Les nombres premiers de la forme n^2+1 passés au crible (de Selberg)

par Vincent Devinck

professeur CPGE : BCPST1 au lycée Thuillier Amiens

RÉSUMÉ. *Nous présentons la méthode du crible de Selberg sur l'ensemble des nombres premiers de la forme n^2+1. Nous donnons une majoration du nombre de tels nombres premiers plus petits qu'une borne fixée et nous en déduisons que ces nombres premiers, qu'ils soient en nombre fini ou non, sont dans un certain sens, rares.*

MOTS-CLÉS : *crible de Selberg, série de Dirichlet, fonctions multiplicatives*

1. Introduction

L'intérêt majeur de cet article est de mettre en œuvre sur un exemple une *méthode de crible*. D'un point de vue général, une méthode de crible est un procédé ayant pour objectif d'estimer le cardinal d'un certain sous-ensemble de $\mathbb{N}$ qui nous intéresse (et que l'on ne sait pas calculer). Le prétexte est ici l'étude des nombres premiers s'écrivant sous la forme n^2+1 (où $n \in \mathbb{N}$). Ces nombres premiers dits *quadratiques* font l'objet de travaux très profonds, notamment en théorie analytique des nombres et plus particulièrement dans la récente théorie des cribles que nous illustrerons ici avec le crible de Selberg. La célèbre conjecture concernant ces nombres premiers s'énonce comme suit :

« *Il existe une infinité de nombres premiers de la forme n^2+1.* »

Nous nous attachons ici à estimer la proportion de nombres premiers de la forme annoncée parmi les entiers plus petits qu'une borne donnée. Notre objectif est de démontrer le théorème suivant.

Théorème 1. *Pour tout nombre réel $X \geqslant 1$, la proportion de nombres premiers de la forme n^2+1 plus petits que X est majorée, à une constante multiplicative près, par $\dfrac{1}{\ln X}$:*

$$\operatorname{card}\left\{1 \leqslant n \leqslant \sqrt{X}\,;\, n^2+1 \textit{ premier}\right\} \ll \frac{\sqrt{X}}{\ln X}$$

Pour deux fonctions f et g définies sur $[1,+\infty[$ à valeurs positives, la notation $f \ll g$ signifie qu'il existe une constante $C > 0$ telle que pour tout $t \in [1,+\infty[$, on ait $f(t) \leqslant Cg(t)$. La notation $\ll$ de Vinogradov a le même sens que la notation $\mathcal{O}$ de Landau et nous utiliserons librement dans le présent article les deux notations. Nous désignerons par (m,n) (respectivement $[m,n]$) le plus grand diviseur (respectivement multiple) commun des entiers m et n. Nous noterons encore $\lfloor x \rfloor$ la partie entière du nombre réel x. Dans tout l'article, la lettre p fera systématiquement référence à un nombre premier et on notera $d \mid n$ pour dire que l'entier d divise l'entier n. Le lecteur rencontrera constamment des sommes du type $\displaystyle\sum_{\substack{d \leqslant t \\ d \in \mathcal{A}}} a_d$: cela signifie qu'on somme les nombres a_d pour les indices $d \in \{1,\dots,t\}$ et vérifiant une certaine propriété décrite par l'ensemble $\mathcal{A}$.

Le théorème 1 ne permet en aucun cas de répondre à la question de l'infinitude des nombres premiers de la forme n^2+1 mais il fournit tout de même une information non triviale sur leur proportion. En effet, on démontrera qu'ils sont *rares* (corollaire 4) dans l'ensemble des nombres premiers qui est, quant à lui, infini. Signalons que des méthodes probabilistes conduisent à penser qu'il existe une infinité de tels nombres premiers et que l'ordre de grandeur avancé dans le théorème 1 semble être le bon. En effet, Hardy et Littlewood ont conjecturé en 1922 qu'il existerait une constante $c > 0$ telle que

$$\operatorname{card}\left\{1 \leqslant n \leqslant \sqrt{X}\,;\, n^2+1 \text{ premier}\right\} \underset{X \to +\infty}{\sim} \frac{c\sqrt{X}}{\ln X}$$

Pour établir l'estimation annoncée dans le théorème 1, nous mettons en œuvre le crible de Selberg, procédé général qui repose essentiellement sur l'optimisation d'une forme quadratique. La méthode de crible, qui a été largement développée au vingtième siècle et qui est encore aujourd'hui employée dans de profonds travaux, tient son origine dans le crible d'Erathostène qui permettait de détecter les nombres premiers dans l'ensemble des entiers. Nous renvoyons le lecteur au livre [4] pour plus d'informations sur les méthodes de cribles. De tous les cribles, le plus simple dans son exécution est celui développé par A. Selberg entre 1942 et 1947. Il permet d'obtenir des bornes supérieures pour de nombreux problèmes attenant notamment à des conjectures de premier rang. Par exemple :

- La conjecture des nombres premiers jumeaux stipule qu'il existe une infinité de nombres premiers p tels que $p+2$ est également premier (de tels nombres premiers sont alors

dits *jumeaux*). Le crible de Selberg permet de majorer le nombre de tels couples de nombres premiers plus petits qu'une borne donnée :

$$\big\{p \leqslant X\,;\, p \text{ et } p+2 \text{ premiers}\big\} \ll \frac{X}{(\ln X)^2}$$

- La conjecture de Goldbach : « Pour tout entier naturel n pair, il existe deux nombres premiers p et q tels que $n = p + q$. » À l'aide du crible de Selberg, on peut majorer le nombre de représentations d'un entier comme somme de deux nombres premiers : pour tout entier $N \geqslant 2$,

$$\big\{p \leqslant N\,;\, p \text{ et } N-p \text{ premier}\big\} \ll \frac{N}{(\ln N)^2} \prod_{p \mid N} \left(1 + \frac{1}{p}\right)$$

Dans un premier temps, nous allons rappeler quelques résultats élémentaires d'arithmétique concernant les résidus quadratiques modulo p (section 2) ; comme nous étudions la primalité de n^2+1, nous devons savoir pour quels nombres p premiers -1 est un carré dans $\mathbb{Z}/p\mathbb{Z}$. Dans un second temps, nous exposerons la notion de fonction multiplicative, qui est essentielle dans la mise en œuvre du crible de Selberg, puis celle de série de Dirichlet et de représentation d'une telle série en produit eulérien qui nous permettra *in fine* d'aboutir à la majoration en $\dfrac{\sqrt{X}}{\ln X}$ du théorème 1 (section 3). La section 4 est consacrée à l'exécution du crible de Selberg dans notre contexte.

2. Éléments d'arithmétique modulaire

2.1. *Préliminaires*

Nous rappelons sans preuve le petit théorème de Fermat.

Lemme 1 (petit théorème de Fermat)**.** *Soient p un nombre premier et a un entier relatif.*

1. *On a $a^p \equiv a \mod p$.*
2. *Si a n'est pas divisible par p, alors $a^{p-1} \equiv 1 \mod p$.*

On en déduit le résultat suivant.

Corollaire 1 (théorème de Wilson)**.** *Pour tout nombre premier p, on a $(p-1)\,! \equiv -1 \mod p$.*

Démonstration. C'est clair si $p = 2$. Soit p un nombre premier impair. Considérons l'ensemble

$$A = \big\{\overline{1}, \overline{2}, \ldots, \overline{p-1}\big\}$$

des éléments inversibles de $\mathbb{Z}/p\mathbb{Z}$ et plus exactement leur produit $P = \overline{1} \cdot \overline{2} \cdots \overline{p-1}$. Dans l'ensemble A, deux éléments exactement sont leur propre inverse : $\overline{1}$ et $\overline{p-1} = \overline{-1}$. En effet,

si $\overline{x}$ est son propre inverse, alors $\overline{x}^2 = \overline{1}$ et donc p divise $x^2 - 1 = (x-1)(x+1)$. Il s'ensuit donc que p divise $x-1$ ou p divise $x+1$, d'où le résultat. Chacun des autres éléments de A peut être associé à son (unique) inverse, le produit des deux donnant la classe $\overline{1}$, de sorte que le produit P est égal au produit $\overline{1} \cdot \overline{p-1} = -\overline{1}$. Par conséquent, $(p-1)\,! \equiv -1 \mod p$, ce qui démontre le corollaire. □

2.2. *Résidus quadratiques modulo* p

Nous rappelons la notion de résidu quadratique modulo un nombre premier.

Définition 1. *Soit p un nombre premier. On dit qu'un entier a est un résidu quadratique modulo p s'il existe un entier x tel que $x^2 \equiv a \mod p$.*

Si $p = 2$, alors -1 est un résidu quadratique modulo 2 car $-1 \equiv 1 \mod 2$. Nous allons maintenant déterminer pour quels nombres premiers p (impairs) la classe de -1 est un résidu quadratique modulo p et nous allons chercher le nombre de façons de le représenter comme carré dans chaque groupe $\mathbb{Z}/p\mathbb{Z}$. Nous nous intéressons d'abord à ce deuxième problème.

Proposition 1. *Soit p un nombre premier impair et $a \in [\![0, p-1]\!]$. Si a est un résidu quadratique modulo p, alors l'équation $x^2 \equiv a \mod p$ a une seule solution dans $\mathbb{Z}/p\mathbb{Z}$ si $a = 0$ et deux solutions sinon.*

Démonstration. Si $a = 0$, alors $x^2 \equiv 0 \mod p$ si et seulement si p divise x^2, c'est-à-dire si et seulement si p divise x, soit $x \equiv 0 \mod p$. Supposons que $a \in [\![1, p-1]\!]$ est tel que l'équation ait une solution x. Alors $-x$ est aussi solution et $x \not\equiv -x \mod p$. En effet, sinon p diviserait $2x$, ce qui est absurde car p est impair et $x \not\equiv 0 \mod p$. Mais comme $\mathbb{Z}/p\mathbb{Z}$ est un corps, il n'y a pas d'autre solution. □

Proposition 2. *Soient p un nombre premier impair et $a \in [\![1, p-1]\!]$. Alors a est un résidu quadratique modulo p si et seulement si $a^{\frac{p-1}{2}} \equiv 1 \mod p$.*

Démonstration. Si a est un résidu quadratique modulo p, alors il existe un entier x tel que $a \equiv x^2 \mod p$; donc $a^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \mod p$ d'après le petit théorème de Fermat.

Supposons que a ne soit pas un résidu quadratique modulo p. Nous allons montrer que les éléments de l'ensemble $A = \left\{\overline{1}, \overline{2}, \ldots, \overline{p-1}\right\}$ peuvent être regroupés par paires $\{\overline{s}, \overline{t}\}$ (avec $\overline{s} \neq \overline{t}$) telles que $\overline{s} \cdot \overline{t} = \overline{a}$. En effet, pour toute $\overline{s} \in A$, alors $\overline{t} = \left(\overline{s}\right)^{-1} \cdot \overline{t} \in A$ et $\overline{s} \cdot \overline{t} = \overline{a}$. De plus, si $\overline{s} = \overline{t}$, alors $s^2 \equiv a \mod p$, ce qui contredit la définition de a. Comme il y a $\dfrac{p-1}{2}$ paires de cette forme dans A, on a $\overline{1} \cdot \overline{2} \cdots \overline{p-1} = \left(\overline{a}\right)^{\frac{p-1}{2}}$, ce qui fournit $a^{\frac{p-1}{2}} \equiv -1 \mod p$ d'après le théorème de Wilson. □

On déduit de cette proposition les nombres premiers p pour lesquels -1 est un carré dans $\mathbb{Z}/p\mathbb{Z}$.

