THÉORIE DES NOMBRES

Cours à l'Université de Rennes 1 (2007–2008)

Antoine Chambert-Loir

Antoine Chambert-Loir
IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex.
E-mail : antoine.chambert-loir@univ-rennes1.fr
Url : http://perso.univ-rennes1.fr/antoine.chambert-loir

Version du 28 septembre 2016
La version la plus à jour est disponible sur le Web à l'adresse http://webusers.imj-prg.fr/antoine.chambert-loir/enseignement/2007-08/h4/coursh4.pdf

SOMMAIRE

CHAPITRE 1

NOMBRES PREMIERS

§1.1. Rappels : factorisation, théorème d'Euclide

A. Définition, crible d'Ératosthène

Le premier chapitre de ce cours de théorie des nombres concerne les *nombres premiers* ; rappelons qu'on dit qu'un nombre entier $p > 1$ est un nombre premier s'il n'est divisible par aucun autre nombre entier positif que 1 et lui-même. On notera $\mathscr{P}$ l'ensemble des nombres premiers.

Il convient de remarquer que tout nombre entier strictement supérieur à 1 est multiple d'un nombre premier : c'est clair s'il est lui-même premier et dans le cas contraire, c'est un multiple de deux nombres entiers strictement plus petits qui sont par récurrence multiples d'un nombre premier. On peut encore raffiner cette remarque : tout nombre entier n strictement supérieur à 1 qui n'est pas premier est divisible par un nombre premier au plus égal à $\sqrt{n}$.

Cette remarque faite, on peut alors commencer à énumérer les éléments de $\mathscr{P}$ à l'aide du *crible d'Eratosthène* qui consiste à écrire les nombres entiers $2, \ldots$, jusqu'à un certain point, disons 20 :

$$2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20$$

puis à raisonner de la façon suivante. L'entier 2 est premier, donc ses multiples $4,6,\ldots$ ne le sont pas ; on les raye :

$$\mathbf{2},3,\cancel{4},5,\cancel{6},7,\cancel{8},9,\cancel{10},11,\cancel{12},13,\cancel{14},15,\cancel{16},17,\cancel{18},19,\cancel{20}.$$

Après 2, le premier entier non rayé, en l'occurence 3, est forcément un nombre premier : dans le cas contraire, il aurait été multiple d'un nombre premier déjà détecté ; on le sélectionne et on raye ses multiples $6,9,\ldots$:

$$\mathbf{2},\mathbf{3},\cancel{4},5,\cancel{6},7,\cancel{8},\cancel{9},\cancel{10},11,\cancel{12},13,\cancel{14},\cancel{15},\cancel{16},17,\cancel{18},19,\cancel{20}.$$

De même, 5 est premier, on raye ses multiples :

$$\mathbf{2},\mathbf{3},\cancel{4},\mathbf{5},\cancel{6},7,\cancel{8},\cancel{9},\cancel{10},11,\cancel{12},\cancel{13},\cancel{14},\cancel{15},\cancel{16},17,\cancel{18},19,\cancel{20}.$$

On constate cependant qu'on n'a rien rayé de nouveau. En effet, $5^2 > 20$ donc les nombres entiers inférieurs à 25 sont ou bien premiers, ou bien sont divisibles par un nombre premier au plus égal à 5. Par le même argument, tous ceux qui restent sont des nombres premiers, d'où le début de l'énumération :

$$\mathscr{P} \cap [1,20] = \{2,3,5,7,11,13,17,19\}.$$

B. Lemme d'Euclide

L'outil incontournable pour l'étude de l'arithmétique des nombres entiers est la *division euclidienne* : si a et b sont des nombres entiers relatifs, avec $b \neq 0$, il existe un unique couple (q,r) formé d'entiers relatifs tels que $0 \leqslant r < |b|$ et $a = bq + r$. On dit que q est le quotient de la division euclidienne de a par b, et que r est le reste. On résumera souvent ces propriétés en disant : « soit $a = bq + r$ la division euclidienne de a par b », ce qui sous-entendra que q est le quotient et r le reste.

La démonstration de ce fait, par récurrence sur a, est supposée connue.

LEMME 1.1.1 (Euclide). — *Soit a et b des nombres entiers et soit p un nombre premier qui divise le produit ab. Alors p divise a, ou p divise b.*

Démonstration. — Supposons que p ne divise pas a et définissons un entier n comme le plus petit entier strictement positif tel que p divise an. Par hypothèse, on a $n > 1$. Puisque p divise évidemment ap, on a $n \leqslant p$. Soit $p = nq + r$ la division euclidienne de p par n ; comme $ar = a(p - nq) = ap - (an)q$ est multiple de p et que $0 \leqslant r < n$, la minimalité de n implique que $r = 0$. Autrement dit, n divise p. Puisque p est un nombre premier et que $n > 1$, on a nécessairement $n = p$.

Puisque p divise ab, on a aussi $p \leqslant b$. Soit $b = pq + r$ la division euclidienne de b par p. Comme précédemment, $ar = ab - p(aq)$ est multiple de p et puisque $0 \leqslant r < p = n$, la minimalité de n entraîne $r = 0$. Autrement dit, p divise b, ce qu'il fallait démontrer. □

C'est en substance la démonstration originelle d'Euclide et un peu d'algèbre permet de la simplifier. L'ensemble des entiers relatifs n de $\mathbf{Z}$ tels que p divise an que l'on a introduit dans la démonstration est stable par somme et différence (et n'est pas vide) : c'est un sous-groupe de $\mathbf{Z}$ (pour l'addition). Or, la propriété de division euclidienne a pour conséquence que tout sous-groupe non-nul G de $\mathbf{Z}$ est l'ensemble des multiples d'un unique élément strictement positif de G, à savoir son plus petit élément strictement positif.

On peut alors reprendre la démonstration précédente :

Démonstration (en termes de sous-groupes de $\mathbf{Z}$). — L'ensemble G des entiers k de $\mathbf{Z}$ tels que p divise ak est un sous-groupe de $\mathbf{Z}$, non réduit à 0 puisqu'il contient p ; il est donc de la forme $n\mathbf{Z}$, avec $n > 0$. Supposons que p ne divise pas a ; on a alors $n > 1$.

Mais comme $p \in G$, p est multiple de n, donc $n = p$. Par ailleurs, b appartient à G par hypothèse, donc b est multiple de p. □

Si a et b sont des entiers (pas tous deux nuls), le plus petit élément strictement positif du sous-groupe $\langle a, b\rangle$ de $\mathbf{Z}$ est appelé leur plus grand diviseur commun (pgcd). Il se calcule via l'algorithme d'Euclide ; une extension de cet algorithme l'écrit explicitement comme combinaison de a et b :

PROPOSITION 1.1.2 (Algorithme d'Euclide étendu). — *Soit a et b des entiers strictement positifs. On définit une suite (a_n, u_n, v_n) de triplets d'entiers relatifs en posant*

$$(a_0, u_0, v_0) = (a, 1, 0), \quad (a_1, u_1, v_1) = (b, 0, 1),$$

et, si $a_n \neq 0$,

$$(a_{n+1}, u_{n+1}, v_{n+1}) = (a_{n-1} - q_n a_n, u_{n-1} - q_n u_n, v_{n-1} - q_n v_n),$$

où q_n désigne le quotient de la division euclidienne de a_{n-1} par a_n, de sorte que a_{n+1} soit son reste. Cette suite est finie : la suite (a_n) est strictement décroissante. Si $a_{n+1} = 0$, alors a_n est le pgcd de a et b et l'on a $a_n = au_n + bv_n$.

Une relation de la forme $d = au + bv$, où $d = \mathrm{pgcd}(a, b)$, est appelée relation de Bézout pour a et b. Si $\mathrm{pgcd}(a, b) = 1$, on dit que *a et b sont premiers entre eux.*

Démonstration. — L'inégalité $0 \leqslant a_{n+1} < a_n$ découle de la définition d'une division euclidienne, si bien que la suite (a_n) est strictement décroissante ; comme c'est une suite d'entiers naturels, elle est par conséquent finie. Pour $n = 0$ ou 1, on a l'égalité $a_n = au_n + bv_n$; si cette égalité est vraie jusqu'au rang n et que $a_n \neq 0$, alors

$$\begin{aligned} au_{n+1} + bv_{n+1} &= a(u_n - q_n u_{n-1}) + b(v_n - q_n u_{n-1}) \\ &= (au_n + bv_n) - q_n(au_{n-1} + bv_{n-1}) = a_n - q_n a_{n-1} = a_{n+1}, \end{aligned}$$

donc elle est vraie au rang $n + 1$.

Si $a_{m+1} = 0$, posons $d = a_m$. Il reste à démontrer que d est le pgcd de a et b. Montrons que d divise a_n pour tout n par récurrence descendente sur n ; c'est vrai pour $n = m$ et $n = m + 1$. Si c'est vrai pour tout entier au moins égal à n, l'égalité $a_{n-1} = a_{n+1} + a_n q_n$ montre que c'est aussi vrai pour $n - 1$. En particulier, d divise $a_0 = a$ et $a_1 = b$. Inversement, la relation $d = a_m = au_m + bv_m$ montre que tout diviseur commun à a et b divise d, et est en particulier inférieur ou égal à d. Par suite, d est le plus grand diviseur commun de a et b.

Nous avons vu que $d = au_m + bv_m$ appartient au sous-groupe $\langle a, b\rangle$. Inversement, tout élément strictement positif du groupe $\langle a, b\rangle$ est divisible par d donc est au moins égal à d. □

Le lemme d'Euclide s'exprime en termes de l'anneau quotient $\mathbf{Z}/p\mathbf{Z}$ de la façon suivante :

PROPOSITION 1.1.3. — *Soit n un entier tel que $n > 1$.*

1) Soit a un nombre entier. Les propriétés suivantes sont équivalentes :

(i) les entiers a et n sont premiers entre eux ;

(ii) la classe de a modulo n est inversible dans $\mathbf{Z}/n\mathbf{Z}$;

(iii) la classe de a modulo n est simplifiable dans $\mathbf{Z}/n\mathbf{Z}$.

2) Les propriétés suivantes sont équivalentes :

(i) L'entier n est un nombre premier ;

(ii) l'anneau $\mathbf{Z}/n\mathbf{Z}$ est un corps ;

(iii) l'anneau $\mathbf{Z}/n\mathbf{Z}$ est intègre ;

Démonstration. — 1) Supposons les entiers a et n premiers entre eux et soit $1 = au + vn$ une relation de Bézout. On a donc $\bar{a}\bar{u} = 1$ dans $\mathbf{Z}/n\mathbf{Z}$, ce qui démontre que $\bar{a}$ est inversible dans $\mathbf{Z}/n\mathbf{Z}$, donc (i) entraîne (ii). L'implication (ii)⇒(iii) est évidente. Supposons maintenant que a et n ne soient pas premiers entre eux et notons d leur pgcd. On peut ainsi écrire $n = dm$, avec $1 \leqslant m < n$; comme a est multiple de d, am est multiple de n, d'où $\bar{a}\bar{m} = 0$ dans $\mathbf{Z}/n\mathbf{Z}$. Puisque $\bar{m} \neq 0$, $\bar{a}$ n'est pas simplifiable, ce qui démontre l'implication (iii)⇒(i) (ou plutôt sa contraposée).

2) Si n est un nombre premier, n est premier avec tout entier a tel que $1 \leqslant a < n$; par suite, tout élément non nul de $\mathbf{Z}/n\mathbf{Z}$ est inversible et $\mathbf{Z}/n\mathbf{Z}$ est un corps. Cela démontre l'implication (i)⇒(ii) et l'implication (ii)⇒(iii) est évidente.

Inversement, si n n'est pas premier, il existe des entiers a et b tels que $1 < a, b < n$ et $n = ab$. Le produit $\bar{a}\bar{b}$ des classes de a et b modulo n est donc nul bien qu'aucune des deux ne soit nulle. Autrement dit, $\mathbf{Z}/n\mathbf{Z}$ n'est pas intègre. □

Si $n \geqslant 1$, on note $\varphi(n)$ le nombre des entiers a dans l'intervalle $[1, n]$ qui sont premiers avec n.

On rappelle aussi le lemme chinois :

PROPOSITION 1.1.4. — *Soit $m_1, \ldots, m_n$ des entiers naturels non nuls ; posons $m = \prod_{i=1}^{n} m_i$. L'homomorphisme de $\mathbf{Z}$ dans $\prod_{i=1}^{n}(\mathbf{Z}/m_i\mathbf{Z})$ qui applique un entier a sur la famille $(a_i \mod m_i)$ passe au quotient et induit un homomorphisme*

$$\varphi : \mathbf{Z}/m\mathbf{Z}) \to \prod_{i=1}^{n}(\mathbf{Z}/m_i\mathbf{Z}).$$

Les conditions suivantes sont équivalentes :

(i) Les entiers $m_1, \ldots, m_n$ sont premiers entre eux deux à deux ;

(ii) L'homomorphisme φ est surjectif ;

(iii) L'homomorphisme φ est injectif ;

(iv) L'homomorphisme φ est un isomorphisme.

C. Factorisation

L'intérêt que l'on porte aux nombres premiers vient que tout nombre entier (strictement positif) s'écrit de manière essentiellement unique comme produit de nombres premiers :

THÉORÈME 1.1.1. — *Pour tout nombre entier naturel $n \geqslant 1$, il existe une unique suite finie $(p_1, \dots, p_m)$, croissante, de nombres premiers telle que $n = p_1 \dots p_m$.*

Démonstration. — Démontrons ce résultat par récurrence sur n.

Par convention, l'entier $n = 1$ est produit des termes de la suite vide () n'ayant aucun terme.

Supposons le résultat vrai pour tout entier $< n$ et soit p_1 le plus petit nombre premier qui divise n. On peut écrire $n = p_1 n'$, où n' est un nombre entier tel que $1 \leqslant n' < n$. Par récurrence, il existe une suite croissante $(p_2, \dots, p_m)$ de nombres premiers tels que $n' = p_2 \dots p_m$, d'où $n = p_1 \dots p_m$. En outre, p_2 divise n donc $p_1 \leqslant p_2$ d'après le choix de p_1, ce qui entraîne que la suite $(p_1, \dots, p_m)$ est croissante. Cela prouve l'*existence* d'une telle suite pour n.

Établissons son *unicité*. Soit en effet $n = q_1 \dots q_r$ une suite croissante de nombres premiers telle que $n = q_1 \dots q_r$. Puisque p_1 est le plus petit nombre premier divisant n, on a $p_1 \leqslant q_1$. D'autre part, le nombre premier p_1 divisant le produit $q_1 \dots q_r$ divise l'un des facteurs en vertu du lemme d'Euclide, donc est égal à l'un des facteurs, disons $p_1 = q_i$. On a alors $q_1 \leqslant q_i = p_1 \leqslant q_1$, ce qui implique $p_1 = q_1$. En simplifiant, on a $n' = q_2 \dots q_r = p_2 \dots p_m$. Par récurrence, les suites $(q_2, \dots, q_r)$ et $(p_2, \dots, p_m)$ sont égales ; on a donc $r = m$ et $(q_1, \dots, q_r) = (p_1, \dots, p_m)$. □

On peut aussi regrouper le produit précédent, nombre premier par nombre premier et en déduire l'énoncé : *Pour tout nombre entier naturel $n \geqslant 1$, il existe une unique suite strictement croissante $(p_1, \dots, p_m)$ de nombres premiers, et pour tout i, un entier strictement positif a_i, tel que $n = p_1^{a_1} \dots p_m^{a_m}$.* Enfin, il est agréable d'autoriser des exposants nuls, voire négatifs en faisant des fractions, de faire abstraction de l'ordre des facteurs, d'autoriser un signe, ce qui amène à énoncer le résultat sous la forme équivalente : *Pour tout nombre rationnel non nul x, il existe une unique famille $(a_p)_{p \in \mathscr{P}}$ d'entiers naturels, nuls sauf un nombre fini d'entre eux, et un unique élément $\varepsilon \in \{\pm 1\}$ tels que $x = \varepsilon \prod_{p \in \mathscr{P}} p^{a_p}$.*

L'entier a_p est appelé *valuation p-adique* du nombre rationnel x et est noté $\mathrm{ord}_p(x)$. On pose par convention $\mathrm{ord}_p(0) = +\infty$. La valuation p-adique vérifie les propriétés suivantes : pour x et $y \in \mathbf{Q}$, on a

$$\mathrm{ord}_p(xy) = \mathrm{ord}_p(x) + \mathrm{ord}_p(y) \quad \text{et} \quad \mathrm{ord}_p(x+y) \geqslant \min(\mathrm{ord}_p(x), \mathrm{ord}_p(y)).$$

On observera en outre que $\mathrm{ord}_p(x+y) = \min(\mathrm{ord}_p(x), \mathrm{ord}_p(y))$ lorsque $\mathrm{ord}_p(x) \neq \mathrm{ord}_p(y)$; en effet, si $\mathrm{ord}_p(x) < \mathrm{ord}_p(y)$, on a

$$\mathrm{ord}_p(y) > \mathrm{ord}_p(x) = \mathrm{ord}_p((x+y)-y) \geqslant \min(\mathrm{ord}_p(x+y), \mathrm{ord}_p(y))$$

ce qui impose $\min(\mathrm{ord}_p(x+y), \mathrm{ord}_p(y)) = \mathrm{ord}_p(x+y)$, d'où $\mathrm{ord}_p(x) \geqslant \mathrm{ord}_p(x+y)$ et donc l'égalité.

Pour qu'un nombre rationnel x non nul soit un entier relatif, il faut et il suffit que $\mathrm{ord}_p(x) \geqslant 0$ pour tout x.

D. Caractère infini de l'ensemble des nombres premiers

THÉORÈME 1.1.1 (Euclide). — *L'ensemble $\mathscr{P}$ des nombres premiers est infini.*

Démonstration. — Dans le cas contraire, $\mathscr{P}$ serait un ensemble fini que l'on pourrait énumérer : $\mathscr{P} = \{p_1, \ldots, p_m\}$. Soit alors N l'entier $p_1 \ldots p_m + 1$. Il vérifie $N > 1$ donc est multiple d'un nombre premier, c'est-à-dire de l'un des entiers $p_1, \ldots, p_m$, disons p_i. Par soustraction, l'entier $1 = N - p_1 \ldots p_m$ est aussi multiple de p_i, ce qui est absurde. □

Dans ce chapitre, nous allons préciser ce résultat dans deux directions. La première est quantitative : pour $x > 0$, posons $\pi(x) = \mathrm{card}(\mathscr{P} \cap [1, x])$, le nombre de nombres premiers au plus égaux à x. Puisque $\mathscr{P}$ est infini, on a $\lim_{x\to\infty} \pi(x) = +\infty$. Le *théorème des nombres premiers*, conjecturé par Euler, Legendre, Gauß dès le XVIII$^{\mathrm{e}}$ siècle mais démontré indépendamment par Jacques Hadamard et Charles-Louis de la Vallée-Poussin en 1896, fournit un équivalent de $\pi(x)$ en $+\infty$:

$$\pi(x) \sim \frac{x}{\log x}.$$

La seconde direction est plus qualitative : on peut s'intéresser aux nombres premiers p vérifiant une propriété de congruence, par exemple $p \equiv -1 \pmod 4$. Il est facile de voir que ce sous-ensemble est lui-même infini : en effet, si $p_1, \ldots, p_m$ sont des nombres premiers congrus à -1 modulo 4, considérons l'entier $N = 4p_1 \ldots p_m - 1$. Il est aussi congru à -1 modulo 4 donc n'a que des facteurs premiers impairs, et en a au moins un, disons p, qui est congru à -1 modulo 4. Ce nombre premier p n'est pas l'un des p_i, car sinon, $1 = 4p_1 \ldots p_m - N$ serait multiple de p_i. Cela fournit un autre nombre premier congru à -1 modulo 4 ; on voit de la sorte qu'il y en a une infinité. Le théorème de Dirichlet (1837) que nous démontrerons aussi affirme que dans toute progression arithmétique $\{kn + a, k \in \mathbf{N}\}$, il y a une infinité de nombres premiers dès lors que le premier terme a et la raison n sont premiers entre eux.

Qu'on ne se réjouisse pas trop vite : malgré la beauté de ces deux énoncés, les résultats démontrés concernant les nombres premiers sont largement en deçà des espérances des mathématiciens. On ne sait par exemple toujours pas s'il existe une infinité

de nombres premiers de la forme n^2+1, ou s'il existe une infinité de nombres premiers p tels que $p+2$ soit encore un nombre premier (problème des nombres premiers « jumeaux »)...

Exercices

1) a) Démontrer qu'il existe une infinité de nombres premiers congrus à -1 modulo 6. (*Si $p_1,\dots,p_r$ sont des nombres premiers, considérer un facteur premier de* $6p_1\dots p_r-1$.)

b) Soit n un entier strictement positif. Démontrer qu'il existe une infinité de nombres premiers qui ne sont pas congrus à 1 modulo n.

2) Soit f un polynôme non constant à coefficients entiers.

a) Montrer qu'il existe des entiers n arbitrairement grands tels que $f(n)$ ne soit pas un nombre premier.

b) Montrer que l'ensemble des nombres premiers qui divisent l'une des valeurs $f(n)$ de f, pour $n \geqslant 1$, est infini.

3) a) Montrer que si p est un nombre premier tel que $p \equiv 3 \pmod 4$, il n'existe pas d'entier $a \in \mathbf{Z}$ tel que a^2+1 soit multiple de p. (Indication : Que vaut a^{p-1} modulo p ?)

b) En déduire qu'il existe une infinité de nombres premiers congrus à 1 modulo 4.

4) Le but de cet exercice est de prouver, suivant V.A. Lebesgue, que l'équation $y^2 = x^3+7$ n'a pas de solutions en nombres entiers. On considère une telle solution (x,y).

a) Démontrer que x est impair et que y est pair.

b) À l'aide de la factorisation $x^3+8 = (x+2)(x^2-2x+4)$, démontrer que x^3+8 possède un facteur premier congru à 3 modulo 4.

c) Démontrer que chaque facteur premier de y^2+1 est congru à 1 modulo 4. Conclure.

5) Soit a et b deux entiers strictement positifs. On suppose que pour tout nombre premier p, le reste de la division euclidienne de a par p est inférieur ou égal au reste de la division euclidienne de b par p.

Soit $A = a! = 1\cdot 2\dots(a-1)a$ et $B = (b-a+1)\dots(b-1)b$, de sorte que $B/A = \binom{b}{a}$. Pour tout nombre premier p et tout entier k, on note $\alpha(p^k)$ et $\beta(p^k)$ le nombre de facteurs de A et B respectivement qui sont divisibles par p^k.

a) Montrer que $\mathrm{ord}_p(a) = \sum_{k=1}^{\infty} \alpha(p^k)$ et $\mathrm{ord}_p(b) = \sum_{k=1}^{\infty} \beta(p^k)$.

b) Montrer que $\left|\alpha(p^k) - \beta(p^k)\right| \leqslant 1$.

c) Montrer que $\alpha(p) \geqslant \beta(p)$.

d) Montrer que pour $p > a$, ni A ni B n'est divisible par p.

e) Soit $\kappa(p)$ le plus grand entier k tel que $\beta(p^k) > 0$. Montrer que B/A divise $\prod_{p\leqslant a} p^{\kappa(p)-1}$.

f) En déduire que $a = b$ (on commencera par démontrer que $b < 2a$).

6) Soit n un nombre entier $\geqslant 1$.

a) Montrer que l'on a

$$\varphi(n) = n\prod_{p\mid n}\frac{p-1}{p},$$

où le produit est étendu à l'ensemble des nombres premiers qui divisent n.

b) Soit a un nombre entier relatif et soit $d = \mathrm{pgcd}(a,n)$. Si $d = 1$, montrer que l'ordre de la classe $\bar{a}$ de a modulo n dans le groupe $\mathbf{Z}/n\mathbf{Z}$ est égal à n.

c) Montrer que l'ordre de $\bar{a}$ dans le groupe abélien $\mathbf{Z}/n\mathbf{Z}$ est égale à n/d.

d) Montrer que l'on a $\sum_{d|n}\varphi(d) = n$, où la somme est étendue à l'ensemble des nombres entiers d dans $[1,n]$ qui divisent n.

§1.2. Critères de primalité ou de non-primalité

Au delà de l'intérêt théorique que suscitent les nombres premiers, ils sont devenus un outil important pour certaines applications des mathématiques à la vie courante, notamment la cryptographie ou la signature électronique (Internet...) et les codes correcteurs d'erreurs utilisés en télécommunication numérique.

La production de nombres premiers « aléatoires » de grande taille est ainsi devenue une activité courante, rendant nécessaire le développement de méthodes efficaces permettant d'affirmer qu'un nombre entier donné n est, ou n'est pas, un nombre premier. La méthode naïve, tenter les divisions par tous les entiers inférieurs à $\sqrt{n}$, est rapidement impraticable si n a plus de quelques chiffres. Le bon paramètre est plutôt le logarithme de n, c'est-à-dire (à un facteur près) le nombre de chiffres de n écrit dans une base arbitraire.

Les algorithmes dont nous allons parler ici résultent le plus souvent de théorèmes qui prennent l'une des formes suivantes :

— *si n est premier, tout entier a tel que $1 \leqslant a \leqslant n$ vérifie une certaine propriété $P(a)$*. Dans ce cas, il suffit d'exhiber un entier a tel que la propriété ne soit pas satisfaite pour en déduire que n n'est pas premier ; un tel a sera appelé *témoin de non-primalité*. En revanche, savoir que tout entier a vérifie $P(a)$ n'implique pas que n soit un nombre premier.

— si n est premier, il existe un entier b, dans $[1,n]$ vérifiant une propriété $Q(b)$. On dira que b est un *témoin de primalité*. Le second cas est symétrique : exhiber un entier b vérifiant $Q(b)$ prouve que b est un nombre premier. Il n'est cependant pas raisonnable de tester tous les entiers successivement, à moins de savoir qu'il existe un entier b vérifiant $Q(b)$ qui soit de taille modérée. Si l'on sait qu'une proportion importante d'entiers b est constituée de de témoins de primalité, on peut tester la propriété $Q(b)$ sur des entiers b pris au hasard : si n est premier, la probabilité qu'aucun des entiers choisis ne soit un témoin de primalité, est égale à $(1-p)^{-k}$ si p est la densité des témoins de primalité et k le nombre d'essais, donc décroît très rapidement avec k. Le test obtenu est ainsi appelé *probabiliste*.

A. Critère de Fermat

Ce test repose sur le « petit théorème de Fermat » :

PROPOSITION 1.2.1. — *Soit p un nombre premier ; pour tout entier a qui n'est pas multiple de p, on a $a^{p-1} \equiv 1 \pmod{p}$.*

Démonstration. — L'ensemble des éléments non nuls du corps $\mathbf{Z}/p\mathbf{Z}$ est un groupe pour la multiplication ; son cardinal est $p-1$. Par conséquent, pour tout $\alpha \in (\mathbf{Z}/p\mathbf{Z})^\times$, $\alpha^{p-1} = 1$. Traduit en termes de nombres entiers, on obtient $a^{p-1} \equiv 1 \pmod{p}$ si a est un entier qui n'est pas multiple de p. □

Cette proposition fournit un test de non-primalité : si n est un nombre entier, il suffit d'exhiber un nombre entier a dans l'intervalle $[1, n-1]$ tel que $a^{r-1} \not\equiv 1 \pmod{n}$ pour en déduire que n n'est pas un nombre premier. On l'appelle le *test de Fermat*. On observera que la donnée d'un tel a prouve que n n'est pas premier sans pour autant l'écrire comme produit de deux entiers inférieurs à n.

Toutefois, il existe des nombres entiers n, appelés *nombres de Carmichael*, vérifiant la propriété suivante : tout entier a tel que $1 \leqslant a < n$ qui est premier à n vérifie $a^{n-1} \equiv 1 \pmod{n}$. Pour ces nombres entiers-là, les seuls témoins de non-primalité sont les entiers a qui ne sont pas premiers à n et fournissent ainsi automatiquement une factorisation partielle de n. Le plus petit d'entre eux, découvert justement par Robert Carmichael en 1910, est 561 (voir l'exercice 8). Un théorème démontré par les mathématiciens Alford, Granville et Pomerance en 1994 affirme qu'il existe une infinité de nombres de Carmichael.

Malgré ce défaut, le test de Fermat est utilisé par l'algorithme de cryptage PGP pour décider qu'un nombre entier donné est premier ; la probabilité qu'il échoue étant considérée comme négligeable.

En effet, si n n'est ni un nombre premier ni un nombre de Carmichael, on peut démontrer qu'au moins la moitié des nombres entiers a de $[1, n]$ qui sont premiers à n sont des témoins de non-primalité pour le test de Fermat (exercice 9).

B. Test de Miller-Rabin

C'est un raffinement du test de Fermat qui repose sur la remarque suivante : si p est un nombre premier et a un entier tel que $a^2 \equiv 1 \pmod{p}$, alors $a \equiv \pm 1 \pmod{p}$. En effet, l'hypothèse est que p divise $a^2 - 1 = (a-1)(a+1)$; il divise donc l'un des facteurs d'après le lemme d'Euclide.

PROPOSITION 1.2.1. — *Soit p un nombre premier impair et soit a un nombre premier non multiple de p. Notons $r = \mathrm{ord}_2(p-1)$, $m = (p-1)/2^r$ et soit $s \geqslant 0$ le plus petit entier tel que $a^{m2^s} \equiv 1 \pmod{p}$.*

Si $s > 0$, alors $a^{m2^{s-1}} \equiv -1 \pmod{p}$.

Soit n un nombre entier impair, $r = \mathrm{ord}_2(n-1)$ et $m = (n-1)/2^r$. Un témoin de non-primalité de Miller-Rabin pour n est un entier a tel que $a^{m2^{s-1}} \not\equiv -1 \pmod{n}$ mais $a^{m2^s} \equiv 1 \pmod{n}$. L'intérêt de ce test est qu'il est fiable : le phénomène des nombres

premiers de Carmichael ne se produit plus. Plus précisément, si n est un nombre entier impair tel que $n > 1$, au moins trois quarts des nombres entiers a dans $[1, n]$ sont des témoins de non-primalité de Miller-Rabin (exercice 10).

On dispose ainsi d'un *algorithme de primalité* de nature *probabiliste* : tirons k nombres entiers au hasard dans $[1, n]$; si n n'est pas premier, la probabilité qu'aucun ne soit un témoin de non-primalité est inférieure ou égale à $1/4^k$. Autrement dit, si l'on tire au hasard k nombres entiers et qu'aucun n'est un témoin de non-primalité de Miller-Rabin, il y a de fortes chances pour que n soit premier.

Si l'on admet l'hypothèse de Riemann généralisée, un théorème d'E. Bach (1990) affirme que dès qu'un entier $n > 1$ n'est pas un nombre premier, il y a un témoin de non-primalité inférieur à $2(\log n)^2$. Cela permet d'envisager essayer les entiers dans l'ordre et de tester si ce sont des témoins de non-primalité et fournit, sous l'hypothèse de Riemann généralisée, un algorithme déterministe pour décider de la primalité d'un entier n dont la complexité est polynomiale en $\log n$.

C. Critère de Lucas

Il repose sur les propriétés des groupes abéliens finis.

LEMME 1.2.1. — *Soit G un groupe abélien fini, noté multiplicativement. Soit g et h des éléments de G d'ordres m et n. Si m et n sont premiers entre eux, alors gh est d'ordre mn.*

Démonstration. — Observons que les sous-groupes $\langle g\rangle$ et $\langle h\rangle$ n'ont pas d'autre élément commun que 1. En effet, l'ordre d'un élément commun doit diviser m et n, donc est égal à 1. Par suite, si $(gh)^N = 1$, qu'on peut récrire $g^N = h^{-N}$ puisque G est commutatif. D'après la remarque faite, $g^N = h^{-N} = 1$; ainsi, N est multiple de m et de n, donc de mn. Inversement, $(gh)^{mn} = g^{mn}h^{mn} = 1$. □

PROPOSITION 1.2.2. — *Soit G un groupe abélien fini, noté multiplicativement. Pour toute partie S de G, il existe un élément de G dont l'ordre est le plus petit multiple commun des ordres des éléments de S. En particulier, il existe dans G un élément dont l'ordre est multiple de l'ordre de chaque élément de G.*

Démonstration. — Par récurrence, il suffit de démontrer si G possède deux éléments g et h d'ordres m et n, il possède un élément d'ordre $\mathrm{ppcm}(m, n)$. Comme $\mathrm{ord}_p(\mathrm{ppcm}(m, n)) = \max(\mathrm{ord}_p(m), \mathrm{ord}_p(n))$, on peut écrire $m = m'm''$ et $n = n'n''$, où les facteurs premiers de m' et n' sont ceux tels que $\mathrm{ord}_p(m) \geqslant \mathrm{ord}_p(n)$, ceux de m'' et n'' étant ceux tels que $\mathrm{ord}_p(n) > \mathrm{ord}_p(m)$. Alors, $g^{m''}$ est d'ordre m', tandis que $h^{n'}$ est d'ordre n''. Les entiers m' et n'' sont premiers entre eux, leur produit est égal au plus petit multiple comun de m et n ; d'après le lemme, l'ordre de $g^{m''}h^{n'}$ est égal à $\mathrm{ppcm}(m, n)$. □

Nous pouvons maintenant énoncer le théorème qui donne lieu au critère de primalité de Lucas-Lehmer.

THÉORÈME 1.2.3. — *Soit n un entier >1. Supposons que pour tout facteur premier p de $n-1$, il existe un entier a tel que $a^{n-1} \equiv 1 \pmod{n}$ et $a^{(n-1)/p} \not\equiv 1 \pmod{n}$. Alors, n est un nombre premier.*

Démonstration. — Si a est un entier tel que $a^{n-1} \equiv 1 \pmod{n}$, observons que a et n sont premiers entre eux. L'hypothèse signifie donc qu'il existe, pour tout facteur premier p de $n-1$, un élément a_p du groupe $(\mathbf{Z}/n\mathbf{Z})^\times$ dont l'ordre divse $n-1$ mais ne divise pas $(n-1)/p$. Le plus petit multiple commun des ordres de ces éléments a_p divise $n-1$ mais n'est pas un diviseur strict de $n-1$; il est donc égal à $n-1$. D'après la proposition 1.2.2, le groupe $(\mathbf{Z}/n\mathbf{Z})^\times$ possède un élément d'ordre $n-1$. Cela impose que $(\mathbf{Z}/n\mathbf{Z})^\times$ soit d'ordre au moins $n-1$, donc que n soit un nombre premier. □

Considérons par exemple l'entier $n = 571$; puisque l'on a

$$n-1 = 570 = 2 \times 3 \times 5 \times 19,$$

il suffit de trouver dans le groupe $(\mathbf{Z}/571\mathbf{Z})^*$ quatre éléments d'ordres ne divisant pas $570/p$ pour $p = 2,3,5,19$, autrement dit quatre entiers x_p tels que 571 divise $x_p^{570}-1$, tandis que $x_p^{570/p}$ et 571 soient premiers entre eux.

Lorsque $p = 2$, $570/p = 285$ et $x_2 = -1$ convient. Voyons si $x = 2$ convient : on a $570/3 = 190$, $570/5 = 114$ et $570/19 = 30$. On a respectivement

$$2^{570} \equiv 1, \quad 2^{190} \equiv 461, \quad 2^{114} \equiv 1, \quad 2^{30} \equiv 306 \pmod{571}.$$

En particulier, l'ordre de 2 dans $(\mathbf{Z}/571\mathbf{Z})^*$ divise 570 mais ne divise pas $570/p$ si $p = 3$ ou 19. Il reste à trouver un entier x convenant pour $p = 5$; comme $3^{114} \equiv 481 \pmod{571}$, il convient et cela conclut la démonstration que 571 est un nombre premier.

Le défaut de ce critère est de nécessiter la factorisation de $n-1$, mais des variantes permettent de se contenter d'une factorisation partielle, cf. l'exercice 11. Son intérêt est qu'il produit un *certificat de primalité* : la simple donnée des entiers $x_2 = -1$, $x_3 = 2$, $x_5 = 3$ et $x_{19} = 2$ permet à n'importe qui de vérifier rapidement que 571 est effectivement un nombre premier.

Il est néanmoins utile lorsque n est un nombre de Fermat, c'est-à-dire de la forme $2^{2^r}+1$; on obtient alors le critères de Pépin et de Proth :

COROLLAIRE 1.2.4 (Critères de Pépin–Proth). — *Soit r un nombre entier tel que $r \geqslant 2$. Pour que le nombre de Fermat $F_r = 2^{2^r}+1$. soit un nombre premier, il faut et il suffit que l'on ait l'une des congruences suivantes :*

a) $3^{(F_r-1)/2} \equiv -1 \pmod{F_r}$ *;*

b) $5^{(F_r-1)/2} \equiv -1 \pmod{F_r}$.

Démonstration. — Comme $F_r - 1 = 2^{2^r}$, la suffisance de ces conditions résulte du critère de Lucas. Supposons inversement que F_r soit un nombre premier et vérifions ces deux congruences. Elles signifient en fait que 3 et 5 ne sont pas des carrés modulo F_r et résultent immédiatement de la *loi de réciprocité quadratique* de Gauß compte tenu des congruences $F_r \equiv 1 + (-1)^{2^r} \equiv 2 \pmod 3$ et $F_r \equiv 1 + 4^{2^{r-1}} \equiv 1 + (-1)^{2^{r-1}} \equiv 2 \pmod 5$. □

Voir l'exercice 14 pour une démonstration des cas de la loi de réciprocité quadratique qui ont été utilisés dans la démonstration précédente.

Sous les hypothèses du critère de Lucas, le groupe multiplicatif $(\mathbf{Z}/n\mathbf{Z})^\times$ est cyclique. Il est important de savoir que cette propriété est vérifiée pour tout nombre premier puisque cela montre que les hypothèses de ce critère sont aussi nécessaires. Plus généralement :

THÉORÈME 1.2.5. — *Soit F un corps commutatif et soit G un sous-groupe fini de $F^\times$. Alors G est cyclique.*

Démonstration. — Soit g un élément de G dont l'ordre, disons n, est multiple de l'ordre de tout élément de G. Tout élément x de G vérifie $x^n = 1$. Comme F est un corps commutatif, l'équation polynomiale $x^n = 1$ dans F a au plus n solutions. Comme les n éléments $1, g, g^2, \ldots, g^{n-1}$ du sous-groupe de G engendré par g sont solutions, on a nécessairement $G = \langle g \rangle$. □

Seconde démonstration. — Soit n l'ordre de G. D'après le lemme et une récurrence évidente, il suffit de produire, pour tout facteur premier p de G, un élément g_p de G dont l'ordre est $p^{\mathrm{ord}_p(n)}$. En effet, le produit des g_p sera un élément de G d'ordre n.

Comme F est un corps commutatif, l'équation polynomiale $x^{n/p} = 1$ a au plus n/p solutions ; comme G est d'ordre $n > n/p$, il existe donc un élément x de G tel que $x^{n/p} \neq 1$. Posons alors $y = x^{n/p^r}$. On a $y^{p^r} = 1$ mais $y^{p^{r-1}} = x^{n/p} \neq 1$. Par suite, l'ordre de y est un diviseur de p^r qui ne divise pas p^{r-1} ; c'est donc p^r et on peut poser $g_p = y$. □

L'intérêt de la seconde démonstration est qu'elle se prête au calcul effectif d'un générateur. Montrons-le sur un exemple, disons le calcul d'un générateur du groupe multiplicatif $(\mathbf{Z}/71\mathbf{Z})^\times$. On a $71 - 1 = 70 = 2 \times 5 \times 7$; d'après la démonstration, nous devons trouver des éléments d'ordres 2, 5 et 7.

Il est clair que $g_2 = -1$ est d'ordre 2. Cherchons un élément d'ordre 5 ; il suffit de trouver un élément x tel que $x^{14} \not\equiv 1 \pmod{71}$. Essayons $x = 2$:

$$2^{14} \equiv 4^7 \equiv 4 \cdot (16)^3 \equiv 4 \cdot (16) \cdot (256) \equiv 64 \cdot 43 \equiv -7 \cdot 43 \equiv -301 \equiv -17 \pmod{71}.$$

Par suite, $g_5 = -17$ est d'ordre 5. De même, cherchons un élément x tel que $x^{10} \not\equiv 1 \pmod{71}$ et essayons encore $x = 2$:

$$2^{10} \equiv 16 \cdot 64 \equiv 16 \cdot (-7) \equiv -112 \equiv 30 \pmod{71}$$

si bien que $g_7 = 30$ est d'ordre 7. Leur produit

$$g_2 g_5 g_7 \equiv (-1) \times (-17) \times (30) \equiv 510 \equiv 13 \pmod{71}$$

est ainsi un générateur du groupe multiplicatif $(\mathbf{Z}/71\mathbf{Z})^\times$.

Remarque 1.2.6. — On peut raisonner à l'envers et se demander, un entier a étant donné, pour quels nombres premiers p est-ce que a est un générateur du groupe multiplicatif $(\mathbf{Z}/p\mathbf{Z})^*$. Pour qu'un tel nombre premier impair existe, une condition nécessaire évidente est que a ne soit pas un carré : si p est impair, les carrés de $(\mathbf{Z}/p\mathbf{Z})^*$ forment un sous-groupe strict. De même, $a = -1$ n'est un générateur de $(\mathbf{Z}/p^{\mathbf{Z}})^*$ que si $p = 2$ ou $p = 3$. Une conjecture du mathématicien Emil Artin est que ces deux conditions nécessaires sont *suffisantes* pour qu'il existe une infinité de tels nombres premiers ! Cette conjecture est toujours ouverte : on ne connaît à ce jour aucun nombre entier a pour lequel on sache qu'elle est vérifiée. Cependant, C. Hooley a démontré en 1967 qu'elle est vérifiée si l'hypothèse de Riemann généralisée l'est, tandis que R. Heath-Brown (1985), généralisant des travaux de R. Gupta et M. Ram Murty (1984), a montré qu'elle est vérifiée pour tout nombre premier a sauf au plus 2. Par exemple, des entiers 3, 5 et 7, au moins l'un est un générateur de $(\mathbf{Z}/p\mathbf{Z})^*$ pour une infinité de nombre premiers p.

Exercices

7) a) Soit n et m des nombres entiers. Montrer que le groupe $(\mathbf{Z}/n\mathbf{Z}) \times (\mathbf{Z}/m\mathbf{Z})$ est cyclique si et seulement si n et m sont premiers entre eux.

b) Si n est une puissance de 2, montrer que $(1+4x)^n \equiv 1 + 4nx$ modulo $8n$.

c) Soit a un entier $\geqslant 2$ et $n = 2^a$. Montrer que l'ordre de (la classe de) 5 dans $(\mathbf{Z}/n\mathbf{Z})^*$ est égal à 2^{a-2}. Le groupe $(\mathbf{Z}/n\mathbf{Z})^*$ est-il cyclique ?

d) Soit p un nombre premier impair, a un entier $\geqslant 2$ et $n = p^a$. Si x est un entier, montrer que $x^p \equiv x \pmod{p}$. Montrer par récurrence sur $a \geqslant 2$ que les conditions $x \equiv 1 \pmod{p^{a-1}}$ et $x^p \equiv 1 \pmod{p^a}$ sont équivalentes.

e) Soit x un entier dont la classe modulo p engendre $(\mathbf{Z}/p\mathbf{Z})^*$. Montrer que (la classe de) x engendre $(\mathbf{Z}/n\mathbf{Z})^*$ si et seulement si $x^{p-1} \not\equiv 1 \pmod{p^2}$.

f) Si $x^{p-1} \equiv 1 \pmod{p^2}$, montrer que $(x+p)^{p-1} \not\equiv 1 \pmod{p^2}$. En déduire qu'au moins l'un des deux, x ou $x+p$, est générateur de $(\mathbf{Z}/n\mathbf{Z})^*$.

g) Pour quels nombres entier n le groupe $(\mathbf{Z}/n\mathbf{Z})^*$ est-il cyclique ?

8) a) Soit $n > 1$ un nombre entier sans facteur carré tel que pour tout facteur premier p de n, $p-1$ divise $n-1$. Démontrer que n est un nombre premier ou un nombre de Carmichael.

b) En déduire que 561 est un nombre de Carmichael.

9) Soit n un nombre entier tel que $n > 1$. Soit p un facteur premier de n et soit $r = \mathrm{ord}_p(n)$; on pose $m = n/p^r$.

a) Soit a_p un générateur du groupe cyclique $(\mathbf{Z}/p^r\mathbf{Z})^\times$ et soit a un nombre entier tel que $a \equiv a_p \pmod{p^r}$ et $a \equiv 1 \pmod{m}$. Si $a^{n-1} \equiv 1 \pmod{n}$, démontrer que $r = 1$ et que $p-1$ divise $n-1$.

b) Si $r > 2$ ou si $p-1$ ne divise pas $n-1$, démontrer que pour au moins la moitié des éléments α de $(\mathbf{Z}/n\mathbf{Z})^\times$, $\alpha^{n-1} \neq 1$.

c) Démontrer que si n est un nombre de Carmichael, alors n est sans facteur carré et $p-1$ divise $n-1$ pour tout facteur premier p de n (*théorème de Korselt*, 1898).

d) Démontrer qu'un nombre de Carmichael est impair et a au moins trois facteurs premiers distincts.

e) Soit k un nombre entier tel que les trois entiers $6k+1$, $12k+1$ et $18k+1$ soient des nombres premiers. Montrer que leur produit $(6k+1)(12k+1)(18k+1)$ est un nombre de Carmichael (*Chernick*, 1939).

10) a) Montrer que dans un groupe cyclique G d'ordre pair, l'équation $2g = 0$ possède exactement deux solutions.

b) Soit n un nombre entier impair tel que > 1 et soit u le nombre de facteurs premiers distincts de n. Démontrer que dans le groupe $(\mathbf{Z}/n\mathbf{Z})^\times$, l'équation $\alpha^2 = 1$ possède exactement 2^u solutions.

c) On suppose que n n'est pas premier, c'est-à-dire que $u \geqslant 2$. Démontrer que parmi les nombres entiers a de $[1, n]$ qui sont premiers à n, les témoins de non-primalité de Miller-Rabin de n sont en proportion au moins égale à $1 - 2^{1-u} \geqslant 3/4$.

11) a) Soit N un entier tel que $N \geqslant 2$. Soit p un nombre premier qui divise $N-1$ et soit e le plus grand entier $\geqslant 1$ tel que p^e divise $N-1$. Soit a un entier tel que $a^{N-1} \equiv 1 \pmod{N}$ et tel que $a^{(N-1)/p} - 1$ et N soient premiers entre eux. Soit ℓ un nombre premier qui divise N.

Montrer que (la classe de) a est inversible dans $(\mathbf{Z}/\ell\mathbf{Z})$. Si t désigne son ordre, montrer les divisibilités $p^e | t$ et $t | \ell - 1$. En déduire que $\ell \equiv 1 \pmod{p^e}$ (*critère de Pocklington*).

b) Soit N un entier $\geqslant 2$. On écrit $N = 1 + uv$ et on suppose que pour tout nombre premier p qui divise u, il existe un entier a comme dans la question précédente. Montrer que tout facteur premier ℓ de N est congru à 1 modulo u. Si de plus $v \leqslant u+1$, en déduire que N est un nombre premier (*critère de primalité de Pocklington-Lehmer*).

12) a) Si $2^n + 1$ est un nombre premier, montrer que n est une puissance de 2.

b) On pose $F_n = 2^{2^n} + 1$ (*nombres de Fermat*). Montrer que F_n est premier pour $n = 0, 1, 2, 3, 4$.

c) Si $n \neq m$, montrer que F_n et F_m sont premiers entre eux ; écrire une relation de Bézout.

d) Montrer que 641 divise F_5. Cette divisibilité a été découverte par Euler, contredisant ainsi une affirmation de Fermat selon laquelle les F_n seraient tous des nombres premiers. (*Écrire* $2^{32} + 1 = 16(2^7)^4 + 1$ *et remarquer que* $16 = 641 - 625$.)

e) Soit p un nombre premier qui divise F_n. Montrer que 2 est d'ordre 2^{n+1} dans $(\mathbf{Z}/p\mathbf{Z})^*$. En déduire que $p \equiv 1 \pmod{2^{n+1}}$.

f) (*suite*) En utilisant que 2 est un carré modulo p, démontrer que $p \equiv 1 \pmod{2^{n+2}}$ (Lucas).

13) Soit p un nombre premier impair.

a) Démontrer que dans $(\mathbf{Z}/p\mathbf{Z})^*$, $(p-1)/2$ éléments sont des carrés et et $(p-1)/2$ n'en sont pas.

b) Soit a un entier non multiple de p. Démontrer que $a^{(p-1)/2}$ vaut ± 1 modulo p.

Démontrer que pour que a soit un carré dans $(\mathbf{Z}/p\mathbf{Z})^*$, il faut et il suffit que $a^{(p-1)/2} \equiv 1 \pmod{p}$. (*Les carrés de* $(\mathbf{Z}/p\mathbf{Z})^*$ *vérifient cette condition ; combien d'éléments au plus la vérifient-ils ?*)

c) Pour que -1 soit un carré modulo p, il faut et il suffit que p soit congru à 1 modulo 4.

d) Montrer que $(p-1)! \equiv -1$ modulo p. Si $p \equiv 1 \pmod{4}$, démontrer que $a = ((p-1)/2)!$ vérifie $a^2 \equiv -1 \pmod{p}$. Est-ce un moyen pratique de déterminer une racine carrée de -1 dans $\mathbf{Z}/p\mathbf{Z}$ lorsque p est grand ?

14) Cet exercice propose la démonstration directe de quelques cas de la loi de réciprocité quadratique, à partir de la remarque que pour certaines valeurs de n, $2\cos(2\pi/n)$ vérifie une équation du second degré dont le discriminant est $\pm n$.

a) Soit x une racine primitive huitième de l'unité. Montrer que $x + 1/x$ est solution d'une équation du second degré à coefficients dans $\mathbf{F}_p$ dont le discriminant est 2. En déduire que 2 est un carré modulo p si $p \equiv \pm 1 \pmod{8}$. Inversement, si 2 est un carré modulo p, montrer que x et $1/x$ sont solutions d'une équation du second degré de discriminant -2. En déduire que $p \equiv \pm 1 \pmod{8}$.

b) Écrire l'équation du second degré dont les solutions sont les racines primitives cubiques de l'unité. En déduire que -3 est un carré modulo un nombre premier $p > 4$ si et seulement si $p \equiv 1 \pmod{3}$.

c) Soit p un nombre premier tel que $p > 5$. Soit x une racine primitive cinquième de l'unité dans une extension de $\mathbf{F}_p$. Montrer que $u = x + 1/x$ est solution d'une équation du second degré de discriminant 5. Si $p^2 \equiv 1 \pmod{5}$, en déduire que 5 est un carré modulo p.

d) (*suite*) Supposons inversement que 5 soit un carré modulo p. Montrer que x et $1/x$ sont les solutions de l'équation $t^2 - tu + 1 = 0$. Montrer que x^p est aussi une solution de cette équation et en déduire que l'on a $x^p = x$ ou $x^p = 1/x$, puis que $p \equiv \pm 1 \pmod{5}$.

15) a) Soit n un nombre entier qui n'est pas multiple de 3. Montrer que $4n^2 + 3$ possède un facteur premier p tel que $p \equiv 7 \pmod{12}$.

b) Démontrer qu'il existe une infinité de nombres premiers congrus à 7 modulo 12.

16) Soit p un nombre premier de la forme $4\ell + 1$, où ℓ est un nombre premier. Démontrer que 2 est un générateur du groupe $(\mathbf{Z}/p\mathbf{Z})^*$.

17) Si r est un entier, on pose $M_r = 2^r - 1$ (nombres de Mersenne).

a) Si M_r est un nombre premier, démontrer que r est un nombre premier.

b) Soit p un nombre premier tel que $p > 2$. Montrer que tout facteur premier de M_p est congru à 1 modulo p, et à ± 1 modulo 8.

c) On dit qu'un nombre entier est parfait s'il est égal à la somme de ses diviseurs (stricts, c'est-à-dire sauf lui-même). Démontrer qu'un nombre entier pair n est parfait si et seulement s'il existe un nombre premier p tel que M_p soit un nombre premier et $n = 2^{p-1}M_p$.

§1.3. Le théorème des nombres premiers

Rappelons que l'on note $\mathscr{P}$ l'ensemble des nombres premiers et $\pi(x)$ le cardinal de $\mathscr{P} \cap [1, x]$. L'objectif de ce paragraphe est le théorème suivant, communément appelé *théorème des nombres premiers* :

THÉORÈME 1.3.1. — *On a l'équivalent, lorsque x tend vers $+\infty$,*

$$\pi(x) \sim \frac{x}{\log x}.$$

Je vais essentiellment suivre, en rajoutant quelques détails, la présentation (en à peine 4 pages !) que donne Zagier dans [**8**] d'une démonstration due à D. J. Newman. Comme dans presque toutes les preuves du théorème des nombres premiers, nous aurons à faire usage de la théorie des fonctions holomorphes. *Presque toutes*, car il existe des démonstrations « élémentaires » qui ne recourrent pas à l'analyse complexe ; ces démonstrations sont cependant moins transparentes et plus difficiles à exposer ! La première démonstration élémentaire est due à Erdös et Selberg, une démonstration de Daboussi est exposée dans le bien joli petit livre [**7**].

A. La fonction de zêta de Riemann et sa dérivée logarithmique

Définissons, pour tout nombre complexe s tel que $\Re(s) > 1$,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

La fonction ainsi définie est appelée *fonction zêta de Riemann* ; bien qu'introduite par Euler vers 17 ??, c'est B. Riemann qui, dans un somptueux article publié en 1859, en a révélé toute l'importance dans la théorie des nombres premiers.

Par comparaison à l'intégrale $\int_1^\infty dt/t^a$, pour $a > 1$, on voit que cette série converge normalement dans le demi-plan fermé d'équation $\Re(s) \leqslant a$. Comme chaque terme de la série est une fonction holomorphe dans le demi-plan ouvert $\Re(s) > a$, la fonction zêta de Riemann est une fonction holomorphe sur ce demi-plan, donc sur leur réunion qui est le demi-plan ouvert d'équation $\Re(s) > 1$. Son comportement au bord de ce demi-plan va s'avérer crucial. Pour $a \in \mathbf{R}$, on notera Ω_a l'ensemble des nombres complexes s tels que $\Re(s) > a$.

Le lien avec les nombres premiers provient de la formule :

PROPOSITION 1.3.1. — *Pour tout nombre complexe s tel que $\Re(s) > 1$, on a*

$$\zeta(s) = \frac{1}{1-2^{-s}} \frac{1}{1-3^{-s}} \cdots = \prod_{p \in \mathscr{P}} \frac{1}{1-p^{-s}},$$

où le produit infini converge absolument dans l'ouvert Ω_1, uniformément dans chaque demi-plan Ω_a avec $a > 1$.

Démonstration. — Pour $\Re(s) > 1$, $\left|p^{-s}\right| < 1$ et l'on a

$$\frac{1}{1-p^{-s}} = \sum_{m=0}^{\infty} p^{-ms}.$$

La convergence absolue du produit infini pour $\Re(s) > 1$, et sa convergence uniforme pour $\Re(s) > a$ si $a > 1$, proviennent ainsi du fait que

$$\left|\frac{1}{1-p^{-s}} - 1\right| \leqslant \frac{2}{\left|p^{-s}\right|}$$

et de la convergence absolue (resp. uniforme...) de la série de Riemann.

Soit T un nombre entier. On peut alors développer le produit des facteurs $1/(1-p^{-s})$ pour p premier, $p \leqslant T$: si $\mathscr{P} \cap [1,P] = \{p_1, \ldots, p_t\}$, on obtient

$$\prod_{p \in \mathscr{P} \cap [1,T]} \frac{1}{1-p^{-s}} = \prod_{i \leqslant t} \sum_{m_i=0}^{\infty} p_i^{-m_i s} = \sum_{m_1=0}^{\infty} \cdots \sum_{m_t=0}^{\infty} (p_1^{m_1} \ldots p_t^{m_t})^{-s}.$$

C'est une sous-série de la série qui définit la fonction zêta de Riemann : seuls sont présents les termes n^{-s} où n est un entier dont tous les facteurs premiers sont inférieurs à T, et ces termes apparaissent une fois et une seule d'après le théorème de décomposition en facteurs premiers. En particulier, cette série multiple converge absolument si $\Re(s) > 1$, ce qui justifie le développement fait. En outre,

$$\left|\zeta(s) - \prod_{p \leqslant T} \frac{1}{1-p^{-s}}\right| \leqslant \sum_{n > T} \left|n^{-s}\right|;$$

ce dernier terme tend vers 0 quand T tend vers l'infini en vertu de la convergence de la série de Riemann. Par suite, le produit infini écrit vaut $\zeta(s)$. □

PROPOSITION 1.3.2. — *La fonction ζ s'étend en une fonction méromorphe sur* $\mathbf{C}$. *Ce prolongement a un unique pôle en $s = 1$, simple et de résidu $s = 1$.*

Démonstration. — On se contente de montrer l'existence d'un prolongement à l'ouvert Ω_0. Pour cela, observons que pour $\Re(s) > 1$,

$$\begin{aligned}\zeta(s) - \frac{1}{s-1} &= \sum_{n=1}^{\infty} n^{-s} - \int_1^{\infty} x^{-s}\,dx \\ &= \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx.\end{aligned}$$

Notons $f_n(s)$ l'intégrale figurant au n-ième terme de cette série. Par intégration par parties,

$$\begin{aligned} f_n(s) &= \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx \\ &= \left[\left(\frac{1}{n^s} - \frac{1}{x^s}\right)(n+1-x)\right]_n^{n+1} + s\int_n^{n+1} (n+1-x)x^{-s-1}\, dx \\ &= s\int_n^{n+1} (n+1-x)x^{-s-1}\, dx, \end{aligned}$$

si bien que

$$\left|f_n(s)\right| \leqslant \frac{|s|}{n^{\Re(s)+1}}.$$

La fonction f_n est holomorphe sur Ω_0 et la majoration précédente entraîne que la série $\sum f_n(s)$ converge pour tout $s \in \Omega_0$, uniformément dans tout ouvert Ω_a avec $a > 0$. La somme de cette série définit donc une fonction holomorphe f sur le demi-plan Ω_0. Pour $\Re(s) > 1$, on a $\zeta(s) = \frac{1}{s-1} + f(s)$; cette dernière expression fournit le prolongement voulu. □

Dans la suite, on notera encore ζ le prolongement méromorphe de la fonction zêta de Riemann.

B. Non-annulation de la fonction zêta sur la droite $\Re(s) = 1$

Même s'il est difficile de le concevoir au premier abord, c'est dans cette propriété analytique de la fonction zêta de Riemann que réside le cœur de la démonstration du théorème des nombres premiers.

PROPOSITION 1.3.1. — *Pour tout nombre complexe $s \neq 1$ tel que $\Re(s) \geqslant 1$, on a $\zeta(s) \neq 0$.*

Démonstration. — Lorsque $\Re(s) > 1$, cela résulte de l'expression de $\zeta(s)$ comme produit infini, aucun facteur n'étant nul. Toujours pour $\Re(s) > 1$, la dérivée logarithmique du produit infini $\prod_p 1/(1-p^{-s})$ s'écrit

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{\log p}{p^s - 1} = \sum_p \frac{\log p}{p^s} + \sum_p \frac{\log p}{p^s(p^s-1)}.$$

Comme $\log p = \mathrm{O}(p^\varepsilon)$ pour tout $\varepsilon > 0$, les deux séries du second membre convergent pour $\Re(s) > 1$ et $\Re(s) > 1/2$, par comparaison avec la série de Riemann et définissent des fonctions holomorphes dans les ouverts Ω_1 et $\Omega_{1/2}$ respectivement, que l'on notera Φ et Ψ. (En effet, elles convergent uniformément dans tout Ω_a avec $a > 1$ et $a > 1/2$.)

L'expression $\Phi(s) = -\zeta'(s)/\zeta(s) - \Psi(s)$ montre que Φ possède un prolongement méromorphe dans le demi-plan $\Omega_{1/2}$. Ses pôles proviennent du pôle simple de ζ en $s = 1$, avec résidu 1, et des zéros de ζ dans le demi-plan $\Omega_{1/2}$, avec résidu -1. Autrement dit,

$$\lim_{\sigma\to 1}(\sigma-1)\Phi(\sigma) = 1, \quad \lim_{\sigma\to 1}(\sigma-1)\Phi(\sigma+i\tau) = -m$$

si ζ possède un zéro d'ordre m en $1+i\tau$.

Suppsosons donc que ζ ait un zéro d'ordre m en $s=1+i\tau$ et un zéro d'ordre n en $1+2i\tau$. Comme $\zeta(\bar{s})=\overline{\zeta(s)}$, ζ a aussi un zéro d'ordre m en $s=1-i\tau$ et un zéro d'ordre n en $1-2i\tau$. Observons que pour tout nombre réel $\sigma>0$,

$$\begin{aligned}\sum_{k=-2}^{2}\binom{4}{2+k}\Phi(\sigma+i\tau k) &= \sum_p \frac{\log p}{p^\sigma}\left(p^{-2i\tau}+4p^{-i\tau}+6+4p^{i\tau}+p^{2i\tau}\right)\\ &= \sum_p \frac{\log p}{p^\sigma}\left(p^{-i\tau}+p^{i\tau}\right)^4\\ &= \sum_p \frac{\log p}{p^\sigma}(2\cos(\tau\log p))^4 \geqslant 0.\end{aligned}$$

Lorsque l'on multiplie cette expression par $\sigma-1$, pour σ un nombre réel >1, et que l'on fait tendre σ vers 1, on obtient

$$-n-4m+6-4m-n\geqslant 0,$$

c'est-à-dire $8m+2n\leqslant 6$. Nécessairement, $m=0$: la fonction ζ ne s'annule pas en $1+i\tau$. □

Mettons en exergue un résultat établi au cours de la démonstration :

COROLLAIRE 1.3.2. — *La série $\Phi(s)=\sum_p(\log p)/p^s$ converge pour $\Re(s)>0$, définit une fonction méromorphe dans l'ouvert $\Omega_{1/2}$, et n'a pas de pôle dans un voisinage du demi-plan $\{\Re(s)\geqslant 0\}$.*

C. Un théorème taubérien, et la conclusion

Pour tout nombre réel $x>0$, on pose

$$\theta(x)=\sum_{p\leqslant x}\log p.$$

Nous allons voir que la connaissance du comportement de θ équivaut à celle du comportement de la fonction π, mais elle est plus facile à étudier.

