



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

*Number of solutions to $A^2 + B^2 = C^2 + C$ in a
binade*

Jean-Michel Muller, Jean-Louis Nicolas and Xavier-François Roblot

No 4945

Octobre 2003

———— THÈME 2 ————

A large blue rectangle occupies the bottom right portion of the page. Overlaid on it is a large, light gray 'R' and the text 'Rapport de recherche' in a white, serif font. A horizontal white line is positioned below the text.

*Rapport
de recherche*

Number of solutions to $A^2 + B^2 = C^2 + C$ in a binade

Jean-Michel Muller*, Jean-Louis Nicolas and Xavier-François Roblot†

Thème 2 — Génie logiciel
et calcul symbolique
Projet Arénaire

Rapport de recherche n° 4945 — Octobre 2003 — 36 pages

Abstract: Let us denote by $Q(N, \lambda)$ the number of solutions of the diophantine equation $A^2 + B^2 = C^2 + C$ satisfying $N \leq A \leq B \leq C \leq \lambda N - \frac{1}{2}$. We prove that, for λ fixed and $N \rightarrow \infty$, there exists a constant $\alpha(\lambda)$ such that $Q(N, \lambda) = \alpha(\lambda)N + O_\lambda(N^{7/8} \log N)$. When $\lambda = 2$, $Q(2^{n-1}, 2)$ counts the number of solutions of $A^2 + B^2 = C^2 + C$ with the same number, n , of binary digits; these solutions are interesting in the problem of computing the function $(a, b) \rightarrow \sqrt{a^2 + b^2}$ in radix-2 floating-point arithmetic. By elementary arguments, $Q(N, \lambda)$ can be expressed in terms of four sums of the type

$$S(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ odd}}} \left(\sum_{\substack{1 \leq A \leq f(d) \\ 4A^2 \equiv -1 \pmod{d}}} 1 \right)$$

where u and v are real numbers and $f : [u, v] \rightarrow \mathbb{R}$ is a function. These sums are estimated by a classical, but deep, method of number theory, using Fourier analysis and Kloosterman sums. This method is effective, and, in the case $\lambda = 2$, a precise upper bound for $|Q(N, \lambda) - \alpha(\lambda)N|$ is given.

Key-words: Computer arithmetic, Correct rounding, Diophantine equations.

(Résumé : *tsvp*)

* CNRS, Lab. LIP, projet CNRS-ENS Lyon-Inria Arénaire

† Institut Girard Desargues, UMR 5028

Nombre de solutions dans une binade de l'équation

$$A^2 + B^2 = C^2 + C$$

Résumé : Notons $Q(N, \lambda)$ le nombre de solutions de l'équation diophantienne $A^2 + B^2 = C^2 + C$ satisfaisant $N \leq A \leq B \leq C \leq \lambda N - \frac{1}{2}$. Nous montrons que pour λ fixé et $N \rightarrow \infty$, il existe une constante $\alpha(\lambda)$ telle que $Q(N, \lambda) = \alpha(\lambda)N + O_\lambda(N^{7/8} \log N)$. Lorsque $\lambda = 2$, $Q(2^{n-1}, 2)$ donne le nombre de solutions de $A^2 + B^2 = C^2 + C$ avec le même nombre n , de chiffres binaires; Ces solutions sont utiles pour le calcul de la fonction $(a, b) \rightarrow \sqrt{a^2 + b^2}$ en arithmétique virgule flottante.

Des arguments élémentaires nous permettent d'exprimer $Q(N, \lambda)$ en termes de quatre sommes de la forme

$$S(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ odd}}} \left(\sum_{\substack{1 \leq A \leq f(d) \\ 4A^2 \equiv -1 \pmod{d}}} 1 \right)$$

où u et v sont des nombres réels et $f : [u, v] \rightarrow \mathbb{R}$ est une fonction. Ces sommes sont estimées à l'aide d'une méthode classique mais profonde de théorie des nombres, qui fait appel à l'analyse de Fourier et aux sommes de Kloosterman. Cette méthode est effective, et dans le cas $\lambda = 2$, une borne supérieure précise pour $|Q(N, \lambda) - \alpha(\lambda)N|$ est donnée.

Mots-clé : Arithmétique des ordinateurs, Arrondi correct, Equations diophantiennes.

1 Introduction.

1.1 Présentation du problème.

Le problème informatique original est le suivant (cf. [15], chap. 9) : on désire construire des systèmes (programmes ou circuits intégrés) capables de calculer la fonction $(a, b) \rightarrow \sqrt{a^2 + b^2}$ avec “arrondi correct”, pour a et b compris entre 1 et 2, lorsque le résultat est lui aussi entre 1 et 2. Par “arrondi correct” on entend que le résultat fourni doit toujours être le nombre exactement représentable en virgule flottante avec n bits de mantisse (on dira “nombre machine” pour faire court) le plus proche du résultat exact. Pour ceci, il faut savoir avec quelle précision on doit approcher $\sqrt{a^2 + b^2}$ lors de calculs intermédiaires pour être certain qu’arrondir au plus près l’approximation est équivalent à arrondir au plus près le résultat exact. Ceci revient à déterminer deux nombres machines a et b compris entre 1 et 2 tels que $a^2 + b^2$ soit le plus proche du carré du milieu de deux nombres machines consécutifs. En définissant $A = 2^{n-1}a$ et $B = 2^{n-1}b$, notre problème est ramené à trouver des entiers A , B et C compris entre 2^{n-1} et $2^n - 1$ tels que

$$A^2 + B^2 = (C + 1/2)^2 + \epsilon,$$

où $|\epsilon|$ est le plus petit possible. Il est clair que $|\epsilon|$ vaut au moins $1/4$. Nous cherchons à savoir si le cas $|\epsilon| = 1/4$ se produit. Notre problème est donc de déterminer si pour toute valeur de n il existe des nombres entiers A , B et C compris entre 2^{n-1} et $2^n - 1$ satisfaisant l’équation diophantienne

$$(1.1) \quad A^2 + B^2 = C^2 + C.$$

Désignons par $Q(2^k, 2)$ (cette notation sera justifiée dans l’énoncé du théorème 1 ci-dessous) le nombre de solutions de (1.1) telles que

$$(1.2) \quad 2^k \leq A \leq B \leq C \leq 2^{k+1} - 1.$$

On trouvera dans la table 1 la valeur de $Q(2^k, 2)$ pour $4 \leq k \leq 40$ (on a $Q(1, 2) = 1$ et, pour $1 \leq k \leq 3$, $Q(2^k, 2) = 0$). Cette table montre que $Q(2^k, 2)/2^k$ semble tendre vers une limite. L’objet de cet article est de démontrer :

Theorème 1. *Soit $\lambda > \sqrt{2}$ un nombre réel fixé. Pour N entier, on désigne par $Q(N, \lambda)$ le nombre de solutions de l’équation diophantienne (1.1) vérifiant*

$$(1.3) \quad N \leq A \leq B \leq C \leq \lambda N - \frac{1}{2}.$$

Si l’on pose

$$(1.4) \quad \alpha(\lambda) = \frac{\lambda}{4} - \frac{\lambda}{\pi} \arcsin\left(\frac{1}{\lambda}\right) + \frac{\log(1 + \sqrt{2})}{\pi} - \frac{\log(\lambda + \sqrt{\lambda^2 - 1})}{\pi}$$

on a

$$(1.5) \quad Q(N, \lambda) = \alpha(\lambda)N + \tilde{R}(N, \lambda)$$

avec, lorsque $N \rightarrow \infty$,

$$(1.6) \quad \tilde{R}(N, \lambda) = O_\lambda(N^{7/8} \log N).$$

Pour $\lambda = 2$, on a

$$(1.7) \quad \alpha(2) = \frac{1}{6} + \frac{\log(1 + \sqrt{2})}{\pi} - \frac{\log(2 + \sqrt{3})}{\pi} = 0.02801587455727\dots$$

et, pour $N \geq 1000$,

$$(1.8) \quad |\tilde{R}(N, 2)| \leq N^{7/8}(51 \log N + 211) + N^{1/2}(77 \log N + 283) + 29.$$

Lorsque $\lambda = 2$, nous avons voulu donner une majoration effective du reste $\tilde{R}(N, 2)$. Cette majoration n'est pas très bonne, et le membre de droite de (1.8) n'est inférieur au terme principal $\alpha(\lambda)N$ de (1.5) que pour $N > 1.32 \cdot 10^{42}$. Au prix de calculs encore plus techniques, il serait possible de l'améliorer un peu. Par exemple, dans la majoration (2.19) des sommes de Kloosterman, on peut remplacer (cf. [13], chap. 4) $\tau(m) = \sum_{d|m} 1$ par $2^{\omega(m)}$ où $\omega(m)$ est le nombre de facteurs premiers de m , ce qui permettrait de diminuer le membre de droite de (1.8). Mais, pour ne pas trop alourdir la présentation, nous avons effectué assez grossièrement la plupart des majorations. Cependant, il semble difficile, par cette méthode, d'obtenir pour $\tilde{R}(N, 2)$ une majoration 100 fois meilleure que (1.8).

1.2 Les diviseurs de $4A^2 + 1$.

Ecrivons $C^2 + C = (C + \frac{1}{2})^2 - \frac{1}{4}$; par le changement de variable

$$(1.9) \quad X = 2A, \quad Y = 2B, \quad Z = 2C + 1.$$

l'équation (1.1) devient

$$(1.10) \quad X^2 + Y^2 = Z^2 - 1.$$

En considérant (1.10) modulo 4, on voit que toute solution X, Y, Z de (1.10) est telle que Z est impair et X et Y sont pairs. Donc, toute solution de (1.10) engendre, par (1.9), une solution de (1.1) et les deux équations diophantiennes (1.1) et (1.10) sont équivalentes.

En écrivant (1.10) sous la forme

$$X^2 + 1 = Z^2 - Y^2 = (Z - Y)(Z + Y),$$

on voit que $Z - Y = d$ est un diviseur de $X^2 + 1$. Réciproquement, à un diviseur d de $X^2 + 1$, la résolution du système

$$\begin{cases} Z - Y = d \\ Z + Y = (X^2 + 1)/d \end{cases}$$

fournit

$$(1.11) \quad \begin{cases} Y = \frac{1}{2} \left(\frac{X^2+1}{d} - d \right) \\ Z = \frac{1}{2} \left(\frac{X^2+1}{d} + d \right) \end{cases} \quad \text{soit} \quad \begin{cases} B = \frac{1}{4} \left(\frac{4A^2+1}{d} - d \right) \\ C = \frac{1}{4} \left(\frac{4A^2+1}{d} + d - 2 \right). \end{cases}$$

Notons que, dans (1.11), B et C sont entiers ; en effet les facteurs premiers de $4A^2 + 1$ sont tous congrus à 1 (mod 4) (car -1 est un carré modulo un tel facteur premier) et d , qui divise $4A^2 + 1$, est aussi congru à 1 (mod 4).

Il y a donc, par (1.11), une bijection entre les diviseurs d de $4A^2 + 1$ et les solutions de (1.1), pour A fixé. Dans cette bijection, la condition $B \leq C$ de (1.3) se traduit par $d \geq 1$ et la condition $A \leq B$ par $4A^2 - 4Ad + 1 - d^2 \geq 0$, soit

$$(1.12) \quad \begin{aligned} A \leq B \leq C &\iff A \geq f_2(d) \stackrel{\text{def}}{=} \frac{1}{2} \left(d + \sqrt{2d^2 - 1} \right) \quad \text{et} \quad d \geq 1 \\ &\iff 1 \leq d \leq f_2^{-1}(A) = \sqrt{8A^2 + 1} - 2A. \end{aligned}$$

Parallèlement, la condition $C \leq \lambda N - 1/2$ de (1.3) se traduit par $4A^2 + d^2 - 4\lambda Nd + 1 \leq 0$; autrement dit, le point (A, d) doit être à l'intérieur de la demi ellipse $4A^2 + (d - 2\lambda N)^2 = 4\lambda^2 N^2 - 1$, $A > 0$. Ceci implique $0 < d < 4\lambda N$ et $A < \lambda N$. D'autre part, par (1.11), la condition $B > 0$ entraîne $d \leq 2A < 2\lambda N$, et l'on a

$$(1.13) \quad \begin{aligned} C \leq \lambda N - 1/2 &\iff A \leq f_1(d) \stackrel{\text{def}}{=} \frac{1}{2} \sqrt{d(4\lambda N - d) - 1} \\ &\iff d \geq f_1^{-1}(A) = 2\lambda N - \sqrt{4\lambda^2 N^2 - 4A^2 - 1}. \end{aligned}$$

Les courbes représentatives des fonctions $f_1^{-1}(A)$ et $f_2^{-1}(A)$ se coupent au point (A_0, v_0)

$$(1.14) \quad A_0 = \frac{1}{4} \sqrt{8\lambda^2 N^2 - 2}, \quad v_0 = f_1^{-1}(A_0) = f_2^{-1}(A_0) = 2\lambda N - \frac{1}{2} \sqrt{8\lambda^2 N^2 - 2}.$$

Ainsi, la condition (1.3) équivaut à choisir le point (A, d) tel que $N \leq A \leq A_0$ et $f_1^{-1}(A) \leq d \leq f_2^{-1}(A)$. Si $\lambda \leq \sqrt{2}$, par (1.14), $A_0 < N$ et $Q(N, \lambda) = 0$.

