L'ALGORITHMIQUE DE LA THÉORIE ALGÉBRIQUE DES NOMBRES

par

Karim Belabas

Table des matières

Commençons par une équation difficile... dans un contexte sans difficulté :

Théorème 0.1. — *Soit $n > 2$ et x, y, z trois* polynômes *à coefficients complexes, premiers entre eux. Alors $x^n + y^n = z^n$ implique que les trois polynômes sont des constantes.*

Démonstration. — Supposons qu'il existe trois polynômes (x, y, z) premiers entre eux satisfaisant $x^n + y^n = z^n$, tels que le maximum des trois degrés soit $D > 0$. On les choisit tels que D soit minimal. Posant $\zeta := \exp(2i\pi/n)$, on a

$$z^n = x^n + y^n = (x + y)(x + \zeta y)(x + \zeta^2 y) \cdots (x + \zeta^{n-1} y). \tag{1}$$

Deux facteurs quelconques du produit n'ont pas de diviseur commun, puisqu'un tel facteur diviserait x et y. Leur produit étant une puissance n-ème, ces facteurs sont de la forme βu^n, $\beta \in \mathbb{C}^*$, ce qui s'écrit encore $(\alpha u)^n$. Puisque $n - 1 \geqslant 2$, il existe des polynômes u, v, w premiers entre eux tels que

$$x + y = u^n, \quad x + \zeta y = v^n, \quad x + \zeta^2 y = w^n.$$

En éliminant x et y de ces équations, on trouve $w^n + \zeta u^n = (1+\zeta)v^n$. Les constantes entrent de nouveau sous les puissances et on pose $x' = w$, $y' = \zeta^{1/n} u$, $z' = (1+\zeta)^{1/n} v$, qui vérifient $x'^n + y'^n = z'^n$. Ces nouveaux polynômes sont toujours premiers entre eux ($\zeta + 1 = 0$ implique $n = 2$) et de degré maximal D' satisfaisant $0 < D' \leqslant D/n < D$. Contradiction. □

Il y a essentiellement deux arguments dans la preuve :

– $\mathbb{C}$ contient ζ et chaque complexe à une racine n-ème,

– $\mathbb{C}[X]$ est factoriel (notion de pgcd et écriture comme puissances n-èmes).

Le premier est surtout une étape conceptuelle : même si l'on s'intéresse aux solutions dans $\mathbb{Z}[X]$, raisonner dans $\mathbb{Z}[X]$ est inadapté. On agrandit l'anneau de base pour pouvoir travailler ; mais pas trop, pour ne pas vider le problème de sa substance arithmétique : dans l'anneau de séries formelles $\mathbb{C}[[X]]$, tout devient trivial. Le deuxième argument, par contre, est profond. Si l'on essaie de démontrer Fermat sur $\mathbb{Q}$ en copiant le premier point, on introduit l'anneau $\mathbb{Z}[\zeta]$, qui n'est presque jamais factoriel. La théorie algébrique des nombres s'est créée autour de cette question (et des généralisations de la loi de réciprocité quadratique de Legendre-Gauss), notamment avec l'invention des « nombres idéaux » par Kummer pour restaurer l'unicité de la décomposition en facteurs irréductibles.

Bien sûr, cela n'a pas suffi pas à résoudre le problème général de Fermat. Mais la théorie algébrique des nombres classique[(1)], essentiellement telle qu'axiomatisée par Dedekind puis Hecke, résout de très nombreuses équations concrètes,... avec le renfort des estimations modernes de transcendance (en particulier les minorations de formes linéaires en logarithmes), et du calcul algorithmique des invariants que nous allons introduire. Les méthodes « modulaires » (utilisant les théorèmes de Ribet et Wiles),... qui permettent de démontrer Fermat, se prêtent également à un traitement algorithmique, pour lequel je renvoie au survol de Siksek [**35**].

Dans une première partie, j'introduis le langage de la théorie algébrique des nombres, en démontrant tout ce qui peut se faire en quelques lignes. On se contente de résultats classiques, en évitant les notions et les preuves les plus générales. Voir par exemple Lang [**25**] pour les démonstrations manquantes.

Une deuxième partie, algorithmique, rendra tout ceci effectif. Elle est largement inspirée par Lenstra [**28**] et Cohen [**10**]. À partir de données explicites, quelles quantités sont effectivement calculables ? Comment et sous quelle forme ? En quel temps ? Je n'ai pas cherché à être exhaustif, ni toujours efficace, ni suffisamment explicite pour permettre une implantation directe (dans ce but, se référer à [**17, 10, 6**]), mais à présenter un panorama d'idées et d'algorithmes essentiels, sous une forme que j'espère compréhensible.

(1) Les développements ultérieurs (théorie du corps de classe, théorie d'Iwasawa, structures galoisiennes et cohomologie, K-théorie algébrique...) nous éloigneraient un peu des motivations diophantiennes immédiates et ne seront plus mentionnées.

PARTIE I
THÉORIE ALGÉBRIQUE DES NOMBRES

1. Préliminaires

1.1. $\mathbb{Z}$-modules de type fini. — Un $\mathbb{Z}$-module est un groupe abélien G. Il est *de type fini* s'il est engendré par un nombre fini d'éléments. Il est *libre* s'il admet une base, c'est-à-dire une famille génératrice $(e_i)_{i\leqslant n}$ libre : si $\sum_{i\leqslant n}\lambda_i e_i = 0$, $\lambda_i \in \mathbb{Z}$, alors $\lambda_1 = \cdots = \lambda_n = 0$. Contrairement aux espaces vectoriels, les modules n'admettent pas nécessairement de base, par exemple $G = \mathbb{Z}/2\mathbb{Z}$. Par contre s'il existe des bases, elles ont bien le même cardinal n, appelé *rang* de G et noté $\operatorname{rg} G$. Les changements de bases sont les éléments de $\mathrm{GL}_n(\mathbb{Z})$.

***Théorème 1.1* (base adaptée).** — *Si G est un $\mathbb{Z}$-module libre de type fini et H un sous-$\mathbb{Z}$-module, il existe une base (g_i) de G et $d_n \mid \cdots \mid d_2 \mid d_1$ dans $\mathbb{N}$ tels que $\{d_i g_i : d_i > 0\}$ soit une base de H. Les (d_i) ne dépendent pas de la base de G choisie.*

En particulier :

Corollaire 1.2. — *Un sous-module H d'un $\mathbb{Z}$-module libre de type fini G est libre de type fini. De plus, $\operatorname{rg} H \leqslant \operatorname{rg} G$.*

Plus généralement :

***Corollaire 1.3* (diviseurs élémentaires).** — *Si G est un $\mathbb{Z}$-module de type fini, il existe $g_1, \dots, g_n \in G$ tels que*

$$G = \bigoplus_{i=1}^{n} (\mathbb{Z}/d_i\mathbb{Z}) \cdot g_i,$$

où $d_n \mid \cdots \mid d_2 \mid d_1$ dans $\mathbb{N}$ ne dépendent que de G.

Démonstration. — Considérer M, le module libre sur les générateurs de G, et $\Lambda \subset M$, le noyau de la projection canonique $M \to G$. Dans une base adaptée, interpréter $G = M/\Lambda$. □

On appelle les (d_i) les *diviseurs élémentaires* de G (*cf.* § 10.4). Le groupe fini $G_{\mathrm{tor}} := \oplus_{i,d_i>0}(\mathbb{Z}/d_i\mathbb{Z}) \cdot g_i$ est la *torsion* de G, et $r := \#\{i : d_i = 0\} = \operatorname{rg} G$ son *rang*. Pour p premier, on note $r_p(G) = \dim_{\mathbb{F}_p} G/pG = \#\{i : p \mid d_i\}$ le p-rang de G.

En particulier, tout groupe abélien fini est somme directe de groupes cycliques et un $\mathbb{Z}$-module de type fini est libre si et seulement s'il est sans torsion. Si $H \subset G$ sont deux $\mathbb{Z}$-modules, on note $[G : H] = \#(G/H) \leqslant +\infty$ l'*indice* de H dans G.

Corollaire 1.4. — *Si $\phi : \mathbb{Z}^n \to \mathbb{Z}^n$ est un morphisme dont l'image est de rang n,*

$$[\mathbb{Z}^n : \operatorname{Im} \phi] = |\det \phi| .$$

Démonstration. — On choisit une base (g_i) de $\mathbb{Z}^n$ telle que $(d_i g_i)$ soit une base de $\operatorname{Im} \phi$ (comme ϕ est de rang n, aucun d_i n'est nul), donc $[\mathbb{Z}^n : \operatorname{Im} \phi] = d_1 \cdots d_n$. Le morphisme $\psi : d_i g_i \mapsto \phi(g_i)$ est un changement de base de $\operatorname{Im} \phi$, donc de déterminant ± 1. Comme ϕ est la composition de l'application diagonale $g_i \mapsto d_i g_i$ et de ψ, on obtient $|\det \phi| = d_1 \cdots d_n$. □

1.2. Réseaux, déterminant, discriminant. — Si (E, q) est un espace euclidien de dimension n, un réseau $\Lambda \subset E$ est un sous-$\mathbb{Z}$-module libre de rang n. Si $(e_i)_{i \leqslant n}$ est une base de Λ, le *discriminant* de Λ est le déterminant de la matrice de Gram

$$\operatorname{Gram}(e_1, \ldots, e_n) = \big(\langle e_i, e_j \rangle \big)_{1 \leqslant i,j \leqslant n}.$$

Le *déterminant* $d(\Lambda) = d(\Lambda, q)$ de Λ est la racine carrée du discriminant de Λ. Il est égal au volume du parallélotope fondamental

$$\Big\{ \sum x_i e_i \colon 0 \leqslant x_i < 1 \Big\} .$$

Proposition 1.5. — *Si ϕ est un endomorphisme de E, de matrice A*

$$\operatorname{Gram}(\phi(e_1), \ldots, \phi(e_n)) = {}^t\!A \cdot \operatorname{Gram}(e_1, \ldots, e_n) \cdot A.$$

Corollaire 1.6

(1) *$d(\Lambda)$ ne dépend pas de la base (e_i) choisie.*
(2) *Si $L \subset \Lambda$ est un sous-réseau, $d(L) = d(\Lambda)[\Lambda : L]$.*

(Noter que si $L \subset \Lambda$ est un sous-réseau, l'indice $[\Lambda : L]$ est fini, puisque L et Λ ont même rang n.) En particulier, si $E = \mathbb{R}^n$, muni de la forme euclidienne standard, et $\Lambda = \operatorname{Im} \phi$, pour $\phi : \mathbb{Z}^n \to E$ de rang n, alors $d(\Lambda) = |\det \phi|$.

Nous arrivons au résultat principal de cette partie, dont le seul tort est d'être *non effectif.* (L'algorithme LLL (§ 10.5) en donnera une variante effective.)

***Théorème 1.7* (Minkowski).** — *Si* $C \subset \mathbb{R}^n$ *est convexe, symétrique autour de* 0 $(-C = C)$ *de volume* $V(C) \leqslant +\infty$*, et si* Λ *est un réseau de déterminant* $d(\Lambda) < 2^{-n}V(C)$*, alors* $C \cap \Lambda \neq (0)$*.*

La preuve est classique, je la reproduis pour manifester son caractère non effectif. Une variante pratique : si C est de plus compact, alors l'égalité large $d(\Lambda) \leqslant 2^{-n}V(C)$ suffit pour obtenir la conclusion (exercice).

Démonstration. — On commence par montrer le théorème de Blichfeldt : si Λ est un réseau, $S \subset \mathbb{R}^n$ et $d(\Lambda) < V(S)$, alors il existe $s_1, s_2 \in S$, $s_1 \neq s_2$, tels que $s_1 - s_2 \in \Lambda$. C'est clair : si P est un parallélotope fondamental pour Λ, les $S_x := S \cap (x + P)$, $x \in \Lambda$, forment une partition de S, donc

$$V(S) = \sum_{x \in \Lambda} V(S_x) = \sum_{x \in \Lambda} V(S_x - x) > d(\Lambda) = V(P).$$

Comme les $S_x - x$ sont inclus dans P, ils ne sont pas disjoints.

Le théorème est un corollaire immédiat pour $S = \frac{1}{2}C$, de volume $2^{-n}V(C) > d(\Lambda)$. Par définition, $2s_1, 2s_2 \in C$, par symétrie $-2s_2 \in C$, et par convexité $c := \frac{1}{2}(2s_1 - 2s_2) \in C$. □

À titre d'entraînement :

Exercice 1.8. — Soit $p \equiv 1 \pmod 4$ un nombre premier, $r \in \mathbb{Z}$ tel que $r^2 \equiv -1 \pmod p$ et $\Lambda \subset \mathbb{R}^2$ le réseau engendré par les colonnes de la matrice $\left(\begin{smallmatrix} p & r \\ 0 & 1 \end{smallmatrix}\right)$. Montrer que Λ contient un point (a, b) tel que $0 < a^2 + b^2 < 2p$. En déduire que $p = a^2 + b^2$ est somme de deux carrés. [*Solution.* $C := B(0, \sqrt{2p})$, $V(C) = 2\pi p > 2^2 p = 2^n d(\Lambda)$.]

Exercice 1.9. — Soit p un nombre premier, montrer qu'il existe $r, s \in \mathbb{Z}$ tels que $r^2 + s^2 \equiv -1 \pmod p$ [*il y a* $(p+1)/2$ *valeurs possibles pour* r^2*, autant pour* $-1 - s^2$]. Soit $\Lambda \subset \mathbb{R}^4$ le réseau engendré par les colonnes de

$$\begin{pmatrix} p & 0 & r & s \\ 0 & p & s & -r \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Montrer que Λ contient un point (a,b,c,d) tel que $0 < a^2+b^2+c^2+d^2 < 2p$. [$d(\Lambda) = p^2$, $C = B(0,\sqrt{2p})$, $V(C) = \frac{\pi^2}{2}(2p)^2 > 2^4 d(\Lambda)$.] En déduire que p est somme de 4 carrés. [*Corollaire : tout entier positif est somme de* 4 *carrés.*]

1.3. Nombres p-adiques. — $\mathbb{R}$ est le complété de $\mathbb{Q}$ pour la topologie associée à la valeur absolue usuelle. En d'autres termes, on peut réaliser $\mathbb{R}$ comme l'anneau des suites de Cauchy de $\mathbb{Q}$, modulo l'idéal des suites de limite nulle. C'est un corps, auquel s'étend la valeur absolue de $\mathbb{Q}$, et il est complet pour la topologie associée.

Si p est premier, et $v_p(x)$ désigne la valuation p-adique de x, la formule $|x|_p := p^{-v_p(x)}$ définit de même une valeur absolue sur $\mathbb{Q}$, donc une topologie (cf. § 4.7). Le complété de $\mathbb{Q}$ par rapport à cette topologie p-adique se note $\mathbb{Q}_p$. La valuation v_p et $|\cdot|_p$ s'étendent à $\mathbb{Q}_p$.

On peut voir $\mathbb{Q}_p$ comme le corps de fractions de $\mathbb{Z}_p = \varprojlim_{n>0} \mathbb{Z}/p^n\mathbb{Z}$, c'est-à-dire qu'un élément x de $\mathbb{Z}_p$ est donné par une suite d'approximations $(x_k)_{k\geqslant 1}$, $x_k \in \mathbb{Z}/p^k\mathbb{Z}$, où $x_{k+1} \equiv x_k \pmod{p^k}$ pour tout $k \geqslant 1$. En particulier il y a une projection canonique $\phi_k : \mathbb{Z}_p \to \mathbb{Z}/p^k\mathbb{Z}$, donnée par $x \mapsto x_k$, pour tout $k \geqslant 1$.

***Théorème 1.10* (Lemme de Hensel faible).** — *Soit $F \in \mathbb{Z}_p[X]$ unitaire, $\overline{F}$ sa réduction dans $\mathbb{F}_p[X]$. On suppose que $\overline{F} = \prod_{i=1}^g \overline{F}_i$ dans $\mathbb{F}_p[X]$, où les $\overline{F}_i$ sont unitaires, deux à deux premiers entre eux. Alors il existe $F_1, \dots, F_g \in \mathbb{Z}_p[X]$ unitaires, tels que la réduction de F_i dans $\mathbb{F}_p[X]$ soit $\overline{F}_i$ pour tout $i \leqslant g$, et tels que $F = \prod_{i=1}^g F_i$.*

En d'autres termes une factorisation *sans facteurs carrés* dans $\mathbb{F}_p[X]$ se relève à $\mathbb{Z}_p[X]$. Plus généralement, une factorisation suffisamment précise pour qu'il n'y ait plus de facteurs carrés se relève :

***Théorème 1.11* (Lemme de Hensel).** — *On suppose que $F, F_1, \dots, F_g \in \mathbb{Z}_p[X]$ sont unitaires, et $a_1, \dots, a_g$ dans $\mathbb{Z}_p[X]$ tels que*

$$F \equiv \prod_{i=1}^{g} F_i \pmod{p^e \mathbb{Z}_p[X]} \quad \textit{et}$$
$$p^d \equiv \sum_{i=1}^{g} a_i \prod_{j \neq i} F_j \pmod{p^{d+1} \mathbb{Z}_p[X]},$$

où $d \geqslant 0$, $e > 2d$. Alors il existe des polynômes $\widehat{F_1}, \dots, \widehat{F_g} \in \mathbb{Z}_p[X]$ unitaires, $\widehat{F_i} \equiv F_i \pmod{p^e \mathbb{Z}_p[X]}$ tels que $F = \prod_{i=1}^{g} \widehat{F_i}$.

Démonstration. — On construit par récurrence, pour tout $k \geqslant e$, des polynômes $F_1^{(k)}, \dots, F_g^{(k)}$ de $\mathbb{Z}_p[X]$ tels que

$$F \equiv \prod_{i=1}^{g} F_i^{(k)} \pmod{p^{k+1} \mathbb{Z}_p[X]} \quad \text{et}$$
$$p^d \equiv \sum_{i=1}^{g} a_i \prod_{j \neq i} F_j^{(k)} \pmod{p^{d+1} \mathbb{Z}_p[X]}.$$

On pose ensuite $\widehat{F_i} := \lim_{k \to \infty} F_i^{(k)}$. (La suite est de Cauchy dans $\mathbb{Z}_p[X]_{\deg F_i}$, qui est complet.) Soit donc $u := p^{-k}(F - \prod_i F_i^{(k)})$, et pour $i = 1, \dots, g$, soit v_i le reste de la division euclidienne de ua_i par F_i. On pose $F_i^{(k+1)} := F_i^{(k)} + p^{k-d} v_i$. ☐

La congruence impliquant les a_i est une identité de type Bezout, qui entraîne que les F_i sont premiers entre eux ; en fait, ils le sont déjà modulo p^{d+1}. La version faible est le cas $d = 0$, $e = 1$. Il est crucial pour nos applications que la preuve soit constructive.

2. Corps de nombres

Un *corps de nombres* K est une extension finie de $\mathbb{Q}$, c'est-à-dire un corps contenant $\mathbb{Q}$, de dimension finie comme $\mathbb{Q}$-espace vectoriel ; cette dimension est appelée *degré* de K. Si α est un nombre algébrique, par exemple $\alpha = \sqrt{2}$ ou $\alpha = \exp(2i\pi/m)$ ($m > 0$ entier), la $\mathbb{Q}$-algèbre $\mathbb{Q}(\alpha)$ est un corps de nombres, respectivement de degré 2 et $\phi(m)$.

Réciproquement, le théorème de l'élément primitif dit que si K est un corps de nombres, il existe $\alpha \in K$ tel que $K = \mathbb{Q}(\alpha)$. Si K

est de degré n et $T \in \mathbb{Q}[X]$ est le polynôme minimal de α sur $\mathbb{Q}$, alors T est de degré n, irréductible dans $\mathbb{Q}[X]$, et l'évaluation $Q \mapsto Q(\alpha)$ induit un isomorphisme de $\mathbb{Q}[X]/(T)$ sur K. Rappelons que le polynôme minimal $P_{\min,x}$ de $x \in K$ est le générateur *unitaire* de l'idéal $\{Q \in \mathbb{Q}[X], Q(x) = 0\}$. On conserve ces notations dans la suite du texte.

2.1. Plongements, signature. — Factorisons $T = \prod_{i=1}^{n}(X - \alpha_i)$ sur $\mathbb{C}$; les α_i sont les *conjugués* de α. Soit r_1 le nombre de racines réelles de T, $2r_2$ le nombre de racines non réelles. On ordonne les α_i de façon à ce que $\alpha_1, \dots, \alpha_{r_1} \in \mathbb{R}$, $\alpha_{i+r_1} = \overline{\alpha_{i+r_1+r_2}}$ pour $1 \leqslant i \leqslant r_2$. Chaque α_i définit un plongement $\sigma_i : Q \mapsto Q(\alpha_i)$ du corps abstrait $K = \mathbb{Q}[X]/(T)$ dans $\mathbb{C}$. Tout morphisme de corps de K dans $\mathbb{C}$ est manifestement de cette forme ; les σ_i, $i \leqslant r_1$ sont les *plongements réels* de K et les σ_i, $r_1 < i \leqslant n$ ses *plongements complexes*, (r_1, r_2) est sa *signature*. Bien sûr, $r_1 + 2r_2 = n$.

Chaque σ_i s'étend à $\mathbb{R}[X]/(T)$ et $\sigma := (\sigma_1, \dots, \sigma_{r_1+r_2})$ définit un isomorphisme de $\mathbb{R}$-algèbres de $\mathbb{R}[X]/(T)$ dans $\mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$. En d'autres termes la signature (r_1, r_2) donne la structure de la $\mathbb{R}$-algèbre $K \otimes_{\mathbb{Q}} \mathbb{R}$.

2.2. Trace, norme. — Soit x dans K ; la multiplication par x induit un endomorphisme $m_x : K \to K$ de $\mathbb{Q}$-espace vectoriel. On définit le polynôme caractéristique de x, noté $P_{\mathrm{char},x}$, la trace de x, notée $\mathrm{Tr}(x)$, et la norme de x, notée $\mathrm{N}(x)$, respectivement comme le polynôme caractéristique, la trace et le déterminant de m_x. La trace est un morphisme de $\mathbb{Q}$-espaces vectoriels de $K \to \mathbb{Q}$, la norme un morphisme de groupes $K^* \to \mathbb{Q}^*$. Plus généralement, ces définitions ont un sens dans n'importe quelle k-algèbre commutative de dimension finie sur un corps k (on les utilisera dans $\mathbb{Q}_p[X]/(T)$ au § 15.3).

La matrice de m_α dans la base $(1, \alpha, \dots, \alpha^{n-1})$ est la matrice compagnon de $P_{\mathrm{char},\alpha} = T$. (En particulier $P_{\mathrm{char},\alpha}(\alpha) = 0$, qui résulte aussi de Cayley-Hamilton.) En se ramenant au cas primitif $K = \mathbb{Q}(\alpha)$, on montre plus généralement que

Théorème 2.1. — *On a*

$$P_{\mathrm{char},x}(X) = \prod_{i=1}^{n}(X - \sigma_i(x)) = P_{\min,x}(X)^{\dim_{\mathbb{Q}(x)} K}, \tag{2}$$

$$\mathrm{Tr}(x) = \sum_{i=1}^{n} \sigma_i(x), \tag{3}$$

$$\mathrm{N}(x) = \prod_{i=1}^{n} \sigma_i(x). \tag{4}$$

Ces calculs sont effectués dans $\mathbb{C}$, mais le résultat final appartient au sous-corps $\mathbb{Q}$.

Théorème 2.2. — *La forme bilinéaire* $T : (x, y) \mapsto \mathrm{Tr}(xy)$ *sur* K^2 *est non dégénérée.*

Démonstration. — Si $x \neq 0$, alors $\mathrm{Tr}(x \cdot 1/x) = \mathrm{Tr}(1) = n \neq 0$. □

L'extension de cette forme quadratique à $K \otimes_{\mathbb{Q}} \mathbb{R} \simeq \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ est

$$(x_1, \ldots, x_{r_1}, a_1 + ib_1, \ldots, a_{r_2} + ib_{r_2}) \longmapsto \sum_{i \leqslant r_1} x_i^2 + 2 \sum_{j \leqslant r_2} (a_j^2 - b_j^2), \tag{5}$$

de signature $(r_1 + r_2, r_2)$.

2.3. Théorie de Galois. — On va peu en parler, par manque de temps, et l'utiliser comme hypothèse technique dans des théorèmes ultérieurs. L'extension $K/\mathbb{Q}$ de degré n est *galoisienne* si le corps K admet n automorphismes ; autrement dit, si les conjugués α_i de α sont tous dans K. (Un automorphisme fixe $\mathbb{Q}$, donc est nécessairement de la forme $\alpha \mapsto \alpha_i$). On notera $\mathrm{Gal}(K/\mathbb{Q})$ le groupe des automorphismes de K. Par exemple, un corps quadratique est galoisien, de groupe $\mathbb{Z}/2\mathbb{Z}$, ainsi que les corps cyclotomiques $\mathbb{Q}\big(\exp(2i\pi/n)\big)$, de groupe $(\mathbb{Z}/n\mathbb{Z})^*$.

Remarquons que si $K/\mathbb{Q}$ est galoisienne de groupe de Galois G, on peut considérer K comme sous-corps de $\mathbb{C}$ et les σ_i ne sont rien d'autre que les éléments de G. Plus précisément, pour chaque $i \leqslant n$, il existe un unique $\tau_i \in G$ tel que $\sigma_i = \sigma_1 \circ \tau_i$.

Si $T \in \mathbb{Q}[X]$, le corps de décomposition $K_T \subset \mathbb{C}$ de T est le plus petit corps contenant toutes les racines complexes de T. C'est une extension galoisienne de $\mathbb{Q}$. Par abus de langage, on appelle *groupe de Galois* de T le groupe de Galois de K_T.

