

**Trois conférences  
sur la théorie analytique  
et probabiliste des nombres**

Gérald Tenenbaum

*Journées État de la Recherche  
7-9 décembre 2000  
Université Bordeaux 1*



# Convolutions et équations fonctionnelles : sur la preuve élémentaire de Daboussi

Gérald Tenenbaum

## 1. Introduction

Mystérieuse pendant des millénaires, la répartition des nombres premiers demeure un sujet d'étude actuel, sous-tendu par de redoutables questions méthodologiques, voire philosophiques. Conjecturé pendant plus d'un siècle, le théorème des nombres premiers, soit

$$\pi(x) \sim x/\log x \quad (x \rightarrow \infty),$$

est resté source d'interrogations après ses premières démonstrations par Hadamard et La Vallée-Poussin, en 1896. En effet, le remarquable succès, sur ce problème, des méthodes de l'analyse complexe semblait leur conférer une supériorité intrinsèque sur celles de l'arithmétique élémentaire, auquel le problème, et même l'énoncé de la solution, se rattachent.

La preuve analytique repose de manière essentielle sur l'absence de zéro de la fonction zêta de Riemann sur la droite  $\sigma = 1$ . En outre, les *théorèmes taubériens* développés par Wiener au début des années 1930, et notamment le célèbre théorème d'Ikehara (1931), fournissent un cadre général dans lequel le théorème des nombres premiers et l'assertion sur les zéros de  $\zeta(s)$  apparaissent de manière éclatante comme des propositions équivalentes. Cela a induit l'idée qu'en un certain sens la théorie des fonctions analytiques était plus « profonde » que l'analyse réelle et qu'il était hautement improbable, pour ne pas dire impossible, de parvenir au théorème des nombres premiers par de simples manipulations d'inégalités.

Ainsi, le théorème des nombres premiers apparaissait comme irréductible, pour sa démonstration, au cadre naturel dans lequel on pouvait l'énoncer — une frustration qui engendrait bien des spéculations méta-mathématiques !

En 1921, Hardy exprimait ainsi, lors d'une communication à la société mathématique de Copenhague, l'opinion suivante : « *Aucune preuve élémentaire du théorème des nombres premiers n'est connue et l'on peut se demander s'il est raisonnable d'en attendre une. Nous savons maintenant que le théorème est essentiellement équivalent à un théorème concernant une fonction analytique, à savoir que la fonction zêta de Riemann ne possède pas de zéro sur une certaine droite. Une démonstration d'un tel résultat fondamentalement indépendante des idées de la théorie des fonctions me semble extraordinairement improbable. Il est hasardeux d'affirmer qu'un théorème mathématique ne peut être démontré d'une façon ou d'une autre, mais un point paraît acquis. Nous avons un certain regard sur la logique de la théorie ; nous pensons que certains théorèmes, sont, comme on dit, plus profonds, alors que d'autres gisent plus près de la surface. Quiconque produirait une preuve élémentaire du théorème des nombres premiers démontrerait par là même que ces*

vues sont fausses, que le sujet n'est pas agencé de la manière dont nous le supposons et qu'il est temps de se débarrasser des livres et de réécrire la théorie.»

Ce fut donc un grand choc lorsqu'en 1949 Erdős et Selberg produisirent une preuve élémentaire,<sup>(1)</sup> mais certainement astucieuse et nullement facile, du théorème des nombres premiers — rendant *ipso facto* caduques toutes considérations sur les profondeurs relatives des méthodes de l'analyse complexe et de l'analyse réelle. L'outil essentiel est ici la formule asymptotique

$$\sum_{p \leq x} (\log p)^2 + \sum_{pq \leq x} \log p \log q = 2x \log x + O(x),$$

aujourd'hui célèbre sous le nom d'*identité de Selberg*. Cette relation peut être vue comme une sorte de formule de Stirling de degré supérieur : au lieu d'employer la fonction de von Mangoldt  $\Lambda(n)$  qui vérifie  $\sum_{d|n} \Lambda(d) = \log n$ , on introduit une nouvelle fonction  $\Lambda_2(n)$  satisfaisant à  $\sum_{d|n} \Lambda_2(d) = (\log n)^2$  et l'on évalue sa fonction sommatoire à partir d'un calcul approché élémentaire (et facile!) de  $\sum_{n \leq x} (\log n)^2$ . La procédure est semblable à celle employée par Tchébychev pour estimer (sans toutefois parvenir à une formule asymptotique) la valeur moyenne de  $\Lambda(n)$  à partir d'une forme faible de la formule de Stirling.

Voyons rapidement les détails de l'approche de Tchébychev. Nous supposons connue la définition de la convolution de Dirichlet et nous notons  $\mathbf{1}$  la fonction arithmétique constante égale à 1,  $\mu$  la fonction de Möbius, et  $\delta$  la fonction qui vaut 1 en  $n = 1$  et 0 pour tout  $n > 1$ . On a  $\log = \Lambda * \mathbf{1}$ , donc

$$(1) \quad \sum_{n \leq x} \log n = \sum_{d \leq x} \Lambda(d) \left\lfloor \frac{x}{d} \right\rfloor.$$

Le membre de gauche vaut  $x \log x - x + O(x)$ . En introduisant

$$\chi(u) := [u] - 2[u/2] = \begin{cases} 0 & (0 \leq u < 1), \\ 1 & (1 \leq u < 2), \end{cases}$$

on obtient

$$\sum_{n \leq x} \chi(x/d) \Lambda(d) = x \log 2 + O(\log x),$$

d'où l'on déduit facilement un encadrement pour la fonction de Tchébychev

$$\psi(x) := \sum_{n \leq x} \Lambda(n),$$

soit

$$x \ll \psi(x) \ll x \quad (x \geq 2).$$

En reportant dans (1), on obtient en particulier

$$(2) \quad \sum_{d \leq x} \frac{\Lambda(d)}{d} = \log x + O(1) \quad (x \geq 2).$$

Dans le cas de  $\Lambda_2$ , on montre aisément que

$$(3) \quad \Lambda_2(n) = \Lambda * \Lambda(n) + \Lambda(n) \log n = \begin{cases} (2j-1)(\log p)^2 & (n = p^j), \\ 2 \log p \log q & (n = p^j q^k), \\ 0 & (\omega(n) \neq 1, 2). \end{cases}^{(2)}$$

---

1. Au sens où elle n'utilise pratiquement aucun outil sophistiqué d'analyse à l'exception des propriétés du logarithme.

Un calcul standard, reposant sur la méthode de l'hyperbole de Dirichlet, permet par ailleurs d'établir que, pour des constantes convenables  $a$  et  $b$ ,

$$\sum_{n \leq x} \mathbf{1} * \mathbf{1} * \mathbf{1}(n) = \frac{1}{2}x(\log x)^2 + ax \log x + bx + O(x^{3/4}),$$

ce que nous pouvons réécrire sous la forme

$$H(x) := \sum_{n \leq x} h(n) \ll x^{3/4}$$

avec  $h := 2(\mathbf{1} * \mathbf{1} * \mathbf{1}) + A(\mathbf{1} * \mathbf{1}) + B\mathbf{1} - (\log)^2$  et où  $A$  et  $B$  sont des constantes adéquates. Or, en introduisant la fonction  $\mu$  de Möbius et en faisant appel à l'identité de convolution classique  $\mathbf{1} * \mu = \delta$ , nous pouvons écrire

$$\Lambda_2 = \mu * (\log)^2 = 2(\mathbf{1} * \mathbf{1}) + A\mathbf{1} + B\delta - h * \mu.$$