Corollaire 2. *Soit p un nombre premier impair. Alors -1 est un résidu quadratique modulo p si et seulement si $p \equiv 1 \mod 4$.*

Démonstration. Si p est impair, alors $(-1)^{\frac{p-1}{2}} \equiv 1 \mod p$ si et seulement si $\dfrac{p-1}{2}$ est pair, c'est-à-dire si et seulement si $p \equiv 1 \mod 4$. □

Dans toute la suite, on notera ρ la fonction définie sur $\mathbb{N}^*$ qui à chaque entier d compte le nombre de représentations de -1 comme carré dans $\mathbb{Z}/d\mathbb{Z}$:

$$\rho(d) = \text{card}\big\{x \in [\![0, d-1]\!] \,;\, x^2 \equiv -1 \mod d\big\} \tag{1}$$

D'après ce qui précède, on sait que

$$\text{pour tout nombre premier } p, \qquad \rho(p) = \begin{cases} 1 & \text{si } p = 2 \\ 2 & \text{si } p \equiv 1 \mod 4 \\ 0 & \text{si } p \equiv 3 \mod 4 \end{cases}$$

Dans la suite, nous aurons besoin d'estimer le nombre de représentants de -1 comme carré dans $\mathbb{Z}/d\mathbb{Z}$, dans un intervalle $[1, y]$ où y est un nombre réel supérieur ou égal à 1. On dira qu'un entier d est *sans facteur carré* s'il n'existe pas de nombre premier p tel que p^2 divise d.

Proposition 3. *Pour tout nombre réel $y \geqslant 1$ et pour tout entier naturel d non nul sans facteur carré, le nombre d'entiers $n \in [1, y]$ tel que d divise $n^2 + 1$ est égal à $y\dfrac{\rho(d)}{d} + \mathcal{O}(\rho(d))$, ce que l'on peut réécrire* :

$$\sum_{\substack{n \leqslant y \\ d \mid n^2+1}} 1 = y\frac{\rho(d)}{d} + \mathcal{O}(\rho(d))$$

Démonstration. L'entier d étant sans facteur carré, notons $d = p_1 \dots p_k$ sa décomposition en produit de facteurs premiers distincts. Un entier est divisible par d si et seulement s'il est divisible par chacun des nombres premiers $p_1, \dots, p_k$. On en déduit que si $n \in \mathbb{N}^*$ alors

$$\begin{aligned} (d \text{ divise } n^2+1) &\iff (\forall \ell \in [\![1, k]\!],\ n^2 + 1 \equiv 0 \mod p_\ell) \\ &\iff (\forall \ell \in [\![1, k]\!],\ -1 \text{ est un résidu quadratique modulo } p_\ell) \end{aligned}$$

D'après le corollaire 2, il n'y a donc pas de multiple de d de la forme $n^2 + 1$ si d possède un facteur premier p tel que $p \equiv 3 \mod 4$. Si d ne possède pas de facteur premier de cette forme, alors pour chaque nombre premier p_j, le nombre de façons de représenter $-1 \mod p_j$ comme un carré dans $\mathbb{Z}/p_j\mathbb{Z}$ est égal à $\rho(p_j)$. On en déduit que le nombre d'entiers $n \in [\![0, d-1]\!]$ tels que d divise $n^2 + 1$ est égal à $\rho(d) = \prod_{j=1}^{k} \rho(p_j)$. De même, pour tout entier k tel que $1 \leqslant k \leqslant \left\lfloor \dfrac{y}{d} \right\rfloor - 1$, le nombre d'entiers $n \in [\![kd, (k+1)d-1]\!]$ tels que d divise n^2+1

est égal à $\rho(d)$ (puisque n est tel que $d \mid n^2+1$ si et seulement si l'entier $n-dk$ est tel que $d \mid (n-dk)^2+1$). Enfin il y a au plus $\rho(d)$ entiers $n \in [d\lfloor y/d \rfloor, y]$ tels que d divise n^2+1. On en déduit le résultat annoncé. □

3. Multiplicativité et séries de Dirichlet

On appelle fonction arithmétique toute fonction f définie sur $\mathbb{N}^*$ à valeurs complexes. Dans cette large classe de fonctions, on distingue les fonctions multiplicatives qui préservent la structure multiplicative des entiers.

3.1. *Fonctions multiplicatives*

Définition 2. *Soit $f : \mathbb{N}^* \longrightarrow \mathbb{C}$ une fonction arithmétique. On dit que f est multiplicative si elle vérifie les deux conditions suivantes* :

- $f(1) = 1$;
- *pour tout couple $(m,n) \in (\mathbb{N}^*)^2$ d'entiers premiers entre eux, on a $f(mn) = f(m)f(n)$.*

Une fonction multiplicative f est donc complètement déterminée par ses valeurs sur les puissances de nombres premiers. En effet, si $n = p_1^{\alpha_1} \dots p_k^{\alpha_k}$ est la décomposition de n en produit de facteurs premiers (les nombres premiers $p_1, \dots, p_k$ sont ici distincts), alors la multiplicativité de f entraîne l'égalité

$$f(n) = \prod_{i=1}^{k} f(p_i^{\alpha_i})$$

Dans la suite, nous aurons besoin de la proposition suivante.

Proposition 4. *Soit f une fonction multiplicative. Pour tout $(m,n) \in (\mathbb{N}^*)^2$, on a l'égalité*

$$f(m)f(n) = f((m,n))f([m,n])$$

Démonstration. Soit $(m,n) \in (\mathbb{N}^*)^2$. On note $n = \prod_p p^{v_p(n)}$ (de même pour m) la décomposition en produit de facteurs premiers de n. Ici, $v_p(n)$ désigne la plus grand entier α pour lequel p^α divise n (on l'appelle la valuation p-adique de n). En particulier, si p ne divise pas n, alors $v_p(n) = 0$ et donc le produit est fini. On sait que

$$(m,n) = \prod_p p^{\min(v_p(m), v_p(n))} \qquad \text{et} \qquad [m,n] = \prod_p p^{\max(v_p(m), v_p(n))}$$

et comme $\{v_p(m), v_p(n)\} = \{\max(v_p(m), v_p(n)), \min(v_p(m), v_p(n))\}$ pour tout nombre premier p, il vient

$$\begin{aligned} f(m)f(n) &= \prod_p f\big(p^{v_p(m)}\big) \prod_p f\big(p^{v_p(n)}\big) \\ &= \prod_p \Big(f\big(p^{v_p(m)}\big) f\big(p^{v_p(n)}\big)\Big) \\ &= \prod_p \Big(f\big(p^{\max(v_p(m),v_p(n))}\big) f\big(p^{\min(v_p(m),v_p(n))}\big)\Big) \\ &= \prod_p f\big(p^{\max(v_p(m),v_p(n))}\big) \prod_p f\big(p^{\min(v_p(m),v_p(n))}\big) \\ &= f([m,n])f((m,n)) \end{aligned}$$

par multiplicativité de f, ce qui démontre la proposition. □

3.2. *Premiers exemples*

Nous donnons ici quelques exemples de fonctions multiplicatives.

- ⋆ La fonction constante $\mathbf{1}$ définie par $\mathbf{1}(n) = 1$ pour tout $n \in \mathbb{N}^*$ est multiplicative.
- ⋆ La fonction identité Id définie par $\mathrm{Id}(n) = n$ pour tout $n \in \mathbb{N}^*$ est multiplicative
- ⋆ La fonction indicatrice de $\{1\}$, notée δ_1, qui est donc définie sur $\mathbb{N}^*$ par

  $$\forall n \in \mathbb{N}^*, \quad \delta_1(n) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n > 1 \end{cases}$$

  est multiplicative.
- ⋆ La fonction de Möbius, notée μ, définie sur $\mathbb{N}^*$ par $\mu(1) = 1$ et pour tout entier $n \geqslant 2$,
  - $\mu(n) = 0$ si n a un facteur carré
  - $\mu(n) = 1$ si n est le produit d'un nombre pair de nombres premiers distincts.
  - $\mu(n) = -1$ si n est le produit d'un nombre impair de nombres premiers distincts

  est multiplicative. En effet, si m et n sont deux entiers premiers entre eux, alors ou bien au moins l'un d'entre eux admet un facteur carré et alors mn a aussi un facteur carré et

  $$\mu(mn) = 0 = \mu(m)\mu(n)$$

  ou bien m et n sont le produit de respectivement t et s nombres premiers distincts et alors

  $$\mu(m)\mu(n) = (-1)^s(-1)^t = (-1)^{s+t} = \mu(mn)$$

  car mn est le produit de $t + s$ nombres premiers deux à deux distincts (puisque m et n sont premiers entre eux).

- ⋆ La fonction μ^2 est multiplicative. Elle correspond à la fonction indicatrice des entiers sans facteurs carrés.

- ⋆ Le *caractère de Dirichlet non principal modulo 4*, noté χ, défini par

$$\chi(n) = \begin{cases} 0 & \text{si } n \equiv 0 \bmod 2 \\ 1 & \text{si } n \equiv 1 \bmod 4 \\ -1 & \text{si } n \equiv 3 \bmod 4 \end{cases}$$

  est multiplicatif.

- ⋆ La fonction ρ que nous avons définie en (1) et qui nous sera utile dans la suite est multiplicative. Démontrons-le. Tout d'abord, on a clairement $\rho(1) = 1$. Si $k \in \mathbb{N}^*$, notons

$$C_k = \left\{ x \bmod k \,;\, x^2 \equiv -1 \bmod k \right\}$$

  Soit $(m, n) \in (\mathbb{N}^*)^2$ tel que $(m, n) = 1$. D'après le lemme chinois, l'application

$$\psi : \begin{cases} \mathbb{Z}/mn\mathbb{Z} & \longrightarrow & \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \\ x \bmod mn & \longmapsto & (x \bmod m, x \bmod n) \end{cases}$$

  est un isomorphisme de groupes. Pour tout $x \in \mathbb{Z}$, on a

$$\begin{aligned} x \bmod mn \in C_{mn} &\iff mn \mid x^2 + 1 \iff (m \mid x^2 + 1 \text{ et } n \mid x^2 + 1) \\ &\iff (x \bmod m \in C_m \text{ et } x \bmod n \in C_n) \end{aligned} \tag{2}$$

  où la deuxième équivalence provient du fait que m et n sont premiers entre eux. En particulier,

$$\psi(C_{mn}) \subset C_m \times C_n$$

  Si $(y \bmod m, z \bmod n) \in C_m \times C_n$, alors il existe $x \in \mathbb{Z}$ tel que $x \equiv y \bmod m$ et $x \equiv z \bmod n$ (par surjectivité de ψ). Donc $x \bmod m \in C_m$ et $x \bmod n \in C_n$ (par définition de y et z) et donc, d'après les équivalences précédentes (2), on a $x \bmod mn \in C_{mn}$. Finalement, $\psi(C_{mn}) = C_m \times C_n$. Comme les ensembles C_{mn} et $C_m \times C_n$ sont en bijection, nous avons $\operatorname{card}(C_{mn}) = \operatorname{card}(C_m \times C_n) = \operatorname{card}(C_m)\operatorname{card}(C_n)$ c'est-à-dire $\rho(mn) = \rho(m)\rho(n)$. Finalement, la fonction ρ est multiplicative.

