Emmanuel Kowalski

Rapport de stage

La fonction de Beurling et le grand crible
Applications

Responsables **Mr. Hédi Daboussi**
Mr. Etienne Fouvry

"Crois bien, dit Don Quichotte, que les grains
de ce blé criblés par ses mains
étaient grains de perles"

Cervantès, "Don Quichotte", 1ère partie, chapitre XXXI

Sommaire

Introduction

Ce stage s'est déroulé au sein du département de mathématiques de l'université de Paris Sud, à Orsay, du 15 juin au 15 juillet et du 1er septembre au 15 septembre. Les responsables en étaient Mr. Fouvry et Mr. Daboussi. L'objectif du stage était l'étude du principe et des applications arithmétiques élémentaires de la méthode dite du "grand crible", avec de surcroît un détour par certaines techniques d'analyse de Fourier qui permettent d'obtenir de façon élégante des versions précises des résultats fondamentaux. Mr. Daboussi m'offrait de choisir comme sujet principal soit la partie d'analyse harmonique, avec pour support essentiel l'article de Vaaler [**Val**], soit le grand crible lui-même, l'analyse harmonique étant considérée comme outil de démonstration plutôt que comme but en soi; dans ce cas, la référence principale était l'article de Montgomery [**Mon**]. C'est plutôt cette approche que j'ai choisie, mais j'ai néanmoins développé la partie d'analyse harmonique au delà du minimum nécessaire pour l'arithmétique, en raison de son intérêt intrinsèque.

Ce rapport comporte quatre parties. La première est consacrée à l'exposition des notations employées et à l'énoncé des théorèmes principaux utilisés dans la suite, d'une part en analyse harmonique et complexe, d'autre part en théorie des nombres. D'autres théorèmes, moins importants, seront simplement cités lorsqu'ils seront employés. La seconde partie est celle d'analyse de Fourier : après l'étude d'un certain problème extrémal, elle se termine par un premier exemple d'application. La troisième partie développe le principe du grand crible à l'aide d'une partie des résultats de la seconde partie. Enfin, la dernière partie expose trois applications relativement élémentaires (mais importantes) du grand crible, et s'achève par quelques indications sur des applications moins immédiates, où le grand crible sert d'outil au côté d'autres méthodes.

Bien que fortement axé sur l'arithmétique, ce stage n'en a pas moins été l'occasion pour moi d'aborder d'autres domaines des mathématiques. J'ai mentionné déjà l'analyse harmonique et l'analyse complexe, mais on verra aussi apparaître un peu d'analyse fonctionnelle et de distributions au coin de certaines démonstrations.

I. Rappels et notations

1. Analyse harmonique et analyse complexe

Pour définir séries et transformées de Fourier, au lieu de la fonction exp, il s'avère plus simple d'utiliser la fonction e définie par $e(z)=e^{2i\pi z}$, de sorte que l'on a les formules suivantes :

• *La série de Fourier d'une fonction f périodique de période 1 est donnée par* $\sum c_n e(nz)$, *avec* $c_n=\int_0^1 f(x)e(-nx)dx$

• *La transformée de Fourier d'une fonction* $f\in L^1$ *est définie par*

(1.1) $$\hat{f}(t)=\int f(x)e(-xt)dx$$

d'où l'on tire la formule d'inversion (dans les cas de validité habituels)

(1.2) $$f(x)=\int \hat{f}(t)e(xt)dt$$

et la formule de Plancherel

(1.3) $$\int f\,\bar{g} = \int \hat{f}\bar{\hat{g}}$$

Théorème (formule de Poisson): *Soit* $f\in L^1$, *normalisée* (ie pour tout x $f(x)=\frac{1}{2}[f(x+)+f(x-)]$), *à variation bornée. Alors on a pour tout réel t :*

(1.4) $$\sum_{n=-\infty}^{+\infty} f(n+t) = \sum_{n=-\infty}^{+\infty} \hat{f}(n)e(tn)$$

Indication sur la démonstration

Le membre de droite est la série de Fourier de celui de gauche, et les hypothèses permettent d'assurer la convergence de ces deux séries.

Définition: *Une fonction entière f est dite de* ***type exponentiel*** σ *si*

$$|f(z)|=O(e^{(\sigma+\varepsilon)|z|})$$

pour tout $\varepsilon>0$. *La borne inférieure des* σ *vérifiant cette propriété est l'****ordre*** *de la fonction f.*

On montre que si f est de type exponentiel σ, *f' l'est également.*

L'intérêt des fonctions de type exponentiel en analyse harmonique est du en grande partie au théorème important suivant.

Théorème (Paley - Wiener): *Pour une fonction f, les conditions suivantes sont équivalentes :*

(i) *f est dans* L^2 *et est la restriction sur* $\mathbb{R}$ *d'une fonction entière de type exponentiel* σ, *avec* $\sigma>0$

(ii) *il existe une fonction* $g\in L^2([-2\pi/\sigma,2\pi/\sigma])$ *telle que, pour tout z, on ait:*

$$f(z)=\int_{-2\pi/\sigma}^{2\pi/\sigma} g(t)e(zt)dt$$

N.B. (i) La fonction g de (ii) n'est bien sur autre que la transformée de Fourier de f, qui est donc à support compact.

(ii) Dans de nombreuses applications, on peut se ramener par un changement de variable de la forme u=at à un ordre σ donné. Pour des raisons de notation, on choisit souvent le cas $\sigma=2\pi$, ie g supportée sur [-1,1].

Définition: *On note* E^p *l'ensemble des fonctions entières de type exponentiel* ***au plus*** 2π *dont la restriction à* $\mathbb{R}$ *est dans* L^p.

Propriétés (Polya - Plancherel, 1938) :

(1.5) • Si $p\le q$, alors $E^p\subset E^q$

• Si $f\in E^p$, alors $f'\in E^p$, et même:

(1.6) il existe A tel que pour toute f, $\int|f'|^p\le A\int|f|^p$

(1.7) • Il existe B tel que si $f\in E^p$, alors $\sum_{n=-\infty}^{\infty}|f(n)|^p\le B\int|f|^p<+\infty$

N.B. Dans les deux dernières formules, le fait de se restreindre (via E^p) à des fonctions d'ordre au plus 2π est important: les constantes A et B dépendent de façon croissante de l'ordre de f. En fait, nous n'utiliserons que E^1 et E^2, et alors on peut souvent vérifier plus simplement ces résultats.

Quelque identités

Les formules suivantes seront souvent utilisées :

(1.8) $$\left\{\frac{\sin \pi z}{\pi z}\right\}^2=\int_{-1}^{1}(1-|t|)e(tz)dt$$

N.B. On vérifie là sur un cas particulier le théorème de Paley - Wiener.

(1.9) $$z\left\{\frac{\sin \pi z}{\pi z}\right\}^2=\frac{1}{2i\pi}\int_{-1}^{1}\mathrm{sgn}(t)e(tz)dt$$

N.B. La fonction 'sgn' est définie sur $\mathbb{R}$ par $\mathrm{sgn}(t)=\frac{|t|}{t}$ si $t\neq 0$ et $\mathrm{sgn}(0)=0$. C'est donc une fonction normalisée impaire.

(1.10) $$\sum_{n=-\infty}^{\infty}(z-n)^{-2}=\left\{\frac{\pi}{\sin \pi z}\right\}^2$$

N.B Cette formule *a priori* plus délicate que les précédentes résulte simplement de l'application de la formule de Poisson (1.4) à la fonction f définie par $f(z)=\left\{\frac{\sin \pi z}{\pi z}\right\}^2$, en observant que (1.8) et la formule d'inversion (1.2) montrent que $\hat{f}(n)=0$ pour $|n|\geq 1$, ainsi (1.4) donne (1.10) sur $\mathbb{R}$, et on l'étend par prolongement analytique sur tout $\mathbb{C}$.

2. Théorie des nombres

On note (a,b) le pgcd de a et b, et [x] la partie entière d'un réel x.

Dans toute la partie arithmétique, la lettre 'p' pour un nombre ou une variable désignera **exclusivement** un nombre premier. En particulier, dans les sommations ou produits, une variable de sommation 'p' indiquera que l'on somme sur les nombres premiers (vérifiant telle ou telle condition additionnelle).

On utilisera la notation de Vinogradov $f\ll g$, équivalente à $f=O(g)$. La dépendance de la constante A implicite du "O" vis-à-vis de certains paramètres possibles sera indiquée en indice ou dans le texte: par exemple, $\ll_p$ signifie que la constante varie suivant la valeur de p.

On écrit $d|n^{\infty}$ - notation empruntée au cours de Tennenbaum de l'université de Nancy - pour dire que $\{p \mid p|d\} \subset \{p \mid p|n\}$. Ainsi, $25|15^{\infty}$.

Définition: *On appelle fonction arithmétique toute application d'une partie de* $\mathbb{Z}$ *dans* $\mathbb{C}$. *Une fonction arithmétique f définie sur* $\mathbb{N}^*$ *sera dite* ***multiplicative*** *si elle vérifie* $f(ab)=f(a)f(b)$ *pour* $(a,b)=1$, *et* ***totalement multiplicative*** *si* $f(ab)=f(a)f(b)$ *pour toutes les valeurs de a et b.*

N.B. Une fonction multiplicative est entièrement déterminée par ses valeurs en p^n pour tout les p et tout les n, aussi définit-on souvent une telle fonction en donnant simplement ces valeurs, qui peuvent alors être arbitraires. De plus, on a $f(1)=1$ si $f\neq 0$.

On utilisera la fonction φ, indicateur d' Euler, et en particulier la formule

(1.11) $$\varphi(n)=n\prod_{p|n}(1-p^{-1})$$

Définition (fonction de Möbius): *On note* μ *la fonction arithmétique multiplicative définie par*

$$\text{(i)}\quad \mu(1)=1$$
$$\text{(ii)}\quad \mu(p)=-1$$
$$\text{(iii)}\quad \mu(p^n)=0 \text{ si } n\geq 2$$

et par multiplicativité pour les autres valeurs de n.

Il est immédiat que $|\mu|=\mu^2$ est la fonction caractéristique de l'ensemble des entiers sans facteur carré. Il sera utile de disposer du nombre de tels entiers inférieurs à x, qui est donné assez précisément par

(1.12)
$$\sum_{n\leq x} \mu^2(n) = \frac{6}{\pi^2}\, x+O(\sqrt{x})$$

(cf. **[H&W]** chap. XVIII)

On utilisera une autre fonction arithmétique, la fonction ν, définie par

$$\nu(n)=\text{Card } \{p,\ p|n\} \quad \text{(nombre de diviseurs premiers)}$$

Cette fonction n'est pas multiplicative mais additive, ie $f(ab)=f(a)+f(b)$, si $(a,b)=1$. Par conséquent, 2^{ν} est multiplicative.

Enfin, on notera $\pi(x)=\text{Card } \{p\leq x\}$, le nombre de nombre premiers $\leq x$, et plus généralement $\pi(x;a,b)=\text{Card } \{p\leq x \mid p\equiv b \pmod a\}$. Le théorème de la progression arithmétique de Dirichlet dit que $\pi(x;a,b)\to\infty$ pour $x\to\infty$, si $(a,b)=1$.

Définition: *On appelle* ***caractère de Dirichlet*** *modulo k toute fonction arithmétique* $\chi: \mathbb{Z}\to\mathbb{C}$*, définie par*

$$\chi(n)=\begin{cases}\chi_1(n) \text{ si } (n,k)=1\\ 0 \text{ sinon}\end{cases}$$

où χ_1 *est un caractère du groupe* $(\mathbb{Z}/k\mathbb{Z})^*$ *des entiers inversibles modulo k.*

Un caractère de Dirichlet est complètement multiplicatif et périodique de période k. On note χ_0 *le caractère de Dirichlet correspondant à l'identité, et* $\overline{\chi}$ *l'inverse de* χ *, il est facile de vérifier que* $\overline{\chi}(n)=\overline{\chi(n)}$ (conjugué dans $\mathbb{C}$).