LEMME 1.3.1. — *Si $\theta(x)\sim x$ en $+\infty$, alors $\pi(x)\sim x/\log(x)$.*

Démonstration. — Supposons donc $\theta(x)\sim x$. On a donc

$$\theta(x)=\sum_{p\leqslant x}\log p\leqslant\sum_{p\leqslant x}\log x\leqslant\pi(x)\log x.$$

En particulier,

$$\limsup_{x\to\infty}\frac{\pi(x)}{x/\log x}\leqslant\limsup_{x\to\infty}\frac{\theta(x)}{x}\leqslant 1.$$

Dans l'autre sens, si ε est un nombre réel tel que $0 < \varepsilon < 1$,

$$\begin{aligned}\theta(x) &\geqslant \sum_{x^{1-\varepsilon}\leqslant p\leqslant x} \log p \\ &\geqslant \sum_{x^{1-\varepsilon}\leqslant p\leqslant x} (1-\varepsilon)\log x \\ &\geqslant (1-\varepsilon)\log x\left(\pi(x)-\pi(x^{1-\varepsilon})\right) \\ &\geqslant (1-\varepsilon)\pi(x)\log x + \mathrm{O}(x^{1-\varepsilon/2})\end{aligned}$$

puisque $\pi(x) = \mathrm{O}(x)$. Il en résulte que

$$\liminf \frac{\theta(x)}{x} \geqslant (1-\varepsilon)\liminf_{x\to\infty} \frac{\pi(x)}{x/\log x},$$

soit encore, en faisant tendre ε vers 0,

$$\liminf_{x\to\infty} \frac{\pi(x)}{x/\log x} \leqslant \liminf \frac{\theta(x)}{x} = 1.$$

Le lemme est ainsi démontré. □

Il suffit donc de démontrer l'équivalent $\theta(x) \sim x$, ce qui va passer par une forme apparemment plus faible :

LEMME 1.3.2. — *Si l'intégrale*

$$(*) \qquad \int_1^\infty \frac{\theta(x)-x}{x^2}\,dx$$

converge, alors $\theta(x) \sim x$ au voisinage de $+\infty$.

Démonstration. — Si x est un nombre réel tel que $\theta(x) > \lambda x$, avec $\lambda > 1$, alors

$$\int_x^{\lambda x} \frac{\theta(t)-t}{t^2}\,dt \geqslant \int_x^{\lambda x} \frac{\theta(x)-t}{t^2}\,dt \geqslant \int_x^{\lambda x} \frac{\lambda x - t}{t^2}\,dt = \int_1^\lambda \frac{\lambda - t}{t^2}\,dt > 0.$$

(La première inégalité provient de ce que θ est croissante.) L'existence de tels nombres réels arbitrairement grands contredit donc le critère de Cauchy pour la convergence de l'intégrale de $(\theta(x)-x)/x^2$. Dans l'autre sens, si x est un nombre réel tel que $\theta(x) < \lambda x$, avec $\lambda < 1$, alors

$$\int_{\lambda x}^x \frac{\theta(t)-t}{t^2}\,dt \leqslant \int_{\lambda x}^x \frac{\theta(x)-t}{t^2}\,dt \leqslant \int_{\lambda x}^x \frac{\lambda x - t}{t^2}\,dt = \int_\lambda^1 \frac{\lambda - t}{t^2}\,dt < 0.$$

Là encore, l'existence de tels nombres réels arbitrairement grands contredit le critère de Cauchy. Cela démontre le lemme. □

La relation entre l'intégrale $(*)$ et notre problème vient de la relation suivante, pour $\Re(s) > 1$. Notons $p_1, p_2, \ldots$ la suite (croissante) des nombres premiers ; posons aussi

$p_0 = 1$. Via une transformation d'Abel, on a donc, pour tout nombre complexe s tel que $\Re(s) > 1$,

$$\begin{aligned}\Phi(s) &= \sum_{i=1}^{\infty} \frac{\log p_i}{p_i^s} \\ &= \sum_{i=1}^{\infty} \frac{\theta(p_i) - \theta(p_{i-1})}{p_i^s} \\ &= \sum_{i=1}^{\infty} \theta(p_i) \left(\frac{1}{p_i^s} - \frac{1}{p_{i+1}^s} \right) \\ &= \sum_{i=1}^{\infty} \theta(p_i) \int_{p_i}^{p_{i+1}} \frac{s}{x^{s+1}}\, dx \\ &= \int_2^{\infty} \theta(x) \frac{s}{x^{s+1}}\, dx.\end{aligned}$$

On peut réduire la borne inférieure de cette dernière intégrale à 1 puisque $\theta(x) = 0$ pour $x < 2$, d'où

$$\Phi(s) = s \int_1^{\infty} \theta(x) x^{-s-1}\, dx.$$

Comme $1/(s-1) = \int_1^{\infty} x^{-s}\, dx$, on a donc

$$\Phi(s) - \frac{s}{s-1} = s \int_1^{\infty} \frac{\theta(x) - x}{x^{s+1}}\, dx.$$

Lorsque s tend vers 1, le membre de gauche a une limite car Φ a un pôle simple de résidu 1 en $s = 1$; il s'agit de démontrer que cette limite est obtenue en passant à la limite sous le signe somme.

Via le changement de variables $x = e^t$, on écrit alors

$$\frac{\Phi(s)}{s} - \frac{1}{s-1} = \int_0^{\infty} f(t) e^{-st}\, dt,$$

où $f(t) = \theta(e^t) e^{-t} - 1$. Le résultat voulu découle alors du théorème taubérien suivant, de nature purement analytique.

THÉORÈME 1.3.1 (Newman). — *Soit $f\colon \mathbf{R}_+ \to \mathbf{C}$ une fonction mesurable bornée, de sorte que, pour tout nombre complexe s tel que $\Re(s) > 0$, l'intégrale $g(s) = \int_0^{\infty} f(t) e^{-st}\, dt$, converge et définit une fonction holomorphe dans le demi-plan Ω_0. Supposons que cette fonction g s'étende en une fonction holomorphe, toujours notée g, définie au voisinage du demi-plan $\{\Re(s) \geqslant 0\}$. Alors, l'intégrale $\int_0^{\infty} f(t)\, dt$ existe et vaut $g(0)$.*

Pour conclure, il reste à vérifier que la fonction f est bornée ; cela fait l'objet du dernier lemme de ce paragraphe.

LEMME 1.3.2. — *Il existe un nombre réel $c > 0$ tel que $\theta(x) \leqslant cx$ pour $x \geqslant 2$.*

Démonstration. — Pour $n \in \mathbf{N}$, la formule du binôme entraîne l'inégalité

$$2^{2n} = (1+1)^{2n} \geqslant \binom{2n}{n}.$$

Décomposons ce dernier coeficient binomial en facteurs premiers. Comme on a

$$\binom{2n}{n} = \frac{(n+1)(n+2)\dots(2n)}{1\cdot 2\dots 2n},$$

tout nombre premier p tel que $n < p < 2n$ divise exactement $\binom{2n}{n}$. Par suite,

$$\binom{2n}{n} \geqslant \prod_{n<p<2n} p,$$

d'où,en prenant les logarithmes, l'inégalité $\theta(2n) - \theta(n) \leqslant 2n\log 2$. Soit x un nombre réel et soit n la « partie entière supérieure » de $x/2$; on a donc $n-1 < x/2 \leqslant n$ et $2n-2 < x \leqslant 2n$; par suite

$$\theta(x) - \theta(x/2) \leqslant \theta(2n) - \theta(n-1) \leqslant 2n\log 2 + \theta(n) - \theta(n-1) \leqslant n(2\log 2 + 1) \leqslant x(1+2\log 2)$$

puisque $n < 1 + x/2 \leqslant x$ si $x \geqslant 2$. Cette inégalité est évidemment vérifiée pour x tel que $0 < x < 2$ puisqu'alors $\theta(x) = 0$. Alors, pour $x > 1$,

$$\theta(x) \leqslant \sum_{n=0}^{\infty} \left(\theta(x/2^n) - \theta(x/2^{n+1})\right) \leqslant (1+2\log 2) \sum_{n=0}^{\infty} \frac{x}{2^n} \leqslant 2(1+2\log 2)x.$$

Le lemme est ainsi démontré. □

D. Démonstration du théorème taubérien

Pour tout nombre réel $T > 0$ et tout nombre complexe s, posons $g_T(s) = \int_0^T f(t)e^{-st}\,dt$. La fonction g_T est holomorphe sur $\mathbf{C}$ et il s'agit de démontrer que l'on a $\lim_{T\to\infty} g_T(0) = g(0)$.

Par hypothèse, la fonction f est bornée ; posons donc $B = \max_{|t|\geqslant 0} \left|f(t)\right|$.

Pour $R > 0$ grand et $\delta > 0$ petit, soit Ω l'ouvert du plan formé des nombres complexes z tels que $|z| < R$ et $\Re(z) > -\delta$. Lorsque R est fixé, et $\delta > 0$ est assez petit (dépendant de R), la fonction g est définie et holomorphe au voisinage de $\overline{\Omega}$. Notons C la frontière de Ω et considérons-la comme un lacet. D'après le théorème des résidus,

$$g(0) - g_T(0) = \frac{1}{2i\pi} \int_C \left(g(z) - g_T(z)\right) e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{dz}{z}. \tag{E}$$

Pour $\Re(z) > 0$, $g(z)$ est égale a l'intégrale $\int_0^\infty f(t)e^{-zt}\,dt$ si bien que l'on a

$$\begin{aligned} \left|g(z) - g_T(z)\right| &= \left|\int_T^\infty f(t)e^{-zt}\,dt\right| \\ &\leqslant B\int_T^\infty e^{-\Re(s)t}\,dt = \frac{B}{\Re(z)} e^{-\Re(z)T}. \end{aligned}$$

Toujours pour $\Re(z) > 0$, on a

$$\left| e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{1}{z}\right| = \frac{1}{R}e^{\Re(z)T}\left|\frac{R}{z}+\frac{z}{R}\right| = \frac{2\Re(z)}{R^2}e^{\Re(z)T}$$

si, de plus, $z \in C$, puisqu'alors $|z| = R$. La contribution à l'intégrale (E) du demi-arc de cercle C_+ formé des $z \in C$ tels que $\Re(z) \geqslant 0$ est ainsi majorée par

$$\left|\frac{1}{2i\pi}\int_{C_+}\left|g(z)-g_T(z)\right|e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}\right| \leqslant \frac{\ell(C_+)}{2\pi}\frac{2B}{R^2} = \frac{B}{R}.$$

Soit C_- la partie du lacet C contenue dans le demi-plan $\Re(z) \leqslant 0$. Pour calculer l'intégrale

$$I_T = \frac{1}{2i\pi}\int_{C_-} g_T(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z},$$

on peut remplacer le lacet C_- par le demi-arc de cercle $C'_- = C_+$ car l'intégrande n'a pas de pôle dans le demi-plan $\Re(z) < 0$, g_T étant holomorphe sur $\mathbf{C}$. Sur cet arc de cercle C'_-, on a

$$\left|g_T(z)\right| = \left|\int_0^T f(t)e^{-zt}\,dt\right| \leqslant B\int_0^T e^{-\Re(z)t}\,dt \leqslant \frac{B}{-\Re(z)}e^{-\Re(z)T}.$$

Pour tout $z \in C'_-$, on a comme précédemment la majoration

$$\left| e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{1}{z}\right| = \frac{1}{R}e^{\Re(z)T}\left|\frac{R}{z}+\frac{z}{R}\right| = \frac{2\left|\Re(z)\right|}{R^2}e^{\Re(z)T}$$

Ainsi,

$$I'_T \leqslant \frac{\ell(C'_-)}{2\pi}\frac{2B}{R^2} \leqslant \frac{B}{R}.$$

Majorons enfin l'intégrale

$$I'_T = \frac{1}{2i\pi}\int_{C_-} g(z)e^{zT}\left(1+\frac{z^2}{R^2}\right)\frac{dz}{z}.$$

Posons

$$B' = \max_{z\in C'_-}\left|g(z)\left(1+\frac{z^2}{R^2}\right)\frac{1}{z}\right|.$$

Le lacet C_- est formé de deux bouts d'arcs de cercles de longueurs au plus $R\arcsin(\delta) \leqslant 2R\delta$ et d'un segment de droite contenu dans la droite d'équation $\Re(z) = -\delta$ et de longueur $2\sqrt{R^2-\delta^2} \leqslant 2R$. La partie de l'intégrale I'_T correspondant aux deux arcs de cercles est ainsi majorée par

$$\frac{4R\delta}{2\pi}B',$$

tandis que la partie de l'intégrale correspondant au segment de droite est majorée par

$$\frac{2R}{2\pi}B'e^{-\delta T}.$$

Ainsi,

$$\left|I'_T\right| \leqslant B'\frac{R}{\pi}\left(2\delta+e^{-\delta T}\right).$$

Mettant bout à bout les trois majorations obtenues, nous pouvons donc affirmer que

$$\left|g(0)-g_T(0)\right| \leqslant 2\frac{B}{R}+B'\frac{R}{\pi}\left(2\delta+e^{-\delta T}\right).$$

Soit $\varepsilon>0$ et montrons que pour T assez grand, $\left|g(0)-g_T(0)\right|$ est inférieur à ε. Commençons par poser $R=2B/\varepsilon$ puis $\delta>0$ de sorte que $2R\delta B'/\pi \leqslant \varepsilon$. Comme $e^{-\delta T}$ tend vers 0 en décroissant quand T tend vers l'infini, il existe T_0 tel que $(RB'/\pi)e^{-\delta T_0}<\varepsilon$. Alors, pour $T \geqslant T_0$, on a

$$\left|g(0)-g_T(0)\right| \leq \varepsilon+\varepsilon+\varepsilon = 3\varepsilon.$$

Cela démontre que $\lim_{T\to\infty} g_T(0)=g(0)$ et conclut la démonstration du théorème taubérien ainsi que celle du théorème des nombres premiers !

F. L'hypothèse de Riemann

Nous avons vu que le théorème des nombres premiers repose sur la non-annulation de la fonction ζ de Riemann sur la droite $\{\Re(s)=1\}$. L'*hypothèse de Riemann* est qu'en fait la fonction ζ ne s'annule pas dans le demi-plan $\Omega_{1/2}$. On peut démontrer que cela entraîne le développement asymptotique suivant :

$$\pi(x) \sim \int_2^x \frac{dt}{\log t} + \mathrm{O}(x^{1/2+\varepsilon}),$$

où $\varepsilon>0$ est un nombre réel arbitraire. (L'intégrale du second membre, appelée logarithme intégral, est équivalente à $x/\log x$ au voisinage de $+\infty$.) Inversement, s'il existe un zéro de la fonction zêta de Riemann dont la partie réelle est égale à a avec $\frac{1}{2}<a<1$, alors on ne peut pas choisir $\varepsilon<a-\frac{1}{2}$ dans le développement asymptotique.

Cette hypothèse figure dans l'article de Riemann de 1859 mais n'est toujours pas démontrée. Elle fait l'objet d'une mise à prix d'un million de dollars par la fondation Clay.

L'hypothèse de Riemann généralisée est l'extension de cette conjecture aux fonctions L de Dirichlet qui sont introduites au paragraphe suivant, et plus généralement, aux fonctions zêta de Dedekind associées aux corps de nombres.

Exercices

18) a) En utilisant le fait que la fonction $\zeta(s)$ tend vers $+\infty$ quand s tend vers 1 par valeurs supérieures, démontrer que la somme des inverses des nombres premiers diverge.

19) On note p_n le n-ième nombre premier. En outre, les fonctions θ et ψ sont définies par

$$\theta(x)=\sum_{p\leqslant x}\log p, \qquad \psi(s)=\sum_{p^m\leqslant x}\log p,$$

où les sommes sont prises sur l'ensemble des nombres premiers et des puissances de nombres premiers inférieurs ou égaux à x.

Montrer que les énoncés suivants sont équivalents au théorème des nombres premiers :
a) $p_n \sim n\log n$;
b) $\theta(x) \sim x$;
c) $\psi(x) \sim x$.

20) a) Démontrer l'égalité

$$\binom{2n}{n} = \prod_{p\leqslant 2n} p^{a_p}, \qquad a_p = \sum_{m=1}^{\infty}\left(\lfloor\frac{2n}{p^m}\rfloor - 2\lfloor\frac{n}{p^m}\rfloor\right).$$

b) Démontrer que pour tout nombre premier p, $a_p \leqslant \lfloor \log 2n/\log p\rfloor$.

c) Vérifier que $\binom{2n}{n} \geqslant 2^n$ et en déduire que $\psi(2n) \geqslant n\log 2$. (La fonction ψ, ainsi que la fonction θ, est définie dans l'exercice précédent.)

d) Démontrer alors qu'il existe un nombre réel c tel que $\theta(n) > cn$ pour $n \geqslant 2$.

21) Démontrer que pour tout entier $n \geqslant 1$, $\theta(n) \leqslant 2n\log 2$. (Reprendre en étant plus soigneux les arguments de la démonstration du lemme 1.3.2.)

22) Le but de cet exercice est de démontrer le *postulat de J. Bertrand* : il existe pour tout entier n un nombre premier p tel que $n < p < 2n$.

a) Vérifier que l'assertion est vraie si $n \leqslant 1000$.

On suppose dans la suite que n est un entier qui met en défaut le postulat de Bertrand et on reprend les notations de l'exercice 20.

b) Soit p un facteur premier de $\binom{2n}{n}$; démontrer que $p \leqslant \frac{2}{3}n$.

c) Soit p un nombre premier tel que p^2 divise $\binom{2n}{n}$. Démontrer que $p \leqslant \sqrt{2n}$.

d) Déduire des deux questions précédentes et de l'exercice 21 que l'on a

$$\log\binom{2n}{n} \leqslant \frac{4}{3}n\log 2 + \sqrt{2n}\log(2n).$$

e) Démontrer que $\binom{2n}{n} \geqslant 2^{2n}/2n$ et obtenir une contradiction.

23) Vérifier que le résultat de l'exercice 5 est une conséquence immédiate du postulat de Bertrand démontré dans l'exercice 22.

24) Dans la littérature, la fonction logarithme intégral est plutôt définie par la formule

$$\mathrm{li}(x) = \lim_{\varepsilon\to 0^+}\int_0^{1-\varepsilon}\frac{dt}{\log t} + \int_{1+\varepsilon}^{x}\frac{dt}{\log t}.$$

a) Justifier l'existence de cette limite.

b) Établir, pour tout entier $k \geqslant 0$, le développement asymptotique

$$\mathrm{li}(x) = \frac{x}{\log x} + \cdots + \frac{x}{\log^k x} + \mathrm{O}(x/\log^{k+1} x).$$

§1.4. Le théorème de la progression arithmétique

Le but de ce paragraphe est de prouver le *théorème de la progression arithmétique*, conjecturé par Euler en 1775 et démontré pour la première fois par Dirichlet en 1837. Il s'agit du théorème suivant :

THÉORÈME 1.4.1. — *Soit a et n des entiers naturels premiers entre eux. Pour $x > 0$, notons $\pi(x; n, a)$ le cardinal de l'ensemble des nombres premiers p tel que $p \equiv a \pmod{n}$ et $p \leqslant x$. Alors, lorsque x tend vers $+\infty$, on a l'équivalent*

$$\pi(x; n, a) \sim \frac{1}{\varphi(n)} \frac{x}{\log x}.$$

En particulier, l'ensemble des nombres premiers congrus à a modulo n est infini.

(En fait, Euler et Dirichlet n'étaient concernés que par le caractère infini de l'ensemble des nombres premiers dans une progression arithmétique ; l'assertion indiquée, plus forte que le théorème des nombres premiers, fut démontrée par de la Vallée Poussin.)

Dans ce théorème, l'hypothèse que n et a sont premiers entre eux est cruciale : comme tout nombre entier qui est congru à a modulo n est multiple du pgcd de n et a, il n'y aurait sinon qu'au plus un nombre premier vérifiant la congruence donnée.

Si l'on veut adapter les méthodes analytiques utilisées pour démontrer le théorème des nombres premiers, on est tenté d'introduire des variantes de la fonction zêta, notamment les fonctions d'une variable s données par

$$\zeta_a(s) = \prod_{p \equiv a \pmod{n}} \frac{1}{1 - p^{-s}},$$

dont le produit est la fonction zêta de Riemann à quelques facteurs près. Mais on ne sait pas comment étudier cette fonction holomorphe si l'on ne sait rien des nombres premiers congrus à a modulo n. L'idée remarquable de Dirichlet a consisté à transformer des propriétés multiplicatives en propriétés additives en introduisant les *caractères* du groupe abélien $(\mathbf{Z}/n\mathbf{Z})^*$.

A. Caractères des groupes abéliens finis

Soit G un groupe abélien fini, noté multiplicativement. On appelle *caractère* de G tout homomorphisme de G dans $\mathbf{C}^*$. Comme G est fini, les valeurs prises par un caractère sont des racines de l'unité. L'ensemble des caractères de G est noté $\hat{G}$; c'est un sous-groupe de l'ensemble des fonctions de G dans $\mathbf{C}^*$ et son élément neutre est la fonction 1 constante de valeur 1. Le conjugué $\bar{\chi}$ d'un caractère χ est aussi un caractère, d'ailleurs égal à l'inverse χ^{-1} de χ.

Supposons que G soit cyclique d'ordre n, engendré par un élément g. Tout élément de G est alors de la forme g^a, pour un entier $a \in \mathbf{Z}$ bien défini modulo n. Un caractère de G est ainsi déterminé par l'image de g qui est une racine n-ième de l'unité. Inversement, pour toute racine n-ième de l'unité z, l'application de G dans $\mathbf{C}^*$ qui applique g^a sur z^a, pour $0 \leqslant a < n$, est un caractère de G. Le groupe $\hat{G}$ est ainsi identifié au groupe des racines n-ièmes de l'unité, qui est un groupe cyclique d'ordre n. Plus généralement :

PROPOSITION 1.4.1. — *Soit G un groupe abélien fini ; le groupe $\hat{G}$ a même cardinal que G.*

Comme dans le cas d'un groupe cyclique, on peut démontrer que les groupes G et $\hat{G}$ sont isomorphes, voir l'exercice 27.

La démonstration de cette proposition utilise un lemme, d'intérêt général.

LEMME 1.4.2. — *Soit G un groupe abélien fini et soit H un sous-groupe de G. Pour tout caractère χ_1 de H, il existe un caractère χ de G tel que $\chi_1 = \chi_{|H}$.*

Démonstration. — Démontrons ce lemme par récurrence sur l'indice $(G : H) = \mathrm{card}(G)/\mathrm{card}(H)$. Il n'y a rien à démontrer si $(G : H) = 1$. Considérons sinon un élément $g \in G \setminus H$ et soit m le plus petit entier > 1 tel que $g^m \in H$. Comme G est commutatif, l'ensemble des produits $g^a h$, où $0 \leqslant a < m$ et $h \in H$ est un sous-groupe H' de G. En outre, l'égalité $g^a h = g^b k$ avec $0 \leqslant a \leqslant b < m$ et $h, k \in H$ entraîne $g^{b-a} = kh^{-1}$, d'où $a = b$ et $h = k$. Tout élément de H' s'écrit donc d'une seule manière sous la forme $g^a h$.

Soit z un nombre complexe tel que $z^m = \chi_1(g^m)$. Pour $h' = g^a h \in H'$, posons $\chi'(h') = z^a \chi_1(h)$. L'application $\chi' : H' \to \mathbf{C}^*$ est un caractère. En effet, si $h' = g^a h$ et $k' = g^b k$ sont des éléments de H, on a $h'k' = g^{a+b}hk$. Si $a + b < m$, il vient

$$\chi'(h'k') = \chi(g^{a+b}hk) = z^{a+b}\chi_1(hk) = z^a z^b \chi_1(h)\chi_1(k) = \chi'(h')\chi'(k'),$$

tandis que si $a + b > m$, on a aussi $a + b < 2m$ et

$$\begin{aligned}\chi'(h'k') &= \chi(g^{a+b-m}g^m hk) = z^{a+b-m}\chi_1(g^m hk) = z^{a+b-m}\chi_1(g^m)\chi_1(h)\chi_1(k)\\ &= z^{a+b}\chi_1(h)\chi_1(k) = \chi'(h')\chi'(k').\end{aligned}$$

Comme l'indice $(G : H')$ de H' dans G est égal à $(G : H)/m < (G : H)$, il existe par récurrence un caractère χ de G dont la restriction à H' est égale à χ'. En particulier, $\chi_{|H} = \chi_1$. □

Démonstration de la proposition 1.4.1. — Nous avons déjà vu que cette proposition est vraie lorsque G est cyclique. Dans le cas général, démontrons-la par récurrence sur l'ordre de G. Les cas $\mathrm{card}(G) = 1, 2, 3$ résultent du cas cyclique.

Soit h un élément de G non réduit à l'identité et soit H le sous-groupe cyclique engendré par h. Si m désigne son ordre, on a $m > 1$. Le groupe H étant cyclique, il existe pour toute racine m-ième de l'unité z, un unique caractère $\tilde{\chi}_z$ de H tel que $\tilde{\chi}_z(h) = z$. Prolongeons-les en des caractères χ_z de G. Par ailleurs, le groupe G/H est d'ordre $k = \mathrm{card}(G)/m < \mathrm{card}(G)$ et il existe par récurrence $\mathrm{card}(G)/m$ caractères de ce groupe ; notons-les $\bar{\theta}_1, \ldots, \bar{\theta}_k$. Pour $1 \leqslant i \leqslant k$, la composition θ_i de l'homomorphisme de G dans G/H et de $\bar{\theta}_i$ est un caractère de G dont la restriction à H est constante de valeur 1. Alors, pour tout couple (z, i) formé d'une racine m-ième de l'unité et d'un entier i tel que $1 \leqslant i \leqslant k$, l'application $\chi_z\theta_i$ est un caractère de G. Ces caractères sont deux

à deux distincts. Supposons en effet que $\chi_z\theta_i = \chi_u\theta_j$; par restriction à H, on trouve $z = u$; alors, $\theta_i = \theta_j$, ce qui entraîne par passage au quotient $\bar{\theta}_i = \bar{\theta}_j$, puis $i = j$. Nous avons ainsi construit $mk = \mathrm{card}(G)$ caractères distincts de G.

Inversement, si χ est un caractère de G, sa restriction à H est de la forme $\tilde{\chi}_z$, où z est une racine m-ième de l'unité. Alors, $\chi\chi_z^{-1}$ est un caractère de G qui applique tout élément de H sur 1. Ce caractère définit donc, par passage au quotient, un caractère du groupe G/H. Autrement dit, il existe un unique entier $i \in \{1,\dots,k\}$ tel que $\chi\chi_z^{-1} = \theta_i$, d'où $\chi = \chi_z\theta_i$. Cela démontre que tout caractère de G est parmi ceux que nous avons construit. Par conséquent, $\mathrm{card}(\hat{G}) = \mathrm{card}(G)$. □

Dans la démonstration, les caractères χ_z prennent la valeur $z \neq 1$ en l'élément h de G. Par conséquent :

LEMME 1.4.3. — *Soit G un groupe abélien fini et soit h un élément de G distinct de l'élément neutre. Il existe un caractère χ de G tel que $\chi(h) \neq 1$.*

La proposition suivante explique pourquoi les caractères permettent de transormer propriétés multiplicatives en propriétés additives.

PROPOSITION 1.4.4. — *Soit G un groupe abélien fini. Pour tout caractère χ de G, on a*

$$\sum_{g\in G} \chi(g) = \begin{cases} 0 & \text{si } \chi \neq 1 \text{ ;} \\ \mathrm{card}(G) & \text{si } \chi = 1. \end{cases}$$

De manière analogue, on a pour tout élément g de G :

$$\sum_{\chi\in\hat{G}} \chi(g) = \begin{cases} 0 & \text{si } g \neq 1 \text{ ;} \\ \mathrm{card}(G) & \text{si } g = 1. \end{cases}$$

Démonstration. — Le cas $\chi = 1$ est évident. Supposons $\chi \neq 1$ et soit $h \in G$ tel que $\chi(h) \neq 1$. Comme $g \mapsto hg$ est une permutation de G,

$$\sum_{g\in G} \chi(g) = \sum_{g\in G} \chi(hg) = \chi(h) \sum_{g\in G} \chi(g).$$

Puisque $\chi(h) \neq 1$, il vient nécessairement $\sum_g \chi(g) = 0$.

L'autre égalité se démontre de même en utilisant que si $g \neq 1$, il existe un caractère ψ tel que $\psi(g) \neq 1$: alors,

$$\sum_{\chi\in\hat{G}} \chi(g) = \sum_{\chi\in\hat{G}} (\chi\psi)(g) = \psi(g) \sum_{\chi\in\hat{G}} \chi(g),$$

d'où l'annulation de $\sum_\chi \chi(g)$ pour $g \neq 1$. Lorsque $g = 1$, on a évidemment

$$\sum_{\chi\in\hat{G}} \chi(g) = \mathrm{card}(\hat{G}) = \mathrm{card}(G).$$ □

B. Fonctions L de Dirichlet

Soit N un entier naturel tel que $N > 1$. Le groupe abélien fini qui va nous intéresser maintenant est le groupe $(\mathbf{Z}/N\mathbf{Z})^*$ est entiers modulo N qui sont inversibles. Comme la classe d'un entier a est inversible dans $\mathbf{Z}/N\mathbf{Z}$ si et seulement si a et N sont premiers entre eux, le cardinal de ce groupe $(\mathbf{Z}/N\mathbf{Z})^*$ est égal à $\varphi(N)$.

Si φ est une fonction du groupe $(\mathbf{Z}/N\mathbf{Z})^*$ dans $\mathbf{C}$, on identifiera φ à la fonction de $\mathbf{Z}$ dans $\mathbf{C}$ qui applique un entier a sur l'image par φ de la classe de a modulo N lorsque a est premier à N, et sur 0 sinon. En particulier, les caractères de $(\mathbf{Z}/N\mathbf{Z})^*$ sont identifiés aux fonctions $\chi : \mathbf{Z} \to \mathbf{C}$ vérifiant les propriétés suivantes :

a) $\chi(1) = 1$;

b) si $(a, N) \neq 1$, $\chi(a) = 0$;

c) si a et b sont deux entiers, alors $\chi(ab) = \chi(a)\chi(b)$.

On note χ_0 le caractère principal : c'est la fonction qui applique un entier sur 1 s'il est premier à N et sur 0 sinon. Elle correspond au caractère trivial de $(\mathbf{Z}/N\mathbf{Z})^*$.

Ces fonctions sont appelées *caractères de Dirichlet modulo* N. Ils sont en nombre $\varphi(N)$; de plus, pour tout caractère modulo N, disons χ, on a

$$\sum_{a=0}^{N-1} \chi(a) = \begin{cases} 0 & \text{si } \chi \neq \chi_0 ; \\ \varphi(N) & \text{si } \chi = \chi_0. \end{cases}$$

Pour tout caractère de Dirichlet modulo N, χ, on pose

$$L(\chi, s) = \sum_{n=1}^{\infty} \chi(n) \frac{1}{n^s}.$$

Cette série converge absolument si $\Re(s) > 1$, et la convergence est normale dans tout demi-plan Ω_a avec $a > 1$. Comme pour la fonction ζ de Riemann, le lien avec les nombres premiers provient de la décomposition en produit eulérien :

PROPOSITION 1.4.1. — *Pour tout nombre complexe s tel que $\Re(s) > 1$, on a*

$$L(\chi, s) = \prod_{\substack{p \in \mathscr{P} \\ p \nmid N}} \frac{1}{1 - \chi(p) p^{-s}}.$$

Démonstration. — La démonstration est presque identique à celle de la proposition 1.3.1. Elle résulte de l'unicité de la décomposition en facteurs premiers et de la propriété de multiplicativité satisfaite par les caractères de Dirichlet. □

Lorsque $\chi = \chi_0$, on voit donc que

$$L(\chi_0, s)/\zeta(s) = \prod_{p \mid N} \left(1 - p^{-s}\right)$$

est une fonction holomorphe sur le plan complexe ; en particulier, $L(\chi_0, s)$ a un prolongement méromorphe à $\mathbf{C}$, avec un unique pôle en $s = 1$, simple et de résidu

$$\prod_{p|N}\left(1-\frac{1}{p}\right)=\frac{\varphi(N)}{N}.$$

PROPOSITION 1.4.2. — *Pour tout caractère de Dirichlet modulo N distinct du caractère principal, la fonction $s \mapsto L(\chi, s)$ s'étend en une fonction holomorphe sur $\mathbf{C}$.*

Démonstration. — Contentons-nous de démontrer l'existence d'un prolongement à l'ouvert Ω_0. Nous effectuons une sommation d'Abel. Posons donc $F(n) = \sum_{m=1}^{n} \chi(m)$. Pour $\Re(s) > 1$, on a alors

$$L(\chi,s)=\sum_{n=1}^{\infty}\chi(n)n^{-s}=\sum_{n=1}^{\infty}\left(F(n)-F(n-1)\right)n^{-s}=\sum_{n=1}^{\infty}F(n)\left(n^{-s}-(n+1)^{-s}\right).$$

Comme $\sum_{m=1}^{N} \chi(m) = 0$, on a $F(n) = \sum_{m=1}^{n \bmod N} \chi(m)$ et la fonction F est bornée, majorée par $\varphi(N)$. En outre, lorsque n tend vers l'infini,

$$n^{-s}-(n+1)^{-s}=n^{-s}\left(1-(1+\frac{1}{n})^{-s}\right)=\mathrm{O}(n^{-s-1}).$$

Par suite, la série de terme général $F(n)(n^{-s} - (n+1)^{-s})$ converge absolument pour $\Re(s) > 0$, uniformément dans tout demi-plan Ω_a avec $a > 0$. Sa somme est une fonction holomorphe dans le demi-plan Ω_0 qui fournit le prolongement cherché de la fonction $L(\chi, s)$. □

Une variante de l'argument qui précède fournit une précision importante, à savoir que *pour tout caractère non principal χ, $L(\chi, s)$ est égale à la somme de la série de terme général $\chi(n)/n^s$ pour tout nombre complexe s tel que $\Re(s) > 0$.*

Le reste de la démonstration est très proche de celle du théorème des nombres premiers, le point crucial étant la non-annulation des fonctions L de Dirichlet sur l'axe $\{\Re(s) = 1\}$. Nous verrons cependant que l'existence d'une *infinité* de nombres premiers dans une progression arithmétique (par opposition au dénombrement asymptotique de tels nombres premiers) ne requiert « que » la non-annulation de $L(\chi, 1)$, lorsque χ est un caractère non-principal. Cela explique que le théorème de la progression arithmétique ait précédé celui des nombres premiers de près de 60 ans.

C. Non-annulation des fonctions L sur la droite $\Re(s) = 1$

PROPOSITION 1.4.1. — *Pour tout caractère de Dirichlet modulo N, χ, distinct du caractère principal, et pour tout nombre complexe s tel que $\Re(s) \geqslant 1$, on a $L(\chi, s) \neq 0$.*

Lorsque χ est le caractère principal, la même assertion vaut si l'on suppose de plus que $s \neq 1$.

Démonstration. — Cela résulte de l'expression de $L(\chi,s)$ comme produit infini si $\Re(s)>1$, la difficulté étant concentrée sur l'axe $\Re(s)=1$. Plutôt que de traiter séparément les diverses fonctions L, faisons leur produit et posons

$$Z(s)=\prod_{\chi} L(\chi,s).$$

Pour $\Re(s)>1$, on a

$$-\frac{Z'(s)}{Z(s)}=\sum_{\chi}\sum_{p}\frac{\chi(p)\log p}{p^s-\chi(p)}=\sum_{p}\left(\sum_{\chi}\chi(p)\right)\frac{\log p}{p^s}+\sum_{\chi}\sum_{p}\frac{\chi(p)^2\log p}{p^s(p^s-\chi(p))}.$$

Notons $\Phi(s)$ le premier terme ; si le second définit une fonction holomorphe dans le demi-plan $\Omega_{1/2}$, Φ n'est a priori holomorphe que dans Ω_1 et nous devons prouver qu'elle n'a pas de pôle sur l'axe $\Re(s)=1$. On a en outre

$$\Phi(s)=\varphi(N)\sum_{p\equiv 1 \pmod N}\frac{\log p}{p^s}.$$

Les pôles de Φ sur cet axe sont simples et proviennent du pôle simple de $L(\chi_0,s)$ en $s=1$ et des zéros éventuels des fonctions L associées aux caractères non-principaux. (Le cas du caractère principal résulte de ce qui a été fait pour la fonction zêta.) Fixons un nombre réel $\tau>0$. Notons q,m,n la somme des ordres des zéros des fonctions $L(\chi,s)$ en $s=1$, $s=1+i\tau$ et $s=1+2i\tau$ lorsque χ parcourt l'ensemble des caractères de Dirichlet modulo N. Autrement dit, on a $\operatorname{ord}_{s=1}Z(s)=q$, $\operatorname{ord}_{s=1+i\tau}Z(s)=m$ et $\operatorname{ord}_{s=1+2i\tau}Z(s)=n$. On a $q\geqslant -1$ (en $s=1$, le caractère principal fournit un pôle simple, les autres fonctions L n'ont pas de pôle) et $m,n\geqslant 0$; il s'agit de prouver que $q=-1$ et $m=0$.

Comme $\overline{L(\chi,s)}=L(\bar\chi,\bar s)$, les ordres d'annulation de $Z(s)$ en $s=1-i\tau$ et $s=1-2i\tau$ sont égaux à m et n respectivement. Reprenons l'astuce utilisée pour la fonction zêta de Riemann, observons que pour $\sigma>1$,

$$\sum_{k=-2}^{2}\Phi(\sigma+i\tau k)=\sum_{p\equiv 1}\frac{\log p}{p^\sigma}\big(2\cos(\tau\log p)\big)^4\geqslant 0.$$

Par suite, multipliant par $(\sigma-1)$ et faisant tendre σ vers 1, on obtient

$$-n-4m-6q-4m-n\geqslant 0,$$

d'où $8m+2n\leqslant -6q$. Nécessairement, $q\leqslant 0$. Si $q=-1$ (ce qu'on veut), il vient $m=0$.

Il reste à exclure le cas où $q=0$: dans ce cas, il existe un caractère modulo N, et un seul, χ, tel que $L(\chi,s)$ ait un zéro simple en $s=1$. Alors, $L(\bar\chi,1)=\overline{L(\chi,1)}=0$, donc $\bar\chi=\chi$ et le caractère χ est à valeurs réelles. L'assertion voulue résulte alors de la proposition suivante. □

PROPOSITION 1.4.2. — *Si χ est un caractère modulo N, non principal et réel, on a $L(\chi,1)\neq 0$.*

Démonstration (d'après [**6**]*).* — Pour $t \in [0,1[$, posons

$$f(t) = \sum_{d=1}^{\infty} \chi(d) \frac{t^d}{1-t^d}.$$

La série converge pour tout t dans l'intervalle indiqué, et la convergence est normale sur chaque intervalle $[0,a]$ avec $a<1$. En développant $1/(1-t^d)$ en série, on obtient

$$f(t) = \sum_{d=1}^{\infty} \chi(d) \sum_{m=1}^{\infty} t^{dm} = \sum_{n=1}^{\infty} t^n \sum_{d|n} \chi(d)$$

en regroupant les termes de même valeur $n = dm$.

Posons $c_n = \sum_{d|n} \chi(d)$. Montrons que $c_n \geqslant 0$ pour tout entier n. Si $n = p^a$ est une puissance d'un nombre premier, on a

$$c_{p^a} = 1 + \chi(p) + \chi(p)^2 + \cdots + \chi(p^a) \geqslant 0.$$

Plus précisément, la somme vaut $a+1$ si $\chi(p) = 1$, 1 si $p|N$; si $\chi(p) = -1$, elle vaut 1 ou 0 suivant que a est pair ou impair. En outre, si $n = \prod p_i^{a_i}$ est la décomposition en facteurs premiers de l'entier n, un diviseur d de n est de la forme $\prod p_i^{m_i}$ avec $0 \leqslant m_i \leqslant a_i$, d'où

$$c_n = \sum_{m_1=0}^{a_1} \cdots \sum_{m_r=0}^{a_r} \chi(p_1^{m_1} \ldots p_r^{m_r}) = \prod_{i=1}^{r} c_{p_i^{a_i}} \geqslant 0.$$

Nous avons remarqué que $c_n = 1$ si n est une puissance d'un nombre premier qui divise N ; en particulier, $\sum c_n = +\infty$.

Par suite, lorsque t tend vers 1 par valeurs inférieures, $f(t)$ tend vers $+\infty$. Lorsque $t \to 1$, $t^n/(1-t^n)$ est équivalent à $1/n(1-t)$. Posons ainsi

$$b_n = (1-t)\left(\frac{t^n}{1-t^n} - \frac{1}{n(1-t)}\right) = \frac{t^n}{1+t+\cdots+t^{n-1}} - \frac{1}{n}.$$

Démontrons que la suite (b_n) est croissante. En effet, on a

$$b_{n+1} - b_n = \frac{t^n}{(1+t+\cdots+t^{n-1})(1+t+\cdots+t^n)} - \frac{1}{n(n+1)}.$$

L'inégalité entre moyennes arithmétique et géométriques entraîne par ailleurs

$$1+t+\cdots+t^{n-1} \geqslant n t^{(n-1)/2} \quad \text{et} \quad 1+t+\cdots+t^n \geqslant (n+1)\, t^{(n+1)/2},$$

si bien que

$$b_{n+1} - b_n \geqslant \frac{t^n}{n(n+1)\,t^n} - \frac{1}{n(n+1)} = 0.$$

Supposons par l'absurde que $L(\chi,1) = 0$; alors $\sum_{n=1}^{\infty} \chi(n)/n = 0$ et

$$f(t) = \sum_{n=1}^{\infty} \chi(n)\left(\frac{t^n}{1-t^n} - \frac{1}{n(1-t)}\right) = \frac{1}{1-t} \sum_{n=1}^{\infty} b_n \chi(n).$$

Puisque χ est périodique de période N et que $\sum_{n=1}^{N}\chi(n)=0$, on a $\left|\sum_{d=1}^{n}\chi(d)\right|\leqslant N$ pour tout entier n. Effectuons alors une transformation d'Abel :

$$\left|\sum_{n=1}^{\infty}b_n\chi(n)\right|=\left|\sum_{d=1}^{\infty}\left(\sum_{n=1}^{d}\chi(n)\right)(b_d-b_{d+1})\right|\leqslant N\sum_{d=1}^{\infty}(b_d-b_{d+1})=N(1-t).$$

On a donc $\left|f(t)\right|\leqslant N$ pour tout $t\in[0,1[$ ce qui contredit le fait que $f(t)\to\infty$ pour $t\to 1$. □

D. Conclusion de la démonstration

Soit a un entier qui est premier à N. Pour $\Re(s)>1$, posons

$$\Phi(s;a)=\sum_{p\equiv a\pmod N}\frac{\log p}{p^s}.$$

Posons aussi, pour tout nombre réel $x>0$,

$$\theta(x;a)=\sum_{\substack{p\leqslant x\\ p\equiv a}}\log p.$$

PROPOSITION 1.4.1. — *La série qui définit Φ_a converge pour $\Re(s)>1$; sa somme se prolonge en une fonction méromorphe dans l'ouvert $\Omega_{1/2}$ qui n'a pas de pôle dans un voisinage du demi-plan $\{Re(s)\geqslant 1\}$, excepté un pôle simple en $s=1$, de résidu $1/\varphi(N)$.*

Démonstration. — Si χ est un caractère de Dirichlet modulo N et s un nombre complexe de partie réelle >1, la dérivée logarithmique de $L(\chi,s)$ admet le développement en série

$$-\frac{L'(\chi,s)}{L(\chi,s)}=\sum_p\chi(p)\frac{\log p\,p^{-s}}{1-\chi(p)p^{-s}}=\sum_p\chi(p)\frac{\log p}{p^s}+\sum_p\chi^2(p)\frac{\log p}{p^s(p^s-\chi(p))}.$$

Le membre de gauche est méromorphe dans le demi-plan Ω_0 et n'a pas de pôle dans un voisinage du demi-plan $\{\Re(s)=1\}$, hormis, si χ est le caractère principal χ_0, un pôle simple de résidu 1 en $s=1$ causé par le pôle simple de $L(\chi_0,s)$ en $s=1$. Le second terme du second membre définit une fonction holomorphe $\Psi(\chi,s)$ dans le demi-plan $\Omega_{1/2}$; par suite, le premier terme du second membre définit une fonction méromorphe dans $\Omega_{1/2}$ qui n'a pas de pôle au voisinage du demi-plan fermé $\{\Re(s)=1\}$, sauf en $s=1$ si $\chi=\chi_0$. Par suite, sommant sur l'ensemble des caractères de Dirichlet modulo N, il vient

$$\begin{aligned}-\sum_\chi\frac{L'(\chi,s)}{L(\chi,s)}\chi(a)^{-1}&=\sum_p\left(\sum_\chi\chi(p)\chi(a)^{-1}\right)\frac{\log p}{p^s}+\sum_\chi\Psi(\chi,s)\chi(a)^{-1}\\&=\varphi(N)\sum_{p\equiv a\pmod N}\frac{\log p}{p^s}+\sum_\chi\Psi(\chi,s)\chi(a)^{-1}\\&=\varphi(N)\Phi(s;a)+\sum_\chi\Psi(\chi,s)\chi(a)^{-1}\end{aligned}$$

puisque pour tout nombre entier p, notant b un inverse de a modulo N, on a

$$\sum_{\chi} \chi(p)\chi(a)^{-1} = \sum_{\chi} \chi(pb) = \begin{cases} \varphi(N) & \text{si } pb \equiv 1 \pmod{N}\,; \\ 0 & \text{sinon.} \end{cases}$$

Par conséquent, la fonction $\Phi(s;a)$ s'étend en une fonction méromorphe dans le demi-plan $\Omega_{1/2}$ sans pôle au voisinage du demi-plan fermé $\{\Re(s) = 1\}$ excepté en $s = 1$ où elle a un pôle simple de résidu $1/\varphi(N)$ issu du terme pour $\chi = \chi_0$ (comme a est premier à N, $\chi_0(a) = 1$). □

À ce stade, il reste à recopier ce qu'on avait fait pour le théorème des nombres premiers. On démontre successivement, par les mêmes arguments, que l'intégrale

$$\int_1^{\infty} \left(\theta(x;a) - \frac{x}{\varphi(N)}\right) \frac{dx}{x^2}$$

converge, qu'au voisinage de $+\infty$,

$$\theta(x;a) \sim \frac{1}{\varphi(N)} x$$

et enfin que

$$\pi(x;a) \sim \frac{1}{\varphi(N)} \frac{x}{\log(x)}.$$

Si l'on souhaite éviter l'emploi du théorème taubérien mais ne démontrer que l'existence d'une infinité de nombres premiers congrus à a modulo N, il suffit d'observer que $\Phi(s;a)$ tend vers l'infini quand s tend vers 1, alors qu'il aurait une limite finie si la progression arithmétique considérée était finie.

Exercices

25) Soit G un groupe abélien fini et soit V l'espace vectoriel des fonctions de G dans $\mathbf{C}$ muni du produit scalaire hermitien

$$\langle \varphi, \psi \rangle = \frac{1}{\mathrm{card}(G)} \sum_{g \in G} \bar{\varphi}(g)\psi(g).$$

Démontrer que les caractères de G forment une base orthonormée de V.

26) Soit G un groupe abélien fini.

Soit h un élément de G dont l'ordre m est multiple de l'ordre de tout élément de G et soit H le sous-groupe de G engendré par h. Démontrer qu'il existe un caractère χ de G qui applique h sur une racine primitive m-ième de l'unité dans $\mathbf{C}$.

a) Démontrer que l'image de χ est le sous-groupe de $\mathbf{C}^*$ formé des racines primitives de l'unité.

b) Soit K le noyau de χ. Démontrer que $H \cap K = \{1\}$ puis que G est isomorphe au produit $H \times K$.

27) Soit G un groupe abélien fini.

a) Démontrer que G est un produit de groupes cycliques. (Utiliser l'exercice 26.)

b) Démontrer que les groupes G et $\hat{G}$ sont isomorphes.

CHAPITRE 2

CORPS FINIS

§2.1. Extensions de corps

A. Rappels

Soit K un corps. Une *extension du corps* K est un corps E muni d'un homomorphisme de K dans E. Un tel homomorphisme est injectif, ce qui permet en pratique de considérer que K est un *sous-corps* de E, ou que E est un *sur-corps* de K. Lorsque $K \hookrightarrow E$ est une extension de corps, on considère E comme une K-algèbre et comme un espace vectoriel sur K ; on note $[E : K]$ sa dimension, et on l'appelle le degré de l'extension $K \hookrightarrow E$, mais cette notation peut être trompeuse si K n'est pas un sous-corps de E.

Soit $K \hookrightarrow E$ une extension de corps ; un élément x de E est dit *algébrique* sur K s'il existe un polynôme non nul P à coefficients dans K tel que $P(x) = 0$. (En toute bonne rigueur, si $i\colon K \hookrightarrow E$ est l'homomorphisme qui définit l'extension, il faudrait noter $i(P)(x) = 0$, où $i(P)$ désigne le polynôme à coefficients dans E dont les coefficients sont les images par i des coefficients de P. On s'autorisera dans la suite ce genre d'abus de langage, à moins que cela puisse créer des confusions dangereuses. Si x n'est pas algébrique, on dit qu'il est *transcendant*.

Si $x \in E$ est algébrique, l'ensemble des polynômes $P \in K[X]$ tels que $P(x) = 0$ est un idéal de $K[X]$, non réduit à 0. Il est donc formé des multiples d'un polynôme unitaire M_x, qu'on appelle *le polynôme minimal de* x et qui est le polynôme unitaire de degré minimal P tel que $P(x) = 0$. Le degré de M_x est appelé le degré de x. Ce polynôme M_x est irréductible : si l'on avait une factorisation non triviale $M_x = PQ$, on aurait $P(x)Q(x) = 0$, d'où $P(x) = 0$ ou $Q(x) = 0$ ce qui contredit l'hypothèse que M_x est de degré minimal.

Soit $K \hookrightarrow E$ une extension de corps et soit x un élément de E. L'ensemble $K[x]$ des éléments de E de la forme $P(x)$, où $P \in K[X]$, est un sous-algèbre de E. Si x est transcendant, elle est isomorphe à $K[X]$; elle est donc de dimension infinie sur K et n'est pas un corps. Si x est algébrique, elle est isomorphe à l'anneau quotient $K[X]/(M_x)$. C'est un sous-corps de E.

Une extension $K \hookrightarrow E$ est dite algébrique si tout élément de E est algébrique sur K.

PROPOSITION 2.1.1. — *Toute extension de degré fini est algébrique. Plus généralement, soit $K \hookrightarrow E$ une extension de corps. Toute sous-K-algèbre A de E qui est de dimension finie est un corps et tout élement de A est algébrique sur K.*

Démonstration. — Soit $K \hookrightarrow E$ une extension de degré fini, n, et soit x un élément de E. Les éléments $1, x, \dots, x^n$ de E sont donc linéairement dépendants et il existe des éléments $a_0, \dots, a_n$ de K, non tous nuls, tels que $a_0 + a_1 x + \cdots + a_n x^n = 0$. Par suite, x est algébrique sur K, de degré $\leqslant n$.

Soit $K \hookrightarrow E$ une extension de corps et soit A une sous-algèbre de E qui est un K-espace vectoriel de dimension finie. Soit x un élément non nul de A. La multiplication par x est un endomorphisme injectif du K-espace vectoriel A car A est un anneau intègre. Cet endomorphisme est donc surjectif et, en particulier, x est inversible dans A. Cela démontre que A est un corps. Alors, $K \hookrightarrow A$ est une extension de degré fini. D'après la première partie de la proposition, tout élément de A est algébrique sur K. □

COROLLAIRE 2.1.2. — *Soit $K \hookrightarrow E$ une extension de corps. L'ensemble des éléments de E qui sont algébriques sur K est un sous-corps de E. En particulier, la somme, le produit de deux éléments algébriques de E est algébrique ; l'inverse d'un élément algébrique non nul de E est algébrique.*

Démonstration. — Soit x, y des éléments de E qui sont algébriques sur K ; notons m et n leurs degrés. Soit alors A l'ensemble des éléments de E de la forme $P(x, y)$, où $P \in K[X, Y]$ est de degré $< m$ en X et de degré $< n$ en Y. C'est un sous-K-espace vectoriel de E qui est stable par la multiplication par x et y ; comme il contient 1, c'est une sous-algèbre de E. Elle est engendrée par la famille finie formée des éléments de la forme $x^i y^j$ avec $0 \leqslant i < m$ et $0 \leqslant j < n$, donc est de dimension finie sur K. D'après la proposition précédente, c'est un corps et tout élément de A est algébrique sur K.

En particulier, $x + y$ et xy sont algébriques sur K, de même que $1/x$ si $x \neq 0$.

Il en résulte aussitôt que l'ensemble des éléments de E qui sont algébriques sur K est un sous-corps de E. □

Soit $K \hookrightarrow E$ une extension de corps et soit x, y des éléments de E qui sont algébriques. Si l'on souhaite déterminer explicitement un polynôme annulateur de $x + y$, le théorème de Cayley-Hamilton s'avère très utile. Connaissant des polynômes P et Q qui annulent x et y, il est en effet très aisé d'écrire une matrice $(a_{(i,j)(k,l)})$ de taille mn, indexée par l'ensemble des couples (i, j) avec $0 \leqslant i < m$ et $0 \leqslant j < n$, telle que

$$(x + y) x^k y^l = \sum a_{(i,j)(k,l)} x^i y^j.$$

Soit Π le polynôme caractéristique de la matrice A. D'après la proposition suivante, le polynôme Π annule l'endomorphisme de multiplication par $x+y$. En particulier, $P(x+y)=0$.

PROPOSITION 2.1.3. — *Soit V un espace vectoriel sur un corps K, soit $(v_1,\dots,v_n)$ une famille génératrice de V. Soit f un endomorphisme de V et soit A une matrice de taille n telle que*

$$f(v_j)=\sum_{i=1}^{n} a_{ij}v_i$$

pour tout entier $j\in\{1,\dots,n\}$. Si P est le polynôme caractéristique de A, on a $P(f)=0$.

Démonstration. — Notons $u\colon K^n\to V$ l'homomorphisme qui applique le i-ième vecteur de la base canonique de K^n, e_i, sur v_i. Soit φ l'endomorphisme de K^n de matrice A. On a $u\circ\varphi=f\circ u$. D'après le théorème de Cayley-Hamilton, $P(\varphi)=0$. Par suite, $0=u\circ P(\varphi)=P(f)\circ u$. Comme u est surjectif, $P(f)=0$. □

Donnons un exemple en exhibant un polynôme annulateur de $\sqrt{2}+\sqrt{3}$. Avec les notations précédentes, $x=\sqrt{2}$ et $y=\sqrt{3}$ et la famille génératrice de l'algèbre $K[x,y]$ est $1,\sqrt{2},\sqrt{3},\sqrt{6}$. On a les relations

$$\begin{aligned}(\sqrt{2}+\sqrt{3})\cdot 1&=0+\sqrt{2}+\sqrt{3}+0\\(\sqrt{2}+\sqrt{3})\cdot\sqrt{2}&=2+0+0+\sqrt{6}\\(\sqrt{2}+\sqrt{3})\cdot\sqrt{3}&=3+0+0+\sqrt{6}\\(\sqrt{2}+\sqrt{3})\cdot\sqrt{6}&=0+3\sqrt{2}+2\sqrt{3}+0,\end{aligned}$$

donc on introduit la matrice

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

Son polynôme caractéristique est X^4-10X^2+1 et l'on a bien

$$(\sqrt{2}+\sqrt{3})^4-10(\sqrt{2}+\sqrt{3})^2+1=0.$$

Une autre méthode, plus systématique, consiste à utiliser la théorie des résultants.

B. Extensions de rupture, extensions de décomposition, clôture algébrique

Soit K un corps et soit P un polynôme *irréductible* à coefficients dans K, notons d son degré. Si $d\geqslant 2$, alors P n'a pas de racine dans K. Nous allons construire une extension de K dans laquelle P a une racine et qui est « universelle » pour cette propriété.

Notons L l'anneau $K[X]/(P)$, quotient de l'anneau des polynômes par l'idéal principal engendré par P. Comme l'anneau $K[X]$ est un anneau principal et que P est irréductible cet idéal est un idéal maximal ; autrement dit l'anneau L est un corps. C'est

donc une extension de K. Notons x la classe de X dans L ; la division euclidienne par P entraîne que $(1, x, \dots, x^{d-1})$ est une base de L sur K. Par suite, le degré $[L : K]$ de cette extension est égal à d.

De plus, comme $P(x)$ est la classe du polynôme $P(X)$ dans L, on a $P(x) = 0$; ainsi, x est une racine de P dans le corps L. Enfin, tout élément de L est un polynôme en x. Nous dirons que L, et plus généralement toute extension de K qui est engendrée par une racine de P, est une *extension de K obtenue par adjonction d'une racine de P*, ou une *extension de rupture du polynôme P*, voire encore un *corps de rupture du polynôme P*.

Si E est une extension de K et $f : L \to E$ est un morphisme d'extensions, $f(x)$ est une racine de P dans E. Comme tout élément de L est un polynôme en x, l'homomorphisme f est déterminé par cette racine $f(x)$.

Inversement, soit E une extension de K et y une racine de P dans E. Notons g l'application de $K[X]$ dans E donnée par $g(Q) = Q(y)$ si $Q \in K[X]$; c'est un homomorphisme de K-algèbres. En outre, $P(x) = 0$ si bien que tout élément de l'idéal (P) a pour image 0 par g. Par passage au quotient, on en déduit donc un homomorphisme f de L dans E tel que $f(x) = g(X) = y$. Si E est engendrée par y, alors g est surjectif, donc f est surjectif. Par suite, c'est un isomorphisme d'extensions.

Nous avons ainsi démontré le résultat suivant :

PROPOSITION 2.1.1. — *Soit L une extension de rupture d'un polynôme irréductible $P \in K[X]$, et soit x une racine de P dans L.*

a) Pour tout couple (E, y) formé d'une extension E de K et d'une racine y de P dans E, il existe un et un seul homomorphisme d'extensions $f : L \to E$ tel que $f(x) = y$.

b) Si E est une extension de rupture de P et y une racine de P dans E, cet homomorphisme f est un isomorphisme.

Soit K un corps et soit P un polynôme à coefficients dans K ; on ne suppose pas P irréductible. Si P n'est pas scindé dans K, choisissons un facteur irréductible Q de P de degré $d \geqslant 2$ et soit $K \to L$ une extension de rupture de Q et x une racine de Q dans L. Dans $L[X]$, le polynôme P est multiple de $X - x$; appliquons alors cet argument au polynôme quotient $P(X)/(X - x)$. On construit ainsi par récurrence une extension E de K dans laquelle le polynôme P est scindé. Notons $x_1, \dots, x_d$ ses racines dans E, répétées suivant leur multiplicité, de sorte que $P(X) = a(X - x_1) \dots (X - x_d)$ dans $E[X]$. Quitte à remplacer E par sa sous-extension engendrée par les x_i, on peut supposer que $E = K[x_1, \dots, x_d]$.

Une telle extension, dans laquelle P est scindé et qui est engendrée par les racines de P, est appelée *extension de décomposition du polynôme P*.

Pour toute extension F de K dans laquelle P est scindé, on construit maintenant, par récurrence descendante sur le nombre de racines de P dans K, un homomorphisme d'extensions de E dans F. Si P est scindé dans K, on a $E = K$ et il n'y a rien à démontrer. Soit sinon Q un facteur irréductible de P dans $K[X]$ dont le degré est supérieur ou égal à 2 et soit x une racine de Q dans E. Alors, $K[x]$ est une extension de rupture du polynôme Q. Choisissons arbitrairement une racine y de Q dans F ; il en existe puisque Q divise P est que P est scindé dans F. Il existe alors un homomorphisme d'extensions f_1 de $K[x]$ dans F qui applique x sur y. Comme P a au moins une racine de plus dans $K[x]$ qu'il n'en avait dans K, on en conclut par récurrence que l'on peut prolonger l'homomorphisme f_1 en un homomorphisme d'extensions de E dans F.

Les racines de P dans F sont les images $f(x_1), \ldots, f(x_d)$ par f des racines de P dans E. Par conséquent, si l'on suppose de plus que F est une extension de décomposition de P, F est engendrée par ces racines $f(x_i)$ et f est surjectif ; c'est donc un isomorphisme.

PROPOSITION 2.1.2. — *Soit P un polynôme à coefficients dans K et soit $K \to E$ une extension de décomposition de P. Soit Ω une extension de K ; Pour qu'il existe un homomorphisme d'extensions de E dans Ω, il faut et il suffit que P soit scindé dans Ω. Si, de plus, Ω est une extension de décomposition, un tel homomorphisme est un isomorphisme.*

Soit K un corps. On dit que K est *algébriquement clos* si tout polynôme non constant (à coefficients dans K) a une racine dans K. Rappelons que le corps $\mathbf{C}$ des nombres complexes est algébriquement clos (théorème de d'Alembert-Gauß). Une *clôture algébrique* de K est une extension algébrique Ω de K dans laquelle tout polynôme à coefficients dans K est scindé.