Donc

$$\begin{aligned} \sum_{n \leq x} \Lambda_2(n) &= 2 \sum_{n \leq x} \left\lfloor \frac{x}{n} \right\rfloor + Ax + B + \sum_{d \leq x} \mu(d)H(x/d) \\ (4) \qquad \qquad &= 2x \log x + O\left(x + \sum_{d \leq x} (x/d)^{3/4}\right) \\ &= 2x \log x + O(x), \end{aligned}$$

Le gain (capital) dû à l'introduction de l'exposant 2 est que la technique produit ici directement une formule asymptotique.

En utilisant l'identité de Selberg, Erdős a prouvé élémentairement que le rapport  $p_{n+1}/p_n$  de deux nombres premiers consécutifs tend vers 1. Il a même établi que, pour tout  $\delta > 0$  et  $x > x_0(\delta)$  il existe entre  $x$  et  $x + \delta x$  au moins  $c(\delta)x/\log x$  nombres premiers, où  $c(\delta)$  est une constante positive ne dépendant que de  $\delta$ . Cela fournissait le pendant local de la régularité globale mise en évidence par la formule de Selberg, et il n'est guère surprenant, du moins *a posteriori*, qu'une preuve élémentaire du théorème des nombres premiers ait pu résulter de la conjonction de ces deux informations : Selberg effectua la liaison finale deux jours seulement après qu'Erdős lui ait communiqué sa preuve. Quelque temps plus tard, les deux mathématiciens simplifièrent ensemble l'argument. La nouvelle preuve n'utilisait plus directement le résultat d'Erdős mais reposait fondamentalement sur les mêmes idées.

Ces idées sont présentes, sous une forme ou une autre, dans toutes les démonstrations élémentaires du théorème des nombres premiers. On peut dire *grosso modo* qu'il s'agit de mettre en évidence des équations ou des inéquations fonctionnelles. La signification profonde d'une équation/inéquation fonctionnelle est celle d'une tendance lourde à la régularité : chaque solution est « attirée » vers une des solutions fondamentales dont elle adopte le comportement asymptotique. La technique adaptée consiste à montrer que des estimations préliminaires (souvent grossières) sont suffisantes pour imposer à la solution étudiée d'appartenir au domaine d'attraction requis.

Nous nous proposons ici de donner une brève description de la preuve initiale de Selberg et une description plus complète de la démonstration élémentaire de Daboussi (1984). Nous verrons en particulier que les détails sont sinon plus simples, certainement plus naturels dans le second cas. Ils constituent non seulement un prototype exemplaire d'utilisation des équations fonctionnelles en théorie élémentaire des nombres mais le fondement d'une véritable méthode où apparaissent en pleine lumière certaines des idées actuelles de la discipline : convolutions de fonctions arithmétiques, crible, solutions d'équations différentielles aux différences.

## 2. Preuve d'Erdős–Selberg

Posons  $R(x) := \psi(x) - x$ . Nous devons montrer que

$$(5) \quad R(x) = o(x) \quad (x \rightarrow \infty).$$

De (3), nous déduisons que

$$\sum_{n \leq x} \Lambda(n) \log n = \sum_{n \leq x} \Lambda_2(n) - \sum_{n \leq x} \Lambda * \Lambda(n).$$

En insérant (4), il s'ensuit que

$$\sum_{n \leq x} \Lambda(n) \log n = 2x \log x - \sum_{n \leq x} \Lambda(n) \psi(x/n) + O(x).$$

dans le membre de gauche, nous pouvons remplacer  $\log n$  par  $\log x$  au prix d'une erreur que, grâce à (2), on peut globalement majorer par  $\ll x$ . En introduisant la fonction  $R$  et en faisant de nouveau appel à (2), on obtient ainsi

$$(6) \quad R(x) \log x = - \sum_{n \leq x} \Lambda(n) R(x/n) + O(x).$$

À présent, récrivons (3) sous la forme

$$(7) \quad \Lambda(n) = \frac{\Lambda_2(n)}{\log n} - \frac{\Lambda * \Lambda(n)}{\log n} \quad (n > 1).$$

Une sommation d'Abel fournit alors

$$\psi(x) = 2x - \sum_{m \leq x} \frac{\Lambda * \Lambda(m)}{\log m} + O\left(\frac{x}{\log 2x}\right) \quad (x \geq 1),$$

où nous convenons que le terme général de la somme en  $m$  est nul pour  $m = 1$ . En remplaçant  $x$  par  $x/n$  et en sommant pour  $n \leq x$ , nous obtenons

$$(8) \quad \sum_{n \leq x} \Lambda(n) \psi\left(\frac{x}{n}\right) = 2x \log x - \sum_{m \leq x} \frac{\Lambda * \Lambda(m)}{\log m} \psi\left(\frac{x}{m}\right) + O(x \log_2 x),$$

où  $\log_2$  désigne le logarithme itéré. Maintenant, nous observons que

$$\begin{aligned} \sum_{m \leq x} \frac{\Lambda * \Lambda(m)}{m} &= \sum_{n \leq x} \frac{\Lambda(n)}{n} \sum_{d \leq x/n} \frac{\Lambda(d)}{d} \\ &= \sum_{n \leq x} \frac{\Lambda(n)}{n} \left\{ \log\left(\frac{x}{n}\right) + O(1) \right\} = \frac{1}{2}(\log x)^2 + O(\log x), \end{aligned}$$

d'après (2). Par sommation d'Abel, il suit

$$\sum_{m \leq x} \frac{\Lambda * \Lambda(m)}{m \log m} = \log x + O(1).$$

En reportant dans (8), nous obtenons

$$\sum_{n \leq x} \Lambda(n) R\left(\frac{x}{n}\right) = - \sum_{m \leq x} \frac{\Lambda * \Lambda(m)}{\log m} R\left(\frac{x}{m}\right) + O(x \log_2 x),$$

et donc, en tenant compte de (6),

$$(9) \quad R(x) \log x = \sum_{m \leq x} \frac{\Lambda * \Lambda(m)}{\log m} R\left(\frac{x}{m}\right) + O(x \log_2 x).$$

Posons  $\ell(n) := \Lambda(n) + \Lambda * \Lambda(n)/\log n$  ( $n > 1$ ) et  $\ell(1) = 0$ . Une sommation d'Abel fournit, à partir de (4),

$$(10) \quad L(x) := \sum_{n \leq x} \ell(n) = 2x + O\left(\frac{x}{\log 2x}\right).$$

