Chapitre 4
Arithmétique entière et polynomiale

Dans ce chapitre, nous énonçons et démontrons les propriétés arithmétiques élémentaires partagées par l'anneau des entiers et l'anneau des polynômes à une indéterminée $\mathbb{K}[X]$ (où $\mathbb{K}$ est un corps commutatif). Afin de donner une présentation unifiée, il est commode de mettre en œuvre un certain nombre de notions générales concernant la divisibilité et la factorisation dans des anneaux commutatifs unitaires quelconques, en particulier la notion d'idéal d'un anneau, introduite en 1921 par Emmy Noether (éminente mathématicienne allemande, 1882–1935).

1. Généralités sur les anneaux et la divisibilité

1.1. Définitions principales

1.1.1. Définition. *Un anneau est un triplet $(A, +, \times)$ formé d'un ensemble A et de deux lois de composition interne*

$$\begin{aligned} + &: A \times A \to A, \quad (x, y) \mapsto x + y, \\ \times &: A \times A \to A, \quad (x, y) \mapsto x \times y \end{aligned}$$

ayant les propriétés suivantes :

(a) *$(A, +)$ est une groupe commutatif, c'est-à-dire que la loi $+$ appelée addition est associative, commutative, dotée d'un élément neutre 0_A, et que tout élément x possède un symétrique x', en sorte que $x + x' = 0_A$ (on le note $x' = -x$ et on l'appelle opposé de x)* ;

(b) *la multiplication $\times$ est associative et distributive par rapport à $+$, et possède un élément neutre noté 1_A.*

1.1.2. Remarque. On omettra assez souvent l'indice A dans l'écriture de 0_A et 1_A. Nous avons ici supposé l'existence d'un élément neutre 1_A pour la multiplication, mais certains auteurs n'incluent pas cet axiome et parlent alors d'*anneau unitaire*. Il est à noter qu'on ne suppose pas nécessairement $1_A \neq 0_A$, par exemple $A = \{0\}$ est bien un anneau et 0 y est élément neutre à la fois pour l'addition et la multiplication. En fait, comme les axiomes impliquent $0_A \times x = 0_A$ et $1_A \times x = x$, on voit que l'on a $1_A = 0_A$ si et seulement si l'anneau est réduit à $\{0_A\}$, donc ce cas est unique et très "dégénéré". Un anneau A est appelé un *corps* si $1_A \neq 0_A$ et si tout élément $x \neq 0_A$ possède un inverse x' pour la multplication, c'est-à-dire tel que $x \times x' = x' \times x = 1_A$. Un corps $\mathbb{K}$ possède donc au moins deux éléments $0_{\mathbb{K}}$ et $1_{\mathbb{K}}$. Il peut fort bien se réduire à ces deux éléments, c'est le cas

du corps noté $\mathbb{F}_2 = \{0, 1\}$ (ou encore $\mathbb{Z}/2\mathbb{Z}$, voir plus loin), dans lequel on prend $1 + 1 = 0$.

1.1.3. Définition. *Un morphisme d'anneaux $\varphi : A \to B$ entre deux anneaux A et B (non nécessairement commutatifs) est une application possédant les trois propriétés suivantes :*

(a) $\forall x, y \in A,\ \varphi(x + y) = \varphi(x) + \varphi(y)$;

(b) $\forall x, y \in A,\ \varphi(x \times y) = \varphi(x) \times \varphi(y)$;

(c) $\varphi(1_A) = 1_B$.

Il résulte de (a) que φ est en particulier un morphisme de groupes additifs de $(A, +)$ dans $(B, +)$, ce qui entraîne que l'on a aussi $\varphi(0_A) = 0_B$ [prendre $x = y = 0_A$]. L'axiome (c) sert à éviter l'application inintéressante $\varphi = 0$ et d'autres pathologies.

Dans la suite de ce chapitre, mis à part l'exemple 1.1.4 (e) ci-dessous, on s'intéressera exclusivement à des anneaux $(A, +, \times)$ commutatifs et à des corps $(\mathbb{K}, +, \times)$ commutatifs, c'est-à-dire tels que la multiplication $\times$ soit *commutative*. Pour simplifier les notations, on omettra souvent l'écriture des lois, et on parlera d'un anneau A et d'un corps $\mathbb{K}$ (qui seront implicitement supposés commutatifs, sauf mention explicite du contraire).

1.1.4. Exemples d'anneaux. Ce chapitre sera presque exclusivement consacré aux deux exemples fondamentaux (a) et (b) ci-dessous et à leurs propriétés arithmétiques.

(a) $(\mathbb{Z}, +, \times)$, l'anneau des entiers relatifs.

(b) $(\mathbb{K}[X], +, \times)$, l'anneau des polynômes à coefficients dans un corps $\mathbb{K}$. Un polynôme est une expression formelle $P = \sum_{i=0}^{d} a_i X^i$, $a_i \in \mathbb{K}$, qui peut être codée comme une somme $P = \sum_{i \in \mathbb{N}} a_i X^i$ possédant seulement un nombre fini de coefficients a_i non nuls, ou encore, si on veut, comme une suite infinie de coefficients $(a_0, a_1, \ldots, a_n, \ldots)$ presque tous nuls. Étant donné

$$P = \sum_{i \in \mathbb{N}} a_i X^i, \quad Q = \sum_{i \in \mathbb{N}} b_i X^i,$$

l'addition et la multiplication sont définies par

$$P + Q = \sum_{i \in \mathbb{N}} (a_i + b_i) X^i, \quad P \times Q = \sum_{k \in \mathbb{N}} \left(\sum_{i+j=k} a_i b_j \right) X^k.$$

Au polynôme $P = \sum_{i=0}^{d} a_i X^i$, $a_i \in \mathbb{K}$, on peut associer la fonction polynomiale

$$f_P : \mathbb{K} \to \mathbb{K}, \quad x \mapsto f_P(x) = \sum_{i=0}^{d} a_i x^i,$$

mais il convient de distinguer soigneusement P de f_P : P n'est pas une fonction, mais un objet "formel", élément d'un espace vectoriel de dimension

infinie sur $\mathbb{K}$, dont $(X^i)_{i\in\mathbb{N}}$ est une base. Par exemple, si on prend le corps à 2 éléments $\mathbb{K} = \mathbb{F}_2 = \{0, 1\}$, tous les polynômes $P_i = X^i$, $i \geqslant 1$, fournissent la même fonction polynomiale $f_i : \mathbb{K} \to \mathbb{K}$, $x \mapsto f_i(x) = x^i$ telle que $f_i(0) = 0$ et $f_i(1) = 1$. Par ailleurs, le polynôme non nul $P = X^2 - X = X(X - 1)$ a pour fonction polynôme associée $f_P = 0$ sur $\mathbb{F}_2$!

(c) L'anneau $A = \mathbb{Z}/6\mathbb{Z}$ (la notation sera expliquée plus loin), formé des entiers calculés modulo 6. On prend $A = \{0, 1, 2, 3, 4, 5\}$ et on calcule le résultat des opérations $+$ et $\times$ en considérant que 6 est nul (de sorte que par exemple $3 + 5 = 8 = 6 + 2 = 2$, $3 \times 5 = 15 = 6 + 6 + 3 = 3$). Ceci donne les tables de Pythagore suivantes pour les lois $+$ et $\times$ de l'anneau :

$+$	0	1	2	3	4	5
0	0	1	2	3	4	5
1	1	2	3	4	5	0
2	2	3	4	5	0	1
3	3	4	5	0	1	2
4	4	5	0	1	2	3
5	5	0	1	2	3	4

$\times$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	0	2	4
3	0	3	0	3	0	3
4	0	4	2	4	4	2
5	0	5	4	3	2	1

Dans cet anneau A "étrange", on a donc des éléments non nuls dont le produit est nul, par exemple $2 \times 3 = 0$.

(d) L'ensemble $\mathcal{F}_{\mathbb{R}}$ des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, muni de l'addition et de la multiplication des fonctions est un anneau commutatif qui a aussi la propriété "étrange" ci-dessus : le produit des fonctions non nulles $f(x) = x + |x|$ et $g(x) = x - |x|$ est la fonction $fg = 0$ (puisque $f(x)g(x) = x^2 - |x|^2 = 0$ sur $\mathbb{R}$).

(e) L'ensemble des matrices carrées $\mathcal{M}_{n\times n}(\mathbb{K})$, muni de l'addition et de la multiplication des matrices, est un anneau non commutatif si $n \geqslant 2$ (si $n = 1$, il s'identifie au corps commutatif $\mathbb{K}$). Le chapitre suivant donnera quelques lumières sur ses propriétés.

1.1.5. Exemples de morphismes d'anneaux.

(a) L'application $\varphi : \mathbb{Z} \to \mathbb{F}_2 = \{0, 1\}$ telle que $\varphi(x) = 0$ si x est pair et $\varphi(x) = 1$ si x est impair est un morphisme d'anneaux.

(b) Si $w \in \mathbb{K}$ est un élément fixé, l'application $\varphi_w : \mathbb{K}[X] \to \mathbb{K}$ qui à un polynôme P associe sa valeur $\varphi_w(P) = P(w)$ au point w est un morphisme d'anneaux.

(c) L'ensemble $\mathbb{D}$ des nombres décimaux est un anneau, et l'inclusion $\varphi : \mathbb{Z} \hookrightarrow \mathbb{D}$ est un morphisme d'anneaux.

(d) L'application $\varphi_w : \mathcal{F}_{\mathbb{R}} \to \mathbb{R}$ qui associe à une fonction $f \in \mathcal{F}_{\mathbb{R}}$ sa valeur $\varphi_w(f) = f(w)$ en un point $w \in \mathbb{R}$ est un morphisme d'anneaux.

(e) Pour tout anneau A, on a un morphisme dit *canonique*

$$\varphi_{\text{can}} : \mathbb{Z} \to A, \quad n \mapsto n_A$$

où l'on définit $n_A := 1_A + \cdots + 1_A$ (n fois) si $n \in \mathbb{N}^*$, $(-n)_A := -(n_A)$, l'élément 0_A étant ici encore l'élément neutre de A pour l'addition. L'additivité de φ_{can} est évidente, et la multiplicativité résulte de la distributivité de $\times$ par rapport à $+$ dans A. On prendra garde au fait que l'on peut avoir $n_A = 0$ même si $n \neq 0$, comme c'est le cas pour $n = 6$ dans l'exemple 1.1.4 (e) ; cependant, il est fréquent que l'on omette l'indice A et que l'on désigne par le même symbole n un élément de $\mathbb{Z}$ et son image dans A, sauf si on veut absolument éviter les confusions.

On donne un nom au phénomène apparu dans les exemples 1.1.4 (c) et (d).

1.1.6. Définition. *Soit A un anneau (commutatif, on ne le répétera plus !).*

(a) *On appelle diviseur de 0 dans A tout élément $x \neq 0$ pour lequel il existe $y \neq 0$ tel que $xy = 0$.*

(b) *L'anneau A est dit* intègre *s'il ne possède pas de diviseurs de 0, c'est-à-dire si*

$$\forall x, y \in A, \quad x \neq 0 \ \text{ et } \ y \neq 0 \implies xy \neq 0.$$

ou encore, par contraposition, si

$$\forall x, y \in A, \quad xy = 0 \implies x = 0 \ \text{ ou } \ y = 0.$$

1.1.7. Exemples d'anneaux intègres et non intègres.

(a) Comme il est bien connu, $\mathbb{Z}$ est un anneau intègre.

(b) Tout corps $\mathbb{K}$ est un anneau intègre : en effet, si $x, y \in \mathbb{K}$ sont non nuls, il possèdent des inverses x', y', et donc $x'xyy' = 1$, ce qui implique $xy \neq 0$.

(c) L'anneau $\mathbb{K}[X]$ est également intègre : en effet si on a deux polynômes non nuls $P = \sum_{j=0}^{d} a_j X^j$, $Q = \sum_{j=0}^{\delta} b_j X^j$ de termes dominants non nuls $a_d X^d$ et $b_\delta X^\delta$, alors PQ est de terme dominant non nul $a_d b_\delta X^{d+\delta}$. En particulier on voit que

$$\deg(PQ) = d + \delta = \deg(P) + \deg(Q).$$

Afin que cette égalité soit encore vraie lorsque $P = 0$ ou $Q = 0$, on convient de définir le degré du polynôme nul comme étant $-\infty$, avec la règle $-\infty + \delta = -\infty$ pour tout $\delta \in \mathbb{N} \cup \{-\infty\}$.

(d) Si A est un anneau commutatif, on peut définir l'anneau de polynômes $A[X]$ à coefficients dans A de la même manière que pour un corps, et le raisonnement du (c) implique que l'anneau $A[X]$ est intègre dès que A est lui-même intègre.

(e) Les anneaux $\mathbb{Z}/6\mathbb{Z}$ et $\mathcal{F}_{\mathbb{R}}$ définis ci-dessus ne sont pas intègres. L'anneau non commutatif $\mathcal{M}_{n\times n}(\mathbb{K})$ non plus, si $n \geqslant 2$: par exemple le carré de la matrice $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ est nul.

1.2. Notions liées à la divisibilité

1.2.1. Définitions et notations. *Soit $(A,+,\times)$ un anneau intègre non trivial.*

(a) *On notera $A^* = A\setminus\{0\}$ l'ensemble des éléments non nuls de A. L'hypothèse A intègre signifie que la multiplication est une loi de composition interne sur A^*.*

(b) *Si $x,y \in A^*$, on dit que x* divise *y, et on écrit $x \mid y$, s'il existe $\lambda \in A^*$ tel que $y = \lambda x$. Dans ce cas, on dit aussi que y est un* multiple *de x et que x est un* diviseur *de y. Par extension, on convient aussi de dire que 0 est un multiple de x (on prend $\lambda = 0$).*

Les propriétés suivantes sont à peu près évidentes.

1.2.2. Propriétés.

(a) *(transitivité) si $x,y,z \in A^*$, $x \mid y$ et $y \mid z$ entraîne $x \mid z$.*

(b) *pour $x \in A^*$ et $y,z,\alpha,\beta \in A$, si $x \mid y$ et $x \mid z$ alors $x \mid (\alpha y + \beta z)$, à moins que $\alpha y + \beta z = 0$.*

Démonstration. (a) Si $y = \lambda x$ et $z = \mu y$, alors $z = \mu(\lambda x) = (\mu\lambda)x$ dans A^*.

(b) Si $y = \lambda x$ et $z = \mu x$, alors $\alpha y + \beta z = (\alpha\lambda + \beta\mu)x$. □

1.2.3. Éléments inversibles. *Soit A est un anneau non trivial (i.e. $1_A \neq 0_A$), non nécessairement commutatif. Un élément $u \in A$ est dit inversible s'il existe $u' \in A$ tel que*

$$uu' = u'u = 1_A.$$

On note $A^\times$ l'ensemble des éléments inversibles. C'est un sous-ensemble de A^*.

1.2.4. Proposition. *$(A^\times,\times)$ est un groupe.*

Démonstration. Il est facile de voir que le produit de deux éléments inversibles u,v est inversible et que $(uv)^{-1} = v^{-1}u^{-1}$, donc la multiplication est une loi de composition interne. Par ailleurs 1_A est inversible et $1_A \in A^\times$. Comme la multiplication est associative et que l'inverse u^{-1} d'un inversible est lui aussi inversible, la proposition s'ensuit. □

1.2.5. Exemples.

(a) Dans $\mathbb{Z}$, l'ensemble des éléments inversibles est $\mathbb{Z}^\times = \{1,-1\}$.

(b) Dans $\mathbb{K}[X]$ si on a $PQ = 1$ alors P et Q sont non nuls et $\deg(P)+\deg(Q) = 0$, ce qui implique que P et Q sont des polynômes constants non nuls. Par conséquent

$$\mathbb{K}[X]^\times = \{\text{polynômes constants } a_0 \neq 0\} \simeq \mathbb{K}^*.$$

(c) Considérons

$$A = \{x = a + b\sqrt{2} \,/\, a,b \in \mathbb{Z}\}$$

Comme

$$(a+b\sqrt{2})(a'+b'\sqrt{2}) = (aa'+2bb') + (ab'+a'b)\sqrt{2},$$

on voit aisément que $(A,+,\times)$ est un anneau. Dans cet anneau

$$(1+\sqrt{2})(-1+\sqrt{2}) = 1$$

donc $\pm(1+\sqrt{2})$ et $\pm(-1+\sqrt{2})$ sont inversibles. Par la propriété de groupe de $A^\times$, tous les éléments $u = \pm(1+\sqrt{2})^k$, $k \in \mathbb{Z}$ sont inversibles. On peut montrer qu'il n'y en a pas d'autres.

1.2.6. Éléments irréductibles. *Soit A un anneau intègre non trivial. Un élément $p \in A^*$ est dit irréductible si p n'est pas inversible et si on ne peut pas décomposer p sous la forme*

$$p = xy \quad \textit{avec } x, y \in A^* \textit{ non inversibles},$$

ou, de façon équivalente, si p est non inversible et si

$$\forall x, y \in A^*, \quad p = xy \implies x \textit{ ou } y \textit{ inversible}.$$

On notera qu'on peut toujours décomposer un élément $x \in A^*$ quelconque sous la forme $x = u \times u^{-1}x$ pour tout élément inversible u, ce qui justifie le fait d'exiger une décomposition en éléments non inversibles.

1.2.7. Exemples.

(a) Dans $\mathbb{Z}$, les éléments irréductibles sont exactement les nombres premiers et leurs opposés :

$$\mathcal{P} = \{2, 3, 5, 7, 11, \dots\},$$
$$-\mathcal{P} = \{-2, -3, -5, -7, -11, \dots\}.$$

(b) Dans $\mathbb{K}[X]$, les éléments irréductibles sont nécessairement des polynômes de degré $\geqslant 1$, car les constantes (non nulles) sont inversibles. Il est clair que tout polynôme $P = aX + b$ de degré 1 ($a \neq 0$) est irréductible, puisque si $P = QR$ avec Q, R non inversibles, alors $\deg(Q) \geqslant 1$ et $\deg(R) \geqslant 1$, donc $\deg(P) \geqslant 2$. Notons que $aX + b = a(X + b/a)$ admet la racine $x = -b/a \in \mathbb{K}$.

(c) Il peut se produire que des polynômes de degré $\geqslant 2$ soient irréductibles. Par exemple, dans $\mathbb{R}[X]$, $P = X^2 + 1$ est irréductible, car s'il avait une décomposition $P = QR$, les facteurs seraient nécessairement de degré 1, disons $Q = aX + b$, $R = a'X + b$, $a, a' \in \mathbb{R}^*$, $b, b' \in \mathbb{R}$, ce qui donnerait des racines réelles pour P.

(d) Le même raisonnement montre que $P = X^3 - 2$ est irréductible dans $\mathbb{Q}[X]$. Sinon $P = QR$ avec un facteur de degré 1, disons $Q = aX + b$ avec $a \in \mathbb{Q}^*$, $b \in \mathbb{Q}$, ce qui donnerait une racine rationnelle $x = -b/a$. Mais $P = X^3 - 2$ admet seulement la racine réelle $x = \sqrt[3]{2}$ qui n'est pas rationnelle, cf. 2.4.5.

1.3. Division euclidienne

Donnons d'abord la définition dans un anneau commutatif intègre quelconque.

1.3.1. Définition. *Soit A un anneau commutatif intègre (non trivial). On dit que A est un anneau euclidien, ou que A possède une division euclidienne, s'il existe une fonction $v : A^* \to \mathbb{N}$ (appelée stathme euclidien ou simplement stathme) telle que*

$$\forall a \in A,\ \forall b \in A^*,\ \exists\, q, r \in A \text{ tels que } a = bq + r \text{ avec } r = 0 \text{ ou } r \neq 0,\ v(r) < v(b).$$

Ici a est appelé dividende, b diviseur, q quotient et r reste de la division euclidienne.