Le produit de convolution, que nous allons maintenant définir, permet de construire de nouveaux exemples de fonctions multiplicatives.

3.3. *Produit de convolution et multiplicativité*

Nous introduisons ici la notion de convolution de fonctions arithmétiques.

Définition 3. *Le produit de convolution de deux fonctions arithmétiques f et g est la fonction arithmétique notée $f \star g$ et définie par*

$$\forall n \in \mathbb{N}^*, \qquad (f \star g)(n) = \sum_{d|n} f(d) g\left(\frac{n}{d}\right)$$

où la somme porte sur les entiers $d \geqslant 1$ divisant n.

Remarque. Une autre façon d'écrire le produit de convolution de f et g est :

$$\forall n \in \mathbb{N}^*, \qquad (f \star g)(n) = \sum_{ab=n} f(a) g(b)$$

où la somme est étendue aux couples (a, b) d'entiers naturels non nuls dont le produit vaut n.

Nous allons démontrer que l'ensemble des fonctions multiplicatives est stable pour l'opération de convolution (proposition 5). Nous aurons besoin du résultat élémentaire ci-dessous. Pour tout entier naturel n non nul, on note $\mathcal{D}(n)$ l'ensemble des diviseurs (positifs) de n. Par exemple, $\mathcal{D}(10) = \{1, 2, 5, 10\}$.

Lemme 2. *Soient m et n deux entiers naturels non nuls premiers entre eux. L'application*

$$\vartheta : \left\{ \begin{array}{ccc} \mathcal{D}(m) \times \mathcal{D}(n) & \longrightarrow & \mathcal{D}(mn) \\ (d, \delta) & \longmapsto & d\delta \end{array} \right.$$

est bijective.

Démonstration. C'est évident si m ou n est égal à 1. On suppose désormais que $m \geqslant 2$ et $n \geqslant 2$. On décompose ces nombres en produits de facteurs premiers : il existe $(r, s) \in (\mathbb{N}^*)^2$, des nombres premiers $p_1, \ldots, p_r, q_1, \ldots, q_s$ deux à deux distincts et des entiers naturels non nuls $\alpha_1, \ldots, \alpha_r, \beta_1, \ldots, \beta_s$ tels que

$$m = \prod_{i=1}^{r} p_i^{\alpha_i} \qquad \text{et} \qquad n = \prod_{i=1}^{s} q_i^{\beta_i}$$

On a $\operatorname{card}(\mathcal{D}(m)) = \prod_{i=1}^{r} (\alpha_i + 1)$ et $\operatorname{card}(\mathcal{D}(n)) = \prod_{i=1}^{s} (\beta_i + 1)$. En écrivant mn en produit de facteurs premiers, on trouve que

$$\operatorname{card}(\mathcal{D}(mn)) = \prod_{i=1}^{r} (\alpha_i + 1) \times \prod_{i=1}^{r} (\beta_i + 1) = \operatorname{card}(\mathcal{D}(m)) \times \operatorname{card}(\mathcal{D}(n))$$

Pour démontrer le lemme, il suffit donc de vérifier que ϑ est surjective. Si ℓ est un diviseur de mn, alors pour tout $i \in [\![1, r]\!]$ et tout $j \in [\![1, s]\!]$, il existe $a_i \in [\![0, \alpha_i]\!]$ et $b_j \in [\![0, \beta_j]\!]$ tels que $\ell = p_1^{a_1} \ldots p_r^{a_r} q_1^{b_1} \ldots q_s^{b_s}$. Les entiers $d = p_1^{a_1} \ldots p_r^{a_r}$ et $\delta = q_1^{b_1} \ldots q_s^{b_s}$ sont des diviseurs de m et n respectivement tels que $\vartheta(m, n) = \ell$. □

On en déduit le résultat de stabilité annoncé.

Proposition 5. *Soient $f : \mathbb{N}^* \longrightarrow \mathbb{C}$ et $g : \mathbb{N}^* \longrightarrow \mathbb{C}$ deux fonctions multiplicatives. Alors la fonction $f \star g$ est multiplicative.*

Démonstration. Tout d'abord, $(f \star g)(1) = f(1)g(1) = 1$ et si m et n sont deux entiers naturels non nuls premiers entre eux, alors

$$(f \star g)(mn) = \sum_{\ell | mn} f(\ell) g\left(\frac{mn}{\ell}\right) = \sum_{\ell \in \mathcal{D}(mn)} f(\ell) g\left(\frac{mn}{\ell}\right)$$

D'après le lemme 2, on peut réécrire la somme comme :

$$(f \star g)(mn) = \sum_{(d,\delta) \in \mathcal{D}(m) \times \mathcal{D}(n)} f(d\delta) g\left(\frac{mn}{d\delta}\right) = \sum_{\substack{d|m \\ \delta|n}} f(d\delta) g\left(\frac{mn}{d\delta}\right)$$

Comme les entiers m et n sont premiers entre eux, leurs diviseurs respectifs d et δ d'une part, $\dfrac{m}{d}$ et $\dfrac{n}{\delta}$ d'autre part, sont premiers entre eux. Comme f et g sont des fonctions multiplicatives, on a

$$f(d\delta) = f(d)f(\delta) \qquad \text{et} \qquad g\left(\frac{mn}{d\delta}\right) = g\left(\frac{m}{d}\right) g\left(\frac{n}{\delta}\right)$$

On peut maintenant achever le calcul :

$$\begin{aligned}
(f \star g)(mn) &= \sum_{\substack{d|m \\ d|n}} f(d)f(\delta) g\left(\frac{m}{d}\right) g\left(\frac{n}{\delta}\right) \\
&= \left(\sum_{d|m} f(d) g\left(\frac{m}{d}\right)\right)\left(\sum_{\delta|n} f(\delta) g\left(\frac{n}{\delta}\right)\right) \\
&= (f \star g)(m)(f \star g)(n)
\end{aligned}$$

Finalement, la fonction $f \star g$ est multiplicative. □

La convolution permet donc de construire de nouvelles fonctions multiplicatives, par exemple :

- ⋆ la fonction diviseur τ, qui associe à tout entier naturel n non nul son nombre de diviseurs $\tau(n)$ est multiplicative : $\tau(n) = \displaystyle\sum_{d|n} 1 = (\mathbf{1} \star \mathbf{1})(n)$;
- ⋆ la fonction somme des diviseurs σ qui à tout $n \in \mathbb{N}^*$ associe la somme $\sigma(n)$ des diviseurs de n est multiplicative : $\sigma(n) = \displaystyle\sum_{d|n} d = (\mathbf{1} \star \mathrm{Id})(n)$;

- $\star$ la fonction indicatrice d'Euler φ définie par :

$$\forall n \in \mathbb{N}^*, \qquad \varphi(n) = \operatorname{card}\left\{k \in [\![0, n-1]\!]\,;\, (m,k) = 1\right\}$$

  est multiplicative. On peut vérifier que $\varphi = \mu \star \mathrm{Id}$.

Voici quelques propriétés élémentaires supplémentaires concernant l'opération de convolution.

Proposition 6. *Dans l'ensemble des fonctions arithmétiques, l'opération de convolution $\star$ est associative, commutative et admet pour élément neutre δ_1. De plus, la fonction constante $\mathbf{1}$ est inversible pour cette opération d'inverse la fonction μ :*

$$\mu \star \mathbf{1} = \delta_1 \quad (\textit{formule d'inversion de Möbius})$$

Démonstration. Nous démontrons uniquement la formule d'inversion de Möbius. Les fonctions δ_1 et $\mu \star \mathbf{1}$ étant multiplicatives, il suffit de montrer que la fonction $\mu \star \mathbf{1}$ s'annule sur les puissances des nombres premiers. Soit p un nombre premier. Alors $\mu(p) = -1$ et si $\beta \in \mathbb{N} \setminus \{0,1\}$, alors $\mu(p^\beta) = 0$. Pour tout $\alpha \in \mathbb{N}^*$, on a donc

$$(\mu \star \mathbf{1})(p^\alpha) = \sum_{\beta=0}^{\alpha} \mu(p^\beta) = \mu(1) + \mu(p) = 0$$

d'où le résultat. □

Remarque. Plus généralement, on montre que toute fonction multiplicative est inversible pour $\star$; les fonctions multiplicatives constituent donc un groupe pour cette opération.

Étant donné une fonction arithmétique f, une question naturelle est de déterminer l'ordre de grandeur de sa fonction sommatoire $F(x) := \sum_{n \leqslant x} f(n)$. Nous serons confronté à ce problème pour démontrer le théorème 1. Nous allons pour cela expliciter la série de Dirichlet associée à f.

3.4. *Séries de Dirichlet*

On associe naturellement à une fonction arithmétique sa série de Dirichlet.

Définition 4. *Soit f une fonction arithmétique. La série de Dirichlet associée à f est la fonction $s \longmapsto D(f,s)$ de la variable réelle s définie a priori formellement par*

$$D(f,s) = \sum_{n \geqslant 1} \frac{f(n)}{n^s} \tag{3}$$

La série (3) peut diverger en tout point comme en témoigne la fonction multiplicative f définie par $f(n) = 2^n$. L'exemple de série de Dirichlet par excellence est la fonction zêta de Riemann. Il s'agit de la série associée à la fonction multiplicative $\mathbf{1}$:

$$\zeta(s) = D(\mathbf{1}, s) = \sum_{n=1}^{+\infty} \frac{1}{n^s}$$

La série est ici (absolument) convergente en tout point s de $]1, +\infty[$.

L'importance des fonctions multiplicatives en théorie analytique des nombres réside dans la possibilité de représenter les séries de Dirichlet associées en produit eulérien.