Définition: *Un caractère de Dirichlet* χ *modulo k est dit i****nduit par le caractère*** χ_1 *modulo d, où* d|k*, si on a*

$$\chi(n)=\begin{cases}\chi_1(n) \text{ si } (n,k)=1\\ 0 \text{ sinon}\end{cases}$$

Un caractère χ *est dit* ***primitif*** *s'il n'est induit par aucun caractère modulo d avec* $d<k$.

N.B. Dans de nombreux cas, les résultats relatifs aux caractères de Dirichlet ne prennent une forme simple que dans le cas où le caractère est primitif. Le passage au cas général fait généralement appel au **conducteur** f du caractère χ, qui est le plus grand entier d tel que χ est induit par un caractère primitif modulo d.

Propriétés:

- On a d'abord l'application des formules d'orthogonalité générales pour les caractères d'un groupe fini abélien

$$\sum_{n=1}^{k}\chi(n) = \begin{cases}\varphi(k) \text{ si } \chi=\chi_0 \\ 0 \text{ sinon}\end{cases}$$

(1.13)

$$\sum_{\chi}\chi(n) = \begin{cases}\varphi(k) \text{ si } n\equiv 1 \pmod k \\ 0 \text{ sinon}\end{cases}$$

la sommation étant sur tout les caractères modulo k

- Si χ est primitif, la **somme de Gauss** $\tau(\chi)=\sum_{a=0}^{k-1}\chi(a)e(a/k)$ vérifie pour tout n

(1.14)
$$\tau(\overline{\chi})\chi(n)=\sum_{a=0}^{k-1}\overline{\chi}(a)e(an/k)$$

et on a de plus

(1.15)
$$|\tau(\chi)|=\sqrt{k}$$

Nous utiliserons la notation $\sum_{\chi}'$ pour noter une somme étendue aux seuls caractères primitifs.

Formule de sommation d'Abel

Soit f une fonction C^1 définie sur $[1,+\infty]$, et a_n $(n\geq 1)$ une suite de complexes, soit $A(x)=\sum_{n\leq x}a_n$. On a pour tout x la formule

(1.16)
$$\sum_{n\leq x}a_n f(n) = A(x)f(x) - \int_1^x A(t)f'(t)dt$$

D'autres résultats de théorie des nombres seront utilisés dans la troisième partie, mais il seront cités au moment de leur utilisation: ce sont principalement des relations asymptotiques pour certaines fonctions arithmétiques, d'un caractère moins général que les résultats présentés ici.

II. Un problème extrémal en analyse de Fourier

1. Introduction

Le problème qui est abordé dans cette partie tire son origine de l'observation faite par Beurling (non publié, cf. **[Val]**) que la fonction entière définie par

(2.1) $$b(z)=\left\{\frac{\sin \pi z}{\pi}\right\}^2 \left(\sum_{n=0}^{\infty}(z-n)^{-2} - \sum_{n=-\infty}^{-1}(z-n)^{-2} + 2z^{-1} \right)$$

qui est de type exponentiel et d'ordre 2π, vérifie

(2.2) $$\begin{gathered} b(x)\geq \mathrm{sgn}(x) \text{ pour tout } x \text{ réel} \\ \int b(x)-\mathrm{sgn}(x)dx=1 \end{gathered}$$

Ceci motive l'énoncé du problème suivant : trouver, parmi les fonctions entières de type exponentiel σ donné, celles qui vérifient (2.2), avec f à la place de b, et telles que $\int f(x)-\mathrm{sgn}(x)dx$ soit minimale (cette intégrale étant toujours définie comme intégrale de Lebesgue d'une fonction mesurable **positive**).

On va prouver un théorème du à Beurling (également non publié), à savoir:

Théorème 1: *Soit f une fonction entière de type exponentiel σ, vérifiant la condition* (2.2). *Alors si* $\sigma>0$, *on a*

(2.3) $$\int f(x)-\mathrm{sgn}(x)dx \geq \frac{2\pi}{\sigma}$$

et si $\sigma=0$, *alors* $\int f(x)-\mathrm{sgn}(x)dx = +\infty$.

De plus, si $\sigma>0$, *on a égalité dans* (2.3) *si et seulement si* $f(z)=b(\frac{\sigma}{2\pi}z)$.

L'obtention de ce résultat sera l'objet du **2.**, tandis qu'en **3.**, on verra une première application de ce résultat, ou plus exactement de la fonction b.

2. La fonction de Beurling et sa propriété extrémale

La démonstration du théorème 1 utilisera un lemme pour prouver que la fonction b vérifie effectivement les propriétés annoncées, et un théorème de représentation des fonctions de E^p, qui sera utilisé pour l'unicité. Nous commencerons par ce dernier résultat, qui permet de motiver les fonctions introduites ensuite.

Théorème 2 (Vaaler **[Val]**): *Soit* $f \in E^p$, $p<+\infty$. *Alors pour tout z*

(2.4)
$$f(z)=\left\{\frac{\sin \pi z}{\pi}\right\}^2 \left(\sum_{n=-\infty}^{\infty} f(n)(z-n)^{-2} + \sum_{n=-\infty}^{\infty} f'(n)(z-n)^{-1} \right)$$

et la convergence est uniforme sur les compacts de $\mathbb{C}$. *Si* $p=2$, *on a bien sur la représentation*

(2.5)
$$f(z)=\int_{-1}^{1} \hat{f}(t)e(zt)dt$$

du théorème de Paley-Wiener, et $\hat{f}$ *est donnée presque partout sur* $[-1,1]$ *par*

(2.6)
$$\hat{f}(t)=(1-|t|)u(t)+\frac{1}{2i\pi}\mathrm{sgn}(t)v(t)$$

où u et v sont des fonctions périodiques de $L^2([0,1])$, *dont le développement en série de Fourier est*

(2.7)
$$u(t)= \sum_{n=-\infty}^{+\infty} f(n)e(-nt)$$
$$v(t)= \sum_{n=-\infty}^{+\infty} f'(n)e(-nt)$$

DÉMONSTRATION

Pour p=2, l'existence d'un tel développement est assez facile à motiver: partant de (2.5), on peut développer $\hat{f}$ en série de Fourier (puisque elle est à support compact), et en intégrant terme à terme on peut s'attendre à trouver quelque chose de semblable à (2.4). Cependant, en procédant exactement ainsi on s'expose à des difficultés du fait que développer $\hat{f}$ requiert des fonctions de période 2. Aussi va-t-on, par une astuce préliminaire, se ramener sur [0,1].

Nous supposons d'abord ici p=2 (ou, par (1.5), $p\leq 2$), et donc on a (2.5). Soit u et v les fonctions définies sur [0,1] par

$$u(t)=\hat{f}(t)+\hat{f}(t-1)$$
$$v(t)=2i\pi(t\hat{f}(t)+(t-1)\hat{f}(t-1))$$

et étendues sur $\mathbb{R}$ par 1-périodicité. $\hat{f}$ étant dans L^2, u et v sont clairement dans $L^2([0,1])$. Partant de la définition de u et v, en considérant les cas $x\geq 0$, et $x<0$ séparément, on trouve sans difficulté (2.6).

Les développements en série de Fourier (2.7) viennent de

$$\begin{aligned} f(n) &= \int_{-1}^{1} \hat{f}(t)e(tn)dt = \int_{0}^{1} \hat{f}(t)e(tn)dt + \int_{-1}^{0} \hat{f}(t)e(tn)dt \\ &= \int_{0}^{1} \hat{f}(t)e(tn)dt + \int_{-1}^{0} \hat{f}(t)e(tn)dt \\ &= \int_{0}^{1} \hat{f}(t)e(tn)dt + \int_{0}^{1} \hat{f}(t-1)e((t-1)n)dt \\ &= \int_{0}^{1} (\hat{f}(t) + \hat{f}(t-1))e(tn)dt \\ f(n) &= \int_{0}^{1} u(t)e(tn)dt \end{aligned}$$

Ce résultat, appliqué à f' avec $\hat{f'}(t)=2i\pi t\hat{f}(t)$, donne aussi $f'(n)=\int_0^1 v(t)e(tn)dt$.

On utilise maintenant (1.8) et (1.9), pour écrire:

$$\left\{\frac{\sin \pi z}{\pi}\right\}^2 \frac{f(n)}{(z-n)^2} = f(n)\left\{\frac{\sin \pi(z-n)}{\pi(z-n)}\right\}^2 = \int_{-1}^{1}(1-|t|)e(t(z-n))dt$$

$$\left\{\frac{\sin \pi z}{\pi}\right\}^2 \frac{f'(n)}{(z-n)} = f'(n)\,(z-n)\left\{\frac{\sin \pi(z-n)}{\pi(z-n)}\right\}^2 = \int_{-1}^{1} \frac{1}{2i\pi}\mathrm{sgn}(t)e(t(z-n))dt$$

d'où il vient, en sommant de -N à N

$$\int_{-1}^{1}\left\{(1-|t|)u(t,N)+\frac{1}{2i\pi}\mathrm{sgn}(t)v(t,N)\right\}e(tn)dt = \left\{\frac{\sin \pi z}{\pi}\right\}^2\left(\sum_{n=-N}^{N} f(n)(z-n)^{-2} + \sum_{n=-N}^{N} f'(n)(z-n)^{-1}\right)$$

u(t,N) et v(t,N) désignant les sommes partielles des séries de Fourier de u et v, de -N à N. On va maintenant faire $N \to \infty$ dans les deux membres.

Du côté droit, comme u et v sont dans L^2, les séries $\sum|f(n)|^2$ et $\sum|f'(n)|^2$ convergent (cf. aussi (1.7)), ce qui permet de prouver que les séries qui s'y trouvent convergent uniformément sur tout compact de $\mathbb{C}$: pour cela on utilise l'inégalité de Cauchy - Schwarz, et on observe simplement que si $q\geq 2$ (ici, les cas q=2 et q=4 sont utilisés), la série $\sum_n \left|\frac{\sin \pi z}{z-n}\right|^q$ converge uniformément sur tout compact (si $|z|\leq A$, pour $n\geq N_0$, $|z-n|\geq|n|-A$, etc...).

Du côté gauche, c'est un résultat fondamental de la théorie des séries de Fourier que u(.,N) (resp. v(.,N)) converge vers u (resp. v) en norme L^2, et comme on est sur un ensemble de mesure finie, cela reste vrai en norme L^1.

Ceci entraîne compte tenu de (2.6) et du fait que l'on multiplie u(.,N) et v(.,N) par des fonctions qui sont bornées:

$$\int_{-1}^{1}\left\{(1-|t|)u(t,N)+\frac{1}{2i\pi}\mathrm{sgn}(t)v(t,N)\right\}e(tz)dt \to \int_{-1}^{1}\hat{f}(t)e(tz)dt = f(z)$$

En comparant les 2 côtés, on trouve donc bien (2.4), ce qui conclut le cas $p\leq 2$.

Dans le cas $p>2$, la méthode est à peu près la suivante: on prouve que la fonction r définie par $r(z)=\begin{cases}\frac{f(z)-f(0)}{z} & \text{si } z\neq 0;\\ f'(0) & \text{si } z=0\end{cases}$ est dans E^2, puis on utilise le cas précédent pour développer r et on exploite la définition de r pour se ramener à f en manipulant les séries obtenues. On aboutit alors à la conclusion désirée. Je ne détaillerai pas davantage ce calcul car dans les applications à suivre, seul le premier cas sera utilisé ♦

N.B. Dans **[Val]**, il est en fait prouvé un cas plus général de ce théorème. En particulier, (2.4) est valide si f est seulement **bornée** (sur$\mathbb{R}$) et **impaire**, auquel cas, bien sur, la série en $\frac{1}{z-n}$ doit être interprétée comme limite des sommes partielles symétriques puisqu'elle ne converge pas absolument.