PROPOSITION 2.1.3. — *Soit K un corps et soit Ω une extension de K qui est un corps algébriquement clos. Soit E l'ensemble des éléments de Ω qui sont algébriques sur K ; alors E est une clôture algébrique de K.*

Concernant l'existence et l'unicité des clôtures algébriques, on a le résultat suivant :

THÉORÈME 2.1.4 (Steinitz). — *Tout corps possède une clôture algébrique. En outre, si K est un corps et si Ω_1, Ω_2 sont des clôtures algébriques de K, il existe un isomorphisme d'extensions de Ω_1 dans Ω_2.*

Nous admettrons ce résultat ; il n'est pas difficile mais sa mise en place rigoureuse requiert un peu de théorie des ensembles. Signalons notamment qu'il dépend de l'axiome du choix.

C. Norme, trace

Soit K un corps et soit E une K-algèbre de dimension finie.

La multiplication μ_a par un élément a de E, définit un endomorphisme du K-espace vectoriel E, donné par $x \mapsto ax$.

Le déterminant et la trace (dans une base arbitraire de E) de cet endomorphisme sont appelés *norme* et *trace* de a et sont notés $\mathrm{N}_{E/K}(a)$ et $\mathrm{Tr}_{E/K}(a)$. Comme l'application $a \mapsto \mu_a$ applique la somme $a+a'$ de deux éléments de E sur la somme $\mu_a+\mu_{a'}$ des endomorphismes μ_a et $\mu_{a'}$, on a $\mathrm{Tr}_{E/K}(aa') = \mathrm{Tr}_{E/K}(a) + \mathrm{Tr}_{E/K}(a')$.

Si a appartient à K, on a $\mathrm{Tr}_{E/K}(a) = [E:K]a$ et $\mathrm{N}_{E/K}(a) = a^{[E:K]}$.

LEMME 2.1.1 (Transitivité des norme et trace). — *Soit $K \to L$ une extension de corps de degré fini d et soit E une L-algèbre de dimension finie e. Alors, E est une K-algèbre de dimension de. De plus, pour tout élément $a \in E$, on a les relations*

$$\mathrm{Tr}_{E/K}(a) = \mathrm{Tr}_{L/K}(\mathrm{Tr}_{E/L}(a)) \quad \textit{et} \quad \mathrm{N}_{E/K}(a) = \mathrm{N}_{L/K}(\mathrm{N}_{E/L}(a)).$$

Démonstration. — Soit $(x_1,\dots,x_d)$ une base de L sur K et soit $(y_1,\dots,y_e)$ une base de E sur L. Alors, la famille $(x_i y_j)$ est une base de E sur K.

Soit a un élément de E et soit $A = (a_{ij})$ la matrice dans la base $(y_1,\dots,y_e)$ de la multiplication par a dans le L-espace vectoriel E. La matrice dans la base $(x_i y_j)$ de la multiplication par a du K-espace vectoriel E s'écrit par blocs $d \times d$, $(A_{i,j})$, le bloc $A_{i,j}$ étant la matrice de la multiplication par $a_{i,j}$ dans la base $(x_1,\dots,x_d)$.

On a donc déjà

$$\mathrm{Tr}_{E/K}(a) = \sum_{i=1}^{e} \mathrm{Tr}(A_{i,i}) = \sum_{i=1}^{e} \mathrm{Tr}_{L/K}(a_{i,i}) = \mathrm{Tr}_{L/K}\Big(\sum_{i=1}^{e} a_{i,i}\Big) = \mathrm{Tr}_{L/K}(\mathrm{Tr}_{E/L}(a)).$$

D'autre part, comme les matrices $A_{i,j}$ commutent deux à deux le déterminant de la matrice par blocs $(A_{i,j})$ se calcule par la formule (lemme 2.1.2 ci-dessous) :

$$\det((A_{i,j})) = \det\left(\sum_{\sigma\in\Sigma_e} \varepsilon_\sigma \prod_{i=1}^{e} A_{i,\sigma(i)}\right).$$

La matrice dont le déterminant prend le second membre est celle de la multiplication par $\mathrm{N}_{E/L}(a)$; le second membre est donc égal à $\mathrm{N}_{L/K}(\mathrm{N}_{E/L}(a))$. Par suite, $\mathrm{N}_{E/K}(a) = \det((A_{i,j})) = \mathrm{N}_{L/K}(\mathrm{N}_{E/L}(a))$ comme annoncé. □

Il reste à démontrer le lemme annoncé :

LEMME 2.1.2. — *Soit K un corps ; soit $(A_{i,j})_{1\leqslant i,j\leqslant e}$ une famille de matrices de taille $d \times d$, à coefficients dans K, commutant deux à deux ; et soit A la matrice par blocs $(A_{i,j})$, de taille $de \times de$. Alors, le déterminant de A est égal au déterminant de la matrice*

$$\sum_{\sigma\in\mathfrak{S}_e} \varepsilon_\sigma \prod_{i=1}^{e} A_{i,\sigma(i)}.$$

Démonstration. — □

Soit K un corps, soit $K \to L$ une extension finie et soit a un élément de L. Nous allons calculer norme et trace de a en fonction du polynôme minimal de a et de ses racines dans une extension algébriquement close de K.

PROPOSITION 2.1.3. — *Soit K un corps, soit $K \to L$ une extension finie. Soit a un élément de L et soit $P \in K[X]$ son polynôme minimal, noté $P = X^d + p_1 X^{d-1} + \cdots + p_d$; soit e l'entier tel que $de = [L : K]$. Alors,*

$$\mathrm{Tr}_{L/K}(a) = -ep_1, \quad \mathrm{N}_{L/K} = \left((-1)^d p_d\right)^e.$$

En outre, soit Ω une extension algébriquement close de K et notons $a_1, \ldots, a_d$ les racines de P dans Ω. On a alors

$$\mathrm{Tr}_{L/K}(a) = e\sum_{j=1}^{d} a_j, \quad \mathrm{N}_{L/K} = \prod_{j=1}^{d} a_j^e.$$

Démonstration. — On a l'égalité $[K(a) : K] = d$ et $(1, x, \ldots, x^{d-1})$ est une base de $K(a)$ sur K. Dans cette base, la matrice de la multiplication par a est la matrice compagnon C_P du polynôme P :

$$C_P = \begin{pmatrix} 0 & \ldots & \ldots & 0 & -p_d \\ 1 & \ddots & & \vdots & -p_{d-1} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -p_2 \\ 0 & \ldots & 0 & 1 & -p_1 \end{pmatrix}$$

dont la trace vaut $-p_1$ et le déterminant $(-1)^d p_d$.

On a aussi $[L : K(a)] = [L : K]/[K(a) : K] = e$; soit $(x_1, \ldots, x_e)$ une base de L sur $K(a)$. Dans la base $(x^i x_j)$ de L sur K, la matrice A de la multiplication par a est diagonale par blocs, chaque bloc étant égal à C_P. On a alors $\det(A) = (\det C_P)^e = \left((-1)^d p_d\right)^e$ et $\mathrm{Tr}(A) = e\,\mathrm{Tr}(A) = -ep_1$.

La dernière formule résulte alors des relations coefficients-racines, elles-mêmes conséquences de l'identification des coefficients dans l'égalité

$$P(X) = (X - a_1)\ldots(X - a_d) = X^d + p_1 X^{d-1} + \cdots + p_d.$$

□

D. Résultant, discriminant

Soit K un corps ; on désigne par $K[X]_{<n}$ l'espace vectoriel des polynômes de degrés $< n$. Soit A, B des polynômes à coefficients dans K de degrés $\leqslant n, m$ respectivement ; notons-les $A = a_0 + a_1 X + \cdots + a_n X^n$ et $B = b_0 + b_1 X + \cdots + b_m X^m$. On appelle *résultant* (en degrés (n, m)) des polynômes A et B, et on note $\mathrm{Res}_{n,m}(A, B)$, le déterminant de l'application linéaire de $K[X]_{<m} \times K[X]_{<n}$ dans $K[X]_{<m+n}$ donnée par $(U, V) \mapsto UA + VB$, pris dans les bases évidentes $(X^{m-1}, \ldots, X, 1; X^{n-1}, \ldots, X, 1)$ et $(X^{m+n-1}, \ldots, X, 1)$. Dans ces bases, la matrice de cette application linéaire (appelée *application résultante*) est la transposée de la *matrice de Sylvester* :

$$\begin{pmatrix} a_n & a_{n-1} & \ldots & a_0 & & \\ & \ddots & \ddots & & \ddots & \\ & & a_n & a_{n-1} & \ldots & a_0 \\ b_m & b_{m-1} & \ldots & b_0 & & \\ & \ddots & \ddots & & \ddots & \\ & & b_m & b_{m-1} & \ldots & b_0 \end{pmatrix},$$

où il y a m lignes de a et n ligne de b.

Si l'on échange les rôles de A et B, la matrice de Sylvester subit une permutation des lignes — il suffit de remonter successivement chaque ligne de b de n positions. On a ainsi la relation

$$\mathrm{Res}_{n,m}(A, B) = (-1)^{mn} \mathrm{Res}_{m,n}(B, A). \tag{E}$$

PROPOSITION 2.1.1. — *Si* $\deg(A) < n$ *et* $\deg(B) < m$, *alors* $\mathrm{Res}_{n,m}(A, B) = 0$. *Supposons que l'on a* $\deg(A) = n$ *ou* $\deg(B) = m$ *; alors* $\mathrm{Res}_{n,m}(A, B) = 0$ *si et seulement si* A *et* B *ont un facteur commun.*

Démonstration. — Soit D le pgcd de A et B et notons A_1, B_1 les polynômes tels que $A = DA_1$ et $B = DB_1$. Si U et V sont des polynômes, la relation $UB + VA = 0$ équivaut aux relations $U = WB_1$, $V = -WA_1$, où $W \in K[X]$.

Si $\deg(A_1) < n$ et $\deg(B_1) < m$, le couple (B_1, A_1) est un élément non nul du noyau de l'application résultante, par suite, $\mathrm{Res}_{n,m}(A, B) = 0$. En particulier, $\mathrm{Res}_{n,m}(A, B)$ s'annule lorsque A et B ont un facteur commun, ou lorsque $\deg(A) < n$ et $\deg(B) < m$.

Si $D = 1$ et que l'on a $\deg(A) = n$, alors tout couple (U, V) appartenant au noyau de l'application résultante vérifie $\deg(V) \geqslant n$ ou $V = 0$; cette application est donc injective, d'où donc $\mathrm{Res}_{n,m}(A, B) \neq 0$. On raisonne de même lorsque $D = 1$ et $\deg(B) = m$. □

Donnons maintenant une formule pour le résultant en fonction des racines de A et B dans une extension de K.

PROPOSITION 2.1.2. — *Soit Ω une extension de K dans laquelle on a les factorisations*

$$A(X) = a_n \prod_{i=1}^{n} (X - x_i), \quad B(X) = b_m \prod_{j=1}^{m} (X - y_j).$$

Alors,

$$\mathrm{Res}_{n,m}(a,B) = a_n^m b_m^n \prod_{i=1}^{n}\prod_{j=1}^{m} (x_i - y_j) = a_n^m \prod_{i=1}^{n} B(x_i) = (-1)^{mn} b_m^n \prod_{j=1}^{m} A(y_j).$$

Démonstration. — Notons S la matrice de Sylvester introduite ci-dessus et V la matrice de Vandermonde

$$\begin{pmatrix} y_1^{m+n-1} & \cdots & y_m^{m+n-1} & x_1^{m+n-1} & \cdots & x_n^{m+n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ y_1^2 & y_2^2 & \cdots & y_m^2 & x_1^2 & \cdots & x_n^2 \\ y_1 & y_2 & \cdots & y_m & x_1 & \cdots & x_n \\ 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \end{pmatrix}.$$

On a

$$SV = \begin{pmatrix} y_1^{m-1}A(y_1) & \cdots & y_m^{m-1}A(y_m) & & & \\ \vdots & & \vdots & & 0 & \\ y_1 A(y_1) & & y_m A(y_m) & & & \\ A(y_1) & \cdots & A(y_m) & & & \\ & & & x_1^{n-1}B(x_1) & \cdots & x_n^{n-1}B(x_n) \\ & 0 & & \vdots & & \vdots \\ & & & B(x_1) & \cdots & B(x_n) \end{pmatrix}.$$

Les colonnes sont multiples de $A(y_1),\dots,A(y_m),B(x_1),\dots,B(x_n)$. Par suite, en appliquant la formule classique pour le déterminant de Vandermonde, il vient

$$\det(S)\det(V) = \prod_{j=1}^{m} A(y_j) \prod_{i=1}^{n} B(x_i) \prod_{j'<j} (y_{j'} - y_j) \prod_{i'<i} (x_{i'} - x_i).$$

Par ailleurs,

$$\det(V) = \prod_{i'<i} (y_{i'} - y_i) \prod_{j'<j} (x_{j'} - x_j) \prod_{i,j} (x_i - y_j).$$

En outre, pour tout $i \in \{1,\dots,n\}$ et pour tout $j \in \{1,\dots,m\}$, on a

$$B(x_i) = b_m \prod_{j=1}^{m} (x_i - y_j) \quad \text{et} \quad A(y_j) = a_n \prod_{i=1}^{n} (y_j - x_i) = (-1)^n a_n \prod_{i=1}^{n} (x_i - y_j).$$

Par conséquent, lorsque les x_i et y_j sont deux à deux distincts, on a les égalités

$$\mathrm{Res}_{n,m}(A,B) = \det(S) = b_m^n (-1)^{mn} \prod_{j=1}^{m} A(y_j) = a_n^m \prod_{i=1}^{n} B(x_i),$$

d'où la proposition dans ce cas.

Pour démontrer que ces égalités restent vraies sans cette hypothèse, le plus simple est peut-être d'utiliser le raisonnement algébrique suivant. Plaçons-nous dans le cas où les coefficients de A et B sont des indéterminées : leur résultant apparaît comme un polynôme « universel » $R_{n,m}(a_0,\dots,a_n;b_0,\dots,b_m)$ en ces coefficients. Supposons maintenant que le corps de base soit celui des fractions rationnelles $K(a_n,x_1,\dots,x_n,y_1,\dots,y_m,b_m)$, de sorte que les coefficients dominants et les racines de A et B sont des indéterminées, distinctes par hypothèse. Les coefficients de A et B s'expriment alors en fonction des polynômes symétriques élémentaires :

$$A=\sum_{k=0}^{n}(-1)^{n-k}a_n\sigma_k(x_1,\dots,x_n)X^k,\quad B=\sum_{k=0}^{m}(-1)^{m-k}a_n\sigma_k(y_1,\dots,y_n)X^k.$$

Alors, substituant les coefficients de A et B dans le polynôme $R_{n,m}$, on obtient l'égalité

$$R_{n,m}(a_n\sigma_0(x),\dots,a_n\sigma_n(x);b_m\sigma_0(y),\dots,b_m\sigma_m(x))=\mathrm{Res}_{n,m}(A,B)=a_n^m b_m^n\prod_{i=1}^{n}\prod_{j=1}^{m}(x_i-y_j).$$

C'est une égalité entre polynômes, égalité dans laquelle il est loisible de spécifier des valeurs pour les x_i et les y_j. Le membre de gauche fournit le résultant des deux polynômes considérés et celui de droite l'expression voulue en fonction des racines. □

L'intérêt de ces formules est de permettre un calcul récursif des résultants. En effet, l'expression $\prod_{i=1}^{n}B(x_i)$ peut se calculer en remplaçant B par le reste B_1 de la division euclidienne de B par A. On obtient alors

$$\mathrm{Res}_{n,m}(A,B)=a_n^m\prod_{i=1}^{n}B(x_i)=a_n^m\prod_{i=1}^{n}B_1(x_i)=a_n^{m-m_1}\mathrm{Res}_{n,m_1}(A,B_1),$$

où m_1 désigne le degré de B_1. On peut alors remplacer échanger les rôles de A et B et continuer le calcul.

Le *discriminant* d'un polynôme A de degré n est défini par la formule

$$\mathrm{disc}(A)=(-1)^{n(n-1)/2}\mathrm{Res}_{n,n-1}(P,P')/a_n,$$

où a_n est le coefficient dominant de A. Si $A=a_n\prod_{i=1}^{n}(X-x_i)$, on a ainsi

$$\begin{aligned}\mathrm{disc}(A)&=(-1)^{n(n-1)/2}a^{n-2}\prod_{i=1}^{n}P'(x_i)\\&=a^{2n-2}\prod_{i=1}^{n}\prod_{j<i}(x_i-x_j)^2.\end{aligned}$$

Exemple 2.1.3. — Soit a et b des éléments de K et soit n un entier. Calculons le discriminant du polynôme $A(X)=aX^2+bX+c$. Si $A(X)=a(X-u)(X-v)$,

$$\mathrm{disc}(A)=a^2(u-v)^2=a^2(u+v)^2-4a^2u^2v^2=b^2-4ac.$$

F. Polynômes cyclotomiques

On note Φ_n le polynôme unitaire dont les racines sont les racines primitives n-ièmes de l'unité dans **C**. Autrement dit,

$$\Phi_n(X) = \prod_{\substack{a=1 \\ (a,n)=1}}^{n} (X - \exp(\frac{2i\pi a}{n})).$$

On l'appelle le n-ième polynôme cyclotomique. On a par exemple $\Phi_1 = X - 1$, $\Phi_2 = X + 1$, $\Phi_3 = X^2 + X + 1$, $\Phi_4 = X^2 + 1$... Le degré du polynôme Φ_n est égal à $\varphi(n)$.

PROPOSITION 2.1.1. — *Pour tout entier $n \geqslant 1$, on a $X^n - 1 = \prod_{d|n} \Phi_d(X)$.*

Cette formule redonne le fait que $\sum_{d|n} \varphi(d) = n$.

Démonstration. — L'ordre d'une racine n-ième de l'unité α est un entier d qui divise n, et α est une racine primitive d-ième de l'unité, donc une racine de Φ_d. Inversement, si d divise n, une racine de Φ_d est une racine de $X^n - 1$. Les deux membres de l'égalité ont donc mêmes racines.

En outre, ces racines sont simples. C'est clair pour le polynôme $X^n - 1$. Par ailleurs, les racines des polynômes Φ_d sont simples et une racine de Φ_d, étant une racine primitive d-ième de l'unité, n'est pas une racine d'un autre polynôme cyclotomique. Ainsi, le polynôme au membre de droite est à racines simples.

Les deux membres sont des polynômes unitaires, ayant les mêmes racines complexes, avec les mêmes multiplicités. Ils sont donc égaux. □

PROPOSITION 2.1.2. — *Pour tout entier n, $\Phi_n(X)$ est un polynôme unitaire à coefficients entiers.*

Démonstration. — C'est vrai pour $n = 1, 2, 3, 4$. Supposons que ce soit vrai jusqu'au rang n (non inclus) et effectuons la division euclidienne du polynôme $X^n - 1$ par le produit des polynômes Φ_d, pour d divisant n et $d \neq n$. Comme les polynômes Φ_d sont à coefficients entiers et unitaires, cette division a un quotient Q et un reste R qui sont des polynômes à coefficients entiers. Or, $Q = \Phi_n$ et $R = 0$. En particulier, Φ_n est un polynôme unitaire à coefficients entiers. □

Comme Φ_n est à coefficients entiers, cela fait sens de considérer ses racines dans un corps arbitraire.

PROPOSITION 2.1.3. — *Soit K un corps et soit n un entier $\geqslant 1$. Si K est de caractéristique 0, posons $m = n$; si K est de caractéristique un nombre premier p, posons $m = n/p^r$ où $r = \mathrm{ord}_p(n)$, de sorte que p ne divise par m.*

Les racines de Φ_n dans K sont les racines primitives m-ièmes de l'unité contenues dans K : ce sont les éléments de K^* dont l'ordre est égal à m.

Démonstration. — Traitons d'abord le cas où n est premier à p ; il faut démontrer que les racines de Φ_n dans K sont simples, égales aux racines primitives n-ièmes de l'unité contenues dans K. La formule $X^n - 1 = \prod_{d|n} \Phi_d$ montre que les racines de Φ_n dans K sont simples, et sont des racines n-ièmes de l'unité, car c'est le cas pour le polynôme $X^n - 1$ qui n'a pas de racine commune avec sa dérivée nX^{n-1}. (C'est là qu'on utilise le fait que n n'est pas multiple de la caractéristique de K.)

Soit α une racine de Φ_n et soit d son ordre. C'est un diviseur de n. Par récurrence, α est racine de Φ_d, donc est une racine multiple de $X^n - 1$ à moins que l'on ait $d = n$. Autrement dit, α est une racine primitive n-ième.

Inversement, soit α une racine primitive n-ième. Par récurrence, α n'est racine d'aucun polynôme Φ_d, où $d < n$. La formule $X^n - 1 = \prod_{d|n} \Phi_d$ entraîne donc que α est racine de Φ_d.

Traitons maintenant le cas général ; alors K est un corps de caractéristique un nombre premier p si bien que l'on a l'égalité

$$X^n - 1 = X^{mp^r} - 1 = (X^m - 1)^{p^r}.$$

Par récurrence sur $r \geqslant 1$, on en déduit la formule

$$\Phi_n(X) = \Phi_m(X)^{p^{r-1}(p-1)}.$$

En particulier, Φ_m et Φ_n ont mêmes racines. □

Terminons ce paragraphe sur les polynômes cyclotomiques en établissant un théorème fondamental, dû encore une fois à Gauß.

THÉORÈME 2.1.4. — *Pour tout entier $n \geqslant 1$, le polynôme Φ_n est irréductible dans $\mathbf{Q}[X]$.*

Le cas particulier où n est un nombre premier se démontre généralement à l'aide du critère d'Eisenstein ; voir l'exercice 4.

Démonstration. — Comme il s'agit de polynômes unitaires à coefficients entiers, il nous suffit de démontrer qu'ils sont irréductibles dans $\mathbf{Z}[X]$.

Notons $\zeta = \exp(2i\pi/n)$ et P son polynôme minimal ; il est à coefficients entiers car c'est un facteur irréductible de Φ_n.

Nous devons prouver que $P = \Phi_n$. Par définition, les racines de Φ_n sont les ζ^a, où a parcourt l'ensemble des entiers de $[1, n]$ qui sont premiers à n. Nous allons démontrer que pour tout nombre premier p qui ne divise pas n, $P(\zeta^p) = 0$. Il en résultera, par récurrence sur le nombre de facteurs premiers de a, que pour tout entier a qui est premier à n, ζ^a est racine de P, et donc que $P = \Phi_n$.

Écrivons $X^n - 1 = PR$, avec $R \in \mathbf{Z}[X]$; dérivons cette relation et évaluons-la en ζ^p. On trouve

$$n\zeta^{p(n-1)} = P(\zeta^p)R'(\zeta^p)'_P(\zeta^p)Q(\zeta^p),$$

soit encore

$$n = \zeta^p P(\zeta^p) R'(\zeta^p) + \zeta^p P'(\zeta^p)\zeta^p).$$

Supposons par l'absurde que $P(\zeta^p) \neq 0$; alors $Q(\zeta^p) = 0$ et $\zeta^p P(\zeta^p) R'(\zeta^p) = n$.

Or, le polynôme P est à coefficients entiers si bien que $P(X)^p \equiv P(X^p) \pmod{p}$, autrement dit, $P(X)^p - P(X^p)$ est un polynôme à coefficients entiers dont les coefficients sont multiples de p. Cela entraîne l'existence d'un polynôme S à coefficients entiers tel que $n = pS(\zeta)$. Or, un argument de division euclidienne par P montre que tout élément de $\mathbf{Z}[\zeta]$ s'écrit de manière unique sous la forme $a_0 + a_1\zeta + \cdots + a_{d-1}\zeta^{d-1}$, où d est le degré de P et les a_i des entiers relatifs. Appliquons ceci à S ; on obtient les égalités absurdes $n = pa_j$, car n n'est pas multiple de p. □

Rappelons, pour mémoire, les énoncés de divisibilité utilisés dans la démonstration précédente.

LEMME 2.1.5. — *Soit P et Q des polynômes à coefficients entiers tels que Q divise P dans l'anneau $\mathbf{Q}[X]$ Si les coefficients de Q sont premiers entre eux* (on dit que Q est primitif), *alors Q divise P dans l'anneau $\mathbf{Z}[X]$.*

Démonstration. — Soit $R \in \mathbf{Q}[X]$ tel que $P = QR$. Nous dévons démontrer que R appartient à $\mathbf{Z}[X]$. Soit $a \in \mathbf{N}^*$ un dénominateur commun de ses coefficients, de sorte que $aR \in \mathbf{Z}[X]$, minimal. On a ainsi $aP = Q\,aR$. Soit p un facteur premier de a et considérons l'égalité précédente modulo p. On a ainsi $(Q \mod p)(aR \mod p) = 0$. Le polynôme $Q \mod p$ n'est pas nul car les coefficients de Q sont premiers entre eux. Comme l'anneau des polynômes à coefficients dans $\mathbf{Z}/p\mathbf{Z}$ est un anneau intègre, $aR \equiv 0 \pmod{p}$. Les coefficients de aR sont ainsi tous multiples de p, ce qui entraîne que $(a/p)R$ est à coefficients entiers, mais contredit l'hypothèse de minimalité faite sur a. Donc a n'a pas de facteur premier, c'est-à-dire $a = 1$ et $R \in \mathbf{Z}[X]$. □

COROLLAIRE 2.1.6. — *Soit $P \in \mathbf{Z}[X]$ un polynôme primitif qui est irréductible dans $\mathbf{Z}[X]$. Le polynôme P est irréductible dans $\mathbf{Q}[X]$.*

Démonstration. — Soit Q un facteur de P dans $\mathbf{Q}[X]$; nous devons prouver que $\deg Q = 0$ ou $\deg Q = \deg P$. Quitte à remplacer Q par aQ, où $a \in \mathbf{Z}$ est un dénominateur commun des coefficients de Q, on peut supposer que $Q \in \mathbf{Z}[X]$. Quitte à le remplacer alors par Q/b, où $b \in \mathbf{Z}$ est un facteur commun de ses coefficients, on peut supposer que Q est primitif. Alors, Q divise P dans $\mathbf{Z}[X]$. Comme P est irréductible dans cet anneau, ou bien Q, ou bien P/Q est inversible dans cet anneau. En particulier, $\deg Q = 0$ ou $\deg Q = \deg P$. □

COROLLAIRE 2.1.7. — *Soit $P \in \mathbf{Z}[X]$ un polynôme unitaire à coefficients entiers. Soit $Q \in \mathbf{Q}[X]$ un polynôme unitaire qui divise P. Alors, Q appartient à $\mathbf{Z}[X]$.*

Démonstration. — Comme dans la preuve du corollaire précédent, introduisons des entiers a et b tels que aQ/b soit un polynôme primitif. Il est loisible de choisir a et b premiers entre eux. D'après le lemme, aQ/b divise P dans $\mathbf{Z}[X]$; soit donc $R \in \mathbf{Z}[X]$ tel que $P = (aQ/b)R$. On a donc $bP = aQR$. Comparons les coefficients dominants, on voit que b est multiple de a, donc $a = 1$ puisque a et b sont premiers entre eux. Le polynôme Q/b apaprtient en particulier à $\mathbf{Z}[X]$ et il en est a fortiori de même du polynôme Q. □

Plus généralement, ce lemme et ses corollaires sont valables lorsqu'on remplace l'anneau $\mathbf{Z}$ par un anneau factoriel A et le corps $\mathbf{Q}$ par le corps des fractions de A. Les démonstrations sont les mêmes, à ceci près que lorsqu'on y choisit un élément minimal a de A, il s'agit d'un élément ayant le nombre minimal de facteurs premiers (comptés avec multiplicités).

Exercices

1) Soit $K \to L$ une extension finie de corps et soit $\alpha \in L$ tel que $L = K[\alpha]$ (on dit que c'est une extension *monogène*). Pour tout corps E tel que $K \subset E \subset L$, notons P_E le polynôme minimal de α sur E ; c'est un polynôme unitaire à coefficients dans E.

a) Démontrer que $[L : E] = \deg P_E$.

b) Si E et E' sont des corps tels que $K \subset E' \subset E \subset L$, démontrer que P_E divise $P_{E'}$.

c) Soit E un corps tel que $K \subset E \subset L$ et soit E' le sous-corps de E engendré par les coefficients de P_E. Démontrer que $P_{E'} = P_E$. Conclure que $E' = E$.

d) Démontrer qu'il y a un nombre fini de corps E tels que $K \subset E \subset L$.

2) Soit K un corps, p un nombre premier et a un élément de K. Montrer que le polynôme $X^p - a$ est réductible sur K si et seulement s'il a une racine dans K. (Si $X^p - a = P(X)Q(X)$, que peut valoir $P(0)$?)

3) [*Polynômes d'Artin-Schreier*] Soit p un nombre premier. Soit K un corps de caractéristique p et a un élément de K. On suppose que le polynôme $P = X^p - X - a$ n'a pas de racine dans K. Soit $K \subset L$ une extension de décomposition de P.

a) Si x est une racine de P dans L, montrer que les autres racines de P sont $x = x + 1$, $x + 2$, …, $x + p - 1$. En particulier, P est séparable et $L = K[x]$.

b) Montrer que P est irréductible dans $K[X]$. (Si Q est un diviseur de P de degré d, considérer le terme de degré $d - 1$ de Q.)

c) Démontrer que le polynôme $X^p - X - 1$ est irréductible dans $(\mathbf{Z}/p\mathbf{Z})[X]$.

4) Soit $P = a_0 + a_1 X + \cdots + a_n X^n$ un polynôme primitif à coefficients entiers. Soit p un nombre premier qui ne divise pas a_n.

a) On suppose que modulo p, le polynôme P est irréductible. Démontrer que P est irréductible dans $\mathbf{Z}[X]$ (et donc dans $\mathbf{Q}[X]$ d'après le théorème de Gauß).

b) On suppose que $a_0, \ldots, a_{n-1}$ sont multiples de p (de sorte que $P \equiv a_n X^n \pmod{p}$) mais que a_0 n'est pas multiple de p^2. Démontrer que P est irréductible dans $\mathbf{Z}[X]$ (*critère d'Eisenstein*). En déduire par exemple que le polynôme cyclotomique Φ_p est irréductible (modulo p, $\Phi_p \equiv (X-1)^{p-1}$; faire le changement de variables $Y = X - 1$).

c) Soit Q un polynôme à coefficients entiers dont le coefficient dominant n'est pas multiple de p ; on suppose que modulo p, on a la congruence $P \equiv Q^e$ et que Q est irréductible modulo p. Si le résultant $\mathrm{Res}(P,Q)$ n'est pas multiple de $p^{2\deg Q}$, démontrer que P est irréductible dans $\mathbf{Z}[X]$.

5) Soit P le polynôme $X^4 - X^2 - 2X + 3$.

a) Factoriser P modulo 2 et 3.

b) Démontrer que P est irréductible dans $\mathbf{Z}[X]$.

6) Soit $a_1, \ldots, a_n$ des entiers deux à deux distincts.

a) Démontrer que le polynôme $\prod_{i=1}^n (X - a_i) - 1$ est irréductible dans $\mathbf{Q}[X]$.

b) Si n est impair, démontrer que le polynôme $\prod_{i=1}^n (X - a_i) + 1$ est irréductible dans $\mathbf{Q}[X]$.

c) Donner un exemple de polynôme de la forme $\prod (X - a_i) + 1$ qui n'est pas irréductible.

7) Calculer le résultant des polynômes $X^7 + 1$ et $X^3 + X + 1$. (Utiliser des divisions euclidiennes successives.)

8) a) Si m et n sont des entiers positifs, calculer le résultant des polynômes $X^n - 1$ et $X^m + 1$

b) Donner un argument de nature arithmétique qui explique pourquoi, s'il n'est pas nul, c'est au signe près une puissance de 2.

9) Soit m et n des entiers positifs et soit d leur pgcd.

a) Démontrer que le pgcd des polynômes $X^m - 1$ et $X^n - 1$ est égal à $X^d - 1$. (Appliquer l'algorithme d'Euclide.)

b) Montrer que pour tout couple (a, b) d'entiers premiers entre eux, le pgcd de $a^m - b^m$ et $a^n - b^n$ est égal à $a^d - b^d$. (Même méthode.)

c) Soit A un anneau et soit a, b des éléments de A tels que $A = (a, b)$. Démontrer plus généralement que l'idéal engendré par $a^m - b^m$ et $a^n - b^n$ est l'idéal $a^d - b^d$. (Même méthode.)

10) Soit x un entier $\geqslant 2$. On dit qu'un facteur premier p de $x^n - 1$ est primitif s'il ne divise aucun des entiers $x^m - 1$ pour $1 \leqslant m < n$.

a) Si n et $x - 1$ sont premiers entre eux, montrer que $(x-1)$ et $(x^n - 1)/(x - 1)$ sont premiers entre eux.

b) Si p^k divise exactement $x - 1$, montrer que p^{k+1} divise exactement $x^p - 1$, sauf si $p = 2$ et $x = 1$.

c) Montrer qu'un nombre premier p est un diviseur primitif de $x^n - 1$ si et seulement si p divise $\Phi_n(x)$ et p ne divise pas n.

d) Pour tout nombre premier p, on pose $a_p = \mathrm{ord}_p(x^n - 1)$; soit P et Q les produits des p^{a_p}, où p parcourt respectivement les diviseurs primitifs de $x^n - 1$ et les autres, de sorte que $x^n - 1 = PQ$.

Monterr que P divise $\Phi_n(x)$.

e) On suppose $n \neq 2$. Si $\Phi_n(x) \neq P$, démontrer que $\Phi_n(x) = \ell P$, où ℓ est le plus grand facteur premier de n.

f) Soit n un entier ; notons s le nombre de facteurs premiers distincts de n. Démontrer l'inégalité

$$\left(\frac{x-1}{x}\right)^{2^{s-1}} x^{\varphi(n)} < \Phi_n(x) < x^{\varphi(n)} \left(\frac{x}{x-1}\right)^{2^{s-1}}.$$

g) Démontrer que pour tout entier n distinct de 2 et 6, $P > 1$. Autrement dit, $x^n - 1$ possède un facteur premier primitif.

h) Généraliser l'exercice en remplaçant $x^n - 1$ par $a^n - b^n$, où a et b sont des entiers premiers entre eux (*Birkhoff–Vandiver*, 1904). (Il n'y a alors pas d'exception pour $n = 6$, autres que $(a, b) = (2, 1)$.)

11) Soit n un nombre entier tel que $n > 1$. On rappelle que Φ_n désigne le n-ième polynôme cyclotomique.

a) Soit x un nombre entier et soit p un facteur premier de $\Phi_n(x)$ qui ne divise pas n. Montrer que $p \equiv 1 \pmod{n}$ et que p ne divise pas x.

b) Démontrer qu'il existe une infinité de nombres premiers congrus à 1 modulo n.

12) En utilisant le polynôme $x^2 + x + 1$, démontrer qu'il existe une infinité de nombres premiers congrus à 1 modulo 6.

13) Soit n un entier strictement positif, soit p un nombre premier impair qui ne divise pas n tel que $\Phi_n(2)$ soit multiple de p mais pas de p^2.

a) Montrer que l'ordre de 2 dans $\mathbf{F}_p^*$ divise n.

b) Remarquer que le polynôme $X^n - 1$ modulo p est séparable. En déduire que l'ordre de 2 dans $\mathbf{F}_p^*$ est égal à n.

c) Montrer que $2^n \not\equiv 1 \pmod{p^2}$ puis qu'il en est de même de 2^{p-1}.

§2.2. Corps finis

A. Existence et unicité

Soit K un corps fini ; notons q son cardinal.

L'homomorphisme naturel de $\mathbf{Z}$ dans K ne peut pas être injectif. Soit p le plus petit entier strictement positif tel que $p1_K = 0$. C'est un nombre premier : on a $p > 1$ car $1_K \neq 0_K$, et si $p = ab$, avec $1 < a, b < p$, alors $(a1_K)(b1_K) = 0$, donc $a1_K = 0$ ou $b1_K = 0$, ce qui contredit la minimalité de p. Ce nombre premier p est appelé la *caractéristique* du corps K. Par ailleurs, l'image de cet homomorphisme est un sous-anneau de K, isomorphe au corps $\mathbf{Z}/p\mathbf{Z}$. On peut ainsi considérer que K est un surcorps de $\mathbf{Z}/p\mathbf{Z}$. C'est en particulier un espace vectoriel sur ce sous-corps $\mathbf{Z}/p\mathbf{Z}$; sa dimension est finie car K, étant fini, admet évidemment une famille génératrice finie. Notons f cette dimension et soit $(x_1, \ldots, x_f)$ une base de K sur $(\mathbf{Z}/p\mathbf{Z})$. Comme tout élément de K s'écrit de manière unique sous la forme $\sum_{i=1}^{f} a_i x_i$, avec $a_i \in \mathbf{Z}/p\mathbf{Z}$, on voit que $q = p^f$. Nous avons ainsi démontré la première partie de la proposition suivante :

PROPOSITION 2.2.1. — *La caractéristique p d'un corps fini K est un nombre premier et son cardinal q en est une puissance. Son groupe multiplicatif est un groupe cyclique de cardinal $q-1$. Enfin, pour tout élément x de K, on a $x^q = x$, et même $x^{q-1} = 1$ si $x \neq 0$.*

Démonstration. — Le groupe multiplicatif de K est de cardinal $q-1$. Nous avons déjà démontré qu'il est cyclique. D'après le théorème de Lagrange, on a donc $x^{q-1} = 1$ pour tout $x \in K^*$. Si $x \in K$ n'est pas nul, on a donc $x^q = x\,x^{q-1} = x$; si $x = 0$, cette dernière relation est encore vérifiée, d'où la proposition. □

La détermination explicite d'un générateur du groupe multiplicatif K^* d'un corps fini K dépend de la façon dont on peut manipuler les éléments de K. Toutefois les remarques que nous avions faites pour le cas du corps fini à p éléments continuent de valoir.

Plus généralement :

PROPOSITION 2.2.2. — *Soit K et K' des corps finis et soit $K' \hookrightarrow K$ une extension. Les corps K et K' ont même caractéristique, disons p, et notons $q = p^f$ et $q' = p^{f'}$ les cardinaux de K et K'. Si n désigne le degré $[K : K']$ de l'extension $K' \hookrightarrow K$, on a $f = nf'$.*

Démonstration. — L'homomorphisme naturel de $\mathbf{Z}$ dans K' est le composé de celui de $\mathbf{Z}$ dans K avec l'homomorphisme injectif de K dans K' qui définit l'extension $K \hookrightarrow K'$. Par conséquent, les relations $a1_K = 0$ et $a1_{K'} = 0$ sont équivalentes ce qui entraîne que K et K' ont même caractéristique.

Considérons le corps K' comme un espace vectoriel sur K, sa dimension est finie car K' est fini, donc possède une famille génératrice finie. En considérant une base $(x_1, \ldots, x_n)$ de K' sur K, on voit que $q' = q^n$. Par suite, $p^{f'} = (p^f)^n = p^{fn}$, donc $f = f'n$. □

Comme dans tout corps de caractéristique un nombre premier p, l'application $F\colon x \mapsto x^p$ de K dans lui-même est un homomorphisme de corps. En effet, on a les égalités évidentes $F(0) = 0$, $F(1) = 1$, $F(xy) = F(x)F(y)$, tandis que la propriété d'additivité $F(x+y) = F(x) + F(y)$ résulte de ce que les coefficients binomiaux $\binom{p}{k}$ sont multiples de p si $1 \leqslant k \leqslant p-1$. Cet homomorphisme est appelé *homomorphisme de Frobenius*. Comme K est fini et que F est injectif, il est automatiquement surjectif. En outre, le composé $F^n = F \circ \cdots \circ F$ de n automorphismes de Frobenius est l'application $x \mapsto x^{p^n}$. Si $n = f$, il vient donc $F^f(x) = x^q = x$ pour tout $x \in K$, donc $F^f = \mathrm{id}_K$.

PROPOSITION 2.2.3. — *Si K est un corps fini de cardinal q, l'homomorphisme de Frobenius est un automorphisme de K. De plus, si $q = p^f$, alors $F^f = \mathrm{id}_K$.*

THÉORÈME 2.2.4 (Existence et unicité des corps finis de cardinal donné)

Soit Ω un corps de caractéristique p ; l'ensemble des racines du polynôme $X^q - X$ dans Ω est un sous-corps fini de Ω. Si le polynôme $X^q - X$ est scindé dans Ω, par exemple

si Ω est algébriquement clos, c'est un corps de cardinal q, et c'est l'unique sous-corps de Ω qui soit de cardinal q. Notons-le $\mathbf{F}_q$.

Alors, tout corps fini de cardinal q est isomorphe à ce corps $\mathbf{F}_q$. Plus précisément, l'image de tout homomorphisme de corps de K dans Ω est égale à $\mathbf{F}_q$.

Démonstration. — L'ensemble considéré est l'ensemble des points fixes de l'homomorphisme de corps $x \mapsto x^q$ de Ω dans lui-même. C'est donc un sous-corps de Ω. Comme un polynôme de degré q a au plus q racines, c'est un corps de cardinal inférieur ou égal à q. Mais le polynôme $X^q - X$, étant de dérivée $qX^{q-1} - 1 = -1$ dans $\Omega[X]$, n'a pas de racine multiple. Par conséquent, ce corps est de cardinal q dès que $X^q - X$ est scindé dans Ω.

Si K est un sous-corps de Ω de cardinal q, on a $x^q = x$ pour tout $x \in K$, donc $K \subset \mathbf{F}_q$ puis l'égalité car ces deux corps ont même cardinal.

Soit enfin K un corps de cardinal q. Comme tout élément de K vérifie $x^q = x$, il existe un homomorphisme de corps $\varphi : K \to \Omega$; l'image de φ est un sous-corps de Ω de cardinal q. On a donc $\varphi(K) = \mathbf{F}_q$. En particulier, K est isomorphe à $\mathbf{F}_q$. □

COROLLAIRE 2.2.5. — *Soit K un corps fini de cardinal q, soit n un entier $\geqslant 1$ et soit $K \hookrightarrow L$ une extension de décomposition du polynôme $X^{q^n} - X$. Alors, L est un corps fini de cardinal q^n.*

Démonstration. — Appliquons le théorème précédent avec $\Omega = L$. On obtient dans L un sous-corps L' de cardinal q^n, formé des solutions dans L de l'équation $x^{q^n} = x$. Ce sous-corps est manifestement une extension de décomposition du polynôme $X^{q^n} - X$, d'où $L = L'$. En particulier, $\mathrm{card}(L) = q^n$. □

COROLLAIRE 2.2.6. — *Soit K un corps fini. Pour tout entier $n \geqslant 1$, soit $\mathscr{P}_n$ l'ensemble des polynômes irréductibles unitaires de degré n à coefficients dans K.*

Alors, $\mathscr{P}_n$ n'est pas vide : il existe dans $K[X]$ des polynômes unitaires de tout degré. De plus, pour tout entier $n \geqslant 1$, on a la formule

$$X^{q^n} - X = \prod_{d|n} \prod_{P \in \mathscr{P}_n} P.$$

Démonstration. — Soit $K \subset L$ une extension de corps où L est un corps à q^n éléments ; le théorème précédent en affirme l'existence. Soit x un élément de L tel que $L = K[x]$; il en existe, il suffit par exemple de prendre pour x un générateur du groupe multiplicatif L^*. Le polynôme minimal P de x sur K est alors un polynôme irréductible unitaire à cofficients dans K ; son degré est égal à $[L : K]$, donc est égal à n. Comme de plus $x^{q^n} = x$, le polynôme P divise le polynôme $X^{q^n} - X$. Plus généralement, le même argument montre que P divise $X^{q^m} - X$ pour tout multiple m de n.

Dans l'autre sens, soit P un facteur irréductible du polynôme $X^{q^n} - X$ et soit d son degré. Soit L une extension de K obtenue par adjonction d'une racine x de P. Les inclusions $K \subset L \subset \mathbf{F}_{q^n}$ entraînent que d divise n. Autrement dit, le degré d'un facteur irréductible de $X^{q^n} - X$ divise n.

Par ailleurs, le polynôme $X^{q^n} - X$ n'a pas de facteur multiple dans $K[X]$ car il est premier avec son polynôme dérivé $q^n X^{q^n-1} - 1 = -1$.

Les polynômes irréductibles unitaires dans $K[X]$ dont le degré divise n sont exactement les facteurs irréductibles unitaires du polynôme $X^{q^n} - X$; comme ces polynômes sont deux à deux premiers entre eux, $X^{q^n} - X$ est égal à leur produit. □

B. Extensions de corps finis et théorie de Galois

Nous verrons plus loin qu'inversement, si K et K' sont des corps finis de cardinaux p^f et $p^{f'}$, où f' est multiple de f, alors il existe des homomorphismes de corps de K dans K'.

C. Facteurs irréductibles des polynômes cyclotomiques

PROPOSITION 2.2.1. — *Soit K un corps fini, soit q son cardinal. Soit N un entier naturel non nul et posons $m = N/p^{\mathrm{ord}_p(N)}$ le plus grand quotient de N qui n'est pas multiple de p. Les facteurs irréductibles de Φ_N dans l'anneau $K[X]$ sont tous de même degré, égal à l'ordre de q dans $(\mathbf{Z}/m\mathbf{Z})^*$.*

En particulier, pour tout entier $n \geqslant 1$, les facteurs irréductibles de Φ_{q^n-1} dans $K[X]$ sont de degré n.

Pour $n = 1$, cela entraîne que Φ_{q-1} est scindé dans $K[X]$. Chacune de ses racines est d'ailleurs un générateur du groupe multiplicatif K^*.

Démonstration. — Soit P un facteur irréductible de Φ_n dans $K[X]$ et soit $L = K[X]/(P)$ l'extension de K obtenue par adjonction d'une racine x de P. On a démontré que x est d'ordre m dans L^*. Si $d = \deg P$, alors $[L : K] = d$, donc L est un corps fini de cardinal q^d. En particulier, $x^{q^d-1} = 1$. Par conséquent, m divise $q^d - 1$, c'est-à-dire $q^d \equiv 1 \pmod{m}$. L'ordre de q dans $\mathbf{Z}/m\mathbf{Z}$ est donc un diviseur de d.

Inversement, si $q^e \equiv 1 \pmod{m}$, pour un entier e tel que $1 \leqslant e \leqslant m$, alors $x^{q^e} = x$ et x appartient au sous-corps K_e de L formé des éléments $t \in L$ tels que $F^e(t) = t$. Par suite, $K[x] \subset K_e$, donc $L = K_e$ puisque $L = K[x]$. Comme le cardinal de K_e est égal à q^e, on a $m = e$.

La dernière remarque résulte de ce que l'ordre de q dans $(\mathbf{Z}/(q^k - 1)\mathbf{Z})^*$ est égal à k. En effet, $q, q^2, \ldots, q^{n-1}$ sont tous compris entre 1 et $q^k - 1$, donc ne sont pas congrus à 1 modulo $q^k - 1$, tandis que $q^n \equiv 1 \pmod{q^k - 1}$. □

Pour construire explicitement un corps fini K à $q = p^n$ éléments, il suffit de trouver un polynôme irréductible P de degré n à coefficients dans $\mathbf{Z}/p\mathbf{Z}$ et de poser $K =$

$(\mathbf{Z}/p\mathbf{Z})[X]/(P)$. D'après la proposition précédente, on peut se contenter de trouver un facteur irréductible du polynôme cyclotomique Φ_{p^n-1}. D'un point de vue théorique, on retrouve l'existence d'un corps fini de cardinal p^n. D'un point de vue pratique, l'intérêt de cette construction est double : tout d'abord, dans la base $(1, x, \ldots, x^{n-1})$ de l'algèbre K sur $(\mathbf{Z}/p\mathbf{Z})$, la table de multiplication est assez simple ; ensuite, x est un générateur du groupe multiplicatif K^*. Cependant, on ne perdra pas de vue que la factorisation de polynômes, même à coefficients dans un corps fini, est un problème délicat.

D. Théorème de Wedderburn

THÉORÈME 2.2.1 (Wedderburn). — *Soit A un anneau à division (c'est-à-dire un anneau unitaire, non nul, non nécessairement commutatif dans lequel tout élément non nul a un inverse). Si A est fini, alors A est un corps commutatif.*

On résume parfois cet énoncé en disant que « tout corps fini est commutatif. » D'un point de vue pédagogique, c'est une source de confusion pour l'étudiant qui aurait perdu de vue que dans ses cours, les corps sont toujours supposés finis.

Ce théorème est important ; la démonstration qui suit fait de nouveau la part belle aux polynômes cyclotomiques.

Démonstration. — Notons Q le cardinal de A.

L'homomorphisme naturel de $\mathbf{Z}$ dans A n'est pas injectif ; son image est un sous-anneau de A, isomorphe à $\mathbf{Z}/p\mathbf{Z}$ où p est un nombre premier. Pour $x \in A$, notons A_x l'ensemble des éléments $y \in A$ tels que $xy = yx$ (le *commutant* de x). C'est un sous-anneau de A ; de plus, si $y \in A_x$ n'est pas nul, alors $y^{-1}x = xy^{-1}$, donc $y^{-1} \in A_x$. Cela démontre que A_x est un sous-anneau à division de A.

L'intersection des A_x (le *centre* de A) est un sous-anneau à division Z de A qui contient le sous-anneau $\mathbf{Z}/p\mathbf{Z}$. Nous devons montrer que $Z = A$. Notons q le cardinal de Z.

Considérons alors A comme un espace vectoriel à gauche sur Z ; on voit que le cardinal de A est une puissance de q, disons $Q = q^n$. De même, pour $x \in A$, le cardinal q_x de A_x est une puissance de q, disons $q_x = q^{m(x)}$. On a $m(x) = 1$ si x appartient au centre de A, et $m(x) > 1$ sinon.

Considérons aussi A comme un espace vectoriel à gauche sur l'anneau à division A_x ; ainsi, Q est une puissance de q_x et il existe un entier $n(x)$ tel que $n = n(x)m(x)$.

Faisons opérer A^* sur A par conjugaison. Le stabilisateur d'un élément $x \in A$ est l'ensemble A_x^* des éléments non-nuls de A_x. Pour que l'orbite d'un élément x soit un singleton, il faut et il suffit que x appartienne au centre Z de A. Soit alors $x_1, \ldots, x_e$ des représentants des orbites de A qui ne sont pas des singletons. L'orbite $\mathcal{O}(x_i)$ de x_i est un sous-groupe de A^* isomorphe au groupe quotient $A^*/A_{x_i}^*$. L'équation aux classes,

qui traduit que le cardinal de A est la somme des cardinaux des orbites, s'écrit donc

$$q^n = q + \sum_{i=1}^{r} \frac{q^n - 1}{q^{n(x_i)} - 1}.$$

Le numérateur $q^n - 1$ se décompose en produit des $\Phi_d(q)$, où d parcourt l'ensemble des diviseurs de n. De même, le dénominateur $q^{n(x_i)-1}$ est égal au produit des $\Phi_d(q)$ où d parcourt l'ensemble des diviseurs de $n(x_i)$. Comme $n = n(x_i)m(x_i)$ est un multiple strict de $n(x_i)$, le quotient $(q^n - 1)/(q^{n(x_i)} - 1)$ est multiple de $\Phi_n(q)$.

Par conséquent, $\Phi_n(q)$ divise $q^n - q$. Or, $\Phi_n(q)$ divise aussi $q^n - 1$; il divise donc $q - 1$. Par ailleurs, la définition

$$\Phi_n(q) = \prod_{\substack{a=1 \\ (a,n)=1}}^{n} \big(q - \exp(2i\pi a/n)\big)$$

entraîne que $\big|\Phi_n(q)\big| \geqslant (q-1)^{\varphi(n)}$, l'égalité étant stricte si $n > 1$. Par suite, $n = 1$ et A est commutatif. □

Exercices

14) a) Pour $d = 2, 3, 4, 5$, déterminer un polynôme irréductible P de degré d à coefficients dans $(\mathbf{Z}/2\mathbf{Z})$.

b) Écrire la table de multiplication du corps $K_d = (\mathbf{Z}/2\mathbf{Z})[X]/(P)$ dans la base $1, x, \ldots, x^{d-1}$, où x est la classe de X. Identifier en outre K_2 à un sous-corps de K_4.

c) Calculer l'ordre de x dans K_d^*.

d) Calculer un générateur du groupe multiplicatif K_d^*.

15) Reprendre l'exercice 15 en remplaçant le corps $\mathbf{Z}/2\mathbf{Z}$ par le corps $\mathbf{Z}/p\mathbf{Z}$, où $p = 3$.

16) Reprendre l'exercice 15 en remplaçant le corps $\mathbf{Z}/2\mathbf{Z}$ par le corps $\mathbf{Z}/p\mathbf{Z}$, où $p = 5$.

17) Cet exercice explique comment extraire une racine p-ième dans un groupe cyclique. En particulier, cela fournit une méthode pour extraire une racine carrée d'un élément d'un corps fini, pourvu que cet élément soit un carré.

Soit G un groupe cyclique d'ordre N, noté multiplicativement. Soit p un diviseur de N ; on note s le plus grand entier tel que p^s divise N. On se donne aussi un élément ε d'ordre p^s dans G. Soit g un élément de G.

a) Montrer qu'il existe $h \in G$ tel que $h^p = g$ si et seulement si $g^{N/p} = 1$. On supposera cette hypothèse vérifiée dans la suite de l'exercice.

b) Soit m l'ordre de g. Si $(p, m) = 1$, trouver une racine p-ième de g de la forme g^n, avec $n \in \mathbf{N}$.

c) Soit g un élément de G dont l'ordre est une puissance de p, disons p^r avec $1 \leqslant r \leqslant s$. Montrer qu'il existe un unique entier $u \in \{0, \ldots, p-1\}$ tel que $g^{p^{r-1}} = \varepsilon^{up^{s-1}}$. En déduire que l'ordre de $g\varepsilon^{-up^{s-r}}$ divise p^{r-1}.

d) Soit K un corps fini et soit q son cardinal, supposé impair. On note s le plus petit entier strictement positif tel que 2^s divise $q - 1$ et on se donne un élément ε de K^* d'ordre 2^s. On pose aussi $m = (q-1)/2^s$. Expliquer pourquoi l'algorithme suivant extrait une racine carrée d'un

élément a de K^* qui est un carré. On pose $(x_0, y_0, z_0) = (a^m, a^{(m+1)/2}, \varepsilon)$. Puis, pour $1 \leqslant i \leqslant s-1$, on pose

$$(x_i, y_i, z_i) = \begin{cases} (x_{i-1}, y_{i-1}, z_{i-1}^2) & \text{si } (x_{i-1})^{2^{s-i-1}} = 1\,; \\ (x_{i-1} z_{i-1}^2, y_{i-1} z_{i-1}, z_{i-1}^2) & \text{sinon.} \end{cases}$$

Alors, y_{s-1} est une racine carrée de a. (*Algorithme de Tonelli pour le calcul des racines carrées.*)

Que se passe-t-il si a n'est pas un carré dans K ?

e) Calculer une racine carrée de 3 modulo 13 puis une racine de 2 modulo 113.

18) Soit K un corps fini et soit q son cardinal, supposé impair.

a) Pour $a \in K$, déterminer le nombre de solutions (s, t) de l'équation $s^2 - t^2 = a$. (Distinguer suivant que a est un carré ou non dans K.)

b) Soit a un élément de K^* qui est un carré. Montrer qu'il existe un élément t de K tel que le polynôme $X^2 - tX + a$ soit irréductible. Quelle est la probabilité que t, choisi au hasard, convienne ?

c) Soit alors L l'extension quadratique $K[X]/(X^2 - tX + a)$. Montrer que la classe x de X dans L vérifie $x^{q+1} = a$ puis que $\xi = x^{(q+1)/2}$ est une des deux racines de a dans K. (*Algorithme de Cipolla pour le calcul des racines carrées.*)

19) Soit K un corps fini et soit a un élément de K^* qui est un carré. On note q le cardinal de K et on le suppose impair.

a) Soit x un élément de K. Montrer qu'il existe au moins $(q-1)/2$ éléments b de $K \setminus \{x\}$ tel que $(x+b)/(x-b)$ ne soit pas un carré dans K.

b) Démontrer qu'avec probabilité supérieure ou égale à $1/2$, le reste de la division euclidienne du polynôme $(X+b)^{(q-1)/2} - 1$ par $X^2 - a$ est un facteur de degré 1 de $X^2 - a$. L'unique racine de ce facteur fournit ainsi une racine carrée de a dans K^* (adaptation de l'*algorithme de Berlekamp*).

c) Déterminer une racine carrée de 7 modulo 37.

§2.3. La loi de réciprocité quadratique

A. Symboles de Legendre et de Jacobi

Soit p un nombre premier impair. Pour tout entier a, on définit le *symbole de Legendre* $\left(\frac{a}{p}\right)$ par la formule

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{si } p \text{ divise } a\,; \\ 1 & \text{si } a \text{ est un carré modulo } p\,; \\ -1 & \text{sinon.} \end{cases}$$

THÉORÈME 2.3.1 (Euler). — *Si p est un nombre premier impair et a un nombre entier, on a l'égalité*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

Démonstration. — L'égalité est évidente si p divise a et résulte sinon de ce que le groupe $(\mathbf{Z}/p\mathbf{Z})^*$ est cyclique. On remarque d'abord que $a^{p-1} \equiv 1 \pmod{p}$, si bien que $a^{(p-1)/2} \equiv \pm 1$. Supposons alors que a soit un carré modulo p ; il existe alors $b \in \mathbf{Z}$ tel que $a \equiv b^2 \pmod{p}$ et $a^{(p-1)/2} \equiv b^{p-1} \equiv 1 \pmod{p}$ puisque p ne divise pas b. L'image Le morphisme de groupes $x \mapsto x^2$ a pour noyau $\{\pm 1\}$; son image est donc de cardinal $(p-1)/2$ et est formée des carrés de $(\mathbf{Z}/p\mathbf{Z})^*$ qui sont, comme on vient de le voir, racines du polynôme $X^{(p-1)/2} - 1$. Comme ce polynôme a au plus $(p-1)/2$ racines, ses racines sont exactement les carrés et $a^{(p-1)/2} \equiv -1 \pmod{p}$ si a n'est pas un carré modulo p. □

COROLLAIRE 2.3.2. — *Soit p un nombre premier impair ; pour que -1 soit un carré modulo p, il faut et il suffit que $p \equiv 1 \pmod{4}$.*

COROLLAIRE 2.3.3. — *Soit p un nombre premier impair et a un nombre entier qui n'est pas multiple de p. Le symbole $\left(\frac{a}{p}\right)$ est égal à la signature de la permutation « multiplication par a » dans $(\mathbf{Z}/p\mathbf{Z})^*$.*

Démonstration. — L'ordre multiplicatif de a divise $p-1$; notons-le $(p-1)/d$. Pour que a soit un carré modulo p, il faut et il suffit que d soit pair. Alors, la permutation « multiplication par a » dans $(\mathbf{Z}/p\mathbf{Z})^*$ possède d orbites, toutes de cardinal $(p-1)/d$. La signature de cette permutation est donc égale à $(-1)^{p-1-d} = (-1)^d$, ce qu'il fallait démontrer. □

La loi de réciprocité quadratique est l'énoncé suivant :

THÉORÈME 2.3.4. — *Soit p et q des nombres premiers impairs. Alors*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}.$$

Autrement dit, si p ou q est congru à 1 modulo 4, p est un carré modulo q si et seulement si q est un carré modulo p. Sinon, si p et q sont tous deux congrus à -1 modulo 4, p est un carré modulo q si et seulement si q n'est pas un carré modulo p. Ce théorème, entrevu par Euler et Legendre et dont Gauß a donné la première démonstration, relie de manière surprenante des propriétés arithmétiques modulo deux nombres premiers. Elle possède plus de 200 démonstrations, Gauß en ayant lui-même proposé huit ! Nous nous contenterons d'en expliquer cinq...

Il nous faut auparavant énoncer et démontrer la « loi complémentaire » calculant le symbole $\left(\frac{2}{p}\right)$.

PROPOSITION 2.3.5. — *Soit p un nombre premier impair. On a*

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & si\ p \equiv \pm 1 \pmod 8\ ; \\ -1 & si\ p \equiv \pm 3 \pmod 8. \end{cases}$$

Démonstration. — Soit ζ une racine primitive 8-ième de l'unité dans une extension K du corps $\mathbf{Z}/p\mathbf{Z}$. Elle vérifie $\zeta^4 = -1$ si bien que $(\zeta + \zeta^{-1})^2 = 2$. (Ce calcul n'est pas magique : dans $\mathbf{C}$, on peut prendre $\zeta = \exp(2i\pi/8) = (1+i)/\sqrt{2}$ et $\sqrt{2} = 2\cos(\pi/4) = \zeta + \zeta^{-1}$.) Posons $\alpha = \zeta + \zeta^{-1}$. Comme $\alpha^2 = 2 \neq 0$, on a $\alpha \neq 0$. La question est de savoir si α est, ou non, un élément de $\mathbf{Z}/p\mathbf{Z}$. Pour cela, on calcule α^p et l'on trouve $\alpha^p = \zeta^p + \zeta^{-p}$ puisque $x \mapsto x^p$ est un homomorphisme de corps en caractéristique p. Par suite, si $p \equiv \pm 1 \pmod 8$, on constate que $\alpha^p = \alpha$. En revanche, si $p \equiv \pm 3 \pmod 8$, on a $\alpha^p = \zeta^3 + \zeta^{-3} = -\alpha$ qui est distinct de α puisque $\alpha \neq 0$. □

Le symbole de Jacobi étend la définition du symbole de Legendre $\left(\frac{a}{n}\right)$ en permettant que n ne soit pas un nombre premier.

Soit n un entier naturel impair et soit soit $n = \prod p_i^{m_i}$ sa décomposition de n en facteurs premiers ; le symbole de Jacobi $\left(\frac{a}{n}\right)$ est donné par la formule

$$\left(\frac{a}{n}\right) = \prod_i \left(\frac{a}{p_i}\right)^{m_i}.$$

Il coïncide donc avec le symbole de Legendre lorsque n est un nombre premier. L'intérêt de son introduction réside en ce que la loi de réciprocité quadratique continue à être valide.