Proposition 7. *Soit f une fonction multiplicative. Alors la série de Dirichlet $D(f, s)$ converge absolument si et seulement si la série $\displaystyle\sum_{p\geqslant 2}\sum_{k=1}^{+\infty} \frac{|f(p^k)|}{p^{ks}}$ est convergente. Lorsque la série de Dirichlet $D(f, s)$ converge absolument, on a la représentation en produit eulérien suivante* :

$$D(f, s) = \prod_{p\geqslant 2} \left(\sum_{k=0}^{+\infty} \frac{f(p^k)}{p^{ks}} \right) = \prod_{p\geqslant 2} \left(1 + \sum_{k=1}^{+\infty} \frac{f(p^k)}{p^{ks}} \right)$$

Démonstration. Soit $s \in \mathbb{R}$. Supposons que la série $D(f, s)$ converge absolument. Pour tout nombre premier p et pour tout entier naturel N, on a $\displaystyle\sum_{k=1}^{N} \frac{|f(p^k)|}{p^{ks}} \leqslant D(|f|, s)$ donc la série à termes positifs $\displaystyle\sum_{k\geqslant 1} |f(p^k)|p^{-ks}$ converge. Pour tout $N \geqslant 2$, on a toujours la majoration $\displaystyle\sum_{p\leqslant N}\sum_{k=1}^{+\infty} \frac{|f(p^k)|}{p^{ks}} \leqslant D(|f|, s)$ (puisque les entiers de la forme p^k avec p premier et $k \geqslant 1$ sont deux à deux distincts) donc la série double est bien convergente. Supposons maintenant que la série double soit convergente. En particulier, la série $s_p = \displaystyle\sum_{k\geqslant 1} |f(p^k)|p^{-ks}$ est convergente pour tout nombre premier p. La majoration $1 + x \leqslant \mathrm{e}^x$ (valable pour tout nombre réel x) fournit, pour tout entier $N \geqslant 2$, la majoration

$$\prod_{p\leqslant N} \left(1 + \sum_{k=1}^{+\infty} \frac{|f(p^k)|}{p^{ks}} \right) \leqslant \prod_{p\leqslant n} \mathrm{e}^{s_p} \leqslant \mathrm{e}^{\sum_{p\geqslant 2} s_p}$$

ce qui montre que le produit infini converge (il est la limite d'une suite croissante bornée). Il reste à établir le développement en produit eulérien de $D(f, s)$ en un point de convergence absolue s. Soit $N \geqslant 2$ un entier. Tout élément du produit $\displaystyle\prod_{p\leqslant N}\sum_{k=0}^{+\infty} \frac{f(p^k)}{p^{ks}}$ s'écrit sous la forme

$\dfrac{f(p_1^{k_1})\dots f(p_\ell^{k_\ell})}{p_1^{k_1 s}\dots p_\ell^{k_\ell s}}$ où les p_i sont des nombres premiers distincts inférieurs ou égaux à N et où les k_i sont des entiers naturels. Par multiplicativité de f,

$$\frac{f(p_1)^{k_1}\dots f(p_\ell)^{k_\ell}}{p_1^{k_1 s}\dots p_\ell^{k_\ell s}} = \frac{f(p_1^{k_1}\dots p_\ell^{k_\ell})}{(p_1^{k_1}\dots p_\ell^{k_\ell})^s}$$

Les éléments du produit sont donc les termes de la forme $f(n)/n^s$ où l'entier n est tel que son plus grand facteur premier $P^+(n)$ soit inférieur ou égal à N. On en déduit donc que

$$\left|\sum_{n=1}^{+\infty}\frac{f(n)}{n^s} - \prod_{p\leqslant N}\sum_{k=0}^{+\infty}\frac{f(p^k)}{p^{ks}}\right| = \left|\sum_{\substack{n\geqslant 1\\ P^+(n)>N}}\frac{f(n)}{n^s}\right| \leqslant \sum_{\substack{n\geqslant 1\\ P^+(n)>N}}\frac{|f(n)|}{n^s}$$
$$\leqslant \sum_{n>N}\frac{|f(n)|}{n^s}$$

où la dernière majoration provient de l'inclusion $\{n\in\mathbb{N}^*\,;\,P^+(n)>N\}\subset[N,+\infty[\cap\mathbb{N}$. On obtient le résultat en faisant tendre N vers $+\infty$. □

La fonction ζ de Riemann converge (absolument) dans l'intervalle $]1,+\infty[$ donc on a la représentation en produit eulérien suivante :

$$\forall s>1,\qquad \zeta(s)=\prod_{p\geqslant 2}\left(1+\frac{1}{p^s}+\frac{1}{p^{2s}}+\dots\right)=\prod_{p\geqslant 2}\frac{1}{1-\frac{1}{p^s}}$$

Une autre propriété fondamentale est la dualité qui existe entre la convolution des fonctions multiplicatives d'une part et le produit des séries de Dirichlet d'autre part.

Proposition 8. *Soient f et g deux fonctions multiplicatives dont les séries de Dirichlet convergent absolument en un même point $s_0\in\mathbb{R}$. Alors la série de Dirichlet $D(f\star g,s_0)$ est absolument convergente et*

$$D(f\star g,s_0)=D(f,s_0)\times D(g,s_0)$$

Démonstration. On sait que les séries de Dirichlet $D(f,s_0)$ et $D(g,s_0)$ convergent absolument donc le produit de ces deux séries, en tant que familles sommables, converge aussi absolument. Le terme général de cette série produit est

$$\sum_{\substack{(a,b)\in(\mathbb{N}^*)^2\\ ab=n}}\frac{f(a)g(b)}{a^{s_0}b^{s_0}}=\frac{1}{n^{s_0}}\sum_{\substack{(a,b)\in(\mathbb{N}^*)^2\\ ab=n}}f(a)g(b)=\frac{(f\star g)(n)}{n^{s_0}}\qquad(n\in\mathbb{N}^*)$$

d'où le résultat. □

4. Mise en œuvre du crible de Selberg (démonstration du théorème 1)

La conjecture évoquée dans l'introduction portant sur l'infinitude des nombres premiers de la forme n^2+1 peut se réécrire :

$$\lim_{X\to+\infty} \operatorname{card}\left\{1 \leqslant n \leqslant \sqrt{X}\,;\, n^2+1 \text{ premier}\right\} = +\infty$$

Si $\mathcal{P}$ désigne l'ensemble des nombres premiers et si $\mathbf{1}_{\mathcal{P}}$ est sa fonction indicatrice, alors

$$\operatorname{card}\left\{1 \leqslant n \leqslant \sqrt{X}\,;\, n^2+1 \text{ premier}\right\} = \sum_{n\leqslant\sqrt{X}} \mathbf{1}_{\mathcal{P}}(n^2+1)$$

ce que nous préférerons écrire $\displaystyle\sum_{\substack{n\leqslant\sqrt{X}\\ n^2+1 \text{ premier}}} 1$. Ainsi, démontrer la conjecture revient à disposer d'informations *suffisamment précises* de la fonction $n \longmapsto \mathbf{1}_{\mathcal{P}}(n^2+1)$. Les méthodes de crible consistent à construire des estimations (minoration ou majoration) de telles fonctions. Nous allons en effet voir ci-dessous qu'il est assez aisé de trouver une fonction majorante exploitable de $n \longmapsto \mathbf{1}_{\mathcal{P}}(n^2+1)$ pour le crible de Selberg. La recherche de minoration de notre fonction est beaucoup plus délicate et fait l'objet de méthodes de crible plus avancées (le crible de Bombieri [3] par exemple). De plus, ces minorations sont en général obtenues pour une fonction différente mais assez proche de la fonction initiale $n \longmapsto \mathbf{1}_{\mathcal{P}}(n^2+1)$. Par exemple, H. Iwaniec [5] a démontré en 1978 qu'il existe une infinité d'entiers n tels que n^2+1 ait au plus deux facteurs premiers, ce qui constitue actuellement le meilleur résultat sur le sujet.

4.1. *Présentation du crible*

Pour tout $X \in [1,+\infty[$, considérons la somme

$$S(X,z) = \sum_{n\leqslant\sqrt{X}} \left(\sum_{d\mid n^2+1} \lambda_d \right)^2$$

où :

- z est un paramètre réel dépendant de X que nous choisirons dans la suite (une petite puissance de X) ;
- la somme intérieure porte sur les diviseurs positifs d de n^2+1 inférieurs ou égaux à z ;
- les λ_d sont des nombres réels tels que $\lambda_1 = 1$, $\lambda_d = 0$ si d a un facteur carré et $\lambda_d = 0$ si $d > z$.

Nous expliquons maintenant pourquoi $\left(\sum_{d|n^2+1}\lambda_d\right)^2$ est, dans le cas $n>z$,

une fonction majorante de notre fonction indicatrice $\mathbf{1}_{\mathcal{P}}(n^2+1)$. Si $n\in\left[1,\sqrt{X}\right]$ est tel que n^2+1 est un nombre premier (c'est-à-dire $\mathbf{1}_{\mathcal{P}}(n^2+1)=1$), alors les diviseurs d de n^2+1 sont 1 et n^2+1. Or si $n>z$, alors on a $\lambda_{n^2+1}=0$ (par définition de λ_d pour $d>z$). Ceci entraîne la majoration suivante :

$$\sum_{\substack{z<n\leqslant\sqrt{X}\\ n^2+1 \text{ premier}}} 1 \leqslant \sum_{n\leqslant\sqrt{X}}\left(\sum_{d|n^2+1}\lambda_d\right)^2 \tag{4}$$

Nous n'avons pas encore traité les nombres premiers de la forme n^2+1 pour $n\in[1,z]$. Il y a au plus $\lfloor z\rfloor$ nombres premiers de cette forme donc

$$\sum_{\substack{n\leqslant\sqrt{X}\\ n^2+1 \text{ premier}}} 1 \leqslant \sum_{\substack{z<n\leqslant\sqrt{X}\\ n^2+1 \text{ premier}}} 1+z \tag{5}$$

En combinant les inégalités (4) et (5), on obtient :

$$\sum_{\substack{n\leqslant\sqrt{X}\\ n^2+1 \text{ premier}}} 1 \leqslant S(X,z)+z \tag{6}$$

Dans la suite, nous expliquons comment nous exploitons la somme $S(X,z)$ dans le but d'établir l'estimation annoncée au théorème 1.

