II. Le problème du nombre de classes

au corps $\mathbf{Q}(\sqrt{-d})$. On a en particulier si $d > 4$

$$h(-df^2) \geqslant h(-d)\varphi(f) \tag{10}$$

où φ est la fonction d'Euler.

II. Le problème du nombre de classes

Dans cette partie, nous allons étudier le comportement du nombre de classes lorsque le discriminant tend vers $-\infty$. Compte tenu des formules (8) de I. § 3 et (10) de I. § 5, il est légitime de restreindre notre étude aux discriminants fondamentaux (cf. I. § 3). Dans toute la suite, $-d$ sera un tel discriminant: on aura donc $\tilde{h}(-d) = h(-d)$.

Dans les derniers numéros de son exposé de la classification des formes quadratiques, Gauss émet quelques observations concernant les tables de nombres de classes (il avait constitué lui-même de telles tables, en particulier pour $d \leqslant 3000$); il qualifie de surprenante l'observation suivante [1]: pour chaque entier $h \geqslant 1$, il semble n'y avoir qu'un nombre *fini* de d tels que $h(-d) = h$. Ainsi, pour $h = 1$, ne trouve-t-il dans sa table que les neuf discriminants fondamentaux

$$-3, -4, -7, -8, -11, -19, -43, -67, -163$$

(et en outre les quatre discriminants non fondamentaux $-12, -16, -27, -28$).

Comme nous l'avons dit dans l'introduction, Heilbronn [2] en 1934 a démontré que, conformément à l'observation de Gauss, on a bien

$$\lim_{d \to \infty} h(-d) = +\infty . \tag{11}$$

Des tables étendues de nombres de classes ont été construites par ordinateur. Buell [3] par exemple a publié les valeurs de $h(-d)$ pour $d \leqslant 4\,000\,000$. Parmi les discriminants fondamentaux satisfaisant à cette inégalité, le nombre de ceux pour lesquels $h(-d)$ est égal à 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 est respectivement 9, 18, 16, 54, 25, 51, 31, 131, 34, 87, et

---

[1]) C.-F. Gauss, *Disquisitiones Arithmeticae*, n° 303.

[2]) H. Heilbronn, *On the class numbers in imaginary quadratic fields*, Quarterly J. of Math. (Oxford), 5 (1934), 150-160.

[3]) D. A. Buell, *Small class numbers and extreme values of L-functions of quadratic fields*, Math. of Comp. 31 (1977), 786-796.

le plus grand d correspondant est respectivement 163, 427, 907, 1555, 2683, 3763, 5923, 6307, 10627, 13843.

Cela semble suggérer que tous les discriminants fondamentaux $-d$ pour lesquels $h(-d) \leqslant 10$ figurent dans la table de Buell. Peut-on le prouver? C'est à ce type de question qu'est consacrée la fin de l'exposé. On s'intéresse à ce problème car les discriminants pour lesquels $h(-d)$ est petit possèdent comme nous le verrons des propriétés arithmétiques remarquables. Nous allons commencer par décrire les deux outils essentiels pour l'étude de $h(-d)$, à savoir les nombres de représentations des entiers par les formes quadratiques et les fonctions zêta associées.

Les formes quadratiques de discriminant -3 *et* -4 *ont des automorphismes distincts de* $\pm I$ *dans* $SL_2(\mathbf{Z})$. *Pour éviter les complications techniques qui en résultent, nous supposerons dans la suite* $d \neq 3$ *et* $d \neq 4$ *(donc* $d \geqslant 7$*).*

§ 1. Représentation des entiers par les formes quadratiques

Soit q une forme quadratique de discriminant $-d$ (distinct de -3 et -4). Le *nombre de représentations primitives* d'un entier $n \geqslant 1$ par q, comptées au signe près, est

$$(12) \qquad r_n(q) = \frac{1}{2}\,\mathrm{Card}\,\{(u, v) \in \mathbf{Z}^2 \mid q(u, v) = n \quad \text{et} \quad \mathrm{pgcd}\,(u, v) = 1\}\,.$$

Ce nombre ne dépend que de la classe C de la forme quadratique q, et on le note aussi $r_n(C)$. Soit $ax^2 + bxy + cy^2$ la forme réduite appartenant à C. On a $3a^2 \leqslant 4ac - b^2 < 4c^2$ (l'inégalité est stricte car $d \neq 4$), d'où $a \leqslant \sqrt{d/3}$ et $c > \sqrt{d}/2$. On a $r_a(C) \neq 0$, et si $n \geqslant 1$ est un entier $< c$ tel que $r_n(C) \neq 0$, on a nécessairement $n = a$ et $r_n(C) = 1$ (I. § 2, formule (6)). On en déduit

$$(13) \qquad \sum_{n \leqslant \sqrt{d}/2} r_n(C) \leqslant 1 \leqslant \sum_{n \leqslant \sqrt{d/3}} r_n(C)\,.$$

Introduisons le *nombre total des représentations primitives*, comptées au signe près, de l'entier n par les différentes classes de formes quadratiques de discriminant $-d$:

$$(14) \qquad r_n(-d) = \sum_{C \in Cl(-d)} r_n(C)\,.$$

On déduit de (13) un *encadrement du nombre de classes*

$$(15) \qquad \sum_{n \leqslant \sqrt{d}/2} r_n(-d) \leqslant h(-d) \leqslant \sum_{n \leqslant \sqrt{d/3}} r_n(-d),$$

ce qui montre que l'étude de $h(-d)$ est liée à celle des nombres $r_n(-d)$.

