \wedge P' = 1$ ;
(i bis) *les polynômes $P$ et $P'$ n'ont pas de racine commune ;*
(ii) *les éléments $a_i$ sont distincts.*

Comme $P$ est scindé, on a bien entendu (i)⇔(i bis).

Observons que $P'(a_i) = \prod_{j \neq i}(a_i - a_j)$. Par conséquent, si les éléments $a_1, \ldots, a_n$ ne sont pas distincts, il existe $i$ tel que $P'(a_i) = 0$. Ceci montre la contraposée de (i bis)⇒(ii).

Inversement, si $P$ et $P'$ ont une racine commune, il existe $i$ tel que $P'(a_i) = 0$, ce qui montre qu'il existe $j \neq i$ tel que $a_i = a_j$, ceci montrant la contraposée de (ii)⇒(i bis).

***Lemme 4.30***. — *Soit $k$ un corps. Soit $(P, Q) \in k[X]^2$. Soit $K/k$ une extension. Calculer le pgcd de $P$ et $Q$ en tant qu'éléments de $k[X]$ ou de $K[X]$ donne le même résultat.*

Appliquer l'algorithme d'Euclide dans $k[X]$ ou dans $K[X]$ donne évidemment le même résultat.

***Définition 4.31***. — Soit $k$ un corps. Soit $P \in k[X]$ un polynôme non nul. On dit que $P$ est séparable si $P \wedge P' = 1$, autrement dit, d'après ce qui précède, que dans un corps de décomposition de $P$, $P$ n'a pas de racine multiple, ou encore que dans un corps de décomposition de $P$, $P$ n'a pas de racine commune avec $P'$.

***Proposition 4.32***. — *Soit $k$ un corps. Soit $P \in k[X]$ un polynôme irréductible. Alors $P$ est séparable si et seulement si $P' \neq 0$.*

$P$ est séparable si et seulement si $P$ et $P'$ sont premiers entre eux, c'est-à-dire, $P$ étant irréductible, que $P$ ne divise pas $P'$. Pour des raisons de degré, cela revient à dire que $P'$ est non nul.

### 4.3. Existence et unicité des corps finis

***Proposition 4.33.*** — *Soit $p$ un nombre premier. Soit $n \geq 1$. On pose $q = p^n$. On note $K/\mathbf{F}_p$ un corps de décomposition de $X^q - X \in \mathbf{F}_p[X]$. Alors, $K$ est un corps fini à $q$ éléments.*

Le polynôme $X^q - X \in \mathbf{F}_p[X]$ est évidemment séparable, donc $K$ contient exactement $q$ racines du polynôme $X^q - X$. On considère le morphisme d'anneaux $\psi\colon K \to K$ défini par $\psi = \varphi^{\circ n}$ où $\varphi$ est le morphisme de Frobenius relatif à $\mathbf{F}_p$, donné par $\varphi(a) = a^p$. On a $\psi(a) = a^q$ pour tout $a \in K$. L'ensemble des éléments de $K$ tels que $a = \psi(a)$ est donc un sous-anneau $B$ de $K$. Comme $K$ est engendré comme $\mathbf{F}_p$-algèbre par les éléments de $B$, on a $K = B$ : tous les éléments de $K$ sont des racines de $X^q - X$.

***Proposition 4.34.*** — *Soit $p$ un nombre premier. On se donne deux entiers naturels non nuls $m$ et $n$. Soit $\mathbf{F}$ un corps fini à $p^m$ éléments. Soit $\mathbf{F}'$ un corps fini à $p^n$ éléments. Il existe un plongement $\mathbf{F} \to \mathbf{F}'$ si et seulement si $m$ divise $n$. En particulier, deux corps finis de même cardinal sont isomorphes.*

S'il existe un plongement $\mathbf{F} \to \mathbf{F}'$, on peut appliquer la formule de multiplicativité des degrés : $[\mathbf{F}' : \mathbf{F}_p] = [\mathbf{F}' : \mathbf{F}] \cdot [\mathbf{F} : \mathbf{F}_p]$.

On suppose que $m$ divise $n$. Posons $q = p^m$ et $d = \frac{n}{m}$. On peut supposer que $\mathbf{F}$ et $\mathbf{F}'$ sont contenus dans un même corps fini $\mathbf{F}''$. Les éléments de $\mathbf{F}$ (resp. $\mathbf{F}'$) sont les racines de $X^q - X$ (resp. $X^{q^d} - X$) dans $\mathbf{F}''$. Par conséquent, il suffit de montrer que si $x \in \mathbf{F}''$ est tel que $x^q = x$, alors $x^{q^d} = x$, ce qui s'obtient aussitôt par récurrence sur $d$.

### 4.4. Polynômes irréductibles sur un corps fini

***Définition 4.35.*** — Soit $\mathbf{F}_q$ un corps fini à $q$ éléments. Pour toute extension finie $K/\mathbf{F}_q$ on note $\varphi_q\colon K \to K$ le morphisme de Frobenius relatif à $\mathbf{F}_q$, il est défini par la formule $\varphi_q(a) = a^q$. Il s'agit d'un $\mathbf{F}_q$-automorphisme de $K$.

***Proposition 4.36.*** — *Soit $\mathbf{F}_q$ un corps fini à $q$ éléments. Soit $P \in \mathbf{F}_q[X]$ un polynôme irréductible unitaire de degré $d$. Soit $E/\mathbf{F}_q$ le corps de rupture de $P$, avec $x \in E$ tel que $P(x) = 0$. Le plus petit entier $k \geq 1$ tel que $\varphi_q^{\circ k}(x) = x$ est $d$, on a la relation :*

$$P = \prod_{i=0}^{d-1} (X - \varphi_q^{\circ i}(x))$$

*Le corps $E$ est un corps de décomposition de $P$. En outre, $P$ est séparable.*

Le corps $E$ a $q^d$ éléments, donc $x^{q^d} = x$, c'est-à-dire $\varphi_q^{\circ d}(x) = x$. Le groupe $\mathbf{Z}$ agit sur $E$ de façon à ce que $n \bigstar y = \varphi_q^{\circ n}(y)$. L'orbite de $x$ est évidemment $x, \varphi_q(x), \ldots, \varphi_q^{\circ k-1}(x)$ où $k$ est le plus petit entier non nul tel que $\varphi_q^{\circ k}(x) = x$. Comme $P(\varphi_q(y)) = \varphi_q(P(y))$, si $y \in E$ est une racine de $P$, alors $\varphi_q(y)$ aussi. Ainsi, les éléments $\varphi_q^{\circ i}(x)$ de $E$ pour tout $0 \leq i \leq k-1$ sont des racines distinctes de $P$. Si on pose $Q = \prod_{i=0}^{k-1}(X - \varphi_q^{\circ i}(x)) \in E[X]$, on obtient que $Q$ divise $P$. Les coefficients de $Q$ sont fixés par $\varphi_q$, donc $Q \in \mathbf{F}_q[X]$. Ce polynôme unitaire $Q$ est un diviseur non trivial du polynôme irréductible unitaire $P$, on a donc $Q = P$ (et $k = d$). On vient de voir que $P$ est scindé sur $E$, et $E$ est engendré par la racine $x$ ; par conséquent, $E$ est un corps de décomposition de $P$. Le fait que $P$ soit séparable est évident d'après sa factorisation dans $E[X]$.

***Proposition 4.37.*** — *Soit $\mathbf{F}_q$ un corps fini à $q$ éléments. Soit $E$ une extension finie de $\mathbf{F}_q$ de degré $d$. Alors, $\mathrm{Aut}(E/\mathbf{F}_q)$ est un groupe cyclique d'ordre $d$ engendré par $\varphi_q$.*

D'après la construction des corps finis, on sait déjà que le sous-groupe cyclique de $\mathrm{Aut}(E/\mathbf{F}_q)$ engendré par $\varphi_q$ est d'ordre $d$. Il s'agit de montrer que tous les $\mathbf{F}_q$-automorphismes de $E$ sont dans ce sous-groupe. On peut choisir un élément primitif $x$ de $E/\mathbf{F}_q$. Notons $P$ le polynôme minimal de $x$ sur $\mathbf{F}_q$. L'ensemble des $\mathbf{F}_q$-morphismes $E \to E$ s'identifie à l'ensemble des racines de $P$ dans

$E$. D'après la proposition 4.36, toutes ces racines sont de la forme $\varphi_q^{\circ i}(x)$ pour $0 \leq i \leq d-1$, ce qui permet de conclure.

***Proposition 4.38***. — *Soit $\mathbf{F}_q$ un corps à $q$ éléments. Soit $n \geq 1$ un entier. Alors, dans $\mathbf{F}_q[X]$, $X^{q^n} - X$ est égal au produit des polynômes irréductibles unitaires de $\mathbf{F}_q[X]$ dont les degrés divisent $n$.*

On considère une extension $E/\mathbf{F}_q$ de degré $n$ et l'action de $\mathbf{Z}/n\mathbf{Z}$ sur $E$ définie de sorte que 1 agisse par $\varphi_q$. Deux éléments de $E$ sont dans la même orbite si et seulement s'ils ont même polynôme minimal. Le polynôme minimal d'un élément $x \in E$ a pour degré le degré du corps $[\mathbf{F}_q(x) : \mathbf{F}_q]$, qui divise donc $n = [E : \mathbf{F}_q]$. Inversement, tout polynôme unitaire irréductible $P$ de degré $d$ divisant $n$ admet un corps de rupture $\mathbf{F}_q[X]/(P)$ qui peut s'identifier à un sous-$\mathbf{F}_q$-corps de $E$ (cf. démonstration de la proposition 4.34), donc $P$ est le polynôme minimal d'un élément de $E$.

Bref, on peut identifier l'ensemble des orbites de $E$ sous $\varphi_q$ à l'ensemble des polynômes irréductibles unitaires de degré divisant $n$. En outre, pour une telle orbite $\mathcal{O}$ associée à $P$, on a $P = \prod_{x \in \mathcal{O}}(X - x)$. Comme $X^{q^n} - X$ s'identifie au produit des $(X - x)$ pour $x \in E$, on en déduit aussitôt le résultat.

***Corollaire 4.39***. — *Soit $\mathbf{F}_q$ un corps à $q$ éléments. Pour tout entier $d \geq 1$, on note $N_d$ le nombre de polynômes unitaires irréductibles de degré $d$ dans $\mathbf{F}_q$. Alors, pour tout $n \geq 1$, on a la relation :*

$$q^n = \sum_{d|n} dN_d \ .$$

***Remarque 4.40***. — Pour tout $d \geq 1$, $N_d \geq 1$. En particulier, il existe un polynôme irréductible unitaire $P$ de degré $d$ sur $\mathbf{F}_p$. Si on sait déterminer un tel $P$, on peut construire le corps fini $\mathbf{F}_{p^d}$ en prenant le corps de rupture de $P$ sur $\mathbf{F}_p$. Si on note $x$ la racine privilégiée de $P$ dans ce corps de rupture, tout élément de $\mathbf{F}_{p^d}$ peut être représenté sous la forme $a_0 + a_1 x + \cdots + a_{d-1}x^{d-1}$ où $(a_0, \ldots, a_{d-1})$ est un $d$-uplet d'entiers modulo $p$. Ceci permet donner une représentation informatique des éléments dans un corps fini.

***Définition 4.41***. — Soit $n \in \mathbf{N} - \{0\}$. On peut écrire $n = \prod_{i=1}^k p_i^{e_i}$ où les $p_i$ sont des nombres premiers distincts et $e_i \geq 1$. S'il existe $i$ tel que $e_i \geq 2$, on pose $\mu(n) = 0$, sinon, $\mu(n) = (-1)^k$ (en particulier, $\mu(1) = 1$). $\mu$ est la fonction de Möbius.

***Proposition 4.42***. — *Avec les notations du corollaire 4.39, on a :*

$$nN_n = \sum_{d|n} \mu(d) q^{\frac{n}{d}} \ .$$

Voici une démonstration possible (ce n'est pas la plus savante). On a vu dans la démonstration de la proposition 4.38 que $nN_n$ était égal au nombre d'éléments de $\mathbf{F}_{q^n}$ qui sont exactement de degré $n$ sur $\mathbf{F}_q$, c'est-à-dire ne sont dans aucun sous-corps $\mathbf{F}_{q^d}$ de $\mathbf{F}_{q^n}$ pour $d$ divisant $n$ et $d \neq n$. On peut donc écrire

$$nN_n = q^n - \#\Big(\bigcup_{d|n, d \neq n} \mathbf{F}_{q^d}\Big)$$

Écrivons $n = \prod_{i=1}^k p_i^{e_i}$ avec les $p_i$ premiers distincts et $e_i \geq 1$. Dans la réunion ci-dessus, il suffit de considérer les $d$ de la forme $\frac{n}{p_i}$. Notons $A_i = \mathbf{F}_{q^{\frac{n}{p_i}}}$. On a la formule

$$\#\bigcup_{i=1}^k A_i = \sum_{l=1}^k (-1)^{l-1} \sum_{i_1 < \cdots < i_l} \#(A_{i_1} \cap \cdots \cap A_{i_l}) \ .$$

***Lemme 4.43***. — *Si $d$ et $e$ deux diviseurs de $n$, alors l'intersection de $\mathbf{F}_{q^d}$ et de $\mathbf{F}_{q^e}$ dans $\mathbf{F}_{q^n}$ est $\mathbf{F}_{q^m}$ où $m$ est le pgcd de $d$ et $e$.*

On sait que $\mathbf{F}_{q^m}$ est contenu dans $\mathbf{F}_{q^d}$ et dans $\mathbf{F}_{q^m}$, il est donc contenu dans leur intersection. Inversement, le degré sur $\mathbf{F}_q$ de l'intersection des deux corps doit diviser $d$ et $e$, donc diviser $m$. L'inclusion précédente est donc une égalité.

En appliquant ce lemme, on obtient, avec les notations ci-dessus :

$$\#(A_{i_1} \cap \cdots \cap A_{i_l}) = q^{\frac{n}{p_{i_1} \cdots p_{i_l}}} .$$

L'ensemble d'indices $l, i_1 < \cdots < i_l$ peut être identifié à l'ensemble des diviseurs $d$ de $p_1 \dots p_k$ autres que 1 par les correspondances $d = p_{i_1} \dots p_{i_l}$ et $(-1)^l = \mu(d)$. On obtient ainsi :

$$nN_n = q^n + \sum_{1 \neq d | p_1 \dots p_k} \mu(d) q^{\frac{n}{d}}$$

Ceci se réécrit aussitôt

$$nN_n = \sum_{d | p_1 \dots p_k} \mu(d) q^{\frac{n}{d}}.$$

Si $d$ est un diviseur de $n$ ne divisant pas $p_1 \dots p_k$, on a $\mu(d) = 0$, ce qui donne la formule voulue.

# CHAPITRE 5

# CODES CORRECTEURS D'ERREURS

*Ce chapitre a été écrit par Christophe Ritzenthaler.*

Les codes correcteurs ont été introduits pour corriger les erreurs de transmission ou de lecture de données numériques, ou les erreurs survenant au cours de leur inscription sur un support physique (bande, CD) ou encore lorsque les données subissent une altération sur le support de stockage. Voici quelques domaines où ils sont appliqués :

- transmissions spatiales ;
- disque compact et DVD ;
- communications par internet.

La première mission d'un code est de détecter des erreurs. Par exemple le numéro de sécurité sociale et le numéro de compte bancaire comportent chacun une clé à deux chiffres. Cette clé sert de contrôle : à partir du numéro et de la clé, on forme un nombre qui doit être divisible par 97. Si ce n'est pas le cas, c'est que le numéro n'est pas valide : une erreur s'est glissée quand on a saisi ou recopié le numéro (ou sa clé). Le même principe s'applique aux codes barres et au code ISBN qui identifie les publications.

Les codes plus performants permettent aussi de corriger des erreurs, en supposant qu'elles ne sont pas trop nombreuses. Le plus simple est le code de répétition : on envoie trois fois de suite le même message. Si après transmission les trois messages ne sont plus identiques, alors il y a eu au moins une erreur. Si cependant deux des trois sont identiques, il est probable qu'il s'agisse du message d'origine : on peut alors corriger les erreurs de transmission.