PROPOSITION 2.3.6 (Loi de réciprocité pour le symbole de Jacobi)

a) Soit n un entier naturel impair ; on a

$$\left(\frac{-1}{n}\right) = (-1)^{(n-1)/2} = \begin{cases} 1 & si\ n \equiv 1 \pmod 4 \\ -1 & si\ n \equiv 3 \pmod 4. \end{cases}$$

b) Soit n un entier naturel impair ; on a

$$\left(\frac{2}{n}\right) = \begin{cases} 1 & si\ n \equiv \pm 1 \pmod 8\ ; \\ -1 & si\ n \equiv \pm 3 \pmod 8. \end{cases}$$

c) Soit m et n des entiers naturels impairs, premiers entre eux. On a l'égalité

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{(m-1)(n-1)/4}.$$

Démonstration. — a) Notons $n = \prod p_i^{n_i}$ la décomposition de n en facteurs premiers. Par définition,

$$\left(\frac{-1}{n}\right) = \prod_i \left(\frac{-1}{p_i}\right)^{n_i}.$$

Les facteurs qui ne valent pas 1 valent -1 et sont ceux pour lesquels $p_i \equiv 3 \pmod 4$ et n_i est impair ; autrement dit, $\left(\frac{-1}{n}\right) = (-1)^\nu$ où ν est le nombre d'indices i tels que $p_i \equiv 3 \pmod 4$ et n_i est impair. On a donc $n \equiv (-1)^\nu \pmod 4$ et $(n-1)/2$ est pair (respectivement impair) quand $\left(\frac{-1}{n}\right) = 1$ (respectivement -1).

b) Raisonnons de manière analogue. Soit $n = \prod p_i^{n_i}$ la décomposition de n en facteurs premiers et soit ν le nombre d'indices i tels que $p_i \equiv \pm 3 \pmod 8$. On a de même $\left(\frac{2}{n}\right) = (-1)^\nu$. D'autre part, comme $3^2 = 9 \equiv 1 \pmod 8$, $n \equiv \pm 1 \pmod 8$ si ν est pair et $n \equiv \pm 3 \pmod 8$ sinon. La formule en découle.

c) Notons $m = \prod p_i^{m_i}$, $n = \prod q_j^{n_j}$ les décompositions de m et n en facteurs premiers distincts. Par définition du symbole de Jacobi et multiplicativité du symbole de Legendre, on a

$$\left(\frac{m}{n}\right) = \prod_j \left(\frac{m}{q_j}\right)^{n_j} = \prod_{i,j} \left(\frac{p_i}{q_j}\right)^{m_i n_j}.$$

De même,

$$\left(\frac{n}{m}\right) = \prod_i \left(\frac{n}{p_i}\right)^{m_i} = \prod_{i,j} \left(\frac{q_j}{p_i}\right)^{m_i n_j}.$$

Par conséquent,

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = \prod_{i,j} \left(\left(\frac{p_i}{q_j}\right)\left(\frac{q_j}{p_i}\right)\right)^{m_i n_j}.$$

C'est un produit de facteurs égaux à ± 1. Les seuls facteurs du produit qui ne valent pas 1 sont ceux pour lesquels $p_i \equiv q_j \equiv 3 \pmod 4$ et où n_i et m_j sont impairs. Notons μ le nombre d'indices i tels que $p_i \equiv 3 \pmod 4$ et $m_i \equiv 1$ et définissons ν de manière analogue. on a ainsi

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right) = (-1)^{\mu\nu}.$$

Cette expression vaut -1 si μ et ν sont impairs, et vaut 1 sinon.

Enfin, $(m-1)(n-1)/4$ est pair si $(m-1)/2$ ou $(n-1)/2$ l'est, c'est-à-dire si m ou n est congru à 1 modulo 4 ; sinon $(m-1)(n-1)/2$ est impair. Comme $m \equiv (-1)^\mu \pmod 4$ et $n \equiv (-1)^\nu \pmod 4$, $(-1)^{(m-1)(n-1)/4}$ vaut -1 si μ et ν sont impairs, et vaut 1 sinon. Nous avons ainsi démontré la loi de réciprocité pour le symbole de Jacobi. □

Cette formule permet de calculer les symboles de Jacobi (et donc aussi ceux de Legendre) par récurrence, via une variante de l'algorithme d'Euclide. Donnons un exemple en calculant $\left(\frac{42}{157}\right)$. On a

$$\left(\frac{42}{157}\right) = \left(\frac{2}{157}\right)\left(\frac{21}{157}\right) = (-1)\left(\frac{157}{21}\right)$$

en utilisant la seconde loi complémentaire pour 157 (qui est congru à -3 modulo 8) et la loi de réciprocité quadratique (en notant que $21 \equiv 1 \pmod 4$). Par suite,

$$\left(\frac{42}{157}\right) = -\left(\frac{13}{21}\right) = -\left(\frac{21}{13}\right) = -\left(\frac{8}{13}\right) = -\left(\frac{2}{13}\right) = 1$$

en utilisant deux fois la loi de réciprocité quadratique ainsi que la seconde loi complémentaire (en observant $13 \equiv -3 \pmod 8$).

Comme 157 est un nombre premier (il n'est divisible par aucun des nombres premiers $2, 3, 5, 7, 11$ et $13^2 > 157$), nous avons en fait calculé un symbole de Legendre. En particulier, 42 est un carré dans $\mathbf{Z}/157\mathbf{Z}$.

Il est en revanche essentiel de remarquer que le symbole de Jacobi $\left(\frac{a}{n}\right)$ est loi de déterminer si a est un carré modulo n. Par exemple $\left(\frac{-1}{21}\right) = 1$ alors que dans l'anneau $\mathbf{Z}/21\mathbf{Z}$ qui est isomorphe à $(\mathbf{Z}/3\mathbf{Z}) \times (\mathbf{Z}/7\mathbf{Z})$, -1 n'est pas un carré (il ne l'est dans aucun des deux facteurs). En fait, si n n'est pas un nombre premier, la moitié des entiers a modulo n pour lesquels $\left(\frac{a}{n}\right) = 1$ ne sont pas des carrés modulo n. Cette remarque est à l'origine du test de primalité de Solovay et Strassen.

B. Trois démonstrations de la loi de réciprocité quadratique

Première démonstration. — Nous avons vu que $\left(\frac{p}{q}\right)$ est la signature de la multiplication par p dans $(\mathbf{Z}/q\mathbf{Z})^*$. La multiplication par p dans $\mathbf{Z}/q\mathbf{Z}$: a une orbite et un élément de plus, donc a même signature. Remplaçons maintenant le groupe $\mathbf{Z}/q\mathbf{Z}$ par celui, isomorphe, des racines q-ièmes de l'unité dans une extension K du corps $\mathbf{Z}/p\mathbf{Z}$. La multiplication par p dans $\mathbf{Z}/q\mathbf{Z}$ est transformée en l'élévation à la puissance p dans ce groupe G ; notons φ cette dernière permutation. La signature d'une permutation est égale à $(-1)^r$, où r est le nombre d'inversions, une fois choisi un ordre sur les éléments permutés. Fixons donc un tel ordre $\prec$. Alors,

$$\left(\frac{p}{q}\right) = \prod_{x<y} \begin{cases} 1 & \text{si } \varphi(x) \prec \varphi(y)\,; \\ -1 & \text{si } \varphi(y) \prec \varphi(x). \end{cases}$$

Par suite, on a l'égalité dans $\mathbf{Z}/p\mathbf{Z}$:

$$\left(\frac{p}{q}\right) = \prod_{x\prec y} \frac{\varphi(x)-\varphi(y)}{x-y},$$

puisque au numérateur et au dénominateur apparaissent, au signe près, toutes les différences $x-y$ avec $x \prec y$, avec autant de signes $-$ que d'inversions. Par conséquent, on a les égalités dans $\mathbf{Z}/p\mathbf{Z}$:

$$\left(\frac{p}{q}\right) \equiv \prod_{x\prec y} \frac{x^p-y^p}{x-y} \equiv \prod_{x\prec y} (x-y)^{p-1} \equiv \left(\prod_{x\neq y} (x-y)(-1)^{q(q-1)/2}\right)^{(p-1)/2} \equiv D^{(p-1)/2},$$

où D est le discriminant du polynôme $Q(X) = X^q - 1$ dans $\mathbf{Z}/p\mathbf{Z}$. La loi de réciprocité quadratique est presque démontrée puisqu'on a réussi à passer d'une condition sur p modulo q (la signature de la multiplication par p dans $\mathbf{Z}/q\mathbf{Z}$) à une condition modulo p, à savoir si D est ou non un carré modulo p.

Il reste à calculer ce discriminant [1]. On a $Q(X) = X^q - 1 = \prod_{y\in G}(X-y)$. Par suite, si x est un élément de G, $\prod_{y\neq x}(x-y) = Q'(y) = qy^{q-1} = q/y$. Finalement,

$$D = (-1)^{q(q-1)/2} \prod_{x\neq y} (x-y) = (-1)^{q(q-1)/2} q^q / \prod_{y\in G} y = (-1)^{q-1)/2} q^q,$$

étant donné que q est impair et que $\prod_{y\in G} y = 1$. On en conclut que

$$\left(\frac{p}{q}\right) \equiv (-1)^{(p-1)(q-1)/4} q^{q(p-1)/2} \equiv (-1)^{(p-1)(q-1)/4} q^{(p-1)/2} \equiv (-1)^{(p-1)(q-1)/4} \left(\frac{q}{p}\right).$$

Il s'agit d'une congruence modulo p entre deux nombres entiers égaux à ± 1. C'est donc en fait une égalité. □

Deuxième démonstration. — Pour tout entier n impair, notons U_n le n-ième polynôme de Tchebychev de seconde espèce caractérisée par la relation

$$U_n(2\cos\theta) = \frac{\sin(n\theta/2)}{\sin(\theta/2)},$$

soit encore

$$U_n(X+\frac{1}{X})X^{(n-1)/2} = \frac{X^n-1}{X-1},$$

qu'on récrit

$$X^{(n-1)/2}(X-1)U_n(X+\frac{1}{X}) = X^n - 1.$$

On a $U_1 = 1$, $U_3 = X+1$ et la relation de récurrence $U_{n+1} = 2XU_n - U_{n-1}$. Cette relation montre que U_n est, pour tout entier impair n, un polynôme unitaire de degré $(n-1)/2$ à coefficients entiers.

1. Ça serait mieux de l'avoir fait avant...

Cette démonstration de la loi de réciprocité quadratique va résulter du calcul du résultant des polynômes U_p et U_q ainsi que de la formule $\mathrm{Res}(P,Q) = \mathrm{Res}(Q,P)(-1)^{\deg P \deg Q}$ si P et Q sont des polynômes.

Tout d'abord, $\mathrm{Res}(U_p, U_q)$, étant le résultant de deux polynômes à coefficients entiers, est un nombre entier.

Si ℓ est un nombre premier, $\mathrm{Res}(U_p, U_q) \pmod{\ell}$ ne s'annule que si U_p et U_q, considérés comme polynômes à coefficients dans $(\mathbf{Z}/\ell\mathbf{Z})$, ont une racine commune (dans une clôture algébrique). Supposons pour commencer que $\ell \neq p$ et $\ell \neq q$. Soit α une racine de U_p ; on peut l'écrire sous la forme $x + 1/x$, et l'on a alors $1 + x + \cdots + x^{p-1} = 0$, si bien que x est racine primitive p-ième de l'unité. De même, une racine β de U_q est de la forme $y + 1/y$ où $1 + y + \cdots + y^{q-1} = 0$ et y est racine primitive q-ième de l'unité. L'égalité $\alpha = \beta$ entraînerait que y est racine du polynôme $T^2 - \alpha T + 1$, donc est égale à x ou $1/x$, donc est d'ordre p ce qui est absurde. Par conséquent, $\mathrm{Res}(U_p, U_q)$ n'est multiple d'aucun nombre premier ℓ distinct de p et q.

Supposons maintenant $\ell = p$. Une racine α de U_p dans une extension de $\mathbf{Z}/p\mathbf{Z}$ est de la forme $x + 1/x$ avec $x^p = 1$, donc $x = 1$ et $\alpha = 2$ Si α était racine de U_q, on pourrait l'écrire $y + 1/y$ où y serait une racine primitive q-ième de l'unité. En résolvant l'équation, on trouve $y = 1$, ce qui est absurde. Autrement dit, U_p et U_q n'ont pas de racine commune dans une extension de $\mathbf{Z}/p\mathbf{Z}$ et p ne divise pas $\mathrm{Res}(U_p, U_q)$.

On raisonne de même lorsque $\ell = q$ et q ne divise pas non plus $\mathrm{Res}(U_p, U_q)$.

En conclusion, $\mathrm{Res}(U_p, U_q)$ est un entier qui n'est divisible par aucun nombre premier. Il est donc égal à ± 1. Nous allons déterminer le signe en calculant $\mathrm{Res}(U_p, U_q)$ modulo p.

Si $\alpha_1, \ldots, \alpha_{(p-1)/2}$ et $\beta_1, \ldots, \beta_{(q-1)/2}$ sont les racines de U_p et U_q respectivement, on a

$$\mathrm{Res}(U_p, U_q) = \prod_{i,j} (\alpha_i - \beta_j).$$

Calculons ce résultant dans le corps $\mathbf{Z}/p\mathbf{Z}$; on a vu que la seule racine de U_p est $\alpha = 2$. Par suite, modulo p,

$$\mathrm{Res}(U_p, U_q) \equiv \prod_j (2 - \beta_j)^{(p-1)/2} \equiv U_q(2)^{(p-1)/2} \equiv q^{(p-1)/2} \pmod{p}.$$

Ainsi, $\mathrm{Res}(U_p, U_q)$ est un entier égal à ± 1 et congru à $\left(\frac{q}{p}\right)$; on a donc $\mathrm{Res}(U_p, U_q) = \left(\frac{q}{p}\right)$.

De même, $\mathrm{Res}(U_q, U_p) = \left(\frac{q}{p}\right)$. La formule $\mathrm{Res}(U_p, U_q) = (-1)^{(p-1)(q-1)/2} \mathrm{Res}(U_q, U_p)$, évidente sur la définition du résultant comme produit de différences des racines de U_p et de U_q entraîne alors l'égalité $\left(\frac{q}{p}\right)\left(\frac{p}{q}\right) = (-1)^{(p-1)(q-1)/2}$. □

Avant de donner la troisième démonstration, rappelons le théorème de Wilson et une variante :

LEMME 2.3.1. — *Soit p un nombre premier impair ; on a les congruences*

$$(p-1)! \equiv -1 \pmod{p} \quad \textit{et} \quad \left(\frac{p-1}{2}\right)!^2 \equiv -(-1)^{(p-1)/2} \pmod{p}.$$

Démonstration. — La première congruence se démontre en regroupant chaque élément a de $\{1,\dots,p-1\}$ avec son inverse modulo p dont il est distinct si et seulement si $a \neq \pm 1$. La seconde en découle, compte tenu des congruences modulo p :

$$(p-1)! \equiv (1\cdot 2\dots\frac{p-1}{2})(\frac{p+1}{2}\dots(p-1)) \equiv \left(\frac{p-1}{2}\right)!\left((-\frac{p-1}{2})\dots(-1)\right) \equiv \left(\frac{p-1}{2}\right)!^2(-1)^{(p-1)/2}.$$

□

Troisième démonstration (Rousseau, 1991). — Soit G le groupe abélien fini quotient du groupe $G_1 = (\mathbf{Z}/p\mathbf{Z})^* \times (\mathbf{Z}/q\mathbf{Z})^*$ par son sous-groupe d'ordre 2, $\{(1,1),(-1,-1)\}$. Nous allons calculer de deux façons différentes le produit Π des éléments de G.

Tout d'abord, les couples (x,y) avec $x \in (\mathbf{Z}/p\mathbf{Z})^*$ et $y \in \{1,\dots,\frac{q-1}{2}\}$ forment un système de représentants dans G_1 des éléments de G. Par suite, Π est l'image dans G de l'élément

$$\begin{aligned}\Pi_1 &= \prod_{x\in(\mathbf{Z}/p\mathbf{Z})^*} \prod_{y=1}^{(q-1)/2} (x,y) \\ &= \left(((p-1)!)^{(q-1)/2}, (\frac{q-1}{2}!)^{p-1}\right) \\ &= \left((-1)^{(q-1)/2}, (-1)^{(p-1)/2}(-1)^{(q-1)(p-1)/4}\right),\end{aligned}$$

grâce au lemme précédent.

D'autre part, lorsque z parcourt l'ensemble des nombres entiers de $\{1,\dots,pq\}$ qui ne sont divisibles ni par p ni par q, les couples ($z \mod p, z \mod q$) parcourent une fois et une seule les éléments de G_1. À l'entier $pq-z$ correspond le couple opposé ($-z \mod p, -z \mod q$). Lorsque z parcourt uniquement $\{1,\dots(pq-1)/2\}$, les couples ($z \mod p, z \mod q$) forment un système de représentants dans G du groupe G_1. Notons c le produit de ces entiers. En rajoutant les multiples de q dans l'intervalle $[1,(pq-1)/2]$, on obtient

$$c = \prod_{\substack{z=1\\(z,p)=1}}^{(pq-1)/2} z \,/\, q\cdot 2q\dots\frac{p-1}{2}q.$$

Les entiers z qui apparaissent au numérateur s'écrivent d'une part $z = pu+v$ avec $0 \leqslant u < \frac{q-1}{2}$ et $1 \leqslant v < p$ et d'autre part $z = p\frac{q-1}{2} + v$ avec $1 \leqslant v \leqslant \frac{p-1}{2}$. Modulo p,

l'entier u n'intervient pas et l'on a donc

$$c \equiv (1\dots(p-1))^{(q-1)/2}(1\dots\frac{p-1}{2}) \,/\, q^{(p-1)/2}(1\dots\frac{p-1}{2}) \equiv \left(\frac{q}{p}\right)(-1)^{(q-1)/2} \pmod{p}.$$

De même,

$$c \equiv \left(\frac{p}{q}\right)(-1)^{(p-1)/2} \pmod{q},$$

si bien que

$$\Pi_2 = (c \mod p, c \mod q) = \left(\left(\frac{q}{p}\right)(-1)^{(q-1)/2}, \left(\frac{p}{q}\right)(-1)^{(p-1)/2}\right)$$

est aussi un représentant de Π dans G_1.

En comparant les expressions de Π_1 et de Π_2, on en déduit que

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4},$$

ce qu'il fallait démontrer. □

C. Sommes de Gauß

DÉFINITION 2.3.1. — *Pour tout entier naturel $n \geqslant 1$, on définit*

$$G_n = \sum_{k=1}^{n} \exp\left(2i\pi\frac{k^2}{n}\right)$$

THÉORÈME 2.3.2 (Gauß). — *Suivant la congruence vérifiée par n modulo 4, on a :*

$$G_n = \begin{cases} \sqrt{n} & si\ n \equiv 1 \pmod 4\ ; \\ 0 & si\ n \equiv 2 \pmod 4\ ; \\ i\sqrt{n} & si\ n \equiv 3 \pmod 4\ ; \\ (1+i)\sqrt{n} & si\ n \equiv 0 \pmod 4. \end{cases}$$

Démonstration. — Soit $f : \mathbf{R} \to \mathbf{C}$ la fonction définie par

$$f(x) = \sum_{k=0}^{n-1} \exp\left(2i\pi\frac{(x+k)^2}{n}\right).$$

Elle est de classe $\mathscr{C}^\infty$ et périodique, de période 1 ; on a bien sûr $G_n = f(0)$. Notons

$$f(x) = \sum c_m \exp(2i\pi m x)$$

sa série de Fourier. Par définition, on a

$$\begin{aligned}
c_m &= \int_0^1 \left(\sum_{k=0}^{n-1} \exp\big(2i\pi \frac{(k+x)^2}{n}\big) \right) \exp(-2im\pi x)\mathrm{d}x \\
&= \sum_{k=0}^{n-1} \int_0^1 \exp\big(\frac{2i\pi}{n}(x+(k-\frac{1}{2}mn)^2\big) dx \exp\big(2i\pi(mk-\frac{1}{4}m^2n)\big) \\
&= \exp(-2i\pi\frac{m^2n}{4}) \sum_{k=0}^{n-1} \int_{k-\frac{1}{2}mn}^{k+1-\frac{1}{2}mn} \exp\big(\frac{2i\pi}{n}u^2\big)\mathrm{d}u \\
&= \exp(-2i\pi\frac{m^2n}{4}) \int_{-\frac{1}{2}mn}^{n-\frac{1}{2}mn} \exp\big(\frac{2i\pi}{n}u^2\big)\mathrm{d}u \\
&= n\exp(-2i\pi\frac{m^2n}{4}) \int_{-\frac{1}{2}m}^{1-\frac{1}{2}mn} \exp\big(2i\pi nv^2\big)\mathrm{d}v.
\end{aligned}$$

Étant de classe $\mathscr{C}^\infty$, la fonction f, est somme de sa série de Fourier ; en particulier

$$G_n = f(0) = \sum_{m\in\mathbf{Z}} c_m.$$

On observe que $\exp(-2i\pi m^2 n/4)$ vaut 1 quand m est pair, et vaut $\exp(-2i\pi n/4)$ quand m est impair. Alors,

$$\begin{aligned}
G_n &= \sum_{m\,\text{pair}} c_m + \sum_{m\,\text{impair}} c_m = \sum_{m\in\mathbf{Z}} c_{2m} + \sum_{m\in\mathbf{Z}} c_{2m+1} \\
&= n\sum_{m\in\mathbf{Z}} \int_{-m}^{1-m} \exp\big(2i\pi nv^2\big)\mathrm{d}v + n\exp(-2i\pi n/4) \sum_{m\in\mathbf{Z}} \int_{-\frac{1}{2}-m}^{\frac{1}{2}-m} \exp\big(2i\pi nv^2\big)\mathrm{d}v.
\end{aligned}$$

Les deux séries de la dernière équation convergent et ont même somme, à savoir l'intégrale (généralisée)

$$\int_{-\infty}^{\infty} \exp(2i\pi nv^2)\mathrm{d}v,$$

intégrale qu'on appelle parfois *intégrale de Gauß* !

Soit en effet des nombres réels a et b de même signe tels que $a < b$, de sorte que 0 n'appartienne pas à l'intervalle $[a,b]$. Alors, une intégration par parties donne

$$\begin{aligned}
\int_a^b \exp(2i\pi nv^2)\mathrm{d}v &= \int_a^b \frac{1}{4i\pi nv} \exp(2i\pi nv^2) 4i\pi nv \mathrm{d}v \\
&= \left[\exp(2i\pi nv^2) \frac{1}{4i\pi nv} \right]_a^b + \int_a^b \frac{1}{4i\pi nv^2} \exp(2i\pi nv^2))\mathrm{d}v
\end{aligned}$$

d'où l'on déduit que $\int_a^b \exp(2i\pi nv^2)\mathrm{d}v$ tend vers 0 quand a et b tendent tous deux vers $+\infty$ ou $-\infty$. Cela vérifie le critère de Cauchy pour l'intégrale généralisée indiquée et entraîne aussi facilement qu'elle est égale aux deux séries qui nous intéressent.

Autrement dit,

$$G_n = n\big(1+\exp(-2i\pi n/4)\big)\int_{-\infty}^{\infty}\exp(2i\pi n v^2)\mathrm{d}v$$
$$= \sqrt{n}\big(1+\exp(-2i\pi n/4)\big)\int_{-\infty}^{\infty}\exp(2i\pi u^2)\mathrm{d}u.$$

Cette relation est vraie pour tout entier n, en particulier pour $n=1$ où elle fournit le calcul de l'intégrale de Gauß :

$$\int_{-\infty}^{\infty}\exp(2i\pi u^2)\mathrm{d}u = \frac{G_1}{1-i} = \frac{1+i}{2}.$$

Il suffit, pour conclure la démonstration du théorème, de disjoindre les cas :

— Si $n\equiv 0 \pmod 4$, $\exp(-2i\pi n/4)=1$ et $G_n=(1+i)\sqrt{n}$;
— si $n\equiv 1 \pmod 4$, $\exp(-2i\pi n/4)=-i$ et $G_n=\sqrt{n}$;
— si $n\equiv 2 \pmod 4$, $\exp(-2i\pi n/4)=-1$ et $G_n=0$;
— enfin, si $n\equiv 3 \pmod 4$, $\exp(-2i\pi n/4)=i$ et $G_n=i\sqrt{n}$. □

Nous allons déduire de cette formule une autre démonstration de la loi de réciprocité quadratique.

LEMME 2.3.3. — *On a l'égalité* $G_{pq}=G_pG_q\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)$.

Démonstration. — Observons que si x et y sont des entiers congrus modulo n, alors $\exp(2i\pi x^2/n)=\exp(2i\pi y^2)/n$. Par suite, pour définir la somme de Gauß, on peut utiliser tout ensemble d'entiers représentant chaque classe de congruence une fois et une seule. Lorsque x et y parcourent $\{1,\dots,p\}$ et $\{1,\dots,q\}$ respectivement, les entiers $qx+py$ sont deux à deux distincts modulo pq. Par suite,

$$G_{pq}=\sum_{x=1}^{p}\sum_{y=1}^{q}\exp\big(2i\pi\frac{(qx+py)^2}{pq}\big)=\sum_{x=1}^{p}\exp\big(2i\pi\frac{qx^2}{p}\big)\sum_{y=1}^{q}\exp\big(2i\pi\frac{py^2}{q}\big).$$

Lorsque x parcourt $\{1,\dots,p-1\}$, $qx^2 \mod p$ parcourt deux fois les carrés non nuls de $\mathbf{Z}/p\mathbf{Z}$ si x est un carré modulo p, et les non-carrés sinon. Dans le premier cas, on a $\sum_{x=1}^{p}\exp\big(2i\pi\frac{qx^2}{p}\big)=G_p$; dans le second, on a

$$\sum_{x=1}^{p}\exp\big(2i\pi\frac{qx^2}{p}\big) = 1+2\sum_{\substack{x=1\\ x\text{ non carré modulo }p}}^{p-1}\exp(2i\pi\frac{x^2}{p})$$
$$= 1-2-2\sum_{\substack{x=1\\ x\text{ carré modulo }p}}^{p-1}\exp(2i\pi\frac{x^2}{p})$$
$$= -G_p$$

compte tenu de l'égalité

$$\sum_{x=1}^{p-1} \exp(2i\pi\frac{x}{p}) = -1 + \sum_{x=1}^{p} \exp(2i\pi\frac{x}{p}) = -1.$$

Par conséquent,

$$\sum_{x=1}^{p} \exp\left(2i\pi\frac{qx^2}{p}\right) = \left(\frac{q}{p}\right) G_p.$$

En utilisant la relation correspondante obtenue en échangeant les rôles de p et q, on a donc le lemme annoncé. □

Le calcul de la somme de Gauß entraîne une démonstration immédiate de la loi de réciprocité quadratique.

Démonstration de la loi de réciprocité quadratique. — Comme p et q sont impairs, G_p, G_q et G_{pq} sont non nuls. On récrit alors la relation du lemme sous la forme

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = \frac{G_{pq}}{\sqrt{pq}} \frac{\sqrt{p}}{G_p} \frac{\sqrt{q}}{G_q}.$$

Si p est congru à 1 modulo 4, alors $pq \equiv q \pmod 4$ et $G_{pq}/\sqrt{pq} = G_q/\sqrt{q}$ et $G_p/\sqrt{p} = 1$; par conséquent, $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = 1$. De même, si q est congru à 1 modulo 4, alors $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = 1$. En revanche, si p et q sont tous deux congrus à 3 modulo 4, $G_p/\sqrt{p} = G_q/\sqrt{q} = i$ mais $pq \equiv 1 \pmod 4$ d'où $G_{pq}/\sqrt{pq} = 1$. Par suite, $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = 1/i^2 = -1$. □

On peut en fait définir une variante de la somme de Gauß dans tout corps qui contient une racine primitive n-ième de l'unité ; elle nous permettra d'ailleurs de donner une cinquième démonstration de la loi de réciprocité quadratique. Soit K un corps et soit $\zeta \in K$ une racine primitive p-ième de l'unité. On définit

$$G(\zeta) = \sum_{x=1}^{p} \zeta^{x^2} = 1 + 2 \sum_{\substack{x=1 \\ x \text{ carré modulo } p}}^{p-1} \zeta^x = \sum_{x \in (\mathbf{Z}/p\mathbf{Z})^*} \left(\frac{x}{p}\right) \zeta^x.$$

(Comme $\zeta^p = 1$, ζ^n ne dépend que de la classe de n modulo p ; si x désigne cette classe, on peut donc noter $\zeta^x = \zeta^n$.) Lorsque $K = \mathbf{C}$ et $\zeta = \exp(2i\pi/p)$, nous avons vu que $G(\zeta)^2 = \pm p$, le signe dépendant de la congruence de p modulo 4. C'est un fait général :

PROPOSITION 2.3.4. — *On a les relations :*

a) $G(\zeta)^2 = (-1)^{(p-1)/2} p$;

b) pour tout $a \in (\mathbf{Z}/p\mathbf{Z})^*$, $G(\zeta^a) = \left(\frac{a}{p}\right) G(\zeta)$.

Démonstration. — Commençons par établir la seconde relation. Soit en effet $a \in (\mathbf{Z}/p\mathbf{Z})^*$; alors,

$$G(\zeta^a) = \sum_{x \in (\mathbf{Z}/p\mathbf{Z})^*} \left(\frac{x}{p}\right) \zeta^{ax} = \sum_{y \in (\mathbf{Z}/p\mathbf{Z})^*} \left(\frac{y/a}{p}\right) \zeta^y = \left(\frac{a}{p}\right) G(\zeta).$$

Par conséquent,

$$G(\zeta)^2 = \left(\frac{-1}{p}\right) G(\zeta) G(\zeta^{-1}) = \left(\frac{-1}{p}\right) \sum_{x,y \in \mathbf{Z}/p\mathbf{Z}} \zeta^{x^2-y^2} = \left(\frac{-1}{p}\right) \sum_{x,y \in \mathbf{Z}/p\mathbf{Z}} \zeta^{(x-y)(x+y)}.$$

Lorsque (x, y) parcourt $(\mathbf{Z}/p\mathbf{Z})^2$, $(x-y, x+y)$ parcourt aussi $(\mathbf{Z}/p\mathbf{Z})^2$, car p est impair. Ainsi,

$$\begin{aligned} G(\zeta)^2 &= \left(\frac{-1}{p}\right) \sum_{u,v \in \mathbf{Z}/p\mathbf{Z}} \zeta^{uv} = \left(\frac{-1}{p}\right) \sum_{u \in \mathbf{Z}/p\mathbf{Z}} \left(\sum_{v \in \mathbf{Z}/p\mathbf{Z}} \zeta^{uv} \right) \\ &= \left(\frac{-1}{p}\right) \sum_{u \in \mathbf{Z}/p\mathbf{Z}} \begin{cases} p & \text{si } u = 0\,; \\ 0 & \text{sinon} \end{cases} \\ &= \left(\frac{-1}{p}\right) p. \end{aligned}$$

Cela conclut la démonstration de la proposition. □

Démonstration de la loi de réciprocité quadratique. — Supposons que K soit un corps de caractéristique q ; on a alors

$$G(\zeta)^q = \left(\sum_x \left(\frac{x}{p}\right) \zeta^x \right)^q = \sum_x \left(\frac{x}{p}\right) \zeta^{qx}$$

car l'élévation à la puissance q est un endomorphisme du corps K et que $\left(\frac{x}{p}\right)^q = \left(\frac{x}{p}\right)$ car q est impair. D'après la proposition, on a donc

$$G(\zeta)^q = G(\zeta^q) = \left(\frac{q}{p}\right) G(\zeta).$$

Alors,

$$\begin{aligned} G(\zeta)^q &= G(\zeta) \left(G(\zeta)^2\right)^{(q-1)/2} = G(\zeta) \left(\left(\frac{-1}{p}\right) p \right)^{(q-1)/2} \\ &= G(\zeta) \left(\frac{-1}{p}\right)^{(q-1)/2} p^{(q-1)/2} \\ &= G(\zeta) \left(\frac{-1}{p}\right)^{(q-1)/2} \left(\frac{p}{q}\right). \end{aligned}$$

Comme $G(\zeta)^2 = \left(\frac{-1}{p}\right) p$ n'est pas nul dans K, $G(\zeta) \neq 0$ et il vient

$$\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)\left(\frac{-1}{p}\right)^{(q-1)/2}.$$

Lorsque p est congru à 1 modulo 4, $\left(\frac{-1}{p}\right) = 1$ et il vient $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$; lorsque q est congru à 1 modulo 4, $(q-1)/2$ est pair et l'on a encore $\left(\frac{q}{p}\right) = \left(\frac{p}{q}\right)$; enfin, lorsque p et q sont tous deux congrus à 3 modulo 4, $\left(\frac{-1}{p}\right) = -1$, $(q-1)/2$ est impair si bien que $\left(\frac{q}{p}\right) = -\left(\frac{p}{q}\right)$. La démonstration est terminée. □

Exercices

20) a) Calculer le symbole de Legendre $\left(\frac{365}{1847}\right)$ à l'aide de la loi de réciprocité quadratique pour le symbole de Legendre. (*Exemple dû à Dirichlet.*)

b) Refaire le calcul en utilisant la loi de réciprocité pour le symbole de Jacobi.

21) a) À quelle condition -2 est-il un carré modulo un nombre premier p ? On explicitera le résultat sous la forme d'une liste de congruences modulo 8.

b) Même question en remplaçant -2 par 6, -5 et 11.

22) a) En utilisant l'interprétation du symbole de Legendre $\left(\frac{a}{p}\right)$ comme la signature d'une permutation, retrouver la valeur de $\left(\frac{2}{p}\right)$ suivant la congruence satisfaite par p modulo 8.

b) Raisonner de même et calculer les symboles $\left(\frac{a}{p}\right)$ si $a = -2$, $a = 3$ ou $a = -3$.

23) Soit p un nombre premier impair. Montrer que les deux nombres complexes $i^{(p-1)/2}(1+i)2^{(p-1)/2}$ et $1+i^p$ appartiennent à $\mathbf{Z}[i]$ (c'est-à-dire sont de parties réelle et imaginaire entières) et sont égaux modulo $p\mathbf{Z}[i]$ (c'est-à-dire que leurs parties réelles et imaginaires sont égales modulo p). En déduire une autre approche du calcul du symbole de Legendre $\left(\frac{2}{p}\right)$.

24) Soit p un nombre premier impair et soit a un nombre entier qui n'est pas multiple de p.

a) Soit μ le nombre des entiers $i \in \{1, \ldots, \frac{1}{2}(p-1)\}$ tels que le reste de la division euclidienne de ai par p soit $> \frac{1}{2}(p-1)$. Démontrer que $\left(\frac{a}{p}\right) = (-1)^\mu$. (*Lemme de Gauß.*)

b) Retrouver ainsi la valeur des symboles de Legendre $\left(\frac{2}{p}\right)$, $\left(\frac{3}{p}\right)$ et $\left(\frac{5}{p}\right)$.

25) Si p et q sont des nombres premiers impairs, on pose

$$S(p,q) = \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{iq}{p} \right\rfloor .$$

a) Démontrer que $S(p,q) + S(q,p) = (p-1)(q-1)/4$. (Observer que $S(p,q)$ et $S(q,p)$ représentent le nombre de points entiers de part et d'autre de la diagonale dans le rectangle de côtés p et q dont l'abscisse est inférieure à $(p-1)/2$ et l'ordonnée est inférieure à $(q-1)/2$.)

b) Démontrer que $S(q,p) \equiv \mu \pmod 2$, où, comme dans l'exercice 24, μ est le nombre des entiers $i \in \{1,\dots,\frac{1}{2}(p-1)\}$ tels que le reste de la division euclidienne de qi par p soit $> \frac{1}{2}(p-1)$.

c) En déduire une autre démonstration de la loi de réciprocité quadratique.

26) Nous reprenons dans cet exercice le calcul de résultants effectué au cours de la deuxième démonstration de la loi de réciprocité quadratique.

a) Si n est un entier impair, démontrer que les racines du polynôme U_n sont les $2\cos(k\pi/n)$, pour $1 \leqslant k \leqslant (n-1)/2$.

b) Démontrer que l'on a

$$\mathrm{Res}(U_n, U_m) = \prod_{k=1}^{(m-1)/2} \frac{\sin(k\pi m/n)}{\sin(k\pi/n)}.$$

c) À l'aide du lemme de Gauß (exercice 24), en déduire que $\mathrm{Res}(U_p, U_q) = \left(\frac{q}{p}\right)$ si p et q sont des nombres premiers impairs, distincts.

27) Soit p un nombre premier impair et soit ω une racine primitive p-ième de l'unité dans $\mathbf{C}$ et soit $S = (\omega^{ij})_{0 \leqslant i,j < p}$ la matrice de Vandermonde de $(1, \omega, \dots, \omega^{p-1})$.

a) Calculer S^2 et en déduire que $\det(S)^2 = (-1)^{p(p-1)/2} p^p$.

b) Montrer que $\det(S) = i^{p(p-1)/2} p^{p/2}$. (*Utiliser le fait que c'est un déterminant de Vandermonde.*)

c) Montrer que les valeurs propres de S sont $\pm\sqrt{p}, \pm i\sqrt{p}$. Soit r, s, t, u les multiplicités de $\sqrt{p}$, $-\sqrt{p}$, $i\sqrt{p}$, $-i\sqrt{p}$. Montrer que $\mathrm{Tr}(S) = \sqrt{p}$ si $p \equiv 1 \pmod 4$ et $\mathrm{Tr}(S) = i\sqrt{p}$ si $p \equiv 3 \pmod 4$.

d) Montrer que l'on a

$$\mathrm{Tr}(S) = \sum_{k=0}^{p-1} \left(\frac{k}{p}\right) \omega^k.$$

(*Détermination du « signe » de la somme de Gauss ;* démontration due à Schur.)

28) On va montrer que l'équation $2x^4 = y^4 - 17z^4 = 0$ n'a pas de solution non nulle $(x,y,z) \in \mathbf{Z}^3$ (*Lind,* 1940). On considère, par l'absurde, une telle solution (x,y,z).

a) Se ramener au cas où x, y et z sont premiers entre eux deux à deux.

b) Montrer que x n'est pas multiple de 17.

c) Montrer que x n'est pas un carré modulo 17.

d) Soit p un nombre premier impair qui divise x. (On a donc $p \neq 17$.) Montrer que 17 est un carré modulo p, puis que p est un carré modulo 17.

e) Montrer que 2 est un carré modulo 17.

f) Montrer que x est un carré modulo 17, d'où la contradiction recherchée.

§2.4. Polynômes irréductibles dans un corps fini

Le but de ce dernier paragraphe est d'exposer quelques méthodes pour factoriser des polynômes à coefficients dans un corps fini, ou pour décider de leur irréductibilité. En théorie, le problème peut sembler sans intérêt, car il n'y a qu'un nombre fini de facteurs possibles : il suffit en principe de les énumérer puis de tester la divisibilité. Toutefois, cette méthode se heurte à un obstacle de complexité important dès que le cardinal du corps fini n'est pas très petit et ne peut donc être appliquée en pratique. Il s'agit donc d'exposer des méthodes efficaces.

A. Critères d'irréductibilité

Soit $\mathbf{F}$ un corps fini de cardinal q et soit $f \in \mathbf{F}[x]$ un polynôme de degré n. Il s'agit pour l'instant de décider si f est irréductible ou non.

Le premier pas consiste à détecter les facteurs multiples en dérivant.

Un polynôme g tel que $g' = 0$ est un polynôme en x^p ; comme tout élément de $\mathbf{F}$ est une puissance p-ième, chacun des monômes de g est une puissance de p, donc g aussi. En particulier, g n'est pas irréductible. Dans l'autre sens, la dérivée d'un polynôme irréductible est non nulle.

Notons $f = \prod_{i=1}^{r} f_i^{e_i}$ la décomposition de f en facteurs irréductibles. Supposons $f' \neq 0$, de sorte qu'au moins un des e_i n'est pas multiple de p. Dans la décomposition

$$f' = \sum_{i=1}^{r} e_i f_i^{e_i-1} f_i' \prod_{j\neq i} f_j^{e_j},$$

le polynôme f_i apparaît avec multiplicité e_i dans chacun des termes d'indice $j \neq i$; il apparaît avec multiplicité $e_i - 1$ dans le terme d'indice i si e_i n'est pas multiple de p. En comparant avec f, on voit donc que

$$\operatorname{pgcd}(f, f') = \prod_{\substack{i=1 \\ p\nmid e_i}}^{r} f_i^{e_i-1} \prod_{\substack{i=1 \\ p\mid e_i}}^{r} f_i^{e_i}.$$

Par suite, $g_1 = f / \operatorname{pgcd}(f, f')$ est le produit des facteurs irréductibles f_i pour lesquels e_i n'est pas multiple de p.

On peut réitérer le procédé avec f/g_1 qui a la même décomposition en facteurs irréductibles que f, les exposants non multiples de p ayant diminué de 1. On obtient alors le produit g_2 des facteurs irréductibles f_i pour lesquels $e_1 \geqslant 2$ mais ni $e_i(e_i - 1)$ n'est pas multiple de p.

Au bout de $p-1$ opérations, l'exposant de f_i dans la décomposition de $f/g_1 \dots g_{p-1}$ est le plus grand multiple de p inférieur ou égal à e_i. Ce polynôme est donc la puissance p-ième d'un polynôme h qu'on peut déterminer explicitement et avec lequel on continue.

PROPOSITION 2.4.1. — *Pour que f soit irréductible, il faut et il suffit que les polynômes $x^{q^e} - x$ et f, pour $1 \leqslant e < n$, et même $1 \leqslant e \leqslant \lfloor n/2 \rfloor$, soient premiers entre eux.*

Démonstration. — Soit ξ une racine de f dans une clôture algébrique $\bar{\mathbf{F}}$ de $\mathbf{F}$. Son polynôme minimal g est irréductible et divise f, donc est un facteur irréductible de f. En outre, le corps $\mathbf{F}[\xi]$ engendré par ξ est un corps fini de cardinal q^e, où $e = \deg(g)$. Il en résulte que $\xi^{q^e} = \xi$, autrement dit les polynômes $x^{q^e} - x$ et f ont une racine commune dans $\bar{\mathbf{F}}$. Cela entraîne que leur pgcd n'est pas égal à 1. Si g est le facteur irréductible de plus petit degré de f, on a même $e \leqslant \lfloor n/2 \rfloor$.

Inversement, si f et $x^{q^e} - x$ ont un facteur irréductible commun, celui-ci sera de degré au plus e, donc distinct de f si $e < n$. □

On peut généraliser ce résultat pour obtenir une factorisation de f en un produit de polynômes f_e, chaque facteur irréductible de f_e étant de degré e. Pour simplifier, on suppose que f est sans facteur multiple.

PROPOSITION 2.4.2. — *Soit $f \in \mathbf{F}[x]$ un polynôme sans facteur multiple. Posons $g_0 = f$; pour $e \geqslant 1$, posons $f_e = \operatorname{pgcd}(x^{p^e} - x, g_{e-1})$ et $g_e = g_{e-1}/f_e$. On a $f = f_1 f_2 \dots$ et pour tout $e \geqslant 1$, chaque facteur irréductible de f_e est degré e.*

Démonstration. — Fixons pour la démonstration une clôture algébrique Ω de $\mathbf{F}$ et notons $\mathbf{F}_{q^e}$ l'unique sous-corps de Ω de cardinal $\mathbf{F}_{q^e}$.

Comme dans la démonstration précédente, les racines de f_1 sont les racines de f qui appartiennent à $\mathbf{F}_q$. Comme les racines de f sont simples, il en est de même de celles de f_1 et le polynôme f_1 est le produit des facteurs linéaires qui divisent f. Les racines de f_2 sont les racines de $g_1 = f/f_1$ qui appartiennent à $\mathbf{F}_{q^2}$; ce sont donc les racines de f qui appartiennent à $\mathbf{F}_{q^2}$ mais pas à $\mathbf{F}_q$ (ces dernières sont racines de f_1 et ne sont plus racines de g_1 car les racines de f sont simples). Par récurrence, les racines de f_e sont les racines de f qui appartiennent à $\mathbf{F}_{q^e}$ mais à aucun des corps $\mathbf{F}_{q^r}$ pour $1 \leqslant r < e$; l'extension de $\mathbf{F}_q$ engendrée par une telle racine est $\mathbf{F}_{q^e}$, si bien que son polynôme minimal est de degré e. □

Le dernier critère, dû à BERLEKAMP, fournit le nombre de facteurs irréductibles de f.

PROPOSITION 2.4.3. — *Supposons que f soit sans facteur carré et notons n son degré. Pour $1 \leqslant e < n$, soit g_e le reste de la division euclidienne du polynôme $x^{qe} - x^e$ par f. Soit k le rang du sous-espace de $\mathbf{F}[x]$ engendré par les polynômes $g_1, \dots, g_{n-1}$. Le nombre de facteurs irréductibles distincts de f est égal à $n - k$.*

Avant de démontrer cette proposition, insistons sur le fait que pour calculer en pratique les g_e, il importe de calculer x^q dans $\mathbf{F}[x]/(f)$ à l'aide de l'algorithme d'exponentiation rapide puis calculer ses puissances.

Démonstration. — Soit A l'algèbre $\mathbf{F}[x]/(f)$ et munissons-la de la base $(1, x, \dots, x^{n-1})$. Soit σ l'application de A dans elle-même donnée par $\sigma(a) = a^q$. C'est un homomorphisme d'algèbres. On a par définition $0 = \sigma(1) - 1$ et $g_e = \sigma(x^e) - x^e$ pour $1 \leqslant e < n$. La proposition équivaut donc à l'énoncé : *le noyau de* $\sigma - \mathrm{id}$ *est un* $\mathbf{F}$*-espace vectoriel de dimension égale au nombre* r *de facteurs irréductibles de* f.

Notons B le noyau de $\sigma - \mathrm{id}$. C'est l'ensemble des éléments $a \in A$ tels que $\sigma(a) = a$; c'est donc une sous-algèbre de A qu'on appelle *la sous-algèbre de Berlekamp de* A. Nous allons montrer qu'elle est isomorphe à $\mathbf{F}^r$.

Écrivons la décomposition de f en produits de facteurs irréductibles, $f = f_1 \dots f_r$. Ils sont premiers entre eux deux à deux ; d'après le lemme chinois, l'algèbre A est donc isomorphe à $\prod_{i=1}^r \mathbf{F}[x]/(f_i)$ et σ s'identifie à l'endomorphisme $(a_1, \dots, a_r) \mapsto (a_1^q, \dots, a_r^q)$. Par suite, l'image de B par l'isomorphisme chinois est l'ensemble des $(a_1, \dots, a_r)$ tels que $a_i^q = a_i$ pour tout i. Comme f_i est irréductible, $\mathbf{F}[x]/(f_i)$ est un corps fini de cardinal $q^{\deg(f_i)}$ et l'ensemble des $a_i \in F[x]/(f_i)$ tels que $a_i^q = a_i$ est le sous-corps à q éléments, c'est-à-dire $\mathbf{F}$. Autrement dit, l'image de B est l'algèbre $\mathbf{F} \times \dots \mathbf{F}$; elle est de dimension r sur $\mathbf{F}$. □

B. Algorithme de CANTOR-ZASSENHAUS (1981)

Il s'agit maintenant, étant donné un polynôme f à coefficients dans un corps fini $\mathbf{F}$, supposé sans racines multiples, de le factoriser. On suppose que tous ses facteurs irréductibles sont de même degré, ce qui est loisible compte tenu des résultats du paragraphe précédent.

Notons $f = f_1 \dots f_r$ la factorisation de f et e le degré des f_i, de sorte que $n = \deg(f) = er$. L'algèbre $A = \mathbf{F}[x]/(f)$ est isomorphe à $(\mathbf{F}_{q^e})^r$.

LEMME 2.4.1. — *Soit* K *un corps fini de cardinal* q.

Supposons q *impair. L'application* $u \mapsto u^{(q-1)/2}$ *de* K *dans lui-même prend une fois la valeur* 0 *(pour* $u = 0$*),* $(q-1)/2$ *fois la valeur* 1 *(lorsque* u *est un carré non nul) et* $(q-1)/2$ *fois la valeur* -1 *(sinon).*

Supposons que $q = 2^n$ *est une puissance de* 2. *L'application* $u \mapsto u + u^2 + u^{2^2} + \dots + u^{2^{n-1}}$ *de* K *dans lui-même prend* 2^{n-1} *fois la valeur* 0 *et* 2^{n-1} *fois la valeur* 1.

Démonstration. — Notons φ cette application.

Dans le cas où q est impair, on a $\varphi(u)^2 = 1$ si $u \neq 0$ et $\varphi(u) = 0$ sinon ; par suite, $\varphi(u)$ vaut ± 1 ou 0. Comme φ est une application polynomiale de degré $(q-1)/2$, elle prend au plus $(q-1)/2$ fois les valeurs -1 et 1, donc exactement $(q-1)/2$ fois. (En fait, $\varphi(u)$ est une généralisation du symbole de Legendre qui vaut 1 si u est un carré non nul, 0 si $u = 0$ et -1 dans les autres cas.)

Lorsque q est une puissance de 2, on a

$$\varphi(u)^2 = u^2 + u^{2^2} + \dots + u^{2^n} = \varphi(u) + u^{2^n} - u = \varphi(u).$$

Par conséquent, $\varphi(u)(\varphi(u)-1)=0$ et $\varphi(u)$ vaut 0 ou 1. En outre, φ prend au plus 2^{n-1} fois chaque valeur, donc exactement 2^{n-1} fois la valeur 0 et 2^{n-1} fois la valeur 1. □

PROPOSITION 2.4.2. — *Soit* $\mathbf{F}$ *un corps fini de cardinal* q*, soit* $f \in \mathbf{F}[x]$ *un polynôme sans facteur carré dont tous les facteurs irréductibles sont de degré* e*.*

a) Supposons que q *soit impair.*

Soit g *un polynôme non nul de degré* $< n$ *et soit* h *le reste de la division par* f *du polynôme* $g^{(q^e-1)/2}$*. Alors,* f *est le produit de* $\operatorname{pgcd}(f,h)$*,* $\operatorname{pgcd}(f,h-1)$ *et* $\operatorname{pgcd}(f,h+1)$*.*

En outre, parmi les q^n-1 *polynômes possibles* g*, seuls* $2((q^e-1)/2)^r$ *fournissent une factorisation triviale.*

b) Supposons que $q=2^m$ *soit une puissance de* 2*. Soit* g *un polynôme de degré* $< n$ *et soit* h *le reste de la division par* f *du polynôme* $g+g^2+\cdots+g^{2^{em-1}}$*. Alors,* f *est le produit de* $\operatorname{pgcd}(f,h)$ *et de* $\operatorname{pgcd}(f,h-1)$*.*

En outre, parmi les q^n *polynômes possibles* g*, seuls* $2(2^{em-1})^r$ *fournissent une factorisation triviale.*

Lorsque q est impair, la probabilité d'obtenir une factorisation triviale est donc égale à

$$2^{1-r}\frac{(q^e-1)^r}{q^{er-1}} \leqslant 2^{1-r}.$$

Lorsque q est une puissance de 2, elle est encore 2^{1-r}. Par conséquent, si $r \geqslant 2$, on peut espérer, en tirant des polynômes g au hasard, factoriser f très rapidement. [2] Il est cependant possible de n'avoir trouvé aucun facteur non trivial au bout de 1000 essais, mais la probabilité est de l'ordre de $2^{1000(1-r)}$, donc ridiculement faible.

Démonstration. — Notons A l'algèbre $\mathbf{F}[x]/(f)$; comme f est produit de r polynômes irréductibles de degré e, deux à deux distincts, le théorème chinois entraîne que A est isomorphe à K^r, où K est un corps à q^e éléments.

Étant donné un polynôme g de degré $< n$, identifié à un élément de A, donc à une famille $(u_1,\dots,u_r)$ d'éléments de K, le calcul de l'énoncé revient à celui de $(\varphi(u_1),\dots,\varphi(u_r))$ dans K^r, où φ est l'application du lemme précédent. Par suite, $\varphi(u_i) \in \{0,1,-1\}$ dans le cas où q est impair, et $\varphi(u_i) \in \{0,1\}$ quand q est une puissance de 2. Cela revient à dire que $h \equiv \varphi(u_i) \pmod{f}_i$, donc que h est multiple des polynômes f_i tels que $\varphi(u_i)=0$, $h-1$ est multiple des polynômes f_i tels que $\varphi(u_i)=1$, et $h+1$ est multiple des polynômes f_i tels que $\varphi(u_i)=-1$. Comme les f_i sont premiers

2. Si $r=1$, cette probabilité est égale à 1, comme il se doit.

entre eux deux à deux, h, $h-1$ et $h+1$ sont en fait multiples des produits correspondants. (Ce dernier polynôme n'intervient pas lorsque q est une puissance de 2.) La première partie de l'énoncé en découle immédiatement.

Pour la seconde, il s'agit de dénombrer les familles $(u_1,\dots,u_r)$ tels que $(\varphi(u_1),\dots,\varphi(u_r))$ soit $(0,\dots,0)$, $(1,\dots,1)$ ou $(-1,\dots,-1)$. On trouve exactement les valeurs indiquées. □

Le cas particulier d'un polynôme scindé dans $\mathbf{F}$ est très important. C'est en effet à lui qu'on se ramène dans l'algorithme de BERLEKAMP exposé ci-dessous, c'est aussi ce qui se passe pour un polynôme de petit degré.

PROPOSITION 2.4.3. — *Soit $\mathbf{F}$ un corps fini de cardinal q et soit $f\in\mathbf{F}[x]$ un polynôme de degré n sans facteur carré qui est scindé dans $\mathbf{F}$.*

a) Supposons q impair. Pour $a\in\mathbf{F}$, posons $h_a=(x+a)^{(q-1)/2} \pmod f$. Alors f est le produit de $\gcd(h_a,f)$, $\gcd(h_a-1,f)$ et $\gcd(h_a+1,f)$. En outre, au plus la moitié des valeurs de a fournit une factorisation triviale.

b) Supposons que $q=2^m$ est une puissance de 2. Pour $a\in\mathbf{F}^*$, posons $h_a=(ax)+(ax)^2+\cdots+(ax)^{2^{m-1}}$. Alors, f est le produit de $\gcd(h_a,f)$ et de $\gcd(h_a-1,f)$. En outre, au plus la moitié des valeurs de a fournit une factorisation triviale.

Dans les deux cas, la probabilité, tirant a au hasard, d'obtenir une factorisation triviale est donc inférieure ou égale à $1/2$.

Démonstration. — Notons $x_1,\dots,x_n$ les racines de f dans $\mathbf{F}$.

Étudions d'abord le cas où q est impair. Dire que a fournit une factorisation triviale, signifie que x_i-a est identiquement nul, carré ou non carré lorsque i parcourt $\{1,\dots,n\}$. Autrement dit, les « mauvaises » valeurs de a sont celles pour lesquelles x_i-a est non nul, et toujours carré ou non carré. En particulier, l'élément $(x_2-a)/(x_1-a)$ de $\mathbf{F}\cup\{\infty\}$ doit toujours être non carré (∞ étant considéré comme un carré). Comme $a\mapsto(x_2-a)/(x_1-a)$ est une homographie, c'est une application bijective de $\mathbf{F}\cup\{\infty\}$ et au plus $(q-1)/2q$ valeurs de a sont mauvaises.

Le cas où q est une puissnace de 2 est analogue. Soit S l'ensemble des 2^{m-1} solutions (dans $\mathbf{F}$) de l'équation $\varphi(u)=u+u^2+\cdot+u^{2^{m-1}}$; c'est un sous-groupe de $\mathbf{F}$ car $\varphi(u+v)=\varphi(u)+\varphi(v)$. En outre, $h_a(x_i)$ vaut 0 si ax_i appartient à S et vaut 1 sinon. En particulier, si $a(x_2-x_1)\notin S$, ax_1 et ax_2 n'appartiennent pas tous deux à S, $h_a(x_1)\neq h_a(x_2)$, donc l'une des racines, x_1 et x_2, est racine de $\gcd(h_a,f)$ tandis que l'autre est racine de $\gcd(h_a-1,f)$. Ainsi, la factorisation n'est triviale que dans au plus $(2^{m-1}-1)$ cas, c'est-à-dire moins de la moitié. □

C. Algorithme de BERLEKAMP (1970)

Revenons au problème de la factorisation d'un polynôme f de degré n, à coefficients dans un corps fini $\mathbf{F}$ de cardinal q, supposé sans racines multiples. Notons

encore A l'algèbre $\mathbf{F}[x]/(f)$ et B la sous-algèbre de Berlekamp, noyau de l'endomorphisme d'espaces vectoriel $a \mapsto a^q - a$. Par l'algèbre linéaire (algorithme de Gauss, ou de Gauss-Jordan), on peut en trouver une $\mathbf{F}$-base $b_1, \ldots, b_r$ représentée par des polynômes de $\mathbf{F}[x]$ de degrés $< n$.

Si $f = \prod_{i=1}^{r} f_i$, le théorème chinois identifie A au produit $K_1 \ldots K_r$ des corps finis $\mathbf{F}[x]/(f_i)$ et identifie B à la sous-algèbre $\mathbf{F}^r$.

Soit b un élément de B. Par définition, il existe pour tout $i \in \{1, \ldots, r\}$ un unique élément $\varphi_i(b) \in \mathbf{F}$ tel que $b \equiv \varphi_i(b) \pmod{f}_i$. Alors, f_i est un facteur de $b - \varphi_i(b)$ si bien que la connaissance des $\varphi_i(b)$ permet de factoriser f. De fait, on a la formule

$$f = \prod_{u \in \mathbf{F}} \operatorname{pgcd}(f, b - u).$$

Pour que l'on obtienne une factorisation non triviale, il faut et il suffit que les $\varphi_i(b)$ ne soient pas tous égaux, c'est-à-dire que b n'appartienne pas à la sous-algèbre diagonale $\mathbf{F}$ de B. Toutefois, la formule donnée est totalement inutilisable dans les applications pratiques où le cardinal q de $\mathbf{F}$ est très grand ; en effet elle suppose de calculer q pgcd, ce qui est prendra bien trop de temps.

Il se pose donc la question de déterminer efficacement les $\varphi_i(b)$. On s'inspire pour cela de l'argument déjà utilisé dans l'algorithme de Cantor-Zassenhaus (historiquement, cela s'est passé dans l'autre sens...).

Supposons d'abord q impair. Comme $b^q - b$ est multiple de f, on a l'égalité

$$f = \operatorname{pgcd}(f, b^{(q-1)/2} - 1) \operatorname{pgcd}(f, b^{(q-1)/2} + 1) \operatorname{pgcd}(f, b).$$

Il s'agit de voir à quelle probabilité cette factorisation est non triviale.

Observons que pour tout i, $\varphi_i(b)^{(q-1)/2}$ vaut 0, 1 ou -1. Il s'agit de trouver b de sorte que ces valeurs ne soient pas toutes égales. Si b est choisi au hasard dans $B \simeq \mathbf{F}^r$, $\varphi_i(b)^{(q-1)/2}$ vaut $(1, \ldots, 1)$ si et seulement si tous les $\varphi_i(b)$ sont des carrés non nuls, $(-1, \ldots, -1)$ si et seulement aucun des $\varphi_i(b)$ n'est un carré, $(0, \ldots, 0)$ si et seulement si $\varphi_i(b) = 0$ pour tout i. Il y a donc $2((q-1)/2)^r + 1$ tels éléments b, qui aboutissent à une factorisation triviale de f, parmi lesquels tous les éléments de la sous-algèbre $\mathbf{F}$. En dehors de ces choix, la factorisation n'est pas triviale, ce qui arrive avec une probabilité supérieure à $1 - 2^{1-r}$.

Pour traiter *le cas où q est pair,* on remplace le polynôme $X^{(q-1)/2}$ par le polynôme $T = X + X^2 + X^4 + \cdots + X^{2^{s-1}}$, où $q = 2^s$ qui vérifie $T(T-1) = T^2 - T = X^q - X$, d'où l'égalité

$$f = \operatorname{pgcd}(f, T(b)) \operatorname{pgcd}(f, T(b) - 1),$$

si b est un élément arbitraire de B. Il s'agit encore de voir si cette factorisation peut être non triviale lorsque b est choisi au hasard. Lorsque b est choisit au hasard dans $B \sim \mathbf{F}^r$, distinct d'un élément de $\mathbf{F}$, $T(b)$ vaut $(0, \ldots, 0)$ ou $(1, \ldots, 1)$ si et seulement si tous les $\varphi_i(b)$ ont même image par T. Or, 0 et 1 ont tous deux 2^{s-1} antécédents ; il y a donc

$2(2^{s-1})^r$ éléments de B dont l'image est $(0,\dots,0)$ ou $(1,\dots,1)$. Ainsi, prenant b au hasard, la probabilité d'obtenir une factorisation non triviale est au moins égale à 2^0 parmi lesquels les $q = 2^s$ éléments de $\mathbf{F}$. Parmi les éléments $q^r - q$ éléments de $B \setminus \mathbf{F}$, $2^{rs-r+1} - q = q^r 2^{1-r} - q$ fournissent une factorisation triviale, et donc $q^r(1-2^{1-r})$ une factorisation non triviale. Là encore, la probabilité d'obtenir une factorisation non triviale est supérieure à $1-2^{1-r}$.

Pour résumer, la méthode est la suivante : compléter la famille (1) en une base de l'espace vectoriel des polynômes $Q \in \mathbf{F}[x]$ de degrés $< n$ tels que $Q^q \equiv Q \pmod f$, espace identifié à B et dont on note r la dimension ; choisir un élément Q au hasard dans $B \setminus \mathbf{F}$ puis calculer le pgcd avec f de $Q^{(q-1)/2} - 1$ (si q est impair) ou de $Q + Q^2 + \cdots + Q^{2^{s-1}}$ (si $q = 2^s$ est pair), d'où, avec probablité au moins $1-2^{1-r}$, un facteur non trivial de f.

