N^{O} d'ordre : 1693.

THÈSE

présentée à

L'UNIVERSITÉ BORDEAUX I

ÉCOLE DOCTORALE DE MATHÉMATIQUES ET INFORMATIQUE

Par **Guillaume HANROT**

POUR OBTENIR LE GRADE DE

DOCTEUR

SPÉCIALITÉ : MATHÉMATIQUES PURES

Résolution effective d'équations diophantiennes : algorithmes et applications.

Soutenue le 28 avril 1997.

Après avis de :

M. WALDSCHMIDT, Professeur	Université Pierre et Marie Curie – Paris 6
B.M.M. DE WEGER, Professeur	Rijksuniversiteit Leiden et Erasmus Universiteit Rotterdam

Devant la commission d'examen formée de :

P. FLAJOLET, Directeur de Recherche	INRIA Rocquencourt	Président
L. HABSIEGER, Chargé de Recherche	A2X (UMR CNRS 9936)	Rapporteur
J-M. DESHOUILLERS, Professeur	Université Victor Segalen Bordeaux 2	Examinateurs
H. IWANIEC, Professeur	Rutgers University et ETH Zürich	
J. MARTINET, Professeur	Université Bordeaux I	
M. WALDSCHMIDT, Professeur	Université Pierre et Marie Curie – Paris 6	
B.M.M. DE WEGER, Professeur	Rijksuniversiteit Leiden et Erasmus Universiteit Rotterdam	

Remerciements

Mes remerciements vont en premier lieu à Jean-Marc Deshouillers, mon directeur de thèse. Il a su faire preuve d'une grande ouverture d'esprit dans les recherches vers lesquelles il m'a orienté, et me suggérer des thèmes en accord avec mes intérêts. J'ai largement mis à profit sa disponibilité et sa vaste culture mathématique ; pendant ces années j'ai énormémement appris à son contact et à son instigation.

Yuri Bilu a été le compagnon de travail quotidien durant une grande partie de cette thèse ; compagnon souvent électronique mais non moins indéfectible, toujours patient et disponible. Merci à lui pour ces années de collaboration ; je ne peux que souhaiter la pérennité de notre association.

Michel Waldschmidt et Benne de Weger ont bien voulu se charger du lourd travail de rapporteur, qu'ils ont accompli scrupuleusement. Je les remercie l'un et l'autre pour leurs remarques constructives, ainsi que pour l'intérêt qu'ils ont manifesté pour ce travail. Merci aussi d'avoir bien voulu faire partie du jury.

Je suis reconnaissant à Laurent Habsieger et à Jacques Martinet d'avoir accepté de participer à ce jury, et pour l'intérêt que l'un et l'autre ont toujours montré pour ces algorithmes. Je profite de l'occasion pour remercier Jacques Martinet de contribuer, en assumant la direction de l'A2X, charge riche en travail mais souvent pauvre en reconnaissance, à en faire un lieu éminemment sympathique et propice à faire des mathématiques et de l'algorithmique.

Merci à Henryk Iwaniec d'avoir bien voulu se déplacer pour participer à ce jury. Je lui suis également redevable d'un fort agréable séjour à Rutgers l'année dernière, qui m'a beaucoup apporté dans la perspective de nouvelles directions de recherche.

Je remercie Philippe Flajolet d'avoir accepté de participer au jury, malgré un emploi du temps chargé ; c'est pour moi l'indication que ces travaux, qui sont de l'algorithmique à objectif mathématique peuvent avoir de l'intérêt aussi pour les informaticiens.

Merci à Henri Cohen et Michel Olivier qui ont répondu avec patience à mes (nombreuses) questions sur `pari` et sur le calcul des systèmes d'unités. Merci

aussi à eux et à leurs acolytes pour avoir mis à disposition de tous le système `pari`, sans lequel la majorité des calculs présentés dans cette thèse n'existeraient qu'à l'état de vaine spéculation.

Merci plus généralement à tout l'A2X, où il règne une excellente ambiance de travail qui est le fait de tous. Je tiens à remercier en particulier Francine Delmer, François Hennecart et Bernard Landreau dont j'ai pu apprécier l'humour, la gentillesse et l'hospitalité.

J'ai effectué la plus grande partie de ce travail à l'Université Bordeaux 2 ; la bonne humeur et la cordialité que j'y ai rencontrées ont rendu ce travail fort agréable.

Et enfin, *last but not least*, je tiens à remercier les amis qui m'ont soutenu pendant cette thèse et grâce auxquels la vie de tous les jours est apparue moins quotidienne. Je pense en particulier à K.B, Laurent, Nicolas, Gilles, David, Richard, Putu et les autres.

J'ai ainsi eu, au cours de ma vie, des tas de contacts avec des tas de gens sérieux. J'ai beaucoup vécu chez les grandes personnes. Je les ai vues de très près. Ça n'a pas beaucoup amélioré mon opinion.

Quand j'en rencontrais une qui me paraissait un peu lucide, je faisais l'expérience sur elle de mon dessin n° 1 que j'ai toujours conservé. Je voulais savoir si elle était vraiment compréhensive. Mais toujours elle me répondait « c'est un chapeau ». Alors je ne lui parlais ni de serpents boas, ni de forêts vierges, ni d'étoiles. Je me mettais à sa portée. Je lui parlais de bridge, de golf, de politique et de cravates. Et la grande personne était bien contente de connaître un homme aussi raisonnable.

Antoine de Saint-Exupéry, *Le Petit Prince.*

À mes parents et grand-parents.

Table des matières

Introduction

La modélisation mathématique d'un problème réel conduit souvent à la résolution d'une "équation", terme vague qui peut recouvrir aussi bien une équation différentielle, algébrique, aux dérivées partielles, transcendante... mais aussi diophantienne, dont les solutions fournissent la réponse au problème posé.

Les équations diophantiennes trouvent leur origine dans l'Antiquité, même si elles sont nommées en l'honneur de Diophante[1](325–409).

De manière élémentaire, une équation diophantienne[2] est une équation que l'on peut former à partir d'inconnues, de nombres entiers, et des deux opérations $+$ et $\times$. Plus formellement, étant donné un polynôme $P \in \mathbb{Z}[X_1, \ldots, X_n]$, on demande de trouver tous les n-uplets $(x_1, \ldots, x_n)$ dans $\mathbb{Z}^n$ (ou dans $\mathbb{Q}^n$; à ce niveau de généralité, il s'agit de la même question, quitte à ajouter une variable) tels que

$$P(x_1, \ldots, x_n) = 0.$$

Historiquement, le premier exemple est sans doute celui de l'équation $x^2+y^2-z^2 = 0$, qui revient à trouver les triangles rectangles dont les côtés sont entiers. L'équation $ax + by - c = 0$, où a, b, c sont fixés, a elle aussi une longue histoire.

Les équations diophantiennes, et surtout l'équation de Fermat $x^p + y^p - z^p = 0$, avec $xyz \neq 0$ ont été un des moteurs du développement de la théorie des nombres moderne ; si l'on énumère tous les outils introduits et utilisés en vue de la résolution de cette conjecture, on trouve la majorité de la théorie algébrique des nombres, des courbes elliptiques, de la géométrie algébrique... et beaucoup d'autres choses.

Parmi les 23 problèmes posés par Hilbert à Paris en 1900 comme "défis" pour le siècle à venir, le $10^{\text{ème}}$ concerne la résolution des équations diophantiennes : Hilbert demandait de construire une méthode générale de résolution ; voici le texte exact du problème, tiré de ses œuvres complètes [Hi].

---

[1] à qui l'on doit le problème de "trouver un triangle rectangle dont la somme de l'aire et de l'hypothénuse soit un carré et dont le périmètre soit un cube", ce qui se ramène à résoudre dans $\mathbb{Q}$ l'équation $y^2 = x^3 + k$.

[2] Notons que cette définition est un peu réductrice en ce qu'elle oublie les équations exponentielles, par exemple l'équation de Thue-Mahler.

> 10. Entscheidung der Lösbarkeit einer diophantischen Gleichung.
> Eine diophantische Gleichung mit irgendwelchen Unbekannten und mit ganzen rationalen Zahlkoeffizienten sei vorgelegt : *man soll ein Verfahren angeben, nach welchen sich mittels einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.*[3]

Matijasevič [Ma70] a prouvé qu'une telle méthode générale ne peut pas exister ; on doit donc se contenter de chercher des algorithmes pour certaines classes, et c'est l'objet de cette thèse.

Deux classes d'équations diophantiennes

De nombreux champs de la théorie des nombres moderne offrent des moyens d'"attaquer" une équation diophantienne. Dans cette thèse, on n'utilisera (à l'exception d'un peu de théorie algébrique des nombres rudimentaire) que les méthodes de transcendance et d'approximation diophantienne.

Avant de décrire les méthodes effectives de résolution des équations diophantiennes, faisons un bref tour d'horizon historique des rapports entre approximation diophantienne et équations diophantiennes.

De Liouville à Siegel

Le premier résultat substantiel dans la théorie de l'approximation diophantienne est dû à Liouville, qui montre dans [Li1844, Li1851] que si α est un nombre algébrique de degré $n \geq 2$, il existe une constante $C(\alpha)$ effective telle que, pour tout $(p, q) \in \mathbb{Z}^2$, on ait

$$\left|\frac{p}{q} - \alpha\right| \geq \frac{C(\alpha)}{q^n}.$$

On voit d'où vient le nom "approximation diophantienne" : on cherche à approcher un nombre réel par un nombre rationnel.

L'étape suivante dans ces résultats est le théorème de Thue [Th09], qui prouve que l'on peut remplacer $C(\alpha)/q^n$ par $C(\alpha, \varepsilon)/q^{n/2+1+\varepsilon}$, pour $\varepsilon > 0$ quelconque. Comme conséquence de ce résultat, il déduit (nous verrons comment) la finitude du nombre de solutions de l'équation diophantienne

$$f_0 Y^n + f_1 Y^{n-1} X + \ldots + f_n X^n = a,$$

où le premier membre est irréductible. L'inconvénient majeur de sa méthode réside dans l'absence d'effectivité pour $C(\alpha, \varepsilon)$. Ceci empêche d'obtenir une borne pour $\max(|X|, |Y|)$, et donc de résoudre complètement l'équation.

[3](Traduction libre) 10. Décidabilité de la résolubilité des équations diophantiennes. Une équation diophantienne à un certain nombre d'inconnues et à coefficients entiers rationnels étant donnée, *on demande de trouver une méthode qui, au moyen d'un nombre fini d'opérations, permet de dire si l'équation est résoluble en entiers rationnels.*

L'équation de Thue ci-dessus sera la première "classe d'équations" que nous allons considérer.

Plus tard, Siegel prouve [Si21] que le résultat de Liouville reste valable en remplaçant n par $2\sqrt{n}$. Toutefois, le problème d'effectivité reste, lui, entier. Par contre, cette amélioration du résultat permet de prouver que le nombre de points entiers sur une courbe de genre $g \geq 1$ est fini ; c'est ce que fait Siegel dans une lettre à Mordell [Si26] ; en particulier, les courbes $y^p = f(x)$, où $p \geq 2$ et où f est séparable de degré au moins 3, n'ont qu'un nombre fini de points entiers. Nous appellerons ces équations hyperelliptiques si $p = 2$, superelliptiques si $p \geq 3$ (encore qu'il faudrait sans doute les appeler elliptiques si $p = 2, \deg f \leq 4$ ou $p = 3 = \deg f$).

L'équation de Thue et les équations (hyper, super)-elliptiques constituent les deux classes d'équations diophantiennes pour lesquelles nous allons donner des algorithmes de résolution.

Théorie de Baker

Plusieurs améliorations du résultat de Siegel interviendront encore dans le courant du siècle, culminant avec le théorème de Roth [Ro55] (on peut prendre $2 + \varepsilon$ au lieu de n dans le théorème de Liouville), sans cependant parvenir à apporter une solution au problème de l'effectivité.

À peu près à la même période, Gelfond s'intéresse à un problème différent[4] : étant donné deux nombres algébriques α et β dont les logarithmes sont $\mathbb{Q}$-linéairement indépendants, il parvient à obtenir un minorant pour la forme en deux logarithmes $|b_1 \log \alpha + b_2 \log \beta|$.

En 1966, Baker [Ba66] obtient le même type de résultats avec n logarithmes au lieu de 2 : il considère alors $\Lambda(b_1, \ldots, b_n) = \sum_{k=1}^{n} b_i \log(\theta_i)$, où les b_i sont entiers et les θ_i algébriques. Baker montre que si le nombre $\Lambda(b_1, \ldots, b_n)$ est non nul, il ne saurait être trop petit. Plus précisément, il existe une constante C explicitement calculable et dépendant uniquement des θ_i et de n telle que

$$\Lambda(b_1, \ldots, b_n) \neq 0 \Rightarrow |\Lambda(b_1, \ldots, b_n)| \geq \exp(-C \log(\max_i |b_i|)).$$

De ce théorème, Baker [Ba68] déduit des résultats de finitude effectifs pour diverses classes d'équations diophantiennes, dont les deux classes mentionnées ci-dessus, c'est-à-dire qu'il donne des bornes pour $\max(|x|, |y|)$, où (x, y) est une solution. Ces bornes dépendent bien évidemment de manière cruciale de la constante C ; plus C sera petite, plus les bornes seront "raisonnables".

La constante initiale de Baker [Ba66] a été largement améliorée par la suite ; parmi les successeurs de Baker, on peut citer Baker lui-même [Ba72] ; puis Shorey [Sh76], Waldschmidt [Wa80] ; Blass, Glass *et al.* ont aussi donné une version dans [BGMMS90] ; enfin Baker et Wüstholz, actuels détenteurs du record

[4] et pour des raisons différentes, liées à des problèmes d'indépendance algébrique de α et e^{α}.

[BW93]. Signalons que Matveev a récemment annoncé un progrès spectaculaire ; nous n'aurons pas recours à son résultat, qui est encore en cours de validation.

Ingrédients effectifs

Il se trouve que cette constante est, dans la pratique, de taille gigantesque — à la notable exception des travaux de Laurent, Mignotte et Nesterenko [LMN95] qui ne s'appliquent toutefois qu'à des formes en deux logarithmes — et malheureusement, on voit mal comment il pourrait en être autrement. Le résultat de Baker est d'une telle généralité qu'il doit englober un certain nombre de cas "pathologiques", et qu'il est donc peu adapté à la plupart des situations pratiques. Concrètement, il ne saurait être question de vérifier un par un tous les (x, y) en deçà de la borne obtenue par la méthode de Baker.

La conséquence première de cette remarque est qu'il va nous falloir exploiter les propriétés numériques de l'équation considérée. En termes plus imagés, après avoir utilisé "l'approximation diophantienne théorique", nous allons faire de "l'approximation diophantienne effective". Pour fixer les idées, prenons le cas des formes linéaires en deux logarithmes. On cherche à minorer $|b_1 \log(\theta_1) + b_2 \log(\theta_2)|$; cela revient à minorer $|b_1 \log(\theta_1)/\log(\theta_2) + b_2|$, et sous cette forme c'est un problème d'approximation de nombres réels par des nombres rationnels ; on sait alors que les "meilleures approximations" (c'est-à-dire les fractions réalisant les minima successifs de $d(b_1 \log(\theta_1)/\log(\theta_2), \mathbb{Z})$) sont obtenues pour des fractions b_1/b_2 bien déterminées, qui sont les réduites du développement en fractions continues du nombre réel $\log(\theta_1)/\log(\theta_2)$.

On voit bien ici que la connaissance d'une borne *a priori*, aussi mauvaise soit-elle, est indispensable à la phase diophantienne effective ; en effet, on peut alors chercher la réduite de plus grand dénominateur inférieur à notre borne, qui donnera le minorant cherché. Ce minorant fournira à son tour une nouvelle borne, que l'on pourra réutiliser, etc. Cette idée d'utiliser des fractions continues est due à Baker et Davenport [BD69], et c'est historiquement la première combinaison d'idées théoriques et effectives dans la filiation de laquelle se situe cette thèse.

Il paraît toutefois difficile d'adapter telle quelle la méthode décrite ci-dessus en dimension $n > 2$. On n'a pas en effet de théorie de l'approximation diophantienne simultanée aussi jolie que la théorie des fractions continues ; en particulier, on n'a plus de manière "canonique" d'obtenir les meilleures approximations successives, et il n'est toujours pas envisageable de traiter individuellement tous les n-uplets en deçà de la borne de Baker.

Heureusement, on dispose quand même d'un outil irremplaçable dans ce contexte, qui est l'algorithme LLL, dû à Lenstra, Lenstra et Lovász [LLL82].

Cet algorithme, qui a véritablement révolutionné la théorie algorithmique des nombres et qui sert dans des domaines divers des mathématiques — voire de la physique — peut (entre autres) servir à produire des relations courtes du type ci-dessus. Contrairement à ce que fournissent les fractions continues, l'algorithme LLL ne produit pas les relations les "plus courtes", mais des relations qui ne sont

pas beaucoup plus longues que les relations les plus courtes ; comme on contrôle ce "pas beaucoup plus longues", on peut en déduire une borne inférieure pour la relation la plus courte, et l'on est à même d'en déduire une borne inférieure pour la combinaison linéaire de logarithmes, qui, remplaçant la borne de Baker, fournit une nouvelle borne supérieure pour $\max_i |b_i|$.

Petit historique des méthodes effectives

Divers mathématiciens se sont simultanément rendus compte que l'approximation "diophantienne effective" pouvait être utilisée en association avec des bornes de type Baker pour résoudre effectivement des équations diophantiennes. Ces idées ont alors été appliquées à divers problèmes : recherche de carrés, cubes et puissances cinquièmes dans diverses récurrences linéaires d'ordre 2 ; résolution d'équations de Thue de degrés 3 et 4, et application à la recherche de points entiers sur des courbes elliptiques.

Le premier à avoir eu l'idée d'utiliser l'algorithme LLL est de Weger [We87, We], qui résout en même temps l'important problème de la stabilité numérique de LLL en en donnant une version fonctionnant entièrement en nombres entiers. On obtient de cette manière une méthode relativement générique pour résoudre une équation diophantienne de type exponentiel ; comme illustration, de Weger résout l'inéquation diophantienne $0 < |x - y| < y^\delta$, où x et y ont tous leurs facteurs premiers dans un ensemble fini, et δ est un réel de $]0, 1[$; en combinant ces idées avec des résultats de minoration des formes linéaires en logarithmes p-adiques, il est à même de résoudre l'équation $x + y = z$, où les facteurs premiers de x, y, z se trouvent encore dans un ensemble fini prescrit.

L'équation de Thue

Le même de Weger, dans un travail commun avec Tzanakis [TW89], donne le premier algorithme "systématique" de résolution de l'équation de Thue. L'algorithme comporte deux étapes limitantes majeures, qui sont

- le calcul d'un système fondamental d'unités d'un corps de nombres $\mathbb{K}$ de degré n,
- la réduction LLL d'un réseau de dimension n.

Dans le premier chapitre de cette thèse, nous décrivons divers ingrédients permettant, sinon de supprimer ces problèmes, du moins d'en réduire l'influence au point que la résolution d'une équation de Thue de degré, disons 8, est devenue quasi-routinière[5] ; des exemples en degré 19 et 33 sont donnés. Il n'est plus nécessaire que de savoir calculer un système d'unités de rang maximal du corps $\mathbb{K}$, et LLL en dimension r (r désigne le rang du groupe des unités du corps $\mathbb{K}$) est remplacé soit par LLL en dimension 3, soit tout simplement par l'algorithme des fractions continues.

[5] du moins tant que les invariants du corps correspondant ne sont pas trop gros.

Notons que le record appartenait précédemment à Voutier [Vo95], qui a résolu plusieurs équations de degré allant jusqu'à 14, et encore a-t-il utilisé des propriétés particulières des équations considérées.

Les équations que Voutier a considérées sont les équations de Thue liées au sous-corps réel maximal d'un corps cyclotomique. On développe au chapitre 2 une nouvelle idée montrant comment exploiter le fait que ces corps ont "souvent" un petit sous-corps ; pour l'illustrer, on résout entre autres une équation de degré 2505.

À terme, combinés avec de nouvelles idées de Voutier, ces nouveaux ingrédients devraient permettre d'achever la résolution du problème des diviseurs primitifs des suites de Lucas-Lehmer.

Équations superelliptiques

La combinaison des idées qu'ils avaient développées pour l'équation de Thue et d'arguments analogues de type p-adique permettent à Tzanakis et de Weger de décrire un algorithme de résolution de l'équation de Thue-Mahler, c'est-à-dire de l'équation de Thue où le second membre n'est plus une constante, mais de la forme $ap_1^{z_1}p_2^{z_2}\dots p_r^{z_r}$, où les p_i sont des nombres premiers et les z_i des inconnues.

Par ailleurs, il existe une méthode relativement systématique permettant de réduire une équation superelliptique à un nombre fini d'équations de Thue (ou de Thue-Mahler si l'on veut les solutions S-entières) ; les deux algorithmes de Tzanakis et de Weger permettent donc de résoudre ce type d'équations. Notons toutefois que cette méthode n'est en général praticable que dans le cas de courbes elliptiques.

Dans ce cas existe justement une autre méthode qui consiste à utiliser non plus le groupe des unités d'un corps de nombres mais le groupe des points rationnels de la courbe elliptique que l'on souhaite étudier ; cette méthode a été décrite et mise simultanément en œuvre par Stroeker et Tzanakis [ST94, Tz95], Gebel, Pethő et Zimmer [GPZ94], et Smart [Sm94] ; voir aussi [GPZ96] pour une version S-entière.

Bilu a suggéré en 1994 [Bi94] une alternative à ces deux méthodes, qui a l'avantage sur la première de fournir moins d'équations aux unités (et donc moins de corps dans lesquels il faudra trouver des unités fondamentales), et sur la seconde de s'appliquer aux équations superelliptiques générales ; nous avons donné une version algorithmique de sa méthode, sans toutefois en modifier le fondement dans [BH96b].

De Weger [We94a] a remarqué que la méthode "alternative" peut être utilisée également pour trouver des points S-entiers, en utilisant bien sûr des formes linéaires en logarithmes p-adiques.

De comparaisons des trois méthodes dues à de Weger [SW94, We94a, We94b], il ressort qu'aucune d'elles n'est franchement supérieure aux autres ; on peut toujours trouver des cas où l'une des méthodes est très efficace, parce que le groupe sous-jacent est facile à calculer, alors que les deux autres sont peu performantes,

parce que les groupes correspondants sont compliqués. Dans l'ensemble, il semble quand même que l'approche "alternative" soit meilleure que la méthode de Thue ; quant à la méthode "elliptique", elle devrait être utilisée dès que le groupe des points rationnels n'est pas trop difficile à déterminer (dans ce contexte, on n'a qu'un seul groupe à calculer). Il est également important que le groupe soit de petit rang (disons au plus 3), sans quoi l'énumération finale se transforme en un problème insurmontable.

On illustrera la méthode alternative par divers exemples. Notons que cette méthode devrait en théorie permettre, avec suffisamment d'efforts de calcul, de résoudre des équations "générales" (mais pas trop), de la forme $y^3 = f(x)$, deg $f = 4$.

Description de la méthode

Dans ce paragraphe, on présente un bref descriptif de la méthode, dont les principes sont communs aux deux classes d'équations que l'on va étudier. Le principe fondamental consiste à utiliser des considérations élémentaires de théorie algébrique des nombres pour réduire l'équation diophantienne de départ à une équation diophantienne exponentielle, qui sera dans notre cas une "équation linéaire aux unités" (*linear unit equation*). On va expliquer selon quels principes on parvient, dans cette thèse, à effectuer cette réduction, puis comment, à partir de cette équation, on résout le problème de départ.

Objets fondamentaux

On va en fait "imiter" les démonstrations effectives de finitude du nombre de solutions ; pour cela, il faut nous ramener à un problème de formes linéaires en logarithmes, c'est-à-dire à prendre le logarithme d'une quantité du type $\theta_1^{b_1} \dots \theta_r^{b_r}$, qui soit en même temps liée à une solution (x, y). On sait minorer une telle quantité, mais il faudra aussi savoir la majorer par quelque chose de très petit, pour que la comparaison du majorant et du minorant fournisse un résultat non-trivial ; il faut donc que $\theta_1^{b_1} \dots \theta_r^{b_r} \approx 1$.

Ce genre de "combinaison linéaire" provient en général de la décomposition d'un élément bien choisi sur une base d'un groupe abélien de type fini ; deux telles familles de groupe viennent à l'esprit : le groupe des unités d'un corps de nombres (théorème de Dirichlet), et le groupe des points rationnels d'une courbe elliptique (théorème de Mordell-Weil).

Nous n'utiliserons dans cette thèse que la première famille de groupes, à l'exception de la fin du chapitre 5 ; la deuxième famille peut s'utiliser dans la recherche de points entiers sur des courbes elliptiques, voir par exemple [ST94], ou la fin du chapitre 5.

Qui dit groupe des unités dit qu'il faut d'abord avoir construit un ou plusieurs corps de nombres. Ce sera la première étape de la méthode : construire une famille "explicite" de corps de nombres $\{\mathbb{K}_i\}$, qui soit en même temps liée de manière

naturelle aux solutions : en d'autres termes, $\{\mathbb{Q}(f(x,y)), (x,y) \text{ solution } \} \subset \{\mathbb{K}_i\}$, où f est une certaine fonction algébrique.

On construit alors pour chaque i une fonction algébrique $\varphi(x,y)$ vivant dans $\mathbb{K}_i$ (on omettra dans la suite la dépendance en y, et parfois même celle en x pour alléger la notation) ; plutôt que d'imposer que cette fonction soit une unité, il est en général plus simple de lui demander d'avoir une norme explicitement calculable ; en effet, l'ensemble des éléments de norme donnée est fini, à l'action multiplicative des unités près. Donc si $\mathbb{K}$ est l'un des $\mathbb{K}_i$, et si $a \in \mathbb{Q}$ est fixé, il existe un ensemble Ξ fini tel que tout $z \in \mathbb{K}$ avec $N_{\mathbb{K}/\mathbb{Q}}(z) = a$ puisse s'écrire $z = \xi\eta$, avec $\xi \in \Xi$ et η unité.

Il nous faut maintenant assurer la condition $\theta_1^{b_1} \dots \theta_r^{b_r} \approx 1$. Pour ceci, on va former des quotients de puissances de conjugués de notre unité η. Il nous faut donc connaître de bonnes approximations des différents conjugués de $\varphi(x,y)$. Par "bonnes approximations", on entend les deux propriétés suivantes ; leur qualité doit s'améliorer suffisamment vite quand x augmente (pour que le majorant de la forme linéaire en logarithmes décroisse plus vite que le minorant), et la forme de l'approximation doit permettre d'éliminer x facilement. En pratique, on recherche des approximations du type

$$|\varphi_i(x,y) - \gamma_i x^{\rho_i}| \ll \frac{1}{x^{l_i}},$$

où les $\varphi_i(x,y)$ sont les différents conjugués de $\varphi(x,y)$ et bien sûr $\rho_i \geq -l_i$. On forme alors $(\varphi_{i_1}(x,y)/\gamma_{i_1})^{\rho_{i_2}} (\varphi_{i_2}(x,y)/\gamma_{i_2})^{-\rho_{i_1}}$, qui par construction est très voisin de 1, et a la forme $\theta_0\theta_1^{b_1} \dots \theta_r^{b_r}$.

En pratique, cet élément très voisin de 1 peut aussi être vu comme provenant d'une équation linéaire aux unités : ainsi, dans le cas de l'équation de Thue, on a $\varphi_2(x,y) - \varphi_3(x,y) = \theta$, où θ est un nombre algébrique fixé. Il suffit alors de diviser par φ_3, qui tend vers l'infini avec x, pour retrouver la situation précédente. On dispose d'une interprétation analogue pour les équations superelliptiques, que l'on réduit à une équation aux unités à 4 termes.

La borne de Baker

La théorie de Baker nous fournit un minorant pour

$$\log((\varphi_{i_1}(x,y)/\gamma_{i_1})^{\rho_{i_2}} (\varphi_{i_2}(x,y)/\gamma_{i_2})^{-\rho_{i_1}}),$$

de la forme $\exp(-C\log(\max_i |b_i|))$. Dans le même temps, on dispose d'un majorant donné par l'approximation de φ : si un nombre complexe est voisin de 1, son logarithme est petit. On va donc obtenir une relation du type

$$\exp(-C\log(\max_i |b_i|)) \ll \frac{1}{|x|^l}.$$

Il faut alors traduire le majorant en terme de b_i. À ces fins, on réutilise l'approximation déjà mentionnée : $\gamma_i x^{\rho_i} \approx \theta_0\theta_1^{b_1} \dots \theta_r^{b_r}$; prenant le logarithme, on va

obtenir un "système" donnant les b_i en fonction de $\log|x|$. Donc les b_i et $\log|x|$ sont du même ordre de grandeur, et l'on doit avoir :

$$\exp(-C\log(\max_i|b_i|)) \ll \exp(-l\max_i|b_i|).$$

Comme prévu, la comparaison du majorant et du minorant fournit une borne pour $\max_i|b_i|$.

Notons que cette idée d'inverser le système est beaucoup plus fondamentale qu'il n'y paraît, et une exploitation pertinente de cette idée est à la base de l'article [BH96]. On y reviendra dans le chapitre 1.

Réduction de la borne et énumération finale

La borne que l'on obtient par ces techniques est inutilisable en l'état. On a donc recours à une phase "diophantienne effective" pour obtenir de nouveaux minorants plus efficaces, que ce soit grâce à l'algorithme LLL ou aux fractions continues.

Les minorants obtenus sont à chaque étape de l'ordre de grandeur d'une puissance de l'inverse de la borne pour $\max_i|b_i|$. Comme on compare ces quantités à $\exp(-l\max_i|b_i|)$, on obtient une nouvelle borne qui dépend logarithmiquement de la précédente ; et rien n'empêche de réitérer le processus... Une suite de valeurs typique est 10^{30}, 50, 5.

On peut alors penser, dès lors que la borne est réduite à quelques unités, que l'énumération finale peut se faire de manière "brutale". Mais si r devient grand (et nous verrons un exemple où $r = 40$), il est impossible d'utiliser cette méthode ; nous verrons comment contourner ce problème dans le chapitre 1.

Il ne reste plus qu'à revenir des b_i à x. Ceci se fait tout simplement en utilisant encore une fois l'approximation $\theta_0\eta_1^{b_1}\dots\eta_r^{b_r} \approx \gamma_i x^{\rho_i}$, qui fournit une bonne approximation de x. Dès que x est assez grand (disons > 10), l'erreur commise est suffisamment petite pour que l'on puisse retrouver x de manière exacte par un simple arrondi. Reste à énumérer quelques valeurs de x.

Plan de la thèse

Les chapitres 1 et 4 de cette thèse mettent en place la méthode ci-dessus dans le cas des deux familles d'équations (Thue et superelliptiques) mentionnées plus haut. Le chapitre 2 introduit une modification dans la méthode du chapitre 1 pour exploiter le fait qu'un corps intervenant dans la résolution d'une équation de Thue ait des sous-corps non triviaux. Cette modification est suggérée par l'étude du problème des diviseurs primitifs des suites de Lucas-Lehmer, que nous présentons au chapitre 3, parmi d'autres exemples. On démontre le

Théorème. *Soit n un entier. On suppose que n vérifie l'une des conditions suivantes :*

– *n est une puissance de nombre premier comprise entre* 31 *et* 67,
– *n est un nombre premier compris entre* 67 *et* 997, *et* $n \equiv 1 \bmod 3$, 5, *ou* 8,
– $n \in \{83, 4001, 5011\}$.

Alors le $n^{\text{ème}}$ terme de toute suite de Lucas ou de Lehmer admet un diviseur primitif.

Enfin, au chapitre 5, nous appliquons les méthodes développées auparavant à la détermination des zéros du cinquième polynôme de Krawtchouk binaire :

Théorème. *Définissons*

$$P_5^n(x) = \frac{n-2x}{24}\left((n-2x)^4 + (n-2x)^2(20-10n) + 15n^2 - 50n + 24\right).$$

Alors les solutions entières (n, x), $0 \leq x < n/2$ de $P_5^n(x) = 0$ sont

$$(1,0), (2,0), (3,0), (3,1), (4,0), (4,1), (10,1), (10,3), (17,3), (36,14), (67,22), (67,28), (289,133), (10882,5292), (48324,24013).$$

Chapitre 1

L'équation de Thue

L'équation de Thue a le mérite de mettre en jeu un attirail technique relativement simple, tout en présentant les mêmes difficultés algorithmiques que les équations superelliptiques.

Ce chapitre va donc nous permettre de revenir de manière plus concrète et illustrée sur la méthodologie de résolution présentée dans l'introduction.

L'algorithme "d'origine" est dû à Tzanakis et de Weger [TW89], mais la version exposée ici est agrémentée de divers raffinements décrits dans [BH96, BW96]. On incorpore aussi une modification qui devrait faire l'objet d'une note [Ha97] : on montre comment, en pratique, se contenter d'un système d'unités de rang maximal plutôt que d'un système fondamental. L'idée d'utiliser un sous-groupe du groupe des unités n'est pas neuve, mais n'a pas, à ma connaissance, été formulée de cette manière, ni mise en œuvre dans la pratique. Nous illustrons l'utilité de cette idée par divers exemples au chapitre 3.

Bon nombre de lemmes, propositions et théorèmes présentés dans ce chapitre font partie du "folklore". J'ai autant que possible essayé d'en fournir des preuves, qui peuvent différer plus ou moins des preuves données à l'origine.

1.1 Préambule et prérequis algorithmiques

On appelle équation de Thue l'équation

$$P(X,Y) = f_0 Y^n + f_1 Y^{n-1} X + \ldots + f_n X^n = a, \tag{1.1}$$

où P est une forme irréductible de degré au moins 3, et a un nombre rationnel fixé.

1.1.1 Problèmes d'algorithmique

Avant d'attaquer la réduction proprement dite à un problème de formes linéaires en logarithmes, il nous faut définir clairement et discuter les problèmes

algorithmiques que nous aurons à résoudre. Dans toute la suite nous noterons $\mathbb{K}$ le corps $\mathbb{Q}(\alpha)$, où α est une racine du polynôme $P(1, Y)$. Il nous faudra

(U*) savoir trouver un système d'unités de rang maximal du corps $\mathbb{K}$,

(N) trouver un ensemble maximal M_a de solutions z non associées de l'équation

$$N_{\mathbb{K}/\mathbb{Q}}(z) = a \tag{1.2}$$

dans l'idéal fractionnaire $I = (1, \alpha)$.

Ce deuxième point mérite quelques éclaircissements. Le groupe $\mathcal{U}_{\mathbb{K}}$ des unités de $\mathbb{K}$ agit naturellement sur M_a par multiplication. Un ensemble maximal de solutions non associées de (1.2) est un ensemble composé d'un élément de chaque classe d'équivalence modulo cette action. D'un point de vue théorique, un tel ensemble est fini et peut être construit explicitement, voir par exemple [BS].

Le premier problème, baptisé (U*) est plus simple à résoudre que le problème (U) de [BH96] qui réclamait la connaissance d'un système d'unités fondamentales. La section suivante discute la détermination de tels systèmes d'unités (qu'ils soient fondamentaux ou maximaux) brièvement. Les références incontournables pour ces problèmes sont [PZ] et [Co].

1.1.1.1 Les unités

Il existe à l'heure actuelle deux grandes méthodes pour déterminer le groupe des unités d'un corps de nombres, la première étant due à Pohst et Zassenhaus [PZ], [PZ82], [PWZ82], et la deuxième trouvant son origine dans un travail de Hafner et McCurley [HM89], étendu et généralisé par Buchmann [Bu88], puis par Cohen, Diaz y Diaz et Olivier [CDO97].

La première méthode fonctionne bien dans les "petits" (degré et discriminant) corps de nombres. Elle consiste à rechercher beaucoup d'éléments de petite norme, et à faire les quotients d'éléments de même norme dans l'espoir d'obtenir des entiers algébriques qui, par construction seraient de norme 1 donc des unités. On poursuit ce travail jusqu'à connaître un système de rang maximal, le rang étant donné par le théorème de Dirichlet.

On recherche alors une majoration de l'indice en utilisant une minoration du régulateur du corps de nombres, et l'on agrandit le système d'unités jusqu'à ce qu'il soit maximal, en "cherchant" des racines $p^{\text{èmes}}$ à adjoindre au groupe construit pour p allant jusqu'à la borne supérieure pour l'indice.

La seconde méthode est conceptuellement très différente, et fournit en même temps le groupe des classes et sa structure, et le groupe des unités. La description qui suit doit beaucoup à un article de Cohen, Diaz y Diaz et Olivier [CDO97].

Dans cette méthode, on cherche en fait à construire une présentation du groupe des classes par générateurs et relations, c'est-à-dire une suite exacte

$$0 \to \Lambda \to \mathbb{Z}^n \to \mathrm{Cl}_{\mathbb{K}} \to 0.$$

À cette fin, on considère un certain nombre n d'idéaux dont on sait qu'ils engendrent le groupe des classes, et l'on recherche des relations entre ces idéaux, c'est-à-dire des produits de puissances de ces idéaux qui soient des idéaux principaux.

Le groupe des classes peut être engendré par les idéaux de norme assez petite, plus petite que la borne de Minkowski, qui dépend essentiellement de la racine carrée du discriminant. Il est déraisonnable d'espérer énumérer tous les idéaux de $\mathbb{K}$ de norme inférieure à cette borne. Toutefois, d'après un résultat de Bach [Ba90], le groupe des classes peut être engendré par les idéaux de norme plus petite que $12\log^2|D|$ (12 peut être remplacé par 6 dans le cas des corps quadratiques), sous l'hypothèse de Riemann généralisée.

Plusieurs méthodes sont employées pour chercher ces relations ; on peut par exemple factoriser des idéaux $\theta\mathbb{Z}_{\mathbb{K}}$, où θ décrit un ensemble d'éléments de petite norme. En utilisant la notion de réduction d'un idéal dans une direction, on peut obtenir des relations aléatoires "à volonté", ce qui permettra d'agrandir le réseau Λ dans les étapes suivantes.

Notons Λ le réseau des relations obtenu. On s'arrête quand le rang de Λ atteint n. On a alors obtenu un groupe fini $\mathbb{Z}^n/\Lambda$ dont l'ordre (qui s'obtient en calculant le déterminant de la matrice dont les colonnes constituent une base du réseau Λ) est un multiple $h'(\mathbb{K})$ du nombre de classes $h(\mathbb{K})$.

Maintenant, si à chaque relation on associe le générateur de l'idéal principal correspondant, ou plutôt son plongement logarithmique, on peut déduire de manière analogue le groupe des unités. Lorsque l'on met la matrice des relations sous forme normale d'Hermite, on va obtenir un certain nombre de relations triviales, c'est-à-dire représentant l'idéal $\mathbb{Z}_{\mathbb{K}}$. Mais si les manipulations effectuées sur la matrice des relations ont aussi été appliquées à la matrice des plongements logarithmiques, on dispose alors d'un générateur de cet idéal, c'est-à-dire d'une unité. De cette manière, avec assez de relations, on obtient un système d'unités de rang maximal.

Si l'on veut obtenir un système fondamental, on peut maintenant calculer le régulateur du système obtenu, qui est un multiple $R'_{\mathbb{K}}$ du régulateur $R_{\mathbb{K}}$ du corps. On calcule alors le produit

$$h(\mathbb{K})R_{\mathbb{K}} = \frac{w(\mathbb{K})\sqrt{|D(\mathbb{K})|}}{2^s(2\pi)^t}\prod_p \frac{1-\dfrac{1}{p}}{\displaystyle\prod_{\mathfrak{p}|p}\left(1-\frac{1}{N\mathfrak{p}}\right)},$$

où $w(\mathbb{K})$ et $D(\mathbb{K})$ sont respectivement le nombre de racines de l'unité et le discriminant de $\mathbb{K}$.

Sous l'hypothèse de Riemann généralisée (que l'on abrégera dans la suite en "GRH"), on peut tronquer le produit eulérien en se limitant aux $\mathfrak{p}$ avec $N\mathfrak{p} < C \log^2 |D(\mathbb{K})|$ (où C est encore 6 ou 12 suivant que le corps est quadratique ou non), et obtenir un nombre z avec

$$\frac{h(\mathbb{K})R_{\mathbb{K}}}{\sqrt{2}} < z < \sqrt{2}h(\mathbb{K})R_{\mathbb{K}}.$$

Dès lors, si $h'(\mathbb{K})R'_{\mathbb{K}} \leq z\sqrt{2}$, on a obtenu un système fondamental, sous l'hypothèse de Riemann pour $\mathbb{K}$; sinon, on recalcule de nouvelles relations.

Cette méthode permet donc en pratique de trouver rapidement un système d'unités dont on est sûr qu'il est de rang maximal (il suffit de calculer un déterminant pour s'en convaincre) mais qui n'est fondamental que sous l'hypothèse de Riemann généralisée. Des méthodes de certification existent, qui permettent de garantir la validité du résultat, mais elles sont très lentes dès que le régulateur et/ou le degré du corps augmente.

Il est donc bien souhaitable, quelle que soit la méthode de calcul des unités choisie, de pouvoir se contenter d'un système de rang maximal. Ceci épargne la seconde phase du calcul dans la méthode de Pohst-Zassenhauss, et permet d'avoir des résultats inconditionnels si l'on utilise la méthode de Buchmann.

La méthode homogène que nous proposons ci-dessous est dans ce dernier cas bien plus efficace que de certifier le système dès lors que la certification pose le moindre problème (gros régulateur, degré élevé...)

1.1.1.2 L'équation aux normes

On peut donner essentiellement deux méthodes pour résoudre l'équation aux normes $N_{\mathbb{K}/\mathbb{Q}}(z) = a$.

La première méthode consiste à décomposer le nombre a en produit d'idéaux premiers dans le corps $\mathbb{K}$, à chercher toutes les combinaisons de ces idéaux qui sont principales et de bonne norme. Cela nécessite de savoir résoudre le problème de l'idéal principal, c'est-à-dire, en pratique, d'utiliser l'algorithme sous GRH décrit à la section suivante. Il est cependant souvent possible d'éviter le recours à GRH. Ainsi, si le groupe calculé (dont le groupe des classes est un sous-groupe) est trivial, $h(\mathbb{K}) = 1$ indépendamment de GRH. De façon analogue, un idéal putativement principal l'est inconditionnellement, et donc si au-dessus de a on ne trouve que des idéaux putativement principaux, on obtient inconditionnellement un ensemble de solutions.

La deuxième méthode, due à Pohst et Zassenhauss [PZ], consiste à chercher des solutions de l'équation aux normes dans des sous-ensembles bien choisis, par exemple des parallélotopes ou ellipsoïdes (où l'on a identifié $\mathbb{Z}_{\mathbb{K}}$ à $\mathbb{Z}^n$ par le choix d'une base d'entiers) ; un choix pertinent permet d'obtenir la totalité des solutions, modulo l'action du groupe des unités.

Cette discussion algorithmique étant close, nous allons maintenant entrer dans le vif du sujet. La présentation adoptée suit *grosso modo* [BH96]. Sauf mention expresse du contraire, la notation $\|\cdot\|$ désignera la distance d'un réel à $\mathbb{Z}$.

1.2 Formes linéaires en logarithmes

1.2.1 Préliminaires

Nous allons maintenant mettre en œuvre notre méthodologie de résolution. L'équation de Thue (1.1) s'écrit

$$P(x,y) = f_0 y^n + f_1 y^{n-1} x + \cdots + f_n x^n = f_0 (y - \alpha^{(1)} x) \cdots (y - \alpha^{(n)} x) = a\,. \quad (1.3)$$

et on note $g(y) = P(1,y)$ et $\alpha = \alpha^{(1)}$. L'écriture ci-dessus incite à considérer le corps $\mathbb{K} = \mathbb{Q}(\alpha)$, qui est explicitement déterminé (du moins à isomorphisme près) par les données du problème.

Le premier point de la réduction à une équation diophantienne exponentielle consistait à construire un corps de nombres ; c'est maintenant chose faite.

On ordonne les $\alpha^{(i)}$ de sorte que $\alpha^{(1)}, \ldots, \alpha^{(s)}$ soient réels et $\alpha^{(s+i)} = \overline{\alpha^{(s+t+i)}}$. Comme il est d'usage, on notera $r = s + t - 1$ le rang du groupe $\mathcal{U}_{\mathbb{K}}$ des unités de $\mathbb{K}$.

Les plongements σ_i sont définis de manière unique par :

$$\begin{array}{rccc} \sigma_i : \mathbb{K} = \mathbb{Q}(\alpha) & \to & \mathbb{Q}(\alpha^{(i)}) \\ \alpha & \mapsto & \alpha^{(i)}. \end{array}$$

Étant donné $\beta \in \mathbb{K}$, on notera indifféremment $\beta^{(i)}$ ou $\sigma_i(\beta)$ le $i^{\text{ème}}$ conjugué de β.