Il est à noter qu'on n'exige pas en général l'unicité du couple (q, r).

1.3.2. Cas des anneaux $\mathbb{Z}$ et $\mathbb{K}[X]$.

(a) $\mathbb{Z}$ possède une division euclidienne* associée au stathme $v(x) = |x|$. Dans ce cas on a une division euclidienne sous la forme plus précise

$$a = bq + r \quad \text{avec} \quad 0 \leqslant r < |b|,$$

et sous cette condition, le couple (q, r) est *unique*. Redémontrons cette propriété. Si $b < 0$, quitte à changer b en $-b$ et q en $-q$, on se ramène au cas $b > 0$. Pour $a \in \mathbb{N}$, on montre alors la propriété par récurrence sur a : elle est vraie si $0 \leqslant a < b$ (prendre $q = 0$ et $r = a$), et si $a \geqslant b$, on peut appliquer l'hypothèse de récurrence à $a' = a - b < a$ pour obtenir $a' = bq' + r$, $0 \leqslant r < b$, ce qui donne

$$a = a' + b = b(q' + 1) + r = bq + r \quad \text{avec} \quad q = q' + 1.$$

Lorsque $a < 0$, on peut appliquer la division à $a'' = a + |a|b \geqslant |a|(b - 1) \geqslant 0$, et la division $a'' = bq'' + r$ donne par soustraction $a = bq + r$ avec $q = q'' - |a|$. Ici, il y a unicité de (q, r), car si $a = bq + r = bq' + r'$ avec $0 \leqslant r, r' < b$, on en déduit que $r' - r = b(q - q')$ est multiple de b, et comme $-b < r' - r < b$, la seule possibilité est $r' - r = 0$, d'où aussi $q - q' = 0$.

(b) $\mathbb{K}[X]$ possède une division euclidienne associée au stathme $v(P) = \deg(P)$: si $B \in \mathbb{K}[X]^*$, tout polynôme $A \in \mathbb{K}[X]$ peut s'écrire de *manière unique*

$$A = BQ + R \quad \text{avec} \quad \deg(R) < \deg(B)$$

(ce qui inclut la possibilité que $R = 0$ avec notre convention que $\deg(0) = -\infty$). La démonstration est similaire à celle faite dans $\mathbb{Z}$: on raisonne par récurrence sur $d = \deg(A)$, le résultat étant évident si $d = \deg(A) < \delta = \deg(B)$ (on prend

* En arithmétique pure, il y a bien d'autres exemples que $\mathbb{Z}$ admettant une division euclidienne, par exemple l'anneau des "entiers de Gauss" $A = \mathbb{Z}[i] = \{x + yi \,/\, x, y \in \mathbb{Z}\}$. Si $a \in A$ et $b \in A^*$, on vérifie aisément que le point $q \in A$ le plus proche du quotient exact $\frac{a}{b} \in \mathbb{C}$ est tel que $r = a - bq$ satisfait $N(r) < N(b)$ où $N(z) = N(x + yi) = x^2 + y^2 \in \mathbb{N}$. On prend ici $v = N$ comme stathme.

alors $Q = 0$, $R = A$). Supposons maintenant $d \geqslant \delta$ et le résultat démontré pour $d' < d$, et écrivons

$$A = a_0 + a_1 X + \cdots + a_d X^d, \quad B = b_0 + b_1 X + \cdots + b_\delta X^\delta, \quad a_d \neq 0,\ b_\delta \neq 0.$$

Si on prend la différence des termes dominants dans

$$A' := A - ((a_d/b_\delta) X^{d-\delta}) B \in \mathbb{K}[X]$$

on trouve $a_d X^d - ((a_d/b_\delta) X^{d-\delta})(b_\delta X^\delta) = 0$, donc $d' = \deg(A') < d$. Par hypothèse de récurrence, on peut écrire $A' = BQ' + R$, ce qui donne alors

$$A = BQ + R \quad \text{avec} \quad Q = (a_d/b_\delta) X^{d-\delta} + Q'.$$

Notons que cette démonstration est précisément l'explicitation théorique de l'algorithme de division des polynômes. Ici encore on a unicité, puisque

$$A = BQ + R = BQ' + R' \quad \text{avec} \quad \deg(R), \deg(R') < \deg(B)$$

implique $R' - R = B(Q - Q')$ avec $\deg(R' - R) < \deg(B)$, d'où $R' - R = 0$ et $Q - Q' = 0$.

1.3.3. Exemple de division dans $\mathbb{K}[X]$. Voici un exemple de division polynomiale ; on procède comme pour la division entière, en appliquant la procédure de récurrence décrite plus haut.

$$\begin{array}{r|l}
3X^5 - 2X^4 + 4X^3 - 5X^2 + 2X + 4 & 2X^2 - 3X + 4 \\
3X^5 - \frac{9}{2}X^4 + 6X^3 & \hline \\
\hline
\frac{5}{2}X^4 - 2X^3 - 5X^2 & \frac{3}{2}X^3 + \frac{5}{4}X^2 + \frac{7}{8}X - \frac{59}{16} \\
\frac{5}{2}X^4 - \frac{15}{4}X^3 + 5X^2 & \\
\hline
\frac{7}{4}X^3 - 10X^2 + 2X & \\
\frac{7}{4}X^3 - \frac{21}{8}X^2 + \frac{7}{2}X & \\
\hline
-\frac{59}{8}X^2 - \frac{3}{2}X + 4 & \\
-\frac{59}{8}X^2 + \frac{177}{16}X - \frac{59}{4} & \\
\hline
-\frac{201}{16}X + \frac{75}{4} &
\end{array}$$

Le degré du dernier reste $\deg(R) = 1$ est inférieur au degré $\deg(B) = 2$ du diviseur $B = 2X^2 - 3X + 4$, donc l'opération est terminée. On en déduit l'égalité

$$\begin{aligned}
&3X^5 - 2X^4 + 4X^3 - 5X^2 + 2X + 4 \\
&\qquad = (2X^2 - 3X + 4)\big(\tfrac{3}{2}X^3 + \tfrac{5}{4}X^2 + \tfrac{7}{8}X - \tfrac{59}{16}\big) + \big(-\tfrac{201}{16}X + \tfrac{75}{4}\big).
\end{aligned}$$

1.3.4. Remarque. Comme on le voit avec l'exemple ci-dessus, la division euclidienne n'est pas possible en général dans un anneau de polynômes $\mathbb{A}[X]$ dont les coefficients sont pris dans un anneau $\mathbb{A}$, car on est amené à utiliser des fractions. En revanche, si le polynôme B est *unitaire* de degré δ, c'est-à-dire de coefficient dominant $b_\delta = 1$,

$$B = X^\delta + b_{\delta-1}X^{\delta-1} + \cdots + b_1 X + b_0, \quad b_j \in \mathbb{A},$$

alors la division euclidienne est possible, car dans ce cas l'étape de récurrence décrite au 1.3.2 (b) ne nécessite plus de fractions. En particulier, si on fait dans $\mathbb{Z}[X]$ une division par un polynôme unitaire, le quotient et le reste sont bien dans $\mathbb{Z}[X]$.

1.4. Racines des polynômes et factorisation

Soit $P \in A[X]$ un polynôme à coefficients dans un anneau intègre A. Si l'on effectue la division par $(X-w)$ avec $w \in A$, ce qui est possible puisque $(X-w)$ est unitaire, le reste $R \in A[X]$ doit vérifier $\deg(R) < \deg(X-a) = 1$, par conséquent ce reste est une constante : $P(X) = (X-w)Q(X)+c$. Mais si on substitue $X = w$, il vient $P(w) = c$. On en déduit aisément :

1.4.1. Théorème. *Pour tout polynôme $P \in A[X]$ non constant et tout élément $w \in A$, il existe un polynôme $Q \in A[X]$ de degré $\deg(Q) = \deg(P) - 1$ tel que*

$$P(X) = (X-w)Q(X) + P(w).$$

En particulier, si $P(w) = 0$, alors P est divisible par $(X-w)$.
[Nota. *Si P est constant, ceci est vrai aussi avec $Q = 0$*].

1.4.2. Corollaire. *Soit $P \in A[X]$ un polynôme admettant des racines 2 à 2 distinctes $w_1, \ldots, w_k$. Alors P est divisible par $(X-w_1)\ldots(X-w_k)$, c'est-à-dire qu'il existe $Q \in A[X]$ tel que*

$$P(X) = (X-w_1)\ldots(X-w_k)Q(X)$$

Démonstration. On raisonne par récurrence sur k, le résultat ayant déjà été démontré pour $k = 1$. Supposons le résultat déjà démontré pour k, et supposons que P admette une autre racine w_{k+1}. Alors

$$0 = P(w_{k+1}) = (w_{k+1}-w_1)\ldots(w_{k+1}-w_k)Q(w_{k+1}).$$

Comme l'anneau A est intègre et que $w_{k+1} - w_j \neq 0$ pour $1 \leqslant j \leqslant k$, on en conclut que $Q(w_{k+1}) = 0$, mais alors $Q(X)$ est divisible par $(X-w_{k+1})$ et par conséquent $P(X)$ est divisible par $(X-w_1)\ldots(X-w_k)(X-w_{k+1})$, ce qui démontre bien la propriété à l'ordre $k+1$. □

Les propriétés qui précèdent sont très utiles également dans les anneaux de polynômes à plusieurs variables. On définit ainsi un polynôme $P \in A[X,Y]$ comme une expression formelle

$$P(X,Y) = \sum_{1\leqslant i\leqslant d,\, 1\leqslant j\leqslant d'} a_{i,j} X^i Y^j, \quad a_{i,j} \in A.$$

Le degré total $\deg(P)$ (resp. les degrés partiels $\deg_X(P)$, $\deg_Y(P)$) désigne le maximum des entiers $i+j$ (resp. i, resp. j) tels qu'il existe un coefficient $a_{i,j} \neq 0$. Si l'on écrit P sous forme factorisée

$$P(X,Y) = \sum_{1\leqslant j\leqslant d'} \left(\sum_{1\leqslant i\leqslant d} a_{i,j} X^i \right) Y^j,$$

on en conclut que $A[X,Y] = A[X][Y]$, c'est-à-dire que $A[X,Y]$ n'est autre que l'anneau des polyômes en l'indéterminée Y dont les coefficients sont des éléments de l'anneau $A[X]$. En appliquant le théorème 1.4.1 à l'anneau $A'[Y]$ avec $A' = A[X]$, on obtient alors la conséquence suivante.

1.4.3. Corollaire. *Soit $P(X,Y) \in A[X,Y]$. Si P est tel que $P(X,X) = 0$, alors $P(X,Y)$ est divisible par $Y - X$ dans $A[X,Y]$.*

Le même raisonnement donne un résultat analogue pour un nombre quelconque d'indéterminées.

1.4.4. Corollaire. *Soit $P \in A[X_1, X_2, \ldots, X_n]$. Si $P(X_1, X_2, \ldots, X_n)$ devient nul lorsqu'on fait la substitution $X_j := X_i$, alors $P(X_1, X_2, \ldots, X_n)$ est divisible par $X_j - X_i$ dans $A[X_1, X_2, \ldots, X_n]$.*

1.4.5. Déterminant de Vandermonde. Nous allons illustrer ce qui précède en calculant la valeur du déterminant dit de Vandermonde

$$\Delta = \begin{vmatrix} 1 & 1 & \ldots & 1 & 1 \\ X_1 & X_2 & \ldots & X_{n-1} & X_n \\ \vdots & \vdots & & \vdots & \vdots \\ X_1^{i-1} & X_2^{i-1} & \ldots\ X_j^{i-1}\ \ldots & X_{n-1}^{i-1} & X_n^{i-1} \\ \vdots & \vdots & & \vdots & \vdots \\ X_1^{n-2} & X_2^{n-2} & \ldots & X_{n-1}^{n-2} & X_n^{n-2} \\ X_1^{n-1} & X_2^{n-1} & \ldots & X_{n-1}^{n-1} & X_n^{n-1} \end{vmatrix},$$

dont le coefficient (i,j) est X_j^{i-1}, et qui définit un polynôme $\Delta \in \mathbb{Z}[X_1, X_2, ..., X_n]$. Comme Δ est une somme de produits de facteurs dont un est pris dans chaque ligne, chaque monôme, tel que le monôme diagonal $X_1^0 X_2^1 \ldots X_n^{n-1}$, est de degré $0 + 1 + \cdots + (n-1) = n(n-1)/2$. Par conséquent Δ est un polynôme homogène de degré $n(n-1)/2$. Or Δ s'annule si on substitue $X_j := X_i$,

$j > i$, et le raisonnement du corollaire 1.4.2 implique que Δ est divisible par le produit $P = \prod_{1\leqslant i<j\leqslant n}(X_j - X_i)$ (on peut appliquer la récurrence faite plus haut, car la substition $X_j := X_i$ annule seulement le facteur $X_j - X_i$). Mais P est également un polynôme homogène de degré $n(n-1)/2$, de sorte que le quotient Δ/P est une constante. Comme par ailleurs P contient le monôme $\prod_{1\leqslant i<j\leqslant n} X_j = \prod_{1\leqslant j\leqslant n} X_j^{j-1}$, le quotient Δ/P est égal à 1. Ceci donne la formule classique

$$\Delta = \prod_{1\leqslant i<j\leqslant n} (X_j - X_i).$$

1.5. Dérivation des polynômes

Soit $P \in A[X]$ un polynôme de degré d sur un anneau intègre A, écrit

$$P(X) = \sum_{j=0}^{d} a_j X^j, \quad a_j \in A.$$

On commence par définir le polynôme dérivé $P'(X)$ de manière purement algébrique – sans avoir à prendre de véritable limite comme dans $\mathbb{R}$ ou $\mathbb{C}$. Soit Y une autre indéterminée. On forme le "taux d'accroissement"

$$\begin{aligned}\tau_P(X,Y) = \frac{P(Y)-P(X)}{Y-X} &= \sum_{j=0}^{d} a_j \frac{Y^j - X^j}{Y-X} \\ &= \sum_{j=0}^{d} a_j (X^{j-1} + X^{j-2}Y + \cdots + XY^{j-2} + Y^{j-1}).\end{aligned}$$

On voit alors que $\tau_P(X,Y) \in A[X,Y]$ est un polynôme de degré total $d-1$.

1.5.1. Définition. *On appelle polynôme dérivé de P le polynôme $P' \in A[X]$ de degré $d-1$ tel que*

$$P'(X) = \tau_P(X,X) = \sum_{j=0}^{d} j a_j X^{j-1}.$$

Il est facile de voir que $\tau_{P+Q}(X,Y) = \tau_P(X,Y) + \tau_Q(X,Y)$ et que

$$\begin{aligned}\tau_{PQ}(X,Y) = \frac{PQ(Y)-PQ(X)}{Y-X} &= \frac{P(Y)-P(X)}{Y-X}Q(Y) + P(X)\frac{Q(Y)-Q(X)}{Y-X} \\ &= \tau_P(X,Y)Q(Y) + P(X)\tau_Q(X,Y),\end{aligned}$$

ce qui, après substitution $Y := X$, fournit les formules usuelles de dérivation :

1.5.2. Proposition. *Pour tous $P, Q \in A[X]$ on a*

$$(P+Q)' = P' + Q', \quad (PQ)' = P'Q + PQ'.$$

1.5.3. Formule de Leibniz. *Pour tous polynômes $P, Q \in A[X]$, la dérivée k-ième du produit PQ est donnée par*

$$(PQ)^{(k)} = \sum_{j=0}^{k} \binom{k}{j} P^{(j)} Q^{(k-j)}$$

en notant $P^{(0)} = P$.

Démonstration. On procède par récurrence sur k, en utilisant les propriétés du triangle de Pascal pour passer de k à $k+1$, à savoir que

$$\binom{k+1}{j} = \binom{k}{j} + \binom{k}{j-1}.$$

Nous laissons le détail de la vérification au lecteur. □

1.5.4. Remarque. Sur certains corps tels que $\mathbb{K} = \mathbb{F}_2 = \{0, 1\}$, la dérivation des polynômes peut présenter des propriétés "bizarres". Ainsi, au polynôme $P = X^2 + X \in \mathbb{F}_2[X]$ est associée la fonction polynôme nulle $f_P = 0 : \mathbb{F}_2 \to \mathbb{F}_2$. Cependant, le polynôme dérivé $P' = 2X + 1 = 1 \in \mathbb{F}_2[X]$ a pour fonction polynôme associée la fonction constante $f_{P'} = 1$.

1.6. Multiplicité des racines d'un polynôme

Soit A un anneau intègre et $P \in A[X]$ un polynôme de degré $d \geqslant 1$ admettant $w \in A$ comme racine. On sait qu'on peut écrire $P(X) = (X - w)P_1(X)$ avec $\deg(P_1) = d - 1$, et le polynôme P_1 peut (ou non) admettre de nouveau w comme racine. Si c'est le cas, on a $P_1(X) = (X - w)P_2(X)$ avec $\deg(P_2) = d - 2$, donc $P(X) = (X - w)^2 P_2(X)$. On peut répéter le raisonnement avec des polynômes P_i de degrés $d - i$ de plus en plus petits jusqu'à ce que $P_i(w) \neq 0$, ce qui se produit nécessairement à une certaine étape, au plus tard lorsque P_i devient de degré 0 (donc constant et non nul).

1.6.1. Définition. *Soit A un anneau intègre et $P \in A[X]$ un polynôme de degré d admettant $w \in A$ comme racine. On dit que w est une racine de multiplicité $m \geqslant 1$ si on peut écrire*

$$P(X) = (X - w)^m Q(X) \quad \text{avec} \quad \deg(Q) = d - m, \quad Q(w) \neq 0.$$

Si w n'est pas racine de P, c'est-à-dire si $P(w) \neq 0$, l'égalité ci-dessous est encore valide si $m = 0$ et $Q = P$; on convient donc de dire que la multiplicité de w dans P est égale à 0. On a toujours

$$m \leqslant d = \deg(P).$$

Nous allons voir que la multiplicité d'une racine peut se déterminer en utilisant les dérivées successives du polynôme P. La dérivée j-ième de $(X - w)^m$ est

$m(m-1)\cdots(m-j+1)(X-w)^{m-j}$ pour $j \leqslant m$ (et est identiquement nulle pour $j > m$). Au point $X = w$, la seule valeur non nulle est celle de la dérivée m-ième qui vaut

$$\left(\frac{d}{dX}\right)^m (X-w)^m = m!\ .$$

La formule de Leibniz appliquée à $P(X) = (X-w)^m Q(X)$ donne

$$P^{(k)}(X) = \sum_{j=0}^{k} \binom{k}{j} m(m-1)\cdots(m-j+1)(X-w)^{m-j} Q^{(k-j)}(X).$$

Pour $P^{(m)}(w)$, on trouve en particulier que le seul terme non nul du membre de droite correspond à $j = m = k$. On obtient par conséquent :

1.6.2. Formule. *Si $P(X) = (X-w)^m Q(X)$, alors*

$$\begin{aligned} P(w) = P'(w) = \ldots = P^{(m-1)}(w) = 0,\\ P^{(m)}(w) = m!\, Q(w). \end{aligned}$$

1.6.3. Corollaire. *Soit $\mathbb{K}$ un corps "de caractéristique 0", c'est-à-dire tel que $n_{\mathbb{K}} \neq 0$ pour tout $n \in \mathbb{N}^*$, par exemple $\mathbb{K} = \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$. Si $P(X) = (X-w)^m Q(X)$ dans $\mathbb{K}[X]$ avec $Q(w) \neq 0$, on a*

$$P(w) = P'(w) = \ldots = P^{(m-1)}(w) = 0, \quad P^{(m)}(w) = m!\, Q(w) \neq 0,$$

autrement dit la multiplicité de w est le plus petit entier $j \in \mathbb{N}$ tel que $P^{(j)}(w) \neq 0$.