Ce théorème 2 motive l'emploi, pour tenter approcher sgn par une fonction de type exponentiel, de la fonction h définie par

(2.8) $$h(z)=\left\{\frac{\sin \pi z}{\pi}\right\}^2\left(\sum_{n=-\infty}^{\infty}\mathrm{sgn}(n)(z-n)^{-2}+2z^{-1}\right)$$

Nous utiliserons aussi les fonctions suivantes:

$$k(z)=\left\{\frac{\sin \pi z}{\pi z}\right\}^2 \qquad\qquad j(z)=\frac{1}{2}h'(z)$$

Lemme 1: h, k *et* j *sont des fonctions entières de type exponentiel* 2π *et* b=h+k *est la fonction de Beurling* (2.1).

DÉMONSTRATION

L'énoncé pour k est trivial. Pour h, il faut d'abord observer que les zéros du sinus annulent les pôles possibles. Pour obtenir le type exponentiel 2π, on peut considérer d'abord le cas $y=\mathrm{Im}(z)\geq 1$, où la série représente une fonction bornée, et si $y<1$, il suffit de considérer pour un $z=x+iy$ donné l'indice n_0 où $|n_0-x|\leq 1/2$, de diviser la série en d'une part le terme d'indice n_0, qui multiplié par sin πz est borné indépendamment de n_0, et d'autre part le reste de la série, qui est majoré par une série convergente, le sin πz en facteur étant $O(e^{2\pi|z|})$. Pour j enfin, c'est une conséquence du résultat sur le type d'une dérivée. Enfin, l'affirmation h+k=b est immédiatement vérifiée ♦

Lemme 2: *On a, pour tout x réel, l'inégalité*

$$|\mathrm{sgn}(x)-h(x)|\le k(x)$$

Ceci entraîne l'inégalité de Beurling $b(x)\ge \mathrm{sgn}(x)$. *De plus,* b-sgn *est intégrable et on a* $\int b(x)-\mathrm{sgn}(x)dx = 1$.

DÉMONSTRATION

Pour le premier point, il suffit de prouver $0\le 1-k(x)\le h(x)\le 1$ pour $x>0$. En effet, alors l'inégalité du lemme est démontrée pour $x>0$, et pour $x<0$, sgn et h étant impaires et k paire:

$$|\mathrm{sgn}(x)-h(x)|=|-(\mathrm{sgn}(-x)-h(-x))|=|\mathrm{sgn}(-x)-h(-x)|\le k(-x)=k(x)$$

L'inégalité $0\le 1-k(x)$ est immédiate. Pour le reste, on utilise l'identité (1.10) qui permet d'écrire, pour $x>0$

$$\begin{aligned} h(x) &= 1+\left\{\frac{\sin \pi x}{\pi}\right\}^2\left(2x^{-1}-x^{-2}-2\sum_{n=1}^{\infty}(x+n)^{-2}\right) \\ &= 1-k(x)+\left\{\frac{\sin \pi x}{\pi}\right\}^2\left(2x^{-1}-2\sum_{n=1}^{\infty}(x+n)^{-2}\right) \end{aligned} \qquad \textbf{(2.9)}$$

mais on a

$$\sum_{n=1}^{\infty}(x+n)^{-2} \le \sum_{n=1}^{\infty}(x+n)(x+n-1) = \sum_{n=1}^{\infty}\{(x+n-1)^{-1}-(x+n)^{-1}\} = x^{-1}$$

et donc le contenu du facteur de $\left\{\frac{\sin \pi x}{\pi}\right\}^2$ dans (2.9) est ≤ 0, ce qui prouve bien l'inégalité $1-k(x)\le h(x)$. Enfin, comme $a^2+b^2\ge 2ab$, on a

$$x^{-2}+2\sum_{n=1}^{\infty}(x+n)^{-2} = \sum_{n=0}^{\infty}\{(x+n)^{-2}+(x+n+1)^{-2}\}\ge 2\sum_{n=1}^{\infty}(x+n)(x+n-1) = 2x^{-1}$$

et avec (2.9) ceci montre bien $h(x)\le 1$. De $|\mathrm{sgn}(x)-h(x)|\le k(x)$ on déduit aussitôt $b(x)\ge \mathrm{sgn}(x)$.

Finalement, k étant intégrable, l'inégalité $|\mathrm{sgn}(x)-h(x)|\le k(x)$ prouve que sgn-h est intégrable, donc sgn-b=sgn-h-k aussi, et

$$\begin{aligned} \int b(x)-\mathrm{sgn}(x)dx &= \int k(x)dx + \int h(x)-\mathrm{sgn}(x)dx = \int k(x)dx, \text{ car h-sgn est impaire} \\ &= 1, \text{ d'après (1.8) et (1.2)} \ \blacklozenge \end{aligned}$$

N.B. On a en fait également prouvé que la fonction $b_1=h-k$ vérifiait

$$\mathrm{sgn}(x)\le b_1(x) \text{ pour tout } x$$
$$\int \mathrm{sgn}(x)-b_1(x)dx = 1$$

Le lemme suivant est seulement technique et nous en omettons la preuve.

Lemme 2: *La fonction* j *est intégrable. Plus précisément, on a*

$$j(x)=O((1+|x|)^{-3}),$$

IDÉE DE DÉMONSTRATION

L'idée est simple: on exploite la définition de h, (1.8) et (1.9) ainsi qu'une dérivation sous ∫ pour écrire j(z) sous la forme (2.5) et on montre que $j\in L^1$ ♦

PREUVE DU THÉORÈME 1

Supposons d'abord $\sigma>0$. Alors, comme on l'a déjà remarqué, on peut se ramener au cas $\sigma=2\pi$ car sgn(ax)=sgn(x) si a>0. Soit donc f une fonction entière de type exponentiel 2π vérifiant (2.2). Si on a $\int f(x)-\mathrm{sgn}(x)dx = +\infty$, (2.4) est démontré, donc on peut supposer que g=f-sgn est intégrable. Soit $m=\frac{1}{2}$ f'.

D'après le lemme 2, b-sgn est intégrable, donc la différence f-b est intégrable, et comme b=h+k avec k intégrable, f-h est intégrable, c'est à dire $f-h\in E^1$. La propriété (1.6) assure donc que $\frac{1}{2}$ (f'-h') = m-j est intégrable, et grâce au lemme 3, m est intégrable.

Nous allons maintenant faire une incursion du côté des distributions pour prouver la formule suivante

$$\hat{g}(t)=\frac{\hat{m}(t)-1}{i\pi t}\text{ , si } t\neq 0$$

En effet, on vérifie (cf. **[Sch]** chap. VII) que tant f que sgn, en tant que distributions, possèdent une transformée de Fourier (distribution), qui pour une distribution T est définie par $\hat{T}(\varphi)=T(\hat{\varphi})$. La formule classique $\hat{f'}=2i\pi t\,\hat{f}$ reste vraie, f et f' étant ici les distributions correspondantes, et on l'applique à g: si t≠0 il vient

$$\hat{g}=\hat{f}-\widehat{\mathrm{sgn}}=\frac{1}{2i\pi t}\{\hat{f'}-(\mathrm{sgn}')\hat{}\,)\}$$

Or sgn=2Y-1 où Y est la distribution de Heaviside, et donc sgn'=2Y'=2δ (distribution de Dirac), donc $(\mathrm{sgn}')\hat{}=2\hat{\delta}=2$, car $\hat{\delta}(\varphi)=\delta(\hat{\varphi})=\hat{\varphi}(0)=\int\varphi(x)dx=1(\varphi)$.

Exploitant cela, on a donc $\hat{g}=\frac{1}{2i\pi t}\{\hat{f'}-(\mathrm{sgn}')\hat{}\,)\}=\frac{\hat{m}-1}{i\pi t}$ et les deux membres de cette égalité étant des fonctions continues, on en déduit le résultat annoncé.

Maintenant, m est une fonction de type exponentiel 2π, et donc $\hat{m}$ est supportée sur [-1,1] par le théorème de Paley - Wiener, d'où pour $|t|\geq 1$, $\hat{g}(t)=-\frac{1}{i\pi t}$.

Mais g est une fonction intégrable, normalisée car f et sgn le sont, à variation bornée (car f' est intégrable et sgn l'est), donc la formule de Poisson (1.4) est applicable et assure qu'en tout x

$$\sum_{n=-\infty}^{+\infty} g(n+x) = \sum_{n=-\infty}^{+\infty} \hat{g}(n)e(xn)$$

Développant le membre de gauche grâce à la formule pour $\hat{g}$, on a donc

$$\sum_{n=-\infty}^{+\infty} g(n+x) = \hat{g}(0) - \sum_{n\neq 0} \frac{e(nx)}{i\pi n}$$

$$= \begin{cases} \hat{g}(0) + 2(x-[x]-1/2) \text{ si } x\notin \mathbb{Z} \\ \hat{g}(0) \text{ sinon} \end{cases}$$

, grâce à un calcul simple de série de Fourier

Comme g est positive, on a donc $\hat{g}(0) + 2(x-[x]-1/2)\geq 0$, ce qui avec $x\to 0^+$ donne $\hat{g}(0)\geq 1$, c'est à dire exactement (2.3).

Supposons maintenant que l'on soit dans un cas d'égalité, c'est à dire $\hat{g}(0)=1$. On a alors

$$\lim_{x\to 0^+} \sum_{n=-\infty}^{+\infty} g(n+x) = 0$$

Comme $g\geq 0$, $0\leq g(n+x) \leq \sum_{n=-\infty}^{+\infty} g(n+x)$, et donc on en déduit que pour tout $l\in \mathbb{Z}$, $f(l)=f(l^+)=sgn(l^+)$, d'où $f(l)=b(l)$. Comme de plus $f\geq sgn$, et que f et sgn sont dérivables pour tout $l\neq 0$, on a pour $l\neq 0$, $f'(l)=0=b'(l)$. Par conséquent si on utilise le théorème 2 pour exprimer f-b qui est dans E^1, on trouve

$$f(z)-b(z) = \frac{f'(0)-2}{z}\left\{\frac{\sin \pi z}{\pi}\right\}^2$$

et comme cette dernière fonction doit être intégrable, il vient $f'(0)=2$, ce qui donne finalement $f=b$.

Pour conclure, dans le cas $\sigma=0$, f est de type exponentiel τ pour tout $\tau>0$, donc (2.3) est vraie pour tout σ, et par conséquent $\int f(x)-sgn(x)dx=+\infty$ ♦

Bien qu'elle ait des applications propres, la fonction b est souvent utilisée par le biais d'une autre fonction introduite par Selberg dans le but de majorer la fonction caractéristique d'un intervalle $E=[\alpha,\beta]$. En effet, on vérifie facilement que si $x\neq\alpha$ et $x\neq\beta$, on peut écrire $\chi_E(x)=\frac{1}{2}\{sgn(\beta-x)+sgn(x-\alpha)\}$, et donc d'après (2.2) la fonction entière de type exponentiel 2π définie par $C_E(z)=\frac{1}{2}\{b(\beta-z)+b(z-\alpha)\}$ vérifie

(2.10)

$$C_E(x) \geq \chi_E(x) \text{ pour tout } x$$
$$\int C_E(x) - \chi_E(x)dx = 1$$

La dernière formule montre même que C_E est intégrable, donc $C_E \in E^1$. Au vu de ces deux propriétés, on pourrait s'attendre à ce que la fonction C_E soit extrémale, de façon similaire au cas de la fonction b. Ce n'est cependant pas le cas, en général. Selberg a prouvé que pour une fonction f de type exponentiel 2π majorant χ_E, on a $\int f(x) - \chi_E(x)dx \geq 1$, **si** $\beta-\alpha$ est un entier. Et même dans ce cas, C_E n'est pas l'unique fonction extrémale. Les solutions furent déterminées par Selberg. Dans le cas où $\beta-\alpha$ n'est pas entier, $\int f(x) - \chi_E(x)dx \geq 1$ est faux en général. Cependant, Logan a déterminé qu'alors il existe une unique fonction extrémale. Cette différence peut s'expliquer: on peut supposer puisque $C_{[\alpha,\beta]}(z) = C_{[0,\beta-\alpha]}(z-\alpha)$, que $\alpha=0$. Si β est entier,il y a simplification dans l'expression de C_E. On a alors

$$C_{[0,\beta]}(z) = \left\{\frac{\sin \pi z}{\pi}\right\}^2 \left(\sum_{n=0}^{\beta} (z-n)^{-2} + z^{-1} - (z-\beta)^{-1} \right)$$

Rien de tel ne se produit dans l'autre cas.