Exercices

29) Soit p un nombre premier impair tel que $\ell = 2p+1$ soit un nombre premier. Soit (x,y,z) une solution de l'équation de Fermat

$$x^p + y^p + z^p = 0.$$

Le but de l'exercice est de démontrer le *théorème de Sophie Germain* (1820) selon lequel p divise xyz : selon la terminologie consacrée, il s'agit de la résolution du premier cas de Fermat pour les exposants qui sont des nombres premiers de Sophie Germain, le second cas étant celui où p ne divise aucun des entiers x, y, z. On raisonne par l'absurde en supposant le contraire.

a) Montrer que $y+z$ et $(y^p+z^p)/(y+z)$ sont des entiers premiers entre eux. En déduire qu'il existe des entiers a et A tels que

$$x = -aA \quad \text{et} \quad y+z = a^p.$$

b) Démontrer de même qu'il existe des entiers b, B, c, C tels que $y = -bB$, $x+z = b^p$, $z = -cC$, $x+y = c^p$. Remarquer que les entiers a, b, c, A, B, C sont premiers entre eux deux à deux.

c) En notant que pour tout nombre entier t, t^p vaut 0, 1 ou -1 modulo ℓ, démontrer que xyz est multiple de ℓ. Si ℓ divise x, montrer en outre que ℓ divise a.

d) Montrer que $A^p \equiv pb^{p-1} \pmod \ell$ puis conclure que p est une puissance p-ième dans $\mathbf{F}_\ell$. En déduire la contradiction recherchée.

e) Donner quelques exemples de nombres premiers p pour lesquels $2p+1$ est un nombre premier.

30) Soit p un nombre premier

a) Soit $\psi : \mathbf{Z}/p\mathbf{Z} \to \mathbf{C}^*$ l'application définie par $\psi(x) = \exp(2i\pi\xi/p)$, où $\xi \in \mathbf{Z}$ est un représentant quelconque de x. Montrer que ψ est un homomorphisme de groupes.

b) Pour tout $a \in \mathbf{F}_p$, calculer $\sum_{x\in\mathbf{F}_p} \psi(ax)$ en fonction de a.

c) Soit $\chi : \mathbf{F}_p^* \to \mathbf{C}^*$ un homomorphisme de groupes. On pose

$$G_a(\chi) = \sum_{x\in\mathbf{F}_p^*} \chi(x)\psi(ax)$$

(somme de Gauß). Montrer que $\big|G_a(\chi)\big| = \sqrt{p}$.

d) Soit $F \in \mathbf{F}_p[X_1, \dots, X_n]$. Montrer que le nombre de solutions dans $(\mathbf{F}_p)^n$ de l'équation $F(x_1, \dots, x_n) = 0$ est égal à

$$\frac{1}{p} \sum_{(x_0, x_1, \dots, x_n) \in (\mathbf{F}_p)^{n+1}} \psi(x_0 F(x_1, \dots, x_n)).$$

e) Soit $a \in \mathbf{F}_p^*$ et soit d un entier tel que $d \geqslant 1$. Montrer que

$$\sum_{\chi \colon \mathbf{F}_p^* \to \mu_d(\mathbf{C})} \chi(a)$$

est égal au nombre de solutions dans $\mathbf{F}_p$ de l'équation $x^d = a$. (*Utiliser le fait que* $\mathbf{F}_p^*$ *est cyclique.*)

f) Soit d un entier strictement positif et soit a un élément de $\mathbf{F}_p^*$. Montrer que

$$\sum_{y \in \mathbf{F}_p} \psi(a y^d) = \sum_{\substack{\chi \colon \mathbf{F}_p^* \to \mu_d(\mathbf{C}) \\ \chi \neq 1}} G_a(\chi).$$

g) Montrer que $\big|G_a(\chi)\big| = \sqrt{p}$.

h) Soit n un entier, soit $a_1, \dots, a_n$ des éléments non nuls de $\mathbf{F}_p$ et $d_1, \dots, d_n$ des nombres entiers strictement positifs. On note N le nombre de solutions dans $(\mathbf{F}_p)^n$ de l'équation

$$a_1 x_1^{d_1} + \cdots + a_n x_n^{d_n} = 0.$$

Montrer que l'on a

$$\big|N - p^{n-1}\big| \leqslant C(p-1)p^{\frac{n}{2}-1},$$

où $C = \prod_{i=1}^n (d_i - 1)$.

i) Le résultat est-il évident si $C = 0$? Observer aussi que l'on peut remplacer d_i par $r_i = \operatorname{pgcd}(d_i, p-1)$ dans la formule qui définit C sans altérer la véracité de l'inégalité précédente.

C'est un cas particulier d'une série d'inégalités concernant le nombre de solutions d'équations polynomiales dans les corps finis, dont les premiers exemples remontent au XIX$^{\text{e}}$ siècle (Gauß), furent démontrées pour les « courbes » et conjecturées en général par André Weil dans les années 1940–1950. Leur démonstration n'a été achevée qu'avec les travaux de Grothendieck, Artin, Verdier dans les années 60, puis Deligne en 1973. C'est précisément cet exemple qui est donné par Weil comme indication de la véracité générale de sa conjecture.

31) Soit p un nombre premier et soit P un polynôme de degré d en n variables à coefficients dans le corps $\mathbf{F}_p$. Soit N le nombre de solutions dans $(\mathbf{F}_p)^n$ de l'équation $P(x_1, \dots, x_n) = 0$.

a) Montrer que l'on a

$$N \equiv \sum_{(x_1, \dots, x_n) \in (\mathbf{F}_p)^n} \big(1 - P(x_1, \dots, x_n)^{p-1}\big) \pmod{p}.$$

b) Calculer, pour $h \in \mathbf{N}$, l'expressions $S_h = \sum_{x \in \mathbf{F}_p} x^h$.

c) Montrer que pour tout $(j_1, \dots, j_n) \in \mathbf{N}^n$ tel que $j_1 + \cdots + j_n < n(p-1)$, on a

$$\sum_{(x_1, \dots, x_n) \in (\mathbf{F}_p)^n} x_1^{j_1} \dots x_n^{j_n} = 0.$$

d) Si $d < n$, montrer que N est multiple de p (*théorème de Chevalley-Warning,* 1936). En particulier, une équation homogène de degré d possède au moins une solution non nulle.

32) Soit $\mathbf{F}$ un corps fini ; on note q son cardinal.

a) Montrer que l'espace vectoriel $\mathbf{F}^2$ est réunion de $q+1$-droites vectorielles, et que $q+1$ droites sont effectivement nécessaires pour recouvrir $\mathbf{F}^2$. Plus généralement, combien faut-il d'hyperplans pour recouvrir $\mathbf{F}^n$?

b) Soit $H_1, \ldots, H_d$ des hyperplans affines de $\mathbf{F}_q^n$ qui ne passent pas par l'origine et qui recouvrent $\mathbf{F}_q^n \setminus \{0\}$. Montrer que $d \geqslant n(q-1)$. Montrer aussi que cette minoration est optimale en exhibant un tel recouvrement avec $d = n(q-1)$. (*Si f_i est une équation de H_i, poser $f = \prod_{i=1}^d f_i$ et considérer la quantité $S(f)$ introduite dans l'exercice 31.*)

CHAPITRE 3

GÉOMÉTRIE DES NOMBRES

§3.1. Formes quadratiques binaires

Dans ce paragraphe, nous présentons quelques résultats classiques concernant la théorie arithmétique des formes quadratiques en deux variables.

La motivation essentielle de cette étude est de trouver des conditions nécessaires ou suffisantes pour qu'un entier n s'écrive sous la forme $ax^2+bxy+cy^2$, pour $(x,y)\in\mathbf{Z}^2$, où a,b,c sont des nombres entiers fixés. Nous allons voir que le résultat est intimement lié à l'ensemble des formes quadratiques qu'on obtient à partir de $ax^2+bxy+cy^2$ par changement de variables dans $\mathrm{SL}_2(\mathbf{Z})$.

A. Minimum d'une forme quadratique binaire

Soit $q(x,y)=ax^2+bxy+cy^2$ une forme quadratique en deux variables, à coefficients (a,b,c) réels. Son *discriminant* est défini par la formule $D=b^2-4ac$; il s'agit du discriminant au sens des polynômes du second degré, c'est-à-dire -4 fois le déterminant de la matrice $\left(\begin{smallmatrix} a & b/2 \\ b/2 & c \end{smallmatrix}\right)$ de la forme quadratique q. On dit que q est non dégénérée si $D\neq 0$. On dit que q est définie positive si $q(x,y)>0$ pour tout couple (x,y) de nombres réels distinct de $(0,0)$; cela équivaut aux inégalités $a>0$ et $D<0$.

THÉORÈME 3.1.1. — *Soit $q(x,y)=ax^2+bxy+cy^2$ une forme quadratique définie positive, de discriminant $D=b^2-4ac<0$. Alors, il existe un couple $(x,y)\in\mathbf{Z}^2$, distinct de $(0,0)$, tel que*

$$0<q(x,y)\leqslant\sqrt{|D|/3}.$$

Démonstration. — Comme q est définie positive, $a>0$ et $D<0$. La décomposition de Gauß

$$q(x,y)=a\left(x+\frac{b}{2a}y\right)^2+(c-\frac{b^2}{4a})y^2=a\left(x+\frac{b}{2a}y\right)^2-\frac{1}{4a}Dy^2$$

montre que $q(x,y)$ tend vers l'infini quand x ou y tend vers l'infini. Par suite, $q(x,y)$ atteint sa borne inférieure sur l'ensemble $\mathbf{Z}^2\setminus\{(0,0)\}$; soit $(\xi,\eta)\in\mathbf{Z}^2$ un couple non nul où q soit minimale.

Observons que ξ et η sont premiers entre eux. En effet, si d désigne leur pgcd, on a $q(\xi,\eta) = d^2 q(\xi/d, \eta/d)$ et $(\xi/d, \eta/d)$ est un couple en lequel q prend la valeur $q(\xi,\eta)/d^2$ qui est $< q(\xi,\eta)$ si $d \neq 1$. D'après le théorème de Bézout, il existe donc des entiers u et v tels que $u\xi - v\eta = 1$. Faisons dans q le changement de variables

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \xi & v \\ \eta & u \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Comme la matrice $\left(\begin{smallmatrix} \xi & v \\ \eta & u \end{smallmatrix}\right)$ est dans $\mathrm{Mat}_2(\mathbf{Z})$, elle induit une application de $\mathbf{Z}^2$ dans lui-même ; puisque son déterminant est égal à 1, son inverse est aussi dans $\mathrm{Mat}_2(\mathbf{Z})$ et ce changement de variables induit un automorphisme du groupe abélien $\mathbf{Z}^2$. Après ce changement de variables, q s'écrit $q(x,y) = Q(X,Y) = AX^2 + BXY + CY^2$, avec $A = q(\xi,\eta)$. De plus, le discriminant de la forme Q est égal à celui de q car le déterminant du changement de variables est de carré 1. Effectuons alors une réduction de Gauß en écrivant

$$Q(X,Y) = A\left(X + \frac{B}{2A}Y\right)^2 - \frac{1}{4A}DY^2.$$

Par hypothèse, $Q(X,Y) \geqslant A$ si $(X,Y) \neq (0,0)$. Prenons par exemple $Y = 1$ et choisissons $X \in \mathbf{Z}$ de sorte que $\left|X + \frac{B}{2A}Y\right| \leqslant 1/2$; il suffit de prendre pour X l'entier le plus proche de $-B/2A$. Comme $D < 0$, on obtient

$$Q(X,Y) \leqslant \frac{1}{4}4 + \frac{1}{4A}D,$$

d'où, puisque $Q(X,Y) \geqslant A$, l'inégalité $\frac{3}{4}A \leqslant \frac{1}{4A}D$, soit encore $A^2 \leqslant D/3$, ce qu'il fallait démontrer. □

Voyons-en tout de suite quelques applications.

PROPOSITION 3.1.2. — *a) Pour qu'un nombre premier p s'écrive sous la forme $x^2 + y^2$, avec $(x,y) \in \mathbf{Z}^2$, il faut et il suffit que $p = 2$ ou $p \equiv 1 \pmod 4$.*

b) Pour qu'un nombre premier p s'écrive sous la forme $x^2 + 2y^2$, avec $(x,y) \in \mathbf{Z}^2$, il faut et il suffit que $p = 2$ ou que p soit congru à 1 ou 3 modulo 8.

c) Pour qu'un nombre premier p s'écrive sous la forme $x^2 + 3y^2$, avec $(x,y) \in \mathbf{Z}^2$, il faut et il suffit que $p = 3$ ou $p \equiv 1 \pmod 3$.

Démonstration. — a) Si $p = x^2 + y^2$, x et y ne sont pas multiples de p (si p divise x, alors p divise $y^2 = p - x^2$, donc p divise y et p^2 divise aussi $p = x^2 + y^2$, ce qui est absurde). On a alors $x^2 \equiv -y^2$ dans $\mathbf{Z}/p\mathbf{Z}$ et le quotient $\bar{x}/\bar{y}$ vérifie $(\bar{x}/\bar{y})^2 \equiv -1$; autrement dit, -1 est un carré dans $\mathbf{Z}/p\mathbf{Z}$ et cela entraîne $p = 2$ ou $p \equiv 1 \pmod 4$.

Établissons la réciproque. On a déjà $2 = 1^2 + 1^2$; supposons maintenant que p soit congru à 1 modulo 4. Il existe alors un entier a tel que $a^2 \equiv -1 \pmod 4$;

introduisons la forme quadratique $q(x,y) = x^2 + (ax+py)^2$. (C'est, après changement de variables, la restriction de la forme x^2+y^2 à l'ensemble des couples (x,y) de $\mathbf{Z}^2$ qui vérifient $y \equiv a \pmod 4$.) On a

$$q(x,y) = (1+a^2)x^2 + 2apxy + p^2y^2$$

et le discriminant de q est égal à $4a^2p^2 - 4p^2(1+a^2) = 4p^2$. D'après le théorème précédent, il existe un couple $(x,y) \in \mathbf{Z}^2$, non nul, tel que $q(x,y) \leqslant p\sqrt{4/3} < 2p$.

Puisque $(x,y) \neq (0,0)$, $q(x,y) \neq 0$. On a en outre $q(x,y) \equiv x^2(1+a^2) \equiv 0 \pmod p$: $q(x,y)$ est multiple de p, or le seul multiple de p non nul inférieur à $p\sqrt{4/3}$ est p. On a donc $q(x,y) = p$. En particulier, $p = x^2 + (ax+py)^2$ est somme de deux carrés d'entiers.

b) On a $2 = 0^2 + 2\cdot 1^2$; supposons désormais que p soit impair. Supposons $p = x^2+2y^2$, avec $(x,y) \in \mathbf{Z}^2$. Comme précédemment, y n'est pas multiple de p (sinon $0 \equiv x^2 \pmod p$, donc x serait aussi multiple de p et $p = x^2+2y^2$ serait multiple de p^2 !). Modulo p, la relation $p = x^2+2y^2$ devient alors $\bar{x}^2 = -2\bar{y}^2$ et $\bar{x}/\bar{y}$ est une racine carrée de -2. D'après la loi de réciprocité quadratique, cela équivaut à $p \equiv 1,3 \pmod 8$.

Pour démontrer la réciproque, introduisons la forme quadratique $q(x,y) = 2x^2 + (ax+py)^2$, où a est un entier tel que $a^2 \equiv -2 \pmod p$. (Compte tenu de la loi de réciprocité quadratique, l'existence d'un tel entier provient de la congruence vérifiée par p.) Son discriminant est $-8p^2$; il existe donc un couple non nul $(x,y) \in \mathbf{Z}^2$ tel que $q(x,y) \leqslant p\sqrt{8/3} < 2p$. Par construction, $q(x,y) \equiv 0 \pmod p$ et $q(x,y) > 0$; nécessairement, $q(x,y) = p$ et $p = (ax+py)^2 + 2x^2$ est de la forme x^2+2y^2.

c) Supposons $p = x^2+3y^2$; alors p ne divise pas y puis -3 est un carré modulo p. D'après la loi de réciprocité quadratique, cela entraîne $p = 3$ ou $p \equiv 1 \pmod 3$.

Inversement, on a $3 = 0^2 + 3\cdot 1^2$; lorsque $p \neq 3$, introduisons la forme quadratique $q(x,y) = 3x^2 + (ax+py)^2$, où a est un entier tel que $a^2 \equiv -3 \pmod p$. Le discriminant de cette forme est égal à $-12p^2$. Il existe ainsi un couple non nul $(x,y) \in \mathbf{Z}^2$ tel que $q(x,y) \leqslant p\sqrt{4\cdot 3/3} = 2p$, d'où l'on déduit que $q(x,y) = p$ ou $q(x,y) = 2p$. Il faut exclure ce dernier cas. Or, si $2p = 3x^2 + (ax+py)^2$, il vient $(ax+py)^2 \equiv 2p \equiv 2 \pmod 3$, ce qui est absurde puisque 2 n'est pas un carré modulo 3. Nécessairement, $q(x,y) = p$ et p est de la forme x^2+3y^2. □

C'est ici que la méthode atteint ses limites. Par exemple, il est facile de vérifier que 23 ne s'écrit pas sous la forme x^2+5y^2 bien que -5 soit un carré modulo 23. Pour autant, voir l'exercice 6.

B. Formes quadratiques réduites

Au paragraphe précédent, nous avons pu voir l'utilité d'écrire une forme quadratique $q(x,y) = ax^2 + bxy + cy^2$ dans une autre base, le changement de base étant cependant donné par une matrice de $\mathrm{SL}_2(\mathbf{Z})$.

Nous dirons ainsi que deux formes quadratiques binaires q et q' sont *équivalentes* s'il existe une matrice $A \in \mathrm{SL}_2(\mathbf{Z})$ telle que $q(x,y) = q'(A\cdot(x',y'))$. Cette notion est évidemment une relation d'équivalence ; deux formes équivalentes ont même discriminant.

Nous dirons enfin qu'une forme définie positive $q(x,y) = ax^2+bxy+cy^2$, est *réduite* si ses coefficients vérifient $|b| \leqslant a \leqslant c$ et si l'on a de plus $b \geqslant 0$ si $c = a$ ou $a = |b|$.

PROPOSITION 3.1.1. — *Supposons que $q(x,y) = ax^2+bxy+cy^2$ soit une forme réduite et notons $D = b^2 - 4ac$ son discriminant. Alors, $a \leqslant \sqrt{D/3}$. De plus, pour tout couple $(x,y) \in \mathbf{Z}^2$ distinct de $(0,0)$ et $(\pm 1, 0)$, on a $q(x,y) \geqslant a$. L'inégalité est toujours stricte à moins que l'on ne soit dans l'un des cas suivants :*

- *On a l'égalité $c = a$ et $(x,y) = (0,\pm 1)$;*
- *On a les égalités $c = a = b$ (autrement dit, q est multiple de la forme x^2+xy+y^2) et $(x,y) = \pm(1,1)$.*

Démonstration. — Observons que l'on a les inégalités $b^2 \leqslant a^2 \leqslant ac$, d'où $3b^2 = 4b^2 - b^2 \leqslant -D$. De plus, $12ac = -3D + 3b^2 \leqslant -4D$ et, comme $a \leqslant c$, $a \leqslant \sqrt{D/3}$.

Par ailleurs, on a pour tout $(x,y) \in \mathbf{R}^2$ l'inégalité $\left|xy\right| \geqslant \min(x^2, y^2)$ si bien que pour tout $(x,y) \in \mathbf{Z}^2$ non nul,

$$q(x,y) \geqslant (a - |b| + c)\min(x^2, y^2).$$

Si aucun de x, y n'est nul, on en déduit que $q(x,y) \geqslant a - |b| + c \geqslant a$, l'égalité ne pouvant avoir lieu que si $a = c = |b|$; dans ce cas, $b \geqslant 0$, donc $a = b = c$ et $q(x,y) = a(x^2+xy+y^2)$ ne vaut a que pour les couples $(x,y) = \pm(1,1)$, $(x,y) = \pm(0,1)$ et $(x,y) = \pm(1,0)$. Si $x = 0$ (mais $y \neq 0$), on a $q(x,y) = cy^2 \geqslant c$, l'égalité n'ayant lieu que si $c = a$, en le couple $(x,y) = \pm(0,1)$; si $y = 0$ (mais $x \neq 0$), on a $q(x,y) = ax^2 \geqslant a$, avec égalité pour $(x,y) = \pm(1,0)$. □

L'intérêt de la notion de forme réduite provient du théorème suivant, dû à Gauß.

THÉORÈME 3.1.2. — *Toute forme quadratique définie positive est équivalente à une forme réduite et une seule.*

Démonstration. — Soit q une forme quadratique définie positive. Comme dans la démonstration du théorème 3.1.1, on peut, quitte à effectuer un changement de base dans $\mathrm{SL}_2(\mathbf{Z})$, supposer que $q(1,0)$ est le minimum des valeurs de q aux couples non nuls $(x,y) \in \mathbf{Z}^2$; cela modifie q en une forme équivalente, disons $ax^2 + 2bxy + cy^2$.

Pour $u \in \mathbf{Z}$, écrivons alors

$$q(x,y) = a(x+uy)^2 + (b-2au)xy + (c-au^2)y^2 = a(x+uy)^2 + (b-2au)(x+uy)y + c'y^2,$$

avec $c' = q(-u,1) = au^2 - bu + c$. Le changement de variables $x' = x+uy$, $y' = y$ est dans $\mathrm{SL}_2(\mathbf{Z})$; en choisissant u égal à l'entier le plus proche de $b/2a$ (choisi inférieur à $b/2a$ s'il y en a deux), on se ramène au cas où l'on a de plus $-a < b \leqslant a$. Comme a est le minimum de q, on a aussi $a \leqslant q(0,1) = c$, d'où les inégalités $|b| \leqslant a \leqslant c$.

Si $|b| = a$, on a de plus l'inégalité $b \geqslant 0$. Supposons que l'on ait $a = c$, c'est-à-dire que q soit équivalente à $ax^2 + bxy + ay^2$ mais que $b < 0$. Comme le changement de variables $x = y'$ et $y = -x'$ transforme cette dernière forme en $ax^2 - bxy + ay^2$, on voit que q est aussi équivalente à $ax^2 - bxy + ay^2$ qui est une forme réduite. □

C. Représentation des nombres premiers par des formes réduites

Nous pouvons maintenant aborder le problème de la représentation d'un nombre premier par une forme quadratique binaire, définie positive et à coefficients entiers, c'est-à-dire de la forme $q(x,y) = ax^2 + bxy + cy^2$ avec $a,b,c \in \mathbf{Z}$.

Si q est une forme à coefficients entiers et $U \in \mathrm{Mat}_2(\mathbf{Z})$, alors $q(U \cdot (x,y))$ est encore à coefficients entiers (changement de variables linéaire dans un polynôme homogène de degré 2). Par conséquent, une forme équivalente à une forme à coefficients entiers est à coefficients entiers.

Le discriminant $D = b^2 - 4ac$ d'une forme entière vérifie la congruence $D \equiv b^2 \equiv 0,1 \pmod 4$. Si $D \equiv 0 \pmod 4$, cela signifie que b est pair (on dit que q est pair), si $D \equiv 1 \pmod 4$, que b est impair (on dit que q est impaire). Comme les discriminants de deux formes équivalentes sont égaux, cette propriété est conservée par équivalence ; autrement dit, une forme équivalente à une forme paire est paire, une forme équivalente à une forme impaire est impaire.

Inversement, si D est un entier qui vérifie cette congruence, il existe une forme quadratique entière de discriminant D, par exemple la forme $q(x,y) = x^2 + \frac{D}{4}y^2$ si D est multiple de 4, et $q(x,y) = x^2 + xy + \frac{D-1}{4}$ si $D \equiv 1 \pmod 4$. Cette forme est appelée la forme principale.

Nous dirons qu'un nombre entier n est représenté par la forme q s'il existe un couple $(x,y) \in \mathbf{Z}^2$ tel que $q(x,y) = n$, et qu'il est *proprement représenté* par q s'il existe un tel couple formé d'entiers premiers entre eux.

Supposons que n soit représenté par q et soit $(x,y) \in \mathbf{Z}^2$ tel que $q(x,y) = n$. On a nécessairement $n \geqslant 0$ car q est définie positive ; de plus, si $n = 0$, il vient $x = y = 0$. On supposera donc dans la suite que $n > 0$ et $(x,y) \neq (0,0)$. Soit d le pgcd de x et y et posons $x' = x/d$, $y' = y/d$. Alors, $n = q(x,y) = d^2 q(x',y')$ et n/d^2 est un entier qui est proprement représenté par q.

Il se trouve que la propriété, pour un entier n, d'être proprement représenté par une forme q est est plus maniable que la notion naïve ; les calculs précédents montrent qu'il n'y a pas grand dommage à se cantonner à cette notion.

THÉORÈME 3.1.1. — *Soit D un entier < 0 et soit n un nombre entier strictement positif. Pour qu'il existe une forme quadratique binaire, définie positive, à coefficients entiers et de discriminant D qui représente proprement n, il faut et il suffit que D soit un carré dans $\mathbf{Z}/4n\mathbf{Z}$.*

Le discriminant D d'une forme entière est congru à 0 ou 1 modulo 4 ; cette congruence est bien vérifiée si n est un entier tel que D soit un carré dans $\mathbf{Z}/n\mathbf{Z}$.

Démonstration. — Soit q une telle forme quadratique et soit (ξ,η) un couple d'entiers premiers entre eux tels que $q(\xi,\eta) = n$. Puisque ξ et η sont premiers entre eux, on peut effectuer un changement de coordonnées dans $\mathrm{SL}_2(\mathbf{Z})$ et supposer que $(\xi,\eta) = (1,0)$. Alors, q est de la forme $q(x,y) = ax^2 + bxy + cy^2$, et $a = q(1,0) = n$. Le discriminant de q est donné par la formule $D = b^2 - 4ac$; autrement dit, $b^2 = D + 4nc \equiv D \pmod{4n}$.

Inversement, supposons que D soit un carré dans $\mathbf{Z}/4n\mathbf{Z}$ et choisissons des entiers b et c tels que $D = b^2 - 4nc$; alors, la forme quadratique $q(x,y) = nx^2 + bxy + cy^2$ est entière, de discriminant D, définie positive et représente proprement n puisque $q(1,0) = n$. □

Pour que ce théorème soit réellement utile, il faut ramener l'ensemble des formes q dont le théorème affirme l'existence à un ensemble contrôlable. Pour cela, la première remarque est que deux formes équivalentes représentent proprement les mêmes entiers. Par suite, dans le théorème précédent, on peut rajouter la condition que la forme quadratique soit réduite.

Enfin, on a la propriété fondamentale :

PROPOSITION 3.1.2. — *Pour tout entier $D < 0$, l'ensemble des classes d'équivalence de formes quadratiques binaires, entières et de discriminant D, est fini.*

Démonstration. — Puisque toute forme est équivalente à une forme réduite, il suffit de démontrer que l'équation $D = b^2 - 4ac$ n'a qu'un nombre fini de solutions $(a,b,c) \in \mathbf{Z}^3$ vérifiant les conditions $|b| \leqslant a \leqslant c$ qui assurent que la forme quadratique $ax^2 + 2bxy + cy^2$ est réduite.

On a déjà vu que ces relations entraînent l'inégalité $|b| \leqslant \sqrt{D/3}$; il n'y a donc qu'un nombre fini d'entiers b possibles. Alors, b étant fixé, l'équation $4ac = -D + b^2$ n'a elle aussi qu'un nombre fini de solutions (a,c) (majoré par le nombre de diviseurs de $-D + b^2$), *a fortiori* qu'un nombre fini de solutions telles que $|b| \leqslant a \leqslant c$. □

Au moins lorsque $-D$ n'est pas trop grand, il n'est pas difficile d'exhiber la liste des formes réduites de discriminant D. Traitons par exemple les cas $D = -3$, $D = -4$, -8, -20 et -23.

Exemple 3.1.3 (Formes réduites de discriminant $D = -3$). — L'inégalité $|b| \leqslant \sqrt{D/3}$ implique $b = 0$ ou $b = \pm 1$, mais $b = 0$ ne convient pas car une forme de discriminant -3 est impaire. L'équation pour (a, c) est alors $4ac = 4$, d'où $ac = 1$ et l'inégalité $0 < a \leqslant c$ entraîne $a = c = 1$. Par suite, $q(x, y) = x^2 \pm xy + y^2$. Par conséquent, la forme $x^2 + xy + y^2$ est la seule forme réduite de discriminant -3.

Exemple 3.1.4 (Formes réduites de discriminant $D = -4$). — Soit $q = ax^2 + bxy + cy^2$ une forme réduite de discriminant -4 ; alors q est paire, donc b est pair. L'inégalité $|b| \leqslant \sqrt{D/3}$ implique $|b| \leqslant 1$, d'où $b = 0$. Ensuite, $4ac = 4$, d'où $a = c = 1$. Il n'y a donc qu'une forme réduite de discriminant 1, la forme $x^2 + y^2$.

Par suite, cette forme représente un nombre premier p si et seulement si -1 est un carré modulo p, c'est-à-dire $p = 2$ ou $p \equiv 1 \pmod 4$. On retrouve un résultat déjà démontré.

Exemple 3.1.5 (Formes réduites de discriminant $D = -20$). — Soit $q = ax^2 + bxy + cy^2$ une forme réduite entière de discriminant -20 ; elle est paire. L'inégalité $|b| \leqslant \sqrt{D/3}$ implique $b \in \{0, -2, 2\}$. Pour $b = 0$, il vient $ac = D = 5$, d'où $a = 1$ et $c = 5$; pour $b = \pm 2$, on trouve $4ac = 20 + 4 = 24$, donc $ac = 6$ et, compte tenu de l'inégalité $|b| \leqslant a \leqslant c$, la seule solution $a = 2$ et $c = 3$; come $|b| = a$, seule la solution $b = 2$ convient. Les formes réduites de discriminant -20 sont donc les formes $x^2 + 5y^2$ et $2x^2 + 2xy + 3y^2$.

Exemple 3.1.6 (Formes réduites de discriminant $D = -23$). — Enfin, soit $q = ax^2 + bxy + cy^2$ une forme réduite de discriminant -23 ; c'est une forme impaire. L'inégalité $|b| \leqslant \sqrt{D/3}$ entraîne $|b| \leqslant 2$, donc $b = \pm 1$. Il vient alors $4ac = 24$, donc $ac = 6$ puis, comme $a \leqslant c$, $(a, c) = (1, 6)$ ou $(a, c) = (2, 3)$. Là première solution impose en outre $b = 1$; les formes réduites de discriminant -23 sont donc les trois formes $x^2 + xy + 6y^2$, $2x^2 + xy + 3y^2$ et $2x^2 - xy + 3y^2$.

Exemple 3.1.7 (Formes réduites de discriminant $D = -19$). — Supposons enfin que l'on ait $D = -19$. L'inégalité $|b| \leqslant \sqrt{D/3}$ entraîne $|b| \leqslant 2$, donc $b = \pm 1$. Lorsque $b = \pm 1$, on a $4ac = 20$ soit $ac = 5$ puis $a = 1$ et $c = 5$. Finalement la seule forme réduite de discriminant -19 est la forme principale $x^2 + xy + 5y^2$.

Exemple 3.1.8 (Formes réduites de discriminant $D = -76$). — Supposons enfin que l'on ait $D = -76$. L'inégalité $|b| \leqslant \sqrt{D/3}$ entraîne $|b| \leqslant 5$, donc $b = 0$ ou $b = \pm 2$. Si $b = 0$, il vient $4ac = 76$, d'où $ac = 19$ et $(a, c) = (1, 19)$. Si $b = \pm 2$, on a $4ac = 80$, donc $ac = 20$ et, compte tenu de l'inégalité $|b| \leqslant a \leqslant c$, les solutions $(a, c) = (2, 10)$ et $(a, c) = (4, 5)$. Les formes réduites de discriminant 19 sont ainsi les formes $x^2 + 19y^2$, $2x^2 + 2xy + 10y^2$ et $4x^2 \pm 2xy + 5y^2$.

Pour tout entier $D < 0$, notons $h(D)$ le nombre de formes quadratiques entières, réduites et de discriminant D.

Exercices

1) a) Démontrer la relation

$$(x^2+y^2)(X^2+Y^2) = (xX - Yy)^2 + (xY + yX)^2.$$

b) En déduire que l'ensemble S des entiers non nuls qui s'écrivent sous la forme x^2+y^2, avec $(x,y) \in \mathbf{Z}^2$, est stable par multiplication.

c) Si $n \in S$, démontrer que $n > 0$ et que pour tout nombre premier p congru à 3 modulo 4, $\mathrm{ord}_p(n)$ est pair.

d) Inversement, si n est un entier strictement positif tel que pour tout nombre premier p, $\mathrm{ord}_p(n)$ soit pair si $p \equiv 3 \pmod 4$, démontrer que $n \in S$.

2) Adapter l'exercice 1 et déterminer l'ensemble des entiers > 0 qui s'écrivent sous la forme $x^2 + 2y^2$.

3) Adapter l'exercice 1 et déterminer l'ensemble des entiers > 0 qui s'écrivent sous la forme $x^2 + 3y^2$.

4) Soit p un nombre premier congru à 1 modulo 4 et soit a un entier tel que $a^2 \equiv -1 \pmod p$.

a) Démontrer qu'il existe deux couples distincts (x,y) et (x',y') d'entiers compris entre 0 et $\sqrt{p}$ tels que $x - ay \equiv x' - ay' \pmod p$.

b) En déduire qu'il existe des entiers u et v vérifiant $|u|, |v| \leqslant \sqrt{p}$ tels que $u \equiv av \pmod p$.

c) Démontrer que $u^2 + v^2 = p$.

5) (Cet exercice fait suite à l'exercice 4.) Soit n un entier strictement positif.

a) Combien l'équation $a^2 \equiv -1 \pmod n$ a-t-elle de solutions ?

b) Soit a un entier tel que $a^2 \equiv -1 \pmod n$. Démontrer qu'il existe des entiers u et v tels que $|u|, |v| \leqslant \sqrt{n}$ et $v \equiv au \pmod n$. En déduire que $u^2 + v^2 = n$.

c) Combien y a-t-il de couples (u,v) d'entiers premiers entre eux tels que $n = u^2 + v^2$? Et si l'on enlève la condition de coprimalité ?

6) Soit p un nombre premier distinct de 5.

a) Écrire la congruence modulo 20 sur p qui équivaut à ce que -5 soit un carré modulo p.

b) Si p s'écrit x^2+5y^2, démontrer que $p \equiv 1,9 \pmod{20}$. Si $2p$ s'écrit x^2+5y^2, démontrer que $p \equiv 3,7 \pmod{20}$.

c) On suppose que $p \equiv 1,3,7,9 \pmod{20}$. Soit $a \in \mathbf{Z}$ tel que $a^2 \equiv -5 \pmod p$ et soit $q(x,y)$ la forme quadratique $(ay + px)^2 + 5y^2$. Considérer un couple $(x,y) \in \mathbf{Z}^2$, non nul, où q soit minimale et vérifier que $q(x,y) = p$ ou $2p$.

d) En déduire que les conditions nécessaires de la question b) pour que p (resp. $2p$) s'écrive $x^2 + 5y^2$ sont aussi suffisantes.

7) Adapter l'exercice 1 et déterminer l'ensemble des entiers > 0 qui s'écrivent sous la forme $x^2 + 5y^2$. (On utilisera aussi le résultat de l'exercice 6.)

8) La démonstration que nous avons donnée du fait que toute forme est équivalente à une forme réduite requiert de savoir en déterminer le minimum. Voici une autre méthode, de nature algorithmique, qui fonctionne l'on se restreint aux formes à coefficients entiers.

a) Soit $q(x,y) = ax^2 + bxy + cy^2$ une forme entière. (On ne suppose pas qu'elle soit définie positive.)

Si $|b| > |a|$, trouver un changement de variables de la forme $x' = x + uy$, $y' = y$ qui transforme q en une forme $q'(x,y) = a'x^2 + b'xy + c'y^2$ pour laquelle $|b'| < |b|$.

b) De même, si $|b| > c$, trouver un changement de variables qui fait diminuer b.

c) En déduire par récurrence que toute forme entière est équivalente à une forme entière $q = ax^2 + bxy + cy^2$ pour laquelle $|b| \leqslant |a| \leqslant |c|$.

Lorsque q est définie positive, on retrouve (essentiellement) la notion de forme réduite.

d) Calculer la forme réduite équivalente à la forme $11x^2 + 11x + 4$. Même question avec la forme $6x^2 + 7xy + 3y^2$.

e) Démontrer que l'ensemble des classes d'équivalence de formes entières de discriminant donné est fini.

9) Soit $q(x,y) = ax^2 + 2bxy + cy^2$ une forme quadratique binaire, à coefficients réels, définie positive. Soit τ l'unique racine de partie imaginaire > 0 de l'équation $q(1,\tau) = 0$.

a) Montrer que l'action de $\mathrm{SL}_2(\mathbf{Z})$ par changement de variables linéaire sur q correspond à l'action de $\mathrm{SL}_2(\mathbf{Z})$ par homographies sur τ.

b) En quelles inégalités sur τ se traduit le fait que q soit réduite ?

c) Démontrer qu'il existe une matrice $g \in \mathrm{SL}_2(\mathbf{Z})$ telle que $g(\tau)$ vérifie $\left|g(\tau)\right| \geqslant 1$ et $|\Re\tau| \leqslant 1/2$ (*domaine fondamental du demi-plan* $\Im\tau > 0$ *sous l'action des homographies de* $\mathrm{SL}_2(\mathbf{Z})$).

10) Soit $q(x,y) = ax^2 + bxy + cy^2$ une forme quadratique binaire à coefficients entiers dont le discriminant $D = b^2 - 4ac$ est strictement positif mais n'est pas un carré.

a) On pose $m = \inf_{(x,y)\in\mathbf{Z}^2\setminus\{0\}} \left|q(x,y)\right|$. Démontrer que $m > 0$ et qu'il existe un vecteur $(x,y) \in \mathbf{Z}^2$ tel que $\left|q(x,y)\right| = m$.

b) Démontrer que $m \leqslant \sqrt{D}/2$.

c) Si $m = \sqrt{D}/2$, démontrer que q est équivalente (par un changement de variables dans $\mathrm{SL}_2(\mathbf{Z})$) à la forme $m(x^2 - y^2)$.

11) Soit $q(x,y) = ax^2 + bxy + cy^2$ une forme quadratique binaire à coefficients entiers dont le discriminant $D = b^2 - 4ac$ est strictement positif mais n'est pas un carré.

a) Soit m la borne inférieure des $q(x,y)$, où (x,y) parcourt les couples non nuls de $\mathbf{Z}^2$ tels que $q(x,y) > 0$.

Démontrer que $m > 0$ et qu'il existe un vecteur $(x,y) \in \mathbf{Z}^2$ tel que $q(x,y) = m$.

b) Démontrer que $m \leqslant \sqrt{D}$.

§3.2. Sous-groupes discrets de $\mathbf{R}^n$, groupes abéliens de type fini

A. Forme de Hermite d'une matrice à coefficients entiers

DÉFINITION 3.2.1. — *Soit* $A \in \mathrm{Mat}_{m,n}(\mathbf{Z})$ *une matrice* $m \times n$ *à coefficients entiers. On dit que* A *est sous* forme normale de Hermite *s'il existe un entier* r *tel que* $1 \leqslant r \leqslant n$ *et une suite* $(m_1, \ldots, m_r)$ *d'entiers tels que* $1 \leqslant m_1 < \cdots < m_r \leqslant m$ *de sorte que les coefficients* $(a_{i,j})$ *de* A *vérifient les propriétés suivantes :*

> *a) Pour tout entier* j *tel que* $r < j \leqslant n$ *et tout entier* $i \in \{1, \ldots, m\}$, *on a* $a_{i,j} = 0$ *(seules les* r *premières colonnes de* A *sont non nulles) ;*

b) Pour tout entier j tel que $1 \leqslant j \leqslant r$ et tout entier i tel que $1 \leqslant i < m_j$, on a $a_{i,j} = 0$ (les r premières colonnes sont échelonnées) ;

c) Pour tout entier j tel que $1 \leqslant j \leqslant r$, on a $a_{m_j,j} > 0$ (les premiers coefficients non nuls de chaque colonne, appelés pivots*, sont strictement positifs) ;*

d) Pour tout entier j tel que $1 \leqslant j \leqslant r$ et tout entier k tel que $1 \leqslant k < m_j$, on a $0 \leqslant a_{m_j,k} < a_{m_j,j}$ (les coefficients de la ligne de pivot m_j sont positifs ou nuls et sont strictement inférieurs au pivot $a_{m_j,j}$).

Cette notion, introduite par Hermite en 1851, tire son intérêt du théorème suivant.

THÉORÈME 3.2.2. — *Soit A une matrice $m \times n$ à coefficients entiers. Il existe une unique matrice H de même taille, sous forme normale de Hermite et une matrice $U \in \mathrm{GL}_n(\mathbf{Z})$ telle que $A = HU$.*

La matrice H est appelée *forme normale de Hermite* de la matrice A.

Démonstration. — L'existence d'un tel couple (H,U) se démontre par récurrence, en effectuant sur les colonnes de la matrice A une successions d'*opérations élémentaires*. Introduisons quelques notations.

Si (i,j) est un couple d'entiers distincts dans $\{1,\dots,n\}$ et $a \in \mathbf{Z}$, on note $E_{i,j}(a)$ la matrice $n \times n$ dont la diagonale est formée de 1 et tous les autres coefficients sont nuls sauf celui de coordonnées (i,j) qui est égal à a. Elle appartient à $\mathrm{SL}_n(\mathbf{Z})$, son inverse étant $E_{i,j}(-a)$. La méthode va consister à partir du couple $(H,U) = (A, \mathrm{I}_n)$, formé d'une matrice entière et d'une matrice de $\mathrm{SL}_n(\mathbf{Z})$, et à multiplier à droite H par une matrice $E_{i,j}(a)$; l'égalité évidente $HU = (HE_{i,j}(a))(E_{i,j}(-a)U)$ montre comment obtenir un autre couple de même même nature et de même produit.

D'autre part, pour tout entier i, on note S_i la matrice diagonale $n \times n$ dont tous les coefficients diagonaux sont égaux à 1 sauf celui de coordonnées (i,i) qui est égal à -1.

Si H est une matrice ayant n colonnes, la matrice $HE_{i,j}(a)$ est obtenue à partir de H en ajoutant à la colonne C_j de H la colonne C_i multipliée par a (en abrégé, $C_j \leftarrow C_j + aC_i$). La matrice HS_i est obtenue en multipliant la colonne i par -1 (ce qu'on note $C_i \leftarrow -C_i$).

Si U est une matrice ayant n lignes, la matrice $E_{i,j}(-a)U$ est obtenue à partir de U en ajoutant à la ligne L_i la ligne L_j multipliée par $(-a)$ (en abrégé, $L_i \leftarrow L_i - aL_j$) ; la matrice S_iU se déduit de U en multipliant la ligne i par -1 (noté $L_i \leftarrow -L_i$).

Cela montre que l'on passe d'un couple (H,U) au suivant par des opérations élémentaires sur les colonnes de H et sur les lignes de U.

Commençons par démontrer qu'il existe une matrice $H \in \mathrm{Mat}_{m,n}(\mathbf{Z})$, échelonnée et à pivots positifs, et un produit de matrices élémentaires $U \in \mathrm{GL}_n(\mathbf{Z})$ tels que $A = HU$. On procède par récurrence sur la taille $m+n$ de A.

Démontrons alors par récurrence sur la somme des valeurs absolue des coefficients de la première ligne de A que l'on peut supposer que $a_{1,1} \geqslant 0$ et $a_{1,j} = 0$ si $2 \leqslant j \leqslant n$.

Si cette propriété n'est pas vérifiée, notons j l'indice de colonne du coefficient non nul de la première ligne qui est de plus petite valeur absolue. Pour $k \neq j$, l'opération $C_k \leftarrow C_k - qC_j$, sur les colonnes de A, où $q = \lfloor a_{1,k}/a_{1,j} \rfloor$ est la partie entière du quotient de $a_{1,k}$ par $a_{1,j}$ ramène la matrice A à une matrice dont le coefficient $a_{1,k}$ vérifie $0 \leqslant a_{1,k} < |a_{1,j}|$. Par récurrence, on se ramène ainsi au cas où $a_{1,j}$ est le seul coefficient non nul de la première ligne ; notons ε son signe. Si $j \neq 1$, les opérations $C_1 \leftarrow C_1 + \varepsilon C_j$ puis $C_j \leftarrow C_j - \varepsilon C_1$ ramènent au cas où $j = 1$ et $a_{1,j} > 0$. Supposons $j = 1$ et $\varepsilon = -1$. Si $n \geqslant 2$, on effectue successivement les opérations $C_2 \leftarrow C_2 + C_1$, $C_1 \leftarrow C_1 - C_2$ qui nous ramènent au cas $j = 2$. Si $n = 1$, la seule possibilité est de multiplier la première colonne par -1.

Nous sommes donc réduits à traiter le cas où, à l'exception du coefficient $a_{1,1}$ qui est positif ou nul, tous les coefficients la première ligne de A sont nuls, on écrit $A = \begin{pmatrix} a_{1,1} & 0 \\ * & A' \end{pmatrix}$ puis $A' = H'U'$, où $H' \in \mathrm{Mat}_{m-1,n-1}(\mathbf{Z})$ est échelonnée à pivots positifs. Posons $U = \begin{pmatrix} 1 & 0 \\ 0 & U' \end{pmatrix}$; c'est un produit de matrices élémentaires : le même que U' où les indices de lignes et de colonnes sont augmentés de 1 et un calcul par blocs montre que la matrice $H = AU^{-1}$ est échelonnée à pivots positifs.

Il reste à mettre sous forme normale de Hermite une matrice échelonnée H autrement dit à faire en sorte que sur la ligne d'un pivot, tous les coefficients soient inférieurs au pivot.

Supposons cette propriété satisfaite pour les lignes des pivots d'indices de colonnes $> j$ et soit m_j l'indice de ligne du pivot de colonne j. Pour tout entier i tel que $1 \leqslant i < j$, remplaçons la colonne C_k de H par la colonne $C_k - qC_j$, où $q = \lfloor a_{k,j}/a_{m_j,j} \rfloor$ est la partie entière du quotient de $a_{k,j}$ par $a_{m_j,j}$. Tous les coefficients de la ligne m_j appartiennent alors à l'intervalle $[0, \ldots, a_{m_j,j} - 1]$ et ceux des lignes des pivots ultérieurs ($a_{m_t,k}$, pour tout entier t tel que $j < t \leqslant r$ et tout entier k tel que $1 \leqslant k < m_t$) n'ont pas été modifiés puisque, la matrice H étant échelonnée, $a_{m_t,j} = 0$ si $t > j$.

Cela conclut, par récurrence, la démonstration qu'il existe une matrice $U \in \mathrm{GL}_n(\mathbf{Z})$ et une matrice sous forme normale de Hermite $H \in \mathrm{Mat}_{m,n}(\mathbf{Z})$ telles que $A = HU$.

Démontrons maintenant que si (H, U) et (H', U') sont deux couples de matrices tels que $HU = H'U'$, où H et $H' \in \mathrm{Mat}_{m,n}(\mathbf{Z})$ sont sous forme normale de Hermite et $U, U' \in \mathrm{GL}_n(\mathbf{Z})$, alors $H = H'$. Quitte à remplacer U par $U(U')^{-1}$, on suppose que $H' = HU$ et l'on doit démontrer que $H = H'$.

Notons r le nombre de pivots de H et $m_1, \ldots, m_r$ les indices de lignes correspondants. Notons $(a_{i,j})$, $(a'_{i,j})$ et $(u_{i,j})$ les coefficients des matrices H, H' et U ; notons aussi $L_1, \ldots, L_n$ les lignes de la matrice U. Si $m_t \leqslant i < m_{t+1}$, la t-ième ligne de la matrice HU est donnée par $L'_i = a_{i,1}L_1 + \cdots + a_{i,t}L_t$. Puisque $H' = HU$ est échelonnée, il vient déjà que les lignes $L_1, \ldots, L_r$ de U le sont aussi : $u_{i,j} = 0$ si $1 \leqslant i \leqslant r$ et $j > i$.

De plus, la positivité de $a_{m_t,t}$ et $a'_{m_t,t}$ implique que les r premiers coefficients diagonaux de U sont positifs ou nuls. On peut donc écrire $U = \begin{pmatrix} U_1 & 0 \\ U_3 & U_4 \end{pmatrix}$, où U_1 est triangulaire inférieure. Démontrons que $U_1 = \mathrm{I}_r$.

On a pour commencer l'égalité $1 = \det(U) = \det(U_1)\det(U_4)$ si bien que $\det(U_1) = \pm 1$. Puisque $\det(U_1)$ est égal au produit des coefficients diagonaux de U_1, ceux-ci sont égaux à 1. Alors, la forme de la matrice HU montre que les pivots de $H' = HU$ sont aux mêmes positions que ceux de H, chacun étant égal au pivot de H correspondant. Nous allons démontrer par récurrence sur $t \in \{1, \dots, r\}$ que $L_t = (0, \dots, 0, 1, 0, \dots, 0)$, le 1 étant en colonne t.

Supposons que ce soit vrai aux rangs $< t$ et démontrons-le au rang t. On a déjà prouvé que $u_{t,t} = 1$; considérons, s'il en existe, un entier j dans $\{1, \dots, t-1\}$ tel que $u_{t,j} \neq 0$ et maximal. (En particulier, $a'_{m_t,k} = a_{m_t,k}$ si $j < k \leqslant t$.) Par hypothèse, on a alors $a'_{m_t,j} = a_{m_t,t}u_{t,j} + a_{m_t,j}$. Comme $a_{m_t,t} = a'_{m_t,t}$, les inégalités $0 \leqslant a_{m_t,j} < a_{m_t,t}$ et $0 \leqslant a'_{m_t,j} < a'_{m_t,t}$ impliquent $\left|a_{m_t,j} - a'_{m_t,j}\right| < a_{m_t,t}$; puisque cette différence est un multiple entier de $a_{m_t,t}$, elle est nulle, d'où $u_{t,j} = 0$. Cela démontre que L_t est de la forme annoncée, puis, par récurrence sur t, que $U_1 = \mathrm{I}_r$.

Alors, $H' = HU = H$, ce qu'il fallait démontrer. □

Remarques 3.2.3. — a) L'entier r est égal au rang de la matrice A. Sauf si $r = n$, on n'a pas unicité de la matrice U.

b) Supposons $r = n$ et $A \in \mathrm{SL}_n(\mathbf{Z})$. Comme $1 = \det(A) = \det(H)$, la matrice H est de déterminant 1 ; comme elle est échelonnée, ses coefficients diagonaux sont donc égaux à 1 et les autres sont nécessairement nuls. Autrement dit, $H = \mathrm{I}_n$ et $A = U$. La démonstration a démontré que A est produit de matrices élémentaires de types $E_{i,j}(a)$ et S_i. L'usage des matrices S_i n'est en fait pas nécessaire : *le groupe* $\mathrm{SL}_n(\mathbf{Z})$ *est engendré par les matrices* $E_{i,j}(1)$. Notons $E \subset \mathrm{SL}_n(\mathbf{Z})$ le sous-groupe engendré par ces matrices ; il contient les matrices $E_{i,j}(a)$ pour tout $a \in \mathbf{Z}$. Pour tout k, $S_k E_{i,j}(a) S_k^{-1}$ est égal à $E_{i,j}(a)$ si $k \notin \{i, j\}$, à $E_{i,j}(-a)$ sinon ; par conséquent, les matrices S_k appartiennent au normalisateur de E. Dans une expression de U comme produit de matrices élémentaires, on peut de proche en proche les repousser à la fin. Comme $\det(A) = 1$, le nombre de telles matrices qui restent à la fin est pair et il suffit de montrer que le produit de deux d'entre elles appartient à E. Or, la suite de matrices 2×2,

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

où chacune se déduit de la précédente par une opération élémentaire du type $C_i \leftarrow C_i + aC_j$ démontre que c'est bien le cas, au moins si $n = 2$. Dans le cas général, il suffit d'effectuer les opérations correspondantes sur les colonnes concernées.

c) L'énoncé et sa démonstration s'étendent au cas d'un anneau euclidien R arbitraire. L'hypothèse de positivité sur les pivots doit être remplacée par l'hypothèse que les pivots appartiennent à un système de représentants d'éléments de R modulo l'action par multiplication des éléments inversibles de R ; quant à l'hypothèse que les coefficients $a_{m_j,k}$ (avec $1 \leqslant k < j$) d'une ligne de pivot sont positifs ou nuls et inférieurs au pivot, elle est remplacée par celle que $a_{m_j,k}$ appartient à un système de représentants des restes des divisions euclidiennes par $a_{m_j,j}$.

d) La démonstration donnée est algorithmique : elle fournit un moyen effectif de calculer des matrices H et U. Lorsque les coefficients sont réels, les erreurs d'arrondis la rendent peu praticable. La situation n'est guère meilleure lorsque A est à coefficients entiers. En effet, même si un calcul exact est alors possible, l'algorithme que nous avons décrit peut faire apparaître, au cours du calcul, des coefficients de taille démesurément grande comparée à celle des coefficients de A. Hafner et McHurley ont donné l'exemple d'une matrice 20×20 de coefficients entre 0 et 10 pour laquelle ce calcul requiert des nombres entiers de 5000 chiffres décimaux ! Signalons qu'il existe aujourd'hui des algorithmes efficaces.

B. Application aux groupes abéliens

LEMME 3.2.1. — *Soit $A \in \mathrm{Mat}_n(\mathbf{R})$. Pour que l'on ait $A(\mathbf{Z}^n) \subset \mathbf{Z}^n$, il faut et il suffit que $A \in \mathrm{Mat}_n(\mathbf{Z})$; pour que l'on ait $A(\mathbf{Z}^n) = \mathbf{Z}^n$, il faut et il suffit que l'on ait $A \in \mathrm{GL}_n(\mathbf{Z})$, c'est-à-dire que $A \in \mathrm{Mat}_n(\mathbf{Z})$ et $\det(A) \in \{\pm 1\}$.*

Démonstration. — Si $A(\mathbf{Z}^n) \subset \mathbf{Z}^n$, les colonnes de A, qui sont les images des vecteurs de base de $\mathbf{R}^n$, appartiennent à $\mathbf{Z}^n$; la réciproque est évidente.

Si $A \in \mathrm{GL}_n(\mathbf{Z})$, alors A et A^{-1} appartiennent à $\mathrm{Mat}_n(\mathbf{Z})$, d'où les deux inclusions $A(\mathbf{Z}^n) \subset \mathbf{Z}^n$ et $A^{-1}(\mathbf{Z}^n) \subset \mathbf{Z}^n$. Par suite, $A(\mathbf{Z}^n) = \mathbf{Z}^n$. La réciproque se démontre de même : si $A(\mathbf{Z}^n) = \mathbf{Z}^n$, on a déjà $A \in \mathrm{Mat}_n(\mathbf{Z})$; d'autre part, l'image de A contenant n vecteurs linéairement indépendants, $A(\mathbf{R}^n) = \mathbf{R}^n$ et A est aussi injective. Par hypothèse, les antécédents des éléments de $\mathbf{Z}^n$ appartiennent à $\mathbf{Z}^n$, donc $A^{-1}(\mathbf{Z}^n) \subset \mathbf{Z}^n$, puis $A^{-1} \in \mathrm{Mat}_n(\mathbf{Z})$.

Enfin, si $A \in \mathrm{GL}_n(\mathbf{Z})$, donc $A^{-1} \in \mathrm{Mat}_n(\mathbf{Z})$, on a $1 = \det(A \cdot A^{-1}) = \det(A)\det(A^{-1})$, d'où $\det(A) \in \{\pm 1\}$. Inversement, si $A \in \mathrm{Mat}_n(\mathbf{Z})$ vérifie $\det(A) \in \{\pm 1\}$, les formules de Cramer ($A^{-1} = (\det A)^{-1}\,{}^{\mathrm{t}}\mathrm{Com(A)}$) entraînent que A^{-1} appartient à $\mathrm{Mat}_n(\mathbf{Z})$, d'où $A \in \mathrm{GL}_n(\mathbf{Z})$. □

THÉORÈME 3.2.2. — *Soit G un sous-groupe abélien de $\mathbf{Z}^m$. Alors G est un groupe abélien libre de rang fini : il existe un entier $r \in \{0, \ldots, n\}$ et des vecteurs $u_1, \ldots, u_r \in \mathbf{R}^m$ appartenant à G tels que tout élément de G s'écrivent de manière unique sous la forme $\sum_{j=1}^{r} a_j u_j$, avec $(a_1, \ldots, a_r) \in \mathbf{Z}^r$.*

Démonstration. — Soit A la matrice de taille $m \times n$ dont les colonnes sont les vecteurs $v_1, \dots, v_n$, soit H sa forme normale de Hermite et soit $U \in \mathrm{GL}_n(\mathbf{Z})$ une matrice telle que $A = HU$. Par hypothèse, $G = A(\mathbf{Z}^n)$. Comme on a $U(\mathbf{Z}^n) = \mathbf{Z}^n$, il vient $G = H(\mathbf{Z}^n)$. Notons r le rang de H et $u_1, \dots, u_r$ ses r premières colonnes. Comme les autres colonnes sont nulles, on a aussi $G = (u_1 \dots u_r)(\mathbf{Z}^r)$: tout vecteur de G est combinaison linéaire à coefficients entiers de $u_1, \dots, u_r$. Comme ces r vecteurs sont échelonnés, ils sont linéairement indépendants dans $\mathbf{R}^m$. Autrement dit, tout vecteur de G est de manière unique combinaison linéaire à coefficients entiers de $u_1, \dots, u_r$, ce qui démontre que G est un groupe abélien libre, de base $(u_1, \dots, u_r)$. □

COROLLAIRE 3.2.3. — *Soit G un groupe abélien engendré par un nombre fini $\{g_1, \dots, g_n\}$ de ses éléments. Si G est sans torsion, alors G est un groupe abélien libre : il existe un entier $r \in \{0, \dots, n\}$ et des éléments $h_1, \dots, h_r$ de G tels que tout élément de G s'écrive de manière unique sous la forme $\sum_{i=1}^{r} a_i h_i$, avec $(a_1, \dots, a_r) \in \mathbf{Z}^r$.*

Démonstration. — □

Soit A une matrice de taille $m \times n$ à coefficients entiers. Pour tout entier r tel que $1 \leqslant r \leqslant \min(m, n)$, on définit $\mathscr{F}_r(A)$ comme le groupe abélien engendré par les mineurs $r \times r$ extraits de A.

Soit $U \in \mathrm{Mat}_n(\mathbf{Z})$ une matrice à coefficients entiers ; par multilinéarité du déterminant, on a $\mathscr{F}_r(AU) \subset \mathscr{F}_r(A)$. En particulier, si $U \in \mathrm{GL}_n(\mathbf{Z})$ est inversible, on a $\mathscr{F}_r(AU) = \mathscr{F}_r(A)$. De même, pour toute matrice $V \in \mathrm{Mat}_m(\mathbf{Z})$, on a $\mathscr{F}_r(VA) \subset \mathscr{F}_r(A)$, et $\mathscr{F}_r(VA) = \mathscr{F}_r(A)$ si V appartient à $\mathrm{GL}_m(\mathbf{Z})$.

En outre, le développement d'un mineur $r \times r$ suivant une ligne montre que $\mathscr{F}_r(A)$ est contenu dans $\mathscr{F}_{r-1}(A)$. Autrement dit, les groupes abéliens $\mathscr{F}_0(A) = \mathbf{Z}$, $\mathscr{F}_1(A), \dots$ obéissent aux inclusions $\mathbf{Z} \supset \mathscr{F}_1(A) \supset \dots$; ils sont appelés *idéaux de Fitting* de la matrice A. [1]

Si A est de rang r, les idéaux de Fitting $\mathscr{F}_s(A)$ sont nuls pour $s > r$; en outre, $\mathscr{F}_r(A) \neq 0$.

THÉORÈME 3.2.4. — *Soit $A \in \mathrm{Mat}_{m,n}(\mathbf{Z})$ avec $n \geqslant m$. Pour qu'il existe une matrice $B \in \mathrm{GL}_n(\mathbf{Z})$ dont A constitue les m premières lignes, il faut et il suffit que l'idéal de Fitting $\mathscr{F}_m(A)$ soit égal à $\mathbf{Z}$, c'est-à-dire que les mineurs $m \times m$ extraits de A soient premiers entre eux.*

Démonstration. — Soit $B \in \mathrm{Mat}_n(\mathbf{Z})$ une matrice carrée dont A constitue les m premières lignes et soit H la forme normale de Hermite de B et U une matrice de $\mathrm{GL}_n(\mathbf{Z})$

1. Les propriétés que nous avons démontrées sont valables, avec la même preuve, lorsque l'anneau $\mathbf{Z}$ est remplacé par un anneau commutatif arbitraire ; les $\mathscr{F}_r(A)$ sont alors définis comme les idéaux engendré par les mineurs $r \times r$ extraits de A, d'où la terminologie.

telle que $B = HU$. Soit $S \in \mathrm{Mat}_{m,n}(\mathbf{Z})$ la matrice formée des m premières lignes de H ; on a donc $A = SU$.

Si B appartient à $\mathrm{GL}_n(\mathbf{Z})$, la relation $\det(B) = \det(H)\det(U)$ entraîne que $\det(H) \in \{\pm 1\}$, donc $H \in \mathrm{GL}_n(\mathbf{Z})$. Comme H est échelonnée, ses coefficients diagonaux sont donc égaux à ± 1. En particulier, le mineur principal $m \times m$ de S est donc égal à ± 1. On a donc $\mathscr{F}_m(S) = \mathbf{Z}$, d'où $\mathscr{F}_m(A) = \mathscr{F}_m(SU) = \mathscr{F}_m(S) = \mathbf{Z}$.