1.6.4. Exemples. (a) Dans $\mathbb{Q}[X]$, on considère

$$P(X) = X^4 + X^3 - 30\,X^2 + 76\,X - 56.$$

On s'aperçoit que $P(2) = 16+8-120+152-56 = 0$. Pour trouver la multiplicité, on peut calculer les dérivées successives

$$\begin{aligned} &P'(X) = 4\,X^3 + 3\,X^2 - 60\,X + 76, && P'(2) = 0,\\ &P''(X) = 12\,X^2 + 6\,X - 60, && P''(2) = 0,\\ &P'''(X) = 24\,X + 6, && P'''(2) = 54 \neq 0, \end{aligned}$$

donc $x = 2$ est une racine triple, c'est-à-dire de multiplicité 3. Le quotient $P(X)/(X-2)^3$ est unitaire de degré 1 et on voit que l'on a la factorisation

$$P(X) = X^4 + X^3 - 30\,X^2 + 76\,X - 56 = (X-2)^3(X+7),$$

car le coefficient constant du quotient est $(-56)/(-2)^3 = 7$.

(b) Dans $\mathbb{C}[X]$, on considère

$$P(X) = X^5 - 3\,X^4 + 2\,X^3 - 6\,X^2 + X - 3.$$

Nous avons $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, d'où $P(i) = i - 3 - 2i + 6 + i - 3 = 0$. On calcule alors

$$\begin{aligned} &P'(X) = 5\,X^4 - 12\,X^3 + 6\,X^2 - 12\,X + 1, \quad && P'(i) = 0, \\ &P''(X) = 20\,X^3 - 36\,X^2 + 12\,X - 12, && P''(i) = 24 - 8i \neq 0. \end{aligned}$$

Par conséquent $x = i$ est racine double de P. Mais comme P est à coefficients réels, on voit en prenant les conjugués que $P(-i) = P'(-i) = 0$ et $P''(-i) = 24 + 8i \neq 0$, de sorte que $-i$ est aussi racine double. On en conclut que $P(X)$ est divisible par $(X-i)^2(X+i)^2 = (X^2+1)^2$. Le quotient $P(X)/(X^2+1)^2$ est unitaire de degré 1 et de coefficient constant -3, d'où la factorisation

$$P(X) = X^5 - 3\,X^4 + 2\,X^3 - 6\,X^2 + X - 3 = (X-i)^2(X+i)^2(X-3).$$

1.7. Théorème de d'Alembert-Gauss

Le théorème de d'Alembert-Gauss, connu aussi sous le nom de *théorème fondamental de l'algèbre*, stipule que tout polynôme $P \in \mathbb{C}[X]$ non constant admet au moins une racine. Les nombres complexes ont été introduits par le mathématicien Bombelli entre 1560 et 1572, mais ce n'est pas avant d'Alembert, en 1746, qu'on se pose le problème d'une preuve rigoureuse du théorème fondamental de l'algèbre. En 1815, Gauss parvient finalement à une preuve complète, utilisant une récurrence algébrique très subtile sur le degré du polynôme ; elle repose d'une part sur la possibilité de résoudre les équations du second degré complexes, ce qui se ramène à des calculs de racines carrées complexes, et d'autre part sur le fait que tout polynôme réel $P \in \mathbb{R}[X]$ de degré impair admet une racine réelle. Pour cette dernière affirmation, on utilise le théorème des valeurs intermédiaires en observant que pour $x \in \mathbb{R}$, $P(x)$ varie entre $-\infty$ et $+\infty$, ou entre $+\infty$ et $-\infty$, en fonction du signe du coefficient dominant. Nous donnerons ici une preuve plus simple et plus directe s'appuyant sur des idées d'Argand (1806) et de Cauchy (1821), qui repose seulement sur le fait qu'une fonction réelle continue sur un segment y atteint son infimum en un point ; les détails de cette preuve restent tout de même assez subtils.

Commençons par des préliminaires généraux. Soit $P = \sum_{k=1}^{d} a_k X^k \in \mathbb{C}[X]$ un polynôme de degré $d \geqslant 1$, avec $a_d \neq 0$. Soit $r \geqslant 0$ fixé. Pour $\theta \in [0, 2\pi]$, la fonction $\theta \mapsto |P(re^{i\theta})|$ est continue, et atteint donc son infimum

$$m(r) = \inf_{\theta \in [0,2\pi]} |P(re^{i\theta})| = \min_{|z|=r} |P(z)|$$

en un point du cercle $|z| = r$. On a

$$\begin{aligned} |P(z)| &= \left|a_d z^d + a_{d-1} z^{d-1} + \cdots + a_1 z + a_0\right| \\ &\geqslant |a_d||z|^d - (|a_{d-1}||z|^{d-1} + \cdots + |a_1||z| + |a_0|), \end{aligned}$$

ce qui entraîne le lemme suivant.

1.7.1. Lemme. *On a la minoration*

$$m(r) \geqslant |a_d| r^d - (|a_{d-1}| r^{d-1} + \cdots + |a_1| r + |a_0|) \quad \textit{avec} \quad |a_d| > 0,$$

par conséquent $\lim_{r \to +\infty} m(r) = +\infty$. □

1.7.2. Lemme. *La fonction $r \mapsto m(r)$ est continue sur $[0, +\infty[$.*

Démonstration. Pour des complexes $z, w \in \mathbb{C}$ vérifiant $|z| \leqslant R$ et $|w| \leqslant R$, on a

$$w^k - z^k = (w - z)(z^{k-1} + z^{k-2} w + \cdots + z w^{k-2} + w^{k-1})$$

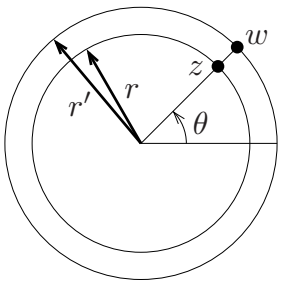
et chacun des k termes $z^j w^{k-1-j}$ est majoré par R^{k-1}. Il s'ensuit par conséquent que $|w^k - z^k| \leqslant k R^{k-1} |w - z|$ et

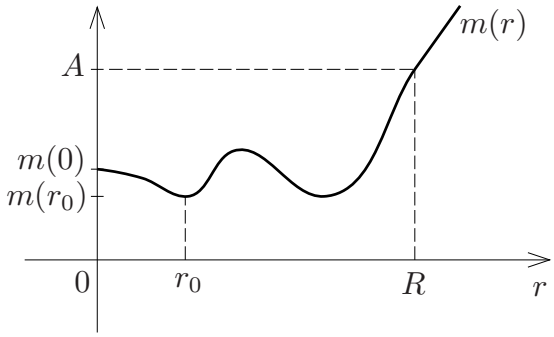
$$|P(w) - P(z)| = \left| \sum_{k=0}^{d} a_k (w^k - z^k) \right| \leqslant |w - z| \sum_{k=1}^{d} k |a_k| R^{k-1}.$$

Maintenant si on prend l'infimum de $|P(z)|$ sur les cercles de rayons $r, r' \in [0, R]$ et si on applique l'inégalité précédente à $z = r e^{i\theta}$ et $w = r' e^{i\theta}$ en observant que $|w - z| = |r' - r|$, on voit en passant à l'infimum que

$$|m(r') - m(r)| \leqslant |r' - r| \sum_{k=1}^{d} k |a_k| R^{k-1}.$$

Ceci entraîne que $r \mapsto m(r)$ est continue sur $[0, R]$ pour tout $R > 0$, d'où la conclusion. □





Choisissons $A > m(0)$. Comme $\lim_{r \to +\infty} m(r) = +\infty$, il existe R suffisamment grand pour que $m(r) \geqslant A > m(0)$ pour $r \in [R, +\infty]$. Ceci implique que la fonction continue $r \mapsto m(r)$ atteint son infimum sur $[0, R]$, et on sait que cet infimum est atteint en un certain point $r_0 \in [0, R]$ (pas nécessairement unique).

1.7.3. Corollaire. *Il existe un point $z_0 = r_0 e^{i\theta_0}$ du cercle $|z| = r_0$ en lequel*

$$|P(z_0)| = m(r_0) = \inf_{r \in [0, +\infty[} m(r) = \inf_{z \in \mathbb{C}} |P(z)|.$$ □

Nous avons besoin d'un dernier lemme, énoncé par Argand en 1806 (mais publié seulement en 1814), qui est le point crucial de la démonstration.

1.7.4. Lemme. *Supposons $P \in \mathbb{C}[X]$ non constant. Soit $z_1 \in \mathbb{C}$ un point tel que $P(z_1) \neq 0$. Alors z_1 ne peut pas être un minimum local pour $z \mapsto |P(z)|$, autrement dit, pour tout $\delta > 0$, il existe $w \in \mathbb{C}$ tel que $|w| < \delta$ et $|P(z_1 + w)| < |P(z_1)|$.*

Notons que ce lemme est violemment faux sur $\mathbb{R}$: le polynôme $P(X) = X^2 + 1$ est non constant, positif, mais passe par un minimum non nul $P(0) = 1$ en $x = 0$.

Démonstration. Posons $Q(X) = P(z_1 + X)/P(z_1)$ de sorte que $Q(0) = 1$, et pour tout $w \in \mathbb{C}$ écrivons

$$Q(w) = 1 + b_k w^k + \cdots + b_d w^d$$

où $b_k \neq 0$ est le coefficient non nul d'indice $k \geqslant 1$ minimal (il existe puisque les polynômes P et Q sont non constants). On va s'arranger pour avoir un terme $b_k w^k$ qui soit réel négatif (et petit). Écrivons $b_k = \beta e^{i\varphi}$ avec $\beta = |b_k| > 0$ et prenons $w = r e^{i(\pi-\varphi)/k}$ avec $r \in]0, \delta[$. On a alors $|w| = r$ et

$$b_k w^k = \beta e^{i\varphi}\, r^k e^{i(\pi-\varphi)} = \beta r^k e^{i\pi} = -\beta r^k,$$

donc

$$Q(w) = 1 - \beta r^k + b_{k+1} w^{k+1} + \cdots + b_d w^d.$$

Choisissons r assez petit pour que $1 - \beta r^k > 0$. En prenant le module, ceci implique

$$\begin{aligned} \frac{|P(z_1 + w)|}{|P(z_1)|} = |Q(w)| &\leqslant 1 - \beta r^k + |b_{k+1}| r^{k+1} + \cdots + |b_d| r^d \\ &= 1 - r^k(\beta - |b_{k+1}| r - \cdots - |b_d| r^{d-k}) < 1 \end{aligned}$$

pour $|w| = r$ assez petit tel que $|b_{k+1}| r + \cdots + |b_d| r^{d-k} < \beta/2$ (et $\beta r^k < 1$). Le lemme est démontré. □

1.7.5. Théorème de d'Alembert-Gauss. *Tout polynôme $P \in \mathbb{C}[X]$ non constant admet au moins une racine $z_1 \in \mathbb{C}$.*

Démonstration. On considère le point $z_1 = z_0$ donné par le corollaire 1.7.3, en lequel $|P(z_1)| = \inf_{z \in \mathbb{C}} |P(z)|$. On veut démontrer que $P(z_1) = 0$. Or, si $P(z_1) \neq 0$, on peut appliquer le lemme 1.7.4 pour trouver un point $z_1' = z_1 + w$ tel que $|P(z_1')| < |P(z_1)|$. C'est une contradiction. Par conséquent $P(z_1) = 0$. □

1.7.6. Théorème. *Soit $P(X) = \sum_{j=0}^{d} a_j X^j \in \mathbb{C}[X]$ de degré d, i.e. $a_d \neq 0$. Alors on peut le factoriser en facteurs de degré 1 sous la forme*

$$P(X) = a_d \prod_{j=1}^{s} (X - z_j)^{m_j}, \quad \sum_{j=1}^{s} m_j = d = \deg P,$$

où $z_1, \ldots, z_s \in \mathbb{C}$ sont les racines complexes 2 à 2 distinctes et les $m_j \in \mathbb{N}^*$ leurs multiplicités.

Démonstration. On raisonne par récurrence sur d. Si $d = 0$, il n'y a rien à démontrer, on a $P(X) = a_0$ et un produit vide égal à 1, avec $s = 0$. Si $d = 1$, on a $P(X) = a_1X + a_0 = a_1(X - z_1)$ où $z_1 = -a_0/a_1$. Supposons $d \geqslant 2$, et le résultat déjà démontré pour $d-1$. D'après le théorème de d'Alembert-Gauss, $P(X)$ possède une racine complexe z_1, et on peut alors factoriser $P(X)$ sous la forme

$$P(X) = (X - z_1)Q(X).$$

Or $Q(X)$ est un polynôme de degré $d-1$ de terme dominant a_dX^{d-1}, on peut alors le factoriser entièrement sous la forme indiquée (avec somme des multiplicités égale à $d-1$). Le résultat s'ensuit pour $P(X)$. □

2. Idéaux et éléments d'arithmétique

2.1. Idéaux, idéaux principaux, anneaux principaux

On introduit d'abord de manière générale la notion d'idéal d'un anneau A : si les éléments de l'anneau sont vus par analogie comme des "scalaires", la définition est voisine de celle de sous-espace vectoriel d'un espace vectoriel.

2.1.1. Définition. *Soit A un anneau. Un sous-ensemble non vide I de A est appelé* idéal *de A si*

(i) $\forall x, y \in I, \quad x + y \in I$ (*stabilité pour* $+$) ;

(ii) $\forall \lambda \in A, \ \forall x \in I, \quad \lambda x \in A$ (*stabilité par la multiplication scalaire de A*).

Il est équivalent de demander que I soit stable par "combinaisons linéaires", autrement dit, la conjection des axiomes (i) *et* (ii) *est équivalente à l'axiome ci-dessous :*

(iii) $\forall \lambda, \mu \in A, \ \forall x, y \in I, \quad \lambda x + \mu y \in A$ (*stabilité par combinaisons linéaires*).

L'hypothèse que l'idéal I soit non vide, combinée à l'axiome (ii) avec $\lambda = 0$ implique que I contient nécessairement l'élément 0 de A.

2.1.2. Exemples. (a) Si $g \in A$, l'ensemble

$$\langle g \rangle = \{ \lambda g \,/\, \lambda \in A \}$$

des multiples de g, aussi noté gA, est un idéal de A ; par exemple, dans $\mathbb{Z}$,

$$\langle 7 \rangle = 7\mathbb{Z} = \{ \ldots, -21, -14, -7, 0, 7, 14, 21, \ldots \}$$

est un idéal. En particulier $\langle 1 \rangle = A$ est un idéal (appelé "idéal unité" de A), de même que $\langle 0 \rangle = \{0\}$ ("idéal zéro").

(b) Plus généralement, si $g_1, \ldots g_N \in A$, l'ensemble des combinaisons linéaires

$$\langle g_1, \ldots, g_N \rangle = \left\{ \sum_{i=1}^{N} \lambda_i g_i \,/\, \lambda_i \in A \right\}$$

est évidemment stable par combinaisons linéaires, donc c'est un idéal de A. Encore plus généralement, on peut considérer une famille finie ou infinie $(g_i)_{i\in S}$ et poser

$$\langle g_i\rangle_{i\in S} = \Big\{ \sum_{\text{finies}} \lambda_i g_i \,/\, i\in S,\ \lambda_i\in A\Big\}.$$

On l'appelle idéal engendré par la famille $(g_i)_{i\in S}$. On dit aussi que $(g_i)_{i\in S}$ est un *système de générateurs* d'un idéal I donné si on a précisément $I=\langle g_i\rangle_{i\in S}$.

(c) Dans l'anneau $\mathbb{K}[X,Y]$ des polynômes à 2 indéterminées, l'ensemble $I_{(0,0)}$ des polynômes P tels que $P(0,0)=0$ est un idéal (puisque cet ensemble est à l'évidence stable par combinaisons linéaires, si $P_1,P_2\in I_{(0,0)}$, alors $P=Q_1P_1+Q_2P_2$ vérifie bien $P(0,0)=0$, donc $P\in I_{(0,0)}$). L'idéal $I_{(0,0)}$ consiste en les polynômes $P=\sum_{i,j} c_{i,j}X^iY^j$ ayant un coefficient constant $c_{0,0}=0$, on peut alors écrire

$$P=\Big(\sum_{i\neq 0,j} c_{i,j}X^{i-1}Y^j\Big)X+\Big(\sum_{i=0,j\neq 0} c_{0,j}Y^{j-1}\Big)Y$$

et on en conclut que

$$I_{(0,0)}=\langle X,Y\rangle.$$

(d) On se place dans l'anneau $\mathbb{K}[X]$ des polynômes sur l'un des corps $\mathbb{K}=\mathbb{Q}$, $\mathbb{R}$ ou $\mathbb{C}$. Soient $w_1,\ldots,w_s\in\mathbb{K}$ des points 2 à 2 distincts et $I_{w_1,\ldots,w_s}$ l'ensemble des polynômes P tels que $P(w_1)=\ldots=P(w_s)=0$. Alors $I_{w_1,\ldots,w_s}$ est un idéal de $\mathbb{K}[X]$, et en fait les résultats de la section précédente montrent que

$$I_{w_1,\ldots,w_s}=\langle (X-w_1)\ldots(X-w_s)\rangle$$

consiste précisément en l'ensemble des polynômes

$$P=Q(X-w_1)\ldots(X-w_s)$$

qui sont multiples de $(X-w_1)\ldots(X-w_s)$. Plus généralement, si on se donne des multiplicités $m_1,\ldots,m_s\in\mathbb{N}$ et qu'on considère l'ensemble $I_{w_1(m_1),\ldots,w_s(m_j)}$ des polynômes P ayant une multiplicité m_i en w_i, c'est-à-dire tels que $P^{(j)}(w_i)=0$ pour tout i et tout $j=0,1,\ldots,m_i-1$, alors

$$I_{w_1(m_1),\ldots,w_s(m_s)}=\langle (X-w_1)^{m_1}\ldots(X-w_s)^{m_s}\rangle$$

est l'idéal engendré par $(X-w_1)^{m_1}\ldots(X-w_s)^{m_s}$.

2.1.3. Définition. *Soit A un anneau.*

(a) *Un idéal I de A est dit* principal *s'il admet un seul générateur, autrement dit, si on peut trouver $g\in I$ tel que $I=\langle g\rangle$ coïncide avec l'ensemble des multiples de g.*

(b) *L'anneau A lui-même est dit* principal *si A est intègre et si tout idéal I de A est principal.*

2.1.4. Exemples.

(a) Les idéaux $I_{w_1,\dots,w_s}$ et $I_{w_1(m_1),\dots,w_s(m_s)}$ de $\mathbb{K}[X]$ sont des idéaux principaux.

(b) L'idéal $I_{(0,0)} = \langle X, Y\rangle$ de $\mathbb{K}[X,Y]$ n'est pas principal. En effet si on avait $I_{(0,0)} = \langle G\rangle$ avec $G \in \mathbb{K}[X,Y]$, il existerait des polynômes $P, Q \in \mathbb{K}[X,Y]$ tels que $X = PG$ et $Y = QG$, donc $G \neq 0$, $G \mid X$ et $G \mid Y$. Mais la condition $G \mid X$ implique $\deg_Y(G) \leqslant \deg_Y(X) = 0$ et la condition $G \mid Y$ implique $\deg_X(G) \leqslant \deg_X(Y) = 0$. Par conséquent G serait une constante non nulle $c \in \mathbb{K}^*$ et on aurait $G(0,0) = c \neq 0$, ce qui est contradictoire puisque $G \in I_{(0,0)}$. Il en résulte que l'anneau $\mathbb{K}[X,Y]$ *n'est pas un anneau principal.*

Nous démontrons maintenant un théorème très important.