## 5.1. Codes et distance de Hamming

Les messages transmis sont supposés découpés en blocs (ou mots) de longueur $n$ écrits avec l'alphabet $\{0,1\}$. Un *code* (binaire) est un sous-ensemble $C$ de l'ensemble $\{0,1\}^n$ de tous les mots possibles. On dit que $n$ est la longueur de $C$.
La *distance de Hamming* entre deux mots $x = (x_1, ..., x_n)$ et $y = (y_1, ..., y_n)$, que l'on notera $d(x,y)$, est le nombre d'indices $i$ tels que $x_i \neq y_i$. C'est bien une distance sur $\{0,1\}^n$. La *distance minimum* du code $C$ est le minimum des $d(x,y)$ pour $x$ et $y$ des mots différents de $C$ (on suppose que $C$ a au moins 2 mots !). On la notera toujours $d$.

***Exemple 1***. — *Considère* $C = \{c_0, c_1, c_2, c_3\}$ *avec*

$$c_0 = (00000), \quad c_1 = (10110), \quad c_2 = (01011), \quad c_3 = (11101).$$

*C'est un code de longueur* 5 *et de distance* $d = 3$.

Le mot $c \in C$ est émis et, après d'éventuelles erreurs de transmission, le mot $r \in \{0,1\}^n$ est reçu. On décode le mot $r$ selon le principe du maximum de vraisemblance, c.-à-d. qu'on le décode comme un mot de $C$ à distance minimum de $r$. On dit que $C$ est $t$-correcteur (ou corrige $t$ erreurs) quand toute erreur portant sur au plus $t$ bits est corrigée correctement. On voit donc que le code $C$

est $t$-correcteur si et seulement si les boules fermées (dans $\{0,1\}^n$ muni de la distance de Hamming) de centres les éléments de $C$ et de rayon $t$ sont disjointes, ou encore si et seulement si la distance minimum $d$ de $C$ vérifie $d \geq 2t+1$.

Il est souvent difficile de calculer la distance minimale et plus encore de décoder un mot sans structure additionnelle ; c'est pourquoi on préfère travailler avec des codes linéaires (voir le paragraphe 5.2).

***Remarque 5.1***. — Supposons que $\#C = 2^k$. Le rapport $k/n$ s'appelle le taux de transmission du code. Soit

$$C(p) = 1 + p\log_2 p + (1-p)\log_2(1-p).$$

Un théorème de Shannon montre que pour une probabilité d'erreur $p < 0.5$ à chaque bit, il existe un code dont le taux de transfert est inférieur à $C(p)$, mais arbitrairement proche de $C(p)$, qui peut corriger tout message avec une probabilité arbitrairement proche de 1. Par exemple si $p = 0.001$, on a $C(p) = 0.986$ donc en ajoutant à peine 15 bits pour $k = 1000$ bits transmis on peut arriver à une probabilité de correction arbitrairement proche de 1. Cependant le théorème de Shannon ne dit pas comment construire de tels codes.

## 5.2. Codes linéaires

**5.2.1. Définitions.** — On note $\mathbb{F}_2$ le corps à deux éléments 0 et 1. Les mots de longueur $n$ sont les éléments de $\mathbb{F}_2^n$, que l'on écrira comme des vecteurs lignes. Un code linéaire de longueur $n$ est un sous-espace vectoriel $C \subset \mathbb{F}_2^n$. La lettre $k$ désignera toujours la dimension de $C$ (comme espace vectoriel sur $\mathbb{F}_2$). Le nombre de mots du code $C$ est $2^k$. Le *poids* d'un mot $x = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$, noté $w(x)$, est le nombre d'indices $i$ tels que $x_i \neq 0$. Comme $d(x,y) = w(x-y)$, la distance minimum $d$ d'un code linéaire $C$ est le minimum des poids $w(x)$ pour $x \in C$ non nul. (On suppose que $C$ n'est pas le code nul.) On regroupe les trois paramètres $n, k$ et $d$ d'un code linéaire $C$ en disant que $C$ est de type $(n,k,d)$.

L'exemple le plus simple est probablement celui où on répète 3 fois l'unique bit qu'on envoie. Le code $C$ est formé des deux triplets $(0,0,0)$ et $(1,1,1)$. Il est de type $(3,1,3)$ ; il permet de détecter deux erreurs et d'en corriger une.

Un autre exemple consiste à prendre pour $C$ l'hyperplan défini par $x_1 + \ldots + x_n = 0$. Cela signifie qu'étant donné un message $(x_0, \ldots, x_{n-1})$, on lui ajoute un bit de contrôle $x_n$ (défini par la condition $x_1 + \ldots + x_n = 0$). L'intérêt de ce procédé est que coder et décoder sont des opérations très simples. Ce code est de type $(n, n-1, 2)$ ; il permet seulement de détecter une erreur.

La borne suivante, due à R. Singleton en 1964, quantifie le fait qu'on ne peut pas avoir à la fois le beurre (une capacité de correction importante, i.e. $d$ grand) et l'argent du beurre (un nombre de mots de code important, i.e. $k$ grand), pour une longueur $n$ fixée.

***Lemme 5.2*** ([**4**, Prop.9.1],[**8**, 18.4]). — *On a toujours $d + k \leq n + 1$ (borne de Singleton).*

*Démonstration.* — Notons $p : \mathbb{F}_2^n \to \mathbb{F}_2^{n-d+1}$ la projection sur les $n-(d-1)$ premières coordonnées. Alors pour tout $x \in \ker p$ on a $w(x) \leq d-1$, donc $C \cap \ker p = \{0\}$. Le code $p(C)$ est donc de dimension $k$ et de longueur $n-d+1$, d'où $k \leq n-d+1$. □

***Remarque 5.3***. — Il existe d'autres bornes, par exemple celle de Hamming. Si $B$ est la boule de centre 0 et de rayon $r$ dans $\mathbb{F}_2^n$, elle contient $R = \sum_{k=0}^{r} \binom{n}{k}$ mots. Ainsi si $C$ est un code de longueur $n$ et de distance minimale $d = 2r+1$ alors $C$ contient au plus $2^n/R$ mots.

***Lemme 5.4***. — *Si $C$ est un code linéaire de type $(n,k,d)$, on définit le code étendu $\tilde{C}$ comme le code formé des mots $(x_1, \ldots, x_{n+1}) \in \mathbb{F}_2^{n+1}$ tels que $(x_1, \ldots, x_n) \in C$ et $\sum_{i=1}^{n+1} x_i = 0$. Le type de $\tilde{C}$ est $(n+1, k, d+1)$ si $d$ est impair et $(n+1,k,d)$ si $d$ est pair.*

*Démonstration.* — Le seul paramètre à déterminer est la distance minimale. Soit $x$ un mot de poids $d$ de $C$. Si $d$ est pair alors $y = (x,0) \in \tilde{C}$ et $w(y) = d$. Si $d$ est impair alors $y = (x,1) \in \tilde{C}$ et $w(y) = d+1$. □

**5.2.2. Matrice génératrice.** — On peut déterminer un sous-espace vectoriel (et donc un code) par une base. Soit $C$ un code linéaire. Une *matrice génératrice* de $C$ est une matrice dont les lignes forment une base de $C$. Une matrice génératrice $G$ est donc de taille $k \times n$ et de rang $k$. Si $m$ est un vecteur ligne de $\mathbb{F}_2^k$, le produit $mG$ est un mot du code $C$ et l'application $m \mapsto mG$ est un isomorphisme de $\mathbb{F}_2^k$ sur $C$ (que l'on peut voir comme une opération de codage). Si la matrice $G$ est de la forme $(I, P)$ par blocs, où $I$ est la matrice identité de taille $k$ et $P \in M_{k,n-k}(\mathbb{F}_2)$, on dit que le codage est *systématique*. Les $k$ premiers bits d'un mot de code portent alors l'information (on y recopie le vecteur de $\mathbb{F}_2^k$), les $n-k$ suivants sont de la redondance.

***Exemple 2.*** — *La matrice*

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$

*est une matrice génératrice pour le code*

$$(00000), (10000), (11010), (11101), (01010), (01101), (00111), (10111).$$

On dit que deux codes linéaires de même longueur sont *équivalents* si l'un s'obtient à partir de l'autre par une permutation des coordonnées. On peut vérifier que deux codes équivalents ont même type. Etant donné une matrice génératrice $G$ d'un code $C$, on peut toujours mettre $G$ sous forme ligne-échelle réduite par des opérations élémentaires sur ses lignes, ce qui correspond à changer de base de $C$. En permutant les colonnes on obtient alors une matrice de la forme $(I, P)$. Cela montre que tout code est équivalent à un code qui possède un codage systématique.

***Exemple 3.*** — *Soit $G$ la matrice génératrice*

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

*On additionne la ligne* 1 *aux lignes* 2 *et* 3 *(ce qui est ici plus rapide que de suivre l'algorithme du pivot de Gauss). On obtient*

$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

*Il ne reste plus qu'à permuter la première et la troisième colonnes, ce qu'on aurait pu faire avant d'agir sur les lignes.*

**5.2.3. Matrice de contrôle.** — On peut ausi se donner un sous-espace vectoriel par un système d'équations indépendantes. Soit $C$ un code linéaire. Une *matrice de contrôle* de $C$ est la matrice d'un système d'équations linéaires homogènes indépendantes dont l'espace des solutions est $C$. Autrement dit, une matrice de contrôle $H$ est de taille $(n-k) \times n$ et de rang $n-k$, et $C = \{x \in \mathbb{F}_2^n, H^t x = 0\}$ ; rappelons que $x \in \mathbb{F}_2^n$ est vu dans tout ce chapitre comme une matrice ligne. Les mots du code correspondent ainsi aux relations de dépendance linéaire entre les colonnes de $H$. Si $C$ est donné sous forme systématique par la matrice génératrice $G = (I_k, P)$, alors on peut prendre comme matrice de contrôle $H = (-{}^tP, I_{n-k})$ (le signe $-$ est superflu en caractéristique 2). Supposons que $c \in C$ est le mot du code envoyé et $r \in \mathbb{F}_2^n$ le mot reçu. La différence $e = r - c$ est le vecteur d'erreur. Son poids $w(e)$ est le nombre de bits erronés dans le mot reçu. Soit $H$ une matrice de contrôle de C. Le *syndrome* du mot reçu $r$ est le vecteur $s \in \mathbb{F}_2^{n-k}$ défini par ${}^ts = H^t r = H^t e$. Le syndrome est nul si et seulement si $r \in C$. Le syndrome définit un isomorphisme du quotient $\mathbb{F}_2^n / C$ sur $\mathbb{F}_2^{n-k}$. Si le syndrome est non nul, on corrige le mot reçu $r$ en appliquant le principe du maximum de vraisemblance : on soustrait à $r$ un mot de poids minimum dans sa classe modulo $C$, c.-à-d. un mot de poids minimum parmi ceux ayant même syndrome que $r$. Dans le cas où $w(e)$ est strictement inférieur à $d/2$, alors $e$ est l'unique mot de poids minimum dans la classe de $r$ modulo $C$ et on récupère bien le mot de code émis.

***Remarque 5.5***. — La matrice de contrôle peut être vue comme la matrice génératrice du code dual

$$C^{\perp} = \{y \in \mathbb{F}_2^n, \forall c \in C, y.c = 0\}$$

où . est le produit scalaire canonique. *Attention, le corps de base est ici* $\mathbb{F}_2$ *et non* $\mathbf{R}$*, donc il n'y a pas de notion de forme bilinéaire positive. On appelle abusivement produit scalaire toute forme bilinéaire symétrique non dégénérée.*

***Proposition 5.6***. — *Soit* $H$ *une matrice de contrôle du code* $C$*. La distance minimum* $d$ *de* $C$ *est caractérisée par les propriétés suivantes :*

- $d-1$ *colonnes de* $H$ *sont toujours linéairement indépendantes.*
- *Il y a un système de* $d$ *colonnes de* $H$ *qui est lié.*

*Démonstration.* — Supposons que $d-1$ colonnes de $H$ sont toujours linéairement indépendantes. Soit $c = (c_1, \ldots, c_n)$ un mot de code. On a $H^t c = 0$. Si $w(c) < d$ alors on a relation entre moins de $d$ colonnes de $H$. Inversement un mot $c$ de poids $d$ donne une relation entre $d$ colonnes de $H$. □

***Exemple 4***. — *Donnons un exemple de décodage par syndrome. Soit* $C$ *le code donné par la matrice de contrôle*

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

*On calcule tout d'abord des représentants de poids minimal pour chacune des classes de* $\mathbb{F}_2^6/C$ *ainsi que le syndrome associé. On obtient :*

| *Représentant* | *Syndrome* |
|---|---|
| 000000 | 000 |
| 100000 | 110 |
| 010000 | 101 |
| 001000 | 011 |
| 000100 | 100 |
| 000010 | 010 |
| 000001 | 001 |
| 100001 | 111 |

*Supposons que* $u = 100011$ *est reçu. Son syndrome est* $H^t u = 101$*. Pour décoder* $u$ *il faut donc lui soustraire* 010000*.*

## 5.3. Quelques codes linéaires

**5.3.1. Codes cycliques.** — Un code linéaire $C \subset \mathbb{F}_2^n$ est dit *cyclique* quand il est stable par l'automorphisme de décalage cyclique

$$\begin{array}{rcl} T : \mathbb{F}_2^n & \to & \mathbb{F}_2^n \\ (x_1, \ldots, x_n) & \mapsto & (x_2, \ldots, x_n, x_1) \end{array}$$

On identifie $\mathbb{F}_2^n$ à l'algèbre $\mathbb{F}_2[X]/(X^n - 1)$ par

$$(x_1, \ldots, x_n) \mapsto x_1 X^{n-1} + \cdots + x_{n-1} X + x_n.$$

On désigne ici par la même lettre l'indéterminée $X$ et son image dans le quotient. L'endomorphisme $T$, modulo cette identification, est l'endomorphisme de multiplication par $X$. Par définition, un code cyclique est donc un sous-espace vectoriel stable par multiplication par $X$. De manière équivalente, c'est un sous-espace vectoriel stable par multiplication par n'importe quel polynôme en $X$. Donc, un code linéaire $C$ de longueur $n$ est cyclique si et seulement si $C$ est un idéal de $\mathbb{F}_2[X]/(X^n - 1)$.

L'homorphisme de passage au quotient induit une bijection entre l'ensemble des idéaux de l'anneau $\mathbb{F}_2[X]/(X^n - 1)$ et l'ensemble des idéaux de $\mathbb{F}_2[X]$ qui contiennent $(X^n - 1)$. Puisque $\mathbb{F}_2[X]$ est principal, ce sont exactement les idéaux engendrés par les diviseurs (que l'on prend unitaires

pour assurer l'unicité) de $X^n-1$ dans $\mathbb{F}_2[X]$. Le diviseur unitaire $g$ de $X^n-1$ ainsi associé à un code cyclique $C$ s'appelle le *polynôme générateur* de $C$. Si $g \neq X^n-1$ (dans le cas contraire $C$ est nul), le code $C$ est engendré (comme espace vectoriel sur $\mathbb{F}_2$) par $g, Xg, ..., X^{n-1-\deg(g)}g$. La dimension de $C$ est dans tous les cas $k = n - \deg(g)$.
Le procédé de codage systématique $\mathbb{F}_2^k \to C$ d'un code cyclique de polynôme générateur $g$ est donné par la division euclidienne par $g$ : le vecteur $(x_1, \ldots, x_k) \in \mathbb{F}_2^k$ est codé par le polynôme $c = c_I - c_R$, où $c_I = x_1X^{n-1} + \ldots + x_kX^{n-k}$, et $c_R$ (de degré $< n-k$) est le reste de la division euclidienne de $c_I$ par $g$. (Le polynôme $c_I$ porte l'information, et $c_R$ la redondance).
On suppose à partir de maintenant que $n$ est premier avec la caractéristique de $\mathbb{F}_2$, c.-à-d. impair. Cette hypothèse entraîne que le polynôme $X^n-1$ est séparable : il a $n$ racines distinctes dans tout corps de décomposition sur $\mathbb{F}_2$. Notons $K$ un corps de décomposition de $X^n-1$ sur $\mathbb{F}_2$. Comme $K^*$ est cyclique, son sous-groupe formé par les $n$ racines de $X^n-1$ l'est aussi. On fixe un générateur $\alpha$ de ce sous-groupe ; on dit que c'est une racine primitive $n$-ième de l'unité dans $K$. La proposition 4.36 montre que le polynôme minimal $P$ de $\alpha$ sur $\mathbb{F}_2$ a pour degré l'ordre $r$ de 2 dans le groupe multiplicatif $(\mathbf{Z}/n\mathbf{Z})^*$, et que ses racines sont $\alpha, \alpha^2, \alpha^4, \ldots, \alpha^{2^{r-1}}$. On a $K = \mathbb{F}_2[\alpha] = \mathbb{F}_2[X]/P = \mathbb{F}_{2^r}$ . Le polynôme générateur $g$ est déterminé par ses racines dans $K$, qui forment un sous-ensemble de l'ensemble $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ des racines $n$-ièmes de l'unité. Ces racines sont importantes car un mot $c = (x_1, \ldots, x_n) \in \mathbb{F}_2^n$ est dans le code si, et seulement si, le polynôme associé $x_1X^{n-1} + \ldots + x_n$ s'annule en toutes les racines de $g$. Soit $\Sigma$ un sous-ensemble de $\mathbf{Z}/n\mathbf{Z}$. Le polynôme $g_\Sigma = \prod_{i\in\Sigma}(X - \alpha^i)$ est un diviseur de $X^n-1$ à coefficients dans $\mathbb{F}_2$ si et seulement si $\Sigma$ est stable par multiplication par 2 (se souvenir que $\mathbb{F}_2$ est l'ensemble des éléments de $K$ laissés fixes par l'élévation au carré). En conclusion, on a une bijection entre les codes cycliques de longueur $n$ et les sous-ensembles de $\mathbf{Z}/n\mathbf{Z}$ stables par multiplication par 2. La configuration des racines du polynôme générateur nous renseigne sur la distance minimale du code cyclique.