De plus, comme  $\ell \geq 0$ , on déduit de (6) et (9) que

$$\begin{aligned} 2|R(x)| \log x &\leq \sum_{n \leq x} \ell(n) |R(x/n)| + O(x \log_2 x) \\ &= \sum_{n \leq x} L(n) \left\{ \left| R\left(\frac{x}{n}\right) \right| - \left| R\left(\frac{x}{n+1}\right) \right| \right\} + O(x \log_2 x). \end{aligned}$$

En faisant alors appel à (10) et en opérant une sommation d'Abel inverse pour évaluer le terme principal, nous en déduisons

$$(11) \quad |R(x)| \log x \leq \sum_{n \leq x} |R(x/n)| + O(x \log_2 x).$$

Le terme d'erreur issu de celui de (10) a été estimé grâce à la majoration

$$\left| R\left(\frac{x}{n}\right) \right| - \left| R\left(\frac{x}{n+1}\right) \right| \leq \psi\left(\frac{x}{n}\right) - \psi\left(\frac{x}{n+1}\right) + \frac{x}{n(n+1)},$$

qui résulte de la croissance de  $\psi$ . En estimant la somme de (11) par une intégrale, nous obtenons finalement l'inégalité fondamentale

$$(12) \quad |R(x)| \leq \frac{x}{\log x} \int_1^x |R(t)| \frac{dt}{t^2} + O\left(\frac{x \log_2 x}{\log x}\right).$$

Montrer (5) à partir de (12) est une tâche relativement aisée dans son principe. On utilise deux ingrédients, soit

$$(13) \quad \int_a^b \frac{R(t)}{t^2} dt \ll 1 \quad (a > 1, b > 1),$$

et

$$(14) \quad |R(v) - R(u)| \leq v - u + O\left(\frac{v}{\log v}\right) \quad (1 < u < v).$$

La première estimation découle de (2), la seconde de (4), en minorant  $\Lambda_2(n)$  par  $\Lambda(n) \log n$  puisque

$$0 \leq \psi(v) - \psi(u) \leq \sum_{u < n \leq v} \frac{\Lambda_2(n)}{\log n} \leq 2(v - u) + O\left(\frac{v}{\log 2v}\right).$$

On majore l'intégrale de (12) en scindant le domaine d'intégration en intervalles à bornes entières où la fonction est de signe constant. Les contributions des intervalles « courts » sont estimées en faisant appel à (13), celles des intervalles « longs » en utilisant (14) puisque la fonction est  $\ll \log u$  en une borne  $u$ . On conclut en montrant que, si  $\beta > \alpha := \limsup |R(x)|/x$  et  $\alpha > 0$ , alors il existe une constante  $\delta = \delta(\alpha) > 0$  telle que

$$\int_1^x |R(t)| \frac{dt}{t^2} \leq (1 - \delta)\beta \log x + o(\log x) \quad (x \rightarrow \infty).$$

Par un argument standard, cela implique  $\alpha \leq \{1 - \delta(\alpha)\}\alpha$ , une contradiction.

On note toutefois que les calculs nécessaires à la conclusion de cette preuve demeurent dans leur ensemble quelque peu artificiels.

### 3. Preuve de Daboussi, revisitée

Daboussi utilise le critère

$$\psi(x) \sim x \Leftrightarrow M(x) = o(x) \quad (x \rightarrow \infty)$$

où  $M(x) := \sum_{n \leq x} \mu(n)$  est la fonction sommatoire de la fonction de Möbius. Bornons-nous à établir ici que la condition est effectivement suffisante. Nous devons donc montrer que, sous l'hypothèse  $M(x) = o(x)$ , la fonction sommatoire de  $\Lambda - \mathbf{1}$  est  $o(x)$ . Cela repose sur une manipulation analogue à celle que nous avons employée pour prouver la formule asymptotique de Selberg, et en particulier sur identité de convolution

$$\begin{aligned} \Lambda - \mathbf{1} &= (\log -\mathbf{1} * \mathbf{1}) * \mu = (\log -\mathbf{1} * \mathbf{1} + 2\gamma\mathbf{1}) * \mu - 2\gamma\delta \\ &= f * \mu - 2\gamma\delta, \end{aligned}$$

où la fonction  $f$  satisfait à

$$F(x) := \sum_{n \leq x} f(n) = O(\sqrt{x}),$$

d'après la formule de Dirichlet

$$\sum_{n \leq x} \mathbf{1} * \mathbf{1}(n) = x(\log x + 2\gamma - 1) + O(\sqrt{x}) \quad (x \geq 1)$$

et celle de Stirling sous la forme faible

$$\sum_{n \leq x} \log n = x \log x - x + O(1 + \log x) \quad (x \geq 1).$$

Nous allons montrer que

$$H(x) := \sum_{n \leq x} f * \mu(n) = o(x)$$

en utilisant l'estimation précédente pour  $F(x)$  et en faisant appel au principe de l'hyperbole dû à Dirichlet. Pour chaque  $y > 2$  fixé, on peut en effet écrire

$$H(x) = \sum_{n \leq x/y} \mu(n)F(x/n) + \sum_{m \leq y} f(m)M(x/m) - F(y)M(x/y).$$

Sous l'hypothèse  $M(x) = o(x)$ , il suit donc, pour chaque  $y$  fixé,

$$\begin{aligned} \limsup_{x \rightarrow \infty} \left| H(x)/x \right| &\leq \limsup_{x \rightarrow \infty} \frac{1}{x} \sum_{n \leq x/y} |F(x/n)| \\ &\ll \limsup_{x \rightarrow \infty} \frac{1}{x} \sum_{n \leq x/y} \sqrt{x/n} \ll \frac{1}{\sqrt{y}}. \end{aligned}$$

Cela implique bien  $H(x) = o(x)$  puisque  $y$  peut être choisi arbitrairement grand.

Désignons par  $P^+(n)$  (resp.  $P^-(n)$ ) le plus grand (resp. le plus petit) facteur premier d'un entier  $n > 1$  et convenons que  $P^+(1) = 1$ ,  $P^-(1) = +\infty$ . Dans la théorie analytique moderne, les ensembles d'entiers

$$\begin{aligned} S(x, y) &:= \{n \leq x : P^+(n) \leq y\}, \\ T(x, y) &:= \{n \leq x : P^-(n) > y\}, \end{aligned}$$

jouent des rôles de plus en plus importants. La source de ce phénomène réside principalement dans le fait que, pour  $y$  donné, chaque entier  $n \leq x$  peut être décomposé de manière unique sous la forme d'un produit  $n = ab$  avec  $a \in S(x, y)$ ,  $b \in T(x, y)$ . On exploite ensuite cette factorisation avec l'idée générale que le nombre  $a$  se comporte essentiellement comme un entier ordinaire, alors que  $b$  possède les caractéristiques d'un nombre premier. *En d'autres termes,  $S(x, y)$  et  $T(x, y)$  constituent respectivement des modèles simplifiés de l'ensemble des entiers et de celui des nombres premiers n'excédant pas  $x$ .* La version présentée ici de la démonstration élémentaire de Daboussi fait appel aux propriétés de base de ces ensembles. Nous notons

$$\Psi(x, y) := |S(x, y)|, \quad \Phi(x, y) := |T(x, y)|.$$

La plupart des estimations qui suivent s'expriment de manière naturelle en fonction de la quantité

$$u := (\log x) / \log y.$$

Nous adopterons systématiquement cette notation. On a pour  $x \geq 1$ ,  $y \geq 2$

$$\Psi(x, y) \ll xe^{-u/2}, \tag{15}$$

$$\Phi(x, y) = x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) + O(xe^{-u/2} \log y). \tag{16}$$