Il n'existe à ma connaissance aucune formule simple permettant pour une classe C donnée de calculer $r_n(C)$. Par contre, Gauss a obtenu le résultat remarquable suivant [1]):

THÉORÈME. *Pour tout entier* $n \geqslant 1$, $r_n(-d)$ *est le nombre de* b (mod. $2n$) *tels que* $b^2 \equiv -d$ (mod. $4n$).

La démonstration de Gauss est très élégante: Soit (q_i) un système de représentants des classes de formes quadratiques de discriminant $-d$. Si b est un entier tel que b^2 s'écrive $-d + 4nc$, la forme quadratique $nx^2 + bxy + cy^2$ a pour discriminant $-d$ et s'écrit $q_i(ux+wy, vx+ty)$ pour un unique indice i et une certaine matrice $\begin{pmatrix} u & w \\ v & t \end{pmatrix} \in SL_2(\mathbf{Z})$. On a $q_i(u, v) = n$, et (u, v) est déterminé au signe près par b (mod. $2n$) car I et $-I$ sont les seuls automorphismes de q_i dans $SL_2(\mathbf{Z})$. Inversement, chaque représentation primitive de n par l'une des formes q_i s'obtient par ce procédé à partir d'un unique b (mod. $2n$) tel que $b^2 \equiv -d$ (mod. $4n$).

En décomposant $\mathbf{Z}/4n\mathbf{Z}$ en ses composantes primaires, on obtient la forme équivalente suivante de l'énoncé précédent:

COROLLAIRE. *Pour que* $r_n(-d) \neq 0$, *il faut et il suffit que* n *soit de la forme* $d' p_1^{\alpha_1} \ldots p_m^{\alpha_m}$, *avec* d' *un diviseur de* d *sans facteurs carrés,* $p_1, \ldots, p_m$ *des nombres premiers deux à deux distincts modulo lesquels* $-d$ *est un carré non nul, et* $\alpha_1, \ldots, \alpha_m$ *des entiers* $\geqslant 1$. *On a alors* $r_n(-d) = 2^m$.

De la formule (15) et du corollaire ci-dessus, on peut retenir le principe suivant:

PRINCIPE. *Si* d *est grand et* $h(-d)$ *est petit, il y a peu de petits entiers* n *qui soient représentés par une forme quadratique de discriminant* $-d$, *et peu de petits nombres premiers modulo lesquels* $-d$ *est un carré.*

Illustrons ceci dans le cas particulier où $d = 163$. On a $h(-163) = 1$ et $x^2 + xy + 41y^2$ est la seule forme quadratique réduite de discriminant

[1]) C.-F. GAUSS, *Disquisitiones Arithmeticae*, n° 167, 168 et 180.

-163. D'après le début de ce paragraphe, on a $r_n(-163) = 0$ pour $2 \leqslant n \leqslant 40$. Par suite, -163 n'est un carré modulo aucun des nombres premiers $\leqslant 39$, et le corollaire au théorème ci-dessus implique que si $r_n(-163) \neq 0$ et $n < 41^2$, nécessairement n est premier. Ceci explique pourquoi la suite (découverte par Euler): 41, 43, 47, 53, 61, ..., formée par les valeurs de $x^2 + x + 41$ pour $x \geqslant 0$ ne comporte que des nombres premiers jusqu'à 1601 $(=39^2+39+41)$.

§ 2. Fonctions zêta

Il est fructueux de réinterpréter les résultats du paragraphe précédent en introduisant des *séries de Dirichlet génératrices*: pour toute forme quadratique q de discriminant $-d$, la série de Dirichlet

$$\zeta(q, s) = \frac{1}{2} \sum_{(u,v) \in \mathbf{Z}^2 - \{(0,0)\}} q(u, v)^{-s} \tag{16}$$

converge absolument pour $\mathrm{Re}\,(s) > 1$ et l'on a

$$\zeta(q, s) = \zeta(2s) \sum_{n=1}^{\infty} r_n(q) n^{-s} \tag{17}$$

où $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ est la fonction zêta de Riemann. Comme $\zeta(q, s)$ ne dépend que de la classe C de q, on l'écrit aussi $\zeta(C, s)$.

La fonction $\zeta(q, s)$ jouit de remarquables propriétés analytiques: la fonction

$$\Lambda(q, s) = 2d^{s/2}(2\pi)^{-s}\Gamma(s)\zeta(q, s) \tag{18}$$

admet un *prolongement méromorphe à* $\mathbf{C}$, avec pour seuls pôles des *pôles simples en* 0 *et* 1 de résidus -1 et 1, et vérifie *l'équation fonctionnelle* $\Lambda(q, 1-s) = \Lambda(q, s)$. En effet, la fonction thêta

$$\theta(q, t) = \sum_{(n,m) \in \mathbf{Z}^2} \exp\left(-q(n, m)2\pi t/\sqrt{d}\right) \tag{19}$$

satisfait d'après la formule sommatoire de Poisson à l'équation fonctionnelle

$$\theta(q, t^{-1}) = t\theta(q, t)\,; \tag{20}$$

on a, par échange de la somme et de l'intégrale,

$$\Lambda(q, s) = \int_0^{\infty} [\theta(q, t) - 1]\, t^{s-1} dt\,, \tag{21}$$

et l'on en déduit l'expression suivante de $\Lambda(q, s)$, sur laquelle le prolongement méromorphe, les pôles et leurs résidus, et l'équation fonctionnelle sont évidents

$$\Lambda(q, s) = \frac{1}{s(s-1)} + \int_1^\infty [\theta(q, t)-1]\,(t^{s-1}+t^{-s})dt\,. \tag{22}$$