4.2. *Décomposition de la somme* $S(X,z)$

Nous commençons par développer le carré intérieur. La somme $S(X,z)$ se réécrit alors

$$S(X,z)=\sum_{n\leqslant\sqrt{X}}\sum_{\substack{d_1|n^2+1\\ d_2|n^2+1}}\lambda_{d_1}\lambda_{d_2}$$

où la somme porte sur les entiers d_1 et d_2 inférieurs ou égaux à z qui divisent n^2+1. Nous permutons ensuite les deux symboles $\sum$, ce qui nous donne

$$S(X,z)=\sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}}\lambda_{d_1}\lambda_{d_2}\sum_{\substack{n\leqslant\sqrt{X}\\ d_1|n^2+1\\ d_2|n^2+1}}1=\sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}}\lambda_{d_1}\lambda_{d_2}\sum_{\substack{n\leqslant\sqrt{X}\\ [d_1,d_2]|n^2+1}}1$$

car n^2+1 est divisible par les entiers d_1 et d_2 si et seulement si il est divisible par leur plus petit commun multiple $[d_1,d_2]$. Le caractère arithmétique du crible de Selberg apparaît dans la somme intérieure. Il s'agit de disposer d'informations précises sur la représentation de $-1 \bmod\ [d_1,d_2]$ comme un carré. Nous avons déjà étudié ce problème dans la proposition 3 : la somme intérieure vaut

$$\sum_{\substack{n\leqslant\sqrt{X}\\ [d_1,d_2]|n^2+1}} 1 = \sqrt{X}\,\frac{\rho([d_1,d_2])}{[d_1,d_2]} + \mathcal{O}(\rho([d_1,d_2]))$$

L'intérêt majeur est ici d'avoir introduit de la multiplicativité par l'intermédiaire de la fonction ρ. En réinjectant cette estimation dans l'expression de $S(X,z)$ précédemment obtenue, il vient

$$S(X,z) = \sqrt{X}\sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}} \lambda_{d_1}\lambda_{d_2}\frac{\rho([d_1,d_2])}{[d_1,d_2]} + \mathcal{O}\left(\sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1,d_2])\right)$$

Finalement, la somme peut s'écrire sous la forme $S(X,z) = \sqrt{X}\,F(z)+E(z)$ où nous avons posé

$$F(z) = \sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}} \lambda_{d_1}\lambda_{d_2}\frac{\rho([d_1,d_2])}{[d_1,d_2]} \qquad \text{et} \qquad E(z) = \mathcal{O}\left(\sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1,d_2])\right)$$

Dans la suite, nous nous attachons à estimer la *partie principale* $\sqrt{X}F(z)$ puis le *terme d'erreur* $E(z)$ de la somme $S(X,z)$. Nous verrons que pour un choix adapté du paramètre z, le terme $E(z)$ est effectivement négligeable devant $\sqrt{X}F(z)$.

4.3. *Étude de la partie principale* $F(z)$

La première étape consiste à écrire la forme quadratique $F(z)$ sous forme *diagonale*.

4.3.1. *Forme diagonale de* $F(z)$

L'idée générale est de séparer les variables d_1 et d_2 qui apparaissent dans $\dfrac{\rho([d_1,d_2])}{[d_1,d_2]}$. Dans toute la suite, nous noterons f la fonction arithmétique définie par

$$\forall d\in\mathbb{N}^*, \qquad f(d) = \begin{cases} \dfrac{d}{\rho(d)} & \text{si } d \text{ n'a pas de facteur premier } p\equiv 3 \bmod 4\\ 0 & \text{sinon}\end{cases}$$

Ainsi définie, la fonction f est multiplicative et d'après la proposition 4, nous pouvons écrire que

$$\forall (d_1,d_2)\in(\mathbb{N}^*)^2, \qquad \frac{\rho([d_1,d_2])}{[d_1,d_2]} = \frac{\rho(d_1)}{d_1}\times\frac{\rho(d_2)}{d_2}\times f((d_1,d_2))$$

ce qui fournit l'écriture de $F(z)$ suivante :

$$F(z) = \sum_{\substack{d_1 \leqslant z \\ d_2 \leqslant z}} \lambda_{d_1} \frac{\rho(d_1)}{d_1} \lambda_{d_2} \frac{\rho(d_2)}{d_2} f((d_1, d_2))$$

Nous devons encore séparer les variables d_1 et d_2 dans $f((d_1, d_2))$. Pour cela, considérons la fonction multiplicative g définie par $g = \mu \star f$. On sait d'après la formule d'inversion de Möbius que $\mathbf{1} \star g = f$. Par conséquent,

$$f((d_1, d_2)) = (\mathbf{1} \star g)((d_1, d_2)) = \sum_{\delta \mid (d_1, d_2)} g(\delta) = \sum_{\substack{\delta \mid d_1 \\ \delta \mid d_2}} g(\delta)$$

car δ divise (d_1, d_2) si et seulement si δ divise d_1 et d_2. On permute ensuite les deux sommes :

$$\begin{aligned} F(z) &= \sum_{\delta \leqslant z} g(\delta) \left(\sum_{\substack{d_1 \leqslant z \\ \delta \mid d_1}} \lambda_{d_1} \frac{\rho(d_1)}{d_1} \right) \left(\sum_{\substack{d_2 \leqslant z \\ \delta \mid d_2}} \lambda_{d_2} \frac{\rho(d_2)}{d_2} \right) \\ &= \sum_{\delta \leqslant z} g(\delta) \left(\sum_{\substack{d \leqslant z \\ \delta \mid d}} \lambda_d \frac{\rho(d)}{d} \right)^2 \end{aligned}$$

où la somme intérieure porte sur les multiples d de δ inférieurs ou égaux à z. Nous avons ainsi obtenu la forme diagonale de $F(z)$ cherchée :

$$F(z) = \sum_{\delta \leqslant z} g(\delta) y_\delta^2 \qquad \text{où} \qquad y_\delta = \sum_{\substack{d \leqslant z \\ \delta \mid d}} \lambda_d \frac{\rho(d)}{d} \tag{7}$$

L'idée est ensuite de trouver les nombres y_δ qui minimisent la forme quadratique $F(z)$ sous la condition de minimisation $\lambda_1 = 1$ (c'est la seule condition essentielle sur les λ_d que nous avons imposée). Pour traduire cette condition de minimisation sur les coefficients y_δ, nous aurons besoin d'exprimer les λ_ℓ en fonction des y_δ.

4.3.2. *Expression des λ_ℓ en fonction des y_δ*

Rappelons que si d a un facteur premier $p \equiv 3 \bmod 4$, alors $\rho(d) = 0$. On déduit donc de (7) que si δ a un tel facteur premier, alors $y_\delta = 0$. De même, si δ a un facteur carré alors $y_\delta = 0$ (par définition des λ_d).

Lemme 3. *On passe des y_δ aux nombres λ_ℓ grâce à la relation* :

$$\lambda_\ell \frac{\rho(\ell)}{\ell} = \sum_{\substack{\delta \leqslant z \\ \ell \mid \delta}} \mu\left(\frac{\delta}{\ell}\right) y_\delta$$

Démonstration. Cette égalité repose essentiellement sur la formule d'inversion de Möbius. En effet, en remplaçant y_δ par son expression (7) et en permutant les deux sommes, il vient

$$\begin{aligned}\sum_{\substack{\delta\leqslant z\\ \ell\mid\delta}}\mu\left(\frac{\delta}{\ell}\right)y_\delta &= \sum_{\substack{\delta\leqslant z\\ \ell\mid\delta}}\mu\left(\frac{\delta}{\ell}\right)\sum_{\substack{d\leqslant z\\ \delta\mid d}}\lambda_d\frac{\rho(d)}{d}\\ &= \sum_{\substack{d\leqslant z\\ \ell\mid d}}\lambda_d\frac{\rho(d)}{d}\sum_{\substack{\delta\in\mathcal{D}(d)\\ \ell\mid\delta}}\mu\left(\frac{\delta}{\ell}\right)\end{aligned}$$

où la dernière somme porte sur les diviseurs δ de d qui sont divisibles par ℓ. Or

$$(\ell \mid \delta \text{ et } \delta \mid d) \iff (t = \delta/\ell \text{ est un entier divisant } d/\ell)$$

Ainsi, en faisant le changement d'indice $t = \delta/\ell$ dans la somme intérieure, on obtient

$$\sum_{\substack{\delta\in\mathcal{D}(d)\\ \ell\mid\delta}}\mu\left(\frac{\delta}{\ell}\right) = \sum_{t\mid(d/\ell)}\mu(t) = \delta_1\left(\frac{d}{\ell}\right) = \begin{cases} 1 & \text{si } d = \ell\\ 0 & \text{si } d \neq \ell\end{cases}$$

ce qui démontre le lemme. □

La relation obtenue dans le lemme 3 ne donnant aucune information sur λ_ℓ si ℓ a un facteur premier $p \equiv 3 \bmod 4$, nous poserons $\lambda_\ell = 0$ dans ce cas.

4.3.3. *Détermination de g*

Nous allons déterminer explicitement la fonction g. Comme nous travaillons uniquement avec des entiers sans facteurs carrés, nous allons expliciter g seulement pour ces entiers. Par multiplicativité, il suffit d'évaluer g sur les nombres premiers. Pour tout nombre premier p, on a

$$g(p) = (\mu \star f)(p) = \mu(1)f(p) + \mu(p)f(1) = f(p) - 1$$

En utilisant l'expression de la fonction ρ, on trouve donc :

$$g(p) = \begin{cases} 1 & \text{si } p = 2\\ \dfrac{p-2}{2} & \text{si } p \equiv 1 \bmod 4\\ -1 & \text{si } p \equiv 3 \bmod 4\end{cases}$$

4.3.4. *Minimisation de la forme quadratique $F(z)$*

La condition de minimisation de la forme quadratique $F(z)$ est $\lambda_1 = 1$. D'après le lemme 3, nous devons donc minimiser F sur l'hyperplan affine :

$$\sum_{\delta\leqslant z}\mu(\delta)y_\delta = 1 \tag{8}$$

Posons $\psi(z) = \sum_{\delta \leqslant z} \mu(\delta) y_\delta$. Le résultat suivant assure l'existence d'un minimum pour F sur $\psi^{-1}(\{1\})$ et va nous permettre dans la suite d'expliciter les coordonnées y_δ de ce minimum.

Proposition 9. *1. La fonction F admet un minimum sur l'hyperplan affine $\psi^{-1}(\{1\})$ noté $Y_0 = (y_\delta)_{\delta \leqslant z}$.*

2. Il existe un nombre réel ξ tel que la différentielle $DF(Y_0)$ de F au point Y_0 s'écrive $DF(Y_0) = \xi\psi$.

Démonstration. 1. Notons $\|\cdot\|$ la norme euclidienne sur $\mathbb{R}^{\lfloor z \rfloor}$. Comme la fonction g est à valeurs strictement positives sur les entiers n'ayant que des facteurs premiers $p \not\equiv 3 \bmod 4$, il est clair que

$$\lim_{\|Y\| \to +\infty} F(Y) = +\infty$$

Par conséquent, la fonction F admet un minimum global Y_0 sur $\mathbb{R}^{\lfloor z \rfloor}$. C'est donc aussi un minimum global de F sur $\psi^{-1}(\{1\})$.

2. La fonction F est différentiable sur $\mathbb{R}^{\lfloor z \rfloor}$ en tant que forme quadratique. Comme $DF(Y_0)$ et ψ sont des formes linéaires non nulles, il suffit de démontrer que Ker(ψ) est inclus dans Ker$(DF(Y_0))$. Soit $Y \in$ Ker(ψ). Pour tout $t > 0$, on a

$$\frac{F(Y_0 + tY) - F(Y_0)}{t} \leqslant 0$$

car Y_0 est le minimum (global) de F sur $\psi^{-1}(\{1\})$ et car $Y_0 + tY \in \psi^{-1}(\{1\})$. Pour des raisons analogues, on a aussi $\dfrac{F(Y_0 + tY) - F(Y_0)}{t} \geqslant 0$ pour tout $t < 0$. On déduit de cela que

$$DF(Y_0)(Y) = \lim_{t \to 0} \frac{F(Y_0 + tY) - F(Y_0)}{t} = 0$$

et donc $Y \in$ Ker$(DF(Y_0))$.