***Proposition 5.7*** **([4**, Prop.9.4**])**. — *Si $\Sigma$ contient $s$ entiers consécutifs $a+1, a+2, \ldots, a+s$ modulo $n$, alors le code cyclique de polynôme générateur $g_\Sigma$ est nul ou a une distance minimum supérieure ou égale à $s+1$.*

La démonstration est une application des propriétés du déterminant de Vandermonde.

On peut démontrer que le polynôme cyclotomique $\phi_n \in \mathbf{Z}[X]$ factorise sur $\mathbb{F}_2$ en produit de facteurs irréductibles de même degré $r$, et que le polynôme $P$ ci-dessus est un de ces facteurs. Par exemple on a $\phi_{15}(X) = X^8-X^7+X^5-X^4+X^3-X+1$, qui se factorise en $(X^4+X+1)(X^4+X^3+1)$ sur $\mathbb{F}_2$, avec $r = 4$.

**5.3.2. Codes BCH.** — Les codes BCH (Bose-Chaudhuri-Hocquenghem) sont des codes cycliques particuliers. La famille des codes BCH contient les codes de Reed-Solomon qui servent pour la lecture des CD (voir [**4**, p.238]). Nous ne considérerons ici que des codes BCH binaires primitifs. Leur longueur $n$ est de la forme $n = 2^r - 1$. Alors $r$ est l'ordre de 2 dans le groupe multiplicatif $(\mathbf{Z}/n\mathbf{Z})^*$, et le corps $K$ engendré par les racines $n$-ièmes de l'unité sur $\mathbb{F}_2$ est $\mathbb{F}_{2^r}$. Tous les calculs de décodage vont se faire sur ce corps $K$. Choisissons une racine primitive $n$-ième de l'unité $\alpha$ dans $K$. Concrètement, on se donne $\alpha$ par son polynôme minimal $P$ sur $\mathbb{F}_2$. Tout élément du groupe multiplicatif $K^*$ s'écrit de manière unique sous la forme $\alpha^i$ avec $0 \leq i < n$, et il s'écrit aussi de manière unique comme combinaison linéaire à coefficients dans $\mathbb{F}_2$ de $1, \alpha, \ldots, \alpha^{r-1}$. On peut voir la table de correspondance entre ces deux représentations pour $K = \mathbb{F}_{16}$, avec $\alpha$ vérifiant $\alpha^4 + \alpha + 1 = 0$ dans [**4**, p.213].
On appelle *code BCH* de longueur $n = 2r-1$ et de distance prescrite $\delta$ (avec $\delta$ entier tel que $0 < \delta \leq n$) le code cyclique de polynôme générateur $g_\Sigma$, où $\Sigma$ est le plus petit sous-ensemble de $\mathbf{Z}/n\mathbf{Z}$ contenant $1, 2, \ldots, \delta-1$ et stable par multiplication par 2. Autrement dit, un polynôme $c = x_1X^{n-1} + \cdots + x_n \in \mathbb{F}_2[X]$ appartient à ce code si et seulement si

$$c(\alpha) = c(\alpha^2) = ... = c(\alpha^{\delta-1}) = 0.$$

On peut trouver des exemples de codes BCH explicites pour $n = 15$ dans [**4**, p.240].

# CHAPITRE 6

# FACTORISATION DANS $\mathbf{F}_q[X]$

L'objet de ce chapitre est d'expliquer comment on peut procéder pour factoriser un polynôme $P \in \mathbf{F}_q[X]$ sous la forme d'un produit de polynômes irréductibles.

Une question voisine consiste, étant donné un polynôme non constant, à pouvoir affirmer soit que $P$ est irréductible, soit que $P$ ne l'est pas et d'exhiber un facteur non trivial. Il suffit de répondre à cette question pour répondre à la première : si on peut exhiber des facteurs non triviaux, on peut applique récursivement l'algorithme de factorisation à des polynômes non constants $Q$ et $R$ tels que $P = QR$ qui auront été déterminés.

## 6.1. Réduction au cas des polynômes sans facteur carré

### 6.1.1. Polynômes sans facteurs carrés, polynômes séparables. —

***Définition 6.1***. — Soit $k$ un corps. Soit $P \in k[X]$, $P \neq 0$. On dit que $P$ est sans facteur carré s'il n'existe pas $Q \in k[X]$ non constant tel que $Q^2$ divise $P$. Autrement dit, aucun polynôme irréductible n'apparaît avec multiplicité au moins 2 dans la décomposition de $P$.

***Définition 6.2***. — Soit $k$ un corps. On dit que $k$ est parfait si pour tout polynôme irréductible $P \in k[X]$, alors $P$ est séparable (c'est-à-dire que $P' \neq 0$).

***Proposition 6.3***. — *Soit $k$ un corps parfait. Soit $P \in k[X]$, $P \neq 0$. Alors, $P$ est séparable si et seulement si $P$ est sans facteurs carrés.*

Si $P$ admet un facteur carré, il existe $Q$ non constant et $R$ tels que $P = Q^2R$, d'où $P' = Q(2Q'R + QR')$, ainsi $Q$ divise le pgcd de $P$ et de $P'$, $P$ n'est pas séparable.

Inversement, supposons que $P$ n'ait pas de facteur carré. On peut écrire $P = \prod_{i=1}^{k} P_i$ où les $P_i$ sont des polynômes irréductibles étrangers deux-à-deux. Montrons que $P_i$ ne divise pas $P'$ (de sorte que $P$ et $P'$ seront premiers entre eux, c'est-à-dire que $P$ sera séparable). Il vient aussitôt que $P'$ est congru à $P_i' \prod_{j \neq i} P_j$ modulo $P_i$. Pour conclure, il suffit de montrer que $P_i$ ne divise pas $P_i' \prod_{j \neq i} P_j$, ce qui est assuré par le fait que $P_i'$ soit premier à $P_i$.

***Théorème 6.4***. — *Soit $k$ un corps. Alors, $k$ est un corps parfait si et seulement si : soit $k$ est de caractéristique 0, soit $k$ est de caractéristique $p > 0$ et le morphisme de Frobenius $x \longmapsto x^p$ définit un automorphisme de $k$.*

Il est évident que dans un corps de caractéristique zéro, le polynôme dérivé d'un polynôme irréductible est non nul. Supposons que $k$ soit de caractéristique $p > 0$. Supposons que le morphisme de Frobenius soit surjectif, montrons que $k$ est parfait. Soit $P \in k[X]$ un polynôme tel que $P' = 0$. La condition $P' = 0$ signifie que $P$ est en fait un polynôme en $X^p$. Les coefficients de $P$ admettent des racines $p$-ièmes dans $k$, on peut donc trouver des éléments $a_i \in k$ tels que $P = \sum_i a_i^p X^p$. Posons $Q = \sum_i a_i X$. On a $P = Q^p$, donc $P$ n'est pas irréductible. Ainsi, tout polynôme irréductible dans $k[X]$ est séparable. Inversement, supposons qu'on ait $a \in k$ tel que $X^p - a$ n'ait pas de racine dans $k$ et montrons que $k$ n'est pas parfait. Considérons une extension finie $L$ de $k$ scindant $X^p - a$.

Notons $b$ une racine $p$-ième de $a$ dans $L$. Comme le morphisme de Frobenius est un morphisme de corps (donc injectif), l'élément $b$ est l'unique racine de $X^p - a$ dans $L$, d'où $X^p - a = (X-b)^p$. Montrons que $P = X^p - a \in k[X]$ est irréductible. Si $Q \in L[X]$ est un polynôme unitaire divisant $P$ (et qui n'est ni 1 ni $P$), alors, en raisonnant dans $L[X]$, on obtient qu'il existe $1 \leq i \leq p-1$ tel que $Q = (X-b)^i$. Ce polynôme $Q$ n'appartient pas à $k[X]$. En effet, le coefficient de $X^{i-1}$ dans $Q$ est $-ib$ (opposé de la somme des racines), qui n'appartient pas à $k$. Ainsi, $P$ est irréductible, mais $P' = 0$, ce qui montre que $k$ n'est pas parfait.

***Corollaire 6.5***. — *Tout corps fini $k$ est parfait, et un polynôme $P \in k[X]$ est séparable si et seulement s'il est sans facteur carré.*

### 6.1.2. Algorithme. —

***Proposition 6.6***. — *Soit $k$ un corps parfait. Pour factoriser des polynômes $P$ dans $k[X]$, on peut supposer que $P$ est sans facteur carré.*

***Lemme 6.7***. — *Soit $k$ un corps parfait de caractéristique $p$. Soit $P \in k[X]$ tel que $P' = 0$. Alors il existe un polynôme $Q \in k[X]$ tel que $P = Q^p$.*

On peut écrire $P = \sum_i a_i X^{ip}$ puisque les coefficients de $X^j$ pour $p$ ne divisant pas $j$ sont nuls. Comme $k$ est parfait, on peut extraire une racine $p$-ième $b_i$ de $a_i$ dans $k$ et poser $Q = \sum_i b_i X^i$. On a bien $P = Q^p$.

***Remarque 6.8***. — Si $k = \mathbf{F}_p$, pour tout $a \in k$, on a $a^p = a$, il est donc trivial d'extraire la racine $p$-ième d'un élément. Si $k = \mathbf{F}_q$ avec $q = p^d$, alors $a^{p^{d-1}}$ est une racine $p$-ième de $a$.

Supposons qu'on dispose d'un algorithme de factorisation dans le cas de polynômes sans facteur carré et donnons un algorithme permettant de déterminer une factorisation dans le cas général :

***Algorithme pour se ramener aux polynômes sans facteur carré 6.9***
*Soit $k$ un corps parfait. Soit $P \in k[X]$, $P$ non constant. On calcule $P'$. Si $P' = 0$, le lemme 6.7 et la remarque subséquente fournissent un polynôme $Q \in k[X]$ tel que $P = Q^p$, on peut renvoyer factoriser $Q$ récursivement pour obtenir la factorisation de $P$. Sinon, on calcule $Q = P \wedge P'$. Si $Q$ est constant (et non nul puisque $P' \neq 0$), $P$ est sans facteur carré, on peut lui appliquer l'algorithme de factorisation des polynômes sans facteur carré. Sinon, $Q$ est un diviseur non trivial de $P$, on peut factoriser (récursivement) $Q$ et $P/Q$ pour obtenir une factorisation de $P$.*

## 6.2. Algorithme de Berlekamp

On fixe une fois pour toutes un corps fini $\mathbf{F}_q$ et un polynôme unitaire séparable $P \in \mathbf{F}_q[X]$ de degré $d \geq 1$. D'après la proposition 6.3, il existe un entier $e \geq 1$ et des polynômes unitaires irréductibles $P_1, \ldots, P_e \in \mathbf{F}_q[X]$ distincts tels que $P = P_1 P_2 \ldots P_e$.

### 6.2.1. Détermination du nombre de facteurs irréductibles. —

***Définition 6.10***. — On note $A$ la $\mathbf{F}_q$-algèbre $A = \mathbf{F}_q[X]/(P)$.

D'après le théorème chinois, on a :

***Proposition 6.11***. — *On a un isomorphisme canonique de $\mathbf{F}_q$-algèbres :*

$$A = \mathbf{F}_q[X]/(P) \simeq K_1 \times K_2 \times \cdots \times K_e \ ,$$

*où pour tout $i$, $K_i$ est le corps de rupture de $P_i$ sur $\mathbf{F}_q$.*

***Définition 6.12***. — On note $\varphi_q \colon A \to A$ l'application $a \longmapsto a^q$. Il s'agit d'un endomorphisme $\mathbf{F}_q$-linéaire de l'anneau $A$. On note $A'$ le sous-anneau de $A$ formé des éléments $a \in A$ tels que $\varphi_q(a) = a$.

***Proposition 6.13***. — *Le sous-anneau $A'$ de $A$ est un $\mathbf{F}_q$-espace vectoriel de dimension $e$.*

On dispose d'un isomorphisme $A \simeq K_1 \times \cdots \times K_e$ de $\mathbf{F}_q$-algèbres. Soit $a = (a_1, \ldots, a_e)$ un élément de $K_1 \times \cdots \times K_e$. On a $a^q = (a_1^q, \ldots, a_e^q)$. Dire que $a^q = a$ revient à demander que les éléments $a_i$ appartienne au sous-corps $\mathbf{F}_q$ du corps $K_i$ pour tout $i$. L'ensemble $A'$ s'identifie donc à $\mathbf{F}_q^e$ ; sa dimension comme $\mathbf{F}_q$-espace vectoriel est $e$.

***Définition 6.14***. — On note $x$ la classe de $X$ dans le quotient $A = \mathbf{F}_q[X]/(P)$. La famille $(1, x, x^2, \ldots, x^{d-1})$ est une base du $\mathbf{F}_q$-espace vectoriel $A$, on note $\Phi_q$ la matrice du $\mathbf{F}_q$-endomorphisme $\varphi_q \colon A \to A$ dans la base $(1, x, x^2, \ldots, x^{d-1})$.

La proposition à retenir constitue le principe de l'algorithme de factorisation de Berlekamp :

***Algorithme de Berlekamp pour déterminer le nombre de facteurs irréductibles 6.15***
*Soit $P$ un polynôme séparable de degré $d \geq 1$ sur $\mathbf{F}_q$. On considère la matrice $\Phi_q$ à coefficients dans $\mathbf{F}_q$ dont la $i$-ème colonne (pour $0 \leq i \leq d-1$) est constituée des coefficients du polynôme $X^{iq} \mod P$. Le nombre de facteurs irréductibles $e$ de $P$ est la dimension du noyau de la matrice $\Phi_q - I_d$ (noyau dont on peut déterminer efficacement une base grâce à l'algorithme du pivot de Gauß).*

Le noyau de $\Phi_q - I_d$ s'identifie à l'ensemble des $a \in A$ tels que $a^q = a$, c'est-à-dire aux éléments de $A'$ qui est bien un $\mathbf{F}_q$-espace vectoriel de dimension $e$ d'après ce qui précède.

***Remarque 6.16***. — Il existe diverses méthodes plus ou moins sophistiquées pour déterminer la matrice $\Phi_q$. Une façon de faire consiste à déterminer $W = X^q \mod P$, puis à déterminer les coefficients de la matrice $\Phi_q$ colonne par colonne : si $W^i \mod P$ a été déterminé, en multipliant le résultat par $W$ et en prenant le reste de la division euclidienne par $P$, on obtient $W^{i+1} \mod P$.

**6.2.2. Détermination d'un facteur non trivial.** — Si $e = 1$, le polynôme $P \in \mathbf{F}_q[X]$ est irréductible. Si $e \geq 2$, il reste à expliciter un facteur non trivial de $P$ :

***Proposition 6.17***. — *Soit $P$ un polynôme séparable de degré $d \geq 1$ sur $\mathbf{F}_q$. Soit $Q \in \mathbf{F}_q[X]$ (non constant, et de degré $< d$) un vecteur propre de $\Phi_q$ pour la valeur propre 1. Alors, il existe $\lambda \in \mathbf{F}_q$ tel que le pgcd de $Q + \lambda$ et $P$ soit un facteur non trivial de $P$.*

Ainsi, en calculant le pgcd de $Q + \lambda$ et de $P$ pour $\lambda$ parcourant $\mathbf{F}_q$ avec l'algorithme d'Euclide, on obtient un facteur non trivial de $P$.