Soit K le corps quadratique imaginaire $\mathbf{Q} + \mathbf{Q}i\sqrt{d}$. On peut déduire du dictionnaire entre formes quadratiques de discriminant $-d$ et $\mathcal{O}(-d)$-idéaux fractionnaires (I., § 4) que l'on a

$$\zeta_K(s) = \sum_{C \in Cl(-d)} \zeta(C, s) = \zeta(2s) \sum_{n=1}^{\infty} r_n(-d)n^{-s} \tag{23}$$

où ζ_K est la fonction zêta du corps K (définie par $\zeta_K(s) = \sum_{\mathfrak{a}} N\mathfrak{a}^{-s}$, où $\mathfrak{a}$ parcourt l'ensemble des idéaux non nuls de l'anneau $\mathcal{O}(-d)$. Cette fonction ζ_K jouit de propriétés analytiques analogues à celles des fonctions $\zeta(C, s)$: en particulier, d'après ce qui précède, elle a un pôle simple en 1 de résidu

$$\mathrm{Res}_{s=1}\,\zeta_K(s) = \pi d^{-1/2}h(-d)\,. \tag{24}$$

Cette formule joue un rôle fondamental pour l'étude de $h(-d)$ par voie analytique.

Notons χ le caractère de Dirichlet $n \mapsto \left(\frac{-d}{n}\right)$. Le théorème de Gauss du § 2, ou plutôt son corollaire, traduit alors l'égalité entre séries de Dirichlet

$$\sum_{n=1}^{\infty} r_n(-d)n^{-s} = \prod_{p \text{ premier}} \left(\frac{1+p^{-s}}{1-\chi(p)p^{-s}}\right) \tag{25}$$

ou encore, compte tenu de (24), l'égalité

$$\zeta_K(s) = \zeta(s)L(\chi, s) \tag{26}$$

où $L(\chi, s)$ est la série de Dirichlet $\sum_{n=1}^{\infty} \chi(n)n^{-s}$. Cette égalité équivaut à la décomposition de ζ_K en produit eulérien, décomposition que l'on prouve de nos jours directement en utilisant la factorisation des idéaux dans l'anneau de Dedekind $\mathcal{O}(-d)$.

En utilisant (25) et (26), nous allons reformuler le principe énoncé à la fin du paragraphe précédent.

PRINCIPE. *Supposons d grand et $h(-d)$ petit. Alors, on a $\chi(p) = -1$ pour la plupart des petits nombres premiers p. Si $\lambda: \mathbf{N} - \{0\} \to \{-1, 1\}$ est la fonction qui à un produit de r nombres premiers (non nécessairement*

distincts) associe $(-1)^r$, *on a* $\lambda(n) = \chi(n)$ *pour la plupart des petits nombres entiers* n. *La fonction* $\zeta_K(s)$ *doit ressembler à la fonction* $\zeta(2s)$.

Ces énoncés sont volontairement vagues. Les rendre précis est souvent le nœud des démonstrations de minorations de $h(-d)$ lorsque d tend vers ∞.

§ 3. Ce que l'on espère sur le comportement de $h(-d)$

On peut montrer que *en moyenne* (en un sens qui demande à être précisé, ce que je ne ferai pas ici), $h(-d)$ est équivalent à une constante non nulle fois $\sqrt{d}$; déjà Gauss connaissait ce type de résultat [1]).

Il n'est pas vrai par contre que $h(-d)/\sqrt{d}$ admette un minorant > 0 ou un majorant lorsque d tend vers $+\infty$: on sait par exemple que $h(-d)/(\sqrt{d}\log\log d)$ ne tend pas vers 0 et que $h(-d)\log\log d/\sqrt{d}$ ne tend pas vers $+\infty$ lorsque d tend vers $+\infty$.

On obtient cependant de façon élémentaire des *majorations* raisonnables de $h(-d)$ (raisonnable signifiant avec l'exposant $\frac{1}{2}$ que l'on attend pour d), de la forme $h(-d) \leqslant C\sqrt{d}\log d$. Par exemple:

Proposition. *On a pour* $d > 4$

$$h(-d) \leqslant \pi^{-1}\sqrt{d}\log d\,. \tag{27}$$

Compte tenu de (24) et (26), il revient au même de montrer que l'on a, en posant $\chi(n) = \left(\frac{-d}{n}\right)$

$$\sum_{n=1}^{\infty} \chi(n)/n \leqslant \log d\,.$$

Or, pour tout nombre réel $x > 0$, la somme $M(x) = \sum_{n\leqslant x}\chi(n)$ est majorée par $N(x) = \inf([x], [(d-1)/2])$, et l'on a donc, en intégrant par parties

$$\sum_{n=1}^{\infty}\frac{\chi(n)}{n} = \int_{1-}^{\infty}\frac{dM(x)}{x} = \int_1^{\infty}\frac{M(x)}{x^2}\,dx \leqslant \int_1^{\infty}\frac{N(x)}{x^2}\,dx$$
$$= \int_{1-}^{\infty}\frac{dN(x)}{x} = \sum_{n\leqslant[(d-1)/2]} 1/n \leqslant \log d\,.$$

[1]) C.-F. Gauss, *Disquisitiones Arithmeticae*, n° 302.