3. Anneau des entiers

Soit $x \in K$; x est un *entier algébrique* (ou simplement un *entier* s'il n'y a pas risque de confusion) s'il satisfait les conditions équivalentes suivantes :

(1) $P_{\mathrm{min},x} \in \mathbb{Z}[X]$,
(2) $P_{\mathrm{char},x} \in \mathbb{Z}[X]$,
(3) il existe $Q \in \mathbb{Z}[X]$, Q unitaire, tel que $Q(x) = 0$,
(4) $\mathbb{Z}[x]$ est un $\mathbb{Z}$-module de type fini,
(5) il existe $M \subset K$ un $\mathbb{Z}$-module de type fini contenant $\mathbb{Z}[x]$.

On note $\mathcal{O}_K$ l'ensemble des entiers de K. On a $\mathbb{Z} \subset \mathcal{O}_K$; on voit facilement que $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$.

Corollaire 3.1. — *En particulier, si $x \in \mathcal{O}_K$, $\mathrm{Tr}\, x \in \mathbb{Z}$, $\mathrm{N}\, x \in \mathbb{Z}$.*

Démonstration. — C'est une conséquence de la deuxième caractérisation. □

Théorème 3.2. — *$\mathcal{O}_K$ est un anneau intègre, de corps de fractions K.*

Démonstration. — La quatrième caractérisation montre que $\mathcal{O}_K \subset K$ est stable par addition et multiplication : si $\mathbb{Z}[x], \mathbb{Z}[y]$ sont de type fini, $\mathbb{Z}[x, y] \supset \mathbb{Z}[x+y], \mathbb{Z}[xy]$ aussi. Donc $\mathcal{O}_K$ est un anneau. Étant contenu dans un corps, il est intègre.

Si $x \in K$, soit $d \in \mathbb{Z} \subset \mathcal{O}_K$ le dénominateur commun des coefficients de $P_{\mathrm{min},x}$. Alors $dx \in \mathcal{O}_K$ puisque son polynôme minimal $d^n P_{\mathrm{min},x}(X/d)$ est dans $\mathbb{Z}[X]$. Donc $K = \mathrm{Frac}\, \mathcal{O}_K$. □

Ceci montre que $\mathcal{O}_K$ contient une $\mathbb{Q}$-base de K, $(e_1, \ldots, e_n)$. Soit $(f_1, \ldots, f_n)$ la base duale par rapport à la forme bilinéaire non dégénérée du Théorème 2.2.

Proposition 3.3. — $\mathcal{O}_K \subset \langle f_1, \ldots, f_n \rangle_{\mathbb{Z}}$.

Démonstration. — Si $x \in \mathcal{O}_K$, $x = \sum x_i f_i$, $x_i \in \mathbb{Q}$. Pour tout i, $xe_i \in \mathcal{O}_K$ d'après le Théorème 3.2, donc $\mathrm{Tr}(xe_i) = x_i \in \mathbb{Z}$ (Corollaire 3.1). □

Corollaire 3.4. — *$\mathcal{O}_K$ est un $\mathbb{Z}$-module libre de rang n.*

Démonstration. — Le $\mathbb{Z}$-module engendré par les f_i est de type fini et sans torsion, donc libre. Donc $\mathcal{O}_K$ est de type fini, libre de rang $\leqslant n$ (Proposition 3.3 et Corollaire 1.2). Il contient une $\mathbb{Q}$-base de K, donc est de rang exactement n. □

Donc il existe une base $(w_1, \dots, w_n)$ tel que tout $x \in \mathcal{O}_K$ s'écrive de façon unique $x = \sum x_i w_i$, $x_i \in \mathbb{Z}$. Ceci implique en particulier que $\mathcal{O}_K / d\mathcal{O}_K$ est de cardinal d^n pour tout entier $d > 0$.

Exemple. — Un entier $d \in \mathbb{Z}$ est dit sans facteurs carrés s'il n'existe pas p premier tel que $p^2 \mid d$. Soit donc $d \in \mathbb{Z}$ sans facteurs carrés, $T = X^2 - d$, α une racine de T et $K = \mathbb{Q}(\alpha)$. Si $x = \frac{1}{2}(u + v\alpha) \in \mathbb{Q}(\alpha)$, $u, v \in \mathbb{Q}$, on calcule $P_{\mathrm{char},x} = X^2 - uX + (u^2 + dv^2)/4$. Donc $x \in \mathcal{O}_K$ si et seulement si $u \in \mathbb{Z}$, $u^2 - dv^2 \in 4\mathbb{Z}$; on en déduit $dv^2 \in \mathbb{Z}$, soit $v \in \mathbb{Z}$ (car d est divisible par le carré du dénominateur de v). Les seuls carrés modulo 4 étant 0 et 1, on en déduit que u, v sont pairs si $d \equiv 2, 3 \pmod 4$ et de même parité si $d \equiv 1 \pmod 4$. Soit

$$\mathcal{O}_K = \begin{cases} \mathbb{Z}[\alpha] & \text{si } d \equiv 2, 3 \pmod 4, \\ \mathbb{Z}[(1+\alpha)/2] & \text{si } d \equiv 1 \pmod 4. \end{cases}$$

4. Idéaux

4.1. Produit d'idéaux, idéaux fractionnaires, idéaux inversibles. — Soit R un anneau intègre, de corps de fractions K. Un idéal *non nul* de R est dit *entier*. Si $\mathfrak{a}, \mathfrak{b}$ sont deux idéaux entiers, on définit leur produit $\mathfrak{a}\mathfrak{b}$ comme l'idéal engendré par les produits ab, $a \in \mathfrak{a}$, $b \in \mathfrak{b}$. C'est une opération associative dont R est l'élément neutre.

On appelle *idéal fractionnaire* de K un sous R-module $\mathfrak{a}$ de K tel que $d\mathfrak{a}$ soit entier pour un $d \in K^*$. (En particulier un idéal fractionnaire est non nul.) La multiplication des idéaux s'étend aux idéaux fractionnaires : si $\mathfrak{a}, \mathfrak{b}$ sont entiers, on pose $(\mathfrak{a}/a)(\mathfrak{b}/b) = \mathfrak{a}\mathfrak{b}/ab$. Un idéal fractionnaire $\mathfrak{a}$ de K est *inversible* s'il existe un idéal fractionnaire $\mathfrak{b}$ tel que $\mathfrak{a}\mathfrak{b} = R$; on note $\mathfrak{b} = \mathfrak{a}^{-1}$.

Un anneau intègre est dit *de Dedekind* si tout idéal fractionnaire est inversible.

4.2. Norme d'idéaux. — On note $\mathrm{N}\,\mathfrak{a} = \#(\mathcal{O}_K/\mathfrak{a}) \leqslant +\infty$ la *norme de* $\mathfrak{a}$; en particulier, si $\mathfrak{a} = (a)$ est principal, le Corollaire 1.4 appliqué à $\phi = m_a$ donne $\mathrm{N}\,\mathfrak{a} = |\mathrm{N}(a)|$.

Proposition 4.1. — *Si $\mathfrak{a} \subset \mathcal{O}_K$ est un idéal non nul, le quotient $\mathcal{O}_K/\mathfrak{a}$ est fini.*

Démonstration. — Si $a \in \mathfrak{a} \smallsetminus \{0\}$, la surjection $\mathcal{O}_K/(a) \twoheadrightarrow \mathcal{O}_K/\mathfrak{a}$ montre que $\mathrm{N}\,\mathfrak{a} \mid \mathrm{N}(a)$, qui est un entier non nul. □

Corollaire 4.2. — *Un idéal premier non nul de $\mathcal{O}_K$ est maximal.*

Démonstration. — Un anneau intègre fini est un corps. (La multiplication par un élément non nul est un endomorphisme injectif, donc surjectif par finitude.) □

4.3. Théorème fondamental de l'arithmétique

Théorème 4.3. — *Si K est un corps de nombres, $\mathcal{O}_K$ est un anneau de Dedekind. De plus, on a*

$$\mathfrak{a}^{-1} = (\mathcal{O}_K : \mathfrak{a}) := \{x \in K : x\mathfrak{a} \subset \mathcal{O}_K\}\,.$$

La démonstration n'est pas difficile mais assez technique, on va seulement l'esquisser. On voit facilement que $\mathfrak{a}' := (\mathcal{O}_K : \mathfrak{a})$ est un idéal fractionnaire et $\mathfrak{a}\mathfrak{a}' \subset \mathcal{O}_K$, mais l'inclusion réciproque n'a rien d'évident. L'étape cruciale consiste à montrer que le théorème est vrai pour $\mathfrak{a} = \mathfrak{p}$ maximal. Dans ce cas on montre d'abord l'existence de $a \in \mathfrak{a}' \smallsetminus \mathcal{O}_K$, qui vient d'un argument de maximalité, utilisant le caractère noethérien de $\mathcal{O}_K$ et le fait que tout idéal premier non nul est maximal. Comme $\mathfrak{a}'\mathfrak{p}$ est un idéal entier et $\mathfrak{p}$ est maximal, la double inclusion $\mathfrak{p} \subset \mathfrak{a}'\mathfrak{p} \subset \mathcal{O}_K$ implique $\mathfrak{a}'\mathfrak{p} = \mathcal{O}_K$ (et on a fini), ou $\mathfrak{a}'\mathfrak{p} = \mathfrak{p}$. Mais ce dernier cas est impossible puisque $\mathfrak{p}$ est un $\mathbb{Z}$-module de type fini, donc $a\mathfrak{p} \subset \mathfrak{p}$ impliquerait $a \in \mathcal{O}_K$. On conclut en montrant que tout idéal entier est produit d'idéaux premiers, avec le même type d'argument noethérien que celui omis ci-dessus.

Plus généralement, la démonstration prouve qu'un anneau intègre R est de Dedekind si et seulement s'il est

– noethérien (tout R-idéal a un nombre fini de générateurs),

– intégralement clos (si $x \in \operatorname{Frac} R$ est racine d'un polynôme unitaire de $R[X]$ alors $x \in R$),

– de dimension de Krull 1 (tout idéal premier non nul est maximal).

La preuve a un corollaire important, généralisation du théorème fondamental de l'arithmétique sur $\mathbb{Z}$, qui montre que $\mathcal{O}_K \subset K$ est une généralisation appropriée de $\mathbb{Z} \subset \mathbb{Q}$:

Corollaire 4.4. — *Si K est un corps de nombres, tout idéal fractionnaire $\mathfrak{a}$ se factorise de manière unique sous la forme $\mathfrak{a} = \prod_{\mathfrak{p}} \mathfrak{p}^{e_{\mathfrak{p}}}$, où $\mathfrak{p}$ parcourt l'ensemble des idéaux maximaux de $\mathfrak{a}$, $e_{\mathfrak{p}} \in \mathbb{Z}$, tous nuls sauf un nombre fini. (On convient que le produit vide est l'élément neutre de la multiplication, $\mathcal{O}_K$. En particulier $\mathfrak{p}^0 = \mathcal{O}_K$.)*

Démonstration. — Il suffit de le démontrer pour un idéal entier. Seule l'unicité reste à voir : on choisit un idéal entier $\mathfrak{a}$, qui a une décomposition de longueur minimale parmi les idéaux entiers admettant plusieurs décompositions. Soit $\mathfrak{a} = \mathfrak{p}_1 \cdots \mathfrak{p}_r = \mathfrak{q}_1 \cdots \mathfrak{q}_s$, où les $\mathfrak{p}_i$ et les $\mathfrak{q}_j$ sont maximaux et r minimal. Alors $\mathfrak{q}_1 \cdots \mathfrak{q}_s \subset \mathfrak{p}_1$ et, $\mathfrak{p}_1$ étant premier, l'un des $\mathfrak{q}_i$ est contenu dans $\mathfrak{p}_1$; $\mathfrak{q}_i$ étant maximal, ils sont égaux. En multipliant par $\mathfrak{p}_1^{-1}$, on obtient un idéal entier admettant deux décomposition, dont l'une de longueur $< r$. Contradiction. □

L'existence d'une décomposition est vraie sous des hypothèses plus faibles. Le point crucial est l'*unicité*.

4.4. Valuation et divisibilité. — Pour $\mathfrak{a} = \prod \mathfrak{p}^{e_{\mathfrak{p}}}$ fractionnaire et $\mathfrak{p}$ maximal, on définit la valuation $\mathfrak{p}$-adique de $\mathfrak{a}$, $v_{\mathfrak{p}}(\mathfrak{a}) := e_{\mathfrak{p}}$. Pour $x \in K^*$ on pose $v_{\mathfrak{p}}(x) := v_{\mathfrak{p}}(x\mathcal{O}_K)$, et $v_{\mathfrak{p}}(0) = +\infty$. Noter que

$$v_{\mathfrak{p}}(x+y) \geqslant \min(v_{\mathfrak{p}}(x), v_{\mathfrak{p}}(x)), \tag{6}$$

avec égalité si $v_{\mathfrak{p}}(x) \neq v_{\mathfrak{p}}(y)$. Finalement on écrit pour $x, y \in K$, $\mathfrak{a}$ un idéal fractionnaire :

$$x \equiv y \pmod{\mathfrak{a}}, \quad \text{si} \quad v_{\mathfrak{p}}(x-y) \geqslant v_{\mathfrak{p}}(\mathfrak{a}), \forall\, \mathfrak{p}.$$

Par exemple, $\frac{1}{2} \equiv \frac{5}{2} \pmod{2\mathbb{Z}}$ dans $\mathbb{Q}$.

Grâce au théorème fondamental, on transporte aux idéaux fractionnaires le vocabulaire de la divisibilité dans $\mathbb{Z}$: $\mathfrak{a} \mid \mathfrak{b}$ si $v_{\mathfrak{p}}(\mathfrak{a}) \leqslant v_{\mathfrak{p}}(\mathfrak{b})$, $\forall\, \mathfrak{p}$; on pose

$$\operatorname{pgcd}(\mathfrak{a}, \mathfrak{b}) := \prod_{\mathfrak{p}} \mathfrak{p}^{\min(v_{\mathfrak{p}}(\mathfrak{a}), v_{\mathfrak{p}}(\mathfrak{b}))},$$

$$\operatorname{ppcm}(\mathfrak{a}, \mathfrak{b}) := \prod_{\mathfrak{p}} \mathfrak{p}^{\max(v_{\mathfrak{p}}(\mathfrak{a}), v_{\mathfrak{p}}(\mathfrak{b}))},$$

d'où on tire $\mathrm{pgcd}(\mathfrak{a},\mathfrak{b})\cdot\mathrm{ppcm}(\mathfrak{a},\mathfrak{b}) = \mathfrak{ab}$. À titre d'exercice, démontrer les deux lemmes suivants :

Lemme 4.5. — *Si* $\mathfrak{a},\mathfrak{b}$ *sont des idéaux fractionnaires,* $\mathfrak{a} \mid \mathfrak{b} \Leftrightarrow \mathfrak{b} \subset \mathfrak{a}$.

Lemme 4.6. — *Si* $\mathfrak{a},\mathfrak{b}$ *sont des idéaux entiers et* $\mathfrak{a}+\mathfrak{b}$ *le plus petit idéal contenant* $\mathfrak{a}$ *et* $\mathfrak{b}$*, alors* $\mathfrak{a}+\mathfrak{b} = \mathrm{pgcd}(\mathfrak{a},\mathfrak{b})$.

On dit que $\mathfrak{a},\mathfrak{b}$ sont premiers entre eux si $\mathrm{pgcd}(\mathfrak{a},\mathfrak{b}) = \mathcal{O}_K$. Le lemme chinois s'adapte facilement :

Lemme 4.7. — *Si* $\mathfrak{a},\mathfrak{b}$ *sont entiers et premiers entre eux, on a un isomorphisme d'anneaux* $\mathcal{O}_K/\mathfrak{ab} \simeq \mathcal{O}_K/\mathfrak{a} \times \mathcal{O}_K/\mathfrak{b}$.

Démonstration. — Comme $\mathfrak{ab} \subset \mathfrak{a}\cap\mathfrak{b}$ l'application $\mathcal{O}_K/\mathfrak{ab} \to \mathcal{O}_K/\mathfrak{a}\times \mathcal{O}_K/\mathfrak{b}$ est bien définie. Si $x \in \mathcal{O}_K$ est tel que $\mathfrak{a} \mid (x)$ et $\mathfrak{b} \mid (x)$, alors $\mathfrak{ab} \mid (x)$ par coprimalité et unicité de la décomposition en produit d'idéaux maximaux ; ceci assure l'injectivité. Pour la surjectivité, il existe $a \in \mathfrak{a}$, $b \in \mathfrak{b}$, tels que $a + b = 1$ (car $\mathfrak{a}+\mathfrak{b} = \mathcal{O}_K$) ; donc si $(\alpha,\beta) \in \mathcal{O}_K/\mathfrak{a}\times\mathcal{O}_K/\mathfrak{b}$, alors $b\alpha + a\beta \in \mathcal{O}_K/\mathfrak{ab}$ est un antécédent. □

En particulier, les cardinaux des deux membres sont les mêmes. Plus généralement

Proposition 4.8. — *Si* $\mathfrak{a},\mathfrak{b}$ *sont deux idéaux entiers, alors on a* $\mathrm{N}\,\mathfrak{ab} = \mathrm{N}\,\mathfrak{a}\cdot\mathrm{N}\,\mathfrak{b}$.

Démonstration. — En itérant le lemme chinois, il reste à montrer que $\mathrm{N}\,\mathfrak{p}^{k+1} = \mathrm{N}\,\mathfrak{p}\,\mathrm{N}\,\mathfrak{p}^k$ pour $\mathfrak{p}$ maximal. Soit $a \in \mathfrak{p}^{k+1} \smallsetminus \mathfrak{p}^k$; la multiplication par a induit un morphisme de groupe additifs de $\mathcal{O}_K/\mathfrak{p} \to \mathfrak{p}^k/\mathfrak{p}^{k+1}$. Si x est dans le noyau, $\mathfrak{p}^{k+1} \mid (xa)$ soit $\mathfrak{p} \mid (x)$; l'injectivité suit. Pour la surjectivité, soit y un représentant de $\overline{y} \in \mathfrak{p}^k/\mathfrak{p}^{k+1}$; comme $y \in \mathfrak{p}^{k+1} + (a) = \mathfrak{p}^k$, la congruence $ax \equiv y \pmod{\mathfrak{p}^{k+1}}$ a une solution $x \in \mathcal{O}_K$. □

On prolonge la norme aux idéaux fractionnaires : $\mathrm{N}(\mathfrak{a}/a) := \mathrm{N}\,\mathfrak{a}/\left|\mathrm{N}(a)\right|$. Elle reste bien entendu multiplicative.

Proposition 4.9. — *Le nombre d'idéaux entiers de norme* $\leqslant C$ *est fini.*

Démonstration. — La norme étant un entier positif, il suffit de démontrer qu'il existe un nombre fini d'idéaux entiers $\mathfrak{a}$ de norme donnée

$N = \#(\mathcal{O}_K/\mathfrak{a})$. Le théorème de Lagrange donne $N \in \mathfrak{a}$ soit $\mathfrak{a} \mid N\mathcal{O}_K$. Or $N\mathcal{O}_K = \prod_{\mathfrak{p}} \mathfrak{p}^{e_\mathfrak{p}}$ a exactement $\prod_{\mathfrak{p}}(e_\mathfrak{p} + 1) < +\infty$ diviseurs. □

4.5. Décomposition des premiers. — Un cas particulier important du théorème fondamental nous dit que chaque nombre premier p se décompose sous la forme

$$p\mathcal{O}_K = \prod_{\mathfrak{p}|p} \mathfrak{p}^{e(\mathfrak{p}/p)}, \tag{7}$$

et on écrit $\mathrm{N}\,\mathfrak{p} = |\mathcal{O}_K/\mathfrak{p}| = p^{f(\mathfrak{p}/p)}$. On dira que $\mathfrak{p}$ est *au-dessus* de p. L'exposant $e(\mathfrak{p}/p)$ est le *degré de ramification* de $\mathfrak{p}/p$. L'exposant $f(\mathfrak{p}/p)$ est le *degré résiduel* de $\mathfrak{p}/p$. Déterminer la décomposition des nombres premiers est la clé de la factorisation effective d'un idéal fractionnaire $\mathfrak{a}$: en effet, factorisant séparément numérateur et dénominateur, on peut supposer $\mathfrak{a}$ entier. Comme on vient de le voir $\mathfrak{a} \mid (\mathrm{N}\,\mathfrak{a})\mathcal{O}_K$, il suffit donc de factoriser l'entier naturel $\mathrm{N}\,\mathfrak{a}$ et de calculer $v_\mathfrak{p}(\mathfrak{a})$ pour chaque $\mathfrak{p} \mid p \mid \mathrm{N}\,\mathfrak{a}$.

***Théorème 4.10* (Kummer).** — *Soit* $K = \mathbb{Q}(X)/(T)$, $T \in \mathbb{Z}[X]$ *unitaire et* $\theta = X \pmod{T}$. On suppose que $p \nmid [\mathcal{O}_K : \mathbb{Z}[\theta]]$. *Alors « la factorisation de* T *mod* p *reflète celle de* $p\mathcal{O}_K$ *». Plus précisément, supposons*

$$T \equiv \prod_i P_i^{e_i} \pmod{p\mathbb{Z}[X]},$$

où les P_i *sont unitaires, irréductibles et* 2 *à* 2 *distincts modulo* p*. Alors*

$$p\mathcal{O}_K = \prod_i \mathfrak{p}_i^{e_i},$$

où les $\mathfrak{p}_i := p\mathcal{O}_K + P_i(\theta)\mathcal{O}_K$ *sont maximaux,* 2 *à* 2 *distincts et de degré résiduel* $\deg P_i$.

Le Théorème 15.5 explique comment tester l'hypothèse $p \nmid [\mathcal{O}_K : \mathbb{Z}[\theta]]$, vraie pour tout p sauf un nombre fini. En particulier elle est vérifiée quand $\overline{T}$ est sans facteurs carrés (les e_i sont tous égaux à 1). Le Théorème 15.7 nous dira quoi faire quand elle n'est pas satisfaite.

On dit que p est *non ramifié* si $e(\mathfrak{p}/p) = 1$ pour tout $\mathfrak{p} \mid p\mathcal{O}_K$. La discussion qui précède a montré que c'est le cas générique : un nombre fini de p sont ramifiés. On s'intéressera aux valeurs possibles pour $f(\mathfrak{p}/p)$ au § 7.2. Pour l'instant, notons que la multiplicativité de la norme, appliquée à (7), implique l'identité

$$\sum_{\mathfrak{p}|p} e(\mathfrak{p}/p) \cdot f(\mathfrak{p}/p) = n. \tag{8}$$

4.6. Le morphisme de Frobenius

Proposition 4.11. — *Si $K/\mathbb{Q}$ est galoisienne de groupe G, les idéaux maximaux $\mathfrak{p}$ divisant un même premier p sont permutés transitivement par G.*

Démonstration. — Supposons que $\mathfrak{p}$ et $\mathfrak{q}$ sont deux maximaux divisant un même premier p, appartenant à deux orbites distinctes, $o(\mathfrak{p}) \neq o(\mathfrak{q})$. Soit $x \in \mathfrak{p} \smallsetminus o(\mathfrak{q})$ (existe par le lemme chinois). Puisque $x \in \mathfrak{p}$, $p \mid \mathrm{N}\,\mathfrak{p} \mid \mathrm{N}(x)$; d'autre part, $\mathrm{N}(x) = \prod_{\sigma\in G} \sigma(x)$. Comme $\mathfrak{q} \mid p\mathcal{O}_K$, $\mathfrak{q} \mid \mathrm{N}(x)\mathcal{O}_K$, soit $\prod \sigma(x) \in \mathfrak{q}$. L'idéal $\mathfrak{q}$ étant premier, il existe $\sigma \in G$ tel que $\sigma(x) \in \mathfrak{q}$, soit $x \in \sigma^{-1}\mathfrak{q}$. Contradiction. □

Si $K/\mathbb{Q}$ est galoisienne, on en déduit que $e = e(\mathfrak{p}/p)$ et $f = f(\mathfrak{p}/p)$ dépendent uniquement de p, pas de $\mathfrak{p}$, et l'équation (8) s'écrit $ef\#\{\mathfrak{p} \mid p\} = n$. On note

$$D(\mathfrak{p}) = \{\sigma \in G \colon \sigma\mathfrak{p} = \mathfrak{p}\},$$

donc $D(\sigma\mathfrak{p}) = \sigma D(\mathfrak{p})\sigma^{-1}$ pour tout $\sigma \in G$. De $|G| = n$ et de la Proposition 4.11, on déduit $\#D(\mathfrak{p}) = ef$. Le groupe G, donc aussi $D(\mathfrak{p})$, stabilise $\mathcal{O}_K$; on montre que la projection canonique $\mathcal{O}_K \to \mathcal{O}_K/\mathfrak{p}$ induit un morphisme *surjectif* de $D(\mathfrak{p})$ dans $\mathrm{Aut}(\mathcal{O}_K/\mathfrak{p})$, qui est un groupe de cardinal f, engendré par le morphisme de Frobenius $x \mapsto x^p$. Le noyau $I(\mathfrak{p})$ est donc de cardinal e ; en particulier, p est non ramifié si et seulement si $I(\mathfrak{p}) = \{\mathrm{Id}\}$.

Dans ce cas, on note $\mathrm{Frob}_{\mathfrak{p}}$ l'unique $\sigma \in D(\mathfrak{p}) \subset G$ se réduisant sur le Frobenius $x \mapsto x^p$. Le *symbole de Frobenius* $\mathrm{Frob}_p \subset G$ désigne l'ensemble des $\mathrm{Frob}_{\mathfrak{p}}$ pour $\mathfrak{p} \mid p$. De même que $D(\sigma\mathfrak{p}) = \sigma D(\mathfrak{p})\sigma^{-1}$, on vérifie que $\mathrm{Frob}_{\sigma\mathfrak{p}} = \sigma\,\mathrm{Frob}_{\mathfrak{p}}\,\sigma^{-1}$, pour $\sigma \in G$. Donc Frob_p est une classe de conjugaison de G.