On suppose donc $\lambda > \sqrt{2}$ et l'on a

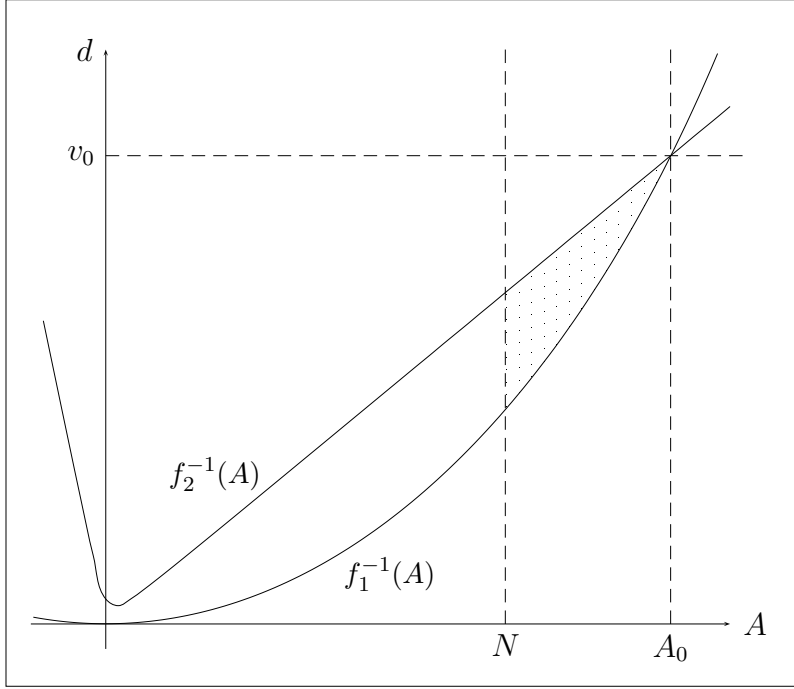
$$(1.15) \quad Q(N, \lambda) = \sum_{N \leq A \leq A_0} \sum_{\substack{d | 4A^2 + 1 \\ f_1^{-1}(A) \leq d \leq f_2^{-1}(A)}} 1.$$

Permutons l'ordre des sommations. Les valeurs extrêmes prises par d sont

$$(1.16) \quad u_0 = f_1^{-1}(N) = 2\lambda N - \sqrt{4(\lambda^2 - 1)N^2 - 1}$$

et v_0 défini par (1.14), car les fonctions $f_1^{-1}(A)$ et $f_2^{-1}(A)$ sont croissantes sur l'intervalle $[N, A_0]$. D'autre part, pour d fixé, A doit vérifier $A \geq \max(N, f_2(d))$ et, en posant

$$(1.17) \quad u_1 = f_2^{-1}(N) = \sqrt{8N^2 + 1} - 2N,$$

FIG. 1 – les valeurs possibles pour (A, d)

l'égalité (1.15) devient

$$(1.18) \quad Q(N, \lambda) = \sum_{u_0 \leq d \leq v_0} \sum_{\substack{N \leq A \leq f_1(d) \\ 4A^2 + 1 \equiv 0 \pmod{d}}} 1 - \sum_{u_1 \leq d \leq v_0} \sum_{\substack{N \leq A < f_2(d) \\ 4A^2 + 1 \equiv 0 \pmod{d}}} 1 = \widehat{Q}(N, \lambda) + \varepsilon(N, \lambda)$$

avec

$$(1.19) \quad \widehat{Q}(N, \lambda) = \sum_{u_0 \leq d \leq v_0} \sum_{\substack{N \leq A \leq f_1(d) \\ 4A^2 + 1 \equiv 0 \pmod{d}}} 1 - \sum_{u_1 \leq d \leq v_0} \sum_{\substack{N \leq A \leq f_2(d) \\ 4A^2 + 1 \equiv 0 \pmod{d}}} 1$$

et

$$(1.20) \quad \varepsilon(N, \lambda) = \sum_{u_1 \leq d \leq v_0} \sum_{\substack{A = f_2(d) \\ 4A^2 + 1 \equiv 0 \pmod{d}}} 1.$$

Définissons

$$(1.21) \quad S(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \left(\sum_{\substack{1 \leq A \leq f(d) \\ 4A^2 \equiv -1 \pmod{d}}} 1 \right).$$

L'égalité (1.19) s'écrit

$$(1.22) \quad \widehat{Q}(N, \lambda) = S(u_0, v_0; f_1) - S(u_0, v_0; N-1) - S(u_1, v_0; f_2) + S(u_1, v_0; N-1)$$

où $N-1$ désigne la fonction constante égale à $N-1$. Par (1.22), l'évaluation de $\widehat{Q}(N, \lambda)$ se ramène à l'évaluation de sommes $S(u, v; f)$.

1.3 Les sommes $S(\mathbf{u}, \mathbf{v}; \mathbf{f})$.

Nous démontrerons au paragraphe 3 le théorème

Theorème 2. Soit u, v deux nombres réels satisfaisant $7 \leq u \leq v$ et f une fonction continûment dérivable de $[u, v]$ dans \mathbb{R} . On pose

$$(1.23) \quad \|f\| = \|f\|_{[u,v]} = \max_{u \leq t \leq v} |f(t)|,$$

$$(1.24) \quad F(t) = \frac{f(t)}{t}$$

et

$$(1.25) \quad \mathcal{M} = \mathcal{M}(u, v; f) = \|tF'(t)\| + \frac{2}{5} = \max_{u \leq t \leq v} |tF'(t)| + \frac{2}{5}.$$

Alors la somme $S(u, v; f)$ définie en (1.21) peut s'écrire

$$(1.26) \quad S(u, v; f) = S_0(u, v; f) - S_1(u, v; f)$$

avec

$$(1.27) \quad S_0(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \frac{\rho(d)f(d)}{d}$$

où ρ est défini par (2.11) et $S_1(u, v; f)$ satisfait l'inégalité

$$(1.28) \quad |S_1(u, v; f)| \leq \frac{19}{2}(\log v + 4)v^{7/8}\sqrt{\mathcal{M}}.$$

La démonstration du théorème 2 suit le chapitre 2 du livre de C. Hooley [11], où, pour prouver que le plus grand facteur premier de $\prod_{n \leq x} (n^2 + 1)$ est, pour x assez grand, supérieur à $x^{11/10}$, l'auteur évalue une somme voisine de $S(u, v; x)$. L'exposant 11/10 est amélioré dans [4]. Un résultat plus général est démontré dans [10], où $n^2 + 1$ est remplacé par $n^2 + D$, avec $n^2 + D$ irréductible. Ce résultat pourrait être adapté pour compter les racines de l'équation diophantienne $X^2 + D = Z^2 - Y^2$ dans un intervalle donné.

1.4 Les solutions $\mathbf{A} = \mathbf{B}$.

Si l'on impose $A = B$ dans (1.1), on a, par (1.9), $X = Y$, et l'équation équivalente (1.10) s'écrit

$$Z^2 - 2X^2 = 1;$$

c'est une équation de Pell-Fermat, dont les solutions sont X_n, Z_n avec $Z_n + \sqrt{2}X_n = (3 + 2\sqrt{2})^n$. On a

$$(1.29) \quad X_n = \frac{(3 + 2\sqrt{2})^n - (3 - 2\sqrt{2})^n}{2\sqrt{2}} \leq \frac{(3 + 2\sqrt{2})^n}{2\sqrt{2}}$$

et, pour $n \geq 1$,

$$(1.30) \quad \begin{aligned} X_n &= \frac{(3 + 2\sqrt{2})^n}{2\sqrt{2}} \left(1 - \left(\frac{3 - 2\sqrt{2}}{3 + 2\sqrt{2}} \right)^n \right) \\ &\geq \frac{(3 + 2\sqrt{2})^n}{2\sqrt{2}} \left(1 - \frac{3 - 2\sqrt{2}}{3 + 2\sqrt{2}} \right) \geq \frac{(3 + 2\sqrt{2})^n}{3}. \end{aligned}$$

Par la définition (1.20) de $\varepsilon(N, \lambda)$, la condition $A = f_2(d)$ est, par (1.12), équivalente à $A = B$. Comme la fonction f_2 définie en (1.12) est croissante, la fonction $\varepsilon(N, \lambda)$ compte le nombre de solutions de l'équation (1.1) avec $A = B$ telle que A soit compris entre $f_2(u_1) = N$ et $f_2(v_0) = A_0 < \frac{\lambda N}{\sqrt{2}}$ par (1.17) et (1.14). On a donc, par (1.29) et (1.30)

$$(1.31) \quad 0 \leq \varepsilon(N, \lambda) = \sum_{2N \leq X_n \leq 2A_0} 1 \leq \sum_{2N \leq X_n < \lambda\sqrt{2}N} 1 \leq \frac{\log \frac{3\lambda}{4}}{\log(3 + 2\sqrt{2})} + 1$$

et, en particulier, pour $\lambda = 2$,

$$(1.32) \quad 0 \leq \varepsilon(N, \lambda) \leq 1.$$

1.5 Les solutions paramétriques.

On peut donner une représentation paramétrique des solutions de l'équation (1.10) (cf. [5], th. 38 et 39 et [6], p. 210). En effet, si $\alpha, \beta, \gamma, \delta$ sont des entiers vérifiant

$$(1.33) \quad \alpha\delta - \beta\gamma = 1$$

il est facile de voir que

$$(1.34) \quad X = \alpha\beta + \gamma\delta, Y = \frac{1}{2}(\alpha^2 + \gamma^2 - \beta^2 - \delta^2), Z = \frac{1}{2}(\alpha^2 + \gamma^2 + \beta^2 + \delta^2)$$

est une solution de (1.10).

Réciproquement, pour voir que toute solution de (1.10) est de cette forme, écrivons la décomposition en facteurs premiers $X^2 + 1 = \prod_{j=1}^J p_j^{\alpha_j}$, où les nombres premiers p_j sont congrus à 1 modulo 4. Dans l'anneau \mathcal{G} des entiers de Gauss (c'est-à-dire l'ensemble des nombres complexes $a + bi$ où a et b sont dans \mathbb{Z}), chaque p_j se factorise sous la forme $p_j = \pi_j \bar{\pi}_j$. Le nombre $X + i$ se factorise dans \mathcal{G} sous la forme

$$(1.35) \quad X + i = i^k \prod_{j=1}^J \pi_j^{\alpha_j}, \quad 0 \leq k \leq 3.$$

En effet, les deux nombres π_j et $\overline{\pi_j}$ ne peuvent apparaître simultanément dans (1.35), sinon $p_j = \pi_j \overline{\pi_j}$ diviserait (dans \mathbb{Z}) X et 1, ce qui est impossible.

Soit d un diviseur de $X^2 + 1$. La décomposition en facteurs premiers de d s'écrit $d = \prod_{j=1}^J p_j^{\beta_j}$ avec $0 \leq \beta_j \leq \alpha_j$. Posons

$$(1.36) \quad \Delta = i^k \prod_{j=1}^J \pi_j^{\beta_j} = \beta + i\delta, \quad \Delta' = \prod_{j=1}^J \pi_j^{\alpha_j - \beta_j} = \alpha - i\gamma.$$

On a, par (1.35),

$$(1.37) \quad X + i = \Delta \Delta' = (\beta + i\delta)(\alpha - i\gamma) = \alpha\beta + \gamma\delta + i(\alpha\delta - \beta\gamma)$$

qui nous donne (1.33) et la valeur de X dans (1.34). On obtient la valeur de Y et Z dans (1.34) en prenant la norme dans (1.36) :

$$d = |\Delta|^2 = \beta^2 + \delta^2, \quad \frac{X^2 + 1}{d} = |\Delta'|^2 = \alpha^2 + \gamma^2$$

et en reportant dans (1.11).

Soit $\theta \in \mathbb{Z}$. On pose

$$\alpha = 10\theta + 8, \quad \beta = 4\theta + 3, \quad \gamma = 5, \quad \delta = 2.$$

On vérifie (1.33) et en substituant dans (1.34), on obtient

$$X = 40\theta^2 + 62\theta + 34, \quad Y = 42\theta^2 + 68\theta + 38, \quad Z = 58\theta^2 + 92\theta + 51$$

d'où, par (1.9),

$$(1.38) \quad A = 20\theta^2 + 31\theta + 17, \quad B = 21\theta^2 + 34\theta + 19, \quad C = 29\theta^2 + 46\theta + 25.$$

Pour $\theta > 0$, la solution (1.38) vérifie $20\theta^2 \leq A$ et $C + 1/2 \leq 29(\theta + 1)^2$. Elle sera comptée dans $Q(N, \lambda)$ pour $\sqrt{\frac{N}{20}} \leq \theta \leq \sqrt{\frac{\lambda N}{29}} - 1$. On obtient ainsi la minoration

$$(1.39) \quad Q(N, \lambda) \geq \left(\sqrt{\frac{\lambda}{29}} - \sqrt{\frac{1}{20}} \right) \sqrt{N} - 2.$$

Lorsque $\lambda = 2$, (1.39) entraîne, pour $N \geq 2630 > \left(\frac{2}{0.039}\right)^2$,

$$(1.40) \quad Q(N, 2) \geq 0.039\sqrt{N} - 2 > 0.$$

L'inégalité (1.40) et la table 1 montrent que l'équation (1.1) a, pour tout $k \geq 4$, des solutions vérifiant (1.2).

1.6 Structure de l'article.

Dans le §2, nous redémontrons des résultats classiques utiles dans la suite, notamment sur la fonction arithmétique ρ (définie par (2.11)) et les sommes de Kloosterman. Nous donnons également deux algorithmes de construction de la table 1 donnée en annexe. Dans le §3, nous démontrons le théorème 2, et dans le §4, le théorème 1.