De nombreuses recherches actuelles concernent ces quantités, et nous en rendrons compte dans le prochain exposé. Cependant, les preuves des estimations fondamentales (15) et (16) sont très simples. La première peut être déduite en quelques lignes de la convergence de la série

$$\sum_{P^+(n) \leq y} 1/n^\sigma = \prod_{p \leq y} (1 - 1/p^\sigma)^{-1}$$

pour tout  $\sigma > 0$ ,<sup>(3)</sup> et la seconde résulte d'une simple application de la formule d'inversion de Möbius

$$\Phi(x, y) = \sum_{d \leq x, P^+(d) \leq y} \mu(d) [x/d].$$

La formule (16) en découle en remplaçant  $[x/d]$  par  $(x/d) + O(1)$  pour  $d \leq x$  et en notant que (15) fournit, par sommation d'Abel, la majoration

$$\sum_{d > x, P^+(d) \leq y} \mu(d)/d \ll e^{-u/2} \log y.$$

L'idée de Daboussi est très audacieuse : déduire le comportement asymptotique de  $M(x)$  de celui de

$$M(x, y) := \sum_{n \in S(x, y)} \mu(n)$$

---

3. C'est la version originelle de la fameuse *méthode de Rankin* [6].

pour  $y$  fixé. Cela signifie essentiellement que  $S(x, y)$  constitue d'une certaine façon un modèle pertinent de  $S(x, x)$ . Il est à noter que

$$(17) \quad |M(x, y)| \leq 2^{\pi(y)} \quad (x \geq 0), \quad M(x, y) = 0 \quad (x \geq N_y := \prod_{p \leq y} p).$$

Le point de départ est l'inégalité suivante.

$$(18) \quad \limsup_{x \rightarrow \infty} \frac{|M(x)|}{x} \leq \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \int_1^{\infty} \frac{|M(x, y)|}{x^2} dx.$$

La majoration de (17) ou (15) impliquent la convergence de l'intégrale. Décomposons canoniquement chaque entier  $n$  n'excédant pas  $x$  sous la forme  $n = ab$  avec  $a \in S(x, y)$ ,  $b \in T(x, y)$ . On a alors nécessairement  $(a, b) = 1$  donc  $\mu(n) = \mu(a)\mu(b)$ . Il suit

$$M(x) = \sum_{b \in T(x, y)} \mu(b)M(x/b, y),$$

et donc

$$\begin{aligned} |M(x)| &\leq \sum_{b \in T(x, y)} |M(x/b, y)| \\ &= \sum_{n \leq x} |M(n, y)| \left\{ \Phi\left(\frac{x}{n}, y\right) - \Phi\left(\frac{x}{n+1}, y\right) \right\}. \end{aligned}$$

Utilisons maintenant (16) pour évaluer l'expression entre accolades. La contribution du terme principal est

$$\prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sum_{n \leq x} |M(n, y)| x \int_n^{n+1} \frac{dt}{t^2} \leq x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \int_1^{\infty} \frac{|M(t, y)|}{t^2} dt.$$

Cette majoration est bien compatible avec le résultat annoncé et il reste seulement à montrer que la contribution des termes d'erreur est  $o(x)$  pour chaque  $y$  fixé. D'après (16), cette contribution est, en notant  $\varepsilon(y) = 1/(2 \log y)$ ,

$$\ll \sum_{n \leq x} |M(n, y)| \left(\frac{x}{n}\right)^{1-\varepsilon(y)} (\log y) \ll 2^{\pi(y)} (\log y) \sum_{n \leq N_y} \left(\frac{x}{n}\right)^{1-\varepsilon(y)} \ll C(y) x^{1-\varepsilon(y)}.$$

Cela achève la preuve de (18).

D'après une formule classique due à Mertens, on a

$$\prod_{p \leq y} \left(1 - \frac{1}{p}\right) \sim e^{-\gamma} / \log y \quad (y \rightarrow \infty).$$

Soit  $\alpha := \limsup |M(x)/x|$ . Nous allons montrer que, si  $\beta > \alpha$  et  $\alpha > 0$ , il existe un  $\delta = \delta(\alpha) > 0$  tel que l'intégrale de (18) soit  $\leq (1 - \delta)\beta e^{\gamma} \log y + o(\log y)$  lorsque  $y \rightarrow \infty$ . La méthode est semblable à celle qui sert dans la preuve de Selberg, mais les détails techniques sont plus simples et plus naturels. Par exemple, l'analogie de (14) est ici l'inégalité triviale

$$|M(u, y) - M(v, y)| \leq v - u \quad (u, v \in \mathbb{N}, 0 < u \leq v).$$

De même, on déduit facilement de l'identité de convolution

$$\sum_{n \leq x} \mu(n) [x/n] = 1 \quad (x \geq 1)$$

que

$$\sup_{x \in \mathbb{R}} \left| \sum_{n \leq x} \mu(n)/n \right| \leq 1.$$

Or, par interversion de sommations, on a pour  $a > 0$ ,  $b > 0$ ,

$$\int_a^b M(x) \frac{dx}{x^2} = \frac{M(a)}{a} - \frac{M(b)}{b} + \sum_{a < n \leq b} \frac{\mu(n)}{n}.$$

On obtient donc

$$\sup_{a, b \in \mathbb{R}^+} \left| \int_a^b M(x) \frac{dx}{x^2} \right| \leq 4.$$

Cela permet, par un argument identique à celui utilisé dans la preuve de Selberg, de montrer l'existence d'un  $\delta$  satisfaisant aux propriétés requises tel que

$$(19) \quad \int_1^y \frac{|M(x, y)|}{x^2} dx = \int_1^y \frac{|M(x)|}{x^2} dx \leq (1 - \delta)\beta \log y + o(\log y) \quad (y \rightarrow \infty).$$

Il reste à estimer l'intégrale relative au domaine  $x > y$ . Il est à noter que l'on a

$$\int_1^\infty \Psi(x, y) \frac{dx}{x^2} = \sum_{P^+(n) \leq y} \frac{1}{n} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right)^{-1} \sim e^\gamma \log y \quad (y \rightarrow \infty)$$

et

$$\int_1^y \Psi(x, y) \frac{dx}{x^2} = \int_1^y [x] \frac{dx}{x^2} = \log y + O(1),$$

donc

$$(20) \quad \int_y^\infty \Psi(x, y) \frac{dx}{x^2} = (e^\gamma - 1) \log y + o(\log y) \quad (y \rightarrow \infty).$$

Notre hypothèse est que, pour  $x_0 = x_0(\beta)$  convenable, on a

$$|M(x, y)| = |M(x)| \leq \beta x = \beta \Psi(x, y) \quad (x_0 < x \leq y).$$

Au vu de (20), il nous suffit donc d'établir que cette estimation ne se détériore pas sensiblement lorsque  $x > y$  : en effet, on déduira alors de (18) et (19) que  $\beta \leq e^{-\gamma}(1 - \delta)\beta + e^{-\gamma}(e^\gamma - 1)\beta$ , d'où  $\beta = 0$ .