C'est la fonction de Selberg C_E qui sera utilisée dans la troisième partie.

3. Une première application

Comme première application de la fonction de Beurling b, nous allons prouver ici une variante de l'inégalité de Hilbert, due à Montgomery et Vaughan, dont la démonstration devient très simple (par rapport à la première, cf. **[Mon]**).

Rappelons d'abord que l'inégalité de Hilbert originale est

$$\left| \sum_{n \neq m} \frac{a_n \bar{a_m}}{n-m} \right| \leq \pi \sum |a_n|^2$$

la constante π est la meilleure possible et fut obtenue par Schur en 1911 (Hilbert avait obtenu ce résultat avec la constante 2π).

N.B. Cette inégalité revient au calcul de la norme d'une forme bilinéaire.

Théorème 3 (Montgomery - Vaughan, 1974): *Soient* $\lambda_1,...,\lambda_N$ *des réels vérifiant* $|\lambda_n - \lambda_m| \geq \delta > 0$ *si* $i \neq j$. *Alors, pour tout* $(a_n)_{1 \leq n \leq N} \in \mathbb{C}^N$, *on a*

$$\left| \sum_{n \neq m} \frac{a_n \bar{a_m}}{\lambda_n - \lambda_m} \right| \leq \pi \delta^{-1} \sum |a_n|^2$$

DÉMONSTRATION (Vaaler, **[Val]**)

Notons f=b-sgn=k+h-sgn, qui est positive et intégrable, et f_δ la fonction définie par $f_\delta(z)=\delta f(\delta z)$, de sorte que $\hat{f}_\delta(t)=\hat{f}(\delta^{-1}t)$. Comme dans la démonstration du théorème 2, on montre que $\hat{f}(t)=-\frac{1}{i\pi t}$ si $|t|\geq 1$. On a $\hat{f}(0)=1$. Par conséquent il vient

$$0\leq \int \left|\sum_{n=1}^{N} a_n e(-\lambda_n x)\right|^2 f_\delta(x)\,dx = \sum_{n,m} a_n \overline{a_m}\, \hat{f}_\delta(\lambda_n-\lambda_m)$$

$$= \sum_n |a_n|^2 \hat{f}_\delta(0) - \sum_{n\neq m} a_n \overline{a_m}\,\frac{\delta}{i\pi(\lambda_n-\lambda_m)}$$

donc

$$\sum_{n\neq m} \frac{a_n \overline{a_m}}{i(\lambda_n-\lambda_m)} \leq \pi\delta^{-1}\sum_n |a_n|^2$$

Maintenant, d'après la remarque faite à la fin de la démonstration du lemme 2, on peut considérer $f_1=\text{sgn}-b_1$, et procéder de même: on aboutira à la même inégalité dans l'autre sens et avec un facteur -1 du côté droit (car $\hat{f}_1(t)= + \frac{1}{i\pi t}$ pour $|t|\geq 1$). Ces deux bornes réunies démontrent le théorème ♦

N.B. Il est clair qu'en prenant $\lambda_n=n$, et en faisant $N \to \infty$, on retrouve l'inégalité de Hilbert, car alors $\delta=1$.

Ceci achève la partie d'analyse. D'autres applications de la fonction b (probabilités, distribution uniforme...), ainsi que des généralisations à la majoration de fonctions plus générales (fonctions normalisées à variation bornée, fonctions périodiques...) par des fonctions entières de type exponentiel se trouvent dans **[Val]**. Une autre approche, qui permet non seulement des majorations, mais aussi des résultats similaires à ceux de Beurling (unicité du majorant ou minorant) dans certains cas particuliers, se trouve dans **[GV]**, avec encore d'autres applications, y compris une autre démonstration du théorème 3 .

III. Le principe analytique du grand crible

1. Introduction et historique des méthodes de crible

Historiquement, la notion de "crible" en arithmétique remonte au crible d'Erathosthène (environ 200 av. JC). Mais la signification moderne de ce terme est beaucoup plus récente, puisque ce n'est qu'au début du $20^{ème}$ siècle qu'apparaissent les premières "méthodes de cribles", outils théoriques puissants d'étude en théorie des nombres.

La formulation du principe des cribles modernes peut se décrire ainsi: d'un ensemble fini d'entiers, fixé, on retire tout les éléments appartenant à certaines classes de congruences modulo p, où p varie dans un ensemble fini de nombres premiers. Cela fait, on cherche à obtenir une estimation du nombre d'entiers restants. Par exemple, si de {2...n} on enlève tout les nombres divisibles (ie $\equiv 0$) par un nombre premier $p \leq \sqrt{n}$, on obtient comme ensemble criblé $\{p \mid \sqrt{n} < p \leq n\}$, et une bonne estimation permet donc d'approcher $\pi(n)-\pi(\sqrt{n})$.

C'est à Viggo Brun que l'on doit les premiers travaux théoriques sur les cribles. Jusque là, le crible d'Erathosthène, même formalisé par Legendre (1780) n'avait guère eu d'applications que calculatoires: dresser des tables, calculer $\pi(x)$... Le crible de Brun lui permet, vers 1919, de démontrer deux théorèmes frappants concernant deux des plus célèbres (et des plus difficiles) conjectures de la théorie des nombres, sur lesquelles jusqu'alors on ne savait pratiquement rien:

- La conjecture de Goldbach (1742) énonce que tout nombre pair $n>4$ est somme de deux nombres premiers. Brun prouve qu'il existe un entier B fixé tel que tout entier pair est somme de deux nombres ayant au plus B facteurs premiers (comptés avec multiplicité).
- Une autre célèbre conjecture affirme qu'il existe une infinité de nombres premiers p tels que p+2 soit aussi premier (nombres premiers jumeaux). Brun prouve que la somme des inverses des nombres premiers jumeaux est convergente (on sait depuis Euler que c'est faux pour l'ensemble des nombres premiers, et Dirichlet a montré (1837) que c'est également faux pour p dans une progression arithmétique an+b, si (a,b)=1).

C'est le début du développement des méthodes de cribles: crible de Brun, crible combinatoire, crible de Selberg, grand crible introduit par Linnik (1941), etc... Chacun de ces cribles est en quelque sorte une technique particulière qui, sous certaines hypothèses sur la nature du crible, permet d'estimer plus ou moins bien le nombre d'éléments non criblés.

C'est le grand crible qui est notre sujet ici. Pour expliquer ce qui le caractérise, notons E l'ensemble initial, P celui des nombres premiers, et pour chaque $p \in P$, soit $\omega(p)$ le nombre de

classes de congruences retirées. Alors que les "petits" cribles sont conçus pour le cas où $\omega(p)$ reste borné quand p tend vers l'infini (comme dans l'exemple ci-dessus, où $\omega(p)=1$), le grand crible permet de traiter les cas où $\omega(p)$ peut tendre vers l'infini (par exemple, si on crible par les carrés modulo p, $\omega(p)=(p-1)/2$). Cependant, nous verrons que le grand crible reste à même de fournir d'excellents résultats même dans des contextes de "petit" crible.

Nous allons maintenant énoncer puis prouver l'inégalité fondamentale, ce qui sera l'occasion de discuter davantage la nature et l'évolution du grand crible.

2. L'inégalité fondamentale, discussion

Le grand crible prend son origine dans un article de Linnik, en 1941. Il fut ensuite essentiellement développé par Rényi à partir de 1947, mais resta très complexe jusqu'à la décennie 1965-1975 où de nombreux travaux permirent d'en simplifier beaucoup la formulation, tout en en accroissant amplement l'efficacité. En particulier, en 1966, Davenport et Halberstam montrèrent comment le ramener à une simple inégalité analytique très générale, et depuis c'est cette approche qui a presque toujours été retenue.

Avant d'énoncer cette inégalité, une définition est utile:

Définition: *Pour tout x réel on note* $\|x\|=\text{Min}\{|x-n| \mid n\in\mathbb{Z}\}$. *On dit qu'une suite finie* $(x_1,...,x_n)$ *de réels est* δ**-bien espacée modulo 1** *si* $\|x_i-x_j\|\geq\delta$ *pour tout (i,j) tel que* $i\neq j$.

N.B. On a vu une notion proche dans le théorème 3. La fonction ||.|| est périodique de période 1 et $\|x\|=|x|$ si $x\in[-1/2,1/2]$.

Soit maintenant S un polynôme trigonométrique de période 1 donné par

$$S(x)=\sum_{n=M+1}^{M+N}a_n e(nx)$$

et soient $\xi_1,...,\xi_R$ des réels δ-bien espacés modulo 1. L'inégalité du grand crible est une inégalité de la forme

(3.1)
$$\sum_{r=1}^{R}|S(\xi_r)|^2\leq\Delta(N,\delta)\sum_{n=M+1}^{M+N}|a_n|^2$$

et le problème est de prouver cette inégalité, qui doit être valable pour tout choix des ξ_i (soumis à la condition précisée) et des a_i, pour un Δ aussi petit que possible.

N.B. • L'absence du paramètre M dans Δ n'a aucune incidence sur la qualité du résultat puisque on peut écrire $S(x)=e(Nx)T(x)$ où T est un polynôme trigonométrique de la même forme que S, avec M=0. Quand au paramètre R, il faut noter que δ en dépend fortement: par exemple, on a $\delta \leq \frac{1}{R}$.

• Il est utile de remarquer qu'on peut, un peu comme dans l'inégalité de Hilbert, interpréter (3.1) comme une estimation de la norme de l'application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^R$, qui au vecteur $(a_{N+1},...,a_{M+N})$ associe $(\sum_{n=M+1}^{M+N} a_n e(n\xi_1),..., \sum_{n=M+1}^{M+N} a_n e(n\xi_R))$.

Avant de prouver une "bonne" version de l'inégalité (3.1), on peut aisément déterminer des minorations pour Δ, qui permettent de juger ensuite de la qualité du résultat obtenu.

Tout d'abord, si l'on pose $a_n=e(-n\xi)$ et R=1, alors $|S(\xi)|^2=N^2=N\sum|a_n|^2$, donc $\Delta\geq N$. D'ailleurs si R=1, (3.1) avec $\Delta=N$ est valide: il s'agit alors simplement de la classique inégalité de Cauchy - Schwarz.

Plus subtilement, supposons les ξ_i également espacés: $||\xi_i-\xi_j||=\delta=R^{-1}$ si $i\neq j$. On observe alors que $\int_0^1\sum|S(\xi_r+\xi)|^2d\xi=R\int_0^1|S(\xi)|^2d\xi=R\sum|a_n|^2$, donc pour au moins un ξ, on a $\sum|S(\xi_r+\xi)|^2\geq R\sum|a_n|^2$, et par conséquent, $\Delta\geq R=\delta^{-1}$.

La recherche de valeurs admissibles pour Δ a été l'objet d'une grande compétition entre 1965 et 1975. Le résultat de Davenport et Halberstam en 1966 était $\Delta\leq 2{,}2\,\mathrm{Max}(2N,\delta^{-1})$. Gallagher trouva $\Delta\leq\pi N+\delta^{-1}$, Montgomery obtint $\Delta\leq N+2\delta^{-1}$, puis avec Vaughan - et à l'aide de leur théorème 3 -, $\Delta\leq N+\delta^{-1}$ (résultat qui était considéré en 1972 par Huxley comme une "conjecture optimiste"). Finalement, Selberg en 1974 (non publié cependant) prouva à l'aide de la fonction C_E du II que $\Delta\leq N-1+\delta^{-1}$ et, comme on le verra, cette dernière borne est d'une certaine façon la meilleure. Toutes ces valeurs de Δ furent obtenues à l'aide d'autant de méthodes différentes. Nous présenterons deux preuves, celle de Gallagher, très simple, et celle de Selberg qui donne donc le meilleur résultat actuellement connu.

Théorème 4 (Gallagher 1967): (3.1) *est valide avec* $\Delta=\pi N+\delta^{-1}$.