Nous remercions chaleureusement A. Schinzel pour nous avoir indiqué le livre de C. Hooley [11] ainsi que la solution paramétrique (1.34). Nous avons plaisir à remercier E. Fouvry, R. Heath-Brown, J. Rivat pour leurs remarques positives. Nous remercions également les systèmes de calcul formel MAPLE [14] et PARI/GP [17] que nous avons largement utilisé, ainsi que le Max Planck Institut de Bonn où une partie des idées de cet article a été développée pendant que le deuxième auteur était invité en mai 2002.

2 Résultats généraux.

2.1 Sommes de deux carrés.

Soit d un entier positif impair. On désigne par \mathcal{S}_d l'ensemble des représentations primitives de d par la forme quadratique $r^2 + 4s^2$, avec $r > 0$ et $s \in \mathbb{Z}$

$$(2.1) \quad \mathcal{S}_d = \{(r, s), r \in \mathbb{N}^*, s \in \mathbb{Z}, r^2 + 4s^2 = d, (r, 2s) = 1\}$$

et par \mathcal{E}_d l'ensemble des classes d'équivalence modulo d satisfaisant la congruence $4\nu^2 \equiv -1 \pmod{d}$:

$$(2.2) \quad \mathcal{E}_d = \{\nu \in (\mathbb{Z}/d\mathbb{Z})^*, 4\nu^2 \equiv -1 \pmod{d}\}.$$

Nous allons redémontrer le résultat classique (cf. [19], articles 86 et 90 et [11], p. 33) :

Proposition 1. *L'application Θ définie par $\Theta(r, s) = r^{-1}s \pmod{d}$, où r^{-1} désigne l'inverse de r dans $(\mathbb{Z}/d\mathbb{Z})^*$ est une bijection de \mathcal{S}_d dans \mathcal{E}_d . On peut écrire $\nu = \Theta(r, s)$ sous la forme*

$$(2.3) \quad \frac{\nu}{d} \equiv \frac{\overline{4s}}{r} + \frac{s}{r(r^2 + 4s^2)} \pmod{1}$$

où l'on a noté $\overline{4s}$ un inverse de $4s$ modulo r ou bien sous la forme

$$(2.4) \quad \frac{\nu}{d} \equiv \frac{\bar{r}}{4s} - \frac{r}{4s(r^2 + 4s^2)} \pmod{1}$$

où, cette fois, nous notons \bar{r} un inverse de r modulo $4s$.

Démonstration.

(i) Θ est bien définie. Si r n'était pas premier avec d , soit p un facteur premier commun à r et d . Comme d est impair, $p \neq 2$, et p divise $4s^2 = d - r^2$, donc p divise s , et r et s ne seraient pas premiers entre eux. Ainsi r^{-1} est bien défini et $r^2 + 4s^2 = d$ entraîne $4s^2 r^{-2} \equiv -1 \pmod{d}$ donc $r^{-1}s \in \mathcal{E}_d$ par (2.2).

(ii) Θ est surjective. Soit $\nu \in \mathcal{E}_d$. On obtient r et s tels que $\Theta(r, s) = \nu$ par l'algorithme de Cornacchia (cf. [16], [2], [1]) : on développe $\frac{2\nu}{d}$ en fraction continue et l'on choisit la réduite $\frac{p_n}{q_n}$ telle que $q_n \leq \sqrt{d} < q_{n+1}$. Les deux nombres $R = q_n$ et $S = 2\nu q_n - dp_n$ vérifient $R^2 + S^2 = d$, $(R, S) = 1$ et $SR^{-1} \equiv 2\nu \pmod{d}$.

Comme $d = R^2 + S^2$ est impair, l'un des nombres R et S est pair, l'autre est impair. Si S est pair, on pose $s = S/2$, $r = R$. On a $(r, s) \in \mathcal{S}_d$ et $\Theta(r, s) = sr^{-1} \equiv \nu \pmod{d}$. Si R est pair, on pose $s = -R/2$, $r = S$; on a $(r, s) \in \mathcal{S}_d$ et

$$\Theta(r, s) = r^{-1}s = -S^{-1}R/2 \equiv -\frac{1}{2}(2\nu)^{-1} \equiv -\frac{1}{2}(-2\nu) \equiv \nu \pmod{d}.$$

(iii) Θ est injective. Soit deux éléments (r, s) et (r', s') de \mathcal{S}_d vérifiant $\Theta(r, s) = \Theta(r', s')$, c'est-à-dire

$$(2.5) \quad r's - rs' \equiv 0 \pmod{d}.$$

On a

$$(2.6) \quad d^2 = (r^2 + 4s^2)(r'^2 + 4s'^2) = (rr' + 4ss')^2 + 4(r's - rs')^2.$$

Mais (2.6) et (2.5) entraînent que $r's - rs' = 0$ c'est-à-dire

$$(2.7) \quad \frac{s}{r} = \frac{s'}{r'} = w$$

(car r et r' , impairs, ne sont pas nuls) et l'on a

$$d = r^2 + 4s^2 = r'^2 + 4s'^2 = r'^2(1 + 4w^2) = r'^2(1 + 4w^2)$$

ce qui implique $r = r'$ puis, par (2.7), $s = s'$.

Il reste à prouver (2.3) et (2.4) Soit $(r, s) \in \mathcal{S}_d$; comme $(r, 2s) = 1$, il existe $z_1, z_2 \in \mathbb{Z}$, tels que

$$(2.8) \quad rz_1 - 4sz_2 = 1.$$

Posons

$$(2.9) \quad \nu = rz_2 + sz_1.$$

En prenant les modules dans la relation

$$(r + 2si)(z_1 + 2z_2i) = (rz_1 - 4sz_2) + 2(rz_2 + sz_1)i$$

on obtient

$$4\nu^2 + 1 \equiv 0 \pmod{d}.$$

Enfin, de (2.8) et (2.9), il suit

$$(2.10) \quad r\nu = r^2z_2 + s(1 + 4sz_2) = s + z_2d \equiv s \pmod{d}$$

donc $\nu = \Theta(r, s)$. Il vient ensuite, par (2.8) et (2.9)

$$\begin{aligned} \frac{\nu}{d} &= \frac{rz_2 + sz_1}{r^2 + 4s^2} = \frac{r^2z_2 + rsz_1}{r(r^2 + 4s^2)} = \frac{s + (r^2 + 4s^2)z_2}{r(r^2 + 4s^2)} \\ &= \frac{z_2}{r} + \frac{s}{r(r^2 + 4s^2)} \equiv \frac{\overline{4s}}{r} + \frac{s}{r(r^2 + 4s^2)} \pmod{1} \end{aligned}$$

ce qui démontre (2.3). De même, on peut écrire

$$\begin{aligned} \frac{\nu}{d} &= \frac{4rsz_2 + 4s^2z_1}{4s(r^2 + 4s^2)} = \frac{-r + (r^2 + 4s^2)z_1}{4s(r^2 + 4s^2)} \\ &= \frac{z_1}{4s} - \frac{r}{4s(r^2 + 4s^2)} \equiv \frac{\overline{r}}{4s} - \frac{r}{4s(r^2 + 4s^2)} \pmod{1} \end{aligned}$$

ce qui prouve (2.4). \square

Nous désignerons par $\rho(n)$ le nombre de solutions de la congruence

$$(2.11) \quad \nu^2 \equiv -1 \pmod{n}.$$

On a pour p premier et $\alpha \geq 1$

$$(2.12) \quad \rho(p^\alpha) = \begin{cases} 1 & \text{si } p = 2 \text{ et } \alpha = 1 \\ 0 & \text{si } p = 2 \text{ et } \alpha \geq 2 \\ 0 & \text{si } p \equiv 3 \pmod{4} \\ 2 & \text{si } p \equiv 1 \pmod{4} \end{cases}$$

et par le théorème des restes des chinois, la fonction arithmétique ρ est multiplicative. Par la proposition 1, pour d impair, on a

$$(2.13) \quad \rho(d) = \text{Card}(\mathcal{E}_d) = \text{Card}(\mathcal{S}_d).$$

2.2 Construction de la table.

Une première méthode consiste à utiliser la formule (1.15). Par (1.14), on a

$$(2.14) \quad N \leq A \leq A_0 = \sqrt{\frac{\lambda^2 N^2}{2} - \frac{1}{8}}.$$

Pour chaque valeur de A vérifiant l'inégalité (2.14), on factorise $4A^2 + 1$, et l'on recherche ses diviseurs. Pour chaque diviseur d de $4A^2 + 1$, on calcule B et C par (1.11) et si (1.3) est satisfaite, on compte cette solution. Cette méthode a pour inconvénient la factorisation de $4A^2 + 1$ qui, pour A grand, devient coûteuse.

Un meilleur algorithme est basé sur la proposition 1. Par une remarque déjà faite juste après (1.11), les facteurs premiers de d , diviseur de $4A^2 + 1$, sont tous congrus à 1 modulo 4, ce qui implique par (2.12), que $\rho(d) > 0$. Par la proposition 1 et (2.13), il s'ensuit que d admet $\rho(d)$ représentations primitives

$$(2.15) \quad d = r^2 + 4s^2, \quad (r, 2s) = 1, \quad r \geq 1, \quad s \in \mathbb{Z}.$$

Par (1.18), on sait que d est compris entre u_0 , défini par (1.16), et v_0 , défini par (1.14). Par une double boucle en r et s , on engendre les nombres d de la forme (2.15) et vérifiant $u_0 \leq d \leq v_0$. Observons que chaque d va être engendré $\rho(d)$ fois.

Ensuite, pour chaque valeur du couple (r, s) , on calcule ν par (2.8) et (2.9) et l'on recherche les nombres A satisfaisant $A \equiv \nu \pmod{d}$ et (2.14). (Notons que, pour $\lambda = 2$, il y a au plus un tel A car, par (1.16), $d \geq u_0 > \sqrt{2N^2 - 1/8} - N$). Comme dans la première méthode, à partir de A et d , on calcule B et C par (1.11) et on incrémente le compteur de solutions si (1.3) est satisfaite.

Cet algorithme est linéaire en N . Il a été implémenté en PARI/C. Le temps d'exécution est de ??? ¹

2.3 Lemmes analytiques.

Lemme 1. Soit $\tau(n) = \sum_{d|n} 1$. Pour tout $x \geq 1$, on a

$$\sum_{1 \leq n \leq x} \tau(n) \leq x(\log x + 1).$$

Démonstration. On a

$$\sum_{1 \leq n \leq x} \tau(n) = \sum_{1 \leq n \leq x} \sum_{d|n} 1 = \sum_{d \leq x} \left\lfloor \frac{x}{d} \right\rfloor \leq x \sum_{d \leq x} \frac{1}{d}.$$

Puis on utilise l'inégalité

$$(2.16) \quad \sum_{d \leq x} \frac{1}{d} \leq 1 + \int_1^x \frac{dt}{t} = 1 + \log x. \quad \square$$

Lemme 2. Soit h un nombre entier non nul et $\tau(n) = \sum_{d|n} 1$. Pour tout $x \geq 1$, on a

$$\sum_{1 \leq n \leq x} \tau(n) \{(h, n)\}^{1/2} \leq x(\log x + 1) \psi(h)$$

où $\psi(h)$ est défini par

$$(2.17) \quad \psi(h) = \sum_{d|h} \frac{\tau(d)}{\sqrt{d}}.$$

Démonstration. En utilisant l'inégalité $\tau(dd') \leq \tau(d)\tau(d')$, valable pour tout couple d'entiers d et d' et le lemme 1, il vient

$$\begin{aligned} \sum_{1 \leq n \leq x} \tau(n) \{(h, n)\}^{1/2} &\leq \sum_{d|h} \sqrt{d} \sum_{\substack{n \leq x \\ d|n}} \tau(n) \leq \sum_{d|h} \sqrt{d} \tau(d) \sum_{d' \leq x/d} \tau(d') \\ &\leq \sum_{d|h} \frac{\tau(d)}{\sqrt{d}} x \left(\log \left(\frac{x}{d} \right) + 1 \right) \leq x(\log x + 1) \psi(h). \quad \square \end{aligned}$$

¹à compléter par X. R.

