JOURNAL DE THÉORIE DES NOMBRES DE BORDEAUX

HENRI COHEN

Calcul du nombre de classes d'un corps quadratique imaginaire ou réel, d'après Shanks, Williams, McCurley, A. K. Lenstra et Schnorr

Journal de Théorie des Nombres de Bordeaux, tome 1, n° 1 (1989), p. 117-135

<http://www.numdam.org/item?id=JTNB_1989__1_1_117_0>

Séminaire de Théorie des Nombres,
Bordeaux, **1**, *117–135 (1989)*

Calcul du nombre de classes d'un corps quadratique imaginaire ou réel

d'après Shanks, Williams, McCurley, A. K. Lenstra et Schnorr

par HENRI COHEN

Résumé — Dans cette note nous décrivons différentes méthodes utilisées en pratique pour calculer le nombre de classes d'un corps quadratique imaginaire ou réel ainsi que pour calculer le régulateur d'un corps quadratique réel. En particulier nous décrivons *l'infrastructure* de Shanks ainsi que la méthode sous-exponentielle de McCurley.

Abstract — *In this paper we briefly describe practical methods to compute the class number of real and imaginary quadratic fields and of the regulator of real quadratic fields. In particular we describe Shanks' infrastructure method and McCurley's subexponential algorithm for computing class numbers.*

Depuis l'introduction par Gauß du groupe des classes d'un corps quadratique (en langage moderne), le calcul de la structure et du cardinal de ces groupes a toujours été un problème intéressant. Ces dernières années, l'intérêt pour ce problème a été ravivé par plusieurs choses : les conjectures de H. W. Lenstra et de l'auteur qui prédisent quantitativement le comportement asymptotique de ces structures ; certaines méthodes de factorisation qui utilisent (parfois implicitement) ces groupes de classes, aussi bien de corps quadratiques imaginaires (Shanks, H. W. Lenstra-Schnorr) que de corps quadratiques réels (méthode SQUFOF de Shanks).

Le but principal de cet exposé est de décrire quelques unes des méthodes découvertes ces dernières années pour résoudre le problème du calcul des nombres de classes. Donnons dès à présent les principaux résultats.

On suppose la validité de l'hypothèse de Riemann généralisée (GRH). On note $h(D)$ et $R(D)$ le nombre de classes et le régulateur du corps quadratique $\mathbf{Q}(\sqrt{D})$.

• Si $D < 0$, $h(D)$ peut être calculé en temps probabiliste $O(L(|D|)^{\alpha+\epsilon})$ pour tout $\epsilon > 0$, où $\alpha = \sqrt{9/8}$ et où $L(x) = \exp(\sqrt{\log x \log\log x})$ (McCurley, A. K. Lenstra-Schnorr [5],[6]).

Manuscrit reçu le 22 mars 1989

• Si $D > 0$, $h(D)$ et $R(D)$ peuvent être calculés en temps $O(D^{1/5+\epsilon})$ pour tout $\epsilon > 0$ (essentiellement dû à Shanks ([8], [9])).

Notons que GRH est nécessaire non seulement pour obtenir des estimations raisonnables du temps d'exécution des algorithmes, mais également pour assurer l'exactitude du calcul de $h(D)$. D'autre part dans le cas $D > 0$, on ne sait pas calculer le produit $h(D)R(D)$ plus vite qu'en calculant $h(D)$ et $R(D)$ séparément.

Le résultat pour $D < 0$ semble être la limite des moyens actuels, en dehors d'une amélioration de l'exposant α qui pourra probablement être descendu à 1 : en effet, la connaissance de $h(D)$ permet de factoriser D très rapidement (voir [8]), et tous les algorithmes rapides de factorisation connus ont ce type de temps d'exécution (avec l'exposant 1). Ceci est dû au théorème de de Bruijn, Canfield-Erdős-Pomerance sur les nombres "lisses" : si on pose

$$\psi(x,y) = |\{n \leq x \ , \ p \mid n \implies p \leq y\}|,$$

alors quand $x \to \infty$ on a :

$$\frac{1}{x}\psi(x, L(x)^{\alpha}) = L(x)^{-\frac{1}{2\alpha}+o(1)}$$

Par contre, le résultat pour $D > 0$ semble plutôt montrer notre ignorance des phénomènes liés aux unités. Par exemple, il semblerait plausible que l'on arrive à calculer $h(D)R(D)$ plus rapidement que $h(D)$ et $R(D)$ séparément, et pourquoi pas en temps $L(D)^{\alpha}$.

Le plan de cet article est le suivant : dans les trois prochains paragraphes, nous exposons trois méthodes pour calculer $h(D)$. Bien qu'il soit nécessaire à chaque fois de distinguer le cas $D > 0$ du cas $D < 0$, ces méthodes sont analogues et ont des temnps d'exécution comparables dans les deux situations si on suppose le régulateur $R(D)$ déjà calculé. Dans le paragraphe 5, nous exposons la méthode de McCurley pour calculer $h(D)$ quand $D < 0$. Enfin, dans le sixième et dernier paragraphe, nous exposons la méthode de Shanks, améliorée depuis par Williams et Lenstra, pour calculer le régulateur $R(D)$.

Notons que le calcul d'une *table* de $h(D)$ pose des problèmes un peu différents, et que dans ce cas l'utilisation de méthodes individuelles comme celles présentées ici est rarement la meilleure solution.

Signalons enfin que la plupart de ces méthodes se généralisent à des corps de nombres quelconques, au prix de beaucoup d'efforts. Voir par exemple la thèse de Buchmann [1].