On peut remarquer que (1.3) peut se lire $f_0 \sigma_1(y - \alpha x) \ldots \sigma_n(y - \alpha x) = a$, de sorte que

$$N_{\mathbb{K}/\mathbb{Q}}(y - \alpha x) = \frac{a}{f_0}. \quad (1.4)$$

Posant donc $\varphi(x,y) = y - \alpha x$, on voit que l'on a construit une fonction de x et y dont on contrôle la norme dès que le couple (x,y) est solution. La deuxième contrainte de la réduction à une équation exponentielle est donc remplie.

Il nous faudra au passage savoir résoudre le problème (N), lié à l'équation (1.2) pour le second membre a/f_0 et l'idéal $(1,\alpha)$ (qui n'est plus un idéal entier dès que $f_0 \not| \operatorname{pgcd}(f_1, \ldots, f_n)$).

La factorisation (1.3) permet déjà d'écarter le cas $s = 0$. En effet, en minorant chacun des facteurs du membre gauche par sa partie imaginaire, il vient :

$$|f(x,y)| \geq c_0 |x|^n\,, \quad (1.5)$$

avec $c_0 = a\left(f_0\left|\text{Im }\alpha^{(1)}\right|\cdots\left|\text{Im }\alpha^{(n)}\right|\right)^{-1}$, et les solutions de (1.1) sont facilement trouvées par énumération directe. Dans la suite, on supposera donc que $s \geq 1$; ceci entraîne en particulier que les seules racines de l'unité de $\mathbb{K}$ sont $\{-1, 1\}$.

Soit maintenant Log la détermination principale du logarithme, c'est-à-dire que l'on a $-\pi \leq \text{Im Log}\, x < \pi$. Rappelons quelques propriétés élémentaires de ce logarithme ; la troisième, notamment, servira à maintes reprises dans cette thèse.

Lemme 1.1. *Soit $x, y \in \mathbb{C}, n \in \mathbb{N}$. Alors*

$$|\text{Log}\,(xy)| \leq |\text{Log}\,(x)| + |\text{Log}\,(y)|,$$

$$|\text{Log}\,(x^n)| \leq n|\text{Log}\,(x)|,$$

$$|x-1| < 1/2 \Rightarrow |\text{Log}\, x| \leq 1,39|x-1|. \tag{1.6}$$

Preuve. Remarquons que si $\text{Arg}\, x + \text{Arg}\, y \in [(2k-1)\pi, (2k+1)\pi[$, $\text{Arg}\,(xy) = \text{Arg}\, x + \text{Arg}\, y - 2k\pi$. En particulier, $|\text{Arg}\,(xy)| \leq |\text{Arg}\, x + \text{Arg}\, y| \leq |\text{Arg}\, x| + |\text{Arg}\, y|$.

Mais alors, comme

$$\begin{aligned} |\text{Log}\,(xy)| &= \sqrt{(\log|x| + \log|y|)^2 + \text{Arg}\,(xy)^2} \\ &\leq \sqrt{(\log|x| + \log|y|)^2 + (\text{Arg}\,(x) + \text{Arg}\,(y))^2}, \end{aligned}$$

le résultat découle de l'inégalité triangulaire pour $\|\cdot\|_2$.

Le second point se déduit facilement du premier par récurrence.

Pour le troisième, on écrit :

$$|\text{Log}\, x| = \left|\sum_{n\geq 1}(-1)^{n-1}\frac{(x-1)^n}{n}\right| \leq 2|x-1|\sum_{n\geq 1}\frac{(1/2)^n}{n} \leq 2\log 2|x-1|,$$

et $2\log 2 < 1,39$. □

1.2.2 L'approximation de $y - \alpha^{(i)}x$

En regardant bien (1.1), on peut remarquer la chose suivante. Si x est assez grand, on s'attend à ce que les facteurs du membre de gauche soient assez grands, c'est ce qui s'est produit pour $s = 0$. Mais inversement, on impose à leur produit d'être borné en taille, ce qui ne peut avoir lieu que si l'un au moins des facteurs est petit. Dans ce cas, il y a un i_0 tel que $y \approx \alpha^{(i_0)}x$, et les autres facteurs sont de l'ordre de grandeur de x.

Cette remarque conduit à énoncer la proposition suivante :

Proposition 1.2. *Soit*

$$X_0 = \begin{cases} \left(\dfrac{2^{n-1}|a|}{\min_{1\le i\le t}|g'(\alpha^{(s+i)})|\cdot\min_{1\le i\le t}|\mathrm{Im}\ \alpha^{(s+i)}|}\right)^{1/n} & si\ \ t\ge 1\,,\\ 1 & si\ \ t=0\,,\end{cases}$$

$$\begin{aligned} c_1 &= \frac{2^{n-1}|a|}{\min_{1\le i\le s}|g'(\alpha^{(i)})|}\,, \qquad c_2 = \min_{1\le i<j\le n}\left|\alpha^{(i)}-\alpha^{(j)}\right|\,,\\ c_3 &= 1,39c_1c_2^{-1}\,, \qquad c_4=(n-1)c_3\,,\\ X_1 &= \max\left(X_0,\,\left(2c_1c_2^{-1}\right)^{1/n}\right)\,. \end{aligned}$$

Soit (x,y) une solution entière de (1.1).

(i) Si $|x|>X_0$ alors, pour un $i_0\in\{1,\dots,s\}$ on a

$$\left|\frac{y}{x}-\alpha^{(i_0)}\right|\le\frac{c_1}{|x|^n}\,. \tag{1.7}$$

Posons alors

$$\psi_i=\begin{cases}(\alpha^{(i_0)}-\alpha^{(i)}), & i\ne i_0,\\ \dfrac{a}{g'(\alpha^{(i_0)})}, & i=i_0.\end{cases} \tag{1.8}$$

(ii) Si $|x|>X_1$, alors

$$\left|\mathrm{Log}\,\frac{y-\alpha^{(i)}x}{\psi_i x}\right|\le\frac{c_3}{|x|^n},\quad (i\ne i_0). \tag{1.9}$$

(iii) Si $|x|>X_1$, alors

$$\left|\mathrm{Log}\,\frac{y-\alpha^{(i_0)}x}{\psi_{i_0}x^{1-n}}\right|\le\frac{c_4}{|x|^n}. \tag{1.10}$$

Preuve. On définit i_0 par $|y-\alpha^{(i_0)}x|=\min\limits_i|y-\alpha^{(i)}x|$. On a alors

$$|f_0|\prod_i|y-\alpha^{(i)}x|=|a|\ ; \tag{1.11}$$

par ailleurs $|y-\alpha^{(i)}x|\ge|x||\alpha^{(i_0)}-\alpha^{(i)}|-|y-\alpha^{(i_0)}x|$, soit encore, par définition de i_0,

$$|y-\alpha^{(i)}x|\ge\frac{|x|}{2}|\alpha^{(i_0)}-\alpha^{(i)}|.$$

On obtient *(i)* en reportant cette minoration dans (1.11). Supposons alors que pour un certain x, on a $i_0>s$. Il vient,

$$\frac{c_1}{|x|^{n-1}}\ge|y-\alpha^{(i_0)}x|\ge|x||\mathrm{Im}\ \alpha^{(i_0)}|,$$

c'est-à-dire que $|x| \leq X_0$.

Pour prouver *(ii)*, on écrit tout simplement que

$$\frac{y/x - \alpha^{(i)}}{\alpha^{(i_0)} - \alpha^{(i)}} - 1 = \frac{y/x - \alpha^{(i_0)}}{\alpha^{(i_0)} - \alpha^{(i)}},$$

et on utilise *(i)* et la définition de c_2. On obtient alors

$$\left| \frac{y - \alpha^{(i)}x}{x(\alpha^{(i_0)} - \alpha^{(i)})} - 1 \right| \leq \frac{c_1}{c_2|x|^n}$$

Il suffit alors d'appliquer (1.6), puisque le choix de X_1 garantit que le majorant est plus petit que $1/2$.

Pour *(iii)*, on écrit

$$\left| \mathrm{Log} \frac{\prod_{i \neq i_0}(y - \alpha^{(i)}x)}{x^{n-1}\prod_{i \neq i_0}\psi_i} \right| \leq \frac{c_4}{|x|^n},$$

en utilisant le Lemme 1.6 et *(ii)* pour tous les $i \neq i_0$. Il suffit alors de remarquer que

$$\prod_{i \neq i_0}(y - \alpha^{(i)}x) = \frac{P(x,y)}{f_0(y - \alpha^{(i_0)}x)}$$

et que

$$\prod_{i \neq i_0}\psi_i = \frac{g'(\alpha^{(i_0)})}{f_0}$$

pour conclure, puisque (x, y) vérifie (1.1). □

Remarque 1.3. On vient de construire une approximation des différents conjugués de notre $\varphi(x, y)$, de la forme $\gamma_i x^{\rho_i}$, ne dépendant que des données de l'équation. Toutes les exigences pour la réduction à un problème de formes linéaires en logarithmes sont donc remplies.

Remarque 1.4. Remarquons également que ce qui précède, joint au résultat de Thue, suffit pour établir la finitude du nombre de solutions. Thue montre en effet le

Théorème 1.5 (Thue, [Th09]). *Soit α un irrationnel algébrique ; alors pour tout $\varepsilon > 0$, l'inéquation*

$$\left| \alpha - \frac{p}{q} \right| \leq \frac{1}{q^{n/2+1+\varepsilon}}$$

n'a qu'un nombre fini de solutions.

La finitude découle donc de (1.7), car $n \geq 3$.

En pratique, la valeur de X_1 obtenue est très petite, et les solutions pour $|x| \leq X_1$ sont trouvées par énumération directe. Notons que le point *(i)* entraîne que pour x assez grand, $|y/x - \alpha^{(i_0)}| \leq 1/(2x^2)$, et que y/x est donc une réduite du développement en fraction continue de α. Ce point sera exploité ultérieurement.

Quitte à renuméroter les racines réelles, on supposera dans la suite que $i_0 = 1$, et donc que $\alpha^{(i_0)} = \alpha$. Bien entendu, en appliquant l'algorithme, il faut faire ce qui suit pour tout i_0 de $\{1, \ldots, s\}$ (voir section 1.6).

1.2.3 Unités

On utilise maintenant la remarque (1.4). Si M_{a/f_0} est l'ensemble construit en appliquant (N) à a/f_0 et à l'idéal $(1, \alpha)$, et $\eta_1, \ldots, \eta_r$ le système d'unités de rang maximal obtenu par (U*), il existe un élément μ de M_{a/f_0} tel que $(y - \alpha x)/\mu$ soit une unité ; il existe donc également un $(r+1)$-uplet $(b_0, \ldots, b_r)$ d'entiers tel que

$$(y - \alpha x)^{b_0} = \mu^{b_0} \eta_1^{b_1} \ldots \eta_r^{b_r}. \tag{1.12}$$

Il s'agit en quelque sorte d'un changement de variables ; nous allons maintenant oublier x et y jusqu'à la fin de l'algorithme, et travailler uniquement avec le $(r+1)$-uplet $(b_0, \ldots, b_r)$.

Remarque 1.6. On peut toujours imposer $b_0 \geq 1$. C'est ce que l'on supposera.

Remarque 1.7. Notons que la variable b_0 a été introduite pour relâcher l'hypothèse (U) en (U*), c'est-à-dire pour qu'il ne soit plus nécessaire d'exiger que le système d'unités soit fondamental. Dans certains cas, on peut se dispenser de la variable b_0 : quand le système d'unités est fondamental, ou plus généralement quand on connaît l'indice du système d'unités.

Le nombre de variables a augmenté de 2 à $r+1$, mais c'est le prix à payer pour transformer notre problème de départ en un problème "linéaire".

De (1.7), on tire

$$\left| \mathrm{Log} \left(\frac{\alpha - \alpha^{(3)}}{\alpha - \alpha^{(2)}} \frac{y - \alpha^{(2)} x}{y - \alpha^{(3)} x} \right) \right| \leq \frac{2c_3}{|x|^n}, \tag{1.13}$$

mais $|\mathrm{Log}\,(z^n)| \leq n |\mathrm{Log}\,(z)|$ entraîne

$$\left| \mathrm{Log} \left[\left(\frac{\alpha - \alpha^{(3)}}{\alpha - \alpha^{(2)}} \frac{y - \alpha^{(2)} x}{y - \alpha^{(3)} x} \right)^{b_0} \right] \right| \leq b_0 \frac{2c_3}{|x|^n}. \tag{1.14}$$

Il reste à exploiter la remarque (1.12) pour obtenir un problème de formes linéaires en logarithmes de nombres algébriques ; ainsi, il existe un nombre entier b_{r+1} tel que

$$\left| b_0 \mathrm{Log} \frac{\alpha - \alpha^{(3)}}{\alpha - \alpha^{(2)}} + b_1 \mathrm{Log} \frac{\eta_1^{(3)}}{\eta_1^{(2)}} + \ldots + b_r \mathrm{Log} \frac{\eta_r^{(3)}}{\eta_r^{(2)}} + b_{r+1} \cdot i\pi \right| \leq b_0 \frac{2c_3}{|x|^n}. \tag{1.15}$$

Remarque 1.8. On pourrait croire plus simple de considérer le logarithme du module, ce qui éviterait en particulier d'introduire b_{r+1}. Malheureusement, pour pouvoir appliquer la borne de Baker, il faut garantir la non-nullité de la forme linéaire en logarithmes correspondante. Ce n'est pas possible en général pour la partie réelle de la forme (1.15) (excepté dans le cas totalement réel, voir plus bas). Pour ce qui est de la réduction, en revanche, on utilisera presque toujours la forme donnée par le logarithme du module, qui comporte une variable de moins.

1.3 La borne de Baker

1.3.1 Bornes inférieures pour les formes linéaires en logarithmes

Tout le contenu de cette section (et, indirectement, l'intégralité de la méthode) repose sur une borne inférieure pour les formes linéaires en logarithmes. Nous énonçons le meilleur résultat actuellement disponible.

Théorème 1.9 (Baker-Wüstholz). [BW93, p. 20] *Soit $\beta_0, \ldots, \beta_r$ des nombres complexes algébriques distincts de 0 et de 1, et $b_0, b_1, \ldots, b_{r+1}$ des entiers. On pose $B'' = \max(e, \max_i |b_i|)$.*
Soit également

$$d \geq [\mathbb{Q}(\beta_0, \ldots, \beta_r) : \mathbb{Q}], \tag{1.16}$$

$$h_i \geq \max\left(h(\beta_i), d^{-1}|\log \beta_i|, d^{-1}\right) \quad (0 \leq i \leq r), \tag{1.17}$$

où $h(\cdot)$ est la hauteur logarithmique absolue. Alors si

$$\Lambda := b_0 \log \beta_0 + b_1 \log \beta_1 + \cdots + b_r \log \beta_r + b_{r+1}\pi i, \tag{1.18}$$

n'est pas nul, on a

$$|\Lambda| \geq \exp(-c_9 \log B''), \tag{1.19}$$

où

$$c_9 = 18\pi \cdot 32^{r+4}(r+3)!(r+2)^{r+3}d^{r+3}\log(2d(r+2))h_0 \cdots h_r.$$

Remarque 1.10. Les paramètres n, $h'(\alpha_1), \ldots, h'(\alpha_n)$, $h'(L)$ du théorème original de [BW93] correspondent dans le Théorème 1.9 respectivement à $r+2$, $h_0, \ldots, h_r$, π/d, $\log B''$.

L'énoncé de [BW93] a été modifié, pour permettre des inégalités dans (1.16) et (1.17). Il est souvent bien plus facile (et surtout rapide) de trouver une borne supérieure pour le degré d'un corps de nombres ou pour la hauteur d'un nombre algébrique que de les calculer précisément. En particulier, si l'on souhaite implanter les algorithmes décrits dans cette thèse, il est recommandé d'utiliser les inégalités

$$h(a+b) \leq h(a) + h(b) + \log 2, \qquad h(ab^{\pm 1}) \leq h(a) + h(b),$$

ceci ayant peu d'incidence sur la borne de Baker initiale, qui n'intervient elle-même que par son ordre de grandeur.

Remarque 1.11. Signalons que d'éventuelles améliorations de la borne de Baker et Wüstholz n'auraient pas de grandes retombées sur la méthode, à moins qu'elles ne soient drastiques. Elles permettraient simplement de diminuer la précision des calculs, ce qui accélérerait d'autant la résolution, mais le nombre d'opérations à effectuer serait sensiblement le même.

Le Théorème 1.9 nous permet de disposer d'un minorant, sous réserve que l'on puisse garantir que la forme linéaire est non nulle. Posons

$$\xi = \frac{\alpha - \alpha^{(3)}}{\alpha - \alpha^{(2)}} \frac{y - \alpha^{(2)}x}{y - \alpha^{(3)}x}.$$

La forme linéaire correspondante est $\mathrm{Log}\,(\xi^{b_0})$. Elle ne peut donc être nulle que si ξ est une racine de l'unité d'ordre divisant b_0, donc en particulier inférieur à b_0. On distingue alors trois cas :

- $\xi = 1$, auquel cas on voit facilement que $y = \alpha x$, ce qui est impossible, car P est irréductible ;
- $\xi^{b_0} \neq 1$, ce qui est le cas en particulier si ξ n'est pas une racine de l'unité d'ordre inférieur à b_0. On minore alors $|\mathrm{Log}\,(\xi_0)|$ au moyen du Théorème 1.9.
- ξ est une racine de l'unité d'ordre inférieur à b_0, et $\xi \neq 1$. Mais alors,

$$|\mathrm{Log}\,\xi| \geq \frac{2\pi}{b_0}, \tag{1.20}$$

minoration que l'on utilise en lieu et place de la borne de Baker.

Remarque 1.12. Ce même argument permet de garantir la non-nullité de la forme linéaire donnée par $\log|\xi|$ dans le cas où $\mathbb{K}$ est totalement réel, puisqu'il suffit de garantir que $\xi \neq -1$. Ceci permet d'améliorer un peu le minorant pour la forme linéaire en logarithmes, puisqu'elle comporte une variable de moins ; dans la section suivante, cela permet d'améliorer le majorant pour ladite forme, car on contrôle beaucoup moins bien b_{r+1} que les autres b_i. L'un dans l'autre, on obtiendra donc une meilleure borne pour $\max_i |b_i|$, mais cette amélioration ne porte que sur la borne de Baker initiale, et n'a donc que peu d'intérêt dans les faits.

La section suivante va donner un majorant pour la forme linéaire, en traduisant la majoration (1.15), qui est en terme de x, en terme de $B := \max_{1\leq i\leq r} |b_i|$. La comparaison du majorant et du minorant donné par le Théorème 1.9 ou par (1.20) nous fournira alors une borne pour B.

1.3.2 B et $\log|x|$

Deux cas doivent être traités à part, celui de b_0 et celui de b_{r+1}. On va donc définir

$$B = \max_{1\leq i\leq r} |b_i|, \qquad B' = \max_{0\leq i\leq r} |b_i|, \qquad B'' = \max_{0\leq i\leq r+1} |b_i|.$$

Commençons par évoquer le cas de b_0.

1.3.2.1 b_0

Dans le cas où le système d'unités $(\eta_1, \ldots, \eta_r)$ est fondamental, la situation est simple puisque l'on peut prendre $b_0 = 1$. Dans le cas contraire, on peut toujours imposer

$$b_0 \leq [\mathfrak{U}_{\mathbb{K}} :< \eta_1, \ldots, \eta_r >]$$

(on peut même imposer à b_0 de diviser l'indice ; c'est inutile ici, car on ne sait que majorer cet indice).

Estimer b_0, c'est donc donner un majorant de l'indice $[\mathfrak{U}_{\mathbb{K}} :< \eta_1, \ldots, \eta_r >]$.

Rappelons la définition du régulateur d'un système d'unités de rang maximal :

Définition 1.13. *Soit $\eta_1, \ldots, \eta_r$ un système d'unités de rang maximal r. On note*

$$\varepsilon_i = \begin{cases} 1, & i \leq s \\ 2, & i > s. \end{cases}$$

Alors la matrice $\left(\varepsilon_i \log|\eta_j^{(i)}|\right)_{1\leq i,j\leq r}$ est inversible, et la valeur absolue de son déterminant est appelée régulateur du système d'unités $\eta_1, \ldots, \eta_r$; on le notera $R(\eta_1, \ldots, \eta_r)$.

Les régulateurs de tous les systèmes fondamentaux sont égaux. Cette valeur commune est appelée régulateur de $\mathbb{K}$. On le notera $R_{\mathbb{K}}$.

On obtient alors la majoration de l'indice par la proposition suivante :

Proposition 1.14. *Soit $\eta_1, \ldots, \eta_r$ un système d'unités de rang maximal r. On a alors :*

$$[\mathfrak{U}_{\mathbb{K}} :< \eta_1, \ldots, \eta_r >] = \frac{R(\eta_1, \ldots, \eta_r)}{R_{\mathbb{K}}}. \tag{1.21}$$

Il nous reste donc à donner une borne inférieure pour $R_{\mathbb{K}}$. De telles bornes inférieures peuvent être obtenues via des procédés analytiques, mais les résultats obtenus ont le défaut d'utiliser extrêmement peu d'invariants du corps (la signature), et donc d'être valides pour une large classe de corps, donc forcément mauvais dans la plupart des exemples concrets. Plus précis, quoique plus délicat à mettre en œuvre, est le résultat suivant :

Théorème 1.15. *On note γ_r la constante d'Hermite en dimension r ; soit*

$$M^* = \min\left\{\sum_{j=1}^{n}(\log|\varepsilon_j|)^2 \ ;\ \varepsilon \in \mathcal{U}_{\mathbb{K}} - \{-1,1\}\right\},\ et\quad C = M^* - n + 1.$$

Posons

$$M_0 = \frac{1}{4}\left(\log\left(\frac{C}{n} + \left(\frac{C^2}{n^2} - 1\right)^{1/2}\right)\right)^2.$$

Alors on a

$$R_{\mathbb{K}} \geq (M_0^r \gamma_r^r 2^t n^{-1})^{1/2}.$$

Preuve. [PoWi97] pour le cas $n = 5$ (où l'on sait faire un peu mieux), [Fi97] sinon. Les formules analytiques mentionnées ci-dessus peuvent être trouvées dans [Zi81], [Fr89] ou encore [CF91]. □

Il est toujours également possible d'utiliser la minoration $R_{\mathbb{K}} \geq 0,2$ pour tout corps de nombres, prouvée dans [Fr89]. Ceci est raisonnable tant que l'indice obtenu reste assez petit (disons quelques centaines).

Définition 1.16. *La borne obtenue pour b_0 par la méthode décrite ci-dessus sera notée $\mathcal{B}$.*

Elle joue un rôle différent des bornes pour les autres $|b_i|$; en particulier, si la taille de la borne de Baker est (relativement) peu importante, la borne $\mathcal{B}$ conditionne la réduction de la borne de Baker ; il convient donc a priori d'utiliser le meilleur résultat de minoration de régulateur possible. On pourrait croire que la réduction de la borne $\mathcal{B}$ se fera simultanément à celle de B. Il n'en est rien, et pour cause : la réduction de B utilise essentiellement le fait que $\max_{1\leq i\leq r}|b_i|$ tend vers l'infini avec x — ce résultat est l'objet du paragraphe suivant — mais ce n'est pas le cas de b_0.

1.3.2.2 $|b_i|$ et $\log|x|$, $1 \leq i \leq r$

C'est le moment d'introduire une remarque qui va nous être fort utile dans la phase "diophantienne effective", c'est-à-dire pour la réduction et l'énumération finale.

Remarquons que par (1.12), on a :

$$(y - \alpha x)^{b_0} = \mu^{b_0}\eta_1^{b_1}\dots\eta_r^{b_r}. \tag{1.22}$$

ce qui veut dire, en passant au logarithme du module et en l'écrivant pour les différents conjugués, que

$$b_1\log|\eta_1^{(i)}| + \dots + b_r\log|\eta_r^{(i)}| = b_0\log\left|\frac{y - \alpha^{(i)}x}{\mu^{(i)}}\right|. \tag{1.23}$$

On définit alors

$$\rho_i = \begin{cases} 1, & i \neq i_0, \\ 1-n, & i = i_0. \end{cases} \tag{1.24}$$

Avec cette notation, il vient :

$$b_1 \log|\eta_1^{(i)}| + \ldots + b_r \log|\eta_r^{(i)}| = b_0\rho_i \log|x| + b_0 \log\left|\frac{\psi_i}{\mu^{(i)}}\right| + b_0 \log\left|\frac{y-\alpha^{(i)}x}{\psi_i x^{\rho_i}}\right|. \tag{1.25}$$

Soit alors $A = [a_{ij}]_{1\leq i,j\leq r}$ l'inverse de la matrice $\left[\log|\eta_i^{(j)}|\right]_{1\leq i,j\leq r}$. Cette matrice est bien inversible, car son déterminant est à un facteur $\pm\min(1, 2^{1-t})$ près le régulateur du corps de nombres $\mathbb{K}$.

En multipliant le "système linéaire" (1.25) à gauche par A, on trouve

$$b_i = b_0\delta_i \log|x| + b_0\lambda_i + b_0\epsilon_i, \tag{1.26}$$

où

$$\delta_i = \sum_{j=1}^{r} a_{ij}\rho_j, \quad \lambda_i = \sum_{j=1}^{r} a_{ij} \log\left|\frac{\psi_j}{\mu^{(j)}}\right|, \epsilon_i = \sum_{j=1}^{r} a_{ij} \log\left|\frac{y-\alpha^{(j)}x}{\psi_j x^{\rho_j}}\right|.$$

Définissons

$$X_2 := \max\left(X_1, \left(10c_4 \max_{1\leq i\leq r} \sum_{j=1}^{r} |a_{ij}|\right)^{1/n}\right).$$

Comme d'après (1.9) et (1.10), on a $|\epsilon_i| \leq \dfrac{c_4}{|x|^n} \displaystyle\sum_{1\leq j\leq r} |a_{ij}|$, pour $|x| \geq X_2$ on a donc $|\epsilon_i| \leq 0,1$.

Proposition 1.17. *Si $|x| \geq X_2$, on a*

$$\max_{1\leq i\leq r} |b_i| \leq b_0 c_5 \log|x| + b_0 c_6, \tag{1.27}$$

avec $c_5 = \max_i |\delta_i|$ et $c_6 = \max_i |\lambda_i| + 0,1$.

Remarque 1.18. On peut bien sûr faire varier le 10 dans la définition de X_2 ; ainsi si n est très grand, on pourra choisir un bien plus grand nombre ici pour améliorer la constante c_6, car cela aura peu d'influence sur X_2.

Toutefois, la constante c_5 est la seule **cruciale** pour la réduction ultérieure. Toute amélioration de c_5 se traduit par une amélioration correspondante du processus de réduction ; dans la version donnée ici, la constante c_5 semble optimale. Cela explique en partie l'amélioration du processus de réduction constatée dans [BH96].

On va maintenant déduire de ce qui précède une estimation pour b_{r+1}.

1.3.2.3 $|b_{r+1}|$ et $\log|x|$

Rappelons que b_{r+1} est la constante introduite en (1.15) pour pallier le défaut d'additivité du logarithme complexe.

Par définition, on a

$$\left|b_0\mathrm{Log}\,\frac{\alpha-\alpha^{(3)}}{\alpha-\alpha^{(2)}}+b_1\mathrm{Log}\,\frac{\eta_1^{(3)}}{\eta_1^{(2)}}+\ldots+b_r\mathrm{Log}\,\frac{\eta_r^{(3)}}{\eta_r^{(2)}}+b_{r+1}\cdot i\pi\right|\leq b_0\frac{2c_3}{|x|^n}. \qquad (1.28)$$

Mais alors la partie imaginaire du premier membre admet la même majoration ; comme $|\mathrm{Im}\,\mathrm{Log}\,z|\leq\pi$ pour tout $z\in\mathbb{C}$,

$$|b_{r+1}|\leq b_0+|b_1|+\ldots+|b_r|+\frac{b_0c_3}{\pi|x|^n}\leq b_0+r(b_0c_5\log|x|+b_0c_6)+0,23b_0.$$

puisque par le choix de X_1, on a $2c_3\pi^{-1}|x|^{-n}\leq 1,39(2\pi)^{-1}\leq 0,23$.

On a donc

$$|b_{r+1}|\leq b_0c_7\log|x|+b_0c_8,$$

avec $c_7=rc_5$, et $c_8=1,23+rc_6$.

Remarque 1.19. Il est possible d'améliorer un peu la constante c_7 en calculant explicitement les parties imaginaires des $\mathrm{Log}\,(\eta_i^{(3)}/\eta_i^{(2)})$. Cela n'a pas grande importance, car cette constante n'influe que sur le calcul de la borne de Baker initiale.

Les résultats de cette section se résument donc en

$$B''\leq b_0\max(1,c_7\log|x|+c_8)\leq\mathcal{B}\max(1,c_7\log|x|+c_8). \qquad (1.29)$$

1.3.3 Une borne supérieure pour B''.

Dans le cas où $c_7\log|x|+c_8<1$, on a déjà la borne $B''\leq\mathcal{B}$.

Dans le cas contraire, en inversant (1.29), il vient :

$$-n\log|x|\leq-\frac{c_{10}B''}{b_0}+\frac{nc_8}{c_7}$$

avec $c_{10}=n/c_7$. En reportant ceci dans le majorant de (1.15), il vient

$$\left|b_0\mathrm{Log}\,\frac{\alpha-\alpha^{(3)}}{\alpha-\alpha^{(2)}}+b_1\mathrm{Log}\,\frac{\eta_1^{(3)}}{\eta_1^{(2)}}+\ldots+b_r\mathrm{Log}\,\frac{\eta_r^{(3)}}{\eta_r^{(2)}}+b_{r+1}i\pi\right|$$
$$\leq b_0c_{11}\exp(-c_{10}B''/b_0).$$

où $c_{11}=2c_3\exp(nc_8/c_7)$.

Dans le cas où la forme linéaire est nulle, on utilise la minoration (1.20)[1] ; il vient

$$B'' \leq \frac{\mathcal{B}}{c_{10}} \log\left(\frac{\mathcal{B}c_{11}}{2\pi}\right).$$

Dans le cas contraire, on utilise maintenant le Théorème 1.9.

Comparant minorant et majorant, on obtient

$$\exp(-c_9 \log \max(B'', e)) \leq \mathcal{B}c_{11} \exp(-c_{10}B''/\mathcal{B}). \tag{1.30}$$

De même que l'on a écarté temporairement le cas $B'' \leq \mathcal{B}$, on peut écarter le cas $B'' \leq e$. On a alors majoré une fonction de B'' par une autre fonction qui décroît beaucoup plus vite. En conséquence B'' est borné ; la borne effective découle du lemme suivant ([PW87]) :

Lemme 1.20. *Soit z et C_1, C_2 des nombres réels positifs, avec $C_1 \geq e$. Si $z \leq C_1 \log z + C_2$, alors $z \leq z_0 := 2(C_1 \log C_1 + C_2)$.*

Preuve. Soit $f(z) = z - C_1 \log(z) - C_2$. Alors $f'(z) = 1 - C_1/z$, et f est croissante pour $z > C_1$. Il suffit donc de montrer que $f(z_0) > 0$. De $C_1 \geq e$, on tire $2 \log C_1 \leq C_1$, d'où $2C_1 \log C_1 \leq C_1^2$, et $C_1 \log(2C_1 \log C_1) \leq 2C_1 \log C_1$. Pour a et b strictement positifs, on a $\log(a+b) = \log(a) + \log(1+b/a) \leq \log(a) + b/a$; d'où

$$\begin{aligned} C_1 \log(2(C_1 \log C_1 + C_2)) &< C_1 \log(2C_1 \log C_1) + \frac{C_1 C_2}{C_1 \log C_1} \\ &< 2C_1 \log C_1 + C_2, \end{aligned}$$

d'où le résultat. □

Du Lemme 1.20 et de (1.30), on déduit, en rajoutant les cas $B'' \leq e$ et $B'' \leq \mathcal{B}$, le

Théorème 1.21. *Si $\mathcal{B}c_{10}^{-1}c_9 > e$, on a*

$$\begin{aligned} B' \leq B'' &\leq \max\left(e, \mathcal{B}, \frac{\mathcal{B}}{c_{10}} \log\left(\frac{\mathcal{B}c_{11}}{2\pi}\right), 2\frac{\mathcal{B}}{c_{10}}\left(c_9 \log\left(\frac{\mathcal{B}c_9}{c_{10}}\right) + \log\left(\mathcal{B}c_{11}\right)\right)\right) \\ &=: B_0. \end{aligned} \tag{1.31}$$

[1] Dans ce cas, il faut supprimer le facteur b_0 de la borne supérieure, car on minore $|\mathrm{Log}\,\xi|$ et non $|\mathrm{Log}\,\xi^{b_0}|$, avec les notations de (1.20).

1.4 La réduction de la borne

1.4.1 Introduction

La borne (1.31) est malheureusement bien loin d'être praticable, à cause de la taille de la constante c_9. Pour les équations les plus banales (degré 3 ou 4, coefficients raisonnables, invariants du corps petits), on se retrouve très facilement avec des bornes de l'ordre d'une vingtaine ou une trentaine de chiffres. Quant à des équations de degré 20 ou 30, inutile d'espérer obtenir une borne ayant moins de 80 chiffres.

Malgré cela, on s'aperçoit que souvent, toutes les solutions correspondent à des b_i très petits. La mauvaise qualité de la borne obtenue via le Théorème 1.9 est une conséquence de sa généralité ; il faut donc exploiter les propriétés numériques de l'équation, et plus précisément de la forme linéaire en logarithmes, pour transformer cette borne initiale en borne "utilisable".

Pour arriver à nos fins, comme le majorant de (1.30) est à peu près optimal, il faut améliorer le minorant. À cette fin, on va remplacer la borne de Baker par une version effective : le fait que l'on soit maintenant capable d'affirmer que les b_i sont bornés va nous permettre de donner un nouveau minorant bien plus réaliste, sans bien sûr énumérer tous les $(r+1)-$uplets possibles.

Oublions un instant le contexte, et intéressons nous au problème suivant : soit $\theta_0, \ldots, \theta_r$ des nombres réels, et $f_0, \ldots, f_r$ des nombres entiers. Comment trouver

$$\min_{|f_i| \leq B', 0 \leq i \leq r} \left| \sum_{i=0}^{r} f_i \theta_i \right| ?$$

Ce problème revient à trouver des approximations rationnelles simultanées des différents nombres θ_i ; dans le cas où $r = 1$, on n'a que deux valeurs ; cela revient à résoudre un problème de fractions continues. Nous verrons que l'on peut se ramener à ce cas si b_0 est connu (par exemple si le système d'unités est fondamental).

Dans le cas général, deux algorithmes existent pour trouver de bonnes approximations diophantiennes simultanées.

1.4.1.1 L'algorithme LLL

Dans cette section, n est un entier fixé ; on notera $\| \cdot \|_2$ la norme euclidienne standard sur $\mathbb{R}^n$, et $d(\cdot, \mathbb{Z})$ la distance à $\mathbb{Z}$.

Soit Λ un réseau de dimension n, c'est-à-dire un $\mathbb{Z}$-sous module libre de rang n de $\mathbb{R}^n$, engendré par les vecteurs $(e_1, \ldots, e_n)$. Le réseau Λ est donc isomorphe à $\mathbb{Z}^n$ par

$$\begin{array}{rccc} f : & \mathbb{Z}^n & \to & \Lambda \\ & (a_1, \ldots, a_n) & \mapsto & \sum_{1 \leq i \leq n} a_i e_i. \end{array}$$

On identifiera Λ et $\mathbb{Z}^n$ dans la suite, c'est-à-dire que la notation $(a_1, \ldots, a_n)$ désignera le vecteur $f(a_1, \ldots, a_n)$. On peut munir le module Λ de la forme quadratique $q((a_1, \ldots, a_n)) = \sum_{i,j} a_i a_j (e_i|e_j)$, où $(\cdot|\cdot)$ est le produit scalaire canonique sur $\mathbb{Z}^n$.

Dans un espace vectoriel, on peut trouver des bases orthonormées pour une forme quadratique non dégénérée donnée ; dans un module, ce n'est plus en général le cas. Un des problèmes centraux de la théorie algorithmique des réseaux consiste justement à trouver des bases constituées de vecteurs les plus courts et les plus orthogonaux possibles. L'algorithme LLL décrit dans [LLL82] fournit une réponse partielle à ce problème :

Définition 1.22. *Soit Λ un réseau de dimension n, $(\boldsymbol{b}_i)_{1\leq i\leq n}$ une base de Λ. Soit $(\boldsymbol{b}_i^*)_{1\leq i\leq n}$ la base orthogonale déduite de $(\boldsymbol{b}_i)$ par le procédé de Gram-Schmidt, c'est-à-dire*

$$\boldsymbol{b}_i^* = \boldsymbol{b}_i - \sum_{j=1}^{i-1} \mu_{ij} \boldsymbol{b}_j^* \text{ ; on note } \mu_{ij} = (\boldsymbol{b}_i, \boldsymbol{b}_j^*)/(\boldsymbol{b}_j^*, \boldsymbol{b}_j^*).$$

La base $(\boldsymbol{b}_i)_{1\leq i\leq n}$ est dite LLL-réduite si

$$|\mu_{ij}| \leq 1/2, \quad 1 \leq j < i \leq n,$$

et

$$\|\boldsymbol{b}_i^* + \mu_{ii-1}\boldsymbol{b}_{i-1}^*\|_2^2 \geq 3\|\boldsymbol{b}_{i-1}^*\|_2^2/4, \quad 1 < i \leq n.$$

Théorème 1.23. *Il existe un algorithme en temps polynomial pour déterminer une base LLL-réduite d'un réseau Λ donné. Cette base vérifie de plus les propriétés suivantes :*

$$\|\boldsymbol{b}_1\|_2^2 \leq 2^{n-1}\|\boldsymbol{x}\|_2^2$$

pour tout $\boldsymbol{x} \in \Lambda - \{(0, \ldots, 0)\}$, ou, plus généralement

$$\|\boldsymbol{b}_j\|_2^2 \leq 2^{n-1} \max(\|\boldsymbol{x}_1\|_2^2, \ldots, \|\boldsymbol{x}_t\|_2^2),$$

pour toute famille $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t)$ de vecteurs linéairement indépendants avec $j \leq t$.

Le fait que l'on soit capable de minorer la longueur du plus court vecteur non nul est cruciale. C'est ce qui nous servira dans toute la suite.

Le détail de l'algorithme LLL et ses raffinements ne seront pas décrits ici. Le lecteur intéressé peut se reporter, par exemple, à [Co].

Notons que l'algorithme LLL peut, *a priori*, traiter n'importe quel sous-réseau de $\mathbb{R}^n$. Toutefois, appliquer LLL à un réseau réel est source de gros problèmes de stabilité numérique. Nous allons voir comment se ramener à un réseau entier ; il existe alors une variante de LLL due à de Weger [We87], ne faisant que des calculs entiers.

1.4.1.2 L'algorithme de Fincke et Pohst

L'algorithme de Fincke et Pohst [FP85] a un objectif légèrement différent ; étant donné une borne M, il trouve tous les vecteurs du réseau dont la norme est au plus M. En particulier, soit il trouve le vecteur non nul le plus court et sa norme, soit le vecteur non nul le plus court est de norme au moins M.

Le principe de cet algorithme est très simple. Pour les finesses et détails d'implantation, voir [Co]. On considère la norme d'un vecteur du réseau comme une forme quadratique définie positive en n variables, que l'on peut, quitte à changer de base, supposer diagonale. On cherche donc les $\boldsymbol{x} = (x_1, \ldots, x_n)$ tels que $Q(\boldsymbol{x}) = q_{11}x_1^2 + \ldots + q_{nn}x_n^2 \leq C$. Cette condition impose $|x_1| \leq \sqrt{C/q_{11}}$. Pour chacune des valeurs de x_1, on a $|x_2| \leq \sqrt{(C - q_{11}x_1^2)/q_{22}}$, etc.

Cet algorithme est *a priori* plus adapté à nos besoins que l'algorithme LLL, puisqu'il fournit le vecteur non nul le plus court. Toutefois il faut noter que sa complexité est exponentielle en fonction des données d'entrée. On lui préférera donc LLL pour les premières phases de réduction, pour lesquelles des minorations grossières sont suffisantes. Une fois épuisées les ressources de LLL, il est possible d'utiliser Fincke-Pohst pour une ou deux étapes supplémentaires de réduction.

Par contre, Fincke-Pohst sera une des méthodes pour réaliser l'énumération des solutions plus petites que la borne réduite, et probablement la plus efficace dans le cas (rare) ou il existe une "grosse" solution.

L'inconvénient principal de Fincke-Pohst est sa complexité ; toutefois, la remarque (1.26) nous permettra de ne l'appliquer qu'à un réseau de dimension 3, ce qui ne pose guère de problèmes.

1.4.2 Mise en œuvre de la réduction

Nous allons présenter, dans l'ordre chronologique, les quatres méthodes de réduction ayant été utilisées à ce jour ; elles sont dues successivement à Tzanakis et de Weger [TW89], Mignotte et de Weger (sur une idée de Bilu)[MW94], Bilu et moi-même [BH96], et enfin Bennett et de Weger [BW96]. Toutes reposent sur le même principe, à savoir transformer le problème de petites valeurs de formes linéaires en logarithmes en un problème de vecteurs courts dans un certain réseau ; seule la mise en œuvre diffère, à travers le choix du réseau. Toutes ces méthodes ont été adaptées ici pour tenir compte de la présence du b_0. Cela est sans incidence majeure, sauf dans le cas de la troisième méthode, pour laquelle on exposera donc le cas $b_0 = 1$ (le cas $b_0 \neq 1$ apparaîtra comme un cas particulier de la quatrième méthode).

1.4.2.1 La méthode de Tzanakis et de Weger

Dans cette méthode, on étudie la forme linéaire en logarithmes

$$\left|\sum_{i=0}^{r} b_i \log \beta_i\right| \leq b_0 c_{12} \exp(-c_{13} B/b_0), \tag{1.32}$$

où $c_{12} = 2c_3 \exp(nc_6/c_5)$, $c_{13} = n/c_5$,

$$\beta_0 = \left|\frac{\alpha^{(2)} - \alpha}{\alpha^{(3)} - \alpha} \cdot \frac{\mu^{(3)}}{\mu^{(2)}}\right|, \quad \beta_i = \left|\frac{\eta_i^{(3)}}{\eta_i^{(2)}}\right|.$$

On a de plus les contraintes $B = \max_{1\leq i\leq r} |b_i| \leq B_0$, $b_0 \leq \mathcal{B}$.

Remarque 1.24. Dans le cas où $r = 1$, $\left|\eta_1^{(3)}/\eta_1^{(2)}\right| = 1$, il faut considérer les arguments des éléments, plutôt que le module, et l'on se retrouve avec une forme linéaire en 3 variables

$$\left|b_0 \mathrm{Arg}\left(\frac{\alpha^{(2)} - \alpha}{\alpha^{(3)} - \alpha} \cdot \frac{\mu^{(3)}}{\mu^{(2)}}\right) + b_1 \mathrm{Arg}\,\frac{\eta_1^{(3)}}{\eta_1^{(2)}} + 2\pi b_2\right| \leq c_{11} \exp(-c_{10} B''),$$

qui se traite de la même manière que la précédente.

On note dans la suite $\lfloor u \rceil$ l'entier le plus proche d'un réel u, et arrondissant inférieurement en cas d'ambiguïté.

On considère le réseau engendré par les colonnes de la matrice

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ \lfloor C \log \beta_0 \rceil & \lfloor C \log \beta_1 \rceil & \dots & \lfloor C \log \beta_{r-1} \rceil & \lfloor C \log \beta_r \rceil \end{pmatrix}.$$

où C est un grand entier dont le choix sera discuté plus loin.

Appliquant LLL au réseau ci-dessus, on obtient une base LLL-réduite. La longueur du premier vecteur de cette base sera notée l_1. On a alors le

Théorème 1.25. *Supposons que*

$$l_1 \geq 2^{r/2} \sqrt{(r-1)B_0^2 + \mathcal{B}^2 + \left(\frac{rB_0 + \mathcal{B}}{2}\right)^2}. \tag{1.33}$$

Alors pour tout $(r+1)$-uplet d'entiers $(b_0, \dots, b_r)$, on a

$$\left|\sum_{i=0}^{r} b_i \log \beta_i\right| \geq \frac{1}{C}\left(\sqrt{2^{-r} l_1^2 - (r-1)B_0^2 - \mathcal{B}^2} - \frac{rB_0 + B}{2}\right).$$

Preuve. Le Théorème 1.23 nous montre que pour tout vecteur $\boldsymbol{b} = (b_0, \ldots, b_r)$, on a

$$\|A\boldsymbol{b}\|_2 \geq 2^{-r/2} l_1.$$

Il nous faut donc majorer $\|A\boldsymbol{b}\|_2$ en terme de la forme linéaire en logarithmes (1.32).