DÉMONSTRATION

On emploie un lemme, qui appartient en fait à l'analyse fonctionnelle et est du à Sobolev.

Lemme: si $f\in C^1([0,1])$, on a $|f(\frac{1}{2})|\leq\int_0^1|f|+\frac{1}{2}|f'|$.

En effet, on vérifie que $f(\frac{1}{2})=\int_0^1 f(x)dx + \int_0^{1/2} xf'(x)dx + \int_{1/2}^1 (x-1)f'(x)dx$, d'où le lemme en prenant les valeurs absolues.

Maintenant, on pose $f(x)=S(x)^2$, et on utilise le lemme en changeant de variable pour ramener l'intervalle [0,1] à $[\xi_r-\delta/2,\xi_r+\delta/2]$. On trouve ainsi

$$|S(\xi_r)|^2 \le \delta^{-1}\int_{\xi_r-\delta/2}^{\xi_r+\delta/2}|S(x)|^2dx + \int_{\xi_r-\delta/2}^{\xi_r+\delta/2}|S(x)S'(x)|dx$$

On somme maintenant sur r, en observant que la 1-périodicité de S permet de se ramener sur [0,1], et que le bon espacement des ξ implique que les intervalles d'intégration sont distincts pour des r distincts. Donc il vient

$$\sum_{r=1}^{R}|S(\xi_r)|^2 \le \delta^{-1}\int_0^1|S(x)|^2dx + \int_0^1|S(x)S'(x)|dx$$

Il ne reste qu'à appliquer l'identité de Parseval à S et à S' - et pour la deuxième intégrale d'abord l'inégalité de Cauchy - Schwarz - pour avoir

$$\sum_{r=1}^{R}|S(\xi_r)|^2 \le \delta^{-1}\sum_{n=M+1}^{M+N}|a_n|^2 + 2\pi(\text{Max }|n|)\sum_{n=M+1}^{M+N}|a_n|^2$$

Comme on a vu que M pouvait être choisi arbitrairement, on observe finalement qu'on peut le prendre pour avoir $\text{Max }|n| \le \frac{N}{2}$, d'où finalement le résultat annoncé ♦

Théorème 5 (Selberg): (3.1) *est valide avec* $\Delta=N-1+\delta^{-1}$.

DÉMONSTRATION

C'est donc une application des résultats de la deuxième partie, plus précisément des propriétés (2.10) de la fonction de Selberg.

On introduit l'intervalle $E=[\delta(M+1),\delta(M+N)]$, et on note f la fonction entière de type exponentiel $2\pi/\delta$ définie par $f(z)=C_E(\delta z)$. On a ainsi $f \ge \chi_{[M+1,M+N]}$, et de plus $f \in E^1$ et pour tout réel t il vient $\hat{f}(t)=\delta^{-1}\hat{C}_E(\delta^{-1}t)$.

Par (1.5), $f \in E^2$, et le théorème de Paley - Wiener dit donc que $\hat{f}$ est supportée sur $[-\delta,\delta]$. Enfin, on calcule:

$$\begin{aligned}\hat{f}(0)&=\delta^{-1}\hat{C}_E(0)=\delta^{-1}\int C_E(x)dx\\ &=\delta^{-1}(1+\delta(M+N)-\delta(M+1)), \text{ d'après (2.10)}\\ &=N-1+\delta^{-1}\end{aligned}$$

Maintenant, nous allons exploiter la remarque faite après l'énoncé de (3.1). On sait que pour toute application linéaire u, on a $||u||=||{}^t u||$. Il en découle que (3.1) est équivalente à l'inégalité duale

(3.2) $$\sum_{n=M+1}^{M+N}|s(n)|^2 \le \Delta(N,\delta)\sum_{r=1}^{R}|b_r|^2$$

où s désigne le polynôme trigonométrique **presque périodique** donné par

$$s(x)=\sum_{r=1}^{R}b_r e(x\xi_r)$$

l'inégalité (3.2) devant être vraie pour tout choix des b_r. Nous allons montrer (3.2). Comme $f\ge\chi_{[M+1,M+N]}$, il vient

$$\sum_n|s(n)|^2 \le \sum_{n=-\infty}^{\infty} f(n)|s(n)|^2$$

On développe alors les modules et on inverse l'ordre de sommation, d'où

(3.3) \I\FC\∑(n,,|s(n)|2) ≤ \I\SU(r,s,,)\O(br,\D\BA1()¯) bs \B\BC(\I\SU(n=-∞,∞, f(n)e((ξs-ξr)n)))

On applique ensuite la formule de Poisson (1.4) à $\hat{f}$. Bien que l'on n'ait pas prouvé que $\hat{f}$ soit à variation bornée, le fait qu'elle soit continue à support compact contenu dans [-1,1] (car $\delta\le 1$), prouve que la série $\sum_{n=-\infty}^{\infty}\hat{f}(n-x)$ se réduit en fait pour chaque x à un nombre fini de termes non nuls (au plus 3), et représente donc une fonction continue dont on vérifie que les coefficients de Fourier sont $c_n=f(n)$. Comme $\sum|c_n|<+\infty$, cette fonction est partout somme de sa série de Fourier. D'où

$$\sum_{n=-\infty}^{\infty} f(n)e((\xi_s-\xi_r)n) = \sum_{n=-\infty}^{\infty}\hat{f}(n-(\xi_s-\xi_r))$$

Or si $s\ne r$, on a pour tout n, $|n-(\xi_s-\xi_r)|\ge||\xi_s-\xi_r||\ge\delta$, et donc $\hat{f}(\xi_s-\xi_r)=0$.

Par conséquent dans (3.3) les termes avec $r\ne s$ sont tous nuls et on peut écrire

$$\sum_n|s(n)|^2 \le \sum_{r=1}^{R}|b_r|^2\,\hat{f}(0)$$

Comme $\hat{f}(0)=N-1+\delta^{-1}$, on a prouvé (3.2) et donc (3.1) ♦

Discussion de l'inégalité du grand crible

• Si les ξ sont également espacés, $R^{-1}\sum_r |S(\xi_r)|^2$ est une somme de Riemann qui approche $\int_0^1 |S(\xi)|^2 d\xi$ et donc (3.1) a alors une interprétation simple. Dans le cas général, cette idée peut aider à la compréhension de la nature de cette inégalité.

• L'inégalité (3.1), dans la version fournie par le théorème 5, est très forte. On le verra dans les applications, on peut déjà en juger par la comparaison du Δ obtenu avec les bornes "triviales" trouvées précédemment: on sait que $\Delta \geq \delta^{-1}$ et que $\Delta \geq N$, et on a trouvé $\Delta \leq N-1+\delta^{-1}$. Ajoutons que la majoration la plus triviale avec l'inégalité de Cauchy - Schwarz (et sans utiliser l'hypothèse d'espacement) donne (3.1) avec $\Delta = RN \leq \delta^{-1}N$. Ceci dit, dans les applications, le fait d'avoir N-1 au lieu de N s'avère pratiquement sans importance, et même dans l'ensemble on obtient quasiment les mêmes résultats avec chacune des variantes de (3.1) indiquées plus haut. Tout au plus le fait d'avoir N au lieu de aN permet-il d'assurer des constantes plus "propres". Mais on travaille souvent pour avoir des résultats en O(...) et une constante multiplicative de plus n'est pas très importante.

• Bien que nous ayons prouvé le théorème 5 par le biais de l'analyse de Fourier, faisant ainsi intervenir des résultats profonds d'analyse harmonique, le grand crible doit être considéré comme une méthode "élémentaire", au sens habituel en théorie des nombres, à savoir ne faisant pas intervenir l'analyse complexe. En effet, le théorème 4 l'est entièrement et il existe des preuves élémentaires de (3.1) avec $\Delta = N+\delta^{-1}$, reposant seulement sur des généralisations de l'inégalité de Bessel dans les espaces de Hilbert (cf. Bombieri **[Bom]**).

• On peut ajouter que la pertinence du choix de (3.1) comme principe du grand crible est confirmé par le fait qu'on puisse en déduire les deux formes du grand crible: la forme additive (majoration du nombre d'élément d'un ensemble criblé), mais également la plus subtile (et plus profonde, cf. Bombieri **[Bom]**) forme multiplicative (majoration de sommes portant sur les caractères de Dirichlet) introduite par Rényi dés 1948.

• Le théorème 5 donne d'une certaine façon le meilleur résultat possible. Ceci fut prouvé par Bombieri et Davenport en 1969. En effet, il existe des polynômes S avec $N \to \infty$ et $\delta \to 0$, $N\delta \to \infty$ (conditions habituellement vraies dans les applications), pour lesquels on a égalité dans (3.1) avec $\Delta = N-1+\delta^{-1}$.

Plus précisément, soit h un entier (grand), L un entier (grand) et soit

$$S(x) = \sum_{n=-L}^{L} e((2h+1)nx)$$

On a donc $N=2(2h+1)L+1$ et $\sum_n |a_n|^2=2L+1$. Prenons $\xi=\pm \frac{r}{2h+1}$, pour $0\le r\le h$, de sorte que $R=2h+1$ et $\delta=\frac{1}{2h+1}$ (d'où $N\delta=2L+\frac{1}{2h+1}$). On vérifie aussitôt que $S(\xi)=2L+1$ pour chacun de ces ξ. Par conséquent

$$\sum_r |S(\xi r)|^2=(2h+1)(2L+1)^2=(2h+1)(2L+1)\sum_n |a_n|^2$$

et on a $N-1+\delta^{-1}=2(2h+1)L+2h+1=(2h+1)(2L+1)=\Delta$.

• L'inégalité (3.1) est également intéressante par les généralisations qu'elle permet, et qui sont souvent très utiles. Un certain nombre d'entre elles sont décrites dans l'article de Montgomery **[Mon]**. Elles sont assez diverses. Par exemple, il est possible de prouver des inégalités similaires pour d'autres classes de fonctions, par exemple pour les polynômes algébriques ou les séries de Dirichlet. D'un autre côté, (3.1) a été généralisée en dimension supérieure, pour traiter des problèmes dans les corps de nombres. Une autre variante due à Montgomery et Vaughan consiste à introduire, au lieu de la seule quantité δ, les δ_r tels que l'on ait $||\xi_r-\xi_s||\ge\delta_r$ pour tout $s\ne r$. On obtient ainsi un résultat un peu plus fin, qui a été mis à profit dans certaines applications:

(3.4)
$$\sum_r |S(\xi_r)|^2 (N+1{,}5\delta_r^{-1})^{-1} \le \sum_{n=M+1}^{M+N} |a_n|^2$$

3. Formes additives et multiplicatives du grand crible

Les applications de l'inégalité du grand crible (3.1) à la théorie des nombres se font en prenant pour les ξ_r les rationnels de la forme $\frac{a}{q}$, où $(a,q)=1$ et $q\le Q$. Q est ici un paramètre. Ces éléments, une fois ordonnés, forment la **suite de Farey** d'ordre Q.

Exemple: pour Q=5, la suite correspondante est

$$\frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, 1$$

On notera (bien que cela ne sera pas utilisé dans la suite) que si $\frac{a}{b}$, $\frac{e}{f}$ et $\frac{c}{d}$ sont trois termes successifs de la suite de Farey, on a les propriétés caractéristiques $ad-bc=1$, et $\frac{e}{f}=\frac{a+c}{b+d}$.

(cf. **[H&W]** chap. III pour une preuve et d'autres propriétés).

Si $\frac{a}{q} \neq \frac{a'}{q'}$, on a $||a/q-a'/q'||=||(aq'-a'q)/qq'||\geq 1/qq' \geq Q^{-2}$, de sorte que l'on peut appliquer (3.1) avec $\delta=Q^{-2}$, obtenant une inégalité d'abord prouvée par Bombieri en 1965 (avec un autre Δ)

(3.5) $$\sum_{q\leq Q} \sum_{\substack{1\leq a\leq q \\ (a,q)=1}} |S(a/q)|^2 \leq (N-1+Q^2) \sum_{n=M+1}^{M+N} |a_n|^2$$

N.B. On voit ici l'utilisation de (3.1) sans que l'on exprime exactement R: on a, en effet, $R=\sum_{a\leq Q} \varphi(a)$ qui ne s'écrit pas de façon "simple". Cependant, on prouve assez facilement (cf. **[H&W]** chap. XVIII) que $R=3Q^2/\pi^2+O(Q\ln Q)$, donc $R \ll Q^2$.