2.1.5. Théorème. *Tout anneau euclidien A est principal. En particulier $\mathbb{Z}$ est principal, et les anneaux de polynômes $\mathbb{K}[X]$ sur les corps sont principaux.**

Démonstration. Par hypothèse, il existe un stathme $v : A^* \to \mathbb{N}$ et un algorithme de division euclidienne :

$$\forall a \in A,\ \forall b \in A^*,\ \ \exists q, r \in A,\ \text{tels que } a = bq + r \ \text{ avec } r = 0 \text{ ou } r \neq 0,\ v(r) < v(b).$$

Soit I un idéal de A. Si $I = \{0\}$, alors $I = \langle 0\rangle$ est principal. Si $I \neq \{0\}$ on choisit un élément $g \in I \setminus \{0\}$ tel que le stathme $v(g) \in \mathbb{N}$ prenne la valeur minimale $\min_{x \in I\setminus\{0\}} v(x)$ (ce qui est toujours possible puisqu'on est dans les entiers naturels). Soit $x \in I$ quelconque. On effectue la division euclidienne de x par g :

$$\exists q, r \in A,\ \ x = gq + r \ \text{ avec } r = 0 \text{ ou } r \neq 0,\ v(r) < v(g).$$

Mais alors $r = x - gq = 1 \times x + (-q) \times g \in I$ car $x, g \in I$. D'après le choix de g on ne peut pas avoir $r \in I \setminus \{0\}$, $v(r) < v(g)$, c'est donc que $r = 0$ et que $x = qg$ est multiple de g. Par conséquent $I \subset \langle g\rangle$. Mais comme $g \in I$, on a aussi $\langle g\rangle = \{\lambda g \,/\, \lambda \in A\} \subset I$ et donc $I = \langle g\rangle$. □

2.1.6. Lien avec la divisibilité. Dans un anneau A intègre, les notions de divisibilité peuvent se relier de manière simple à l'inclusion et l'égalité des idéaux principaux.

(a) Si $x, y \in A^*$, alors $\langle x\rangle \subset \langle y\rangle$ si et seulement si $y \mid x$.

En effet, si $y \mid x$, soit $x = \alpha y$, tout multiple $\lambda x = \lambda\alpha y$ est aussi multiple de y, donc $\langle x\rangle \subset \langle y\rangle$. Réciproquement, si $\langle x\rangle \subset \langle y\rangle$, on a $x \in \langle y\rangle$, donc x est multiple de y, i.e. $y \mid x$.

(b) Si $x, y \in A^*$, alors $\langle x\rangle = \langle y\rangle$ si et seulement s'il existe $u \in A^\times$ tel que $y = ux$

(u étant donc inversible). On dit alors que x, y sont multiplicativement équivalents.

* En revanche, il n'est pas vrai que "A principal $\Longrightarrow$ $A[X]$ principal". Par exemple $A = \mathbb{K}[Y]$ est principal, mais $A[X] = \mathbb{K}[Y][X] = \mathbb{K}[X,Y]$ n'est pas principal.

En effet, d'après (a), si $\langle x\rangle = \langle y\rangle$, il existe $\alpha, \beta \in A^*$ tels que $x = \alpha y$ et $y = \beta x$, donc $x = \alpha\beta x$, et comme A est intègre, on en déduit que $\alpha\beta = 1$. On a donc bien $y = ux$ (et $x = u^{-1}y$) avec $u = \beta$ et $u^{-1} = \alpha$. La réciproque est claire.

(c) Si $x \in A^*$, alors $\langle x\rangle = \langle 1\rangle = A$ si et seulement si $x \in A^\times$, i.e. x inversible.

C'est un cas particulier du (b).

Une conséquence de ce qui précède est que dans toutes les questions liées à la divisibilité, on considère comme équivalents des éléments qui ne diffèrent que par un élément inversible. Par exemple, dans $\mathbb{Z}$, on considère 7 et -7 comme équivalents, et lorsqu'on a affaire à des éléments irréductibles p (nombres premiers), on pourra toujours choisir plutôt $p > 0$, à équivalence près. De même, l'ensemble des inversibles de $\mathbb{K}[X]$ consiste en les constantes $\alpha \in \mathbb{K}^*$, et si on a affaire à un polynôme irréductible $P \in \mathbb{K}[X]$, on pourra toujours diviser par son coefficient dominant de façon à se ramener à un polynôme unitaire.

2.2. Opérations sur les idéaux

Soit A un anneau et $I_1, I_2, \ldots, I_k$ des idéaux de A. On leur associe alors des nouveaux idéaux comme suit.

2.2.1. Intersection. $I = I_1 \cap I_2 \cap \ldots \cap I_k$ *est un idéal de* A.

Démonstration. Notons d'abord que $0 \in I$, donc I n'est pas vide. Soient $\lambda, \mu \in A$ et $x, y \in I$. Alors pour tout $\ell = 1, 2, \ldots, k$ on a $x, y \in I_\ell$ donc $\lambda x + \mu y \in I_\ell$. Par conséquent $\lambda x + \mu y \in I$, et I est bien un idéal.

2.2.2. Exemple. Dans $\mathbb{Z}$, l'intersection $\langle 4\rangle \cap \langle 10\rangle$ représente les entiers qui sont à la fois multiples de 4 et de 10, on voit qu'il s'agit donc des multiples de 20. Par conséquent

$$\langle 4\rangle \cap \langle 10\rangle = \langle 20\rangle.$$

2.2.3. Somme. *On définit*

$$I_1 + I_2 + \cdots + I_k = \big\{x_1 + x_2 + \cdots + x_k \,/\, x_\ell \in I_\ell\big\}.$$

Alors $I_1 + I_2 + \cdots + I_k$ *est un idéal de* A.

Démonstration. Comme les I_ℓ sont stables par combinaisons linéaires, la stabilité de la somme par combinaisons linéaires est également évidente. □

2.2.4. Exemple. Dans $\mathbb{Z}$, la somme $\langle 4\rangle + \langle 10\rangle$ est constitué d'entiers de la forme $4n + 10p$ qui sont tous pairs, donc $\langle 4\rangle + \langle 10\rangle \subset \langle 2\rangle$. Mais d'autre part $2 = (-2) \times 4 + 10 \in \langle 4\rangle + \langle 10\rangle$ donc $\langle 2\rangle \subset \langle 4\rangle + \langle 10\rangle$. Ceci implique

$$\langle 4\rangle + \langle 10\rangle = \langle 2\rangle.$$

2.2.5. Produit. (*La définition est plus subtile* !). *Pour des idéaux* I, J, *on pose*

$$IJ = \Big\{\sum_\ell x_\ell\, y_\ell \,/\, x_\ell \in I,\ y_\ell \in J\Big\}.$$

Il est facile de voir que IJ est bien un idéal (on a $0 \in IJ$, et IJ est stable par combinaisons linéaires). Plus généralement, on pose

$$I_1 I_2 \cdots I_k = \Big\{ \sum_\ell x_{1,\ell}\, x_{2,\ell} \cdots x_{k,\ell} \,/\, x_{1,\ell} \in I_1, x_{2,\ell} \in I_2, \ldots, x_{k,\ell} \in I_k \Big\},$$

les sommes étant toujours prises finies.

2.2.6. Exemple. Il n'est pas difficile de voir que dans un anneau A quelconque, un produit d'idéaux principaux est donné par la formule

$$\langle g \rangle \langle h \rangle = \langle gh \rangle,$$

par exemple $\langle 4 \rangle \langle 10 \rangle = \langle 40 \rangle$ dans $\mathbb{Z}$. Plus généralement, pour des idéaux non nécessairement principaux, on a (exercice !)

$$\langle g_i \rangle_{i \in S}\, \langle h_j \rangle_{j \in T} = \langle g_i h_j \rangle_{(i,j) \in S \times T}. \qquad \square$$

2.2.7. Réunion. En général, ce n'est pas un idéal ! Par exemple, dans $\mathbb{Z}$

$$\langle 4 \rangle \cup \langle 10 \rangle$$

contient 4 et 10, mais ne contient pas 14 qui n'est ni multiple de 4 ni multiple de 10, donc il n'y a pas stabilité pour l'addition. $\square$

2.2.8. Réunion croissante. Pour la réunion, il y a tout de même un cas intéressant, celui d'une suite croissante infinie d'idéaux

$$I_0 \subset I_1 \subset \cdots \subset I_k \subset \cdots, \quad k \in \mathbb{N}.$$

Alors la réunion $I = \bigcup_{k \in \mathbb{N}} I_k$ est bien un idéal. En effet, si on prend une combinaison linéaire $\lambda x + \mu y$ d'éléments $x, y \in I$ avec $\lambda, \mu \in A$, alors x appartient à un certain I_k et y à un certain I_ℓ, mais alors $x, y \in I_m$ avec $m = \max(k, \ell)$. Donc $\lambda x + \mu y \in I_m \subset I$. $\square$

2.2.9. Exemple*. Illustrons par un exemple une situation où on a une telle réunion croissante. Soit A l'anneau des fonctions $f : \mathbb{R}_+ \to \mathbb{R}$ qui peuvent s'écrire comme des somme finies

$$\mathbb{R}_+ \ni x \longmapsto f(x) = \sum\nolimits_\ell a_\ell\, x^{b_\ell}, \quad a_\ell \in \mathbb{R},\ b_\ell \in \mathbb{Q}_+.$$

Les exposants b_ℓ sont donc ici des rationnels positifs ou nuls, et non des entiers comme dans le cas des polynômes. Il est facile de voir que $(A, +, \times)$ est un anneau intègre (exercice !). Prenons pour I l'idéal des fonctions f telles que $f(0) = 0$: ce sont les fonctions dont le coefficient a_0 du terme constant $a_0\, x^0$ est nul, puisque tous les autres termes $a_\ell\, x^{b_\ell}$ avec $b_\ell > 0$ s'annulent en 0. D'autre part, soit $(\varepsilon_k)_{k \in \mathbb{N}}$ une suite strictement décroissante de rationnels positifs convergeant vers 0, par exemple $\varepsilon_k = 2^{-k}$. Pour $f \in I$, on a un plus petit exposant $b_m > 0$ qui intervient dans

l'écriture de f avec un coefficient $a_m \neq 0$, et alors f est divisible par la fonction $x \mapsto x^{\varepsilon_k}$ dès que $k \in \mathbb{N}^*$ est pris assez grand pour que $\varepsilon_k \leqslant b_m$, puisqu'alors tous les quotients $a_\ell\, x^{b_\ell}/x^{\varepsilon_k} = a_\ell\, x^{b_\ell-\varepsilon_k}$ sont d'exposants $b_\ell - \varepsilon_k \geqslant b_m - \varepsilon_k \geqslant 0$ dans $\mathbb{Q}_+$. Ceci montre que I est la réunion des idéaux principaux

$$I_k = \langle x \mapsto x^{\varepsilon_k} \rangle.$$

Cette réunion d'idéaux est strictement croissante, car $x^{\varepsilon_{k+1}}$ divise x^{ε_k}, sans que le quotient $x^{\varepsilon_k}/x^{\varepsilon_{k+1}} = x^{\varepsilon_k-\varepsilon_{k+1}}$ définisse une fonction inversible dans A. Il résulte du lemme ci-dessous que I n'est pas un idéal principal, et donc que A n'est pas un anneau principal.

2.2.10. Lemme (Emmy Noether). *Si A possède une suite infinie strictement croissante d'idéaux*

$$I_0 \subsetneq I_1 \subsetneq \cdots \subsetneq I_k \subsetneq I_{k+1} \subsetneq \cdots\ ,$$

alors l'idéal $I = \bigcup_{k\in\mathbb{N}}$ n'est pas principal, et donc l'anneau A n'est pas principal. Par conséquent, un anneau principal ne peut pas posséder une telle suite infinie strictement croissante d'idéaux.

Démonstration. Supposons $I = \langle g \rangle$. Alors $g \in I$ appartiendrait à un certain idéal I_k, et on aurait donc $I = \langle g \rangle \subset I_k$, par suite $I_{k+1} \subset I \subset I_k$, contradiction. □

2.3. PPCM, PGCD et algorithme d'Euclide

On suppose dans toute cette section que A est un *anneau principal*. Étant donné des éléments $x_1, \ldots, x_s \in A^*$, l'idéal intersection $\langle x_1 \rangle \cap \cdots \cap \langle x_s \rangle$ consiste en l'ensemble des multiples communs aux x_j. Cet idéal n'est pas réduit à $\{0\}$, puisqu'il contient $x_1 x_2 \cdots x_s$. Comme l'anneau A est principal, il existe un élément $m \in A^*$ tel que

$$\langle x_1 \rangle \cap \cdots \cap \langle x_s \rangle = \langle m \rangle,$$

et les multiples communs à $x_1, \ldots, x_s$ sont donc précisément les multiples de m.

2.3.1. Définition du ppcm. *Dans un anneau principal A, on appelle* plus petit commun multiple *de $x_1, \ldots, x_s \in A^*$, noté $m = \mathrm{ppcm}(x_1, \ldots, x_s)$ tout élément $m \in A^*$ tel que*

$$\langle x_1 \rangle \cap \cdots \cap \langle x_s \rangle = \langle m \rangle.$$

On observera que m n'est défini de manière unique qu'à un facteur inversible près $u \in A^\times$: l'élément $\widetilde{m} = um$ conviendrait aussi puisque $\langle \widetilde{m} \rangle = \langle m \rangle$. En fait, seul l'idéal $\langle m \rangle$ est défini de manière unique. Pour pallier cette absence d'unicité, on pourra convenir dans $\mathbb{Z}$ de toujours choisir $m > 0$, et dans $\mathbb{K}[X]$ de prendre un polynôme $M \in \mathbb{K}[X]$ qui soit unitaire, mais c'est un choix purement "esthétique". Dans le cas de deux éléments $x, y \in A^*$, on peut écrire

$$\langle x \rangle \cap \langle y \rangle = \langle \mathrm{ppcm}(x, y) \rangle. \tag{2.3.2}$$

L'opération ppcm correspond simplement à l'intersection des idéaux; comme l'intersection des idéaux est commutative et associative, il en est de même pour le ppcm :

2.3.3. Propriété. *Le* ppcm *est commutatif et associatif, c'est-à-dire*

$$\forall x, y \in A^*, \quad \operatorname{ppcm}(x,y) = \operatorname{ppcm}(y,x),$$
$$\forall x, y, z \in A^*, \quad \operatorname{ppcm}(x,y,z) = \operatorname{ppcm}(x, \operatorname{ppcm}(y,z)) = \operatorname{ppcm}(\operatorname{ppcm}(x,y), z)$$

(*les égalités ayant lieu à des éléments inversibles près*).

Pour calculer un ppcm d'un nombre quelconque d'éléments, on peut donc procéder dans un ordre quelconque et se ramener à des ppcm de 2 éléments seulement. On verra plus tard plusieurs méthodes de calcul pratique du ppcm (via le pgcd).

2.3.4. Cas du pgcd. Étant donné $x_1, \dots, x_s \in A^*$, l'idéal somme $\langle x_1 \rangle + \cdots + \langle x_s \rangle$ est la même chose que l'idéal engendré par les x_j :

$$\langle x_1 \rangle + \cdots + \langle x_s \rangle = \big\{ \lambda_1 x_1 + \cdots + \lambda_s x_s \,/\, \lambda_j \in A \big\} = \langle x_1, \dots, x_s \rangle.$$

Cet idéal n'est évidemment pas réduit à $\{0\}$, et comme l'anneau A est principal, il existe un élément $d \in A^*$ tel que

$$\langle x_1 \rangle + \cdots + \langle x_s \rangle = \langle d \rangle.$$

En particulier $x_j \in \langle d \rangle$, ce qui implique que d est un diviseur commun à $x_1, \dots, x_s$. Mais réciproquement, si $d' \in A^*$ est un diviseur commun à $x_1, \dots, x_s$, c'est-à-dire $x_j = k_j d'$, alors tout élément de l'idéal somme $\sum_{j=1}^p \lambda_j x_j = \big(\sum_{j=1}^p \lambda_j k_j\big) d'$ est multiple de d', et en particulier d doit être un multiple de d'. L'élément $d \in A^*$ est donc le "plus grand commun diviseur" des x_j, ou "plus grand" doit se comprendre au sens de la relation de divisibilité : $d' \mid d$.

2.3.5. Définition du pgcd. *Dans un anneau principal A, on appelle* plus grand commun diviseur *de $x_1, \dots, x_s \in A^*$, noté $d = \operatorname{pgcd}(x_1, \dots, x_s)$ tout élément $d \in A^*$, défini à un facteur inversible près $u \in A^\times$, tel que*

$$\langle x_1 \rangle + \cdots + \langle x_s \rangle = \langle d \rangle.$$

Dans le cas de deux éléments $x, y \in A^*$, on peut écrire

$$\langle x \rangle + \langle y \rangle = \langle \operatorname{pgcd}(x,y) \rangle, \tag{2.3.6}$$

et comme l'addition des idéaux est commutative et associative, on en déduit aussitôt :

2.3.7. Propriété. *Le pgcd est commutatif et associatif, c'est-à-dire*

$$\forall x, y \in A^*, \quad \operatorname{pgcd}(x,y) = \operatorname{pgcd}(y,x),$$
$$\forall x, y, z \in A^*, \quad \operatorname{pgcd}(x,y,z) = \operatorname{pgcd}(x, \operatorname{pgcd}(y,z)) = \operatorname{pgcd}(\operatorname{pgcd}(x,y), z)$$

(*les égalités ayant lieu à des éléments inversibles près*).

Une conséquence immédiate de la définition du pgcd est l'identité dite de Bézout ou de Bachet-Bézout (du nom des mathématiciens Français Claude-Gaspard Bachet de Méziriac, 1581–1638, et Étienne Bézout, 1730–1783) :

2.3.8. Identité de Bézout. *Si* $d = \mathrm{pgcd}(x_1, \ldots, x_s)$ *avec* $x_1, \ldots, x_s \in A^*$, *il existe des éléments* $\lambda_1, \ldots, \lambda_s \in A$ *tels que*

$$\lambda_1 x_1 + \cdots + \lambda_s x_s = d.$$

2.3.9. Théorème et définition. *On dit que les éléments* $x_1, \ldots, x_s \in A^*$ *sont* premiers entre eux dans leur ensemble *si*

$$\mathrm{pgcd}(x_1, \ldots, x_s) = 1.$$

Pour cela, il faut et il suffit que

$$\exists\, \lambda_1, \ldots, \lambda_s \in A \quad \textit{tels que} \quad \lambda_1 x_1 + \cdots + \lambda_s x_s = 1.$$

Démonstration. En effet, l'existence d'éléments λ_j comme ci-dessus est bien équivalente au fait que l'idéal engendré $\langle x_1 \rangle + \cdots + \langle x_s \rangle$ coïncide avec l'idéal unité $\langle 1 \rangle = A$. □

2.3.10. Remarque. Si $\mathrm{pgcd}(x_1, \ldots, x_s) = d$ alors on peut "simplifier" le diviseur commun d en écrivant $x_1 = dx'_1, \ldots, x_s = dx'_s$ et l'identité de Bézout 2.3.5 pour $x_1, \ldots, x_s$ implique $\lambda_1 x'_1 + \cdots + \lambda_s x'_s = 1$ après simplification, donc

$$\mathrm{pgcd}(x'_1, \ldots, x'_s) = 1.$$

2.3.11. Attention. Dans $\mathbb{Z}$, on a $\mathrm{pgcd}(6, 10, 15) = 1$ (puisque par exemple $6 + 10 - 15 = 1$), mais $\mathrm{pgcd}(6, 10) = 2$, $\mathrm{pgcd}(6, 15) = 3$, $\mathrm{pgcd}(10, 15) = 5$, il n'y a donc pas équivalence entre le fait que $x_1, \ldots, x_s$ soient premiers entre eux dans leur ensemble, ou premiers entre eux deux à deux (ce qui est une hypothèse bien plus forte d'après 2.3.7).