En particulier, si G est abélien, les classes de conjugaison ont un unique élément et $p \mapsto \mathrm{Frob}_p \in G$ associe un automorphisme de K à tout premier non ramifié. Ainsi, si $K = \mathbb{Q}(\zeta_n)$, $G \simeq (\mathbb{Z}/n\mathbb{Z})^*$ est abélien et Frob_p est donné par la classe de p dans $(\mathbb{Z}/n\mathbb{Z})^*$. D'après le théorème de la progression arithmétique, tout élément de G est un Frobenius. C'est un phénomène général sur lequel on s'étendra au § 7.2.

4.7. Valeurs absolues et places. — On appelle *valeur absolue* de K un morphisme $|\cdot| : K^* \to \mathbb{R}_+^*$, étendu à K par $|0| = 0$, et satisfaisant une inégalité triangulaire faible : il existe $C \geqslant 1$ tel que $|x+y| \leqslant C(|x|+|y|)$ pour tout $x, y \in K$. Une valeur absolue définit une métrique sur K, *via* $d(x,y) = |x-y|$, donc une topologie. Deux valeurs absolues sont *équivalentes* si elles définissent la même topologie.

Théorème 4.12. — *Deux valeurs absolues* $|\cdot|_1$ *et* $|\cdot|_2$ *sont équivalentes si et seulement s'il existe* $t > 0$ *tel que* $|\cdot|_1 = |\cdot|_2^t$.

Par exemple, si $\mathfrak{p}$ est maximal, $|x|_{\mathfrak{p}} := \mathrm{N}\,\mathfrak{p}^{-v_{\mathfrak{p}}(x)}$ est une valeur absolue. Plus généralement, $c^{v_{\mathfrak{p}}(x)}$ serait une valeur absolue (équivalente à celle-ci) pour $c < 1$. L'inégalité triangulaire est ici renforcée, puisque (6) donne
$$|x+y|_{\mathfrak{p}} \leqslant \max(|x|_{\mathfrak{p}}, |y|_{\mathfrak{p}});$$
on dit que $|\cdot|_{\mathfrak{p}}$ est *non archimédienne*, par référence à l'axiome d'Archimède sur l'ordre de $\mathbb{R}$. En effet ce dernier énonce que si $x, y \in \mathbb{R}^*$, il existe $n \in \mathbb{Z}$, tel que $|nx| > |y|$. Or $|z|_{\mathfrak{p}} \leqslant 1$ pour tout $z \in \mathcal{O}_K \supset \mathbb{Z}$, soit $|nx|_{\mathfrak{p}} \leqslant |x|_{\mathfrak{p}}$ pour tout $x \in K$, $n \in \mathbb{Z}$.

Une autre classe d'exemples provient des plongements de K. Si $\sigma : K \hookrightarrow \mathbb{C}$ est l'un des r_1+2r_2 plongements de K dans $\mathbb{C}$, $|x| = |\sigma(x)|^t$ est une valeur absolue pour tout $t > 0$. Celles-ci sont *archimédiennes* (vérifient l'axiome d'Archimède). Les valeurs absolues associées à deux plongements complexes conjugués sont équivalentes. On choisit la normalisation suivante
$$|x|_\sigma := \begin{cases} |\sigma(x)| & \text{si } \sigma \text{ est réel,} \\ |\sigma(x)|^2 & \text{si } \sigma \text{ est complexe.} \end{cases}$$
Noter que $|\cdot|_\sigma$ pour σ complexe ne vérifie pas l'inégalité triangulaire. (C'est l'unique raison pour laquelle on a affaibli celle-ci.)

Un dernier exemple parfaitement inintéressant est fourni par la valeur absolue triviale : $|x| = 1$ pour $x \in K^*$. On appelle *place* de K une classe d'équivalence de valeurs absolues non triviales.

***Théorème 4.13* (Ostrowski).** — *Les $|\cdot|_{\sigma_i}$, pour $i \leqslant r_1 + r_2$ et les $|\cdot|_{\mathfrak{p}}$, pour $\mathfrak{p}$ maximal sont un système de représentants des places de K.*

En contemplant (4) on obtient alors la *formule du produit* : pour tout $x \in K^*$, on a $\prod_v |x|_v = 1$, où v parcourt l'ensemble des représentants normalisés des places de K.

5. Géométrie des nombres

5.1. Tailles. — Une « taille » H sur un ensemble E associe à chaque $x \in E$ un réel positif $H(x)$ tel que $\{x \in E\colon H(x) < C\}$ soit fini pour tout C. On veut introduire des considérations de taille sur K. Si la norme convient pour les idéaux entiers (Proposition 4.9), l'exemple des $(2-\sqrt{3})^n$, $n \in \mathbb{N}$, tous distincts de norme 1 montre qu'il n'en est pas de même sur K. Le même exemple montre que le module complexe, en considérant $K \subset \mathbb{C}$, ne suffit pas non plus.

Définition 5.1. — La *hauteur* naïve $H(\alpha)$ de $\alpha \in K$ est définie par

$$H(\alpha) := \prod_v \max\{1, |\alpha|_v\},$$

où v parcourt l'ensemble des places de K.

On montre que $H(\alpha)$ ne dépend pas du corps de nombres K contenant $\mathbb{Q}(\alpha)$. Elle a aussi une traduction agréable quand $K = \mathbb{Q}$:

Proposition 5.2. — *Soit $a/b \in \mathbb{Q}$, fraction réduite au plus petit dénominateur. Alors $H(a/b) = \max(|a|, |b|)$.*

Démonstration. — Les places de $\mathbb{Q}$ sont les p premiers et la place archimédienne associée à la valeur absolue usuelle. Si $a = 0$, $b = \pm 1$ et le résultat est évident. Sinon, le produit sur les p premiers vaut $\prod_{p, v_p(a/b)<0} p^{-v_p(a/b)} = |b|$, que l'on multiplie par $\max(1, |a/b|)$. □

Proposition 5.3. — *Pour tout $C \geqslant 1$, le nombre de $x \in K$ tels que $H(x) < C$ est fini.*

Démonstration. — On peut écrire $x\mathcal{O}_K = \mathfrak{a}/\mathfrak{b}$ où $\mathfrak{a}, \mathfrak{b}$ sont entiers premiers entre eux. Le produit sur les v non archimédiennes et archimédiennes sont tous deux bornés par C. Le premier montre que $\mathrm{N}\,\mathfrak{b} < C$ (nombre fini de possibilités), le deuxième que $y = \mathrm{N}\,\mathfrak{b}x$, qui appartient à $\mathcal{O}_K$, a tous ses plongements bornés par C^2 ; les coefficients de $P_{\mathrm{char},y} \in \mathbb{Z}[X]$ sont donc bornés. On a donc un nombre fini de possibilités pour $P_{\mathrm{char},y}$, donc pour y. □

La fonction H est la « bonne » notion théorique ; cependant, d'un point de vue calculatoire, elle est peu commode. On préférerait une notion de type norme infinie, par exemple $H_\infty(x) := \max |x|_v$ où v parcourt seulement l'ensemble des places *archimédiennes*. On ne contrôle plus les dénominateurs, mais $H_\infty(x)$ ne dépend toujours pas de K et peut jouer le rôle de taille sur $\mathcal{O}_K$. D'un strict point de vue algorithmique, on préfère une norme L^2 pour des raisons qui apparaîtront aux § 15.6 et § 15.7, l'algorithme LLL en particulier. On considère donc $K \otimes_{\mathbb{Q}} \mathbb{R} \simeq \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$ comme un espace euclidien muni de la forme naturelle

$$T_2 : (x_1, \dots, x_{r_1}, z_1, \dots, z_{r_2}) \longmapsto \sum_{i \leqslant r_1} x_i^2 + 2 \sum_{j \leqslant r_2} |z_j|^2 ,$$

cf. (5) ; noter que si $r_2 = 0$, T_2 est la forme trace du Théorème 2.2. Explicitement, pour $x \in K$,

$$T_2(x) = \sum_{k=1}^{n} |\sigma_k(x)|^2 .$$

Proposition 5.4. — *Pour $C > 0$, le nombre de $x \in \mathcal{O}_K$ tels que $T_2(x) < C$ est fini.*

Démonstration. — $\mathcal{O}_K$ est discret (Corollaire 3.4) et la boule $T_2(x) \leqslant C$ est compacte. □

***Théorème 5.5* (Kronecker).** — *Soit $x \in \mathcal{O}_K$, $x \neq 0$. Alors $T_2(x) \geqslant n$ avec égalité si et seulement si x est une racine de l'unité.*

Démonstration. — L'inégalité est celle de la moyenne arithmético-géométrique : $|\mathrm{N}(x)|^{2/n} \leqslant \frac{1}{n} T_2(x)$, qui implique $T_2(x) \geqslant n$ puisque $|\mathrm{N}(x)| \geqslant 1$. On a égalité si et seulement si les $|\sigma_k(x)|^2$ sont tous égaux, nécessairement à 1 puisque leur somme est n. L'ensemble $\{x^k, k > 0\} \subset \mathcal{O}_K$ est borné pour T_2, donc fini d'après la Proposition. Donc x est une racine de l'unité. □

5.2. Discriminant. — On utilise les notations du § 1.2. Par extension des scalaires, $E = (K \otimes_{\mathbb{Q}} \mathbb{R}, T_2)$ est un espace euclidien. Si Λ est un sous-$\mathbb{Z}$-module libre de rang n de K, on peut le considérer comme réseau de E. On a défini au § 1.2 son déterminant $d(\Lambda)$ et son discriminant $\Delta_\Lambda = d(\Lambda)^2$.

Définition 5.6. — On note $\Delta_K = \Delta_{\mathcal{O}_K}$ le discriminant de $\mathcal{O}_K$.

Ceci définit le discriminant comme carré d'un volume, on aurait pu le définir algébriquement grâce aux identités suivantes : si $(e_1, \dots, e_n)$ est une $\mathbb{Z}$-base de Λ, $S = (\sigma_j(e_i))_{i,j \leqslant n}$, et $T = (\mathrm{Tr}(e_i e_j))_{i,j \leqslant n}$ la matrice de la forme trace, on a

$$\mathrm{Gram}(e_1, \dots, e_n) = S \cdot \overline{{}^tS}, \quad T = \Big(\sum_k \sigma_k(e_i)\sigma_k(e_j)\Big)_{i,j \leqslant n} = S \cdot {}^tS.$$

En particulier, $\Delta_K = |\det T|$. Comme T est à coefficients entiers, $\Delta_K \in \mathbb{N}$. Plus généralement :

Proposition 5.7. — *Soit $\mathcal{O}$ un sous-$\mathbb{Z}$-module d'indice fini de $\mathcal{O}_K$, et $\Delta_{\mathcal{O}}$ son discriminant. Alors $\Delta_{\mathcal{O}} \in \mathbb{N}$ et $\Delta_{\mathcal{O}} = \Delta_K [\mathcal{O}_K : \mathcal{O}]^2$.*

Démonstration. — Résulte de ce qui précède et du Corollaire 1.6. □

5.3. Applications du théorème de Minkowski. — Notre premier corollaire dit simplement qu'un idéal entier contient un élément non nul de norme proche du minimum :

Corollaire 5.8. — *Si K est un corps de nombres de degré n et signature (r_1, r_2), alors tout idéal entier $\mathfrak{a}$ contient un $a \neq 0$ tel que*

$$|\mathrm{N}(a)| \leqslant c_K \cdot \mathrm{N}\,\mathfrak{a}, \quad \textit{où } c_K := \Big(\frac{4}{\pi}\Big)^{r_2} \frac{n!}{n^n} \sqrt{\Delta_K}.$$

Noter que $a \in \mathfrak{a}$ implique $\mathrm{N}\,\mathfrak{a} \mid \mathrm{N}(a)$.

Démonstration. — On plonge K dans $K \otimes_{\mathbb{Q}} \mathbb{R} \simeq \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} \simeq \mathbb{R}^n$ par

$$x \longmapsto \big(\sigma_1(x), \dots, \sigma_{r_1}(x), \mathrm{Re}\,\sigma_{r_1+1}(x), \mathrm{Im}\,\sigma_{r_1+1}(x), \dots, \\ \mathrm{Re}\,\sigma_{r_1+r_2}(x), \mathrm{Im}\,\sigma_{r_1+r_2}(x)\big),$$

muni de la forme euclidienne standard. On en considère $\mathfrak{a}$ et $\mathcal{O}_K$ comme des réseaux, qui vérifient $d(\mathfrak{a}) = d(\mathcal{O}_K)\,\mathrm{N}\,\mathfrak{a}$. Par rapport à la structure euclidienne utilisée pour définir $\sqrt{\Delta_K} = d(\mathcal{O}_K, T_2)$, on a remplacé $T_2(x) = \sum_{i=1}^{r_1+2r_2} |\sigma_i(x)|^2$ par $\sum_{i=1}^{r_1+r_2} |\sigma_i(x)|^2$, soit $d(\mathcal{O}_K) = 2^{-r_2}\sqrt{\Delta_K}$.

On pose

$$C_t := \Big\{(x_1, \dots, x_{r_1}, z_1, \dots, z_{r_2}) \in \mathbb{R}^{r_1} \times \mathbb{C}^{r_2} : \sum_{i \leqslant r_1} |x_i| + 2 \sum_{j \leqslant r_2} |z_j|^2 \leqslant t\Big\},$$

de volume $V(C_t) = 2^{r_1}(\frac{\pi}{2})^{r_2} t^n/n!$. Il existe un point non nul de $\mathfrak{a}$ dans C_t si $V(C_t) \geqslant 2^n d(\mathfrak{a})$: on peut prendre $t^n = (\frac{4}{\pi})^{r_2} n! \sqrt{\Delta_K}\, \mathrm{N}\,\mathfrak{a}$, en utilisant $r_1 + 2r_2 = n$. D'autre part, par l'inégalité de la moyenne arithmético-géométrique, un point de C_t vérifie

$$\prod_i |x_i| \prod_j |z_j|^2 \leqslant \left(\frac{t}{n}\right)^n. \qquad \square$$

Corollaire 5.9. — *Tout idéal fractionnaire $\mathfrak{A}$ de K s'écrit $\mathfrak{A} = (\alpha)\mathfrak{a}$, où $\alpha \in k^*$ et $\mathfrak{a}$ est un idéal entier tel que $\mathrm{N}\,\mathfrak{a} \leqslant c_K$.*

Démonstration. — Soit $\mathfrak{B} := \mathfrak{A}^{-1} = \frac{1}{d}\mathfrak{b}$, où $\mathfrak{b}$ est entier. Donc $\mathfrak{b}$ contient $b \neq 0$ tel que $|\mathrm{N}(b)| \leqslant c_K \,\mathrm{N}\,\mathfrak{b}$; l'idéal $(b)\mathfrak{b}^{-1}$ est entier, de norme $\leqslant c_K$. On pose $\alpha = d/b$. $\square$

Appliquant le Corollaire 5.8 à $\mathfrak{a} = \mathcal{O}_K$, on remarque que le membre de gauche est un entier naturel $\geqslant 1$. Une rapide étude de fonction donne :

Corollaire 5.10. — *Si $n > 1$, i.e. $K \neq \mathbb{Q}$, alors $\Delta_K > 1$. Si $n \to +\infty$, alors $\Delta_K \to +\infty$.*

Les mêmes techniques permettent de prouver :

***Théorème 5.11* (Hermite).** — *Soit $X > 0$. Les corps de nombres K tels que $\Delta_K < X$ sont en nombre fini.*

Démonstration (idée). — Grâce au dernier corollaire, on peut supposer K de degré n fixé. Supposons d'abord $r_1 > 0$. Par Minkowski, on construit $\alpha \in \mathcal{O}_K$ tel que tous les $\sigma_i(\alpha)$ sont de module < 1, à l'exception d'un d'entre eux, de module borné en fonction de Δ_K. D'après la Proposition 5.4, il existe un nombre fini de tels α. Par construction, la racine distinguée de $P_{\mathrm{char},\alpha}$ est simple donc $K = \mathbb{Q}(\alpha)$ (Théorème 2.1). Si $r_1 = 0$, on procède de même avec deux racines distinguées de module > 1, complexes conjuguées et toutes deux simples. $\square$

Si $N_n(X)$ désigne le nombre de corps de nombres K de degré $n > 1$ satisfaisant $\Delta_K \leqslant X$, on pense que $N_n(X) \sim \alpha_n X$ quand $X \to \infty$, pour un certain $\alpha_n > 0$. Le meilleur résultat connu dans cette direction, obtenu par Ellenberg et Venkatesh [**15**], dit que pour tout $\varepsilon > 0$, on a

$$\limsup_{X\to+\infty} \frac{\log N_n(X)}{\log X} < n^{\varepsilon}, \quad \liminf_{X\to+\infty} \frac{\log N_n(X)}{\log X} \geqslant \frac{1}{2} + \frac{1}{n^2},$$

6. Groupe des classes, unités

Revenons à nos préoccupations diophantiennes ; on a une bonne théorie de la divisibilité pour les idéaux. Si $\mathcal{O}_K$ est principal, et que l'on maîtrise ses unités $U(K)$, on peut revenir des idéaux aux éléments de K.

6.1. Groupe des classes. — D'après le théorème fondamental, l'ensemble $I(K)$ des idéaux fractionnaires de K est un groupe abélien. Le *groupe des classes d'idéaux* de K, noté $\mathrm{Cl}(K)$, est le quotient de $I(K)$ par le sous groupe $P(K)$ des idéaux fractionnaires principaux.

Théorème 6.1. — *Le groupe* $\mathrm{Cl}(K)$ *est fini.*

Démonstration. — On applique le Corollaire 5.9. Les idéaux entiers de norme bornée étant en nombre fini (Proposition 4.9), il en est de même du nombre de classes d'idéaux. □

Le cardinal de $\mathrm{Cl}(K)$, noté $h(K)$, est le *nombre de classes* de K.

Le groupe des classes mesure une obstruction (à la principalité de $\mathcal{O}_K$) ; idéalement, il est trivial. C'est expérimentalement relativement fréquent, mais la conjecture naturelle est toujours ouverte :

***Conjecture 6.2* (Gauss-Hasse).** — *Il existe une infinité de corps de nombres* $K \subset \mathbb{C}$ *tels que* $h(K) = 1$.

Pour toute signature (r_1, r_2) différente de $(1, 0)$ $(K = \mathbb{Q})$ et $(0, 1)$ (K quadratique imaginaire), on pense qu'il existe une infinité de K partageant cette signature tels que $h(K) = 1$. Il est nécessaire d'exclure $(0, 1)$:

Théorème 6.3 **(Heegner, Stark, Baker (indépendamment))**
Un corps quadratique imaginaire $K = \mathbb{Q}(\sqrt{-\Delta_K})$ vérifie $h(K) = 1$ si et seulement si

$$\Delta_K \in \{3, 4, 7, 8, 11, 19, 43, 67, 163\}$$

C'est un théorème difficile (qui serait une conséquence facile de l'hypothèse de Riemann, voir le Théorème de Brauer-Siegel 7.6) ; par contre, il est relativement simple de montrer que la liste est complète à une exception près.

6.2. Unités, S-unités, et régulateurs. — Les *unités* $U(K) := \mathcal{O}_K^*$ de $\mathcal{O}_K$ forment un $\mathbb{Z}$-module de type fini dont la structure est donnée par le théorème de Dirichlet :

$$U(K) = (\mathbb{Z}/w\mathbb{Z}) \cdot \zeta \oplus \bigoplus_{i=1}^{r_1+r_2-1} \mathbb{Z} \cdot \eta_i.$$

En d'autre termes chaque unité s'écrit $\zeta^{a_0} \prod \eta_i^{a_i}$ où $a_0 \in \mathbb{Z}/w\mathbb{Z}$ et les $a_i \in \mathbb{Z}$ sont uniquement déterminés. Plus généralement, soit S_∞ l'ensemble des $r_1 + r_2$ places archimédiennes et S un ensemble fini de places contenant S_∞. On définit

$$U_S(K) = \left\{ x \in K : |x|_{\mathfrak{p}} = 1, \forall\, \mathfrak{p} \notin S \right\},$$

le groupe des S-*unités* de K. En particulier $U_{S_\infty}(K) = U(K)$.

Théorème 6.4 **(Dirichlet, généralisé par Chevalley et Hasse)**
$U_S(K)$ est somme directe du groupe cyclique $\mu(K)$ des racines de l'unité de K et d'un $\mathbb{Z}$-module libre de rang $|S| - 1$.

Démonstration (vague idée). — Comme le corps cyclotomique $\mathbb{Q}(\zeta_n)$ est de degré $\phi(n) \to +\infty$ quand $n \to +\infty$, $\mu(K)$ est fini. Il est donc cyclique (comme sous-groupe fini du groupe multiplicatif d'un corps). On considère $\Phi : U_S(K) \to \mathbb{R}^S$, qui à x associe $(\log |x|_v)_{v \in S}$. C'est un morphisme de groupe dont l'image est incluse dans l'hyperplan $\sum_v x_v = 0$ (formule du produit). Le Théorème 5.5 donne $\operatorname{Ker} \Phi = \mu(K)$. On montre ensuite que l'image de Φ, isomorphe à $U_S(K)/\mu(K)$, est un réseau de cet hyperplan (c'est assez long). $\square$

Soit $(\eta_i)_{1 \leqslant i < |S|}$ une base de $U_S(K)/\mu(K)$. Le *S-régulateur* $R_S(K)$ de K est la valeur absolue du déterminant de la matrice

$(\log |\eta_i|_v)_{v \in S \smallsetminus \{v_0\}}$, où $v_0 \in S$ est choisie arbitrairement, et le déterminant de la matrice vide est 1 par définition si S ne contient qu'un élément. (Exercice : la définition ne dépend pas du choix de v_0.) On pose $R(K) := R_{S_\infty}(K)$.

6.3. Heuristiques. — On a très peu de résultats généraux sur $\mathrm{Cl}(K)$, mais on dispose d'un modèle probabiliste convaincant, dû à Cohen-Lenstra [**11**] (corps $K/\mathbb{Q}$ quadratiques) et Cohen-Martinet [**12**] (extensions K/k arbitraires). Voici le cas le plus simple :

Conjecture 6.5. — *Si f est une fonction de l'ensemble des classes d'isomorphismes de groupes abéliens finis dans $\mathbb{R}^+$, si G parcourt l'ensemble des classes d'isomorphismes de groupes abéliens finis, et K les classes d'isomorphismes de corps quadratiques imaginaires, on définit*

$$\mathbb{E}_0(f) := \lim_{X \to +\infty} \frac{\sum_{K, \Delta_K < X} f(\mathrm{Cl}(K))}{\sum_{K, \Delta_K < X} 1} \leqslant +\infty,$$

$$\mathbb{E}'_0(f) := \lim_{X \to +\infty} \frac{\sum_{G, |G| < X} \frac{1}{|\mathrm{Aut}\, G|} f(G)}{\sum_{G, |G| < X} \frac{1}{|\mathrm{Aut}\, G|} 1} \leqslant +\infty.$$

Alors, si f est « raisonnable »,

$$\mathbb{E}_0(f) = \mathbb{E}'_0(f).$$

Intuitivement, le modèle dit qu'en moyenne, le groupe des classes d'un corps quadratique imaginaire est le groupe (abélien, fini) G avec probabilité $1/|\mathrm{Aut}\, G|$. C'est un poids omniprésent en arithmétique, en particulier dans les « formules de masse », qui a le mérite d'expliquer la prépondérance des groupes cycliques parmi les groupes de classes calculés en pratique. Une première généralisation du modèle assimile les groupes de classes d'extensions galoisiennes $K/\mathbb{Q}$ dont le rang des unités $\mathrm{rg}\, U(K)$ est r à un quotient $G/\langle \sigma_1, \ldots, \sigma_r \rangle$ (dans la conjecture citée, $r = 0$).

Les fonctions liées au 2-Sylow G_2 de G (plus généralement aux p-Sylow G_p pour p divisant l'ordre de $\mathrm{Gal}(K/\mathbb{Q})$) sont *a priori* déraisonnables : G_2 est soumis à des contraintes plus fortes et ne suit pas le modèle naïf. Par exemple si K est quadratique imaginaire,

$$r_2(\mathrm{Cl}(K)) = \omega(\Delta_K) - 1,$$

où $\omega(n)$ désigne le nombre de diviseurs premiers distincts de n (formule des genres, due à Gauss). Des fonctions f raisonnables sont par exemple, $f(G) = |G_p|$ (p impair), la fonction caractéristique de l'ensemble des G tels que G/G_2 soit cyclique, etc.

Pour les fonctions f arithmétiquement raisonnables, $\mathbb{E}'_0(f)$ existe et se calcule facilement. Dans les rares cas où $\mathbb{E}_0(f)$ est connue, elle coïncide avec la prédiction. Un exemple célèbre est $f(G) = 3^{r_3(G)}$ (théorème de Davenport-Heilbronn [**13**]), dont la valeur moyenne est connue pour les groupes de classes des corps quadratiques. L'accord avec le petit nombre de tables existantes est acceptable, mais la convergence apparente des densités est lente (une ou deux décimales correctes).

7. Théorie analytique des nombres

7.1. Fonctions L de Hecke. — On encode tout ce que nous venons de voir dans la fonction zêta de Dedekind de K :

$$\zeta_K(s) := \sum_{\mathfrak{a}} \mathrm{N}\,\mathfrak{a}^{-s},$$

où $\mathfrak{a}$ parcourt l'ensemble des idéaux entiers et $\mathrm{Re}(s) > 1$. (Exercice : vérifier la convergence.) Bien sûr, $\zeta_{\mathbb{Q}}(s) = \zeta(s)$ est la fonction zêta de Riemann.