Il est possible d'obtenir des *minorations* raisonnables de $h(-d)$ *si l'on admet l'hypothèse de Riemann généralisée.* Ainsi par exemple, en suivant une démonstration de Hecke, publiée par Landau [1]), on obtient:

PROPOSITION. *Si la fonction zêta* ζ_K *du corps* $K = \mathbf{Q} + \mathbf{Q}i\sqrt{d}$ *n'admet aucun zéro réel* $> 1 - (2/\log d)$, *on a*

$$h(-d) \geqslant \frac{2}{\pi e}\sqrt{d}/\log d\,. \tag{28}$$

Soit $\alpha \in]1/2, 1[$ tel que ζ_K ne s'annule pas dans l'intervalle $]\alpha, 1[$. On a alors $\zeta_K(\alpha) \leqslant 0$, c'est-à-dire $\sum_{C \in Cl(-d)} \Lambda(C, \alpha) \leqslant 0$ (formule (23)). Or il résulte de la formule (22) que $\Lambda(C, \alpha) + (\alpha(1-\alpha))^{-1}$ est positif pour toute classe $C \in Cl(-d)$, et même supérieur à $2\int_1^\infty e^{-2\pi t/\sqrt{d}}(t^{\alpha-1} + t^{-\alpha})dt$ lorsque C est la classe neutre. On a par conséquent

$$h(-d) \geqslant 2\alpha(1-\alpha)\int_1^\infty e^{-2\pi t/\sqrt{d}}(t^{\alpha-1} + t^{-\alpha})dt\,.$$

Le second membre de (28) est majoré par 1 pour $d \leqslant 800$, par 2 pour $d \leqslant 5000$, par 3 pour $d \leqslant 15000$. Il nous suffit donc de démontrer la proposition pour $d \geqslant 15000$. Prenons alors α égal à $1 - (2/\log d)$; remarquons que

$$\int_1^\infty e^{-2\pi t/\sqrt{d}} t^{-\alpha} dt \geqslant \int_1^6 e^{-2\pi t/\sqrt{d}} t^{-1} dt \geqslant e^{-12\pi/\sqrt{d}} \log 6 \geqslant 1,3 \geqslant$$

$$1/\alpha = \int_0^1 t^{\alpha-1} dt \geqslant \int_0^1 e^{-2\pi t/\sqrt{d}} t^{\alpha-1} dt\,,$$

d'où

$$h(-d) \geqslant 2\alpha(1-\alpha)\int_0^\infty e^{-2\pi t/\sqrt{d}} t^{\alpha-1} dt = 2\alpha(1-\alpha)\,(\sqrt{d}/2\pi)^\alpha \Gamma(\alpha)\,.$$

L'application $x \mapsto x(2\pi)^{-x}\Gamma(x)$ étant décroissante sur $\left]\frac{1}{2}, 1\right[$, on en déduit

[1]) E. LANDAU, *Über die Klassenzahl imaginär-quadratischer Zahlkörper*, Göttingen Nachrichten (1918), 285-295.

$$h(-d) \geqslant \frac{1}{\pi}(1-\alpha)d^{\alpha/2} = \frac{2}{\pi e}(\sqrt{d}/\log d).$$

Si nous sommes entrés dans les détails de cette démonstration, c'est pour bien illustrer les deux points suivants:

1) Nous voyons à l'œuvre le principe général énoncé à la fin du § 2, qui dit que si d est grand et $h(-d)$ est petit, $\zeta_K(s)$ doit ressembler à $\zeta(2s)$: en effet $\zeta_K(s)$ admet un pôle en 1, alors que $\zeta(2s)$ est holomorphe pour $\mathrm{Re}\,(s) > \frac{1}{2}$; mais si d est grand et $h(-d)$ petit, l'existence du pôle pour ζ_K doit être contrebalancée par l'existence d'un zéro de ζ_K proche de 1, d'après la proposition ci-dessus.

2) Si l'hypothèse de Riemann généralisée était démontrée, les questions posées dans l'introduction de cette deuxième partie seraient résolues: ainsi par exemple il résulterait de la proposition que tous les discriminants fondamentaux $-d$ pour lesquels $h(-d) \leqslant 30$ figurent dans la table de Buell.

§ 4. Minorations non effectives de $h(-d)$

Comme nous l'avons vu au paragraphe précédent, $h(-d)$ est grand lorsque d est grand et que la fonction $L(\chi_d, s)$, où $\chi_d(n) = \left(\frac{-d}{n}\right)$, n'a pas de zéro voisin de 1. Supposons alors que $h(-d)$ et $h(-d')$ soient petits pour deux grandes valeurs de d et d' (en un sens que l'on peut préciser, ce que je ne ferai pas ici). Les fonctions $L(\chi_d, s)$ et $L(\chi_{d'}, s)$ ont alors chacune un zéro voisin de 1, et l'on en déduit que la fonction zêta du corps biquadratique $\mathbf{Q}[i\sqrt{d}, i\sqrt{d'}]$ a deux zéros voisins de 1. Des estimées élémentaires permettent d'en déduire une contradiction. Cette méthode montre que $h(-d)$ ne peut être petit que pour au plus un grand d. Elle est une variante de celle utilisée par Heilbronn pour montrer que

$$\lim_{d\to\infty} h(-d) = \infty, \tag{29}$$

et a été utilisée par Siegel [1]) pour préciser à quelle allure $h(-d)$ tend vers $+\infty$: Siegel montre que pour tout $\varepsilon > 0$, il existe un entier $d(\varepsilon)$ tel que: $h(-d) \geqslant \sqrt{d}^{1-\varepsilon}$ pour $d \geqslant d(\varepsilon)$.