Lemme 3. Soit $\psi(h)$ défini par (2.17). Pour tout $x \geq 1$, on a

$$\sum_{1 \leq h \leq x} \psi(h) \leq 7x.$$

Démonstration. On a

$$\sum_{1 \leq h \leq x} \psi(h) = \sum_{1 \leq h \leq x} \sum_{d|h} \frac{\tau(d)}{\sqrt{d}} = \sum_{d \leq x} \frac{\tau(d)}{\sqrt{d}} \left\lfloor \frac{x}{d} \right\rfloor \leq x \sum_{d=1}^{\infty} \frac{\tau(d)}{d^{3/2}} = x\zeta(3/2)^2$$

et $\zeta(3/2)^2 = 6.8245\dots$ \square

Lemme 4. Soit $r(n)$ le nombre de façons d'écrire n comme somme de deux carrés. Pour tout $x \geq 1$, on a

$$\sum_{1 \leq n \leq x} r(n) = \pi x + \widehat{R}_1(x) \quad \text{avec } |\widehat{R}_1(x)| \leq 9\sqrt{x}.$$

Démonstration. L'argument classique (cf. [9], théorème 339) de comptage des points à coordonnées entières dans le disque centré à l'origine et de rayon \sqrt{x} donne

$$\pi(\sqrt{x} - \sqrt{2})^2 \leq 1 + \sum_{1 \leq n \leq x} r(n) \leq \pi(\sqrt{x} + \sqrt{2})^2.$$

Il s'ensuit que

$$|\widehat{R}_1(x)| \leq 2\sqrt{2}\pi\sqrt{x} + (2\pi - 1) \leq 9\sqrt{x}$$

pour $x \geq 2500$. Le calcul direct de $r(1)+r(2)+\dots+r(n)$ montre que, pour $1 \leq x \leq 2500$, on a $\widehat{R}_1(x) \leq 2.29\sqrt{x}$. W. Sierpinski a amélioré la majoration de $|\widehat{R}_1(x)|$, cf. [18]. Le record est actuellement détenu par [12]. \square

Lemme 5. Soit ρ la fonction multiplicative définie par (2.11) et (2.12). Alors, pour tout $x \geq 1$, on a

$$(i) \quad \Upsilon(x) = \sum_{1 \leq n \leq x} \rho(n) = \frac{3}{2\pi}x + \widehat{R}_2(x) \quad \text{avec } |\widehat{R}_2(x)| \leq \left(\frac{9}{8} \log x + 4\right) \sqrt{x},$$

$$(ii) \quad \Upsilon^*(x) = \sum_{\substack{1 \leq n \leq x \\ n \text{ impair}}} \rho(n) = \frac{1}{\pi}x + \widehat{R}_3(x) \quad \text{avec } |\widehat{R}_3(x)| \leq (4 \log x + 14) \sqrt{x},$$

et pour $x \geq 7$, on a

$$(iii) \quad \Upsilon^*(x) \leq \frac{3x}{7}.$$

Démonstration. (i) On sait que la fonction $\frac{1}{4}r(n)$ définie dans le lemme 4 est multiplicative et vaut $\frac{1}{4}r(2^\alpha) = 1$ si $\alpha \geq 1$, $\frac{1}{4}r(p^\alpha) = 0$ si $p \equiv 3 \pmod{4}$ et α impair, $\frac{1}{4}r(p^\alpha) = 1$ si $p \equiv 3 \pmod{4}$ et α pair, $\frac{1}{4}r(p^\alpha) = \alpha + 1$ si $p \equiv 1 \pmod{4}$ (cf. [9], théorème 278).

Désignons par μ la fonction de Möbius. Un raisonnement classique montre que la fonction $\frac{1}{4} \sum_{d^2|n} \mu(d)r(n/d^2)$ est également multiplicative et que sa valeur sur les puissances de nombres premiers coïncide avec celle de ρ . On a donc

$$\rho(n) = \frac{1}{4} \sum_{d^2|n} \mu(d)r(n/d^2).$$

Par le lemme 4, on a

$$\begin{aligned} \Upsilon(x) &= \sum_{n \leq x} \frac{1}{4} \sum_{d^2|n} \mu(d)r(n/d^2) = \frac{1}{4} \sum_{d \leq \sqrt{x}} \mu(d) \sum_{d' \leq x/d^2} r(d') \\ &= \frac{1}{4} \sum_{d \leq \sqrt{x}} \mu(d) \left(\pi \frac{x}{d^2} + \widehat{R}_1 \left(\frac{x}{d^2} \right) \right) = \frac{3x}{2\pi} + \widehat{R}_2(x) \end{aligned}$$

avec

$$\widehat{R}_2(x) = -\frac{\pi x}{4} \sum_{d > \sqrt{x}} \frac{\mu(d)}{d^2} + \frac{1}{4} \sum_{d \leq \sqrt{x}} \widehat{R}_1 \left(\frac{x}{d^2} \right).$$

Par le lemme 4 et (2.16), on obtient

$$\begin{aligned} |\widehat{R}_2(x)| &\leq \frac{\pi x}{4} \left(\frac{1}{x} + \int_{\sqrt{x}}^{+\infty} \frac{dt}{t^2} \right) + \frac{1}{4} \sum_{d \leq \sqrt{x}} 9 \frac{\sqrt{x}}{d} \\ &\leq \frac{\pi \sqrt{x}}{2} + \frac{9}{4} \sqrt{x} \left(1 + \frac{1}{2} \log x \right) \leq \left(\frac{9}{8} \log x + 4 \right) \sqrt{x}. \end{aligned}$$

(ii) Compte tenu de la valeur de $\rho(2^\alpha)$ dans (2.12), on a

$$\begin{aligned} \sum_{\substack{1 \leq n \leq x \\ n \text{ pair}}} \rho(n) &= \sum_{\substack{1 \leq n \leq x \\ n \equiv 2 \pmod{4}}} \rho(n) = \sum_{\substack{1 \leq n \leq x/2 \\ n \text{ impair}}} \rho(n) \\ &= \sum_{1 \leq n \leq x/2} \rho(n) - \sum_{\substack{1 \leq n \leq x/2 \\ n \text{ pair}}} \rho(n). \end{aligned}$$

En réitérant l'argument, on obtient

$$\Upsilon^*(x) = \sum_{j=0}^k (-1)^j \Upsilon \left(\frac{x}{2^j} \right)$$

où k est l'entier défini par $2^k \leq x < 2^{k+1}$. En appliquant (i), il vient

$$\Upsilon^*(x) = \sum_{j=0}^k (-1)^j \left(\frac{3x}{2^{j+1}\pi} + \widehat{R}_2 \left(\frac{x}{2^j} \right) \right) = \frac{x}{\pi} + \widehat{R}_3(x)$$

avec

$$\widehat{R}_3(x) = -\frac{3}{2\pi} \sum_{j=k+1}^{+\infty} (-1)^j \frac{x}{2^j} + \sum_{j=0}^k (-1)^j \widehat{R}_2\left(\frac{x}{2^j}\right).$$

On a

$$\begin{aligned} |\widehat{R}_3(x)| &\leq \frac{3}{2\pi} \sum_{j=k+1}^{+\infty} \sqrt{\frac{x}{2^j}} + \left(\frac{9}{8} \log x + 4\right) \sum_{j=0}^k \sqrt{\frac{x}{2^j}} \\ &\leq (2 + \sqrt{2}) \left(\frac{9}{8} \log x + 4\right) \sqrt{x} \leq (4 \log x + 14) \sqrt{x}. \end{aligned}$$

(iii) La fonction $\frac{4 \log t + 14}{\sqrt{t}}$ est décroissante pour $t \geq 1$ et elle est inférieure à $1/10$ pour $t \geq 500000$; (iii) résulte donc de (ii) pour $x \geq 500000$. Le calcul numérique de $\Upsilon^*(x)$ pour x entier inférieure à 500000 complète la preuve de (iii). \square

Lemme 6. Soient deux nombres réels u et v satisfaisant $1 \leq u \leq v$. Soit F une fonction continûment dérivable de $[u, v]$ dans \mathbb{R} , et ρ défini par (2.11). On pose $\|F\| = \|F\|_{[u,v]} = \max_{u \leq t \leq v} |F(t)|$. Alors on a

$$\sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \rho(d) F(d) = \frac{1}{\pi} \int_u^v F(t) dt + \widehat{R}_0(u, v; F)$$

avec

$$\left| \widehat{R}_0(u, v; F) \right| \leq (4 \log v + 14) \sqrt{v} \left(2 \|F\| + \frac{2}{3} v \|F'\| \right).$$

Démonstration. Par le lemme 5 (ii), on a, en utilisant l'intégrale de Stiltjes

$$\sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \rho(d) F(d) = \int_{u^-}^v F(t) d[\Upsilon^*(t)] = \frac{1}{\pi} \int_u^v F(t) dt + \widehat{R}_0(u, v; F)$$

avec, par intégration par parties,

$$\widehat{R}_0(u, v; F) = \int_{u^-}^v F(t) d[\widehat{R}_3(t)] = F(v) \widehat{R}_3(v) - F(u) \widehat{R}_3(u) - \int_u^v F'(t) \widehat{R}_3(t) dt$$

et

$$\left| \widehat{R}_0(u, v; F) \right| \leq 2 \|F\| (4 \log v + 14) \sqrt{v} + \|F'\| (4 \log v + 14) \int_u^v \sqrt{t} dt$$

et comme $\int_u^v \sqrt{t} dt \leq \int_0^v \sqrt{t} dt = \frac{2}{3} v^{3/2}$ cela achève la preuve du lemme 6.

2.4 Sommes de Kloosterman.

Soit $a \in \mathbb{Z}$, $b \in \mathbb{Z}$ et m un entier positif. Si ℓ est premier avec m , on notera $\bar{\ell}$ un entier vérifiant $\ell\bar{\ell} \equiv 1 \pmod{m}$. La somme de Kloosterman $K(a, b; m)$ est définie par

$$(2.18) \quad K(a, b; m) = \sum_{\substack{0 < \ell \leq m \\ (\ell, m) = 1}} \exp \frac{2i\pi(a\ell + b\bar{\ell})}{m}.$$

Grâce aux travaux de A. Weil, on sait majorer cette somme, et l'on a l'inégalité (cf. [21], [11], p. 35 ou [13], p. 61)

$$(2.19) \quad |K(a, b; m)| \leq \tau(m)\sqrt{m}\{(a, b, m)\}^{1/2}, \quad a, b, m \in \mathbb{Z}, m \geq 1,$$

où $\tau(m) = \sum_{d|m} 1$ désigne le nombre de diviseurs de m et (a, b, m) le pgcd des trois nombres a, b et m . On peut déduire de (2.19) une majoration de la somme de Kloosterman incomplète :

Lemme 7 (cf. [11], Lemma 6, p. 36 et [3], p. 433). *Soit $m \geq 1$ un entier, et $b \in \mathbb{Z}$. Soit ξ et ξ' deux nombres entiers vérifiant $1 \leq \xi \leq m$ et $1 \leq \xi' \leq m$. On a*

$$(2.20) \quad \left| \sum_{\substack{\xi \leq \ell \leq \xi' \\ (\ell, m) = 1}} \exp \frac{2i\pi b\bar{\ell}}{m} \right| \leq \tau(m)\sqrt{m}\{(b, m)\}^{1/2}\gamma(m)$$

où

$$(2.21) \quad \gamma(m) = 1 + \frac{2}{\pi} \log \frac{4m}{\pi} = 1.153\dots + 0.6366\dots \log m.$$

Démonstration. Si $\xi' < \xi$, la somme dans (2.20) est nulle et le lemme est démontré; on peut donc supposer $\xi \leq \xi'$.

En observant que le crochet ci-dessous vaut 1 si $n \equiv \ell \pmod{m}$ (ce qui équivaut à $n = \ell$) et 0 sinon, il vient

$$\begin{aligned} \sum_{\substack{\xi \leq \ell \leq \xi' \\ (\ell, m) = 1}} \exp \frac{2i\pi b\bar{\ell}}{m} &= \sum_{\substack{0 < \ell \leq m \\ (\ell, m) = 1}} \exp \frac{2i\pi b\bar{\ell}}{m} \sum_{\xi \leq n \leq \xi'} \left[\frac{1}{m} \sum_{1 \leq a \leq m} e^{\frac{2i\pi a(n-\ell)}{m}} \right] \\ &= \sum_{1 \leq a \leq m} K(-a, b; m) \left\{ \frac{1}{m} \sum_{\xi \leq n \leq \xi'} \exp \frac{2i\pi an}{m} \right\}, \end{aligned}$$

ce qui, en utilisant (2.19), entraîne

$$(2.22) \quad \left| \sum_{\substack{\xi \leq \ell \leq \xi' \\ (\ell, m) = 1}} \exp \frac{2i\pi b\bar{\ell}}{m} \right| \leq \tau(m)\sqrt{m}\{(b, m)\}^{1/2} \sum_{1 \leq a \leq m} \left\{ \left| \frac{1}{m} \sum_{\xi \leq n \leq \xi'} \exp \frac{2i\pi an}{m} \right| \right\}.$$

La quantité dans l'accolade vaut, pour $a = m$, $|\xi' - \xi + 1|/m \leq 1$, et pour $1 \leq a < m$, elle vaut

$$\left| \frac{\exp \frac{2i\pi a \xi}{m} - \exp \frac{2i\pi a (\xi'+1)}{m}}{m (\exp \frac{2i\pi a}{m} - 1)} \right| \leq \frac{1}{m \sin \frac{\pi a}{m}}.$$

La somme en a de (2.22) est donc majorée par

$$(2.23) \quad 1 + \frac{1}{m} \sum_{a=1}^{m-1} \frac{1}{\sin \frac{\pi a}{m}}.$$

Mais une fonction f convexe sur l'intervalle $[a-1/2, a+1/2]$ satisfait $f(a) \leq \int_{a-1/2}^{a+1/2} f(t) dt$. En remarquant que la fonction $t \mapsto \frac{1}{\sin \frac{\pi t}{m}}$ est convexe sur l'intervalle $[1/2, m-1/2]$, on a

$$(2.24) \quad \begin{aligned} \sum_{a=1}^{m-1} \frac{1}{\sin \frac{\pi a}{m}} &\leq \int_{1/2}^{m-1/2} \frac{1}{\sin \frac{\pi t}{m}} dt = \frac{m}{\pi} \left[\log \tan \frac{\pi t}{2m} \right]_{1/2}^{m-1/2} \\ &= \frac{2m}{\pi} \log \cot \frac{\pi}{4m} \leq \frac{2m}{\pi} \log \frac{4m}{\pi} \end{aligned}$$

et (2.20) résulte de (2.22), (2.23) et (2.24). \square

Lemme 8 (cf. [11], p. 38). *Nous conservons les notations du lemme 7. Soit une fonction complexe $\Phi(t)$ continûment dérivable définie sur l'intervalle réel $[0, m]$. On suppose que, pour tout $t \in [0, m]$, on a $|\Phi(t)| \leq M_0$ et $|\Phi'(t)| \leq M_1$. Alors, on a*

$$\left| \sum_{\substack{\xi \leq \ell \leq \xi' \\ (\ell, m)=1}} \exp \frac{2i\pi b \bar{\ell}}{m} \Phi(\ell) \right| \leq \tau(m) \sqrt{m} \{(b, m)\}^{1/2} \gamma(m) [|\xi' - \xi| M_1 + M_0].$$