Commençons par une estimation simple de  $\Psi(x, y)$ . En classant les entiers  $n$  de  $S(x, y)$  selon la taille de leur plus grand facteur premier (i.e. en écrivant  $n = mp$  avec  $P^+(m) \leq p \leq y$ ) on obtient immédiatement l'équation de Buchstab :

$$\Psi(x, y) = 1 + \sum_{p \leq y} \Psi(x/p, p) \quad (x \geq 1, y \geq 1).$$

Si l'on approche formellement  $\Psi(x, y)$  par  $x\varrho(u)$  où  $\varrho$  est une fonction à déterminer telle que  $\varrho(u) = 1$  pour  $0 \leq u \leq 1$ , on déduit de la formule de Mertens (2) (il est bien sûr interdit d'employer ici le théorème des nombres premiers pour approcher la somme en  $p$  par une fonction continue) que

$$(21) \quad \varrho(u) - \varrho(v) = \int_v^u \frac{\varrho(t-1)}{t} dt$$

où l'on constate que la variable  $t - 1$  provient de l'approximation de

$$\Psi(x/p, p) \approx \frac{x}{p} \varrho\left(\frac{\log(x/p)}{\log p}\right) = \frac{x}{\log p} \varrho\left(\frac{\log x}{\log p} - 1\right)$$

dans la somme en  $p$ .

La fonction  $\varrho$  peut être définie par (21) à partir de la condition initiale et l'on peut justifier le raisonnement heuristique précédent pour établir que l'on a, pour tout  $\varepsilon > 0$  et uniformément pour  $1 \leq x^\varepsilon \leq y$ ,  $u = (\log x)/\log y$ ,

$$(22) \quad \Psi(x, y) = x\varrho(u) \left\{ 1 + O\left(\frac{1}{\log 2x}\right) \right\}.$$

Cette formule a été essentiellement prouvée par Dickman en 1930 et la fonction  $\varrho$  porte son nom. De nombreux travaux actuels concernent des améliorations de (22), tant dans le domaine de validité que dans la qualité de l'approximation. Nous reviendrons sur ce sujet au prochain exposé.

Les solutions d'équations différentielles aux différences, comme la fonction de Dickman, interviennent dans les problèmes de théorie des nombres parce qu'elles sont des versions continues de quantités arithmétiques qui vérifient des équations analogues discrètes. C'est en particulier le cas des fonctions de la théorie du crible.

On déduit immédiatement de (22) et (15) que

$$\int_1^\infty \varrho(u) du = e^\gamma - 1.$$

L'inégalité nécessaire pour achever la preuve de Daboussi, soit

$$(23) \quad \int_y^\infty |M(x, y)| \frac{dx}{x^2} \leq \beta(e^\gamma - 1) \log y + o(\log y) \quad (y \rightarrow \infty),$$

découle donc de l'estimation suivante, valable pour tous  $\varepsilon > 0$ ,  $\beta > \alpha$  : on a uniformément pour  $1 \leq x^\varepsilon \leq y$ ,  $u = (\log x)/\log y$ ,

$$(24) \quad |M(x, y)| \leq \beta x \varrho(u) \left\{ 1 + O\left(\frac{1}{\log 2x}\right) \right\}.$$

Pour établir (24), il serait tentant d'utiliser directement l'équation de Buchstab pour  $M(x, y)$ , soit

$$M(x, y) = 1 - \sum_{p \leq y} M(x/p, p - 1).$$

Pendant la majoration de récurrence qui en résulte, à savoir

$$|M(x, y)| \leq |M(x, x^{1/k})| + \sum_{y < p \leq x^{1/k}} |M(x/p, p - 1)|,$$

n'a pas  $x\varrho(u)$  pour solution asymptotique.

On contourne cette difficulté technique en évaluant la somme pondérée

$$\begin{aligned} \sum_{n \in S(x, y)} \mu(n) \log n &= \sum_{p \leq y} \sum_{\substack{m \in S(x/p, y) \\ p \nmid m}} \mu(mp) \log p \\ &= - \sum_{p \leq y} \{M(x/p, y) + O(x/p^2)\} \log p \\ &= - \sum_{p \leq y} M(x/p, y) \log p + O(x). \end{aligned}$$

Observant par ailleurs que

$$\left| \sum_{n \in S(x,y)} \mu(n) \log(x/n) \right| \leq \sum_{n \leq x} \log(x/n) \ll x,$$

et donc  $\sum_{n \in S(x,y)} \mu(n) \log n = M(x, y) \log x + O(x)$ , on déduit de ce qui précède que

$$M(x, y) \log x = - \sum_{p \leq y} M(x/p, y) \log p + O(x).$$

En particulier, il existe une constante absolue  $C$  telle que l'on ait pour  $x \geq 1$ ,  $y \geq 1$

$$(25) \quad |M(x, y)| \log x \leq \sum_{p \leq y} |M(x/p, y)| \log p + Cx.$$

Comme on a similairement

$$\begin{aligned} \Psi(x, y) \log x &= \sum_{n \in S(x,y)} \log n + O(x) \\ &= \sum_{nd \in S(x,y)} \Lambda(d) + O(x) \\ &= \sum_{d \in S(x,y)} \Lambda(d) \Psi(x/d, y) + O(x) \\ &= \sum_{p \leq y} \Psi(x/p, y) \log p + O(x), \end{aligned}$$

on voit que  $x\rho(u)$  est nécessairement solution asymptotique de l'inégalité fonctionnelle. On peut effectivement prouver (24) à partir de (25) par une récurrence simple en n'utilisant que le théorème de Mertens. Cela achève la preuve élémentaire de Daboussi.

Cette démonstration fournit ainsi une méthode d'attaque des problèmes de théorie des nombres en proposant un modèle simplifié de l'ensemble des entiers. Daboussi et d'autres en ont tiré de nombreuses conséquences avec des applications à des problèmes classiques de théorie analytique des nombres : sommes d'exponentielles sur les nombres premiers, valeurs moyennes de fonctions multiplicatives, répartition des fonctions additives, etc. Cette démarche naturelle, où les détails techniques sont en un certain sens réduits au minimum, dégage le rôle d'objets fondamentaux de la théorie élémentaire et réduit la traditionnelle frustration des mathématiciens devant une preuve astucieuse mais dont les idées essentielles demeurent cachées.