Plus généralement, soit $\chi : I(K) \to \mathbb{C}^*$ un morphisme de groupes, trivial sur les idéaux principaux. Par passage au quotient, χ définit un caractère du groupe abélien fini $\mathrm{Cl}(K)$. On définit

$$L(s, \chi, K) = \sum_{\mathfrak{a}} \chi(\mathfrak{a})\,\mathrm{N}\,\mathfrak{a}^{-s},$$

où $\mathfrak{a}$ parcourt l'ensemble des idéaux entiers de K. En particulier, si χ_0 est le caractère trivial, $L(s, \chi_0, K) = \zeta_K(s)$.

Théorème 7.1

(1) *Si* $\mathrm{Re}(s) > 1$,

$$L(s, \chi, K) = \prod_{\mathfrak{p}} (1 - \chi(\mathfrak{p})\,\mathrm{N}\,\mathfrak{p}^{-s})^{-1}$$

admet un produit eulérien, où $\mathfrak{p}$ *parcourt l'ensemble des idéaux maximaux de* $\mathcal{O}_K$.

(2) $(s-1)L(s,\chi,K)$ *admet un prolongement holomorphe au plan complexe. On garde les notations* $L(s,\chi,K)$ *et* ζ_K *pour les fonctions méromorphes associées.*

(3) $\zeta_K(s)$ *a un pôle simple en* $s=1$ *de résidu*

$$r_K := 2^{r_1}(2\pi)^{r_2}\frac{h(K)R(K)}{w(K)\sqrt{\Delta_K}},$$

où $h(K)$, $R(K)$, $w(K)$ *sont respectivement le nombre de classes, le régulateur, et le nombre de racines de l'unité de* K.

(4) *Si* $\chi \neq \chi_0$, $L(s,\chi,K)$ *est une fonction entière.*

(5) *Soit* $A := 2^{-r_2}\pi^{-n/2}\Delta_K^{1/2}$ *et* $\gamma(s) := A^s\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}$. *La fonction*

$$\xi(s,\chi,K) := s(s-1)\gamma(s)\zeta_K(s),$$

est entière et vérifie l'équation fonctionnelle

$$\xi(s,\chi,K) = W(\chi)\cdot\xi(1-s,\overline{\chi},K),$$

où $W(\chi)$ *est un complexe de module* 1. *On a* $W(\chi_0)=1$.

Les fonctions L associées aux caractères de $\mathrm{Cl}(K)$ sont d'un genre très particulier, et ne généralisent pas les séries L de Dirichlet, associées à un caractère de $(\mathbb{Z}/N\mathbb{Z})^*$. Une généralisation possible est la suivante : on remplace $\mathrm{Cl}(K)$ par le *groupe des classes de rayon* $\mathfrak{f}$, $\mathrm{Cl}_\mathfrak{f} := I_\mathfrak{f}(K)/\,\mathrm{Cl}_\mathfrak{f}(K)$, où $\mathfrak{f}$ est un idéal entier, $I_\mathfrak{f}(K)$ est le sous-groupe de $I(K)$ formé des idéaux premiers à $\mathfrak{f}$, et $P_\mathfrak{f}(K)$ le sous-groupe des idéaux fractionnaires principaux (α), où $\alpha \equiv 1 \pmod{\mathfrak{f}}$ et $\sigma(\alpha) > 0$ pour toutes les plongements *réels* de K. Le groupe $\mathrm{Cl}_\mathfrak{f}(K)$ est fini, et c'est une extension de $\mathrm{Cl}(K)$ par

$$\big((\mathcal{O}_K/\mathfrak{f})^* \times \{-1,1\}^{r_1}\big)/U(K),$$

où $U(K) \to (\mathcal{O}_K/\mathfrak{f})^*$ est la projection canonique et $U(K) \to \{-1,1\}^{r_1}$ est le vecteur des signes des plongements réels. En particulier, si $K=\mathbb{Q}$, $\mathfrak{f}=q\mathbb{Z}$, $\mathrm{Cl}_\mathfrak{f}(K) = (\mathbb{Z}/q\mathbb{Z})^*$.

Étant donné un caractère de $\mathrm{Cl}_\mathfrak{f}$, on définit $L(s,\chi,K) = \sum_\mathfrak{a}\chi(\mathfrak{a})\,\mathrm{N}\,\mathfrak{a}^{-s}$, en posant $\chi(\mathfrak{a}) = 0$ si $\mathfrak{a}$ et $\mathfrak{f}$ ne sont pas premiers entre eux. Cette fonction L vérifie des propriétés analogues à celles du Théorème 7.1, avec une équation fonctionnelle plus compliquée.

Les groupes $\mathrm{Cl}_{\mathfrak{f}}(K)$ jouent un rôle important dans la théorie (classique) du *corps de classe*, qui étudie les extensions abéliennes de K, c'est-à-dire les extensions galoisiennes de groupe de Galois abélien. Toute extension abélienne de K se réalise comme sous-corps d'une extension $H_{\mathfrak{f}}(K)/K$ pour un certain idéal entier $\mathfrak{f}$, dont le groupe de Galois est $\mathrm{Cl}_{\mathfrak{f}}(K)$. Si $K = \mathbb{Q}$, $\mathfrak{f} = q\mathbb{Z}$, on a $H_{\mathfrak{f}}(K) = \mathbb{Q}(\zeta_q)$ et on retrouve le théorème de Kronecker-Weber : toute extension abélienne de $\mathbb{Q}$ est contenue dans un corps cyclotomique.

7.2. Densités d'idéaux premiers. — La fonction ζ de Riemann permet de montrer le théorème des nombres premiers (essentiellement équivalent à $\zeta(1+it) \neq 0$ pour $t \in \mathbb{R}$), les séries L le théorème de la progression arithmétique (essentiellement équivalent à $L(1,\chi) \neq 0$). Tous deux se généralisent en utilisant essentiellement les mêmes arguments analytique ($\zeta_K(1+it) \neq 0$ et $L(1,\chi,K) \neq 0$) :

***Théorème 7.2* (des idéaux premiers).** — *Soit K un corps de nombres et $\pi_K(X)$ le cardinal de*

$$\{\mathfrak{p} \subset \mathcal{O}_K : \mathfrak{p} \text{ maximal, } \mathrm{N}\,\mathfrak{p} \leqslant X\}.$$

Quand $X \to +\infty$, $\pi_K(X) \sim X/\log X$.

La démonstration est analogue à celle du théorème des nombres premiers. Le théorème de la progression arithmétique est plus difficile, et a joué un rôle historique important dans le développement de la théorie du corps de classe :

***Théorème 7.3* (Chebotarëv).** — *Soit $K/\mathbb{Q}$ une extension galoisienne de groupe G, et C une classe de conjugaison de G. On note $\pi_K(X,C)$ le nombre de premiers $p \leqslant X$ non ramifiés tels que* $\mathrm{Frob}_p \in C$. *Quand $X \to +\infty$,*

$$\pi_K(x,C) \sim \frac{|C|}{|G|} \cdot \pi_K(X).$$

Pour une démonstration, voir **[25]** ou **[36]** (ce dernier donne les grandes lignes d'une preuve n'utilisant pas la théorie du corps de classes).

Le théorème de Chebotarëv généralise le théorème de Dirichlet : $K = \mathbb{Q}(\zeta_q)$ est une extension galoisienne de $\mathbb{Q}$ de groupe $G \simeq (\mathbb{Z}/q\mathbb{Z})^*$, qui est abélien. Donc $|C| = 1$ et le théorème de Chebotarëv dit que les nombres premiers se répartissent de manière asymptotiquement équiprobable entre les différentes classes de $(\mathbb{Z}/q\mathbb{Z})^*$.

Mentionnons un dernier corollaire utile, reposant sur l'interprétation suivante : soit $f \in \mathbb{Z}[X]$ unitaire, irréductible de degré n dans $\mathbb{Q}[X]$ et K un corps de décomposition de f. L'extension $K/\mathbb{Q}$ est galoisienne de groupe de Galois $G \subset S_n$ (G permute les n racines de f dans K). Supposons que la réduction $\overline{f}$ dans $\mathbb{F}_p[X]$ est sans facteurs carrés, ce qui implique que p est non ramifié dans K. Alors la classe de conjugaison de Frob_p dans G détermine le type de décomposition de f modulo p (en fait, la classe de conjugaison dans S_n le détermine déjà). Plus précisément :

– si $\overline{f} = f_1 \cdots f_g$ dans $\mathbb{F}_p[X]$, où les f_i sont irréductibles, distincts, l'ensemble $T = \{\deg f_1, \ldots, \deg f_g\}$ est le *type de décomposition* de $\overline{f}$. (Par abus de langage, on écrira $\overline{f} \in T$.)

– si $\sigma \in S_n$, $\sigma = c_1 \cdots c_g$ où les c_i sont des cycles disjoints de S_n, la classe de conjugaison de σ est déterminée par $\{\lg(c_1), \ldots, \lg(c_g)\}$, où $\lg(c)$ désigne la longueur du cycle c.

Ces deux ensembles d'entiers sont identiques si $\sigma = \mathrm{Frob}_p$, sous l'hypothèse que les f_i sont distincts (exercice).

***Corollaire 7.4* (Frobenius, Hecke).** — *Soient f, n, K, G comme ci-dessus ; soit T un type de décomposition de f, aussi vu comme classe de conjugaison de S_n. Alors*

$$\lim_{x \to +\infty} \frac{\#\{p \leqslant x \colon \overline{f} \in T\}}{\#\{p \leqslant x\}} = \frac{|T \cap G|}{|G|}.$$

Le résultat de densité ci-dessus est dû à Hecke. Frobenius utilisait la densité analytique d'un ensemble $\mathcal{P}$,

$$\delta(\mathcal{P}) := \lim_{s \to 1^+} \frac{1}{\log(s-1)} \sum_{p \in \mathcal{P}} p^{-s},$$

qui est une notion plus faible : si la densité naturelle existe, la densité analytique aussi et elles sont égales. La réciproque est fausse.

7.3. L'hypothèse de Riemann. — L'*hypothèse de Riemann* (généralisée, GRH) dit que $L(s, \chi, K) \neq 0$ si $\mathrm{Re}(s) > 1/2$; elle n'est bien sûr connue pour aucun K ou χ. Elle a des conséquences remarquables sur les termes d'erreurs dans les théorèmes de densité du paragraphe précédent (voir [**24**]), qui sont obtenus à l'aide de formules intégrales du type

$$(9)\quad \sum_{\mathrm{N}\,\mathfrak{a}\leqslant x} \chi(\mathfrak{a})\Lambda(\mathfrak{a}) = \frac{1}{2i\pi}\int_{\operatorname{Re} s=c} -\frac{L'}{L}(s,\chi,K)\,x^{-s}\,\frac{ds}{s}, \quad (x>0,\ x\notin\mathbb{N}),$$

où $c > 1$ et $\Lambda(\mathfrak{a}) = \log \mathrm{N}\,\mathfrak{p}$ si $\mathfrak{a} = \mathfrak{p}^k$ est une puissance d'un idéal maximal, et 0 sinon. Il « suffit » ensuite de déplacer un contour : le terme x^{-s} est un $O(x^{-\operatorname{Re} s})$ et l'intégrale devient négligeable par rapport aux résidus que l'on récupère en déplaçant la droite d'intégration vers la gauche. Il est clair qu'une bonne localisation des singularités de l'intégrande dans le demi-plan $\operatorname{Re} s > 0$ ($s = 1$ si $\chi = \chi_0$ et les zéros de L) permet d'estimer plus finement les résidus.

En particulier, si GRH est vraie, il existe $p = O(\log \Delta_K)^2$ satisfaisant $\mathrm{Frob}_p \in C$ dans le théorème de Chebotarëv, où la constante implicite est effective (voir [**23**]). On peut donc trouver un tel p en un temps raisonnable, et en tout cas précisément borné, par une recherche exhaustive $p = 2, 3, 5, \ldots$.

Une conséquence du Corollaire 5.9 est que les (classes des) idéaux maximaux de $\mathcal{O}_K$ de norme $\leqslant c_K = O(\sqrt{\Delta_K})$ engendrent $\mathrm{Cl}(K)$. GRH permet beaucoup mieux :

***Théorème 7.5* (Bach [4]).** — *Sous GRH, les idéaux maximaux de K de norme inférieure à* $12(\log \Delta_K)^2$ *engendrent* $\mathrm{Cl}(K)$.

Démonstration (vague idée). — Supposons que les idéaux de norme $< x$ engendrent un sous-groupe *strict* H de $\mathrm{Cl}(K)$. Alors il existe un caractère $\chi \neq \chi_0$ de $\mathrm{Cl}(K)$ tel que $\chi_{|H} = \chi_{0|H}$: il suffit de relever un caractère non trivial de $\mathrm{Cl}(K)/H \neq \{1\}$. Au vu de (9),

$$\int_{\operatorname{Re} s=c} \frac{L'}{L}(s,\chi,K)\,x^{-s}\,\frac{ds}{s} = \int_{\operatorname{Re} s=c} \frac{L'}{L}(s,\chi_0,K)\,x^{-s}\,\frac{ds}{s}.$$

D'après le Théorème 7.1, en $s = 1$, $L(s,\chi_0,K)$ a un pôle simple alors que $L(s,\chi,K)$ est régulière. Sous GRH, aucun des deux intégrandes n'a de pôle de partie réelle $\frac{1}{2} < \operatorname{Re}(s) < 1$, et on en tire une contradiction du type $O(x^{1/2}) = x + O(x^{1/2})$. Malheureusement, les constantes implicites dépendent de K !

Pour rendre ceci rigoureux, il faut donc étudier la dépendance en K des termes d'erreurs. En fait, on utilise une « Formule Explicite » [**25**, Chap. 17], provenant de la dérivée logarithmique du produit de Weierstrass de la fonction d'ordre 1

$$\xi(s,\chi,K) = e^{b_0+b_1 s}\prod_{\rho}\left(1-\frac{s}{\rho}\right)e^{s/\rho},$$

et qui généralise (9) en remplaçant la fonction caractéristique de $[0,x[$ par une fonction test f très générale. Elle énonce une égalité entre

$$\sum_{\mathfrak{a}}\chi(\mathfrak{a})\Lambda(\mathfrak{a})\,f(\mathrm{N}\,\mathfrak{a}).$$

et une somme sur les zéros de $L(s,\chi,K)$, modulo quelques termes parasites. L'essentiel du travail consiste à optimiser la fonction test. $\square$

Donnons une dernière application, liée aux estimations du nombre de classes :

***Théorème 7.6* (Brauer-Siegel).** — *Si K parcourt l'ensemble des corps de nombres de degré n fixé, on a, pour tout $\varepsilon > 0$,*

$$\Delta_K^{-\varepsilon} \ll_\varepsilon \frac{h(K)R(K)}{\sqrt{\Delta_K}} \ll_\varepsilon \Delta_K^{\varepsilon}.$$

En particulier, $\log\big(h(K)R(K)\big) \sim \log\sqrt{\Delta_K}$ *quand* $\Delta_K \to \infty$ *et que le degré de K reste fixe.*

La borne supérieure est élémentaire et effective ; inconditionnellement, la borne inférieure n'est *pas* effective, on ne peut donc pas démontrer le Théorème 6.3 ainsi. Mais elle le devient sous GRH ! Plus précisément, la constante dépend de Δ_{K_0}, où $K_0/\mathbb{Q}$ est une clôture galoisienne d'une extension de degré n telle qu'il existe $s_0 \in\,]\frac{1}{2},1[$ tel que $\zeta_{K_0}(s_0) = 0$. Sous GRH, K_0 n'existe pas.

La seule minoration effective de $h(K)$ actuellement connue inconditionnellement (Goldfeld-Gross-Zagier, voir [**31**]) concerne les corps quadratiques *imaginaires*, pour lesquels $R(K) = 1$, et elle est nettement inférieure à $\log\Delta_K$ (!) :

$$h(K) > \frac{\log\Delta_K}{1700}\prod_{p\mid\Delta_K}\left(1-\frac{4\sqrt{p}}{p-1}\right).$$

8. Cahier des charges

Notre tâche principale est de calculer les objets suivants, associés à un corps de nombres K, défini par le polynôme minimal d'un élément primitif T :

(1) plongements, signature,
(2) $\mathbb{Z}$-base de $\mathcal{O}_K$, factorisation des idéaux,
(3) $\mathrm{Cl}(K)$, $U(K)$.

On peut y rajouter de nombreux problèmes annexes, qui surgissent naturellement

(4) tables de corps de nombres ordonnés par discriminants,
(5) énumération des $x \in K$, $H(x) < C$,
(6) factorisation des polynômes de $K[X]$,
(7) calcul de $\mathrm{Aut}\, K$, du corps de décomposition K_T, de son groupe de Galois,
(8) calcul approché de $\zeta_K(s)$, $s \in \mathbb{C}$.
etc.

PARTIE II
ALGORITHMIQUE

9. Introduction

9.1. Complexité. — On ne définira pas précisément ce qu'est un *algorithme* (qui impliquerait la spécification d'un modèle de calcul, de la représentation des données, d'un programme et de la relation souhaitée entre entrée et sortie de celui-ci), qu'on assimilera à un programme informatique. La *taille* d'une donnée est le nombre de bits utilisés pour la coder, dans la représentation choisie. Le *temps de calcul* est le nombre de pas de programme exécutés avant l'arrêt.

Nous considérerons des algorithmes *probabilistes*, qui ont le droit de faire des choix aléatoires, et des algorithmes *déterministes*, qui ne l'ont pas. Informellement, un algorithme est *bon* s'il s'exécute en un temps polynomial en la taille combinée des entrées et de la sortie du programme[(2)]. Pour un algorithme probabiliste, ceci signifie que, pour une entrée fixée, la moyenne des temps de calculs sur toutes les exécutions possibles de l'algorithme est polynomiale en la taille T ; l'incertitude porte uniquement sur le temps de calcul, pas sur la correction du résultat. Un problème est *facile* si on connaît un bon algorithme pour le résoudre, et *difficile* sinon. À défaut de bon algorithme, on apprécie qu'il soit sous-exponentiel, en temps $o(\exp(\varepsilon T))$ pour tout $\varepsilon > 0$.

C'est une définition naïve : on ne se prononce pas sur la difficulté intrinsèque du problème et, de fait, la classe des problèmes « difficiles » a tendance à se résorber. À l'inverse, une modification anodine d'un problème facile peut le rendre difficile[(3)] ; donc, sans spécification précise du modèle de calcul, on bâtit sur du sable. Mais on peut déjà dire des choses intéressantes à partir de cette approche naïve. (Voir [**3**], et surtout [**32**], pour une approche plus formelle.) Il est aussi utile de distinguer plus finement entre bons, ou moins bons, algorithmes. Dans ce survol, on ne s'en préoccupera pas, ou à peine.

Quelques problèmes faciles :

– primalité sur $\mathbb{Z}$ (Agrawal-Kayal-Saxena [**2**]),

– factorisation sur $\mathbb{Q}[X]$ (Lenstra-Lenstra-Lovász [**27**]) ou sur $\mathbb{F}_q[X]$ (Berlekamp [**8**]),

– construction d'un corps fini $\mathbb{F}_{p^n}$ (Galois, Adleman-Lenstra [**1**]).

Quelque problèmes difficiles :

– factorisation sur $\mathbb{Z}$,

– calcul d'une $\mathbb{Z}$-base de $\mathcal{O}_K$,

– groupe de Galois d'un polynôme de $\mathbb{Q}[X]$,

– calcul de $\mathrm{Cl}(K)$ ou $U(K)$.

Si on rejette les algorithmes probabilistes, alors la factorisation sur $\mathbb{F}_q[X]$ et la construction d'un corps fini deviennent difficiles, mais le test d'irréductibilité dans $\mathbb{F}_q[X]$ reste facile. Même le calcul d'une

(2) Ainsi la taille associée au problème de la recherche des sous-ensembles de $\{1,\dots,n\}$ est de l'ordre de 2^n, pas $\log n$. La taille de l'entrée est $\log n$, celle de la sortie $\sum_{k=0}^{n}\binom{n}{k}k\log n = 2^{n-1}n\log n$.

(3) Par exemple exiger un algorithme déterministe, ou demander un bon comportement sur des données creuses (on mesure alors la taille en fonction du nombre de coefficients non nuls d'un polynôme, par exemple, et non plus en fonction de son degré).

racine carrée ou la construction d'un corps quadratique $\mathbb{F}_{p^2}$ sont difficiles dans le cadre déterministe. *Si* GRH est vraie, construire $\mathbb{F}_{p^n}$ redevient facile ; plus précisément, on dispose d'un bon algorithme déterministe (utilisant essentiellement le théorème de Chebotarëv) qui, pour chaque valeur de (n,p) fixée, soit construit un polynôme irréductible de degré n sur $\mathbb{F}_p$, soit prouve que GRH est fausse.

9.2. Un exemple. — Nous allons « calculer » des groupes abéliens : $\mathrm{Cl}(K)$, $U(K)$... Qu'entend-on par là ? Commençons par un exemple simple : soit p un nombre premier, on veut « calculer » $G = (\mathbb{Z}/p\mathbb{Z})^*$ le groupe multiplicatif du corps fini associé. Quelques descriptions possibles :

– théorique : $G \simeq \mathbb{Z}/(p-1)\mathbb{Z}$ est cyclique.

– pseudo-effective : $G = \mathbb{Z}/(p-1)\mathbb{Z} \cdot g$, où on fixe un générateur g pour fixer un isomorphisme. La recherche de g se fait en temps fini en calculant l'ordre des éléments successifs de G.

Complexité : la taille des données est $\log_2 p$. En supposant que la factorisation de $p-1$ est connue, on a un bon algorithme pour calculer cet ordre. Sous GRH (pour $K = \mathbb{Q}$ et les caractères modulo p), il existe un générateur de représentant $O(\log^2 p)$. On a donc un bon algorithme en testant $2, 3, \ldots$ Seul le caractère polynomial de l'algorithme est conditionnel. Si la factorisation de $p-1$ est inconnue, même le calcul de l'ordre d'un élément est difficile. Par exemple, vérifier qu'un g donné répond bien à la question est difficile.

– effective : comme la précédente mais, pour que l'isomorphisme soit effectif, il faut savoir résoudre le problème du logarithme discret dans G : étant donné $a \in G$, trouver l'unique $x := \log_g a \in \mathbb{Z}/(p-1)\mathbb{Z}$ tel que $g^x = a$.

Plus généralement pour décrire un groupe abélien de type fini, on le représente sous la forme

$$G = \bigoplus_{i=1}^{g} (\mathbb{Z}/d_i\mathbb{Z}) g_i, \quad \text{où } d_1 \mid \cdots \mid d_g.$$

Le problème du logarithme discret correspondant est le plus souvent difficile.

9.3. Notations. — Dans la suite, on fixe $K/\mathbb{Q}$ un corps de nombres, $n := \dim_{\mathbb{Q}} K$. On supposera que K est donné par le polynôme minimal

T d'un élément primitif[4]. En d'autres termes $K = \mathbb{Q}[X]/(T)$; on note $\theta = X \pmod{T}$, soit $K = \mathbb{Q}(\theta)$. On supposera que $T \in \mathbb{Z}[X]$. La lettre p désigne toujours un nombre premier, et $\mathfrak{p}$ un idéal maximal de $\mathcal{O}_K$ au dessus de p.

Après quelques préliminaires, dont le plus important est sans conteste l'algorithme LLL (§ 10.5), on s'intéressera aux plongements du corps K : complexes (§ 11), p-adiques (§ 12), ainsi que dans d'autres corps de nombres (§ 13), ce qui se traduit par de classiques problèmes de factorisation de polynômes, dans $\mathbb{C}[X]$, $\mathbb{Q}_p[X]$ (en particulier $\mathbb{F}_p[X]$), ou $\mathbb{Q}[X]$. Dans une deuxième partie, on s'intéressera à l'arithmétique de K : ses ordres (§ 14) et son anneau d'entiers $\mathcal{O}_K$ (§ 15). Finalement, on parlera de groupes de classes et d'unités (§ 16).

10. Préliminaires

10.1. Opérations élémentaires. — Les opérations élémentaires sont faciles dans $\mathbb{Z}$: addition, multiplication, division euclidienne, pgcd (algorithme d'Euclide). Il en est de même dans le corps de fractions $\mathbb{Q}$, ainsi que dans les quotients $\mathbb{Z}/N\mathbb{Z}$, manipulés *via* un système complet de représentants $(0, \dots, N-1$, ou bien les représentants de valeur absolue minimale). À ceci près que dans $\mathbb{Z}/N\mathbb{Z}$ la division euclidienne est remplacée par l'inversion, qui échoue sur un diviseur de 0, et se réalise à l'aide de l'algorithme d'Euclide étendu. Plus généralement, une relation de Bezout entre $a, b \in \mathbb{N}$ s'obtient grâce à une suite d'opérations sur les lignes du système

$$\begin{pmatrix} u_i & v_i \\ u_{i+1} & v_{i+1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix}, \quad \text{avec } \begin{pmatrix} u_0 & v_0 \\ u_1 & v_1 \end{pmatrix} = \mathrm{Id}, \ \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix}.$$

Si $x_{i+1} = 0$, l'algorithme s'arrête et la première ligne est une relation de Bezout, sinon on pose la division euclidienne $x_i = qx_{i+1} + r$ et $(L_1, L_2) \leftarrow (L_2, L_1 - qL_2)$. Ceci effectue simplement l'algorithme d'Euclide sur le membre de droite, en conservant la trace des opérations dans une matrice auxiliaire, comme il est d'usage en algèbre

[4] D'autres points de vue sont possibles, qui ont tous leur intérêt, par exemple $K \overset{\sigma}{\hookrightarrow} \mathbb{C}$ (on fixe un plongement explicite), K corps de décomposition ou donné par une tour d'extensions, K compositum de sous-corps, corps fixe donné par théorie de Galois, extension donnée par théorie du corps de classe...

linéaire. En inversant l'identité ci-dessus, on obtient

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} |v_{i+1}| & |v_i| \\ |u_{i+1}| & |u_i| \end{pmatrix} \begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix},$$

d'où on déduit des majorations de toutes les quantités en jeu, et un coût quadratique $O(\log \max(a, b))^2$ pour l'algorithme. À ce stade, on remarque qu'il suffit de calculer une seule colonne de la matrice $\left(\begin{smallmatrix} u_i & v_i \\ u_{i+1} & v_{i+1} \end{smallmatrix}\right)$, puisque la connaissance du pgcd et d'un coefficient de Bezout permet de déterminer l'autre.