C'est l'inégalité (3.5) qui sera maintenant exploitée pour prouver les deux théorèmes suivants.

Le premier est ce qu'on appelle la forme additive du grand crible, la seule en fait à correspondre au terme de "crible". On se place dans un contexte semblable à celui indiqué dans l'introduction pour définir les cribles: ici, $E=\{M+1,...,M+N\}$, et $P=\{p \mid p\leq Q\}$. On note encore $\omega(p)$ $(0\leq\omega(p)\leq p-1)$ le nombre de classes de congruences "retirées" lors du crible, et E° l'ensemble criblé. Soit $Z=\text{Card } E^\circ$: nous allons obtenir une majoration non triviale de Z.

Théorème 6 (Montgomery 1968) *On a*

(3.6) $$Z \leq \frac{N-1+Q^2}{L} \quad \textit{avec} \quad L=\sum_{q\leq Q} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p-\omega(p)}$$

N.B. • Le terme L, on le voit, ne fait intervenir des contributions que pour les nombres sans facteur carré. Ceci s'interprète par la relative indépendance des diverses classes de congruences modulo a et b si (a,b)=1, démontrée par le théorème chinois. Il est bon de noter que ces nombres sont en proportion **positive** parmi les entiers ainsi que le montre la formule (1.12).

• Dans certaines cas, il est utile de considérer une autre approche du crible: on suppose E° **donné**, et on détermine $\omega(p)$ comme étant le nombre de classes de congruences modulo p ne contenant aucun élément de E° (ie à partir du résultat du crible, on détermine par quoi on a criblé). Il est clair que ceci est équivalent à faire un crible par ces classes de congruences.

DÉMONSTRATION

On note $J(q)=\prod_{p|q} \frac{\omega(p)}{p-\omega(p)}$, de sorte que $L=\sum_{q\leq Q} \mu^2(q)J(q)$.

Soit E° l'ensemble criblé, donc Z=Card E°.

On va prouver un résultat plus général que celui annoncé: pour toute suite finie $(a_n)_{M+1\leq n\leq M+N}$ vérifiant $a_n=0$ si $n\notin E°$, on a

(3.7) $$|S(0)|^2L \leq (N-1+Q^2) \sum_{n=M+1}^{M+N}|a_n|^2$$

Ceci entraîne effectivement (3.6): soit $a_n=\chi_{E°}(n)$, $M+1\leq n\leq M+N$. On a alors $\sum|a_n|^2=Z=S(0)$. L'application de (3.7) donne donc $Z^2L\leq(N+Q^2)Z$, d'où le résultat.

Reste à prouver (3.7): soit donc (a_n) satisfaisant à la condition annoncée. Compte tenu de l'inégalité (3.5), il suffit alors de démontrer que pour tout q≤Q, on a

(3.8) $$\sum_{\substack{a\leq q\\(a,q)=1}} |S(a/q)|^2 \geq \mu^2(q)J(q)|S(0)|^2$$

En effet, si on somme alors sur q≤Q, on trouve bien $|S(0)|^2L\leq(N-1+Q^2)\sum|a_n|^2$, c'est à dire (3.7).

Pour prouver l'inégalité (3.8), on observe tout d'abord que si $\mu(q)=0$, elle est trivialement vérifiée. Reste à traiter le cas où q est sans facteur carré. Pour cela, on va procéder par récurrence sur le nombre de facteurs premiers de q.

Tout d'abord, considérons le cas où q est premier, soit q=p.

La condition de sommation du côté gauche de (3.8), (a≤q et (a,q)=1) devient alors simplement (1≤a≤p-1). Soit $S(p,h)=\sum_{\substack{M+1\leq n\leq M+N\\ n\equiv h\ (p)}} a_n$, de sorte que $\sum_{h=1}^{p}S(p,h)=S(0)$.

On calcule alors

$$\sum_{a=1}^{p}|S(a/p)|^2 = \sum_{a=1}^{p}\left|\sum_{n=N+1}^{M+N}a_n e(an/p)\right|^2 = \sum_{a=1}^{p}\left|\sum_{h=1}^{p}S(p,h)e(ah/p)\right|^2$$

On développe alors les carrés en somme double sur h et k, et on inverse l'ordre de sommation. La somme la plus intérieure (sur a) est alors

$$\sum_{a=1}^{p} e(a(h-k)/p) = \begin{cases} p \text{ si } h-k\equiv 0 \ (p) \\ 0 \text{ sinon} \end{cases}$$

On trouve donc $\sum_{a=1}^{p}|S(a/p)|^2 = p\sum_{h=1}^{p}|S(p,h)|^2$, d'où

(3.9) $$\sum_{a=1}^{p-1}|S(a/p)|^2 = p\sum_{h=1}^{p}|S(p,h)|^2 - |S(0)|^2$$

Or $|S(0)|^2=\left|\sum_{h=1}^{p}S(p,h)\right|^2$ et il y a au moins $\omega(p)$ termes nuls dans cette somme par hypothèse, donc par Cauchy - Schwarz il vient

$$|S(0)|^2 \le (p-\omega(p))\sum_{h=1}^{p}|S(p,h)|^2$$

En insérant cela dans (3.9), on trouve $\sum_{a=1}^{p-1}|S(a/p)|^2 \ge J(p)|S(0)|^2$, ce qui prouve (3.8) dans le cas où q est premier.

Pour achever la récurrence, on observe en premier lieu que J est multiplicative. De plus, si (3.8) est vraie pour une valeur de q, son application au polynôme T construit avec $a_n e(n\beta)$ à la place de a_n , ie $T(x)=S(x+\beta)$, conduit à

$$\sum_{\substack{a\le q \\ (a,q)=1}} |S(a/q+\beta)|^2 \ge \mu^2(q)J(q)|S(\beta)|^2$$

Par conséquent, si q=pr, avec p premier et (p,r)=1, comme (3.8) est par hypothèse de récurrence vraie pour p et r, il vient

$$\sum_{\substack{a\le q \\ (a,q)=1}} |S(a/q)|^2 = \sum_{\substack{a\le p \\ (a,p)=1}} \sum_{\substack{b\le r \\ (b,r)=1}} |S(a/p+b/r)|^2 \text{ , grâce au théorème chinois}$$

$$\ge J(q) \sum_{\substack{b\le r \\ (b,r)=1}} |S(b/r)|^2 \ge J(q)J(r)|S(0)|^2 = J(s)|S(0)|^2$$

en appliquant successivement la remarque précédente à p et r.

D'après les observations faites au début de la démonstration, ceci achève la preuve du théorème 6 ♦

N.B. • Dans les applications, Q est un paramètre qu'on choisit de manière à avoir une estimation optimale, c'est à dire L le plus grand possible. L'expérience montre que ceci se produit généralement en prenant Q de l'ordre de $N^{1/2}$, ou bien légèrement inférieur (si $Q=o(N^{1/2})$, $N+Q^2=N(1+o(1))$ et on gagne pour Q grand un facteur 2).

• Il est clair au vu du théorème 6 que dans les applications à des cas "concrets", l'une des difficultés sera d'estimer assez précisément l'expression L.

• A partir de cette démonstration, il est possible de prouver un autre résultat intéressant. Grâce à $\sum_{h=1}^{p} S(p,h)=S(0)$ on trouve sans peine la relation

$$p\sum_{h=1}^{p}|S(p,h)-1/pS(0)|^2 = p\sum_{h=1}^{p}|S(p,h)|^2 -|S(0)|^2$$

Si on somme alors (3.9) sur $p\leq Q$, on trouve donc par (3.5)

(3.10)
$$\sum_{p\leq Q} p\sum_{h=1}^{p}|S(p,h)-1/pS(0)|^2 \leq (N-1+Q^2)\sum_{n=M+1}^{M+N}|a_n|^2$$

Ce résultat est moins précis que (3.6) parce qu'on ne somme que sur les nombres premiers. Il est cependant souvent utile. Il a de plus une certaine interprétation probabiliste: $\frac{S(0)}{p}$ est en effet la moyenne de S(p,h), pour $1\leq h\leq p$, et (3.10) majore donc la somme des variances des S(p,h). On peut ainsi vérifier que pour "la plupart" des nombres premiers $p\leq Q$, les éléments criblés sont "bien" distribués dans les classes de congruences modulo p. C'est Rényi qui a abordé le premier cet aspect probabiliste du grand crible, qui fut également développé par Roth.

• Le théorème 6 ne fournit qu'une majoration du nombre d'éléments non criblés. Il est évident qu'on préfèrerait obtenir aussi une minoration, voire même un équivalent pour $N\to\infty$ de Z. Le grand crible ne peut fournir un tel résultat, ce qui peut se comprendre en notant qu'on utilise en fait, pour obtenir le théorème 6, que peu d'information sur la nature arithmétique du crible effectué: seul $\omega(p)$ intervient. Cependant une minoration peut aisément être obtenue pour (3.2), en utilisant une fonction minorant χ_E au lieu d'une fonction majorante. Malheureusement, il n'y a plus de correspondance entre cette inégalité et celle qui lui correspond par dualité. On peut avec d'autres cribles se rapprocher de cela. En particulier, le crible de Selberg permet d'obtenir une minoration, toutefois plus complexe et moins exploitable que la majoration. Notons également que dans les cas où l'équivalent est connu (par exemple

inégalité de Brun-Titchmarsh), il s'avère que la majoration fournie par le grand crible est à une constante multiplicative près de l'ordre de grandeur exact.

Le second aspect du grand crible est sa forme multiplicative, dont la base est contenue dans le théorème 7. Pour tout complexes $(a_n)_{M+1\le n\le M+N}$, on note encore S le polynôme trigonométrique introduit au début, et de plus on note pour tout q et tout caractère de Dirichlet χ modulo q

$$T(\chi) = \sum_{n=M+1}^{M+N} a_n \chi(n)$$

Théorème 7 (Gallagher 1967): *Soit* $(a_n)_{M+1\le n\le M+N}$ *des complexes quelconques. On a alors*

(3.11)
$$\sum_{q\le Q} \frac{q}{\varphi(q)} \sideset{}{'}\sum_{\chi} |T(\chi)|^2 \le (N-1+Q^2) \sum_{n=M+1}^{M+N} |a_n|^2$$

DÉMONSTRATION

Encore une fois, compte tenu de (3.5), il suffit de montrer

$$\sideset{}{'}\sum_{\chi} |T(\chi)|^2 \le \frac{\varphi(q)}{q} \sum_{\substack{a\le q\\(a,q)=1}} |S(a/q)|^2$$

pour q≤Q, et ensuite de sommer pour obtenir (3.11).

Soit χ un caractère de Dirichlet primitif modulo q. A l'aide de l'identité (1.14) on trouve

$$\tau(\bar\chi)T(\chi) = \sum_{a=0}^{q-1} \bar\chi(a) S(a/q)$$

On prend maintenant le carré du module des deux côtés, on utilise (1.15) et on somme sur tout les caractères primitifs modulo q, d'où

$$q \sideset{}{'}\sum_{\chi} |T(\chi)|^2 = \sideset{}{'}\sum_{\chi} \left| \sum_{a=0}^{q-1} \bar\chi(a) S(a/q) \right|^2 \le \sum_{\chi} \left| \sum_{a=0}^{q-1} \bar\chi(a) S(a/q) \right|^2$$

On procède alors comme d'habitude: on développe les modules et on somme sur χ en premier: à cause des relations d'orthogonalité (1.13) et de $\chi(a)=0$ si $(a,q)\neq 1$, il vient

$$q \sideset{}{'}\sum_{\chi} |T(\chi)|^2 \le \varphi(q) \sum_{\substack{a\le q\\(a,q)=1}} |S(a/q)|^2$$

ce qui permet de conclure ♦

IV. Applications élémentaires du grand crible

Cette partie contient quelques unes des plus simples applications de la forme additive du grand crible, qui permettent toutefois d'apprécier l'efficacité de cette méthode. Les applications de la forme multiplicative, bien que très importantes, sont à la fois plus complexes et moins directes, elles font aussi appel à d'autres résultats difficiles, aussi ne pourront-elles pas être abordées ici.