J'espère vous avoir convaincus que l'approche élémentaire de la théorie des nombres premiers peut aussi dévoiler de la structure.

## Bibliographie

- [1] H. Daboussi, Sur le théorème des nombres premiers, *C. R. Acad. Sc. Paris* **298** (1984), Série I, no. 8, 161–164.
- [2] H. Daboussi, On a convolution method, in : E. Aparicio, C. Calderón & J.C. Peral (eds.), *Congreso de Teoria de los Números*, 110–137, Universidad del País Vasco, 1989.
- [3] P. Erdős, On a new method in elementary number theory which leads to an elementary proof of the prime number theorem, *Proc. Nat. Acad. Sci. (Washington)* **35** (1949), 374–384.
- [4] J. Hadamard, Sur la distribution des zéros de la fonction  $\zeta(s)$  et ses conséquences arithmétiques, *Bull. Soc. Math. de France* **14** (1896), 199–220.
- [5] C. de La Vallée-Poussin, Recherches analytiques sur la théorie des nombres, 1, *Ann. Soc. Scient. Bruxelles* **20** (1896), 183–256.

- [6] R. Rankin, The difference between consecutive prime numbers, *J. London Math. Soc.* **13**, 242–247.
- [7] A. Selberg, An elementary proof of the prime number theorem, *Ann. Math.* **50** (1949), 305–313.

Gérald Tenenbaum  
Institut Élie Cartan  
Université Henri Poincaré–Nancy 1  
BP 239  
54506 Vandœuvre Cedex  
France

# Intégration complexe : la méthode du col en théorie analytique des nombres

Gérald Tenenbaum

## 1. Introduction

La preuve d'Euler de l'infinitude de l'ensemble des nombres premiers est peut-être le premier exemple de l'utilisation d'une série de Dirichlet pour obtenir un renseignement arithmétique sur les coefficients. On a

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1},$$

donc

$$\log \zeta(s) = \sum_{p, \nu \geq 1} \frac{1}{\nu p^{\nu s}} = \sum_p \frac{1}{p^s} + O(1) \quad (s > 1).$$

Comme  $\zeta(1+) = +\infty$ ,  $\sum_p 1/p$  diverge et il y a une infinité de nombres premiers.

Avec Riemann, les séries de Dirichlet

$$F(s) = \sum_{n \geq 1} \frac{a_n}{n^s}$$

sont considérées comme des fonctions de variable complexe. Les propriétés des coefficients sont alors déduites des valeurs de la somme par intégration complexe. La formule de Perron est une variante adaptée à ce contexte de la formule d'inversion de Fourier : on a

$$\sum_{n \leq x} a_n = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F(s) x^s \frac{ds}{s} \quad (x \in \mathbb{R}^+ \setminus \mathbb{N})$$

pour tout nombre réel positif  $c$  où la série converge.

La technique la plus classique consiste à déformer le contour d'intégration pour tenir compte des singularités d'un éventuel prolongement analytique de la somme. Ainsi la preuve classique du théorème des nombres premiers consiste à montrer que

$$\frac{-\zeta'(s)}{\zeta(s)} = \sum_{n \geq 1} \frac{\Lambda(n)}{n^s}$$

est prolongeable en fonction méromorphe sur le demi-plan fermé  $\sigma = \Re s \geq 1$ , ayant pour seule singularité un pôle simple en  $s = 1$ , de résidu 1. On déduit alors du théorème des résidus et de majorations convenables de ce prolongement que

$$\psi(x) = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{-\zeta'(s)}{\zeta(s)} x^s \frac{ds}{s} = x + o(x) \quad (x \rightarrow \infty).$$

La même méthode fonctionne pour des séries ayant des pôles d'ordre plus grand que 1, par exemple si  $\tau_k(n)$  est le coefficient général du développement de  $\zeta(s)^k$  en série de Dirichlet, on obtient

$$\sum_{n \leq x} \tau_k(n) = x P_{k-1}(\log x) + O(x^{1-\delta_k})$$

où  $P_{k-1}$  est un polynôme de degré  $k-1$  et  $\delta_k > 0$ . Par exemple tout choix  $\delta_k < 2/(k+2)$  est admissible. Sous l'hypothèse de Riemann, on peut choisir tout  $\delta_k < \frac{1}{2}$ .

On peut aussi appliquer cette méthode à des séries présentant des singularités logarithmiques. Ainsi  $\zeta(s)^z$  se comporte au voisinage de  $s=1$  comme  $(s-1)^z$  et n'est donc holomorphe que dans un domaine simplement connexe ne contenant pas le point 1. Désignant par  $\tau_z(n)$  le coefficient générique de cette série, on peut ainsi établir que, pour tout  $z \in \mathbb{C}$ ,

$$\sum_{n \leq x} \tau_z(n) = \frac{x}{\Gamma(z)} (\log x)^{z-1} \{ P_N(1/\log x) + O(1/(\log x)^{N+1}) \}$$

où  $P_N$  est un polynôme de degré  $N$  dépendant de  $z$ .

Enfin, il est également possible d'adapter la technique au cas d'une ou plusieurs singularités essentielles. Ainsi la série de Dirichlet d'Oppenheim (1927)

$$F(s) = \prod_{m \geq 2} \left(1 - \frac{1}{m^s}\right)^{-1}$$

dont le coefficient générique  $Q(n)$  est égal au nombre de façon de décomposer  $n$  en produit de facteurs  $> 1$ ,<sup>(1)</sup> possède une singularité du type  $\exp\{1/(s-1)\}$  en  $s=1$ . On obtient par la méthode d'intégration complexe

$$\sum_{n \leq x} Q(n) = x \frac{e^{2\sqrt{\log x}}}{2\sqrt{\pi}(\log x)^{3/4}} \left\{ 1 + O\left(\frac{1}{\sqrt{\log x}}\right) \right\}.$$

Dans certaines circonstances, il peut cependant être judicieux de confiner l'intégration complexe au domaine de convergence. C'est évidemment le cas si aucun prolongement analytique n'est connu, mais il peut également se produire que l'on sache établir l'existence d'un prolongement sans pour autant disposer de renseignements suffisants pour les majorations nécessaires à l'application du théorème des résidus.