On a

$$A\boldsymbol{b} = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{r-1} \\ \sum_{0\leq i\leq r} b_i \lfloor C \log \beta_i \rceil \end{pmatrix},$$

d'où

$$\|A\boldsymbol{b}\|_2^2 = \sum_{0\leq i\leq r-1} b_i^2 + \left(\sum_{0\leq i\leq r} b_i \lfloor C \log \beta_i \rceil \right)^2.$$

On majore alors b_0^2 par $\mathcal{B}^2$ et chacun des autres termes b_i^2 par B_0^2, il vient :

$$\left| \sum_{0\leq i\leq r} b_i \lfloor C \log \beta_i \rceil \right| \geq \sqrt{2^{-r} l_1^2 - \mathcal{B}^2 - (r-1) B_0^2}.$$

Mais

$$\left| \sum_{0\leq i\leq r} b_i \lfloor C \log \beta_i \rceil - C \sum_{0\leq i\leq r} b_i \log \beta_i \right| \leq \sum_{0\leq i\leq r} b_i/2 \leq \frac{\mathcal{B} + r B_0}{2},$$

d'où

$$\left| \sum_{0\leq i\leq r} b_i \log \beta_i \right| \geq \frac{1}{C} \left(\sqrt{2^{-r} l_1^2 - (r-1) B_0^2 - \mathcal{B}^2} - \frac{r B_0 + \mathcal{B}}{2} \right).$$

□

Définissons alors

$$l_0 := \frac{1}{C} \left(\sqrt{2^{-r} l_1^2 - (r-1) B_0^2 - \mathcal{B}^2} - \frac{r B_0 + \mathcal{B}}{2} \right).$$

Corollaire 1.26. *Supposons (1.33) vérifiée. Alors*

$$\max_{1\leq i\leq r} |b_i| \leq \frac{\mathcal{B}}{c_{13}} \log \frac{\mathcal{B} c_{12}}{l_0}.$$

Preuve. Il suffit de comparer le majorant de (1.32) et le minorant du Théorème 1.25. □

Reste à discuter le choix de C. Le bon choix est celui qui donnera un l_0 maximal ; heuristiquement, il faut choisir C le plus petit possible de sorte que (1.32) soit vérifiée. Si l'on suppose que tous les vecteurs de la base réduite sont du même ordre de grandeur, on obtient comme déterminant du réseau quelque chose de l'ordre de l_1^{r+1}, puisque la base est relativement orthogonale. Comme le discriminant est en fait de l'ordre de C, il faut alors choisir $C \approx l_1^{r+1}$. La condition sur l_1 imposant à celui-ci d'être de l'ordre d'une petite constante fois B_0, il apparaît que le bon choix heuristique pour C est de l'ordre de κB_0^{r+1}, avec κ de l'ordre de 10 ou 100. En pratique, on essaie avec une première valeur de κ ; si la condition (1.33) n'est pas vérifiée, on recommence en augmentant κ.

Remarque 1.27. Quand le corps n'a qu'un plongement réel, il faut être un peu précautionneux : si les racines $\alpha^{(2)}$ et $\alpha^{(3)}$ sont choisies imaginaires conjuguées, les β_i valent 1 et la forme linéaire (1.32) est nulle. Il convient donc dans ce cas d'utiliser la partie imaginaire de la forme linéaire (1.15), ce qui a l'inconvénient d'augmenter la dimension du réseau d'une unité.

Remarque 1.28. Avec ce choix de C, on voit que la nouvelle borne pour B dépend du logarithme de la précédente, ce qui explique que numériquement la réduction est très efficace.

Remarque 1.29. Le procédé décrit ci-dessus peut être réitéré avec la nouvelle valeur obtenue pour B_0. Notons que, heuristiquement, la valeur limite que l'on peut espérer obtenir est approximativement la solution en B de l'équation

$$B = \frac{\mathcal{B}}{c_{13}} \left(r \log B + \log(\mathcal{B} c_{12}) \right). \tag{1.34}$$

Dans le cas où $\mathcal{B} = 1$, la réduction est en règle générale très efficace, et fournit ultimement des bornes de l'ordre de quelques dizaines d'unités.

Remarque 1.30. On peut ici exploiter une idée se trouvant dans l'article [TW92], qui permet de tirer parti du fait que l'on contrôle mieux b_0 que les autres b_i pour diminuer un peu la valeur de C, et donc la précision nécessaire : il suffit, dans la matrice A, de remplacer le 1 supérieur gauche par $\lfloor B_0/\mathcal{B} \rfloor$; on peut alors choisir C de l'ordre de $\mathcal{B} B_0^r$, et la réduction s'améliore également un peu.

Remarque 1.31. Il va de soi qu'il est possible de conserver la même méthode quand le système d'unités est fondamental ; toutefois il vaut mieux procéder de la manière suivante : au lieu d'estimer la longueur du plus court vecteur du réseau, on estime la distance entre le point ${}^t(0, 0, \ldots, [C \log(\beta_0)])$ et le réseau. Cela peut encore se faire au moyen de LLL, via le

Lemme 1.32 ([We87]). *Soit $\boldsymbol{x} = (x_i)$ un vecteur de $\mathbb{Z}^n$, et $A = (\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$ une base LLL-réduite d'un réseau Λ. Posons $\boldsymbol{s} = (s_i) = A^{-1}(x_i)$. Alors*

$$d(\boldsymbol{x}, \Lambda) \geq 2^{(1-n)/2} d(s_{i^*}, \mathbb{Z}) \|\boldsymbol{b}_1\|_2,$$

où i^* est le plus grand entier i tel que $d(s_i, \mathbb{Z}) \neq 0$.

Preuve. On considère la base $(\boldsymbol{b}_1^*, \ldots, \boldsymbol{b}_n^*)$, orthogonalisée de Gram-Schmidt de la base $(\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n)$. Les $\boldsymbol{b}_i$ sont donnés dans la base $(\boldsymbol{b}_i^*)$ par

$$\boldsymbol{b}_i = \boldsymbol{b}_i^* + \sum_{j=1}^{i-1} \mu_{ij} \boldsymbol{b}_j^*,$$

où $\mu_{ij} = (\boldsymbol{b}_i, \boldsymbol{b}_j^*)/\|\boldsymbol{b}_j^*\|_2^2$. On définira dans la suite $\mu_{ii} = 1$.

Soit $\boldsymbol{k} \in \mathbb{Z}^n$, et formons $\|A\boldsymbol{k} - \boldsymbol{x}\|_2 = \|A(\boldsymbol{k} - \boldsymbol{s})\|_2$:

$$\begin{aligned} A(\boldsymbol{k} - \boldsymbol{s}) &= \sum_{i=1}^{n} (k_i - s_i) \boldsymbol{b}_i, \\ &= \sum_{i=1}^{n} \sum_{j=1}^{i} (k_i - s_i) \mu_{ij} \boldsymbol{b}_j^*, \\ &= \sum_{j=1}^{n} \left(\sum_{i=j}^{n} \mu_{ij} (k_i - s_i) \right) \boldsymbol{b}_j^*. \end{aligned}$$

Par suite,

$$\|A(\boldsymbol{k} - \boldsymbol{s})\|_2^2 = \sum_{j=1}^{n} \left(\sum_{i=j}^{n} \mu_{ij} (k_i - s_i) \right)^2 \|\boldsymbol{b}_j^*\|_2^2.$$

Posons $i_1 := \max\{i : k_i \neq s_i\}$. Il est clair que $i_1 \geq i^*$. Si $i_1 > i^*$, on a $|k_{i_1} - s_{i_1}| \geq 1 \geq d(s_{i^*}, \mathbb{Z})$, et sinon $|k_{i_1} - s_{i_1}| \geq d(s_{i^*}, \mathbb{Z})$. Dans tous les cas, on a donc

$$\|A(\boldsymbol{k} - \boldsymbol{s})\|_2^2 \geq d(s_{i^*}, \mathbb{Z})^2 \|\boldsymbol{b}_{i_1}^*\|_2^2.$$

On utilise maintenant [LLL82, (1.7)] pour conclure. □

On dispose alors de résultats tout à fait semblables à ceux ci-dessus, sauf que la dimension du réseau a diminué d'une unité, ce qui rend cette méthode préférable dans le cas où b_0 est connu.

1.4.2.2 La méthode de Mignotte et de Weger

Cette modification de la méthode précédente repose sur la remarque suivante ; en réduisant le réseau ci-dessus, on se contente de contraindre une forme linéaire en logarithmes à être petite. Mais le problème qui nous intéresse est bien plus précis, puisqu'il impose à $r-1$ formes indépendantes d'être simultanément petites, ce qui est une exigence très forte. Ces $r-1$ formes sont celles mentionnées plus haut, en remplaçant le couple $(2, 3)$ par $(2, i)$ pour $3 \leq i \leq r+1$.

On utilise donc les $r-1$ formes

$$\left|\sum_{i=0}^{r} b_i \log \beta_i^{(j)}\right| \leq b_0 c_{12} \exp(-c_{13} B/b_0), \quad 1 \leq j \leq r-1, \tag{1.35}$$

avec des $\beta_i^{(j)}$ analogues aux β_i de la section précédente.

La matrice dont les colonnes définissent le réseau est donc

$$A = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \lfloor C \log \beta_0^{(1)} \rceil & \lfloor C \log \beta_1^{(1)} \rceil & \dots & \lfloor C \log \beta_{r-1}^{(1)} \rceil & \lfloor C \log \beta_r^{(1)} \rceil \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \lfloor C \log \beta_0^{(r-1)} \rceil & \lfloor C \log \beta_1^{(r-1)} \rceil & \dots & \lfloor C \log \beta_{r-1}^{(r-1)} \rceil & \lfloor C \log \beta_r^{(r-1)} \rceil \end{pmatrix}.$$

où, de nouveau, le choix de C sera discuté ultérieurement.

L'équivalent du Théorème 1.25 est le

Théorème 1.33. *Soit l_1 le plus court vecteur d'une base LLL-réduite du réseau engendré par les colonnes de A. Alors si*

$$l_1 > 2^{r/2} \sqrt{\left(\frac{(r-1)(rB_0 + \mathcal{B})}{2}\right)^2 + \mathcal{B}^2 + B_0^2},$$

on a, pour tout $(r+1)$-uplet $(b_0, \dots, b_r)$,

$$\left|\sum_{i=0}^{r} b_i \log \beta_i^{(j)}\right| > \frac{1}{C}\left(\sqrt{\frac{2^{-r} l_1^2 - \mathcal{B}^2 - B_0^2}{r-1}} - \frac{rB_0 + \mathcal{B}}{2}\right).$$

Preuve. En tous points semblable à celle du Théorème 1.25. □

On a un corollaire analogue à 1.26, en remplaçant l_0 par l_0'.

Le bon choix de C, par les mêmes arguments heuristiques que précédemment, est de l'ordre de $B_0^{(r+1)/(r-1)}$.

Dans les faits, la réduction est bien plus efficace que par la méthode précédente, mais présente l'inconvénient d'être bien plus lente dès que r augmente.

1.4.2.3 Réduction au cas de la dimension 2 ou 3.

Les méthodes décrites ci-dessus ont toutes deux un important inconvénient, qui est d'imposer la réduction LLL d'un réseau de dimension r dont les coefficients sont très grands (rappelons que C est plus grand que B_0, qui initialement a plusieurs dizaines de chiffres). Un tel problème dépasse rapidement la capacité des machines quand le degré devient grand. On montre dans cette section comment

il est possible de se ramener au cas de la dimension 2 si b_0 est connu ; on verra dans la section suivante comment se ramener au cas de la dimension 3 si l'on ne connaît qu'un majorant de b_0.

Le principe consiste tout simplement à exploiter de nouveau le fait que l'on ait $r-1$ relations en r variables, mais cette fois-ci à en extraire une relation en 2 variables.

Cela peut se faire au moyen de l'identité (1.26), qui dit en particulier que si $|x| > X_2$, on a

$$|b_i - b_0\delta_i \log|x| - b_0\lambda_i| \leq b_0 \frac{c_{14}}{|x|^n}, \tag{1.36}$$

où $c_{14} = c_4 \max\limits_{1\leq i\leq r} \sum\limits_{j=1}^{r} |a_{ij}|$.

Pour éliminer x entre deux de ces relations, disons la $i_1^{\text{ème}}$ et la $i_2^{\text{ème}}$, on introduit les quantités

$$\overline{\delta}_i = \delta_i \delta_{i_1}^{-1}, \quad \overline{\lambda}_i = \lambda_i - \overline{\delta}_i \lambda_{i_1}, \quad 1 \leq i \leq r. \tag{1.37}$$

Pour des raisons de stabilité numérique, le meilleur choix de i_1 est celui qui correspond au maximum des $|\delta_i|$.

On a alors la proposition suivante :

Proposition 1.34. *On a*

$$\left|b_{i_2} - \overline{\delta}_{i_2} b_{i_1} - \overline{\lambda}_{i_2} b_0\right| \leq (1 + |\overline{\delta}_{i_2}|) b_0 \frac{c_{14}}{|x|^n}.$$

Preuve. Ceci résulte de la chaîne d'inégalités suivante

$$\begin{aligned}
\left|b_{i_2} - \overline{\delta}_{i_2} b_{i_1} - \overline{\lambda}_{i_2} b_0\right| &\leq |b_{i_2} - b_0\delta_{i_2}\log|x| - b_0\lambda_{i_2}| + \\
&\quad |b_0\delta_{i_2}\log|x| + b_0\lambda_{i_2} - \overline{\delta}_{i_2} b_{i_1} - \overline{\lambda}_{i_2} b_0|, \\
&\leq b_0 \frac{c_{14}}{|x|^n} + |\overline{\delta}_{i_2}||b_{i_1} - b_0\delta_{i_1}\log|x| - b_0\lambda_{i_1}|, \\
&\leq (1 + \overline{\delta}_{i_2}) b_0 \frac{c_{14}}{|x|^n},
\end{aligned}$$

ce qui conclut la preuve. □

Appliquant alors (1.27), il vient :

$$\left|b_{i_2} - \overline{\delta}_{i_2} b_{i_1} - \overline{\lambda}_{i_2} b_0\right| \leq b_0 c_{15} \exp(-c_{13} B/b_0). \tag{1.38}$$

avec $c_{15} = (1 + |\overline{\delta}_{i_2}|) c_{14} \exp(n c_6/c_5)$.

Étudions maintenant le cas où b_0 est connu, par exemple $b_0 = 1$; le cas où b_0 est inconnu app araîtra comme un cas particulier de la section suivante. En

appliquant les techniques de Tzanakis et de Weger à l'identité ci-dessus, on se retrouve à résoudre un problème de réduction d'un réseau de dimension 2 ; un tel problème peut être résolu de manière exacte par LLL qui coïncide dans ce cas avec l'algorithme de Gauss ; comme de plus la première ligne de la matrice est $(1, 0)$, on n'a affaire à ni plus ni moins qu'un problème de fractions continues. Dans ce contexte, on peut revenir à la méthode de Baker et Davenport, voir [BD69].

Soit q un entier de l'ordre de κB_0. On a alors :

$$\left|qb_{i_2} - q\overline{\delta}_{i_2}b_{i_1} - q\overline{\lambda}_{i_2}\right| \leq qc_{15}\exp(-c_{13}B).$$

On minore le membre gauche par sa distance à $\mathbb{Z}$. On a

$$\begin{aligned} \|q\overline{\delta}_{i_2}b_{i_1} + q\overline{\lambda}_{i_2}\| &\leq qc_{15}\exp(-c_{13}B), \\ \|q\overline{\lambda}_{i_2}\| - b_{i_1}\|q\overline{\delta}_{i_2}\| &\leq qc_{15}\exp(-c_{13}B). \end{aligned}$$

Maintenant, on choisit pour q le dénominateur q_0 de la plus grande réduite du développement en fractions continues de $\overline{\delta}_{i_2}$ avec $q \leq \kappa B_0$. On a alors le

Théorème 1.35. *Si* $l_0'' := \|q_0\overline{\lambda}_{i_2}\| - B_0\|q_0\overline{\delta}_{i_2}\| > 0$, *on a*

$$B \leq c_{13}^{-1}\log\left(\frac{q_0c_{15}}{l_0''}\right).$$

Remarque 1.36. Notons que le choix de q est libre, mais il faut que $B_0\|q_0\overline{\delta}_{i_2}\|$ soit petit, donc que $\|q_0\overline{\delta}_{i_2}\|$ soit de l'ordre de $1/q_0$, ce qui impose plus ou moins le choix d'une réduite du développement en fractions continues.

Notons au passage que $\overline{\delta}_{i_2}$ doit être connu assez précisément pour pouvoir majorer $B_0\|q_0\overline{\delta}_{i_2}\|$. Ainsi, si $\tilde{\delta}$ est la valeur approchée de $\overline{\delta}_{i_2}$, on a

$$|B_0\|q_0\overline{\delta}_{i_2}\| - B_0\|q_0\tilde{\delta}\|| \leq \kappa B_0^2|\overline{\delta}_{i_2} - \tilde{\delta}|,$$

ce qui impose de connaître δ avec une précision de l'ordre de B_0^2. Ceci est cohérent avec ce que l'on aurait obtenu en réduisant par les méthodes des sections précédentes un réseau de $\mathbb{Z}^2$.

Pour une discussion sur le calcul de $\overline{\delta}_{i_2}$ à une telle précision, voir la section 1.7.

Cette méthode de réduction est d'efficacité similaire à la précédente, mais beaucoup plus rapide ; on a en effet remplacé l'algorithme LLL (en dimension éventuellement grande) par le simple algorithme des fractions continues, qui est très rapide. Dans la pratique, l'usage de cette méthode (ou de celle qui suit) est donc vivement recommandée. Toutefois, dans certaines situations (liées à d'autres équations diophantiennes, voir la méthode des courbes elliptiques au chapitre 5), on aura parfois recours, faute de mieux, à la méthode de Tzanakis et de Weger.

Décrivons maintenant la dernière méthode, qui combine les idées des deux sections précédentes.

1.4.2.4 La méthode de Bennett et de Weger

De manière schématique, la méthode de Tzanakis et de Weger consiste à réduire une forme en $r+1$ variables ; la méthode de Mignotte et de Weger consiste à réduire $r-1$ formes en $r+1$ variables ; la méthode de la section précédente consiste à réduire une forme en deux variables ; et l'idée qui suit consiste à réduire plusieurs formes en deux variables.

Pour simplifier les notations, supposons par exemple que $i_1 = 1$. On va alors réduire les relations (1.38) pour $i_2 = 2, \dots, r$.

On considère de nouveau un grand C, et le réseau engendré par les colonnes de la matrice

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \lfloor -C\overline{\lambda}_2 \rceil & \lfloor -C\overline{\delta}_2 \rceil & C & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lfloor -C\overline{\lambda}_r \rceil & \lfloor -C\overline{\delta}_r \rceil & 0 & \dots & C \end{pmatrix}.$$

Le bon choix de C est cette fois, pour les mêmes raisons que précédemment, de l'ordre de $B_0^{(r+1)/(r-1)}$.

On a le

Théorème 1.37. *Soit l_1 la longueur du plus court vecteur d'une base LLL-réduite engendrée par les colonnes de A. Alors, si*

$$l_1 > 2^{r/2} \sqrt{\left(\frac{(r-1)(B_0 + \mathcal{B})}{2} \right)^2 + \mathcal{B}^2 + B_0^2} =: l_0'$$

on a, pour tout $(r+1)$-uplet $(b_0, \dots, b_r)$,

$$\left| \sum_{i=0}^{r} b_i \log \beta_i^{(j)} \right| > \frac{1}{C} \left(\sqrt{\frac{2^{-r} l_1^2 - \mathcal{B}^2 - B_0^2}{r-1}} - \frac{B_0 + \mathcal{B}}{2} \right).$$

Remarque 1.38. C'est dans la pratique la méthode la plus efficace, avec des exigences de précision relativement modestes (par rapport aux autres). En compensation, quand r augmente, LLL en dimension r devient vite assez lent ; mon conseil serait de ne pas utiliser les r formes disponibles, mais de se contenter de 2 à 5 d'entre elles, suivant le degré — le 2 se rapporte aux degrés élevés ; plus le degré est élevé, plus C en général sera grand et plus il sera difficile d'arriver à réduire le réseau correspondant, même si celui-ci est de dimension modeste.

Remarque 1.39. On peut remarquer que $\mathcal{B}$, lui, n'a été réduit dans aucune de ces méthodes. C'est normal vu la manière dont on a procédé, c'est-à-dire en écrivant que si x est petit, les b_i sont petits ; pour b_0, c'est le contraire qui se produit. Il est toutefois possible de réduire la borne pour $\mathcal{B}$ au moyen d'une des deux inégalités du lemme suivant :

Lemme 1.40. *(i) Supposons que* $|x| > (20c_{14})^{1/n}$*. Alors si* $|\overline{\lambda}_{i_2}| > 0,1$*, on a*

$$b_0 \leq B\frac{1+|\overline{\delta}_{i_2}|}{|\overline{\lambda}_{i_2}|-0,1}.$$

(ii) Supposons que

$$|x| > \max\left(X_2, \min_i \exp\left(\frac{(2B/\mathcal{B})+0,1+|\lambda_i|}{|\delta_i|}\right)\right).$$

Alors

$$b_0 \leq \mathcal{B}/2.$$

Preuve. *(i)* est une conséquence facile de la proposition 1.34, et *(ii)* de (1.26).□

Dans la pratique, l'une comme l'autre de ces deux inégalités n'ont que peu d'intérêt ; la première s'applique très rarement, et n'améliore pas $\mathcal{B}$ quand elle s'applique, et la seconde donne une borne sur $|x|$ trop grande pour être exploitable.

Quel que soit le procédé de réduction utilisé, on notera B^* la borne obtenue après plusieurs étapes de réduction, et $\mathcal{B}^*$ la borne (éventuellement réduite) pour b_0.

1.4.3 Mauvaise réduction

On explique sommairement dans cette section les raisons qui peuvent faire que les diverses méthodes de réduction de la borne échouent, et comment y remédier.

1.4.3.1 Mauvaise réduction, dimension ≥ 3, cas inhomogène

Ce problème intervient dans le cas $b_0 = 1$, quand $\log\beta_0$ est combinaison linéaire des $\log\beta_i$. On "devine" alors à l'aide de LLL une telle combinaison linéaire :

$$|\log\beta_0 - \sum_i a_i \log\beta_i| \leq \varepsilon,$$

de sorte que l'on cherche en fait à minorer

$$\left|\sum_{i=1}^{r}(b_i-a_i)\log\beta_i\right| - \varepsilon,$$

sous la contrainte $\max_i |b_i| \leq B_0 + \max_i |a_i|$. On peut donc supposer que l'on est ramené à un problème homogène.

1.4.3.2 Mauvaise réduction, dimension ≥ 3

Ce phénomène peut intervenir pour diverses raisons ; en particulier, il y a un problème dès que les nombres $\log \beta_i$ de (1.32) sont linéairement dépendants sur $\mathbb{Q}$, car dans ce cas, le vecteur le plus court trouvé par LLL sera la relation de dépendance linéaire correspondante.

Dans ce cas, en utilisant LLL, on "devine" une relation de dépendance linéaire $|\sum_{k=1}^{n} a_k \log \beta_k| \leq \varepsilon_1$, avec les a_k entiers, et l'on remplace l'un des $\log \beta_l$ tels que $a_l \neq 0$ par $-\sum_{k \neq l}(a_k \log \beta_k)/a_l$, ce qui donne, dans (1.32) :

$$\left| \sum_{i=0, i \neq l}^{r} (a_l b_i - b_l a_i) \log \beta_i \right| \leq a_l (b_0 c_{12} \exp(-c_{13} B / b_0) + \varepsilon_1).$$

On pose alors $b'_i = a_l b_i - b_l a_i$, on a $|b'_i| \leq 2 \max_k |a_k| B_0$, et on est ramené à un problème analogue à (1.32), avec une dimension de moins. En conséquence, soit la réduction finit par fonctionner, soit on peut supposer que l'on est ramené au cas de la dimension 2.

1.4.3.3 Mauvaise réduction, dimension 2

Dans la section précédente, on s'est ramené à la dimension 2, cas homogène. Pour être complet, il nous faut aussi discuter le cas inhomogène de la dimension 2 qui peut intervenir lors de la méthode des fractions continues.

La mauvaise réduction intervient essentiellement lorsque $\overline{\delta}_i, \overline{\lambda}_i$ et 1 sont $\mathbb{Q}$-linéairement dépendants. Notons que le cas où $\overline{\delta}_i$ est seul rationnel ne pose pas de problème, pas plus que le cas où $\overline{\delta_i}$ et $\overline{\lambda}_i$ sont tous deux irrationnels, qui se traite par des méthodes analogues à celles de la section précédente.

Premier cas : $\overline{\lambda}_i$ rationnel, $\overline{\delta}_i$ irrationnel.

On n'a dans les faits de problème à ce stade que quand $\overline{\lambda}_i$ est un rationnel de petit dénominateur, voire un entier ; écrivons alors $\overline{\lambda}_i = p/q + \varepsilon$. Quitte à multiplier par q et à minorer $|\overline{\lambda}_i + b_i \overline{\delta}_i + b_1|$ par sa distance à $\mathbb{Z}$, on peut supposer $p/q = 0$. On a donc à minorer $|b_i \overline{\delta}_i + b_1|$ sous la condition $|b_i| \leq q B_0$, ce qui se fait très facilement en calculant la réduite p_k/q_k du développement en fractions continues de $\overline{\delta}_i$ dont le dénominateur est immédiatement inférieur à qB_0 ; le minorant cherché est alors $|q_k \overline{\delta}_i + p_k|$.

Second cas : $\overline{\lambda}_i$ et $\overline{\delta}_i$ sont rationnels

On va voir que ce cas se réduit au cas $r = 1$. Posons $\overline{\lambda}_i = p_i/q_i + \varepsilon_i$, $\overline{\delta}_i = p'_i/q'_i + \varepsilon'_i$.

Soit $Q = \max_i(q_i, q'_i)$ et $E = \max_i(\varepsilon_i, \varepsilon'_i)$. Alors pour tout i on a

$$|q'_i q_i b_i - p'_i q_i b_{i_1} + p_i q'_i| \leq 2Q^2 \left(\frac{(1 + |\overline{\delta}_i|) c_{14}}{|x|^n} + E \right)$$

de sorte que si $x \geq X_2' := \left(\frac{2c_{14}Q^2}{1-4Q^2E}\right)^{1/n}$, on a

$$q_i' q_i b_i = p_i' q_i b_{i_1} + p_i q_i', \qquad (1 \leq i \leq r). \tag{1.39}$$

En particulier, il n'y a pas de solution plus grande que X_2' dès lors qu'il existe un i tel que $q_i \not| q_i'$.

Supposons alors que $q_i \mid q_i'$ pour tout i. D'après le théorème de Bézout, il existe ξ_i et χ_i tels que $q_i'\chi_i - p_i'\xi_i = p_i q_i'/q_i$ pour tout i, et

$$b_i = \chi_i + t_i p_i' \tag{1.40}$$
$$b_{i_1} = \xi_i + t_i q_i' \tag{1.41}$$

Soit q' le ppcm des q_i', et soit ξ défini par $\xi \equiv \xi_i \pmod{q'}$ pour tout i. S'il n'existe pas de tel ξ, l'équation n'a pas de solution plus grande que X_2'.

Quitte à changer χ_i, on peut supposer que (1.40–1.41) restent vraies, avec ξ_i remplacé par ξ dans (1.41).

Par suite, $t_1 q_1' = \ldots = t_r q_r'$; on définit alors t par $t_j q_j' = tq'$. Alors $b_i = \chi_i + tq' p_i'/q_i'$, et

$$\beta_0 \beta_1^{b_1} \ldots \beta_r^{b_r} = H_0 H_1^t, \tag{1.42}$$

où $H_0 = \beta_0 \beta_1^{\chi_1} \ldots \beta_r^{\chi_r}$ et $H_1 = \beta_1^{q'p_1'/q_1'} \ldots \beta_r^{q'p_r'/q_r'}$.

Il reste alors à distinguer trois cas :

- $|H_1| \neq 1$, $|H_0| \neq 1$.
  On a alors la minoration $|t \log|H_1| + \log|H_0|| \geq \log|H_1| \| \log|H_0|/\log|H_1| \|$, où $\|\cdot\|$ est la distance à $\mathbb{Z}$.
- $|H_1| \neq 1$, $|H_0| = 1$.
  Si $t \neq 0$, on a alors la minoration $|t \log|H_1|| \geq \log|H_1|$.
- $|H_1| = |H_0| = 1$. On s'intéresse dans ce cas à l'argument du produit (1.42). Il existe un entier t' tel que $\operatorname{Arg}(H_0 H_1^t) = \operatorname{Arg}(H_0) + t\operatorname{Arg}(H_1) + 2\pi t'$. Pour minorer ceci, on utilise les méthodes de la section 1.4.2.3.
  On échoue encore si $\operatorname{Arg} H_0/(2\pi)$ et $\operatorname{Arg} H_1/(2\pi)$ sont rationnels ; mais dans ce cas H_0 et H_1 sont des racines de l'unité, donc ± 1, et ceci peut être vérifié de manière exacte dans le compositum $\sigma_2(\mathbb{K})\sigma_3(\mathbb{K})$. On a donc $H_0 H_1^t = \pm(\pm 1)^t$. Cependant, ce produit ne peut pas être égal à 1, et on a donc

  $$2 = |H_0 H_1^t - 1| \leq \frac{c_4 b_0}{|x|^n},$$

  qui donne immédiatement une (petite) borne sur $|x|$.

Ceci conclut la description du processus de réduction.

1.5 L'énumération finale

On dispose maintenant d'une nouvelle borne pour B de taille raisonnable. De nombreuses possibilités s'offrent alors pour l'énumération finale. Cette section a pour ambition, sans être exhaustive, de présenter la majorité des méthodes utilisées à ce jour. On décrit d'abord l'énumération des $\boldsymbol{b}$, puis ensuite l'énumération des x.

1.5.1 Énumérer les $\boldsymbol{b}$

La première possibilité venant à l'esprit est l'énumération brutale, qui donne $\mathcal{B}^* \cdot (2B^* + 1)^r$ possibilités à tester ; ceci est totalement déraisonnable dès que r dépasse 2 ou 3.

1.5.1.1 Énumération systématique

Il est bien plus efficace de ré-utiliser la proposition 1.34 de la manière suivante :

Proposition 1.41. *Soit* $1/2 > \varepsilon > 0$ *et*

$$X_3 := \max\left(X_2, \left((1 + \max_{i \neq i_1} |\delta_i|)c_{14}\mathcal{B}^*\varepsilon^{-1}\right)^{1/n}\right).$$

Alors soit $|x| \leq X_3$*, soit*

$$\left|b_{i_2} - \overline{\delta}_{i_2} b_{i_1} - \overline{\lambda}_{i_2} b_0\right| \leq \varepsilon.$$

En particulier,

$$\left\|\overline{\delta}_{i_2} b_{i_1} - \overline{\lambda}_{i_2} b_0\right\| \leq \varepsilon, \tag{1.43}$$

où $\|\cdot\|$ *représente la distance à* $\mathbb{Z}$*, et* b_{i_2} *est parfaitement déterminé par le choix de* (b_0, b_{i_1})*.*

Il suffit ainsi d'énumérer tous les (b_0, b_{i_1}) vérifiant $b_0 \leq \mathcal{B}^*, |b_i| \leq B^*$; pour chacun d'entre eux, dès qu'il existe un i_2 tel que le critère (1.43) ne soit pas vérifié, on peut passer au couple suivant. Sinon, on calcule le $(r+1)-$uplet $(b_0, \ldots, b_r)$ comme dans la section précédente, et l'on vérifie s'il donne naissance à une solution de la manière indiquée plus loin.

Ce procédé met en jeu une énumération de $\mathcal{B}^*(2B^* + 1)$ termes ; il doit donc être écarté lorsque la réduction n'a pas permis de descendre en deçà de quelques centaines (c'est-à-dire essentiellement quand on n'a pu réduire la borne provenant des minorations de régulateurs).

1.5.1.2 Petits vecteurs de Λ

La manière en un sens la plus "logique" et la plus économique d'effectuer cette énumération (encore que ce qui précède puisse être plus rapide quand la borne finale est très petite) consiste à réutiliser les idées de la réduction de la borne ; on cherche toujours des vecteurs très courts d'un réseau ; mais l'algorithme de Fincke-Pohst sait trouver tous les vecteurs dont la norme n'excède pas une quantité donnée. On se fixe donc une borne l, et on cherche tous les vecteurs du réseau engendré par les colonnes de

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \lfloor -C\overline{\lambda}_{i_2} \rceil & \lfloor -C\overline{\delta}_{i_2} \rceil & C \end{pmatrix}$$

dont la norme est plus petite que l. À l'aide du triplet (b_0, b_{i_1}, b_{i_2}) correspondant, on reconstitue le $(r+1)$-uplet b_i, et l'on vérifie alors s'il donne naissance à une solution, comme décrit plus loin.

Reste à montrer que les vecteurs de "grande norme" du réseau ci-dessus ne peuvent donner naissance qu'à de "petites" solutions. Ce résultat est "dual" des Théorèmes 1.25 et 1.33. On l'énonce dans le cas du réseau donné ci-dessus.

Proposition 1.42. *Soit $(b_0, \ldots, b_r)$ un $(r+1)-$uplet correspondant à une solution (x, y) ; on suppose que le vecteur du réseau correspondant est de norme plus grande que l, où $l > \sqrt{5\mathcal{B}^{*2} + 2B^*\mathcal{B}^* + 5B^{*2}}/2$. Alors on a*

$$|x| \leq \left(\frac{\mathcal{B}^* c_{15}(1 + \max_i |\overline{\delta}_i|)C}{\sqrt{l^2 - \mathcal{B}^{*2} - B^{*2}} - (\mathcal{B}^* + B^*)/2} \right)^{1/n}.$$

Preuve. Il suffit de combiner les arguments de la preuve du Théorème 1.25, du Corollaire 1.26 en utilisant la majoration donnée par la Proposition 1.34 au lieu de (1.32). □

Remarque 1.43. Notons que la remarque ci-dessus peut très fréquemment être utilisée avec profit en prenant pour l la borne inférieure obtenue lors de la dernière étape de réduction ; la borne sur $|x|$ est en pratique très petite, tout particulièrement quand le degré est grand, cela même quand celle obtenue pour B est encore grande (ce qui n'arrive déjà en pratique que quand la borne $\mathcal{B}$ est mauvaise) ; voir le cas $p = 83$ pour l'équation cyclotomique réelle, ou la borne finale sur B est de l'ordre de 10^{25}, mais la borne sur $|x|$ est de l'ordre de 50.

La bonne qualité de la borne obtenue pour $|x|$ provient entre autres du fait que la borne supérieure pour la forme linéaire en terme de $|x|$ est bien plus précise que la borne en terme de B.

Cette idée de tirer de la réduction une borne pour $|x|$ plutôt que pour B se trouve dans [We87], mais semble être un peu tombée dans l'oubli ultérieurement. Pourtant, dès que le degré est plus grand que 4, c'est de loin la méthode la plus efficace, et il est rarement utile d'avoir recours à Fincke-Pohst.

Notons que l'on peut, comme dans la section précédente, varier à l'infini (ou presque) le choix du réseau sur lequel on travaille. Vu la complexité de l'algorithme de Fincke-Pohst, il semble toutefois préférable de se limiter à des réseaux de petite dimension.

1.5.1.3 Cribler les b

Une autre possibilité, exploitée dans [TW92] et [Sm95], qui concerne le cas $b_0 = 1$, consiste à fabriquer des relations de congruence de la manière suivante.

On choisit un nombre premier q ayant au moins trois idéaux premiers de $\mathbb{K}$ distincts $\mathfrak{q}_i$, $i = 1, 2, 3$, de degrés résiduels 1 au-dessus de lui. L'existence d'une infinité de tels premiers est garantie par le théorème de Čebotarev. On choisit alors des entiers m_i tels que $\alpha \equiv m_i \bmod \mathfrak{q}_i$, et des entiers $M_i, E_i^{(1)}, \ldots, E_i^{(r)}$ tels que $\mu \equiv M_i \bmod \mathfrak{q}_i$, $\eta_k \equiv E_i^{(k)} \bmod \mathfrak{q}_i$. On a alors

$$y - m_i x \equiv M_i {E_i^{(1)}}^{b_1} \ldots {E_i^{(r)}}^{b_r} \bmod \mathfrak{q}_i.$$

Tous les termes étant entiers, cette équation a en fait lieu modulo q. On peut alors éliminer x et y des trois équations ci-dessus, ce qui va imposer des conditions de congruence sur les b_i permettant d'éliminer une partie des r-uplets ne correspondant pas à des solutions. Une fois que l'on a criblé modulo suffisamment de nombres premiers, on effectue une recherche exhaustive parmi les r-uplets restants.

1.5.2 Des b_i à x

Il faut encore passer des $(r+1)$-uplets b_i à (x, y). On peut par exemple utiliser la proposition suivante :

Proposition 1.44. *Soit* $|x| > \max\left(X_3, (2c_1)^{1/(n-1)}, (2c_1/c_2)^{1/(n-1)}\right)$ *correspondant au* $(r+1)$*-uplet* $(b_0, \ldots, b_r)$*. Alors on a*

$$|x| = \left\lfloor \frac{\left(|\mu^{(i)}| \prod_{1 \le i \le k} |\eta_k^{(i)}|^{b_k}\right)^{1/b_0}}{|\alpha^{(i)} - \alpha^{(1)}|} \right\rceil, y = \lfloor \alpha^{(1)} x \rceil.$$

Preuve. On a

$$||\alpha^{(i)} - \alpha^{(1)}||x| - |y - \alpha^{(i)} x|| \le |y - \alpha^{(1)} x| \le \frac{c_1}{|x|^{n-1}}$$

donc

$$\left||x| - \left|\frac{y-\alpha^{(i)}x}{\alpha^{(i)}-\alpha^{(1)}}\right|\right| \leq \frac{|y-\alpha^{(1)}x|}{|\alpha^{(i)}-\alpha^{(1)}|} \leq \frac{c_1}{c_2|x|^{n-1}},$$

et l'hypothèse sur x garantit alors le résultat, sachant que $y-\alpha^{(i)}x$ est donné par (1.12). Le deuxième point est assuré dès que $c_1/|x|^{n-1} < 1/2$. □

1.5.3 Énumérer x

Au cours de l'algorithme, les "petites" valeurs de x ont souvent été laissées de côté ; quelques indications suivent sur la manière de vérifier si un x donné peut correspondre à une solution.

1.5.3.1 La borne pour $|x|$ est grande

On a alors en général recours à une remarque faite au tout début de ce chapitre :

Proposition 1.45. *Soit (x,y) une solution de (1.1) ; on suppose que l'on a $|x| \geq (2c_1)^{1/(n-2)}$. Alors y/x est une réduite du développement en fractions continues de α.*

Preuve. On a, par la proposition (1.2) *(i)*,

$$\left|\frac{y}{x} - \alpha\right| \leq \frac{c_1}{|x|^n} \leq \frac{1}{2x^2},$$

la dernière inégalité étant valide du fait du choix de x. Il est alors bien connu que y/x est une réduite du développement en fractions continues de α. □

Ceci permet en général de se ramener à une borne de taille très raisonnable pour x.

Remarque 1.46. Il convient de noter que y/x et $(-y)/(-x)$ doivent tous deux être considérées comme des réduites du développement en fractions continues.

1.5.3.2 La borne pour $|x|$ est petite

Soit X^* la borne en question. Étant donné un x tel que $|x| \leq X^*$, on peut obtenir quelques contraintes sur les y tels que (x,y) soit solution :

Lemme 1.47. *Soit (x,y) solution de (1.1), avec $x \neq 0$. Alors*

(i)

$$\left|\frac{y}{x}\right| \leq |a/f_0|^{1/n} + \max_i |\alpha^{(i)}|$$

et

(ii)

$$y \mid (a - f_n x^n).$$

Preuve. Pour *(i)*, remarquons que

$$\prod_{1 \le i \le n} \left| \frac{y}{x} - \alpha^{(i)} \right| \le \left| \frac{a}{f_0} \right|$$

donc il existe un i tel que

$$\left| \left| \frac{y}{x} \right| - |\alpha^{(i)}| \right| \le \left| \frac{y}{x} - \alpha^{(i)} \right| \le \left| \frac{a}{f_0} \right|^{1/n},$$

ce qui conduit facilement à la conclusion ; pour *(ii)*, la relation

$$y(f_0 y^{n-1} + f_1 y^{n-2} x + \ldots + f_{n-1} x^{n-1}) = a - f_n x^n$$

donne le résultat. Si $a - f_n x^n = 0$, le deuxième point est vide, mais $y \mid f_{n-1} x^{n-1}$, etc. □

Ceci nous fournit deux contraintes sur y que l'on peut exploiter soit séparément, soit simultanément (en n'énumérant que les petits diviseurs de $a - f_n x^n$).

1.6 L'algorithme

On résume ici sous forme d'algorithme tout ce qui précède. Toutes les variantes décrites ne sont bien sûr pas présentes ici. On suppose que l'on dispose initialement des valeurs approchées des conjugués de α (avec une précision arbitraire), un système d'unités de rang maximal de $\mathbb{K}$, avec un majorant $\mathcal{B}$ de son indice dans le groupe des unités et de l'ensemble M_{a/f_0}. L'algorithme tient compte du fait que seule une partie des calculs dépend effectivement du second membre de l'équation ; ainsi, il n'est pas besoin de recalculer les unités ou la matrice A si seul le second membre change.

1. Calculer la matrice A et son inverse, avec une très grande précision.
2. Calculer X_0, X_1, $c_1 - c_4$.
3. $i_0 \leftarrow 1$.
4. Calculer les $\overline{\delta}_i$ avec une très grande précision (que l'on peut estimer en calculant le facteur constant de la borne de Baker), ainsi que les constantes c_5, c_7, c_9, c_{10}, c_{13}, c_{14}, X_2.
5. Pour chaque $\mu \in M_{a/f_0}$, calculer les $\overline{\lambda}_i$, c_6, c_8, c_{11}, c_{12}, c_{15}.
6. Réduire la borne B. Si la précision de $\overline{\delta}_i$ est insuffisante, augmenter la précision, recalculer $\overline{\delta}_i$ et aller en 3.

7. Essayer de réduire la borne pour B. Si cela réussit, retourner en 6.
8. Calculer la borne supérieure X_4 pour $|x|$ donnée par (1.42), avec $l = l_0$. Si $X_4 \leq X_3$, aller en 10.
9. Fixer un x de taille raisonnable, et calculer la longueur l correspondante. Utiliser Fincke-Pohst pour détecter les solutions pour lesquelles la norme du vecteur est plus grande que l.
10. $i_0 \leftarrow i_0 + 1$; si $i_0 \leq s$ aller en 4.
11. Vérifier les solutions pour tous les $|x| < X_4$.

1.7 Un détail numérique

Concluons ce chapitre par une justification ayant trait à la précision des calculs. Cette discussion a été repoussée jusqu'à maintenant pour ne pas rompre la continuité de la description de la méthode.

Seuls les $\overline{\delta}_i$ et les $\overline{\lambda}_i$ ont réellement besoin d'être connus avec une très grande précision. Mais leur calcul comporte une inversion de matrice, dont on sait bien qu'elle peut être extrêmement instable.

Le lemme suivant permet de contrôler l'erreur commise ; étant donnée une matrice $A = [a_{ij}]$, on note $\|A\|_\infty = \max_{i,j} |a_{ij}|$.

Lemme 1.48. *Soit R, $\tilde{R}$, A, $\tilde{A}$ des matrices $r \times r$ à coefficients réels et ε_1, ε_2 des nombres réels positifs, vérifiant les hypothèses suivantes*

$$AR = I\,, \tag{1.44}$$
$$\|R - \tilde{R}\|_\infty \leq \varepsilon_1\,, \tag{1.45}$$
$$\|\tilde{A}\tilde{R} - I\|_\infty \leq \varepsilon_2\,, \tag{1.46}$$
$$r\|\tilde{A}\|_\infty\varepsilon_1 + \varepsilon_2 \leq \frac{1}{2r}\,, \tag{1.47}$$

où I est la matrice identité d'ordre r. Alors

$$\|A - \tilde{A}\|_\infty \leq \varepsilon_3\,, \tag{1.48}$$

où $\varepsilon_3 = 2r^2\|\tilde{A}\|_\infty \Big(r\|\tilde{A}\|_\infty\varepsilon_1 + \varepsilon_2\Big)$.