□

Remarque. Le deuxième point de la proposition 9 est une version faible du théorème des multiplicateurs de Lagrange.

Pour palier à une petite difficulté technique, nous noterons dans la suite h la fonction multiplicative définie par

$$\forall q \in \mathbb{N}^*, \qquad h(q) = \frac{(\mathbf{1} \star \chi)(q)}{\prod\limits_{\substack{p \mid q \\ p > 2}} (p-2)}$$

où la fonction χ est le caractère de Dirichlet non principal modulo 4 défini par

$$\forall q \in \mathbb{N}^*, \qquad \chi(q) = \begin{cases} 0 & \text{si } q \equiv 0 \mod 2 \\ 1 & \text{si } q \equiv 1 \mod 4 \\ -1 & \text{si } q \equiv 3 \mod 4 \end{cases}$$

La fonction h correspond à l'inverse de g sur les entiers sans facteurs premiers $p \equiv 3 \bmod 4$. Elle est nulle sinon (à cause du numérateur $\mathbf{1} \star \chi$) ; nous verrons que c'est la façon de traduire le fait que y_δ est nul si δ a un facteur premier $p \equiv 3 \bmod 4$. Notons encore G la fonction sommatoire de la fonction multiplicative $\mu^2 h$, c'est-à-dire définie par

$$G(z) = \sum_{q \leqslant z} \mu^2(q) h(q)$$

que nous aurons besoin d'estimer dans la suite. Nous donnons ci-dessous les expressions des coefficients y_δ (et λ_ℓ) minimisant la forme quadratique $F(z)$.

Corollaire 3. *Sous la condition* (8)*, la forme quadratique F est minimale pour*

$$y_\delta = \frac{\mu(\delta) h(\delta)}{G(z)}$$

Par conséquent,

$$\lambda_\ell = \mu(\ell) f(\ell) h(\ell) \frac{G_\ell(z)}{G(z)} \qquad \textit{où} \qquad G_\ell(z) = \sum_{\substack{q \leqslant (z/\ell) \\ (q,\ell)=1}} \mu^2(q) h(q)$$

Démonstration. On sait d'après la proposition 9 qu'il existe un nombre réel ξ tel que $DF(Y_0) = \xi D\psi(Y_0)$. Pour tout $\delta \leqslant z$ n'ayant pas de facteur premier $p \equiv 3 \bmod 4$, on a donc

$$\frac{\partial F}{\partial y_\delta}(Y_0) = \xi \frac{\partial \psi}{\partial y_\delta}(Y_0)$$

c'est-à-dire $2g(\delta) y_\delta = \xi \mu(\delta)$, soit encore

$$y_\delta = \frac{\xi}{2} \times \frac{\mu(\delta)}{g(\delta)} = \frac{\xi}{2} \times \mu(\delta) h(\delta)$$

cette égalité étant aussi valable si δ a un facteur premier $p \equiv 3 \bmod 4$ par définition de h. En remplaçant la valeur de y_δ dans la condition d'optimisation (8), nous obtenons la valeur de ξ :

$$\frac{2}{\xi} = \sum_{\delta \leqslant z} \mu^2(\delta) h(\delta) = G(z)$$

puis celle de y_δ annoncée. Le lemme 3 permet d'en déduire l'expression des coefficients λ_ℓ :

$$\lambda_\ell = f(\ell) \sum_{\substack{\delta \leqslant z \\ \ell \mid \delta}} \mu\left(\frac{\delta}{\ell}\right) y_\delta = \frac{f(\ell)}{G(z)} \sum_{\substack{\delta \leqslant z \\ \ell \mid \delta}} \mu\left(\frac{\delta}{\ell}\right) \mu(\delta) h(\delta)$$

Les multiples $\delta \leqslant z$ de ℓ sont de la forme $\delta = q\ell$ où l'entier q vérifie $1 \leqslant q \leqslant (z/\ell)$. Donc

$$\lambda_\ell = \frac{f(\ell)}{G(z)} \sum_{q \leqslant (z/\ell)} \mu(q)\mu(q\ell)h(q\ell)$$

Seuls les entiers q premiers à ℓ apportent une contribution non nulle dans la somme précédente (si $(q,\ell) > 1$, alors $q\ell$ a un facteur carré et $\mu(q\ell) = 0$). Si $(q,\ell) = 1$, alors la multiplicativité des fonctions arithmétiques μ et g nous donne

$$\lambda_\ell = \frac{\mu(\ell)f(\ell)h(\ell)}{G(z)} \sum_{\substack{q \leqslant (z/\ell) \\ (q,\ell)=1}} \mu^2(q)h(q) = \mu(\ell)f(\ell)h(\ell)\frac{G_\ell(z)}{G(z)}$$

d'où le résultat. □

4.4. *Estimation de la fonction sommatoire* G

D'après le corollaire 3, le terme principal $F(z)$ (obtenu sous forme diagonale en (7)) se réécrit :

$$F(z) = \sum_{\delta \leqslant z} g(\delta)\frac{\mu^2(\delta)h(\delta)^2}{G(z)^2} = \frac{1}{G(z)}$$

car $g(\delta)h(\delta) = 1$ si δ n'a pas de facteur premier $p \equiv 3 \bmod 4$ et par définition de $G(z)$. Pour démontrer le théorème 1, il faut disposer d'une estimation de $F(z)$, ce qui revient à estimer la fonction sommatoire G. Pour cela, nous allons utiliser la *méthode de convolution* qui va nous donner l'*ordre moyen* de la fonction multiplicative $\mu^2 h$ (proposition 10). Nous rappelons plus loin l'idée générale de la méthode de convolution ; pour plus d'informations à ce sujet, nous renvoyons le lecteur à [2].

4.4.1. *Intermède : deux petites estimations*

Nous aurons besoin des estimations préliminaires suivantes. C'est l'occasion d'illustrer dans la preuve du lemme 4 la méthode de *sommation par parties* très utilisée pour estimer la fonction sommatoire d'une fonction arithmétique.

Lemme 4. *Pour tout nombre réel $X \geqslant 1$, on a*

$$\sum_{n \leqslant X} \frac{1}{n} = \ln X + \mathcal{O}(1)$$

Démonstration. Soit $X \in [1, +\infty[$. Pour chaque entier $n \in [1, X]$, on peut écrire $\dfrac{1}{n} = \displaystyle\int_n^X \frac{\mathrm{d}t}{t^2} + \frac{1}{X}$. Donc

$$\sum_{n \leqslant X} \frac{1}{n} = \sum_{n \leqslant X} \int_n^X \frac{\mathrm{d}t}{t^2} + \frac{\sum_{n \leqslant X} 1}{X}$$

Nous voulons permuter les symboles $\displaystyle\sum_{n\leqslant X}$ et $\displaystyle\int_n^X$. Pour cela, on écrit l'intégrale comme :

$$\int_n^X \frac{\mathrm{d}t}{t^2} = \int_1^X \mathbf{1}_{[n,X]}(t)\frac{\mathrm{d}t}{t^2}$$

On a alors

$$\sum_{n\leqslant X}\int_n^X \frac{\mathrm{d}t}{t^2} = \int_1^X \left(\sum_{n\leqslant X}\mathbf{1}_{[n,X]}(t)\right)\frac{\mathrm{d}t}{t^2}$$

Pour tout nombre réel $t \in [1, X]$, la somme $\displaystyle\sum_{n\leqslant X}\mathbf{1}_{[n,X]}(t)$ est égale au nombre d'entiers $n \in [1, X]$ tels que $n \leqslant t$ (puisque pour chaque entier n, l'indicatrice $\mathbf{1}_{[n,X]}(t)$ vaut 1 si $n \leqslant t$ et 0 sinon). La permutation des symboles $\displaystyle\sum_{n\leqslant X}$ et $\displaystyle\int_n^X$ fournit donc :

$$\sum_{n\leqslant X}\int_n^X \frac{\mathrm{d}t}{t^2} = \int_1^X \left(\sum_{n\leqslant t} 1\right)\frac{\mathrm{d}t}{t^2} = \int_1^X \lfloor t\rfloor\frac{\mathrm{d}t}{t^2}$$

En utilisant l'estimation grossière $\lfloor t\rfloor = t + \mathcal{O}(1)$ (puisque $0 \leqslant t - \lfloor t\rfloor < 1$), il vient

$$\begin{aligned}\sum_{n\leqslant X}\frac{1}{n} &= \int_1^X (t+\mathcal{O}(1))\frac{\mathrm{d}t}{t^2} + \frac{X+\mathcal{O}(1)}{X}\\ &= \int_1^X \frac{\mathrm{d}t}{t} + 1 + \mathcal{O}\left(\int_1^X \frac{\mathrm{d}t}{t^2} + \frac{1}{X}\right)\\ &= \ln X + \mathcal{O}(1)\end{aligned}$$

d'où le lemme. □

Lemme 5. *Pour tout nombre réel $X \geqslant 1$, on a l'estimation*

$$\sum_{n\leqslant X}\frac{\chi(n)}{n} = \frac{\pi}{4} + \mathcal{O}\left(\frac{1}{X}\right)$$

Démonstration. Rappelons que le caractère de Dirichlet χ non principal modulo 4 est défini par $\chi(2k) = 0$ et $\chi(2k+1) = (-1)^k$ pour tout entier naturel k. Par conséquent,

$$\sum_{n\leqslant X}\frac{\chi(n)}{n} = \sum_{0\leqslant k\leqslant \frac{X-1}{2}}\frac{\chi(2k+1)}{2k+1} = \sum_{0\leqslant k\leqslant \frac{X-1}{2}}\frac{(-1)^k}{2k+1}$$

On a donc affaire aux sommes partielles d'une série alternée convergente. La somme de cette série vaut $\displaystyle\sum_{k=0}^{+\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4}$ (en considérant par exemple le développement en série entière de la fonction $\arctan$) et le reste vérifie

$$\left| \sum_{k> \frac{X-1}{2}} \frac{(-1)^k}{2k+1} \right| \leqslant \frac{1}{2 \times \frac{X-1}{2} + 1} = \frac{1}{X}$$

ce qui établit le lemme. □

4.4.2. *Retour aux séries de Dirichlet : estimation de $G(z)$*

La méthode de convolution va nous permettre de trouver l'ordre moyen de la fonction multiplicative $\mu^2 h$, c'est-à dire de trouver un développement asymptotique de sa fonction sommatoire G. Examinons de plus près la série de Dirichlet de cette fonction. Formellement,

$$D(\mu^2 h, z) = \prod_{p \geqslant 2} \left(1 + \frac{h(p)}{p^s}\right) = \left(1 + \frac{1}{2^s}\right) \prod_{p \geqslant 3} \left(1 + \frac{(\mathbf{1} \star \chi)(p)}{(p-2)p^s}\right)$$

Cette série de Dirichlet converge absolument sur $]0, +\infty[$ (ceci est une conséquence de la proposition 7 et du fait que la série $\displaystyle\sum_{p \geqslant 2} \frac{1}{p^{s+1}}$ converge si et seulement si $s > 0$). L'idée de la méthode de convolution consiste à *comparer* cette série une série de Dirichlet D qui lui est *proche* et dont on dispose d'informations plus précises (notamment sur l'ordre moyen de la fonction multiplicative associée). Ceci aura pour effet d'augmenter l'abscisse de convergence absolue de la série de Dirichlet quotient $D(\mu^2 h, s)/D(s)$.