À partir d'une relation de Bezout, le lemme chinois est effectif. Les anneaux de polynômes $\mathbb{F}_p[X]$ et leurs quotients se traitent de même, où les représentants choisis sont de degrés minimal. On peut traiter de même $\mathbb{Q}[X]$ et ses quotients, mais l'explosion des coefficients fait que l'algorithme d'Euclide naïf n'est plus un algorithme utilisable. (Il reste bon, mais ça n'a rien d'évident, voir [**17**, Chap. 6].) Une méthode modulaire s'impose : calcul du pgcd dans $\mathbb{F}_p[X]$ modulo suffisamment de petits premiers et reconstruction du pgcd global par lemme chinois.

À l'exception du pgcd dans $\mathbb{Q}[X]$, les algorithmes naïfs pour les problèmes ci-dessus sont tous au pire quadratiques en la taille des données. Les méthodes utilisant la multiplication par transformation de Fourier rapide (§ 11.3) et l'inversion par itération de Newton (10) sont asymptotiquement quasi-linéaires, voir [**17**].

10.2. Exponentiation binaire. — La technique est simple mais utile : si $n = \sum_{i=0}^{k} \varepsilon_i 2^i > 0$, $\varepsilon_i \in \{0, 1\}$, on calcule

$$x^n = \prod_{\substack{i=0 \\ \varepsilon_i = 1}}^{k} x^{2^i},$$

où les x^{2^i} sont obtenus par mise au carré successives, soit $O(\log n)$ multiplications au lieu de $n - 1$. Notons une application typique, en dehors des tests de primalité : $\mathrm{pgcd}(X^p - X, T)$ dans $k[X]$ se calcule comme $\mathrm{pgcd}(A - X, T)$, où $A \in k[X]$ est un représentant de la classe de X^p dans $k[X]/(T)$, calculé en $O(\log p)$ multiplications n'impliquant que des polynômes de degré inférieur à $\deg T$.

Des variantes gagnent sur la constante implicite, par exemple

$$y_0 := x^{e_k}, \quad y_i := y_{i-1}^2 x^{e_{k-i}} \quad \text{pour } 1 \leqslant i \leqslant k,$$

alors $y_k = x^n$, qui rappelle le schéma de Horner. Dans le cas typique ci-dessus, les multiplications par x^{e_i} sont des multiplications par X, peu coûteuses.

10.3. Factorisation, primalité dans $\mathbb{Z}$. — La primalité est facile, la factorisation difficile. Ces techniques sont complexes, mais bien et fréquemment décrites dans la littérature. Voir [**34**] et [**10, 26**] respectivement pour une présentation des tests de primalité et des algorithmes de factorisation.

10.4. Formes normales d'Hermite (HNF) et de Smith (SNF)

Ce sont des familles de matrices, généralisant les formes normales de Gauss-Jordan des espaces vectoriels aux $\mathbb{Z}$-modules (plus généralement aux A-modules pour A principal). L'algorithme qui permet de s'y ramener (si la relation de Bezout est effective dans A, par exemple si A est euclidien) généralise le pivot de Gauss. L'inversion d'un pivot est en général impossible, mais l'algorithme d'Euclide permet d'y suppléer : si $au + bv = \mathrm{pgcd}(a, b) = d$ est une relation de Bezout, la multiplication à droite par la matrice inversible $\left(\begin{smallmatrix} b/d & u \\ -a/d & v \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ met la matrice ligne (a, b) sous forme normale $(0, d)$.

Par une suite d'opérations élémentaires de ce type sur ses colonnes, une matrice $m \times n$ à coefficients dans $\mathbb{Z}$ peut être mise sous forme échelonnée, dite *forme normale d'Hermite* (HNF), en annulant successivement tous les coefficients *à gauche* d'un pivot dans une ligne donnée, avant de poursuivre sur la matrice extraite dont on a supprimé la ligne et la colonne du pivot. Il n'est en général pas possible d'annuler les coefficients *à droite* d'un pivot q, tout au plus peuvent-ils être choisis dans un système fixé de représentants de $\mathbb{Z}/q\mathbb{Z}$. Formellement, la matrice $(0|H) \in M_{m\times n}(\mathbb{Z})$ est une HNF si $H = (h_{i,j})$ est une matrice $m \times r$ de rang $r \leqslant n$ telle qu'il existe une fonction $f : [1, r] \to [1, m]$ strictement croissante telle pour $1 \leqslant j \leqslant r$, on ait

(1) $q_j := h_{f(j),j} > 0$ et $h_{i,j} = 0$ si $i > f(j)$,

(2) $0 \leqslant h_{f(j),k} < q_j$ si $k > j$.

($f(j)$ indique la ligne où se trouve le pivot $q_j > 0$ de la colonne j de H, les coefficients à droite du pivot sont réduits modulo q_j.) Les matrices HNF $m \times n$ forment un système complet de représentants de $M_{m\times n}(\mathbb{Z})/\mathrm{GL}_n(\mathbb{Z})$. En d'autres termes, le $\mathbb{Z}$-module engendré par les colonnes d'une matrice a une base (échelonnée) canonique, donnée par les colonnes non nulles de sa HNF.

En autorisant de plus les opérations élémentaires sur les lignes, on applique un algorithme de pivot à la ligne L_i d'une matrice (opérations sur les colonnes), puis à la colonne C_j (opérations sur les lignes) ; en itérant à i et j fixés, le pivot $q_{i,j}$ diminue strictement jusqu'à être le seul élément non nul de L_i et C_j. En poursuivant sur les autres lignes et colonnes, on obtient cette fois une matrice diagonale, à un bloc de 0 près. Toujours par des opérations inversibles sur lignes et colonnes, on peut remplacer un couple (a, b) de pivots consécutifs par $(\mathrm{ppcm}(a, b), \mathrm{pgcd}(a, b))$, et ainsi supposer que les pivots successifs $(d_1, \dots, d_n)$ sont positifs ou nuls et vérifient $d_n \mid \cdots \mid d_1$. Formellement, les matrices $m \times n$ de forme $(0|D)$ ou $\left(\begin{smallmatrix} 0 \\ D \end{smallmatrix}\right)$ sont des *Formes Normales de Smith* (SNF) si D est une matrice diagonale vérifiant la condition ci-dessus. Les matrices SNF $m \times n$ forment un système complet de représentants de $\mathrm{GL}_m(\mathbb{Z})\backslash M_{m\times n}(\mathbb{Z})/\mathrm{GL}_n(\mathbb{Z})$. Nous venons essentiellement de démontrer le Théorème 1.1 de la base adaptée (reste à voir l'unicité des d_i).

Lemme 10.1. — *Soit $M \in M_{n\times n}(\mathbb{Z})$ et $(d_1, \dots, d_n)$ la diagonale de sa SNF. Alors $G = \mathbb{Z}^n/\operatorname{Im} M \simeq \oplus_{i=1}^n (\mathbb{Z}/d_i\mathbb{Z})$.*

Démonstration. — Soit $D = UMV$ la SNF de M, où $U, V \in \mathrm{GL}_n(\mathbb{Z})$. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{Z}^n$; ses projections engendrent G. Les seules relations entre les e_i sont de la forme $(e_1, \dots, e_n)MY = 0$, $Y \in \mathbb{Z}^n$. Soit

$$(g_1, \dots, g_n) = (e_1, \dots, e_n)U^{-1}$$

une nouvelle famille génératrice de G ; on a $(g_1, \dots, g_n)X = 0$ si et seulement si $X \in \operatorname{Im} D$. Donc

$$G = \Big(\sum_i \mathbb{Z} \cdot g_i\Big)\Big/ \operatorname{Im} D = \sum_{i=1}^n (\mathbb{Z}/d_i\mathbb{Z}) \cdot \overline{g_i}. \qquad \square$$

Se restreindre à $M_{n\times n}(\mathbb{Z})$ au lieu de $M_{m\times n}(\mathbb{Z})$ est sans importance : on peut remplacer M par les colonnes non nulles de sa HNF pour assurer $n \leqslant m$, puis concaténer des colonnes nulles pour assurer $n = m$.

Contrairement à Euclide sur $\mathbb{Z}$ (mais conformément à Euclide sur $\mathbb{Q}[X]$), les algorithmes naïfs évoqués ci-dessus ne sont pas bons : les

$O(nm^2)$ opérations dans $\mathbb{Z}$ associées à l'algorithme de pivot font exploser les coefficients. De bons algorithmes existent, reposant toujours sur des techniques modulaires (voir Storjohann [**37**]).

Un des principes de l'algorithmique arithmétique est de transformer les problèmes arithmétiques en algèbre linéaire sur des $\mathbb{Z}$-modules ; l'algorithme d'Euclide pour le pgcd est l'exemple type. L'algèbre linéaire sur un corps fini ou sur $\mathbb{Z}$ est facile. Par là j'entends plus précisément la résolution de systèmes linéaires, le calcul d'une base de noyaux, images ou conoyaux, et celui de polynôme minimal ou caractéristique. Sur un corps fini cela suit du pivot de Gauss, sur $\mathbb{Z}$ de la forme normale d'Hermite. En particulier, il est facile de transformer un système fini de $\mathbb{Z}$-générateurs d'un $\mathbb{Z}$-module libre en une $\mathbb{Z}$-base. Si A et B sont des sous-$\mathbb{Z}$-modules libres de $\mathbb{Z}^n$, vu comme $\mathbb{Z}$-algèbre, on en déduit de bons algorithmes pour le calcul de $A\cap B$, $A+B$, AB, $(A:B)=\{z\in\mathbb{Z}^n : zB\subset A\}$.

Exemple. — Soit à résoudre le système linéaire $XM = Y$, $M \in M_{m\times n}(\mathbb{Z})$, d'inconnue $X\in M_{\ell\times m}(\mathbb{Z})$. On pose $MU=(0|H)$ la HNF de M, où $U\in \mathrm{GL}_n(\mathbb{Z})$ et H est de rang $r\leqslant n$. L'équation est équivalente au système triangulaire $X(0|H) = YU$, qu'il est facile de résoudre par substitution.

10.5. Réduction de bases de réseaux, LLL. — Parmi les bases d'un réseau, certaines sont plus agréables que d'autres : ainsi la base canonique de $\mathbb{Z}^2$ est une base de vecteurs courts (pour la forme euclidienne standard), mais si $au-bv=1$ est une relation de Bezout, les vecteurs $\binom{a}{b},\binom{v}{u}$ forment une base de $\mathbb{Z}^2$, dont les vecteurs peuvent être rendus arbitrairement longs. La réduction d'une base est simplement son remplacement par une base dont les vecteurs sont plus courts. On peut définir une notion de base optimale (Minkowski), dont les vecteurs sont aussi courts que possibles, mais on ne connaît pas de bon algorithme pour en déterminer une à partir d'une base arbitraire. Essentiellement, on en vient à tester successivement tous les vecteurs du réseau dans une boule et leur nombre est exponentiel en la dimension. Nous allons définir une notion de réduction adaptée au traitement algorithmique.

Soit $(\mathbb{R}^n, q)$ un espace euclidien et Λ un réseau donné par une base $(b_i)_{i\leqslant n}$. Soit (b_i^*) la base orthogonale de $\mathbb{R}^n$ donnée par le procédé de

Gram-Schmidt, telle que

$$b_1^* = b_1, \quad b_i^* = b_i - \sum_{j<i} \mu_{i,j} b_j^*, \quad \mu_{i,j} = \frac{b_i \cdot b_j^*}{b_j^* \cdot b_j^*}, \quad \text{pour } 1 \leqslant j < i \leqslant n.$$

Lemme 10.2. — $\prod_{i=1}^n q(b_i^*) = d(\Lambda)^2$ *ne dépend que de* Λ.

Démonstration. — La matrice de Gram de (b_i^*) est diagonale, donc son déterminant est $\prod q(b_i^*)$. D'autre part la matrice de changement de $\mathbb{R}$-base de (b_i) à (b_i^*) est triangulaire de déterminant 1. On conclut avec la Proposition 1.5. □

Définition 10.3. — Soit $c \in \,]1/4, 1[$. Une famille libre de vecteurs $(b_1, \ldots, b_n)$ est dite LLL-réduite pour la constante de Lovász c si

(1) $|\mu_{i,j}| \leqslant \frac{1}{2}$, pour $1 \leqslant j < i \leqslant n$.
(2) $q(b_i^* + \mu_{i,i-1} b_{i-1}^*)/q(b_{i-1}^*) \geqslant c$, pour $1 < i \leqslant n$.

La première condition $|\mu_{i,j}| \leqslant 1/2$ pour $j < i$ est facile à mettre en œuvre et suit d'une « réduction de b_i modulo $b_1, \ldots, b_{i-1}$ ». Supposons qu'elle soit vérifiée pour $\ell < j < i$ (initialement $\ell = i$), alors la substitution $b_i \leftarrow b_i - \lceil \mu_{i,\ell} \rfloor b_\ell$

– ne modifie pas les $\mu_{t,j}$ pour $t < i$
– ne modifie pas les $\mu_{i,j}$ pour $j > \ell$
– remplace $\mu_{i,\ell}$ par $\mu_{i,\ell} - \lceil \mu_{i,\ell} \rfloor$ dont la valeur absolue est $\leqslant 1/2$.

La deuxième condition compare les longueurs des projections de b_i (numérateur) et b_{i-1} (dénominateur) sur l'orthogonal de $(b_1, \ldots, b_{i-2})$. Intuitivement, on veut placer en premier le plus « petit » des deux vecteurs (par rapport aux vecteurs restant à réduire), puisqu'il réduira « mieux » les vecteurs suivants.

Pour simplifier, on prend $c = 3/4$ pour la suite. La condition (2) est équivalente à $q(b_i^*) \geqslant \left(\frac{3}{4} - \mu_{i,i-1}^2\right) q(b_{i-1}^*)$. Avec la première, elle implique $q(b_{i-1}^*) \leqslant 2q(b_i^*)$, et en fait

$$\frac{q(b_i)}{q(b_i^*)} = 1 + \sum_{j<i} \mu_{i,j}^2 \frac{q(b_j^*)}{q(b_i^*)} \leqslant 1 + \frac{1}{4} \sum_{j<i} 2^{i-j} = \frac{2^i + 1}{2}$$

Ce quotient étant $\geqslant 1$, $q(b_i)$ et $q(b_i^*)$ sont proches. On en déduit facilement :

Proposition 10.4. — *Soit Λ un réseau de l'espace euclidien $(\mathbb{R}^n, q)$ dont (b_i) une base* LLL-*réduite. Alors, pour toute toute suite de vecteurs indépendants $x_1, \ldots, x_n$ de Λ, on a*

$$q(b_j) \leqslant 2^{n-1} \max \{q(x_1), \ldots, q(x_j)\}, \quad \textit{pour } j \leqslant n.$$

En particulier le premier vecteur b_1 est essentiellement aussi court que possible, à un facteur 2^{n-1} près. Trouver un *plus court* vecteur est difficile.

Bien sûr, l'intérêt de la notion est l'existence d'un bon *algorithme* de calcul d'une base LLL-réduite, dû à A. Lenstra, H. Lenstra et Lovász [**27**]. L'algorithme est très simple. Comme celui d'Euclide qu'il généralise c'est une succession de réductions et d'échanges : supposons que $(b_1, \ldots, b_{k-1})$ soit LLL-réduite (c'est initialement le cas pour $k = 2$).

Réduction. — On remplace b_k par une combinaison linéaire $b_k - \sum_{i<k} a_i b_i$ comme ci-dessus de façon à ce que $|\mu_{k,j}| \leqslant \frac{1}{2}$ pour tout $j < k$.

Échange

– si $q(b_k^* + \mu_{k,k-1} b_{k-1}^*) \geqslant \frac{3}{4} q(b_{k-1}^*)$, alors $(b_1, \ldots, b_k)$ est LLL-réduite et on incrémente k. Si $k > n$, l'algorithme s'arrête.

– sinon, on échange b_{k-1} et b_k et on décrémente k.

Intuitivement, la réduction diminue la taille de b_k, et les échanges mélangent les vecteurs en tentant de favoriser les réductions suivantes. Il est clair que $(b_1, \ldots, b_n)$ reste une base du réseau tout au long de l'algorithme, et qu'elle est LLL-réduite quand il s'arrête. Il n'est pas clair qu'il s'arrête, ni *a fortiori* que ce soit un bon algorithme.

***Théorème 10.5* (Lenstra-Lenstra-Lovász).** — *C'est un bon algorithme.*

Démonstration (idée). — Soit Λ_i le réseau (de $\mathbb{R}^i$) engendré par $(b_1, \ldots, b_i)$. On note

$$D_i = \prod_{j \leqslant i} q(b_j^*) = d(\Lambda_i)^2 > 0, \quad D := \prod_{i<n} D_i.$$

D ne peut changer que lors d'un échange, où (b_{k-1}^*, b_k^*) est remplacé par (s, t), avec $s = b_k^* + \mu_{k,k-1} b_{k-1}^*$. Puisque D_i ne dépend que de Λ_i et pas de la base utilisée pour le définir (Lemme 10.2), on a $q(s)q(t) = q(b_{k-1}^*)q(b_k^*)$ et l'échange laisse les D_i invariants, à l'exception de D_{k-1}

qui est multiplié par $q(s)/q(b^*_{k-1}) < \frac{3}{4}$. Donc D est multiplié par un facteur $< \frac{3}{4}$.

Soit m_i la longueur d'un plus petit vecteur non nul de $\Lambda_i \subset \mathbb{R}^i$. D'après le théorème de Minkowski, $m_i^i v_i \leqslant 2^i d(\Lambda_i)$ où v_i est le volume de la boule unité de $\mathbb{R}^i$ (et $m_i^i v_i$ celui de la boule de rayon m_i). On en déduit que

$$d(\Lambda_i) \geqslant v_i \cdot (m_i/2)^i \geqslant v_i \cdot (m_n/2)^i,$$

qui ne dépend que de i et de Λ, mais plus des (b_i). On en déduit que D est minoré par une constante strictement positive ne dépendant que du réseau, ce qui permet de borner le nombre d'échanges avant l'arrêt de l'algorithme. Reste à compter le nombre d'opérations et à estimer la taille de leurs opérandes en termes des données. □

Plus précisément :

Théorème 10.6 **(Nguyen-Stehlé [30]).** — *Soit $\Lambda \subset \mathbb{Z}^d$ un réseau, donné par une famille génératrice de n vecteurs de longueur inférieure à B. Une base* LLL*-réduite de Λ se calcule en temps $O(d^4 n(d + \log B)\log B)$.*

11. Factorisation dans $\mathbb{C}[X]$

Notre but est de calculer les plongements du corps de nombres $K = \mathbb{Q}[X]/(T)$ ce qui revient à approcher les racines complexes de T. Les méthodes itératives de type Newton sont bien connues, mais il est difficile de garantir un temps d'exécution, ni l'exactitude du résultat en présence d'erreurs d'arrondi. J'en tire prétexte pour présenter l'utilisation d'algorithmes numériques (Newton, lemme chinois approché, intégration complexe par discrétisation, FFT, méthode de Graeffe) dans notre contexte algébrique.

Les racines complexes sont une fonction continue du polynôme, mais le problème est déjà mal conditionné par rapport au polynôme lui-même :

$$P = (x - 1/10)^{20}, \quad \text{et} \quad \widehat{P} = P - 10^{-20}$$

coïncident jusqu'à 20 chiffres après la virgule, mais $\widehat{P}$ admet la racine 0 qui est à distance 0.1 de l'unique racine de P. Dans le contexte de la résolution d'équations diophantiennes, il faut borner rigoureusement la distance des racines approchées aux racines de P, et pouvoir la rendre arbitrairement petite.

11.1. L'algorithme de Schönhage. — On présente ici les grandes lignes de l'algorithme de Schönhage, développé par Gourdon [**18**]. Étant donné un polynôme $P \in \mathbb{C}[X]$ unitaire de degré n, et un paramètre d'erreur ε, l'algorithme retourne n complexes $z_1, \dots, z_n$ tels que

$$|P - (X - z_1) \cdots (X - z_n)| < \varepsilon \, |P| \,,$$

où $|\cdot|$ est la norme L^1 ; on en déduira un encadrement des racines au § 11.6. Ce type de résultat est particulièrement adapté au cas de données approchées : plutôt que d'estimer l'erreur commise sur le résultat, on estime la distance des données réelles à des données virtuelles qui produiraient ce résultat exact.

C'est un bon algorithme, quoique assez technique : pour des raisons d'efficacité, les calculs sont approchés, et il faut démontrer de nombreux résultats de perturbation, ainsi qu'estimer au plus près de nombreux paramètres.

Si P est de degré 1, il n'y a rien à faire. Sinon, l'algorithme détermine d'abord un cercle de séparation Γ, dont l'intérieur contient $k < n$ racines de P, disons $u_1, \dots, u_k$ (idéalement $k \approx n/2$ et Γ est éloigné des racines de P), puis il approche les sommes de Newton associées :

$$s_m = u_1^m + \cdots + u_k^m = \frac{1}{2i\pi} \oint_\Gamma \frac{P'(z)}{P(z)} z^m \, dz, \quad 1 \leqslant m \leqslant k,$$

par intégration numérique. Grâce aux formules de Newton, on reconstruit une valeur approchée F_0 de $F = \prod_{i \leqslant k} (X - u_i)$ à partir des valeurs approchées des s_i, et on applique récursivement l'algorithme à F_0 et $G_0 = P/F_0$ qui approche $G = P/F$.

Voici un exemple typique d'utilisation non triviale : soit $T_{300} \in \mathbb{Z}[X]$ le 300-ème polynôme de Chebyshev, de norme infinie $\approx 3 \cdot 10^{113}$, on veut les racines de $T_{300} + i$ (toutes de module < 1, proches de l'axe réel). Le temps de calcul sur une machine à 1.6GHz est de $\approx$ 28 secondes pour 28 décimales garanties, $\approx$ 75 secondes pour 1000 décimales.

11.2. Itération de Newton-Schönhage. — En pratique, on applique une stratégie mixte : l'approximation grossière $P \approx F_0 G_0$ est raffinée par une itération de Newton, fondée sur une nouvelle intégration complexe. Plus précisément, on cherche des termes correcteurs f et g tels que $F_1 = F_0 + f$ et $G_1 = G_0 + g$ soient de meilleures approximations que F_0 et G_0, c'est-à-dire telles que $P - F_0 G_0 = f G_0 + g F_0$ (ce qui implique que $P - F_1 G_1 = -fg$ est du second ordre). On peut obtenir f et g par l'algorithme d'Euclide, mais ce dernier est instable et difficile à contrôler numériquement.

On introduit plutôt H, le représentant de degré minimal de l'inverse de G_0 dans $\mathbb{C}[X]/(F_0)$, qui existe puisque les racines de G_0 sont à l'extérieur de Γ et celles de F_0 à l'intérieur. Il suffit ensuite de poser $f = HP \bmod F_0$ et de calculer $G_0 + g$ comme quotient de la division euclidienne de P par $F_0 + f$. Au lieu de l'algorithme d'Euclide pour le calcul d'un inverse dans $\mathbb{C}[X]/(F_0)$, on utilise la formule

$$H(X) = \frac{1}{2i\pi} \oint_\Gamma \frac{1}{F_0 G_0(z)} \frac{F_0(X) - F_0(z)}{X - z} \, dz.$$

En effet, c'est bien un polynôme de degré $< k = \deg F_0$ et un calcul de résidu montre que $H(z) = 1/G_0(z)$ si z est l'une des k racines de F_0 à l'intérieur de Γ. Par intégration numérique, on obtient une approximation grossière H_0 de H que l'on raffine par l'itération de Newton

$$H_{m+1} = H_m(2 - H_m G_0) \pmod{F_0},$$

qui converge quadratiquement vers H. Cette dernière est un analogue de l'itération classique utilisée pour calculer l'inverse dans $\mathbb{R}$, qui suit de la méthode de Newton appliquée à la fonction $f(x) = \frac{1}{x} - y$:

$$x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)} = x_m(2 - x_m y). \tag{10}$$

11.3. Intégration numérique et FFT. — Par translation et homothétie, on ramène les intégrales sur Γ au cercle unité. On discrétise le cercle par les racines de l'unité d'ordre 2^N et il suffit d'évaluer nos fractions rationnelles en ces points (et de borner l'erreur commise en fonction de la distance estimée de Γ aux racines de P).

On utilise la transformée de Fourier rapide (FFT, pour Fast Fourier Transform), qui permet d'évaluer un polynôme Q de degré $n < 2^N$ en toutes les racines 2^N-èmes de l'unité en $O(N 2^N)$ multiplications au lieu de 2^{2N} par l'algorithme de Horner appliqué 2^N fois. Soit $Q_{\text{pair}}(X^2)$ la partie paire de Q et $XQ_{\text{impair}}(X^2)$ sa partie impaire. La FFT vient de l'identité

$$Q(\omega^j) = Q_{\text{pair}}\big((\omega^2)^j\big) + \omega^j Q_{\text{impair}}\big((\omega^2)^j\big),$$

qui ramène essentiellement le calcul d'une transformée de longueur L à deux transformée de longueur $L/2$.