Preuve. En combinant (1.45) et (1.46), on obtient

$$\|\tilde{A}R - I\|_\infty \leq \varepsilon_4\,, \tag{1.49}$$

où ε_4 est le membre gauche de (1.47). On écrit $\tilde{A}R = I + E$ avec $\|E\|_\infty \leq \varepsilon_4$. Par récurrence il est facile de prouver que $\|E^\nu\|_\infty \leq r^{\nu-1}\varepsilon_4^\nu$, où $\nu = 1, 2, \ldots$. Ainsi

$$\|(I+E)^{-1}\|_\infty \leq 1 + \sum_{\nu=1}^{\infty} \|E^\nu\|_\infty \leq 1 + \frac{\varepsilon_4}{1 - r\varepsilon_4} \leq 1 + 2\varepsilon_4 \leq 2\,.$$

Par suite,

$$\|A\|_\infty = \|(I+E)^{-1}\tilde{A}\|_\infty \leq r\|(I+E)^{-1}\|_\infty\|\tilde{A}\|_\infty \leq 2r\|\tilde{A}\|_\infty\,.$$

Enfin, on a donc

$$\|A-\tilde{A}\|_\infty = \|EA\|_\infty \leq r\|E\|_\infty\|A\|_\infty \leq 2r^2\|\tilde{A}\|_\infty\varepsilon_4 = \varepsilon_3\,,$$

ce qui conclut la preuve. □

Divers exemples sont exposés au chapitre 3, avec les données numériques correspondantes.

Chapitre 2

L'équation de Thue : cas d'un corps composé

Ce chapitre expose une amélioration de la méthode du chapitre précédent quand le corps $\mathbb{L}$ engendré par une racine α de $P(1,Y)$ est composé ; plus précisément, il faut que le corps $\mathbb{L}$ ait un sous-corps $\mathbb{K}$ de degré au moins 3.

Ce chapitre est largement inspiré du travail [BH97], auquel est incorporée la modification permettant d'utiliser un système d'unités qui soit seulement maximal. Celle-ci n'est pas dépourvue d'intérêt dans ce contexte, contrairement à ce que l'on pourrait croire au premier abord. Voir l'exemple $N = 4001$ du chapitre suivant.

2.1 Préliminaires

2.1.1 Notations

On considère de nouveau l'équation

$$P(X,Y) = f_0Y^n + f_1XY^{n-1} + \ldots + f_nX^n = a, \tag{2.1}$$

où P est une forme irréductible de degré au moins égal à 3, et a un nombre rationnel, et on pose encore $g(Y) = P(1,Y)$. Soit α une racine de g et le corps $\mathbb{L} = \mathbb{Q}(\alpha)$. On suppose que $\mathbb{L}$ admet un sous-corps $\mathbb{K}$ de degré au moins égal à 3. Soit

$$n = [\mathbb{L} : \mathbb{Q}], \quad m = [\mathbb{K} : \mathbb{Q}], \quad l = [\mathbb{L} : \mathbb{K}].$$

Soit σ_i, $i = 1, \ldots, m$ les plongements de $\mathbb{K}$ dans $\mathbb{C}$ et τ_{ik}, $k = 1, \ldots, l$ les plongements de $\mathbb{L}$ dans $\mathbb{C}$ au-dessus de σ_i.

On notera plutôt $\alpha^{(ik)}$ le nombre $\tau_{ik}(\alpha)$. Écrivons $m = s+2t$, où s est le nombre de plongements réels et t la moitié du nombre de plongements imaginaires ; r sera le rang du groupe $\mathcal{U}_{\mathbb{K}}$ du groupe des unités de $\mathbb{K}$.

Quitte à réordonner les σ_i, on peut toujours supposer que $\sigma_1, \ldots, \sigma_s$ sont réels et que $\sigma_{s+t+i} = \overline{\sigma_{s+i}}$.

2.1.2 Prérequis algorithmiques

De la même façon que précédemment, il nous faudra savoir trouver un système d'unités de rang maximal, et savoir résoudre l'équation aux normes. L'importante amélioration viendra du fait que ces problèmes devront être résolus dans $\mathbb{K}$, et non plus dans $\mathbb{L}$:

($\mathrm{U}^*_{\mathbb{K}}$) Trouver un système d'unités de rang maximal du corps $\mathbb{K}$.

($\mathrm{N}_{\mathbb{K}}$) Trouver un ensemble maximal M_{a/f_0} de solutions z non associées de l'équation

$$N_{\mathbb{K}/\mathbb{Q}}(z) = \frac{a}{f_0}$$

dans l'idéal fractionnaire $I = (1, \alpha)$.

L'idée fondamentale de tout ce qui suit est que l'équation de Thue

$$N_{\mathbb{L}/\mathbb{Q}}(y - \alpha x) = \frac{a}{f_0}$$

peut en fait s'écrire

$$N_{\mathbb{K}/\mathbb{Q}}\left(N_{\mathbb{L}/\mathbb{K}}(y - \alpha x)\right) = \frac{a}{f_0}.$$

Définissant $\varphi = N_{\mathbb{L}/\mathbb{K}}(y - \alpha x)$, on voit que φ est un élément de $\mathbb{K}$ dont la norme est connue ; il est donc possible d'écrire φ en terme des unités de $\mathbb{K}$; il est alors seulement nécessaire de disposer de ces dernières.

2.2 Réduction aux formes linéaires en logarithmes

Soit (x, y) une solution de (2.1). Si $s = 0$, on conclut en appliquant (1.5). Dans la suite, on suppose donc que $s \geq 1$.

2.2.1 L'approximation de $\varphi^{(i)}$.

Définissons

$$\varphi^{(i)} = \prod_{1 \leq k \leq l} (y - \alpha^{(ik)} x), \qquad (1 \leq i \leq m). \tag{2.2}$$

On a la proposition suivante, exact analogue de la Proposition 1.2 :

Proposition 2.1. *Soit*

$$X_0 = \begin{cases} \left(\dfrac{2^{n-1}\cdot|a|}{\min_{\alpha^{(ik)}\notin\mathbb{R}}|g'(\alpha^{(ik)})|\cdot\min_{\alpha^{(ik)}\notin\mathbb{R}}|\mathrm{Im}\ \alpha^{(ik)}|}\right)^{1/n} & si\ \ t\geq 1\,, \\ 1 & si\ \ t=0\,, \end{cases}$$

$$\begin{aligned} c_1 &= \frac{2^{n-1}\cdot|a|}{\min_{\alpha^{(ik)}\in\mathbb{R}}|g'(\alpha^{(ik)})|}\,, \qquad c_2 = \min_{(i,k)\neq(i',k')}\left|\alpha^{(ik)}-\alpha^{(i'k')}\right|\,, \\ c_3 &= 1,39c_1c_2^{-1}\,, \qquad\qquad c_4 = lc_3, \quad c_5=(m-1)c_4, \\ X_1 &= \max\left(X_0,\ \left(2c_1c_2^{-1}\right)^{1/n}\right)\,. \end{aligned}$$

Soit (x,y) une solution entière de (1.1).

(i) Si $|x|>X_0$ alors, pour un $\alpha^{(i_0k_0)}\in\mathbb{R}$ on a

$$\left|\frac{y}{x}-\alpha^{(i_0k_0)}\right|\leq\frac{c_1}{|x|^n}\,. \tag{2.3}$$

Posons alors

$$\psi^{(i)}=\begin{cases} \displaystyle\prod_{1\leq k\leq l}(\alpha^{(ik)}-\alpha^{(i_0k_0)}), & i\neq i_0, \\ \displaystyle a\left(\prod_{i\neq i_0}\psi^{(i)}\right)^{-1}, & i=i_0. \end{cases} \tag{2.4}$$

(ii) Si $|x|>X_1$, alors

$$\left|\mathrm{Log}\,\frac{\varphi^{(i)}}{\psi^{(i)}x^l}\right|\leq\frac{c_4}{|x|^n}\quad(i\neq i_0), \tag{2.5}$$

(iii) Si $|x|>X_1$, alors

$$\left|\mathrm{Log}\,\frac{\varphi^{(i_0)}}{\psi^{(i_0)}x^{(1-m)l}}\right|\leq\frac{c_5}{|x|^n}. \tag{2.6}$$

Preuve. *(i)* est identique au premier point de la Proposition 1.2. Pour *(ii)* et *(iii)*, on applique les points *(ii)* et *(iii)* de la Proposition 1.2, et on somme sur les k concernés. □

On a donc obtenu une approximation de φ_i de la forme souhaitée, c'est-à-dire $\gamma_i x^{\rho_i}$.

Toutes les remarques faites au chapitre précédent se transposent telles quelles dans ce contexte ; ainsi quand $|x|$ est assez grand, y/x est une réduite du développement en fractions continues de $\alpha^{(i_0k_0)}$, etc.

Pour pouvoir ultérieurement appliquer la borne de Baker, il nous faut garantir que l'on peut construire une forme linéaire de logarithmes non nulle, c'est-à-dire que les $\varphi^{(i)}/\psi^{(i)}$ ne sont pas tous égaux.

Proposition 2.2. *Les nombres*

$$\varphi^{(i)}/\psi^{(i)}, \quad (i \neq i_0),$$

ne sont pas tous égaux.

Preuve. On suppose que tous les $\varphi^{(i)}/\psi^{(i)}$ sont égaux. Définissons

$$P_i(X) = \prod_{1 \leq k \leq l} (X - \alpha^{(ik)}).$$

Alors l'hypothèse s'écrit :

$$\frac{P_i(\theta)}{P_i(\alpha^{(i_0 k_0)})} = \frac{P_{i'}(\theta)}{P_{i'}(\alpha^{(i_0 k_0)})} \quad (i, i' \neq i_0), \tag{2.7}$$

où $\theta = y/x$.

Notons $\mathbb{K}^{\text{gal}}$ la clôture normale de $\mathbb{K}$. On regarde maintenant l'action de $G = \text{Gal}\left(\mathbb{K}^{\text{gal}}\left(\alpha^{(i_0 k)}\right)/\mathbb{K}^{\text{gal}}\right)$ sur l'identité (2.7). Ce groupe opère transitivement sur l'ensemble des $\alpha^{(i_0 k)}$, où k décrit $\{1, \ldots, m\}$. Par construction, les P_i sont à coefficients dans $\mathbb{K}^{\text{gal}}$, et sont donc fixés par G, de même que θ qui est rationnel. On a donc :

$$\frac{P_i(\theta)}{P_i(\alpha^{(i_0 k)})} = \frac{P_{i'}(\theta)}{P_{i'}(\alpha^{(i_0 k)})} \quad (i, i' \neq i_0, k \in \{1, \ldots, l\}). \tag{2.8}$$

Posons alors $\phi = P_i(\alpha^{(i_0 k)})/P_{i'}(\alpha^{(i_0 k)})$. Le polynôme $P_i(T) - \phi P_{i'}(T)$ est de degré au plus l et a les $l+1$ racines distinctes $\theta, \alpha^{(i_0 1)}, \ldots, \alpha^{(i_0 l)}$, il est donc identiquement nul. Comme son coefficient dominant est $1 - \phi$, on a donc $P_i = P_{i'}$, ce qui est impossible. □

2.2.2 De x aux b_i

On introduit maintenant les b_i comme coefficients de la décomposition de l'unité provenant de φ sur un système η_i, puis l'on borne $\max_i |b_i|$.

2.2.2.1 Une unité

Soit $\eta_1, \ldots, \eta_r$ une solution du problème ($\text{U}^*_{\mathbb{K}}$), à savoir un système d'unités de rang maximal du corps $\mathbb{K}$, et M_{a/f_0} un ensemble maximal de solutions non-associées de l'équation $N_{\mathbb{K}/\mathbb{Q}}(\theta) = a$ (une solution de ($\text{N}_{\mathbb{K}}$).)

On notera $\mathcal{B}$ le majorant obtenu pour l'indice $[\mathcal{U}_{\mathbb{K}} :< \eta_1, \ldots, \eta_r >]$ par les méthodes de la section 1.3.2.1

Comme $N_{\mathbb{K}/\mathbb{Q}}(\varphi) = a/f_0$, il existe un $(r+1)$-uplet d'entiers $(b_0, b_1, \ldots, b_r)$ et un élément μ de $\pm M_{a/f_0}$ tel que :

$$\varphi^{b_0} = \pm\mu^{b_0}\eta_1^{b_1}\dots\eta_r^{b_r}. \tag{2.9}$$

Définissons

$$\rho_i = \begin{cases} l, & i \neq i_0, \\ (1-m)l, & i = i_0. \end{cases}$$

Notons alors $A = [a_{ij}]_{1\le i,j\le r}$ l'inverse de la matrice $[\log|\eta_j^{(i)}|]_{1\le i,j\le r}$; de (2.9) et de la Proposition 2.1, *(ii)* et *(iii)*, on tire

$$b_i = \sum_{1\le j\le r} a_{ij}\log|\varphi^{(j)}/\mu^{(j)}| = b_0\delta_i\log|x| + b_0\lambda_i + b_0\epsilon_i,$$

avec

$$\delta_i = \sum_{j=1}^{r} a_{ij}\rho_j, \quad \lambda_i = \sum_{j=1}^{r} a_{ij}\log|\psi^{(j)}/\mu^{(j)}|, \quad |\epsilon_i| \le \max_{1\le i\le r}\left(c_5\sum_{j=1}^{r}|a_{ij}|\right) =: c_6.$$

On peut déduire de ces identités la proposition suivante :

Proposition 2.3. *Supposons que* $|x| \ge X_2 := \max\left(X_1, (2\cdot 10^{10}c_6)^{1/n}\right)$. *Alors*

$$B := \max(|b_1|,\dots,|b_r|) \le b_0(c_7\log|x| + c_8),$$

avec $c_7 = \max_{1\le i\le r}|\delta_i|$ *et* $c_8 = \max_{1\le i\le r}|\lambda_i| + 10^{-10}$.

2.2.2.2 Une borne pour $\max_i |b_i|$

D'après la Proposition 2.2, on peut trouver deux entiers i_1 et i_2 de $\{2,\dots,m\}$ tels que :

$$\frac{\psi^{(i_2)}\varphi^{(i_1)}}{\psi^{(i_1)}\varphi^{(i_2)}} \neq 1.$$

On a par ailleurs, en utilisant la Proposition 2.1, (ii) :

$$\left|\operatorname{Log}\frac{\psi^{(i_2)}\varphi^{(i_1)}}{\psi^{(i_1)}\varphi^{(i_2)}}\right| \le \frac{2c_4}{|x|^n},$$

soit encore

$$\left|\operatorname{Log}\left(\frac{\psi^{(i_2)}\varphi^{(i_1)}}{\psi^{(i_1)}\varphi^{(i_2)}}\right)^{b_0}\right| \le \frac{2b_0c_4}{|x|^n}.$$

Mais, en posant

$$\beta_0 = \frac{\psi^{(i_2)}}{\psi^{(i_1)}}\cdot\frac{\mu^{(i_1)}}{\mu^{(i_2)}}, \quad \beta_j = \frac{\eta_j^{(i_1)}}{\eta_j^{(i_2)}}, \quad (1\le j\le r),$$

on a alors

$$|b_0\mathrm{Log}\,\beta_0 + b_1\mathrm{Log}\,\beta_1 + \ldots + b_r\mathrm{Log}\,\beta_r + b_{r+1}\cdot i\pi| \leq \frac{2b_0c_4}{|x|^n}, \tag{2.10}$$

où b_{r+1} est un entier.

Comme au premier chapitre, en comparant les parties imaginaires, on obtient

$$|b_{r+1}| \leq b_0 + |b_1| + \ldots + |b_r| + \frac{2b_0c_4}{\pi|x|^n} \leq b_0(1,23 + rB)$$

Si l'on pose alors $c_{10} = rc_8 + 1,23$ et $c_{11} = rc_7$, il vient

$$B'' := \max_{0\leq i\leq r+1} |b_i| \leq b_0 \max(1, c_{11}\log|x| + c_{10})$$

Supposons que $B'' \geq \mathcal{B}$. En injectant ce qui précède dans (2.10), il vient

$$|b_0\mathrm{Log}\,\beta_0 + b_1\mathrm{Log}\,\beta_1 + \ldots + b_r\mathrm{Log}\,\beta_r + b_{r+1}\cdot i\pi| \leq b_0c_{13}\exp(-c_{12}B''/b_0), \tag{2.11}$$

avec $c_{12} = nc_{11}^{-1}$, $c_{13} = 2c_4\exp(nc_{10}/c_{11})$.

De nouveau, si l'on pose

$$\xi = \frac{\psi^{(i_2)}\varphi^{(i_1)}}{\psi^{(i_1)}\varphi^{(i_2)}},$$

il faut distinguer deux cas :

- $\xi^{b_0} \neq 1$, ce qui est le cas en particulier quand ξ n'est pas une racine de l'unité d'ordre inférieur à b_0. Dans ce cas, le Théorème 1.9 s'applique, et il existe une constante c_9 explicite avec

$$|b_0\mathrm{Log}\,\beta_0 + b_1\mathrm{Log}\,\beta_1 + \ldots b_r\mathrm{Log}\,\beta_r + b_{r+1}\cdot i\pi| \geq \exp(-c_9\log\max_{0\leq i\leq r+1}|b_i|). \tag{2.12}$$

- ξ est une racine de l'unité d'ordre inférieur ou égal à b_0. Dans ce cas, comme ξ est différent de 1 par le choix de i_1 et de i_2, on a

$$|\mathrm{Log}\,\xi| \geq \frac{2\pi}{\mathcal{B}}.$$

Rapprochant ceci de (2.11), on obtient via le Lemme 1.20 :

Théorème 2.4.

$$B \leq B'' \leq B_0 := \max\left(e, \mathcal{B}, \frac{\mathcal{B}}{c_{12}}\log\left(\frac{\mathcal{B}c_{13}}{2\pi}\right), 2\mathcal{B}(c_{12}^{-1}c_9\log(\mathcal{B}c_{12}^{-1}c_9) + c_{13})\right).$$

Remarque 2.5. La borne obtenue de cette manière est beaucoup plus petite que la borne obtenue par la méthode directe ; en effet, le facteur important dans la constante c_9 n'est pas tant le degré d du corps de nombres (qui n'a pas changé) que le nombre r de termes (qui est maintenant au plus égal à $m-1$).

Les techniques de réduction sont exactement les mêmes qu'au chapitre précédent ; le contexte est également le même. On ne redécrit pas la procédure de réduction de la borne de Baker. Dans la suite, on suppose donc connues une borne B^* pour $\max_{1\le i\le r}|b_i|$ et une borne $\mathcal{B}^*$ pour b_0. Par ailleurs, l'énumération finale, elle aussi, se fait de manière analogue à celle du chapitre précédent.

Il ne nous reste essentiellement plus qu'à décrire comment revenir de la famille b_i à x.

2.2.3 Des b_i à x

On utilise de nouveau l'approximation $|\varphi^{(i)}/\psi^{(i)}| \approx |x^l|$ pour $i \neq i_0$. On a, d'après la Proposition 2.1, *(ii)* :

$$\left|\mathrm{Log}\left|\frac{\varphi^{(i)}}{\psi^{(i)}x^l}\right|\right| \le \frac{c_4}{|x|^n},$$

d'où

$$\left|\mathrm{Log}\,\frac{\omega^{(i)}}{|x|}\right| \le \frac{c_4 b_0}{l|x|^n},$$

où $\omega^{(i)} = |\varphi^{(i)}/\psi^{(i)}|^{1/l}$. Maintenant, $|\exp(x)-1| \le 1,3|x|$ pour $|x| < 1/2$, d'où

$$\left|\omega^{(i)} - |x|\right| \le \frac{1,3c_4}{l|x|^{n-1}}.$$

En conséquence, si $|x| \ge \left(\dfrac{2,6c_4}{l}\right)^{1/(n-1)} =: X_3$, on a

$$|x| = \lfloor \omega^{(i)} \rceil.$$

De même qu'au premier chapitre, y est alors obtenu comme $\lfloor \alpha^{(i_0k_0)}x \rceil$.

Remarque 2.6. Notons que si $\varphi^{(i)}/\psi^{(i)}$ n'est déterminé qu'à une racine $2b_0^{\text{ème}}$ de l'unité près, son module est lui parfaitement connu si l'on connaît $(b_0,\dots,b_r)$.

Remarque 2.7. On peut également jouer sur le facteur $2,6$; en l'augmentant, on peut écarter les $\omega^{(i)}$ qui ne sont pas assez voisins d'un entier.

2.3 L'algorithme

On expose une version où la réduction intervient simultanément pour tous les (k_0,μ). Dans le cas $b_0 = 1$, on peut ainsi se contenter de réduire un seul réseau (ou de calculer un seul développement en fractions continues), puisque $\overline{\delta}_i$ ne dépend pas de (k_0,μ).

1. Calculer s et t par l'algorithme de Sturm. Si $s = 0$, calculer la borne (1.5), énumérer les solutions inférieures à cette borne, et fin.
2. Calculer un système d'unités de rang maximal de $\mathbb{K}$ et une borne pour l'indice de ce système.
3. Calculer les constantes c_1–c_5, X_0, X_1, la matrice $[\log|\eta_i^{(j)}]$, les hauteurs logarithmiques absolues des unités, de α et des solutions de l'équation aux normes. En déduire une constante c_9 "uniforme", en utilisant les inégalités
$$\begin{aligned} h(\beta_0) &\leq 2h(\mu) + 2l(2h(\alpha) + \log(2)) \\ h(\beta_j) &\leq 2h(\eta_j), \quad (1 \leq j \leq r). \end{aligned}$$
4. Calculer la matrice A avec une très grande précision (que l'on peut estimer à l'aide de c_9), en utilisant la section 1.7. En déduire les constantes c_6, X_2, X_3.
5. $i_0 \leftarrow 1$.
6. Calculer les $\overline{\delta}_i$, les constantes c_7, c_{11}, c_{12}, c_{15}.
7. Calculer les $\psi^{(i)}$, et en déduire les $\overline{\lambda}_i$, et les constantes c_8, c_{10}, c_{13}, c_{14}, ainsi que la borne B_0.
8. Réduire la borne B_0 simultanément pour tous les k_0, en réitérant la réduction jusqu'à ce qu'une étape supplémentaire de réduction n'améliore plus significativement la borne.
9. Tenter de réduire la borne $\mathcal{B}$. En cas de réussite, aller en 8.
10. Calculer la borne pour x donnée par (1.42), avec $l = l_0$. Si elle est plus petite que X_3, aller en 12.
11. Fixer un X_4 de taille raisonnable, et calculer la longueur l correspondante. Utiliser Fincke-Pohst pour détecter les solutions pour lesquelles la norme du vecteur est plus grande que l, ce pour tous les k_0.
12. $i_0 \leftarrow i_0 + 1$; si $i_0 \leq s$ aller en 5.
13. Énumérer les $|x| < X_4$; regrouper les solutions. Fin.

Chapitre 3

Exemples et applications

3.1 Remarques générales

La mise en œuvre de la méthode nécessite le calcul d'un certain nombre de constantes ; certaines se déduisent facilement des autres, d'autres nécessitent des calculs plus importants. Ne sont données ici que les valeurs des constantes les plus significatives.

Les constantes dépendent pour la plupart de $i_0 \in \{1, \ldots, s\}$ et de $\mu \in M$. Pour chacune de ces constantes, nous donnons ici la pire des valeurs obtenues.

Les programmes utilisés ont été implantés en langage C, en utilisant la librairie PARI, versions 1.39 et 1.915.

3.2 $x^{19} + 2y^{19} = \pm 1, \pm 2$.

Posons $\alpha = \sqrt[19]{2}$. Le corps $\mathbb{K}=\mathbb{Q}(\alpha)$ n'a aucun sous-corps non trivial ; il nous faut donc utiliser les techniques du chapitre 1.

La signature de $\mathbb{K}$ est $(1, 9)$; en particulier, on a donc $r = 9$. Le corps $\mathbb{K}$ a pour discriminant $-2^{18}19^{19}$; la base de puissances $(\alpha^k)_{0\leq k\leq 18}$ constitue une base d'entiers.

On est dans la situation où b_0 est connu ; en effet un système d'unités fondamentales est donné dans le livre de Pohst [Po93] ; voici les coefficients de ces unités sur la base de puissances :

	1	α	α^2	α^3	α^4	α^5	α^6	α^7	α^8	α^9	α^{10}	α^{11}	α^{12}	α^{13}	α^{14}	α^{15}	α^{16}	α^{17}	α^{18}
η_1	-1	-1																	
η_2	-1	1	0	0	0	0	0	0	0	0	-1								
η_3	1	-1	1	0	0	0	0	0	0	0	1	-1							
η_4	1	1	0	0	1	0	0	0	0	0	0	0	-1	0	0	0	-1		
η_5	1	-1	0	-1	0	-2	0	0	1	0	1	0	1	0	0	-1	0	-1	
η_6	-1	1	-1	0	1	-1	0	1	-1	0	1	0	-1	0	1	-1	0	1	-1
η_7	1	-2	0	2	-1	-1	1	1	-2	0	1	-1	-1	1	0	-1	0	1	-1
η_8	-1	2	1	-3	2	-1	-2	1	0	-2	2	-1	-1	1	-1	-2	1	-2	-1
η_9	1	-3	2	-1	1	0	-2	1	-1	1	0	-1	0	0	0	1	-1		

Le premier 2 se décompose comme

$$2\mathbb{Z}_{\mathbb{K}} = \mathfrak{p}^{19}$$

où $\mathfrak{p} = \alpha\mathbb{Z}_{\mathbb{K}}$. Par suite,

$$N_{\mathbb{K}/\mathbb{Q}}(\mu) = 2$$

n'a qu'une seule solution modulo les unités, par exemple $-\alpha$.

Voici les valeurs des principales constantes, convenablement arrondies :

$$\begin{array}{lll} c_1 = 14310 & c_2 = 0,341 & c_3 = 12817 \\ c_5 = 2,39 & c_6 = 1,09 & c_{12} = 277579 \\ c_{14} = 1,54 \cdot 10^9 & c_{15} = 7,976 & \widehat{X}_3 = 2 \end{array}$$

La borne de Baker vaut $B_0 = 2,32 \cdot 10^{92}$. Après la première réduction avec $\kappa = 10$ on obtient $B_0 = 29$. À l'issue de la seconde réduction, avec $\kappa = 100$ on obtient $B_0 = 4$.

Les coefficients de la matrice $[\log \left|\eta_i^{(j)}\right|]_{1\le i,j\le r}$ ont été calculées avec une erreur d'au plus 2×10^{-202}. Les coefficients de l'inverse A ont été trouvés avec une erreur d'au plus 5×10^{-200}. Ceci permet d'obtenir δ avec une précision 2×10^{-199}. Comme $\kappa \le 100$, les calculs sont corrects.

Après énumération, on a obtenu les quatre solutions suivantes $(1,-1)$, $(-1,1)$, $(0,1)$, $(0,-1)$ pour l'équation $y^{19}+2x^{19} = \pm 1$, et les deux solutions $(1,0)$, $(-1,0)$ pour l'équation $y^{19} + 2x^{19} = \pm 2$. On a donc le

Théorème 3.1. *Les seules solutions de l'équation* $y^{19} + 2x^{19} = \pm 1, \pm 2$ *sont*

$$(x,y) = (1,-1), (-1,1), (0,1), (0,-1), (1,0), (-1,0).$$

L'ensemble du calcul a nécessité 11,7 secondes (Sun SPARC 10).

3.3 $y^4 + xy^3 - 1500x^2y^2 + 23756x^3y - 81536x^4 = \pm 1$.

Posons $P(X,Y) = Y^4 + XY^3 - 1500X^2Y^2 + 23756X^3Y - 81536X^4$, et intéressons-nous aux équations de Thue $P(X,Y) = \pm 1$.

Ces équations, qui sont sans intérêt propre, fournissent un exemple où le calcul d'un système d'unités fondamentales est à peu près sans espoir par les méthodes actuelles.

Le corps correspondant intervient naturellement dans la section 3.4 comme sous-corps de $\mathbb{Q}(\cos(2\pi/4001))$.

Le calcul des unités du corps $\mathbb{K} = \mathbb{Q}(\alpha)$, où α est une racine du polynôme $x^4+x^3-1500x^2+23756x-81536$, par la méthode de Buchmann prend de l'ordre de 10 secondes ; mais la certification et le calcul des unités par la méthode de Pohst et Zassenhaus ont tous deux été interrompus au bout d'une journée entière de calcul.

Le corps $\mathbb{K}$ est totalement réel, donc $r = 3$, le discriminant $d_{\mathbb{K}}$ vaut 4001^2. Si α est une racine de $P(1,Y)$, une base d'entiers du corps $\mathbb{Q}(\alpha)$ est donnée par

$\tau_1, \tau_2, \tau_3, \tau_4$, où l'expression des τ_i sur la base de puissances est donnée par le tableau suivant :

	1	α	α^2	α^3
τ_1	1	0	0	0
τ_2	0	1	0	0
τ_3	$-2/5$	$1/2$	$1/10$	0
τ_4	$-4/25$	$-47/100$	$3/200$	$1/200$

Sur cette base d'entiers, un système d'unités de rang maximal est donné par

	η_1	η_2	η_3
τ_1	2579620139	-68133221488165211383	1305916649079360678869
τ_2	-534883224	31300841079878935930	-328312134982958131010
τ_3	44573602	-16828180003143035894	97359743058696947252
τ_4	89147204	70329602822392822O4	80531563055553911358

Le régulateur de ce système (qui est fondamental sous l'hypothèse de Riemann généralisée) est à peu près 164174,5.

J'ai utilisé la borne inférieure fournie par Kash 1.7 pour le régulateur de $\mathbb{K}$; renseignements pris, cette borne résulte d'une version nettement affaiblie du Théorème 1.15.

Ladite borne donne $R_{\mathbb{K}} \geq 44,8$, soit $b_0 \leq 3664$. Les constantes significatives sont données ci-dessous :

$$\begin{array}{lll} c_1 = 0,002 & c_2 = 8,78 & c_5 = 0,07 \\ c_6 = 0,28 & B_0 = 2,3 \cdot 10^{41} & B^* = 1783 \\ X_5 = 16 & & \end{array}$$

La méthode exposée au chapitre 1 permet alors d'en déduire le

Théorème 3.2. *L'équation* $y^4 + xy^3 - 1500x^2y^2 + 23756x^3y - 81536x^4 = \pm 1$ *a pour seules solutions* $(-1,0)$ *et* $(1,0)$.

Le temps de calcul total est de 12,3 secondes.

3.4 Diviseurs primitifs des suites de Lucas et Lehmer

Soit α et β deux nombres algébriques tels que $\alpha+\beta$ pour les suites de Lucas, $(\alpha+\beta)^2$ pour les suites de Lehmer, et α/β soient des entiers premiers entre eux et que α/β ne soit pas une racine de l'unité. Les suites de Lucas (u_n) et de Lehmer (v_n) associées à α et β sont définies par

$$u_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}, \quad v_n = \begin{cases} \dfrac{\alpha^n - \beta^n}{\alpha - \beta} & \text{si } n \text{ est impair} \\ \dfrac{\alpha^n - \beta^n}{\alpha^2 - \beta^2} & \text{si } n \text{ est pair.} \end{cases}$$

Un nombre p est dit diviseur primitif d'un terme u_n d'une suite de Lucas si $p \,|\, u_n$ et p ne divise pas $(\alpha - \beta)^2 u_2 \dots u_{n-1}$. Pour les suites de Lehmer, p doit diviser u_n, mais pas $(\alpha^2 - \beta^2)^2 u_3 \dots u_{n-1}$.

La question de l'existence d'un diviseur primitif pour les suites de Lucas et de Lehmer est abordée dans divers travaux depuis le début du siècle : Birkhoff et Vandiver [BV04] ont prouvé que le $n^{\text{ème}}$ terme d'une suite de Lucas, avec $n > 6$, avait toujours un diviseur primitif, dès lors que α et β sont entiers. Carmichael [Ca13] a montré le même résultat pour $(\alpha, \beta) \in \mathbb{R}^2$, dès que $n > 12$. Ward [Wa55] a étendu ce dernier résultat aux suites de Lehmer. Ceci conduit à faire la conjecture suivante :

Conjecture 3.3 ([Vo95]). *Pour $n > 30$, le $n^{\text{ème}}$ élément d'une suite de Lucas ou de Lehmer a toujours un diviseur primitif.*

Dans ce dernier travail sont énumérées toutes les suites dont le $n^{\text{ème}}$ terme $(5 \leq n \leq 30)$ n'a pas de diviseur primitif.

Divers résultats sont connus qui vont dans la direction de cette conjecture ; Stewart [St77] a montré que le $n^{\text{ème}}$ terme d'une suite de Lucas avait un diviseur primitif dès que $n > e^{452} 2^{67}$, et de même pour les suites de Lehmer dès que $n > e^{452} 4^{67}$. Plus récemment, Voutier [Vo97] a montré le même résultat dès que $h(\alpha/\beta) < 4$.

Nous allons voir que les méthodes développées dans les chapitres précédents peuvent permettre d'avancer dans la direction opposée, c'est-à-dire de montrer que pour n petit, le $n^{\text{ème}}$ terme d'une suite de Lucas ou de Lehmer a toujours un diviseur primitif.

Soit $\phi_n(X)$ le $n^{\text{ème}}$ polynôme cyclotomique, et $\Phi_n(X, Y) = X^n \phi_n(Y/X)$. Notons $P^+(n)$ le plus grand diviseur premier de n.

Le lemme suivant est extrait de [Vo95] ; il reprend un critère établi par Stewart dans [St77], et réduit l'existence d'un diviseur primitif pour le $n^{mboxème}$ terme d'une suite de Lucas ou de Lehmer à la résolution de certaines équations de Thue.

Lemme 3.4. *Soit $n > 4$, $n \neq 6, 12$. Alors u_n a un diviseur primitif si et seulement si $\Phi_n(\alpha, \beta) \neq \pm 1, \pm P^+(n/(n,3))$.*

On a alors

$$\Phi_n(\alpha, \beta) = \prod_{j=1,(j,n)=1}^{n} (\alpha - \exp(2j\pi/n)\beta) = \prod_{j=1,(j,n)=1}^{n/2} (\alpha^2 + \beta^2 - 2\alpha\beta \cos(2j\pi/n)).$$

Par hypothèse, $\alpha+\beta$ et $\alpha\beta$ sont entiers ; il en est donc de même de $\alpha^2+\beta^2$, et l'on voit qu'à une suite de Lucas ou de Lehmer dont le $n^{\text{ème}}$ terme est dépourvu de diviseur primitif, on peut associer une solution de l'une des quatre équations de Thue

$$\prod_{j=1,(j,n)=1}^{n/2} (Y - 2X\cos(2j\pi/n)) = \pm 1, \pm P^+(n/(n,3)).$$

Dans les sections qui suivent, on s'intéressera à cette équation pour différentes valeurs de n : puissances de nombres premiers dans l'intervalle [31,67], nombres premiers compris entre 67 et 997 vérifiant certaines conditions de congruence, et quelques cas isolés : 83, 4001, 5011.

3.4.1 Le cas $31 \leq p^\alpha \leq 67$

Les sous-corps réels maximaux des corps cyclotomiques dont l'ordre est une puissance d'un nombre premier se trouvant dans cet intervalle ont la particularité d'avoir un système d'unités fondamentales privilégiées (qui sont les normes des unités de $\mathbb{Q}(\zeta_{p^\alpha})$).

On a en effet le résultat suivant.

Proposition 3.5. *Soit p un nombre premier, avec $31 \leq p^\alpha \leq 67$.*

1. *Les $(p^\alpha-3)/2$ nombres*

$$\frac{\sin(k\pi/p^\alpha)}{\sin(\pi/p^\alpha)}, \quad k = 2, \ldots, (p^\alpha-1)/2,$$

   sont des unités indépendantes du corps $\mathbb{K} = \mathbb{Q}(\cos(2\pi/p^\alpha))$, et l'indice du sous-groupe engendré par ces unités dans le groupe des unités de $\mathbb{K}$ est égal à $h(\mathbb{K})$; en particulier si $31 \leq p^\alpha \leq 67$, ce système d'unités est fondamental.
2. *Supposons p impair, $p^\alpha \neq 3$. Posons $\alpha = 2\cos(2\pi/p)$. Alors l'ensemble $M = \{2-\alpha\}$ est un système complet de solutions non associées de l'équation $N_{\mathbb{K}/\mathbb{Q}}(\beta) = p$ dans l'idéal $\mathbb{Z}_\mathbb{K}$.*
3. *Supposons $p = 2$. Alors l'ensemble $M = \{\alpha\}$ est un système complet de solutions non associées de l'équation $N_{\mathbb{K}/\mathbb{Q}}(\beta) = p$ dans l'idéal $\mathbb{Z}_\mathbb{K}$.*

Preuve. La première partie de *1.* est classique. Voir par exemple [BS, Ch. V, §5, Th. 2]. Pour la deuxième assertion, on a $h(\mathbb{K}) = 1$ pour p^α dans l'intervalle considéré, d'après Masley [Ma78].

En ce qui concerne la deuxième partie, il est classique que p est totalement ramifié dans $\mathbb{K}$; on a $(p) = \mathfrak{p}^{\phi(p^\alpha)/2}$, et n'importe quelle solution de $N_{\mathbb{K}/\mathbb{Q}}(\beta) = p$ constitue donc un système complet de solutions non-associées de l'équation aux normes ci-dessus. □

On peut alors appliquer la méthode du chapitre 1 avec $b_0 = 1$.

Au vu des tables données en appendice, on a alors le

Théorème 3.6. *Pour* $31 \leq p^\alpha \leq 67$*, les seules solutions des équations*

$$F_{p^\alpha}(X,Y) = \pm 1, \pm P^+(n/(n,3))$$

sont $(0,\pm 1)$*,* $(\pm 1,0)$*,* $\pm(1,1)$*,* $\pm(1,-1)$*,* $\pm(-1,2)$*,* $\pm(1,2)$*,*

et en conséquence, pour p^α dans cet intervalle, le p^α-ème terme d'une suite de Lucas ou de Lehmer a toujours un diviseur primitif (en effet, les valeurs de α et β données par les solutions des équations ci-dessus sont 0 ou des racines de l'unité).

3.4.2 Le cas $67 \leq p \leq 1000$, $p = 5011$.

Le titre de cette section est quelque peu trompeur, puisque nous ne pourrons en fait pas traiter complètement les premiers compris entre 67 et 1000.

Soit p un premier compris entre 67 et 1000, ou $p = 5011$. On va appliquer les méthodes du chapitre 2 (que l'on aurait aussi bien pu appliquer à la section précédente, avec un faible gain d'efficacité, voir le cas $p = 67$).

Il nous faut donc trouver de petits sous-corps de $\mathbb{K} = \mathbb{Q}(\cos(2\pi/p))$. Mais $\mathrm{Gal}(\mathbb{K}/\mathbb{Q}) = \mathbb{Z}/\left((p-1)/2\right)\mathbb{Z}$, donc $\mathbb{K}$ a un sous-corps de degré n dès lors que $p = 1 \bmod (2n)$. En particulier, si n est un nombre impair, il faut et il suffit que $p = 1 \bmod n$.

On se limite ici à $n = 3, 4, 5$. Ceci élimine les nombres premiers 83, 107, 149, 167, 173, 179, 197, 227, 239, 263, 269, 293, 317, 347, 359, 383, 389, 419, 443, 467, 479, 503, 509, 557, 563, 587, 599, 647, 653, 659, 677, 683, 719, 743, 773, 797, 827, 839, 863, 887, 947, 983 (à savoir 42 sur 150).

Il nous faut déterminer une équation du sous-corps correspondant, c'est-à-dire le polynôme minimal d'un élément primitif de ce sous-corps.

Proposition 3.7. *Soit p un nombre premier et n un entier tel que $n \mid (p-1)$ si n impair, $2n \mid (p-1)$ si n pair.*

1. *Il existe un sous-corps de $\mathbb{Q}(\cos(2\pi/p))$ de degré n.*
2. *Soit $a \bmod p$ une racine primitive modulo p. Les nombres*

$$\xi_i = \sum_{k=1}^{(p-1)/2n} 2\cos(2a^{2nk+i}\pi/p), \quad 1 \leq i \leq n$$

sont les conjugués d'un générateur d'un tel sous-corps.

Preuve. Une racine primitive $a \bmod p$ étant choisie, l'identification entre $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ et $\mathbb{Z}/\left((p-1)/2\right)\mathbb{Z}$ se fait au moyen de

$$\begin{array}{rccl} \varphi & : & \mathbb{Z}/\left((p-1)/2\right)\mathbb{Z} & \to & \mathrm{Gal}(\mathbb{K}/\mathbb{Q}) \\ & & k & \mapsto & \sigma_k : \cos(2a\pi/p) \mapsto \cos(2a^{2k}\pi/p). \end{array}$$

Par la théorie de Galois, les sous-corps de degré n de $\mathbb{K}$ sont les sous-corps fixés par les sous-groupes d'indice n de $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$; sous les conditions de divisibilité de l'énoncé, $\{\sigma_{kn}, k \in \{1, \ldots, (p-1)/(2n)\}\}$ est l'unique sous-groupe d'indice n de $\mathbb{K}$. On vérifie alors facilement que $\sigma_{kn}(\xi_i) = \xi_i$, et ce pour tout (i, k). Reste à prouver que ξ_i engendre effectivement le sous-corps de degré n de $\mathbb{Q}(\cos(2\pi/p))$; ce sous-corps $\mathbb{K}$ est engendré par les coefficients du polynôme minimal de $\cos(2\pi/p)$ sur $\mathbb{K}$. Mais ces coefficients sont des polynômes en les ξ_i à coefficients rationnels. Donc $\mathbb{K} = \mathbb{Q}(\xi_1, \ldots, \xi_n)$. Maintenant, l'extension $\mathbb{K}/\mathbb{Q}$ est cyclique, donc l'un au moins des ξ_i engendre $\mathbb{K}$ sur $\mathbb{Q}$. Comme ils sont tous conjugués, on a $\mathbb{K} = \mathbb{Q}(\xi_i)$ pour tout i. □

La démonstration précédente fournit en même temps un procédé de construction du corps ; on calcule de bonnes approximations numériques des différents conjugués et on forme les fonctions symétriques élémentaires de ces conjugués (ou on cherche une dépendance algébrique pour l'un des conjugués en utilisant LLL), et on obtient ainsi un polynôme P candidat à définir notre extension $\mathbb{K}_0$. On calcule alors le groupe de Galois de l'extension définie par P. Si ce groupe est abélien, le théorème de Kronecker-Weber nous affirme qu'il est contenu dans une extension cyclotomique, et comme le corps est totalement réel, il est contenu dans un $\mathbb{Q}(\cos(2\pi/f))$. Il suffit alors de calculer le discriminant ; si celui-ci vaut p^{n-1}, on a $f = p$, et on a bien trouvé l'équation de $\mathbb{K}_0$.

Il ne reste alors plus qu'à mettre en œuvre la technique du chapitre 2. Les tables correspondantes se trouvent dans l'annexe A.

Notons que pour les degrés que l'on envisage d'atteindre, le calcul de c_1, de c_2 et l'évaluation du polynôme peuvent se révéler difficiles. Les lemmes suivants montrent comment contourner le problème.

Lemme 3.8. *Soit l et p deux entiers tels que $1 \leq l \leq (p-1)/2$.*

$$\Psi(l) = \prod_{k=1, k\neq l}^{(p-1)/2} |2\cos(2k\pi/p) - 2\cos(2l\pi/p)|\,.$$

Soit

$$p_0 = \left\lfloor \mathrm{Acos}\left(\frac{\sqrt{3}}{3}\right)\frac{p}{\pi} \right\rfloor.$$

Alors

$$\min_{1\leq l\leq (p-1)/2} |\Psi(l)| = \min(|\Psi(p_0)|, |\Psi(p_0+1)|),$$

et c'est $|\Psi(p_0)|$ si et seulement si

$$\sin(2p_0\pi/p)\sin(p_0\pi/p) \geq \sin(2(p_0+1)\pi/p)\sin((p_0+1)\pi/p).$$

Preuve. Remarquons d'abord que le lemme résout le problème de l'évaluation de c_1, puisque quand l décrit $[1,(N-1)/2]$, $\Psi(l)$ décrit les $g'(\alpha^{(ik)})$, avec les notations du chapitre 2.