1. L'inégalité de Brun - Titchmarsh

Le problème du comportement asymptotique de la suite des nombres premiers fut énoncé d'abord par Legendre et Gauss vers 1780. En 1896, conformément à leurs conjectures, Hadamard et de la Vallée Poussin démontrèrent indépendamment que $\pi(x,a,b) \sim \frac{1}{\varphi(a)} \frac{x}{\ln x}$, ou plus précisément $\pi(x,a,b) \sim \frac{li(x)}{\varphi(a)}$, li dénotant le logarithme intégral $li(x)=\int_2^{\infty} \frac{dt}{\ln t}$ (li(x) s'avère être une bien meilleure approximation de π que $\frac{x}{\ln x}$, toutefois, $li(x) \sim \frac{x}{\ln x}$).

Assez naturellement, et en relation avec les applications de ces résultats, se pose le problème d'obtenir des estimations **uniformes** en a, c'est à dire où les constantes implicites sont les mêmes soit pour tout a, soit pour a dans un ensemble assez vaste. Le meilleur résultat actuel (1938 cependant) est le théorème de Siegel - Walfisz, conséquence d'un résultat profond sur les zéros des fonctions L, qui assure que l'équivalence ci-dessus est vraie uniformément pour $a \leq (\ln x)^H$. Ici, H peut être un réel positif quelconque, mais pour des valeurs de H différentes on a évidemment des constantes différentes dans les O(...).

Le théorème suivant donne un résultat plus faible, mais plus uniforme en a et valide sur tout intervalle.

Théorème 8 (inégalité de Brun - Titchmarsh): *pour tout* $\varepsilon>0$, *il existe* $\delta>0$ *tel que pour tout* $(x,y)\in\mathbb{R}^{+2}$ *on a*

(4.1) $$\pi(x+y;a,b)-\pi(x;a,b) \leq (2+\varepsilon)\frac{y}{\varphi(q)\ln \frac{y}{a}}$$

uniformément pour $a<\delta y$, *et tout b tel que* $(a,b)=1$.

DÉMONSTRATION

C'est une application directe du théorème 6. Soient a et b tels que (a,b)=1.

On choisit de cribler l'ensemble E des n tels que $x<an+b\leq x+y$ (donc un intervalle de $\mathbb{N}$ de longueur au plus $\frac{y}{a}+1$), par $\{p \mid p\leq Q\}$. Pour déterminer les classes de congruences à retirer, considérons un nombre premier $p=an+b$, avec $n\in E$. Si $p>Q$, il est clair que pour tout $p'\leq Q$, on a $p'\nmid an+b$, donc $an+b\not\equiv 0\ (p')$. Si a est inversible modulo p', c'est à dire si p' ne divise pas a, cette condition équivaut à exclure une classe de congruence. Nous choisirons donc $\omega(p')=1$ si $p'\nmid a$ et $\omega(p')=0$ sinon. De cette façon, l'ensemble criblé contient donc les nombres premiers de la forme an+b compris entre x et x+y **et qui sont >Q** (mais pas forcément seulement ceux-ci). Pour avoir donc au moins tout ces nombres premiers, il faut tenir compte de ceux qui éventuellement seraient ≤Q et auraient été éliminés lors du crible.

Mais on peut en tout cas écrire $\pi(x+y;a,b)-\pi(x;a,b)\leq Z+Q$, en estimant de la façon la plus grossière qui soit l'importance du second cas. Par le théorème 6 il vient donc

$$\pi(x+y;a,b)-\pi(x;a,b) \leq \frac{N-1+Q^2}{L} + Q$$

pour $L = \sum_{q\leq Q} \mu^2(q) \prod_{p|q} \frac{\omega(p)}{p-\omega(p)}$ et $N \leq \frac{y}{a}+1$, et avec les valeurs de $\omega(p)$ dans ce cas, on trouve

$$L = \sum_{\substack{q\leq Q \\ (q,a)=1}} \mu^2(q) \prod_{p|q} \frac{1}{p-1}$$

(puisque si $(q,a)>1$, il existe $p\leq Q$ tel que $p|q$ et $p|a$, donc $\omega(p)=0$...), et par conséquent grâce à (1.11)

$$L= \sum_{\substack{q\leq Q \\ (a,q)=1}} \frac{\mu^2(q)}{\varphi(q)}$$

Il nous faut donc minorer cette quantité. Pour cela nous observons que pour tout n on a

$$\frac{n}{\varphi(n)} = \prod_{p|n} \frac{1}{1-\frac{1}{p}} = \sum_{d|n^\infty} \frac{1}{d}$$

en développant $(1-p^{-1})^{-1}=\sum_{n\geq 0} p^{-n}$.

Donc

$$\frac{a}{\varphi(a)} L = \sum_{\substack{q\leq Q \\ (a,q)=1}} \frac{\mu^2(q)}{q} \frac{q}{\varphi(q)} \frac{a}{\varphi(a)} = \sum_{\substack{q\leq Q \\ (a,q)=1}} \frac{\mu^2(q)}{q} \sum_{d|q^\infty} \frac{1}{d} \sum_{e|a^\infty} \frac{1}{e}$$

Soit un $n \leq Q$. On peut écrire $n=bm$, avec $b|a^\infty$ et $(a,m)=1$, et ensuite $m=qm_1$, avec $\mu(q)\neq 0$ et $m_1|q^\infty$. Par conséquent, $n=bqm_1$ intervient dans la somme obtenue en développant l'expression ci-dessus. On en déduit

$$\frac{a}{\varphi(a)} L \geq \sum_{n\leq Q} \frac{1}{n} \geq \ln Q$$

la dernière inégalité étant classique.

L'inégalité obtenue grâce au crible implique donc

$$\pi(x+y,a,b)-\pi(x;a,b) \leq \frac{y+aQ^2}{\varphi(a)\ln Q} + Q$$

Reste seulement à choisir Q. On prend ici $Q=\sqrt{y/a}\,\ln (y/a)^{-1}$. Par conséquent

$$y+aQ^2=y(1+(\ln y/a)^{-2})$$

$$\varphi(a)\ln Q=\frac{1}{2}\varphi(a)\ln y/a\left(1 - \frac{2\ln\ln y/a}{\ln y/a}\right)$$

Comme de plus $Q=o(\sqrt{y/a})$, on obtient donc le résultat voulu en faisant $y/a\to\infty$

(Le choix de $Q=\sqrt{y/a}$ qui peut paraître plus naturel donne le résultat avec un 4 à la place du 2) ♦

N.B. • Il est clair que l'on peut obtenir une expression plus précise du reste en le développant. Mais ce travail est inutile: en effet, en utilisant la variante du grand crible donnée par (3.4), Montgomery et Vaughan (1973) ont prouvé que

$$\pi(x+y;a,b)-\pi(x;a,b) \leq \frac{2y}{\varphi(q)\ln \frac{y}{a}}$$

à la seule condition que $a<y$.

• Il existe de nombreuses versions de l'inégalité de Brun - Titchmarsh qui, comme son nom l'indique, fut d'abord prouvée à l'aide du crible de Brun. Le résultat de Montgomery et Vaughan indiqué ci-dessus est le meilleur connu à l'heure actuelle, toutes méthodes confondues.

• Le gros avantage de l'inégalité de Brun - Titchmarsh par rapport au théorème de Siegel - Walfisz réside dans sa validité uniforme sur une étendue beaucoup plus grande: $a<y$ au lieu de $a\leq(\ln y)^H$.

• On observera ici un exemple frappant d'application efficace du grand crible dans un contexte de petit crible, puisque $\omega(p)\leq 1$ pour tout p.

Voici maintenant un exemple d'application de l'inégalité de Brun - Titchmarsh, qui est aussi un exemple du grand crible au sens premier du terme: $\omega(p) \to \infty$. Une conjecture d'Artin énonce que tout entier positif qui n'est pas un carré est une racine primitive modulo p (ie engendre le groupe multiplicatif des éléments inversibles modulo p) pour une infinité de nombres premiers p. Notons que l'on sait (Gauss) que pour tout nombre premier p, il existe $\varphi(p-1)$ racines primitives modulo p. L'exclusion des carrés se base sur le fait que pour un entier n qui est carré modulo p, $n^{(p-1)/2}\equiv 1$ (critère d'Euler).

Le grand crible permet de prouver aisément le résultat suivant

Théorème 9 (Gallagher 1967): *dans tout intervalle de longueur* N, *le nombre d'entiers qui ne sont racine primitive pour aucun* $p\leq Q$ *est* $\ll (N+Q^2)\frac{\ln Q}{Q}$.

DÉMONSTRATION

L'énoncé du théorème indique clairement la marche à suivre. On crible $\{M+1,...,M+N\}$ par les $\varphi(p-1)$ classes des racines primitives modulo p pour tout $p\leq Q$, et le cardinal de l'ensemble criblé est exactement le nombre recherché, noté Z(Q).

D'après le théorème 6, on a $Z(Q)\leq\frac{N-1+Q^2}{L}$.

Le terme L étant ici fort difficile à estimer, nous utiliserons l'inégalité (médiocre)

$L\geq \sum_{p\leq Q} \frac{\varphi(p-1)}{p-\varphi(p-1)} \geq \sum_{p\leq Q} \frac{\varphi(p-1)}{p-1} = L'$. Ceci revient en gros à employer (3.10): on ne garde que les termes de L pour q premier.

Maintenant, par Cauchy - Schwarz, $\pi(Q)^2\leq L' \sum_{p\leq Q} \frac{p-1}{\varphi(p-1)}$ et on va majorer cette dernière somme.

L'identité $n=\sum_{d|n}\varphi(d)$, et la multiplicativité de φ, entraînent pour tout entier n $\frac{n}{\varphi(n)} = \sum_{d|n} \frac{1}{\varphi(d)}$

(en effet, grâce à (1.11), $\frac{n}{\varphi(n)}$ ne dépend que de $\prod_{p|n} p$ et donc on peut supposer n sans facteur carré et employer la multiplicativité de φ car d et n/d sont toujours premiers entre eux)

Par conséquent $L'=\sum_{p\le Q}\sum_{d|p-1}\frac{1}{\varphi(d)}=\sum_{d\le Q}\frac{1}{\varphi(d)}\pi(Q;d,1)$

On sépare alors la somme en deux parties: pour $d\le Q^{1/2}$, on a $\pi(Q;d,1)\ll\frac{Q}{\varphi(d)\ln Q}$, et pour $Q\ge d>Q^{1/2}$, trivialement $\pi(Q;d,1)\le\frac{Q}{d}$. D'où

$$L'\ll\frac{Q}{\ln Q}\sum_{d\le Q^{1/2}}\frac{1}{\varphi(d)^2}+Q\sum_{Q\ge d>Q^{1/2}}\frac{1}{d\varphi(d)}$$

Or, les deux séries $\sum_n\frac{1}{\varphi(n)^2}$ et $\sum_n\frac{1}{n\varphi(n)}$ sont convergentes à cause de l'estimation (cf. **[H&W]** chap. XXII)

$$\varphi(n)\ge n\prod_{p\le n}(1-p^{-1})\gg\frac{n}{\ln n}$$

De plus, on vérifie aisément que $\sum_{n>Q}\frac{1}{n\varphi(n)}\ll\frac{1+\ln Q}{Q}$, et on obtient donc

$$L'\ll\frac{Q}{\ln Q}$$

en majorant la première somme partielle par celle de la série et en étendant la seconde à $Q\to\infty$.

Finalement, $L'\gg(\pi(Q))^2\frac{\ln Q}{Q}\gg\frac{Q}{\ln Q}$, puisque $\pi(Q)\sim\frac{Q}{\ln Q}$ d'après le théorème des nombres premiers.