Montrons la proposition. On utilise l'isomorphisme canonique $f \colon A \xrightarrow{\sim} K_1 \times \cdots \times K_e$. Posons $(a_1, \ldots, a_e) = f(Q)$. L'hypothèse selon laquelle $\Phi_q(Q) = Q$ signifie que pour tout $i$, $a_i$ appartient au sous-corps $\mathbf{F}_q$ de $K_i$. Dire que $Q$ n'est pas constant signifie que les $a_i$ ne sont pas tous égaux entre eux. Posons $\lambda = -a_1$. Par construction, $f(Q + \lambda)$ est non nul mais au moins une (et l'occurrence la première) de ses coordonnées s'annule. Ainsi, $P_1$ divise $Q + \lambda$ ; le pgcd de $P$ et $Q + \lambda$ est un facteur non trivial de $P$.

# CHAPITRE 7

# RÉSULTANTS

## 7.1. Définitions

***Définition 7.1***. — Soit $A$ un anneau. Soit $n \in \mathbf{N}$. On note $A[X]_{\leq n}$ ou $A[X]_{<n+1}$ le sous-$A$-module de $A[X]$ formé des polynômes de degré $\leq n$. La base canonique de $A[X]_{\leq n}$ est $(1, X, X^2, \ldots, X^n)$. La base « anticanonique » est $X^n, \ldots, X, 1$.

***Définition 7.2***. — Soit $A$ un anneau. Soit $P, Q \in A[X] \setminus \{0\}$ ; notons $m = \deg P$ et $n = \deg Q$ et supposons $(m, n) \neq (0, 0)$. On note $\mathrm{Res}(P, Q) \in A$ le déterminant du morphisme qui à $(U, V) \in A[X]_{<n} \times A[X]_{<m}$ associe $UP + VQ \in A[X]_{<m+n}$ dans les bases « anticanoniques », à savoir $((X^{n-1}, 0), \ldots, (X, 0), (1, 0), (0, X^{m-1}), \ldots, (0, X), (0, 1))$ pour le $A$-module de départ et $(X^{m+n-1}, X^{m+n-2}, \ldots, X, 1)$ pour le $A$-module d'arrivée.

***Proposition 7.3***. — *Soit $P = v_m X^m + \cdots + v_1 X + v_0 \in A[X]$ et $Q = w_n X^n + \cdots + w_1 X + w_0 \in A[X]$, avec $v_m$ et $w_n$ non nuls. Alors,* $\mathrm{Res}(P, Q)$ *est le déterminant de la matrice de Sylvester* $\mathrm{Syl}(P, Q)$ *(carrée de taille $m + n$) :*

$$\begin{pmatrix}
v_m & v_{m-1} & \ldots & v_0 & 0 & 0 & \ldots & 0 \\
0 & v_m & v_{m-1} & \ldots & v_0 & 0 & \ldots & 0 \\
0 & 0 & \ddots & \ddots & \ddots & \ddots & 0 & 0 \\
0 & \ldots & 0 & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & \ldots & \ldots & 0 & v_m & v_{m-1} & \ldots & v_0 \\
w_n & \ldots & \ldots & \ldots & w_1 & w_0 & 0 & \ldots \\
0 & w_n & \ddots & \ddots & \ddots & w_1 & w_0 & 0 \\
0 & 0 & w_n & \ldots & \ldots & \ldots & w_1 & w_0
\end{pmatrix}$$

La matrice de Sylvester est la transposée de la matrice du morphisme défini plus haut dans les bases considérées.

***Proposition 7.4***. — *Soit $A$ un anneau (intègre). Soit $P \in A[X]$ de degré $m \geq 1$ ; notons $v_m$ son coefficient dominant et supposons $P' \neq 0$. On note* $\mathrm{Disc}\, P = \frac{1}{v_m}(-1)^{\frac{m(m-1)}{2}} \mathrm{Res}(P, P')$. *Si* $\deg P' = m - 1$ *il s'agit d'un élément de $A$, appelé discriminant de $P$.*

Lorsque $P' \neq 0$ et $\deg P' < m - 1$, la normalisation choisie ici ne garantit pas que $\mathrm{Disc}\, P \in A$ en général ; c'est sans importance pour les applications que nous avons en vue.

Sur la première colonne de la matrice de Sylvester $\mathrm{Syl}(P, P')$, tous les coefficients, sauf peut-être deux, sont nuls. Les deux autres sont les coefficients de $X^m$ dans $P$ et de $X^{m-1}$ dans $P'$, à savoir $v_m$ et $mv_m$. On peut définir une matrice $\mathrm{Syl}'(P, P')$ en remplaçant respectivement les coefficients $v_m$ et $mv_m$ de la première colonne de $\mathrm{Syl}(P, P')$ par 1 et $m$. Si $\deg P' = m - 1$ alors au signe près, le discriminant de $P$ est le déterminant de cette matrice $\mathrm{Syl}'(P, P')$ (ceci a un sens même si

l'anneau $A$ n'est pas intègre). Par construction, ce déterminant est un élément de $A$ satisfaisant :

$$v_m \operatorname{Disc} P = (-1)^{\frac{m(m-1)}{2}} \operatorname{Res}(P, P') .$$

## 7.2. Propriétés fondamentales du résultant

Sauf mention du contraire, dans toute la suite, $P$ et $Q$ sont des polynômes à coefficients dans un anneau $A$ de degrés respectifs $m$ et $n$. On suppose $(m, n) \neq (0, 0)$.

***Proposition 7.5***. — *Il existe $U \in A[X]_{<n}$ et $V \in A[X]_{<m}$ tels que $UP + VQ = \operatorname{Res}(P, Q)$.*

Notons $M$ la transposée de la comatrice de la transposée de la matrice de Sylvester $\operatorname{Syl}(P, Q)$. On a

$$\operatorname{Res}(P, Q) \cdot \mathbf{I}_{m+n} = {}^t \operatorname{Syl}(P, Q) \cdot M .$$

Si on note $W$ la dernière colonne de la matrice $M$, on obtient :

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \operatorname{Res}(P, Q) \end{pmatrix} = {}^t \operatorname{Syl}(P, Q) \cdot W .$$

La traduction en termes d'applications linéaires de l'existence d'un vecteur colonne $W$ satisfaisant cette relation est qu'il existe $(U, V)$ vérifiant les conditions de degré ci-dessus tels que $UP + VQ = \operatorname{Res}(P, Q)$.

***Théorème 7.6***. — *Supposons que $A$ est un corps $k$. Alors les conditions suivantes sont équivalentes :*

(i) $\operatorname{Res}(P, Q) \neq 0$ ;
(ii) *Pour tout $R \in k[X]_{<m+n}$, il existe $U \in k[X]_{<n}$ et $V \in k[X]_{<m}$ tels que $UP + VQ = R$ ;*
(iii) *Il existe $U \in k[X]_{<n}$ et $V \in k[X]_{<m}$ tels que $UP + VQ = 1$ ;*
(iv) *$P$ et $Q$ sont premiers entre eux.*

*Si $E/k$ est une extension de $k$ sur laquelle $P$ ou $Q$ est scindé, on peut rajouter la condition suivante :*

(v) *$P$ et $Q$ n'ont pas de racines communes dans $E$.*

La non-annulation du déterminant $\operatorname{Res}(P, Q)$ de la matrice de Sylvester équivaut à la surjectivité de l'application linéaire associée, c'est-à-dire que (i) $\iff$ (ii). Ensuite, on a bien sûr, (ii) $\implies$ (iii) $\implies$ (iv). Montrons (iv) $\implies$ (ii). Pour $R \in k[X]_{<m+n}$, il existe évidemment $U$ et $V$ dans $k[X]$ tels que $UP + VQ = R$ ; il reste à montrer que l'on peut s'arranger pour que $U$ et $V$ vérifient les conditions de degré énoncées dans (ii). Soit $\widetilde{U} = U - AQ \in k[X]_{<n}$ le reste de la division euclidienne de $U$ par $Q$. On pose $\widetilde{V} = V + AP$. On vérifie aussitôt que $\widetilde{U}P + \widetilde{V}Q = R$. Comme $\widetilde{V}Q = R - \widetilde{U}P$ et que $R$ et $\widetilde{U}P$ sont de degrés $< m+n$, il vient que $\widetilde{V} \in k[X]_{<m}$. Pour l'équivalence (i) $\iff$ (v), on peut supposer que $P$ est scindé sur $E$, et le résultat découle de la formule de la proposition 7.17 ci-dessous. On peut aussi démontrer directement (iv) $\iff$ (v) en invoquant le fait que $P$ et $Q$ ont le même pgcd qu'on les voie dans $k[X]$ ou dans $E[X]$ (car ce pgcd s'obtient par l'algorithme d'Euclide, qui se déroule de la même façon sur $k$ et sur $E$).

***Proposition 7.7***. — *Soit $\varphi\colon A \to B$ un morphisme d'anneaux. On note $P^\varphi$ (resp. $Q^\varphi$) le polynôme de $B[X]$ obtenu en appliquant $\varphi$ aux coefficients de $P$ (resp. $Q$).*
*Si $\deg P^\varphi = \deg P$ et $\deg Q^\varphi = \deg Q$ alors $\operatorname{Res}(P^\varphi, Q^\varphi) = \varphi(\operatorname{Res}(P, Q))$.*

Par exemple si $k$ est un corps, $A = k[Y]$ et $B = k$, on peut utiliser ce résultat lorsque $\varphi$ est la spécialisation en $y_0 \in k$, définie par $\varphi(R) = R(y_0)$ ; on a alors $P^\varphi(X) = P(X, y_0)$ (voir §7.5.3). Une autre application de la proposition 7.7 (combinée au théorème 7.6) est la suivante.

***Proposition 7.8***. — *Soit $p$ un nombre premier. Soit $P, Q \in \mathbf{Z}[X] \setminus \{0\}$. Supposons que $p$ ne divise ni le coefficient dominant de $P$, ni celui de $Q$ ; notons $\overline{P}$ et $\overline{Q}$ les réductions modulo $p$ de $P$ et $Q$ respectivement. Alors la classe modulo $p$ de* $\operatorname{Res}(P,Q)$ *est* $\operatorname{Res}(\overline{P},\overline{Q})$. *En particulier, les assertions suivantes sont équivalentes : $p$ divise* $\operatorname{Res}(P,Q)$ *;* $\operatorname{Res}(\overline{P},\overline{Q}) = 0$ *; $\overline{P}$ et $\overline{Q}$ ne sont pas premiers entre eux dans $\mathbf{F}_p[X]$.*

Un autre corollaire du théorème 7.6 est le suivant.

***Corollaire 7.9***. — *Soit $k$ un corps. Soit $P \in k[X]$ un polynôme tel que $P' \neq 0$. Alors, $P$ est séparable si et seulement si* $\operatorname{Res}(P,P') \neq 0$, *c'est-à-dire que* $\operatorname{Disc} P \neq 0$.

En utilisant en outre la proposition 7.7 on obtient une preuve du corollaire suivant, dans le cas où $p$ ne divise pas le degré de $P$.

***Corollaire 7.10***. — *Soit $P \in \mathbf{Z}[X]$ un polynôme de degré $m \geq 1$. Soit $p$ un nombre premier. On suppose que $p$ ne divise pas le coefficient dominant de $P$. On note $\overline{P} \in \mathbf{Z}/p\mathbf{Z}[X]$ la réduction de $P$ modulo $p$. Alors, $p$ divise* $\operatorname{Disc} P$ *si et seulement si $\overline{P}$ n'est pas séparable.*

Ce corollaire est le point crucial qui servira au chapitre 8. Lorsque $p$ divise le degré de $P$, on le déduit de la proposition suivante, qui généralise la proposition 7.7.

***Proposition 7.11***. — *Soit $\varphi\colon A \to B$ un morphisme d'anneaux. On note $P^\varphi$ (resp. $Q^\varphi$) le polynôme de $B[X]$ obtenu en appliquant $\varphi$ aux coefficients de $P$ (resp. $Q$). Supposons que* $\deg P^\varphi = \deg P$ *et notons $v_m$ le coefficient dominant de $P$. Alors on a*

$$\varphi(\operatorname{Res}(P,Q)) = \begin{cases} 0 \text{ si } Q^\varphi = 0 \text{ et } \deg P \geq 1 \\ \varphi(v_m)^n \text{ si } Q^\varphi = 0 \text{ et } \deg P = 0 \\ \varphi(v_m)^{n-\deg(Q^\varphi)} \operatorname{Res}(P^\varphi, Q^\varphi) \text{ si } Q^\varphi \neq 0. \end{cases}$$

En effet si $Q^\varphi = 0$ et $\deg P \geq 1$ alors la dernière ligne de $\varphi(\operatorname{Syl}(P,Q))$ est identiquement nulle, d'où $\varphi(\operatorname{Res}(P,Q)) = \det \varphi(\operatorname{Syl}(P,Q)) = 0$. Si $P$ est constant (et non nul) alors $\operatorname{Res}(P,Q) = v_m^n$ et le résultat est évident. Enfin, si $Q^\varphi \neq 0$ on constate que les $n - \deg Q^\varphi$ premières colonnes de $\varphi(\operatorname{Syl}(P,Q))$ ont chacune $\varphi(v_m)$ comme coefficient diagonal, et seulement des zéros en-dessous ; après développement par rapport à ces colonnes il reste $\det \operatorname{Syl}(P^\varphi, Q^\varphi)$.

Les propositions 7.11 et 7.14 montrent notamment qu'on peut généraliser comme suit la proposition 7.8.

***Proposition 7.12***. — *Soit $p$ un nombre premier. Soit $P, Q \in \mathbf{Z}[X] \setminus \{0\}$. Supposons que $p$ ne divise pas à la fois le coefficient dominant de $P$ et celui de $Q$. Notons $\overline{P}$ et $\overline{Q}$ les réductions modulo $p$ de $P$ et $Q$ respectivement ; supposons $\overline{P} \neq 0$ et $\overline{Q} \neq 0$. Alors les assertions suivantes sont équivalentes : $p$ divise* $\operatorname{Res}(P,Q)$ *;* $\operatorname{Res}(\overline{P},\overline{Q}) = 0$ *; $\overline{P}$ et $\overline{Q}$ ne sont pas premiers entre eux dans $\mathbf{F}_p[X]$.*

## 7.3. Formules

On considère encore $P, Q \in A[X] \setminus \{0\}$ ; on note $m = \deg P$ et $n = \deg Q$ et on suppose $(m,n) \neq (0,0)$.

***Proposition 7.13***. — *Soient $\lambda, \mu \in A \setminus \{0\}$ tels que* $\deg(\lambda P) = \deg P$ *et* $\deg(\mu Q) = \deg Q$. *Alors on a* $\operatorname{Res}(\lambda P, \mu Q) = \lambda^n \mu^m \operatorname{Res}(P,Q)$. *En particulier,* $\operatorname{Res}(\lambda, Q) = \lambda^n$ *et* $\operatorname{Res}(P,\mu) = \mu^m$.

***Proposition 7.14***. — $\operatorname{Res}(Q,P) = (-1)^{mn} \operatorname{Res}(P,Q)$.

Calculer $\operatorname{Res}(Q,P)$ revient à calculer le déterminant intervenant dans le calcul de $\operatorname{Res}(P,Q)$ en changeant la base du $A$-module de départ : on inverse les vecteurs de base associés aux variables $U$ et $V$ ; le nombre d'inversions de cette permutation est $mn$.

***Proposition 7.15***. — *Soit $\lambda \in A$. On a* $\operatorname{Res}(P(X+\lambda), Q(X+\lambda)) = \operatorname{Res}(P,Q)$.

On peut calculer $\mathrm{Res}(P(X+\lambda), Q(X+\lambda))$ en utilisant des bases autres que celles considérées dans la définition 7.2. La matrice du morphisme $(U,V) \longmapsto UP(X+\lambda) + VQ(X+\lambda)$ dans les bases $(((X+\lambda)^{n-1},0),\dots,(X+\lambda,0),(1,0),(0,(X+\lambda)^{m-1}),\dots,(0,X+\lambda),(0,1))$ et $((X+\lambda)^{m+n-1},\dots,X+\lambda,1)$ de $A[X]_{<n} \times A[X]_{<m}$ et $A[X]_{<m+n}$ est égale à la matrice du morphisme $(U,V) \longmapsto UP+VQ$ dans les bases de la définition 7.2 ; son déterminant est donc $\mathrm{Res}(P,Q)$. Or, manifestement, les matrices de changement de base entre les bases considérées ici et celles de la définition 7.2 sont triangulaires et leurs diagonales ne contiennent que des 1. On en déduit l'égalité voulue.