On a

$$\Psi(l) = 2^{(p-3)} \prod_{k \neq l} \left| \sin\left(\frac{(k+l)\pi}{p}\right) \sin\left(\frac{(k-l)\pi}{p}\right) \right|.$$

En conséquence,

$$\begin{aligned} \frac{\Psi(l+1)}{\Psi(l)} &= \frac{\sin\left(\frac{2l\pi}{p}\right) \sin\left(\frac{(p-1+2l)\pi}{2p}\right) \sin\left(\frac{l\pi}{p}\right)}{\sin\left(\frac{(2l+2)\pi}{p}\right) \sin\left(\frac{(l+1)\pi}{p}\right) \sin\left(\frac{(p+1-2l)\pi}{2p}\right)} \\ &= \frac{\sin\left(\frac{2l\pi}{p}\right) \sin\left(\frac{l\pi}{p}\right)}{\sin\left(\frac{(2l+2)\pi}{p}\right) \sin\left(\frac{(l+1)\pi}{p}\right)}. \end{aligned}$$

Il suffit alors d'étudier la fonction $f(x) = \sin(x\pi/p)\sin(2x\pi/p)$, dont le maximum est atteint pour $x = \mathrm{Acos}\left(\frac{\sqrt{3}}{3}\right)\frac{p}{\pi}$, pour conclure. □

Lemme 3.9. *On a*

$$c_2 = 4\sin(\pi/p)\sin(2\pi/p).$$

Preuve. Soit k et l deux entiers distincts de $[1,(p-1)/2]$. Alors

$$|2\cos(2k\pi/p) - 2\cos(2l\pi/p)| = 4|\sin((k+l)\pi/p)\sin((k-l)\pi/p)|.$$

Comme p est impair, on ne peut avoir simultanément $k+l = p-1$ et $k-l = 1$. On a donc $c_2 \leq 4\sin(\pi/p)\sin(2\pi/p)$, atteint pour $k = (p-1)/2$, $l = (p-3)/2$.□

Lemme 3.10. *Soit* $\phi_p(x) = \dfrac{x^p - 1}{x - 1}$ *le* $p^{\text{ème}}$ *polynôme cyclotomique (p est premier). Alors, pour* $x \neq 0$,

$$F_p(x,y) = \left(\frac{2x^2}{y + \sqrt{y^2 - 4x^2}}\right)^{(p-1)/2} \phi_p\left(\frac{y + \sqrt{y^2 - 4x^2}}{2x}\right).$$

Preuve. L'identité

$$F_p(1, u + 1/u) = u^{-\frac{p-1}{2}} \phi_p(u)$$

s'inverse en

$$F_p(1, y/x) = \left(\frac{y+\sqrt{y^2-4x^2}}{2x}\right)^{-\frac{p-1}{2}} \phi_p\left(\frac{y+\sqrt{y^2-4x^2}}{2x}\right),$$

pour $x \neq 0$. On conclut en écrivant que $F_p(x, y) = x^{(p-1)/2} F_p(1, y/x)$. □

3.4.3 $p = 83$.

Le résultat de Masley utilisé plus haut pour affirmer que les unités cyclotomiques sont fondamentales n'est vrai que sous l'hypothèse de Riemann généralisée pour $67 < p^\alpha \leq 97$. Pour $p = 83$, $(p-1)/2 = 41$ est un nombre premier, donc les techniques du chapitre 2 ne peuvent s'appliquer. On utilise donc le chapitre 1, avec un b_0 ; la borne inférieure pour le régulateur de $\mathbb{Q}(\cos(2\pi/83))$ est issue de [CoFr91] et vaut 85,4. Ceci montre que l'indice des unités cyclotomiques dans le groupe des unités est au plus 3493839744905263432648 51.

$$\begin{array}{lll} c_1 = 3,4 \cdot 10^{13} & c_2 = 0,01 & c_5 = 2,8 \\ c_6 = 2,1 & B_0 = 2,4 \cdot 10^{283} & B^* = 6,3 \cdot 10^{24} \\ X_5 = 46 & & \end{array}$$

Les solutions sont les mêmes que plus haut, ce qui signifie que le 83ème terme d'une suite de Lucas et Lehmer a toujours un diviseur primitif. Le temps de calcul total est d'environ 23 minutes, dont 5 minutes pour les étapes de réduction (méthode de Bennett et de Weger, avec 2 formes linéaires en 3 variables), et 5 minutes pour l'inversion de la matrice.

3.4.4 $p = 4001$.

Dans le cas $p = 4001$, en raisonnant de la même manière que dans la section 3.4.2, on trouve que le corps $\mathbb{K}_0$ est engendré sur $\mathbb{Q}$ par une racine du polynôme $x^4 + x^3 - 1500x^2 + 23756x - 81536$. Comme vu dans la section 3.2, le calcul d'un système d'unités fondamentales de $\mathbb{K}_0$ est désespéré ; on peut alors penser à considérer le sous-corps $\mathbb{K}_0'$ de degré 5, qui est engendré par une racine de $x^5 + x^4 - 1600x^3 - 20325x^2 + 123999x + 321199$, mais le régulateur est de l'ordre de 900000. Il semble donc bien qu'on ne puisse se dispenser de l'utilisation de b_0.

La borne pour le régulateur utilisée était la même que dans la section 3.2.

Voici les valeurs des principales constantes :

$$\begin{array}{lll} c_1 = 1,8 \cdot 10^{602} & c_2 = 0,000004 & c_7 = 32 \\ c_8 = 1,6 & B_0 = 2,2 \cdot 10^{52} & B^* = 3,3 \cdot 10^8 \\ X_3 = 2 & & \end{array}$$

Le temps de calcul total a été de 11 minutes 38 secondes.

De tous les résultats de ces différentes sections, on déduit le théorème suivant, annoncé dans l'introduction.

Théorème 3.11. *Soit n un entier. On suppose que n vérifie l'une des conditions suivantes :*

- *n est une puissance de nombre premier comprise entre* 31 *et* 67*,*
- *n est un nombre premier compris entre* 67 *et* 997*, et $n \equiv 1 \bmod 3, 5$, ou 8,*
- *$n \in \{83, 4001, 5011\}$.*

Alors le $n^{\text{ème}}$ terme de toute suite de Lucas ou de Lehmer admet un diviseur primitif.

Chapitre 4

Équations superelliptiques

Ce chapitre est consacré à la description d'une méthode "systématique" de résolution d'équations diophantiennes dites superelliptiques, à savoir :

$$ay^p = f(x), \tag{4.1}$$

où a est un entier non nul, $p \geq 3$, et $f(x) \in \mathbb{Z}[x]$ un polynôme séparable à coefficients entiers de degré $n \geq 2$. On évoquera brièvement le cas $p = 2$ à la fin de ce chapitre.

Le contenu de ce chapitre reprend [BH96b] ; c'en est essentiellement une traduction, à quelques permutations de sections près, et avec quelques omissions pour éviter les redites. En particulier, tout ce chapitre résulte d'un travail commun avec Yuri Bilu, et développe une version "pratique" de la méthode décrite par ce dernier dans [Bi94].

Je n'ai pas cru bon d'inclure la modification permettant de relâcher l'hypothèse sur le système d'unités, qui se transpose de la même manière que dans les chapitres précédents. Son intérêt est de toute façon moindre que dans les chapitres 1 et 2, pour deux raisons :

- l'approximation de $\varphi(x)$ est bien moins précise (en $|x|^{-1}$ au lieu de $|x|^{-n}$), ce qui fait que la mauvaise qualité du processus de réduction de la borne a des conséquences bien plus importantes,
- il nous faudra résoudre une équation aux normes dans le corps où l'on a besoin de connaître les unités ; si l'on veut utiliser le groupe des classes à ces fins, il faudra souvent en passer par l'hypothèse de Riemann généralisée ou la certification.

4.1 Introduction

La méthode "traditionnelle" pour résoudre ces équations consiste à les transformer en un certain nombre d'équations de Thue, qui, conformément à la méthode des chapitres 1 et 2, sont ensuite elles-mêmes réduites à un certain nombre d'équations linéaires aux unités.

Malheureusement, cette méthode impose de résoudre un grand nombre d'équations de Thue, dont les coefficients peuvent être très grands. Voir par exemple [SW94].

Nous présentons ici une méthode différente ; en bref, on réduit directement l'équation superelliptique à un problème de formes linéaires en logarithmes, sans utiliser d'équations de Thue.

Toutefois, comme dans la méthode de Thue, la méthode "explose" assez vite quand le degré augmente ; si le polynôme f est irréductible, la limite de la méthode est probablement $p = 3$, $\deg f = 4$, et encore ne pourra-t-on pas traiter toutes les équations de ce type. Toutefois, si f est réductible, la situation est plus agréable.

4.2 Notations

Dans la suite du chapitre, p sera un nombre premier fixé, et $f(x) \in \mathbb{Z}[x]$ un polynome séparable fixé de degré $n \geq 2$.

Soit ζ une racine primitive $p^{\text{ème}}$ de l'unité, et posons $\mathcal{P}=\{0, \ldots, p-1\}$.

On fixe également un plongement de $\overline{\mathbb{Q}}$ dans $\mathbb{C}$, ce qui permet de parler sans ambiguïté de la valeur d'un nombre algébrique.

Les symboles $\operatorname{Log} z$ et $z^{1/p}$ désigneront les déterminations principales des deux fonctions correspondantes, c'est-à-dire que $-\pi < \operatorname{Im} \operatorname{Log} z \leq \pi$ et $-\pi/p < \operatorname{Arg} z^{1/p} \leq \pi/p$. Le symbole $z^{-1/p}$ désignera $(z^{1/p})^{-1}$. On a en particulier le lemme suivant :

Lemme 4.1. *Soit z_1 et z_2 deux nombres complexes tels que $\operatorname{Re} z_1, \operatorname{Re} z_2 > 0$; on a alors $(z_1 z_2^{\pm 1})^{1/p} = z_1^{1/p} z_2^{\pm 1/p}$.*

Preuve. Le membre droit est bien une racine $p^{\text{ème}}$ de $z_1 z_2^{\pm 1}$; par ailleurs, comme $\operatorname{Re} z_1 > 0$, $-\pi/2p \leq \operatorname{Arg} z_1^{1/p} \leq \pi/2p$, et de même pour z_2. Par suite, $-\pi/p < \operatorname{Arg} z_1^{1/p} z_2^{\pm 1/p} \leq \pi/p$, ce qui prouve l'identité. □

On notera

$$\mathrm{Sol} = \left\{ x \in \mathbb{Z} \ : \ (a^{-1} f(x))^{1/p} \in \mathbb{Z} \right\}.$$

On appellera "solutions" les éléments de Sol.

On se donne également deux racines distinctes α et β de $f(x)$. On pose

$$c_1 = \max\left(\overline{|\alpha|}, \overline{|\beta|} \right), \qquad X_1 = 3c_1,$$

où $\overline{|\alpha|}$ est le plus grand des modules des conjugués de α sur $\mathbb{Q}$; de même pour $\overline{|\beta|}$.

On supposera dans toute la suite jusqu'à la phase d'énumération des petits x que la solution x est, en valeur absolue, plus grande que X_1. Ceci permettra d'éviter un certain nombre de problèmes ; en particulier

$$x \neq 0, \alpha, \beta \quad \text{et} \quad \operatorname{Arg} \frac{x-\alpha}{x-\beta} \neq \pi, \tag{4.2}$$

propriétés que nous utiliserons sans y faire spécialement référence.

Les prérequis algorithmiques de ce chapitre sont bien plus importants que ceux du chapitre précédent ; en particulier, les problèmes suivants devront être résolus :

(DP) Savoir décomposer un idéal fractionnaire donné en produit d'idéaux premiers,

(U) déterminer le groupe des racines de l'unité et un système d'unités fondamentales,

(IP) déterminer si un idéal fractionnaire donné est principal et, si oui, en donner un générateur,

(GC) calculer le groupe des classes, construire un système de représentants des classes d'idéaux et trouver la classe d'un idéal fractionnaire donné.

La réalisation de ces problèmes d'algorithmique constitue en pratique l'étape limitante (en temps) de la méthode. Les problèmes (U) et (GC) ont été discutés au chapitre 1 ; le problème (IP) se résout simultanément ; le problème (DP), quant à lui, est de nature plus facile ; on a déjà eu recours à sa solution, due à Buchmann et Lenstra, pour le calcul du groupe des classes et des unités.

Notons que toutes ces opérations devront être effectuées dans un corps qui sera, a priori, de degré $pn(n-1)/2$ sur $\mathbb{Q}$. Toutefois, le corps en question est composé ; on ne peut que souhaiter que les algorithmes de résolution des problèmes ci-dessus viennent à en tenir compte dans un proche avenir.

4.3 Une famille de corps de nombres

Dans cette section, on construit la famille de corps de nombres du premier point de la méthode générale de réduction à une équation diophantienne exponentielle.

4.3.1 Idéaux exclusifs

Définition 4.2. *Un idéal premier* $\mathfrak{p}$ *de* $\mathbb{Q}(\alpha)$ *est dit exclusif si*

$$\operatorname{Ord}_{\mathfrak{p}}(a) > 0, \qquad \textit{ou } \operatorname{Ord}_{\mathfrak{p}}(\alpha) < 0, \qquad \textit{ou encore } \operatorname{Ord}_{\mathfrak{p}}\left(f'(\alpha)\right) > 0,$$

où la notation $\operatorname{Ord}_{\mathfrak{p}}(\gamma)$ *désigne, comme c'est l'usage, le plus grand entier* m *tel que* $\gamma \in \mathfrak{p}^m$.

Notons qu'il n'y a qu'un nombre fini d'idéaux exclusifs ; si f est unitaire, ce sont les idéaux premiers au-dessus de a et de $f'(\alpha)$.

En fait, les idéaux exclusifs sont les seuls à jouer un rôle significatif ; c'est ce que montre la proposition suivante :

Proposition 4.3. *Soit $\mathfrak{p}$ un idéal premier de $\mathbb{Q}(\alpha)$ tel que $\mathrm{Ord}_{\mathfrak{p}}(\alpha) \geq 0$. Alors pour tout x solution, on a soit*

$$0 \leq \mathrm{Ord}_{\mathfrak{p}}(x-\alpha) \leq \mathrm{Ord}_{\mathfrak{p}}(f'(\alpha)),$$

soit

$$0 \leq (\mathrm{Ord}_{\mathfrak{p}}(a) - \mathrm{Ord}_{\mathfrak{p}}(x-\alpha))_{\,(\mathrm{mod}\ p)} \leq \mathrm{Ord}_{\mathfrak{p}}(f'(\alpha)),$$

où $(b)_{\,(\mathrm{mod}\ p)}$ désigne le reste de la division euclidienne de b par p — avec la convention que ce reste est positif quel que soit le signe de b. En particulier, si $\mathfrak{p}$ est non exclusif, on a $p \,|\, \mathrm{Ord}_{\mathfrak{p}}(x-\alpha)$.

Preuve. Soit $\mathfrak{p}$ un idéal premier du corps $\mathbb{Q}(\alpha)$. Soit $\mathcal{O}_{\mathfrak{p}}$ le localisé de $\mathbb{Z}_{\mathbb{K}}$ en $\mathfrak{p}$. Définissons $f_\alpha(x) = f(x)/(x-\alpha)$. Comme $f'(\alpha)$ est le résultant du polynôme $x-\alpha$ et $f_\alpha(x)$, qui sont à coefficients dans $\mathcal{O}_{\mathfrak{p}}$, il existe $(A,B) \in \mathcal{O}_{\mathfrak{p}}^2$ tels que

$$A(x-\alpha) + Bf_\alpha(x) = f'(\alpha). \tag{4.3}$$

Si maintenant (x,y) est une solution de (4.1), deux cas sont possibles :
– $\mathrm{Ord}_{\mathfrak{p}}(x-\alpha) \leq \mathrm{Ord}_{\mathfrak{p}}(f'(\alpha))$, qui est le premier cas de la proposition,
– $\mathrm{Ord}_{\mathfrak{p}}(x-\alpha) > \mathrm{Ord}_{\mathfrak{p}}(f'(\alpha))$;
dans ce cas $\mathrm{Ord}_{\mathfrak{p}}\left(f_\alpha(x)\right) \leq \mathrm{Ord}_{\mathfrak{p}}\left(f'(\alpha)\right)$ d'après (4.3) et il vient

$$\begin{aligned} \mathrm{Ord}_{\mathfrak{p}}(x-\alpha) &= \mathrm{Ord}_{\mathfrak{p}}\left(f(x)\right) - \mathrm{Ord}_{\mathfrak{p}}\left(f_\alpha(x)\right) \\ &= \mathrm{Ord}_{\mathfrak{p}}\left(ay^p\right) - \mathrm{Ord}_{\mathfrak{p}}\left(f_\alpha(x)\right) \\ &\equiv \mathrm{Ord}_{\mathfrak{p}}\left(a\right) - \mathrm{Ord}_{\mathfrak{p}}\left(f_\alpha(x)\right) \pmod{p}, \end{aligned}$$

et le résultat s'ensuit. □

4.3.2 L'ensemble Ξ

On peut maintenant déduire de ceci que dès que x est solution, le nombre $x-\alpha$ est "connu" à une puissance $p^{\text{ème}}$ près.

Lemme 4.4. *Il existe un ensemble $\Xi(\alpha)$ fini, effectivement constructible, tel que pour tout $x \in \mathrm{Sol}$, il existe $\xi \in \Xi$ et $\lambda \in \mathbb{Q}(\alpha)$ avec*

$$x - \alpha = \xi\lambda^p. \tag{4.4}$$

Preuve. La preuve est constructive, en ce sens que nous allons explicitement décrire la construction de l'ensemble Ξ avant de montrer que l'ensemble construit convient.

Si l'on regarde la décomposition de $x-\alpha$ en facteurs premiers, la Proposition 4.3 montre qu'on peut mettre dans λ tout ce qui concerne les idéaux non-exclusifs ; reste à construire ξ à partir des unités et des idéaux exclusifs. C'est l'idée de la construction, qui se complique un peu du fait de la non principalité éventuelle de $\mathbb{Z}_{\mathbb{K}}$.

On décompose l'idéal principal (α) en $(\alpha)_0/(\alpha)_\infty$, où $(\alpha)_0$ et $(\alpha)_\infty$ sont deux idéaux entiers premiers entre eux du corps $\mathbb{Q}(\alpha)$. Soit $\mathfrak{p}_1,\dots,\mathfrak{p}_k$ les idéaux exclusifs tels que $\mathrm{Ord}_{\mathfrak{p}_i}\alpha \geq 0$. On considère les idéaux du type

$$I = I(b_1,\dots,b_k) = (\alpha)_\infty^{-1}\mathfrak{p}_1^{b_1}\cdots\mathfrak{p}_k^{b_k}, \tag{4.5}$$

où $(b_1,\dots,b_k)$ décrit l'ensemble des k-uplets d'éléments de $\mathcal{P}$ vérifiant les conditions de la proposition 4.3.

Pour chacun de ces idéaux, soit $\Delta(I)$ un ensemble maximal d'idéaux I_1 non-équivalents deux à deux du corps $\mathbb{Q}(\alpha)$ tels que II_1^p soit principal pour tout $I_1 \in \Delta$.

On rappelle que deux idéaux de $\mathbb{Q}(\alpha)$ sont dits équivalents si leur quotient est principal.

En particulier, si $h_{\mathbb{K}} = 1$, on peut prendre $\Delta(I) = \{\mathbb{Z}_{\mathbb{K}}\}$ pour tout I. Si $h_{\mathbb{K}}$ est premier à p, soit n un inverse de $-p$ mod $h_{\mathbb{K}}$, on peut prendre $\Delta(I) = \{I^n\}$.

On fixe un générateur ξ_0 pour chacun des idéaux principaux II_1^p, où I est du type (4.5) et I_1 dans $\Delta(I)$; soit Ξ_0 l'ensemble de ces ξ_0.

Soit alors ω un générateur du groupe des racines de l'unité de $\mathbb{Q}(\alpha)$ et $\eta_1,\dots,\eta_r$ un système d'unités fondamentales de $\mathbb{Q}(\alpha)$. On pose

$$\Xi = \left\{\xi_0\omega^{b_0}\eta_1^{b_1}\cdots\eta_r^{b_r}\ ;\ \xi_0 \in \Xi_0, b_0,\dots,b_r \in \mathcal{P}\right\}.$$

Montrons maintenant que cet ensemble Ξ convient ; la Proposition 4.3 montre que l'idéal $(x-\alpha)$ s'écrit II_0^p, où I est un idéal du type (4.5). Soit alors I_1 dans $\Delta(I)$ un idéal équivalent à I_0 et $\xi_0 \in \Xi_0$ un élément engendrant II_1^p. Alors $(x-\alpha) = (\xi_0)(\lambda_0)^p$, avec $\lambda_0 \in \mathbb{Q}(\alpha)$. On a donc (4.4) avec $\xi \in \Xi$ et $\lambda \in \mathbb{Q}(\alpha)$. □

Remarque 4.5. Notons que si f est irréductible et unitaire, on peut éliminer de l'ensemble Ξ tous les éléments tels que $N_{\mathbb{Q}(\alpha)/\mathbb{Q}}(\xi)/a$ ne soit pas une puissance $p^{\text{ème}}$.

4.3.3 Corps admissibles

On peut maintenant finir ladite construction. Posons $\mathbb{K}_0 = \mathbb{Q}(\alpha,\beta)$.

Définition 4.6. *Un corps de nombres $\mathbb{K}$ est admissible pour une solution x s'il existe $k \in \mathcal{P}$ tel que $\mathbb{K} = \mathbb{K}_0\left(\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}\right)$. Une famille de corps de nombres $\{\mathbb{K}\}$ est un système complet de corps admissibles s'il contient un corps admissible pour chaque x solution.*

Tous les conjugués de $\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}$ au-dessus de $\mathbb{K}_0$ sont parmi les nombres $\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}$, où $k \in \mathcal{P}$. Par suite, tout corps isomorphe à $\mathbb{K}$ au-dessus de $\mathbb{K}_0$ est admissible pour x dès que $\mathbb{K}$ est admissible pour x. C'est faux si les corps ne sont isomorphes qu'au-dessus de $\mathbb{Q}$.

Un système complet de corps admissibles sera dit minimal si deux corps distincts ne sont pas isomorphes au-dessus de $\mathbb{K}_0$.

Proposition 4.7. *Il existe un système complet fini de corps admissibles effectivement constructible.*

Preuve. D'après la théorie de Kummer, soit il existe k tel que $\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p} \in \mathbb{K}_0$, soit $\left[\mathbb{K}_0\left(\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}\right) : \mathbb{K}_0\right] = p$ pour tout $k \in \mathcal{P}$, et les p corps $\mathbb{K}_0\left(\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}\right)$ sont isomorphes au-dessus de $\mathbb{K}_0$.

Soit maintenant $\Xi(\alpha)$ l'ensemble construit au Lemme 4.4 et $\Xi(\beta)$ l'ensemble correspondant pour la racine β. On définit alors

$$M_0 = \{\xi'/\xi'' : \xi' \in \Xi(\alpha), \xi'' \in \Xi(\beta)\} \subset \mathbb{Q}(\alpha, \beta).$$

Si α et β sont conjugués sur $\mathbb{Q}$ (c'est en particulier le cas si f est irréductible) et que $\tau : \mathbb{Q}(\alpha) \to \mathbb{Q}(\beta)$ est l'isomorphisme envoyant α sur β, il suffit de prendre

$$M_0 = \{\xi/\tau(\xi) : \xi \in \Xi(\alpha)\}. \tag{4.6}$$

On définit alors

$$M = \left\{\mu \in M_0 : \mu \text{ n'est pas une puissance } p^{\text{ème}} \text{ dans } \mathbb{K}_0\right\}. \tag{4.7}$$

Alors la famille $\left\{\mathbb{K}_0, \mathbb{K}_0\left(\mu^{1/p}\right), \mu \in M\right\}$ forme un système complet de corps admissibles. Pour obtenir un système minimal, il suffit de chercher les corps isomorphes. □

4.4 Une unité

Dans toute cette section, on fixe un corps $\mathbb{K}$ appartenant au système complet de corps admissibles construit précédemment. On notera $k(x)$ l'entier tel que $\zeta^{k(x)}\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p} \in \mathbb{K}$.

4.4.1 Notations

Soit $m = [\mathbb{K} : \mathbb{Q}] = s + 2t$, où $\sigma_1, \ldots, \sigma_s : \mathbb{K} \to \mathbb{R}$ sont les plongements réels de $\mathbb{K}$ et $\sigma_{s+1}, \ldots, \sigma_{s+2t} : \mathbb{K} \to \mathbb{C}$ les plongements imaginaires, avec $\sigma_{s+t+i} = \overline{\sigma}_{s+i}$.

On notera α_i et β_i plutôt que $\sigma_i(\alpha)$ ou $\sigma_i(\beta)$.

4.4.2 Le vecteur k

Cette section répond (encore que cela ne soit pas vraiment une réponse) au problème suivant. Le plongement σ_i envoie α sur α_i et β sur β_i. Il envoie donc $\zeta^{k(x)}\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}$ sur $\zeta^{k_i(x)}\left(\dfrac{x-\alpha_i}{x-\beta_i}\right)^{1/p}$, où $k_i(x)$ est un entier de $\mathcal{P}$.

On ne peut pas déterminer explicitement k_i ; il va falloir nous contenter d'énumérer tous les k_i possibles ; toutefois, il est quand même possible de restreindre l'énumération à un ensemble plus petit que $\mathcal{P}^n$.

On étudie d'abord le prolongement des plongements de $\mathbb{K}_0$:

Proposition 4.8. *Si $[\mathbb{K} : \mathbb{K}_0] = p \geq 3$, chaque plongement de $\mathbb{K}_0$ a exactement un prolongement réel à $\mathbb{K}$, et $(p-1)/2$ paires de prolongements complexes conjugués ; en particulier, dans ce cas $s = s_0$ et $t = pt_0 + (p-1)s_0/2$, où s_0 et $2t_0$ sont respectivement les nombres de plongements réels et imaginaires du corps $\mathbb{K}_0$.*

Preuve. Définir un plongement prolongeant un plongement de $\mathbb{K}_0$, c'est choisir un vecteur k_i. Comme deux prolongements diffèrent d'une racine de l'unité, on voit qu'il y a au plus un prolongement réel ; il y en a exactement un qui correspond au choix $k = 0$ si $\dfrac{x-\alpha}{x-\beta}$ est positif, $k = (p-1)/2$ si $\dfrac{x-\alpha}{x-\beta} < 0$. Remarquons cependant que le choix de $|x| > X_1$ impose $\dfrac{x-\alpha}{x-\beta} > 0$.

Dans le cas imaginaire, on voit bien que l'argument de $\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}$ n'est pas un multiple de π/p, donc que tout prolongement est voué à être imaginaire. □

Les conditions triviales sur les k_i s'écrivent alors

$$k_1(x) = \cdots = k_s(x) = 0, \tag{4.8}$$

$$k_i(x) + k_{i+t}(x) \equiv 0 \bmod p \quad (s < i \leq s+t). \tag{4.9}$$

et la condition imposée par la Proposition (4.8) s'écrit comme suit :

$$\begin{array}{l}\text{si } \sigma_{i_1},\dots,\sigma_{i_p} \text{ sont les } p \text{ prolongements}\\ \text{distincts d'un plongement fixé de } \mathbb{K}_0,\\ \text{alors } \left\{k_{i_1}(x),\dots,k_{i_p}(x)\right\} = \mathcal{P}.\end{array} \tag{4.10}$$

Les conditions (4.8)–(4.10) réduisent le nombre de possibilités pour $\boldsymbol{k}$ à $\left(2^{\frac{p-1}{2}}\left(\frac{p-1}{2}\right)!\right)^{s_0}(p!)^{t_0}$.

Il y a moyen d'imposer une condition supplémentaire ; celle-ci n'a toutefois vraiment d'intérêt que quand le nombre de $\mu \in M$ donnant naissance au corps $\mathbb{K}$ est petit ; en effet, l'utilisation de cette condition impose d'énumérer tous les $\mu \in M$.

Soit $\mu \in M$ tel que $\mathbb{K} \cong \mathbb{K}_0(\mu^{1/p})$. Alors il existe un κ tel que $\mathbb{K}_0\left(\zeta^\kappa \mu^{1/p}\right) = \mathbb{K}$. On définit $\mu_i = \sigma_i(\mu)$ et $l_i \in \mathcal{P}$ par $\sigma_i\left(\zeta^\kappa \mu^{1/p}\right) = \zeta^{l_i}\mu_i^{1/p}$.

Soit alors $x \in \mathrm{Sol}$ tel que

$$\zeta^k\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\mu^{-1/p} \in \mathbb{K}_0 \quad \text{pour un } k \in \mathcal{P}. \tag{4.11}$$

Il s'ensuit que $\zeta^{k(x)}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\zeta^{-\kappa}\mu^{-1/p} \in \mathbb{K}_0$ également.

Proposition 4.9. *Supposons que* $\sigma_i\big|_{\mathbb{K}_0} = \sigma_{i'}\big|_{\mathbb{K}_0}$, *c'est-à-dire que* $\alpha_i = \alpha_{i'}$ *et* $\beta_i = \beta_{i'}$. *Alors pour tout* $x \in \mathrm{Sol}(\mathbb{K})$ *tel que l'on ait (4.11), on a*

$$k_i(x) - k_{i'}(x) \equiv l_i - l_{i'} \bmod p. \tag{4.12}$$

Preuve. Soit

$$\lambda = \zeta^{k(x)}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\zeta^{-\kappa}\mu^{-1/p}, \quad \lambda_i = \sigma_i(\lambda).$$

Comme λ et μ sont dans $\mathbb{K}_0$, on a $\lambda_i = \lambda_{i'}$ et $\mu_i = \mu_{i'}$. Par conséquent

$$\begin{aligned}\zeta^{k_i(x)}\left(\frac{x-\alpha_i}{x-\beta_i}\right)^{1/p} &= \zeta^{l_i}\mu_i^{1/p}\lambda_i,\\ &= \zeta^{l_i-l_{i'}}\zeta^{l_{i'}}\mu_{i'}^{1/p}\lambda_{i'},\\ &= \zeta^{l_i-l_{i'}}\zeta^{k_{i'}(x)}\left(\frac{x-\alpha_{i'}}{x-\beta_{i'}}\right)^{1/p},\\ &= \zeta^{l_i-l_{i'}}\zeta^{k_{i'}(x)}\left(\frac{x-\alpha_i}{x-\beta_i}\right)^{1/p},\end{aligned}$$

ce qui prouve la proposition. □

Étant donné la valeur de $k_i(x)$ pour un i donné, la condition (4.12) définit de manière unique $k_{i'}(x)$ pour les p valeurs de i' satisfaisant $\sigma_i\big|_{\mathbb{K}_0} = \sigma_{i'}\big|_{\mathbb{K}_0}$.

Combiné avec (4.8)–(4.9), ceci laisse donc uniquement p^{t_0} possibilités pour $\boldsymbol{k}(x)$. Il y a donc au plus

$$p^{t_0} \left| \left\{ \mu \in M \ : \ \mathbb{K}_0\left(\mu^{1/p}\right) \cong \mathbb{K} \right\} \right| \tag{4.13}$$

possibilités pour $\boldsymbol{k}(x)$, ce qui peut parfois être inférieur à $\left(2^{\frac{p-1}{2}}\left(\frac{p-1}{2}\right)!\right)^{s_0}(p!)^{t_0}$.

4.5 $\varphi(x)$

On définit maintenant $\varphi(x)$.

$$\varphi(x) = (x-\beta)\left(\zeta^{k(x)}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} - 1\right)^p. \tag{4.14}$$

Proposition 4.10. *Il existe un ensemble fini effectivement constructible $\Theta_0 \subset \mathbb{K}$ tel que pour tout $x \in \mathrm{Sol}(\mathbb{K})$, il existe $\theta_0 \in \Theta_0$ et η une unité de $\mathbb{K}$ avec*

$$\varphi(x) = \theta_0 \eta. \tag{4.15}$$

Preuve. Pour tout idéal premier $\mathfrak{P}$ de $\mathbb{K}$, on définit

$$u_1(\mathfrak{P}) = \max\left(0, -\mathrm{Ord}_{\mathfrak{P}}(\alpha), -\mathrm{Ord}_{\mathfrak{P}}(\beta)\right), \tag{4.16}$$

$$u_2(\mathfrak{P}) = \max\left(0, \mathrm{Ord}_{\mathfrak{P}}(\alpha-\beta)\right). \tag{4.17}$$

Les entiers naturels $u_1(\mathfrak{P})$ et $u_2(\mathfrak{P})$ sont tous deux nuls sauf pour un nombre fini d'idéaux premiers $\mathfrak{P}$. En particulier, si $f(x)$ est unitaire, $u_1(\mathfrak{P}) = 0$ pour tout $\mathfrak{P}$.

Soit Θ_0 un ensemble maximal de θ_0 deux à deux non associés de $\mathbb{K}$ satisfaisant

$$-u_1(\mathfrak{P}) \le \mathrm{Ord}_{\mathfrak{P}}(\theta_0) \le p u_2(\mathfrak{P}) + (p-1)u_1(\mathfrak{P}) \tag{4.18}$$

pour tout $\mathfrak{P}$. On rappelle que deux éléments de $\mathbb{K}$ sont dits associés si leur quotient est une unité.

Un tel ensemble peut être construit au moyen des opérations (DP) et (IP) dans $\mathbb{K}$.

Soit maintenant x une solution. Alors

$$\varphi(x)\tilde{\varphi}(x) = (\beta-\alpha)^p \tag{4.19}$$

où

$$\tilde{\varphi}(x) = \prod_{\substack{k' \in \mathcal{P} \\ k' \neq k(x)}} (x-\beta)\left(\zeta^{k'}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} - 1\right)^p$$

Mais pour un certain $k'' \in \mathcal{P}$, on a

$$\varphi(x) = \left((x-\alpha)^{1/p} - \zeta^{k''}(x-\beta)^{1/p}\right)^p \tag{4.20}$$

$$\tilde{\varphi}(x) = \prod_{\substack{k' \in \mathcal{P} \\ k' \neq k''}} \left((x-\alpha)^{1/p} - \zeta^{k'}(x-\beta)^{1/p}\right)^p \tag{4.21}$$

On fixe maintenant un idéal premier $\mathfrak{P}$ du corps $\mathbb{K}$. Soit $|\cdot|_{\mathfrak{P}}$ l'extension de la valuation $\mathfrak{P}$-adique sur $\mathbb{K}$ au corps $\mathbb{K}\left(\zeta, (x-\alpha)^{1/p}, (x-\beta)^{1/p}\right)$.

Au vu de (4.20) et (4.21) il est bien clair que

$$\begin{aligned} |\varphi(x)|_{\mathfrak{P}} &\leq \max\left(1, |\alpha|_{\mathfrak{P}}, |\beta|_{\mathfrak{P}}\right), \\ |\tilde{\varphi}(x)|_{\mathfrak{P}} &\leq \left(\max\left(1, |\alpha|_{\mathfrak{P}}, |\beta|_{\mathfrak{P}}\right)\right)^{p-1}. \end{aligned}$$

Par suite, vu (4.19),

$$-u_1(\mathfrak{P}) \leq \mathrm{Ord}_{\mathfrak{P}}\left(\varphi(x)\right) \leq pu_2(\mathfrak{P}) + (p-1)u_1(\mathfrak{P}), \tag{4.22}$$

donc

$$\varphi(x) \text{ est associé à un élément } \theta_0 \in \Theta_0, \tag{4.23}$$

ce qui conclut la preuve. □

Remarque 4.11. Dans le cas où $[\mathbb{K} : \mathbb{K}_0] = p$, la preuve peut être considérablement simplifiée et l'ensemble Θ_0 diminué en remarquant que

$$N_{\mathbb{K}/\mathbb{K}_0}\left(\varphi(x)\right) = \varphi(x)\tilde{\varphi}(x) = (\beta-\alpha)^p. \tag{4.24}$$

Il suffit donc de prendre pour Θ_0 un ensemble maximal de solutions non-associées de l'équation aux normes $N_{\mathbb{K}/\mathbb{K}_0}(z) = (\beta-\alpha)^p$, ou encore d'exclure de l'ensemble construit dans la preuve de la proposition tous les θ_0 tels que les idéaux

$$\left(N_{\mathbb{K}/\mathbb{K}_0}\left(\theta_0\right)\right) \quad \text{et} \quad \left((\beta-\alpha)^p\right) \tag{4.25}$$

soient distincts.

Soit Ω le groupe des racines de l'unité du corps $\mathbb{K}$, et $\eta_1, \ldots, \eta_r$ un système d'unités fondamentales – on peut encore une fois relâcher cette hypothèse. Définissons

$$\Theta = \{\theta_0\omega \ : \ \theta_0 \in \Theta_0, \omega \in \Omega\}.$$

Alors pour tout $x \in \mathrm{Sol}(\mathbb{K})$, il existe $\theta(x) \in \Theta$ tel que

$$\varphi(x) = \theta(x)\eta_1^{b_1(x)} \cdots \eta_r^{b_r(x)}, \tag{4.26}$$

où $\boldsymbol{b}(x) = (b_1(x), \ldots, b_r(x)) \in \mathbb{Z}^r$.

4.6 L'approximation de $\varphi(x)$

On fixe maintenant un corps admissible $\mathbb{K}$, un élément θ de Θ, et un vecteur $\boldsymbol{k}$ vérifiant les conditions (4.8)–(4.10) (plus éventuellement la condition qui suit celles-ci, mais alors il faut aussi fixer un $\mu \in M$). On notera $\mathrm{Sol}(\mathbb{K}, \boldsymbol{k}, \theta)$ le sous-ensemble des éléments de Sol correspondant à ce choix de $\mathbb{K}$, $\boldsymbol{k}$ et θ.

On pose alors $\varphi_i(x) = \sigma_i(\varphi(x)) = (x-\beta_i)\left(\zeta_i\left(\dfrac{x-\alpha_i}{x-\beta_i}\right)^{1/p} - 1\right)^p$,

où $\zeta_i = \zeta^{k_i}$. Définissons également

$$\begin{array}{ll} \theta_i = \sigma_i(\theta), & \eta_{ij} = \sigma_i(\eta_j), \\ \rho_i = \begin{cases} 1-p, & k_i = 0, \\ 1, & k_i \neq 0, \end{cases} & \gamma_i = \begin{cases} \left(\frac{\beta_i - \alpha_i}{p}\right)^p, & k_i = 0, \\ (\zeta_i - 1)^p, & k_i \neq 0, \end{cases} \end{array} \tag{4.27}$$

pour $1 \leq i \leq m$.
Soit $A = [a_{ij}]_{1\leq i,j\leq r}$ l'inverse de la matrice $[\log|\eta_{ij}|]_{1\leq i,j\leq r}$.
Pour $1 \leq i \leq r$ on pose

$$\delta_i = \sum_{j=1}^{r} a_{ij}\rho_j, \qquad \lambda_i = \sum_{j=1}^{r} \left(a_{ij} \log|\gamma_j \theta_j^{-1}|\right). \tag{4.28}$$

On définit enfin les constantes

$$\begin{array}{ll} c_2 = 4\left|2^{-1/p} + (2p)^{-1} - 1\right|, & c_3 = 2\left(1 - 2^{-1/p}\right), \quad c_4 = 2\sin(\pi/p), \\ c_5 = \max\limits_{1\leq i\leq m}(|\alpha_i| + |\beta_i|), & c_6 = \max\limits_{1\leq i\leq m} \dfrac{|\alpha_i|^2 + |\beta_i|^2}{|\alpha_i - \beta_i|}, \\ X_2 = \max\left(X_1, 2c_3c_4^{-1}c_5, 2pc_2c_6\right), & c_7 = 1,39p\max\left(c_3c_4^{-1}c_5, pc_2c_6\right), \\ c_8 = \max\limits_{1\leq j\leq r}|\delta_j|, & c_9 = \dfrac{1}{20} + \max\limits_{1\leq j\leq r}|\lambda_j|, \\ c_{10} = c_7 \max\limits_{1\leq i\leq r}\sum\limits_{j=1}^{r}|a_{ij}|, & X_3 = \max(X_2, 20c_{10}). \end{array}$$

Proposition 4.12. *On suppose que $x \in \mathrm{Sol}(\mathbb{K}, \boldsymbol{k}, \theta)$ et que $|x| > X_2$. Alors*

$$\mathrm{Log}\left(\frac{\varphi_i(x)}{\gamma_i x^{\rho_i}}\right) \leq \frac{c_7}{|x|}, \quad (1 \leq i \leq m), \tag{4.29}$$

$$|b_i(x) - \delta_i \log|x| + \lambda_i| \leq \frac{c_{10}}{|x|}, \quad (1 \leq i \leq r). \tag{4.30}$$

De plus, si $|x| > X_3$,

$$\max_{1\leq i\leq r}|b_i(x)| \leq c_8 \log|x| + c_9. \tag{4.31}$$

Preuve. Soit z un nombre complexe vérifiant $|z| \leq 1/2$. Alors

$$|(1+z)^{1/p} - 1 - p^{-1}z| \leq c_2|z|^2, \tag{4.32}$$
$$|\mathrm{Log}\,(1+z)| \leq 1,39|z|. \tag{4.33}$$

Le deuxième point est extrait du Lemme 1.6. Quant au premier point, il se prouve de manière analogue : on considère $\psi(z) = z^{-2}\left((1+z)^{1/p} - 1 - p^{-1}z\right)$, et on a

$$|\psi(z)| = \left|\sum_{\nu=2}^{\infty} \binom{1/p}{\nu} z^{\nu-2}\right| \leq \sum_{\nu=2}^{\infty} \left|\binom{1/p}{\nu}\right| (1/2)^{\nu-2} = |\psi(-1/2)| = c_2.$$

On peut aussi montrer de la même manière que

$$|(1+z)^{1/p} - 1| \leq c_3|z|. \tag{4.34}$$

Maintenant, comme $|x| > X_1$, de (4.1) on déduit que

$$\varphi_i(x) = x\left(\zeta_i(1-\alpha_i x^{-1})^{1/p} - (1-\beta_i x^{-1})^{1/p}\right)^p. \tag{4.35}$$

Quand $k_i \neq 0$, on a $\zeta_i \neq 1$ et de plus $|\zeta_i - 1| > c_4$. Par suite,

$$\left|\frac{\zeta_i(1-\alpha_i x^{-1})^{1/p} - (1-\beta_i x^{-1})^{1/p}}{\zeta_i - 1} - 1\right| \leq \frac{c_3 c_5}{c_4|x|},$$

et, comme $|x| > X_2$, le second membre est plus petit que $1/2$, et

$$\left|\mathrm{Log}\,\frac{\varphi_i(x)}{(\zeta_i - 1)^p\, x}\right| \leq \frac{1,39pc_3c_5}{c_4|x|}$$

c'est-à-dire le cas $k_i \neq 0$ de (4.29).

Si $k_i = 0$, on a :

$$\left|\frac{(1-\alpha_i x^{-1})^{1/p} - (1-\beta_i x^{-1})^{1/p}}{\left(\dfrac{\beta_i - \alpha_i}{p}\right)^p x^{-p}} - 1\right| \leq \frac{pc_2c_6}{|x|},$$

et de même que précédemment,

$$\mathrm{Log}\left(\frac{\varphi_i(x)}{\gamma_i x^{1-p}}\right) \leq \frac{1,39p^2c_2c_6}{|x|},$$

et (4.29) est établie aussi pour les i tels que $k_i = 0$.

Dès lors, le deuxième point se prouve de la même manière que l'identité analogue du chapitre 1. □

En conclusion, montrons que dans le cas où

$$\big|\{i\colon k_i(x)=0\}\big| \neq m/p, \tag{4.36}$$

on peut donner une borne supérieure pour les solutions.

Corollaire 4.13. *Supposons que (4.36) est réalisée, et faisons*

$$\rho' = m - p\,\big|\{i\colon k_i(x)=0\}\big|\,.$$

Alors tout $x \in \mathrm{Sol}\,(\mathbb{K}, \boldsymbol{k}, \theta)$ *vérifie*

$$|x| \leq X_4 := \max\Big(X_2,\ 3mc_7,\ e^{1/(3|\rho'|)}\,\big|\gamma_1\cdots\gamma_m\big|^{-1/\rho'}\,\big|\theta_1\cdots\theta_m\big|^{1/\rho'}\Big)\,. \tag{4.37}$$

Preuve. D'une part,

$$\varphi_1(x)\cdots\varphi_m(x) = N_{\mathbb{K}/\mathbb{Q}}(\varphi(x)) = \pm\theta_1\cdots\theta_m. \tag{4.38}$$

D'autre part,

$$\log\left|\frac{\varphi_1(x)\cdots\varphi_m(x)}{\gamma_1\cdots\gamma_m x^{\rho_1+\cdots+\rho_m}}\right| \leq \frac{mc_7}{|x|}.$$

Comme $\rho_1+\cdots+\rho_m = \rho'$, le résultat s'ensuit.

Remarque 4.14. Au vu de (4.10), (4.36) ne peut se produire que si $\mathbb{K} = \mathbb{K}_0$; en revanche, si $m \not\equiv 0 \bmod p$, la condition (4.36) est toujours vérifiée pour $\mathbb{K}_0$.

4.7 Une borne pour $\max_i |b_i|$

De la théorie de Baker et des résultats de la section précédente, on va maintenant déduire une borne supérieure pour $\max_i |b_i|$.

Vu (4.10), dans le cas où $[\mathbb{K} : \mathbb{K}_0] = p \geq 3$, il existe i_1 et i_2 parmi $\{1,\dots,m\}$ tels que

$$\alpha_{i_1} = \alpha_{i_2},\quad \beta_{i_1} = \beta_{i_2},\quad k_{i_1} \neq 0,\quad k_{i_1} \neq 0,\quad k_{i_1} \neq k_{i_2}. \tag{4.39}$$

Dans le cas où $\mathbb{K} = \mathbb{K}_0$, on se contente d'exiger que parmi les quatre nombres

$$\alpha_{i_1},\ \alpha_{i_2},\ \beta_{i_1},\ \beta_{i_2} \tag{4.40}$$

il y en ait au moins trois distincts.