2.3.12. Distributivité de la multiplication par rapport au ppcm et pgcd. *Pour tous* $a \in A^*$ *et* $x_1, \ldots, x_s \in A^*$, *on a*

(a) $\mathrm{ppcm}(ax_1, \ldots, ax_s) = a\ \mathrm{ppcm}(x_1, \ldots, x_s)$,

(b) $\mathrm{pgcd}(ax_1, \ldots, ax_s) = a\ \mathrm{pgcd}(x_1, \ldots, x_s)$.

Démonstration. (a) Il s'agit de voir que $\langle ax_1 \rangle \cap \ldots \cap \langle ax_s \rangle = \langle a \rangle \big(\langle x_1 \rangle \cap \ldots \cap \langle x_s \rangle \big)$. Or $\langle ax_1 \rangle \cap \ldots \cap \langle ax_s \rangle$ consiste en les éléments $y \in A$ tels qu'il existe $\lambda_1, \ldots, \lambda_s \in A$ vérifiant

$$y = \lambda_1 a x_1 = \ldots = \lambda_s a x_s.$$

Mais comme A est intègre, on peut simplifier les égalités par a, ce qui donne

$$y = az \quad \text{avec} \quad z = \lambda_1 x_1 = \ldots = \lambda_s x_s \in \langle x_1\rangle \cap \ldots \cap \langle x_s\rangle.$$

Ceci implique $\langle ax_1\rangle \cap \ldots \cap \langle ax_s\rangle \subset \langle a\rangle\big(\langle x_1\rangle \cap \ldots \cap \langle x_s\rangle\big)$. L'inclusion inverse est évidente.

(b) L'affirmation découle de la distributivité de l'addition par rapport à la multiplication, qui implique aussitôt

$$a(\lambda_1 x_1 + \cdots + \lambda_s x_s) = \lambda_1(ax_1) + \ldots + \lambda_s(ax_s)$$

(en tenant compte aussi de l'associativité et de la commutativité de $\times$) et donc

$$\langle a\rangle\big(\langle x_1\rangle + \cdots + \langle x_s\rangle\big) = \langle ax_1\rangle + \cdots + \langle ax_s\rangle. \qquad \square$$

2.3.13. Algorithme d'Euclide. On suppose ici que A est un *anneau euclidien*, muni d'un stathme $v : A^* \to \mathbb{N}$. On cherche à calculer le pgcd de deux éléments $a, b \in A^*$. Soit $v_1 = \min(v(a), v(b))$ le minimum de la valeur du stathme pour a et b. On peut toujours ordonner les éléments en sorte que $v(b) \leqslant v(a)$, de façon que $v_1 = v(b)$ [on choisit pour b le "plus petit" des deux éléments]. On effectue alors la division euclidienne de a par b. Ceci donne des éléments $q, r \in A$ tels que

$$a = bq + r \quad \text{avec } r = 0 \text{ ou } r \neq 0,\ v(r) < v(b).$$

Or les combinaisons linéaires de a, b peuvent s'écrire

$$\lambda a + \mu b = \lambda(bq + r) + \mu b = (\lambda q + \mu)b + \lambda r,$$

et inversement les combinaisons linéaires de b, r peuvent s'écrire

$$\lambda' b + \mu' r = \lambda' b + \mu'(a - bq) = \mu' a + (\lambda' - \mu' q)b.$$

Ceci implique $\langle a\rangle + \langle b\rangle = \langle b\rangle + \langle r\rangle$, c'est-à-dire

$$\operatorname{pgcd}(a, b) = \operatorname{pgcd}(b, r).$$

Si $r = 0$, on a $b \mid a$ et $\operatorname{pgcd}(a, b) = b$, tandis que si $r \in A^*$, on peut procéder par récurrence en posant $r_{-1} = a$, $r_0 = b$, $q_0 = q$ et $r_1 = r$, ce qui donne $r_{-1} = q_0 r_0 + r_1$ avec $v(r_1) = v(r) < v(b) = v(r_0)$. L'égalité $\operatorname{pgcd}(a, b) = \operatorname{pgcd}(b, r)$ implique

$$\operatorname{pgcd}(r_{-1}, r_0) = \operatorname{pgcd}(r_0, r_1), \quad \text{et on a } \ v(r_1) < v(r_0).$$

On peut alors procéder inductivement. Tant que les restes r_i obtenus sont non nuls, on produit ainsi des couples successifs $(r_{i-1}, r_i)_{i\geqslant 0}$ en partant de (a, b) et en effectuant des divisions euclidiennes

$$r_{i-1} = q_i r_i + r_{i+1}, \quad \text{avec } r_{i+1} = 0 \text{ ou } r_{i+1} \neq 0,\ v(r_{i+1}) < v(r_i).$$

On obtient ainsi une suite strictement décroissante de valeurs $v(r_i)$ avec

$$r_{i+1} = 0 \ \text{ ou } \ v(r_{i+1}) < v(r_i) < \cdots < v(r_1) < v(r_0).$$

Comme il s'agit d'une suite strictement décroissante d'entiers naturels, le procédé s'arrête nécessairement, ce qui implique qu'on finit par obtenir $r_{i+1} = 0$ à une certaine étape, par conséquent $r_i \mid r_{i-1}$ et

$$\operatorname{pgcd}(a,b) = \operatorname{pgcd}(r_{-1}, r_0) = \operatorname{pgcd}(r_0, r_1) = \ldots = \operatorname{pgcd}(r_{i-1}, r_i) = r_i.$$

On retiendra la règle suivante :

2.3.14. Règle. *Dans l'algorithme d'Euclide,* $\operatorname{pgcd}(a,b)$ *est égal au dernier reste* r_i *non nul calculé (ou à* $r_0 = b$*, si* $b \mid a$*).* □

2.3.15. Retour sur l'identité de Bézout. Un autre mérite de l'algorithme d'Euclide est de permettre le calcul effectif d'éléments $\lambda, \mu \in A$ tels que $\lambda a + \mu b = d$. En effet, d'après la règle ci-dessus, on a

$$\begin{aligned} d = r_i &= r_{i-2} - q_{i-1} r_{i-1} \\ &= r_{i-2} - q_{i-1}(r_{i-3} - q_{i-2} r_{i-2}) \\ &= (q_{i-1} q_{i-2} + 1) r_{i-2} - q_{i-1} r_{i-3} \\ &= (q_{i-1} q_{i-2} + 1)(r_{i-4} - q_{i-3} r_{i-3}) - q_{i-1} r_{i-3} \ \ldots , \end{aligned}$$

et on peut ainsi remonter jusqu'à $r_{-1} = a$, $r_0 = b$.

2.3.16. Exemples. (a) Dans $\mathbb{Z}$, on demande de calculer $d = \operatorname{pgcd}(1662, 1356)$ et de déterminer des entiers $\lambda, \mu \in \mathbb{Z}$ tels que $\lambda\, 1662 + \mu\, 1356 = d$. On effectue pour cela des divisions successives en prenant à chaque fois les restes obtenus comme nouveaux diviseurs :

$$\begin{aligned} 1662 &= 1 \times 1356 + 306, \\ 1356 &= 4 \times 306 + 132, \\ 306 &= 2 \times 132 + 42, \\ 132 &= 3 \times 42 + 6, \\ 42 &= 7 \times 6 + 0. \end{aligned}$$

Le pgcd est le dernier reste non nul, donc $\operatorname{pgcd}(1662, 1356) = 6$, et en effet on a bien $1662 = 6 \times 277$, $1356 = 6 \times 226$ avec $\operatorname{pgcd}(277, 226) = 1$. Pour trouver les coefficients λ et μ, on "remonte" dans les divisions en écrivant

$$\begin{aligned} 6 &= 132 - 3 \times 42 = 132 - 3 \times (306 - 2 \times 132) \\ &= -3 \times 306 + 7 \times 132 = -3 \times 306 + 7 \times (1356 - 4 \times 306) \\ &= 7 \times 1356 - 31 \times 306 = 7 \times 1356 - 31 \times (1662 - 1 \times 1356) \\ &= -31 \times 1662 + 38 \times 1356. \end{aligned}$$

Une solution possible est donc $(\lambda,\mu) = (-31, 38)$. Cette solution n'est pas unique : il est facile de voir que pour tout entier $k \in \mathbb{Z}$, le couple $(\lambda,\mu) = (-31 - 226\,k, 38 + 277\,k)$ vérifie encore

$$\begin{aligned}\lambda\,1662 + \mu\,1356 &= (-31 - 226\,k) \times 1662 + (38 + 277\,k) \times 1356 \\ &= 6 + k(-226 \times 6 \times 277 + 277 \times 6 \times 226) = 6.\end{aligned}$$

(b) Considérons maintenant un exemple de calcul de $\operatorname{pgcd}(A, B)$ pour des polynômes A, B de l'anneau $\mathbb{K}[X]$, à savoir le pgcd de

$$A = X^a - 1, \quad B = X^b - 1, \quad a, b \in \mathbb{N}^*.$$

On supposera par exemple $a \geqslant b$, et on utilise l'algorithme d'Euclide pour calculer $d = \operatorname{pgcd}(a, b)$. Pour cela, on commence par effectuer une division $a = bq + r$, $0 \leqslant r < b$. On a l'identité

$$X^{bq} - 1 = (X^b)^q - 1 = (X^b - 1)(X^{b(q-1)} + X^{b(q-2)} + \cdots + X^b + 1),$$

ce qui donne

$$X^a - 1 = X^{bq+r} - 1 = (X^{bq} - 1)X^r + (X^r - 1),$$

par conséquent

$$X^a - 1 = (X^b - 1)(X^{b(q-1)} + X^{b(q-2)} + \cdots + X^b + 1)X^r + (X^r - 1),$$

soit $A = BQ + R$, $\deg(R) < \deg(B)$, avec

$$Q = (X^{b(q-1)} + X^{b(q-2)} + \cdots + X^b + 1)X^r, \quad R = X^r - 1.$$

D'après l'algorithme d'Euclide appliqué dans l'anneau $\mathbb{K}[X]$, on en conclut que

$$\operatorname{pgcd}(X^a - 1, X^b - 1) = \operatorname{pgcd}(A, B) = \operatorname{pgcd}(B, R) = \operatorname{pgcd}(X^b - 1, X^r - 1).$$

Si (r_{i-1}, r_i) est la suite produite par les divisions successives jusqu'à l'obtention d'un reste $r_{i+1} = 0$, on obtient $R_{i+1} = X^{r_{i+1}} - 1 = 0$, donc $\operatorname{pgcd}(a, b) = r_i$ et $\operatorname{pgcd}(X^a - 1, X^b - 1) = X^{r_i} - 1$, d'où la formule amusante

$$\operatorname{pgcd}(X^a - 1, X^b - 1) = X^{\operatorname{pgcd}(a,b)} - 1.$$

2.4. Décomposition en facteurs irréductibles

2.4.1. Un exemple d'anneau non factoriel. Pour donner un exemple de situation où les choses "se passent mal", considérons l'anneau noté $\mathbb{Z}[i\sqrt{5}]$ tel que

$$\mathbb{Z}[i\sqrt{5}] = \left\{a + bi\sqrt{5} \,/\, a, b \in \mathbb{Z}\right\} \subset \mathbb{C}.$$

Il s'agit bien d'un anneau, puisque la multiplication est une loi interne (comme l'est aussi trivialement l'addition) :

$$(a+bi\sqrt{5})(a'+b'i\sqrt{5}) = (aa'+5bb') + (ab'+ba')i\sqrt{5}.$$

Pour $z = a+bi\sqrt{5} \in \mathbb{Z}[i\sqrt{5}]$, on a $\overline{z} = a-bi\sqrt{5} \in \mathbb{Z}[i\sqrt{5}]$, et on définit

$$N(z) = z\overline{z} = a^2+5b^2 \in \mathbb{N}.$$

Si $z, z' \in \mathbb{Z}[i\sqrt{5}]$, il vient $N(zz') = N(z)N(z')$. Lorsque $u \in \mathbb{Z}[i\sqrt{5}]^\times$ est inversible d'inverse u', la relation $uu' = 1$ implique $N(u)N(u') = 1$ et la seule possibilité est que $N(u) = 1$. Réciproquement, si $N(u) = 1$, il vient $u\overline{u} = 1$, et u est inversible d'inverse $\overline{u}$. On voit alors que les éléments inversibles de l'anneau sont les $u = a+bi\sqrt{5}$ tels que $a^2+5b^2 = 1$. La seule possibilité dans les entiers est $a = \pm 1$, $b = 0$, donc $\mathbb{Z}[i\sqrt{5}]^\times = \{1,-1\}$. Maintenant, l'élément 6 de l'anneau admet les décompositions

$$6 = 2\times 3 = (1+i\sqrt{5})\times(1-i\sqrt{5}),$$

et nous affirmons que les quatre éléments $2, 3, 1+i\sqrt{5}, 1-i\sqrt{5}$ sont irréductibles. En effet

$$N(2) = 4, \quad N(3) = 9, \quad N(1+i\sqrt{5}) = N(1-i\sqrt{5}) = 6,$$

mais il n'existe pas d'éléments $z = a+bi\sqrt{5}$, $a, b \in \mathbb{Z}$ tels que $N(z) = a^2+5b^2 = 2$ ou $N(z) = a^2+5b^2 = 3$, donc les seules "décompositions" possibles de $2, 3$, $1+i\sqrt{5}, 1-i\sqrt{5}$ comportent nécessairement l'un des facteurs inversibles ± 1. L'anneau $\mathbb{Z}[i\sqrt{5}]$ a la propriété surprenante que l'élément 6 possède deux décompositions en facteurs irréductibles totalement différentes !* □

Lorsque l'anneau considéré est principal, la décomposition en facteurs irréductibles a les propriétés attendues. On commence pour cela par démontrer quelques résultats préliminaires, énoncés par Gauss dans ses Disquisitiones arithmeticae (1801), mais probablement déjà connus antérieurement par des mathématiciens comme Fermat (1607–1665).

2.4.2. Lemme (Gauss). *On considère un anneau principal A et des éléments $a, b_1, \ldots, b_s, b, c, p \in A^*$.*

(a) *On suppose que $a \mid bc$ et* $\operatorname{pgcd}(a,b) = 1$. *Alors $a \mid c$.*

(b) *On suppose que* $\operatorname{pgcd}(a,b_1) = 1, \ldots, \operatorname{pgcd}(a,b_s) = 1$. *Alors* $\operatorname{pgcd}(a, b_1 \ldots b_s) = 1$.

* Au niveau des idéaux, les choses se passent en revanche beaucoup mieux. Si on introduit les idéaux non principaux

$$I = \langle 2, 1+i\sqrt{5}\rangle, \quad J = \langle 2, 1-i\sqrt{5}\rangle, \quad K = \langle 3, 1+i\sqrt{5}\rangle, \quad L = \langle 3, 1-i\sqrt{5}\rangle,$$

le lecteur vérifiera facilement (exercice !) que

$$\langle 2\rangle = IJ, \quad \langle 3\rangle = KL, \quad \langle 1+i\sqrt{5}\rangle = IK, \quad \langle 1-i\sqrt{5}\rangle = JL, \quad \langle 6\rangle = IJKL.$$

C'est de là historiquement que vient la terminologie de "nombres idéaux", permettant de rétablir un substitut adéquat à la décomposition bancale en facteurs irréductibles, de la même manière qu'on avait inventé les "imaginaires" pour suppléer à l'inexistence de $\sqrt{-1}$ dans $\mathbb{R}$.

(c) *Si p est irréductible et $p \nmid a$, alors* $\operatorname{pgcd}(p,a) = 1$.

(d) *Si p est irréductible et $p \mid b_1 \dots b_s$, alors il existe j tel que $p \mid b_j$.*

Démonstration. (a) Il est clair que $a \mid ac$ et par hypothèse $a \mid bc$, donc

$$a \mid \operatorname{pgcd}(ac,bc), \quad \text{et d'autre part} \quad \operatorname{pgcd}(ac,bc) = c\operatorname{pgcd}(a,b) = c.$$

(b) Par récurrence sur s, il suffit de le voir pour $s = 2$, disons pour $b_1 = b$ et $b_2 = c$. Or si $\operatorname{pgcd}(a,b) = 1$ et $\operatorname{pgcd}(a,c) = 1$, il existe $\lambda, \mu, \lambda', \mu' \in A$ tels que

$$\lambda a + \mu b = 1, \quad \lambda' a + \mu' c = 1,$$

par conséquent

$$1 = (\lambda a + \mu b)(\lambda' a + \mu' c) = (\lambda\lambda' a + \mu\lambda' b + \lambda\mu' c)a + (\mu\mu')bc = 1,$$

ce qui montre que $\operatorname{pgcd}(a,bc) = 1$.

(c) Si p est irréductible, les seules "décompositions" possibles de p sont de la forme $p = u(u'p)$ avec $u, u' \in A^\times$, $uu' = 1$, donc les seuls diviseurs de p sont les inversibles u et les éléments de la forme $u'p$. Mais par hypothèse $p \nmid a$ équivaut à dire que $u'p \nmid a$, donc seuls restent les éléments inversibles u comme diviseurs communs possibles à p et a. Ceci montre bien que $\operatorname{pgcd}(p,a) = 1$.

(d) Si p ne divisait aucun des b_j, on aurait $\operatorname{pgcd}(p,b_j) = 1$ d'après (c), et donc $\operatorname{pgcd}(p, b_1 \dots b_s) = 1$ d'après (b), contradiction. On peut aussi déduire (d) de (a) et (c) par récurrence sur s (la propriété étant triviale si $s = 1$) : si $p \nmid b_1$ alors $\operatorname{pgcd}(p,b_1) = 1$ et comme $p \mid b_1(b_2 \dots b_s)$, la propriété (a) implique $p \mid b_2 \dots b_s$, d'où la conclusion par hypothèse de récurrence. □

2.4.3. Définition. *Soit A un anneau intègre. Une partie $\mathcal{P} \subset A^*$ sera appelé sous-ensemble "représentatif" des éléments irréductibles de A si $\mathcal{P}$ contient exactement un élément p dans chaque classe d'équivalence $\{up\}$ d'éléments multiplicativement équivalents, de sorte que $\mathcal{P} \ni p \mapsto \langle p \rangle$ soit une bijection de $\mathcal{P}$ sur l'ensemble des idéaux principaux de A.*

Par exemple, dans $\mathbb{Z}$, on choisira pour $\mathcal{P}$ l'ensemble des nombres premiers $p > 0$, et dans $\mathbb{K}[X]$ on choisira l'ensemble des polynômes irréductibles qui sont unitaires.

2.4.4. Définition. *Soit A un anneau intègre et $\mathcal{P} \subset A^*$ un ensemble "représentatif" des éléments irréductibles. On dit que A est* factoriel *si tout élément $x \in A^*$ peut se décomposer sous la forme d'un produit*

$$x = u\, p_1^{m_1} p_2^{m_2} \cdots p_s^{m_s}, \quad u \in A^\times, \quad p_j \in \mathcal{P}, \quad m_j \in \mathbb{N}^*$$

(les $p_j \in \mathcal{P}$ étant deux à deux distincts et s étant éventuellement nul si $x = u$ est inversible), et si de plus la décomposition est unique *à l'ordre près des facteurs $p_j^{m_j}$.*

2.4.5. Théorème. *Tout anneau principal A est factoriel.*

En résumé, pour un anneau intègre A quelconque, on a la chaîne d'implications

$$\text{A euclidien} \;\Rightarrow\; \text{A principal} \;\Rightarrow\; \text{A factoriel},$$

et on peut montrer par des exemples (que nous n'étudierons pas ici) que les implications réciproques ne sont pas vraies.