Dans une telle situation, la méthode du col, due à Riemann (1863) et Debye (1909), peut se révéler un outil analytique très performant. Le principe général consiste à approcher une intégrale complexe

$$\int_{s_1}^{s_2} e^{\Phi(s)} ds$$

dépendant d'un paramètre  $x$ , lorsque  $x \rightarrow \infty$ . Si  $\Phi$  est holomorphe, la valeur de l'intégrale est indépendante du chemin joignant  $s_1$  à  $s_2$ . Il est logique de supputer que cette valeur sera dominée par les maximums locaux du module de l'intégrande, qui sont nécessairement des zéros de  $\Phi'$ . Supposons pour simplifier qu'il n'y ait qu'un seul point critique, disons  $s_0$ . Si le chemin choisi traverse le point  $s_0$  de façon que  $s-s_0$  ait un argument proche de

$$\frac{1}{2}\{\pi - \arg \Phi''(s_0)\},$$

un développement limité montre que la valeur absolue de l'intégrande décroît comme

$$e^{-|s-s_0|^2 |\Phi''(s_0)|/2}$$

---

1. Deux représentations ne différant que par l'ordre des facteurs sont identifiées.

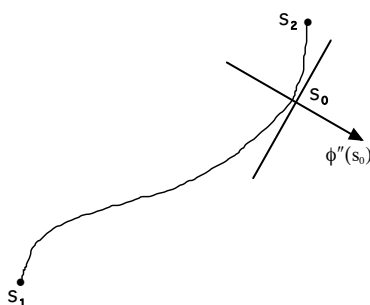


FIGURE.— *Point-selle.*

au voisinage de  $s_0$  et que cette décroissance est la plus rapide possible. On peut alors espérer que l'intégrale tout entière est dominée par un petit voisinage du col  $s_0$ .

Il est à noter que si  $s_0$  et  $\Phi''(s_0)$  sont réels, ce qui est le cas si par exemple  $e^\Phi$  est une série de Dirichlet à coefficients réels dont la dérivée s'annule sur l'axe réel, alors le chemin optimal croise l'axe réel normalement : c'est donc une situation particulièrement favorable à l'emploi de la formule de Perron.

Cela étant, il n'est en général pas facile de trouver le meilleur chemin d'intégration. J'ai eu récemment à évaluer par la méthode du col, pour  $k \in \mathbb{N}^*$ , la transformée de Laplace de  $t \mapsto \exp(-t^{2k})$ , qui se comporte essentiellement comme  $e^{-c_k \tau^{2k/(2k-1)}}$ . J'ai dû batailler deux jours durant pour trouver un chemin d'intégration convenable, qui se trouve finalement être une droite horizontale...

## 2. Entiers friables

Nous avons déjà rencontré l'ensemble

$$S(x, y) := \{n \leq x : P^+(n) \leq y\}$$

des nombres  $y$ -friables n'excédant pas  $x$ . La série de Dirichlet associée

$$\zeta(s, y) := \prod_{p \leq y} \left(1 - \frac{1}{p^s}\right)^{-1}$$

est méromorphe dans tout le plan complexe et ses pôles sont tous situés sur la droite  $\sigma = 0$ . Cette fonction est particulièrement bien adaptée à un traitement par la méthode du col car les variations de  $x^s \zeta(s, y)$  en fonction de  $s$  sont très violentes, au moins lorsque  $y$  est assez petit devant  $x$ , ce qui représente le cas le plus intéressant. Comme  $\zeta(s, y)$  est à coefficients réels et que le point critique  $\alpha$  solution de l'équation

$$\frac{-\zeta'(\alpha, y)}{\zeta(\alpha, y)} = \sum_{p \leq y} \frac{\log p}{p^\alpha - 1} = \log x$$

est également réel, on peut choisir pour chemin d'intégration une droite verticale. Posons

$$\varphi(\sigma, y) = \sum_{p \leq y} \log(1 - 1/p^\sigma), \quad \varphi''(\sigma, y) = \frac{d^2 \varphi(\sigma, y)}{d\sigma^2} = \sum_{p \leq y} \frac{p^\sigma (\log p)^2}{(p^\sigma - 1)^2}.$$

Nous obtenons entre autres le résultat suivant. Comme précédemment, on pose  $\Psi(x, y) := |S(x, y)|$  et l'on désigne par  $\varrho$  la fonction de Dickman.

**Théorème 1 [6].** *On a, uniformément dans le domaine  $x \geq y \geq 2$ ,*

$$(1) \quad \Psi(x, y) = \frac{x^\alpha \zeta(\alpha, y)}{\alpha \sqrt{2\pi \varphi''(\alpha, y)}} \left\{ 1 + O\left(\frac{1}{u} + \frac{\log y}{y}\right) \right\},$$

$$\varphi''(\alpha, y) = \left(1 + \frac{\log x}{y}\right) \log x \log y \left\{ 1 + O\left(\frac{1}{\log(u+1)} + \frac{1}{\log y}\right) \right\}.$$

De plus, si  $0 < \varepsilon < \frac{1}{2}$ ,  $y \geq (\log x)^{1+\varepsilon}$ , on a

$$(2) \quad \Psi(x, y) = x \varrho(u) \exp \left\{ O\left(\frac{\log(u+1)}{\log y} + \frac{u}{L_\varepsilon(y)}\right) \right\}.$$

L'approximation « abstraite » (1), dépendant d'une quantité définie par une équation transcendante, possède de nombreuses applications effectives.

La première conséquence à tirer des résultats issus de la méthode du col consiste à insérer dans les formules une approximation régulière de  $\alpha$ . On a

$$\alpha = \frac{\log(1 + y/\log x)}{\log y} \left\{ 1 + O\left(\frac{\log_2 y}{\log y}\right) \right\} \quad (x \geq y \geq 2)$$

et des formules plus précises peuvent être établies dans des sous-domaines. C'est ainsi que l'on déduit (2) de (1). On a en particulier

$$(3) \quad \Psi(x, y) = x \varrho(u) \left\{ 1 + O\left(\frac{\log(u+1)}{\log y}\right) \right\}$$

dans le domaine

$$(H_\varepsilon) \quad e^{(\log_2 x)^{5/3+\varepsilon}} \leq y \leq x.$$

Cette formule, initialement établie par Hildebrand [5] par une autre méthode, fournit le plus grand domaine connu pour une approximation régulière de  $\Psi(x, y)$  par  $x \varrho(u)$ . On peut montrer que toute extension significative de ce domaine de validité revient à repousser les régions sans zéro de la fonction zêta de Riemann. En particulier, montrer que (3) est valable pour  $y \geq (\log x)^{2+\varepsilon}$  équivaut à l'hypothèse de Riemann.