***Proposition 7.16***. — *Soit $\lambda \in A$. On a* $\mathrm{Res}((X-\lambda)P, Q) = \mathrm{Res}(P,Q)\cdot\mathrm{Res}(X-\lambda, Q)$. *En outre,* $\mathrm{Res}(X-\lambda, Q) = Q(\lambda)$.

D'après la proposition 7.15, on peut supposer que $\lambda = 0$. Sur la dernière colonne de la matrice de Sylvester $\mathrm{Syl}(XP,Q)$, tous les coefficients sont nuls sauf peut-être le dernier qui vaut $Q(0)$. En développant le déterminant de $\mathrm{Syl}(XP,Q)$ par rapport à cette colonne, on obtient aussitôt l'égalité $\mathrm{Res}(XP,Q) = \mathrm{Res}(P,Q)\cdot Q(0)$. L'égalité $\mathrm{Res}(X,Q) = Q(0)$ est évidente.

***Proposition 7.17***. — *Soit $(a_1,\dots,a_m) \in A^m$. On suppose que $P = (X-a_1)\dots(X-a_m)$. Alors,* $\mathrm{Res}(P,Q) = \prod_{i=1}^m Q(a_i)$.

Cette formule se déduit aussitôt de la proposition 7.16 par récurrence sur $m$.

***Proposition 7.18***. — *Soit $(a_1,\dots,a_m) \in A^m$. Soit $(b_1,\dots,b_n) \in A^n$. On suppose que $P = (X-a_1)\dots(X-a_m)$ et $Q = (X-b_1)\dots(X-b_n)$. Alors,* $\mathrm{Res}(P,Q) = \prod_{1\le i\le m, 1\le j\le n}(a_i - b_j)$.

Cela résulte aussitôt de la proposition 7.17.

***Proposition 7.19***. — *Soit $(a_1,\dots,a_m) \in A^m$. On pose $P = (X-a_1)\dots(X-a_m)$ et on suppose $P' \neq 0$. Alors,* $\mathrm{Disc}\, P = \prod_{i<j}(a_i - a_j)^2$.

En utilisant la formule de la proposition 7.17 et l'égalité $P'(a_i) = \prod_{j\neq i}(a_i - a_j)$ pour tout $i$, on obtient que $\mathrm{Res}(P,P') = \prod_{i\neq j}(a_i - a_j)$, ce qui permet de conclure.

## 7.4. Calcul efficace du résultant

On considère toujours $P, Q \in A[X] \setminus \{0\}$ ; on note $m = \deg P$ et $n = \deg Q$, et on suppose $(m,n) \neq (0,0)$.

***Proposition 7.20***. — *On suppose $m \ge n \ge 1$. Soit $S \in A[X]_{\le m-n}$ tel que $P + SQ \neq 0$. Alors* $\mathrm{Res}(P,Q) = (-1)^{(m-r)n} w_n^{m-r}\,\mathrm{Res}(P+SQ, Q)$ *où $r = \deg(P+SQ)$ et $w_n$ est le coefficient dominant de $Q$.*

Notons $\alpha, \beta\colon A[X]_{<n} \times A[X]_{<m} \to A[X]_{<m+n}$ les applications $A$-linéaires définies par $\alpha(U,V) = UP + VQ$ et $\beta(U,V) = U(P+SQ) + VQ$. On peut considérer l'endomorphisme $\gamma$ de $A[X]_{<n} \times A[X]_{<m}$ défini par $\gamma(U,V) = (U, V + US)$. Il est de déterminant 1 car sa matrice dans la base anticanonique de $A[X]_{<n} \times A[X]_{<m}$ est triangulaire à diagonale de 1. La relation $\beta = \alpha \circ \gamma$ implique donc $\det \mathrm{Mat}\,\beta = \det \mathrm{Mat}\,\alpha$, où les matrices sont prises dans les bases considérées dans la définition 7.2. Par définition on a $\mathrm{Res}(P,Q) = \det \mathrm{Mat}\,\alpha$. Dans le cas particulier $r = m$ on a de même $\mathrm{Res}(P+SQ, Q) = \det \mathrm{Mat}\,\beta$, ce qui termine la preuve de la proposition. En général on a seulement $0 \le r \le m$, et en développant successivement par rapport aux $m-r$ premières colonnes de la transposée de $\mathrm{Mat}\,\beta$ on obtient $\det \mathrm{Mat}\,\beta = ((-1)^n w_n)^{m-r}\,\mathrm{Res}(P+SQ, Q)$. Ceci termine la preuve de la proposition 7.20.

***Proposition 7.21***. — *On suppose que $A$ est un corps. Supposons $m = \deg P \ge n = \deg Q \ge 1$. Notons $R$ le reste de la division euclidienne de $P$ par $Q$. Si $R = 0$, alors* $\mathrm{Res}(P,Q) = 0$, *sinon* $\mathrm{Res}(P,Q) = (-1)^{mn} w_n^{m-r}\,\mathrm{Res}(Q,R)$ *où $w_n$ est le coefficient dominant de $Q$ et $r = \deg R$.*

Si $R = 0$ alors $P$ et $Q$ ont un facteur commun non trivial (car $m \geq n > 0$), donc $\mathrm{Res}(P, Q) = 0$. Sinon on applique les propositions 7.20 et 7.14 :

$$\mathrm{Res}(P, Q) = (-1)^{(m-r)n} w_n^{m-r} \mathrm{Res}(R, Q) = (-1)^{mn} w_n^{m-r} \mathrm{Res}(Q, R).$$

***Remarque 7.22.*** — Cette proposition fournit un algorithme pour déterminer le résultant de deux polynômes. Sa complexité est manifestement du même ordre que celle de l'algorithme d'Euclide, c'est-à-dire $O((\deg P)(\deg Q))$ opérations arithmétiques dans le corps de base. A titre de comparaison, le calcul direct du résultant vu comme déterminant de la matrice de Sylvester a un coût de $O(n^3)$ opérations, avec $n = \max(\deg P, \deg Q)$ (voir le chapitre 3).

## 7.5. Exemples

**7.5.1. Une première courbe paramétrée : le cercle unité.** — Les résultants permettent d'éliminer des variables dans des systèmes d'équations algébriques. Un exemple est celui des courbes paramétrées comme celle-ci :

$$\begin{cases} x = \frac{1-t^2}{1+t^2} \\ y = \frac{2t}{1+t^2} \end{cases}$$

Notons $C \subset \mathbf{C}^2$ l'ensemble des points dont les coordonnées $(x, y)$ s'obtiennent ainsi pour un paramètre $t \in \mathbf{C} - \{\pm i\}$. Autrement dit, $C$ est l'ensemble des couples $(x, y) \in \mathbf{C}$ tels qu'il existe $t \in \mathbf{C} - \{\pm i\}$ solution commune des deux équations :

$$\begin{cases} (1+t^2)x = 1 - t^2 \\ (1+t^2)y = 2t \end{cases}$$

Pour $t = \pm i$, il n'y a pas de solutions. Il revient donc au même de chercher $t$ dans $\mathbf{C} - \{\pm i\}$ ou dans $\mathbf{C}$ tout entier. Pour $(x, y) \in \mathbf{C}$ tels que $x \neq -1$ et $y \neq 0$, on considère les polynômes $P_{x,y} = (x+1)T^2 + x - 1$ et $Q_{x,y} = yT^2 - 2T + y$ ; alors $\deg P_{x,y} = \deg Q_{x,y} = 2$. Comme $\mathbf{C}$ est algébriquement clos, d'après le théorème 7.6 on a équivalence entre l'annulation de $\mathrm{Res}(P_{x,y}, Q_{x,y})$ et le fait que $P_{x,y}$ et $Q_{x,y}$ admettent une racine commune $t$ (c'est-à-dire $(x, y) \in C$). Le calcul donne $\mathrm{Res}(P_{x,y}, Q_{x,y}) = 4(x^2 + y^2 - 1)$. Donc pour $x \neq -1$ et $y \neq 0$, on a $(x, y) \in C$ si, et seulement si, $4(x^2 + y^2 - 1) = 0$. Reste à étudier les cas particuliers $x = -1$ et $y = 0$. Pour $y = 0$, on a $(x, 0) \in C$ si, et seulement si, x=1 ; et $R(x, 0) = 0$ équivaut à $x \in \{-1, 1\}$. Pour $x = -1$ il n'existe aucun $y \in C$ tel que $(x, y) \in C$ (car on a $1 - t^2 \neq -1 - t^2$ pour tout $t$), alors que $R(-1, y) = 0$ équivaut à $y = 0$. Finalement, $C$ est donc l'ensemble des $(x, y) \in \mathbf{C}^2$ différents de $(-1, 0)$ et tels que $4(x^2 + y^2 - 1) = 0$ : c'est le cercle unité privé du point $(-1, 0)$. On aurait aussi pu utiliser la proposition 7.11 pour voir que seul le cas $(x, y) = (-1, 0)$ nécessitait vraiment d'être traité séparément.

Plus conceptuellement, on peut considérer $P = (X+1)T^2 + X - 1$ et $Q = YT^2 - 2T + Y$ comme des éléments de $A[T]$ où $A = \mathbf{C}[X, Y]$. Soit $R = \mathrm{Res}(P, Q) \in A = \mathbf{C}[X, Y]$. Pour tout couple $(x, y) \in \mathbf{C}^2$ tel que $x \neq -1$ et $y \neq 0$, le résultant de $P_{x,y}$ et $Q_{x,y}$ est $R(x, y)$ car $\deg P_{x,y} = \deg P = 2$ ou $\deg Q_{x,y} = \deg Q = 2$.

**7.5.2. Une deuxième courbe paramétrée : la cardioïde.** — On fixe une constante $a \in \mathbf{C}^\times$. On considère la courbe paramétrée suivante :

$$\begin{cases} x = 2a\frac{1-t^2}{(1+t^2)^2} \\ y = \frac{4at}{(1+t^2)^2} \end{cases}$$

On cherche à déterminer une équation de cette courbe plane. Un point $(x, y) \in \mathbf{C}^2$ du plan étant donné, la question est de savoir s'il existe $t \in \mathbf{C} - \{\pm i\}$ vérifiant le système d'équations ci-dessus, système que l'on peut réécrire sous une forme polynomiale :

$$\begin{cases} x(1+t^2)^2 - 2a(1-t^2) = 0 \\ y(1+t^2)^2 - 4at = 0 \end{cases}$$

(Si on fait $t = \pm i$, on n'obtient pas de solutions supplémentaires.) On peut considérer ces équations comme étant définies par l'annulation de deux polynômes $P$ et $Q$ dans $\mathbf{C}[X,Y][T]$. On a $\deg P = \deg Q = 4$ ; le coefficient dominant de $P$ est $X$, et celui de $Q$ est $Y$.

Un point $(x,0)$ appartient à la courbe si, et seulement si, $x = 2a$ (car le paramètre $t$ vaut alors 0). De même, un point $(0,y)$ appartient à la courbe si, et seulement si, $y = \pm a$ (avec $t = \pm 1$). Soit $(x,y) \in \mathbf{C}^2$ tel que $x \neq 0$ et $y \neq 0$. Alors on a $\deg P(x,y,T) = \deg P(X,Y,T) = 4$ (et idem pour $Q$) donc en posant $R = \mathrm{Res}(P,Q) \in \mathbf{C}[X,Y]$ on a $R(x,y) = \mathrm{Res}(P(x,y,T), Q(x,y,T))$ d'où l'équivalence (puisque $\mathbf{C}$ est algébriquement clos) entre :

- $R(x,y) = 0$ ;
- Il existe $t \in \mathbf{C}$ tel que $P(x,y,t) = Q(x,y,t) = 0$.

Il reste à voir si cette équivalence est encore vraie lorsque $x$ ou $y$ est nul. Pour cela on commence par calculer le résultant : on trouve $R = 256a^4\left[(X^2+Y^2-aX)^2 - a^2(X^2+Y^2)\right]$. On a donc $R(x,0) = 256a^4x^3(x-2a)$ qui s'annule en 0 et en $2a$ ; le point $(2a,0)$ appartient bien à la courbe, mais pas le point $(0,0)$. De même, on a $R(0,y) = 256a^4y^2(y-a)(y+a)$ qui s'annule en 0 et en $\pm a$. Finalement, la courbe est exactement l'ensemble des éléments $(x,y) \in \mathbf{C}^2 - \{(0,0)\}$ tels que $R(x,y) = 0$.

On aurait aussi pu utiliser la proposition 7.11 pour voir que seul le cas $(x,y) = (0,0)$ nécessitait vraiment d'être traité séparément.

**7.5.3. Intersection de deux courbes planes.** — Considérons dans $\mathbf{C}[X,Y]$ les deux polynômes suivants :

$$P(X,Y) = 2XY^2 + 4XY + 5 \qquad \text{et} \qquad Q(X,Y) = XY - X - 1.$$

Notons $\mathcal{C}$ l'ensemble des couples $(x,y) \in \mathbf{C}^2$ tels que $P(x,y) = Q(x,y) = 0$, et $R \in \mathbf{C}[X]$ le résultant de $P$ et $Q$ vus dans $(\mathbf{C}[X])[Y]$. On a $R(X) = X(X+2)(6X+1)$.

Montrons que pour tout $(x,y) \in \mathcal{C}$ on a $R(x) = 0$. Soit $(x,y) \in \mathcal{C}$. Notons $\varphi : \mathbf{C}[X] \to \mathbf{C}$ l'homomorphisme de $\mathbf{C}$-algèbres consistant à évaluer en $x$. Pour tout $S = S(X,Y) \in \mathbf{C}[X,Y]$, on note $S^\varphi(Y) = S(x,Y) \in \mathbf{C}[Y]$ l'image de $S$ par l'homomorphisme $\mathbf{C}[X,Y] \to \mathbf{C}[Y]$ consistant à appliquer $\varphi$ à chaque coefficient de $S$ vu dans $(\mathbf{C}[X])[Y]$. Le coefficient dominant de $P$ dans $(\mathbf{C}[X])[Y]$ est $2X$ ; celui de $Q$ est $X$. Si $x = 0$ la conclusion $R(0) = 0$ est triviale ; supposons désormais $x \neq 0$. Alors les coefficients dominants de $P$ et $Q$ ont des images non nulles par $\varphi$ donc on a

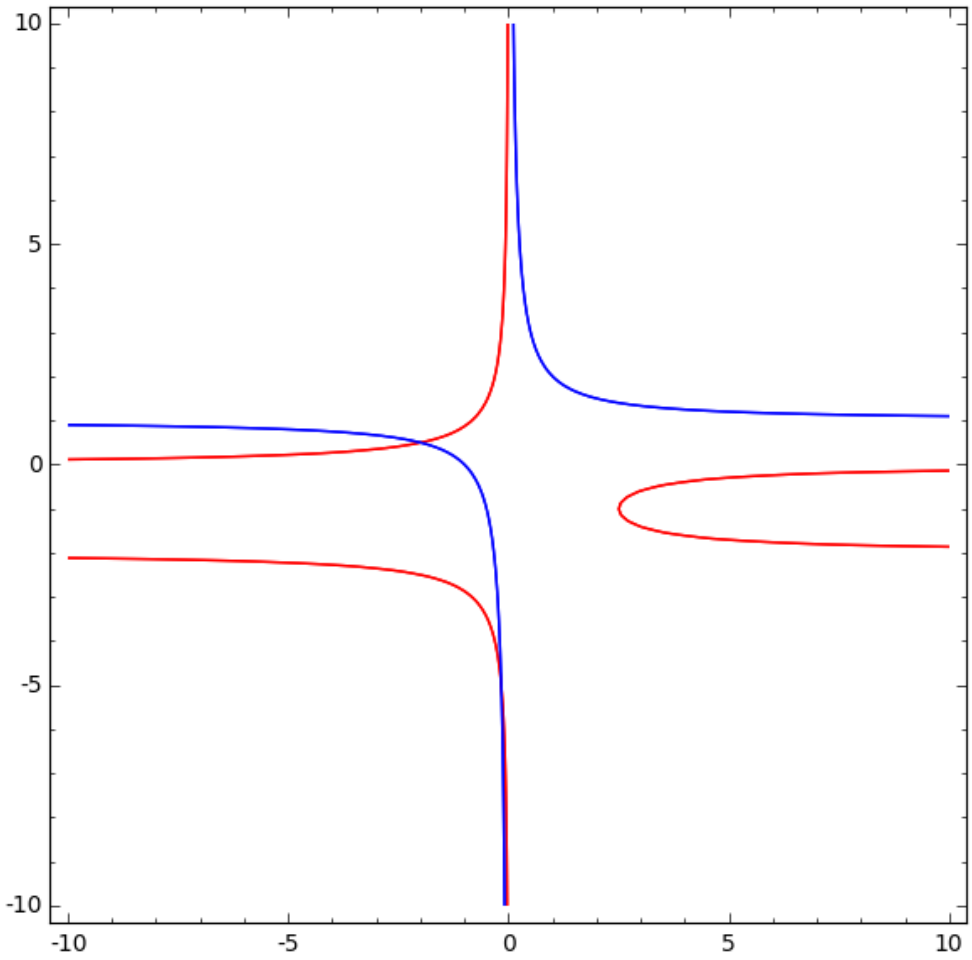
$$\mathrm{Res}(P^\varphi, Q^\varphi) = \varphi(\mathrm{Res}(P,Q)) = \varphi(R) = R(x).$$

Or par hypothèse le réel $y$ vérifie $P(x,y) = Q(x,y) = 0$ donc les polynômes $P^\varphi$ et $Q^\varphi$ ont une racine commune $y$. Cela implique $\mathrm{Res}(P^\varphi, Q^\varphi) = 0$, c'est-à-dire $R(x) = 0$.