Démonstration. Comme dans le lemme 7, nous pouvons supposer $\xi \leq \xi'$. Il est commode de poser, pour $0 \leq t \leq m$

$$(2.25) \quad g(t) = \sum_{\substack{\xi \leq \ell \leq t \\ (\ell, m)=1}} \exp \frac{2i\pi b \bar{\ell}}{m}$$

et, par le lemme 7, on a

$$(2.26) \quad |g(t)| \leq \tau(m) \sqrt{m} \{(b, m)\}^{1/2} \gamma(m).$$

Par la méthode de la sommation d'Abel, on écrit :

$$(2.27) \quad \begin{aligned} \sum_{\substack{\xi \leq \ell \leq \xi' \\ (\ell, m)=1}} \exp \frac{2i\pi b \bar{\ell}}{m} \Phi(\ell) &= g(\xi) \Phi(\xi) + \sum_{\ell=\xi+1}^{\xi'} (g(\ell) - g(\ell-1)) \Phi(\ell) \\ &= \sum_{\ell=\xi}^{\xi'-1} g(\ell) (\Phi(\ell) - \Phi(\ell+1)) + g(\xi') \Phi(\xi'). \end{aligned}$$

On applique le théorème des accroissements finis à $\Phi(\ell+1) - \Phi(\ell)$, et le lemme suit de (2.27) et (2.26). \square

3 Démonstration du théorème 2.

3.1 Le polynôme de Bernoulli.

Soit $d \neq 1$ un nombre impair ; avec la définition de \mathcal{E}_d donnée par (2.2), en choisissant ν entre 1 et $d - 1$, on a

$$(3.1) \quad \sum_{\substack{1 \leq A \leq f(d) \\ 4A^2 \equiv -1 \pmod{d}}} 1 = \sum_{\nu \in \mathcal{E}_d} \sum_{\substack{1 \leq A \leq f(d) \\ A \equiv \nu \pmod{d}}} 1 = \sum_{\nu \in \mathcal{E}_d} \left(\left\lfloor \frac{f(d) - \nu}{d} \right\rfloor + 1 \right).$$

On remplace dans (3.1) la partie entière par son expression en fonction du polynôme de Bernoulli : $B_1(t) = t - [t] - 1/2$; il vient

$$(3.2) \quad \sum_{\substack{1 \leq A \leq f(d) \\ 4A^2 \equiv -1 \pmod{d}}} 1 = \sum_{\nu \in \mathcal{E}_d} \left(-B_1 \left(\frac{f(d) - \nu}{d} \right) + \frac{f(d) - \nu}{d} - \frac{\nu}{d} + \frac{1}{2} \right).$$

En associant les deux éléments ν et $d - \nu$ de \mathcal{E}_d (qui sont distincts, car d est impair), les deux derniers termes de (3.2) disparaissent, et, par (2.13), on obtient

$$(3.3) \quad \sum_{\substack{1 \leq A \leq f(d) \\ 4A^2 \equiv -1 \pmod{d}}} 1 = \frac{f(d)\rho(d)}{d} - \sum_{\nu \in \mathcal{E}_d} B_1 \left(\frac{f(d) - \nu}{d} \right).$$

On peut ainsi écrire (1.26) où $S_0(u, v; f)$ a été défini en (1.27) et

$$(3.4) \quad S_1(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \sum_{\nu \in \mathcal{E}_d} B_1 \left(\frac{f(d) - \nu}{d} \right).$$

Il nous faut maintenant majorer $|S_1(u, v; f)|$.

3.2 La formule de J. Vaaler.

Pour majorer $S_1(u, v; f)$ nous commençons par approcher la fonction $B_1(t)$ par un polynôme trigonométrique. Nous noterons $e(t) = e^{2i\pi t}$, et nous utiliserons le lemme ci-dessous de J. Vaaler.

Lemme 9. Soit $\omega \in \mathbb{N}$, $h \in \mathbb{Z}$, $1 \leq |h| \leq \omega$ et

$$(3.5) \quad 0 < b_\omega(h) \stackrel{\text{def}}{=} \pi \frac{|h|}{\omega + 1} \left(1 - \frac{|h|}{\omega + 1} \right) \cot \left(\pi \frac{|h|}{\omega + 1} \right) + \frac{|h|}{\omega + 1} < 1.$$

Alors, en posant

$$(3.6) \quad B_\omega^*(t) = -\frac{1}{2i\pi} \sum_{1 \leq |h| \leq \omega} \frac{b_\omega(h)}{h} e(ht)$$

on peut écrire

$$(3.7) \quad B_1(t) = B_\omega^*(t) + R_\omega^*(t)$$

avec, pour tout $t \in \mathbb{R}$,

$$(3.8) \quad |R_\omega^*(t)| \leq \frac{1}{2\omega + 2} \sum_{0 \leq |h| \leq \omega} \left(1 - \frac{|h|}{\omega + 1}\right) e^{(ht)} = \frac{\sin^2 \pi(\omega + 1)t}{2(\omega + 1)^2 \sin^2 \pi t}.$$

Démonstration. Pour $t \notin \mathbb{Z}$, (3.8) est l'inégalité (7.14) de Vaaler (cf. [20], cf. aussi le théorème A.6 de [7] et [8]). Pour $t \in \mathbb{Z}$, les deux membres de (3.8) sont égaux à $1/2$, et le résultat est encore vrai. \square

En utilisant (3.7), pour une valeur de ω que l'on précisera plus tard, (3.4) devient

$$(3.9) \quad S_1(u, v; f) = S_2(u, v; f) + S_3(u, v; f)$$

avec, par (3.6),

$$(3.10) \quad \begin{aligned} S_2(u, v; f) &= \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \sum_{\nu \in \mathcal{E}_d} B_\omega^* \left(\frac{f(d) - \nu}{d} \right) \\ &= -\frac{1}{2i\pi} \sum_{1 \leq |h| \leq \omega} \frac{b_\omega(h)}{h} \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} e \left(\frac{hf(d)}{d} \right) \sum_{\nu \in \mathcal{E}_d} e \left(-\frac{h\nu}{d} \right) \end{aligned}$$

et

$$(3.11) \quad S_3(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \sum_{\nu \in \mathcal{E}_d} R_\omega^* \left(\frac{f(d) - \nu}{d} \right).$$

Par (3.8), on a

$$(3.12) \quad |S_3(u, v; f)| \leq S_4(u, v; f)$$

en posant

$$(3.13) \quad S_4(u, v; f) = \frac{1}{2\omega + 2} \sum_{0 \leq |h| \leq \omega} \left(1 - \frac{|h|}{\omega + 1}\right) \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} e \left(\frac{hf(d)}{d} \right) \sum_{\nu \in \mathcal{E}_d} e \left(-\frac{h\nu}{d} \right).$$

Il est commode de poser, pour $h \neq 0$,

$$(3.14) \quad P = P(h, u, v, f) = \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} e \left(\frac{hf(d)}{d} \right) \sum_{\nu \in \mathcal{E}_d} e \left(-\frac{h\nu}{d} \right)$$

de telle sorte que (3.10) et (3.13) deviennent

$$(3.15) \quad S_2(u, v; f) = -\frac{1}{2i\pi} \sum_{1 \leq |h| \leq \omega} \frac{b_\omega(h)}{h} P(h, u, v, f)$$

et

$$(3.16) \quad S_4(u, v; f) = \frac{1}{2\omega + 2} \left(\sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \rho(d) + \sum_{1 \leq |h| \leq \omega} \left(1 - \frac{|h|}{\omega + 1}\right) P(h, u, v, f) \right).$$

Il nous faut maintenant majorer $|P(h, u, v, f)|$.

3.3 Evaluation de $P(\mathbf{h}, \mathbf{u}, \mathbf{v}, \mathbf{f})$.

Pour évaluer $P(h, u, v, f)$, nous utiliserons la bijection Θ entre \mathcal{S}_d et \mathcal{E}_d étudiée dans la proposition 1. Par cette bijection, on déduit de (3.14)

$$(3.17) \quad P(h, u, v, f) = \sum_{\substack{u \leq r^2 + 4s^2 \leq v \\ r \text{ impair}, r > 0, (r, s) = 1}} e \left(\frac{hf(r^2 + 4s^2)}{r^2 + 4s^2} - \frac{\nu h}{r^2 + 4s^2} \right)$$

où $\nu = \Theta(r, s)$ est donné par l'une des formules (2.3) ou (2.4). Nous coupons la somme (3.17) en quatre parties suivant que $|s| < r$ ou $|s| > r$ et $s > 0$ ou $s < 0$ (Remarquons que $|s| = r$ ou $s = 0$ entraînerait $r = (r, s) = 1$ ce qui est incompatible avec $r^2 + 4s^2 \geq u$, pour $u > 5$). Ainsi dans la somme P_1 , on impose $1 \leq s < r$ ce qui entraîne

$$u \leq r^2 + 4s^2 < 5r^2 \quad \text{et} \quad r^2 < r^2 + 4s^2 \leq v$$

et encore

$$(3.18) \quad \sqrt{u/5} < r < \sqrt{v};$$

pour r fixé dans l'intervalle (3.18), les inégalités $u \leq r^2 + 4s^2 \leq v$ et $1 \leq s < r$ entraînent l'existence (cf. Fig. 2) de deux nombres entiers $\xi_1 = \xi_1(r)$ et $\xi'_1 = \xi'_1(r)$ vérifiant $1 \leq \xi_1 \leq r - 1$ et $1 \leq \xi'_1 \leq r - 1$ et tels que

$$u \leq r^2 + 4s^2 \leq v \quad \text{et} \quad 1 \leq s < r \quad \iff \quad \xi_1 \leq s \leq \xi'_1.$$

Les valeurs de ξ_1 et ξ'_1 sont données par :

$$\xi_1 = 1 \quad \text{si} \quad r \geq \sqrt{u} \quad \text{et} \quad \xi_1 = \left\lceil \sqrt{\frac{u - r^2}{4}} \right\rceil \quad \text{si} \quad r < \sqrt{u}$$

$$\xi'_1 = r - 1 \quad \text{si} \quad r \leq \sqrt{\frac{v}{5}} \quad \text{et} \quad \xi'_1 = \left\lfloor \sqrt{\frac{v - r^2}{4}} \right\rfloor \quad \text{si} \quad r > \sqrt{\frac{v}{5}}.$$

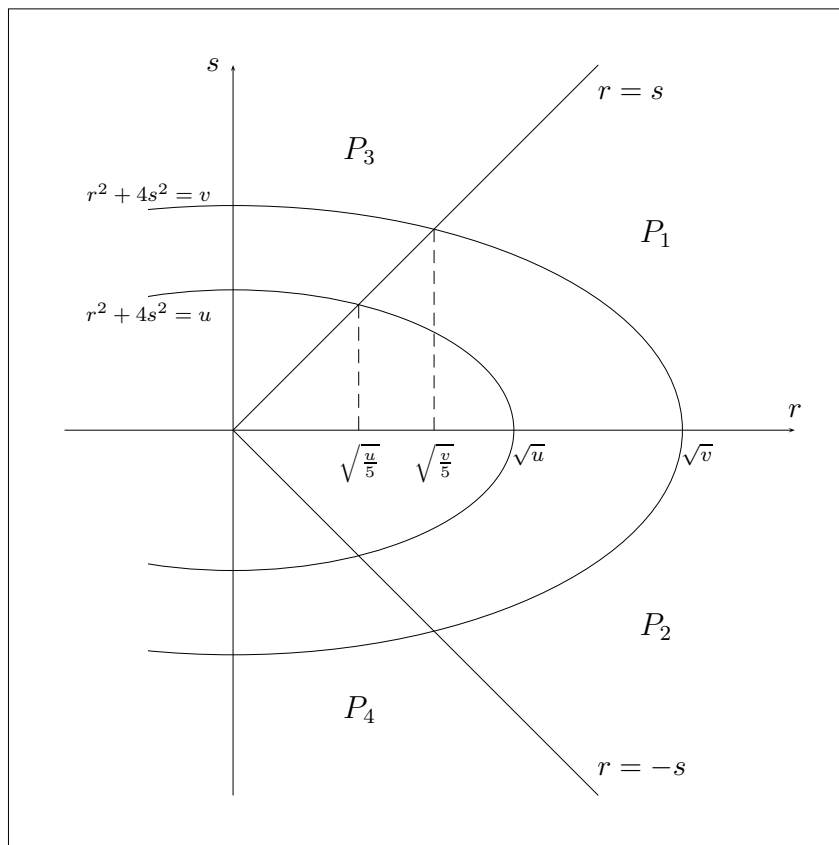


FIG. 2 – Découpage en quatre zones

On a ainsi à partir de (3.17) et (2.3)

$$(3.19) \quad P_1 = P_1(h, u, v, f) = \sum_{\substack{\sqrt{u/5} < r < \sqrt{v} \\ r \text{ impair}}} \sum_{\substack{\xi_1 \leq s \leq \xi'_1 \\ (r,s)=1}} e\left(\frac{-h\bar{4}s}{r}\right) \Phi_1(s)$$

avec, par (1.24),

$$\Phi_1(s) = \exp\left(2i\pi h F(r^2 + 4s^2) - \frac{2i\pi h s}{r(r^2 + 4s^2)}\right).$$