2. Utilisation des formes quadratiques réduites.

2.1. Corps quadratiques imaginaires.

Soit $K = \mathbf{Q}(\sqrt{D})$ un corps quadratique imaginaire, où $D < 0$ est le discriminant du corps. Rappelons qu'une forme quadratique

$$ax^2 + bxy + cy^2$$

avec $a, b, c \in \mathbf{Z}$, $a > 0$, $(a, b, c) = 1$ de discriminant $D = b^2 - 4ac$ est *réduite* si $|b| \leq a \leq c$ et si quand $|b| = a$ ou $a = c$ on a en plus $b \geq 0$. Il revient au même de dire que

$$\tau = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{-b + i\sqrt{|D|}}{2a} \in \mathfrak{F},$$

le domaine fondamental habituel de $\mathfrak{H}/SL_2(\mathbf{Z})$. Par abus de notation, nous écrirons (a, b, c) pour la forme quadratique ayant ces coefficients.

Il est facile de montrer qu'à toute classe d'idéaux de $\mathbf{Q}(\sqrt{D})$ correspond exactement une forme réduite, donc que $h(D)$ est égal au nombre de formes réduites. Pour compter ces formes, on remarque que les conditions de réduction impliquent que $|D| = 4ac - b^2 \geq 4a^2 - a^2$, donc que $a < \sqrt{|D|/3}$. On remarque de plus que $b \equiv D \pmod 2$. Pour chaque couple de valeurs possibles de (a, b) on calcule alors $c = (b^2 - D)/(4a)$ et on vérifie que $c \in \mathbf{Z}$ et que les autres conditions de réduction sont satisfaites.

Prenons par exemple $D = -23$. On a alors $a \leq 2$ donc $a = 1$ ou $a = 2$. D'autre part b doit être impair. Les conditions de réduction montrent immédiatement qu'il y a exactement 3 formes réduites, à savoir $(1, 1, 6)$, $(2, 1, 3)$ et $(2, -1, 3)$, donc $h(-23) = 3$.

Tel qu'il est décrit ci-dessus, le temps d'exécution de cet algorithme est $O(D)$, ce qui est assez lent. Toutefois cet algorithme peut être utilisé efficacement pour la construction de tables, et d'autre part il est possible de l'améliorer au prix de grosses complications, et de faire descendre son temps d'exécution à $O(D^{1/2+\epsilon})$. Comme nous verrons de meilleurs algorithmes par la suite, je ne donne pas le détail de ces améliorations.

2.2. Corps quadratiques réels.

Nous conservons les notations ci-dessus, sauf que maintenant le discriminant D est positif. La plus grande part des difficultés que l'on rencontre dans ce cas provient de la non nullité du rang du groupe des unités. Dans le cas qui nous préoccupe, nous allons voir que le comptage des formes réduites va être remplacé par le comptage des *cycles* de formes réduites.

Dand le cas $D > 0$, on dit qu'une forme quadratique

$$ax^2 + bxy + cy^2$$

avec a, b, $c \in \mathbb{Z}$, $(a,b,c) = 1$ de discriminant $D = b^2 - 4ac$ est *réduite* si $b > 0$ et si on a

$$\sqrt{D} - 2\inf(|a|,|c|) < b < \sqrt{D}.$$

Si on pose

$$\tau = \frac{-b + \sqrt{b^2 - 4ac}}{2a} = \frac{-b + \sqrt{D}}{2a},$$

il revient au même de dire que $|\tau| < 1$ et $|\overline{\tau}| > 1$, où $\overline{\tau}$ désigne le conjugué réel de τ. (Plus généralement, on peut définir la notion d'idéal réduit dans tout corps de nombres).

Dans le cas quadratique imaginaire, à toute classe d'idéaux correspond exactement une forme réduite. Ce n'est plus le cas ici : à toute classe d'idéaux correspond un *cycle* de formes quadratiques réduites, ce cycle correspondant simplement à une période du développement en fraction continue de $|\overline{\tau}|$. La présence de ces cycles correspond évidemment aux unités, et on peut obtenir l'unité fondamentale (ou le régulateur) en même temps si on le désire.

En fait, le changemnt de (a,b,c) en $(-c,b,-a)$ correspondant à l'unité triviale -1, il ne faut prendre que les cycles modulo cette équivalence. Le nombre de tels (classes d'équivalence de) cycles est alors égal au nombre de classes $h(D)$.

Notons qu'il est facile de montrer que la définition de la réduction implique que a, b et c sont inférieurs à $\sqrt{D}$. Ceci contraste avec le cas imaginaire où c peut être beaucoup plus grand. Cette remarque a des conséquences algorithmiques importantes : dans beaucoup d'algorithmes ayant trait de près ou de loin avec les corps quadratiques réels, on peut en général s'arranger pour que les calculs soient effectués avec des nombres inférieurs à $\sqrt{D}$. Ceci représente un gain de temps considérable, au minimum d'un facteur 4, et beaucoup plus quand le passage de $\sqrt{D}$ à D fait passer de la simple précision à la multiprécision. Un exemple notable est l'algorithme de factorisation SQUFOF de Shanks qui est d'une remarquable efficacité quand $D < 10^{16}$ sur une calculette.