Géométriquement, notons $p$ la projection orthogonale sur l'axe des abscisses, dans le plan $\mathbf{R}^2$ muni de sa structure euclidienne canonique ; autrement dit on a $p((x,y)) = x$. Plus généralement on note $p((x,y)) = x$ pour tout $(x,y) \in \mathbf{C}^2$. On a montré que l'image de $\mathcal{C}$ par $p$ est contenue dans l'ensemble des points de la forme $(x,0)$, où $x$ est une racine de $R$. Cela permet de déterminer exactement $\mathcal{C}$ : on sait que les abscisses des points de $\mathcal{C}$ sont des racines de $R$, donc appartiennent à l'ensemble fini $\{0,-2,-1/6\}$. Pour chacune de ces racines $x$, on factorise les polynômes $P(x,Y)$ et $Q(x,Y)$ pour voir si ils ont une racine commune. On constate que c'est le cas pour $-2$ et $-1/6$, mais pas pour 0. On obtient ainsi que $\mathcal{C}$ est l'ensemble formé par les deux points $(-2,1/2)$ et $(-1/6,-5)$.

Il faut prendre garde au fait que l'implication "Si $(x,y) \in \mathcal{C}$ alors $R(x) = 0$" démontrée ci-dessus n'est pas une équivalence : on a vu que 0 fait exception à la réciproque. Une première explication de ce phénomène est que 0 est une racine commune des coefficients dominants de $P$ et $Q$. Ceci implique $R(0) = 0$, indépendamment de l'existence de $y \in \mathbf{C}$ tel que $(x,y) \in \mathcal{C}$. Une autre explication, plus géométrique et qui dépasse le cadre du cours, est la présence d'une asymptote verticale commune aux courbes définies par $P$ et $Q$, d'équation $x = 0$. Intuitivement, il y a bien un point d'abscisse 0 sur $\mathcal{C}$, mais il est "à l'infini" dans la direction verticale. L'allure des courbes d'équations $P(x,y) = 0$ (en rouge, avec trois branches et deux asymptotes horizontales d'équations $y = 0$ et $y = -2$) et $Q(x,y) = 0$ (en bleu, avec deux branches et une asymptote horizontale d'équation $y = 1$) est donnée sur la figure ci-dessous.

En calculant le résultant $S$ de $P$ et $Q$, vus dans $(\mathbf{C}[Y])[X]$, ce phénomène ne se serait pas produit car les coefficients dominants de $P$ et $Q$ dans $(\mathbf{C}[Y])[X]$ sont respectivement $2Y^2+4Y$ et $Y-1$ : ils n'ont aucune racine commune. On obtiendrait donc $1/2$ et $-5$ comme seules racines de $S$ : ce sont exactement les ordonnées des éléments de $\mathcal{C}$.



# CHAPITRE 8

# LEMME DE HENSEL, BORNE DE MIGNOTTE, FACTORISATION DANS $\mathbf{Z}[X]$

## 8.1. Lemme de Hensel

### 8.1.1. Résolution d'équations. —

***Lemme 8.1* (Hensel).** — *Soit $P \in \mathbf{Z}[X]$ un polynôme unitaire. Soit $p$ un nombre premier. Soit $n \geq 1$. Soit $x \in \mathbf{Z}$ tel que $P(x) \equiv 0 \mod p^n$ et $P'(x) \not\equiv 0 \mod p$. Alors, il existe $\tilde{x} \in \mathbf{Z}$ tel que $P(\tilde{x}) \equiv 0 \mod p^{2n}$ et $\tilde{x} \equiv x \mod p^n$. En outre, la classe de congruence de $\tilde{x}$ modulo $p^{2n}$ est unique.*

On cherche $\tilde{x}$ sous la forme $\tilde{x} = x + p^n y$. Le polynôme $P(x+X)$ est de la forme $P(x) + P'(x)X + R(X)$ où $R(X)$ est un polynôme dont les coefficients de 1 et de $X$ sont nuls. Par conséquent, pour $y \in \mathbf{Z}$, $R(p^n y)$ est multiple de $p^{2n}$. Ainsi, pour tout $y \in \mathbf{Z}$, on a :

$$P(\tilde{x}) = P(x + p^n y) \equiv P(x) + p^n y P'(x) \mod p^{2n} .$$

Comme $P(x) \equiv 0 \mod p^n$, on peut écrire $P(x) = p^n v$ avec $v \in \mathbf{Z}$. On a ainsi $P(\tilde{x}) \equiv 0 \mod p^{2n}$ si et seulement si $v + yP'(x) \equiv 0 \mod p^n$. Comme $p$ ne divise pas $P'(x)$, il existe $z \in \mathbf{Z}$ tel que $zP'(x) \equiv 1 \mod p^n$. La condition se réécrit $y \equiv -vz \mod p^n$. Ainsi, $\tilde{x} = x - p^n vz$ convient et est l'unique solution modulo $p^{2n}$ se réduisant sur $x$ modulo $p^n$.

***Remarque 8.2.*** — En itérant le résultat du lemme, on obtient facilement qu'une solution $x$ de $P$ modulo $p^n$ (par exemple une solution modulo $p$) telle que $p$ ne divise pas $P'(x)$ se relève de manière unique en une solution modulo $p^m$ pour tout $m \geq n$.

On peut réinterpréter le lemme de Hensel ainsi :

***Proposition 8.3.*** — *Soit $P \in \mathbf{Z}[X]$ un polynôme unitaire. Soit $p$ un nombre premier. Soit $x \in \mathbf{Z}$ tel que $p$ divise $P(x)$ mais ne divise pas $P'(x)$. Soit $N \geq 1$. On définit une suite $(x_n)_{n \in \mathbf{N}}$ à valeurs dans $\mathbf{Z}/p^N\mathbf{Z}$ en notant $x_0$ la classe de $x$ modulo $p^N$ et par la relation de récurrence :*

$$x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)} .$$

*Alors, à partir d'un certain rang (et en tout cas pour $n \geq \log_2 N$), la suite $(x_n)$ est constante et $P(x_n) = 0$ dans $\mathbf{Z}/p^N\mathbf{Z}$.*

Il convient d'expliquer en quoi cet énoncé a un sens. Tout d'abord, par récurrence, on obtient aussitôt que pour tout $n \geq 0$, $x_n \equiv x_0 \mod p$, que $P(x_n) \equiv P(x_0) \equiv 0 \mod p$, que $P'(x_n) \not\equiv 0 \mod p$ et qu'ainsi, $P'(x_n)$ est inversible dans l'anneau $\mathbf{Z}/p^N\mathbf{Z}$ de sorte que $x_{n+1}$ a bien un sens. La démonstration du lemme de Hensel montre que $P(x_n) \equiv 0 \mod p^{2^n}$. Par conséquent, la suite $(x_n)$ devient constante dès que pour $2^n \geq N$, etc.

On peut rapprocher le lemme de Hensel tel qu'il est interprété dans la proposition 8.3 de la méthode de Newton :

Soit $f : I \to \mathbf{R}$ une fonction de classe $C^2$ définie sur un intervalle compact $I$. On suppose que $f$ admet un zéro $x$ dans cet intervalle et on cherche à en déterminer de bonnes approximations. On

choisit un point de départ $x_0 \in I$ et on définit une suite $(x_n)$ par la formule $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ (c'est la même formule que plus haut !).

On suppose que la suite $(x_n)_{n\in\mathbf{N}}$ est bien définie, c'est-à-dire que l'on aura toujours $x_n \in I$ et $f'(x_n) \neq 0$.

***Proposition 8.4***. — *Avec les notations et hypothèses ci-dessus, si on suppose que $f'$ ne s'annule pas sur $I$ (en particulier $f'(x) \neq 0$) et pourvu que $x_0$ ait été choisi assez proche de $x$, alors $x_n$ converge vers $x$ et, essentiellement, la précision double à chaque itération.*

D'après la formule de Taylor-Lagrange, pour tout $n \geq 0$, il existe un $\xi \in I$ tel que

$$0 = f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(\xi)}{2}(x - x_n)^2 \ .$$

En divisant par $f'(x_n)$, on obtient :

$$x - x_{n+1} = -\frac{f''(\xi)}{2f'(x_n)}(x - x_n)^2 \ .$$

Comme $f'$ ne s'annule pas sur l'intervalle compact $I$, on peut minorer $f'$ en valeur absolue sur $I$ et la valeur absolue $f''$ est majorée sur $I$. Il existe donc une constante $C > 0$ telle que pour tout $n$, on ait :

$$|x - x_{n+1}| \leq C\,|x - x_n|^2 \ .$$

Notons $u_n = \log C + \log|x - x_n|$, on a $u_{n+1} \leq 2u_n$. Pourvu que $u_0 < 0$ (c'est-à-dire $|x - x_0| < \frac{1}{C}$), $u_n$ tend vers $-\infty$ avec même $u_n \leq u_0 2^n$, ce qui permet de conclure.

### 8.1.2. Factorisation. —

***Lemme 8.5***. — *Soit $P \in \mathbf{Z}[X]$ un polynôme unitaire de degré $d$. Soit $p$ un nombre premier. Soit $n \geq 1$. Soit $(\overline{Q}, \overline{R}) \in \mathbf{Z}/p^n\mathbf{Z}[X]$ deux polynômes unitaires de degrés respectifs $q$ et $r$ tels que si on note $\overline{P}$ la classe de $P$ dans $\mathbf{Z}/p^n\mathbf{Z}[X]$ on ait $\overline{P} = \overline{QR}$. On suppose que $\mathrm{Res}(\overline{Q}, \overline{R})$ n'est pas congru à 0 modulo $p$. Alors, pour tout $m \geq n$, il existe d'uniques polynômes unitaires $Q$ et $R$ dans $\mathbf{Z}/p^m\mathbf{Z}[X]$ de degrés $q$ et $r$ se réduisant modulo $p^n$ sur $\overline{Q}$ et $\overline{R}$ tels que l'on ait $P \equiv QR \mod p^m$.*

Il s'agit bien d'une généralisation de la première version du lemme de Hensel puisque la recherche de zéros correspond au cas particulier où $Q$ est de la forme $X - x$ ; les hypothèses sont cohérentes puisqu'alors $\mathrm{Res}(\overline{Q}, \overline{R}) = R(x) = P'(x)$.

Comme pour l'autre version du lemme de Hensel, il suffit de traiter le cas $m = 2n$. On choisit des polynômes unitaires $\tilde{Q}$ et $\tilde{R}$ dans $\mathbf{Z}/p^{2n}\mathbf{Z}[X]$ se réduisant sur $\overline{Q}$ et $\overline{R}$ modulo $p^n$. On a $\mathrm{Res}(Q, R) \not\equiv 0 \mod p$, c'est-à-dire $\mathrm{Res}(\tilde{Q}, \tilde{R}) \not\equiv 0 \mod p$. Notons $A = \mathbf{Z}/p^{2n}\mathbf{Z}$. Le fait que le résultant de $\tilde{Q}$ et de $\tilde{R}$ soit inversible dans $A$ est équivalent à demander que la matrice de Sylvester de $\tilde{Q}$ et de $\tilde{R}$ soit inversible dans $M_d(A)$, c'est-à-dire que l'application qui à $U \in A[X]_{<r}$ et $V \in A[X]_{<q}$ associe $U\tilde{Q} + V\tilde{R} \in A[X]_{<d}$ soit un isomorphisme. On considère la différence $\Delta = P - \tilde{Q}\tilde{R} \in A[X]_{<d}$. Comme $\overline{P} = \overline{QR}$ dans $\mathbf{Z}/p^n\mathbf{Z}[X]$, il vient que les coefficients de $\Delta$ sont congrus à 0 modulo $p^n$. Ce qui précède montre qu'il existe d'uniques $U$ et $V$ comme ci-dessus tels que $\Delta = U\tilde{Q} + V\tilde{R} \in A[X]_{<d}$ et qu'en outre les coefficients de $U$ et $V$ soient des multiples de $p^n$. Posons $Q = \tilde{Q} + V$ et $R = \tilde{R} + U$. Comme les coefficients de $U$ et $V$ sont multiples de $p^n$, on a $UV = 0$ dans $A[X]$, ce qui montre que $QR = \tilde{Q}\tilde{R} + U\tilde{Q} + V\tilde{R} = \tilde{Q}\tilde{R} + \Delta = P$, ce qui montre l'existence de $Q$ et $R$, l'unicité se montrant de manière similaire.

***Remarque 8.6***. — Si on implémente l'idée de la démonstration pour appliquer plusieurs fois le calcul, mieux vaut garder en mémoire une « relation de Bézout » (c'est-à-dire des polynômes $\overline{A}$ et $\overline{B}$ de degrés appropriés tels que $\overline{AQ} + \overline{BR} = 1$) et chercher à déterminer non seulement $Q$ et $R$ vérifiant des congruences modulo une puissance double de $p$ mais aussi $A$ et $B$ tels que $AQ + BR = 1$ en « modifiant » des relèvements arbitraires de $\overline{A}$ et de $\overline{B}$. Pour $n = 1$, on peut utiliser l'algorithme d'Euclide étendu pour obtenir la relation de Bézout.

## 8.2. Borne de Mignotte

***Définition 8.7.*** — Soit $P = a_0 + a_1 X + \cdots \in \mathbf{C}[X]$. On pose $\|P\|_\infty = \max_i |a_i|$, $\|P\|_1 = \sum_i |a_i|$ et $\|P\|_2 = \sqrt{\sum_i |a_i|^2}$. Si $P = a_n \prod_{k=1}^n (X - z_k)$, on pose $M(P) = |a_n| \prod_{k=1}^n \max(1, |z_k|)$ (si $P = 0$, on pose $M(P) = 0$).

***Théorème 8.8* (Mignotte).** — *Soit $P = \sum_i a_i X^i \in \mathbf{C}[X]$. Soit $Q = \sum_i b_i X^i \in \mathbf{C}[X]$. On note $n = \deg P$, $m = \deg Q$. On suppose que $Q$ divise $P$. Alors,*

$$\|Q\|_\infty \leq \|Q\|_1 \leq 2^m M(Q) \leq \left|\frac{b_m}{a_n}\right| 2^m \|P\|_2 \ .$$

***Corollaire 8.9.*** — *Soit $P \in \mathbf{Z}[X]$. Soit $Q \in \mathbf{Z}[X]$. On suppose que $Q$ divise $P$ dans $\mathbf{Z}[X]$. Alors,*

$$\|Q\|_\infty \leq 2^{\deg Q} \|P\|_2 \ .$$

En effet, avec les notations du théorème 8.8, $b_m$ divise $a_n$, donc $\left|\frac{b_m}{a_n}\right| \leq 1$.