Ce choix est possible dès que $[\mathbb{K}_0 : \mathbb{Q}] \geq 3$, ce que nous supposerons donc dorénavant.

On définit maintenant

$$\Phi(x) = \frac{\gamma_{i_2}^{\rho_{i_1}} \varphi_{i_1}(x)^{\rho_{i_2}}}{\gamma_{i_1}^{\rho_{i_2}} \varphi_{i_2}(x)^{\rho_{i_1}}},$$

qui doit être très voisin de 1, en vertu de (4.29). Avant d'appliquer le Théorème 1.9, il nous faut étudier l'équation

$$\Phi(x) = 1. \tag{4.41}$$

4.7.1 L'équation $\Phi(x) = 1$

4.7.1.1 $[\mathbb{K} : \mathbb{K}_0] = p$

Dans le cas où $[\mathbb{K} : \mathbb{K}_0] = p \geq 3$, le choix de i_1 et de i_2 garantit que $k_{i_1} \neq 0$ et $k_{i_2} \neq 0$; par suite $\rho_{i_1} = \rho_{i_2} = 1$, et l'équation $\Phi(x) = 1$ se réécrit

$$\zeta_{i_1} \left(1 - \alpha_{i_1} x^{-1}\right)^{1/p} - \left(1 - \beta_{i_1} x^{-1}\right)^{1/p} = \zeta^{\kappa} \left(\zeta_{i_2} \left(1 - \alpha_{i_2} x^{-1}\right)^{1/p} - \left(1 - \beta_{i_2} x^{-1}\right)^{1/p}\right),$$

où κ est un élément de $\mathcal{P}$. En utilisant le fait que $\alpha_{i_1} = \alpha_{i_2}$ et $\beta_{i_1} = \beta_{i_2}$, il vient

$$x = \frac{\alpha_{i_1} \left(\zeta^{k_{i_1}} - \zeta^{\kappa + k_{i_2}}\right)^p - \beta_{i_1} \left(1 - \zeta^{\kappa}\right)^p}{\left(\zeta^{k_{i_1}} - \zeta^{\kappa + k_{i_2}}\right)^p - \left(1 - \zeta^{\kappa}\right)^p},$$

dès lors que le dénominateur de cette expression est non nul.

4.7.1.2 $\mathbb{K} = \mathbb{K}_0$

Si $\mathbb{K} = \mathbb{K}_0$, soit $\mathcal{K}$=$\mathbb{C}\left(T, \left(\frac{T-\alpha_{i_1}}{T-\beta_{i_1}}\right)^{1/p}, \left(\frac{T-\alpha_{i_2}}{T-\beta_{i_2}}\right)^{1/p}\right)$, qui est une extension finie de $\mathbb{C}(T)$. On résout $\Phi(x) = 1$ en calculant la fraction rationnelle à coefficients complexes

$$N_{\mathcal{K}/\mathbb{C}(T)}(\Phi(T) - 1) \tag{4.42}$$

dont on cherche ensuite les zéros.

On peut prouver que la fraction rationnelle ci-dessus n'est pas nulle, de sorte que chercher ses zéros a un sens ; ceci découle du lemme suivant, conséquence facile de la théorie de Kummer pour le corps $\mathbb{C}(T)$:

Lemme 4.15. *Soit $\varsigma_1, \ldots, \varsigma_\nu$ des nombres complexes non nuls deux à deux distincts et $F(T, T_1, \ldots, T_\nu) \in \mathbb{C}[T, T_1, \ldots, T_\nu]$ un polynôme non nul vérifiant*

$$F\left(T, (1 - \varsigma_1 T)^{1/p} \ldots (1 - \varsigma_\nu T)^{1/p}\right) \equiv 0.$$

Alors $\deg_{T_i} F \geq p$ *pour* $1 \leq i \leq \nu$.

Supposons maintenant que $\Phi(x)$ soit constante ; alors selon que ρ_{i_1} et ρ_{i_2} sont égaux ou non, on a soit

$$\zeta_{i_1}(1-\alpha_{i_1}x^{-1})^{1/p} - (1-\beta_{i_1}x^{-1})^{1/p} = c\left(\zeta_{i_2}(1-\alpha_{i_2}x^{-1})^{1/p} - (1-\beta_{i_2}x^{-1})^{1/p}\right),$$

soit

$$x\left(\zeta_{i_1}(1-\alpha_{i_1}x^{-1})^{1/p} - (1-\beta_{i_1}x^{-1})^{1/p}\right)\left(\zeta_{i_2}(1-\alpha_{i_2}x^{-1})^{1/p} - (1-\beta_{i_2}x^{-1})^{1/p}\right)^{p-1} = c.$$

où l'on a supposé x assez grand et $\rho_{i_1} = 1, \rho_{i_2} = 1-p$ dans le cas où les deux diffèrent.

Maintenant, au moins trois des radicaux sont distincts, par le choix de i_1 et de i_2. Donc après simplifications, il reste au moins deux termes distincts non nuls, qui interviennent avec un degré strictement inférieur à p, et l'application du lemme ci-dessus permet de conclure.

Il reste encore une difficulté, qui tient au fait que si la fraction rationnelle (4.42) peut être calculée à une précision arbitraire, on ne peut en donner une valeur exacte, et cela perturbe les valeurs des racines. Le lemme suivant montre comment contrôler cette perturbation.

Lemme 4.16. *Soit $P(x) = a_0x^N + a_1x^{N-1} + \ldots + a_N$ et $Q(x) = b_0x^N + b_1x^{N-1} + \ldots + b_N$ des polynômes à coefficients complexes. Soit ε un nombre positif vérifiant la condition*

(*) quelles que soient z et z' racines de Q, soit $|z-z'| \le \varepsilon/2$, soit $|z-z'| \ge 2\varepsilon$.

Posons

$$\delta = \delta(\varepsilon) = \min_{1\le i\le N} \frac{|b_0|\prod_{j=1}^N \left||z_j - z_i| - \varepsilon\right|}{\sum_{j=0}^N (|z_i|+\varepsilon)^j}, \tag{4.43}$$

où $z_1, \ldots, z_N$ sont les racines de Q, comptées avec multiplicités. On suppose que

$$|a_i - b_i| < \delta \quad (0 \le i \le N).$$

Alors pour toute racine z de P il existe une racine z' de Q telle que $|z - z'| < \varepsilon$.

Preuve. On définit une relation d'équivalence sur l'ensemble des racines de Q par $z \sim z'$ si $|z - z'| \le \varepsilon/2$ (la transitivité provient de (*)). Soit $m_1, \ldots, m_k$ les cardinalités des classes d'équivalence, (de sorte que $m_1 + \cdots + m_k = N$; les m_i sont des sortes de "quasi-multiplicités"), et réordonnons les racines de manière à ce que $z_1, \ldots, z_k$ soient des représentants des différentes classes. En particulier,

$$|z_i - z_j| \ge 2\varepsilon, \quad (1 \le i < j \le k). \tag{4.44}$$

Soit maintenant i tel que $1 \le i \le k$. Alors Q a m_i racines dans le disque $\Delta_i := \{z \in \mathbb{C} \,:\, |z - z_i| < \varepsilon\}$. Pour tout z sur le cercle $\Gamma_i := \{z \in \mathbb{C} \,:\, |z - z_i| = \varepsilon\}$ on a

$$|Q(z)| = \left|b_0(z-z_1)\ldots(z-z_N)\right| \ge |b_0| \prod_{j=1}^N \left||z_j - z_i| - \varepsilon\right|.$$

Par suite, pour tout $z \in \Gamma_i$

$$|P(z) - Q(z)| < \delta \sum_{j=0}^{N} (|z_i| + \varepsilon)^j \leq |Q(z)|.$$

D'après le théorème de Rouché, P a donc également m_i racines dans Δ_i.

D'après (4.44), les disques Δ_i sont disjoints deux à deux ; comme $m_1 + \cdots + m_k = N$, toute racine de P est dans l'un des Δ_i. □

Il peut ne pas être superflu d'indiquer comment utiliser ce lemme ; supposons que l'on ait à trouver toutes les racines entières de P, dont on connaît une approximation Q avec une précision δ_0.

Une fois calculées les racines $z_1, \ldots, z_N$ de Q, on décide quels sont les entiers qui sont très voisins d'une racine de Q. Notons $\mathcal{S}$ l'ensemble de ces entiers.

Pour tout $z \in \mathbb{C}$, on note $\rho(z)$ la distance de z à $\mathbb{Z}\backslash\mathcal{S}$.

On pose alors $\varepsilon_0 = 0,1 \min(\rho(z_1), \ldots, \rho(z_N))$. Si ε_0 vérifie $(*)$, on prend $\varepsilon = \varepsilon_0$; sinon, $\varepsilon_1 = 0,1 \min_{|z-z'|\geq\varepsilon_0} |z - z'|$, où z et z' décrivent l'ensemble des racines de $\mathbb{Q}$. Si ε_1 vérifie $(*)$, on prend $\varepsilon = \varepsilon_1$; dans le cas contraire, posons $\varepsilon_2 = 0,1 \min_{|z-z'|\geq\varepsilon_1} |z - z'|$, etc. En pratique, on trouve toujours un ε convenable en deux ou trois étapes. On calcule alors δ. Si $\delta > \delta_0$, les éléments de $\mathcal{S}$ sont les seules racines entières possibles. Sinon (ce n'est jamais arrivé en pratique), on recalcule l'approximation Q de P avec une meilleure précision.

4.7.2 La borne de Baker

Si l'on exclut alors les x trouvés dans la section précédente, on peut trouver une borne pour $\max_{1\leq i\leq r} |b_i|$. On suppose donc

$$\Phi(x) \neq 1. \tag{4.45}$$

On sait, d'après (4.29), que

$$\operatorname{Log} \Phi(x) \leq \frac{c_{12}}{|x|}. \tag{4.46}$$

Par ailleurs,

$$\Phi(x) = \vartheta_0 \vartheta_1^{b_1(x)} \cdots \vartheta_r^{b_r(x)}, \tag{4.47}$$

où

$$\vartheta_0 = \frac{\gamma_{i_2}^{\rho_{i_1}} \theta_{i_1}^{\rho_{i_2}}}{\gamma_{i_1}^{\rho_{i_2}} \theta_{i_2}^{\rho_{i_1}}}, \qquad \vartheta_j = \frac{\eta_{i_1 j}^{\rho_{i_2}}}{\eta_{i_2 j}^{\rho_{i_1}}}. \tag{4.48}$$

En prenant le logarithme, on voit qu'il existe un entier $b_{r+1}(x) \in \mathbb{Z}$ tel que

$$0 < |\log \Phi(x)| = |\log \vartheta_0 + b_1(x) \log \vartheta_1 + \cdots + b_r(x) \log \vartheta_r + b_{r+1}(x)\pi i| \leq c_{12}|x|^{-1}. \tag{4.49}$$

Mais d'après le Théorème 1.9, il existe une constante c_{11} telle que

$$|\log \Phi(x)| \geq \exp\left(-c_{11} \log B(x)\right), \tag{4.50}$$

où

$$B(x) := \max\left(b_1(x) \dots b_r(x), b_{r+1}(x), e\right) \leq c_{13} \log|x| + c_{14}, \tag{4.51}$$

avec $c_{13} = rc_8$ et $c_{14} = \max\left(1 + \pi^{-1}c_{12} + rc_9, e\right)$.

Il s'ensuit le

Théorème 4.17. *On a $B(x) \leq B_0 := 2(c_{15} \log c_{15} + c_{16})$, où $c_{15} = c_{11}c_{13}$, $c_{16} = c_{13}c_{12} + c_{14}$.*

Preuve. De (4.50) et (4.46), on tire

$$\log|x| \leq c_{11} \log B(x) + \log c_{12}. \tag{4.52}$$

Par suite, $B(x) \leq c_{15} \log B(x) + c_{16}$. Il suffit alors d'appliquer le Lemme 1.20. □

Remarque 4.18. Dans le cas où $r = 1$, notons que l'inégalité (4.49), où l'on a pris les logarithmes des modules, permet de résoudre l'équation quand $|\vartheta_1| \neq 1$, sans utiliser la borne de Baker. En effet, on a une identité du type

$$|\alpha + \beta b_1| \leq \frac{c_{12}}{|x|^n}.$$

Dès lors que $|x| \geq 2c_{12}/\beta$, ceci impose $b_1 = \lfloor -\alpha/\beta \rceil$. On peut alors calculer le x et le y correspondants, et vérifier s'ils correspondent bien à une solution ou non.

Cette situation ne peut pas se produire dans les chapitres 1 et 2 : $r = 1$ n'est possible que pour un corps quadratique réel, auquel cas $P(X, Y)$ est réductible, pour un corps cubique avec $s = 1$, auquel cas $|\vartheta_1| = 1$ (puisque c'est le quotient de deux conjugués), ou pour un corps quartique totalement imaginaire, auquel cas on dispose d'une très bonne borne sur $|x|$ par un argument élémentaire, voir (1.5) et ce qui suit.

4.8 Des b_i à x

La réduction agit de la même manière qu'au chapitre 1, à ceci près qu'elle est un peu moins efficace ; en effet, l'approximation de $\varphi(x)$ est en $|x|^{-1}$, et non plus $|x|^{-n}$; ceci a pour effet de diminuer la valeur de la constante dans le majorant en $O(\exp(-C \max_i |b_i|))$.

L'énumération des vecteurs $\boldsymbol{b}$ procède également des mêmes idées ; on se contente de décrire ici comment revenir des b_i aux x.

Fixons encore un corps admissible $\mathbb{K}$, un élément θ de Θ, et un vecteur $\boldsymbol{k}$.

Choisissons un i tel que $k_i = 0$. Pour $x \in \mathbb{Z}$, on pose

$$\omega_i := \gamma_i^{-1}\varphi_i(x) - \gamma_i',$$

où $\gamma_i' = \frac{\beta_i - \zeta_i\alpha_i}{\zeta_i - 1}$.

On a aussi besoin des constantes suivantes :

$$c_{19} = 4(3/2)^p - 4 - 2p, \quad c_{20} = c_2(|\alpha_i|^2 + |\beta_i|^2)/|\zeta_i - 1|, \quad c_{21} = p^{-1}|\gamma_i'| + c_{20}X_3^{-1},$$
$$c_{22} = c_{21}X_3^{-1}, \qquad c_{23} = c_{19}c_{21} + pc_{20}, \qquad X_5 = \max(X_3, 2c_{22}, 2c_{23}).$$

Lemme 4.19. *Si $|x| > X_5$, alors*

$$|x - \omega_i| < \min(1/2, c_{23}/(|\omega_i| - 1/2)). \tag{4.53}$$

Preuve. Pour $|z| \leq 1/2$ on a

$$|(1+z)^p - 1 - pz| \leq c_{19}|z|^2, \tag{4.54}$$

ce qui se prouve de la même manière que (4.34). D'après (4.54), (4.34) et (4.35), pour $|x| > X_3$, on a

$$\begin{aligned} \varphi_i(x) &= \gamma_i x\left(1 + \gamma_i' x^{-1}/p + c_{20}x^{-2}\nu\right)^p \\ &= \gamma_i x\left(1 + \gamma_i' x^{-1} + \nu' c_{23}x^{-2})\right) \end{aligned}$$

où $|\nu| \leq 1$, $|\nu'| \leq 1$, et donc

$$|x - \omega_i| \leq c_{23}|x|^{-1}. \tag{4.55}$$

Comme $|x| > 2c_{23}$, on a $|x - \omega_i| < 1/2$. En particulier, $|\omega_i| < |x| + 1/2$, ce qui, joint à (4.55), montre que $|x - \omega_i| < c_{23}/(|\omega_i| - 1/2)$. □

Corollaire 4.20. *Soit $\boldsymbol{b} = (b_1, \ldots, b_r)$ un r-uplet d'entiers, et posons*

$$\omega_i := \gamma_i^{-1}\theta_i\eta_{i1}^{b_1}\cdots\eta_{ir}^{b_r} - \gamma_i'. \tag{4.56}$$

Alors si $\boldsymbol{b}$ correspond à une solution $|x| > X_5$, on a

$$|\omega_i| > X_5 - 1/2 \quad \textit{et} \quad \|\omega_i\| < c_{23}/(|\omega_i| - 1/2).$$

Si les deux conditions du corollaire sont vérifiées, on a alors $x = \lfloor \omega_i \rceil$.

4.9 L'algorithme

On résume l'algorithme décrit jusque là dans le cas où $p \geq 3$.

1. Construire un système complet minimal de corps admissibles.
2. Choisir un corps admissible $\mathbb{K}$ n'ayant pas encore été traité ; si tous les corps ont été traités, aller à l'étape 10.
3. Construire l'ensemble Θ correspondant, et les vecteurs $\boldsymbol{k}$ possibles.
4. Fixer Θ et $\boldsymbol{k}$; si tous les couples $(\Theta, \boldsymbol{k})$ ont été traités, aller à l'étape 2.
5. Si (4.36) est réalisé, calculer X_4 et aller à l'étape 4. Sinon, calculer X_3, et aller à l'étape 6.
6. Construire la fonction $\Phi(x)$ et trouver les solutions entières de $\Phi(x) = 1$. Pour chacune de ces solutions, vérifier s'il s'agit bien d'une solution de (4.1).
7. Calculer la borne de Baker B_0.
8. Réduire la borne de Baker jusqu'à B_0'.
9. Énumérer les vecteurs $\boldsymbol{b}$ possibles, et retrouver les x correspondants ; vérifier s'ils sont bien solutions. Aller à l'étape 4.
10. Calculer X_6, le maximum de tous les X_4 calculés à l'étape 5, et tous les X_5 calculés à l'étape 9.
11. Pour tout $x \in \mathbb{Z}$ tel que $|x| \leq X_6$, vérifier si x est solution de (4.1).
12. Rassembler toutes les solutions obtenues lors des étapes 6, 9, et 11. Fin.

4.10 (α, β)-symétrie

4.10.1 Discussion

Discutons maintenant la portée des exigences algorithmiques formulées au début de ce chapitre. Il nous faut effectuer les opérations (DP), (U) et (GC) dans les corps $\mathbb{Q}(\alpha)$ et $\mathbb{Q}(\beta)$, et les opérations (DP), (U) et (IP) dans chacun des corps admissibles. Ce dernier point est, en pratique, le plus difficile à réaliser ; de fait, dans le cas général, les corps admissibles (sauf $\mathbb{K}_0$) sont de degré $pn(n-1)$.

Même dans le cas $(p, n) = (3, 4)$ on doit accomplir les opérations "multiplicatives" ci-dessus dans des corps de degré 36, ce qui est bien au-delà des possibilités de la théorie algorithmique des nombres aujourd'hui.

Il y a moyen de se protéger contre ce "cas le pire", en utilisant une idée analogue à celle du chapitre 2. On va exploiter l'existence éventuelle d'un sous-corps $\mathbb{K}$ du corps admissible $\mathbb{K}'$ qui ne soit pas $\mathbb{K}_0$, et se ramener à ce sous-corps en remplaçant $\varphi(x)$ par $N_{\mathbb{K}'/\mathbb{K}}(\varphi(x))$.

4.10.2 Préliminaires

On considère toujours l'équation $y^p = f(x)$. On dira qu'il y a (α, β)-*symétrie* s'il existe un automorphisme σ de $\mathbb{Q}(\alpha, \beta)$ permutant α et β.

Les racines α et β de f seront dans ce cas dites symétriques.

Dans ce cas, on a un sous-corps $\mathbb{Q}(\alpha+\beta, \alpha\beta) = \mathbb{Q}(\alpha, \beta)^\sigma$, dont on va exploiter l'existence comme on l'a fait au chapitre 2 ; les corps admissibles vont de la même manière voir leur degré divisé par 2.

Lemme 4.21. *Soit $\mathbb{K}_0 \subset \mathbb{K}_1 \subset \mathbb{K}_2$ une tour de corps de nombres vérifiant les propriétés suivantes :*

– *$[\mathbb{K}_1 : \mathbb{K}_0] = 2$ et $[\mathbb{K}_2 : \mathbb{K}_1] = p$.*

– *$\mathbb{K}_2 = \mathbb{K}_1(\nu)$, où $\nu^p = \mu \in \mathbb{K}_1$ et μ est conjugué à μ^{-1} sur $\mathbb{K}_0$.*

Alors

(a) $[\mathbb{K}_0(\nu + \nu^{-1}) : \mathbb{K}_0] = p$.

Soit $\mathbb{K}$ un corps de nombres intermédiaire entre $\mathbb{K}_0$ et $\mathbb{K}_2$, tel que $[\mathbb{K} : \mathbb{K}_0] = p$. Alors

(b) si $\zeta \notin \mathbb{K}_0$ alors $\mathbb{K} = \mathbb{K}_0(\nu + \nu^{-1})$;

(c) si $\zeta \in \mathbb{K}_0$, $\mathbb{K}$ est l'un des p corps distincts $\mathbb{K}_0(\zeta^k\nu + \zeta^{-k}\nu^{-1})$, où $k \in \mathcal{P}$;

(d) le compositum $\mathbb{K}_1\mathbb{K}$ est $\mathbb{K}_2$.

Preuve. Il est bien clair que $\mathbb{K}_2 = \mathbb{K}_0(\nu)$. De plus, ν et ν^{-1} sont conjugués au dessus de $\mathbb{K}_0$. On définit $\tau : \mathbb{K}_2 \to \mathbb{K}_2$ par $\tau(\nu) = \nu^{-1}$. Alors $\nu + \nu^{-1}$ est stable par τ, donc le degré $[\mathbb{K}_0(\nu+\nu^{-1}) : \mathbb{K}_0]$ divise p. Ceci prouve le point (a), puisque $\nu + \nu^{-1} \notin \mathbb{K}_0$ (sinon, $[\mathbb{K}_2 : \mathbb{K}_0] = [\mathbb{K}_0(\nu) : \mathbb{K}_0] \leq 2$, ce qui n'est pas).

Le point (d) est clair : comme p est impair, $\mathbb{K}_1 \not\subset \mathbb{K}$, donc $\mathbb{K} \subsetneq \mathbb{K}_1\mathbb{K} \subseteq \mathbb{K}_2$, et la seule possibilité est $\mathbb{K}_1\mathbb{K} = \mathbb{K}_2$.

Pour prouver (b) et (c), on remarque que les corps $\mathbb{K}$ correspondent à des involutions non triviales de $\mathbb{K}_2$ au dessus de $\mathbb{K}_0$, c'est-à-dire à des automorphismes $\tau : \mathbb{K}_2 \to \mathbb{K}_2$ tels que $\tau\big|_{\mathbb{K}_0} = \mathrm{id}_{\mathbb{K}_0}$ et τ d'ordre 2. Les conjugués de ν au-dessus de $\mathbb{K}_0$ se trouvent parmi les nombres

$$\nu, \zeta\nu \dots \zeta^{p-1}\nu, \nu^{-1}, \zeta\nu^{-1} \dots \zeta^{p-1}\nu^{-1}. \tag{4.57}$$

Comme $[\mathbb{K}_0(\nu) : \mathbb{K}_0] = 2p$, tous les nombres de la liste ci-dessus sont des conjugués de $\mathbb{K}_0$ sur ν.

Trois cas sont à distinguer ;

(b)′ $\zeta \notin \mathbb{K}_1$. Alors, parmi les conjugués de ν au-dessus de $\mathbb{K}_0$, seul ν^{-1}appartient à $\mathbb{K}_2$. Par suite, la seule involution de $\mathbb{K}_2$ au-dessus de $\mathbb{K}_0$ est celle envoyant ν sur ν^{-1}, et $\mathbb{K} = \mathbb{K}_0(\nu + \nu^{-1})$.

(b)″ $\zeta \in \mathbb{K}_1 \setminus \mathbb{K}_0$. Dans ce cas, tous les nombres (4.57) sont dans $\mathbb{K}_2$. Par suite $\mathbb{K}_2/\mathbb{K}_0$ est normale et le groupe $G := \mathrm{Gal}(\mathbb{K}_2/\mathbb{K}_0)$ est engendré par $\tau : \nu \to \nu^{-1}$ et $\sigma : \nu \to \zeta\nu$. Clairement, $\tau\big|_{\mathbb{K}_1}$ est la seule involution non triviale de $\mathbb{K}_1/\mathbb{K}_0$, donc $\tau(\zeta) = \zeta^{-1}$. Par suite, $\tau\sigma = \sigma\tau$, G est abélien, et τ est de nouveau la seule involution de $\mathbb{K}_2/\mathbb{K}_0$.

(c) $\zeta \in \mathbb{K}_0$. Dans ce cas, $\tau\sigma = \sigma^{-1}\tau$,il y a p involutions distinctes $\tau_k := \sigma^k \tau \sigma^{-k}$, où $k \in \mathcal{P}$. De fait, les involutions τ_k sont deux à deux distinctes puisque les nombres $\tau_k(\nu) = \zeta^{-2k}\nu$ sont distincts quand k décrit $\mathcal{P}$. Mais, par le théorème de Sylow, tous les sous-groupes à deux éléments de G sont conjugués. Comme il y a au plus $[G : \{1, \tau\}]$ tels sous-groupes, il n'y a pas d'autres involutions que $\tau_0 = \tau, \tau_1, \ldots, \tau_{p-1}$.

L'involution τ_k correspond au corps $\mathbb{K}_0(\zeta^k\nu + \zeta^{-k}\nu^{-1})$.

Ceci conclut la preuve. □

4.10.3 Corps admissibles

Dans la situation d'(α, β)-symétrie on est amené à définir les corps admissibles de manière un peu différente.

Soit $\tau : \mathbb{Q}(\alpha, \beta) \to \mathbb{Q}(\alpha, \beta)$ donné par $\tau(\alpha) = \beta$, $\tau(\beta) = \alpha$. On définit $\mathbb{K}_0 = \mathbb{Q}(\alpha + \beta, \alpha\beta)$, $\mathbb{K}_0' = \mathbb{Q}(\alpha, \beta)$, c'est-à-dire que $\mathbb{K}_0 = \mathbb{K}_0'^{\tau}$.

Définition 4.22. *Un corps de nombres est admissible pour une solution x s'il existe un $k \in \mathcal{P}$ tel que* $\mathbb{K} = \mathbb{K}_0\left(\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p} + \zeta^{-k}\left(\dfrac{x-\beta}{x-\alpha}\right)^{1/p}\right)$.

On définit de la même manière que dans le cas général les systèmes complets et systèmes complets minimaux de corps admissibles.

Soit M l'ensemble fini de $\mathbb{Q}(\alpha, \beta)$ construit dans (4.6) et (4.7).

Proposition 4.23. *Les corps $\mathbb{K}_0$ et $\mathbb{K}_0\left(\mu^{1/p} + \mu^{-1/p}\right)$, où μ décrit M, forment un système complet de corps admissibles.*

Preuve. Supposons tout d'abord que $\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p} \in \mathbb{K}_0'$ pour un certain k de $\mathcal{P}$. Montrons que dans ce cas K_0' est admissible pour x.

On a $\tau\left(\zeta^k\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p}\right) = \zeta^{k'}\left(\dfrac{x-\beta}{x-\alpha}\right)^{1/p}$ pour un certain k' de $\mathcal{P}$. Par suite,

$$\zeta^{k+k'} = N_{\mathbb{K}_0'/\mathbb{K}_0}\left(\zeta^k\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) \in \mathbb{K}_0.$$

Si $\zeta \notin \mathbb{K}_0$, $k' = -k$ et

$$\zeta^k\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} + \zeta^{-k}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p} = \mathrm{Tr}_{\mathbb{K}_0'/\mathbb{K}_0}\left(\zeta^k\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) \in \mathbb{K}_0. \quad (4.58)$$

Quand $\zeta \in \mathbb{K}_0$ on définit $k'' \in \mathcal{P}$ par $2k'' = k - k' \bmod p$, et

$$\tau\left(\zeta^{k''}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) = \zeta^{k''-k}\tau\left(\zeta^{k}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) = \zeta^{k''-k}\zeta^{k'}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p} = \zeta^{-k''}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p},$$

et on obtient (4.58) avec k'' au lieu de k.

Supposons maintenant que $\left[\mathbb{K}_0'\left(\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) \,:\, \mathbb{K}_0'\right] = p$. Alors

$$\zeta^{k}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\mu^{-1/p} \in \mathbb{K}_0'$$

pour un certain $k \in \mathcal{P}$ et $\mu \in M$. Montrons alors que $\mathbb{K} = \mathbb{K}_0\left(\mu^{1/p} + \mu^{-1/p}\right)$ est admissible pour x. D'après le Lemme 4.21, (a), on a

$$\left[\mathbb{K}_0\left(\zeta^{k}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} + \zeta^{-k}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p}\right) \,:\, \mathbb{K}_0\right] = [\mathbb{K} \,:\, \mathbb{K}_0] = p.$$

Si $\zeta \notin \mathbb{K}_0$, le point (b) du même lemme montre que

$$\mathbb{K}_0\left(\zeta^{k}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} + \zeta^{-k}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p}\right) = \mathbb{K}.$$

Si ζ est dans $\mathbb{K}_0$ alors, par le point (c), le corps $\mathbb{K}$ coïncide avec l'un des p corps $\mathbb{K}_0\left(\zeta^{k'}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} + \zeta^{-k'}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p}\right)$, où $k' \in \mathcal{P}$. Ceci achève la preuve de la proposition. □

4.10.4 Une unité

Fixons alors un corps admissible $\mathbb{K}$ du système construit à l'instant, et définissons m, s, t et σ_i comme dans le premier paragraphe.

En particulier,

$$\sigma_i(\gamma) = \overline{\sigma_{i+t}(\gamma)} \quad (s+1 \le i \le s+t) \tag{4.59}$$

pour tout $\gamma \in \mathbb{K}$.

Notons $\mathbb{K}'$ le compositum $\mathbb{K}_0'\mathbb{K}$. Alors $[\mathbb{K}' \,:\, \mathbb{K}] = 2$ et il y a exactement deux prolongements de σ_i à $\mathbb{K}'$. Fixons l'un d'entre eux, que nous noterons σ_i également ; l'autre est alors $\sigma_i\tau$, où τ est l'involution non-triviale de $\mathbb{K}'/\mathbb{K}$.

Ces prolongements peuvent être choisis de sorte à satisfaire (4.59) pour $\gamma \in \mathbb{K}'$.

On peut alors définir $\alpha_i = \sigma_i(\alpha)$ et $\beta_i = \sigma_i(\beta)$.

On dira qu'un plongement réel de $\mathbb{K}$ (resp. $\mathbb{K}_0$) est *stable* si ses prolongements à $\mathbb{K}'$ (resp. $\mathbb{K}'_0$) sont égalements réels ; dans le cas contraire, le plongement sera dit *instable*.

On ordonne alors $\sigma_i, \ldots, \sigma_s$ de telle sorte que $\sigma_1, \ldots, \sigma_{s'}$ soient stables, tandis que $\sigma_{s'+1}, \ldots, \sigma_s$ sont instables.

On a alors la

Proposition 4.24. *Tout plongement réel stable de $\mathbb{K}_0$ a exactement un prolongement réel à $\mathbb{K}$, lui aussi stable, et $(p-1)/2$ couples de prolongements complexes conjugués.*

Tout plongement réel instable de $\mathbb{K}_0$ a p prolongements réels à $\mathbb{K}$, tous instables.

En particulier, $s' = s'_0$, $s - s' = p(s_0 - s'_0)$ et $t = pt_0 + \frac{p-1}{2}s'_0$, où s_0, s'_0 et $2t_0$ sont le nombre de plongements réels, réels stables et complexes de $\mathbb{K}_0$.

On note de nouveau $\mathrm{Sol}(\mathbb{K})$ l'ensemble des $x \in \mathrm{Sol}$ tels que $\mathbb{K}$ soit admissible pour x. Pour tout x élément de $\mathrm{Sol}(\mathbb{K})$, il existe $k(x) \in \mathcal{P}$ tel que $\zeta^{k(x)}\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p} \in \mathbb{K}'$. Comme les σ_i ont été prolongés à K', on peut encore définir le vecteur $\boldsymbol{k} = (k_1(x), \ldots, k_m(x))$ par

$$\sigma_i\left(\zeta^{k(x)}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) = \zeta^{k_i(x)}\left(\frac{x-\alpha_i}{x-\beta_i}\right)^{1/p}.$$

Les conditions (4.9) et (4.10) subsistent. En revanche, (4.8) doit être affaibli en

$$k_1(x) = \cdots = k_{s'}(x) = 0.$$

Le nombre de vecteurs $\boldsymbol{k}$ est alors $\left(2^{\frac{p-1}{2}}\left(\frac{p-1}{2}\right)!\right)^{s'_0}(p!)^{t_0+s_0-s'_0}$.

4.10.5 $\varphi(x)$

On définit $\varphi(x)$ comme précédemment, c'est-à-dire par

$$\varphi(x) = (x-\beta)\left(\zeta^{k(x)}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} - 1\right)^p.$$

Si $\varphi(x)$ ne vit pas dans $\mathbb{K}$, cependant, $\varphi^2(x)$, lui, vit dans $\mathbb{K}$. De fait,

$$\begin{aligned}\varphi^2(x) &= (x-\beta)\left(\zeta^{k(x)}\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} - 1\right)^p (x-\alpha)\left(\zeta^{-k(x)}\left(\frac{x-\beta}{x-\alpha}\right)^{1/p} - 1\right)^p \\ &= N_{\mathbb{K}'/\mathbb{K}}(\varphi(x)) \in \mathbb{K}.\end{aligned}$$

Il suffit alors d'appliquer le même type de techniques que dans les sections qui précèdent, en remplaçant $\varphi(x)$ par $\varphi(x)^2$, et en modifiant les constantes correspondantes.

Voici la liste des modifications à apporter :

1. Remplacer respectivement $\varphi(x)$, $\varphi_i(x)$, $\Phi(x)$ par $\varphi^2(x)$, $\varphi_i^2(x)$, $\Phi^2(x)$ dans les équations (4.15), (4.22), (4.23), (4.24), (4.26), (4.38), (4.41), (4.45), (4.47), (4.49), (4.50).
2. Modifier $u_1(\mathfrak{P})$ et $u_2(\mathfrak{P})$ comme suit : $u_1(\mathfrak{P}) = \max(0, -\mathrm{Ord}_{\mathfrak{P}}(\alpha\beta))$ et $u_2(\mathfrak{P}) = \max\left(0, \mathrm{Ord}_{\mathfrak{P}}\left((\alpha-\beta)^2\right)\right)$.
3. Dans (4.24) et (4.25), remplacer $(\beta-\alpha)^p$ par $(\beta-\alpha)^{2p}$.
4. Modifier δ_i et λ_i comme suit :

$$\delta_i = 2\sum_{j=1}^{r} a_{ij}\rho_j \quad \text{et} \quad \lambda_i = \sum_{j=1}^{r} \left(a_{ij} \log|\gamma_j^2\theta_j^{-1}|\right).$$

5. Modifier les constantes suivantes :

$$\begin{aligned} &c_{10} = 2c_7 \max_{1\le i\le r} \sum_{j=1}^{r} |a_{ij}|, \quad c_{14} = \max\left(1 + 2\pi^{-1}c_{12} + rc_9, e\right), \\ &c_{16} = 2c_{13}c_{12} + c_{14}, \\ &X_4 = \max\left(X_2,\ 3mc_7,\ e^{1/(3|\rho'|)} \left|\gamma_1\cdots\gamma_m\right|^{-1/\rho'} \left|\theta_1\cdots\theta_m\right|^{1/(2\rho')}\right). \end{aligned}$$

6. Dans la section 4.7, il suffit maintenant de supposer que $[\mathbb{K} : \mathbb{Q}] \ge 2$. Pour expliquer ceci, rappelons que l'hypothèse $[\mathbb{K} : \mathbb{Q}] \ge 3$ servait dans le cas $\mathbb{K} = \mathbb{K}_0$ à garantir l'existence de i_1 et i_2 tels que parmi les nombres (4.40) il y en ait au moins trois distincts.

   Dans le cas de l'(α, β)-symétrie, il suffit d'avoir $[\mathbb{K}_0 : \mathbb{Q}] \ge 2$, parce que pour tout couple (i_1, i_2), il y a au moins trois nombres distincts parmi les nombres de (4.40). (Sinon, on aurait $\alpha_{i_1} = \beta_{i_1}$ et $\alpha_{i_2} = \beta_{i_2}$, ou $\alpha_{i_1} = \beta_{i_2}$ et $\alpha_{i_2} = \beta_{i_1}$. Dans tous les cas, $\alpha_{i_1} + \alpha_{i_2} = \beta_{i_1} + \beta_{i_2}$ et $\alpha_{i_1}\alpha_{i_2} = \beta_{i_1}\beta_{i_2}$, c'est-à-dire que $\sigma_{i_1} = \sigma_{i_2}$.)
7. Définir ϑ_0 comme $\vartheta_0 = (\gamma_{i_2}^{2\rho_{i_1}}\theta_{i_1}^{\rho_{i_2}})/(\gamma_{i_1}^{2\rho_{i_2}}\theta_{i_2}^{\rho_{i_1}})$.
8. Dans (4.49), (4.51), et (4.52), changer c_{12} en $2c_{12}$.
9. Réécrire (4.56) en

$$\omega_i := \pm\gamma_i^{-1}\left(\theta_i\eta_{i1}^{b_1}\cdots\eta_{ir}^{b_r}\right)^{1/2} - \gamma_i', \tag{4.60}$$

   ce qui donne deux possibilités pour ω_i, qui doivent toutes deux être prises en compte.

4.11 Une remarque conclusive

Dans le cas où l'on a (α, β)-symétrie, on est amené à travailler dans des corps de degré $pn(n-1)/2$.

En fait, on peut montrer que si les corps admissibles sont de degré $pn(n-1)$ (c'est-à-dire si $\mathbb{Q}(\alpha, \beta)$ est de degré $n(n-1)$, il y a toujours symétrie. Par suite, le cas le pire est le cas $\deg(\mathbb{K}) = n(n-1)/2$.

Proposition 4.25. *Soit $f(x) \in \mathbb{Q}(x)$ ayant au moins deux racines distinctes. Alors,*

(i) soit il existe deux racines distinctes α, β de f avec (α, β)-symétrie,

(ii) soit il existe deux racines distinctes α, β de f telles que $[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] \leq n(n-1)/2$, où $n = \deg f$.

Preuve. Quitte à remplacer f par un de ses facteurs irréductibles, on peut supposer f irréductible, sauf dans le cas où f est scindé sur $\mathbb{Q}$, cas dans lequel la proposition est triviale.

Soit G le groupe de Galois de f sur $\mathbb{Q}$.

Si $|G|$ est pair, il existe $\tau \in G$ d'ordre 2, qui donne l'(α, β)-symétrie.

Si $|G|$ est impair, fixons une racine α de f. Alors $g(x) = f(x)/(x - \alpha)$ est réductible sur $\mathbb{Q}(\alpha)$, sans quoi $n(n-1) \,|\, |G|$, impossible. Soit alors g_1 un facteur de g irréductible sur $\mathbb{Q}(\alpha)$ de degré minimal ; on a $\deg(g_1) \leq (n-1)/2$, et pour toute racine β de g_1, on a alors le deuxième point.

Ceci conclut la preuve. □

4.12 Le cas $p = 2$

On expose ici succinctement les modifications à effectuer dans le cas $p = 2$.

4.12.1 Quelques modifications

On fixe de nouveau α et β deux racines distinctes de f, et on suppose d'abord que $[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] \geq 3$. Il est facile de voir dans ce cas que la méthode de la section 4.5 continue à s'appliquer. En effet, nous n'avons utilisé le fait que $p \geq 3$ qu'à deux endroits ; dans la Proposition 4.8, et dans la section 4.7. Seul le second point est important ; on ne peut effectivement plus satisfaire (4.39) quand $p = 2$. Ceci dit, comme $[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] \geq 3$, on peut toujours trouver i_1 et i_2 tels que trois des nombres (4.40) soient distincts, ce qui permet de calculer la borne de Baker.

Dans le cas $[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] \leq 2]$, il nous faut utiliser une racine supplémentaire γ de f. On peut supposer que $[\mathbb{Q}(\alpha, \gamma) : \mathbb{Q}] \leq 2$ et que $[\mathbb{Q}(\beta, \gamma) : \mathbb{Q}] \leq 2]$, sans quoi en réarrangeant les racines, on revient au cas $[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] \geq 3$.

Par suite, $[\mathbb{Q}(\alpha,\beta,\gamma) : \mathbb{Q}] \leq 2$. En réarrangeant les racines, on peut supposer en plus que

(i) soit $\alpha,\beta,\gamma \in \mathbb{Q}$,

(ii) soit α et β sont quadratiques et conjugués sur $\mathbb{Q}$, et $\gamma \in \mathbb{Q}(\alpha)$.

Le premier cas est simple ; on peut le réduire à deux équations de Pell simultanées, voir par exemple [Za87]. Dans le second cas, faisons $\mathbb{K}_0 = \mathbb{Q}(\alpha)$, et disons qu'un corps de nombres est admissible pour une solution x si $\mathbb{K} = \mathbb{K}_0\left(\left(\frac{x-\alpha}{x-\gamma}\right)^{1/2}\right)$. On construit facilement un système complet de corps admissibles, qui comprend $\mathbb{K}_0$ et un certain nombre d'extensions quadratiques de $\mathbb{K}_0$.

Fixons maintenant un de ces corps admissibles, et posons

$$\varphi(x) = (x-\gamma)\left(\left(\frac{x-\alpha}{x-\gamma}\right)^{1/2} + 1\right)^2.$$

Alors on a de nouveau (4.26), où $\theta(x)$ appartient à un ensemble fini effectivement constructible.

Si $\mathbb{K} = \mathbb{K}_0$, on n'a qu'un nombre fini de possiblités pour $\phi(x)\sigma(\phi(x))$, où σ est l'automorphisme non-trivial de $\mathbb{K}_0$. Cependant, $\sigma\left(\left(\frac{x-\alpha}{x-\gamma}\right)^{1/2}\right) = \pm\left(\frac{x-\beta}{x-\gamma'}\right)^{1/2}$, where $\gamma' = \sigma(\gamma)$. Utilisant la théorie de Kummer, on prouve que

$$\varphi(x)\sigma(\varphi(x)) = (x-\gamma)\left(\left(\frac{x-\alpha}{x-\gamma}\right)^{1/2} + 1\right)^2 (x-\gamma')\left(\pm\left(\frac{x-\beta}{x-\gamma'}\right)^{1/2} + 1\right)^2$$

est non-constante, ce qui permet de calculer l'ensemble $\mathrm{Sol}(\mathbb{K}_0)$.

Si $[\mathbb{K} : \mathbb{K}_0] = 2$, on pose

$$\Phi(x) = \frac{(x-\gamma)\left(\left(\frac{x-\alpha}{x-\gamma}\right)^{1/2} + 1\right)^2}{(x-\gamma')\left(\left(\frac{x-\beta}{x-\gamma'}\right)^{1/2} + 1\right)^2},$$

et on peut alors calculer la borne de Baker et continuer comme dans la section 4.5. Voir par exemple [We94a, We94b] pour un algorithme similaire dans le cas $\deg f = 3$.

4.12.2 (α,β)-symétrie

La construction de l'(α,β)-symétrie est similaire ; il est toutefois nécessaire d'adapter un peu les preuves.

Lemme 4.26. *Soit $\mathbb{K}_0 \subset \mathbb{K}_1 \subset \mathbb{K}_2$ une tour de corps de nombres vérifiant les propriétés suivantes*

$[\mathbb{K}_1 : \mathbb{K}_0] = 2$ *et* $[\mathbb{K}_2 : \mathbb{K}_1] = 2$.

$\mathbb{K}_2 = \mathbb{K}_1(\nu)$, *où* $\nu^2 = \mu \in \mathbb{K}_1$ *et* μ *est conjugué à* μ^{-1} *sur* $\mathbb{K}_0$, $\mu \neq -1$.

Alors

(a) $[\mathbb{K}_0(\nu + \nu^{-1}) : \mathbb{K}_0] = 2$.

Soit $\mathbb{K}$ un corps de nombres intermédiaire entre $\mathbb{K}_0$ et $\mathbb{K}_2$, tel que $[\mathbb{K} : \mathbb{K}_0] = 2$, et $\mathbb{K} \neq \mathbb{K}_1$. Alors

(b) $\mathbb{K}$ est l'un des 2 corps distincts $\mathbb{K}_0(\nu \pm \nu^{-1})$.

(c) Le compositum $\mathbb{K}_1\mathbb{K}$ est $\mathbb{K}_2$.