Donnons un exemple très simple de calcul de $h(D)$. Prenons $D = 60$, donc $K = \mathbb{Q}(\sqrt{15})$. On a $\lfloor\sqrt{D}\rfloor = 7$ et b doit être pair, donc les valeurs possibles de b sont 0, 2, 4, 6. Ceci donne respectivement $ac = -15$, -14, -11, -6. Sachant que $|a| \leq 7$ et $|c| \leq 7$ il nous reste les formes $(3,0,-5)$,

$(-5,0,3)$, $(2,2,-7)$, $(-7,2,2)$, $(2,6,-3)$, $(-3,6,2)$, $(1,6,-6)$, $(-6,6,1)$, après élimination des formes trivialement équivalentes. Parmi ces 8 formes, les 4 premières ne vérifient pas l'inégalité $\sqrt{D} - 2\inf(|a|,|c|) < b$ donc sont à éliminer ; enfin, les développements en fraction continue de $\dfrac{6+\sqrt{D}}{4}$ et de $\dfrac{6+\sqrt{D}}{2}$ montrent que l'on a les deux cycles $\{(1,6,-6),(-6,6,1)\}$ et $\{(2,6,-3),(-3,6,2)\}$, donc $h(60) = 2$.

Comme dans le cas imaginaire, on peut améliorer cet algorithme pour obtenir un temps d'exécution en $O(D^{1/2+\epsilon})$ pour tout $\epsilon > 0$. Vu que le nombre de formes réduites a aussi cet ordre de grandeur, il est clair que l'on ne peut pas espérer faire mieux par simple comptage.

3. Utilisation de la formule analytique de Dirichlet.

Nous poserons

$$L_D(s) = \sum_{n\geq 1} \left(\frac{D}{n}\right) n^{-s}$$

et prolongée analytiquement.

3.1. Corps quadratiques imaginaires.

Nous supposerons $D < -4$ (on a bien sûr $h(-3) = h(-4) = 1$). Dans ce cas on a la formule de Dirichlet :

$$L_D(1) = \frac{\pi h(D)}{\sqrt{|D|}}.$$

Une sommation convenable fournit la formule équivalente

$$h(D) = \frac{1}{D} \sum_{1\leq r<|D|} r\left(\frac{D}{r}\right).$$

Toutefois l'utilisation de cette formule ou de formules similaires donne un algorithme en $O(|D|^{1+\epsilon})$, donc très peu efficace.

Une amélioration considérable peut être obtenue en utilisant *l'équation fonctionnelle* de $L_D(s)$: si l'on pose

$$\Lambda_D(s) = |D/\pi|^{(s+1)/2}\Gamma((s+1)/2)L_D(s),$$

alors on a

$$\Lambda_D(1-s) = \Lambda_D(s).$$

Cette équation est essentiellement équivalente au fait que $\Lambda_D(s-1)$ est la transformée de Mellin d'une fonction theta, à savoir

$$\Theta_D(\tau) = \sum_{n\geq 1} n\left(\frac{D}{n}\right) q^{n^2}.$$

Ceci conduit à la formule suivante. Posons

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt$$

(fonction d'erreur complémentaire). Alors pour $D < -4$ on a :

$$h(D) = \sum_{n\geq 1} \left(\frac{D}{n}\right) \left(\operatorname{erfc}\left(n\sqrt{\frac{\pi}{|D|}}\right) + \frac{\sqrt{|D|}}{\pi n} e^{-\frac{\pi n^2}{|D|}} \right). \tag{1}$$

Remarquons que la fonction $\operatorname{erfc}(x)$ est très facile à calculer à l'aide des formules suivantes :

Si x est petit (par exemple $x \leq 2$)

$$\operatorname{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}} \sum_{k\geq 0} (-1)^k \frac{x^{2k+1}}{k!(2k+1)},$$

et si x est grand (par exemple $x \geq 2$)

$$\operatorname{erfc}(x) = \frac{e^{-x^2}}{x\sqrt{\pi}} \left(1 - \cfrac{1/2}{2 + X - \cfrac{1\cdot 3/2}{4 + X - \cfrac{2\cdot 5/2}{6 + X - \ddots}}} \right),$$

où $X = x^2 - 1/2$.

Ceci implique en particulier que

$$\operatorname{erfc}\left(n\sqrt{\frac{\pi}{|D|}}\right) \sim \frac{\sqrt{|D|}}{\pi n} e^{-\frac{\pi n^2}{|D|}}$$

et que

$$\operatorname{erfc}(x) < \frac{e^{-x^2}}{x\sqrt{\pi}}.$$

La convergence de la série (1) étant exponentielle et $h(D)$ étant un entier, il est clair que le temps de calcul de $h(D)$ par cette série est $O(|D|^{1/2+\epsilon})$ pour tout $\epsilon > 0$. Par exemple, il est facile de montrer que $h(D)$ est l'entier le plus proche de la somme partielle de la série tronquée à $n = \left\lfloor \sqrt{|D|\log|D|/(2\pi)} \right\rfloor$. Ceci est donc un deuxième algorithme pour calculer $h(D)$ en temps $O(|D|^{1/2+\epsilon})$, mais nettement plus simple que le précédent. Toutefois nous sommes encore bien loin de ce qu'on peut faire de mieux.

3.2. Corps quadratiques réels.

Dans le cas où $D > 0$, on a des formules analogues. Tout d'abord, la formule de Dirichlet s'écrit

$$L_D(1) = \frac{2h(D)R(D)}{\sqrt{D}},$$

où $R(D) = \log\varepsilon$ est le régulateur, $\varepsilon > 1$ étant l'unité fondamentale.

Mettons de côté provisoirement le problème du calcul de $R(D)$: on sait que l'utilisation du développement en fraction continue de $\sqrt{D}/2$ si $D \equiv 0 \pmod 4$, ou de $(1+\sqrt{D})/2$ si $D \equiv 1 \pmod 4$ fournit aisément $R(D)$ en un temps $O(D^{1/2+\epsilon})$. Nous verrons une meilleure méthode au paragraphe 6.