***Remarque 8.10.*** — Si $P$ et $Q$ sont des éléments de $\mathbf{C}[X]$, on a $M(PQ) = M(P)M(Q)$.

Un des ingrédients importants de la démonstration du théorème 8.8 est le suivant :

***Théorème 8.11* (Landau).** — *Soit $P \in \mathbf{C}[X]$. Alors, $M(P) \leq \|P\|_2$.*

***Lemme 8.12.*** — *Soit $P \in \mathbf{C}[X]$. Soit $z \in \mathbf{C}$. $\|(X - z)P\|_2 = \|(\overline{z}X - 1)P\|_2$.*

On peut écrire $P = \sum_i a_i X^i$ et poser $a_{-1} = 0$. Le coefficient de $X^i$ dans $(X - z)P$ (resp. $(\overline{z}X - 1)P$) est $a_{i-1} - z a_i$ (resp. $\overline{z} a_{i-1} - a_i$).

On en déduit d'une part :

$$\begin{aligned}
\|(X - z)P\|_2^2 &= \sum_i [(a_{i-1} - z a_i)(\overline{a_{i-1}} - \overline{z a_i})] \\
&= \sum_i |a_{i-1}|^2 + |z|^2 \sum_i |a_i|^2 - z \sum_i a_i \overline{a_{i-1}} - \overline{z} \sum_i a_{i-1} \overline{a_i} \\
&= (1 + |z|^2) \|P\|_2^2 - z \sum_i a_i \overline{a_{i-1}} - \overline{z} \sum_i a_{i-1} \overline{a_i} \ ,
\end{aligned}$$

et d'autre part :

$$\begin{aligned}
\|(\overline{z}X - 1)P\|_2^2 &= \sum_i [(\overline{z} a_{i-1} - a_i)(z \overline{a_{i-1}} - \overline{a_i})] \\
&= |z|^2 \sum_i |a_{i-1}|^2 + \sum_i |a_i|^2 - z \sum_i a_i \overline{a_{i-1}} - \overline{z} \sum_i a_{i-1} \overline{a_i} \\
&= (1 + |z|^2) \|P\|_2^2 - z \sum_i a_i \overline{a_{i-1}} - \overline{z} \sum_i a_{i-1} \overline{a_i} \ ,
\end{aligned}$$

ce qui permet de conclure.

Démontrons le théorème de Landau. On peut supposer que $P$ est unitaire, et écrire $P = (X - z_1) \dots (X - z_k)(X - u_1) \dots (X - u_l)$ où les racines complexes de $P$ vérifient $|z_i| > 1$ et $|u_i| \leq 1$. Notons $Q = (X - z_1) \dots (X - z_k)(\overline{u_1}X - 1) \dots (\overline{u_l}X - 1)$. En appliquant $l$ fois le lemme 8.12, on obtient que $\|P\|_2 = \|Q\|_2$. On peut conclure en utilisant l'inégalité évidente $\|Q\|_2 \geq |Q(0)| = |z_1 \dots z_k| = M(P)$.

Démontrons le théorème de Mignotte. On peut supposer que $P$ et $Q$ sont unitaires. L'inégalité $\|Q\|_\infty \leq \|Q\|_1$ est évidente. Montrons maintenant l'inégalité $M(Q) \leq \|P\|_2$. Notons $R$ le polynôme unitaire tel que $P = QR$. On a évidemment $M(R) \geq 1$, d'où $M(P) = M(Q)M(R) \geq M(Q)$. D'après le théorème de Landau, il vient $M(Q) \leq M(P) \leq \|P\|_2$. Il reste à établir l'inégalité

$\|Q\|_1 \leq 2^m M(Q)$. On peut écrire $Q = (X - z_1)\dots(X - z_m) = \sum_i b_i X^i$. Pour tout $0 \leq i \leq m$, on a

$$(-1)^i b_{m-i} = \sum_{J \subset \{1,\dots,m\}, \#J = i} \prod_{j \in J} z_j \ .$$

Il en résulte que $|b_{m-i}| \leq \begin{pmatrix} m \\ i \end{pmatrix} M(Q)$, d'où

$$\|Q\|_1 = \sum_i |b_{m-i}| \leq \left( \sum_i \begin{pmatrix} m \\ i \end{pmatrix} \right) M(Q) = 2^m M(Q) \ .$$

### 8.3. Factorisation dans $\mathbf{Z}[X]$ et $\mathbf{Q}[X]$

On commence par un algorithme de factorisation pour des polynômes unitaires à coefficients entiers et sans facteur carré (autrement dit, en tant qu'élément de $\mathbf{Q}[X]$, $P$ est séparable, ou encore $\operatorname{Disc} P \neq 0$).

***Algorithme de factorisation dans $\mathbf{Z}[X]$ 8.13.*** — *Soit $P \in \mathbf{Z}[X]$. On suppose que $P$ est unitaire et que $\operatorname{Disc} P \neq 0$. On va déterminer la factorisation de $P$ dans $\mathbf{Z}[X]$ en produits de polynômes unitaires irréductibles.*

*On choisit un nombre premier $p$ ne divisant pas $\operatorname{Disc} P$. On détermine un entier $M \geq 1$ tel que $M \geq 2^{\deg P} \|P\|_2$, puis un entier $d \geq 1$ tel que $p^d \geq 2M + 1$.*

*On détermine la factorisation $Q_1 \dots Q_k$ dans $\mathbf{F}_p[X]$ de la réduction modulo $p$ de $P$ en utilisant l'algorithme de Berlekamp (les $Q_i \in \mathbf{F}_p[X]$ sont irréductibles et unitaires).*

*En utilisant le lemme de Hensel, on détermine $\tilde{Q}_1, \dots, \tilde{Q}_k \in \mathbf{Z}/p^d\mathbf{Z}[X]$ unitaires se réduisant sur les $Q_i$ modulo $p$ et tels que si on note $\tilde{P}$ la réduction de $P$ modulo $p^d$, on ait $\tilde{P} = \tilde{Q}_1 \dots \tilde{Q}_k$.*

*Pour toute partie $I \subset \{1, \dots, k\}$ de cardinal compris entre 1 et $k-1$, on calcule $\tilde{Q}_I = \prod_{i \in I} \tilde{Q}_i \in \mathbf{Z}/p^d\mathbf{Z}[X]$. On note $Q_I \in \mathbf{Z}[X]$ un représentant (unitaire) de $\tilde{Q}_I$ tel que $\|Q_I\|_\infty \leq \left\lfloor \frac{p^d}{2} \right\rfloor$. On teste si $Q_I$ divise $P$, si oui, on a déterminé un facteur non trivial de $P$ dans $\mathbf{Z}[X]$ (et si aucun $Q_J$ pour $J$ contenu strictement dans $I$ ne vérifie déjà cela, alors $Q_I$ est irréductible et on peut continuer à examiner les parties de $\{1, \dots, k\}$ contenues dans le complémentaire de $I$ pour déterminer les autres facteurs irréductibles).*

*Si aucun polynôme $Q_I$ ne fournit un facteur non trivial de $P$, alors $P \in \mathbf{Z}[X]$ est irréductible.*

Remarquons que l'hypothèse selon laquelle $p$ ne divise pas $\operatorname{Disc} P$ fait que la réduction de $P$ modulo $p$ est un polynôme séparable ; l'algorithme de Berlekamp s'applique à lui et les hypothèses du lemme de Hensel seront aussi satisfaites. Manifestement, si l'algorithme décrit ci-dessus termine en renvoyant un facteur non trivial de $P$ dans $\mathbf{Z}[X]$, ce résultat est juste. Pour conclure, il faut montrer inversement que si $Q \in \mathbf{Z}[X]$ est un diviseur non trivial (que l'on peut supposer unitaire) de $P$, alors il existe une partie $I$ telle que $Q = Q_I$. La borne de Mignotte nous dit que si on tel polynôme $Q$ existe, alors ses coefficients sont plus petits que $M$ en valeur absolue ; en particulier, il suffit de connaître sa réduction $\tilde{Q}$ dans $\mathbf{Z}/p^d\mathbf{Z}[X]$. Si on réduit $\tilde{Q}$ modulo $p$, on obtient nécessairement un sous-produit $\prod_{i \in I} Q_i$ de la factorisation $Q_1 \dots Q_k$ de $P$ modulo $p$ pour une partie $I$ de $\{1, \dots, k\}$ (de cardinal différent de 0 et $k$). D'après l'unicité du résultat dans le lemme de Hensel, on doit avoir $\tilde{Q} = \prod_{i \in I} \tilde{Q}_i = \tilde{Q}_I$.

Il reste à expliquer comment utiliser l'algorithme 8.13 pour factoriser dans $\mathbf{Z}[X]$ et $\mathbf{Q}[X]$. Tout d'abord, pourvu que l'on sache factoriser dans $\mathbf{Z}$, le problème de la factorisation dans $\mathbf{Z}[X]$ et dans $\mathbf{Q}[X]$ sont essentiellement équivalents. Comme dans le cas des polynômes à coefficients dans $\mathbf{F}_q[X]$, pour factoriser dans $\mathbf{Q}[X]$, on peut supposer que le polynôme est sans facteurs carrés (c'est-à-dire séparable), autrement dit que le discriminant du polynôme est non nul. On s'est ramené à la factorisation de $P \in \mathbf{Q}[X]$ unitaire et séparable. Pour appliquer l'algorithme 8.13, il reste à montrer que l'on peut supposer que $P \in \mathbf{Z}[X]$. On utilise le lemme suivant :

***Lemme 8.14***. — *Soit $P \in \mathbf{Q}[X]$ un polynôme unitaire. Il existe un entier $n \geq 1$ tel que $n^{\deg P} P(\frac{1}{n}X)$ appartienne à $\mathbf{Z}[X]$ et soit unitaire.*

Si $P = \sum_i a_i X^i$ et $d = \deg P$, on obtient $n^d P(\frac{1}{n}X) = \sum_i a_i n^{d-i} X^i = X^d + n a_{d-1} X^{d-1} + n^2 a_{d-2} X^{d-2} + \cdots + n^d a_0$. Il est évident que l'on peut trouver un entier $n \geq 1$ convenable.

Le problème de la factorisation dans $\mathbf{Q}[X]$ de $P$ et de $P(\frac{1}{n}X)$ étant équivalents, on peut effectivement se ramener à la situation de l'algorithme 8.13.

# CHAPITRE 9

# TRANSFORMÉE DE FOURIER (DISCRÈTE ET RAPIDE)

Le but de ce chapitre est d'expliquer comment calculer efficacement le produit de deux polynômes ou de deux entiers.

## 9.1. Racines de l'unité

***Définition 9.1***. — Soit $n \geq 1$. Soit $A$ un anneau dans lequel $n$ est inversible. Soit $\omega \in A$. On dit que $\omega$ est une racine primitive $n$-ième de l'unité si $\omega^n = 1$ et que pour tout $1 \leq i \leq n-1$, $1 - \omega^i$ soit un élément inversible de $A$.

Pourvu que $A$ ne soit pas l'anneau nul, si $\omega$ est une racine primitive $n$-ième de l'unité dans $A$, alors le sous-groupe de $A^\times$ engendré par $\omega$ est d'ordre $n$ ; la réciproque est fausse en général, mais vraie si $A$ est un corps.

Si $\omega$ est une racine primitive $n$-ième de l'unité, alors $\omega^i - \omega^j$ est inversible si $i \not\equiv j \mod n$.

***Proposition 9.2***. — *Soit $A$ un anneau. Soit $n \geq 2$. Soit $\omega$ une racine primitive $n$-ième de l'unité.*

(i) *Pour tout diviseur $m$ de $n$, $\omega^{\frac{n}{m}}$ est une racine primitive $m$-ième de l'unité dans $A$.*

(ii) *Pour $j \not\equiv 0 \mod n$, $\sum_{i=0}^{n-1} \omega^{ji} = 0$.*

La propriété (i) est triviale. La propriété (ii) résulte de la relation

$$(1 - \omega^j) \sum_{i=0}^{n-1} \omega^{ji} = 1 - \omega^{nj} = 0 \, ,$$

et du fait que $1 - \omega^j$ soit inversible.

***Remarque 9.3***. — Si $A = \mathbf{C}$, $\omega = e^{\frac{2i\pi}{n}}$ est une racine primitive $n$-ième de l'unité, de même que $\omega^{-1} = \overline{\omega}$ ; plus généralement, les racines primitives $n$-ièmes sont les $\omega^i$ avec $i \in \mathbf{Z}/n\mathbf{Z}^\times$. Si $A = \mathbf{F}_q$ est un corps fini à $q$ éléments (et que la caractéristique de $\mathbf{F}_q$ ne divise pas $n$), alors $A$ admet une racine primitive $n$-ième de l'unité si et seulement si $n$ divise $q - 1$.

***Proposition 9.4***. — *Soit $A$ un anneau dans lequel 2 est inversible. Soit $k \geq 0$. On pose $n = 2^{k+1}$. Soit $\omega \in A$. Alors, $\omega$ est une racine primitive $n$-ième de l'unité si et seulement si $\omega^{\frac{n}{2}} = -1$.*

Montrons le résultat pour $k = 0$. Soit $\omega \in A$ une racine primitive deuxième de l'unité. On a $\omega^2 = 1$ et $1 - \omega$ inversible. La relation $\omega^2 = 1$ se traduisant par $(1 + \omega)(1 - \omega) = 0$, l'inversibilité de $1 - \omega$ conduit à $\omega = -1$. Réciproquement, $-1$ est bien une racine primitive deuxième de l'unité puisque l'on a supposé 2 inversible.

Soit $k \geq 0$. Si $\omega$ est une racine primitive $2^{k+1}$-ième de l'unité, d'après la proposition 9.2, $\omega^{2^k}$ est une racine primitive deuxième de l'unité. D'après le cas précédent, on a bien $\omega^{2^k} = -1$. Montrons la réciproque par récurrence sur $k$. Le cas $k = 0$ a déjà été traité. Supposons $k \geq 1$ et $\omega^{2^k} = -1$. Posons $\omega' = \omega^2$. On a $\omega'^{2^{k-1}} = -1$. Par hypothèse de récurrence, $\omega'$ est une racine primitive $2^k$-ième de l'unité. Posons $m = 2^k$. Le fait que $\omega'$ soit une racine primitive $m$-ième de l'unité signifie que

$1-\omega'^i$ est inversible pour $1 \leq i \leq m-1$. Observons que $1-\omega'^i = (1-\omega^i)(1+\omega^i) = (1-\omega^i)(1-\omega^{m+i})$. Ainsi, $1-\omega^i$ et $1-\omega^{m+i}$ sont inversibles pour $1 \leq i \leq m-1$. Comme $1-\omega^m = 2$ est inversible, on obtient bien ainsi que pour tout $1 \leq j \leq 2m-1$, $1-\omega^j$ est inversible, c'est-à-dire que $\omega$ est une racine primitive $2^{k+1}$-ième de l'unité.

***Remarque 9.5***. — Soit $A$ un anneau dans lequel 2 est inversible. Si $n = 2^{k+1}$ avec $k \geq 0$, alors la classe de $X$ dans l'anneau quotient $A[X]/(X^{\frac{n}{2}}+1)$ est une racine primitive $n$-ième de l'unité.

## 9.2. Transformée de Fourier discrète

On fixe un entier $n \geq 1$, $\omega$ une racine primitive $n$-ième de l'unité dans un anneau $A$.

***Définition 9.6***. — La transformée de Fourier discrète est l'application $A$-linéaire $\mathrm{TFD}_\omega \colon A[X]_{<n} \to A^n$ qui à un polynôme $P$ de degré au plus $n-1$ associe le $n$-uplet $(P(1), P(\omega), \ldots, P(\omega^{n-1}))$.

***Remarque 9.7***. — Les morphismes évaluations en $1, \omega, \ldots, \omega^{n-1}$ sont définis sur tout $A[X]$. L'idéal engendré par $X^n-1$ est bien sûr dans le noyau de ces morphismes d'évaluations. Comme on a un isomorphisme canonique $A[X]_{<n} \simeq A[X]/(X^n-1)$, on peut considérer la transformée de Fourier discrète comme un morphisme $\mathrm{TFD}_\omega \colon A[X]/(X^n-1) \to A^n$.