Inversement, supposons que $\mathscr{F}_m(A) = \mathbf{Z}$; soit S la forme normale de Hermite de A et soit $U \in \mathrm{GL}_n(\mathbf{Z})$ telle que $A = SU$. Si $\mathscr{F}_m(A) = \mathbf{Z}$, il en est de même de $\mathscr{F}_m(S)$. Or, S est de rang m, échelonnée, donc $\mathscr{F}_m(S)$ est engendré par le mineur principal de S — les autres sont nuls ! Par conséquent, les coefficients diagonaux de S sont égaux à ± 1. Soit $H \in \mathrm{Mat}_n(\mathbf{Z})$ la matrice dont les m premières lignes sont celles de S et les $n - m$ dernières sont celles de la matrice I_n. C'est une matrice triangulaire supérieure, à coefficients diagonaux ± 1 ; elle appartient donc à $\mathrm{GL}_n(\mathbf{Z})$, de même que la matrice HU. Un calcul par bloc entraîne alors que les m premières lignes de HU sont celles de A, d'où le théorème. □

COROLLAIRE 3.2.5. — *Soit $u \in \mathbf{Z}^n$; pour qu'il existe une matrice $A \in \mathrm{GL}_n(\mathbf{Z})$ dont u soit le premier vecteur colonne, il faut et il suffit que les coordonnées de u soient premières entre elles.*

Démonstration. — Quitte à transposer les matrices en jeu, ce qui échange lignes et colonnes, c'est le cas $m = 1$ du théorème. □

C. Sous-groupes discrets de $\mathbf{R}^n$ et réseaux

Rappelons qu'on dit qu'un espace topologique X est discret si, pour tout point x de X, la partie $\{x\}$ est ouverte dans X. Cela revient à dire que toute partie de X est ouverte et fermée. Une partie S d'un espace topologique X est discrète si, pour la topologie induite, c'est un espace topologique discret ; cela revient à dire que pour tout point x de S, la partie $\{x\}$ de S est ouverte *dans* S, autrement dit, qu'il existe un ouvert U de X tel que $U \cap S = \{x\}$.

PROPOSITION 3.2.1. — *Soit G un sous-groupe de $\mathbf{R}^n$. Les conditions suivantes sont équivalentes :*

(i) L'espace G est discret dans $\mathbf{R}^n$.

(ii) pour $r > 0$ assez petit, l'intersection de G et de la boule de centre 0 et de rayon r est réduite au point 0 ;

(iii) pour tout R, l'intersection de G et de la boule de rayon R n'a qu'un nombre fini de points.

Démonstration. — L'implication (i)⇒(ii) résulte de la définition. Inversement, soit $r > 0$ tel que la boule de centre 0 et de rayon r dans $\mathbf{R}^n$ ne rencontre G qu'en 0. Soit $g \in G$;

pour tout point x de $G \cap B(g,r)$, $x-g$ appartient à $G \cap B(0,r)$, donc $x = g$ si bien que $G \cap B(g,r) = \{g\}$. Par suite, G est discret, d'où (i).

Avec ces notations, les boules de centres $g \in G$ et de rayon $r/2$ sont disjointes : si $x \in B(g,r/2) \cap B(g',r/2)$, alors $\|g'-g\| < r$, donc $g' \in B(g,r) \cap G$, d'où $g' = g$. Par conséquent, si $B(0,R)$ contient N points, la boule $B(0,R+r/2)$ contient N boules disjointes de rayon r. Comme elles ont même volume, on a donc $N\,\mathrm{vol}(B(0,r)) \leqslant \mathrm{vol}(B(0,R+r/2))$; en particulier, $N \neq +\infty$. Cela démontre (iii).

Inversement, supposons que $G \cap B(0,R)$ soit fini. Si $G = \{0\}$, alors G est discret. Sinon, soit g un élément non nul de G et posons $R = \|g\| + 1$. Comme $G \cap B(0,R)$ privé de 0 est fini (par hypothèse) et non vide (il contient g), l'ensemble des normes $\|g\|$ des points $g \neq 0$ de $G \cap B(0,R)$ admet un plus petit élément r. On a donc $G \cap B(0,r) = \{0\}$. Cela démontre (i). □

DÉFINITION 3.2.2. — *Soit V un espace vectoriel réel de dimension finie. On appelle* réseau *de V tout sous-groupe de V formé des combinaisons linéaires à coefficients entiers d'une famille libre de V.*

On appelle *rang* d'un réseau la dimension de l'espace vectoriel réel qu'il engendre.

Soit G un réseau d'un espace vectoriel de dimension finie V. Soit $(v_1,\dots,v_r)$ une famille libre de V qui engendre G comme sous-groupe de V. Alors, $G \subset \mathrm{vect}(v_1,\dots,v_r)$ donc $\mathrm{vect}(G) \subset \mathrm{vect}(v_1,\dots,v_r)$; inversement, comme G contient $v_1,\dots,v_r$, $\mathrm{vect}(G)$ contient $\mathrm{vect}(v_1,\dots,v_r)$. On a donc $\mathrm{vect}(G) = \mathrm{vect}(v_1,\dots,v_r)$ et $\dim \mathrm{vect}(G) = r$. Cela démontre que le rang d'un réseau est égal au cardinal de toute famille libre qui engendre ce réseau.

THÉORÈME 3.2.3. — *Soit G un sous-groupe discret de $\mathbf{R}^n$. Alors, G est un réseau de $\mathbf{R}^n$. Il existe un entier $r \in \{0,\dots,n\}$ et une famille libre $(g_1,\dots,g_r)$ de $\mathbf{R}^n$ qui est une $\mathbf{Z}$-base de G.*

Inversement, observons que le groupe abélien engendré par une famille libre de $\mathbf{R}^n$ est un sous-groupe discret : quitte à effectuer un changement de base, on peut supposer que cette famille libre est formée des r premiers vecteurs de la base canonique de $\mathbf{R}^n$; alors, $G = \mathbf{Z}^r \times \{0\}^{n-r}$ est un sous-groupe discret de $\mathbf{R}^n$.

Nous allons en fait démontrer un résultat plus précis qui fournit un procédé constructif de base.

THÉORÈME 3.2.4. — *Soit G un sous-groupe discret de $\mathbf{R}^n$, soit V le sous-espace vectoriel réel de $\mathbf{R}^n$ engendré par G et soit $(g_1,\dots,g_r)$ une base de V formée d'éléments de G. Pour $i \in \{0,\dots,r\}$, posons $V_i = \mathrm{vect}(g_1,\dots,g_i)$.*

Pour tout $i \in \{1,\dots,r\}$, il existe un élément $v_i \in V_i \cap G$ tel que l'on ait $\|v-w\| \geqslant \|v_i - w'\|$ pour tous $w, w' \in V_{i-1}$ et tout $v \in G \cap (V_i \setminus V_{i-1})$. Alors, $(v_1,\dots,v_r)$ est une base

de V formée d'éléments de G et tout élément de G est combinaison linéaire de $v_1, \dots, v_r$. En particulier, G est un réseau de rang r.

Démonstration. — On va démontrer par récurrence sur i l'existence de v_i et le fait que $V_i \cap G$ (qui est manifestement un sous-groupe discret de $\mathbf{R}^n$ contenu dans V_i) est un réseau de base $(v_1, \dots, v_i)$.

Pour $i = 0$, on a $V_0 = 0$, donc $V_0 \cap G = \{0\}$ est un réseau de $\mathbf{R}^n$ de base la famille vide.

Soit alors $i \in \{1, \dots, r\}$; supposons construite une base $(v_1, \dots, v_{i-1})$ de V_{i-1} telle que $V_{i-1} \cap G$ soit un réseau de base $(v_1, \dots, v_{i-1})$. Démontrons l'existence d'un vecteur $v_i \in G \cap V_i \setminus V_{i-1}$ tel que $\|v_i - w\| \leqslant \|v - w'\|$ pour tous $w, w' \in V_{i-1}$ et tout $v \in G \cap V_i \setminus V_{i-1}$.

Soit (g_n) une suite d'éléments de $G \cap V_i \setminus V_{i-1}$ et (w_n) une suite d'éléments de V_{i-1} telles que $\|g_n - w_n\|$ tende vers la borne inférieure m des $\|g - w\|$ pour $g \in G \cap V_i \setminus V_{i-1}$ et $w \in V_{i-1}$.

Décomposons w_n dans la base $(v_1, \dots, v_{i-1})$ de V_{i-1}, $w_n = \sum_{k<i} x_{k,n} v_k$. Posons $h_n = \sum_{k<i} \lfloor x_{k,n} \rfloor v_k$. Quitte à remplacer g_n par $g_n - h_n$ et w_n par $w_n - h_n$, on peut supposer que $0 \leqslant x_{k,n} < 1$ pour tous k, n. La suite (w_n) est donc bornée, de même que la suite (g_n) puisque $\|g_n\| \leqslant \|g_n - w_n\| + \|w_n\|$ et que $\|g_n - w_n\|$ tend vers m. Comme les g_n appartiennent à G, qui est un sous-groupe discret de $\mathbf{R}^n$, ils ne peuvent donc prendre qu'un nombre fini de valeurs. Quitte à considérer une sous-suite de la suite (g_n), on peut supposer qu'elle est constante. Notons g sa valeur et démontrons que l'on peut poser $v_i = g$. Il est clair que $g \in G \cap V_i$ mais que $g \notin V_{i-1}$ car c'est le cas de chaque g_n.

Par compacité de l'intervalle $[0, 1]$, on peut, quitte à en considérer successivement des sous-suites, supposer que les $x_{k,n}$ convergent tous ; alors, la suite (w_n) vers un élément w de V_{i-1}. On a donc $\|g - w_n\| \to m$, donc $\|g - w\| \geqslant m$, donc l'égalité $\|g - w\| = m$ par définition de m, ce qui démontre l'existence de v_i.

Comme $v_i \notin V_i \setminus V_{i-1}$, on a

$$\dim V_i \geqslant \dim \operatorname{vect}(v_1, \dots, v_i) \geqslant \dim \operatorname{vect}(v_1, \dots, v_{i-1}) + 1 \geqslant \dim V_{i-1} + 1;$$

comme, par définition, $\dim V_i = \dim V_{i-1} + 1$, chacune de ces inégalités est une égalité et $(v_1, \dots, v_i)$ est une base de V_i.

Soit maintenant $v \in V_i$ et faisons l'observation suivante concernant $\|v - w\|$ lorsque w parcourt V_{i-1}. Notons $x_1, \dots, x_i$ les coordonnées de v dans la base $(g_1, \dots, g_i)$, de sorte que $v = x_1 g_1 + \cdots + x_i g_i$; démontrons que $d(v, V_{i-1}) = |x_i|\, d(g_i, V_{i-1})$. En effet, pour tout $w \in V_{i-1}$, on a $\|v - w\| = \|x_i g_i - w'\|$, avec $w' = w - \sum_{k<i} x_k g_k$. Ainsi, $w' \in V_{i-1}$, de même que w'/x_i si $x_i \neq 0$, si bien que $\|v - w\| \geqslant |x_i|\, d(g_i, V_{i-1})$. L'autre inégalité se démontre de façon similaire : si (w_n) est une suite de vecteurs de V_{i-1} telle que $\|g_i - w_n\|$ converge vers $d(g_i, V_{i-1})$, alors $\|v - w'_n\|$ converge vers $|x_i|\, d(g_i, V_{i-1})$, où w'_n est défini par $w'_n = x_i w_n + \sum_{k<n} x_k g_k$. Par conséquent, $|x_i|\, d(g_i, V_{i-1}) \geqslant d(v, V_{i-1})$.

Supposons de plus que v appartienne aussi à G et démontrons qu'il est combinaison linéaire à coefficients entiers de $v_1, \dots, v_i$. Notons $\alpha_1, \dots, \alpha_i$ les coordonnées de v_i dans

la base $(g_1,\dots,g_i)$, de sorte que $v_i = \alpha_1 g_1 + \cdots + \alpha_i g_i$. Comme $v_i \notin V_{i-1}$, $\alpha_i \neq 0$; posons $q = \lfloor x_i/\alpha_i \rfloor$ et $v' = v - qv_i = \sum x'_k g_k$ avec $x'_k = x_k - q\alpha_k$. Alors, $v' \in V_i \cap G$ et

$$d(v', V_{i-1}) = |x'_i|\, d(g_i, V_{i-1}) = \frac{x'_i}{\alpha_i} d(v_i, V_{i-1}).$$

Puisque $0 \leqslant x'_i < \alpha_i$ et $d(v_i, V_{i-1}) > 0$, on a $d(v', V_{i-1}) < d(v_i, V_{i-1})$. Par définition de v_i, cela entraîne $v' \in V_{i-1}$. Par récurrence, v' est une combinaison linéaire à coefficients entiers de $v_1,\dots,v_{i-1}$: soit $q_1,\dots,q_{i-1}$ des entiers tels que $v' = q_1 v_1 + \cdots + q_{i-1} v_{i-1}$. Posons aussi $q_i = q$. On a donc $v = \sum_{k\leqslant i} q_k v_k$, d'où la conclusion voulue. □

COROLLAIRE 3.2.5. — *Soit G un sous-groupe discret de $\mathbf{R}^n$. Pour que G soit un réseau de $\mathbf{R}^n$ de rang n, il faut et il suffit qu'il existe une partie bornée P de $\mathbf{R}^n$ telle que $G + P = \mathbf{R}^n$ (autrement dit, que G soit* cocompact*).*

Démonstration. — Soit $(g_1,\dots,g_r)$ une base de G.

Si G est un réseau de rang n, on a $r = n$ et on peut prendre pour P le parallélépipède de côtés $g_1,\dots,g_n$, c'est-à-dire l'ensemble des $\sum_{i=1}^n x_i g_i$, avec $x_i \in [0,1]$ pour tout i.

Inversement, soit P une partie compacte de $\mathbf{R}^n$ telle que $G + P = \mathbf{R}^n$. Soit v un vecteur quelconque de $\mathbf{R}^n$. Pour tout entier $m \geqslant 1$, on peut écrire $mv = g_m + p_m$, où $g_m \in G$ et $p_m \in P$; comme P est borné, on en déduit que $v = \frac{1}{m} g_m + \mathrm{O}(1/m)$ lorsque m tend vers $+\infty$. En outre, $\frac{1}{m} g_m$ appartient au sous-espace vectoriel V engendré par $(g_1,\dots,g_r)$. Passant à la limite, il vient $v \in V$ (regarder les coordonnées de v après avoir complété $(g_1,\dots,g_r)$ en une base ; se rappeler éventuellement qu'un sous-espace de $\mathbf{R}^n$ est fermé). Comme v est arbitraire, $V = \mathbf{R}^n$ et $(g_1,\dots,g_r)$ engendre $\mathbf{R}^n$. Puisque c'est une famille libre, c'est une base de $\mathbf{R}^n$ et G est un réseau. □

D. Volumes et indices

Soit G un réseau de $\mathbf{R}^n$ et soit $(v_1,\dots,v_n)$ une base de G qui est une famille libre de $\mathbf{R}^n$. Posons $V(G) = |\det(v_1,\dots,v_n)|$, le déterminant étant pris par rapport à la base canonique de $\mathbf{R}^n$. C'est le volume du parallélépipède de $\mathbf{R}^n$ de côtés $v_1,\dots,v_n$.

Si $(u_1,\dots,u_n)$ est une autre base de G, il existe une matrice $A \in \mathrm{GL}_n(\mathbf{Z})$ telle que $(u_1,\dots,u_n) = (v_1,\dots,v_n)$. Par conséquent,

$$\det(u_1,\dots,u_n) = \det(A)\det(v_1,\dots,v_n) = \pm\det(v_1,\dots,v_n).$$

En particulier, $|\det(u_1,\dots,u_n)| = |\det(v_1,\dots,v_n)|$ ce qui montre que $V(G)$ ne dépend pas du choix de la base $(v_1,\dots,v_n)$ et justifie la notation adoptée. On appelle $V(G)$ le *covolume* du réseau G. On a par exemple, $V(\mathbf{Z}^n) = 1$. Si G est un réseau et λ un nombre réel non nul, l'ensemble λG des λg, pour $g \in G$, est un réseau et $V(\lambda G) = |\lambda|^n V(G)$. En effet, si $(u_1,\dots,u_n)$ est une base de G, $(\lambda u_1,\dots,\lambda u_n)$ est une base de λG.

Soit G' un sous-groupe de G. C'est en particulier un sous-groupe discret de $\mathbf{R}^n$ et, d'après le théorème 3.2.3, c'est un groupe abélien libre engendré par une famille libre $(u_1,\dots,u_r)$ de $\mathbf{R}^n$. Pour que G' engendre $\mathbf{R}^n$, il faut et il suffit que l'on ait $r=n$.

PROPOSITION 3.2.1. — *Le groupe quotient G/G' est fini si et seulement si $r=n$, c'est-à-dire si et seulement si G' est aussi un réseau de $\mathbf{R}^n$. Le cardinal de G/G' est alors égal à $V(G')/V(G)$.*

Démonstration. — □

Exercices

12) On dit qu'une matrice de taille $n\times m$ à coefficients dans un corps K est sous forme échelonnée réduite (par lignes) si a) ses lignes sont échelonnées ; b) le premier coefficient non nul d'une ligne (un pivot) est égal à 1 ; c) dans la colonne d'un pivot, les autres coefficients sont nuls.

a) Démontrer que toute matrice $A\in \mathrm{Mat}_{n,m}(K)$ peut être mise sous forme échelonnée réduite par des opérations élémentaires sur les lignes.

b) Si $A\in \mathrm{Mat}_{n,m}(K)$, démontrer qu'il existe une matrice sous forme échelonnée réduite H, et une seule, ainsi qu'une matrice inversible $U\in \mathrm{GL}_n(K)$, telle que $A=UH$.

13) Pour $x\in\mathbf{R}$, on note $\lfloor x\rfloor$ sa partie entière et $\{x\}=x-\lfloor x\rfloor$ sa partie fractionnaire.

Soit α un nombre réel irrationnel.

a) Pour $q\in\mathbf{N}^*$, montrer que les éléments $0,\{\alpha\},\{2\alpha\},\dots,\{q\alpha\}$ de $[0,1]$ sont deux à deux distincts.

b) À l'aide du principe des tiroirs, en déduire qu'il existe des entiers i et j tels que $0\leqslant i<j\leqslant q$ et tels que $\big|\{j\alpha\}-\{i\alpha\}\big|\leqslant 1/q$.

c) Montrer qu'il existe, pour tout entier $Q>0$ une fraction irréductible p/q telle que $\left|\alpha-\frac{p}{q}\right|\leqslant 1/qQ$ (*Dirichlet*).

14) On note $(e_1,\dots,e_n)$ la base canonique de $\mathbf{R}^n$. Soit G un sous-groupe de $\mathbf{Z}^n$ qui n'est pas contenu dans un hyperplan de $\mathbf{R}^n$.

a) Pour $i\in\{1,\dots,n\}$, démontrer qu'il existe des vecteurs $f_i\in G$ de la forme $x_1e_1+\cdots+x_ie_i$, où $x_1,\dots,x_i$ sont des entiers et $x_i\neq 0$.

b) Pour $i\in\{1,\dots,n\}$, soit f_i un vecteur comme ci-dessus, choisi de sorte que x_i soit strictement positif et minimal.

Démontrer que $(f_1,\dots,f_n)$ est une base de G.

§3.3. Théorèmes de Hermite et Minkowski

A. Minimas d'une forme quadratique définie positive

Soit q une forme quadratique sur $\mathbf{R}^n$ et soit Q sa matrice dans la base canonique. On a donc $q(x)={}^txQx$ pour tout vecteur colonne $x\in\mathbf{R}^n$. Définissons le discriminant de q, noté $\mathrm{disc}(q)$, comme le déterminant de Q.

Supposons q définie positive, c'est-à-dire $q(x) > 0$ pour tout $x \in \mathbf{R}^n$, $x \neq 0$. Alors, $\operatorname{disc}(q) > 0$.

Le premier énoncé généralise en toute dimension le théorème 3.1.1.

THÉORÈME 3.3.1 (Hermite, 1854). — *Il existe un vecteur $u \in \mathbf{Z}^n$, $u \neq 0$, tel que $q(u) \leqslant \left(\frac{4}{3}\right)^{(n-1)/2} \operatorname{disc}(q)^{1/n}$.*

Plus généralement, il existe n vecteurs $u_1, \dots, u_n \in \mathbf{Z}^n$, linéairement indépendants, tels que

$$q(u_1) \dots q(u_n) \leqslant \left(\frac{4}{3}\right)^{n(n-1)/2} \operatorname{disc}(q).$$

Démonstration. — Comme q est définie positive, $q(x)$ tend vers l'infini lorsque $\|x\| \to \infty$; en particulier, il existe un vecteur non nul $u \in \mathbf{Z}^n$ en lequel $q(u)$ est minimale.

Les coordonnées de u sont premières entre elles : sinon, il existerait un entier $d > 1$ et un vecteur $u' \in \mathbf{Z}^n$ tel que $u = du'$; alors, $q(u) = d^2 q(u')$, d'où $0 < q(u') < q(u)$ ce qui contredit le choix de u. D'après le corollaire 3.2.5, il existe une matrice $U \in \mathrm{GL}_n(\mathbf{Z})$ dont u est la première colonne. La forme quadratique $q(Uy)$ est définie positive, de même discriminant que q ; puisque $U \in \mathrm{GL}_n(\mathbf{Z})$, elle atteint son minimum sur $\mathbf{Z}^n \setminus \{0\}$ en $(1, 0, \dots, 0)$.

Notons $(b_{i,j})$ la matrice de la forme quadratique $q(Uy)$; on a donc

$$q(Uy) = \sum_{i,j=1}^{n} b_{i,j} y_i y_j = b_{1,1} \left(y_1 + \sum_{i=2}^{n} \frac{b_{1,i}}{b_{1,1}} y_i \right)^2 + q'(y'),$$

où q' est une forme quadratique sur $\mathbf{R}^{n-1}$ en la variable $y' = (y_2, \dots, y_n)$.

Si l'on pose $y_1' = y_1 + \sum_{i=2}^{n} \frac{b_{1,i}}{b_{1,1}} y_i$, le changement de variables $(y_1, \dots, y_n) \mapsto (y_1', y_2, \dots, y_n)$ transforme la forme $q(Uy)$ en la forme $b_{1,1}(y_1')^2 + q'(y')$ dont le discriminant est $b_{1,1} \operatorname{disc}(q')$. Comme ce changement de variables est de déterminant 1, on a on a $b_{1,1} \operatorname{disc}(q'(y')) = \operatorname{disc}(q(Uy)) = \operatorname{disc}(q)$.

Par récurrence, il existe un vecteur $y' = (y_2, \dots, y_n) \in \mathbf{Z}^{n-1}$, non nul, tel que $q'(y') \leqslant \left(\frac{4}{3}\right)^{(n-2)/2} \operatorname{disc}(q')^{1/(n-1)}$. Soit y_1 l'entier le plus proche de $-\sum_{i=2}^{n} \frac{b_{1,i}}{b_{1,1}} y_i$ et posons $y = (y_1, y_2, \dots, y_n)$. Alors,

$$q(Uy) \leqslant b_{1,1} \frac{1}{4} + \left(\frac{4}{3}\right)^{(n-2)/2} \operatorname{disc}(q')^{1/(n-1)}.$$

En outre $q(Uy) \geqslant b_{1,1}$. On en déduit que

$$\frac{3}{4} b_{1,1} \leqslant \left(\frac{4}{3}\right)^{(n-2)/2} \left(\operatorname{disc}(q)/b_{1,1}\right)^{1/(n-1)}.$$

Mettant cette expression à la puissance $n-1$, on trouve

$$\left(\frac{3}{4} b_{1,1}\right)^{n-1} b_{1,1} \leqslant \left(\frac{4}{3}\right)^{(n-1)(n-2)/2} \operatorname{disc}(q),$$

d'où

$$b_{1,1}^n \leqslant (4/3)^{(n-1)(n-2)/2+(n-1)} \operatorname{disc}(q) = (4/3)^{n(n-1)/2} \operatorname{disc}(q).$$

La première partie du théorème en résulte.

Démontrons aussi la seconde assertion par récurrence. Il existe ainsi des vecteurs $u'_2, \dots, u'_n \in \mathbf{Z}^{n-1}$, linéairement indépendants, tels que

$$q'(u'_2) \dots q'(u'_n) \leqslant \left(\frac{4}{3}\right)^{(n-1)(n-2)/2} \operatorname{disc}(q').$$

Pour $i \in \{2, \dots, n\}$, définissons un vecteur $u_i \in \mathbf{Z}^n$ comme suit : si u'_i a pour coordonnées $(y_2, \dots, y_n)$, choisissons pour y_1 l'entier le plus proche de $-\sum_{i=2}^n \frac{b_{1,i}}{b_{1,1}} y_i$ et posons $u_i = (y_1, \dots, y_n)$. Alors, pour tout $i \in \{2, \dots, n\}$,

$$q(Uu_i) \leqslant b_{1,1} \frac{1}{4} + q'(u'_i) \leqslant \frac{1}{4} q(Uu_i) + q'(u'_i)$$

et $q(Uu_i) \leqslant \frac{4}{3} q'(u'_i)$. On a ainsi

$$\begin{aligned} q(Uu_1) q(Uu_2) \dots q(Uu_n) &\leqslant b_{1,1} \left(\frac{4}{3}\right)^{n-1} q'(u'_2) \dots q'(u'_n) \\ &\leqslant \left(\frac{4}{3}\right)^{n-1} \left(\frac{4}{3}\right)^{(n-1)(n-2)/2} \operatorname{disc}(q') b_{1,1} \\ &= \left(\frac{4}{3}\right)^{n(n-1)/2} \operatorname{disc}(q). \end{aligned}$$

Les vecteurs $u'_2, \dots, u'_n$ sont linéairement indépendants, donc les vecteurs $u_2, \dots, u_n$ aussi. Comme seule la première coordonnée de u_1 n'est pas nulle, u_1 n'appartient pas à l'espace engendré par les vecteurs $u_2, \dots, u_n$ et la famille $(u_1, \dots, u_n)$ est libre. Le théorème est ainsi démontré. □

Le théorème précédent concerne le comportement d'une forme quadratique définie positive vis à vis du réseau $\mathbf{Z}^n$. Dans la pratique, il convient d'en énoncer une variante où interviennent un réseau et une forme quadratique arbitraires.

COROLLAIRE 3.3.2. — *Soit G un réseau de $\mathbf{R}^n$ et soit q une forme quadratique définie positive sur $\mathbf{R}^n$.*

Il existe un élément $g \in G$, non nul, tel que $q(g) \leqslant (4/3)^{(n-1)/2} (\operatorname{disc}(q) V(G)^2)^{1/n}$.

Il existe des éléments $g_1, \dots, g_n$ de G, linéairement indépendants, tels que

$$q(g_1) \dots q(g_n) \leqslant \left(\frac{4}{3}\right)^{n(n-1)/2} \operatorname{disc}(q) V(G)^2.$$

Démonstration. — Soit $(v_1, \dots, v_n)$ une base du réseau G et considérons la forme quadratique q' sur $\mathbf{R}^n$ donnée par

$$q'(x_1, \dots, x_n) = q(x_1 v_1 + \dots + x_n v_n).$$

Son discriminant est égal à $\mathrm{disc}(q)\det(v_1,\dots,v_n)^2 = \mathrm{disc}(q)V(G)^2$. Le corollaire en découle aussitôt. □

COROLLAIRE 3.3.3. — *Tout réseau G de $\mathbf{R}^n$ contient un point non nul g tel que $\|g\| \leqslant (4/3)^{(n-1)/4}V(G)^{1/n}$.*

Démonstration. — C'est le cas particulier où q est la forme quadratique donnée par $q(v) = \|v\|^2$. □

On appelle *constante de Hermite* le plus petit nombre réel γ_n tel que tout réseau G de $\mathbf{R}^n$ contienne un point non nul g tel que $\|g\|^2 \leqslant \gamma_n V(G)^{2/n}$. On a donc $\gamma_n \leqslant (4/3)^{(n-1)/2}$.

Si cette majoration est une égalité pour $n = 1$ ou 2 (considérer le réseau de base 1 et $e^{2i\pi/3}$ dans $\mathbf{C} = \mathbf{R}^2$), elle est très grossière lorsque n grandit. On peut en effet démontrer que lorsque n tend vers l'infini,

$$\frac{1}{2\pi e} \leqslant \frac{\gamma_n}{n} \leqslant \frac{1{,}744}{2\pi e};$$

la croissance de γ_n est donc linéaire et non exponentielle ! Signalons qu'on en connaît la valeur exacte jusque $n = 8$, mais pas au-delà !

B. Sommes de quatre carrés

Dans ce paragraphe, on démontre par les méthodes de géométrie des nombres le théorème de Lagrange selon lequel tout entier positif est somme de quatre carrés.

LEMME 3.3.1. — *Considérons r formes linéaires $\ell_1,\dots,\ell_r$ sur $\mathbf{Z}^n$ et des entiers strictement positifs $d_1,\dots,d_r$. L'ensemble G des éléments $x \in \mathbf{Z}^n$ tels que $\ell_i(x) \equiv 0 \pmod{d_i}$ est un réseau de $\mathbf{R}^n$ tel que $V(G) \leqslant d_1 \dots d_r$.*

Démonstration. — D'après la proposition 3.2.1, il suffit de démontrer que le groupe $\mathbf{Z}^n/G$ est fini de cardinal au plus $d_1 \dots d_r$. Or, l'application $L\colon \mathbf{Z}^n \to \prod_{i=1}^r (\mathbf{Z}/d_i\mathbf{Z})$ donnée par $L(x) = (\ell_i(x) \pmod{d_i})$ est un homomorphisme de groupes abélien dont le noyau est égal à G. Par suite, $\mathbf{Z}^n/G$ est isomorphe à un sous-groupe de $\prod_{i=1}^r (\mathbf{Z}/d_i\mathbf{Z})$. Il est en particulier fini et son cardinal divise le produit $d_1 \dots d_r$. □

LEMME 3.3.2. — *Pour tout nombre premier p, il existe des entiers a et b tels que $a^2 + b^2 + 1$ soit multiple de p.*

Démonstration. — En effet, il y a $1 + (p-1)/2 = (p+1)/2$ éléments de la forme a^2, avec $a \in \mathbf{Z}/p\mathbf{Z}$, et autant d'éléments de la forme $1 - b^2$, avec $b \in \mathbf{Z}/p\mathbf{Z}$. Puisque $(p+1)/2 + (p+1)/2 = p+1 > p$, il existe a et b dans $\mathbf{Z}/p\mathbf{Z}$ tels que $a^2 = 1 - b^2$, d'où le lemme. □

THÉORÈME 3.3.3 (Lagrange, 18??) — *Tout entier positif est somme de quatre carrés : pour tout entier $n > 0$, il existe des nombres entiers x, y, z, t tels que $n = x^2 + y^2 + z^2 + t^2$.*

Démonstration. — Il suffit de traiter le cas où n est sans facteur carré. Pour tout facteur premier p de n, choisissons des entiers a_p et b_p tels que $a_p^2 + b_p^2 + 1$ soit multiple de p. Considérons alors l'ensemble G des éléments (x, y, z, t) de $\mathbf{Z}^2$ vérifiant les congruences suivantes :

$$x = a_p z + b_p t \pmod{p}, \qquad y = b_p z - a_p t \pmod{p}.$$

D'après le lemme 3.3.1, G est un réseau de $\mathbf{R}^4$ tel que $V(G)$ est majoré par le produit des carrés des facteurs premiers de n ; autrement dit, $V(G) \leqslant n^2$.

D'après le théorème d'Hermite, il existe un élément non nul (x, y, z, t) de G tel que

$$x^2 + y^2 + z^2 + t^2 \leqslant (4/3)^{3/2} V(G)^{1/2} < 2n,$$

puisque $(4/3)^3 = 64/27 \leqslant 3 < 4$. Posons $N = x^2 + y^2 + z^2 + t^2$ et démontrons que $N = n$. Pour tout facteur premier p de n,

$$\begin{aligned} N = x^2 + y^2 + z^2 + t^2 &\equiv (a_p z + b_p t)^2 + (b_p z - a_p t)^2 + z^2 + t^2 \\ &\equiv (a_p^2 + b_p^2 + 1) z^2 + (b_p^2 + a_p^2 + 1) t^2 \pmod{p}, \end{aligned}$$

donc $x^2 + y^2 + z^2 + t^2$ est multiple de p. Par conséquent, N est multiple de n. Puisque $0 < N < 2n$, on a nécessairement $N = n$, ce qui conclut la démonstration du théorème de Lagrange. □

C. Les théorèmes de Minkowski

Les théorèmes de Minkowski auquel ce paragraphe est consacré généralisent du cas d'une norme euclidienne au cas d'une norme quelconque les théorèmes de Hermite. La méthode inaugurée par Minkowski, mélangeant géométrie et théorie des nombres, était d'une très grande innovation au point qu'elle provoqua l'apparition d'une nouvelle branche de l'arithmétique : la *géométrie des nombres*.

PROPOSITION 3.3.1. — *Soit $\|\cdot\|$ une norme sur $\mathbf{R}^n$ et soit V le volume de la boule unité B. Soit r un entier naturel. Si $V \geqslant r2^n$, il existe dans $\mathbf{Z}^n$ des vecteurs $\pm u_1, \ldots, \pm u_n$ de normes $\leqslant 1$ et deux à deux distincts.*

Démonstration. — Supposons d'abord que $V > r2^n$. Soit f la fonction indicatrice de la boule de centre 0 et de rayon 1/2 et posons, pour $x \in \mathbf{R}^n$, $F(x) = \sum_{u \in \mathbf{Z}^n} f(x - u)$. Comme f est à support compact, c'est une somme finie ; plus précisément, les termes d'indice u sont nuls dès que $\|u\| \geqslant \|x\| + 1$. Intégrons F sur le parallélépipède $P = [0,1]^n$; on trouve

$$\int_P F = \sum_{u \in \mathbf{Z}^n} \int_P f(x-u)\,dx = \sum_{u \in \mathbf{Z}^n} \int_{u+P} f(x)\,dx = \int_{\mathbf{R}^n} f(x)\,dx = \mathrm{vol}(B(0,1/2)) = V/2^n > r.$$

Comme le volume de P est égal à 1, il existe au moins un point x de P tel que $F(x) > r$; comme F est à valeurs entières, on a donc $F(x) \geqslant r+1$ et il existe des éléments distincts $u_0, \ldots, u_r$ de $\mathbf{Z}^n$ tels que $f(x-u_i) \geqslant 1$, c'est-à-dire $\|x - u_i\| \leqslant 1/2$. Ordonnons les u_i suivant l'ordre lexicographique de leurs coordonnées. Alors, pour tout $i \in \{1, \ldots, r\}$, $u_i - u_0$

est un point de $\mathbf{Z}^n$ tel que $\|u_i - u_0\| \leqslant 1$. Ces vecteurs $u_i - u_0$ sont non nuls et deux-à-deux distincts ; de plus, si $u_i - u_0 = -(u_j - u_0)$, alors $u_0 = (u_i + u_j)/2$, ce qui contredit le fait que u_0 est le vecteur dont les coordonnées sont les plus petites dans l'ordre lexicographique. Les $2r$ vecteurs $\pm(u_i - u_0)$ sont donc non nuls, deux à deux distincts, à coordonnées entières et de normes $\leqslant 1$. Cela conclut la démonstration lorsque $V > r2^n$.

Supposons maintenant que l'on ait l'égalité $V = r2^n$ et considérons les normes $\alpha\|\cdot\|$, où α est un nombre réel tel que $\alpha > 1$. Comme le volume de la boule unité pour cette nouvelle norme est $\alpha^n V > r2^N$, il existe $2r$ vecteurs $\pm v_{1,\alpha}, \ldots, \pm v_{r,\alpha}$, non nuls et deux à deux distincts dans $\mathbf{Z}^n$ tels que $\|v_{i,\alpha}\| \leqslant \alpha$ pour tout i. Faisons tendre α vers 1 ; comme les parties fermées bornées de $\mathbf{R}^n$ sont compactes, on peut trouver une suite (α_n) de nombres réels tels que $\alpha_n > 1$ et $\alpha_n \to 1$ de sorte que pour tout i, la suite (v_{i,α_n}) convergent ; notons v_i sa limite ; on a $\|v_i\| \leqslant 1$. Une suite convergente de vecteurs à coordonnées entières est stationnaire ; on a donc $v_{i,\alpha_n} = v_i$ pour n assez grand. En particulier, les $2r$ vecteurs $\pm v_1, \ldots, \pm v_r$ sont non nuls et deux à deux distincts. Cela termine la démonstration. □

THÉORÈME 3.3.2. — *Soit G un réseau de $\mathbf{R}^n$; soit $\|\cdot\|$ une norme sur $\mathbf{R}^n$, B sa boule unité et* $\mathrm{vol}(B)$ *son volume. Il existe un vecteur non nul $u \in G$ tel que $\|u\| \leqslant 2(V(G)/\mathrm{vol}(B))^{1/n}$.*

Démonstration. — Soit A la matrice d'une base $(u_1, \ldots, u_n)$ de G et considérons la norme $\|\cdot\|'$ sur $\mathbf{R}^n$ donnée par $\|x\|' = \alpha\|Ax\|$, où $\alpha = (\mathrm{vol}(B)/V(G))^{1/n}/2$. La boule unité B' pour cette norme est égale à $\alpha^{-1}A^{-1}B$, son volume $\mathrm{vol}(B')$ est donc égal à $\mathrm{vol}(B)/\alpha^n|\det A| = 2^n V(G)/|\det A|$. Par ailleurs, on a $V(G) = |\det A|$ par définition, d'où $\mathrm{vol}(B') = 2^n$. D'après la proposition, il existe un vecteur non nul $x \in \mathbf{Z}^n$ tel que $\|x\|' \leqslant 1$. Alors, le vecteur $u = x_1u_1 + \cdots + x_nu_n$ de G vérifie

$$\|u\| = \|Ax\| = \|x\|'/\alpha = 2(V(G)/\mathrm{vol}(B))^{1/n}.$$

Le théorème est ainsi démontré. □

Dans le cas où la norme est euclidienne, il convient de comparer le théorème de Minkowski avec celui d'Hermite. Commençons par rappeler le volume d'un ellipsoïde :

LEMME 3.3.3. — *Soit q une forme quadratique définie positive sur $\mathbf{R}^n$ et soit B l'ensemble des $x \in \mathbf{R}^n$ tels que $q(x) \leqslant 1$. Alors,* $\mathrm{vol}(B) = \pi^{n/2}\Gamma(1 + n/2)^{-1}\,\mathrm{disc}(q)^{-1/2}$.

Démonstration. — La forme q définit un produit scalaire sur $\mathbf{R}^n$. Soit A la matrice d'une base orthonormée pour q, de sorte que $q(Ax) = \|x\|^2$, où $\|\cdot\|$ désigne la norme euclidienne usuelle sur $\mathbf{R}^n$. Si B_1 désigne la boule unité de $\mathbf{R}^n$ (pour la norme euclidienne), on a donc $B = A(B_1)$ et $\mathrm{vol}(B) = |\det A|\,\mathrm{vol}(B_1)$. D'autre part, $\mathrm{disc}(q) = |\det A|^{-2}$ car le discriminant de la forme quadratique $\|\cdot\|^2$ est égal à 1.

Il suffit donc de démontrer que $\mathrm{vol}(B_1) = \pi^{n/2}/\Gamma(1+n/2)$. Pour cela, on calcule $I_n = \int_{\mathbf{R}^n} \exp(-\|x\|^2)\,dx$. Un changement de variables en coordonnées polaires fournit

$$I_n = S_n \int_0^\infty \exp(-r^2) r^{n-1}\,dr,$$

où S_n est la surface de la sphère unité de $\mathbf{R}^n$. On a donc

$$I_n = S_n \int_0^\infty \exp(-x) x^{(n-1)/2} \frac{dx}{2\sqrt{x}} = \frac{1}{2} S_n \int_0^\infty \exp(-x) x^{n/2-1}\,dx = \frac{1}{2} S_n \Gamma(n/2),$$

où Γ désigne la fonction Γ d'Euler. D'autre part, le volume de la boule unité se calcule comme suit :

$$\mathrm{vol}(B_1) = S_n \int_0^1 r^{n-1}\,dr = \frac{1}{n} S_n.$$

Par suite,

$$\mathrm{vol}(B_1) = \frac{2}{n} I_n \frac{1}{\Gamma(n/2)} = I_n/\Gamma(1+\frac{n}{2}).$$

Enfin, le théorème de Fubini assure que $I_n = (I_1)^n$; de même, $I_2 = (I_1)^2$ d'où $I_n = (I_2)^{n/2}$. En dimension 1 et 2, le volume de la boule unité vaut respectivement 2 et π. Par suite,

$$I_1 = 2\Gamma(1+\frac{1}{2}) = 2\Gamma(3/2) = \Gamma(1/2) \quad \text{et} \quad I_2 = \pi\Gamma(2) = \pi.$$

On en déduit $\Gamma(1/2) = \sqrt{\pi}$ et $\mathrm{vol}(B_1) = \pi^{n/2}/\Gamma(1+n/2)$. □

La constante c_{M} figurant au membre de droite du théorème de Minkowski est donc égale à

$$\begin{aligned} c_{\mathrm{M}} - 2V(G)/\mathrm{vol}(B)^{1/n} &= 2V(G)\pi^{-1/2}\Gamma(1+n/2)^{1/n}\,\mathrm{disc}(q)^{1/2n} \\ &= V(G)\,\mathrm{disc}(q)^{1/2n}\left(\frac{2}{\sqrt{\pi}}\Gamma(1+n/2)^{1/n}\right). \end{aligned}$$

D'après la formule de Stirling,

$$\Gamma(1+n/2) = (n/2)\Gamma(n/2) \sim (n/2)^{1+n/2} e^{-n/2}\sqrt{\pi n}$$

et il en résulte

$$\Gamma(1+n/2)^{1/n} \sim (n/2)^{1/2} e^{-1/2} \sim \frac{\sqrt{n}}{\sqrt{2e}}.$$

Autrement dit, lorsque n tend vers l'infini,

$$c_{\mathrm{M}}^2 \sim V(G)^2\,\mathrm{disc}(q)^{1/n}\frac{4n}{2\pi e}$$

et le théorème de Minkowski implique l'inégalité

$$\gamma_n \leqslant \frac{4n}{2\pi e} + \mathrm{o}(n).$$

Ce résultat est d'une bien meilleure qualité que ce que fournit le théorème de Hermite qui se contente d'affirmer que $\gamma_n \leqslant (4/3)^{(n-1)/2}$.

Cependant, on observera que lorsque $n = 2$, le théorème de Minkowski fournit la majoration $\gamma_2 \leqslant 4/\pi\Gamma(3) = 2/\pi$ alors que l'on a $\gamma_2 = 2/\sqrt{3}$.

De manière analogue au théorème de Hermite, Minkowski a aussi démontré l'existence d'une base de $\mathbf{R}^n$ formé de vecteurs entiers d'un réseau et de normes petites. La démonstration étant nettement plus délicate, nous nous contentons de l'énoncer :

THÉORÈME 3.3.4. — *Soit G un réseau de $\mathbf{R}^n$; soit $\|\cdot\|$ une norme sur $\mathbf{R}^n$, B sa boule unité et* vol(B) *son volume. Il existe une base $(u_1, \ldots, u_n)$ de $\mathbf{R}^n$ formée de vecteurs de G tels que*

$$\|u_1\| \ldots \|u_n\| \leqslant 2^n \,\mathrm{vol}(B)/V(G).$$

Exercices

15) Soit $Q \in \mathrm{Mat}_3(\mathbf{Z})$ une matrice symétrique définie positive à coefficients entiers et de déterminant 1. Soit q la forme quadratique correspondante.

a) Démontrer qu'il existe une base (e_1, e_2, e_3) de $\mathbf{R}^3$ formée de vecteurs de $\mathbf{Z}^3$ tels que $q(e_1)q(e_2)q(e_3) \leqslant 2$.

b) Démontrer qu'il existe une base (e'_1, e'_2, e'_3) de $\mathbf{Z}^3$ dans laquelle la matrice de q est l'identité.

16) a) Soit x, y, z trois nombres entiers et $n = x^2 + y^2 + z^2$. Démontrer que n n'est pas de la forme $4^a(8b + 7)$, pour $a, b \in \mathbf{N}$.

b) Inversement, soit n un entier qui n'est congru ni à 4 ni à 7 modulo 8. Démontrer, en considérant les diverses congruences possibles pour n modulo 8 qu'il existe un entier $b > 0$ tel que $p = bn - 1$ soit un nombre premier et $\left(\dfrac{-b}{p}\right) = 1$.

c) Construire alors une matrice symétrique, définie positive, à coefficients entiers de la forme

$$\begin{pmatrix} a_{11} & a_{12} & 1 \\ a_{12} & a_{22} & 0 \\ 1 & 0 & n \end{pmatrix}$$

et dont le déterminant soit égal à 1.

d) En utilisant le résultat de l'exercice 15, démontrer que n est la somme de trois carrés de nombres entiers. (*Théorème de Dirichlet*).

17) Soit $\|\cdot\|$ une norme sur $\mathbf{R}^n$, soit B sa boule unité et vol(B) son volume.

a) Soit $N(t)$ le nombre de points $x \in \mathbf{Z}^n$ tels que $\|x\| \leqslant t$. Lorsque t tend vers $+\infty$, démontrer l'équivalent $N(t) \sim t^n \,\mathrm{vol}(B)$. (En notant $P = [0,1]^n$, observer que les cubes $x + P$, où $x \in \mathbf{Z}^n$ vérifie $\|x\| \leqslant t - n$ sont contenus dans tB.)

b) On suppose que vol$(B) > 2^n$. Soit $T = \lceil 2t \rceil$ (plus petit entier supérieur ou égal à $2t$) ; si t est assez grand, démontrer qu'il existe des points distincts $x = (x_1, \ldots, x_n)$ et $x' = (x'_1, \ldots, x'_n) \in \mathbf{Z}^n \cap tB$ tels que $x_i \equiv x'_i \pmod{T}$ pour tout i. En déduire qu'il existe un point $y \in \mathbf{Z}^n \cap B$ différent de 0. (*Démonstration du théorème de Minkowski due à Mordell, 1935*)

CHAPITRE 4

ANNEAUX D'ENTIERS ALGÉBRIQUES

§4.1. Nombres algébriques, entiers algébriques

A. Définitions ; premières propriétés

DÉFINITION 4.1.1. — *On dit qu'un nombre complexe z est un* nombre algébrique *s'il existe un polynôme non nul $P \in \mathbf{Q}[X]$ tel que $P(z) = 0$. Un nombre complexe qui n'est pas algébrique est appelé* transcendant. *On dit qu'un nombre complexe z est un* entier algébrique *s'il existe un polynôme $P \in \mathbf{Z}[X]$, unitaire (donc non nul), tel que $P(z) = 0$.*

Les nombres complexes $\sqrt{2}$, $3 + i\sqrt{5}$, etc. sont des nombres algébriques, puisqu'il sont racines des polynômes $X^2 - 2$ et $(X-3)^2 + 5 = X^2 - 6X + 14$ respectivement. Ce sont même des entiers algébriques.

En revanche, les nombres réels e (base des logarithmes néperiens) ou π (surface d'un cercle de rayon 1) sont des nombres transcendants. On observera que ces deux nombres transcendants ne sont pas définis par l'algèbre ou l'arithmétique, mais par l'analyse : ce sont les valeurs de séries ($\sum_{n=1}^{\infty} 1/n!$) ou d'intégrales ($4\int_0^1 \sqrt{1-x^2}\,dx$).

PROPOSITION 4.1.2. — *La somme, le produit de deux nombres algébriques est un nombre algébrique ; la somme, le produit de deux entiers algébriques est un entier algébrique. L'inverse d'un nombre algébrique non nul est un nombre algébrique.*

En conséquence, les nombres algébriques forment donc un sous-corps du corps $\mathbf{C}$ des nombres complexes que l'on note $\bar{\mathbf{Q}}$; les entiers algébriques forment un sous-anneau de $\bar{\mathbf{Q}}$ qui est noté $\bar{\mathbf{Z}}$. Plus généralement, si K est un sous-corps de $\mathbf{C}$, on note $\mathbf{Z}_K$ l'anneau des éléments de K qui sont des entiers algébriques.

Démonstration. — La première partie a déjà été démontrée (corollaire 2.1.2) en utilisant le théorème de Cayley-Hamilton. Vérifions que cette démonstration s'adapte au cas des entiers algébriques.

Soit x et y des entiers algébriques, annulés par des polynômes unitaires à coefficients entiers P et Q. Par récurrence, il existe pour tout couple (m,n) d'entiers naturels, un polynôme $P_{m,n} \in \mathbf{Z}[X,Y]$, de degré au plus $\deg(P)-1$ en X et de degré au plus $\deg(Q)-1$ en Y, tel que $x^m y^n = P_{m,n}(x,y)$. (Il suffit de prendre pour $P_{m,n}$ le produit des restes des divisions euclidiennes du polynôme X^m par P et du polynôme Y^n par Q.) Si $z = x+y$ ou $z = xy$, on en déduit ainsi l'existence d'une matrice carrée $A = (a_{(i,j),(k,l)})$ de taille $\deg(P)\deg(Q)$, à coefficients entiers, indexée par l'ensemble des couples (i,j) tels que $0 \leqslant i < \deg(P)$ et $0 \leqslant j < \deg(Q)$, telle que

$$z x^k y^l = \sum a_{(i,j),(k,l)} x^i y^j .$$

Le polynôme caractéristique de la matrice A est à coefficients entiers et est unitaire. D'après la proposition 2.1.3, il annule z, donc z est un entier algébrique. □

Le *polynôme minimal* d'un nombre algébrique z est le polynôme unitaire à coefficients rationnels de plus petit degré qui annule z. (Il n'y en a qu'un seul : si $P(z) = Q(z) = 0$, P et Q étant de mêmes degrés, $P-Q$ est un polynôme qui annule z et est de degré strictement inférieur ; c'est donc le polynôme nul.) Son degré est appelé *degré du nombre algébrique*.

Le polynôme minimal d'un nombre algébrique est un polynôme irréductible de $\mathbf{Q}[X]$. Les polynômes à coefficients entiers qui annulent un nombre algébrique sont exactement les multiples de son polynôme minimal.

PROPOSITION 4.1.3. — *Pour qu'un nombre algébrique soit un entier algébrique, il faut et il suffit que son polynôme minimal soit à coefficients entiers.*

Démonstration. — Un sens est évident : si le polynôme minimal P d'un nombre algébrique z est à coefficients entiers, la relation $P(z) = 0$ entraîne que z est un entier algébrique.

La réciproque est plus subtile et repose sur des résultats démontrés au paragraphe F. Soit z un nombre algébrique, soit P son polynôme minimal et soit Q un polynôme unitaire à coefficients entiers tel que $Q(z) = 0$. D'après le corollaire 2.1.7, le polynôme P est à coefficients entiers, ce qu'il fallait démontrer. □

COROLLAIRE 4.1.4. — *Soit z un nombre algébrique et P son polynôme minimal, noté $P = X^n + a_{n-1}X^{n-1} + \cdots + a_0$.*

Soit d un entier relatif non nul ; pour que dz soit un entier algébrique, il faut et il suffit que que $d^i a_{n-i}$ soit un entier relatif, pour tout entier i tel que $1 \leqslant i \leqslant n$.

Supposons que z soit un entier algébrique non nul. Pour que $1/z$ soit un entier algébrique, il faut et il suffit que $a_0 = \pm 1$.

Démonstration. — Le polynôme $d^n P(X/d)$ est unitaire, irréductible de degré n ; c'est le polynôme minimal de dz. De plus, on a

$$d^n P(X/d) = X^n + d a_{n-1} X^{n-1} + \cdots + d^n a_0,$$

d'où la première assertion.

Si $z \neq 0$, on a $a_0 \neq 0$. Le polynôme $Q = a_0^{-1} X^n P(1/X)$ est irréductible, unitaire, de degré n, et s'annule en $1/z$. Son terme constant est a_0^{-1} ; pour que Q soit à coefficients entiers, il est nécessaire que l'on ait $a_0 = \pm 1$. Si c'est le cas, Q est à coefficients entiers, donc $1/z$ est un entier algébrique. Inversement, si $1/z$ est un entier algébrique, Q est à coefficients entiers et $a_0 = \pm 1$. □

Avec les notations du corollaire, observons que si d est un dénominateur commun de $a_0, \ldots, a_{n-1}$, alors dz est un entier algébrique. De plus, l'entier d est le plus petit entier naturel non nul tel que le polynôme $dP(X)$ appartienne à $\mathbf{Z}[X]$; ses coefficients sont des entiers relatifs premiers entre eux.

COROLLAIRE 4.1.5. — *Soit z un nombre rationnel. Pour que z soit un entier algébrique, il faut et il suffit que ce soit un entier relatif.*

Démonstration. — En effet, si $z \in \mathbf{Q}$, son polynôme minimal est $X - z$. □

D'après le lemme ci-dessous, le polynôme minimal d'un nombre algébrique n'a que des racines simples dans $\mathbf{C}$; ces racines sont appelées les les *conjugués* du nombre algébrique considéré.

LEMME 4.1.6. — *Soit F un corps de caractéristique zéro et soit Ω une extension algébriquement close de F. Soit P un polynôme irréductible à coefficients dans F. L'ensemble des racines de P dans Ω est de cardinal* $\deg(P)$.

Démonstration. — Pour simplifier les notations, on suppose que P est unitaire. Il s'agit de démontrer que P n'a pas de racine multiple. Une telle racine serait une racine commune à P et P'. Posons $D = \operatorname{pgcd}(P, P')$; c'est un polynôme à coefficient dans F qui divise P et est unitaire. Comme P est irréductible a donc $D = P$ ou $D = 1$. Comme F est de caractéristique 0, P' est de degré $\deg(P) - 1$: si le terme de plus haut degré de P est aX^n (avec $a \neq 0$), on a $na \neq 0$ donc le terme de plus haut degré de P' est naX^{n-1} et P' est de degré $n-1$. Cela empêche que P' soit multiple de P si bien que $D = 1$. Autrement dit, P et P' n'ont pas de racine commune. □

DÉFINITION 4.1.7. — *Soit z un nombre algébrique ; notons n son degré, $P = X^n + a_{n-1} X^{n-1} + \cdots + a_0$ son polynôme minimal et $z_1, \ldots, z_n$ ses conjugués. Le dénominateur de z, noté* $\operatorname{den}(z)$*, est le plus petit entier $d > 1$ tel que dz soit un entier algébrique. On appelle alors taille de z l'expression :*

$$\mathrm{t}(z) = \max(\operatorname{den}(z), |z_1|, \ldots, |z_n|).$$

PROPOSITION 4.1.8. — *Soit T un nombre réel. L'ensemble des nombres algébriques de degré n et de taille au plus égale à T est fini.*

Démonstration. — Soit z un tel nombre algébrique, posons $d = \mathrm{den}(z)$ et soit $z_1, \dots, z_n$ ses conjugués. On a donc $|d| \leqslant T$ et $|z_i| \leqslant T$ pour $1 \leqslant i \leqslant n$. Le polynôme minimal de dz est égal au polynôme $\prod_{i=1}^{n}(X - dz_i)$; notons-le $X^n + a_1 X^{n-1} + \cdots + a_n$. Comme dz est un entier algébrique, c'est un polynôme à coefficients entiers.

Pour $1 \leqslant k \leqslant n$, $(-1)^k a_k$ est la k-ième fonction symétrique élémentaire des dz_i ; par suite, $|a_k| \leqslant \sum_{i_1<\cdots<i_k} |dz_{i_1}| \dots |dz_{i_k}| \leqslant \binom{n}{k} e^{2k}$. Ces inégalités montrent que les coefficients du polynôme minimal de dz sont bornés indépendamment de z ; comme ils sont à coefficients entiers, ces polynômes appartiennent à un ensemble fini indépendant de z. Par suite, dz appartient à un ensemble fini indépendant de z, et il en est de même de z car d ne peut prendre qu'un nombre fini de valeurs. □

B. Corps quadratiques

Le début de ce paragraphe est consacré à déterminer les entiers algébriques d'un corps quadratique, c'est-à-dire un sous-corps de $\mathbf{C}$ qui est une extension de degré 2 du corps $\mathbf{Q}$.

PROPOSITION 4.1.1. — *Soit K un corps quadratique. Il existe un unique entier relatif d sans facteur carré tel que $K = \mathbf{Q}(\sqrt{d})$ si $d > 0$ et $K = \mathbf{Q}(i\sqrt{-d})$ si $d < 0$.*

Posons $\delta = \sqrt{d}$ si $d > 0$ et $\delta = i\sqrt{-d}$ sinon. Tout élément z de K s'écrit de manière unique sous la forme $a + b\delta d$, avec $a, b \in \mathbf{Q}$. S'il n'appartient pas à $\mathbf{Q}$, le polynôme minimal d'un tel élément est $X^2 - 2aX + a^2 - db^2$.

Il est pratique de noter $\sqrt{d}$ l'élément δ introduit ci-dessus, c'est-à-dire $\delta = \sqrt{d}$ si $d > 0$ et $\delta = i\sqrt{-d}$ si $d < 0$; on ne s'en privera pas.

Démonstration. — Soit z un élément de $K \setminus \mathbf{Q}$. Comme $\mathbf{Q} \subsetneq \mathbf{Q}(z) \subset K$, on a $[\mathbf{Q}(z) : \mathbf{Q}] \geqslant 2 = [K : \mathbf{Q}]$, d'où $K = \mathbf{Q}(z)$. En particulier, z est degré 2 sur $\mathbf{Q}$ et son polynôme minimal est de la forme $X^2 + aX + b$, avec $a, b \in \mathbf{Q}$. Alors, $(z + \frac{1}{2}a)^2 = -b + \frac{1}{4}a^2$. Posons $D = -b + \frac{1}{4}a^2$; c'est un nombre rationnel. Il n'est pas nul car sinon, on aurait $z = -\frac{1}{2}a$, d'où $z \in \mathbf{Q}$. Notons $D = \varepsilon \prod p_i^{n_i}$ sa décomposition en facteurs premiers, avec $\varepsilon \in \{\pm 1\}$ et $n_i \in \mathbf{Z} \setminus \{0\}$. Posons $m_i = 1$ si n_i est impair et $m_i = 0$ sinon et posons alors $d = \varepsilon \prod p_i^{m_i}$. Alors, d est sans facteur carré ; de plus, $D/d = \prod p_i^{n_i - m_i} = (\prod p_i^{(n_i - m_i)/2})^2$ est un carré, disons $D = dr^2$ avec $r \in \mathbf{Q}$.

Posons $\delta = \sqrt{d}$ si $d > 0$ et $\delta = i\sqrt{-d}$ si $d < 0$ et vérifions que $K = \mathbf{Q}(\delta)$. Tout d'abord, $(z + \frac{1}{2}a)^2 = dr^2 = (\delta r)^2$. Par suite, $\delta = \pm(z + \frac{1}{2}a)/r$ appartient à K ; en outre, $\delta \notin \mathbf{Q}$ car cela entraînerait $z = -\frac{1}{2}a \pm r\delta \in \mathbf{Q}$. Reprenant l'argument initial, on voit que $K = \mathbf{Q}(\delta)$.

Alors, $(1, \delta)$ est une famille de K libre sur $\mathbf{Q}$, donc une base de K car $[K : \mathbf{Q}] = 2$. Tout élément z de K s'écrit donc de façon unique $a + b\delta$, avec $a, b \in \mathbf{Q}$. Si $z \notin \mathbf{Q}$, on a $b \neq 0$, et réciproquement. Supposons $z \notin \mathbf{Q}$. Alors, $(z - a)^2 = b^2\delta^2 = db^2$, si bien que z est racine

du polynôme $X^2 - 2aX + a^2 - db^2$. L'autre racine de ce polynôme est $a - b\delta$. Par suite, ce polynôme n'a pas de racine dans $\mathbf{Q}$; comme il est de degré 2, il est irréductible et c'est le polynôme minimal de z.

Démontrons finalement l'unicité d'un tel entier d : supposons $K = \mathbf{Q}(\delta) = \mathbf{Q}(\delta')$, avec $\delta^2 = d$ et $(\delta')^2 = d'$, d et d' étant des entiers relatifs sans facteurs carrés. On peut donc écrire $\delta' = a + b\delta$, avec $a, b \in \mathbf{Q}$ et $d' = (\delta')^2 = (a^2 + b^2\delta^2) + 2ab\delta = (a^2 + b^2 d) + 2ab\delta$. Par suite, $2ab = 0$ et $a^2 + b^2 d = d'$. Comme $\delta' \notin \mathbf{Q}$, on a $b \neq 0$, d'où $a = 0$ et $d' = b^2 d$. D'après le lemme ci-dessous, b est un entier relatif. On peut écrire de même $d = (1/b)^2 d'$ donc $1/b$ est aussi un entier relatif. On a donc $b = \pm 1$, $b^2 = 1$ et $d = d'$. □

LEMME 4.1.2. — *Soit d un entier relatif non nul sans facteur carré. Si x est un nombre rationnel tel que $dx^2 \in \mathbf{Z}$, c'est un entier relatif.*

Démonstration. — Posons $n = dx^2$ et écrivons $x = p/q$, où p et q sont des entiers relatifs premiers entre eux ; il vient $p^2 d = q^2 n$, si bien que q^2 divise $p^2 d$ en étant premier avec p^2. Par suite, q^2 divise d. Comme d est sans facteur carré, $q^2 = 1$ et $x \in \mathbf{Z}$. □

COROLLAIRE 4.1.3. — *Soit d un entier relatif sans facteur carré. Pour qu'un élément $z = a + b\sqrt{d}$ de $\mathbf{Q}(\sqrt{d})$ soit un entier algébrique, il faut et il suffit que l'on ait :*

— *$a, b \in \mathbf{Z}$ si $d \not\equiv 1 \pmod 4$;*

— *$2a, 2b \in \mathbf{Z}$, $a - b \in \mathbf{Z}$ si $d \equiv 1 \pmod 4$.*

Démonstration. — On doit déterminer les couples (a, b) de nombres rationnels tels que $2a \in \mathbf{Z}$ et $a^2 - db^2 \in \mathbf{Z}$.