Une sommation convenable fournit la formule équivalente

$$h(D)R(D) = -\sum_{r=1}^{\lfloor (D-1)/2 \rfloor} \left(\frac{D}{r}\right) \log\sin\left(\frac{r\pi}{D}\right).$$

Comme dans le cas imaginaire, l'utilisation de l'équation fonctionnelle de $L_D(s)$ conduit à une amélioration considérable. Si l'on pose

$$\Lambda_D(s) = (D/\pi)^{s/2}\Gamma(s/2)L_D(s),$$

alors on a à nouveau

$$\Lambda_D(1-s) = \Lambda_D(s).$$

Ici Λ_D est la transformée de Mellin de la fonction theta

$$\Theta_D(\tau) = \sum_{n\geq 1} \left(\frac{D}{n}\right) q^{n^2},$$

et nous conduit à la formule suivante. Posons

$$E_1(x) = \int_x^\infty \frac{e^{-t}}{t}\,dt$$

(intégrale exponentielle). Alors on a :

$$(2) \qquad 2h(D)R(D) = \sum_{n\geq 1}\left(\frac{D}{n}\right)\left(\frac{\sqrt{D}}{n}\operatorname{erfc}\left(n\sqrt{\frac{\pi}{D}}\right) + E_1\left(\frac{\pi n^2}{D}\right)\right).$$

Nous avons vu ci-dessus comment calculer $\operatorname{erfc}(x)$ efficacement. On peut faire de même pour $E_1(x)$ à l'aide des formules suivantes :

Si x est petit (par exemple $x \leq 4$)

$$E_1(x) = -\gamma - \log(x) + \sum_{k\geq 1}(-1)^{k-1}\frac{x^k}{k!k},$$

où $\gamma = 0.57721566490153286\ldots$ est la constante d'Euler, et si x est grand (par exemple $x \geq 4$)

$$E_1(x) = \frac{e^{-x}}{x}\left(1 - \cfrac{1}{2+x-\cfrac{1\cdot 2}{4+x-\cfrac{2\cdot 3}{6+x-\ddots}}}\right),$$

Ceci implique en particulier que

$$E_1(x) \sim \frac{e^{-x}}{x}$$

et que

$$E_1(x) < \frac{e^{-x}}{x}.$$

La convergence de la série étant exponentielle et $h(D)$ **étant un entier**, il est clair que le temps de calcul de $h(D)$ quand $R(D)$ est connu est encore en $O(D^{1/2+\epsilon})$. Comme dans le cas imaginaire, on pourrait très facilement

expliciter le nombre de termes que l'on doit prendre dans la série avant de tronquer.

3.3. Remarque.

Il est à noter que les méthodes décrites dans ce paragraphe, qui sont connues depuis fort longtemps, s'appliquent en principe dans tous les cas où l'on a une série de Dirichlet avec équation fonctionnelle. Par exemple Friedman [4] utilise cette méthode pour trouver les plus petits régulateurs de tous les corps de nombres. Comme autre exemple, citons le cas des valeurs aux entiers critiques des fonctions L attachées à des formes modulaires (les cas ci-dessus correspondant aux fonctions Θ_D). Dans le cas où la fonction est de poids entier, il n'y a même pas de fonction spéciale du type erfc ou E_1. Par exemple pour une courbe elliptique de Weil E de conducteur N dont l'équation fonctionnelle a le signe $+$, on a :

$$L_E(1) = 2\sum_{n\geq 1} \frac{a(n)}{n} e^{-2\pi n/\sqrt{N}},$$

où $\sum a(n)q^n$ est la forme modulaire de poids 2 correspondant à E (i.e. $L_E(s) = \sum a(n)n^{-s}$).

Notons que d'après la conjecture de Birch et Swinnerton-Dyer, ceci permet de calculer l'ordre du groupe de Tate-Shafarevitch Ш quand le rang est nul, ce groupe jouant ici un rôle analogue au rôle joué par le groupe de classes dans le cas des corps de nombres.

4. Utilisation de la structure de groupe.

L'idée supplémentaire nécessaire pour aller plus vite que les algorithmes précédents est d'utiliser aussi la structure de groupe sur les classes d'idéaux, ou ce qui revient au même, sur les classes de formes quadratiques. Aussi surprenant que cela puisse paraître, il a fallu attendre 1969 pour que cette idée naturelle apparaisse. Ceci est dû à Shanks [8]. C'est d'autant plus étonnant que la loi de groupe est connue, sous une forme légèrement différente, depuis 1800.

La première idée de Shanks est de trouver une valeur approchée de $L_D(1)$ grâce à l'utilisation de la formule analytique de Dirichlet et du produit Eulerien de la fonction L_D. La deuxième idée, beaucoup plus novatrice et utilisable dans de nombreux autres contextes, est sa méthode dite "baby step giant step". Il est préférable de regarder séparément les cas $D < 0$ et $D > 0$ bien que les principes soient quasiment identiques.

4.1. Corps quadratiques imaginaires.

Pour un P convenable à choisir ultérieurement, on pose

$$\tilde{h} = \left\lfloor \frac{\sqrt{|D|}}{\pi} \prod_{p \le P} \left(1 - \frac{\left(\frac{D}{p}\right)}{p}\right)^{-1} \right\rfloor .$$

On approche ainsi le produit Eulerien et non la série de Dirichlet, ce qui donne un gain assez faible mais qui n'est pas à négliger.