Ce type d'idées permet aussi de multiplier rapidement polynômes[5] et entiers (voir [**17**]), et peut être utilisé par les autres phases de l'algorithme, en particulier l'itération de Graeffe.

11.4. Le cercle de séparation. — Notons $u_1, \dots, u_n$ les racines de $P = a_n X^n + \dots + a_0$ ordonnées par modules croissants et $\rho_1(P) \leqslant \dots \leqslant \rho_n(P)$ leurs modules. On suppose $n \geqslant 2$. Pour déterminer un cercle Γ convenable, il suffit de savoir estimer les $\rho_k(P)$.

En effet, par translation de vecteur a_{n-1}/a_n, on place le barycentre des racines à l'origine. Par homothétie de rapport $1/\rho_n(P)$ on ramène la racine de plus grand module au voisinage du cercle unité. Un petit dessin montre que l'un des 4 polynômes translatés $P(X+\lambda)$, pour $\lambda = \pm 2, \pm 2i$ vérifie $\rho_n \geqslant 2$, $\rho_1 \leqslant \big|2 - e^{i\pi/4}\big|$, soit $\rho_n/\rho_1 > 1.35$.

On s'est donc ramené au cas où ρ_n/ρ_1 n'est pas trop proche de 1. Le cercle Γ est centré en 0, reste à choisir son rayon R. La quantité intervenant dans les termes d'erreurs est δ tel qu'il n'y ait aucune racine de P dans la couronne $Re^{-\delta} < |z| < Re^{\delta}$, et il faut la maximiser. Si j maximise le rapport de deux modules consécutifs ρ_{j+1}/ρ_j, on pose donc $R = \sqrt{\rho_j \rho_{j+1}}$. En pratique, on ne calcule pas tous les ρ_j, $j \leqslant n$, mais on utilise une dichotomie qui en calcule $O(\log n)$ pour produire un cercle de même qualité.

11.5. Estimation de $\rho_k(P)$, la méthode de Graeffe. — On note Graeffe(P) le polynôme Q tel que $Q(X^2) = P(X)P(-X)$: c'est un polynôme de même degré que P, dont les racines sont les carrés des racines de P. On construit une suite de polynômes (P_m) par la récurrence $P_0 = P$, $P_{m+1} = \text{Graeffe}(P_m)$ et on écrit $P_m = \sum_{k \leqslant n} a_k^{(m)} X^k$.

[5] À partir de ses valeurs aux racines 2^N-èmes de l'unité, on reconstruit un polynôme de degré $< 2^N$ par interpolation de Lagrange, qui est une transformée de Fourier inverse dans ce cadre.

On a

$$\frac{a_k^{(m)}}{a_n^{(m)}} = \sum_{i_1 < \cdots < i_{n-k}} (u_{i_1} \cdots u_{i_{n-k}})^{2^m}.$$

Supposons dans un premier temps que les modules des racines sont 2 à 2 *distincts*. Alors

$$\frac{a_k^{(m)}}{a_n^{(m)}} \sim (u_{k+1} \cdots u_n)^{2^m}, \quad \text{et} \quad \lim_{m \to +\infty} \left| a_{k-1}^{(m)} / a_k^{(m)} \right|^{2^{-m}} = \rho_k(P).$$

Bien sûr si les ρ_j sont proches, la convergence peut être lente. Pour obtenir une borne rigoureuse, on applique des inégalités générales grossières aux P_m, dont on déduit des bornes arbitrairement précises pour P lorsque m est assez grand.

***Lemme 11.1* (de Cauchy).** — *Soit $P = a_0 + \cdots + a_n X^n \in \mathbb{C}[X]$, alors*

$$\rho_n(P) \leqslant 2 \max_{i<n} |a_i/a_n|^{1/n-i}.$$

Démonstration. — Si $(|z|/2)^{n-i} > |a_i/a_n|$ pour tout i, on a

$$\sum_{i<n} |a_i| \, |z|^i < |a_n| \, |z|^n \sum_{i<n} 2^{n-i} < |a_n| \, |z|^n.$$

Donc $P(z) = 0$ est impossible. □

Voir Henrici [**21**, § 6.4] pour des renforcements conséquents du lemme suivant :

***Lemme 11.2*.** — *Soit $P = a_0 + \cdots + a_n X^n \in \mathbb{C}[X]$, $n \geqslant 2$, et $k \in \mathbb{N}$ tel que $|a_k| = \max_{i \leqslant n} |a_i|$ et $a_0 a_n \neq 0$. Alors*

$$\rho_k(P) \leqslant 2n, \quad \rho_{k+1}(P) \geqslant \frac{1}{2n}.$$

Démonstration. — Il suffit de démontrer la première inégalité : la deuxième suit de la première appliquée au polynôme réciproque $P^*(X) = X^n P(1/X)$ de P, qui vérifie $\rho_{n-k}(P^*) = \rho_{k+1}(P)^{-1}$. (L'hypothèse $a_0 a_n \neq 0$ assure que les ρ_i sont non nuls.)

Le cas $k = n$ suit de l'inégalité de Cauchy. Le cas général pour $k < n$ suit de l'inégalité de Cauchy appliquée à $Q =$

$P/(X-u_n)\cdots(X-u_{k+1}) = b_0 + \cdots + b_k X^k$, qui vérifie $\rho_k(Q) = \rho_k(P) =: \rho$. Plus précisément, on a

$$\begin{aligned}\frac{1}{(X-u_n)\cdots(X-u_{k+1})} &= (-1)^{n-k}(u_n\cdots u_{k+1})^{-1}\prod_{\ell=k+1}^{n}\Bigg(\sum_{j_\ell\geqslant 0}\left(\frac{X}{u_\ell}\right)^{j_\ell}\Bigg)\\ &=: \sum_{i\geqslant 0} s_i X^i \in \mathbb{C}[[X]],\end{aligned}$$

où $|s_i| \leqslant \rho^{-(n-k+i)}\#\big\{(j_{k+1},\dots,j_n)\in\mathbb{N}^{n-k} \colon j_{k+1}+\cdots+j_n = i\big\}$. Ce dernier cardinal vaut $\binom{n-k-1+i}{i}$, mais la majoration triviale $(i+1)^{n-k}$ suffit. Supposons $\rho > 2n$, soit $(i+1)\rho^{-1} < 1$ et $|s_i| < \rho^{-i}$ pour tout $i < k \leqslant n$. On a $b_k = a_k$ et

$$|b_i| = \sum_{j\leqslant i} s_j a_{i-j} \leqslant |a_k| \sum_j \rho^{-j} < 2\,|b_k|\,,$$

soit $\rho = \rho_k(Q) \leqslant 4$ (Cauchy) et $4 > 2n$. Contradiction. □

Lemme 11.3. — *Soit $P = a_0 + \cdots + a_n X^n \in \mathbb{C}[X]$, $n \geqslant 1$. On peut choisir $r > 0$ pour que $Q = P(rX) = b_0 + \cdots + b_n X^n$ soit tel qu'il existe ℓ et h vérifiant*

$$\ell < k \leqslant h, \quad |b_\ell| = |b_h| \geqslant |b_j|\,, \quad \textit{pour } j \leqslant n.$$

En particulier

$$\frac{1}{2n} \leqslant \rho_\ell(Q) \leqslant \rho_k(Q) \leqslant \rho_h(Q) \leqslant 2n.$$

Démonstration. — Soit $\mathcal{C}$ l'enveloppe convexe supérieure des points $M_j = (j, \log|a_j|)$. On considère les j tels que $M_j \in \mathcal{C}$ et on choisit $\ell :=$ le plus grand tel $j < k$, et $h :=$ le plus petit tel $j \geqslant k$. On choisit enfin $r = |a_\ell/a_h|^{1/(h-\ell)}$. Les hypothèses du Lemme 11.2 sont donc vérifiées pour les indices ℓ et h et $P = Q$. □

Étant donné $P \in \mathbb{C}[X]$ non nul, un entier $1 \leqslant k \leqslant n$, et un paramètre d'erreur $\delta > 0$, on veut déterminer un rayon $R > 0$ tel que

$$Re^{-\delta} \leqslant \rho_k(P) \leqslant Re^{\delta}.$$

On définit les suites de polynômes P_m, Q_m, par

$$P_0 = P, \quad Q_m = P_m(r_m X), \quad P_{m+1} = \mathrm{Graeffe}(Q_m),$$

où Q_m, r_m proviennent du Lemme 11.3 appliqué à $P = P_m$. Les racines de Q_m sont les

$$(u_i/R_m)^{2^m}, \quad R_m = r_0 r_1^{1/2} \cdots r_m^{1/2^m}.$$

L'inégalité

$$\frac{1}{2n} \leqslant \rho_{\ell_m}(Q_m) \leqslant \rho_k(Q_m) \leqslant \rho_{h_m}(Q_m) \leqslant 2n,$$

donne

$$R_m \cdot (2n)^{-1/2^m} \leqslant \rho_k(P) \leqslant R_m \cdot (2n)^{1/2^m}.$$

Il suffit donc de choisir m tel que $2^m \geqslant \log(2n)/\delta$ et de retourner $R = R_m$.

11.6. Encadrement des racines. — À partir de l'estimation

$$|P - (X - z_1) \cdots (X - z_n)| < \varepsilon^n,$$

on peut estimer les racines $(u_1, \ldots, u_n)$ de P par un argument d'homotopie :

Théorème 11.4 **(Ostrowski).** — *Soit* $\rho = \max(1, |z_1|, \ldots, |z_n|)$. *En réordonnant au besoin les* u_i*, on a pour tout* $i \leqslant n$*, l'inégalité*

$$(1 - 4\rho\varepsilon)\,|u_i - z_i| < 4\rho\varepsilon.$$

Démonstration. — Soit z une racine de $\widehat{P} = (X - z_1) \cdots (X - z_n)$. Considérons la famille continue de polynômes $H_t = tP + (1 - t)\widehat{P}$, $t \in [0, 1]$. Par continuité des racines de H_t, il existe un chemin continu $t \mapsto d_t$ tel que $d_0 = 0$ et $z + d_t$ est racine de H_t pour tout t ; en particulier, $z + d_1 = u$ est racine de $H_1 = P$. Ainsi

$$\left|\widehat{P}(z + d_t)\right| = t\left|(\widehat{P} - P)(z + d_t)\right| < \varepsilon^n(\rho + |d_t|)^n.$$

Par continuité, $|d_t|$ prend au moins toutes les valeurs dans $[0, |u - z|]$. Par ailleurs,

$$\left|\widehat{P}(z + d_t)\right| = \prod_i |(z + d_t) - z_i| \geqslant \prod_i \big|\,|d_t| - |z - z_i|\,\big|,$$

soit, pour tout $d \in [0, |u - z|]$,

$$\prod_i (d - |z - z_i|) \leqslant \varepsilon^n(\rho + d)^n \leqslant \varepsilon^n(\rho + |u - z|)^n.$$

D'après la propriété de minimax des polynômes de Chebyshev (Lemme 11.5), il existe d dans $[0, |u-z|]$ tel que le membre de gauche soit $\geqslant (|u-z|/4)^n$ et le résultat suit. □

Lemme 11.5. — *Soit* $T_n(X)$ *(*$= \cos(n \arccos(X))$ *pour* $X \in [-1,1]$*) le* n*-ème polynôme de Chebyshev et* $\mathcal{P}$ *l'ensemble des polynômes unitaires de* $\mathbb{R}[X]$ *de degré* $\leqslant n$*. Alors* $t_n := 2^{1-n}T_n \in \mathcal{P}$ *et, pour tout* $P \in \mathcal{P}$*, on a*

$$\max_{x\in[a,b]} |P(x)| \geqslant \left(\frac{b-a}{2}\right)^n \max_{x\in[-1,1]} |t_n(x)| = 2\left(\frac{b-a}{4}\right)^n.$$

Démonstration. — t_n est le polynôme unitaire de degré n de plus petite norme uniforme sur $[-1,1]$, qui vaut 2^{1-n} (voir par exemple [**14**, § 3.1]). Grâce à l'application linéaire $x \mapsto \frac{a+b}{2} + x\frac{b-a}{2}$ de $[-1,1]$ dans $[a,b]$, on écrit

$$\max_{x\in[a,b]} |P(x)| = \Big(\frac{b-a}{2}\Big)^n \max_{x\in[-1,1]} |Q(x)|, \quad \text{où } Q(X) := P\Big(X + \frac{a+b}{a-b}\Big).$$

□

11.7. Factorisation dans $\mathbb{R}[X]$. — Une factorisation approchée dans $\mathbb{R}[X]$ s'obtient à partir d'une factorisation dans $\mathbb{C}[X]$ suffisamment précise pour identifier les paires de racines conjuguées. Le cas particulier de la recherche de *racines* réelles d'un polynôme de $\mathbb{R}[X]$ est important, mais nettement plus simple puisque les méthodes de dichotomie sont directement disponibles (voir Rouillier-Zimmermann [**33**]). L'algorithme de Sturm ([**10**, § 4.1]) fournit le *nombre* de racines réelles dans un intervalle $]a,b]$.

12. Factorisation dans $\mathbb{Q}_p[X]$

Nous examinons maintenant un algorithme numérique p-adique typique, fondé sur le lemme de Hensel. L'avantage sur les méthodes complexes est que les résultats de perturbations sont en général évidents, et qu'il n'y a pas d'erreur d'arrondi. Par contre, les divisions par p induisent une perte de précision, et le champ d'application est limité aux contextes algébriques.

Le lemme de Hensel permet de relever une factorisation suffisamment précise pour être sans facteurs carrés. Comment obtenir cette dernière ? On va déjà voir comment procéder modulo p, c'est-à-dire dans $\mathbb{F}_p[X]$, et considérer d'abord le cas particulier de la recherche de racines dans $\mathbb{F}_p$. Ces méthodes s'adaptent sans mal au cas d'un corps fini $\mathbb{F}_q$ général, avec un petit effort supplémentaire en caractéristique 2.

12.1. Principe. — Pour factoriser un élément N d'un anneau euclidien R, on recherche (ou on fabrique) un diviseur de 0 dans $R/(N)$, c'est-à-dire un $\overline{a}$ non nul tel que $m_{\overline{a}}$ soit non injective. Alors, si a relève $\overline{a}$, $d := \operatorname{pgcd}(a, N)$ est un facteur non trivial de N, fourni par l'algorithme d'Euclide. On invoque ensuite récursivement l'algorithme sur chacun des deux facteurs d et N/d. Par exemple, la plupart des méthodes modernes de factorisation dans $\mathbb{Z}$ construisent une identité $x^2 \equiv y^2 \pmod N$, d'où on tire un diviseur de 0 potentiel $x - y$. (Les mauvais jours, $x = \pm y$ et on a travaillé pour rien.)

12.2. Racines dans $\mathbb{F}_p[X]$. — Soit donc $T \in \mathbb{F}_p[X]$, qu'on peut supposer scindé à racines simples en le remplaçant par $\operatorname{pgcd}(T, X^p - X)$ (§ 10.2). On pose $n = \deg T$. On veut trouver un diviseur de 0 dans la $\mathbb{F}_p$-algèbre $\mathbb{F}_p[X]/(T) \simeq (\mathbb{F}_p)^n$, dont tous les éléments vérifient

$$0 = x^p - x = \prod_{i \in \mathbb{F}_p} (x - i).$$

Calculer le produit partiel jusqu'à trouver 0 n'est pas un bon algorithme si $n \ll p$: au pire p multiplications, pour une taille $n \log p$. Si $p = 2$, cette méthode est satisfaisante. Sinon, on choisit x au hasard dans $\mathbb{F}_p[X]/(T) \simeq (\mathbb{F}_p)^n$ et on écrit

$$0 = x(x^t - 1)(x^t + 1), \quad t := (p-1)/2.$$

Si les trois termes sont non nuls, on a obtenu un diviseur de 0. Les mauvais cas correspondent à $x = 0$ (1 cas), $x^t - 1 = 0$ (t^n cas), ou $x^t + 1 = 0$ (t^n cas). On est donc malchanceux avec probabilité

$$\frac{1 + 2t^n}{p^n} \leqslant \frac{1}{2^{n-1}} \leqslant \frac{1}{2}, \quad \text{si } n \geqslant 2.$$

L'espérance du nombre d'essais avant de trouver un facteur de T est inférieure à 2 : c'est un bon algorithme.

12.3. Factorisation dans $\mathbb{F}_p[X]$. — Soit $T \in \mathbb{F}_p[X]$, qu'on peut supposer sans facteurs carrés : si $T' = 0$, alors $T = Q(X^p) = Q(X)^p$ et on remplace T par Q, sinon $D = \mathrm{pgcd}(T, T')$ est un diviseur strict de T ; on remplace successivement T par T/D (sans facteurs carrés) puis D (de degré $< \deg T$).

Soit ϕ l'endomorphisme de la $\mathbb{F}_p$-algèbre

$$\mathbb{F}_p[X]/(T) \simeq \prod_{i=1}^{r} \mathbb{F}_p[X]/(T_i), \quad \text{si } T = \prod_{i=1}^{r} T_i,$$

donné par $a \mapsto a^p - a$ et $V := \operatorname{Ker} \phi \simeq (\mathbb{F}_p)^r$. Alors $\dim_{\mathbb{F}_p} V = r$ est le nombre de diviseurs irréductibles de T. Si $r = 1$, T est irréductible ; sinon, soit $\alpha \in V \smallsetminus \mathbb{F}_p$ et $P_{\min,\alpha}$ son polynôme minimal, qui est scindé à racines simples, de degré $\geqslant 2$ puisque $\alpha \notin \mathbb{F}_p$. Si $z \in \mathbb{F}_p$ est une des racines (§ 12.2), $\alpha - z$ est un diviseur de 0 de $\mathbb{F}_p[X]/(T)$.

12.4. L'algorithme Round 4

Définition 12.1

– Un polynôme unitaire $Q \in \mathbb{Z}_p[X]$ est *de type Eisenstein* si

$$Q = b^k + p(qb + r),$$

où $b, q, r \in \mathbb{Z}_p[X]$, b est irréductible modulo p, $\deg r < \deg b$ et r est non nul modulo p.

– Soit $T \in \mathbb{Z}_p[X]$ unitaire ; on dit que $\alpha \in \mathbb{Q}_p[X]$ *certifie* (l'irréductibilité de) T si le polynôme caractéristique de $\overline{\alpha}$ dans $\mathbb{Q}_p[X]/(T)$ est de type Eisenstein.

Par exemple, un polynôme d'Eisenstein est de type Eisenstein avec $b = X$. Un polynôme de type Eisenstein est irréductible ; en effet, une décomposition serait de la forme $T = (b^{k_1} + pa_1)(b^{k_2} + pa_2)$, ce qui interdit $r \not\equiv 0 \pmod p$ si $k_1 k_2 \neq 0$.

***Proposition 12.2* (Zassenhaus, *cf.* Ford-Pauli-Roblot [16])**

(1) *T est irréductible dans $\mathbb{Q}_p[X]$ si et seulement s'il existe $\alpha \in \mathbb{Q}_p[X]$ certifiant T.*

(2) *Supposons qu'il existe $\overline{\alpha}$ dans la $\mathbb{Q}_p$-algèbre $\mathbb{Q}_p[X]/(T)$ dont le polynôme caractéristique $P_{\mathrm{char},\overline{\alpha}}$ appartient à $\mathbb{Z}_p[X]$ et admet deux facteurs irréductibles distincts modulo p. Alors T est réductible dans $\mathbb{Q}_p[X]$.*

Démonstration (de (2)*).* — Soit $\alpha \in \mathbb{Q}_p[X]$ relevant $\overline{\alpha}$. Par le lemme de Hensel, on peut écrire $P_{\mathrm{char},\overline{\alpha}} = P_1 P_2$ dans $\mathbb{Z}_p[X]$. Alors $T = \mathrm{pgcd}(T, P_1(\alpha)) \cdot \mathrm{pgcd}(T, P_2(\alpha))$ est une décomposition non triviale. □

La suffisance dans (1) est évidente : si T est réductible, l'algèbre $\mathbb{Q}_p[X]/(T)$ est une somme directe et les polynômes caractéristiques de tous ses éléments sont réductibles. La réciproque est plus technique (il faut étudier les extensions finies de $\mathbb{Q}_p$, en particulier les extensions non ramifiées), mais constructive. L'algorithme Round 4 (Zassenhaus) s'efforce de construire α certifiant T, en calculant successivement les chiffres de son développement p-adique. S'il échoue, il produit une factorisation dans $\mathbb{Q}_p[X]$ comme indiqué dans la démonstration de (2), puis essaie récursivement de certifier les facteurs.

L'algorithme précis est technique (voir [**16**]) : les calculs sont effectués à précision finie, c'est-à-dire dans $\mathbb{Z}/p^k\mathbb{Z}$ et non dans $\mathbb{Z}_p$. Donc il faut borner tous les dénominateurs susceptibles d'apparaître pour choisir une valeur de k suffisante pour maîtriser les erreurs d'arrondi.

13. Factorisation dans $\mathbb{Q}[X]$

13.1. L'algorithme naïf. — Après ce qui précède, il suffit de connaître une borne sur la taille des facteurs pour obtenir un algorithme de factorisation dans $\mathbb{Q}[X]$. Il y a plusieurs possibilités, par exemple [**5, 29**]. La plus simple est la suivante :

***Théorème 13.1* (Landau).** — *Si $F \in \mathbb{C}[X]$ est de degré m et $A \mid F$, alors $\|A\|_\infty \leqslant 2^m \|F\|_2$.*

Démonstration. — On factorise $F = a_m \prod_{i \leqslant m}(X - \alpha_i)$ et on définit la mesure de Mahler :

$$M(F) = |a_m| \prod_{i\,:\,|\alpha_i|>1} |\alpha_i| \,.$$

La majoration $\|A\|_\infty \leqslant 2^m M(F)$ suit de l'expression des coefficients de A comme fonction symétrique de ses racines (qui sont racines de F). Un calcul immédiat montre que, pour tout $\alpha \in \mathbb{C}$ et $C \in \mathbb{C}[X]$, les polynômes $(X - \alpha)C$ et $(\overline{\alpha}X - 1)C$ ont même norme L^2. Donc F à la même norme que

$$|a_m| \prod_{i\colon |\alpha_i|>1} (X - \alpha_i) \prod_{i\colon |\alpha_i|\leqslant 1} (\overline{\alpha_i} X - 1),$$

dont on minore la norme par la valeur absolue $M(F)$ de son coefficient constant. Soit $M(F) \leqslant \|F\|_2$. □

Soit donc $F \in \mathbb{Q}[X]$ que l'on désire factoriser. On peut supposer successivement que F est dans $\mathbb{Z}[X]$, unitaire (par changement de variable ; le lemme de Gauss dit alors que ses facteurs unitaires sont dans $\mathbb{Z}[X]$), et sans facteurs carrés. Le théorème fournit $C > 0$ tel que tous les diviseurs de F ont leurs coefficients dans $[-C, C]$.

On cherche un premier p tel que F modulo p reste sans facteurs carrés dans $\mathbb{F}_p[X]$. Il y a un nombre fini de premiers à éviter : les diviseurs du résultant de F et F', qui est non nul[(6)]. On peut alors factoriser F dans $\mathbb{Z}_p[X]$ grâce au lemme de Hensel (faible), c'est-à-dire écrire,

$$F \equiv \prod_{i=1}^{r} F_i \pmod{p^k \mathbb{Z}_p[X]}.$$

où les $F_i \in \mathbb{Z}_p[X]$ sont irréductibles modulo p et connus modulo p^k, et où on choisit k tel que $p^k > 2C$.

Les diviseurs de F dans $\mathbb{Q}[X]$ sont de la forme $F_S := \prod_{i\in S} F_i$ où $S \subset \{1, \dots, r\}$. On calcule une approximation $\widehat{F_S} \equiv F_S \pmod{p^k}$, $\widehat{F_S} \in \mathbb{Z}[X]$, dont les coefficients sont dans $]-p^k/2, p^k/2]$. D'après le théorème de Landau, $F_S \in \mathbb{Z}[X]$ si et seulement si $F_S = \widehat{F_S}$ et il suffit de tester si $\widehat{F_S}$ divise F dans $\mathbb{Z}[X]$ pour le déterminer.

Jusqu'à présent, tout ceci est un bon algorithme. Mais il reste 2^r possibilités pour S, et il est facile de construire des exemples où $r \geqslant \frac{1}{2} \deg F$, quel que soit p. Par exemple, le polynôme minimal de $\sqrt{2} + \sqrt{3} + \cdots + \sqrt{p_k}$, où p_k est le k-ème nombre premier, est irréductible de degré 2^k, définit une extension galoisienne de groupe de Galois $(\mathbb{Z}/2\mathbb{Z})^k$, et le théorème de Chebotarëv (Théorème 7.3) implique qu'il se décompose dans $\mathbb{F}_p[X]$ en produit de facteurs tous linéaires ou

(6) Mieux, il est bornée polynomialement en terme des données, il existe donc un petit nombre premier qui ne le divise pas : le théorème des nombres premiers (Théorème 7.2 pour $K = \mathbb{Q}$) implique que $\sum_{p\leqslant x} \log p \sim x$ donc un entier divisible par tous les premiers $< x$ est minoré par $\exp(x - o(x))$.

tous quadratiques, avec probabilité 2^{-k} et $1 - 2^{-k}$ respectivement. Par contre, si le groupe de Galois d'un corps de décomposition K de F contient une classe de conjugaison favorable (un produit d'un petit nombre de cycles), et si Δ_K est raisonnablement petit, le § 7.3 indique qu'on trouvera rapidement un premier p qui l'exhibe, ou bien un contre-exemple à GRH.