Preuve. Considérons les nombres

$$\nu, -\nu, 1/\nu, -1/\nu. \tag{4.61}$$

qui sont des conjugués de ν sur $\mathbb{K}_0$. Comme $[\mathbb{K}_2 : \mathbb{K}_0] = 4$, c'est la liste complète des conjugués de ν dans $\mathbb{K}_2$, et l'extension $\mathbb{K}_2/\mathbb{K}_0$ est normale. Le groupe $G := \mathrm{Gal}(\mathbb{K}_2/\mathbb{K}_0)$ est abélien engendré par $\theta\colon \nu \to \nu^{-1}$ et $\sigma\colon \nu \to -\nu$. Comme θ et σ sont deux involutions non triviales, $G = (\mathbb{Z}/2\mathbb{Z})^2$. Il y a donc trois sous-corps de $\mathbb{K}_2$, l'un correspondant à σ, qui est $\mathbb{K}_0(\nu^2) = \mathbb{K}_1$, le corps $\mathbb{K}_0(\nu+1/\nu)$ (correspondant à θ) et le corps $\mathbb{K}_0(\nu - 1/\nu)$ qui correspond à $\sigma\theta$. Les trois assertions du lemme s'ensuivent immédiatement. □

On va maintenant étudier les corps admissibles. Remarquons que $x \geq X_1$ permet d'identifier $\left(\dfrac{x-\beta}{x-\alpha}\right)^{1/p}$ et $\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p^{-1}}$.

Soit $\tau : \mathbb{Q}(\alpha,\beta) \to \mathbb{Q}(\alpha,\beta)$ donné par $\tau(\alpha) = \beta$, $\tau(\beta) = \alpha$. On définit $K_0 = \mathbb{Q}(\alpha+\beta, \alpha\beta)$, $K_0' = \mathbb{Q}(\alpha,\beta)$, c'est-à-dire que $\mathbb{K}_0 = \mathbb{K}_0'^{\tau}$.

Définition 4.27. *Un corps de nombres est admissible pour une solution x si on a* $\mathbb{K} = \mathbb{K}_0\left(\left(\dfrac{x-\alpha}{x-\beta}\right)^{1/p} + \left(\dfrac{x-\beta}{x-\alpha}\right)^{1/p}\right)$.

On définit de la même manière que dans le cas général les systèmes complets et systèmes complets minimaux de corps admissibles.

Soit M l'ensemble fini de $\mathbb{Q}(\alpha,\beta)$ construit précedemment.

Proposition 4.28. *Les corps $\mathbb{K}_0'$ et $\mathbb{K}_0\left(\sqrt{\mu} + \dfrac{1}{\sqrt{\mu}}\right)$, où μ décrit M, forment un système complet de corps admissibles.*

Preuve. Supposons tout d'abord que $\left(\frac{x-\alpha}{x-\beta}\right)^{1/p} \in \mathbb{K}_0'$. Dans ce cas $\mathbb{K}_0'$ est admissible pour x.

Supposons maintenant que $\left[\mathbb{K}_0'\left(\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right) \ : \ \mathbb{K}_0'\right] = p$. Alors

$$\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\frac{1}{\sqrt{\mu}} \in \mathbb{K}_0'$$

pour un certain $\mu \in M$.

Montrons alors que $\mathbb{K} = \mathbb{K}_0\left(\sqrt{\mu}+\frac{1}{\sqrt{\mu}}\right)$ est admissible pour x.

L'hypothèse $|x| \geq X_1$ interdit à $\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}$ d'être égal à i. Par suite, d'après le lemme, (a), on a

$$\left[\mathbb{K}_0\left(\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}+\left(\frac{x-\beta}{x-\alpha}\right)^{1/p}\right) \ : \ \mathbb{K}_0\right] = [\mathbb{K} \ : \ \mathbb{K}_0] = 2.$$

On est dans la situation du lemme avec $\mathbb{K}_1 = \mathbb{K}_0'$, $\mathbb{K}_2 = \mathbb{K}_1\left(\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\right)$. Étudions l'action de $\theta : \left(\frac{x-\alpha}{x-\beta}\right)^{1/p} \mapsto \left(\frac{x-\beta}{x-\alpha}\right)^{1/p}$ et de $\sigma : \left(\frac{x-\alpha}{x-\beta}\right)^{1/p} \mapsto -\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}$ (les deux fixant $\mathbb{K}_0$) sur $\sqrt{\mu}$.

Par construction, voir la Proposition 4.4, on sait que

$$\frac{x-\alpha}{x-\beta} = \mu\left(\frac{\lambda}{\tau(\lambda)}\right)^2,$$

Par suite,

$$\sqrt{\mu} = \pm\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}\frac{\tau(\lambda)}{\lambda},$$

Comme la restriction de θ à $\mathbb{K}_0'$ coïncide avec τ, on a alors

$$\theta(\sqrt{\mu}) = \frac{1}{\sqrt{\mu}}, \quad \sigma(\sqrt{\mu}) = -\sqrt{\mu}.$$

On vérifie alors facilement que $\mathbb{K} = \mathbb{K}_2^{\theta} = \mathbb{K}_0\left(\left(\frac{x-\alpha}{x-\beta}\right)^{1/p}+\left(\frac{x-\beta}{x-\alpha}\right)^{1/p}\right)$. Ceci achève la preuve de la proposition. □

Chapitre 5

Équations superelliptiques ; exemples et applications

Dans ce chapitre, on va mettre en œuvre les techniques décrites au chapitre précédent pour résoudre diverses équations diophantiennes.

D'autres illustrations peuvent être trouvées dans des travaux de de Weger [We94a, We94b].

5.1 Zéros entiers des polynômes de Krawtchouk

Le problème traité dans cette section m'a été suggéré par Laurent Habsieger, que je tiens à remercier.

Définition 5.1. *Le $k^{\text{ème}}$ polynôme de Krawtchouk binaire est le coefficient de z^k dans le développement de la fonction*

$$(1-z)^x(1+z)^{n-x}.$$

En règle générale, n est fixé. Toutefois dans ce chapitre, on fixera $k = 4$ ou 5, et on fera varier x et n.

Les polynômes de Krawtchouk et leurs zéros (plus précisément, le fait qu'ils soient entiers ou non) interviennent dans de nombreux problèmes de combinatoire et de théorie des codes. Un descriptif de ces problèmes, ainsi qu'une bibliographie exhaustive se trouvent dans [KL96].

Ces polynômes peuvent facilement être calculés à l'aide des relations de récurrence suivantes :

$$(k+1)P_{k+1}^n(x) = (n-2x)P_k^n(x) - (n-k+1)P_n^{k-1}(x),$$

avec $P_0(x) = 1$, $P_1(x) = n - 2x$.

On va ici s'intéresser à la conjecture suivante sur les zéros de P_4^n et de P_5^n, énoncée dans [DG85] :

Conjecture 5.2.

1. *Les solutions* (n, x), $0 \leq x < n/2$ *de* $P_4^n(x) = 0$ *sont*

$$(1,0),(2,0),(2,1),(3,0),(3,1),(8,1),(8,3),(17,7),(66,30),\\(1521,715),(15043,7476).$$

2. *Les solutions* (n, x), $0 \leq x < n/2$ *de* $P_5^n(x) = 0$ *sont*

$$(1,0),(2,0),(3,0),(3,1),(4,0),(4,1),(10,1),(10,3),(17,3),(36,14),\\(67,22),(67,28),(289,133),(10882,5292).$$

Remarquons que la première partie de cette conjecture est une conséquence du travail de Stroeker et de Weger [SW96].

Pour résoudre la deuxième équation, nous allons réduire celle-ci à une équation hyperelliptique, que nous résoudrons suivant la méthode du chapitre 4. Au cours de la réduction à une équation hyperelliptique, nous remarquerons l'existence d'automorphismes non-triviaux sur la courbe $P_5^n(x) = 0$ qui montreront que la liste de zéros donnés dans la deuxième conjecture est incomplète : nous prouvons le

Théorème 5.3. *Les solutions* (n, x), $0 \leq x < n/2$ *de* $P_5^n(x) = 0$ *sont*

$$(1,0),(2,0),(3,0),(3,1),(4,0),(4,1),(10,1),(10,3),(17,3),(36,14),(67,22),\\(67,28),(289,133),(10882,5292),(48324,24013).$$

5.1.1 Réduction à des équations hyperelliptiques

En utilisant les relations de récurrence données plus haut, il vient

$$P_n^4(x) = \frac{1}{6}\left((n-2x)^4 + (n-2x)^2(8-6n) + 3n^2 - 6n\right),$$
$$P_n^5(x) = \frac{n-2x}{24}\left((n-2x)^4 + (n-2x)^2(20-10n) + 15n^2 - 50n + 24\right).$$

On pose alors $U = (n-2x)$ et $V = n$; en réarrangeant les termes, on obtient :

$$3V^2 - 6V(U^2+1) + U^4 + 8U^2 = 0,$$
$$15V^2 - 10V(U^2+5) + U^4 + 20U + 24 = 0.$$

On complète maintenant les carrés, et on s'arrange pour avoir un polynôme f unitaire ; on trouve alors les deux équations suivantes :

$$6Y^2 = X^4 - 4X^2 + 24, \qquad 10Y^2 = X^4 - 20X^2 + 424. \tag{5.1}$$

La correspondance entre les racines du quatrième polynôme et la première équation est

$$\begin{array}{ccc} & (n,x) & \mapsto \quad (2(n-2x), 2(n-1-(n-2x)^2)) \\ \left(\frac{2Y+X^2+4}{4}, \frac{2Y+X^2-2X+4}{8}\right) & \leftarrow & (X,Y) \end{array} \tag{5.2}$$

tandis que la correspondance pour le cinquième polynôme est

$$\begin{array}{ccc} & (n,x) & \mapsto \quad \big(2(n-2x), 2(3n-(n-2x)^2-5)\big) \\ \left(\frac{2Y+X^2+20}{12}, \frac{2Y+X^2-6X+20}{24}\right) & \leftarrow & (X,Y) \end{array} \tag{5.3}$$

On voit bien grâce aux expressions ci-dessus que la connaissance des solutions entières de (5.1) permet de connaître les solutions entières de $P_i^n(x)$.

5.1.2 La racine manquante

On peut maintenant remarquer que si le seul automorphisme évident de la courbe de départ était $(n,x) \to (n-x,x)$, il n'en est pas de même de la courbe associée, qui a les automorphismes $(X,Y) \to (\pm X, \pm Y)$. En particulier, pour le cinquième polynôme, l'automorphisme involutif $(X,Y) \to (X,-Y)$ de la courbe elliptique correspondante donne l'automorphisme involutif

$$(n,x) \to \left(\frac{2(n-2x)^2-3n+10}{3}, \frac{2(n-2x)^2+6x-6n+10}{6}\right),$$

qui une fois appliqué à $(10882, 5292)$ donne $(48324, 24013)$.

Les racines de ces polynômes vérifiant $0 \leq x < n/2$ sont couplées de la manière suivante :

$$\{(0,0),(2,1)\},\{(1,0),(3,1)\},\{(2,0),(8,3)\},\{(17,7),(3,0)\},$$
$$\{(66,30),(8,1)\},\{(1521,715),(15043,7476)\}$$

pour le quatrième polynôme,

$$\{(1,0),(3,1)\},\{(2,0),(4,1)\},\{(3,0),(19/3,5/3)\},\{(4,0),(10,3)\},\{(17,3),(67,28)\},$$
$$\{(36,14),(10,1)\},\{(67,22),(289,133)\},\{(10882,5292),(48324,24013)\}$$

pour le cinquième.

Remarquons qu'on a ajouté $(0,0)$ à la liste pour le quatrième polynôme et une racine non-entière pour le cinquième, afin de compléter les couples.

5.1.3 Résolution par la méthode "alternative"

Le nom de "méthode alternative" a été attribué par de Weger[We94a, We94b] à la méthode du chapitre 4, parce qu'elle offre une alternative aux méthodes de Thue et des courbes elliptiques.

On ne donne ici que quelques détails sur la résolution par la méthode alternative ; des compléments numériques se trouvent en annexe B, et la totalité des données numériques sont disponibles sur demande. On n'a pas explicitement exploité l'(α, β)-symétrie, à part lors de la détermination de l'ensemble Θ, comme mentionné plus bas.

5.1.3.1 Corps admissibles

L'ensemble $\Xi(\alpha)$ comporte 1024 éléments, et donc aussi l'ensemble Ξ. On élimine de $\Xi(\alpha)$ les éléments dont la norme n'est pas 10 fois un carré ; on identifie également dans Ξ deux éléments dont le quotient est un carré de $\mathbb{K}_0'$, puisqu'ils donnent naissance à des corps isomorphes. On obtient ainsi un ensemble $\tilde{\Xi}$ qui comporte 12 éléments.

Les corps admissibles sont au nombre de 5, $\mathbb{K}_0'$ et les 4 corps dont des générateurs sont donnés par les zéros des polynômes :

$$X^8 + 4X^6 + 50X^4 + 604X^2 + 1681, X^8 + 10X^6 + 26X^4 + 16X^2 + 64,$$
$$X^8 - 18X^6 + 74X^4 + 72X^2 + 16, X^8 + 16X^6 + 390X^4 + 16X^2 + 1.$$

5.1.3.2 Constantes "uniformes"

On donne ici les constantes qui ne dépendent pas du corps admissible choisi.

$$\begin{array}{cccc} c_1 = 4,6 & X_1 = 13 & c_2 = 0,18 & c_3 = 0,59 \\ c_4 = 2 & c_5 = 9,1 & c_6 = 4,6 & c_7 = 7,4 \\ X_2 = 13 & c_{12} = 29,6 & c_{19} = 1 & c_{20} = 3,6 \\ c_{21} = 0,02. & & & \end{array}$$

Dans la suite, les expressions de la forme $[a_0, a_1, \ldots, a_n]$ représentent un élément du corps considéré exprimé sur la base d'entiers donnée pour ce corps.

5.1.3.3 Le corps $\mathbb{K}_0$

On note toujours α une racine de $X^4 - 20X^2 + 424$. Le corps $\mathbb{K}_0$ a pour discriminant 27136, une base d'entiers est donnée par $[1, \alpha, \alpha^2/18 + 4/9, \alpha^3/36 + 2\alpha/9]$.

Le groupe des unités est de rang 1, engendré par $[17, -2, -14, 6]$. La torsion est d'ordre 4 et $i = [1, 0, -1, 0]$. Le groupe des classes est isomorphe à $\mathbb{Z}/3\mathbb{Z}$.

Comme $r = 1$, on peut obtenir de manière systématique une borne sur les éléments de $\mathrm{Sol}(\mathbb{K}_0)$ pour chacun des éléments de Θ ; on obtient $|x| \leq 189$, et la vérification des petites valeurs correspondante est englobée dans la vérification finale.

5.1.3.4 Le corps $\mathbb{K}_1$

Notons α_1 une racine du polynôme $X^8 + 4X^6 + 50X^4 + 604X^2 + 1681$. Le corps $\mathbb{K}_1 = \mathbb{Q}(\alpha_1)$ a pour discriminant 294544998400. Voici les coefficients d'une base d'entiers sur la base de puissances $(\alpha_1^k)_{0\leq k\leq 7}$:

	1	α_1	α_1^2	α_1^3	α_1^4	α_1^5	α_1^6	α_1^7
τ_1	1							
τ_2	0	1						
τ_3	0	0	1					
τ_4	0	0	0	1				
τ_5	−1/16	0	−3/8	0	1/16			
τ_6	0	−1/16	0	−3/8	0	1/16		
τ_7	−37/48	0	1/48	0	−1/48	0	1/48	
τ_8	−37/96	−2881/3936	1/96	−647/3936	−1/96	−37/3936	1/96	1/3936

Le groupe des unités est de rang 3, engendré par les éléments

$$[-9, 11, -2, 3, -1, 2, -8, 14], [59, 23, 8, 7, 1, 4, -8, 30], [0, 2, 0, 1, 0, 1, -1, 2].$$

La torsion est engendrée par $i = [8, 0, 1, 0, 0, 0, 1, 0]$. Le groupe des classes est isomorphe à $\mathbb{Z}/6\mathbb{Z}$.

La constante c_8, indépendante de θ, vaut $0,52$. On a par ailleurs $c_{10} \leq 11,9$.

On a $(\beta - \alpha)\mathbb{Z}_{\mathbb{K}_1} = \mathfrak{p}^{14}\mathfrak{q}_1\mathfrak{q}_2$, où $\mathfrak{p}$ est l'idéal premier au-dessus de 2 et $\mathfrak{q}_1$ et $\mathfrak{q}_2$ sont deux des quatre idéaux au-dessus de 53.

L'ensemble Θ_0 est donc composé de 3 éléments, à savoir les générateurs des idéaux $\mathfrak{p}^{14}\mathfrak{q}_1^a\mathfrak{q}_2^{2-a}$, $a = 0, 1, 2$; pour chacun de ces θ il y a quatre vecteurs $\boldsymbol{k}$ possibles.

On donne les pires valeurs des principales constantes dans chaque cas en annexe B.

L'ensemble $\mathrm{Sol}(\mathbb{K}_1)$ est $\{\pm 46\}$. La constante X_6 relative à $\mathbb{K}_1$ vaut 238.

5.1.3.5 Le corps $\mathbb{K}_2$

Notons α_2 une racine du polynôme $X^8 + 10X^6 + 26X^4 + 16X^2 + 64$. Le discriminant du corps $\mathbb{K}_2 = \mathbb{Q}(\alpha_2)$ vaut 18409062400. Voici les coefficients d'une base d'entiers sur la base de puissances $(\alpha_2^k)_{0\leq k\leq 7}$:

	1	α_2	α_2^2	α_2^3	α_2^4	α_2^5	α_2^6	α_2^7
τ_1	1							
τ_2	0	1						
τ_3	0	0	1					
τ_4	0	0	0	1				
τ_5	0	0	0	0	1/2			
τ_6	0	1/2	0	1/2	0	1/4		
τ_7	1/3	0	5/12	0	1/12	0	1/24	
τ_8	0	11/3	0	5/24	0	1/24	0	1/48

Le groupe des unités est de rang 3, engendré par les éléments

$$[4,-4,3,-2,2,0,3,2],[-1,4,2,-2,-2,2,-6,8],[4,6,6,2,4,3,5,3].$$

La torsion est engendrée par $i=[0,0,1,0,0,0,-1,0]$. Le groupe des classes est isomorphe à $\mathbb{Z}/3\mathbb{Z}$.

On a $(\beta-\alpha)\mathbb{Z}_{\mathbb{K}_2}=\mathfrak{p}_1^7\mathfrak{p}_2^7\mathfrak{q}_1\mathfrak{q}_2$, où $\mathfrak{p}_1,\mathfrak{p}_2$ sont deux des idéaux premiers au-dessus de 2, $\mathfrak{q}_1$ et $\mathfrak{q}_2$ sont deux des quatre idéaux au-dessus de 53. Il y a donc 45 éléments θ possibles, qui sont les générateurs des idéaux $\mathfrak{p}_1^a\mathfrak{p}_2^{14-a}\mathfrak{q}_1^b\mathfrak{q}_2^{2-b}$, $a=0,\dots,14; b=0,1,2$.

On peut toutefois considérer l'opération des automorphismes au-dessus du plongement de $\mathbb{K}_0$ envoyant α sur β. Deux automorphismes prolongent ce dernier, qui sont $\sqrt{\xi/\sigma(\xi)}\mapsto\pm1/\sqrt{\xi/\sigma(\xi)}$. Celui des deux correspondant au signe + doit laisser invariant $\varphi(x)$ (on le voit, par exemple, dans la preuve de la Proposition 4.28). Notons cet automorphisme σ_1. On a alors $\sigma_1(\mathfrak{p}_1)=\mathfrak{p}_2$ et $\sigma_1(\mathfrak{p}_2)=\mathfrak{p}_1$, ce qui impose $a=7$. Par contre, les deux idéaux au-dessus de 53 sont invariants par σ_1. Il y a donc bien exactement 3 éléments θ à considérer. Notons que l'on vient, de manière "cachée", d'utiliser l'(α,β)-symétrie pour réduire l'ensemble Θ.

La constante c_8 (indépendante de θ) vérifie $c_8\leq 0,51$, et on a $c_{10}\leq 15,4$.

L'ensemble $\mathrm{Sol}(\mathbb{K}_2)$ est vide ; le nombre X_6 relatif à $\mathbb{K}_2$ vaut 308.

5.1.3.6 Le corps $\mathbb{K}_3$

Notons α_3 une racine du polynôme $X^8-18X^6+74X^4+72X^2+16$. Le discriminant du corps $\mathbb{K}_3=\mathbb{Q}(\alpha_3)$ vaut 1491134054400. Voici les coefficients d'une base d'entiers sur la base de puissances $(\alpha_3^k)_{0\leq k\leq 7}$:

	1	α_3	α_3^2	α_3^3	α_3^4	α_3^5	α_3^6	α_3^7
τ_1	1							
τ_2	0	1						
τ_3	0	0	1					
τ_4	0	0	0	1				
τ_5	0	0	0	0	1/2			
τ_6	0	0	0	0	0	1/2		
τ_7	0	0	1/2	0	0	0	1/4	
τ_8	0	0	0	1/4	0	1/4	0	1/8

Le groupe des unités est de rang 3, engendré par les éléments

$$[-3,86,-19,180,9,-92,-1,18],[-129,10,-266,68,126,-32,-14,6],$$
$$[1,0,4,-1,5,-3,-1,1].$$

La torsion est engendrée par $i=[9,0,19,0,-9,0,1,0]$. Le groupe des classes est isomorphe à $\mathbb{Z}/6\mathbb{Z}$.

On a $(\beta-\alpha)\mathbb{Z}_{\mathbb{K}_3}=\mathfrak{p}_1^7\mathfrak{p}_2^7\mathfrak{q}_1\mathfrak{q}_2$, où $\mathfrak{p}_1,\mathfrak{p}_2$ sont deux des idéaux premiers au-dessus de 2, $\mathfrak{q}_1$ et $\mathfrak{q}_2$ sont deux des quatre idéaux au-dessus de 53. Il y a donc

45 éléments θ possibles, qui sont les générateurs des idéaux $\mathfrak{p}_1^a\mathfrak{p}_2^{14-a}\mathfrak{q}_1^b\mathfrak{q}_2^{2-b}$, $a = 0,\dots,14; b = 0,1,2$.

Un raisonnement analogue à celui fait pour le corps $\mathbb{K}_2$ montre que l'on peut en fait se limiter au cas $a = 7$.

La constante c_8 (indépendante de θ) vérifie $c_8 \leq 0,4$, et on a $c_{10} \leq 5,91$. On a $\mathrm{Sol}(\mathbb{K}_3) = \{\pm 16, \pm 596\}$, et $X_6 = 118$.

5.1.3.7 Le corps $\mathbb{K}_4$

Notons α_4 une racine du polynôme $X^8 + 16X^6 + 390X^4 + 16X^2 + 1$. Le discriminant du corps $\mathbb{K}_4 = \mathbb{Q}(\alpha_4)$ vaut 23858144870400. Voici les coefficients d'une base d'entiers sur la base de puissances $(\alpha_4^k)_{0\leq k\leq 7}$:

	1	α_4	α_4^2	α_4^3	α_4^4	α_4^5	α_4^6	α_4^7
τ_1	1							
τ_2	0	1						
τ_3	0	0	1					
τ_4	1/2	1/2	1/2	1/2				
τ_5	1/18	0	4/9	0	1/18			
τ_6	−1/2	−4/9	−1/2	−1/18	0	1/18		
τ_7	11/36	0	−1/36	0	1/36	0	1/36	
τ_8	0	11/36	0	−1/36	0	1/36	0	1/36

Le groupe des unités est de rang 3, engendré par les éléments

$$[0,1,0,0,0,0,0,0], [-383,-412,-361,1126,15,374,2,50],$$
$$[-665,-766,-631,2024,45,672,6,90].$$

La torsion est engendrée par $i = [1,0,-15,0,-15,0,-2,0]$. Le groupe des classes est isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/12\mathbb{Z}$.

On a $(\beta-\alpha)\mathbb{Z}_{\mathbb{K}_4} = \mathfrak{p}^{14}\mathfrak{q}_1\mathfrak{q}_2$, où $\mathfrak{p}$ est l'idéal premier au-dessus de 2, $\mathfrak{q}_1$ et $\mathfrak{q}_2$ sont deux des quatre idéaux au-dessus de 53. Il y a donc 3 éléments θ possibles, qui sont les générateurs des idéaux $\mathfrak{p}^{14}\mathfrak{q}_1^a\mathfrak{q}_2^{2-a}$, $a = 0,1,2$.

On a $c_8 \leq 0,7$, et on a $c_{10} \leq 9,92$. L'ensemble $\mathrm{Sol}(\mathbb{K}_4)$ vaut $\{\pm 2, \pm 22\}$, et $X_6 = 306$.

5.1.3.8 Conclusion

Un examen des $|x| \leq 308$ montre que les solutions manquantes sont $(\pm 4, \pm 6)$, $(\pm 6, \pm 10), (\pm 8, \pm 18)$.

On a donc le

Théorème 5.4. *Les solutions de l'équation* $10y^2 = x^4 - 20x^2 + 424$ *sont* $(\pm 2, \pm 6)$, $(\pm 4, \pm 6)$, $(\pm 6, \pm 10)$, $(\pm 8, \pm 18)$, $(\pm 22, \pm 150)$, $(\pm 16, \pm 78)$, $(\pm 46, \pm 666)$, $(\pm 596, \pm 112326)$.

Au vu des formules (5.3), le Théorème 5.3 est donc prouvé.

Notons que ces 3 solutions manquantes auraient pu être trouvées respectivement dans $\mathrm{Sol}(\mathbb{K}_3)$, $\mathrm{Sol}(\mathbb{K}_1)$, $\mathrm{Sol}(\mathbb{K}_2)$.

5.1.4 Résolution par la méthode des courbes elliptiques

5.1.4.1 Introduction

Comme je l'ai déjà mentionné plus haut, la méthode des courbes elliptiques consiste à voir l'équation diophantienne correspondante comme l'équation d'une courbe elliptique E, et à écrire un point entier donné de E comme combinaison linéaire de générateurs du groupe $E(\mathbb{Q})$ qui jouent le rôle des unités fondamentales dans les méthodes décrites dans cette thèse.

Je n'expose pas les notions de base concernant l'arithmétique des courbes elliptiques, voir par exemple [Si].

Étant donné une courbe elliptique définie sur $\mathbb{C}$, on dispose d'un paramétrage de cette courbe par $\mathbb{C}/\Lambda$, où Λ est un certain réseau de $\mathbb{Z}^2$, donné par

$$\begin{array}{rccc} F: & \mathbb{C}/\Lambda & \to & E(\mathbb{C}) \subset \mathbb{P}^2(\mathbb{C}) \\ & z & \mapsto & [\wp(z), \wp'(z), 1] \end{array}.$$

Ce paramétrage est de plus un morphisme de groupes, c'est-à-dire qu'il va nous servir de fonction exponentielle ; sa fonction réciproque est donnée par une intégrale elliptique[1]. Si l'on écrit un modèle de Weierstraß de la courbe comme $y^2 = f(x)$, avec f de degré 3, cette réciproque est donnée par

$$F^*(P) = \int_O^P \frac{dz}{\sqrt{f(z)}} \pmod{\Lambda},$$

qui, convenablement normalisée, va nous servir de fonction "logarithme".

Pour ce type de fonctions "logarithme", on dispose de minorations de formes linéaires en logarithmes, voir [Da95], et même de formes en logarithmes p-adiques dans le cas des courbes de rang 1, voir [RU96].

Les idées ci-dessus sont à la base de [ST94]. Malheureusement, si l'on part d'un modèle quartique $U^2 = F(V)$, avec F de degré 4, la situation est un peu plus complexe.

La principale difficulté vient de ce qu'en se restreignant à $E(\mathbb{R})$, les choses ne sont plus aussi simples ; dès que f a ses trois zéros réels, la courbe $E(\mathbb{R})$ aura deux composantes connexes, et le paramétrage diffère selon que l'on considère la composante connexe compacte ou la composante connexe non bornée. Tant que l'on travaille au départ dans un modèle de Weierstraß, cela ne pose pas de problème, car les points entiers de la composante connexe compacte sont trouvés par une simple énumération.

Dans le cas présent, toutefois, il va falloir transformer notre quartique pour disposer du paramétrage ci-dessus ; cela va nous amener à faire une transformation birationnelle qui, si elle préserve le caractère rationnel des points, ne préserve pas le caractère entier. En particulier, cette fois, on ne peut plus se débarrasser aussi facilement de la composante connexe bornée.

[1] C'est de là que vient le nom "courbe elliptique".

Heureusement, il se trouve que l'on peut trouver U_0 tel que les points (U, V) pour lesquels $|U| > U_0$ ont tous, dans le modèle de Weierstraß, une image qui est située dans la même composante connexe. Des calculs soigneux permettent de déterminer dans quel cas on se trouve, et de choisir alors intelligemment le paramétrage et la forme linéaire correspondante ; une version de l'algorithme ainsi adaptée à un modèle quartique de la courbe elliptique se trouve dans [Tz95]. Nous suivons la méthode décrite dans cet article.

5.1.4.2 Modèles de Weierstraß

Il n'est plus utile dans ce contexte d'imposer au second membre d'être unitaire ; en revanche, le terme constant doit être un carré. Une transformation linéaire appropriée conduit à l'équation

$$V^2 = 2/5U^4 - 8/5U^3 + 2/5U^2 + 12/5U + 9 =: Q(U). \tag{5.4}$$

Par symétrie, il suffit de trouver les solutions de cette équation vérifiant $U > 1$.

Cette équation définit une courbe elliptique dont un modèle de Weierstraß est donné par

$$Y^2 + 4/5XY - 48/5Y = X^3 + 6/25X^2 - 72/5X - 432/125, \tag{5.5}$$

ou encore

$$y^2 = x^3 - 1372/75x + 14864/675 =: q(x). \tag{5.6}$$

Suivant le modèle dans lequel les points seront donnés, on adoptera la notation (U, V), (X, Y), (x, y). Les transformations birationnelles correspondantes sont données en annexe.

5.1.4.3 Invariants de la courbe

Grâce au programme `mwrank` de John Cremona, on découvre que $E(\mathbb{Q})$ est de rang 3, engendré par les points

$$(x, y) = (74/15, 36/5), (56/15, -12/5), (262/75, 108/125)$$

et que la torsion est d'ordre 2, engendrée par $(x, y) = (4/3, 0)$. Notons que les générateurs donnés pour la partie libre de $E(\mathbb{Q})$ sont tous dans la composante connexe $E_0(\mathbb{R})$ non bornée de $E(\mathbb{R})$.

Le discriminant Δ de la courbe vaut 44509824000000, et son j-invariant 1291315424/347733. En utilisant le Théorème 1.1 de [Si90] pour le modèle de Weierstraß minimal $y_1^2 = x_1^3 - x_1^2 + 11433x_1 - 340263$, on voit que $\hat{h}(P) - \frac{1}{2}h(x_1(P)) \leq 1/12(\log(\Delta) + \log(j)) + 1/2\log 2 + 1,07 \leq 4,721$, où $\hat{h}$ est la hauteur de Néron-Tate et h la hauteur de Weil.

5.1.4.4 Mise en œuvre de la méthode

Dans le cas présent, les critères donnés dans [Tz95] montrent que la partie de la courbe quartique vérifiant $|U| > U_0 = 12,07$ s'envoie par les transformations birationnelles données en annexe sur la composante connexe non bornée de E.

Soit alors P un point de $E(\mathbb{Q})$. Écrivons $P = -n_1P_1 - n_2P_2 - n_3P_3 + T$, où T est de torsion.

D'après [Tz95, (7)], la forme linéaire en logarithmes est donnée dans ce cas par

$$\begin{aligned}\Phi(P) &= \frac{1}{\omega}\int_U^{+\infty}\frac{du}{\sqrt{Q(u)}},\\ &= \phi(P) - \phi(P_0),\\ &= -(\phi(P_0) + n_1\phi(P_1) + n_2\phi(P_2) + n_3\phi(P_3) + n_0/2),\end{aligned}$$

où n_0 est un entier tel que le second membre soit dans $[0,1]$, U est l'abscisse du point correspondant à P dans le modèle (5.4), P_0 est le point $((2+18\sqrt{10})/15, (120-12\sqrt{10})/25)$, et $\omega = 1,49400...$ est la période réelle de la courbe elliptique dans les modèles (5.5), (5.6).

Il nous faut maintenant majorer notre forme linéaire en fonction de $|U|$, ce qui, en utilisant l'inégalité sur la hauteur de Weil donnée plus haut, fournira une majoration en terme de $\exp(-\max_i |n_i|)$:

$$\begin{aligned}\int_U^{+\infty}\frac{du}{\sqrt{Q(u)}} &= \int_U^{\infty}\frac{du}{\sqrt{2/5}\sqrt{((u-1)^2-5/2)^2+(9/2)^2}},\\ &\leq \sqrt{5/2}\int_U^{+\infty}\frac{du}{|(u-1)^2-5/2|}.\end{aligned}$$

Comme $U \geq 12$, on peut enlever la valeur absolue de la dernière expression, et $(u-1)^2 - 5/2 \geq 0,97(u-1)^2$ pour $u \geq 12$, d'où

$$\Phi(P) \leq \frac{0,97\sqrt{5/2}}{\omega U}.$$

Maintenant, on utilise l'expression de $X_1(P)$, abscisse de P dans le modèle de Weierstraß minimal :

$$X_1(P) = 25X(P) + 3 = 25\frac{6\sqrt{Q(U)} + 3U^2 + 2/5U + 18}{U^2},$$

et donc $h(X_1(P)) \leq \max(2\log(5) + \log(6\sqrt{Q(U)} + 3U^2 + 2/5U + 18), 2\log(U))$.

Mais $6\sqrt{Q(U)}+3U^2+2/5U+18 \leq 6,8U^2$, et donc $h(X_1(P)) \leq 5,14+2\log(U)$.

Calculant les valeurs propres de la matrice $(< P_i, P_j >)_{1\leq i,j\leq 3}$, on voit que $\hat{h}(P) \geq 0,84M^2$. Il vient :

$$\begin{aligned} |\omega\Phi(U)| &\leq 1,62\exp(-\log(U)) && (5.7)\\ &\leq 1,62\exp\left(-\frac{h(X_1(P))}{2}+2,57\right) && (5.8)\\ &\leq 21,2\exp(4,721-0,84M^2) && (5.9)\\ &\leq 2377\exp(-0,84M^2). && (5.10)\end{aligned}$$

Posons alors $N = \max(|n_0|, |n_1|, |n_2|, |n_3|)$. Par définition de n_0, on a $N \leq 6M + 2$.

On va alors appliquer le résultat de David [Da95], qui dit que si $\Phi(P) \neq 0$, on a

$$|\omega\Phi(P)| \geq \exp(-8,1 \cdot 10^{39} \log N(\log\log N + 21)^5). \tag{5.11}$$

Comparant (5.7) et (5.11), il vient

$$M^2 \leq 9,3 + 9,6 \cdot 10^{39} \log N(\log\log N + 21)^5,$$

on en déduit $M \leq M_0 := 2,3 \cdot 10^{22}$.

5.1.5 Réduction

Comme[2] au chapitre 1, on considère le réseau engendré par les colonnes de la matrice.

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ [C\phi(P_1)] & [C\phi(P_2)] & [C\phi(P_3)] & C/2 \end{pmatrix},$$

avec C pair. Le plus court vecteur de la base LLL-réduite de ce réseau (avec $C = 10^{110}$ vérifie $\|b_1\|_2 \geq 2,24 \cdot 10^{27}$. La distance du point ${}^t(0, 0, 0, [-C\Phi(P_0)])$ à ce même réseau vérifie $d \geq 3,97 \cdot 10^{26}$, par le Lemme 1.30. Par des techniques analogues à celle du chapitre 1, on peut déduire la minoration :

$$\begin{aligned} |m_1\phi(P_1) + m_2\phi(P_2) + m_3\phi(P_3) + m_0/2 - \Phi(P_0)| &\geq \frac{1}{C}\sqrt{d^2 - 3M_0^2} - 3M_0 - 1\\ &= 3,96 \cdot 10^{-84}. \end{aligned}$$

Comme cette quantité est également dominée par $1590\exp(-0,84M^2)$, on voit que $M \leq 15$.

[2] si ce n'est que l'on n'a plus de conjugaison et que l'on ne peut plus utiliser qu'une seule forme en $r + 1$ variables

Une étape supplémentaire avec $C = 10^{15}$ donne $M \leq 6$. On énumère alors les 4394 points possibles, dont on calcule les coordonnées dans le modèle quartique ; les points entiers trouvés (et le point à l'infini) sont donnés dans le tableau suivant

n_0	n_1	n_2	n_3	(U,V)	n_0	n_1	n_2	n_3	(U,V)
0	−2	0	0	$[-10,75]$	1	−2	0	0	$[12,-75]$
0	−2	1	−1	$[9,39]$	1	−2	1	−1	$[-7,-39]$
0	−2	1	1	$[299,56163]$	1	−2	1	1	$[-297,-56163]$
0	−1	0	−1	$[-2,5]$	1	−1	0	−1	$[4,-5]$
0	−1	0	0	$[3,3]$	1	−1	0	0	$[-1,-3]$
0	−1	1	−1	$[2,3]$	1	−1	1	−1	$[0,-3]$
0	−1	1	0	$[-3,9]$	1	−1	1	0	$[5,-9]$
0	−1	2	−1	$[-22,333]$	1	−1	2	−1	$[24,-333]$
0	0	−1	0	$[24,333]$	1	0	−1	0	$[-22,-333]$
0	0	0	−1	$[5,9]$	1	0	0	−1	$[-3,-9]$
0	0	0	0	∞	1	0	0	0	$[2,-3]$
0	0	1	−1	$[-1,3]$	1	0	1	−1	$[3,-3]$
0	0	1	0	$[4,5]$	1	0	1	0	$[-2,-5]$
0	1	0	−2	$[-297,56163]$	1	1	0	−2	$[299,-56163]$
0	1	0	0	$[-7,39]$	1	1	0	0	$[9,-39]$
0	1	1	−1	$[12,75]$	1	1	1	−1	$[-10,-75]$

auquel il convient d'ajouter les éventuelles solutions avec $U \leq 12$ (qui sont en fait déjà dans la table ci-dessus). Ceci fournit une deuxième preuve du théorème 5.3.

5.2 Deux exemples superelliptiques

Nous donnons ici deux exemples destinés à montrer que la méthode du chapitre 4 fonctionne également bien pour des équations superelliptiques. Les exemples choisis sont volontairement simples (les deux groupes de Galois sont $\mathbb{D}_4$) pour que les corps admissibles mis en jeu soient de petit degré (6, en l'occurence). Les équations du type $y^3 = f(x)$ avec $\deg f = 4$, $\mathrm{Gal}\,(f) = S_4$, qui donnent naissance à des corps admissibles de degré 18, sont probablement la limite de la méthode actuelle — et il paraît peu vraisemblable que l'on puisse certifier les unités dans ces corps, si tant est que l'on arrive à les obtenir.

Précisons que la situation est tout de suite beaucoup plus agréable si f n'est plus irréductible.

5.2.1 $28y^3 = x^4 - 20x^2 - 32x + 28$

Pour cette équation, comme pour la suivante, on ne donnera pas autant de détails que dans la section précédente, mais juste une liste de corps admissibles, le cardinal de Θ, le nombre de vecteurs $\boldsymbol{k}$, ainsi que les principales constantes.

Les corps admissibles sont au nombre de 5, $\mathbb{K}_0 = \mathbb{Q}(\alpha_0)$, où α_0 est une racine de $x^2+12x+28$, et les 4 corps de degré 6 engendrés par des racines des polynômes suivants :

$$x^6 - 18x^4 - 8x^3 + 63x^2 + 168x - 112, x^6 - 18x^4 - 24x^3 + 63x^2 + 168x + 112$$
$$x^6 + 6x^4 - 4x^3 - 63x^2 + 84x - 28, x^6 - 6x^4 - 4x^3 + 9x^2 + 12x - 4.$$

Les constantes suivantes ne dépendent pas du corps admissible choisi :

$$\begin{array}{cccc} c_1 = 5,1 & X_1 = 15 & c_2 = 0,16 & c_3 = 0,42 \\ c_4 = 1,8 & c_5 = 5,95 & c_6 = 9,7 & c_7 = 19,3 \\ X_2 = 15 & c_{12} = 116 & c_{19} = 7/2 & c_{20} = 2,35 \\ c_{21} = 1,12 & & & \end{array}$$

Pour chacun des corps, il existe 19 éléments θ ; comme $m \not\equiv 0 \pmod{p}$, le corps $\mathbb{K}_0$ se traite via le Corollaire 4.13, qui fournit la majoration $|x| \leq 115$.

Voici les pires valeurs des principales constantes pour les autres corps admissibles ; les deux seules solutions $(-2,1)$ et $(0,1)$ ont été trouvées lors de l'énumération finale et correspondent respectivement aux corps $\mathbb{K}_2$ et $\mathbb{K}_3$.

$$\begin{array}{ccc} c_8 = 2,46 & c_9 = 3,07 & c_{10} = 43,7 \\ B_0 = 8,3 \cdot 10^{35} & B_{\text{red}} = 28 & X_6 = 873 \end{array}$$

On a donc le

Théorème 5.5. *Les solutions de l'équation diophantienne* $28y^3 = x^4 - 20x^2 - 32x + 28$ *sont* $(x,y) = (-2,1), (0,1)$.

Le temps de calcul total est de 6 minutes.

5.2.2 $y^3 = x^4 - x^3 - 3x^2 + x + 1$

Les corps admissibles sont au nombre de 5, $\mathbb{K}_0 = \mathbb{Q}(\alpha_0)$, où α_0 est une racine de $x^2 - x - 1$, et les 4 corps de degré 6 engendrés par des racines des polynômes suivants :

$$x^6 + 6x^4 - x^3 + 9x^2 - 3x - 1, x^6 + 3x^4 - x^3 - 9x^2 - 9x - 1,$$
$$x^6 - 3x^4 - 3x^3 - 9x^2 - 18x - 9, x^6 + 3x^4 - 5x^3 - 9x^2 + 5.$$

Les constantes suivantes ne dépendent pas du corps admissible choisi :

$$\begin{array}{cccc} c_1 = 2,1 & X_1 = 6 & c_2 = 0,16 & c_3 = 0,42 \\ c_4 = 1,8 & c_5 = 2,58 & c_6 = 1,8 & c_7 = 3,57 \\ X_2 = 6 & c_{12} = 22 & c_{19} = 7/2 & c_{20} = 0,43 \\ c_{21} = 0,37 & & & \end{array}$$

Pour chacun des corps, il existe 4 éléments θ ; comme $m \not\equiv 0 \pmod{p}$, le corps $\mathbb{K}_0$ se traite via le Corollaire 4.13, qui fournit la majoration $|x| \leq 21$.

Voici les pires valeurs des principales constantes pour les autres corps admissibles ; les seules solutions $(-1,-1)$,$(0,1)$,$(1,-1)$, $(2,-1)$ (correspondant aux corps $\mathbb{K}_4$, $\mathbb{K}_1$, $\mathbb{K}_2$, $\mathbb{K}_3$) ont été trouvées lors de l'énumération finale.

$$\begin{array}{ccc} c_8 = 8,4 & c_9 = 3,22 & c_{10} = 16,4 \\ B_0 = 6,2 \cdot 10^{29} & B_{\mathrm{red}} = 52 & X_6 = 327 \end{array}$$

Le temps de calcul total est de moins d'une minute.

Théorème 5.6. *Les seules solutions de l'équation* $y^3 = x^4 - x^3 - 3x^2 + x + 1$ *sont* $(x,y) = (-1,-1), (0,1), (1,-1), (2,-1)$.

Bibliographie

[Ba66] A. Baker, Linear forms in the logarithms of algebraic numbers I, *Mathematika* **13** (1966), 204–216 ; II, ibid. **14** (1967), 102–107 ; III, ibid. **14** (1967), 220–228 ; IV, ibid. **15** (1968), 204–216.

[Ba68] A. Baker, Contributions to the theory of Diophantine equations. I. On the representation of integers by binary forms, *Philos. Trans. Roy. Soc. London Ser. A* **263** (1968), 173–191 ; II. The Diophantine equation $y^2 = x^3 + k$, ibid. **263** (1968), 193–208.

[Ba72] A. Baker, A sharpening of the bound for linear forms in logarithms, I, *Acta Arith.*, **21** (1972), 117–129 ; II, ibid. **24** (1973), 33-36 ; III, ibid. **27** (1975), 247–252.

[Ba90] E. Bach, Explicit bounds for primality testing and related problems, *Math. Comp.* **55** (1990), 355–380.

[BD69] A. Baker, H. Davenport, The equations $3x^2-2 = y^2$ and $8x^2-7 = z^2$, *Quart. J. Math. Oxford (2)* **20** (1969), 129–137.