---

[1]) C. L. Siegel, *Über die Classenzahl quadratischer Zahlkörper*, Acta Arithmetica 1 (1936), 83-86.

Il n'est malheureusement pas possible de calculer $d(\varepsilon)$ car cet entier dépend de l'hypothétique grand discriminant exceptionnel pour lequel $h(-d)$ serait petit.

On peut cependant obtenir par les méthodes précédentes un énoncé « effectif à au plus une exception près ». Cela a été fait par Tatuzawa [1]) en explicitant les constantes dans la démonstration de Siegel: si $0 < \varepsilon < \frac{1}{2}$, on a

$$h(-d) \geqslant \frac{0{,}655}{\pi} \varepsilon d^{\frac{1}{2}-\varepsilon} \tag{30}$$

pour $d > \sup(e^{1/\varepsilon}, e^{11,2})$ à au plus une exception près. On en déduit par exemple, en prenant $\varepsilon = 1/15$, que tous les discriminants fondamentaux $-d$ pour lesquels $h(-d) \leqslant 10$, *à au plus une exception près*, figurent dans la table de Buell et par suite sont de valeur absolue $\leqslant 13843$.

§ 5. Les cas $h = 1$ et $h = 2$ [2])

D'après le paragraphe précédent, il existe au plus un discriminant fondamental $-d$ tel que $h(-d) = 1$ et qui ne figure pas parmi les neuf déjà connus de Gauss. La question de savoir si un tel d existe est restée longtemps ouverte et est devenue célèbre sous le nom de *problème du dixième discriminant* (ou *du dixième corps quadratique imaginaire*).

En 1952, Heegner publie une preuve de la non-existence du dixième discriminant reposant sur la théorie des formes modulaires, mais cette preuve fut jugée incomplète à l'époque.

En 1966, Stark et Baker prouvent indépendamment la non-existence du dixième discriminant. Dans sa preuve, Stark ramène ce problème à la détermination des solutions entières des équations $8x^6 \pm 1 = y^2$ et $x^6 \pm 1 = 2y^2$. Ces équations apparaissent déjà dans le travail de Heegner. En fait, deux ans plus tard, Stark et Birch reprennent en détail les arguments de Heegner et montrent la validité de sa démonstration.

La méthode de Baker utilise les minorations effectives de formes linéaires en logarithmes de nombres algébriques. Elle a l'avantage de s'étendre au problème du nombre de classes 2, et a permis à Baker et Stark de majorer

---

[1]) T. Tatuzawa, *On a theorem of Siegel*, Jap. J. of Math., 21 (1951), 163-178.

[2]) Pour un exposé plus détaillé des questions abordées dans ce paragraphe, avec références bibliographiques, on pourra consulter par exemple l'exposé de M. Waldschmidt au Séminaire de Théorie des nombres de Paris en 1973 (exposé 12).

de façon effective les d pour lesquels $h(-d) = 2$; les bornes obtenues sont très grandes (Stark obtient par exemple $|d| < 10^{1100}$), mais Stark d'une part, Montgomery et Weinberger de l'autre, ont mis au point des méthodes qui permettent par un calcul sur ordinateur utilisant les zéros de la fonction zêta de Riemann (pour Stark) ou de séries $L(\chi, s)$ (pour Montgomery et Weinberger) de vérifier que, en dessous des bornes précédentes, tous les d pour lesquels $h(-d) = 2$ sont $\leqslant 427$.

Pour l'instant, aucune des méthodes précédentes n'a pu être appliquée au problème du nombre de classes h pour $h \geqslant 3$.

§ 6. Courbes elliptiques et fonctions L

Nous allons maintenant parler un peu des courbes elliptiques, car elles jouent un rôle fondamental dans la suite de l'histoire du problème de Gauss.

Considérons une équation de la forme

$$(W) \qquad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\,,$$

où les a_i sont dans $\mathbf{Q}$. La cubique projective E définie par l'équation homogène associée a un unique point à l'infini 0. Lorsque E est non singulière, on dit que E (ou plutôt que le couple $(E, 0)$) est une *courbe elliptique définie sur* $\mathbf{Q}$, et que (W) en est une *équation de Weierstrass*. Un changement de variables

$$(C) \qquad \begin{aligned} x &= u^2x' + r \\ y &= u^3y' + sx' + t \qquad (u, r, s, t \text{ dans } \mathbf{Q}, u \neq 0) \end{aligned}$$

conduit à une autre équation de Weierstrass (W') de E. On dit que l'équation (W) est *minimale* si les coefficients a_i sont entiers et si les équations (W') déduites de (W) par un changement de variables (C) avec u, r, s, t entiers et $u \neq \pm 1$, ne sont pas à coefficients entiers.

Une courbe elliptique E définie sur $\mathbf{Q}$ admet une équation minimale et toute autre équation minimale s'en déduit par un changement de variables (C) avec $u = \pm 1$ et r, s, t dans $\mathbf{Z}$.