Démonstration du théorème 2.4.5. Démontrons d'abord *l'existence* de la décomposition 2.4.4. Supposons par l'absurde qu'on ait un élément $x_0 \in A^*$ non décomposable en facteurs irréductibles. Alors x_0 n'est ni inversible ni irréductible, sinon on aurait $x_0 = u$, resp. $x_0 = up$ avec $p \in \mathcal{P}$ et $u \in A^\times$. Par conséquent x_0 peut s'écrire comme un produit $x_0 = x_1 x_1'$ d'éléments $x_1, x_1' \in A^*$ non inversibles. Nécessairement l'un au moins des éléments x_1, x_1' est non décomposable en facteurs irréductibles (sinon $x_0 = x_1 x_1'$ le serait !). Quitte à échanger x_1, x_1', on peut supposer que x_1 est non décomposable. Par récurrence, on construit ainsi une suite $x_{k-1} = x_k x_k'$, $k \in \mathbb{N}^*$, avec x_k non décomposable et x_k' non inversible. On obtient alors une suite infinie strictement croissante d'idéaux

$$\langle x_0\rangle \subsetneq \langle x_1\rangle \subsetneq \cdots \subsetneq \langle x_{k-1}\rangle \subsetneq \langle x_k\rangle \subsetneq \cdots\ ,$$

ce qui contredit le lemme de Noether 2.2.10. Cette contradiction montre que tout élément $x \in A^*$ est bien décomposable en facteurs irréductibles.

En ce qui concerne *l'unicité*, on raisonne par récurrence sur la "multiplicité totale" $m = \sum_{1\leqslant j\leqslant s} m_j$ de l'une des décompositions $x = u\, p_1^{m_1} p_2^{m_2} \cdots p_s^{m_s}$. Si $m = 0$, i.e. $s = 0$, on voit que $x = u$ est inversible, et dans ce cas x ne peut être divisible par aucun facteur irréductible q (sinon $x = u = \lambda q$, $\lambda \in A^*$, et donc $1 = (u^{-1}\lambda)q$, ce qui est contradictoire puisque q n'est pas inversible) ; lorsque $m = 0$, $x = u$ est donc la seule décomposition possible. En général, supposons qu'on ait un élément x possédant deux décompositions

$$x = u\, p_1^{m_1} p_2^{m_2} \cdots p_s^{m_s} = v\, q_1^{r_1} q_2^{r_2} \cdots q_t^{r_t}, \quad p_j, q_j \in \mathcal{P}, \quad m_j, r_j \in \mathbb{N}^*, \quad u, v \in A^\times$$

et supposons l'unicité déjà démontrée pour $m-1$ avec $m = \sum m_j \geqslant 1$. Alors en particulier p_1 divise le produit $v\, q_1^{r_1} q_2^{r_2} \cdots q_t^{r_t}$. D'après le lemme de Gauss, p_1 doit diviser l'un des facteurs. Mais p_1 qui est irréductible ne peut diviser l'élément inversible v. Donc p_1 divise l'un des q_j, et l'irréductibilité de q_j implique alors $q_j = w p_1$ avec $w \in A^\times$ inversible. Puisque $p_1, q_j \in \mathcal{P}$ et que $\mathcal{P}$ est représentatif, on en déduit $p_1 = q_j$. Quitte à permuter les facteurs q_j, on peut supposer $p_1 = q_1$. Après simplification, on trouve

$$u\, p_1^{m_1-1} p_2^{m_2} \cdots p_s^{m_s} = v\, q_1^{r_1-1} q_2^{r_2} \cdots q_t^{r_t}.$$

Comme l'unicité est supposée vraie pour $m-1$ par hypothèse de récurrence, on conclut après permutation des facteurs que $t = s$, $q_j = p_j$ et $r_j = m_j$. □

2.4.6. Remarque. Dans l'exemple de l'anneau $A = \mathbb{Z}[i\sqrt{5}]$, on peut voir que la décomposition en facteurs irréductibles existe bien, même si elle n'est pas

unique : ceci résulte du fait que si $x \in A^*$ s'écrit comme un produit $\prod y_j$ alors $N(x) = \prod N(y_j) \in \mathbb{N}^*$ et lorsqu'on a atteint des facteurs ayant des valeurs $N(y_j)$ minimales, les y_j sont nécessairement irréductibles. En revanche, dans le cas de l'anneau A introduit en 2.2.9, la fonction $f(x) = x^b$ vérifie par exemple $f(x) = (x^{b/2})^2 = \ldots = (x^{b/2^n})^{2^n}$ et on ne peut jamais atteindre d'éléments irréductibles, donc la décomposition en facteurs irréductibles n'existe pas !

2.4.7. Application. (a) Soit $\mathcal{P} \subset \mathbb{N}^*$ l'ensemble des nombres premiers usuels. Alors tout nombre rationnel $x \in \mathbb{Q}^*$ peut s'écrire de manière unique

$$(*) \qquad x = \pm \prod_{p \in \mathcal{P}} p^{m_p}, \quad m_p \in \mathbb{Z}$$

avec une suite d'entiers relatifs $(m_2, m_3, m_5, m_7, \ldots)$ presque tous nuls. En effet si on écrit $x = \frac{a}{b}$ avec $a \in \mathbb{Z}^*$ et $b \in \mathbb{N}^*$, l'existence de la décomposition $(*)$ provient de la décomposition de a et b en facteurs premiers. Pour l'unicité, on utilise le fait qu'une égalité $\prod_{p\in\mathcal{P}} p^{\alpha_p} = \prod_{p\in\mathcal{P}} p^{\beta_p}$ avec les α_p, β_p presque tous nuls se ramène à $\prod_{p\in\mathcal{P}} p^{\gamma_p} = 1$ avec $\gamma_p = \beta_p - \alpha_p \in \mathbb{Z}$, puis à

$$\prod_{p\in\mathcal{P}'} p^{\gamma_p} = \prod_{p\in\mathcal{P}''} p^{-\gamma_p}, \quad \mathcal{P}' = \{p \in \mathcal{P} \,/\, \gamma_p > 0\}, \quad \mathcal{P}'' = \{p \in \mathcal{P} \,/\, \gamma_p < 0\}.$$

Comme $\mathcal{P}'$, $\mathcal{P}''$ sont disjoints, l'égalité ci-dessus n'est possible que si $\mathcal{P}' = \mathcal{P}'' = \emptyset$, ce qui implique $\gamma_p = 0$ pour tout $p \in \mathcal{P}$. La notation standard est

$$x = \varepsilon \prod_{p \in \mathcal{P}} p^{v_p(x)}$$

où $\varepsilon = \pm 1$ et où $m_p = v_p(x) \in \mathbb{Z}$ s'appelle la *valuation p-adique* de x. Une autre façon d'interpréter le résultat ci-dessus est de dire qu'on a un isomorphisme de groupes

$$(\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}^{(\mathcal{P})}, +) \longrightarrow (\mathbb{Q}^*, \times), \quad (a, m_2, m_3, m_5, \ldots) \longmapsto (-1)^a \prod_{p\in\mathcal{P}} p^{m_p}$$

où $a = 0$ ou 1 modulo 2 et $\mathbb{Z}^{(\mathcal{P})}$ désigne l'ensemble des familles d'entiers relatifs $(m_p)_{p\in\mathcal{P}}$ presque tous nuls.

(b) Étant donné $x \in \mathbb{Q}_+^*$ et un entier $q \geqslant 2$, on se demande quelle est la condition nécessaire et suffisante pour que $\sqrt[n]{x} \in \mathbb{Q}$. Pour que ce soit le cas, il faut et il suffit que $x = y^n$ avec $y = \sqrt[n]{x} \in \mathbb{Q}_+^*$. Cette égalité se traduit sous la forme

$$\prod_{p\in\mathcal{P}} p^{v_p(x)} = \left(\prod_{p\in\mathcal{P}} p^{v_p(y)} \right)^n = \prod_{p\in\mathcal{P}} p^{n\, v_p(y)},$$

il faut donc que $v_p(x) = n\, v_p(y)$ soit multiple de n pour tout $p \in \mathcal{P}$; on peut alors prendre y tel que $v_p(y) = \frac{1}{n} v_p(x) \in \mathbb{Z}$. La condition nécessaire et suffisante

est donc que toutes les valuations p-adiques $v_p(x)$ soient multiples de n. Ainsi $\sqrt[3]{729/500} \notin \mathbb{Q}$, car $729/500 = 2^{-2}3^6 5^{-3}$ et l'exposant $-2 = v_2(729/500)$ n'est pas multiple de $n = 3$.

2.4.8. Généralisation. Soit A un anneau factoriel quelconque et $\mathcal{P}$ un ensemble représentatif d'éléments irréductibles. Tout élément $x \in A^*$ s'exprime sous forme d'un unique produit

$$x = \varepsilon(x) \prod_{p\in\mathcal{P}} p^{v_p(x)}, \quad v_p(x) \in \mathbb{N},$$

avec $\varepsilon(x) \in A^\times$ inversible. Il est alors facile de vérifier la proposition suivante.

2.4.9. Proposition. *Soient A un anneau factoriel A, et $x, y \in A^*$. Alors $x \mid y$ si et seulement si pour tout $p \in \mathcal{P}$ on a $v_p(x) \leqslant v_p(y)$ (i.e. si l'exposant de p dans la factorisation en irréductibles de x est inférieur à l'exposant correspondant de la factorisation de y).*

De là on tire aussi que pour $x, y \in A^*$ donnés, les éléments z qui sont multiples à la fois de x et y (resp. qui divisent à la fois x et y) sont ceux tels que $v_p(z) \geqslant \max(v_p(x), v_p(y))$ (resp. tels que $v_p(z) \leqslant \min(v_p(x), v_p(y))$. Ceci permet d'étendre comme suit les formules de calcul du pgcd et du ppcm à tous les anneaux factoriels (et plus seulement aux anneaux principaux).

2.4.10. Théorème et définition. *Dans un anneau factoriel, le ppcm et le pgcd peuvent se définir pour tous $x, y \in A^*$ par les formules*

(a) $$\mathrm{ppcm}(x,y) = \prod_{p\in\mathcal{P}} p^{\max(v_p(x),v_p(y))};$$

(b) $$\mathrm{pgcd}(x,y) = \prod_{p\in\mathcal{P}} p^{\min(v_p(x),v_p(y))}.$$

(c) *On a alors l'identité* $\langle \mathrm{pgcd}(x,y)\rangle\langle \mathrm{ppcm}(x,y)\rangle = \langle xy\rangle$.

Démonstration. Les formules (a) et (b) ont déjà été justifiées dans la discussion préliminaire. Pour vérifier (c), on applique (a) et (b), ce qui donne

$$\langle \mathrm{pgcd}(x,y)\rangle\langle \mathrm{ppcm}(x,y)\rangle = \left\langle \prod_{p\in\mathcal{P}} p^{\min(v_p(x),v_p(y)+\max(v_p(x),v_p(y))} \right\rangle.$$

Mais pour des entiers quelconques u, v, il est clair que $\min(u,v)+\max(u,v) = u+v$. Ceci donne

$$\begin{aligned}\langle \mathrm{pgcd}(x,y)\rangle\langle \mathrm{ppcm}(x,y)\rangle &= \left\langle \prod_{p\in\mathcal{P}} p^{v_p(x)+v_p(y)} \right\rangle = \left\langle \prod_{p\in\mathcal{P}} p^{v_p(x)} \prod_{p\in\mathcal{P}} p^{v_p(y)} \right\rangle \\ &= \langle xy\rangle.\end{aligned}$$

□

2.5. Éléments irréductibles de $\mathbb{Q}[X]$, $\mathbb{R}[X]$ et $\mathbb{C}[X]$

On considère ici l'anneau des polynômes $\mathbb{K}[X]$ à une indéterminée à coefficients dans un corps commutatif $\mathbb{K}$. D'après ce que nous avons vu au paragraphe précédent, tout polynôme $F \in \mathbb{K}[X]^*$ de degré $d \geqslant 0$ se décompose de manière unique sous la forme

$$F = a_d\, P_1^{m_1} P_2^{m_2} \cdots P_s^{m_s}, \quad P_j \in \mathcal{P} \ \text{ 2 à 2 distincts}, \ m_j \in \mathbb{N}^*,$$

où $\mathcal{P}$ désigne l'ensemble des polynômes irréductibles unitaires de $\mathbb{K}[X]$, et où $a_d \in \mathbb{K}^*$ est le coefficient dominant de F. L'ensemble des polynômes irréductibles dépend beaucoup du corps sur lequel on se place.

2.5.1. Anneau $\mathbb{C}[X]$. Dans ce cas, on sait que F se scinde en facteurs de degré 1, les éléments de $\mathcal{P}$ sont les $X - w$, $w \in \mathbb{C}$, et F se factorise sous la forme

$$F(X) = a_d(X - w_1)^{m_1}(X - w_2)^{m_2} \cdots (X - w_s)^{m_s}$$

où $w_1, \ldots, w_s \in \mathbb{C}$ sont les racines distinctes, et $m_1, \ldots, m_s$ leurs multiplicités. On a $d = \deg(P) = \sum_{1 \leqslant j \leqslant s} m_j$.

2.5.2. Anneau $\mathbb{R}[X]$. Soit

$$F(X) = a_d X^d + \cdots + a_1 X + a_0, \quad a_j \in \mathbb{R}, \quad a_d \neq 0.$$

Le polynôme F peut avoir des racines réelles r_j, mais il peut aussi avoir des racines complexes $w_j \in \mathbb{C} \smallsetminus \mathbb{R}$. Dans ce cas

$$F(\overline{w}_j) = a_d \overline{w}_j^d + \cdots + a_1 \overline{w}_j + a_0 = \overline{F(w_j)} = 0.$$

Ceci implique que les racines complexes non réelles $w_j \in \mathbb{C} \smallsetminus \mathbb{R}$ se regroupent par paires conjuguées $w_j, \overline{w}_j$. On peut aussi observer aussi que $w_j, \overline{w}_j$ ont les mêmes multiplicités, car les dérivées vérifient

$$F^{(\alpha)}(\overline{w}_j) = \overline{F^{(\alpha)}(w_j)} \quad \text{pour tout } \alpha \in \mathbb{N}.$$

Chaque produit $(X - w_j)(X - \overline{w}_j)$ s'écrit comme un trinôme du second degré $X^2 + \beta_j X + \gamma_j$ dont les coefficients $\beta_j = -w_j - \overline{w}_j = -2\operatorname{Re} w_j$ et $\gamma_j = |w_j|^2$ sont réels. Le discriminant $\Delta = \beta_j^2 - 4\gamma_j$ est nécessairement < 0 (sinon les racines seraient réelles). Réciproquement, un tel trinôme du second degré de discriminant négatif est bien irréductible dans $\mathbb{R}[X]$. On en déduit que la décomposition de F en facteurs irréductibles est de la forme

$$F(X) = a_d \prod_{1 \leqslant j \leqslant s} (X - r_j)^{m_j} \prod_{1 \leqslant j \leqslant t} (X^2 + \beta_j X + \gamma_j)^{k_j}$$

où les r_j sont les racines réelles, et où les trinômes $X^2 + \beta_j X + \gamma_j$ sont des trinômes de discriminants $\Delta_j = \beta_j^2 - 4\gamma_j < 0$ admettant des racines complexes conjuguées $w_j, \overline{w}_j \in \mathbb{C} \smallsetminus \mathbb{R}$. On a ici $d = \deg(F) = \sum m_j + 2\sum k_j$.

2.5.3. Exemple. Le polynôme $F(X) = X^4 + 1 \in \mathbb{R}[X]$ qui est de degré 4 sans racines réelles est nécessairement réductible dans $\mathbb{R}[X]$ (puisque les polynômes irréductibles y sont de degrés 1 et 2 seulement). On voit en fait que

$$\begin{aligned} X^4 + 1 &= (X^2+1)^2 - 2X^2 = (X^2+1)^2 - (\sqrt{2}\,X)^2 \\ &= (X^2 + \sqrt{2}\,X + 1)(X^2 - \sqrt{2}\,X + 1). \end{aligned}$$

Il s'agit de la décomposition en facteurs irréductibles, car les deux trinômes du second degré ont pour discriminant $\Delta = 2 - 4 = -2$. On en déduit ainsi qu'on a quatre racines complexes deux à deux conjuguées

$$w_1 = \frac{-\sqrt{2} + i\sqrt{2}}{2}, \quad \overline{w}_1 = \frac{-\sqrt{2} - i\sqrt{2}}{2}, \quad w_2 = \frac{\sqrt{2} + i\sqrt{2}}{2}, \quad \overline{w}_2 = \frac{\sqrt{2} - i\sqrt{2}}{2}.$$

2.5.4. Anneau $\mathbb{Q}[X]$. La situation est ici beaucoup plus compliquée, on montrera plus loin qu'il y a dans $\mathbb{Q}[X]$ des polynômes irréductibles de tous degrés, par exemple $X^d - p$ si p est un nombre premier. Nous nous contenterons de quelques résultats élémentaires, car la théorie générale nécessite des outils plus avancés dont nous ne disposons pas dans ce cours.

2.5.5. Proposition. *Un polynôme $F \in \mathbb{Q}[X]$ de degré 2 ou 3 est irréductible si et seulement si F n'admet pas de racine dans $\mathbb{Q}$.*

Démonstration. Ce fait a déjà été remarqué : si un tel polynôme est réductible dans $\mathbb{Q}[X]$, alors l'un des facteurs au moins est de degré 1 et possède donc une racine rationnelle. □

Nous donnons maintenant une méthode permettant de détecter les racines rationnelles. Remarquons que pour un polynôme $F \in \mathbb{Q}[X]$, on peut toujours multiplier les coefficients par un dénominateur commun de façon à se ramener à un polynôme $F \in \mathbb{Z}[X]$.

2.5.6. Proposition. *Soit $F \in \mathbb{Z}[X]$ un polynôme de degré d,*

$$F(X) = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_1 X + a_0, \quad a_j \in \mathbb{Z}, \quad a_d \neq 0.$$

Supposons aussi $a_0 \neq 0$ (sinon $F(X)$ admet la racine $x = 0$ et on peut factoriser). Si $F(X)$ admet une racine rationnelle $x \in \mathbb{Q}^*$ écrite sous forme d'une fraction réduite $x = k/\ell$, c'est-à-dire telle que $\operatorname{pgcd}(k, \ell) = 1$, alors $k \mid a_0$ et $\ell \mid a_d$.

En pratique, on n'a donc qu'à chercher les diviseurs de a_0 et a_d, et cela ne laisse qu'un nombre fini de possibilités de racines $x \in \mathbb{Q}^*$ à tester. On notera en particulier que si le polynôme F est unitaire ($a_d = 1$), alors $\ell = \pm 1$, donc les racines rationnelles sont nécessairement dans $\mathbb{Z}$.