[BGMMS90] J. Blass, A.M.W. Glass, D.K. Manski, D.B. Meronk, R.P. Steiner, Constants for lower bounds for linear forms in logarithms of algebraic numbers II. The homogeneous rational case, *Acta Arith.* **55** (1990), 15–22.

[BH96] Yu. Bilu, G. Hanrot, Solving Thue Equations of High Degree, *J. Number Th.* **60** (1996), 373–392.

[BH96b] Yu. Bilu, G. Hanrot, Solving Superelliptic Diophantine Equations by Baker's method, *Compositio Math.*, à paraître.

[BH97] Yu. Bilu, G. Hanrot, Thue Equations with Composite Fields, soumis.

[Bi94] Yu. Bilu, Solving Superelliptic Diophantine Equations by Baker's method, prépublication.

[BS] Z. I. Borevich, I. R. Shafarevich, "Number Theory", Academic Press, New York, 1966.

[Bu88] J. Buchmann, Computing class groups and regulators in subexponential time, Séminaire de Théorie des Nombres de Paris 1988-89, 27–39, Birkhaüser.

[BV04] G.D. Birkhoff, H.S. Vandiver, On the integral divisors of $a^n - b^n$, *Ann. Math. (2)* **5** (1904), 173–180.

[BW93] A. BAKER AND G. WÜSTHOLZ, Logarithmic forms and group varieties, *J. Reine Angew. Math.* **442** (1993), 19–62.

[BW96] M. BENNETT, B.M.M. DE WEGER, On the Diophantine Equation $|ax^n - by^n| = 1$, *Math. Comp.*, à paraître.

[Ca13] R. D. CARMICHAEL, On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$, *Ann. Math. (2)* **15** (1913), 30–70.

[CDO97] H. COHEN, F. DIAZ Y DIAZ, M. OLIVIER, Subexponential algorithms for class group and units computation, *J. Symb. Comp.*, à paraître.

[CF91] A. COSTA, E. FRIEDMAN, Ratios of regulators in totally real extensions of number fields, *J. Number Th.* **37** (1991), 288–297.

[Co] H. COHEN, "A Course in Computational Algebraic Number Theory", Graduate Texts in Math., Vol. 138, Springer, 1993.

[Da95] S. DAVID, Minorations de formes linéaires de logarithmes elliptiques, *Mém. Soc. Math. France Nouv. Ser.* **62**.

[DG85] P. DIACONIS, R.L. GRAHAM, The Radon transform on $\mathbb{Z}_2^k$, *Pac. J. Math.* **118** (1985), 323–345.

[Fi97] C. FIEKER, communication privée.

[FP85] U. FINCKE, M. POHST, Improved methods for calculating vectors of short length in a lattice, including a complexity analysis, *Math. Comp.* **44** (1985), 463–471.

[Fr89] E. FRIEDMANN, Analytic formulas for the regulator of a number field, *Invent. Math.* **98** (1989), 599–622.

[GPZ94] J. GEBEL, A. PETHŐ, H. ZIMMER, Computing integral points on elliptic curves, *Acta Arith.* **68** (1994), 171-192.

[GPZ96] J. GEBEL, A. PETHŐ, H. ZIMMER, Computing S-integral Points on Elliptic Curves, in H. Cohen (Ed.), "Algorithmic Number Theory", Proceedings of the Second International Symposium ANTS II.

[Ha97] G. HANROT, Solving Thue Equations without the Full Unit Group, Prépublication A2X 97-1.

[HM89] J. HAFNER, K. MCCURLEY, A rigorous subexponential algorithm for computation of class groups, *J. Amer. Math. Soc.* **2** (1989), 837–850.

[Hi] D. HILBERT, Gesammelte Abhandlungen, Band 3, Chelsea publishing company (1933).

[KL96] I. KRASIKOV, S. LITSYN, On Integral Zeros of Krawtchouk Polynomials, *J. Comb. Th., Ser. A*, **74** (1996), 71-99.

[Li1844] J. LIOUVILLE, Remarques relatives à des classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques, *C. R. Acad. Sc. Paris* **18** (1844), 883–885 et 910–911.

[Li1851] J. LIOUVILLE, Sur des classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques, *J. Math. Pures App.* **16** (1851), 133–142.

[LLL82] A.K. LENSTRA, H.W. LENSTRA, JR., L. LOVÁSZ, Factoring polynomials with rational coefficients, *Math. Ann.* **261** (1982), 515–534.

[LMN95] M. LAURENT, M. MIGNOTTE, Y. NESTERENKO, Formes linéaires en deux logarithmes et déterminants d'interpolation, *J. Number Th.* **65** (1995), 285–321.

[Ma70] YU. MATIJASEVIČ, Enumerable sets are diophantine, *Dokl. Akad. Nauk. SSSR* **191** (1970) (Russe) ; version anglaise améliorée : *Soviet Math. Doklady* **12** (1971), 249–54.

[Ma78] J.M. MASLEY, Class numbers of real cyclic number fields with small conductor, *Compositio Math.* **37** (1978), 297–319.

[MP96] M. MIGNOTTE, A. PETHŐ, On the system of diophantine equations $x^2 - 6y^2 = -5$ and $x = 2z^2 - 1$, *Math. Scand.* **76** (1995), 50–60.

[MW94] M. MIGNOTTE, B.M.M. DE WEGER, On the Diophantine equations $x^2 + 74 = y^5$ and $x^2 + 86 = y^5$, *Glasgow Math. J.* **38** (1996), 77–87.

[Pe90] A. PETHŐ, Computational Methods For the Resolution of Diophantine Equations, *in* R.A. Mollin (ed.), Number Theory : Proc. First Conf. Can. Number Th. Ass., Banff, 1988, de Gruyter, 1990, 477–492.

[Po93] M. POHST, "Computational Algebraic Number Theory", DMV Seminar, Vol. 21, Birkhäuser, Basel, 1993.

[PW97] M. POHST, K. WILDANGER, Tables of unit groups and class groups of quintic fields and a regulator bound, *Math. Comp.*, à paraître.

[PW87] A. PETHŐ, B.M.M. DE WEGER, Products of Prime powers in Binary Recurrence Sequences Part I : The Hyperbolic Case, with an Application to the Generalized Ramanujan–Nagell Equation, *Math. Comp.* **47** (1987), 713–727.

[PWZ82] M. POHST, P. WEILER, H. ZASSENHAUS, On effective computation of fundamental units II, *Math. Comp.* **38** (1982), 293–329.

[PZ] M. POHST, H. ZASSENHAUS, "Algorithmic Algebraic Number Theory", Cambridge University Press, 1989.

[PZ82] M. POHST, H. ZASSENHAUS, On effective computation of fundamental units I, *Math. Comp.* **38** (1982), 275–291.

[RU96] G. RÉMOND, F. URFELS, Approximation diophantienne de logarithmes elliptiques p-adiques, *J. Number Th.* **57**, (1996), 133–169.

[Ro55] K.F. ROTH, Rational approximations to algebraic numbers, *Mathematika*, **2** (1955), 1–20.

[Sh76] T.N. SHOREY, On linear forms in the logarithms of algebraic numbers, *Acta Arith.* **30** (1976), 27–42.

[Si21] C.L. SIEGEL, Approximation algebraischer Zahlen, *Math. Zeit.* **10** (1921), 173–213.

[Si26] C.L. SIEGEL, The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \ldots + k$ (Extract from a letter to Prof. L.J. Mordell unter dem Pseudonym X), *J. London Math. Soc.* **1** (1926), 66–68.

[Si] J.H. SILVERMAN, "The Arithmetic of Elliptic Curves", Graduate Texts in Mathematics Vol 106, Springer-Verlag, 1986.

[Si90] J.H. SILVERMAN, The difference between the Weil height and the canonical height on elliptic curves, *Math. Comp.* **55** (1990), 723–743.

[Sm94] N.P. SMART, S-integral points on elliptic curves, *Math. Proc. Camb. Philos. Soc.* **116** (1994), 391–399.

[Sm95] N.P. SMART, The solution of triangularly connected decomposable form equations, *Math. Comp.* **64** (1995), 819–840.

[Sp] V.G. SPRINDŽUK, "Classical Diophantine Equations in Two Unknowns", Lecture Notes in Mathematics, Vol. 1559, Springer, 1994.

[St77] C.L. STEWART, On divisors of Fermat, Fibonacci, Lucas and Lehmer sequences, *Proc. London. Math. Soc. (3)* **35** (1977), 425–447.

[ST] T.N. SHOREY, R. TIJDEMAN, "Exponential Diophantine Equations", Cambridge University Press, Cambridge, 1986.

[ST94] R. STROEKER, N. TZANAKIS, Solving elliptic Diophantine equations by estimating linear forms in elliptic logarithms, *Acta Arith.* **67** (1994), 177–196.

[SW94] R. STROEKER, B.M.M. DE WEGER, On some Diophantine Equations that defy Thue's method : the case of Ochoa's curve, *Exper. Math.* **2** (1994), 209–220.

[SW96] R. STROEKER, B.M.M. DE WEGER, On a quartic Diophantine Equation, *Proc. Edinburgh Math. Soc.* **39** (1996), 97–115.

[Th09] A. THUE, Über Annäherungswerte algebraischer Zahlen, *J. Reine Angew. Math.* **135** (1909), 284–305.

[Tz95] N. TZANAKIS, Solving elliptic diophantine equations by estimating linear forms in elliptic logarithms. The case of quartic equations, *Acta Arith.* **75** (1996), 165–190.

[TW89] N. TZANAKIS, B.M.M. DE WEGER, On the Practical Solution of the Thue Equation, *J. Number Th.* **31** (1989), 99–132.

[TW92] N. TZANAKIS, B.M.M. DE WEGER, How to explicitly solve a Thue–Mahler Equation, *Compositio Math.* **84** (1992), 223–288.

[Vo95] P.M. VOUTIER, Primitive divisors of Lucas and Lehmer sequences *Math. Comp.*, **64** (1995), 869–888.

[Vo97] P.M. VOUTIER, Primitive divisors of Lucas and Lehmer sequences, II, *J. Th. Nombres Bordeaux* **8** (1996), 251–275.

[Wa55] M. WARD, The intrinsic divisors of Lehmer numbers, *Ann. Math. (2)* **62** (1955), 230–236.

[Wa80] M. WALDSCHMIDT, A lower bound for linear forms in logarithms, *Acta Arith.* **37** (1980), 257–283.

[We] B.M.M. DE WEGER, Algorithms for diophantine equations, CWI-Tract n° 65, Centre for Math. and Comp. Sci., Amsterdam, 1989.

[We87] B.M.M. DE WEGER, Solving exponential diophantine equations using lattice basis reduction algorithms, *J. Number Th.* **26** (1987), 325–367.

[We92] B.M.M. DE WEGER, A hyperelliptic diophantine equation related to imaginary quadratic number fields with class number 2, *J. Reine Angew. Math.* **427** (1992), 137–156.

[We94a] B.M.M. DE WEGER, Integral and S-integral solutions of a Weierstrass equation, rapport 9452/B, Econometric Institute, Erasmus University, Rotterdam (1994).

[We94b] B.M.M. DE WEGER, Solving elliptic Diophantine equations avoiding Thue equations and elliptic logarithms, rapport 9469/B, Econometric Institute, Erasmus University, Rotterdam (1994).

[Za87] D. ZAGIER, Large integral points on elliptic curves, *Math. Comp.* **48** (1987), 425–436.

[Zi81] R. ZIMMERT, Ideale kleiner Norm in Idealklassen und eine Regulatorabschätzung, *Invent. Math.* **62** (1981), 367–380.

Annexe A

L'équation cyclotomique réelle

Seules les constantes ne pouvant être recalculées à partir d'autres constantes figurent ci-dessous. Par ailleurs, les bornes obtenues pour $|x|$ à l'issue de la première étape de fractions continues étaient suffisamment petites pour nous dispenser de l'énumération finale. On donne donc juste ces bornes pour $|x|$. La constante B_1 est celle obtenue à l'issue de la première étape de réduction.

Certaines des constantes figurant ci-dessous dépendent du i_0 choisi ; on a fait figurer ci-dessous une valeur admissible "uniforme" pour chacune des constantes, c'est-à-dire une valeur supérieure (ou inférieure, c'est selon) au maximum (resp. minimum) des valeurs trouvées pour chacune des constantes.

L'ensemble des constantes, ainsi que les programmes concernés sont disponibles par courrier électronique[1].

Les temps de calcul (trois dernières colonnes) sont donnés en millisecondes, sur un Pentium Pro 200 MHz, Linux 2.0.28, 256 Mo RAM.

Enfin, ces programmes utilisent lourdement la bibliothèque PARI ; les temps donnés se rapportent à la version 1.915.

A.1 Cas $31 \leq p^\alpha \leq 67$.

Le contenu de cette table a trait au chapitre 3, section 3.1. Notons que les temps de calcul donnés ne sont probablement pas les meilleurs, en ce sens que la précision utilisée n'a pas été ajustée pour être la plus petite possible ; ainsi les grands "sauts" de temps de calcul proviennent-ils d'une augmentation de la précision lorsque celle-ci était devenue insuffisante.

Par ailleurs, les temps de calcul (qui proviennent du timer de PARI) ne sont pas d'une grande fiabilité, ce qui peut expliquer certaines différences mineures (de l'ordre de la seconde).

Enfin, la nette amélioration du temps de calcul pour le cas $p = 67$ par rapport à [BH96] a (au moins) trois explications : une machine plus rapide, le fait que

[1] `hanrotmath.u-bordeaux.fr`

le léger changement de présentation évite de multiplier à chaque étape par une matrice T (voir [BH96]), et enfin l'énumération finale est rendue inutile par la borne obtenue pour $|x|$ après une seule étape de réduction.

Les trois dernières colonnes de la table se rapportent au temps de calcul (respectivement de A^{-1}, de réduction, et total).

p^α	a	c_1	c_2	c_5	c_6	B_0	B_1	X_5	A^{-1}	*Red.*	T*ps*
31	± 1	1619	0,08	3,8	2,1	$1,7 \cdot 10^{85}$	57	3	0,9	0,95	17,7
31	± 31	50170	0,08	3,8	1,91	$2,6 \cdot 10^{85}$	58	4	0,9	0,9	18,5
2^5	± 1	16	0,2	3,6	2,1	$1,6 \cdot 10^{46}$	58	3	0,18	0,66	8,00
2^5	± 2	32	0,2	3,6	2,0	$1,9 \cdot 10^{46}$	59	3	0,18	0,65	8,34
37	± 1	10900	0,05	4,2	2,1	$4,7 \cdot 10^{103}$	64	2	1,7	1,2	26,4
37	± 37	403050	0,05	4,2	2,1	$7,0 \cdot 10^{103}$	64	3	1,7	1,2	26,8
41	± 1	39300	0,04	3,9	2,1	$1,4 \cdot 10^{116}$	60	2	2,3	1,5	33,0
41	± 41	$1,6 \cdot 10^6$	0,04	3,9	2,0	$2,2 \cdot 10^{116}$	60	3	2,3	1,5	33,6
43	± 1	75080	0,04	3,5	2,1	$3,5 \cdot 10^{122}$	53	2	2,7	1,6	36,5
43	± 43	$3,22 \cdot 10^6$	0,04	3,5	1,9	$5,2 \cdot 10^{122}$	53	3	2,7	1,6	37,2
47	± 1	$2,75 \cdot 10^5$	0,03	4,1	2,1	$3,2 \cdot 10^{135}$	63	2	3,6	1,8	44,4
47	± 47	$1,3 \cdot 10^7$	0,03	7,8	2,2	$4,9 \cdot 10^{135}$	63	3	3,6	1,8	45,3
7^2	± 1	83000	0,03	5,9	2,1	$9,9 \cdot 10^{122}$	90	2	2,1	1,24	28,7
7^2	± 7	$5,8 \cdot 10^6$	0,03	5,9	2,2	$1,5 \cdot 10^{123}$	90	3	2,1	1,24	29,4
53	± 1	$2 \cdot 10^6$	0,02	4,3	2,1	$2,0 \cdot 10^{155}$	66	2	5,2	2,3	58,2
53	± 53	$1,1 \cdot 10^8$	0,02	4,3	2,1	$3,1 \cdot 10^{155}$	66	3	5,2	2,2	58,8
59	± 1	$1,5 \cdot 10^7$	0,02	3,3	2,1	$2,6 \cdot 10^{175}$	52	2	9,0	3,2	91,7
59	± 59	$8,3 \cdot 10^8$	0,02	3,3	1,9	$3,8 \cdot 10^{175}$	52	3	9,0	3,2	92,3
61	± 1	$2,8 \cdot 10^8$	0,02	3,8	2,1	$1,9 \cdot 10^{182}$	59	2	12,0	4,1	120
61	± 61	$1,7 \cdot 10^{10}$	0,02	3,8	2,1	$2,8 \cdot 10^{182}$	59	2	12,0	4,1	122
2^6	± 1	2040	0,07	4,9	2,2	$2,5 \cdot 10^{92}$	73	2	0,9	0,9	16,4
2^6	± 2	4080	0,07	4,9	2,1	$3,1 \cdot 10^{92}$	74	3	0,9	0,9	16,8
67	± 1	$2,0 \cdot 10^9$	0,01	3,76	2,1	$7,2 \cdot 10^{202}$	59	2	22,5	6,7	212,5
67	± 67	$1,4 \cdot 10^{10}$	0,01	13,0	3,3	$9,5 \cdot 10^{202}$	60	2	22,3	6,8	215,1

A.2 $67 \leq p \leq 1000$

Dans cette table, les résultats concernant les nombres premiers 251, 431, 491, 701, 911, 971 ne sont valables que sous l'hypothèse de Riemann généralisée ; le temps nécessaire à la certification étant en rapport direct avec la taille du régulateur, nous n'avons pas cru bon de certifier ces exemples, qui sont de toutes façons justiciables de la méthode de la section suivante. Notons que tous les temps de calcul supérieurs à la minute correspondent à des régulateurs plus grands que 5000.

p	n	c_1	c_2	c_7	c_8	B_0	B_1	X_5	Tps	Tps UF
67	3	$1,4 \cdot 10^{10}$	0,017	4,7	2,0	$2,1 \cdot 10^{28}$	7,9	7	4,3	1,0
71	5	$5,3 \cdot 10^{10}$	0,015	11,	2,9	$9,3 \cdot 10^{38}$	15,	7	10,2	5,1
73	3	$1,1 \cdot 10^{11}$	0,014	5,4	2,0	$3,4 \cdot 10^{28}$	8,7	6	4,6	1,3
79	3	$8,5 \cdot 10^{11}$	0,012	12,	3,8	$2,4 \cdot 10^{28}$	19,	6	4,5	1,1
89	4	$2,8 \cdot 10^{13}$	0,0099	4,3	1,9	$1,2 \cdot 10^{35}$	6,8	5	106,5	102,1
97	3	$4,4 \cdot 10^{14}$	0,0083	15,	4,3	$9,0 \cdot 10^{28}$	20,	5	4,6	1,0
101	5	$1,8 \cdot 10^{15}$	0,0077	13,	3,2	$2,2 \cdot 10^{40}$	17,	4	10,9	5,6
103	3	$3,5 \cdot 10^{15}$	0,0074	6,7	2,2	$3,4 \cdot 10^{29}$	9,6	4	4,6	1,3
109	3	$2,8 \cdot 10^{16}$	0,0066	5,0	1,9	$6,3 \cdot 10^{29}$	7,2	4	5,6	1,9
113	4	$1,2 \cdot 10^{17}$	0,0061	4,8	1,9	$7,6 \cdot 10^{35}$	8,0	4	215,2	210,3
127	3	$1,5 \cdot 10^{19}$	0,0048	4,1	1,2	$2,5 \cdot 10^{30}$	9,7	4	7,2	2,4
131	5	$5,7 \cdot 10^{19}$	0,0045	14,	2,9	$2,3 \cdot 10^{41}$	17,	4	19,4	12,7
137	4	$4,6 \cdot 10^{20}$	0,0042	16,	3,7	$5,2 \cdot 10^{35}$	19,	3	7,7	2,5
139	3	$9,1 \cdot 10^{20}$	0,0040	19,	5,3	$9,2 \cdot 10^{29}$	25,	3	5,2	1,3
151	3	$5,9 \cdot 10^{22}$	0,0034	8,0	2,4	$4,2 \cdot 10^{30}$	17,	6	6,3	1,1
157	3	$4,7 \cdot 10^{23}$	0,0032	5,4	1,7	$8,9 \cdot 10^{30}$	11,	6	7,8	2,5
163	3	$3,8 \cdot 10^{24}$	0,0029	22,	4,9	$2,6 \cdot 10^{30}$	44,	6	6,4	1,1
181	3	$2,0 \cdot 10^{27}$	0,0024	9,7	2,7	$1,4 \cdot 10^{31}$	19,	5	7,4	1,5
191	5	$6,1 \cdot 10^{28}$	0,0021	22,	4,3	$7,0 \cdot 10^{42}$	47,	6	24,8	14,6
193	3	$1,3 \cdot 10^{29}$	0,0021	8,5	2,8	$2,6 \cdot 10^{31}$	17,	5	7,9	2,1
199	3	$9,8 \cdot 10^{29}$	0,0019	13,	2,4	$2,2 \cdot 10^{31}$	22,	5	9,2	3,0
211	3	$6,3 \cdot 10^{31}$	0,0017	12,	2,2	$3,6 \cdot 10^{31}$	20,	4	7,5	1,4
223	3	$4,0 \cdot 10^{33}$	0,0015	6,0	1,1	$9,6 \cdot 10^{31}$	10,	4	8,0	1,5
229	3	$3,2 \cdot 10^{34}$	0,0015	6,3	2,0	$1,3 \cdot 10^{32}$	11,	4	9,1	2,5
233	4	$1,3 \cdot 10^{35}$	0,0014	7,2	2,0	$2,8 \cdot 10^{38}$	13,	5	770,4	761,1
241	3	$2,1 \cdot 10^{36}$	0,0013	15,	2,6	$6,6 \cdot 10^{31}$	23,	4	8,9	2,1
251	5	$6,6 \cdot 10^{37}$	0,0012	6,6	2,0	$3,7 \cdot 10^{45}$	12,	5	27,7	15,4
257	4	$5,3 \cdot 10^{38}$	0,0011	13,	2,0	$4,5 \cdot 10^{38}$	21,	4	15,3	5,8
271	3	$6,8 \cdot 10^{40}$	0,0010	9,3	2,0	$2,5 \cdot 10^{32}$	15,	4	10,0	2,4
277	3	$5,4 \cdot 10^{41}$	0,0010	15,	3,2	$2,0 \cdot 10^{32}$	23,	4	9,6	2,0
281	4	$2,2 \cdot 10^{42}$	0,00099	8,5	2,1	$1,1 \cdot 10^{39}$	14,	4	893,8	883,2
283	3	$4,3 \cdot 10^{42}$	0,00098	8,4	1,7	$4,0 \cdot 10^{32}$	13,	4	9,1	1,2
307	3	$1,8 \cdot 10^{46}$	0,00083	5,0	2,0	$1,4 \cdot 10^{33}$	8,5	3	10,8	2,5

p	n	c_1	c_2	c_7	c_8	B_0	B_1	X_5	Tps	Tps UF
311	5	$7,1 \cdot 10^{46}$	0,00081	12,	2,1	$4,1 \cdot 10^{45}$	20,	4	58,0	43,4
313	3	$1,5 \cdot 10^{47}$	0,00080	37,	6,6	$1,7 \cdot 10^{32}$	54,	3	9,7	1,2
331	3	$7,3 \cdot 10^{49}$	0,00072	25,	3,8	$3,8 \cdot 10^{32}$	35,	3	11,1	2,1
337	3	$5,8 \cdot 10^{50}$	0,00069	13,	2,6	$9,8 \cdot 10^{32}$	18,	3	11,5	2,4
349	3	$3,7 \cdot 10^{52}$	0,00064	40,	7,0	$3,4 \cdot 10^{32}$	57,	3	11,3	1,6
353	4	$1,5 \cdot 10^{53}$	0,00063	7,6	1,3	$1,2 \cdot 10^{40}$	11,	4	2270,8	2257,5
367	3	$1,9 \cdot 10^{55}$	0,00058	9,1	2,5	$2,6 \cdot 10^{33}$	13,	3	11,5	1,4
373	3	$1,6 \cdot 10^{56}$	0,00056	17,	3,0	$1,3 \cdot 10^{33}$	23,	3	13,1	2,7
379	3	$1,3 \cdot 10^{57}$	0,00054	19,	3,4	$1,4 \cdot 10^{33}$	26,	3	12,6	2,2
397	3	$6,2 \cdot 10^{59}$	0,00050	18,	3,2	$2,1 \cdot 10^{33}$	24,	3	12,8	1,8
401	4	$2,5 \cdot 10^{60}$	0,00049	26,	3,4	$3,6 \cdot 10^{39}$	36,	3	21,5	6,2
409	3	$4,0 \cdot 10^{61}$	0,00047	16,	2,8	$3,3 \cdot 10^{33}$	20,	3	14,1	2,5
421	3	$2,6 \cdot 10^{63}$	0,00044	19,	3,2	$3,4 \cdot 10^{33}$	24,	3	14,3	2,5
431	5	$8,2 \cdot 10^{64}$	0,00042	6,5	1,0	$1,5 \cdot 10^{48}$	9,5	3	57,7	37,5
433	3	$1,7 \cdot 10^{65}$	0,00042	12,	2,3	$5,8 \cdot 10^{33}$	15,	3	13,5	1,3
439	3	$1,3 \cdot 10^{66}$	0,00040	5,5	1,0	$1,6 \cdot 10^{34}$	7,4	3	14,9	2,6
449	4	$4,2 \cdot 10^{67}$	0,00039	14,	2,0	$4,5 \cdot 10^{40}$	18,	3	992,0	974,8
457	3	$6,7 \cdot 10^{68}$	0,00037	7,2	1,1	$1,6 \cdot 10^{34}$	9,2	3	15,0	2,1
461	5	$2,7 \cdot 10^{69}$	0,00037	21,	2,9	$8,1 \cdot 10^{46}$	28,	3	100,2	78,1
463	3	$5,4 \cdot 10^{69}$	0,00036	26,	3,7	$4,1 \cdot 10^{33}$	33,	3	15,7	2,4
487	3	$2,2 \cdot 10^{73}$	0,00033	17,	2,5	$1,1 \cdot 10^{34}$	20,	3	16,8	2,8
491	5	$8,8 \cdot 10^{73}$	0,00032	15,	2,8	$3,5 \cdot 10^{47}$	21,	3	36,3	13,1
499	3	$1,4 \cdot 10^{75}$	0,00031	5,7	1,1	$3,5 \cdot 10^{34}$	7,3	3	17,2	2,8
521	4	$2,9 \cdot 10^{78}$	0,00029	17,	2,6	$1,1 \cdot 10^{41}$	21,	3	577,0	556,6
523	3	$5,8 \cdot 10^{78}$	0,00028	15,	2,2	$1,9 \cdot 10^{34}$	18,	2	16,4	1,4
541	3	$3,0 \cdot 10^{81}$	0,00026	26,	4,2	$1,5 \cdot 10^{34}$	32,	2	18,5	2,8
547	3	$2,4 \cdot 10^{82}$	0,00026	38,	5,9	$9,1 \cdot 10^{33}$	47,	2	18,3	2,5
569	4	$4,8 \cdot 10^{85}$	0,00024	22,	2,8	$1,8 \cdot 10^{41}$	26,	3	485,2	462,8
571	3	$9,6 \cdot 10^{85}$	0,00024	29,	3,6	$2,0 \cdot 10^{34}$	33,	2	19,6	3,0
577	3	$7,7 \cdot 10^{86}$	0,00023	15,	2,0	$4,2 \cdot 10^{34}$	17,	2	19,5	2,9
593	4	$2,0 \cdot 10^{89}$	0,00022	12,	1,8	$8,1 \cdot 10^{41}$	14,	3	6061,0	6037,4
601	3	$3,2 \cdot 10^{90}$	0,00021	13,	2,3	$6,1 \cdot 10^{34}$	15,	2	19,5	1,6
607	3	$2,6 \cdot 10^{91}$	0,00021	64,	8,2	$1,2 \cdot 10^{34}$	74,	2	19,1	1,2
613	3	$2,1 \cdot 10^{92}$	0,00021	17,	2,9	$5,8 \cdot 10^{34}$	19,	2	20,8	2,8
617	4	$8,1 \cdot 10^{92}$	0,00020	12,	2,2	$8,7 \cdot 10^{41}$	14,	2	4387,8	4363,5
619	3	$1,7 \cdot 10^{93}$	0,00020	16,	2,3	$5,8 \cdot 10^{34}$	18,	2	20,2	1,9
631	3	$1,1 \cdot 10^{95}$	0,00019	29,	3,9	$3,9 \cdot 10^{34}$	33,	2	22,0	3,1
641	4	$3,3 \cdot 10^{96}$	0,00019	53,	7,3	$1,3 \cdot 10^{41}$	64,	2	37,4	11,9
643	3	$6,6 \cdot 10^{96}$	0,00019	6,0	1,0	$2,3 \cdot 10^{35}$	7,1	2	22,5	3,3
661	3	$3,4 \cdot 10^{99}$	0,00018	15,	2,2	$1,2 \cdot 10^{35}$	16,	2	21,9	1,9
673	3	$2,2 \cdot 10^{101}$	0,00017	26,	3,6	$6,4 \cdot 10^{34}$	29,	2	23,5	3,0
691	3	$1,2 \cdot 10^{104}$	0,00016	8,7	1,5	$2,3 \cdot 10^{35}$	10,	2	23,5	2,3
701	5	$3,6 \cdot 10^{105}$	0,00016	25,	2,7	$4,3 \cdot 10^{48}$	29,	3	49,2	13,9

p	n	c_1	c_2	c_7	c_8	B_0	B_1	X_5	Tps	Tps UF
709	3	$5,7 \cdot 10^{106}$	$0,00015$	$72,$	$7,8$	$3,1 \cdot 10^{34}$	$79,$	2	$23,7$	$1,8$
727	3	$2,9 \cdot 10^{109}$	$0,00014$	$6,7$	$1,1$	$4,7 \cdot 10^{35}$	$7,6$	2	$25,7$	$3,2$
733	3	$2,4 \cdot 10^{110}$	$0,00014$	$9,1$	$1,5$	$3,8 \cdot 10^{35}$	$10,$	2	$25,2$	$2,7$
739	3	$1,9 \cdot 10^{111}$	$0,00014$	$9,8$	$1,8$	$3,7 \cdot 10^{35}$	$11,$	2	$25,4$	$2,4$
751	3	$1,2 \cdot 10^{113}$	$0,00013$	$35,$	$4,3$	$9,7 \cdot 10^{34}$	$38,$	2	$26,0$	$2,3$
757	3	$9,5 \cdot 10^{113}$	$0,00013$	$23,$	$3,2$	$1,7 \cdot 10^{35}$	$25,$	2	$25,2$	$1,4$
761	4	$3,8 \cdot 10^{114}$	$0,00013$	$36,$	$3,7$	$5,3 \cdot 10^{41}$	$39,$	2	$197,7$	$165,9$
769	3	$6,1 \cdot 10^{115}$	$0,00013$	$22,$	$2,6$	$2,2 \cdot 10^{35}$	$23,$	2	$27,2$	$2,7$
787	3	$3,2 \cdot 10^{118}$	$0,00012$	$24,$	$3,8$	$2,2 \cdot 10^{35}$	$26,$	2	$26,5$	$1,4$
809	4	$6,4 \cdot 10^{121}$	$0,00012$	$28,$	$2,7$	$3,0 \cdot 10^{42}$	$30,$	2	$848,2$	$814,1$
811	3	$1,3 \cdot 10^{122}$	$0,00012$	$11,$	$1,4$	$6,6 \cdot 10^{35}$	$11,$	2	$28,8$	$2,7$
821	5	$4,1 \cdot 10^{123}$	$0,00011$	$40,$	$3,7$	$6,3 \cdot 10^{48}$	$44,$	2	$734,8$	$692,4$
823	3	$8,2 \cdot 10^{123}$	$0,00011$	$25,$	$2,6$	$3,0 \cdot 10^{35}$	$26,$	2	$28,4$	$1,7$
829	3	$6,6 \cdot 10^{124}$	$0,00011$	$28,$	$2,8$	$2,9 \cdot 10^{35}$	$29,$	2	$30,0$	$3,1$
853	3	$2,7 \cdot 10^{128}$	$0,00010$	$24,$	$4,4$	$4,0 \cdot 10^{35}$	$27,$	2	$30,8$	$3,1$
857	4	$1,1 \cdot 10^{129}$	$0,00010$	$70,$	$7,4$	$7,2 \cdot 10^{41}$	$77,$	2	$49,0$	$12,0$
859	3	$2,2 \cdot 10^{129}$	$0,00010$	$32,$	$4,1$	$3,0 \cdot 10^{35}$	$34,$	2	$31,3$	$3,0$
877	3	$1,1 \cdot 10^{132}$	$0,00010$	$87,$	$11,$	$1,2 \cdot 10^{35}$	$91,$	2	$31,0$	$2,0$
881	4	$4,4 \cdot 10^{132}$	$0,00010$	$17,$	$2,5$	$1,4 \cdot 10^{43}$	$18,$	2	$6109,1$	$6070,9$
883	3	$8,8 \cdot 10^{132}$	$0,00010$	$28,$	$3,4$	$4,5 \cdot 10^{35}$	$29,$	2	$32,6$	$3,2$
907	3	$3,6 \cdot 10^{136}$	$0,000095$	$36,$	$4,1$	$4,4 \cdot 10^{35}$	$37,$	2	$33,3$	$3,0$
911	5	$1,5 \cdot 10^{137}$	$0,000095$	$12,$	$2,0$	$9,3 \cdot 10^{50}$	$14,$	2	$70,3$	$22,7$
919	3	$2,3 \cdot 10^{138}$	$0,000093$	$8,1$	$2,0$	$2,2 \cdot 10^{36}$	$9,5$	2	$34,4$	$3,7$
929	4	$7,4 \cdot 10^{139}$	$0,000091$	$22,$	$2,7$	$1,3 \cdot 10^{43}$	$22,$	2	$2073,6$	$2032,4$
937	3	$1,2 \cdot 10^{141}$	$0,000089$	$92,$	$11,$	$1,9 \cdot 10^{35}$	$95,$	2	$33,9$	$2,1$
941	5	$4,7 \cdot 10^{141}$	$0,000089$	$82,$	$6,5$	$8,0 \cdot 10^{48}$	$87,$	2	$540,6$	$489,4$
953	4	$3,1 \cdot 10^{143}$	$0,000086$	$28,$	$3,3$	$1,0 \cdot 10^{43}$	$29,$	2	$902,9$	$860,2$
967	3	$3,9 \cdot 10^{145}$	$0,000084$	$19,$	$2,5$	$1,4 \cdot 10^{36}$	$19,$	2	$35,0$	$1,9$
971	5	$1,6 \cdot 10^{146}$	$0,000083$	$14,$	$2,0$	$2,8 \cdot 10^{51}$	$15,$	2	$82,5$	$30,4$
977	4	$1,3 \cdot 10^{147}$	$0,000082$	$77,$	$7,9$	$2,6 \cdot 10^{42}$	$80,$	2	$60,4$	$16,5$
991	3	$1,6 \cdot 10^{149}$	$0,000080$	$22,$	$2,4$	$1,4 \cdot 10^{36}$	$22,$	2	$36,7$	$2,1$
997	3	$1,3 \cdot 10^{150}$	$0,000079$	$6,9$	$2,0$	$4,6 \cdot 10^{36}$	$8,3$	2	$38,8$	$4,7$
5011	3	$1.9 \cdot 10^{754}$	$3,1 \cdot 10^{-6}$	57.8	3.4	$5,8 \cdot 10^{40}$	46	2	$479,8$	$6,6$

A.3 $p \in \{251, 431, 491, 701, 911, 971\}$

On regroupe ici les résultats obtenus par la méthode du chapitre 2, avec un $b_0 \neq 1$ pour les nombres premiers ci-dessus. On a aussi inclus le premier 881, à titre de comparaison entre les deux méthodes. La méthode de réduction utilisée est celle de Bennett et de Weger, avec $m - 1$ formes. Cela peut expliquer que la borne réduite pour X soit meilleure que celle obtenue par les fractions continues.

p	m	c_1	c_2	c_7	c_8	B_0	$\mathcal{B}$	B_1	X_5	Tps
251	5	$6,6 \cdot 10^{37}$	$0,001$	$6,6$	$2,0$	$7,5 \cdot 10^{48}$	1876	$3,2 \cdot 10^{8}$	2	61
431	5	$8,2 \cdot 10^{64}$	$0,0004$	$6,5$	$1,0$	$1,6 \cdot 10^{52}$	9838	$6 \cdot 10^{8}$	2	119
491	5	$8,8 \cdot 10^{73}$	$0,0003$	$15,0$	$2,8$	$2,4 \cdot 10^{50}$	659	$6,9 \cdot 10^{6}$	2	98
701	5	$3,6 \cdot 10^{105}$	$0,0001$	$24,2$	$2,7$	$9,9 \cdot 10^{50}$	218	$1,1 \cdot 10^{6}$	2	136
881	4	$4,4 \cdot 10^{132}$	$0,0001$	$17,0$	$2,5$	$4,5 \cdot 10^{45}$	311	$1,5 \cdot 10^{6}$	2	94
911	5	$1,5 \cdot 10^{137}$	$0,00009$	$12,0$	$2,0$	$1,4 \cdot 10^{55}$	13074	$2 \cdot 10^{10}$	2	206
971	5	$1,6 \cdot 10^{146}$	$0,00008$	$13,7$	$2,0$	$4,1 \cdot 10^{55}$	13447	$2,3 \cdot 10^{10}$	2	230

Annexe B

Données numériques pour les équations superelliptiques

On trouve ci-dessus quelques tables numériques omises dans le chapitre 5. Les solutions en gras ont été trouvées à l'étape 9 (énumération des $\boldsymbol{b}$), les autres ont été trouvées lors de l'énumération des petits x. Elles ne sont pas apparues à l'étape 9 parce que $|x - \omega_i| > 1/2$. On les a toutefois notées à l'endroit où elles auraient pu apparaître.

B.1 Le cinquième polynôme de Krawtchouk

B.1.1 Le corps $\mathbb{K}_1$

θ	c_9	c_{10}	B_0	B_{red}	$\text{Sol}(\mathbb{K}, \boldsymbol{k}, \theta)$
$[96, -68, 12, -20, 0, -12, 54, -84]$	$1,14$	$12,0$	$6,1 \cdot 10^{28}$	7	$\mathbf{\{\pm 46\}}, \{\pm 6\}$
$[0, 160, 0, 44, 0, 28, -94, 188]$	$0,67$	$12,0$	$2,8 \cdot 10^{28}$	5	
$[96, 68, 12, 20, 0, 12, -30, 84]$	$1,14$	$12,0$	$6,1 \cdot 10^{28}$	7	$\mathbf{\{\pm 46\}}, \{\pm 6\}$

B.1.2 Le corps $\mathbb{K}_2$

θ	c_9	B_0	B_{red}	$\text{Sol}(\mathbb{K}, \boldsymbol{k}, \theta)$
$[8, 2, 0, -6, 0, 8, 0, 22]$	$1,05$	$4,5 \cdot 10^{28}$	7	$\{\pm 8\}$
$[0, 10, 0, -2, 0, 0, 0, 6]$	$0,67$	$2,4 \cdot 10^{28}$	7	
$[-8, 2, 0, -6, 0, 8, 0, 22]$	$1,05$	$4,5 \cdot 10^{28}$	8	$\{\pm 8\}$

B.1.3 Le corps $\mathbb{K}_3$

θ	c_9	B_0	B_{red}	$\text{Sol}(\mathbb{K}, \boldsymbol{k}, \theta)$
$[-72, 86, -152, 160, 72, -82, -8, 16]$	$0,81$	$1,5 \cdot 10^{29}$	5	$\{\mathbf{\pm 16, \pm 596}, \pm 4\}$
$[0, 46, 0, 124, 0, -62, 0, 12]$	$0,58$	$8,8 \cdot 10^{28}$	4	
$[-72, -86, -152, -160, 72, 82, -8, -16]$	$0,81$	$1,5 \cdot 10^{29}$	5	$\{\mathbf{\pm 16, \pm 596}, \pm 4\}$

B.1.4 Le corps $\mathbb{K}_4$

θ	c_9	B_0	B_{red}	$\text{Sol}(\mathbb{K}, \boldsymbol{k}, \theta)$
$[546, 564, 542, -1624, 0, -540, 0, -72]$	$0,80$	$2,7 \cdot 10^{29}$	10	$\mathbf{\{\pm 2, \pm 22\}}$
$[-270, -324, -270, 808, 0, 268, 0, 36]$	$0,80$	$6,4 \cdot 10^{28}$	10	
$[-538, -564, -542, 1624, 0, 540, 0, 72]$	$0,80$	$6,4 \cdot 10^{28}$	10	$\mathbf{\{\pm 2, \pm 22\}}$

B.2 Méthode des courbes elliptiques

Rappelons que les 3 modèles sont donnés par $V^2 = 2/5U^4 - 8/5U^3 + 2/5U^2 + 12/5U + 9$, $Y^2 + 4/5XY - 48/5Y = X^3 + 6/25X^2 - 72/5X - 432/125$, et $y^2 = x^3 - 1372/75x + 14864/675$.

Les transformations, calculées à l'aide d'Apecs, sont données par

$$(X, Y) = \left(\frac{12U + 30V + 90}{5U^2}, \frac{900V + 2700 + 360U + 36U^2}{25U^3} \right)$$

$$(U, V) = \left(\frac{150X + 36}{25Y}, \frac{-1875Y^2 + 3750X^3 + 1800X^2 - 1500XY + 216X - 360Y}{625Y^2} \right).$$

$$(x, y) = (X + 2/15, -Y - 2X/5 + 24/5) \quad (X, Y) = (x - 2/15, -2x/5 - y + 364/75),$$

$$(U, V) = \left(\frac{-900x - 96}{60x + 150y - 728}, \frac{2700x^2(25x + 4) - 270y(125y - 1224) - 809056}{120x(15x + 75y - 364) + 150y(75y - 128) + 264992} \right),$$

$$(x, y) = \left(\frac{2U^2 + 36U + 90V + 270}{15U^2}, \frac{24U^3 - 12U^2 - 12VU - 108U - 180V - 540}{5U^3} \right).$$

Annexe C

Liste des équations résolues

C.1 Équations de Thue

$y^{19} + 2x^{19} = \pm 1, \pm 2$

Cette équation a pour solutions $(1, -1)$, $(-1, 1)$, $(0, 1)$, $(0, -1)$, $(1, 0)$, $(-1, 0)$ (Théorème 3.1, section 3.2).

$y^4 + xy^3 - 1500x^2y^2 + 23756x^3y - 81536x^4 = \pm 1.$

Cette équation a pour solutions $(-1, 0)$ et $(1, 0)$. (Théorème 3.2, section 3.3).

Équations liées au problème de Lucas-Lehmer

Soit n un nombre entier ; l'équation liée au problème des diviseurs primitifs pour le $n^{\text{ème}}$ terme d'une suite de Lucas et de Lehmer est donnée par

$$\prod_{\substack{1 \leq k \leq n/2 \\ (k,n)=1}} \left(y - 2\cos\left(\frac{2k\pi}{n}\right) x \right) = \pm 1, \pm P^+(n/(n,3)).$$

Pour n une puissance de nombre premier dans l'intervalle $[31, 67]$, n un nombre premier congru à 1 modulo 3, 5 ou 8 dans l'intervalle $[67, 997]$ et $n \in \{83, 4001, 5011\}$, cette équation a pour solutions $(0, \pm 1)$, $(\pm 1, 0)$, $\pm(1, 1)$, $\pm(1, -1)$, $\pm(-1, 2)$, $\pm(1, 2)$ (Théorème 3.11, section 3.4). Le cas $n = 5011$ donne une équation de degré 2505, record actuel.

C.2 Équations hyper- et superelliptiques

$10y^2 = x^4 - 20x^2 + 424$

Les solutions de cette équation sont $(\pm 2, \pm 6)$, $(\pm 4, \pm 6)$, $(\pm 6, \pm 10)$, $(\pm 8, \pm 18)$, $(\pm 22, \pm 150)$, $(\pm 16, \pm 78)$, $(\pm 46, \pm 666)$, $(\pm 596, \pm 112326)$ (Théorème 5.4, section 5.1).

$28y^3 = x^4 - 20x^2 - 32x + 28$

Les solutions de cette équation sont $(-2, 1)$ et $(0, 1)$ (Théorème 5.5, section 5.2.1).

$y^3 = x^4 - x^3 - 3x^2 + x + 1$

Les solutions de cette équation sont $(-1, -1)$,$(0, 1)$,$(1, -1)$, $(2, -1)$ (Théorème 5.6, section 5.2.2).