Supposons désormais (W) minimale. Si l'on pose

$$\begin{aligned} X &= x + (a_1^2/12) + (a_2/3) \\ Y &= y + (a_1/2)x + (a_3/2)\,, \end{aligned}$$

l'équation (W) s'écrit $Y^2 = X^3 - (c_4/48)X - (c_6/864)$. Un calcul élémentaire montre que c_4, c_6 et $\Delta = (c_4^3 - c_6^2)/1728$ s'expriment comme polynômes

universels à coefficients entiers en a_1, a_2, a_3, a_4, a_6, donc sont *entiers*. Ces entiers *ne dépendent pas du choix de* (W), mais seulement de la courbe elliptique. On dit que Δ est le *discriminant minimal* de E.

Soit $E(\mathbf{Q})$ l'ensemble des points rationnels de E (i.e. les solutions $(x, y) \in \mathbf{Q}^2$ de l'équation (W), auxquelles on ajoute le point à l'infini 0). Il existe une unique structure de *groupe abélien* sur $E(\mathbf{Q})$, d'élément neutre 0, pour laquelle trois points de $E(\mathbf{Q})$ ont une somme nulle si et seulement si ce sont les points d'intersection (avec multiplicités) de E et d'une droite du plan projectif.

Pour obtenir des informations sur les solutions rationnelles de l'équation (W), on est amené à étudier le groupe $E(\mathbf{Q})$. Je pense qu'il n'est pas exagéré de prétendre que la majeure partie des travaux effectués et des notions introduites dans la théorie des courbes elliptiques ont pour but ultime de décrire $E(\mathbf{Q})$. Un théorème important dans cette direction est le théorème de Mordell-Weil: *le groupe* $E(\mathbf{Q})$ *est de type fini*, et est par suite isomorphe à $F \times \mathbf{Z}^r$ où F est un groupe fini et r un entier $\geqslant 0$ (que nous appellerons le rang de $E(\mathbf{Q})$). On a des informations précises sur F à la suite de travaux de Mazur (par exemple, on sait que F est d'ordre $\leqslant 16$); par contre, r reste pour l'instant mystérieux (on ne sait même pas s'il peut prendre des valeurs arbitrairement grandes, bien que l'on pense que tel est le cas).

Comme les coefficients de l'équation (W) sont entiers, on peut réduire cette équation modulo un nombre premier p, puis compter le nombre de ses solutions (x, y) dans $(\mathbf{Z}/p\mathbf{Z})^2$. Ce nombre ne dépend pas du choix de (W), mais seulement de E. D'après un théorème de Hasse, il est de la forme $p - a_p$ où a_p satisfait à l'inégalité

$$|a_p| < 2\sqrt{p}\,. \tag{31}$$

La *fonction* L_E *de Hasse-Weil* associée à la courbe elliptique est par définition la série de Dirichlet

$$L_E(s) = \prod_{p \mid \Delta} (1 - a_p p^{-s})^{-1} \prod_{p \nmid \Delta} (1 - a_p p^{-s} + p^{1-2s})^{-1}\,. \tag{32}$$

Ce produit converge pour $\mathrm{Re}\,(s) > 3/2$ d'après (31). Un cas particulier de conjectures générales sur les fonctions L associées à des variétés algébriques est:

CONJECTURE 1. *La fonction* $\Lambda_E(s) = (2\pi)^{-s}\Gamma(s)L_E(s)$ *admet un prolongement holomorphe à* $\mathbf{C}$, *borné dans toute bande verticale, et il existe* $\varepsilon_E \in \{-1, 1\}$ *et un entier* $N_E \geqslant 1$ *tels que* $\Lambda_E(2-s) = \varepsilon_E N_E{}^{s-1}\Lambda_E(s)$.

Posons $L_E(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ et définissons sur le demi-plan de Poincaré $\{\tau \in \mathbf{C} \mid \mathrm{Im}(\tau) > 0\}$ une fonction f_E par

$$f_E(\tau) = \sum_{n=1}^{\infty} a_n e^{2\pi i n \tau}. \tag{33}$$

La théorie de Hecke, qui s'appuie sur la transformation de Mellin $\Lambda_E(s) = \int_0^{\infty} f_E(iy)\, y^{s-1} dy$, permet de montrer l'équivalence entre la conjecture 1 et la suivante:

CONJECTURE 1'. *Il existe* $\varepsilon_E \in \{-1, 1\}$ *et un entier* $N_E \geqslant 1$ *(les mêmes qu'avant) tels que* $f_E(-1/N_E\tau) = -\varepsilon_E N_E \tau^2 f_E(\tau)$.

On dispose de conjectures étendant la conjecture 1 aux séries $L_E(\chi, s) = \sum_{n=1}^{\infty} a_n \chi(n) n^{-s}$, avec χ caractère de Dirichlet. Généralisant le travail de Hecke, Weil [1]) a montré que ces conjectures pour tous les χ (ou même seulement pour une famille assez grande de χ) équivalent à la suivante sur f_E:

CONJECTURE 2 (Taniyama-Weil) [2]). *La fonction* f_E *satisfait à la conjecture* 1' *et est une forme modulaire parabolique de poids* 2 *pour* $\Gamma_0(N_E)$.