On calcule la dérivée logarithmique de Φ_1 :

$$\frac{\Phi'_1(s)}{\Phi_1(s)} = \frac{2i\pi h}{r} \left(8rsF'(r^2 + 4s^2) - \frac{r^2 - 4s^2}{(r^2 + 4s^2)^2}\right)$$

et l'on a, en observant que $|r^2 - 4s^2| \leq r^2 + 4s^2$ et $4|rs| \leq r^2 + 4s^2$

$$|\Phi'_1(s)| = \left|\frac{\Phi'_1(s)}{\Phi_1(s)}\right| \leq \frac{2\pi|h|}{r} \left(2(r^2 + 4s^2)|F'(r^2 + 4s^2)| + \frac{1}{u}\right).$$

On pose $b = -h\bar{4}$, où, rappelons le, $\bar{4}$ est un inverse de 4 modulo r . On a $(b, r) = (h, r)$ (car r est impair), $|\xi'_1 - \xi_1| \leq r$ et le lemme 8 donne

$$\left| \sum_{\substack{\xi_1 \leq s \leq \xi'_1 \\ (r,s)=1}} e\left(\frac{-h\bar{4}s}{r}\right) \Phi_1(s) \right| \leq \tau(r)\sqrt{r}\{(h, r)\}^{1/2}\gamma(r)M(h, r)$$

avec, par (1.23)

$$M(h, r) = 2\pi|h| \left(2\|tF'(t)\| + \frac{1}{u}\right) + 1.$$

En notant que $u \geq 7$ et en utilisant (1.25), il vient

$$M(h, r) \leq 4\pi|h| \left(\|tF'(t)\| + \frac{1}{2u} + \frac{1}{4\pi}\right) \leq 4\pi|h|\mathcal{M}.$$

En reportant dans (3.19) on obtient

$$(3.20) \quad |P_1| \leq 4\pi|h|\mathcal{M} \sum_{\substack{\sqrt{u/5} < r < \sqrt{v} \\ r \text{ impair}}} \tau(r)\sqrt{r}\{(h, r)\}^{1/2}\gamma(r).$$

Dans la somme P_2 , on met les termes de (3.17) avec $-r < s \leq -1$. On a

$$\begin{aligned}
 P_2 = P_2(h, u, v, f) &= \sum_{\substack{\sqrt{u/5} < r < \sqrt{v} \\ r \text{ impair}}} \sum_{\substack{-\xi'_1 \leq s \leq -\xi_1 \\ (r,s)=1}} e\left(\frac{-h\overline{4s}}{r}\right) \Phi_1(s) \\
 (3.21) \qquad \qquad \qquad &= \sum_{\substack{\sqrt{u/5} < r < \sqrt{v} \\ r \text{ impair}}} \sum_{\substack{\xi_1 \leq s \leq \xi'_1 \\ (r,s)=1}} e\left(\frac{h\overline{4s}}{r}\right) \Phi_1(-s).
 \end{aligned}$$

Par application du lemme 8, on trouve pour $|P_2|$ la même majoration que pour $|P_1|$, et l'on a, par (3.20) :

$$(3.22) \qquad |P_1| + |P_2| \leq 8\pi|h|\mathcal{M} \sum_{\substack{\sqrt{u/5} < r < \sqrt{v} \\ r \text{ impair}}} \tau(r)\sqrt{r}\{(h, r)\}^{1/2}\gamma(r).$$

La somme P_3 contient les termes de la somme (3.17), avec ν donné par (2.4), satisfaisant $1 \leq r < s$. Il existe deux nombres entiers ξ_3 et ξ'_3 tels que

$$(3.23) \qquad 1 \leq \xi_3 \leq s \qquad \text{et} \qquad 1 \leq \xi'_3 \leq s$$

tels que

$$(3.24) \qquad P_3 = P_3(h, u, v, f) = \sum_{\sqrt{u/5} < s < \frac{\sqrt{v}}{2}} \sum_{\substack{\xi_3 \leq r \leq \xi'_3 \\ (r,4s)=1}} e\left(\frac{-h\overline{r}}{4s}\right) \Phi_3(r)$$

avec

$$\Phi_3(r) = \exp\left(2i\pi hF(r^2 + 4s^2) + \frac{i\pi hr}{2s(r^2 + 4s^2)}\right).$$

On a comme précédemment,

$$|\Phi'_3(r)| = \left| \frac{\Phi'_3(r)}{\Phi_3(r)} \right| \leq \frac{\pi|h|}{s} \left((r^2 + 4s^2)|F'(r^2 + 4s^2)| + \frac{1}{2u} \right).$$

On applique le lemme 8 avec $m = 4s$ et $b = -h$; par (3.23), on a $|\xi'_3 - \xi_3| \leq s$ et il s'ensuit

$$\left| \sum_{\substack{\xi_3 \leq r \leq \xi'_3 \\ (r,4s)=1}} e\left(\frac{-h\overline{r}}{4s}\right) \Phi_3(r) \right| \leq \pi|h|\mathcal{M}\tau(4s)\sqrt{4s}\{(h, 4s)\}^{1/2}\gamma(4s).$$

Par (3.24), on déduit que

$$|P_3| \leq \pi|h|\mathcal{M} \sum_{\substack{4\sqrt{u/5} < m < 2\sqrt{v} \\ 4|m}} \tau(m)\sqrt{m}\{(h, m)\}^{1/2}\gamma(m).$$

La même majoration est valable pour la somme P_4 , où l'on compte les termes de (3.17) avec $s < 0$ et $1 \leq r \leq -s$, et l'on déduit de (3.22)

$$|P| = |P_1 + P_2 + P_3 + P_4| \leq 8\pi|h|\mathcal{M} \sum_{m < 2\sqrt{v}} \tau(m)\sqrt{m}\{(h, m)\}^{1/2}\gamma(m).$$

Par (2.21) et le lemme 2, il suit

$$\begin{aligned} |P| &\leq 8\pi|h|\mathcal{M}\gamma(2\sqrt{v})\sqrt{2}v^{1/4} \sum_{m \leq 2\sqrt{v}} \tau(m)\{(h, m)\}^{1/2} \\ &\leq 8|h|\mathcal{M}\sqrt{2} \left(\log v + \pi + 2 \log \frac{8}{\pi} \right) (\log v + 2 + 2 \log 2)v^{3/4}\psi(|h|) \\ (3.25) \quad &\leq 8|h|\mathcal{M}\sqrt{2}(\log v + 4)^2v^{3/4}\psi(|h|) \end{aligned}$$

où $\psi(|h|)$ est défini en (2.17).

3.4 Choix de ω .

Par (3.15) et (3.5), on a

$$|S_2(u, v; f)| \leq \frac{1}{2\pi} \sum_{1 \leq |h| \leq \omega} \frac{1}{|h|} |P|$$

et par (3.16) et le lemme 5 (iii), on a

$$\begin{aligned} |S_4(u, v; f)| &\leq \frac{3v}{14(\omega + 1)} + \frac{1}{2\omega + 2} \sum_{1 \leq |h| \leq \omega} |P| \\ &\leq \frac{3v}{14(\omega + 1)} + \frac{1}{2} \sum_{1 \leq |h| \leq \omega} \frac{1}{|h|} |P| \end{aligned}$$

de sorte que, par (3.9) et (3.12) on a

$$\begin{aligned} |S_1(u, v; f)| &\leq |S_2(u, v; f)| + |S_4(u, v; f)| \\ &\leq \frac{3v}{14(\omega + 1)} + \frac{1}{2} \left(\frac{1}{\pi} + 1 \right) \sum_{1 \leq |h| \leq \omega} \frac{1}{|h|} |P|. \end{aligned}$$

Par (3.25) et le lemme 3, en tenant compte des valeurs négatives de h , il suit

$$(3.26) \quad |S_1(u, v; f)| \leq \frac{3v}{14(\omega + 1)} + \left(\frac{1}{\pi} + 1 \right) 8\sqrt{2}\mathcal{M}(\log v + 4)^2v^{3/4}(7\omega).$$

On pose

$$(3.27) \quad \omega_0 = \frac{1}{14} \sqrt{\frac{3\pi}{4\sqrt{2}\mathcal{M}(\pi + 1)}} \frac{v^{1/8}}{\log v + 4}$$

et l'on choisit ω

$$(3.28) \quad \omega = \lfloor \omega_0 \rfloor .$$

Par (3.26), il vient

$$\begin{aligned} |S_1(u, v; f)| &\leq \frac{3v}{14\omega_0} + 56\sqrt{2} \left(\frac{1}{\pi} + 1 \right) (\log v + 4)^2 v^{3/4} \omega_0 \mathcal{M} \\ &= 4\sqrt{3\sqrt{2}(1 + 1/\pi)(\log v + 4)v^{7/8}\sqrt{\mathcal{M}}} \end{aligned}$$

et comme $4\sqrt{3\sqrt{2}(1 + 1/\pi)} = 9.4599\dots$, cela achève la démonstration du théorème 2. \square

4 Démonstration du théorème 1.

Dans ce paragraphe, pour les calculs numériques, nous supposons

$$(4.1) \quad \lambda = 2 \quad \text{et} \quad N \geq 1000.$$

Compte tenu de (1.26) nous écrivons (1.22)

$$(4.2) \quad \widehat{Q}(N, \lambda) = Q_0(N, \lambda) - Q_1(N, \lambda)$$

avec, pour $i = 0$ ou 1

$$(4.3) \quad Q_i(N, \lambda) = S_i(u_0, v_0; f_1) - S_i(u_0, v_0; N - 1) - S_i(u_1, v_0; f_2) + S_i(u_1, v_0; N - 1).$$

Nous évaluerons successivement $Q_0(N, \lambda)$ et $Q_1(N, \lambda)$; mais il nous faut d'abord définir quelques nouvelles variables.

4.1 Les quantités approchées.

Nous utiliserons les trois quantités linéaires en N

$$(4.4) \quad \widehat{v}_0 = \lambda(2 - \sqrt{2})N, \quad \widehat{u}_0 = 2(\lambda - \sqrt{\lambda^2 - 1})N, \quad \widehat{u}_1 = 2(\sqrt{2} - 1)N,$$

comme des approximations des nombres v_0 , u_0 et u_1 définis en (1.14), (1.16) et (1.17). En utilisant les relations

$$(4.5) \quad \sqrt{a} \leq \sqrt{a+t} \leq \sqrt{a} + \frac{t}{2\sqrt{a}}, \quad a > 0, t \geq 0$$

et

$$(4.6) \quad \sqrt{a} - \frac{t}{\sqrt{a}} = \sqrt{a} \left(1 - \frac{t}{a} \right) \leq \sqrt{a} \sqrt{1 - \frac{t}{a}} = \sqrt{a-t} \leq \sqrt{a}, \quad a > 0, 0 \leq t \leq a$$

on obtient les inégalités (rappelons que $\lambda > \sqrt{2}$ et $N \geq 1$)

$$(4.7) \quad \widehat{v}_0 \leq v_0 = 2\lambda N - \frac{1}{2}\sqrt{8\lambda^2 N^2 - 2} \leq \widehat{v}_0 + \frac{1}{2\sqrt{2}\lambda N} < \widehat{v}_0 + \frac{1}{4N} < \widehat{v}_0 + 1 \leq 2\lambda N,$$

$$(4.8) \quad \widehat{u}_0 < u_0 = 2\lambda N - \sqrt{4(\lambda^2 - 1)N^2 - 1} \leq \widehat{u}_0 + \frac{1}{2\sqrt{\lambda^2 - 1}N} < \widehat{u}_0 + \frac{1}{2N} < \widehat{u}_0 + 1$$

et

$$(4.9) \quad \widehat{u}_1 < u_1 = \sqrt{8N^2 + 1} - 2N \leq \widehat{u}_1 + \frac{1}{4\sqrt{2}N} < \widehat{u}_1 + 1.$$

Nous approcherons les fonctions

$$(4.10) \quad f_1(t) = \frac{1}{2}\sqrt{t(4\lambda N - t) - 1} \quad \text{et} \quad f_2(t) = \frac{1}{2}\left(t + \sqrt{2t^2 - 1}\right)$$

définies en (1.13) et (1.12) par

$$(4.11) \quad \widehat{f}_1(t) = \frac{1}{2}\sqrt{t(4\lambda N - t)} \quad \text{et} \quad \widehat{f}_2(t) = \frac{1 + \sqrt{2}}{2}t.$$

Par (4.6), on a, pour $1 \leq t \leq 4\lambda N - 1$, $N \geq 1$ et $\lambda > \sqrt{2}$

$$(4.12) \quad \widehat{f}_1(t) - \frac{1}{4} \leq \widehat{f}_1(t) - \frac{1}{2\sqrt{4\lambda N - 1}} \leq \widehat{f}_1(t) - \frac{1}{2\sqrt{t(4\lambda N - t)}} \leq f_1(t) \leq \widehat{f}_1(t)$$

et, toujours par (4.6), on a, pour $t \geq 1$

$$(4.13) \quad \widehat{f}_2(t) - \frac{1}{2} \leq \frac{1}{2}\left(t + \sqrt{2}t - \frac{1}{\sqrt{2}t}\right) \leq f_2(t) \leq \widehat{f}_2(t).$$

4.2 Les termes secondaires de reste.

Pour estimer $Q_0(N, \lambda)$ défini par (4.3), on utilise (1.27) et le lemme 6

$$(4.14) \quad S_0(u, v; f) = \sum_{\substack{u \leq d \leq v \\ d \text{ impair}}} \frac{\rho(d)f(d)}{d} = J(u, v; f) + \widehat{R}_0(u, v; F)$$

où l'on a posé $F(t) = f(t)/t$ et

$$(4.15) \quad J(u, v; f) = \frac{1}{\pi} \int_u^v F(t) dt = \frac{1}{\pi} \int_u^v \frac{f(t)}{t} dt.$$

Nous écrivons

$$(4.16) \quad Q_0(N, \lambda) = Q_2(N, \lambda) + Q_3(N, \lambda)$$

avec

$$(4.17) \quad Q_2(N, \lambda) = J(u_0, v_0; f_1) - J(u_0, v_0; N-1) - J(u_1, v_0; f_2) + J(u_1, v_0; N-1)$$

et

$$(4.18) \quad \begin{aligned} Q_3(N, \lambda) &= \widehat{R}_0(u_0, v_0; F_1) - \widehat{R}_0\left(u_0, v_0; \frac{N-1}{t}\right) \\ &\quad - \widehat{R}_0(u_1, v_0; F_2) + \widehat{R}_0\left(u_1, v_0; \frac{N-1}{t}\right). \end{aligned}$$

Dans ce paragraphe nous majorerons les quatre termes de reste $\widehat{R}_0(u, v; F)$ et la valeur absolue de leur somme $Q_3(N, \lambda)$ tandis que, dans le paragraphe suivant, nous traiterons les termes principaux $J(u, v; f)$ et $Q_2(N, \lambda)$.