Soit maintenant $p > 2$ un nombre premier petit ($p < \sqrt{|D|/4}$ par exemple) tel que $\left(\frac{D}{p}\right) = +1$. Il existe alors une forme réduite $F_1 = (p, b, c)$ de discriminant D. L'idée fondamentale pour trouver $h = h(D)$ est que l'on doit avoir

$$F_1^h = I,$$

où I est la forme identité (qui est la seule forme réduite du type $(1, b, c)$). L'opération de multiplication sur les formes quadratiques, due à Gauß, peut être effectuée rapidement grâce aux algorithmes COMPOS [8] ou NUCOMP [10] de Shanks.

Si $\delta = \tilde{h} - h$ et $G = F_1^{\tilde{h}}$ on a donc $G = F_1^{\delta}$. Réciproquement, si l'on trouve un (petit) $\delta \ge 0$ tel que $G = F_1^{\pm\delta}$, on en déduit que $F_1^{\tilde{h} \pm \delta} = I$, donc que l'ordre de F_1 divise $\tilde{h} \pm \delta$. Si l'ordre de F_1 (qu'il est maintenant facile de calculer) n'est pas trop petit, cela implique que $h = \tilde{h} \pm \delta$. Sinon, on peut recommencer avec une deuxième forme F_2 obtenue à partir d'un nouveau nombre premier p.

Reste à expliquer comment trouver des coïncidences du type

$$G = F^{\pm\delta}.$$

Notons tout d'abord que si $G = (a, b, c)$ alors $G^{-1} = (a, -b, c)$ ce qui divise par 2 le travail. Mais d'autre part pour $D < 0$ on doit s'attendre à ce que $h(D)$ soit de l'ordre de grandeur de $\sqrt{|D|}$, donc δ aussi à moins d'avoir poussé le produit Eulerien vraiment loin. L'idée consistant à calculer les puissances successives $F^1, F^2, \ldots$ jusqu'à l'obtention d'une coïncidence est donc raisonnable puisqu'elle fournit à nouveau un algorithme en $O(|D|^{1/2+\epsilon})$. Mais l'apport fondamentalement novateur de Shanks est l'introduction de la méthode "baby step giant step".

L'idée est la suivante. On suppose que l'on sait déterminer une borne supérieure A pour la valeur δ cherchée (on aura $A = O(|D|^{1/2})$). On pose alors $k = \left\lceil \sqrt{A} \right\rceil$, $F_g = F^k$, et on calcule $F_g, F_g^2, \ldots F_g^k$. En temps $O(k \log k)$

on ordonne ces formes d'une façon où d'une autre. Si $\pm\delta = kq + r$ est la division euclidienne de $\pm\delta$ par k, on a $0 \leq r < k$ et $|q| \leq k$. La coïncidence $G = F^{\pm\delta}$ peut donc se réécrire

$$F_g^q = GF^{-r}$$

. Elle peut donc se tester en calculant les k formes GF^{-r} pour $0 \leq r < k$ et en les comparant aux formes précalculées et surtout triées F_g^q (Noter que si les formes n'avaient pas été triées, chaque comparaison aurait pris un temps en $O(k)$, alors que triées, une recherche dichotomique ne prend que $O(\log k)$).

Un certain nombre de détails doivent bien sûr être précisés, mais il est clair que cet algorithme necessite un temps en $O(A^{1/2+\epsilon}) = O(|D|^{1/4+\epsilon})$, ce qui est bien mieux que tous les précédents. De plus, si l'on suppose GRH, on peut obtenir un temps en $O(|D|^{1/5+\epsilon})$, en choisissant la borne P du produit Eulerien égale à $|D|^{1/5}$.

4.2. Corps quadratiques réels.

Sans entrer dans les détails, disons brièvement ce qui se passe dans le cas des corps quadratiques réels. Tout d'abord, il faut noter que si l'on ne fait aucune hypothèse de type GRH, les méthodes en $O(D^{1/2+\epsilon})$ décrites aux paragraphes précédents sont les plus rapides connues. Désormais nous supposerons donc GRH.

L'idée est maintenant la suivante. Le produit $h(D)R(D)$ est de l'ordre de $O(\sqrt{D})$, mais ici en moyenne $h(D)$ est très petit (voir par exemple les heuristiques de [2]). L'idée initiale de Shanks est donc à nouveau d'utiliser la formule de Dirichlet, ici

$$h(D) = \frac{\sqrt{D}}{2R(D)} \prod_p \left(1 - \frac{\left(\frac{D}{p}\right)}{p}\right)^{-1},$$

mais d'utiliser aussi le fait que $h(D)$ est un petit entier. On calcule par exemple le produit partiel par bloc de 500 nombres premiers, et on s'arrête lorsque 6 valeurs successives donnent un résultat proche du même entier à moins de 0.1 près.

Cette méthode et ces valeurs (500, 6, 0.1) sont a priori totalement empiriques. Toutefois on doit noter les faits suivants :

• La méthode est très rapide et en pratique donne toujours le bon résultat.

• On peut obtenir une version rigoureuse de cette méthode en utilisant GRH (voir [11]).

• Cette dernière version, couplée avec la méthode "baby step giant step" de Shanks donne un algorithme qui permet de calculer $h(D)$ en temps $O(D^{1/5+\epsilon})$ à partir de la connaissances de $R(D)$, qui lui aussi peut être calculé en temps $O(D^{1/5+\epsilon})$, d'où le résultat annoncé dans l'introduction. Je renvoie à [11] pour tous les détails.