[La dernière assertion signifie que $f_E((a\tau+b)/(c\tau+d)) = (c\tau+d)^2 f(\tau)$ si $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ appartient au sous-groupe $\Gamma_0(N_E)$ de $SL_2(\mathbf{Z})$ formé par les matrices telles que N_E divise c, et que la fonction $\tau \mapsto f(\tau)\,\mathrm{Im}\,\tau$ est bornée sur le demi-plan de Poincaré.]

Une courbe elliptique E définie sur $\mathbf{Q}$ qui satisfait à la conjecture 2 est appelée *courbe elliptique modulaire* ou *courbe de Weil*. On sait que si la courbe E est à multiplications complexes, elle est de Weil. D'autre part, étant donnée une courbe elliptique E, il existe des algorithmes permettant de déterminer si elle est ou non une courbe de Weil. Cela a été appliqué à de nombreux exemples et toutes les courbes elliptiques étudiées se sont avérées être des courbes de Weil, conformément aux conjectures.

---

[1]) A. WEIL, *Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen*, Math. Ann. 168 (1967), 149-156.

[2]) Lorsque cette conjecture est satisfaite, f_E est une newform au sens d'Atkin-Lehner, d'après un théorème de W. Li; l'entier N_E est le conducteur géométrique de la courbe elliptique E, d'après un théorème de Carayol; en particulier, les facteurs premiers de N_E sont les mêmes que ceux du discriminant minimal de E.

Birch et Swinnerton-Dyer ont émis une autre conjecture, stupéfiante car elle relie la fonction L_E, définie à partir des nombres de solutions de l'équation (W) sur les corps finis, au rang r de $E(\mathbf{Q})$ qui fournit une information sur les solutions rationnelles de l'équation (W). Cette conjecture suppose implicitement la conjecture 1 satisfaite:

CONJECTURE 3 (Birch et Swinnerton-Dyer). *Le rang r de $E(\mathbf{Q})$ est égal à l'ordre du zéro de la fonction L_E au point* 1.

(Birch et Swinnerton-Dyer donnent en outre une expression conjecturale de $\lim_{s\to 1}(s-1)^r L_E(s)$.)

§ 7. LE THÉORÈME DE GOLDFELD

Un pas décisif vers la solution effective du problème du nombre de classes a été franchi par Goldfeld en 1976. L'idée à la base de son travail est la suivante: Supposons que nous connaissions une série de Dirichlet $\sum_{n=1}^{\infty} a_n n^{-s}$ telle que pour tout caractère de Dirichlet $\chi: n \mapsto \left(\frac{-d}{n}\right)$ avec $-d$ discriminant fondamental, la série $\sum_{n=1}^{\infty} a_n \chi(n) n^{-s}$ ait un comportement analytique très différent de la série $\sum_{n=1}^{\infty} a_n \lambda(n) n^{-s}$ où λ est la fonction multiplicative introduite à la fin de II, § 2. On peut alors espérer d'après le principe de II, § 2, montrer de façon effective que lorsque d est grand, $h(-d)$ ne peut être petit.

De fait, Goldfeld montre [1]) qu'il suffit de *connaître une seule courbe elliptique E* définie sur $\mathbf{Q}$ telle que

— E soit une *courbe de Weil*;

— la fonction L_E ait un zéro *au moins triple* au point 1,

et d'appliquer l'idée précédente à la série de Dirichlet L_E pour obtenir des minorations effectives de nombres de classes. Celles-ci sont bien moins bonnes que celles que donne l'hypothèse de Riemann généralisée (cf. § 3):

---

[1]) D. M. GOLDFELD, *The conjecture of Birch and Swinnerton-Dyer and the class number of quadratic fields*, Journées Arithmétiques de Caen, Astérisque 41-42 (1977), 219-227.

on obtient par exemple [1]) pour $h(-d)$ *impair* une inégalité de la forme

$$h(-d) \geqslant c_E \log d \tag{34}$$

où c_E est une constante dépendant de la courbe elliptique E choisie, et susceptible d'être calculée. (Plus généralement, si $h(-d)$ est de la forme $2^t h'$ avec h' impair, on a une inégalité analogue à (34) à condition de remplacer c_E par une nouvelle constante $c_E(t)$ qui dépend de t, par exemple $c_E(t) = c_E e^{-3\sqrt{t}}$, et de supposer d premier à N_E; cette dernière condition peut même être omise si l'on choisit E convenablement comme l'ont remarqué Gross et Zagier.)

Comment trouver E remplissant les deux conditions énoncées ci-dessus? On commence par choisir une courbe elliptique E telle que le groupe $E(\mathbf{Q})$ ait un rang impair $r \geqslant 3$ (il y en a une infinité et on peut en expliciter à volonté). On vérifie qu'elle est de Weil (soit parce qu'elle est à multiplications complexes, soit par un calcul sur ordinateur) et que le signe ε_E de l'équation fonctionnelle de L_E est -1 (par le calcul). La fonction L_E a alors un zéro d'ordre ρ *impair* en 1, et si l'on croit en la conjecture de Birch et Swinnerton-Dyer, ρ doit être égal à r, donc $\geqslant 3$. Malheureusement, cette conjecture n'est pas démontrée. Peut-on s'en passer et dans le cas particulier choisi, prouver directement l'inégalité $\rho \geqslant 3$? Puisque ρ est impair, cela revient à montrer que $L'_E(1) = 0$. Il est possible d'obtenir par calcul sur ordinateur des valeurs approchées de $L'_E(1)$, mais a priori même si celles-ci sont très petites on ne peut conclure à la nullité de $L'_E(1)$.