Le premier terme $\widehat{\mathbf{R}}_0(\mathbf{u}_0, \mathbf{v}_0; \mathbf{F}_1)$. Posons par (4.10)

$$(4.19) \quad F_1(t) = \frac{f_1(t)}{t} = \frac{\sqrt{t(4\lambda N - t) - 1}}{2t}.$$

Avec l'aide de MAPLE, nous dérivons :

$$(4.20) \quad F_1'(t) = \frac{-2\lambda Nt + 1}{2t^2 \sqrt{t(4\lambda N - t) - 1}} = \frac{-2\lambda Nt + 1}{4t^2 f_1(t)},$$

$$(4.21) \quad F_1''(t) = \frac{4t^2 \lambda N(3\lambda N - t - 3/t) + 3t^2 + 2}{2t^3(t(4\lambda N - t) - 1)^{3/2}}.$$

Pour $\lambda > \sqrt{2}$, $N \geq 1$ et $t \geq 1$, on a $2\lambda Nt > 1$ et, par (4.20), on voit que $F_1'(t) < 0$. La quantité $3\lambda N - t - 3/t$ est minimale en $t = \sqrt{3}$ et vaut $3\lambda N - 2\sqrt{3} > 0$; par (4.21), $F_1''(t) > 0$ pour tout $t > 0$ et donc

$$(4.22) \quad F_1(t) = \frac{f_1(t)}{t} \text{ est décroissante et convexe pour } 1 \leq t \leq v_0.$$

Par (4.8) et (4.4), si l'on choisit $N \geq \lambda$, on a $u_0 \geq \widehat{u}_0 = \frac{2N}{\lambda + \sqrt{\lambda^2 - 1}} \geq \frac{2N}{2\lambda} \geq 1$; par (4.22), F_1 est décroissante sur $[u_0, v_0]$ et par (1.16) on obtient

$$(4.23) \quad \|F_1\| = \|F_1\|_{[u_0, v_0]} = \frac{f_1(u_0)}{u_0} = \frac{N}{u_0} \leq \frac{N}{\widehat{u}_0} = \frac{\lambda + \sqrt{\lambda^2 - 1}}{2} < 1.87$$

tandis que, par (4.22), (4.20), (1.16), (4.8) et (4.4), on a

$$(4.24) \quad \begin{aligned} \|F_1'\| &= \|F_1'\|_{[u_0, v_0]} = -F_1'(u_0) = \frac{2\lambda N u_0 - 1}{4u_0^2 f_1(u_0)} = \frac{2\lambda N u_0 - 1}{4u_0^2 N} \\ &< \frac{2\lambda N u_0}{4u_0^2 N} = \frac{\lambda}{2u_0} < \frac{\lambda}{2\widehat{u}_0} = \frac{\lambda(\lambda + \sqrt{\lambda^2 - 1})}{4N} < \frac{1.87}{N}. \end{aligned}$$

Par le lemme 6, on obtient, pour $\lambda = 2$

$$(4.25) \quad |\widehat{R}_0(u_0, v_0; F_1)| \leq (4 \log v_0 + 14) \sqrt{v_0} \left(3.74 + 1.25 \frac{v_0}{N} \right).$$

Le deuxième terme $\widehat{R}_0(\mathbf{u}_0, \mathbf{v}_0; \frac{N-1}{t})$. On a en utilisant (4.8) et (4.4)

$$(4.26) \quad \left\| \frac{N-1}{t} \right\|_{[u_0, v_0]} = \frac{N-1}{u_0} < \frac{N}{u_0} \leq \frac{N}{\widehat{u}_0} = \frac{\lambda + \sqrt{\lambda^2 - 1}}{2} < 1.87$$

et

$$\left\| -\frac{N-1}{t^2} \right\|_{[u_0, v_0]} = \frac{N-1}{u_0^2} \leq \frac{N}{u_0^2} \leq \frac{N}{\widehat{u}_0^2} = \frac{(\lambda + \sqrt{\lambda^2 - 1})^2}{4N} < \frac{3.49}{N}.$$

Par le lemme 6, il suit pour $\lambda = 2$

$$(4.27) \quad \left| \widehat{R}_0 \left(u_0, v_0; \frac{N-1}{t} \right) \right| \leq (4 \log v_0 + 14) \sqrt{v_0} \left(3.74 + 2.33 \frac{v_0}{N} \right).$$

Le troisième terme $\widehat{R}_0(\mathbf{u}_1, \mathbf{v}_0; \mathbf{F}_2)$. Par (4.10), on a pour $t \geq 1$

$$(4.28) \quad F_2(t) = \frac{f_2(t)}{t} = \frac{1}{2} \left(\frac{t + \sqrt{2t^2 - 1}}{t} \right) \leq \frac{1 + \sqrt{2}}{2}$$

et donc

$$(4.29) \quad \|F_2\| = \|F_2\|_{[u_1, v_0]} \leq \frac{1 + \sqrt{2}}{2} \leq 1.21.$$

Pour $t \geq 1$, on a $2t^2 - 1 \geq 1$ et

$$F_2'(t) = \frac{1}{2t^2 \sqrt{2t^2 - 1}} \leq \frac{1}{2t^2}.$$

Donc, par (4.9) et (4.4)

$$(4.30) \quad \|F_2'\| = \|F_2'\|_{[u_1, v_0]} \leq \frac{1}{2u_1^2} \leq \frac{1}{2\widehat{u}_1^2} = \frac{1}{8(\sqrt{2} - 1)^2 N^2} \leq \frac{0.73}{N^2}.$$

Par le lemme 6, il suit pour $\lambda = 2$

$$(4.31) \quad |\widehat{R}_0(u_1, v_0; F_2)| \leq (4 \log v_0 + 14) \sqrt{v_0} \left(2.42 + 0.49 \frac{v_0}{N^2} \right).$$

Le quatrième terme $\widehat{R}_0(\mathbf{u}_1, \mathbf{v}_0; \frac{N-1}{t})$. On a par (4.9) et (4.4)

$$(4.32) \quad \left\| \frac{N-1}{t} \right\|_{[u_1, v_0]} = \frac{N-1}{u_1} < \frac{N}{u_1} < \frac{N}{\widehat{u}_1} = \frac{1}{2(\sqrt{2} - 1)} = \frac{1 + \sqrt{2}}{2} < 1.21$$

et

$$\left\| -\frac{N-1}{t^2} \right\|_{[u_1, v_0]} = \frac{N-1}{u_1^2} \leq \frac{N}{u_1^2} \leq \frac{N}{\widehat{u}_1^2} = \frac{1}{4(\sqrt{2} - 1)^2 N} < \frac{1.46}{N}.$$

Par le lemme 6, on a pour $\lambda = 2$

$$(4.33) \quad \left| \widehat{R}_0 \left(u_1, v_0; \frac{N-1}{t} \right) \right| \leq (4 \log v_0 + 14) \sqrt{v_0} \left(2.42 + 0.98 \frac{v_0}{N} \right).$$

On a donc, par (4.18), en ajoutant les inégalités (4.25), (4.27), (4.31) et (4.33),

$$|Q_3(N, 2)| \leq (4 \log v_0 + 14) \sqrt{v_0} \left(12.32 + 4.56 \frac{v_0}{N} + 0.49 \frac{v_0}{N^2} \right).$$

Maintenant par (4.7) et (4.4), on a, pour $N \geq 1000$ et $\lambda = 2$,

$$(4.34) \quad v_0 \leq \widehat{v}_0 + \frac{1}{4N} = \lambda(2 - \sqrt{2})N \left(1 + \frac{1}{4\lambda(2 - \sqrt{2})N^2} \right) \leq 1.18N,$$

ce qui entraîne $\log v_0 \leq \log N + \log 1.18 \leq \log N + 0.17$ et $v_0/N^2 \leq 1.18/N \leq 0.002$. On obtient ainsi

$$(4.35) \quad |Q_3(N, 2)| \leq (77 \log N + 283) \sqrt{N}.$$

4.3 Les termes principaux.

Considérons un quadruplet de nombres réels $\{\widehat{u}, u, \widehat{v}, v\}$ satisfaisant

$$(4.36) \quad \frac{N}{\lambda} \leq \widehat{u} \leq u \leq \widehat{u} + 1 < \widehat{v} \leq v \leq \widehat{v} + 1 \leq 2\lambda N$$

et un couple de fonctions $\{f, \widehat{f}\}$ continues de l'intervalle $[\widehat{u}, v]$ dans \mathbb{R} et satisfaisant

$$(4.37) \quad \|\widehat{f} - f\|_{[\widehat{u}, v]} \leq 1 \quad \text{et} \quad \|f\|_{[\widehat{u}, v]} \leq \|\widehat{f}\|_{[\widehat{u}, v]} \leq \lambda(1 + \sqrt{2})N.$$

Alors on a

$$(4.38) \quad \int_u^v \frac{f(t)}{t} dt = \int_{\widehat{u}}^{\widehat{v}} \frac{\widehat{f}(t)}{t} dt + \widehat{J}$$

avec

$$\widehat{J} = \int_{\widehat{u}}^{\widehat{v}} \frac{f(t) - \widehat{f}(t)}{t} dt - \int_{\widehat{u}}^u \frac{f(t)}{t} dt + \int_{\widehat{v}}^v \frac{f(t)}{t} dt$$

et par (4.36) et (4.37)

$$(4.39) \quad \begin{aligned} \widehat{J} &\leq \log \frac{\widehat{v}}{\widehat{u}} + \lambda(1 + \sqrt{2})N \left(\log \frac{u}{\widehat{u}} + \log \frac{v}{\widehat{v}} \right) \\ &\leq \log \frac{2\lambda N}{N/\lambda} + \lambda(1 + \sqrt{2})N \left(\frac{u - \widehat{u}}{\widehat{u}} + \frac{v - \widehat{v}}{\widehat{v}} \right) \\ &\leq \log(2\lambda^2) + \lambda(1 + \sqrt{2})N \left(\frac{1}{N/\lambda} + \frac{1}{N/\lambda} \right) \\ &= \log(2\lambda^2) + 2\lambda^2(1 + \sqrt{2}) \leq 21.40 \quad (\text{pour } \lambda = 2). \end{aligned}$$

Nous appliquerons quatre fois l'égalité (4.38) complétée par l'inégalité (4.39). Observons d'abord que, par (4.8), (4.7), (4.9) et (4.4), les quadruplets $\{\widehat{u}_0, u_0, \widehat{v}_0, v_0\}$ et $\{\widehat{u}_1, u_1, \widehat{v}_0, v_0\}$ satisfont (4.36) : l'inégalité $\widehat{u}_0 + 1 \leq \widehat{u}_1 + 1 < \widehat{v}_0$ est vérifiée, à partir de (4.4), pour $N > \frac{2+\sqrt{2}}{2(\lambda-\sqrt{2})}$. Observons ensuite que les couples de fonctions $\{f_1, \widehat{f}_1\}$, $\{f_2, \widehat{f}_2\}$ et $\{N-1, N\}$ vérifient (4.37) sur l'intervalle $[\widehat{u}_0, v_0]$ en utilisant (4.11), (4.12), (4.13) et (4.7).