4.3. Remarque.

La méthode "baby step giant step" de Shanks peut être utilisée dans tout contexte où l'on désire calculer l'ordre d'un groupe, et que l'on a seulement une borne supérieure ou une approximation de cet ordre. Par exemple, dès que le nombre premier p n'est pas trop petit, cette méthode est très efficace pour calculer le nombre de points d'une courbe elliptique modulo p. En effet, dans ce cas on sait que ce nombre N_p vérifie $|p+1-N_p| < 2\sqrt{p}$, ce qui fournit un algorithme en $O(p^{1/4+\epsilon})$, bien meilleur que l'algorithme banal consistant à compter les points en ajoutant des symboles de Legendre.

5. L'algorithme de McCurley.

Nous allons maintenant décrire un algorithme, dû initialement à McCurley, pour calculer le nombre de classes d'un corps quadratique imaginaire en temps $O(L(|D|)^{\alpha}$, donc beaucoup plus rapide que les précédents. On ne connait pas d'analogue de cet algorithme pour les corps quadratiques réels, même en supposant le régulateur calculé.

Dans l'algorithme de Shanks, ce qui nous a permis d'avancer a été l'utilisation de la relation fondamentale

$$F^h = I.$$

L'idée de McCurley est que l'obtention de *plusieurs* relations entre plusieurs formes est plus facile, et d'autre part que cela peut conduire aussi au calcul de $h(D)$. Donnons maintenant un aperçu de la méthode. Pour une description plus détaillée voir Hafner et McCurley [5].

Tout d'abord, il est indispensable de faire appel à la méthode ECM des courbes elliptiques de Lenstra, et plus précisément au "résultat" suivant :

Si tous les facteurs premiers d'un nombre n sont inférieurs ou égaux à P, n peur être factorisé en temps $O(L(P)^{\sqrt{2}})$.

Nous avons mis résultat entre guillemets car d'une part ceci n'est que probabiliste, et d'autre part l'analyse du temps d'exécution nécessite l'utilisation d'une hypothèse très plausible mais non démontrée de théorie analytique des nombres.

L'algorithme se déroule de la façon suivante. Tout d'abord on choisit une borne P de l'ordre de $O(L(|D|)^{\alpha})$ où α est à optimiser. Pour tout p premier tel que $p \leq P$ et $\left(\frac{D}{p}\right) = 1$, il existe une forme réduite $F_p = (p, b_p, c_p)$. On calcule la forme **réduite** :

$$(a,b,c) = \prod_{p \leq P,\ (\frac{D}{p})=1} F_p^{e_p}$$

où les e_p sont des exposants choisis essentiellement au hasard. Puisque (a,b,c) est réduite, on a $a \leq \sqrt{|D|/3}$ et on peut donc factoriser a en temps $O(L(|D|)^{1/\sqrt{2}})$. Si tous les facteurs premiers de a sont inférieurs ou égaux à P et si $(a,D) = 1$ on garde la forme (a,b,c), et sinon on la rejette.

Si $a = \prod_{p \leq P} p^{v_p}$, il est facile de montrer que

$$(a,b,c) = \prod_{p \leq P} F_p^{\varepsilon_p v_p},$$

où $\varepsilon_p = \pm 1$ est défini par la congruence

$$b \equiv \varepsilon_p b_p \pmod{2p}$$

(le facteur 2 n'est nécessaire que pour $p = 2$. Noter que $b_p^2 \equiv b^2 \equiv D \pmod{4p}$). On en déduit donc la relation

$$\prod_{p \leq P, (\frac{D}{p})=1} F_p^{e_p - \varepsilon_p v_p} = I.$$

Si $\mathcal{P} = \{p \leq P,\ \left(\frac{D}{p}\right) = 1,\ p \text{ premier}\}$, et $n = |\mathcal{P}|$, on a donc un vecteur

$$(e_p - \varepsilon_p v_p)_{p \in \mathcal{P}} \in \mathbb{Z}^n$$

de relations entre les F_p.

En prenant de nouveaux exposants e_p, on obtient ainsi autant de relations que l'on veut. Si l'on suppose GRH, on peut prouver qu'il existe une constante absolue effectivement caculable c telle que pour $p \leq c \log^2 |D|$ les F_p engendrent le groupe des classes. Comme P a été choisi beaucoup plus grand que cela, l'application

$$\begin{aligned} \varphi : \mathbb{Z}^n &\longrightarrow Cl(\mathbb{Q}(\sqrt{D})) \\ (x_p)_{p \in \mathcal{P}} &\longmapsto \prod_{p \in \mathcal{P}} F_p^{x_p} \end{aligned}$$

est un morphisme surjectif de groupes. Son noyau Λ est un sous réseau de $\mathbb{Z}^n$ et on a

$$\mathbb{Z}^n/\Lambda \simeq Cl(\mathbb{Q}(\sqrt{D})) \quad \text{et} \quad \det\Lambda = h(D).$$

Les relations trouvées ci-dessus sont par définition des éléments de Λ. Elles ne sont pas forcément linéairement indépendantes, mais en cas de dépendance il suffit de choisir de nouveaux exposants e_p au hasard, jusqu'à ce qu'on obtienne ainsi un système de rang n.

Le déterminant du système obtenu est alors un multiple de $\det\Lambda = h(D)$. Pour obtenir $h(D)$ il suffit d'ajouter encore quelques relations. On obtient ainsi plusieurs multiples de $h(D)$, dont on calcule le PGCD. Sous des hypothèses raisonnables, on peut *démontrer* que l'on obtiendra ainsi la valeur de $h(D)$ et non l'un de ses multiples [5]. Ceci peut d'ailleurs se tester facilement : on peut montrer qu'il existe une constante absolue c effectivement calculable telle que, si

$$\tilde{h} = \frac{\sqrt{|D|}}{\pi} \prod_{p \le c\log^2|D|} \left(1 - \frac{\left(\frac{D}{p}\right)}{p}\right)^{-1},$$

alors on a

$$\left|\log_2 \frac{\tilde{h}}{h(D)}\right| \le \frac{1}{2}.$$

On voit donc que si h est le multiple trouvé de $h(D)$, on peut tester si $h = h(D)$ en testant l'inégalité

$$h < \tilde{h}\sqrt{2}.$$

En faisant des hypothèses heuristiques raisonnables, il est clair que l'algorithme précédent a un temps d'exécution en $O(L(|D|)^{\beta})$.