***Théorème 9.8***. — *On identifie canoniquement $A[X]_{<n}$ à $A^n$ faisant correspondre $(a_0, \ldots, a_{n-1})$ à $a_0 + a_1X + \cdots + a_{n-1}X^{n-1}$. On peut donc considérer les transformée de Fourier discrète $\mathrm{TFD}_\omega$ et $\mathrm{TFD}_{\omega^{-1}}$ à des endomorphismes de $A^n$. On a alors, $\mathrm{TFD}_{\omega^{-1}} \circ \mathrm{TFD}_\omega = n\,\mathrm{id}_{A^n}$, autrement dit la transformée de Fourier discrète $\mathrm{TFD}_\omega$ est un isomorphisme et son inverse $\mathrm{TFD}_\omega^{-1}$ est $\frac{1}{n}\mathrm{TFD}_{\omega^{-1}}$.*

Notons $V_\omega$ (resp. $V_{\omega^{-1}}$) la matrice de $\mathrm{TFD}_\omega$ (resp. $\mathrm{TFD}_{\omega^{-1}}$) dans les bases canoniques. On a

$$V_\omega = \begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega & \omega^2 & \ldots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \ldots & \omega^{2(n-1)} \\ 1 & \vdots & \vdots & \ldots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \ldots & \omega^{(n-1)(n-1)} \end{pmatrix}.$$

Le terme général est $\omega^{ij}$ pour $0 \leq i, j \leq n-1$.

Il vient aussitôt que $V_{\omega^{-1}} \cdot V_\omega = (\sigma_{i,j})_{0 \leq i,j \leq n-1}$ avec

$$\sigma_{i,j} = \sum_{k=0}^{n-1} \omega^{-ki}\omega^{kj}\ .$$

Si $i = j$, on a évidemment $\sigma_{i,j} = n$. Si $i \neq j$, on a $\sigma_{i,j} = 0$ d'après la proposition 9.2.

***Corollaire 9.9***. — *Si $P$ et $Q$ sont des polynômes à coefficients dans $A$, alors*

$$\mathrm{TFD}_\omega(PQ) = \mathrm{TFD}_\omega(P) \cdot \mathrm{TFD}_\omega(Q)\ ,$$

*où l'on a noté $\cdot$ le produit terme à terme dans $A^n$. Si le degré de $PQ$ est strictement inférieur à $n$, on a*

$$PQ = \frac{1}{n}\,\mathrm{TFD}_{\omega^{-1}}(\mathrm{TFD}_\omega(P) \cdot \mathrm{TFD}_\omega(Q))\ ,$$

*avec l'abus de notation signalé plus haut.*

Multiplier terme à terme des éléments de $A^n$ nécessite uniquement $n$ multiplications dans $A$. Ainsi, si on arrive à calculer des transformées de Fourier discrète plus efficacement qu'en $O(n^2)$ opérations, on obtiendra un algorithme de multiplication de polynômes plus efficace que l'algorithme naïf.

## 9.3. Transformée de Fourier rapide (FFT)

On présente ici un algorithme permettant de calculer efficacement la transformée discrète dans le cas où $n$ est une puissance de deux.

On suppose que $n = 2^k$. Notons $\omega$ une racine $n$-ième primitive de l'unité. Pour tout diviseur $d$ de $n$, on note $\omega_d = \omega^{\frac{n}{d}}$ : c'est une racine primitive $d$-ième de l'unité. Si $k = 0$, il n'y a essentiellement rien à faire. Pour expliquer le principe de l'algorithme, on suppose $k \geq 1$, de sorte que l'on peut écrire $n = 2m$. On va montrer comment calculer $\mathrm{TFD}_{\omega_n}$ en combinant le résultat de deux appels à $\mathrm{TFD}_{\omega_m}$.

Soit $P \in A[X]_{<n}$, $P = \sum_{i=0}^{n-1} a_i X^i$. On pose $Q_0 = \sum_{j=0}^{m-1} a_{2j} X^j$ et $Q_1 = \sum_{j=0}^{m-1} a_{2j+1} X^j$. On a la relation $P(X) = Q_0(X^2) + XQ_1(X^2)$. Ainsi, pour tout $0 \leq i \leq n-1$, on a :

$$P(\omega_n^i) = Q_0(\omega_m^i) + \omega_n^i Q_1(\omega_m^i) \ .$$

En utilisant le fait que $\omega_n^m = -1$, on peut écrire deux relations pour tout $0 \leq i \leq m-1$ :

$$P(\omega_n^i) = Q_0(\omega_m^i) + \omega_n^i Q_1(\omega_m^i) \qquad P(\omega_n^{m+i}) = Q_0(\omega_m^i) - \omega_n^i Q_1(\omega_m^i) \ .$$

Calculer les valeurs $Q_0(\omega_m^i)$ et $Q_1(\omega_m^i)$ pour $0 \leq i \leq m-1$ revient à calculer $\mathrm{TFD}_{\omega_m}(Q_0)$ et $\mathrm{TFD}_{\omega_m}(Q_1)$. Une fois ce calcul fait, on peut utiliser les relations ci-dessus pour obtenir $\mathrm{TFD}_{\omega}(P)$.

Une manière efficace de procéder à cette combinaison des deux résultats de $\mathrm{TFD}_{\omega_m}$ consiste à stocker dans deux tableaux $t$ et $t'$ de taille $m$ (indexés par les entiers compris entre 0 et $m-1$) les résultats respectifs de $\mathrm{TFD}_{\omega_m}(Q_0)$ et $\mathrm{TFD}_{\omega_m}(Q_1)$ et de remplir de la manière suivante un nouveau tableau $r$ de taille $n$ contenant le résultat $\mathrm{TFD}_{\omega}(P)$. On initialise une variable $x$ à 1. Pour $i$ allant de 0 jusqu'à $m-1$, on calcule $u := x \cdot t'_i$, on fait $r_i := t_i + u$ et $r_{m+i} := t_i - u$ et on fait $x := \omega \cdot x$.

On peut appliquer l'algorithme ci-dessus récursivement. Notons $C_k$ le coût en nombres d'opérations (addition, multiplication) dans $A$ de l'algorithme de transformée de Fourier rapide pour des polynômes de degré $< 2^k$. On a manifestement une relation de récurrence :

$$C_k = 2C_{k-1} + O(2^k) \ .$$

Il en résulte que $C_k = O(k2^k)$, autrement dit le coût est en $O(n \log_2 n)$.

## 9.4. Multiplication rapide

D'après le corollaire 9.9, on peut utiliser la transformée de Fourier discrète pour calculer un produit de polynômes de degré $\leq n$ à coefficients dans un anneau $A$ en $n \log_2 n$ opérations dans cet anneau, pourvu que l'on dispose de racines de l'unités. Plus précisément, on a :

***Proposition 9.10***. — *Soit $A$ un anneau muni d'une racine $n$-ième de l'unité $\omega$, avec $n$ une puissance de deux. On peut calculer une multiplication dans $A[X]/(X^n - 1)$ en utilisant $n$ multiplications dans $A$ et $O(n \log_2 n)$ additions dans $A$ ou multiplications dans $A$ par une puissance de $\omega$.*

***Corollaire 9.11***. — *Soit $A$ un anneau muni d'une racine primitive $2n$-ième de l'unité $\omega$, avec $n$ une puissance de deux. On peut calculer une multiplication dans $A[U]/(U^n + 1)$ en utilisant $n$ multiplications dans $A$ et $O(n \log_2 n)$ additions dans $A$ ou multiplications dans $A$ par une puissance de $\omega$.*

Soit $\omega$ une racine primitive $2n$-ième de l'unité. $\omega^2$ est une racine primitive $n$-ième de l'unité et on a $\omega^n = -1$. D'après la proposition, on sait calculer le produit dans $A[X]/(X^n - 1)$ en $O(n \log_2 n)$ opérations d'addition ou de multiplication par une puissance de $\omega$ et en $n$ multiplications dans $A$. Par ailleurs, on a un isomorphisme de $A$-algèbres $A[X]/(X^n - 1) \xrightarrow{\sim} A[U]/(U^n + 1)$ envoyant $X$ sur $\omega U$ et l'isomorphisme inverse $U$ sur $\omega^{-1}X$. Pour multiplier les polynômes en $U$ modulo $U^n + 1$, on passer à la représentation comme polynôme en $X$ en substituant $U := \omega^{-1}X$, multiplier ces polynômes en $X$, puis revenir à la représentation comme polynôme en $U$ en substituant $X := \omega U$. Le surcoût par rapport à l'algorithme précédent n'est qu'un $O(n)$ de multiplications par des puissances de $\omega$.

***Théorème 9.12.*** — *Soit $A$ un anneau dans lequel $2$ est inversible. On peut calculer des produits dans $A[X]/(X^n+1)$ avec $n=2^k$ en $O(n\log_2 n\log_2\log_2 n)$ opérations dans $A$.*

Si $k\leq 2$, on utilise un algorithme trivial. Sinon, on pose $m=2^{\lfloor \frac{k}{2}\rfloor}$, $t=2^{\lceil \frac{k}{2}\rceil}$ ($n=mt$). On se donne $f$ et $g$ deux éléments de $A[X]_{<n}$. On peut trouver d'uniques éléments $f',g'$ des $A[X,Y]$ tels que $f'(X,X^m)=f$, $g'(X,X^m)=g$ et que $f'$ et $g'$ soient de degré $<m$ par rapport à $X$. On considère les images $f^\star$ et $g^\star$ de $f'$ et $g'$ dans l'anneau quotient $B[Y]/(Y^t+1)$ avec $B=A[X]/(X^{2m}+1)$.

On applique l'algorithme du corollaire 9.11 pour déterminer le produit $h^\star=f^\star\cdot g^\star\in B[Y]/(Y^t+1)$, les opérations de multiplications dans $B$ étant réalisées en appliquant récursivement l'algorithme que l'on décrit ici. Il faut préciser la racine $2t$-ième primitive $\omega$ de l'unité de $B$ que l'on utilise : si $k$ est pair, $m=t$, la classe de $X^2$ est une racine primitive $2t$-ième de l'unité dans $B$ (car $\left(X^2\right)^t=X^{2m}=-1$ dans $B$), si $k$ est impair, $t=2m$, la classe de $X$ est une racine primitive $2t$-ième de l'unité dans $B$ (car $X^t=X^{2m}=-1$ dans $B$). On connaît alors le produit $h^\star$ de $f^\star$ et $g^\star$ dans $A[X]/(X^{2m}+1)[Y]/(Y^t+1)$. Notons $h'$ l'unique élément de $A[X][Y]/(Y^t+1)$ de degré $<2m$ par rapport à $X$ relevant $h^\star$. Le produit $fg$ cherché est le reste de la division euclidienne de $h'(X,X^m)$ par $X^n+1$.

En effet, comme les images de $f'$ et $g'$ dans $A[X,Y]/(Y^t+1)$ sont de degrés $<m$ par rapport à $X$, l'image de $f'g'$ dans $A[X,Y]/(Y^t+1)$ est de degré $<2m$ par rapport à $X$ ; les images de $h'$ et $f'g'$ dans $A[X,Y]/(Y^t+1)$ coïncident, puisqu'ils coïncident modulo $X^{2m}+1$. Au lieu de réduire modulo $X^{2m}+1$, réduisons maintenant par rapport à l'idéal engendré par $Y-X^m$. Comme on a un isomorphisme évident d'anneaux $A[X]/(X^n+1)\simeq A[X,Y]/(Y^t+1,Y-X^m)$, on obtient une certaine égalité dans $A[X]/(X^n+1)$, précisément entre $h'(X,X^m)$ et $f'(X,X^m)g'(X,X^m)=fg$.

Etudions le coût de cet algorithme. Notons $C_k$ le coût de la multiplication dans $A[X]/(X^{2^k}+1)$ en opérations arithmétiques (addition ou multiplication) dans $A$. Faisons plusieurs observations :

- Le calcul de $f',g'$ puis $f^\star,g^\star$ à partir de $f$ et $g$ et celui de $h'$ en fonction de $h^\star$ se fait évidemment en $O(n)$ opérations sur les éléments de $A$.
- Le calcul de $fg$ à partir de $h'$ se fait aussi en $O(n)$ opérations sur les éléments de $A$. Il s'agit en effet de calculer le reste de la division euclidienne par $X^n+1$ d'un polynôme de degré $<2n$ à coefficients dans $A$ : si $h'=\sum_{i=0}^{2n-1}a_iX^i$, le reste est $\sum_{i=0}^{n-1}(a_i-a_{n+i})X^i$.
- L'estimation qui reste à faire est celle du coût de la multiplication de $f^\star$ et de $g^\star$ dans $B[Y]/(Y^t+1)$. On utilise pour cela le corollaire 9.11 : on utilise $t$ multiplications dans $B$ et $O(t\log_2 t)$ additions dans $B$ ou multiplications par des puissances de $\omega$.
- Une addition dans $B$ utilise $2m$ additions dans $A$. Une multiplication dans $B$ par une puissance de $\omega$ se fait en $O(m)$ opérations dans $A$. Le coût des $O(t\log_2 t)$ opérations de ces deux types utilisées ci-dessus est donc un $O(mt\log_2 t)=O(2^kk)$ opérations dans $A$.

On obtient finalement une relation de récurrence :

$$C_k\leq 2^{\lceil \frac{k}{2}\rceil}C_{\lfloor \frac{k}{2}\rfloor+1}+O(2^kk)\ .$$

Notons $S_k=\frac{C_k}{2^k(k-2)}$. On obtient :

$$2^k(k-2)S_k\leq 2^{\lceil \frac{k}{2}\rceil}2^{\lfloor \frac{k}{2}\rfloor+1}\left(\frac{k}{2}-1\right)S_{\lfloor \frac{k}{2}\rfloor+1}+O(2^kk)\ .$$

Divisons par $2^k(k-2)$ (on suppose $k\geq 3$) :

$$S_k\leq S_{\lfloor \frac{k}{2}\rfloor+1}+O(1)\ .$$

On en déduit que $S_k=O(\log_2 k)$, et donc $C_k=O(2^kk\log_2 k)$. Comme $n=2^k$, on a bien un coût en $O(n\log_2 n\log_2\log_2 n)$ opérations dans $A$.

***Théorème 9.13*** **(Schönhage-Strassen, 1971)**. — *On peut multiplier des entiers à $n$ bits en $O(n\log_2 n\log_2\log_2 n)$ opérations élémentaires sur les bits.*

On va présenter un autre algorithme utilisant aussi la FFT. Il ne permet pas de multiplier des entiers de taille arbitraire, mais possède un domaine de validité suffisant pour les multiplications

susceptibles d'être réalisées compte tenu des capacités de stockage actuelles. Il présente aussi l'avantage d'utiliser une représentation des nombres entiers par blocs de 64 bits, qui est la taille des registres sur certains processeurs utilisés couramment de nos jours.

Soit $s \leq 61$ un entier. On va décrire un algorithme pour multiplier des entiers naturels $a$ et $b$ tels que $ab < 2^{64 \cdot 2^s}$. On précalcule une fois pour toutes trois nombres premiers impairs $p_1, p_2, p_3$ distincts et compris entre $2^{63}$ et $2^{64}$ et des racines primitives $2^s$-ièmes $\omega_i$ dans $\mathbf{F}_{p_i}$ (l'existence de ces racines primitives revient à demander que les $p_i$ soient congrus à 1 modulo $2^s$).

On se donne deux entiers naturels $n$ et $n'$, représentés chacun sous une forme $\cdots + 2^{2 \cdot 64} a_2 + 2^{64} a_1 + a_0$ avec $0 \leq a_i < 2^{64} - 1$, autrement dit $P(2^{64})$ et $P'(2^{64})$ où $P$ et $P'$ sont des polynômes à coefficients entiers compris entre 0 et $2^{64} - 1$. On suppose que la somme des degrés de $P$ et $P'$ est $< 2^s$ (et on a intérêt à choisir $s$ le plus petit possible).

Pour $i \in \{1, 2, 3\}$, on calcule l'image du produit $PP'$ dans $\mathbf{F}_{p_i}$ en utilisant l'algorithme de la proposition 9.10. En utilisant le lemme chinois (on a intérêt à précalculer aussi les coefficients appropriés des relations de Bézout), on détermine l'unique polynôme $Q \in \mathbf{Z}[X]$ à coefficients compris entre 0 et $p_1 p_2 p_3 - 1$ tel que $Q \equiv PP' \mod p_1 p_2 p_3$.

On a alors $Q = PP'$. En effet, les coefficients de $PP'$ sont des entiers naturels $\leq 2^s 2^{2 \cdot 64} \leq 2^{61 + 2 \cdot 64} = 2^{3 \cdot 63} < p_1 p_2 p_3$.

Le résultat $nn'$ est $Q(2^{64})$, dont on peut déterminer rapidement la suite des blocs de 64 bits à partir des coefficients de $Q$.