La complexité de notre fonction multiplicative $\mu^2 h$ réside dans la présence du facteur μ^2 qui consiste à ne prendre en compte que les entiers sans facteurs carrés. La convolution va notamment nous permettre de nous « débarasser » de ce facteur.

Pour tout nombre premier $p \geqslant 3$,

$$1 + \frac{(\mathbf{1} \star \chi)(p)}{(p-2)p^s} \quad \text{est } \textit{de l'ordre de} \quad 1 + \frac{(\mathbf{1} \star \chi)(p)}{p \times p^s} + \sum_{k \geqslant 2} \frac{(\mathbf{1} \star \chi)(p^k)}{p^k \times p^{ks}} = \sum_{k=0}^{+\infty} \frac{\left(\frac{\mathbf{1} \star \chi}{\mathrm{Id}}\right)(p^k)}{p^{ks}}$$

d'où l'idée de comparer la série de Dirichlet de $\mu^2 h$ avec celle de la fonction multiplicative $\dfrac{\mathbf{1} \star \chi}{\mathrm{Id}}$.

Notons θ la fonction multiplicative définie par $\mu^2 h = \left(\dfrac{\mathbf{1} \star \chi}{\mathrm{Id}}\right) \star \theta$. Pour tout nombre premier $p \geqslant 3$, on a

$$\sum_{k=0}^{+\infty} \frac{(\mathbf{1} \star \chi)(p^k)}{p^k \times p^{ks}} = 1 + \frac{(\mathbf{1} \star \chi)(p)}{p \times p^s} + \mathcal{O}\left(\frac{1}{p^{2s+2}}\right)$$

où la constante intervenant dans le $\mathcal{O}$ ne dépend pas du nombre premier p. En effet, si $p \equiv 3 \bmod 4$, alors la suite $\big((\mathbf{1}\star\chi)(p^k)\big)_{k\in\mathbb{N}}$ est bornée par 1 et, en utilisant la formule donnant la somme d'une série géométrique convergente,

$$\sum_{k=2}^{+\infty}\frac{(\mathbf{1}\star\chi)(p^k)}{p^k\times p^{ks}}\leqslant\sum_{k=2}^{+\infty}\frac{1}{p^{k(s+1)}}\quad\text{donc}\quad\sum_{k=2}^{+\infty}\frac{(\mathbf{1}\star\chi)(p^k)}{p^k\times p^{ks}}\leqslant\frac{1}{p^{2s+2}}\times\frac{p^{s+1}}{p^{s+1}-1}\leqslant\frac{3}{p^{2s+2}}$$

cette dernière inégalité étant valable pour $s\geqslant -1/2$. Si maintenant $p\equiv 1\bmod 4$ (donc en particulier $p\geqslant 5$), alors pour tout entier naturel k, on a $(\mathbf{1}\star\chi)(p^k)=k+1$ et donc, en utilisant la formule donnant la somme d'une série géométrique dérivée, on a

$$\sum_{k=2}^{+\infty}\frac{(\mathbf{1}\star\chi)(p^k)}{p^k\times p^{ks}}=\frac{p^{2(s+1)}}{(p^{s+1}-1)^2}-1-\frac{2}{p^{s+1}}=\frac{3p^{s+1}-2}{(p^{s+1}-1)^2p^{s+1}}\leqslant\frac{4}{p^{2s+2}}$$

pour $s\geqslant -1/2$. En utilisant la proposition 8 donnant la série de Dirichlet d'une fonction multiplicative qui s'écrit comme un produit de convolution puis en se servant de l'estimation $\dfrac{1}{1+u}=1-u+\mathcal{O}(u^2)$ valable au voisinage de 0, on trouve que la série de Dirichlet de θ s'écrit :

$$\begin{aligned}
D(\theta,s)&=\frac{D(\mu^2h,s)}{D\big(\frac{\mathbf{1}\star\chi}{\mathrm{Id}},s\big)}\\
&=\left(1+\frac{1}{2^s}\right)\left(1-\frac{1}{2^s}\right)\prod_{p\geqslant 3}\left(1+\frac{(\mathbf{1}\star\chi)(p)}{(p-2)p^s}\right)\left(1-\frac{(\mathbf{1}\star\chi)(p)}{p\times p^s}+\mathcal{O}\left(\frac{1}{p^{2s+2}}\right)\right)\\
&=\left(1-\frac{1}{4^s}\right)\prod_{p\geqslant 3}\left(1+2\frac{(\mathbf{1}\star\chi)(p)}{(p-2)p^{s+1}}-\frac{(\mathbf{1}\star\chi)(p)^2}{(p-2)p^{2s+1}}+\mathcal{O}\left(\frac{1}{p^{2s+2}}\right)\right)\\
&=\left(1-\frac{1}{4^s}\right)\prod_{p\geqslant 3}\left(1+2\frac{(\mathbf{1}\star\chi)(p)}{(p-2)p^{s+1}}+\mathcal{O}\left(\frac{1}{p^{2s+2}}\right)\right)
\end{aligned}$$

Donc la série $D(\theta,s)$ converge absolument si et seulement si la série

$$\sum_{p\geqslant 3}\left(2\frac{(\mathbf{1}\star\chi)(p)}{(p-2)p^{s+1}}+\mathcal{O}\left(\frac{1}{p^{2s+2}}\right)\right)$$

est absolument convergente (d'après la proposition 7), c'est-à-dire si et seulement si $2s+2>1$ soit pour $s\in]-1/2,+\infty[$.

Le développement asymptotique de la fonction sommatoire G fait l'objet du résultat suivant.

Proposition 10. *Il existe une constante $C>0$ telle que, pour tout nombre réel $z\geqslant 1$, on ait l'estimation*

$$G(z)=C\ln z+\mathcal{O}(1)$$

Démonstration. Par définition de θ, on a

$$G(z) = \sum_{q \leqslant z} \mu^2(q) h(q) = \sum_{q \leqslant z} \left(\frac{\mathbf{1} \star \chi}{\mathrm{Id}} \star \theta \right)(q) = \sum_{abc \leqslant z} \frac{\chi(b)\theta(c)}{ab}$$

par définition du produit de convolution et en écrivant $q = abc$ comme le produit de trois entiers. Nous allons maintenant utiliser les estimations obtenues dans les lemmes 4 et 5 :

$$\begin{aligned} G(z) = \sum_{c \leqslant z} \theta(c) \sum_{ab \leqslant \frac{z}{c}} \frac{\chi(b)}{ab} &= \sum_{c \leqslant z} \theta(c) \sum_{a \leqslant \frac{z}{c}} \frac{1}{a} \sum_{b \leqslant \frac{z}{ac}} \frac{\chi(b)}{b} \\ &= \sum_{c \geqslant 1} \theta(c) \sum_{a \leqslant \frac{z}{c}} \frac{1}{a} \sum_{b \leqslant \frac{z}{ac}} \frac{\chi(b)}{b} \end{aligned}$$

où la dernière égalité provient du fait que si $c > z$, alors le sommant $\displaystyle\sum_{a \leqslant \frac{z}{c}} \frac{1}{a} \sum_{b \leqslant \frac{z}{ac}} \frac{\chi(b)}{b}$ est nul. Nous pouvons appliquer le lemme 5 à la somme intérieure, ce qui donne :

$$\begin{aligned} G(z) &= \sum_{c \geqslant 1} \theta(c) \sum_{a \leqslant \frac{z}{c}} \frac{1}{a} \left[\frac{\pi}{4} + \mathcal{O}\left(\frac{ac}{z} \right) \right] \\ &= \frac{\pi}{4} \sum_{c \geqslant 1} \theta(c) \sum_{a \leqslant \frac{z}{c}} \frac{1}{a} + \mathcal{O}\left(\sum_{c \geqslant 1} |\theta(c)| \sum_{a \leqslant \frac{z}{c}} \frac{1}{a} \times \frac{ac}{z} \right) \end{aligned}$$

Or le reste dans le $\mathcal{O}$ est inférieur ou égal $\displaystyle\sum_{c \geqslant 1} |\theta(c)|$, qui est la somme d'une série convergente car la série de Dirichlet associée à la fonction θ converge absolument sur l'intervalle $]-1/2, +\infty[$ donc en particulier en 0. Nous avons donc l'estimation :

$$G(z) = \frac{\pi}{4} \sum_{c \geqslant 1} \theta(c) \sum_{a \leqslant \frac{z}{c}} \frac{1}{a} + \mathcal{O}(1)$$

Grâce à l'estimation obtenue dans le lemme 4, il vient :

$$\begin{aligned} G(z) &= \frac{\pi}{4} \sum_{c \geqslant 1} \theta(c) \big[\ln z + \mathcal{O}(1) \big] + \mathcal{O}(1) \\ &= \left(\frac{\pi}{4} \sum_{c \geqslant 1} \theta(c) \right) \ln z + \mathcal{O}\left(1 + \sum_{c \leqslant z} |\theta(c)| \right) \end{aligned}$$

et comme la série $\displaystyle\sum_{c \geqslant 1} |\theta(c)|$ est convergente, la proposition est démontrée avec la constante $C = \dfrac{\pi}{4} \displaystyle\sum_{c \geqslant 1} \theta(c) > 0$. □

On en déduit donc une estimation du terme principal de la somme initiale $S(X,z)$:

$$\sqrt{X}F(z) = \frac{\sqrt{X}}{G(z)} = \frac{\sqrt{X}}{C\ln(z)+\mathcal{O}(1)} = C^{-1}\frac{\sqrt{X}}{\ln(z)} + \mathcal{O}\left(\frac{\sqrt{X}}{\ln(z)^2}\right) \tag{9}$$

4.5. *Estimation du terme d'erreur*

Nous cherchons maintenant à estimer le terme d'erreur que nous avons noté $E(z)$:

$$E(z) = \mathcal{O}\left(\sum_{\substack{d_1\leqslant z\\ d_2\leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1,d_2])\right)$$

Nous commençons par montrer que la suite de poids $(\lambda_d)_{d\in\mathbb{N}^*}$ est bornée. Rappelons que nous avons obtenu au corollaire 3 que

$$\forall d\in\mathbb{N}^*, \qquad \lambda_d = \mu(d)f(d)h(d)\frac{G_d(z)}{G(z)}$$

Proposition 11. *Pour tout entier naturel d non nul, on a $|\lambda_d|\leqslant 1$.*