Dans Lenstra-Schnorr [6], cet algorithme est légèrement modifié et ils peuvent démontrer qu'en supposant seulement GRH (et pas d'autres hypothèses heuristiques), l'algorithme a un temps probabiliste d'exécution qui est $O(L(|D|)^{\alpha+\epsilon})$ pour tout $\epsilon > 0$, avec $\alpha = \sqrt{9/8} = 1.06\ldots$ Comme il a été dit dans l'introduction, il est fort probable que l'on puisse descendre cette constante à 1 ou même à $1/\sqrt{2}$.

Remarques

1. A ma connaissance, aucun test expérimental n'a été effectué pour déterminer la valeur de D à partir de laquelle cet algorithme devient plus

efficace que la méthode du paragraphe 4.1. Si je devais deviner, je dirais pour $D > 10^{30}$.

2. Les algorithmes de McCurley et Lenstra-Schnorr peuvent a priori s'appliquer au calcul rapide de l'ordre d'un groupe abélien fini dès que l'on sait :

- Calculer explicitement dans ce groupe.
- Engendrer "au hasard" des relations entre éléments du groupe.

Ceci est expliqué en détail dans Lenstra-Schnorr[6].

6. Calcul du régulateur $R(D)$.

Nous terminerons cet exposé en expliquant comment on peut calculer rapidement le régulateur $R(D)$ d'un corps quadratique réel, donc avec $D > 0$. Indépendamment de son intérêt propre, rappelons que c'est indispensable pour calculer le nombre de classes.

6.1. La méthode traditionnelle.

Les cycles dont nous avons parlé au début correspondent aux périodes des fractions continues associées aux nombres quadratiques correspondants. Nous nous intéressons ici au cycle principal, correspondant à l'unique forme quadratique réduite représentant 1, à savoir $(1, 2q+r, q^2+qr+(r-D)/4)$, où l'on a posé $r = 0$ si $D \equiv 0 \pmod 4$, $r = 1$ si $D \equiv 1 \pmod 4$, et où $q = \lfloor (\sqrt{D} - r)/2 \rfloor$. On a alors

$$|\overline{\tau}| = \frac{2q + r + \sqrt{D}}{2}.$$

Si on développe ce nombre en fraction continue, la définition même d'une forme réduite montre que le développement sera purement périodique, soit

$$|\overline{\tau}| = [\overline{a_0, a_1, \ldots, a_m}]$$

en employant une notation standard de fractions continues. Les a_i peuvent se calculer récursivement de façon simple, sans calculs flottants, et en utilisant uniquement des entiers de taille inférieure à $\sqrt{D}$, en représentant toujours les nombres comme des nombres quadratiques. Ceci correspond bien sûr exactement au calcul du cycle principal de formes quadratiques ; nous y reviendrons.

Si l'on pose

$$\frac{p_m}{q_m} = [a_0, a_1, \ldots, a_m],$$

on sait alors que

$$\varepsilon = p_m + q_m \tau$$

est l'unité fondamentale, donc $R(D) = \log \varepsilon$ est le régulateur cherché.

La longueur de la période étant $O(D^{1/2+\epsilon})$, ceci fournit un algorithme en temps $O(D^{1/2+\epsilon})$ pour calculer $R(D)$. L'utilisation de la symétrie de $(a_1, \ldots, a_m)$ permet facilement de gagner un facteur 2 ; d'autre part comme $\varepsilon = \exp(R(D))$, il est en fait hors de question de calculer ε exactement sauf quand D est petit. La première idée nouvelle pour améliorer cette situation est la remarque suivante : dans la formule analytique de Dirichlet, seule une approximation du régulateur nous intéresse. Or celui-ci peut se calculer directement par la formule

$$R(D) = \sum_{k=0}^{m} \log\left(\frac{\rho_k}{\rho_{k-1}}\right),$$

où l'on a posé

$$\rho_k = p_k + q_k \tau$$

(et en particulier $\rho_{-1} = 1$). Or les quantités ρ_k/ρ_{k-1} peuvent être calculées très facilement par récurrence en même temps que le développement en fraction continue.

Cette idée n'améliore pas le nombre d'opérations effectuées par l'algorithme, mais améliore considérablement son temps d'exécution puisque tous les calculs peuvent être faits en flottants avec une précision raisonnable, au lieu d'une précision parfois démentielle nécessaire pour le calcul final de ε et de son log. Noter que contrairement aux apparences, la formule que nous venons de donner pour $R(D)$ ne nécessite pas le calcul de m logarithmes mais d'un seul. Nous laissons au lecteur le plaisir de découvrir l'astuce très simple à utiliser.

6.2. Quelques résultats sur les formes réduites.

Avant de voir comment la méthode ci-dessus peut être améliorée, il est nécessaire d'énoncer quelques résultats sur les formes réduites. Nous utiliserons librement l'expression "développement en fraction continue" (en abrégé dfc) aussi bien pour les nombres quadratiques que pour les formes réduites qui leur correspondent.