13.2. LLL et l'algorithme de van Hoeij. — Il existe un bon algorithme de factorisation dans $\mathbb{Q}[X]$, inconditionnel, qui est l'application originelle de l'algorithme LLL [**27**]. On considère un facteur p-adique F_1 de F, une approximation $\widehat{F_1} \in \mathbb{Z}[X]$ de F_1, modulo p^K, et on cherche un facteur strict de F dans $\mathbb{Z}[X]$ que F_1 divise. Pour ceci, on considère le réseau de $\mathbb{R}_{\deg F - 1}[X] \simeq \mathbb{R}^{\deg F}$ engendré par

$$\widehat{F_1}\left\{1, X, \ldots, X^{\deg F - \deg F_1 - 1}\right\} \cup p^K \left\{1, X, \ldots, X^{\deg F - 1}\right\},$$

qui contient tous les polynômes de $\mathbb{Z}[X]$ de degré $< \deg F$ qui sont multiples de F_1. Soit (b_i) une base LLL-réduite de ce réseau. Si p^K est assez grand, on démontre que, soit $\mathrm{pgcd}(b_1, F)$ est un facteur non trivial de F, soit ce dernier est irréductible dans $\mathbb{Q}[X]$.

L'algorithme de van Hoeij [**22, 7**] est un autre bon algorithme, qui utilise aussi LLL, mais sur un réseau plus agréable. On considère le morphisme injectif

$$\begin{aligned} \Phi : \mathbb{Q}_p(X)^*/\mathbb{Q}_p^* &\longrightarrow \mathbb{Q}_p(X) \\ g &\longmapsto F \cdot \frac{g'}{g} \end{aligned}$$

Si g appartient au sous-groupe G_p engendré par les F_i, alors $\Phi(g) \in \mathbb{Z}_p[X]$. Si g appartient au sous-groupe $G \subset G_p$ engendré par les facteurs dans $\mathbb{Q}[X]$, alors $\Phi(g) \in \mathbb{Z}[X]$.

On détermine G à partir d'une base LLL-réduite du réseau engendré par les $\Phi(F_i) \pmod{p^K}$ et les $p^K X^i$. Plus précisément, on dispose d'une borne sur $\|\Phi(g)\|_\infty$ si g est un facteur de F dans $\mathbb{Q}[X]$; les vecteurs de la base trop longs pour provenir de tels facteurs rationnels sont éliminés ; si K est suffisamment grand, les éléments restants de la base sont les générateurs de $\Phi(G)$. On en déduit G, puis les facteurs.

L'intérêt par rapport à l'algorithme précédent est qu'il existe une base LLL-réduite donnée par une matrice de changement de base U dont les coordonnées sont dans $\{0, 1\}$ (ce sont essentiellement les ε_i

tels que $\prod F_i^{\varepsilon_i} \in \mathbb{Z}[X]$). On peut espérer détecter ces vecteurs, ou l'irréductibilité, à une précision p^K bien inférieure à la précision théorique. Ce que le premier algorithme a peu de chance de faire puisque, dans ce cas, les coefficients de U sont ceux d'un facteur, donc potentiellement gigantesques.

13.3. Factorisation dans $K[X]$. — Les idées du paragraphe précédent se généralisent à la factorisation dans $K[X]$, où $K = \mathbb{Q}(\theta)$ est un corps de nombres, ainsi d'ailleurs qu'a la factorisation dans $k[X, Y]$ où k est un corps tel que l'on dispose d'un algorithme de factorisation dans $k[X]$, par exemple un corps fini.

Une méthode plus simple pour factoriser $F \in K[X]$, quoique moins efficace en général, se ramène au cas sans facteurs carrés en remplaçant F par $F/\operatorname{pgcd}(F, F')$, puis considère $F_\lambda = F(X + \lambda\theta)$ pour $\lambda \in \mathbb{Z}$, et $\mathrm{N}(F_\lambda) := \prod_\sigma \sigma(F_\lambda) \in \mathbb{Q}[X]$, où σ parcourt les plongements complexes de K. Pour tout λ sauf un nombre fini (et facilement borné), $\mathrm{N}(F_\lambda)$ est san s facteurs carrés. On choisit un tel λ et on factorise $\mathrm{N}(F_\lambda) = \prod f_i$ dans $\mathbb{Q}[X]$, où les f_i sont irréductibles et distincts. Alors les facteurs irréductibles de F dans $K[X]$ sont les $\operatorname{pgcd}(f_i(X - \lambda\theta), F)$ dans $K[X]$. Voir [**10**, § 3.6] pour une démonstration simple. C'est un bon algorithme si on utilise un bon algorithme de factorisation dans $\mathbb{Q}[X]$. Contrairement aux apparences, ces calculs sont exacts : θ est une variable formelle, classe de Y modulo $T(Y)$ dans $K = \mathbb{Q}[Y]/(T)$; les pgcds se calculent donc par l'algorithme d'Euclide et $\mathrm{N}(F_\lambda) = \operatorname{Res}_Y(F(X + \lambda Y), T(Y)$ est un simple résultant.

Grâce à ce qui précède, on peut maintenant

– calculer $\operatorname{Aut} K$ (trouver les facteurs de degré un de T dans $K[X]$),

– tester si $K \subset L$, plus précisément s'il existe un plongement de $\mathbb{Q}[X]/(T)$ dans le corps de nombres L (T a-t-il une racine dans $L[X]$?),

– tester si $K \simeq L$ (tester si $\deg T = \dim_{\mathbb{Q}} L$ et appliquer le point précédent).

14. Ordres

Il est difficile de calculer $\mathcal{O}_K$ si on ne connaît pas la factorisation de Δ_K. Or, la factorisation dans $\mathbb{Z}$ est difficile. Heureusement, pour de nombreuses applications, on n'a pas besoin de $\mathcal{O}_K$, mais seulement d'un sous-anneau qui en est une approximation raisonnable. On peut citer la décomposition des nombres premiers, la factorisation dans $K[X]$ donc les problèmes d'isomorphisme ou du sous-corps (a-t-on $L \subset K$?), la détermination de l'ensemble des sous-corps, du groupe de Galois de $K/\mathbb{Q}$...

14.1. Définition. — Un *ordre*[7] de K est un sous-anneau, donc contenant 1, qui est un $\mathbb{Z}$-module de rang $n = \dim_{\mathbb{Q}} K$ (de façon équivalente, $\operatorname{Frac}\mathcal{O} = K$). Les ordres de K sont partiellement ordonnés par inclusion.

Proposition 14.1. — *L'anneau $\mathcal{O}_K$ des entiers algébriques de K est son ordre maximal.*

Démonstration. — $\mathcal{O}_K$ est un ordre. Soit $\mathcal{O}$ un ordre de K et $\alpha \in \mathcal{O}$. Alors $\mathbb{Z}[\alpha] \subset \mathcal{O}$ est de type fini, donc α est entier. $\square$

En un certain sens, la notion d'ordre est duale de celle de localisé : un ordre est un sous-anneau de $\mathcal{O}_K$ de type fini comme $\mathbb{Z}$-module, mais non intégralement clos (si $\mathcal{O} \subsetneq \mathcal{O}_K$). Un localisé $S^{-1}\mathcal{O}_K = \{a/s \colon a \in \mathcal{O}_K, s \in S\}$ est un sur-anneau intégralement clos, non de type fini (si $S \not\subset U(K)$). Dans le langage de la géométrie algébrique, cette notion correspond à celle de courbe singulière : sur la courbe $(\operatorname{spec}\mathcal{O}, \mathcal{O})$, un nombre fini de localisés $\mathcal{O}_{\mathfrak{p}}$ ne sont pas des anneaux de valuation discrète.

Les ordres non maximaux sont de braves anneaux intègres (noethériens, de dimension 1) de corps de fractions K, mais ils gardent des comportement pathologiques par rapport à l'anneau de Dedekind $\mathcal{O}_K$. Par exemple, la norme $\mathrm{N}\,\mathfrak{a} := \mathcal{O}/\mathfrak{a}$ d'un idéal non nul est bien définie mais n'est plus multiplicative, il existe des idéaux non inversibles, $\mathfrak{a} \subset \mathfrak{b}$ n'implique pas $\mathfrak{b} \mid \mathfrak{a}$ (il existe un $\mathcal{O}$-idéal $\mathfrak{c}$ tel que $\mathfrak{a} = \mathfrak{b}\mathfrak{c}$).

14.2. Construction. — L'intérêt de la notion est qu'il est facile de construire des ordres. Par exemple, si $T \in \mathbb{Z}[X]$ est *unitaire*, alors

$$\mathcal{O}_T := \left\langle 1, \theta, \ldots, \theta^{n-1} \right\rangle_{\mathbb{Z}} = \mathbb{Z}[\theta]$$

[7] Introduit par Dedekind (*Ordnung*) comme anneau de stabilisateurs $(A : A)$ d'un $\mathbb{Z}$-module de type fini A. Généralise la notion introduite par Gauss pour les formes quadratiques binaires de discriminant fixé (*ordo,-inis* = rangée, file).

est un ordre. Plus généralement.

***Théorème 14.2* (Dedekind).** — *Soit $T(X) = a_0X^n + \cdots + a_n \in \mathbb{Z}[X]$; on définit une suite (T_i) de polynômes de degré $i < n$:*

$$T_0 := a_0 \quad \textit{et} \quad T_{i+1} := XT_i + a_{i+1}, \ \textit{pour } 0 \leqslant i < n.$$

Alors $\mathcal{O}_T := \langle 1, T_1(\theta), \ldots, T_{n-1}(\theta)\rangle_{\mathbb{Z}}$ est un ordre.

La suite (T_i) est un schéma d'évaluation de Horner pour T, en particulier $T_n = T$. Si $a_0 = \pm 1$, on retrouve l'exemple précédent.

Démonstration. — Posons $t_i := T_i(\theta)$ pour $0 \leqslant i \leqslant n$. Pour $1 \leqslant i \leqslant j < n$, on a

$$t_it_j = (\theta t_{i-1} + a_i)t_j = a_it_j + t_{i-1}(t_{j+1} - a_{j+1}) \equiv t_{i-1}t_{j+1} \mod \mathcal{O}.$$

En itérant, on se ramène au cas de t_it_n ou t_0t_j ; le premier est nul, le deuxième égal à $a_0t_j \in \mathcal{O}$. Donc ce $\mathbb{Z}$-module de rang au plus n est un anneau. Il contient $a_0\theta$, donc on a bien $\operatorname{Frac}\mathcal{O} = K$. □

Plus généralement, les ordres apparaissent comme « anneaux de stabilisateurs ». Si A et B sont deux sous-$\mathbb{Z}$-modules de K, on note

$$(A : B) := \{\alpha \in K, \alpha B \subset A\},$$

appelé transporteur de B dans A, ou « A divisé par B ». C'est un $\mathbb{Z}$-module. Si $\mathcal{O} = \mathcal{O}_K$ est de Dedekind, et $\mathfrak{a}$, $\mathfrak{b}$ sont des idéaux fractionnaires, on a $(\mathfrak{a} : \mathfrak{b}) = \mathfrak{a}\mathfrak{b}^{-1}$.

***Proposition 14.3*.** — *Si $A \subset K$ est un $\mathbb{Z}$-module de rang n, $\mathcal{O} = (A : A)$ est un ordre.*

Démonstration. — C'est évidemment un anneau. Tout $\alpha \in \mathcal{O}$ est entier (A est un $\mathbb{Z}$-module de type fini tel que $\alpha A \subset A$!), donc $\mathcal{O} \subset \mathcal{O}_K$ est de type fini. Comme $\mathcal{O}_K A$ est de type fini et A contient une $\mathbb{Q}$-base de K, il existe $d \in \mathbb{N}_{>0}$ tel que $d\mathcal{O}_K A \subset A$, donc $\mathcal{O} \supset d\mathcal{O}_K$ est de rang n. □

***Exemple*.** — On pose $A := \left\langle 1, \theta, \ldots, \theta^{n-1}\right\rangle_{\mathbb{Z}}$ ($\neq \mathbb{Z}[\theta]$ si T n'est pas unitaire !). Alors $(A : A) = \mathcal{O}_{T/\delta}$, où δ est le pgcd des coefficients de T.

14.3. Manipulation des ordres, dénominateurs. — Un ordre $\mathcal{O}$ est donné par une base (w_i) et la table de multiplication correspondante $w_i w_j = \sum a_{i,j,k} w_k$, où $a_{i,j,k} \in \mathbb{Z}$. On donne un élément de $\mathcal{O}$ par ses coordonnées dans la base (w_i), un sous-$\mathbb{Z}$-module de $\mathcal{O}$ par une $\mathbb{Z}$-base, c'est-à-dire par une matrice à coefficients entiers. Si on veut une représentation unique, la HNF donne une base canonique ((w_i) étant fixée).

Pour représenter $x \in K$, on choisit un entier $d \in \mathbb{N}_{>0}$, unique si on l'impose minimal, tel que $dx \subset \mathcal{O}$, et on est ramené au cas précédent. Un sous $\mathbb{Z}$-module de type fini M de K, par exemple un idéal fractionnaire, se manipule de même. Jusqu'à présent, avec l'écriture $K = \mathbb{Q}[X]/(T)$, T unitaire, nous avons ainsi utilisé l'ordre $\mathcal{O} = \mathcal{O}_T$, qui admet la représentation alternative économique $\mathbb{Z}[X]/(T)$.

Choisir un autre ordre que $\mathcal{O}_T$ est souvent plus flexible. Par exemple, pour calculer $\mathcal{O}_K$ (§ 15.1). Pour minimiser les dénominateurs, on choisit $\mathcal{O} = \mathcal{O}_K$, qui n'est en général pas de la forme $\mathcal{O}_T$ (§ 15.4). Si on exprime $\alpha \in \mathcal{O}_K$ dans une $\mathbb{Z}$-base de $\mathcal{O}$, son dénominateur divise l'exposant du groupe additif $\mathcal{O}_K/\mathcal{O}$.

15. L'ordre maximal $\mathcal{O}_K$

Définition 15.1. — Soit m un entier. Un ordre $\mathcal{O} \subset \mathcal{O}_K$ est dit *m-maximal* si m et l'indice $[\mathcal{O}_K : \mathcal{O}]$ sont premiers entre eux.

Le calcul de $\mathcal{O}_K$ est un problème local : à partir d'un ordre $\mathcal{O}$, il suffit de calculer un ordre p-maximal $\mathcal{O}_p \supset \mathcal{O}$ pour chaque p premier, et $\mathcal{O}_K = \sum_p \mathcal{O}_p$ (puisque ce dernier est un ordre p-maximal pour tout p). On peut prendre $\mathcal{O}_p = \mathcal{O}$ si $p^2 \nmid \Delta_{\mathcal{O}}$ (Proposition 5.7), donc il n'y a qu'un nombre fini de premiers à traiter. D'un point de vue algorithmique, la vraie obstruction est globale :

***Théorème 15.2* (Chistov).** — *Les deux problèmes suivants sont de même difficulté :*

– *étant donné un corps de nombres K, trouver $\mathcal{O}_K$,*

– *étant donné un entier D, trouver le plus grand entier $d \mid D$ qui soit sans facteurs carrés.*

Le deuxième problème est actuellement aussi difficile que la factorisation dans $\mathbb{Z}$, donc calculer $\mathcal{O}_K$ est difficile. Pour la même raison, tester si un ordre donné est maximal est difficile : il faut décider si un entier est sans facteurs carrés. Pour s'en convaincre, considérer le cas particulier $\mathbb{Z}[\sqrt{D}] \subset \mathbb{Q}(\sqrt{D})$. Examinons maintenant le problème local.

15.1. L'algorithme Round 2. — Soit $\mathcal{O}$ un ordre et p un nombre premier. On désire calculer l'ordre p-maximal R, $\mathcal{O} \subset R \subset \mathcal{O}_K$, tel que $[R : \mathcal{O}]$ soit une puissance de p. Explicitement

$$R := \{x \in \mathcal{O}_K,\ p^n x \in \mathcal{O} \text{ pour } n \gg 1\}.$$

On note $I_p = \operatorname{rad}(p\mathcal{O})$ l'idéal radical de $p\mathcal{O}$: par définition $I_p/p\mathcal{O}$ est le nilradical (l'idéal des nilpotents) de l'anneau fini $\mathcal{O}/p\mathcal{O}$ et I_p en est le relèvement dans $\mathcal{O}$. C'est aussi l'intersection des idéaux premiers de $\mathcal{O}$ contenant p,

$$I_p := \bigcap_{\mathfrak{p}\,:\, p\in\mathfrak{p}} \mathfrak{p} = \prod_{\mathfrak{p}\,:\, p\in\mathfrak{p}} \mathfrak{p}.$$

Ces idéaux premiers sont en nombre fini et ils sont maximaux (dans $\mathcal{O}$).

Proposition 15.3. — *Il est facile de calculer I_p.*

Démonstration. — Si t est un entier tel que $p^t \geqslant n = \dim_{\mathbb{Q}} K$, alors $\mathcal{O}/p\mathcal{O}$ est un $\mathbb{F}_p$-ev de dimension n et $I_p/p\mathcal{O}$ est le noyau[8] de l'application $\mathbb{F}_p$-linéaire $x \mapsto x^{p^t}$. □

L'algorithme Round 2 est un bijou, issu du théorème suivant :

***Théorème 15.4* (Zassenhaus).** — *Soit $\mathcal{O}' := (I_p : I_p)$. L'ordre $\mathcal{O}$ est p-maximal si et seulement si $\mathcal{O} = \mathcal{O}'$. Dans le cas contraire, on a $p \mid [\mathcal{O}' : \mathcal{O}] \mid p^n$.*

---

[8] Si $p > n$, c'est aussi le noyau de

$$\begin{aligned} \mathcal{O}/p\mathcal{O} &\longrightarrow \operatorname{Hom}(\mathcal{O}/p\mathcal{O}, \mathbb{F}_p) \\ x &\longmapsto (y \mapsto \operatorname{Tr}(xy)) \end{aligned}$$

Démonstration. — Soit u un endomorphisme d'un $\mathbb{F}_p$-ev de dimension finie $n < p$; si $\operatorname{Tr}(u^k) = 0$ pour tout $k \geqslant 1$, alors u est nilpotent (formules de Newton + Cayley-Hamilton par exemple). Donc un élément du noyau appartient à I_p (en choisissant $y = 1, x, x^2, \ldots$). Réciproquement, un nilpotent est de trace nulle. □

Cette deuxième description est préférable quand p est grand : on supprime le coût de la mise à la puissance p.

Idée. — On veut vérifier que les maximaux contenant p sont inversibles ; il suffit de vérifier que leur produit l'est.

Démonstration. — $\mathcal{O}'$ est un ordre contenant $\mathcal{O}$ (car I_p est un $\mathcal{O}$-idéal, Proposition 14.3) ; comme $p \in I_p$, on obtient $p\mathcal{O}' \subset I_p \subset \mathcal{O}$ donc $[\mathcal{O}' : \mathcal{O}] \mid p^n$. D'où $\mathcal{O} = \mathcal{O}'$ si $\mathcal{O}$ est p-maximal.

Réciproquement, si $\mathcal{O} = \mathcal{O}'$, soit R l'ordre p-maximal contenant $\mathcal{O}$ du début de la section. Comme I_p et R sont de type fini, on a $I_p^m \subset p\mathcal{O}$ et $p^m R \subset \mathcal{O}$ pour $m \gg 1$. Donc $RI_p^m \subset \mathcal{O}$ pour m assez grand. Par l'absurde, supposons qu'il existe $m \geqslant 0$ tel que $RI_p^m \not\subset \mathcal{O}$ et choisissons m maximal, puis $\alpha \in RI_p^m \smallsetminus \mathcal{O}$. Alors $\alpha I_p \subset \mathcal{O}$ et donc $\alpha I_p \subset I_p$ (comme I_p est de type fini, il existe k tel que $I_p^k \subset p\mathcal{O}$), soit $\alpha \in (I_p : I_p) = \mathcal{O}' = \mathcal{O}$. Absurde. □

L'algorithme de normalisation est immédiat : on part de $\mathcal{O} = \mathcal{O}_T$; on calcule $I_p/p\mathcal{O}$ et on remplace $\mathcal{O}$ par $\mathcal{O}' := (I_p : I_p)$ tant que $\mathcal{O} \neq \mathcal{O}'$. C'est un bon algorithme, qu'on applique à tous les p premiers tels que $p^2 \mid \Delta_{\mathcal{O}}$. À condition de les connaître !

Il est remarquable que l'on puisse remplacer p par un entier m sans facteurs carrés arbitraire (Buchmann-Lenstra [**9**]). Soit l'algorithme rencontre une impossibilité qui exhibe un facteur de m (un élément non nul mais non inversible de $\mathbb{Z}/m\mathbb{Z}$), soit il produit un ordre m-maximal.

15.2. Le cas particulier de $\mathcal{O}_T$. — Pour chaque premier p, la première étape de l'algorithme Round 2 se traduit plus efficacement par (revoir le Théorème 4.10) :

Théorème 15.5 **(Dedekind).** — *Soit $K = \mathbb{Q}(X)/(T)$, $T \in \mathbb{Z}[X]$ unitaire et $\theta = X \pmod T$, tel que*
$$T \equiv \prod_i P_i^{e_i} \pmod{p\mathbb{Z}[X]},$$
où les P_i sont unitaires, irréductibles et 2 à 2 distincts modulo p. Soit
$$f := \prod P_i, \quad g := \prod P_i^{e_i - 1}, \quad h := (T - fg)/p \in \mathbb{Z}[X].$$
Soit $\delta := \operatorname{pgcd}(\overline{f}, \overline{g}, \overline{h})$ dans $\mathbb{F}_p[X]$, et U un relèvement de $\overline{T}/\delta$, alors
$$\mathcal{O}' = \mathcal{O}_T + \frac{U(\theta)}{p}\mathcal{O}_T.$$
En particulier, $\mathcal{O}_T$ est p-maximal si et seulement si $\delta = 1$.

Démonstration. — On trouve $I_p = p\mathcal{O}_T + g(\theta)\mathcal{O}_T$, et on explicite le calcul de $(I_p : I_p)$. □

Corollaire 15.6. — *Si $T = b^k + p(qb + r)$ est de type Eisenstein, alors $\mathcal{O}_T$ est p-maximal.*

Démonstration. — $f = b$, $g = b^{k-1}$, $h = qb + r$. On a $\operatorname{pgcd}(\overline{f}, \overline{g}, \overline{h}) = (\overline{b}, \overline{r}) = 1$. □

15.3. L'algorithme Round 4. — Round 2 nécessite quand même la manipulation de matrices $n \times n \times n$ (§ 14.3) et de nombreuses itérations quand $v_p([\mathcal{O}_K : \mathcal{O}])$ est grand ($[\mathcal{O}' : \mathcal{O}] \mid p^n$, donc la valuation de l'indice diminue au plus de n à chaque itération). En pratique, on utilise un autre algorithme de normalisation locale, lié à la factorisation dans $\mathbb{Q}_p[X]$ par l'algorithme Round 4. Par rapport à Round 2, il est d'autant plus intéressant que n ou $v_p([\mathcal{O}_K : \mathcal{O}])$ est grand, mais de description moins attrayante. Je ne sais pas si l'on peut obtenir les même garanties pour Round 4 que pour l'algorithme Round 2 modifié par Buchmann et Lenstra, où p est remplacé par m sans facteurs carrés. C'est probable.

On utilise la généralisation suivante du Théorème 4.10, auquel il se réduit si T est sans facteurs carrés modulo p :

Théorème 15.7. — *Soit $T \in \mathbb{Z}[X]$ unitaire irréductible dans $\mathbb{Q}[X]$. Soit $T = F_1 \cdots F_g$ la factorisation de F en produit de polynômes unitaires irréductibles dans $\mathbb{Z}_p[X]$ (ils sont distincts), et $\alpha_i \in \mathbb{Q}_p[X]$ certifiant F_i pour $i = 1, \ldots, g$. Soit p^d un dénominateur commun aux α_i ; pour $x \in \mathbb{Q}_p[X]$, on définit une approximation $\widehat{x} \in \mathbb{Q}[X]$ telle que $x - \widehat{x} \in p^{d+1}\mathbb{Z}_p[X]$ et sa projection $\overline{x}$ dans $\mathbb{Q}[X]/(T)$. Alors, pour tout i,*

- *$F_i \equiv t_i^{e_i} \pmod{p\mathbb{Z}_p[X]}$, où t_i est irréductible modulo p,*
- *$\mathfrak{p}_i = p\mathcal{O}_K + t_i(\overline{\alpha_i})\mathcal{O}_K$ est maximal,*
- *les $\mathfrak{p}_i$ sont distincts, on a $e(\mathfrak{p}_i/p) = e_i$, $f(\mathfrak{p}_i/p) = \deg t_i$ et*

$$p\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_g^{e_g}.$$

Soit $\varepsilon_1, \ldots, \varepsilon_g$ les idempotents orthogonaux associés : si $\sum_{i=1}^g a_i \cdot (T/F_i) = 1$ est une relation de Bezout dans $\mathbb{Q}_p[X]$, on pose $\varepsilon_i =$

$a_i \cdot (T/F_i)$; alors

$$\mathcal{O}_p := \sum_{i=1}^{g} \overline{\varepsilon_i}\, \mathbb{Z}[\overline{\alpha_i}]$$

est p-maximal.

15.4. Diviseurs inessentiels. — En changeant T, on espère obtenir un ordre $\mathcal{O}_T$ qui soit p-maximal. Hélas, dès que $n \geqslant 3$, il peut exister des *diviseurs inessentiels*[9] p du discriminant tels que $p \mid [\mathcal{O}_K : \mathcal{O}_T]$ pour tout T tel que $p \nmid a_0$ (ou, de façon équivalente, pour T unitaire) :

***Théorème 15.8* (Hensel).** — *Soit p un nombre premier fixé. On note*

- *$r(f)$ le nombre de $\mathfrak{p} \mid p$ de degré résiduel f.*
- *$i(f)$ le nombre de $P \in \mathbb{F}_p[X]$ irréductibles unitaires de degré f.*

Alors p est diviseur inessentiel si et seulement s'il existe f tel que $r(f) > i(f)$.