Alors, par (4.15) et (4.38), (4.17) s'écrit

$$(4.40) \quad Q_2(N, \lambda) = \frac{1}{\pi} \left(\int_{\widehat{u}_0}^{\widehat{v}_0} \frac{\widehat{f}_1(t)}{t} dt - \int_{\widehat{u}_0}^{\widehat{v}_0} \frac{N}{t} dt - \int_{\widehat{u}_1}^{\widehat{v}_0} \frac{\widehat{f}_2(t)}{t} dt + \int_{\widehat{u}_1}^{\widehat{v}_0} \frac{N}{t} dt \right) + Q_4$$

avec, par (4.39),

$$(4.41) \quad |Q_4| = |Q_4(N, \lambda)| \leq \frac{4}{\pi} 21.40 \leq 28.$$

Par (4.4), (4.11) et le changement de variable $y = \sqrt{\frac{4\lambda N}{t} - 1}$, on calcule

$$\begin{aligned} \int_{\widehat{u}_0}^{\widehat{v}_0} \frac{\widehat{f}_1(t)}{t} dt &= \frac{1}{2} \int_{2(\lambda-\sqrt{\lambda^2-1})N}^{\lambda(2-\sqrt{2})N} \sqrt{\frac{4\lambda N}{t} - 1} dt = 4\lambda N \int_{1+\sqrt{2}}^{\lambda+\sqrt{\lambda^2-1}} \frac{y^2 dy}{(1+y^2)^2} \\ &= \lambda N \left(\frac{\pi}{4} - \arcsin \frac{1}{\lambda} \right) - N(1+\sqrt{2}) \end{aligned}$$

en notant que $\arctan(\lambda + \sqrt{\lambda^2-1}) = \frac{1}{2}(\pi - \arcsin \frac{1}{\lambda})$ qui, en substituant $\sqrt{2}$ à λ , donne $\arctan(1 + \sqrt{2}) = \frac{3\pi}{8}$.

Par (4.4) et (4.11), (4.40) se réécrit

$$(4.42) \quad Q_2(N, \lambda) = \alpha(\lambda)N + Q_4(N, \lambda)$$

où $\alpha(\lambda)$ a été défini en (1.4).

On peut remarquer, par (4.40) que

$$\alpha(\lambda)N = \frac{1}{\pi} \iint_{\mathcal{D}} \frac{dx dt}{t}$$

où \mathcal{D} est le domaine des points (x, t) satisfaisant $N \leq x \leq \frac{\lambda N}{\sqrt{2}}$ et situés entre l'ellipse $t^2 - 4\lambda Nt + 4x^2 = 0$ et la droite $t = 2(\sqrt{2}-1)x$. En intégrant d'abord en t , on obtient

$$\alpha(\lambda)N = \int_N^{\frac{\lambda N}{\sqrt{2}}} dx \int_{2(\lambda N - \sqrt{\lambda^2 N^2 - x^2})}^{2(\sqrt{2}-1)x} \frac{dt}{\pi t} = \frac{1}{\pi} \int_N^{\frac{\lambda N}{\sqrt{2}}} \log \frac{2(\sqrt{2}-1)x}{2(\lambda N - \sqrt{\lambda^2 N^2 - x^2})} dx.$$

4.4 Le terme principal de reste.

Pour estimer $Q_1(N, \lambda)$ défini par (4.3), nous appliquerons quatre fois le théorème 2. Pour cela, nous supposons que $N \geq 7\lambda$ de façon que, par (4.8), (4.9) et (4.4), on ait $u_0 \geq \widehat{u}_0 \geq 7$ et $u_1 \geq \widehat{u}_1 \geq 7$. Rappelons que $\mathcal{M}(u, v; f)$ est défini en (1.25).

Le premier terme $\mathcal{M}(\mathbf{u}_0, \mathbf{v}_0; \mathbf{f}_1)$. On pose $F_1(t) = f_1(t)/t$. A partir de (4.20), on calcule

$$\frac{d}{dt}(tF_1'(t)) = \frac{2\lambda N t^2(2\lambda N - t - 3/t) + 2t^2 + 1}{2t^2(t(4\lambda N - t) - 1)^{3/2}}.$$

La quantité $2\lambda N - t - 3/t$ est positive pour $\lambda > \sqrt{2}$, $N \geq 2$ et pour tout $t > 0$. En remarquant que $F_1'(t)$ est négative et en utilisant (4.24), il vient

$$\|tF_1'(t)\|_{[u_0, v_0]} = -u_0 F_1'(u_0) < \frac{\lambda}{2}.$$

Il s'ensuit par (1.25) que, pour $\lambda = 2$

$$(4.43) \quad \mathcal{M}(u_0, v_0; f_1) \leq 1 + \frac{2}{5} = 1.4 < (1.19)^2.$$

Le deuxième terme $\mathcal{M}(\mathbf{u}_0, \mathbf{v}_0; \mathbf{N} - \mathbf{1})$. Posons $f(t) = N - 1$, $F(t) = (N - 1)/t$; on a $tF'(t) = -\frac{N-1}{t}$, et par (1.25) et (4.26)

$$(4.44) \quad \mathcal{M}(u_0, v_0; N - 1) = \left\| \frac{N - 1}{t} \right\|_{[u_0, v_0]} + \frac{2}{5} < \frac{\lambda + \sqrt{\lambda^2 - 1}}{2} + \frac{2}{5} < 2.27 < (1.51)^2.$$

Le troisième terme $\mathcal{M}(\mathbf{u}_1, \mathbf{v}_0; \mathbf{f}_2)$. On pose $F_2(t) = f_2(t)/t$. On a par (4.30) et (4.7)

$$\|tF_2'(t)\|_{[u_1, v_0]} \leq \frac{v_0}{8(\sqrt{2} - 1)^2 N^2} \leq \frac{2\lambda N}{8(\sqrt{2} - 1)^2 N^2} = \frac{\lambda}{4(\sqrt{2} - 1)^2 N} \leq \frac{1.46}{N}$$

et, pour $N \geq 1000$,

$$(4.45) \quad \mathcal{M}(u_1, v_0; f_2) \leq \frac{1.46}{N} + \frac{2}{5} \leq 0.41 \leq (0.64)^2.$$

Le quatrième terme $\mathcal{M}(\mathbf{u}_1, \mathbf{v}_0; \mathbf{N} - \mathbf{1})$. Posons $f(t) = N - 1$, $F(t) = (N - 1)/t$; on a $tF'(t) = -\frac{N-1}{t}$, et par (1.25) et (4.32)

$$(4.46) \quad \mathcal{M}(u_1, v_0; N - 1) = \left\| \frac{N - 1}{t} \right\|_{[u_1, v_0]} + \frac{2}{5} \leq \frac{1 + \sqrt{2}}{2} + \frac{2}{5} < 1.61 < (1.27)^2.$$

En conclusion, on peut majorer $|Q_1(N, 2)|$ défini par (4.3) à l'aide de (1.28) et, en ajoutant (4.43), (4.44), (4.45) et (4.46), on obtient

$$(4.47) \quad |Q_1(N, 2)| \leq 9.5(\log v_0 + 4)v_0^{7/8}(1.19 + 1.51 + 0.64 + 1.27) = 43.8(\log v_0 + 4)v_0^{7/8}.$$

En utilisant (4.34), (4.47) entraîne

$$(4.48) \quad |Q_1(N, 2)| \leq 43.8(\log(1.18N) + 4)(1.18N)^{7/8} \leq (211 + 51 \log N)N^{7/8}.$$

4.5 Récapitulation.

Par (1.18), (4.2), (4.16) et (4.42), il vient

$$Q(N, \lambda) = \alpha(\lambda)N + \tilde{R}(N, \lambda)$$

avec

$$(4.49) \quad \tilde{R}(N, \lambda) = -Q_1(N, \lambda) + Q_3(N, \lambda) + Q_4(N, \lambda) + \varepsilon(N, \lambda)$$

et (1.8) résulte de (4.48), (4.35), (4.41) et (1.32).

La méthode utilisée pour obtenir (4.48), (4.35) et (4.41) montre que, pour tout λ fixé, $\lambda > \sqrt{2}$, on a $Q_1(N, \lambda) = O_\lambda(N^{7/8}(\log N))$, $Q_3(N, \lambda) = O_\lambda(N^{1/2}(\log N))$, $Q_4(N, \lambda) = O_\lambda(1)$, et, par (1.31), on a $\varepsilon(N, \lambda) = O_\lambda(1)$; ainsi, par (4.49), (1.6) est démontré, ce qui achève la preuve du théorème 1.

TABLE 1

k	$N = 2^k$	$Q(N, 2)$	$Q(N, 2)/N$	$Q(N, 2) - \alpha(2)N$
4	16	1	0.06250000	0.55
5	32	2	0.06250000	1.10
6	64	2	0.03125000	0.21
7	128	3	0.02343750	-0.59
8	256	7	0.02734075	-0.17
9	512	21	0.04101062	6.66
10	1024	29	0.02832031	0.31
11	2048	60	0.02929688	2.62
12	4096	122	0.02978516	7.25
13	8192	245	0.02990723	15.49
14	16384	460	0.02807617	0.99
15	32768	944	0.02880859	25.98
16	65536	1806	0.02755737	-30.05
17	131072	3639	0.02776337	-33.10
18	262144	7347	0.02802658	2.81
19	524288	14756	0.02814484	67.61
20	1048576	29576	0.02820587	199.23
21	2097152	58698	0.02798939	-55.55
22	4194304	117372	0.02798367	-135.09
23	8388608	235082	0.02802396	67.81
24	16777216	470241	0.02802855	212.62
25	33554432	939804	0.02800834	-252.76
26	67108864	1880297	0.02801861	183.48
27	134217728	3761402	0.02802463	1174.97
28	268435456	7521153	0.02801844	698.94
29	536870912	15040659	0.02801541	-249.12
30	1073741824	30080525	0.02801467	-1291.25
31	2147483648	60153965	0.02801137	-9667.50
32	4294967296	120332158	0.02801701	4893.01
33	8589934592	240660303	0.02801655	5773.02
34	17179869184	481315101	0.02801622	6041.03
35	34359738368	962578706	0.02801473	-39413.94
36	68719476736	1925208869	0.02801547	-27370.87
37	137438953472	3850396279	0.02801532	-76200.75
38	274877906944	7701008772	0.02801610	63812.49
39	549755813888	15401833428	0.02801577	-56491.01
40	1099511627776	30803795356	0.02801589	15517.96

Références

- [1] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993.
- [2] R. Crandall and C. Pomerance, *Prime Numbers A Computational Perspective*, Springer-Verlag, 2001.
- [3] H. Daboussi et J. Rivat, Explicit upper bounds for exponential sums over primes, *Math. Comp.* 70, 2001, 431-447
- [4] J.-M. Deshouillers and H. Iwaniec, On the greatest prime factor of $n^2 + 1$, *Ann. Inst. Fourier*, 32, n°4, 1982, 1-11.
- [5] L.E. Dickson, *Studies in the theory of numbers*, The University of Chicago Press, 1930.
- [6] L.E. Dickson, *History of the theory of numbers*, t. 3, Chelsea Publishing Company, 1971.
- [7] S. Graham and G. Kolesnik, *Van der Corput's Method of Exponential Sums*, London Mathematical Society Lecture Note Series 126, Cambridge University Press, 1991.
- [8] G. Hanrot, J. Rivat, G. Tenenbaum and P. Zimmermann, Density results on floating-point invertible numbers, prépublication.
- [9] G.H. Hardy and E.M. Wright, *An introduction to the theory of numbers*, 4th edition, Oxford at the Clarendon Press, 1964.
- [10] C. Hooley, On the greatest prime factor of a quadratic polynomial, *Acta. Math.* 117, 1967, 281-299.
- [11] C. Hooley, *Applications of sieve methods to the theory of numbers*, Cambridge tracts in Mathematics n° 70, Cambridge University Press, 1976.
- [12] M.N. Huxley, Exponential sums and lattice points, III, à paraître, *Proc. London Math. Soc.*
- [13] H. Iwaniec, *Topics in Classical Automorphic Forms*, Graduate Studies in Mathematics, vol. 17, Amer. Math. Soc., 1991.
- [14] Le système MAPLE de calcul formel, <http://www.maplesoft.com/>
- [15] J.-M. Muller, *Elementary Functions, Algorithms and Implementations*, Birkhäuser, 1997.
- [16] A. Nitaj, L'algorithme de Cornacchia, *Expositiones Mathematicæ* 13, 1995, 358-365.
- [17] Le système PARI/GP, <http://www.parigp-home.de>
- [18] W. Sierpinski, Sur un problème du calcul des fonctions asymptotiques, *Prace mat.-fiz.*, 17, 1906, 77-118; et *Œuvres Choisies*, vol. 1, 73-108.
- [19] H.J.S. Smith, *The collected mathematical papers of Henry John Stephen Smith*, vol. 1, Chelsea Publishing Company, 1965.

- [20] J. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc.*, 12, 1985, 183-216.
- [21] A. Weil, On some exponentials sums, *Proc. Nat. Acad. Sci. U.S.A.*, 34, 1948, 204-207.

Jean-Michel Muller

Projet Arénaire, Laboratoire de l'Informatique du Parallélisme, UMR 5668

Ecole Normale Supérieure de Lyon

46 Allée d'Italie,

F-69364 Lyon cedex 07

e-mail : Jean-Michel.Muller@ens-lyon.fr

Jean-Louis Nicolas, Xavier-François Roblot

Institut Girard Desargues, UMR 5028,

Bâtiment Doyen Jean Braconnier,

Université Claude Bernard (Lyon 1),

21 Avenue Claude Bernard,

F-69622 Villeurbanne, France

e-mail : jlnicola@in2p3.fr, roblot@euler.univ-lyon1.fr



Unit é de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit é de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit é de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit é de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit é de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
<http://www.inria.fr>
ISSN 0249-6399