• Si (a, b, c) est une forme quadratique réduite, son dfc fournit le cycle de toutes les formes quadratiques réduites qui lui sont équivalentes. Le nombre de formes dans ces cycles est variable, mais la valeur du nombre

$$\sum_{k=0}^{m} \log\left(\frac{\rho_k}{\rho_{k-1}}\right)$$

calculée comme ci-dessus sur un cycle quelconque, est toujours la même, à savoir le régulateur $R(D)$.

• Si (a,b,c) est une forme primitive, c'est à dire telle que a, b, c soient premiers entre eux dans leur ensemble, son dfc conduit très rapidement à une forme réduite, en au plus

$$\max(2, 4 + \log(|c|/(2\sqrt{D}))/(2\log((1+\sqrt{5})/2)))$$

étapes (donc en $O(\log D)$). Ceci correspond à une majoration de la prépériode du dfc d'un nombre quadratique.

6.3. L'infrastructure d'un cycle.

Nous en arrivons à une idée extrêmement originale et importante de Shanks : ce qu'il appelle "l'infrastructure" d'un cycle de formes réduites. Très schématiquement, on a envie de dire que l'ensemble de toutes les formes réduites de discriminant D forme un "groupe", que le cycle principal est un "sous-groupe cyclique", les autres cycles étant les classes d'équivalence modulo ce sous-groupe, et le groupe des classes étant donc le "groupe quotient" du groupe des formes réduites par le sous-groupe cyclique des formes principales. Cette intérpretation est d'ailleurs à la base des heuristiques faites avec Lenstra[2] et Martinet[3].

Plusieurs choses sont incorrectes dans ces énoncés, mais le principal défaut est que le produit de deux formes réduites n'est en général pas une forme réduite. La première remarque de Shanks est que c'est quand même presque vrai : grâce au résultat cité au paragraphe 6.2., on sait qu'en utilisant au plus $O(\log D)$ pas de réduction, on peut se ramener à une forme réduite. Cela conduit à écrire

$$f_3 = \phi(f_1 \cdot f_2),$$

où f_1, f_2, f_3 sont trois formes réduites, et où ϕ représente le petit nombre de réductions nécessaires pour réduire le produit $f_1 \cdot f_2$. Bien sûr, étant donnés f_1 et f_2, ϕ et f_3 ne sont pas uniques, mais on suppose le nombre de réductions choisi en $O(\log D)$.

La deuxième remarque de Shanks est certainement la plus importante. Considérons par exemple le cycle principal. Soit f_0 l'unique forme réduite représentant 1 (donnée explicitement au paragraphe 6.1). On définit la **distance** d'une forme f_k à f_0 en posant

$$d(f_k) = \log|\rho_{k-1}|,$$

où ρ_k a été défini en 6.1. (Noter que $d(f_0) = 0$.)

Si $f_3 = \phi(f_1 \cdot f_2)$, Shanks remarque alors que

$$d(f_3) \simeq d(f_1) + d(f_2),$$

donc que cette notion de distance est approximativement additive. En fait on a plus précisément

$$d(f_3) = d(f_1) + d(f_2) - \delta(\phi),$$

où $\delta(\phi)$ se calcule de façon analogue à $d(f)$ en fonction du (court) dfc donnant ϕ. Grâce à cette remarque, on peut maintenant parcourir le développement en fraction continue en faisant des grand sauts ("giant steps"), tout simplement en multipliant par une forme située loin de f_0. On peut alors appliquer une variante de la méthode "baby step giant step" et obtenir ainsi un algorithme en $O(D^{1/4+\epsilon})$ opérations. De même que pour le calcul du nombre de classes, l'utilisation de GRH conduit même à un algorithme en $O(D^{1/5+\epsilon})$ opérations, mais qui n'est pas très utilisable en pratique. Les détails de l'implémentation de cet algorithme sont donnés dans [11]. Voir également [6].

References

1. J. Buchmann, *Zur Komplexität der Berechnung von Einheiten und Klassenzahlen algebraischer Zahlkörper*, Habilitationsschrift, Düsseldorf, Oktober 1987.
2. H. Cohen and H. W. Lenstra, *Heuristics on class groups of number fields*, Number Theory (Noordwijkerhout 1983), Lecture Notes in Math. **1068** (1984), 33–62, Springer-Verlag, Berlin and New York.
3. H. Cohen and J. Martinet, *Class groups of number fields: Numerical heuristics*, Math. Comp. **48** (1987), 123–137.
4. E. Friedman, *Analytic formulas for the regulator of a number field*, (to appear).
5. J. L. Hafner and K. S. McCurley, *A rigorous subexponential algorithm for computation of class groups*, (to appear).
6. A. K. Lenstra and C. P. Schnorr, *On the computation of the order of finite abelian groups using random relations*, (to appear).

7. H. W Lenstra Jr., *On the computation of regulators and class numbers of quadratic fields*, Lond. Math. Soc. Lec. Notes Ser. **56** (1982), 123–150.

8. D. Shanks, *Class numbers, a theory of factorisation and genera*, Proc. Symp. Pure Math. **20** (1971), 415–440.

9. D. Shanks, *The infrastructure of a real quadratic field and its applications*, Proc. Number Theory Conference (1972), 217–224.

10. D. Shanks, *On Gauss and composition*, in "Proceedings of the 1987 conference on Number Theory, Banff" (to appear).

11. A. J. Stephens and H. C. Williams, *Computation of real quadratic fields with class number one*, Math. Comp. **51**, 809–824.

Mots clefs: .

Centre de Recherche en Mathématiques de Bordeaux,Université Bordeaux I
C.N.R.S. U.A. 226
U.F.R. de Mathématiques
351, cours de la Libération
33405 Talence Cedex, FRANCE.