Démonstration. Si $x = k/\ell \in \mathbb{Q}^*$ est racine, alors

$$\ell^d F(k/\ell) = a_d k^d + a_{d-1} k^{d-1} \ell + \cdots + a_1 k \ell^{d-1} + a_0 \ell^d = 0.$$

Ceci donne

$$a_d k^d = -\ell\big(a_{d-1}k^{d-1} + \cdots + a_1 k\ell^{d-2} + a_0\ell^{d-1}\big),$$

donc $\ell \mid a_d k^d$. Comme $\operatorname{pgcd}(k,\ell) = 1$, le lemme de Gauss implique que $\ell \mid a_d$. De même l'égalité

$$a_0\ell^d = -k\big(a_d k^{d-1} + a_{d-1}k^{d-2}\ell + \cdots + a_1 k\ell^{d-1}\big)$$

implique $k \mid a_0$. □

2.5.7. Un exemple de degré 4. Le polynôme $X^4 + 1$ est irréductible dans l'anneau $\mathbb{Q}[X]$. En effet ce polynôme n'a pas de racine rationnelle (ni même réelle), la seule possibilité serait une décomposition en deux facteurs de degré 2 dans $\mathbb{Q}[X]$. Mais on a déjà vu qu'on avait dans $\mathbb{R}[X]$ une décomposition en deux facteurs irréductibles de degré 2, et c'est la seule décomposition possible (si on prend des polynômes unitaires). Il faudrait donc que cette décomposition soit à coefficients rationnels, ce qui n'est pas le cas, car on a le coefficient $\sqrt{2}$ qui apparaît dans les facteurs irréductibles de $X^4 + 1$ dans $\mathbb{R}[X]$.

2.5.8. Autre exemple. Le polynôme $X^4 + 4$ est en revanche *réductible* dans l'anneau $\mathbb{Q}[X]$, bien que n'ayant aucune racine rationnelle. En effet

$$\begin{aligned} X^4 + 4 &= (X^2+2)^2 - 4X^2 = (X^2+2)^2 - (2X)^2 \\ &= (X^2 + 2X + 2)(X^2 - 2X + 2). \end{aligned}$$

[Accessoirement, les racines complexes sont $z = \pm 1 \pm i$.]

2.5.9. Théorème. *Soit $p \in \mathbb{N}^*$ un nombre premier et $d \in \mathbb{N}^*$. Alors $X^d - p$ est irréductible dans $\mathbb{Q}[X]$.*

Démonstration. Les cas $d = 1, 2, 3$ sont triviaux (pour $d = 2, 3$, on applique 2.5.5). En général, pour $d \geqslant 2$, les racines complexes sont données par $z = p^{1/d}\zeta$ où $\zeta^d = 1$, c'est-à-dire $z = p^{1/d}e^{2\pi ik/d}$, $k = 0, 1, \ldots d-1$, et on a dans $\mathbb{C}[X]$ la factorisation

$$X^d - p = \prod_{k=0}^{p-1}(X - p^{1/d}e^{2\pi ik/d}).$$

Si $X^d - p$ avait une décomposition sous la forme $G(X)H(X)$ dans $\mathbb{Q}[X]$ (avec G, H qu'on peut supposer unitaires, et $0 < \deg(G) < d$, $0 < \deg(H) < d$), alors G (disons) serait un produit de certains des facteurs complexes $(X - p^{1/d}e^{2\pi ik_j/d})$, $1 \leqslant j \leqslant \delta = \deg(G)$. Mais alors le coefficient constant $c \in \mathbb{Q}$ de G serait égal à

$$c = \prod_{1\leqslant j\leqslant\delta}(-p^{1/d}e^{2\pi ik_j/d}) = \pm p^{\delta/d}e^{2\pi i\ell/d}, \quad 1 \leqslant \delta < d.$$

Ceci impliquerait $|c| = p^{\delta/d} = \sqrt[d]{p^\delta} \notin \mathbb{Q}$ puisque $d \nmid \delta$. Or $|c| = \pm c \in \mathbb{Q}$. Cette contradiction démontre l'irréductibilité de $X^d - p$ dans $\mathbb{Q}[X]$. □

2.5.10. Exercice. *Déterminer les racines réelles et la décomposition en facteurs irréductibles de $X^d - p$ dans $\mathbb{R}[X]$, suivant que d est pair ou impair. Redémontrer ainsi l'irréductibilité de $X^d - p$.*

2.5.11. Exercice.** En raffinant le raisonnement du théorème 2.5.9, montrer qu'un polynôme $X^d - a$ avec $a \in \mathbb{Q}_+^*$ est irréductible dans $\mathbb{Q}[X]$ si et seulement si pour tout diviseur d' de d tel que $1 < d' \leqslant d$, on a $\sqrt[d']{a} \notin \mathbb{Q}$ (ce qui revient à dire que si $p_1, \ldots, p_s$ sont les facteurs premiers intervenant dans a, on a $\mathrm{pgcd}(v_{p_1}(a), \ldots, v_{p_s}(a), d) = 1$).

2.6. Congruences et anneaux quotients

On introduit d'abord la notion élémentaire de congruence modulo un idéal.

2.6.1. Définition. *Soit A un anneau (commutatif) et I un idéal de A. On dit que des éléments $x, y \in A$ sont congrus modulo I, et on écrit $x \equiv y \mod I$ si $x - y \in I$.*

Il est clair qu'il s'agit d'une relation d'équivalence. En effet on a :

– réflexivité: $x \equiv x \mod I$, puisque $x - x = 0 \in I$;

– symétrie: $x \equiv y \mod I \Leftrightarrow y \equiv x \mod I$, car $\delta = x - y \in I \Leftrightarrow -\delta = y - x \in I$;

– transitivité : $x \equiv y$ et $y \equiv z \mod I \Rightarrow x \equiv z \mod I$, car $x - z = (x-y)+(y-z) \in I$.

Étant donné $x \in A$, on note $\dot{x}$ la classe d'équivalence pour la relation de congruence modulo I. Comme $y \equiv x \mod I$ équivaut à $y - x = t \in I$, il s'agit simplement de

$$\dot{x} = x + I = \{x + t \,/\, t \in I\}. \tag{2.6.2}$$

Maintenant étant donné deux classes $\dot{x}$ et $\dot{y}$ et $x' = x + s \in \dot{x}$, $y' = y + t \in \dot{y}$, $s, t \in I$, on observe que

$$x' + y' = x + y + (s + t) \quad \text{avec } s + t \in I,$$

$$x'y' = (x + s)(y + t) = xy + (xt + ys + st) \quad \text{avec } xt + ys + st \in I,$$

par conséquent $x' + y' \equiv x + y$ et $x'y' \equiv xy \mod I$. Ceci montre qu'il est possible de poser par définition

$$\dot{x} + \dot{y} := (x + y)^{\bullet} = (x' + y')^{\bullet}, \tag{2.6.3}$$

$$\dot{x}\,\dot{y} := (x\,y)^{\bullet} = (x'\,y')^{\bullet}, \tag{2.6.3$'$}$$

puisque les classes définies par les membres de droite ne dépendent pas des représentants x, y ou x', y' choisis.

2.6.4. Théorème et définition. *L'ensemble des classes d'équivalence $\dot{x}$ des éléments $x \in A$ pour la relation de congruence modulo I se note A/I. Avec les lois définies par (2.6.3) et (2.6.3$'$), on obtient une structure d'anneau commutatif $(A/I, +, \times)$, appelé* anneau quotient *de A par I. De plus, l'application naturelle*

$$\pi_{A,I} : A \to A/I, \quad x \mapsto \dot{x}$$

est un morphisme surjectif d'anneaux.

Démonstration. Les explications données plus haut montre que les lois de composition interne $+$ et $\times$ sont bien définies sur A/I. Tous les axiomes des anneaux (associativité, distributivité, ...) vraies dans A passent "automatiquement" à A/I. La classe $\dot{0}_A$ est élément neutre pour $+$, et la classe $\dot{1}_A$ est élément neutre pour $\times$ dans A/I. La propriété de morphisme est évidente par définition, et la surjectivité de $\pi_{A,I}$ aussi. □

2.6.5. Exemples. (a) On a deux cas assez peu intéressants, à savoir $I = \{0\}$, où chaque classe $\dot{x}$ se réduit au singleton $\{x\}$, de sorte que $A/\{0\}$ peut s'identifier à A lui-même, et le cas $I = A$ où on a $\dot{x} = A = \dot{0}$ pour tout $x \in A$, de sorte que $A/A = \{\dot{0}\}$ est l'anneau trivial (avec $\dot{0} = \dot{1}$).

(b) Dans le cas où $I = \langle g \rangle = gA$ est un idéal principal, la classe d'équivalence $\dot{x}$ consiste en les éléments

$$\dot{x} = \{x + \lambda g \,/\, \lambda \in A\}$$

et l'anneau quotient est souvent noté A/gA. Pour $n \in \mathbb{N}^*$, l'anneau $\mathbb{Z}/n\mathbb{Z}$ est ainsi constitué de n éléments, à savoir les classes

$$\dot{0},\ \dot{1},\ \ldots\ ,(n-1)^{\bullet}$$

où la classe $\dot{x} = \dot{r}$ est obtenu en calculant le reste r de la division de x par n, tel que $x = nq + r$. (On a encore ici les cas "inintéressants" $\mathbb{Z}/0\,\mathbb{Z} = \mathbb{Z}/\{0\} \simeq \mathbb{Z}$ et $\mathbb{Z}/1\,\mathbb{Z} = \mathbb{Z}/\mathbb{Z} = \{\dot{0}\}$). L'anneau $\mathbb{Z}/6\mathbb{Z}$ est bien celui défini par les tables de Pythagore du 1.1.4 (c), et l'anneau $\mathbb{Z}/2\mathbb{Z}$ est la même chose que le corps $\mathbb{F}_2$ déjà mentionné à plusieurs reprises.

2.6.6. Application : "preuves" par 9 et 11. Il s'agit de techniques utilisées autrefois – et peut-être encore aujourd'hui ? – par les écoliers des classes primaires pour vérifier leurs opérations arithmétiques. Bien entendu, les maîtres donnaient en général seulement la recette sans trop en expliquer les raisons ...

(a) *"Preuve" par 9.* On travaille dans l'anneau quotient $\mathbb{Z}/9\mathbb{Z}$. On remarque que l'on a $10 \equiv 1 \bmod 9$ et donc $10^k \equiv 1 \bmod 9$ pour tout $k \in \mathbb{N}$. Par conséquent, si $x \in \mathbb{N}$ est un entier écrit en base 10, on voit que

$$x = a_k\,10^k + a_{k-1}\,10^{k-1} + \cdots + a_1\,10 + a_0 \equiv a_k + a_{k-1} + \cdots + a_1 + a_0 \mod 9,$$

autrement dit x est congru à la somme de ses chiffres modulo 9. Par exemple 4570891 est congru à $4+5+7+8+9+1$, et donc à $7+8+1 = 16$ (puisque $9 \equiv 0 \bmod 9$), 16 étant lui-même congru à $1+6 = 7$. Par suite $4570891 \equiv 7 \bmod 9$. Pour "vérifier" le résultat z d'une multiplication xy, on calcule comme ci-dessus les classes $\dot{x}, \dot{y}, \dot{z}$, et on s'assure que $\dot{x}\,\dot{y} = \dot{z}$. Par exemple, pour vérifier que $23 \times 47 = 1081$ on fait $2+3 = 5$, $4+7 = 11 \equiv 2$, $1+0+8+1 \equiv 1$ et on constate qu'on a bien $5 \times 2 \equiv 1 \bmod 9$. Mais si on avait trouvé $z = 1171$, on ne se serait quand même pas aperçu de l'erreur, la "preuve" par 9 est loin d'être infaillible !

(b) *"Preuve" par 11.* L'idée est la même, on travaille cette fois dans l'anneau quotient $\mathbb{Z}/11\mathbb{Z}$. On a $10 \equiv -1 \bmod 11$ et donc $10^k \equiv (-1)^k \bmod 11$, d'où

$$\begin{aligned} x &= a_k\, 10^k + a_{k-1}\, 10^{k-1} + \cdots + a_1\, 10 + a_0 \\ &\equiv (-1)^k a_k + (-1)^{k-1} a_{k-1} + \cdots - a_1 + a_0 \mod 11, \end{aligned}$$

c'est-à-dire que x est congru à la somme alternée de ses chiffres modulo 11. La "preuve" par 10 est beaucoup plus simple, puisqu'elle dépend seulement du dernier chiffre a_0. On peut aussi faire les "preuves" par 7 ou 13 (ou n'importe quel autre entier n), mais c'est plus compliqué ...

2.6.7. Théorème. *Si A est un anneau principal et $I = gA$ est l'idéal engendré par un élément $g \in A^*$, alors il y a équivalence entre les propriétés suivantes :*

(i) *g est un élément irréductible de A ;*

(ii) *A/gA est intègre non trivial ;*

(iii) *A/gA est un corps.*

Démonstration. Il est évident que (iii) $\Rightarrow$ (ii). D'autre part, si g est décomposable sous la forme $g = xy$ avec $x, y \in A^*$ non inversibles, alors g ne peut diviser x ou y (sinon y, resp. x, serait inversible d'inverse x/g, resp. y/g), donc $\dot{x} \neq \dot{0}$, $\dot{y} \neq \dot{0}$ et

$$\dot{x}\,\dot{y} = \dot{g} = \dot{0} \quad \text{dans } A/gA,$$

de sorte que A/gA est non trivial et possède des diviseurs de zéro. Ceci montre par contraposition que (ii) $\Rightarrow$ (i). Il reste à voir que (i) $\Rightarrow$ (iii). Supposons maintenant g irréductible et soit $\dot{x} \neq \dot{0}$ dans A/gA. Ceci signifie que $g \nmid x$, et d'après le lemme de Gauss 2.4.2 (c), on en déduit que $\operatorname{pgcd}(x, g) = 1$. D'après l'identité de Bézout, il existe alors $\lambda, \mu \in A$ tels que $\lambda x + \mu g = 1$, ce qui donne $\dot{\lambda}\,\dot{x} = \dot{1}$, et on voit que A/gA est bien un corps. □

2.6.8. Corollaire. *L'anneau quotient $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement n est un nombre premier. On note $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$; c'est un corps à p éléments, constitué des classes $\dot{0}, \dot{1}, \ldots, (p-1)^{\bullet}$.*

2.6.9. Notion de caractéristique d'un corps. Soit $\mathbb{K}$ un corps, et

$$\varphi_{\mathrm{can}} : \mathbb{Z} \to \mathbb{K}, \quad n \mapsto n_{\mathbb{K}}$$

le morphisme canonique. Le noyau $\operatorname{Ker}(\varphi_{\mathrm{can}}) = \{0\}$ est un idéal de $\mathbb{Z}$, car $\varphi_{\mathrm{can}}(x) = \varphi_{\mathrm{can}}(y) = 0_{\mathbb{K}}$ implique trivialement

$$\varphi_{\mathrm{can}}(\lambda x + \mu y) = \varphi_{\mathrm{can}}(\lambda)\varphi_{\mathrm{can}}(x) + \varphi_{\mathrm{can}}(\mu)\varphi_{\mathrm{can}}(y) = 0_{\mathbb{K}}.$$

Comme $\mathbb{Z}$ est un anneau principal, on a $\operatorname{Ker}(\varphi_{\mathrm{can}}) = p\mathbb{Z}$ pour un certain entier $p \in \mathbb{N}$. On dit alors que $\mathbb{K}$ est un corps de *caractéristique* p. Deux alternatives exclusives l'une de l'autre peuvent se produire.

(a) $p = 0$. Dans ce cas $\mathrm{Ker}(\varphi_{\mathrm{can}}) = \{0\}$, i.e. φ_{can} est injectif, et on peut définir un morphisme canonique de corps en posant

$$\psi_{\mathrm{can}} : \mathbb{Q} \to \mathbb{K}, \quad \frac{a}{b} \mapsto (a_{\mathbb{K}})(b_{\mathbb{K}})^{-1}, \quad a \in \mathbb{Z}, \quad b \in \mathbb{N}^*$$

la vérification facile en est laissée au lecteur. Comme $\psi_{\mathrm{can}}(x) \neq 0$ pour tout $x = a/b \in \mathbb{Q}^*$, on a aussi $\mathrm{Ker}(\psi_{\mathrm{can}}) = \{0\}$, c'est-à-dire que ψ_{can} est injectif et définit un isomorphisme de $\mathbb{Q}$ sur son image $\psi_{\mathrm{can}}(\mathbb{Q}) \subset \mathbb{K}$ (qui est un sous-corps).

(b) $p \in \mathbb{N}^*$, de sorte que $p_K = 0_K$. Comme $1_{\mathbb{K}} \neq 0_{\mathbb{K}}$, on a nécessairement $p > 1$, et d'autre part p doit être un *nombre premier*, sinon on aurait $p = xy$ avec $1 < x, y < p$ et donc $x_{\mathbb{K}}\, y_{\mathbb{K}} = 0_{\mathbb{K}}$ avec $x_{\mathbb{K}}, y_{\mathbb{K}} \neq 0$, ce qui contredirait le fait que $\mathbb{K}$ est un corps. Comme $\varphi_{\mathrm{can}}(x + \lambda p) = \varphi_{\mathrm{can}}(x) = x_{\mathbb{K}}$, on voit que φ_{can} définit par passage au quotient un morphisme de corps

$$\psi_{\mathrm{can}} : \mathbb{Z}/p\mathbb{Z} \to \mathbb{K}, \quad \dot{x} \mapsto x_{\mathbb{K}} \quad (p \text{ premier}),$$

dont le noyau est cette fois réduit à $\{\dot{0}\}$. Par conséquent ψ_{can} est injectif, et définit un isomorphisme de $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ sur $\psi_{\mathrm{can}}(\mathbb{Z}/p\mathbb{Z}) \subset \mathbb{K}$ (qui est un sous-corps).

2.6.10. Théorème. *Soit $\mathbb{K}$ un corps fini. Alors $\mathbb{K}$ est de caractéristique $p > 0$ et $q = \mathrm{card}\,\mathbb{K}$ est une puissance de p, soit $q = p^n$. De plus, pour tout $\alpha \in \mathbb{K}^*$, on a $\alpha^{q-1} = 1_{\mathbb{K}}$, et le polynôme $X^q - X \in \mathbb{K}[X]$ admet la factorisation*

$$X^q - X = \prod_{\alpha \in \mathbb{K}} (X - \alpha).$$

Démonstration. La caractéristique de $\mathbb{K}$ est nécessairement positive, sinon $\mathbb{K}$ contiendrait un sous-corps infini isomorphe à $\mathbb{Q}$. Donc $\mathbb{K}$ contient un sous-corps $\mathbb{K}'$ isomorphe à $\mathbb{F}_p$, et les lois $+$, $\times$ munissent $\mathbb{K}$ d'une structure d'espace vectoriel de dimension finie sur $\mathbb{K}' \simeq \mathbb{F}_p$. Si n est sa dimension, on a alors $\mathbb{K} \simeq (\mathbb{F}_p)^n$, donc $q = \mathrm{card}\,\mathbb{K} = p^n$. Si $\alpha \in \mathbb{K}^*$, l'application $\mathbb{K}^* \to \mathbb{K}^*$, $x \mapsto \alpha x$ est une bijection, dont la bijection inverse est $x \mapsto \alpha^{-1}x$. Comme $\mathrm{card}\,\mathbb{K}^* = q - 1$, le produit (non nul) de tous les éléments de $\mathbb{K}^*$ fournit

$$\prod_{x \in \mathbb{K}^*} x = \prod_{x \in \mathbb{K}^*} \alpha x = \alpha^{q-1} \prod_{x \in \mathbb{K}^*} x.$$

Ceci implique bien que $\alpha^{q-1} = 1$. Comme le polynôme $X^{q-1} - 1$ est de degré $q - 1$ et admet les $q - 1$ éléments de $\mathbb{K}^*$ comme racines, on en déduit que ces racines sont simples et qu'on a la factorisation

$$X^{q-1} - 1 = \prod_{\alpha \in \mathbb{K}^*} (X - \alpha).$$

En multipliant par X, on obtient la factorisation annoncée de $X^q - X$. □

2.6.11. Corollaire : petit théorème de Fermat. *Soit p un nombre premier. Alors*

(a) *pour tout $x \in \mathbb{Z}$ non divisible par p, on a $x^{p-1} \equiv 1 \ mod\ p$;*

(b) *pour tout $x \in \mathbb{Z}$, on a $x^p \equiv x \ mod\ p$.*

(c) *dans $\mathbb{F}_p[X]$, on a la factorisation*

$$X^p - X = \prod_{\alpha \in \mathbb{F}_p} (X - \alpha).$$

□

2.7. Théorème des restes chinois

Le problème est de résoudre des congruences simultanées

$$x \equiv a_1 \mod \langle n_1 \rangle, \quad x \equiv a_2 \mod \langle n_2 \rangle, \quad \dots \quad , \quad x \equiv a_s \mod \langle n_s \rangle,$$

dans $\mathbb{Z}$, ou plus généralement dans un *anneau principal* A. Les références historiques attestent que les mathématiciens chinois se posaient déjà ce genre des questions au début de notre ère, probablement pour traiter des problèmes astronomiques ou calendaires ; les documents trouvés remontent au moins au III$^{\text{e}}$ siècle après J.C. (Sunzi Suanjing, "The mathematical classic of Master Sun"). Il semble que les Grecs aient aussi étudié cette question, peut-être même antérieurement aux Chinois. On commence par traiter le cas plus simple $s = 2$.