La première condition entraîne $a \in \mathbf{Z}$ ou $a + \frac{1}{2} \in \mathbf{Z}$. Plaçons-nous dans le premier cas. Alors, la seconde condition montre que db^2 est un entier, donc b est un entier d'après le lemme ci-dessus. Inversement, si $(a, b) \in \mathbf{Z}^2$, $2a \in \mathbf{Z}$ et $a^2 - db^2 \in \mathbf{Z}$ ce qui démontre que $a + b\delta$ est un entier algébrique.

Plaçons-nous dans le second cas et posons $n = a^2 - db^2$. Alors, $4n = (2a)^2 - d(2b)^2$ ce qui entraîne que $d(2b)^2$ est un entier. Par suite, $2b$ est un entier. Posons $A = 2a$ et $B = 2b$; A est impair et $dB^2 = 4n - A^2$ est impair ; par suite, d et B sont impairs. D'après la première partie de la démonstration, $\frac{A-1}{2} + \frac{B-1}{2}\delta$ est un entier algébrique, donc $z - \left(\frac{A-1}{2} + \frac{B-1}{2}\delta\right) = \frac{1+\delta}{2}$ aussi. Le polynôme minimal de $(1 + \delta)/2$ est $X^2 - X + \frac{1-d}{4}$. Par suite, $d \equiv 1 \pmod 4$. Cela démontre que si $d \not\equiv 1 \pmod 4$, ce cas ne se produit pas et les éléments de K qui sont des entiers algébriques sont les $a + b\delta$ avec $(a, b) \in \mathbf{Z}^2$.

Supposons maintenant que $d \equiv 1 \pmod 4$. Alors, $(1 + \delta)/2$ est un entier algébrique. Ce qui précède démontre que les entiers algébriques sont les éléments de K de la forme $a + b\delta$, avec $(a, b) \in \mathbf{Z}^2$ ou de la forme $\frac{A}{2} + \frac{B}{2}\delta$, où A, B sont des entiers relatifs de même parité. Posant $a = A/2$ et $b = B/2$, cela signifie $2a \in \mathbf{Z}$, $2b \in \mathbf{Z}$ et $a - b = (A - B)/2 \in \mathbf{Z}$. □

Remarque 4.1.4. — Soit d un entier relatif sans facteur carré. Posons $K = \mathbf{Q}(\sqrt{d})$ et soit $\mathbf{Z}_K$ l'ensemble des éléments de K qui sont des entiers algébriques.

Le sous-anneau $\mathbf{Z}[\sqrt{d}]$ engendré par $\sqrt{d}$ dans K est contenu dans $\mathbf{Z}_K$; il est égal à $\mathbf{Z} \oplus \mathbf{Z}\sqrt{d}$.

Dans le cas où $d \not\equiv 1 \pmod 4$, il résulte de la proposition précédente que $\mathbf{Z}_K = \mathbf{Z} \oplus \mathbf{Z}\sqrt{d} = \mathbf{Z}[\sqrt{d}]$.

Dans le cas où $d \equiv 1 \pmod 4$, $(1+\sqrt{d})/2$ est un élément de $\mathbf{Z}_K$ qui n'appartient pas à $\mathbf{Z}[\sqrt{d}]$. Plus généralement, soit $z = a + b\sqrt{d}$ un élément de $\mathbf{Z}_K$. On peut écrire $z = (a-b) + 2b\frac{1+\sqrt{d}}{2}$. Les conditions $2a \in \mathbf{Z}$, $2b \in \mathbf{Z}$ et $a - b \in \mathbf{Z}$ équivalent à $(a-b) \in \mathbf{Z}$ et $2b \in \mathbf{Z}$. Par conséquent, $\mathbf{Z}_K = \mathbf{Z} \oplus \mathbf{Z}\frac{1+\sqrt{d}}{2}$.

Dans les deux cas, l'anneau $\mathbf{Z}_K$ est un groupe abélien libre de rang 2. C'est un cas particulier d'un résultat important que nous démontrerons plus tard.

C. Corps de nombres

Rappelons qu'on appelle corps de nombres un sous-corps de $\mathbf{C}$ qui est une extension finie de $\mathbf{Q}$ (autrement dit, un $\mathbf{Q}$-espace vectoriel de dimension finie). Tout élément d'un corps de nombres est un nombre algébrique.

PROPOSITION 4.1.1. — *Soit K un corps de nombres et soit d son degré. L'ensemble des homomorphismes de corps de K dans $\mathbf{C}$ est de cardinal d.*

Plus généralement, soit K et L deux corps de nombres tels que $L \subset K$. Soit $\varphi : L \to \mathbf{C}$ un homomorphisme de corps. L'ensemble des homomorphismes de corps ψ de K dans $\mathbf{C}$ tels que $\psi|_L = \varphi$ est de cardinal $[K : L]$.

Démonstration. — La première assertion est un cas particulier de la seconde, appliquée avec $L = \mathbf{Q}$ et φ l'inclusion de $\mathbf{Q}$ dans $\mathbf{C}$ qui est, du reste, l'unique homomorphisme de corps de $\mathbf{Q}$ dans $\mathbf{C}$.

On va démontrer la seconde par récurrence sur $[K : L]$. Elle est évidente si $[K : L] = 1$; supposons-la démontrée pour toute inclusion de corps de nombres de degré strictement inférieur à $[K : L]$.

Soit α un élément de K qui n'appartient pas à L. On dispose alors d'une extension intermédiaire $L(\alpha)$ telle que $L \subset L(\alpha) \subset K$. La restriction à $L(\alpha)$ d'un homomorphisme de corps ψ de K dans $\mathbf{C}$ tel que $\psi|_L = \varphi$ est un homomorphisme de corps ψ' de $L(\alpha)$ dans $\mathbf{C}$ dont la restriction à L est φ. Commençons par déterminer ces homomorphismes ψ'. Soit P le polynôme minimal de α sur L, de sorte que $L(\alpha) \simeq L[X]/(P)$. Un homomorphisme de corps $\psi' : L(\alpha) \to \mathbf{C}$ tel que $\psi'|_L = \varphi$ est induit, par passage au quotient, d'un homomorphisme de $L[X]$ dans $\mathbf{C}$ de la forme $a_0 + a_1 X + \ldots \mapsto \varphi(a_0) + \varphi(a_1) z + \ldots$. Que cet homomorphisme passe au quotient signifie exactement que z est racine du polynôme $\varphi(P)$ obtenu en appliquant φ aux coefficients de P.

Le polynôme P est irréductible dans $L[X]$; le polynôme $\varphi(P)$ est donc irréductible dans $\varphi(L)[X]$. Comme $\varphi(L)$ est de caractéristique zéro, le polynôme $\varphi(P)$ a exactement $\deg(P)$ racines (lemm. 4.1.6). Par conséquent, l'ensemble des homomorphismes ψ'

convenables est de cardinal $\deg(P)$. Par l'hypothèse de récurrence, appliquée à l'extension $L(\alpha) \subset L$ et aux homomorphismes $\psi' : L(\alpha) \to \mathbf{C}$, il existe pour chacun d'entre eux exactement $[K : L(\alpha)]$ prolongements à K, d'où au total $[K : L(\alpha)]\deg(P) = [K : L]$ homomorphismes de K dans $\mathbf{C}$ dont la restriction à L est l'homomorphisme φ. Cela conclut la démonstration de la proposition. □

Soit K un corps de nombres, soit d son degré. Soit Φ l'ensemble des homomorphismes de corps de K dans $\mathbf{C}$; on dira aussi *plongements* de K dans $\mathbf{C}$. Parmi les éléments de Φ, certains sont à valeurs dans $\mathbf{R}$, on les appelle *plongement réels*. Les autres prennent des valeurs complexes et on les appelle *plongements complexes*. Si $\varphi : K \hookrightarrow \mathbf{C}$ est un plongement, $z \mapsto \overline{\varphi(z)}$ est encore un plongement de K dans $\mathbf{C}$, noté $\overline{\varphi}$, le plongement conjugué de φ. Dire que φ est réel signifie que $\overline{\varphi} = \varphi$. Ainsi, les plongements complexes se regroupent par paires formées d'un plongement complexe et du plongement complexe conjugué.

Conformément à la tradition, on note r_1 le nombre de plongements réels de K et $2r_2$ le nombre de plongements complexes. Ainsi, $r_1 + 2r_2 = n$ et l'on peut numéroter les plongements de K de la forme $(\varphi_1, \dots, \varphi_{r_1}, \varphi_{r_1+1}, \overline{\varphi_{r_1+2}}, \dots, \varphi_{r_1+r_2}, \overline{\varphi_{r_1+r_2}})$.

Exemple 4.1.2. — Soit K un corps quadratique ; $K = \mathbf{Q}(\sqrt{d})$. Les deux plongements de K dans $\mathbf{C}$ sont donnés par $a + b\sqrt{d} \mapsto a \pm b\sqrt{d}$. Ils sont tous deux réels si $d > 0$ et complexes (conjugués) si $d < 0$.

PROPOSITION 4.1.3. — *Soit K un corps de nombres de degré n, soit $\varphi_1, \dots, \varphi_n$ les plongements de K dans $\mathbf{C}$.*

Soit α un élément de K, soit d son degré et soit $P \in \mathbf{Q}[X]$ son polynôme minimal. Alors, on a $\prod_{i=1}^{n}(X - \varphi_i(\alpha)) = P(X)^{n/d}$.

Démonstration. — Les éléments $\varphi_i(\alpha)$ de $\mathbf{C}$ sont des racines de P ; ils sont donc en nombre au plus d. Inversement, chaque racine β de P fournit un plongement de $\mathbf{Q}(\alpha)$ dans $\mathbf{C}$ tel que $\alpha \mapsto \beta$. Chacun de ces plongements se prolonge en exactement n/d plongements de K dans $\mathbf{C}$, fournissant ainsi les n homomorphismes de K dans $\mathbf{C}$. Par suite, dans la suite $(\varphi_i(\alpha))_{1 \leqslant i \leqslant d}$, chaque racine de P apparaît exactement n/d fois. Si $\beta_1, \dots, \beta_d$ désignent les racines de P, on a donc

$$\prod_{i=1}^{n}(X - \varphi_i(\alpha)) = \prod_{j=1}^{d}(X - \beta_j)^{n/d} = P(X)^{n/d}.$$

□

COROLLAIRE 4.1.4. — *Soit K un corps de nombres de degré n et soit $\varphi_1, \dots, \varphi_n$ les plongements de K dans $\mathbf{C}$. Soit α un élément de K tel que $\varphi_i(\alpha) \neq \varphi_1(\alpha)$ si $i \neq 1$. Alors, $K = \mathbf{Q}(\alpha)$.*

Démonstration. — D'après la proposition précédente, l'ensemble des entiers $i \in \{1,\dots,n\}$ tels que $\varphi_i(\alpha) = \varphi_1(\alpha)$ est de cardinal n/d, où d est le degré de α. Par hypothèse, cet ensemble est de cardinal 1 ; on a donc $n = d$. □

COROLLAIRE 4.1.5 (Théorème de l'élément primitif). — *Soit K un corps de nombres. Il existe un élément $\alpha \in K$ tel que $K = \mathbf{Q}(\alpha)$.*

Démonstration. — Avec les notations du corollaire précédent, il suffit de démontrer qu'il existe un élément $\alpha \in K$ tel que $\varphi_i(\alpha) \neq \varphi_1(\alpha)$ si $i \neq 1$. Soit $i \in \{2,\dots,n\}$; l'ensemble des $\alpha \in K$ tels que $\varphi_i(\alpha) = \varphi_1(\alpha)$ est un sous-$\mathbf{Q}$-espace vectoriel V_i de K. Il est distinct de K car $\varphi_i \neq \varphi_1$. D'après le lemme ci-dessous, leur réunion $V_2 \cup \cdots \cup V_n$ est distincte de K, d'où le corollaire. □

COROLLAIRE 4.1.6. — *Soit F un corps infini, soit V un F-espace vectoriel et soit $V_1,\dots,V_n$ des sous-espaces vectoriels de V distincts de V. Alors $V_1 \cup \cdots \cup V_n$ est distinct de V.*

Démonstration. — Démontrons-la par récurrence sur n. L'assertion est évidente si $n \leqslant 1$. Supposons-la vraie pour $n-1$; il existe donc un vecteur $w \in V$ n'appartenant pas à V_i si $i \geqslant 2$. L'assertion est démontrée si l'on a $w \notin V_1$. Supposons donc $w \in V_1$ et considérons un vecteur $v \in V \setminus V_1$.

Fixons alors $i \in \{1,\dots,n\}$. Soit t et t' des éléments distincts de F tels que $v + tw$ et $v + t'w$ appartiennent à V_i. Alors, $(t - t')w$ appartient à V_i, donc $w \in V_i$ et $i = 1$. Alors, $v = (v + tw) - tw$ appartient à V_1, ce qui est absurde. Par conséquent, il existe au plus un élément $t \in F$ tel que $v + tw$ appartienne à V_i.

Par suite, il existe au plus n éléments t de F tels que $v + tw$ appartienne à l'un des V_i. Comme F est infini, il existe donc $t \in F$ tel que $v + tw$ n'appartienne à aucun des V_i. On a donc $V_1 \cup \cdots \cup V_n \neq V$. □

D. Normes, traces et discriminants

DÉFINITION 4.1.1. — *Soit K un corps de nombres. On appelle* trace *et* norme *d'un élément w de K la trace et le déterminant de l'endomorphisme $\mathbf{Q}$-linéaire de K donné par la multiplication par w. On les note* $\mathrm{Tr}_K(w)$ *et* $\mathrm{N}_K(w)$. *Ce sont des nombres rationnels.*

Pour tout $w \in K$, notons μ_w l'endomorphisme de multiplication par w dans K.

On a $\mu_{w+w'} = \mu_w + \mu_{w'}$. Par suite, $\mathrm{Tr}_K(w + w') = \mathrm{Tr}_K(w) + \mathrm{Tr}_K(w')$. De même, on a $\mu_{ww'} = \mu_w \circ \mu_{w'}$ si bien que $\mathrm{N}_K(ww') = \mathrm{N}_K(w)\,\mathrm{N}_K(w')$. Si $a \in \mathbf{Q}$, on a $\mu_{aw} = a\mu_w$, d'où $\mathrm{Tr}_K(aw) = a\,\mathrm{Tr}_K(w)$ et $\mathrm{N}_K(aw) = a^n\,\mathrm{N}_K(w)$.

PROPOSITION 4.1.2. — *Soit n le degré de K et soit $\varphi_1,\dots,\varphi_n$ les n plongements de K dans $\mathbf{C}$. On a alors, pour tout $w \in K$, les égalités*

$$\mathrm{Tr}_K(w) = \sum_{i=1}^{n} \varphi_i(w) \quad \textit{et} \quad \mathrm{N}_K(w) = \prod_{i=1}^{n} \varphi_i(w).$$

Ce résultat est une conséquence du lemme plus précis suivant.

LEMME 4.1.3. — *Soit $w \in K$ et soit W la matrice de l'endomorphisme de multiplication par w dans une base de K. Alors, W est semblable sur $\mathbf{C}$ à la matrice diagonale* $\mathrm{diag}(\varphi_1(w),\dots,\varphi_n(w))$.

Démonstration. — Soit α un élément primitif de K. Le polynôme minimal de α est $\prod(X-\varphi_i(\alpha))$; comme α est primitif, il est à racines simples et la matrice A de la multiplication par α (dans une base arbitraire fixée) est diagonalisable (sur $\mathbf{C}$), ses valeurs propres étant les $\varphi_i(\alpha)$. Autrement dit, la matrice A est semblable sur $\mathbf{C}$ à la matrice diagonale $D = \mathrm{diag}(\varphi_1(\alpha),\dots,\varphi_n(\alpha))$.

Le cas d'un élément arbitraire de K s'en déduit facilement : si $w \in K$, il existe un polynôme $P \in \mathbf{Q}[X]$ tel que $w = P(\alpha)$. Alors, la matrice W de la multiplication par w dans la base fixée de K est égale à $P(A)$. Le changement de base qui diagonalise A la transforme en la matrice $P(D)$ qui est diagonale, de coefficients $P(\varphi_i(\alpha)) = \varphi_i(P(\alpha)) = \varphi_i(w)$, pour $1 \leqslant i \leqslant n$. □

COROLLAIRE 4.1.4. — *Si w est un entier algébrique de K, alors* $\mathrm{Tr}_K(w)$ *et* $\mathrm{N}_K(w)$ *sont des entiers relatifs.*

Démonstration. — Notons $P = X^d + a_1X^{d-1} + \cdots + a_d$ le polynôme minimal de w. Alors, $\mathrm{Tr}_K(w) = -(n/d)a_1$ et $\mathrm{N}_K(w) = a_d^{n/d}$. Comme d divise n, ce sont des entiers. □

DÉFINITION 4.1.5. — *Soit K un corps de nombres, n son degré. Soit $\mathscr{B} = (z_1,\dots,z_n)$ une base de K. On définit le discriminant de $\mathscr{B}$ par la formule*

$$\mathrm{disc}(\mathscr{B}) = \det(\varphi_i(z_j))^2.$$

(Lorsqu'on change l'ordre des φ_i, le déterminant $\det(\varphi_i(z_j))$ est multiplié par ± 1 ; le discriminant de $\mathscr{B}$ est donc bien défini.)

Exemple 4.1.6. — Soit α un élément primitif de K et soit $\mathscr{B}$ la base $(1,\alpha,\dots,\alpha^{n-1})$. On a donc

$$\mathrm{disc}(\mathscr{B}) = \det(\varphi_i(\alpha^{j-1}))^2 = \det(\varphi_i(\alpha)^{j-1})^2 = V(\varphi_1(\alpha),\dots,\varphi_n(\alpha))^2,$$

où $V(\dots)$ désigne le déterminant de Vandermonde. Par conséquent,

$$\mathrm{disc}(\mathscr{B}) = \prod_{j>i}(\varphi_j(\alpha)-\varphi_i(\alpha))^2.$$

Ainsi, $\mathrm{disc}(\mathscr{B})$ est égal au discriminant du polynôme minimal de α. Il n'est pas nul.

PROPOSITION 4.1.7. — *Soit K un corps de nombres, n son degré. Soit $z_1,\dots,z_n$ des éléments de K et soit D la matrice* $(\mathrm{Tr}_K(z_iz_j))$. *Alors,* $\det(D) = \det(\varphi_i(z_j))^2$. *En particulier, le discriminant d'une base de K est un élément non nul de $\mathbf{Q}$.*

Démonstration. — Notons $\varphi_1, \dots, \varphi_n$ les plongements de K dans $\mathbf{C}$. D'après la proposition précédente, on a

$$\mathrm{Tr}_K(z_i z_k) = \sum_{j=1}^{n} \varphi_j(z_i z_k) = \sum_{j=1}^{n} \varphi_j(z_i)\varphi_j(z_k).$$

Par conséquent, la matrice D est égale à ${}^tAA$, où A est la matrice $(\varphi_i(z_j))$. En particulier, $\det(D) = \det(A)^2$.

Supposons par l'absurde que $(z_1, \dots, z_n)$ soit une base de K de discriminant nul. Alors, la matrice $(\varphi_i(z_j))$ n'est pas inversible ; il existe donc des nombres complexes $(c_1, \dots, c_n)$ tels que $\sum_{i=1}^{n} c_i\varphi_i(z_j) = 0$ pour tout entier j tel que $1 \leqslant j \leqslant n$. L'application $\sum c_i\varphi_i$ de K dans $\mathbf{C}$ étant $\mathbf{Q}$-linéaire et nulle en tout élément d'une base de K, elle est identiquement nulle. Or, $\varphi_1, \dots, \varphi_n$ sont n homomorphismes distincts dans $\mathbf{C}$ et, à ce titre, sont linéairement indépendants sur $\mathbf{C}$ (voir l'exercice **??**). □

COROLLAIRE 4.1.8. — *Soit K un corps de nombres, soit n son degré et soit $\mathscr{B}, \mathscr{B}'$ des bases de K. Alors,*

$$\mathrm{disc}(\mathscr{B}') = \det_{\mathscr{B}}(\mathscr{B}')^2 \,\mathrm{disc}(\mathscr{B}').$$

Démonstration. — Posons $\mathscr{B} = (z_1, \dots, z_n)$, $\mathscr{B}' = (z'_1, \dots, z'_n)$ et soit P la matrice de passage de la base $\mathscr{B}$ à la base $\mathscr{B}'$ telle que $z'_j = \sum_{k=1}^{n} p_{kj} z_k$ pour tout entier $j \in \{1, \dots, n\}$. Soit A et A' les matrices $(\varphi_i(z_j))$ et $(\varphi_i(z'_j))$. Les égalités $\varphi_i(z'_j) = \sum_{k=1}^{n} p_{kj}\varphi_i(z_k)$ entraînent $A' = PA$, d'où $\mathrm{disc}(\mathscr{B}') = \det(A')^2 = \det(P)^2 \det(A)^2 = \det_{\mathscr{B}}(\mathscr{B}')^2 \,\mathrm{disc}(\mathscr{B})$. □

Comme le discriminant de la base donnée par les puissances d'un élément primitif n'est pas nul, il en est de même du discriminant d'une base arbitraire.

Remarque 4.1.9. — La proposition 4.1.7 peut être interprétée en disant que $\mathrm{disc}(\mathscr{B})$ est le discriminant de la forme quadratique $w \mapsto \mathrm{Tr}_K(w^2)$ dans la base $\mathscr{B}$. Cette forme quadratique est donc non dégénérée. De plus, le corollaire précédent n'est autre que la formule classique comparant les discriminants d'une forme quadratique dans deux bases.

Exercices

1) a) Soit $P = a_n X^n + \cdots + a_0$ un polynôme à coefficients entiers, où $a_n \neq 0$. Soit r un nombre rationnel, donné par la fraction irréductible u/v, tel que $P(r) = 0$. Montrer que u divise a_0 et que v divise a_n.

b) Soit x un nombre rationnel. Montrer que $2\cos(\pi x)$ est un entier algébrique.

c) Soit x un nombre rationnel appartenant à $[0, 1]$ tel que $\cos(\pi x)$ soit aussi rationnel. Montrer que x appartient à l'ensemble $\{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$.

2) a) Est-ce que $\dfrac{2}{\sqrt{13}+3}$ est un entier algébrique ?

b) Donner le polynôme minimal sur $\mathbf{Q}$ de $\dfrac{-2+\sqrt{2}+i\sqrt{2}}{2}$. Est-ce un entier algébrique ?

3) Soit P un polynôme unitaire à coefficients entiers. On suppose que toutes les racines complexes de P sont de modulo $\leqslant 1$.

a) Soit α une racine complexe de P. Démontrer que pour tout entier $n \geqslant 0$, α^n est un entier algébrique de taille $\leqslant 1$.

En déduire que $\alpha = 0$ ou α est une racine de l'unité.

b) Si P est irréductible, démontrer que $P = X$ ou P est un polynôme cyclotomique. (*Théorème de Kronecker.*)

4) Soit K le corps de nombres $\mathbf{Q}(\alpha)$, où α est une racine du polynôme $P = X^3 + X - 3$.

a) Démontrer que le polynôme P est irréductible. (Utiliser au moins deux méthodes, par exemple en recherchant des racines rationnelles et en réduisant modulo un nombre premier p). En déduire que K est de degré 3.

b) Calculer les polynômes caractéristiques de la multiplication par α et α^2 dans K. En déduire $\mathrm{N}_K(\alpha^2)$ et $\mathrm{Tr}_K(\alpha^2)$.

c) Démontrer que pour tout $x \in \mathbf{Q}$, la norme de $x - \alpha$ est égale à $P(x)$.

5) Soit P un polynôme à coefficients complexes, noté

$$P = (X - \alpha_1)\dots(X - \alpha_n) = X^n + a_{n-1}X^{n-1} + \cdots + a_0.$$

a) Si $n \geqslant 2$, calculer $\sum_{j=1}^n \alpha_j^2$ en fonction de a_{n-1} et a_{n-2}.

b) Si $n \geqslant 3$, calculer $\sum_{j=1}^n \alpha_j^3$ en fonction de a_{n-1}, a_{n-2} et a_{n-3}.

c) On suppose que $P \in \mathbf{Q}[X]$ est irréductible ; on note α une de ses racines et l'on pose $K = \mathbf{Q}(\alpha)$. Calculer $\mathrm{Tr}_K(\alpha)$, $\mathrm{Tr}_K(\alpha^2)$ et $\mathrm{Tr}_K(\alpha^3)$ en fonction des coefficients de P.

6) a) Soit α un nombre algébrique et soit P son polynôme minimal. On pose $K = \mathbf{Q}(\alpha)$. Démontrer que $\mathrm{disc}(P) = \mathrm{N}_K(P'(\alpha))$.

b) Calculer le discriminant d'un trinôme du second degré $X^2 + pX + q$.

c) Calculer le discriminant d'un trinôme $X^n + pX + q$.

§4.2. L'anneau des entiers d'un corps de nombres ; ordres

A. L'anneau des entiers est un réseau

Soit K un corps de nombres et soit $\mathbf{Z}_K$ l'ensemble des éléments de K qui sont des entiers algébriques ; on dit que c'est l'*anneau des entiers de* K.

Notons $n = [K : \mathbf{Q}]$. Notons $\varphi_1, \dots, \varphi_n$ les n plongements de K dans $\mathbf{C}$ de sorte que $\varphi_1, \dots, \varphi_{r_1}$ soient les plongements réels, que $\varphi_{r_1+1}, \dots, \varphi_{r_1+r_2}$ soient des plongements complexes deux à deux non conjugués et que $\varphi_{r_1+r_2+k} = \overline{\varphi_{r_1+k}}$ si k est un entier tel que $1 \leqslant k \leqslant r_2$. On a donc $n = r_1 + 2r_2$. Soit V l'espace vectoriel réel $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ et soit $\Phi\colon K \to V$ l'application définie par

$$\Phi(z) = (\varphi_1(z), \dots, \varphi_{r_1}(z), \varphi_{r_1+1}(z), \dots, \varphi_{r_1+r_2}(z)).$$

C'est une application $\mathbf{Q}$-linéaire de K dans V ; elle est injective car φ_1 est injective.

THÉORÈME 4.2.1. — *L'image de* $\mathbf{Z}_K$ *par l'application* Φ *est un réseau de rang* n *de l'espace vectoriel* V.

Démonstration. — Démontrons d'abord que l'image de $\mathbf{Z}_K$ est un sous-groupe discret de V. D'après la proposition 3.2.1, il suffit de démontrer que pour tout nombre réel R, l'ensemble des $z \in \mathbf{Z}_K$ tels que $\|\Phi(z)\| \leqslant R$ est fini. Le choix de la norme sur V n'importe pas, nous supposerons que l'on a

$$\left\|(x_1,\dots,x_{r_1},x_{r_1+1},\dots,x_{r_1+r_2})\right\| = \max(|x_1|,\dots,\left|x_{r_1}\right|,\left|x_{r_1+1}\right|,\dots,\left|x_{r_1+r_2}\right|)$$

pour tout $(x_1,\dots,x_{r_1+r_2}) \in \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$. Soit $z \in \mathbf{Z}_K$, soit d son degré et soit $z_1,\dots,z_d$ ses conjugués. Les ensembles $\{z_1,\dots,z_d\}$ et $\{\varphi_1(z),\dots,\varphi_n(z)\}$ coïncident. Par conséquent,

$$\max(|z_1|,\dots,|z_d|) = \max(\left|\varphi_1(z)\right|,\dots,\left|\varphi_n(z)\right|) = \max(\left|\varphi_1(z)\right|,\dots,\left|\varphi_{r_1+r_2}(z)\right|)$$

puisque $\left|\varphi_{r_1+r_2+k}(z)\right| = \left|\varphi_{r_1+k}(z)\right|$ si $1 \leqslant\leqslant n$. De plus, le dénominateur de z est égal à 1, car z est un entier algébrique. Ainsi, si $z \in \mathbf{Z}_K$, sa taille est égale à $\|\Phi(z)\|$. D'après la proposition 4.1.8, il n'y a donc qu'un nombre fini d'entiers algébriques tels que $\|\Phi(z)\| \leqslant R$ et $\Phi(\mathbf{Z}_K)$ est un sous-groupe discret de V.

D'après le théorème 3.2.3, $\Phi(\mathbf{Z}_K)$ est un réseau de V. Son rang est la dimension maximale d'une famille linéairement indépendante de $\Phi(\mathbf{Z}_K)$; il est inférieur ou égal à $\dim_{\mathbf{R}} V = n$. Or, K est un espace vectoriel de dimension n sur $\mathbf{Q}$; considérons-en une base $(\alpha_1,\dots,\alpha_n)$ et soit D le ppcm des dénominateurs des α_i. Alors, $(D\alpha_1,\dots,D\alpha_n)$ est une base de K sur $\mathbf{Q}$ formée d'entiers algébriques. L'image de ces éléments par Φ est une famille libre de cardinal n de $\Phi(\mathbf{Z}_K)$. Par conséquent, $\Phi(\mathbf{Z}_K)$ est de rang $n = [K : \mathbf{Q}]$. □

COROLLAIRE 4.2.2. — *Soit K un corps de nombres, soit n son degré et soit $\mathbf{Z}_K$ l'anneau des entiers de K. En tant que groupe abélien, $\mathbf{Z}_K$ est isomorphe à $\mathbf{Z}^n$.*

B. Ordres

DÉFINITION 4.2.1. — *Soit K un corps de nombres et soit A un sous-anneau de K. On dit que A est un ordre de K s'il vérifie les deux propriétés suivantes :*

— *comme $\mathbf{Q}$-espace vectoriel, A engendre K ;*

— *comme groupe abélien, A est de type fini.*

Exemples 4.2.2. — a) D'après le corollaire précédent, $\mathbf{Z}_K$ est un ordre de K.

b) Soit α un entier algébrique tel que $K = \mathbf{Q}(\alpha)$; posons $n = [K : \mathbf{Q}]$. Démontrons que $\mathbf{Z}[\alpha]$ est un ordre de K.

Puisque A contient une base de K comme $\mathbf{Q}$-espace vectoriel, il engendre K. Soit en outre P le polynôme minimal de α ; c'est un polynôme à coefficients entiers, unitaire, de degré d. Par suite, pour tout entier $k \geqslant 0$, α^k est combinaison linéaire à coefficients entiers de $1,\alpha,\dots,\alpha^{n-1}$; ainsi, $\mathbf{Z}[\alpha]$ est engendré comme groupe abélien par $1,\alpha,\dots,\alpha^{n-1}$.

Remarques 4.2.3. — a) *Tout ordre de K est contenu dans $\mathbf{Z}_K$.* Soit en effet A un ordre de K et soit $z_1,\dots,z_n$ une famille génératrice de A comme groupe abélien. Soit w un

élément de A ; définissons une matrice $M = (m_{ij}) \in \mathrm{Mat}_n(\mathbf{Z})$ de sorte que pour tout $j \in \{1,\dots,n\}$, on ait $wz_j = \sum_{i=1}^n a_{ij} z_i$. Soit P le polynôme caractéristique de M. D'après la proposition 2.1.3, P annule l'endomorphisme μ_w de A donné par la multiplication par w ; en particulier, $P(\mu_w)(1) = 0$, c'est-à-dire $P(w) = 0$. Comme P est un polynôme unitaire à coefficients entiers, w est un entier algébrique, c'est-à-dire $w \in \mathbf{Z}_K$.

b) Inversement, *un sous-groupe de $\mathbf{Z}_K$ qui contient 1 et est stable par produit est un ordre de K dès qu'il contient une base de K*. Sous ces hypothèses en effet, un tel sous-groupe abélien est un sous-anneau de K et engendre K comme espace vectoriel. De plus, étant un sous-groupe d'un groupe isomorphe à $\mathbf{Z}^n$, il est de type fini (REFERENCE).

COROLLAIRE 4.2.4. — *Soit K un corps de nombres, soit n son degré et soit A un ordre de K. Comme groupe abélien, A est isomorphe à $\mathbf{Z}^n$.*

Démonstration. — En effet, A est contenu dans $\mathbf{Z}_K$ d'après la remarque ci-dessus. C'est en particulier un groupe abélien libre, c'est-à-dire isomorphe à $\mathbf{Z}^m$ pour un entier m tel que $0 \leqslant m \leqslant n$. Comme A engendre K comme $\mathbf{Q}$-espace vectoriel, il contient n éléments linéairement indépendants et l'on a $m = n$. □

C. Discriminant d'un ordre

Soit K un corps de nombres et A un ordre de K. Soit $\mathscr{B}$ et $\mathscr{B}'$ des bases du groupe abélien A. On a donc $\det_{\mathscr{B}}(\mathscr{B}') = \pm 1$ si bien que $\mathrm{disc}(\mathscr{B}) = \mathrm{disc}(\mathscr{B}')$.

DÉFINITION 4.2.1. — *La valeur commune de ces discriminants est appelée* discriminant de l'ordre *A et notée* $\mathrm{disc}(A)$. *Par abus de langage, on appelle discriminant de K le discriminant de $\mathbf{Z}_K$.*

PROPOSITION 4.2.2. — *Soit K un corps de nombres, soit A un ordre de K. On a l'égalité* $\mathrm{disc}(A) = \mathrm{disc}(K)\,\mathrm{card}(\mathbf{Z}_K/A)^2$.

Démonstration. — Notons n le degré de K. D'après le théorème de la base adaptée, il existe une base $(w_1,\dots,w_n)$ de $\mathbf{Z}_K$ et des entiers non nuls $d_1,\dots,d_n$ tels que $(d_1 w_1,\dots,d_n w_n)$ soit une base de A. On a donc $\mathrm{disc}(A) = (d_1\dots d_n)^2\,\mathrm{disc}(K)$. En outre, $\mathbf{Z}_K/A$ est un groupe abélien isomorphe au produit des groupes $\mathbf{Z}/d_i\mathbf{Z}$, donc est de cardinal $d_1\dots d_n$. La proposition est ainsi démontrée. □

COROLLAIRE 4.2.3. — *Si* $\mathrm{disc}(A)$ *est sans facteur carré, alors $A = \mathbf{Z}_K$.*

Exemple 4.2.4. — Soit K le corps $\mathbf{Q}(\alpha)$ où α est une racine du polynôme $P = X^3 + X + 1$. Comme P est irréductible modulo 2 (il est de degré 3 et n'a pas de racine dans $\mathbf{Z}/2\mathbf{Z}$), P est irréductible dans $\mathbf{Q}[X]$ et est le polynôme minimal de α. Ainsi, $[K : \mathbf{Q}] = 3$.

De plus, α est un entier algébrique, donc $\mathbf{Z}[\alpha]$ est un ordre de K. Son discriminant est donné par

$$\operatorname{disc}(\mathbf{Z}[\alpha]) = \det\begin{pmatrix} \operatorname{Tr}_K(1) & \operatorname{Tr}_K(\alpha) & \operatorname{Tr}_K(\alpha^2) \\ \operatorname{Tr}_K(\alpha) & \operatorname{Tr}_K(\alpha^2) & \operatorname{Tr}_K(\alpha^3) \\ \operatorname{Tr}_K(\alpha^2) & \operatorname{Tr}_K(\alpha^3) & \operatorname{Tr}_K(\alpha^4) \end{pmatrix}.$$

On a $\operatorname{Tr}_K(1) = 3$ et $\operatorname{Tr}_K(\alpha) = 0$ (terme de degre 2 du polynôme minimal de α). On a $\alpha^4 = -\alpha - \alpha^2$ et $\alpha^5 = -\alpha^2 - \alpha^3 = -\alpha^2 + \alpha + 1$. Dans la base $(1, \alpha, \alpha^2)$ de K, la matrice de la multiplication par α^2 est donc

$$\begin{pmatrix} 0 & -1 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

donc $\operatorname{Tr}_K(\alpha^2) = -2$. Par suite, $\operatorname{Tr}_K(\alpha^3) = \operatorname{Tr}_K(-1-\alpha) = -3$ et $\operatorname{Tr}_K(\alpha^4) = \operatorname{Tr}_K(-\alpha-\alpha^2) = 2$. Ainsi,

$$\operatorname{disc}(\mathbf{Z}[\alpha]) = \det\begin{pmatrix} 3 & 0 & -2 \\ 0 & -2 & -3 \\ -2 & -3 & 2 \end{pmatrix} = 3(-4-9) - 2(-4) = -39 + 8 = -31.$$

Comme $\operatorname{disc}(\mathbf{Z}[\alpha])$ est sans facteur carré, on en déduit que $\mathbf{Z}_K = \mathbf{Z}[\alpha]$.

Dans la suite de ce paragraphe, on relie $\operatorname{disc}(A)$ au volume du domaine fondamental du réseau de $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ associé à A.

Soit $\varphi_1, \ldots, \varphi_{r_1}$ les plongements réels de K et $\varphi_{r_1+1}, \ldots, \varphi_{r_1+r_2}$ des plongements complexes, deux à deux non conjugués. Si n désigne le degré de K, on a donc $n = r_1 + 2r_2$. Soit $\Phi\colon K \to \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ l'application donnée par $w \mapsto (\varphi_j(w))_{1 \leqslant j \leqslant r_1+r_2}$. On a vu que Φ est $\mathbf{Q}$-linéaire, injective. En outre, $\Phi(A)$ est contenu dans $\Phi(\mathbf{Z}_K)$, donc est un sous-groupe discret de $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$. Par ailleurs, A contient une base de K, donc $\Phi(A)$ contient n vecteurs linéairement indépendants sur $\mathbf{Q}$. Cela entraîne que $\Phi(A)$ est un réseau de rang n de l'espace vectoriel $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$.

PROPOSITION 4.2.5. — *Identifions $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ à $\mathbf{R}^n$ par l'application*

$$\theta\colon (x_1, \ldots, x_{r_1}, z_{r_1+1}, \ldots, z_{r_1+r_2}) \mapsto (x_1, \ldots, x_{r_1}, \Re(z_{r_1+1}), \Im(z_{r_1+1}), \ldots, \Re(z_{r_1+r_2}), \Im(z_{r_1+r_2}))$$

et soit P un domaine fondamental de $\Phi(A)$ dans $\mathbf{R}^n$. On a $|\operatorname{disc}(A)| = 4^{r_2} \operatorname{vol}(P)^2$.

Démonstration. — Pour tout entier j tel que $1 \leqslant j \leqslant r_2$, posons $\varphi_{r_1+r_2+j} = \overline{\varphi_{r_1+j}}$. Soit $w_1, \ldots, w_n$ une base de A. Par définition, on a

$$\operatorname{disc}(A) = \det\big(\varphi_j(w_k)\big)^2.$$

Soit D la matrice $(\varphi_j(w_k))$. Si j vérifie $1 \leqslant j \leqslant r_2$, ses lignes d'indices $r_1 + j$ et $r_1 + r_2 + j$ sont conjuguées l'une de l'autre. L'opération $L_{r_1+j} \leftarrow (L_{r_1+j} + L_{r_1+r_2+j})/2$ remplace la ligne L_{r_1+j} par sa partie réelle et divise le déterminant par 2. L'opération $L_{r_1+r_2+j} \leftarrow i(L_{r_1+r_2+j} - L_{r_1+j})$ remplace alors la ligne $L_{r_1+r_2+j}$ par la partie imaginaire de l'ancienne

ligne L_{r_1+j} et multiplie le déterminant par i. Par r_2 échanges de lignes, on obtient donc une matrice D' et $\left|\det(D')\right|$ est égal au volume du domaine fondamental $P = [0,1]w_1 + \cdots + [0,1]w_n$ de $\Phi(A)$ dans $\mathbf{R}^n$. On a donc $\mathrm{vol}(P) = \left|\det(D')\right|$, d'où $\det(D')^2 = \mathrm{vol}(P)^2$. Comme $\det(D') = (-i/2)^{r_2}\det(D)$, on a

$$\mathrm{vol}(P)^2 = (-1/4)^{r_2}\det(D)^2 = (-1/4)^{r_2}\,\mathrm{disc}(A),$$

autrement dit $\mathrm{disc}(A) = (-4)^{r_2}\mathrm{vol}(P)^2$. □

COROLLAIRE 4.2.6. — *Soit K un corps de nombres ; le signe du discriminant de tout ordre de K est égal à $(-1)^{r_2}$, où r_2 est la moitié du nombre de plongements complexes de K.*

Exercices

7) Soit K un corps de nombres de degré n. Parmi toutes les bases $(\alpha_1,\ldots,\alpha_n)$ de K formées d'entiers algébriques, on en choisit une dont la valeur absolume du discriminant, $D = \mathrm{disc}(\alpha_1,\ldots,\alpha_n)$ est minimale.

a) Soit ω un élément de $\mathbf{Z}_K$ tel que l'on puisse écrire $\omega = x_1\alpha_1 + \cdots + x_n\alpha_n$, avec $x_1,\ldots,x_n \in \mathbf{Q}$ et $0 < x_1 < 1$. Démontrer que $(\omega,\alpha_2,\ldots,\alpha_n)$ est une base de K formée d'entiers algébriques. Calculer son discriminant en fonction de D et en déduire une contradiction.

b) Démontrer que $(\alpha_1,\ldots,\alpha_n)$ est une base de $\mathbf{Z}_K$.

8) a) Soit $P = X^3 - X - 1$ et soit α une racine complexe de P. Calculer le discriminant de l'ordre $\mathbf{Z}[\alpha]$ de K. En déduire que $A = \mathbf{Z}_K$.

b) Mêmes questions avec les polynômes $X^4 - X - 1$ et $X^4 - 2 * X^2 - X - 1$.

9) Soit P le polynôme $X^3 - X - 4$ et soit α une racine complexe de P. Soit K le corps de nombres $\mathbf{Q}(\alpha)$.

a) Démontrer que $A = \mathbf{Z}[\alpha]$ est un ordre de K ; calculer son discriminant.

b) Calculer le polynôme minimal de $\omega = (\alpha^2 + \alpha)/2$.

c) Démontrer que $\mathbf{Z} + \mathbf{Z}\alpha + \mathbf{Z}\omega$ est un sous-groupe abélien de $\mathbf{Z}_K$ (c'est en fait un ordre de K). Calculer le discriminant de la base $(1,\alpha,\omega)$ et en déduire que $(1,\alpha,\omega)$ est une base de l'anneau des entiers de K.

10) Soit P le polynôme $X^3 - X - 2$ et α une racine de P. Soit K le corps de nombres $\mathbf{Q}(\alpha)$.

a) Calculer le discriminant de l'ordre $A = \mathbf{Z}[\alpha]$ de K. Si $A \neq \mathbf{Z}_K$, démontrer que le groupe abélien $\mathbf{Z}_K/A$ est isomorphe à $\mathbf{Z}/2\mathbf{Z}$.

b) Calculer les traces de $1, \alpha, \alpha^2$.

c) Soit u, v, w des entiers relatifs ; écrire la matrice de la multiplication par $u + v\alpha + w\alpha^2$ dan la base $(1,\alpha,\alpha^2)$. En déduire que $\mathrm{N}_K(u + v\alpha + w\alpha^2)$ est de même parité que u.

d) Démontrer que $(1+\alpha)/2$, $\alpha^2/2$ et $(1+\alpha+\alpha^2)/2$ ne sont pas des entiers algébriques.

e) On suppose par l'absurde que $\mathbf{Z}_K \neq A$. Démontrer qu'il existe des entiers u, v, w, pas tous pairs, tels que $\omega = (u + v\alpha + w\alpha^2)/2$ appartienne à $\mathbf{Z}_K$. En considérant $\mathrm{Tr}_K(\omega)$, vérifier que u et v sont de même parité puis obtenir une contradiction.

11) Soit p un nombre premier ($p > 2$) et soit ω une racine primitive p-ième de l'unité dans $\mathbf{C}$. On pose $K = \mathbf{Q}(\omega)$.

a) Quel est le polynôme minimal de ω ? Quel est le degré $[K : \mathbf{Q}]$? Démontrer que l'on peut numéroter les $(p-1)$ plongements de K dans $\mathbf{C}$ de sorte que $\varphi_i(\omega) = \omega^i$ pour $i \in \{1, \dots, p-1\}$.

b) Calculer $\mathrm{Tr}_K(\omega)$ et $\mathrm{N}_K(\omega)$. Calculer aussi $\mathrm{N}_K(1-\omega)$.

c) Soit $\alpha \in \mathbf{Z}_K$ et soit $\alpha_1, \dots, \alpha_{p-1}$ ses conjugués ; ce sont des éléments de $\mathbf{Z}_K$. Observer que $\varphi_i((1-\omega)\alpha) = (1-\omega^i)\alpha_i$ pour tout i, puis démontrer que que $\mathrm{Tr}_K(\alpha(1-\omega))$ appartient à $(1-\omega)\mathbf{Z}_K$. En déduire que $\mathrm{Tr}_K(\alpha(1-\omega)) \in p\mathbf{Z}$.

d) On veut démontrer que $\mathbf{Z}_K = \mathbf{Z}[1-\omega]$. Soit $\alpha \in \mathbf{Z}_K$; démontrer qu'il existe des nombres rationnels $z_1, \dots, z_{p-1}$ tels que $\alpha = \sum_{i=1}^{p-1} z_j (1-\omega)^{j-1}$.

e) Démontrer que les z_j sont solution d'un système linéaire à coefficients dans $\mathbf{Z}_K$ dont le déterminant est p^{p-2}. En déduire qu'il existe des entiers relatifs $n_1, \dots, n_{p-1}$ tels que $z_j = n_j / p^{(p-3)/2}$.

f) Soit $\alpha \in \mathbf{Z}_K$; on suppose que $p\alpha \in \mathbf{Z}[1-\omega]$. En considérant $\mathrm{Tr}_K(p\alpha/(1-\omega)^j)$, démontrer par récurrence sur j que z_j est multiple de p pour tout p puis que $\alpha \in \mathbf{Z}[1-\omega]$.

g) Démontrer que $\mathbf{Z}_K = \mathbf{Z}[1-\omega]$.

12) Soit n un entier tel que $|n| > 1$ et qui n'a pas de facteur cubique (c'est-à-dire qui n'est multiple de p^3 pour aucun nombre premier p).

Soit ω la racine cubique de n dans $\mathbf{R}$; on pose $K = \mathbf{Q}(\omega)$.

a) Quel est le polynôme minimal de ω ? Quel est son discriminant ?

b) On suppose que $n \equiv 1 \pmod 9$ et on pose $\alpha = (1 + \omega + \omega^2)/3$. Calculer $\mathrm{Tr}_K(\alpha)$, $\mathrm{N}_K(\alpha)$ ainsi que le polynôme minimal de α. En déduire que α est un entier algébrique.

c) Démontrer que $(1, \omega, \alpha)$ est une base entière de $\mathbf{Z}_K$.

13) Soit n un entier naturel tel que $n \geqslant 2$ et soit p un nombre premier. Soit $P \in \mathbf{Z}[X]$ un polynôme unitaire de degré n qui vérifie le critère d'Eisenstein en p : $P \equiv X^n \pmod p$ et $P(0)$ n'est pas multiple de p^2. On rappelle que P est alors irréductible.

Soit α une racine complexe de P et soit K le corps de nombres $\mathbf{Q}(\alpha)$. Soit A l'ordre $\mathbf{Z}[\alpha]$ de K.

a) Soit M la matrice de la multiplication par α dans la base $(1, \alpha, \dots, \alpha^{n-1})$ de K. En calculant M modulo p, démontrer que pour $(u_0, \dots, u_{n-1}) \in \mathbf{Z}^n$,

$$\mathrm{N}_K(u_0 + u_1\alpha + \cdots + u_{n-1}\alpha^{n-1}) \equiv u_0^n \pmod p.$$

b) Soit $(u_0, \dots, u_{n-1}) \in \mathbf{Z}^n$; on pose $\omega = u_0 + u_1\alpha + \cdots + u_{n-1}\alpha^{n-1}$ et on suppose que ω/p est un entier algébrique. Démontrer que par récurrence que pour tout $i \in \{0, \dots, n-1\}$, u_i est divisible par p.

c) Démontrer que l'indice de A dans $\mathbf{Z}_K$ n'est pas multiple de p.

d) On suppose que $P = X^3 - 5X - 5$. Démontrer que A est l'anneau des entiers de K.

14) Soit K un corps de nombres de degré n, soit $\mathbf{Z}_K$ son anneau d'entiers. On note $\varphi_1, \dots, \varphi_n$ les plongements de K dans $\mathbf{C}$.

a) Soit z un élément de K tel que $\varphi_i(z) = \varphi_1(z)$ pour tout $i \in \{1, \dots, n\}$; démontrer que $z \in \mathbf{Q}$.

b) Soit $\alpha_1, \dots, \alpha_n$ des éléments de $\mathbf{Z}_K$. On pose

$$a = \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \varepsilon(\sigma) = 1}} \prod_{j=1}^{n} \varphi_{\sigma(j)}(\alpha_j) \quad \text{et} \quad b = \sum_{\substack{\sigma \in \mathfrak{S}_n \\ \varepsilon(\sigma) = -1}} \prod_{j=1}^{n} \varphi_{\sigma(j)}(\alpha_j),$$

où les sommes concernent les permutations de $\{1,\dots,n\}$ qui sont respectivement paires et impaires. Démontrer que $\mathrm{disc}(\alpha_1,\dots,\alpha_n) = (a-b)^2$.

c) Démontrer que $a+b$ et ab sont des entiers relatifs

d) En déduire que $\mathrm{disc}(\alpha_1,\dots,\alpha_n)$ est congru à 0 ou 1 modulo 4. En particulier, $\mathrm{disc}(K)$ est congru à 0 ou 1 modulo 4. (*Théorème de Stickelberger*; démonstration d'I. Schur.)

15) Soit K un corps de nombres de degré n; soit $(\alpha_1,\dots,\alpha_n)$ une base de $\mathbf{Z}_K$. Notons $\varphi_1,\dots,\varphi_n$ les plongements de K dans $\mathbf{C}$.

a) Pour $(x_1,\dots,x_n) \in \mathbf{R}^n$, on pose

$$\|(x_1,\dots,x_n)\| = \max_{1\leqslant j\leqslant n} \left(\left|x_1\varphi_j(\alpha_1)+\cdots+x_n\varphi_j(\alpha_n)\right|\right).$$

Démontrer que c'est une norme sur $\mathbf{R}^n$ et que le volume de sa boule unité est égal à $\sqrt{|\mathrm{disc}(K))|}$.

b) À l'aide du théorème de Minkowski, en déduire qu'il existe un élément non nul de $\mathbf{Z}_K$ de norme $< \sqrt{|\mathrm{disc}(K)|}$.

c) Démontrer que $|\mathrm{disc}(K)| > 1$ (*théorème de Minkowski*).

d) En vue d'améliorer cette inégalité, on considère la norme suivante sur $\mathbf{R}^n$:

$$\|(x_1,\dots,x_n)\| = \sum j = 1^n \left|x_1\varphi_j(\alpha_1)+\cdots+x_n\varphi_j(\alpha_n)\right|.$$

Démontrer que sa boule unité a pour volume $2^{r_1}(\pi/2)^{r_2}/n!$.

e) Soit $(z_1,\dots,z_n)$ des nombres complexes. Démontrer l'inégalité arithmético-géométrique :

$$|z_1\dots z_n| \leqslant \left(\frac{1}{n}\sum_{i=1}^{n}|z_i|\right)^n.$$

f) Démontrer que $|\mathrm{disc}(K)| > (\pi/4)^{2r_2}(n^n/n!)^2$ et vérifier que cette minoration améliore la précédente.

§4.3. Le théorème des unités

Soit K un corps de nombres. Dans ce paragraphe, on étudie le groupe des éléments inversibles (ou *unités*) de l'anneau $\mathbf{Z}_K$.

On note n le degré de K, r_1 le nombre de plongements réels de K, r_2 la moitié du nombre de plongements complexes de K et on fixe des plongements $\varphi_1,\dots,\varphi_{r_1+r_2}$ comme précédemment. On note aussi $\Phi\colon K \to \mathbf{R}^{r_1}\times\mathbf{C}^{r_2}$ l'application $w \mapsto (\varphi_j(w))$.

A. L'application logarithmique et son noyau

Notons Λ l'application de K^* dans $\mathbf{R}^{r_1+r_2}$ donnée par

$$\Lambda(w) = (\log\left|\varphi_1(w)\right|,\dots,\log\left|\varphi_{r_1}(w)\right|,\log\left|\varphi_{r_1+1}(w)\right|^2,\dots,\log\left|\varphi_{r_1+r_2}(w)\right|^2).$$

C'est un homomorphisme de groupes qu'on appelle l'*application logarithmique.*

PROPOSITION 4.3.1. — *L'intersection de* $\mathrm{Ker}(\Lambda)$ *avec* $\mathbf{Z}_K$ *est l'ensemble des racines de l'unités contenues dans K ; c'est un groupe fini.*

Démonstration. — Soit $w \in \mathrm{Ker}(\Lambda) \cap \mathbf{Z}_K$; on a donc $\left|\varphi_j(w)\right| = 1$ pour tout j et la taille de w est égale à 1. Comme il n'y a qu'un nombre fini de nombres algébriques de K de taille bornée, $\mathrm{Ker}(\Lambda) \cap \mathbf{Z}_K$ est un ensemble fini.

Si $w \in \mathrm{Ker}(\Lambda) \cap \mathbf{Z}_K$, il existe donc des entiers k et $d > 0$ tels que $w^k = w^{k+d}$; puisque $w \neq 0$, on a $w^d = 1$ et w est une racine de l'unité. Inversement, si w est une racine de l'unité contenue dans K, il existe un entier $d > 0$ tel que $w^d = 1$ et $\left|\varphi_j(w)\right| = 1$ pour tout j ; en outre, w est un entier algébrique. Par suite, $w \in \mathrm{Ker}(\Lambda) \cap \mathbf{Z}_K$. □

B. Le théorème des unités

Notons $\mathbf{U}_K$ le groupe des éléments inversibles de $\mathbf{Z}_K$, qu'on appelle aussi *unités*. Rappelons qu'un élément w de $\mathbf{Z}_K$ est inversible si et seulement si $\mathrm{N}_K(w) = \pm 1$. Ainsi, si $w \in \mathbf{U}_K$, $\Lambda(w)$ est un élément de l'hyperplan H de $\mathbf{R}^{r_1+r_2}$ d'équation $t_1 + \cdots + t_{r_1+r_2} = 0$.

THÉORÈME 4.3.1. — *L'image de* $\mathbf{U}_K$ *par* Λ *est un réseau de rang* $r_1 + r_2 - 1$ *de* H.

COROLLAIRE 4.3.2 (Théorème des unités de Dirichlet). — *Soit K un corps de nombres ; soit μ_K le groupe fini des racines de l'unité de K. Le groupe $\mathbf{U}_K$ est isomorphe au produit $\mu_K \times \mathbf{Z}^{r_1+r_2-1}$.*

Démonstration. — Posons $r = r_1 + r_2 - 1$ et soit $w_1, \ldots, w_r$ des éléments de $\mathbf{U}_K$ tels que $\Lambda(w_1), \ldots, \Lambda(w_r)$ forment une base de $\Lambda(\mathbf{U}_K)$. Démontrons que l'application θ de $\mu_K \times \mathbf{Z}^r$ dans $\mathbf{U}_K$ donnée par $(z, n_1, \ldots, n_r) \mapsto z w_1^{n_1} \ldots w_r^{n_r}$ est un isomorphisme.

Soit $(z, n_1, \ldots, n_r)$ un élément du noyau de θ. On a $\Lambda(\theta(z, n_1, \ldots, n_r)) = n_1\Lambda(w_1) + \cdots + n_r\Lambda(w_r)$. Par suite, $n_1 = \cdots = n_r = 0$ puis $z = 1$. Autrement dit, θ est injectif.

Soit alors $w \in \mathbf{U}_K$. Soit $n_1, \ldots, n_r$ des entiers relatifs tels que $\Lambda(w) = n_1\Lambda(w_1) + \cdots + n_r\Lambda(w_r)$. Alors, $w w_1^{-n_1} \ldots w_r^{-n_r}$ est une unité w' telle que $\Lambda(w') = 0$; on a donc $w' \in \mu_K$ ce qui entraîne que θ est surjectif. □

La fin de ce paragraphe est consacrée à la démonstration du théorème 4.3.1.

Démontrons d'abord que $\Lambda(\mathbf{U}_K)$ est un sous-groupe discret de $\mathbf{R}^{r_1+r_2}$. Il suffit pour cela de vérifier que pour tout nombre réel $c > 0$ l'ensemble des éléments $w \in \mathbf{U}_K$ tels que $\left|\log\left|\varphi_j(w)\right|\right| \leqslant c$ est fini, car cela démontrera que l'ensemble des éléments $(t_1, \ldots, t_{r_1+r_2})$ de $\Lambda(\mathbf{U}_K)$ tels que $\left|t_j\right| \leqslant c$ pour tout j est un ensemble fini.

Or, l'inégalité $\left|\log\left|\varphi_j(w)\right|\right| \leqslant c$ implique $\left|\varphi_j(w)\right| \leqslant e^c$; comme c'est un entier algébrique, la taille d'une telle unité est donc majorée par e^c. On conclut en rappelant qu'il n'y a qu'un nombre fini de nombres algébriques de K de taille majorée par e^c.

Comme $\Lambda(\mathbf{U}_K)$ est contenu dans l'hyperplan H, il reste à démontrer que $\Lambda(\mathbf{U}_K)$ contient $r_1 + r_2 - 1$ éléments linéairement indépendants, c'est-à-dire à construire $r_1 + r_2 - 1$ unités multiplicativement indépendantes.

PROPOSITION 4.3.3. — *Pour tout entier $j \in \{1, \ldots, r_1 + r_2\}$, il existe une unité $w_j \in \mathbf{U}_K$ telle que* $\log\left|\varphi_k(w_j)\right| < 0$ *si* $k \neq j$.

Bien sûr, la relation $\left|\mathrm{N}_K(w_j)\right| = 1$ entraîne alors que $\log\left|\varphi_j(w_j)\right| > 0$.

Démonstration. — Soit $c = (c_1, \dots, c_{r_1+r_2})$ une famille de nombres réels strictement positifs et soit $\|\cdot\|_c$ la norme sur $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ donnée par

$$\left\|(x_1, \dots, x_{r_1}, z_{r_1+1}, \dots, z_{r_1+r_2})\right\|_c = \max\left(|x_1|\, c_1^{-1}, \dots, \left|x_{r_1}\right| c_{r_1}^{-1} \left|z_{r_1+1}\right| c_{r_1+1}^{-1/2}, \dots, \left|z_{r_1+r_2}\right| c_{r_1+r_2}^{-1/2}\right).$$

La boule unité B_c de cet espace vectoriel est l'ensemble des $(x_1, \dots, z_{r_1+r_2})$ tel que $|x_k| \leqslant c_k$ pour $1 \leqslant k \leqslant r_1$ et et $\left|z_{r_1+k}\right| \leqslant c_{r_1+k}^{1/2}$ pour $1 \leqslant k \leqslant r_2$. Sont volume est donc

$$\mathrm{vol}(B_c) = c_1 \dots c_{r_1}(\pi c_{r_1+1}) \dots (\pi c_{r_1+r_2}) = \pi^{r_2}(c_1 \dots c_{r_1+r_2}).$$

Soit P un domaine fondamental du réseau $\Phi(\mathbf{Z}_K)$ de $\mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$. D'après le théorème de Minkowski, il existe, si $\mathrm{vol}(B_c) = 2^n \mathrm{vol}(P)$, un entier algébrique non nul $a_c \in \mathbf{Z}_K$ tel que $\Phi(a_c) \in B_c$. On a en outre

$$\begin{aligned} |\mathrm{N}_K(a_c)| &= \left|\varphi_1(a_c)\right| \dots \left|\varphi_{r_1}(a_c)\right| \left|\varphi_{r_1+1}(a_c)\right|^2 \dots \left|\varphi_{r_1+r_2}(a_c)\right|^2 \\ &\leqslant c_1 \dots c_{r_1+r_2} = \pi^{-r_2}\,\mathrm{vol}(B_c) \\ &\leqslant 2^n \pi^{-r_2}\,\mathrm{vol}(P). \end{aligned}$$

On a toute latitude pour choisir le paramètre c ; on choisit tous les c_k (pour $k \neq j$) égaux à un nombre réel $t > 0$, supposé petit et on pose ensuite $c_j = \pi^{-r_2} 2^n \,\mathrm{vol}(P) / t^{r_1+r_2-1}$. Posons $a_t = a_c$.

Lorsque t tend vers 0, $\left|\varphi_k(a_t)\right|$ tend vers 0 si $k \neq j$ et $\left|\varphi_j(a_t)\right|$ tend vers $+\infty$; cela prouve que l'on a construit une infinité d'entiers algébriques distincts dont les normes sont toutes majorées par un même nombre réel. Comme ces normes sont des entiers, on a ainsi construit une suite infinie d'entiers algébriques de K de même norme N. Le groupe abélien $\mathbf{Z}_K / N\mathbf{Z}_K$ est isomorphe à $(\mathbf{Z}/N\mathbf{Z})^n$; il est en particulier fini. Quitte à considérer une sous-suite de la suite précédente, on voit qu'il existe une suite infinie $(a_{t_m})_m$ d'entiers algébriques non nuls de K de même norme N et qui sont deux à deux congrus modulo N.