Démonstration. Si d a un facteur carré ou si d a un facteur premier $p\equiv 3 \mod 4$, alors $\lambda_d = 0$ et donc l'inégalité est claire. Nous supposons donc dans la suite que $\mu^2(d)=1$ et que d n'a pas de facteur premier $p\equiv 3 \mod 4$. L'idée de la démonstration est de comparer les fonctions G et G_d. On commence par regrouper les entiers $q\leqslant z$ suivant les différentes valeurs du plus grand commun diviseur (q,d) de q et d :

$$G(z) = \sum_{q\leqslant z}\mu^2(q)h(q) = \sum_{t|d}\sum_{\substack{q\leqslant z\\ (q,d)=t}}\mu^2(q)h(q)$$

Un entier q tel que $(q,d)=t$ s'écrit sous la forme $q=t\ell$ où $(\ell,d)=1$ (et $(t,\ell)=1$ car d est sans facteur carré). La multiplicativité de la fonction $\mu^2 h$ nous donne $\mu^2(t\ell)h(t\ell) = \big(\mu^2(t)h(t)\big)\big(\mu^2(\ell)h(\ell)\big)$. En faisant le changement de variable $q=t\ell$ dans la somme intérieure, on obtient donc

$$G(z) = \sum_{t|d}\mu^2(t)h(t)\sum_{\substack{\ell\leqslant(z/t)\\ (\ell,d)=1}}\mu^2(\ell)h(\ell)$$

Or $\dfrac{z}{t}\geqslant\dfrac{z}{d}$ (car $t\leqslant d$) donc, comme tous les sommants sont positifs ou nuls,

$$G(z)\geqslant\left(\sum_{t|d}\mu^2(t)h(t)\right)\left(\sum_{\substack{\ell\leqslant(z/d)\\ (\ell,d)=1}}\mu^2(\ell)h(\ell)\right)$$

On simplifie maintenant la somme entre parenthèses en utilisant à nouveau la multiplicativité de $\mu^2 h$:

$$\sum_{t|d} \mu^2(t)h(t) = \prod_{p|d} \big(1 + \mu^2(p)h(p)\big) = \prod_{p|d} \left(\frac{g(p)+1}{g(p)}\right)$$

puisque h correspond à l'inverse de g sur les entiers sans facteurs premiers $p \equiv 3 \mod 4$. Par définition de la fonction g (qui vérifie la relation $f = \mathbf{1} \star g$), on a $g(p) + 1 = f(p)$ pour tout diviseur p de d. Par multiplicativité des deux fonctions f et g, on obtient $\displaystyle\sum_{t|d} \frac{\mu^2(t)}{g(t)} = \frac{f(d)}{g(d)}$.

On a finalement l'inégalité :

$$G(z) \geqslant \frac{f(d)}{g(d)} G_d(z) \tag{10}$$

La majoration $|\lambda_d| \leqslant 1$ provient de l'inégalité (10) et de l'expression de λ_d rappelée avant la proposition. □

Nous devons ici estimer le reste $E(z) = \mathcal{O}\left(\displaystyle\sum_{\substack{d_1 \leqslant z \\ d_2 \leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1,d_2])\right)$.

Pour tout couple d'entiers $(d_1, d_2) \in (\mathbb{N}^*)^2$ n'ayant ni facteur carré ni facteur premier $p \equiv 3 \mod 4$, on a $\rho((d_1,d_2)) \geqslant 1$ donc, par multiplicativité de ρ (et d'après la proposition 4), on a

$$\rho([d_1,d_2]) = \frac{\rho(d_1)\rho(d_2)}{\rho((d_1,d_2))} \leqslant \rho(d_1)\rho(d_2)$$

On déduit de cette majoration et de la proposition 11 les inégalités

$$\sum_{\substack{d_1 \leqslant z \\ d_2 \leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1,d_2]) \leqslant \left(\sum_{d \leqslant z} |\lambda_d|\,\rho(d)\right)^2 \leqslant \left(\sum_{d \leqslant z} \mu^2(d)\rho(d)\right)^2$$

Si p est un nombre premier, alors il est clair que $\rho(p) \leqslant \tau(p)$ (où τ est la fonction nombre de diviseurs). Par multiplicativité et positivité des fonctions ρ et τ, on obtient la majoration $\rho(d) \leqslant \tau(d)$ pour tout entier naturel d sans facteur carré. Nous obtenons donc la majoration :

$$\sum_{\substack{d_1 \leqslant z \\ d_2 \leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1,d_2]) \leqslant \left(\sum_{d \leqslant z} \tau(d)\right)^2$$

où la dernière somme est étendue à tous les entiers inférieurs ou égaux à z. Il est nécessaire à ce stade de connaître un ordre de grandeur de la fonction diviseur τ.

Lemme 6. *Pour tout $\varepsilon > 0$, on a $\tau(d) \ll_\varepsilon d^\varepsilon$ (la notation $\ll_\varepsilon$ signifiant que la constante dépend de ε).*

Démonstration. Soit $\varepsilon > 0$. Il existe un nombre premier p_0 tel que $p_0^{\varepsilon} \geqslant 2$. Pour tout entier naturel α, nous avons donc $\tau(p_0^{\alpha}) = \alpha + 1 \leqslant 2^{\alpha} \leqslant p_0^{\varepsilon\alpha}$. Si p est un nombre premier tel que $p \geqslant p_0$, alors on a aussi $\tau(p^{\alpha}) \leqslant p^{\varepsilon\alpha}$ et si $p < p_0$, alors il existe une constante $C_p > 0$ telle que pour tout entier naturel α, on ait $\alpha + 1 \leqslant C_p p^{\varepsilon\alpha}$. Par multiplicativité de la fonction τ, on obtient :

$$\tau(d) = \tau\left(\prod_{p^{\alpha}\| d} p^{\alpha}\right) = \left(\prod_{\substack{p^{\alpha}\| d \\ p<p_0}} \tau(p^{\alpha})\right)\left(\prod_{\substack{p^{\alpha}\| d \\ p\geqslant p_0}} \tau(p^{\alpha})\right) \leqslant \left(\prod_{p<p_0} C_p\right) \prod_{p^{\alpha}\| d} p^{\varepsilon\alpha}$$

où la notation $p^{\alpha} \parallel d$ signifie que p^{α} divise d et $p^{\alpha+1}$ ne divise pas d. En notant C_{ε} la constante $\displaystyle\prod_{p<p_0} C_p$, on obtient $\tau(d) \leqslant C_{\varepsilon} d^{\varepsilon}$. □

On déduit du lemme 6 que pour tout $\varepsilon > 0$,

$$E(z) \ll \sum_{\substack{d_1 \leqslant z \\ d_2 \leqslant z}} |\lambda_{d_1}|\,|\lambda_{d_2}|\,\rho([d_1, d_2]) \ll_{\varepsilon} \left(\sum_{d\leqslant z} d^{\varepsilon}\right)^2 \ll_{\varepsilon} z^{2+\varepsilon} \tag{11}$$

4.6. *Choix du paramètre z et fin de la démonstration du théorème 1*

On déduit des estimations (9) et (11) précédemment obtenues que

$$S(X, z) = \sqrt{X} F(z) + E(z) \ll_{\varepsilon} \frac{\sqrt{X}}{\ln(z)} + \frac{\sqrt{X}}{\ln(z)^2} + z^{2+\varepsilon}$$

Pour le choix du paramètre $z = X^{1/5}$, le terme reste $z^{2+\varepsilon}$ est négligeable (pour ε petit) devant le terme principal $\dfrac{\sqrt{X}}{\ln(X)}$. Ceci fournit alors l'estimation $S(X, z) \ll \dfrac{\sqrt{X}}{\ln(X)}$, ce qui termine la démonstration du théorème 1.

L'estimation établie dans le théorème 1 permet d'obtenir une information non triviale sur les nombres premiers quadratiques. Nous allons en effet montrer que ces nombres premiers sont *rares* au sens où la série $\displaystyle\sum_{\substack{n\geqslant 2 \\ n^2+1 \text{ premier}}} \frac{1}{n\ln(n)}$ est convergente (alors que la série de Bertrand $\displaystyle\sum_{n\geqslant 2} \frac{1}{n\ln(n)}$ diverge).

Corollaire 4. *Pour tout $\varepsilon > 0$, la série $\displaystyle\sum_{\substack{n\geqslant 2 \\ n^2+1 \text{ premier}}} \frac{1}{n(\ln n)^{\varepsilon}}$ est convergente.*

Démonstration. Soit $\varepsilon > 0$. Pour tout nombre réel $X \geqslant 2$, on a

$$\sum_{\substack{2\leqslant n\leqslant X\\ n^2+1 \text{ premier}}} \frac{1}{n(\ln n)^{\varepsilon}} = \sum_{\substack{2\leqslant n\leqslant X\\ n^2+1 \text{ premier}}} \left(\int_n^X \frac{\varepsilon + \ln t}{t^2 \ln(t)^{\varepsilon+1}} \,\mathrm{d}t + \frac{1}{X(\ln X)^{\varepsilon}} \right)$$

$$= \int_2^X \left(\sum_{\substack{2\leqslant n\leqslant t\\ n^2+1 \text{ premier}}} 1 \right) \frac{\varepsilon + \ln t}{t^2 \ln(t)^{1+\varepsilon}} \,\mathrm{d}t + \frac{1}{X(\ln X)^{\varepsilon}} \sum_{\substack{2\leqslant n\leqslant X\\ n^2+1 \text{ premier}}} 1$$

En utilisant l'estimation établie, il vient

$$\sum_{\substack{2\leqslant n\leqslant X\\ n^2+1 \text{ premier}}} \frac{1}{n(\ln n)^{\varepsilon}} \ll \int_2^X \frac{t}{\ln t} \times \frac{\varepsilon + \ln t}{t^2 \ln(t)^{1+\varepsilon}} \,\mathrm{d}t + \frac{1}{X(\ln X)^{\varepsilon}} \times \frac{X}{\ln X}$$

$$\ll \int_2^X \frac{\mathrm{d}t}{t \ln(t)^{1+\varepsilon}} + \frac{1}{\ln(X)^{1+\varepsilon}}$$

$$\ll 1$$

car l'intégrale $\displaystyle\int_2^{+\infty} \frac{\mathrm{d}t}{t \ln(t)^{1+\varepsilon}}$ est une intégrale de Bertrand convergente (puisque $\varepsilon > 0$).□

Références

[1] P.T. BATEMAN ET H.G. DIAMOND, *Analytic number theory,* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2004, An introductory course

[2] P. BERMENT ET O. RAMARÉ, *Estimation de l'ordre moyen d'une fonction arithmétique par la méthode de convolution,* RMS, Volume 122-1, 2011

[3] E. BOMBIERI, *Le grand crible dans la théorie analytique des nombres* Astérisque 18, 1987

[4] H.H. HALBERSTAM ET H.E. RICHERT, *Sieve methods,* Academic Press, London-New York, 1974

[5] H. IWANIEC, *Almost-primes represented by quadratic polynomials,* Invent. Math., 1978

[6] G. TENENBAUM, *Introduction à la théorie analytique et probabiliste des nombres,* Société Mathématique de France, Paris, seconde édition, 1995