2.7.1. Théorème des restes chinois. *On se donne $n_1, n_2 \neq 0$ dans un anneau principal A. Si* $\operatorname{pgcd}(n_1, n_2) = 1$, *les congruences simultanées*

$$x \equiv a_1 \mod \langle n_1 \rangle, \quad x \equiv a_2 \mod \langle n_2 \rangle$$

se résolvent comme l'ensemble des $x \in A$ satisfaisant une unique congruence

$$x \equiv x_0 \mod \langle n_1 n_2 \rangle,$$

où x_0 dépend des données a_1, a_2 (et n_1, n_2).

Démonstration. Démontrons d'abord l'existence d'une solution x_0. D'après l'identité de Bézout, on peut trouver $\lambda_1, \lambda_2 \in A$ tels que $\lambda_1 n_1 + \lambda_2 n_2 = 1$ (et si A est euclidien, on peut utiliser l'algorithme d'Euclide pour trouver λ_1, λ_2). Cette relation implique de façon évidente que

$$\begin{aligned} &\lambda_2 n_2 \equiv 1 \mod \langle n_1 \rangle, \quad \lambda_2 n_2 \equiv 0 \mod \langle n_2 \rangle, \\ &\lambda_1 n_1 \equiv 0 \mod \langle n_1 \rangle, \quad \lambda_1 n_1 \equiv 1 \mod \langle n_2 \rangle. \end{aligned}$$

Par conséquent, si on prend la combinaison linéaire

$$x_0 = a_1(\lambda_2 n_2) + a_2(\lambda_1 n_1),$$

on trouve bien

$$x_0 \equiv a_1 \mod \langle n_1 \rangle, \quad x_0 \equiv a_2 \mod \langle n_2 \rangle.$$

Maintenant, si $x \in A$ est une autre solution, on trouve par différence

$$x - x_0 \equiv 0 \mod \langle n_1 \rangle, \quad x - x_0 \equiv 0 \mod \langle n_2 \rangle,$$

c'est-à-dire que $x - x_0$ doit être multiple de n_1 et n_2. Il doit donc être multiple de $\mathrm{ppcm}(n_1, n_2) = n_1 n_2$ (du fait que $\mathrm{pgcd}(n_1, n_2) = 1$, cf. 2.4.10 (c)). Ceci implique $x \equiv x_0 \mod \langle n_1 n_2 \rangle$, et réciproquement de tels éléments x sont bien des solutions du problème, puisque $x \equiv x_0 \mod \langle n_1 n_2 \rangle$ implique $x \equiv x_0 \equiv a_1 \mod \langle n_1 \rangle$ et $x \equiv x_0 \equiv a_2 \mod \langle n_2 \rangle$. □

2.7.2. Formulation "moderne". Une façon beaucoup plus moderne (20$^{\mathrm{e}}$ siècle), et également plus abstraite, de formuler le théorème des restes chinois est d'écrire que si $\mathrm{pgcd}(n_1, n_2) = 1$, on a un isomorphisme d'anneaux

$$\varphi : A/n_1 n_2 A \to A/n_1 A \times A/n_2 A, \quad \overset{\bullet}{x} \mod \langle n_1 n_2 \rangle \mapsto (\overset{\bullet}{x} \mod \langle n_1 \rangle, \overset{\bullet}{x} \mod \langle n_2 \rangle)$$

[Si A_1 et A_2 sont des anneaux, on munit $A_1 \times A_2$ d'une structure d'anneau en posant $(x, y) + (x', y') = (x + x', y + y')$ et $(x, y)(x', y') = (xx', yy')$; on a $1_{A_1 \times A_2} = (1_{A_1}, 1_{A_2})$]. Il est ici évident que φ est un morphisme, et c'est sa bijectivité qui est équivalente au théorème des restes chinois (la surjectivité exprimant l'existence de la solution x_0, et l'injectivité son unicité modulo $\langle n_1 n_2 \rangle$). La preuve du théorème fournit en outre une formule explicite pour l'isomorphisme inverse φ^{-1} : si $\lambda_1 n_1 + \lambda_2 n_2 = 1$, cet inverse est donné par

$$\begin{aligned} \varphi^{-1} : \quad & A/n_1 A \times A/n_2 A \quad \longrightarrow \quad A/n_1 n_2 A, \\ & (\overset{\bullet}{a_1} \mod \langle n_1 \rangle, \overset{\bullet}{a_2} \mod \langle n_2 \rangle) \longmapsto \big(a_1(\lambda_2 n_2) + a_2(\lambda_1 n_1)\big)^{\bullet} \mod \langle n_1 n_2 \rangle. \end{aligned}$$

2.7.3. Exemples. (a) Dans $\mathbb{Z}$, les conguences simultanées

$$x \equiv 2 \mod \langle 4 \rangle, \quad x \equiv 3 \mod \langle 6 \rangle$$

sont incompatibles. En effet la première implique x pair et la seconde x impair. Mais cela ne contredit pas le théorème des restes chinois puisque $\mathrm{pgcd}(4, 6) = 2 \neq 1$.

(b) Soit à résoudre dans $\mathbb{Q}[X]$ les congruences simultanées

$$P \equiv X \mod \langle X^3 - 1 \rangle, \quad P \equiv -X^2 \mod \langle X^5 + 1 \rangle.$$

On commence par calculer $\mathrm{pgcd}(X^3 - 1, X^5 + 1)$ par l'algorithme d'Euclide :

$$\begin{aligned} X^5 + 1 &= (X^3 - 1)X^2 + (X^2 + 1), \\ X^3 - 1 &= (X^2 + 1)X - (X + 1), \\ X^2 + 1 &= (X + 1)(X - 1) + 2, \\ X + 1 &= 2\left(\frac{1}{2}X + \frac{1}{2}\right) + 0, \end{aligned}$$

ce qui implique bien $\operatorname{pgcd}(X^5+1, X^3-1) = 1$ (à l'élément inversible $2 \in \mathbb{Q}^*$ près ; dans le calcul ci-dessus, on a aussi remplacé $-(X+1)$ par $X+1$, -1 étant inversible). On voit que

$$\begin{aligned} 2 &= (X^2+1) - (X+1)(X-1) = (X^2+1) + \big((X^3-1) - (X^2+1)X\big)(X-1) \\ &= (-X^2+X+1)(X^2+1) + (X-1)(X^3-1) \\ &= (-X^2+X+1)\big((X^5+1) - (X^3-1)X^2\big) + (X-1)(X^3-1) \\ &= (-X^2+X+1)(X^5+1) + (X^4-X^3-X^2+X-1)(X^3-1). \end{aligned}$$

(et on doit diviser par 2 pour obtenir la constante 1). D'après la formule vue dans la démonstration du théorème 2.7.1, une solution du problème est alors

$$\begin{aligned} P_0 &= \frac{1}{2}\Big(X(-X^2+X+1)(X^5+1) + (-X^2)(X^4-X^3-X^2+X-1)(X^3-1)\Big) \\ &= \frac{1}{2}\Big(-X^9+2\,X^7+X^6-X^4+X\Big) \equiv X^7 \mod \langle (X^3-1)(X^5+1)\rangle, \end{aligned}$$

et la solution générale est donc

$$P \equiv X^7 \mod \langle (X^3-1)(X^5+1)\rangle.$$

On peut aisément le vérifier a posteriori :

$$X^7 = X(X^3+1)(X^3-1) + X, \quad X^7 = X^2(X^5+1) - X^2.$$

2.7.4. Généralisation. *On se donne $n_1, \dots, n_s \neq 0$ dans un anneau principal A. Si $\operatorname{pgcd}(n_i, n_j) = 1$ pour tous $i \neq j$, les congruences simultanées*

$$x \equiv a_1 \mod \langle n_1\rangle, \quad x \equiv a_2 \mod \langle n_2\rangle, \quad \dots \quad , \quad x \equiv a_s \mod \langle n_s\rangle$$

se résolvent comme l'ensemble des $x \in A$ satisfaisant une unique congruence

$$x \equiv x_0 \mod \langle n_1 n_2 \cdots n_s\rangle,$$

où x_0 dépend des données $a_1, \dots, a_s$ (et $n_1, \dots, n_s$). En d'autres termes, on a un isomorphisme d'anneaux

$$\begin{aligned} &\varphi : A/(n_1\cdots n_s)A \longrightarrow A/n_1A \times \cdots \times A/n_sA, \\ &\dot{x} \mod \langle n_1\cdots n_s\rangle \longmapsto (\dot{x} \mod \langle n_j\rangle)_{1\leqslant j\leqslant s}. \end{aligned}$$

Démonstration. On raisonne par récurrence sur $s \geqslant 3$, le problème ayant déjà été traité pour $s = 2$. Les deux dernières congruences $x \equiv a_{s-1} \mod \langle n_{s-1}\rangle$, $x \equiv a_s \mod \langle n_s\rangle$ équivalent à une seule congruence $x \equiv a'_{s-1} \mod \langle n_{s-1}n_s\rangle$ d'après le cas $s = 2$, et le lemme de Gauss implique que l'on a bien $\operatorname{pgcd}(n_i, n_{s-1}n_s) = 1$ pour $i < s-1$. On peut donc appliquer l'hypothèse de récurrence pour $s-1$ congruences, respectivement modulo $n_1, n_2, \dots, n_{s-2}, n'_{s-1} = n_{s-1}n_s$, et conclure alors que la solution s'exprime sous forme d'une unique congruence $x \equiv x_0 \mod \langle n_1\cdots n_{s-2}(n_{s-1}n_s)\rangle$. □

2.8. Corps de décomposition*

Nous démontrons ici quelques résultats importants en théorie des corps, qui peuvent être obtenus comme des conséquences assez directes du théorème 2.6.7 dans le cas des anneaux de polynômes. Cette section (quoiqu'en principe accessible sans trop de difficultés au moyen des résultats qui précèdent) est hors programme, et réservée aux étudiants souhaitant étendre un peu leur culture mathématique.

2.8.1. Définition. *Étant donné un corps $\mathbb{K}$, on appelle extension de $\mathbb{K}$ tout corps $\mathbb{L}$ qui contient $\mathbb{K}$ comme sous-corps, et on écrira $\mathbb{L} \supset \mathbb{K}$ pour indiquer cette situation.*

Un exemple important bien connu est $\mathbb{C} \supset \mathbb{R}$.

2.8.2. Théorème. *Soit P un polynôme irréductible de degré d d'un anneau $\mathbb{K}[X]$. Alors l'anneau quotient $\mathbb{L} = \mathbb{K}[X]/\langle P\rangle$ est un corps. De plus $\mathbb{K}$ s'identifie à un sous-corps de $\mathbb{L}$, via le morphisme injectif $\mathbb{K} \hookrightarrow \mathbb{L}$, $c \mapsto \dot{c}$. Le corps $\mathbb{L}$ peut ainsi être considéré comme une extension de $\mathbb{K}$. C'est aussi un $\mathbb{K}$-espace vectoriel de dimension $\dim_{\mathbb{K}} \mathbb{L} = d$, admettant comme base*

$$(\dot{1}, \dot{X}, \dots, \dot{X}^{d-1}).$$

Démonstration. Comme $\mathbb{K}[X]$ est principal, le fait que $\mathbb{L}$ soit un corps résulte du théorème 2.6.7. L'application $\varphi : \mathbb{K} \to \mathbb{L}$, $c \mapsto \dot{c}$ est bien injective car $\mathrm{Ker}(\varphi) = \{0\}$: pour tout $c \in \mathbb{K}^*$, $\dot{c} \neq \dot{0}$, car c ne peut être multiple de P qui est de degré $d \geqslant 1$. Supposons P unitaire (ce n'est pas restrictif), et écrivons

$$P(X) = X^d + a_{d-1}X^{d-1} + \cdots + a_1 X + a_0, \quad d \geqslant 1.$$

En prenant les classes modulo $\langle P\rangle$, on en déduit

$$\dot{X}^d = -\dot{a}_0\dot{1} - \dot{a}_1\dot{X} - \cdots - \dot{a}_{d-1}\dot{X}^{d-1},$$

et en multipliant par $\dot{X}^{n-d}$ pour $n \geqslant d$, on voit que

$$\dot{X}^n = -\dot{a}_0\dot{X}^{n-d} - \dot{a}_1\dot{X}^{n-d+1} - \cdots - \dot{a}_{d-1}\dot{X}^{n-1}, \quad \forall n \geqslant d.$$

Ceci entraîne que modulo P, la classe $\dot{G}$ de tout polynôme $G = \sum_{j=0}^{n} b_j X^j \in \mathbb{K}[X]$ est égale à une combinaison linéaire des éléments de la famille

$$(\dot{1}, \dot{X}, \dots, \dot{X}^{d-1}),$$

de sorte que celle-ci est une famille génératrice de $\mathbb{L}$, vu comme espace vectoriel sur $\mathbb{K}$. Mais si la famille était liée, il existerait des coefficients $\lambda_0, \lambda_1, \dots, \lambda_{d-1} \in \mathbb{K}$ non tous nuls tels que

$$\dot{\lambda}_0\dot{1} + \dot{\lambda}_1\dot{X} + \cdots + \dot{\lambda}_{d-1}\dot{X}^{d-1} = \dot{0},$$

ce qui signifie précisément que $\lambda_0 + \lambda_1 X + \cdots + \lambda_{d-1} X^{d-1}$ est divisible par P. Ceci est impossible puisque P est de degré d. Par conséquent la famille considérée est bien une base, et le théorème est démontré. (On remarquera que le cas $d = 1$ n'est pas très intéressant, on a alors $\mathbb{L} \simeq \mathbb{K}$). □

2.8.3. Exemple. Le polynôme $P(X) = X^2 + 1 \in \mathbb{R}[X]$ est irréductible, donc $\mathbb{R}[X]/\langle X^2+1\rangle$ est un corps. Le théorème ci-dessus montre qu'il admet pour base $(\dot{1}, \dot{X})$ sur $\mathbb{R}$ et qu'on a la relation $(\dot{X})^2 = -\dot{1}$. C'est précisément le corps des complexes ! En fait cette méthode est probablement la méthode la plus fondamentale qui existe pour définir $\mathbb{C}$, fournissant un procédé algébrique naturel pour ajouter une racine manquante à l'équation $X^2 + 1 = 0$.

Le procédé précédent s'applique de manière générale à tout corps $\mathbb{K}$ et tout polynôme $G \in \mathbb{K}[X]$. De façon précise, on va montrer :

2.8.4. Théorème. *Pour tout corps $\mathbb{K}$ et tout polynôme $G \in \mathbb{K}[X]$, il existe une extension $\mathbb{L} \supset \mathbb{K}$ dans laquelle le polynôme G est complètement scindé, à savoir qu'on peut trouver des éléments $w_j \in \mathbb{L}$ tels que*

$$G(X) = a_d \prod (X - w_j)^{m_j} \quad \text{dans } \mathbb{L}[X] \supset \mathbb{K}[X],$$

où a_d est le coefficient dominant de G.

Démonstration. On démontre le résultat par récurrence sur $d = \deg(G)$. Si $d = 1$, le polynôme P est déjà scindé, on prend $\mathbb{L} = \mathbb{K}$ et il n'y a rien à montrer. Supposons le résultat déjà démontré pour $d - 1$ et prenons $G \in \mathbb{K}[X]$ de degré $\deg(G) = d$. On commence par le factoriser en ses facteurs irréductibles, soit

$$G = a_d P_1^{k_1} \dots P_s^{k_s}.$$

On pose $\mathbb{L}_1 = \mathbb{K}[Y]/\langle P_1(Y)\rangle$. Ce quotient est une extension $\mathbb{L}_1 \supset \mathbb{K}$ d'après le théorème précédent. Par construction P_1 admet la racine $w_1 = \dot{Y}$ dans $\mathbb{L}_1$, de sorte que $w_1 \in \mathbb{L}_1$ est aussi une racine de $G \in \mathbb{K}[X] \subset \mathbb{L}_1[X]$. Par conséquent, on peut écrire $G(X) = (X - w_1)H(X)$ avec $H \in \mathbb{L}_1[X]$ de degré $d - 1$. L'hypothèse de récurrence entraîne qu'il existe une extension $\mathbb{L} \supset \mathbb{L}_1 \supset \mathbb{K}$ dans laquelle H est complètement scindé, et donc G aussi. Le théorème s'ensuit. □

2.8.5. Remarque. la démonstration précédente construit en fait la plus petite extension $\mathbb{L} \supset \mathbb{K}$ dans laquelle $G \in \mathbb{K}[X]$ puisse se scinder. Un tel corps $\mathbb{L}$ est appelé *corps de décomposition* de G. On peut montrer qu'il est unique à isomorphisme de corps près.

2.8.6. Exercice.**

(a) Soit $\mathbb{K}$ un corps de caractéristique $p > 0$. Montrer par la formule du binôme que $F : x \mapsto x^p$ est un morphisme de corps ("morphisme de Frobenius"). Lorsque $\mathbb{K} = \mathbb{F}_p$, montrer que $F = \mathrm{Id}$ (calculer $F(1), F(2) = F(1+1), \dots$) et retrouver ainsi le petit théorème de Fermat.

(b) Soit de nouveau $\mathbb{K}$ un corps de caractéristique $p > 0$. Déduire de (a) que l'application $F^n = F \circ \cdots \circ F : x \mapsto x^{p^n}$ est aussi un morphisme de corps, puis que $\mathbb{K}_n = \{x \in \mathbb{K} \,/\, F^n(x) = x\}$ est un sous-corps de $\mathbb{K}$ possédant au plus p^n éléments.

(c) Montrer que le corps de décomposition $\mathbb{L}$ du polynôme $G = X^{p^n} - X \in \mathbb{F}_p[X]$ coïncide avec $\mathbb{L}_n$ (qui n'est autre que l'ensemble des racines du polynôme G), et que les racines sont simples [raisonner sur la dérivée]. En déduire que $\mathbb{L}$ est un corps ayant exactement $q = p^n$ éléments.

Ce corps est traditionnellement noté $\mathbb{F}_q$.

Nota : le théorème 2.6.10 entraîne assez aisément que tout corps $\mathbb{K}$ à $q = p^n$ éléments est isomorphe à $\mathbb{F}_q$; on a ainsi déterminé tous les corps finis commutatifs ; mais un théorème célèbre démontré en 1905 par le mathématicien écossais Joseph Wedderburn (1882-1948) énonce qu'il n'y a pas de corps finis non commutatifs. Les corps finis sont très importants en pratique, ils sont par exemple utilisés pour les codes correcteurs d'erreurs en informatique et en téléphonie mobile, tels que le code de Reed-Solomon s'appuyant sur le corps $\mathbb{F}_{256}$.