Il a fallu attendre 1983 et les travaux de Gross et Zagier pour arriver enfin à surmonter cette difficulté et à appliquer le théorème de Goldfeld.

§ 8. Le théorème de Gross et Zagier

Soit E une courbe elliptique définie sur $\mathbf{Q}$ et soit $P \in E(\mathbf{Q})$ un point rationnel de E. Ecrivons l'abscisse $x(n(P))$ du point $P + \ldots + P$ (n termes, la somme étant calculée dans le groupe $E(\mathbf{Q})$) sous forme d'une fraction irréductible a_n/b_n. On montre que l'expression $\frac{1}{2} n^{-2} \log(\sup(|a_n|, |b_n|))$ a une limite $\hat{h}(P)$ lorsque P tend vers $+\infty$, appelée *hauteur de Néron-Tate de P.*

[1]) Cette inégalité, un peu meilleure que celle de Goldfeld, est prouvée par la même méthode dans mon exposé sur la question au Séminaire Bourbaki (Juin 1984, exposé 631).

L'application $P \mapsto \hat{h}(P)$ de $E(\mathbf{Q})$ dans $\mathbf{R}$ est quadratique et positive, et l'on a $\hat{h}(P) = 0$ si et seulement si P est un point *de torsion* du groupe $E(\mathbf{Q})$.

Gross et Zagier ont obtenu en 1983 un très beau théorème [1]) qui donne *une expression de la dérivée en* 1 *de certaines fonctions* L associées à des formes modulaires. Exposons simplement le cas particulier de ce théorème qui nous intéresse pour le problème du nombre de classes: considérons comme au § 7 une courbe elliptique E de Weil, telle que le signe ε_E de l'équation fonctionnelle L_E soit -1, et notons f_E la forme modulaire associée (§ 6); il existe alors une constante réelle calculable non nulle c_E telle que:

Pour tout caractère de Dirichlet quadratique impair χ *de conducteur* $d \geqslant 7$ *tel que* $\chi(N_E) = 1$, *il existe un point* $P \in E(\mathbf{Q})$ *tel que*

$$L'_E(1)L_E(\chi, 1) = c_E\hat{h}(P)\,.$$

Ce théorème peut être utilisé pour résoudre le problème laissé en suspens au paragraphe précédent, à savoir vérifier si $L'_E(1) = 0$: pour cela, on choisit un caractère de Dirichlet χ comme ci-dessus pour lequel $L_E(\chi, 1) \neq 0$ (ceci est toujours possible, d'après un théorème de Waldspurger, et on trouve facilement un tel χ lorsque E est choisie). Comme on dispose de majorations de $L_E(\chi, 1)$, de la valeur approchée de c_E et de minorations des $\hat{h}(P)$ non nuls lorsque P décrit $E(\mathbf{Q})$, il suffit alors pour conclure à la nullité de $L'_E(1)$ de montrer que $L'_E(1)$ est assez petit, ce qu'un calcul sur ordinateur permet de faire.

§ 9. Conclusion

Gross et Zagier ont vérifié que la courbe elliptique d'équation (minimale):

$$y^2 + y = x^3 - x^2 - 450\,823x + 112\,971\,139$$

satisfait aux exigences du § 6. En calculant la constante c_E correspondante (cf. pour cela mon exposé au Séminaire Bourbaki), on obtient

$$h(-d) = 3 \Rightarrow \log d \leqslant 21\,000$$
$$h(-d) = 4 \Rightarrow \log d \leqslant 336\,000$$
$$h(-d) = 5 \Rightarrow \log d \leqslant 35\,000$$
$$h(-d) = 6 \Rightarrow \log d \leqslant 168\,000$$

etc.

---

[1]) B. H. Gross et D. B. Zagier, *Heegner points and derivatives of L-series*, Inv. Math. 84 (1986), 225-320.

D'autres courbes elliptiques de Weil E telles que $E(\mathbf{Q})$ soit de rang 3, trouvées par Mestre,

$$y^2 + y = x^3 - 7x + 6 \qquad (N_E = 5\,077)$$
$$y^2 + y = x^3 - x + 6 \qquad (N_E = 16\,811)$$
$$y^2 + y = x^3 - 19x + 30 \qquad (N_E = 43\,669),$$

permettent d'obtenir de meilleures majorations:

$$h(-d) = 3 \Rightarrow \log d \leqslant 165$$
$$h(-d) = 4 \Rightarrow \log d \leqslant 2\,640$$
$$h(-d) = 5 \Rightarrow \log d \leqslant 275$$
$$h(-d) = 6 \Rightarrow \log d \leqslant 1\,320$$

etc.

Pour achever complètement de résoudre le problème du nombre de classes, il reste en fait à vérifier qu'en-dessous des bornes précédentes les seuls d pour lesquels $h(-d)$ vaut 3, 4, 5, 6, etc. sont ceux qui figurent dans la table de Buell. Il devrait être possible de le faire en reprenant les calculs de Stark et Montgomery-Weinberger évoqués au § 5. Pour l'instant, cela n'a été fait que pour $h = 3$ (par Montgomery et Weinberger), et pour $h = 4$ (par Arno).

(Reçu le 30 mars 1987)

J. Oesterlé

Université Paris VI
UER Mathématiques
4, place Jussieu
75230 Paris Cedex 05

vide-leer-empty