LEMME 4.3.4. — *Soit a et b des éléments non nuls de $\mathbf{Z}_K$ et N un entier relatif. On suppose que $\mathrm{N}_K(a) = \mathrm{N}_K(b) = N$ et que $a - b \in N\mathbf{Z}_K$. Alors, a/b est une unité de K.*

Démonstration. — Posons $w = a/b$ et écrivons $a = b + Nc$, avec $c \in \mathbf{Z}_K$. Comme $N = \mathrm{N}_K(b)$, il existe un entier algébrique $b' \in \mathbf{Z}_K$ tel que $N = bb'$. Alors, $a = b + bb'c = b(1 + b'c)$ et $w = 1 + b'c$. Par suite, w est un entier algébrique. On démontre de même que b/a est un entier algébrique, ce qui conclut la démonstration que w est une unité de K. □

Pour tout entier $m \geqslant 0$, posons $w_m = a_{t_m} / a_{t_1}$. D'après le lemme précédent, w_m est une unité de K pour tout $m \in \mathbf{N}$; ces unités sont deux à deux distinctes. En outre, si $k \in \{1, \dots, r_1 + r_2\}$ est distinct de j, $\left|\varphi_k(w_m)\right|$ tend vers 0. Il existe en particulier une unité $u_j \in \mathbf{U}_K$ telle que $\left|\varphi_k(u_j)\right| < 1$ pour tout entier $k \in \{1, \dots, r_1 + r_2\}$ qui est distinct de j. Cela termine la démonstration de la proposition 4.3.3. □

Pour finir la démonstration du théorème des unités, il suffit de démontrer que les unités $u_1, \dots, u_{r_1+r_2-1}$ construites dans la proposition 4.3.3 sont multiplicativement indépendantes, ou encore, que leurs images par Λ sont linéairement indépendantes dans $\mathbf{R}^{r_1+r_2}$. Considérons la matrice M de taille $(r_1+r_2)\times(r_1+r_2-1)$ formée par ces images et extrayons-en une matrice carrée $M' = (m_{ij})_{1\leqslant i,j\leqslant r_1+r_2-1}$ en ôtant la dernière ligne. Si $i \neq j$, on a $m_{ij} = \log\left|\varphi_i(u_j)\right|$ (lorsque $1 \leqslant j \leqslant r_1$) et $m_{ij} = 2\log\left|\varphi_i(u_j)\right|$ sinon ; dans tous les cas, $m_{ij} < 0$. En outre, pour tout entier j, on a $m_{1j} + \cdots + m_{rj} = -m_{r_1+r_2,j} > 0$. D'après le lemme suivant, M' est inversible donc M est de rang $r_1 + r_2 - 1$, ce qu'il fallait démontrer.

LEMME 4.3.5. — *Soit M une matrice $n \times n$ à coefficients réels. On suppose que $m_{ij} < 0$ si $i \neq j$ et que la somme des coefficients de chaque colonne est strictement positive. Alors, M est inversible.*

Démonstration. — Considérons une relation de dépendance linéaire entre les lignes de M et démontrons qu'elle est triviale. Soit donc $x_1, \dots, x_n$ des nombres réels tels que

$$\sum_{i=1}^{n} x_i m_{ij} = 0 \quad \text{pour } j \in \{1, \dots, n\}.$$

Soit j un entier tel que $x_j = \max(x_1, \dots, x_n)$. Alors

$$0 = \sum_{i=1}^{n} x_i m_{ij} = \sum_{i=1}^{n} (x_i - x_j) m_{ij} + x_j \sum_{i=1}^{n} m_{ij} \geqslant x_j \sum_{i=1}^{n} m_{ij}$$

puisque $(x_i - x_j)m_{ij}$ est positif ou nul pour $i \neq j$, et est nul si $i = j$. Comme $\sum_{i=1}^{n} m_{ij} > 0$, il vient $x_j \leqslant 0$ et tous les x_i sont négatifs ou nuls. Le même argument appliqué à leurs opposés démontre qu'ils sont tous positifs ou nuls. Par suite, $x_1 = \cdots = x_n = 0$, ce qu'il fallait démontrer. □

C. L'équation de Pell-Fermat

Exercices

16) Soit d un entier strictement positif qui n'est pas le carré d'un nombre entier.

a) Montrer que $\sqrt{d}$ est irrationnel.

b) Montrer qu'il existe un entier $n \neq 0$ tel que l'équation $x^2 - dy^2 = n$ possède une infinité de solutions. On fixe un tel entier dans la suite de l'exercice.

c) Montrer qu'il existe des solutions distinctes (x_1, y_1) et (x_2, y_2) de l'équation $x^2 - dy^2 = n$ telles que $x_1 \equiv x_2 \pmod{n}$ et $y_1 \equiv y_2 \pmod{n}$.

d) Écrire la fraction $(x_1 + y_1\sqrt{d})/(x_2 + y_2\sqrt{d})$ sous la forme $x_3 + y_3\sqrt{d}$. Montrer que x_3 et y_3 sont entiers et vérifient $(x_3)^2 - d(y_3)^2 = 1$.

e) Soit $(u, v) \in \mathbf{N}^2$ une solution de l'équation $x^2 - dy^2 = 1$ où $v > 0$ est minimal. Soit $(x, y) \in \mathbf{N}^2$ une autre solution. Si $y > 0$, simplifier $(x + y\sqrt{d})/(u + v\sqrt{d})$ et construire une solution $(x', y') \in \mathbf{N}^2$ telle que $0 \leqslant y' < y$.

f) Montrer que les solutions de l'équation $x^2 - dy^2 = 1$ sont de la forme $\pm(x_n, y_n)$, pour $n \in \mathbf{Z}$, où x_n et y_n sont déterminés par la relation $x_n + y_n\sqrt{d} = (u + v\sqrt{d})^n$ (résolution de l'*équation de Pell*).

17) Soit a et b des entiers positifs tels que $ab + 1$ divise $a^2 + b^2$. Montrer que leur quotient est un carré parfait (*Olympiades internationales de mathématiques,* 1988).

18) Soit P le polynôme $X^3 - X - 5$.

a) Montrer que P a une seule racine réelle, qu'on notera α, et qu'elle est comprise entre 1 et 2.

b) Soit β et $\bar{\beta}$ les deux racines complexes de P. Vérifier que $|\beta| \leqslant 2$.

c) Soit K le corps de nombres $\mathbf{Q}(\alpha)$. En calculant son discriminant, démontrer que l'ordre $\mathbf{Z}[\alpha]$ est égal à $\mathbf{Z}_K$.

d) Démontrer que $2 - \alpha$ est une unité de $\mathbf{Z}_K$.

e) Démontrer qu'il existe un nombre réel $c > 0$ tel que l'on ait, pour tout $(x, y, z) \in \mathbf{R}^3$,

$$\max(|x + y\alpha + z\alpha^2|, |x + y\beta + z\beta^2|) \geqslant c \max(|x|, |y|, |z|).$$

Vérifier par un calcul numérique que $c = 0{,}9$ convient.

f) Soit $\omega = u + v\alpha + w\alpha^2$ une unité de $\mathbf{Z}_K$ telle que $1 < \omega < 4$. On pose $\theta = u + v\beta + w\beta^2$. Démontrer que $\frac{1}{2} < |\theta| < 1$ et en déduire que $\max(|u|, |v|, |w|) \leqslant 4$. Démontrer aussi que $0 < \mathrm{Tr}_K(\omega) < 6$ et $\mathrm{N}_K(\omega) = 1$.

g) Par une recherche exhaustive des triplets (u, v, w) possibles, vérifier qu'il n'y a pas d'unité de $\mathbf{Z}_K$ dans l'intervalle $]1, 4[$.

h) Démontrer que $\mathbf{Z}_K^*$ est le groupe engendré par -1 et $2 - \alpha$.

19) Soit K un corps de nombres, soit $\mathbf{Z}_K$ son anneau d'entiers et soit A un ordre de K.

a) Démontrer qu'il existe un entier $f \geqslant 1$ tel que A contienne $\mathbf{Z} + f\mathbf{Z}_K$.

b) Soit u une unité de $\mathbf{Z}_K$. Démontrer qu'il existe un entier naturel $n \geqslant 1$ tel que l'image de u^n dans $\mathbf{Z}_K/(f)$ soit égale à 1. En déduire que u^n appartient à A.

c) Démontrer que par l'application logarithmique, le groupe des unités de A a pour image un réseau de $\mathbf{R}^{r_1+r_2}$ d'indice fini dans l'image des unités de $\mathbf{Z}_K$.

§4.4. Factorisation ; groupe des classes d'idéaux

A. Anneaux d'entiers euclidiens, principaux

Commençons par quelques rappels.

Soit A un anneau intègre. Rappelons qu'on dit que A est un anneau *euclidien* s'il existe Une *jauge* (ou *jauge euclidienne*, ou encore un *stathme euclidien*) sur A est une fonction $j : A \setminus \{0\} \to \mathbf{N}$ vérifiant la propriété suivante : *pour tout* $a \in A$ *et tout* $b \in A \setminus \{0\}$, *il existe* q *et* r *dans* A *tels que* $a = bq + r$ *et tels que* $j(r) < j(b)$ *ou* $r = 0$. On exige aussi que $j(ab) \geqslant \max(j(a), j(b))$ pour tous a et b dans $A \setminus \{0\}$ mais cette dernière propriété n'est pas essentielle ; en effet, la fonction donnée par $j_{\max}(a) = \min_{b \neq 0} j(ab)$ est encore une jauge qui satisfait à cette condition supplémentaire.

L'anneau $\mathbf{Z}$ des entiers relatifs est un anneau euclidien pour la jauge $j = |\cdot|$; si k est un corps, l'anneau $k[X]$ des polynômes en une variables à coefficients dans k est un anneau euclidien pour la jauge $j = \deg$.

Un anneau euclidien est un *anneau principal*. La démonstration est la même que celle utilisée pour les anneaux $\mathbf{Z}$ et $k[X]$. Soit en effet A un anneau euclidien pour une jauge j et soit I un idéal de A. L'idéal 0 étant principal, supposons $I \neq 0$ et considérons un élément $b \in I \setminus \{0\}$ de jauge minimale. Soit $a \in I$; soit q et $r \in A$ tels que $a = bq + r$ et $r = 0$ ou $j(r) < j(b)$. Comme $b \in I$, $bq \in I$ et $r = a - bq \in I$; l'inégalité $j(r) < j(b)$ et la définition de b entraînent $r = 0$, c'est-à-dire $a \in (b)$. Inversement, tout multiple de b appartient à I, d'où $I = (b)$.

Un anneau principal est un *anneau factoriel*. Là encore, la démonstration est la même que pour les entiers ou les polynômes.

Donnons quelques exemples de corps de nombres K pour lesquels l'anneau $\mathbf{Z}_K$ est un anneau euclidien, la jauge étant la fonction norme $\mathbf{N}_K$.

PROPOSITION 4.4.1. — *Soit K un corps de nombres. Supposons que pour tout élément $z \in K$, il existe un élément $a \in \mathbf{Z}_K$ tel que $\mathrm{N}_K(z - a) < 1$. Alors, N_K définit une jauge euclidienne sur $\mathbf{Z}_K$ et $\mathbf{Z}_K$ est un anneau euclidien.*

Démonstration. — Soit a et b des éléments de $\mathbf{Z}_K$, avec $b \neq 0$; posons $z = a/b$ et soit $q \in \mathbf{Z}_K$ tel que $\mathrm{N}_K(z - q) < 1$. Posons alors $r = a - bq$; c'est un élément de $\mathbf{Z}_K$ tel que $r = b(z - q)$, donc $\mathrm{N}_K(r) < \mathrm{N}_K(b)$. Cela démontre que N_K est une jauge. □

Exemples 4.4.2. — *a*) L'anneau $\mathbf{Z}[i] = \mathbf{Z}_{\mathbf{Q}(i)}$ est euclidien pour la norme.

En effet, si $z = x + iy \in \mathbf{Q}(i)$, posons $a = \xi + i\eta$, où ξ et η sont les entiers les plus proches de x et y. On a donc $|\xi - x| \leqslant 1/2$ et $|\eta - y| \leqslant 1/2$. Alors, $\mathrm{N}_K(z - a) = |\xi - x|^2 + |\eta - y|^2 \leqslant \frac{1}{4} + \frac{1}{4} \leqslant \frac{1}{2}$.

b) Soit j une racine cubique primitive de l'unité. De même, l'anneau $\mathbf{Z}[j] = \mathbf{Z}_{\mathbf{Q}(j)}$ est euclidien pour la norme.

On raisonne de même : soit $z = x + jy \in \mathbf{Q}(i)$ et posons $a = \xi + j\eta$, où ξ et η sont les entiers les plus proches de x et y. On a $\mathrm{N}(z-a) = (x-\xi)^2 + (x-\xi)(y-\eta) + (y-\eta)^2 \leqslant \frac{3}{4} < 1$.

Exemple 4.4.3. — Soit K le corps $\mathbf{Q}(\sqrt{-19})$ et $\mathbf{Z}_K = \mathbf{Z}[\omega]$ son anneau d'entiers, où l'on a posé $\omega = \frac{1+\sqrt{-19}}{2}$. Démontrons que $\mathbf{Z}_K$ ne possède pas de jauge, donc n'est pas un anneau euclidien.

Considérons, par l'absurde, une jauge j sur $\mathbf{Z}_K$. Soit $a \in \mathbf{Z}_K$ un élément qui n'est ni nul ni inversible et tel que $j(a)$ soit minimal. Soit $z \in \mathbf{Z}_K$ et considérons une division euclidienne $z = aq + r$ avec $r = 0$ ou $j(r) < j(a)$. On a donc $r = 0$ ou r inversible. Autrement dit, tout élément de $\mathbf{Z}_K$ est congru modulo a à un élément qui est nul ou inversible. En particulier, $\mathrm{card}(\mathbf{Z}_K/(a)) \leqslant \mathrm{card}(\mathbf{Z}_K^*) + 1$. Comme K est un corps quadratique imaginaire, $\mathbf{Z}_K^*$ est un groupe fini, en l'occurrence réduit à $\{\pm 1\}$; ainsi, $\mathrm{card}(\mathbf{Z}/K/(a)) \leqslant 3$.

Or, $\mathbf{Z}_K/(a)$ est de cardinal $\mathrm{N}_K(a)$. Si l'on écrit $a = x + y\omega$, avec x et $y \in \mathbf{Z}$, on a $\mathrm{N}_K(a) = x^2 + xy + 5y^2$. Si $|y| \geqslant 1$, on a $\mathrm{N}_K(a) \geqslant (x^2 + xy + y^2) + 4y^2 \geqslant 4$; si $|y| = 0$, alors $|x| \geqslant 2$ car $a = x$ n'est ni nul, ni inversible, et $\mathrm{N}_K(a) \geqslant 4$. Cette contradiction démontre que j n'existe pas.

En revanche, on peut démontrer que $\mathbf{Z}_K$ est un anneau principal.

B. Idéaux d'un anneau d'entiers

Soit K un corps de nombres et $\mathbf{Z}_K$ son anneau d'entiers. On cherche à comprendre la relation de divisibilité dans $\mathbf{Z}_K$, en particulier lorsque cet anneau n'est pas factoriel, ou n'est pas principal. La théorie des idéaux a déjà démontré son intérêt, même dans le cas « facile » des anneaux principaux ; pensons par exemple à l'égalité de Bézout. Comme nous le verrons plus loin, il se trouve que les notions d'anneau factoriel et d'anneau principal vont coïncider pour $\mathbf{Z}_K$, mais la théorie des idéaux est l'outil privilégié pour cette étude.

Nous reprenons les notations du paragraphe précédent en notant $n = [K : \mathbf{Q}]$, r_1 et $2r_2$ les nombres de plongements réels, resp. complexes, de K et Φ l'application canonique de K dans l'espace vectoriel $V_K = \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$.

PROPOSITION 4.4.1. — *Soit I un idéal non nul de $\mathbf{Z}_K$. L'anneau $\mathbf{Z}_K/I$ est fini et l'entier* $\mathrm{card}(\mathbf{Z}_K/I)$ *appartient à I.*

Démonstration. — Soit a un élément non nul de I. Il existe un élément b de I tel que $\mathrm{N}_K(a) = ab$. En effet, $\mathrm{N}_K(a)/a$ est le produits de $(n-1)$ entiers algébriques (conjugués de a) donc est un entier algébrique ; comme c'est un élément de K, il appartient à $\mathbf{Z}_K$. Cela démontre que l'entier $\mathrm{N}_K(a)$ appartient à I. Par suite, l'anneau $\mathbf{Z}_K/I$ est quotient de l'anneau $\mathbf{Z}_K/\mathrm{N}_K(a)$; en tant que groupe abélien, ce dernier est isomorphe à $(\mathbf{Z}/\mathrm{N}_K(a)\mathbf{Z})^n$, donc est fini. Par conséquent, $\mathbf{Z}_K/I$ est fini.

Soit N son cardinal ; d'après le théorème de Lagrange, on a $N\,\mathrm{cl}(a) = 0$ pour tout $a \in \mathbf{Z}_K$, cl désignant la classe modulo I. En particulier, prenant $a = 1$, on a $N \equiv 0 \pmod{I}$ c'est-à-dire $N \in I$. □

COROLLAIRE 4.4.2. — *Un idéal premier non nul de $\mathbf{Z}_K$ est maximal.*

Démonstration. — Soit I un idéal premier de $\mathbf{Z}_K$, supposé différent de l'idéal nul. Alors, $\mathbf{Z}_K/I$ est un anneau intègre (car I est premier) et fini. C'est donc un corps (si $a \in \mathbf{Z}_K/I$ n'est pas nul, la multiplication par a est un endomorphisme injectif de l'anneau fini $\mathbf{Z}_K/I$, donc est bijectif, donc a est inversible) ce qui signifie que l'idéal I est un idéal maximal. □

COROLLAIRE 4.4.3. — *L'anneau $\mathbf{Z}_K$ est noethérien : toute suite croissante (I_n) d'idéaux de $\mathbf{Z}_K$ est stationnaire.*

Démonstration. — Considérons, par l'absurde, une suite strictement croissante (I_n) d'idéaux de $\mathbf{Z}_K$. Pour $n \geqslant 1$, $I_n \neq 0$ donc $\mathbf{Z}_K/I_n$ est un ensemble fini. L'inclusion $I_n \subset I_{n+1}$ entraîne qu'il existe un homomorphisme surjectif de $\mathbf{Z}_K/I_n$ sur $\mathbf{Z}_K/I_{n+1}$; comme $I_n \neq I_{n+1}$, cet homomorphisme n'est pas injectif. On a donc $\mathrm{card}(\mathbf{Z}_K/I_{n+1}) < \mathrm{card}(\mathbf{Z}_K/I_n)$.

La suite des nombre entiers $(\mathrm{card}(\mathbf{Z}_K/I_n))_{n\geqslant 1}$ est ainsi strictement décroissante, d'où la contradiction recherchée. □

COROLLAIRE 4.4.4. — *Tout idéal de $\mathbf{Z}_K$, distinct de $\mathbf{Z}_K$, est contenu dans un idéal maximal de $\mathbf{Z}_K$.*

Démonstration. — Soit I un idéal de $\mathbf{Z}_K$ tel que $I \neq \mathbf{Z}_K$. Démontrons par récurrence sur $\mathrm{card}(\mathbf{Z}_K/I)$ qu'il existe un idéal maximal P de $\mathbf{Z}_K$ tel que $I \subset P$.

Si I est maximal, il suffit de poser $P = I$. Sinon, il existe un idéal I_1 de $\mathbf{Z}_K$ tel que $I \subsetneq I_1 \subsetneq \mathbf{Z}_K$ et $\mathrm{card}(\mathbf{Z}_K/I_1) < \mathrm{card}(\mathbf{Z}_K/I)$. Par récurrence, il existe un idéal maximal P de $\mathbf{Z}_K$ tel que $I_1 \subset P$, et l'on a donc $I \subset P$. □

COROLLAIRE 4.4.5. — *Soit I un idéal non nul de $\mathbf{Z}_K$. Alors, $\Phi(I)$ est un réseau de rang n de V_K ; le volume d'un domaine fondamental de ce réseau est égal à*

$$V(\Phi(I)) = \mathrm{card}(\mathbf{Z}_K/I)\,V(\Phi(\mathbf{Z}_K)) = \mathrm{card}(\mathbf{Z}_K/I)2^{-r_2}\sqrt{|\mathrm{disc}(K)|}.$$

De plus, il existe un élément non nul $a \in I$ de taille

$$\mathrm{t}(a) \leqslant (2/\pi)^{r_2/n}\,|\mathrm{disc}(K)|^{1/2n}\,\mathrm{card}(\mathbf{Z}_K/I)^{1/n}.$$

En particulier, il existe un élément non nul $a \in I$ tel que

$$\mathrm{N}_K(a) \leqslant \left(\frac{2}{\pi}\right)^{r_2}\sqrt{|\mathrm{disc}(K)|}\,\mathrm{card}(\mathbf{Z}_K/I).$$

Démonstration. — On a $\Phi(I) \subset \Phi(\mathbf{Z}_K)$; comme $\Phi(\mathbf{Z}_K)$ est un réseau de V_K, il en est de même de $\Phi(I)$. Comme $\mathbf{Z}_K/I$ est fini, $\Phi(I)$ est de rang n et ses domaines fondamentaux ont une volume donné par la formule $V(\Phi(I)) = \mathrm{card}(\mathbf{Z}_K/I)V(\Phi(\mathbf{Z}_K))$.

Munissons $V_K = \mathbf{R}^{r_1} \times \mathbf{C}^{r_2}$ de la norme définie par $\|(z_1,\dots,z_{r_1+r_2})\| = \max(|z_j|)$. Sa boule unité est égale à $[-1,1]^{r_1} \times B(0,1)^{r_2}$ et son volume vaut $2^{r_1}\pi^{r_2}$. D'après le théorème de Minkowski, le réseau $\Phi(I)$ contient un élément non nul $\Phi(a)$ tel que

$$\begin{aligned}\|\Phi(a)\| &\leqslant 2\left(\frac{V(\Phi(I)}{V(B)}\right)^{1/n}\\ &\leqslant \left(2^{n-r_1}\pi^{-r_2}2^{-r_2}\sqrt{|\mathrm{disc}(K)|}\,\mathrm{card}(\mathbf{Z}_K/I)\right)^{1/n}\\ &\leqslant \left(\frac{2}{\pi}\right)^{r_2/n}|\mathrm{disc}(K)|^{1/2n}\,\mathrm{card}(\mathbf{Z}_K/I)^{1/n}.\end{aligned}$$

Cela signifie que tous les conjugués de a sont de valeur absolue majorée par le membre de droite X de l'expression précédente. En particulier, la norme de a est majorée par sa puissance n-ième. En outre, la taille de a est majorée par $\max(1,X)$, car a est un entier algébrique. □

Si I est un idéal non nul de l'anneau $\mathbf{Z}_K$, on appelle *norme de l'idéal* I, et on note $\mathrm{N}(I)$, le cardinal de l'anneau quotient $\mathbf{Z}_K/I$. Notons tout de suite le lien avec la norme d'un élément de K.

PROPOSITION 4.4.6. — *Soit a un élément non nul de $\mathbf{Z}_K$. On a $\mathrm{N}(a\mathbf{Z}_K) = |\mathrm{N}_K(a)|$.*

Plus généralement, pour tout idéal non nul I de $\mathbf{Z}_K$, on a $\mathrm{N}(aI) = |\mathrm{N}_K(a)|\,\mathrm{N}(I)$.

Démonstration. — Soit $(z_1,\dots,z_n)$ une base de I. Alors, $(az_1,\dots,az_n)$ est une base de aI. Compte tenu de la formule donnée pour le volume du réseau défini par un idéal, il suffit de vérifier que $V(\Phi(aI)) = |\mathrm{N}_K(a)|\,V(\Phi(I))$. À son tour, cette dernière formule résultera de l'égalité de discriminants

$$\mathrm{disc}((az_1,\dots,az_n)) = \mathrm{N}_K(a)^2\,\mathrm{disc}((z_1,\dots,z_n)).$$

Or, notant $(\varphi_j)_{1\leqslant j\leqslant n}$ les plongements de K dans $\mathbf{C}$, on a

$$\begin{aligned}\mathrm{disc}((az_1,\dots,az_n)) &= \det(\varphi_i(az_j))^2 = \det((\varphi_i(a)\varphi_i(z_j)))^2\\ &= (\varphi_1(a)\dots\varphi_n(a))^2\det((\varphi_i(z_j)))^2\\ &= \mathrm{N}_K(a)^2\,\mathrm{disc}((z_1,\dots,z_n)).\end{aligned}$$

Cela conclut la démonstration de la proposition. □

COROLLAIRE 4.4.7. — *Pour qu'un idéal non nul I de $\mathbf{Z}_K$ soit principal, il faut et il suffit qu'il existe un élément $a \in I$ tel que $\mathrm{N}_K(a) = \pm\mathrm{N}(I)$; un tel élément est alors un générateur de I.*

Démonstration. — Supposons que I soit un idéal principal et soit a un générateur de I ; on a donc $a \in I$ et $|\mathrm{N}_K(a)| = \mathrm{N}(I)$, ce qu'il fallait démontrer.

Inversement, soit a un élément de I tel que $\mathrm{N}_K(a) = \pm\mathrm{N}(I)$. L'inclusion d'idéaux $(a) \subset I$ et l'égalité $\mathrm{N}((a)) = |\mathrm{N}_K(a)| = \mathrm{N}(I)$ entraînent que $I = (a)$; en particulier, I est principal. □

C. Groupe des classes d'idéaux

Notons $\mathscr{I}_K$ l'ensemble des idéaux non nuls de $\mathbf{Z}_K$ et munissons-le de la relation suivante : si $I, J \in \mathscr{I}_K$, on dit que $I \sim J$ s'il existe des éléments non nuls a et b de $\mathbf{Z}_K$ tels que $aI = bJ$. C'est une relation d'équivalence. Notons $\mathscr{C}_K$ l'ensemble des classes d'équivalence.

Si $\mathbf{Z}_K$ est un anneau principal, tous les éléments de $\mathscr{I}_K$ sont équivalents entre eux. Inversement, Si I est un idéal non nul tel que $I \sim \mathbf{Z}_K$, démontrons que I est un idéal principal. Soit en effet a et $b \in \mathbf{Z}_K$ tels que $aI = b\mathbf{Z}_K = (b)$, avec $a, b \neq 0$. Comme $b \in aI$, il existe $w \in I$ tel que $b = aw$ et l'on a $aI = (aw)$, d'où l'on déduit facilement que $I = (w)$.

PROPOSITION 4.4.1. — *Posons $c = (2/\pi)^{r_2}\sqrt{|\mathrm{disc}(K)|}$. Tout idéal non nul de $\mathbf{Z}_K$ est équivalent à un idéal de norme $\leqslant c$. L'ensemble $\mathscr{C}_K$ des classes d'idéaux de $\mathbf{Z}_K$ est fini.*

Démonstration. — Pour tout entier naturel t tel que $1 \leqslant t \leqslant c$, les idéaux de $\mathbf{Z}_K$ contenant t sont en bijection avec les idéaux de l'anneau fini $\mathbf{Z}_K/(t)$; il n'y en a qu'un nombre fini. Démontrons que tout idéal non nul de $\mathbf{Z}_K$ est équivalent à l'un de ces idéaux.

Soit donc I un idéal non nul de $\mathbf{Z}_K$; d'après le corollaire **??**, il existe un élément non nul a de I tel que $|\mathrm{N}_K(a)| \leqslant c\,\mathrm{N}(I)$. Considérant la suite de groupes abéliens $a\mathbf{Z}_K \subset I \subset \mathbf{Z}_K$, on voit que $I/a\mathbf{Z}_K$ est un groupe de cardinal inférieur ou égal à c. Il existe donc un entier naturel t tel que $1 \leqslant t \leqslant c$ et tel que $tI \subset a\mathbf{Z}_K$. Notons J la partie $a^{-1}tI$ de $\mathbf{Z}_K$; c'est un idéal tel que $tI = aJ$. En particulier, I et J sont équivalents. Comme $a \in I$, $t \in J$. La proposition est ainsi démontrée. □

DÉFINITION 4.4.2. — *On appelle nombre de classes de K le cardinal de $\mathscr{C}_K$, c'est-à-dire le nombre de classes d'équivalences pour la relation d'équivalence introduite dans $\mathscr{I}_K$; on le note h_K.*

Ainsi, $h_K = 1$ signifie exactement que $\mathbf{Z}_K$ est un anneau principal.

L'ensemble $\mathscr{I}_K$ possède une structure de *monoïde* commutatif, associatif et unitaire donnée par le produit des idéaux : si I et J sont des idéaux, l'idéal IJ est l'idéal engendré par les produits ab, où $a \in I$ et $b \in J$. La commutativité vient de ce que $IJ = JI$, l'associativité de la relation $(II')I'' = I(I'I'')$; enfin, l'élément neutre est l'idéal $\mathbf{Z}_K$.

En outre, la relation d'équivalence introduite ci-dessus est compatible au produit des idéaux : si $I \sim I'$ et $J \sim J'$, alors $IJ \sim I'J'$. (En effet, choisissons a, a', b, b' dans $\mathbf{Z}_K$, non nuls, tels que $aI = a'I'$ et $bJ = b'J'$; alors, $abIJ = (aI)(bJ) = (a'I')(b'J') = a'b'I'J'$.)

L'ensemble quotient $\mathscr{C}_K$ hérite donc d'une structure de monoïde commutatif, associatif et unitaire.

Le but de ce paragraphe est d'élucider la structure des monoïdes $\mathscr{I}_K$ et $\mathscr{C}_K$.

THÉORÈME 4.4.3. — a*) Le monoïde $\mathscr{C}_K$ est un groupe fini.*

b*) Dans le monoïde $\mathscr{I}_K$, tout élément est simplifiable : si I, I', J sont des idéaux non nuls de $\mathbf{Z}_K$ tels que $IJ = I'J$, alors $I = I'$.*

On démontre ce théorème par une série de lemmes.

LEMME 4.4.4. — *Soit I et J des idéaux non nuls de $\mathbf{Z}_K$ et soit a un élément non nul de $\mathbf{Z}_K$. On suppose que $IJ = aJ$. Alors, $I = (a)$.*

Démonstration. — Soit $z \in I$, posons $w = z/a$ et démontrons que $w \in \mathbf{Z}_K$. On remarque que $wJ \subset J$. Soit $(z_1, \dots, z_n)$ une base de J. Écrite dans cette base, la matrice M_w de l'endomorphisme de multiplication par w est donc à coefficients entiers ; si P désigne le polynôme caractéristique de M_w, on a donc $P(w) = 0$ donc w est un entier algébrique. Par conséquent, $w \in \mathbf{Z}_K$. Ainsi, $I \subset (a)$.

Pour démontrer l'autre inégalité, notons I' l'ensemble $a^{-1}I$ de $\mathbf{Z}_K$ et prouvons que $I' = \mathbf{Z}_K$. Il est clair que i' est un idéal tel que $I'J = J$. Il existe donc des éléments $a_{ij} \in I'$ tels que $z_j = \sum_{i=1}^n a_{ij} z_i$. Notant $Q = X^n + q_1 X^{n-1} + \cdots + q_n$ le polynôme caractéristique de la matrice (a_{ij}), le théorème de Cayley-Hamilton (ou la définition d'une valeur propre) entraîne que $Q(1) = 0$. Comme les q_i appartiennent à l'idéal I', cela entraîne que $1 \in I'$, d'où $I' = \mathbf{Z}_K$, ce qu'il fallait démontrer. □

LEMME 4.4.5. — *Pour tout idéal non nul I de A, il existe un entier t tel que $1 \leqslant t \leqslant h_K$ et tel que l'idéal I^t soit un idéal principal.*

Démonstration. — Parmi les $h_K + 1$ idéaux, $\mathbf{Z}_K, I, \dots, I^{h_K}$, deux au moins ont même classe dans $\mathscr{C}_K$, disons I^m et I^{m+p}, avec $0 \leqslant m < m+p \leqslant h_K$. Il existe donc des éléments non nuls a et b dans $\mathbf{Z}_K$ tels que $aI^m = bI^{m+p}$, ce qu'on écrit $bI^p I^m = aI^m$. D'après le lemme précédent, $bI^p = (a)$ et I^p est principal. □

Démonstration du théorème. — *a*) Tout d'abord, le monoïde $\mathscr{C}_K$ est commutatif, associatif et unitaire. D'après le lemme précédent, tout élément de $\mathscr{C}_K$ possède un inverse. Cela signifie que $\mathscr{C}_K$ est un groupe abélien. Il est fini d'après la proposition **??**.

b) Soit I, I', J des idéaux non nuls de $\mathbf{Z}_K$ tels que $IJ = I'J$; démontrons que $I = I'$.

Soit J' un idéal non nul de $\mathbf{Z}_K$ tel que JJ' soit un idéal principal, disons (a), avec $a \in \mathbf{Z}_K$. Alors, $IJJ' = I'JJ'$, d'où $aI = aI'$. Multipliant cette égalité par a^{-1}, on a $I = I'$. □

D. Factorisation

Nous démontrons dans ce paragraphe que les anneaux d'entiers de corps de nombres jouissent d'une propriété très proche de celle des anneaux factoriels : les

idéaux d'un tel anneau se décomposent de manière unique en produit d'idéaux maximaux.

On conserve les notations précédentes : K est un corps de nombres de degré n, $\mathscr{I}_K$ est le monoïde des idéaux non nuls de l'anneau des entiers $\mathbf{Z}_K$ et $\mathscr{C}_K$ est le groupe des classes d'idéaux.

Commençons par un lemme qui relie les relation d'inclusion et de divisibilité dans les idéaux. Si un idéal I est produit de deux idéaux J et J', alors I est contenu dans J

LEMME 4.4.1. — *Soit I et J deux idéaux non nuls de $\mathbf{Z}_K$. Pour qu'il existe un idéal J' de $\mathbf{Z}_K$ tel que $I \subset J$, il faut et il suffit qu'il existe un idéal J' de K tel que $I = JJ'$.*

Démonstration. — Supposons d'abord $I = JJ'$. Cela signifie que I est l'idéal engendré par les produits ab, avec $a \in J$ et $b \in J'$. Chacun de ces produits ab est contenu dans J, par définition d'un idéal. Donc l'idéal qu'ils engendrent est aussi contenu dans J, c'est-à-dire $I \subset J$.

Pour la réciproque, nous devrons utiliser l'hypothèse que K est un anneau d'entiers de corps de nombres. Soit J_1 un idéal non nul de $\mathbf{Z}_K$ tel que l'idéal JJ_1 soit un idéal principal ; soit alors $\alpha \in \mathbf{Z}_K$ un générateur de l'idéal JJ_1, de sorte que $JJ_1 = (\alpha)$. Comme $I \subset J$, on a $IJ_1 \subset JJ_1 = \alpha\mathbf{Z}_K$. La partie $\alpha^{-1}IJ_1$ de K est donc contenue dans $\mathbf{Z}_K$ et est un idéal de $\mathbf{Z}_K$. Notons-le J'. On a donc $\alpha J' = IJ_1$. Multiplions cette égalité par l'idéal J ; on trouve $\alpha JJ' = IJ_1J = I(\alpha) = \alpha I$. On peut alors simplifier par α (qui n'est pas nul), d'où l'égalité $JJ' = I$, ce qui démontre le lemme. □

THÉORÈME 4.4.2. — a*) Tout idéal non nul de $\mathbf{Z}_K$ est produit d'idéaux maximaux.*

b*) Si $P_1, \dots, P_t$ et $Q_1, \dots, Q_s$ sont des idéaux maximaux de $\mathbf{Z}_K$ tels que $P_1 \dots P_t = Q_1 \dots Q_s$, alors $t = s$ et il existe une permutation σ de $\{1, \dots, t\}$ tel que $Q_i = P_{\sigma(i)}$ pour tout i.*

Démonstration. — *a*) Soit I un idéal non nul de $\mathbf{Z}_K$. Démontrons par récurrence sur $\mathrm{N}(I)$ qu'il existe des idéaux maximaux $P_1, \dots, P_t$ de $\mathbf{Z}_K$ tels que $I = P_1 \dots P_t$. Supposons le résultat vrai pour les idéaux de norme $< \mathrm{N}(I)$. Si $\mathrm{N}(I) = 1$, c'est-à-dire $I = \mathbf{Z}_K$, on prend $t = 0$ (produit vide). Sinon, $I \neq \mathbf{Z}_K$ et il existe d'après le corollaire 4.4.4 un idéal maximal P_1 de $\mathbf{Z}_K$ tel que $I \subset P_1$. D'après le lemme précédent, il existe un idéal I_1 de $\mathbf{Z}_K$ tel que $I = P_1 I_1$. On a $I \subset I_1$; si l'on avait $I = I_1$, on aurait alors $P_1 = \mathbf{Z}_K$ d'après le corollaire **??**, ce qui est absurde. Par suite, $I \subsetneq I_1$ et $\mathrm{N}(I_1) < \mathrm{N}(I)$. Par récurrence, il existe des idéaux maximaux $P_2, \dots, P_t$ de $\mathbf{Z}_K$ tels que $I_1 = P_2 \dots P_t$ et $I = P_1 I_1 = P_1 P_2 \dots P_t$.

b) Démontrons le résultat voulu par récurrence sur t. Posons $I = P_1 \dots P_t = Q_1 \dots Q_s$. Si $t = 0$, $I = \mathbf{Z}_K$. Si l'on avait $s > 0$, on aurait $I \subset Q_1 \neq \mathbf{Z}_K$, d'où une contradiction, on a donc $s = 0$.

Supposons maintenant $t > 0$. Par hypothèse, $Q_1 \dots Q_s \subset P_t$. D'après le lemme ci-dessous, il existe un entier $j \in \{1, \dots, s\}$ tel que $Q_j \subset P_t$, d'où $Q_j = P_t$ car Q_j est un idéal maximal de $\mathbf{Z}_K$. Quitte à renuméroter les Q_i, on peut supposer que $j = s$. Posons

alors $I_1 = P_1 \dots P_{t-1}$ et $I_1' = Q_1 \dots Q_{s-1}$; on a $I_1 P_t = I_1' Q_s$, d'où $I_1 = I_1'$ d'après le corollaire **??**. Par récurrence, $s-1 = t-1$ et il existe une permutation σ_1 de $\{1,\dots,t-1\}$ tel que $Q_i = P_{\sigma_1(i)}$ pour tout $i \in \{1,\dots,s-1\}$. On a donc $t = s$ et prolongeons σ en une permutation de $\{1,\dots,t\}$ en posant $\sigma(i) = \sigma_1(i)$ pour $i \in \{1,\dots,t-1\}$ et $\sigma(t) = t$. On a ainsi $Q_i = P_{\sigma(i)}$ pour $i \in \{1,\dots,t\}$. □

LEMME 4.4.3. — *Soit Q un idéal premier de $\mathbf{Z}_K$ et $I_1,\dots,I_n$ des idéaux de $\mathbf{Z}_k$ tels que $I_1 \dots I_n \subset Q$. Il existe un entier $j \in \{1,\dots,n\}$ tel que $I_j \subset Q$.*

Démonstration. — Le cas $n = 0$ ne se produit pas car l'hypothèse entraînerait $\mathbf{Z}_K \subset Q$ et donc $Q = \mathbf{Z}_K$, contrairement à l'hypothèse que Q est un idéal premier de $\mathbf{Z}_K$. Si $n = 1$, il suffit de poser $j = 1$.

Par récurrence, il suffit alors de traiter le cas où $n = 2$. Si $I_1 \not\subset Q$ et $I_2 \not\subset Q$, choisissons des éléments $a_1 \in I_1$ et $a_2 \in I_2$ qui n'appartiennent pas à Q. Comme Q est un idéal premier de $\mathbf{Z}_K$, $a_1 a_2 \notin Q$. Mais $a_1 a_2$ appartient à $I_1 I_2$, par définition de l'idéal produit, donc appartient à Q puisque $I_1 I_2 \subset Q$. Cette contradiction démontre que $I_1 \subset Q$ ou $I_2 \subset Q$. □

COROLLAIRE 4.4.4. — *Posons $c = (2/\pi)^{r_2}\sqrt{|\mathrm{disc}(K)|}$. Le groupe $\mathscr{C}_K$ est engendré par les classes d'idéaux maximaux de $\mathbf{Z}_K$ de normes $\leqslant c$. En particulier, pour que $\mathbf{Z}_K$ soit un anneau principal, il faut et il suffit que les idéaux maximaux de normes $\leqslant c$ soient principaux.*

Démonstration. — Posons $c = (2/\pi)^{r_2}\sqrt{|\mathrm{disc}(K)|}$. D'après la proposition **??**, tout idéal non nul I de $\mathbf{Z}_K$ est équivalent à un idéal J de norme $\leqslant c$. Décomposons alors J en produit d'idéaux maximaux, disons $J = P_1 \dots P_t$; pour tout j, $J \subset P_j$ donc $\mathrm{N}(P_j) \leqslant \mathrm{N}(J) \leqslant c$. Par suite, la classe de J, et donc celle de I, appartient au sous-groupe de $\mathscr{C}_K$ engendré par les idéaux maximaux de $\mathbf{Z}_K$ de norme $\leqslant c$. Le corollaire est ainsi démontré. □

COROLLAIRE 4.4.5. — *Soit K un corps de nombres. Si l'anneau d'entiers $\mathbf{Z}_K$ est un anneau factoriel, c'est un anneau principal.*

Démonstration. — Soit P un idéal maximal de $\mathbf{Z}_K$.

Soit a un élément non nul de P tel que $|\mathrm{N}_K(a)|$ soit minimal et observons que a est un élément irréductible. Soit en effet une factorisation $a = a_1 a_2$ de a en produit d'éléments de $\mathbf{Z}_K$. Comme P est un idéal maximal, on a donc $a_1 \in P$ ou $a_2 \in P$; supposons pour fixer les idées que $a_1 \in P$ et démontrons que a_2 est un élément inversible de $\mathbf{Z}_K$.

Compte-tenu de l'égalité $\mathrm{N}_K(a) = \mathrm{N}_K(a_1)\mathrm{N}_K(a_2)$ et de l'inégalité $|\mathrm{N}_K(a_2)| \geqslant 1$, la minimalité de $|\mathrm{N}_K(a)|$ entraîne que $|\mathrm{N}_K(a_1)| = |\mathrm{N}_K(a)|$, d'où $|\mathrm{N}_K(a_2)| = 1$. Par suite, a_2 est un élément inversible de $\mathbf{Z}_K$, ainsi qu'il fallait démontrer.

Comme $\mathbf{Z}_K$ est un anneau factoriel, l'idéal (a) engendré par a est un idéal premier, non nul puisque $a \neq 0$, donc maximal. L'inclusion évidente $a\mathbf{Z}_K \subset P$ entraîne alors que $P = a\mathbf{Z}_K$, autrement dit P est un idéal principal.

D'après le corollaire précédent, tout idéal non nul de $\mathbf{Z}_K$ est principal et $\mathbf{Z}_K$ est un anneau principal. □

E. Exemples

Voyons sur quelques exemples comment utiliser les résultats précédents pour déterminer la structure du groupe de classes d'idéaux.

Nous aurons besoin, un nombre entier naturel n étant donné, de trouver explicitement tous les idéaux maximaux de norme n. Faisons pour cela quelques remarques générales. Soit K un corps de nombres et $\mathbf{Z}_K$ son anneau d'entiers. Soit P un idéal maximal de $\mathbf{Z}_K$.

a) L'anneau quotient $\mathbf{Z}_K/P$ est un corps fini ; son cardinal est donc une puissance d'un nombre premier p. On peut donc supposer que n est de la forme p^f.

b) Comme le corps $\mathbf{Z}_K/P$ est de caractéristique p, p est nul dans ce corps et l'on a $p \in P$.

c) Supposons que l'on connaisse un entier algébrique α de K tel que $\mathbf{Z}_K = \mathbf{Z}[\alpha]$; soit A le polynôme minimal de α ; l'anneau $\mathbf{Z}_K$ est donc isomorphe à $\mathbf{Z}[X]/(A(X))$. L'image réciproque I dans $\mathbf{Z}[X]$ de l'idéal P est un idéal de $\mathbf{Z}[X]$ qui contient p et $A(X)$; on a en outre un isomorphisme d'anneaux $\mathbf{Z}_K/P \simeq \mathbf{Z}[X]/I$. Soit J l'image de I dans l'anneau $\mathbf{F}_p[X]$ par l'homomorphisme de réduction modulo p ; c'est un idéal de $\mathbf{F}_p[X]$ car cet homomorphisme est surjectif ; il existe donc un polynôme $R \in \mathbf{F}_p[X]$ tel que $J = (R(X))$. De plus, on a un isomorphisme d'anneaux $\mathbf{Z}[X]/I \simeq \mathbf{F}_p[X]/(R(X))$; par suite, R est un polynôme irréductible de $\mathbf{F}_p[X]$. Comme $A(X)$ appartient à I, sa réduction modulo p appartient à $J = (R(X))$ ce qui entraîne que R divise A. En remontant les calculs et en notant $\tilde{R}$ un polynôme de $\mathbf{Z}[X]$ arbitraire dont la réduction modulo p est égale à R, on a $P = (p, \tilde{R}(\alpha))$.

Inversement, par cette formule, tout facteur irréductible R de $A \pmod p$ définit un idéal maximal de $\mathbf{Z}_K$ de caractéristique p.

d) Avec les notations précédentes, le corps fini $\mathbf{Z}_K/P$ est isomorphe à $\mathbf{F}_p[X]/(R(X))$. On a donc $\mathrm{N}(P) = p^{\deg R}$.

Récapitulons : *les idéaux maximaux de $\mathbf{Z}_K$ contenant un nombre premier p sont en bijection avec les facteurs irréductibles de degré f du polynôme $A(X) \pmod p$.*

Exemple 4.4.1. — Soit $P = X^3 + X + 1$; comme $P' = 3X^2 + 1$, P définit une fonction strictement croissante sur $\mathbf{R}$ et a donc une unique racine réelle, α. Posons $K = \mathbf{Q}(\alpha)$. C'est un corps de nombres de degré 3 ; on a $\mathbf{Z}_K = \mathbf{Z}[\alpha]$ et $\mathrm{disc}(K) = -31$. Comme P possède une seule racine réelle, on a $r_1 = r_2 = 1$. Posons alors $c = (2/\pi)^{r_2}\sqrt{|\mathrm{disc}(K)|}$; on a $c \approx 3{,}5$. Par conséquent, le groupe des classes $\mathscr{C}_K$ est engendré par les idéaux maximaux de $\mathbf{Z}_K$ de normes 2 ou 3.

Soit M un idéal maximal contenant 2. Dans $\mathbf{F}_2$, le polynôme P n'a pas de racine ; come il est de degré 3, P est irréductible modulo 2. Par suite, $M = (2)$ et $\mathrm{N}(M) = 8$.

Soit M un idéal maximal contenant 3. Dans $\mathbf{F}_3$, 1 est racine de P et l'on a $P = (X-1)(X^2+X-1)$, le polynôme X^2+X-1 étant irréductible. On a donc deux cas :

— soit $M = (3, \alpha - 1)$ et $\mathrm{N}(M) = 3$;

— soit $M = (3, \alpha^2 + \alpha - 1)$ et $\mathrm{N}(M) = 9$.

Cela démontre que $\mathscr{C}_K$ est engendré par la classe de l'idéal $M = (3, \alpha - 1)$. Comme $\mathrm{N}_K(1-\alpha) = P(1) = 3 = \mathrm{N}(M)$, $1-\alpha$ est un générateur de M. Par suite, l'anneau $\mathbf{Z}_K$ est principal.

Exemple 4.4.2. — Soit K le corps quadratique $\mathbf{Q}(\sqrt{-5})$; on a $\mathbf{Z}_K = \mathbf{Z}[\sqrt{-5}]$ et $\mathrm{disc}(\mathbf{Z}_K) = -20$.

Le groupe des classes $\mathscr{C}_K$ est donc engendré par les idéaux maximaux de $\mathbf{Z}_K$ de normes $\leqslant (2/\pi)\sqrt{20} \approx 2{,}8$, autrement dit par les idéaux maximaux de $\mathbf{Z}_K$ de norme 2.

Déterminons ces idéaux. Le polynôme minimal de $\sqrt{-5}$ est égal à X^2+5 et l'on a $X^2+5 \equiv (X+1)^2 \pmod 2$. Ainsi, le seul idéal maximal de $\mathbf{Z}_K$ qui contienne 2 est l'idéal $M = (2, 1+\sqrt{-5})$; sa norme est effectivement égale à 2.

Il reste à déterminer si cet idéal est principal ou non. Cherchons pour cela un élément de norme 2 dans M. Or, si $z = x + y\sqrt{-5} \in \mathbf{Z}_K$, $\mathrm{N}_K(z) = x^2 + 5y^2$. Par suite, si $y \neq 0$, on a $|\mathrm{N}_K(z)| = \mathrm{N}_K(z) \geqslant 5$; l'égalité $\mathrm{N}_K(z) = \pm 2$ entraîne donc $z \in \mathbf{Z}$, mais alors $\mathrm{N}_K(z) = z^2$ n'est pas égale à ± 2. Par conséquent, l'idéal $M = (2, 1+\sqrt{-5})$ n'est pas principal.

L'idéal M^2 est engendré par 4, $2(\alpha+1)$ et $(\alpha+1)^2 = \alpha^2 + 2\alpha + 1 = 2\alpha - 4$. On a alors

$$M^2 = (4, 2\alpha+2, 2\alpha-4) = (4, 2\alpha+2, 2\alpha) = (4, 2\alpha, 2) = (2, 2\alpha) = (2)$$

ce qui démontre que M^2 est un idéal principal.

Par conséquent, le groupe des classes d'idéaux de l'anneau $\mathbf{Z}_K$ est isomorphe à $\mathbf{Z}/2\mathbf{Z}$ et la classe de l'idéal $(2, 1+\sqrt{-5})$ en est un générateur.

F. Multiplicativité des normes des idéaux

LEMME 4.4.1. — *Soit I et J des idéaux non nuls de $\mathbf{Z}_K$ tels que $I + J = \mathbf{Z}_K$. On a $\mathrm{N}(IJ) = \mathrm{N}(I)\,\mathrm{N}(J)$.*

Démonstration. — Soit φ l'application de $\mathbf{Z}_K$ dans $(\mathbf{Z}_K/I) \times (\mathbf{Z}_K/J)$ telle que $\varphi(a) = (\mathrm{cl}_I(a), \mathrm{cl}_J(a))$ pour $a \in \mathbf{Z}_K$. C'est un homomorphisme d'anneaux. Démontrons qu'il est surjectif. Soit a et b des éléments de I et J respectivement tels que $1 = a + b$. Si $z \in \mathbf{Z}_K$, on a donc $\mathrm{cl}_I(az) = 0$, $\mathrm{cl}_J(bz) = 0$ tandis que $\mathrm{cl}_I(bz) = \mathrm{cl}_I((1-a)z) = \mathrm{cl}_I(z)$ et $\mathrm{cl}_J(az) = \mathrm{cl}_J(z)$. Par suite, si $z, w \in \mathbf{Z}_K$, $\varphi(bz + aw) = (\mathrm{cl}_I(z), \mathrm{cl}_J(w))$. Cela entraîne que φ est surjectif. Soit z un élément du noyau de φ, c'est-à-dire un élément de $I \cap J$. Écrivons $z = z(a+b) = az + ab$: comme $z \in J$, il vient $az \in IJ$; comme $z \in I$, on a $bz \in IJ$; en déduit que z appartient à IJ. Inversement, si $z \in IJ$, on a $z \in I$ et $z \in J$, donc $\varphi(z) = 0$.

Il résulte de φ un isomorphisme d'anneaux de $\mathbf{Z}_K/IJ$ sur $(\mathbf{Z}_K/I)\times(\mathbf{Z}_K/J)$. En particulier, $\mathrm{N}(IJ)$, qui est le cardinal de $\mathbf{Z}_K/IJ$, est égal au produit de $\mathrm{N}(I)$ par $\mathrm{N}(J)$. □

LEMME 4.4.2. — *Soit P un idéal maximal de $\mathbf{Z}_K$. Pour tout entier naturel e, on a* $\mathrm{N}(P^e)=\mathrm{N}(P)^e$.

Démonstration. — On démontre ce résultat par récurrence sur e. De la suite d'idéaux $P^{e+1}\subset P^e\subset\mathbf{Z}_K$, on déduit une égalité de cardinaux $\mathrm{card}(\mathbf{Z}_K/P^{e+1})=\mathrm{card}(\mathbf{Z}_K/P^e)\,\mathrm{card}(P^e/P^{e+1})$ et il suffit de démontrer que $\mathrm{card}(P^e/P^{e+1})=\mathrm{card}(\mathbf{Z}_K/P)$.

Soit a un élément de P^e qui n'appartient pas à P^{e+1} ; il en existe car si l'on avait $P^e=P^{e+1}$, l'idéal P^e aurait deux décompositions distinctes en produit d'idéaux maximaux, à savoir P^e et P^{e+1}.

Soit $\varphi\colon\mathbf{Z}_K\to P^e/P^{e+1}$ l'application telle que $\varphi(z)=\mathrm{cl}_{P^{e+1}}(az)$. C'est un homomorphisme de groupes abéliens ; son noyau Q est un sous-groupe de P^e ; c'est même un idéal de $\mathbf{Z}_K$ car si $z\in Q$ et $\lambda\in\mathbf{Z}_K$, $\varphi(\lambda z)=\mathrm{cl}(a\lambda z)=\lambda\,\mathrm{cl}(az)=0$. En outre, Q contient P. Cependant, $\varphi(1)=\mathrm{cl}(a)\neq 0$, donc $Q\neq\mathbf{Z}_K$. Les inclusions $P\subset Q\subsetneq\mathbf{Z}_K$ et l'hypothèse que P est un idéal maximal de $\mathbf{Z}_K$ entraînent alors $Q=P$.

Démontrons que φ est surjectif. Soit J l'idéal $P^{e+1}+(a)$. Il est contient P^{e+1} donc il existe un idéal J' de $\mathbf{Z}_K$ tel que $P^{e+1}=JJ'$. La décomposition de J en produit d'idéaux maximaux est donc de la forme $J=P^f$, où f est un entier naturel tel que $0\leqslant f\leqslant e+1$. De plus, J est contenu dans P^e (car P^{e+1} et (a) le sont), donc $f\geqslant e$. Toutefois, J n'est pas égal à P^{e+1} car $a\notin P^{e+1}$; ainsi, $f\neq e+1$. On a donc $f=e$ et $J=P^e$.

Soit $x\in P^e$; démontrons que $\mathrm{cl}(x)$ appartient à l'image de φ. Comme $x\in J=P^{e+1}+(a)$, il existe donc $z\in P^{e+1}$ et $w\in\mathbf{Z}_K$ tels que $x=z+aw$; alors $\mathrm{cl}(x)=\mathrm{cl}(z)+\mathrm{cl}(aw)=\varphi(w)$, comme il fallait démontrer, et φ est surjective.

Par conséquent, $\mathbf{Z}_K/P$ est un groupe abélien isomorphe à P^e/P^{e+1}. En particulier, $\mathrm{N}(P)=\mathrm{card}(\mathbf{Z}_K/P)=\mathrm{card}(P^e/P^{e+1})$.

Par récurrence, $\mathrm{N}(P^{e+1})=\mathrm{card}(\mathbf{Z}_K/P^e)\,\mathrm{card}(P^e/P^{e+1})=\mathrm{N}(P)^{e+1}$, d'où le lemme. □

THÉORÈME 4.4.3. — *Soit I et J des idéaux non nuls de $\mathbf{Z}_K$. On a l'égalité* $\mathrm{N}(IJ)=\mathrm{N}(I)\,\mathrm{N}(J)$.

Démonstration. — Écrivons $I=P_1^{m_1}\dots P_t^{m_t}$ la décomposition de l'idéal I en produits d'idéaux maximaux, où on suppose que $P_1,\dots,P_t$ sont deux à deux distincts. Nous allons démontrer l'égalité

$$\mathrm{N}(I)=\mathrm{N}(P_1)^{m_1}\dots\mathrm{N}(P_t)^{m_t}\quad;$$

le théorème en découlera en appliquant cette égalité ainsi que les deux égalités analogues pour $\mathrm{N}(J)$ et $\mathrm{N}(IJ)$.

Lorsque $t=1$, l'assertion résulte du lemme précédent. On raisonne en général par récurrence sur t. Posons en effet $I_1=P_1^{m_1}$, $I_2=P_2^{m_2}\dots P_t^{m_t}$. Vérifions que $I_1+I_2=\mathbf{Z}_K$. Comme $P_1^{m_1}\subset I_1+I_2$, il existe un entier e tel que $0\leqslant e\leqslant m_1$ de sorte que $I_1+I_2=P_1^e$.

Par construction, I_2 n'est pas contenu dans P_1 (sinon, P_1 apparaîtrait dans la décomposition de I_2 en produit d'idéaux maximaux) ; en particulier, $I_1 + I_2$ n'est pas contenu dans P_1 ce qui entraîne $e = 0$. Autrement dit, $I_1 + I_2 = \mathbf{Z}_K$.

D'après le lemme **??**, $\mathrm{N}(I) = \mathrm{N}(I_1 I_2) = \mathrm{N}(I_1)\,\mathrm{N}(I_2)$; par récurrence, on a aussi $\mathrm{N}(I_2) = \mathrm{N}(P_2)^{m_2} \ldots \mathrm{N}(P_t)^{m_t}$, tandis que le cas $t = 1$ implique $\mathrm{N}(I_1) = \mathrm{N}(P_1)^{m_1}$. L'égalité voulue en résulte, d'où le théorème. □

Exercices

20) a) Démontrer que l'anneau des entiers du corps quadratique $\mathbf{Q}(\sqrt{-19})$ est un anneau principal.

b) Même question pour les corps quadratiques de discriminants -43, -67 et -163.

21) Soit K un corps quadratique de discriminant Δ ; on pose $\omega = (1+\sqrt{\Delta})/2$ si $\Delta \equiv 1 \pmod 4$ et $\omega = \sqrt{\Delta}$ sinon, de sorte que $\mathbf{Z}_K = \mathbf{Z}[\omega]$.

a) Soit P un idéal maximal de $\mathbf{Z}_K$ contenant un nombre premier p. Démontrer que $\mathrm{N}_K(P) = p$ ou p^2. Si $\mathrm{N}_K(P) = p^2$, vérifier que $P = (p)$ et que p est un élément irréductible de $\mathbf{Z}_K$.

b) On suppose que $\mathrm{N}_K(P) = p$. Démontrer qu'il existe un entier $a \in \{0, \ldots, p-1\}$ tel que $P = (p, a-\omega)$. (*Par la méthode du cours, P est de la forme $(p, x+y\omega)$; observer que y n'est pas multiple de p.*)

c) On suppose que P n'est pas principal. Démontrer que $\left|\mathrm{N}_K(a-\omega+np)\right| \geqslant p^2$ pour tout $n \in \mathbf{Z}$.

d) Si $\Delta = 101$, par exemple, démontrer que $\mathbf{Z}_K$ est un anneau principal.

22) Soit I un idéal non nul d'un anneau d'entiers de corps de nombres dont la norme est un nombre premier. Démontrer que I est un idéal maximal.

23) Soit K un corps de nombres et $\mathbf{Z}_K$ son anneau d'entiers.

a) Soit I un idéal non nul de $\mathbf{Z}_K$ et P un idéal maximal de $\mathbf{Z}_K$. Démontrer qu'il existe un plus grand entier naturel e tel que $I \subset P^e$. On le note $\mathrm{ord}_P(I)$.

b) Soit I, J des idéaux non nuls de $\mathbf{Z}_K$ et soit P un idéal maximal de $\mathbf{Z}_K$. Démontrer que $\mathrm{ord}_P(IJ) = \mathrm{ord}_P(I) + \mathrm{ord}_P(J)$ et que $\mathrm{ord}_P(I+J) = \min(\mathrm{ord}_P(I), \mathrm{ord}_P(J))$.

c) Soit I un idéal non nul de $\mathbf{Z}_K$. Démontrer qu'il existe un élément $a \in I$ tel que pour tout idéal maximal P de $\mathbf{Z}_K$ qui contient I, $a \notin P^{1+\mathrm{ord}_P(I)}$. Soit J l'idéal de $\mathbf{Z}_K$ tel que $(a) = IJ$; démontrer que $I + J = \mathbf{Z}_K$.

24) Soit K un corps de nombres et $\mathbf{Z}_K$ son anneau d'entiers. Soit I un idéal non nul de $\mathbf{Z}_K$ et soit a un élément non nul de I.

a) Démontrer qu'il existe un idéal J de $\mathbf{Z}_K$ et un élément non nul $b \in \mathbf{Z}_K$ tels que $\mathbf{Z}_K = (a) + J$ et $IJ = (b)$. En déduire que l'idéal (a, b) engendré par a et b est contenu dans I.

b) Démontrer que $I = (a, b)$.

BIBLIOGRAPHIE

[1] J. W. S. CASSELS – *An introduction to the geometry of numbers*, Classics in Mathematics, Springer-Verlag, Berlin, 1997, Corrected reprint of the 1971 edition.

[2] C. F. GAUSS – *Recherches arithmétiques*, Jacques Gabay, 1989, Fac-simile de l'édition Courier, Paris, 1807 ; *http://gallica.bnf.fr/Catalogue/noticesInd/FRBNF35030770.htm*.

[3] E. HECKE – *Vorlesungen über die Theorie der algebraischen Zahlen*, Chelsea Publishing Co., Bronx, N.Y., 1970, Second edition of the 1923 original, with an index.

[4] E. LANDAU – *Elementary number theory*, Chelsea Publishing Co., New York, N.Y., 1958, Translated by J. E. Goodman.

[5] P. G. LEJEUNE DIRICHLET – *Lectures on number theory*, History of Mathematics – Sources, vol. 16, American Mathematical Society, London Mathematical Society, 1999.

[6] P. MONSKY – « Simplifying the proof of Dirichlet's theorem », *Amer. Math. Monthly* **100** (1993), no. 9, p. 861–862.

[7] G. TENENBAUM & M. MENDÈS-FRANCE – *Les nombres premiers*, Que Sais-Je ?, vol. 571, Presses Universitaires de France, 1997.

[8] D. ZAGIER – « Newman's short proof of the Prime number theorem », *Amer. Math. Monthly* **104** (1997), no. 8, p. 705–708.

TABLE DES MATIÈRES