Le résultat obtenu en définitive est donc bien $Z(Q)\ll(N+Q^2)\frac{\ln Q}{Q}$ ♦

N.B. • Si on considère l'intervalle $\{1,...,N\}$ et $Q=N^{1/2}$, on trouve que le nombre d'entiers qui ne sont racine primitive modulo p pour aucun $p\le N^{1/2}$ est $\ll N^{1/2}\ln N$, alors même qu'il y a environ $N^{1/2}$ carrés qui ne peuvent être des racines primitives pour aucun p.

• On voit là que bien qu'on ait employé (3.10) au lieu de (3.5), on a obtenu un résultat intéressant. Ceci est propre au fait que l'on soit dans un contexte de grand crible: utiliser (3.10) pour majorer $\pi(x;a,b)$ ne donne qu'une estimation très médiocre.

2. Les nombres premiers jumeaux

Nous avons déjà indiqué la conjecture des nombres premiers jumeaux: il existe une infinité de nombre premiers p tels que p+2 soit aussi premier.

A l'aide du grand crible, nous allons prouver le théorème de Brun.

Théorème 10 (Brun 1919): *la somme des inverses des nombres premiers jumeaux converge.*

En fait, comme c'est souvent le cas avec les cribles, nous allons en fait prouver un résultat sensiblement plus précis.

Théorème 11: *soit J(x) le nombre de nombres premiers p≤x tels que p+2 soit aussi premier. On a l'estimation*

(4.2) $$J(x)\ll\frac{x}{(\ln x)^2}$$

DÉMONSTRATION

C'est encore une application du théorème 6. On crible l'ensemble des entiers inférieurs à x par {p |p≤Q}, en enlevant les nombres congrus à 0 ou -2. Alors les nombres n>Q tels que n et n+2 soient premiers ne sont pas criblés. En effet, soit un tel n: on ne peut avoir ni n≡0 (p) pour un p≤Q, ni n+2≡0.

Par conséquent avec ce crible on a

$$J(x)\le Z+J(Q)\le Z+Q$$

en estimant encore très grossièrement J(Q).

Pour majorer Z, on note que ω(2)=1 (puisque 0=-2 dans $\mathbb{Z}/2\mathbb{Z}$) et ω(p)=2 si p≥3. D'après le théorème 6 il faut donc minorer la quantité L correspondante.

Nous allons ici, en vue du relativement simple théorème 11, le faire de façon peu élégante mais efficace. On vérifie aisément que

$$2L\ge\sum_{q\le Q}\mu^2(q)\prod_{p|q}\frac{2}{p-1}=L'$$

et L' s'écrit encore

$$L'=\sum_{q\le Q}\mu^2(q)\frac{2^{\nu(q)}}{\varphi(q)}$$

que l'on développe de façon similaire à ce qui a été fait dans la preuve de l'inégalité de Brun - Titchmarsh:

$$L' = \sum_{q \le Q} \frac{\mu^2(q)}{q} 2^{\nu(q)} \sum_{d | q^\infty} \frac{1}{d} \ge \sum_{n \le Q} \frac{2^{\nu(n)}}{n}$$

Maintenant, notons pour tout x réel, $N(x) = \sum_{n \le x} 2^{\nu(n)}$, et évaluons N(x) à l'aide de l'identité $2^{\nu(n)} = \sum_{d|n} \mu^2(d)$ et de (1.12)

$$N(x) = \sum_{n \le x} \sum_{d|n} \mu^2(d) = \sum_{d \le x} \mu^2(d) \left[\frac{x}{d}\right] = x \sum_{d \le x} \frac{\mu^2(d)}{d} + O(x)$$

et l'application de la formule d'Abel (1.16) avec f(t)=1/t et de (1.12) à la dernière somme donne finalement

$$N(x) \sim \frac{6}{\pi^2} x \ln x \text{ pour } x \to \infty$$

Appliquons encore la formule d'Abel (1.16) pour évaluer $\sum_{n \le Q} \frac{2^{\nu(n)}}{n}$:

$$\sum_{n \le Q} \frac{2^{\nu(n)}}{n} = \frac{N(Q)}{Q} + \int_1^Q \frac{N(t)}{t^2} dt = O(\ln Q) + \frac{3}{\pi^2} (\ln Q)^2 + O((\ln Q)^2) \sim \frac{3}{\pi^2} (\ln Q)^2$$

Le grand crible donne par conséquent le résultat $J(x) \ll \frac{N+Q^2}{(\ln Q)^2} + Q$, et c'est encore avec $Q = N^{1/2}$ que l'on obtient exactement le théorème 11 ♦

Voyons maintenant comment en déduire simplement le théorème de Brun. Si il n'existe qu'un nombre fini de nombres premiers jumeaux, le résultat est acquis. Sinon, soit j_n le n-ème nombre premier p tel que p+2 soit aussi premier ($j_1=3$, $j_2=5$, $j_3=11$...). On a $J(j_n)=n$, donc le théorème 11 entraîne

$$n \ll \frac{j_n}{(\ln j_n)^2}$$

or $j_n > n$ de façon évidente, et donc

$$\frac{1}{j_n} \ll \frac{1}{n(\ln n)^2}$$

ce qui prouve que $\sum_n \frac{1}{j_n}$ converge.

N.B. • Il ne serait pas difficile d'adapter la démonstration du théorème 10 pour obtenir une constante explicite du O(...). Cependant, ceci ne serait pas très intéressant compte tenu du fait que l'on a évalué en fait L'. Par contre, en calculant sur L même et à l'aide de manipulations un peu plus complexes, on peut aboutir à une estimation très instructive

$$J(x)\leq(8C+o(1))\frac{x}{(\ln x)^2}$$

où C est la constante de Shah et Wilson, $C=2\prod_{p\geq 3}(1-(p-1)^{-2}) = 1,3203236...$ L'intérêt de cette forme de majoration réside dans le fait que divers raisonnements heuristiques ou probabilistes (cf. par exemple **[H&W]** fin chap. XXII) conduisent à la conjecture que $J(x)\sim C\frac{x}{(\ln x)^2}$. Cette conjecture correspond bien aux données extraites des tables de nombres premiers (ce qui bien sur ne prouve rien).

• Numériquement, Fröberg à montré que $\sum_n(j_n^{-1}+(j_n+2)^{-1}) =1,70195+\varepsilon$, avec $|\varepsilon|<10^{-5}$.

• Par rapport à la conjecture des nombres premiers jumeaux, le théorème de Brun est plutôt négatif. Cependant, et également par des méthodes de cribles - pas exclusivement de grand crible - Chen (1973) a prouvé qu'il existe une infinité de nombres premiers p tels que p+2 ait au plus deux facteurs premiers.

3. Autres applications du grand crible

Ce paragraphe est consacré à l'énoncé d'autres applications du grand crible, en particulier des théorèmes les plus profonds pour lesquels il n'est souvent qu'un outil parmi d'autres, mais essentiel.

La forme multiplicative du grand crible est en fait plus "profonde" et plus puissante que la forme additive. En effet, elle est essentielle dans l'étude des fonctions L de Dirichlet - définies par $L(s,\chi)=\sum_{n=1}^{\infty}\frac{\chi(n)}{n^s}$ pour tout caractère de Dirichlet χ -, et de la distribution de leurs zéros, sujet qui est lui-même au coeur de la partie analyse complexe de la théorie analytique des nombres. En particulier, (3.10) est fondamental pour démontrer le très important théorème de Bombieri - Vinogradov

$$\sum_{q\leq Q}\max_{y\leq x}\max_{\substack{a\\(a,q)=1}}\left|\pi(y;q,a)-\frac{li(y)}{\varphi(q)}\right| \ll_A x(\ln x)^{-A}$$

valable uniformément pour $Q \leq x^{1/2}(\ln x)^{-B}$, avec B=B(A). L'utilité de ce théorème réside dans les nombreuse applications où il s'avère pouvoir servir de substitut à la célèbre Hypothèse de Riemann Généralisée selon laquelle tout les zéros des fonctions L ont une partie réelle égale à 1/2.

Concernant les deux conjectures énoncées au début de la troisième partie, de considérables progrès ont été effectués: parallèlement à son théorème sur les nombres premiers jumeaux, Chen a prouvé que tout entier pair assez grand est somme d'un nombre premier et d'un nombre ayant au plus deux facteurs premiers. De plus, si G(x) désigne le nombre d'entiers pairs inférieurs à x qui ne sont pas somme de deux nombres premiers, Montgomery et Vaughan (1975) ont montré qu'il existe un $\delta>0$ tel que $G(x)=O(x^{1-\delta})$. En particulier, G(x)=o(x), c'est à dire que "presque tout" les nombres pairs sont somme de deux nombres premiers. De plus, tout nombre pair est somme d'au plus 26 nombres premiers (Vaughan 1977). Un autre résultat important - mais prouvé par d'autres méthodes - est le théorème de Vinogradov: tout nombre impair assez grand est somme de trois nombres premiers.

Conclusion

La méthode du grand crible, ainsi que l'observait déjà Montgomery en 1978 dans son article d'exposition **[Mon]**, n'est plus de nos jours une méthode "profonde" et difficile en théorie des nombres, contrairement à la situation prévalant dans les premières années de son développement, durant lesquelles seuls quelques spécialistes pouvaient maîtriser les grandes difficultés techniques alors présentes dans son étude. Les travaux effectués entre 1965 et 1975 ont permis de comprendre la nature analytique du grand crible et d'en explorer les limites, qui sont maintenant bien connues. Ainsi, s'il n'est plus possible d'attendre des miracles du grand crible seul, cette méthode n'en reste pas moins un moyen efficace pour explorer de très nombreux problèmes de théorie des nombres, et par ses généralisations l'inégalité (3.1) apparaît toujours comme d'une grande utilité.

D'un point de vue personnel, le bilan que je tire de ce stage est très positif. D'une part il m'a été possible de m'initier aux méthodes de cribles, et cela de la façon la plus agréable qu'il soit puisque le grand crible se distingue en étant à la fois relativement simple et efficace, et d'autre part les divers documents que j'ai étudié de façon plus ou moins approfondie m'ont permis d'améliorer grandement mes connaissances des techniques même les plus élémentaires en théorie des nombres (comment aborder une somme à estimer...). De surcroît, il n'y avait pas de grosse difficulté théorique, et le vocabulaire ainsi que les diverses notations m'étaient souvent déjà sinon familières du moins pas totalement inconnues, ce qui m'a autorisé à me consacrer d'emblée au sujet lui-même.

Je dirai également que la possibilité, propre au sujet, de le prendre à son tout début - démonstration de l'inégalité du grand crible -, et de le conduire jusqu'à des applications non triviales et intéressantes - inégalité de Brun - Titchmarsh, théorème de Brun -, cela sans avoir à admettre des théorèmes importants, m'a fortement motivé, en me donnant au moins l'illusion de comprendre les tenants et les aboutissants du problème.

Références

[Bom] E. Bombieri, "*Le grand crible dans la théorie analytique des nombres* ", Astérisque 18, Société Mathématique de France (1974)

[GV] S.W Graham & Jeffrey D. Vaaler, "*A class of extremal functions for the Fourier transform* ", Transactions of the American Mathematical Society **265** (1981), 283-302

[H&W] G.H. Hardy & E.M. Wright, "*An introduction to the theory of numbers* ", 5ème édition, Oxford University Press (1979)

[Mon] Hugh L. Montgomery, "*The analytic principle of the large sieve* ", Bulletin of the American Mathematical Society **84** (1978) 547-567

[Sch] Laurent Schwartz, "*Théorie des distributions* ", 2ème édition, Hermann 1966

[Val] Jeffrey D. Vaaler, "*Extremal functions in Fourier analysis* ", Bulletin of the American Mathematical Society, Nouvelle Série **12** (1985), 183-216

N.B. • Il y a peu de littérature en français sur le grand crible à part le petit livre de Bombieri. Celui-ci met assez largement l'accent sur la forme multiplicative du grand crible, en particulier pour les applications.

• Pour une bibliographie beaucoup plus complète, voir **[Mon]**.