Démonstration. — La condition est suffisante d'après le critère de Kummer (Théorème 4.10). La réciproque est un calcul local assez précis (fait dans Hasse [**20**]). □

Corollaire 15.9. — *Si p est diviseur inessentiel, alors $p < n =$* $\dim_{\mathbb{Q}} K$.

Démonstration. — D'après (8), $f \cdot r(f) \leqslant n$. D'autre part, soit

$$I := f \cdot i(f) = \sum_{d \mid f} \mu(f/d) p^d.$$

On a $I > 0$ et I est divisible par p, donc $I \geqslant p$. (Si on ne veut pas utiliser l'existence de $\mathbb{F}_{p^f}$, on peut se contenter de $I \geqslant 0$, $I \equiv p^d \pmod{p^{d+1}}$, d minimal tel que f/d soit sans facteurs carrés, qui donne $I \geqslant p^d$.) □

Exemple. — si $n = 3$, seul 2 peut être diviseur inessentiel. En fait, il l'est si et seulement si $2\mathcal{O}_K = \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3$, où $e(\mathfrak{p}_i/2) = f(\mathfrak{p}_i/2) = 1$ pour $i = 1, 2, 3$. On trouve facilement des exemples de ce phénomène

[9] Dedekind écrit *ausserwesentlich* (= d'essence extérieure), Bourbaki (mal inspiré) « facteurs extraordinaires » dans ses notes historiques. La littérature se partage entre diviseurs essentiels ou inessentiels... pour la même notion.

en utilisant le Théorème 15.7 : par exemple $K = \mathbb{Q}[X]/(T)$ où $T = x^3 - x - 8 = x(x-1)(x+1) + 8$.

15.5. Valuations. — L'algorithme Round 4 (§ 15.3 et Théorème 15.7) fournit simultanément la factorisation de T dans $\mathbb{Q}_p[X]$, un ordre p-maximal $\mathcal{O}$ de $\mathbb{Q}[X]/(T)$, et la décomposition de $p\mathcal{O}$ en produit d'idéaux maximaux donnés sous la forme $\mathfrak{p} = p\mathcal{O} + \pi\mathcal{O}$.

Proposition 15.10. — *Si $x \in K^*$, il est facile de calculer $v = v_{\mathfrak{p}}(x)$.*

Démonstration. — On peut supposer $x \in \mathcal{O}$, $x \neq 0$. Il suffit alors d'appliquer la définition : $v = \max\{w : x \in \mathfrak{p}^w\}$.

Une variante plus agréable utilise l'écriture $\mathfrak{p} = p\mathcal{O} + \pi\mathcal{O}$: la multiplication par π dans la $\mathbb{F}_p$-algèbre $\mathcal{O}/p\mathcal{O}$ n'est pas injective (son image est $\mathfrak{p}/p\mathcal{O}$ qui est de codimension $f(\mathfrak{p}/p) \geqslant 1$). Soit $\tau_0 \in \mathcal{O} \smallsetminus p\mathcal{O}$ un relèvement d'un élément non-trivial de son noyau et $\tau := \tau_0/p$. Alors $v = \max\{w, \tau^w x \in \mathcal{O}\}$; en effet, $v_{\mathfrak{p}}(\tau) = -1$ et $v_{\mathfrak{q}}(\tau) \geqslant 0$ pour tout $\mathfrak{q} \neq \mathfrak{p}$. □

La valuation d'un idéal se calcule comme le min des valuations de ses générateurs.

15.6. L'algorithme POLRED. — Un corps de nombres $K = \mathbb{Q}[X]/(T)$ est représenté par une infinité de polynômes $T \in \mathbb{Z}[X]$ différents. Certains d'entre eux sont plus agréables que d'autres. L'algorithme POLRED (Polynomial Reduction) de Cohen et Diaz y Diaz essaie de trouver un joli polynôme T définissant le même corps. Il calcule une base LLL-réduite (b_i) du réseau $(\mathcal{O}_K, T_2)$ et teste si l'un des b_i est primitif en vérifiant si leur polynôme caractéristique est sans facteurs carrés. Soit $\alpha = A(X) \bmod T$ dans K. Son polynôme caractéristique se calcule comme un résultant $P_{\mathrm{char},\alpha}(Y) = \mathrm{Res}_X(T(X), Y - A(X))$, ou comme produit des $Y - \sigma_i(\alpha)$, où les σ_i sont les plongements complexes de K.

POLRED peut échouer : il arrive qu'aucun des b_i ne soit primitif, auquel cas on considère de petites combinaisons linéaires des b_i. En cas de succès, il fournit un polynôme T engendrant le même corps K, dont les racines sont petites ; ainsi donc que ses coefficients. Remarquons qu'en cas d'échec, POLRED fournit des sous-corps de K.

15.7. Réduction LLL. — Soit $\mathfrak{A}$ un idéal fractionnaire non nul. Le premier vecteur d'une base LLL-réduite du réseau $(\mathfrak{A}, T_2)$ est un $\alpha \in \mathfrak{A}$ de norme relativement petite. On récrit

$$\mathfrak{A} = (\alpha)(\mathfrak{A}/\alpha) = (a)(\alpha)\mathfrak{a},$$

où $\mathfrak{a}$ est entier et primitif, $\alpha \in \mathcal{O}_K$ et $a \in \mathbb{Q}^*$. Ces trois composantes dépendent de la variante de l'algorithme LLL utilisé mais

Proposition 15.11. — $\mathrm{N}\,\mathfrak{a}$ *est bornée par une constante ne dépendant que de* K.

Démonstration. — Corollaire 5.8 (Minkowski) et Proposition 10.4 (LLL). □

En particulier, un produit quelconque d'idéaux se simplifie sous la forme $(\alpha)\mathfrak{a}$, où α est un produit d'éléments de K^*, que l'on peut développer si nécessaire, et $\mathfrak{a}$ est un *petit* idéal entier. Il est préférable de conserver α sous forme de produit *formel* : on écrit $\alpha = \prod x_i^{e_i} \in \mathbb{Z}[K^*]$. Par exemple, méditer sur les 30103 chiffres décimaux de la représentation développée des 7 caractères « 2^{100000} ». Dans l'essentiel des applications, on n'utilisera pas α mais sa projection sur un domaine où les calculs sont plus simples : $K \otimes \mathbb{R}$, $K \otimes \mathbb{Q}_p$, $(\mathcal{O}_K/\mathfrak{f})^*$, $K^*/(K^*)^\ell$... On en verra un exemple au § 16.5.

16. Groupe de classes et unités

Dans cette section, on suppose connu l'ordre maximal $\mathcal{O}_K$. On peut définir $\mathrm{Cl}(\mathcal{O})$ et $\mathcal{O}^*$ pour un ordre quelconque, mais en rajoutant des difficultés techniques sans éclaircir le problème initial. Pour montrer que $\mathrm{Cl}(K)$ est calculable en principe, il ne suffit pas d'invoquer la borne de Minkowski (Corollaire 5.8). Il manque une procédure effective pour décider si deux idéaux sont équivalents. « Factoriser suffisamment d'idéaux principaux » n'est pas recevable...

16.1. Calculabilité. — Pour toute place $\mathfrak{p}$ de K, on note $|.|_\mathfrak{p}$ la valeur absolue normalisée associée. Soit S_∞ l'ensemble des places archimédiennes de K et $S \supset S_\infty$ un ensemble fini de places.

***Théorème 16.1* (Lenstra [28]).** — *Soit* $d := (2/\pi)^{r_2}\Delta_K^{1/2}$, $S_0 := \{\mathfrak{p}, \mathrm{N}\,\mathfrak{p} \leqslant d\}$ *et* $S := S_\infty \cup S_0$. *Alors le groupe* $U_S(K)$ *est engendré par* $\left\{\alpha \in U_S(K), H(\alpha) \leqslant d^2\right\}$, *et* $\mathrm{Cl}(K)$ *est engendré par les classes des éléments de* S_0.

La calculabilité de $U(K)$ et $\mathrm{Cl}(K)$ se déduit de la suite exacte :

$$0 \longrightarrow U(K) \longrightarrow U_S(K) \xrightarrow{f} \mathbb{Z}^{S_0} \xrightarrow{g} \mathrm{Cl}(K) \longrightarrow 0,$$

où $f : \alpha \mapsto (v_{\mathfrak{p}}(\alpha))_{\mathfrak{p} \in S_0}$ et $g : (e_{\mathfrak{p}}) \mapsto \prod_{\mathfrak{p}} \mathfrak{p}^{e_{\mathfrak{p}}}$. L'exactitude provient du théorème et il suffit de calculer noyau et conoyau de f, en commençant par construire une $\mathbb{Z}$-base de $U_S(K)$ grâce au Théorème 16.1.

La démonstration est élémentaire, fondée sur le théorème de Minkowski ; l'assertion sur les générateurs de $\mathrm{Cl}(K)$ est immédiate car plus faible que celle du Corollaire 5.9. À première vue, ce théorème est un peu surprenant si $r_1 + r_2 > 1$ (c'est-à-dire $U(K) \neq \mu(K)$) : on s'attend à ce que le *régulateur* $R(K)$ de K soit souvent de la taille de $\sqrt{\Delta_K}$ d'après Brauer-Siegel (Théorème 7.6) et les heuristiques de Cohen-Lenstra-Martinet (§ 6.3) qui prédisent que $h(K)$ est « fréquemment » petit. Donc certains plongements des unités fondamentales devraient être de l'ordre de $\exp(\sqrt{\Delta_K}/n)$, quelle que soit la $\mathbb{Z}$-base choisie pour $U(K)/\mu(K)$. Par exemple, il est bien connu que l'unité fondamentale d'un corps quadratique peut devenir gigantesque, voir le célèbre problème des Bœufs d'Archimède. L'astuce consiste à considérer un système non minimal de générateurs.

Ce n'est *pas* un bon algorithme, car $U_S(K)$ est beaucoup trop gros : son rang est exponentiel en $\log \Delta_K$. En pratique, on utilise le principe du calcul d'indice.

16.2. Calcul d'indice. — Pour calculer un groupe abélien *fini* M par générateurs et relations, la méthode de calcul d'indice nécessite quatre ingrédients :

– Un $\mathbb{Z}$-module libre A_0 dont M est un quotient

$$0 \longrightarrow \Lambda_0? \longrightarrow A_0 \longrightarrow M \longrightarrow 0$$

(le noyau Λ_0 est inconnu).

– Un sous-groupe *de type fini* $A \subset A_0$, muni d'une $\mathbb{Z}$-base $\mathfrak{B}$, dont M reste un quotient :

$$0 \longrightarrow \Lambda_0 \cap A =: \Lambda? \longrightarrow A = \mathbb{Z}^{\mathfrak{B}} \longrightarrow M \longrightarrow 0$$

(le noyau Λ est inconnu). Un élément de Λ_0 appartenant à Λ est dit *friable*, $\mathfrak{B}$ est une *base de factorisation.*

– Un moyen de produire des éléments « bien répartis » dans Λ_0, puis de les plonger dans $\mathbb{Z}^{\mathfrak{B}}$ s'ils appartiennent à Λ (factorisation des éléments friables).

– Une évaluation grossière H de $h := \#M = [A : \Lambda]$, telle que $H < 2h$.

L'algorithme probabiliste suivant détermine alors M : produire des éléments de Λ, engendrant un sous groupe $\widehat{\Lambda} \subset \Lambda$, jusqu'à ce que

$$\widehat{h} := [A : \widehat{\Lambda}] \leqslant H < 2h,$$

ce qui entraîne $\widehat{h} = h$ puisque $\widehat{h}$ est un multiple entier de h, et donc $\widehat{\Lambda} = \Lambda$. On en déduit la structure de $M = \mathbb{Z}^{\mathfrak{B}}/\Lambda$ en calculant la SNF de Λ (Lemme 10.1) :

$$M = \oplus(\mathbb{Z}/d_i\mathbb{Z})g_i, \quad d_1 \mid d_2 \mid \dots, \quad g_i \in A.$$

La solution du logarithme discret dans A, c'est-à-dire l'écriture d'un élément de M comme produit des g_i et d'un élément de Λ est une généralisation simple que l'on considérera ultérieurement.

16.3. Adaptation au cas $M = \mathrm{Cl}(K)$, sous GRH. — Pour calculer $\mathrm{Cl}(K)$ et $U(K)$ en un temps raisonnable, on admet GRH pour pouvoir utiliser la borne de Bach (Théorème 7.5). On pose donc

$$\mathfrak{B} := \left\{\mathfrak{p}, \mathrm{N}\,\mathfrak{p} \leqslant 12(\log \Delta_K)^2\right\},$$

qui engendre $\mathrm{Cl}(K)$ et remplace l'ensemble S_0 de Lenstra. Même sous ces hypothèses fortes, la complexité des algorithmes est au mieux sous-exponentielle en $\log \Delta_K$, et reste heuristique. Elle n'est démontrée, sous GRH, que pour K quadratique imaginaire (Hafner-McCurley [**19**]). La complexité n'est pas seule en jeu : si GRH est fausse, le résultat obtenu peut l'être aussi. Tous les algorithmes permettant de le vérifier inconditionnellement sont exponentiels.

On étend l'idée du calcul d'indice en calculant simultanément $\widehat{\Lambda}$ comme ci-dessus et un sous-groupe $\widehat{U}$ de $U = U(K)$. On désire calculer $M := \mathrm{Cl}(K)$,

– A_0 est le groupe des idéaux fractionnaires de K,

– On produit des élément de Λ_0 de petite norme dans $\mathcal{O}_K$ en utilisant la réduction § 15.7 sur un produit aléatoire d'éléments de $\mathfrak{B}$: si le petit représentant de sa classe d'idéaux est friable, on obtient une relation. Le test de friabilité et la factorisation d'un élément friable s'effectue par divisions successives par les éléments de $\mathfrak{B}$.

– On calcule une approximation numérique du produit hR, par un produit eulérien tronqué convergeant vers le résidu en $s = 1$ de ζ_K :

$$\prod_{p \leqslant Y} \frac{(1-p^{-1})}{\prod_{\mathfrak{p}|p}(1-\mathrm{N}\,\mathfrak{p}^{-1})} \longrightarrow \left(\frac{\zeta_K}{\zeta}\right)(1) = 2^{r_1}(2\pi)^{r_2}\frac{hR}{w\sqrt{\Delta_K}}$$

Toujours sous GRH, l'approximation obtenue pour $Y = O(\log \Delta_K)^2$ est suffisamment précise.

– Dans la formule du résidu, Δ_K est connu puisque $\mathcal{O}_K$ l'est, (r_1, r_2) se calcule avec l'algorithme de Sturm (§ 11.7), et w en énumérant les points tels que $T_2(x) = n$ (Théorème 5.5) ou en testant l'inclusion de corps cyclotomiques $\mathbb{Q}(\zeta_m)$ dans K pour des valeurs convenables de m (§ 13.3).

Les dépendances entre éléments de $\widehat{\Lambda}$, découvertes au moment du calcul de $[A : \widehat{\Lambda}]$, se traduisent par des identités entre idéaux principaux $(\alpha) = (\alpha')$. On en déduit des unités $u := \alpha/\alpha' \in U$; ces éléments u engendrent un sous-groupe $\widehat{U}$ de U. Soit $\widehat{R}$ le régulateur de $\widehat{U}$, qui est un multiple entier du régulateur R ; tout comme ci-dessus, lorsque $hR < 2\widehat{h}\widehat{R}$, alors $U = \widehat{U}$ et $\Lambda = \widehat{\Lambda}$, d'où on tire $\mathrm{Cl}(K)$.

Pour obtenir des générateurs (g_i) explicites, il faut conserver les matrices de changement de base associées à toute cette algèbre linéaire. On obtient les g_i sous forme *factorisée* dans $\mathbb{Z}^{\mathfrak{B}}$, ce qui est une présentation compacte agréable. On peut la développer en utilisant la technique du § 15.7 pour obtenir de petits représentants de leurs classes d'idéaux.

D'après la discussion qui suit le Théorème 16.1, la taille de l'écriture naïve de générateurs de $U(K)$ comme éléments de $\mathbb{Q}[X]/(T)$ est *a priori* exponentielle. Dans l'algorithme ci-dessus, les unités sont obtenues elles-aussi comme produit de S-unités, qui n'appartiennent pas à $U(K)$ individuellement. Cette présentation compacte des unités est de taille raisonnable.

16.4. Logarithme discret. — Nous pouvons donc calculer

$$\mathrm{Cl}(K) = \oplus(\mathbb{Z}/d_i\mathbb{Z})g_i,$$

où l'on sait exprimer les g_i comme produits d'éléments de $\mathfrak{B}$, ou comme *petit* idéal, à un idéal principal explicite près. Réciproquement on peut exprimer un élément de $\mathfrak{B}$ en terme des g_i. Soit maintenant un idéal $\mathfrak{a}$; on veut calculer son logarithme discret dans $\mathrm{Cl}(K)$, c'est-à-dire $(e_i) \in \prod_i(\mathbb{Z}/d_i\mathbb{Z})$ et $\tau \in K^*$, tels que $\mathfrak{a} = (\tau)\prod g_i^{e_i}$.

– Pour calculer les (e_i), multiplier $\mathfrak{a}$ par des produits aléatoires d'idéaux de $\mathfrak{B}$ et LLL-réduire le résultat, jusqu'à ce que la composante non principale soit friable. À un idéal principal près, on sait alors exprimer $\mathfrak{a}$ comme produit d'éléments de $\mathfrak{B}$, donc des g_i.

– Calculer $\mathfrak{a}\prod g_i^{-e_i}$ sous la forme $(\beta)\mathfrak{B}$, $\beta \in \mathbb{Z}[K^*]$ (§ 15.7). Puis réaliser le *petit* idéal *principal* $\mathfrak{B}$ comme (γ), en utilisant la même méthode que ci-dessus, mais cette fois-ci en accumulant les idéaux principaux rencontrés. On pose $\tau := \beta\gamma \in \mathbb{Z}[K^*]$.

16.5. Application : entiers de norme donnée. — Soit $a \in \mathbb{Z}$, $a \neq 0$, et soit à résoudre l'équation $\mathrm{N}(x) = a$, $x \in \mathcal{O}_K$. C'est une équation diophantienne en les $x_i \in \mathbb{Z}$, si on exprime $x = \sum_{i=1}^n x_i w_i$, dans une $\mathbb{Z}$-base (w_i) de $\mathcal{O}_K$. Par exemple $x^2 - 2y^2 = -1$ est de cette forme.

Il suffit de chercher les solutions modulo les unités de norme 1. On détermine d'abord s'il existe une unité de norme -1, en considérant successivement les générateurs de $U(K)$. Remarquons que la présentation compacte $u = \prod x_i^{e_i} \in \mathbb{Z}[K^*]$ permet de tester le signe de $\mathrm{N}(u)$ à faible coût : il est donné par la parité du nombre de plongement réels σ de K tels que $\sigma(u) < 0$ (la norme est le produit des plongements, le produit de deux plongements complexes conjugués est > 0). Ce signe se détecte à partir de valeurs approchées des $\sigma(x_i)$ associés aux e_i impairs.

En factorisant $|a|$ en produit de nombres premiers, puis en décomposant ceux-ci dans $K/\mathbb{Q}$, on détermine l'ensemble (fini) des idéaux entiers de norme $|a|$. Pour un tel idéal $\mathfrak{a}$, on teste s'il est principal ; si $\mathfrak{a} = (\alpha)$, on teste si $a\,\mathrm{N}(\alpha) > 0$, comme indiqué ci-dessus. Sinon, $\mathrm{N}(\alpha) = -a$ et s'il existe une unité u de norme -1, on corrige $\alpha \leftarrow u\alpha$.

Références

[1] L.M. ADLEMAN & H.W. LENSTRA, JR. – « Finding irreducible polynomials over finite fields », *18th ACM Symposium on Theory of Computing* (1986), p. 350–355.

[2] M. AGRAWAL, N. KAYAL & N. SAXENA – « PRIMES is in P », *Ann. of Math. (2)* **160** (2004), no. 2, p. 781–793.

[3] A.V. AHO, J.E. HOPCROFT & J.D. ULLMAN – *The design and analysis of computer algorithms*, Addison-Wesley, 1975, Second printing.

[4] E. BACH – « Explicit bounds for primality testing and related problems », *Math. Comp.* **55** (1990), no. 191, p. 355–380.

[5] B. BEAUZAMY – « Products of polynomials and a priori estimates for coefficients in polynomial decompositions : a sharp result », *J. Symbolic Comput.* **13** (1992), no. 5, p. 463–472.

[6] K. BELABAS – « Topics in computational algebraic number theory », *J. Théor. Nombres Bordeaux* **16** (2004), p. 19–63.

[7] K. BELABAS, M. VAN HOEIJ, J. KLÜNERS & A. STEEL – « Factoring polynomials over global fields », preprint.

[8] E.R. BERLEKAMP – « Factoring polynomials over large finite fields », *Math. Comp.* **24** (1970), p. 713–735.

[9] J. BUCHMANN & H.W. LENSTRA, JR. – « Approximating rings of integers in number fields », *J. Théor. Nombres Bordeaux* **6** (1994), no. 2, p. 221–260.

[10] H. COHEN – *A course in computational algebraic number theory*, 3e éd., Springer-Verlag, 1996.

[11] H. COHEN & H.W. LENSTRA, JR. – « Heuristics on class groups of number fields », in *Number theory, Noordwijkerhout 1983*, Lecture Notes in Math., vol. 1068, Springer, Berlin, 1984, p. 33–62.

[12] H. COHEN & J. MARTINET – « Études heuristiques des groupes de classes des corps de nombres », *J. reine angew. Math.* **404** (1990), p. 39–76.

[13] H. DAVENPORT & H. HEILBRONN – « On the density of discriminants of cubic fields (II) », *Proc. Roy. Soc. Lond. A* **322** (1971), p. 405–420.

[14] J.-P. DEMAILLY – *Analyse numérique et équations différentielles*, Presses Universitaires de Grenoble, 1996.

[15] J. ELLENBERG & A. VENKATESH – « The number of extensions of a number field with fixed degree and bounded discriminant », *Annals of Math*, to appear.

[16] D. FORD, S. PAULI & X.-F. ROBLOT – « A fast algorithm for polynomial factorization over $\mathbb{Q}_p$ », *J. Théor. Nombres Bordeaux* **14** (2002), no. 1, p. 151–169.

[17] J. VON ZUR GATHEN & J. GERHARD – *Modern computer algebra*, Cambridge University Press, New York, 1999.

[18] X. Gourdon – « Algorithmique du théorème fondamental de l'algèbre », Rapport de recherche 1852, INRIA, 1993.

[19] J.L. Hafner & K.S. McCurley – « A rigorous subexponential algorithm for computation of class groups », *J. Amer. Math. Soc.* **2** (1989), no. 4, p. 837–850.

[20] H. Hasse – *Zahlentheorie*, Akademie-Verlag GmbH, 1949.

[21] P. Henrici – *Applied and computational complex analysis*, Wiley-Interscience [John Wiley & Sons], New York, 1974, Volume 1 : Power series—integration—conformal mapping—location of zeros, Pure and Applied Mathematics.

[22] M. van Hoeij – « Factoring polynomials and the knapsack problem », *J. Number Theory* **95** (2002), no. 2, p. 167–189.

[23] J.C. Lagarias, H.L. Montgomery & A.M. Odlyzko – « A bound for the least prime ideal in the Chebotarev density theorem », *Invent. Math.* **54** (1979), no. 3, p. 271–296.

[24] J.C. Lagarias & A.M. Odlyzko – « Effective versions of the Chebotarev density theorem », in *Algebraic number fields : L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975)*, Academic Press, London, 1977, p. 409–464.

[25] S. Lang – *Algebraic number theory*, 2^{e} éd., Graduate Texts in Math., vol. 110, Springer-Verlag, New York, 1994.

[26] A.K. Lenstra & H.W. Lenstra, Jr. (éds.) – *The development of the number field sieve*, Lecture Notes in Math., vol. 1554, Springer-Verlag, Berlin, 1993.

[27] A.K. Lenstra, H.W. Lenstra, Jr. & L. Lovász – « Factoring polynomials with rational coefficients », *Math. Ann.* **261** (1982), no. 4, p. 515–534.

[28] H.W. Lenstra, Jr. – « Algorithms in algebraic number theory », *Bull. Amer. Math. Soc. (N.S.)* **26** (1992), no. 2, p. 211–244.

[29] M. Mignotte – « An inequality about factors of polynomials », *Math. Comp.* **28** (1974), p. 1153–1157.

[30] P. Nguyen & D. Stehlé – « Floating point LLL revisited », proceedings of Eurocrypt'05, to appear.

[31] J. Oesterlé – « Le problème de gauss sur le nombre de classes », *Enseign. Math.* **34** (1988), p. 43–67.

[32] C.H. Papadimitriou – *Computational complexity*, Addison-Wesley, 1994.

[33] F. Rouillier & P. Zimmermann – « Efficient isolation of polynomial real roots », *Journal of Computational and Applied Mathematics* **162** (2003), no. 1, p. 33–50.

[34] R. Schoof – « Four primality algorithms », to appear, `http://www.mat.uniroma2.it/~schoof/millerrabinpom.pdf`.

[35] S. Siksek – « The modular approach to diophantine equations », preprint, `http://igd.univ-lyon1.fr/~webeuler/ihp/LectureNotes_Siksek.dvi`.

[36] P. Stevenhagen & H.W. Lenstra, Jr. – « Chebotarëv and his density theorem. », *Math. Intell.* **18** (1996), no. 2, p. 26–37.

[37] A. Storjohann – « Algorithms for matrix canonical forms », Thèse, ETH Zurich, 2000.

K. Belabas, Universite Bordeaux 1, 351, cours de la Liberation, F-33405 Talence (France) • *E-mail :* `belabas@math.u-bordeaux.fr`
Url : `http://www.math.u-bordeaux.fr/~belabas`