Le second type d'application de la méthode du col consiste à exploiter plus radicalement le succès de la technique d'intégration complexe en introduisant une fonction dont la transformée de Mellin est proche de celle de l'intégrande dans un domaine plus vaste que le simple voisinage du point-selle dont la contribution domine effectivement l'intégrale de Perron. Dans le cas des entiers friables, il se trouve que, si l'on pose  $w := (s-1) \log y$ , la quantité

$$\frac{\zeta(s)}{s} w \widehat{\varrho}(w),$$

où  $\widehat{\varrho}(w) := \int_0^\infty e^{-wt} \varrho(t) dt$  désigne la transformée de Laplace de la fonction de Dickman  $\varrho$ , est une excellente approximation de  $\zeta(s, y)/s$  sur un large segment de la droite d'intégration  $\sigma = \alpha$ . Par un calcul facile de transformation de Mellin (ou de Laplace) inverse, on obtient ainsi que

$$\Lambda(x, y) := x \int_{\mathbb{R}} \varrho(u-v) d(\lfloor y^v \rfloor / y^v) \quad (u := (\log x) / \log y)$$

est une bonne approximation de  $\Psi(x, y)$  dans le domaine  $(H_\varepsilon)$ . Plus précisément, Saias a montré en 1989 que l'on a

$$(4) \quad \Psi(x, y) = \Lambda(x, y) \left\{ 1 + O\left(e^{-(\log y)^{3/5-\varepsilon}}\right) \right\} \quad ((x, y) \in (H_\varepsilon)).$$

Cette formule peut elle-même être rendue plus explicite au prix d'un affaiblissement du terme d'erreur et du domaine de validité : on a

$$\Lambda(x, y) = x \sum_{j=0}^k a_j \frac{\varrho^{(j)}(u)}{(\log y)^j} + O\left(x \frac{\varrho^{(k+1)}(u)}{(\log y)^{k+1}}\right)$$

uniformément sous la condition

$$(5) \quad x \geq 2, \quad (\log x)^{1+\varepsilon} \leq y \leq x, \quad \min_{0 \leq j \leq k, j \leq u} \left( \frac{u-j}{k+1-j} \right) \geq \frac{\log_2 y}{\log y},$$

où les  $a_j$  sont les coefficients de Taylor en  $s = 0$  de  $s\zeta(s+1)/(s+1)$ . La dernière des conditions (5) est en fait nécessaire — cf. [12].

Toutefois, le terme principal abstrait  $\Lambda(x, y)$ , quoiqu'un peu rébarbatif à cause de l'intégrale de Laplace–Stieltjes, se révèle dans certaines circonstances très performant et d'une utilisation plus flexible que les approximation usuelles en fonction de  $\varrho$  et de ses dérivées. On montre par exemple facilement [13], [14], à partir de (4) que

$$(6) \quad \sum_{n \leq x} \log P^+(n) = \alpha x \left\{ \log x - 1 + \gamma + O\left(e^{-(\log x)^{(3/8)-\varepsilon}}\right) \right\}$$

avec  $\alpha := 1 - \int_1^{+\infty} \varrho(v) dv/v^2 \approx 0,62433$ .

La troisième catégorie d'applications concerne les « formules semi-asymptotiques » chères à Erdős. Il s'agit d'obtenir une évaluation du comportement local des quantités étudiées même dans les domaines où ces quantités n'admettent pas d'approximation régulière. Les formules issues de la méthode du col sont particulièrement favorables à ce type de renseignements car les termes principaux sont de la forme

$$e^{g(\alpha)}$$

avec  $g'(\alpha) = o(1)$  en vertu du choix de  $\alpha$  selon la méthode du col. On peut donc s'attendre à ce que les petites variations de  $g$  soient du second ordre relativement aux variations du paramètre du problème. Dans le cas des entiers friables, on obtient ainsi [6]

$$\Psi(2x, y) \sim \Psi(x, y) \Leftrightarrow y \leq (\log x)^{1+o(1)}$$

et

$$\Psi(2x, y) \sim 2\Psi(x, y) \Leftrightarrow (\log y)/\log_2 x \rightarrow +\infty.$$

### 3. Entiers friables sans facteur carré

Une autre situation exemplaire où la méthode du col fournit des résultats de grande qualité est celle du problème des entiers friables sans facteur carré. Il s'agit donc d'évaluer

$$\Psi_1(x, y) := \sum_{n \in S(x, y)} \mu(n)^2.$$

La série de Dirichlet en cause est ici

$$\zeta_1(s, y) := \prod_{p \leq y} (1 + 1/p^s)$$

et l'on est en droit de se demander s'il ne s'agit pas ici d'une variante banale du problème initial, sans intérêt spécifique. Une telle critique est partiellement fondée, dans la mesure où la technique précédemment décrite s'applique sans changement notable dans un large domaine en  $(x, y)$ . Cependant, les deux problèmes présentent certaines différences structurelles très intéressantes.

La plus notable est que  $\Psi_1(x, y)$  possède une interprétation probabiliste en termes de somme de variables indépendantes, ce qui n'est pas le cas de  $\Psi(x, y)$ . Si l'on définit, en effet, une famille de variables de Bernoulli indépendantes  $\xi_p$ , indexée par les nombres premiers  $p \leq y$ , par

$$\mathbb{P}(\xi_p = 0) = \mathbb{P}(\xi_p = 1) = \frac{1}{2},$$

et si l'on pose

$$(7) \quad S_y := \sum_{p \leq y} (\log p) \xi_p,$$

alors

$$\Psi_1(x, y) = 2^{\pi(y)} \mathbb{P}(S_y \leq \log x).$$

Une telle interprétation n'est pas possible dans le cas de  $\Psi(x, y)$ . Ainsi, les résultats probabilistes concernant les sommes de variables aléatoires indépendantes sont applicables au problème de l'évaluation de  $\Psi_1(x, y)$  et, inversement, la technique du point-selle mise en œuvre pour attaquer ce problème est susceptible d'être étendue à des situations analogues apparaissant en théorie des probabilités.

Une seconde différence entre les deux problèmes est que l'équation au point-selle du second, soit

$$(8) \quad \sum_{p \leq y} \frac{\log p}{1 + p^\alpha} = \log x,$$

n'a pas toujours une solution. Une condition nécessaire et suffisante d'existence du col est

$$\vartheta(y) := \sum_{p \leq y} \log p > \log x.$$

En fait, la symétrie des diviseurs de

$$N_y := e^{\vartheta(y)} = \prod_{p \leq y} p$$

autour de  $\sqrt{N_y}$  permet de restreindre l'étude au cas  $\vartheta(y) > 2 \log x$ , c'est-à-dire  $\alpha > 0$ . Cependant, contrairement à la situation de  $\Psi(x, y)$ , le point-selle  $\alpha$  peut tendre très rapidement vers 0, révélant alors des phénomènes nouveaux.

Posons  $\varphi_1(s, y) := \sum_{p \leq y} \log(1 + 1/p^s)$  ( $\sigma := \Re s > 0$ ), où le logarithme complexe est pris en détermination principale, et  $\varphi_1^{(j)}(s, y) := d^j \varphi_1(s, y) / ds^j$ .

Nous désignons par

$$(9) \quad \Phi^*(z) := \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-t^2/2} dt$$

la fonction de répartition décroissante de la loi de Gauss et nous posons

$$(10) \quad G(z) := e^{z^2/2} \Phi^*(z) \quad (z \in \mathbb{R}).$$

On a  $G(z) = \frac{1}{2} + O(z)$  quand  $z \rightarrow 0$ . Lorsque  $z \rightarrow \infty$ , des intégrations par parties successives fournissent un développement asymptotique : à l'ordre deux, on obtient

$$(11) \quad G(z) = \frac{1}{\sqrt{2\pi}z} \left\{ 1 - \frac{1}{z^2} + O\left(\frac{1}{z^4}\right) \right\} \quad (z \geq 1).$$

Nous conservons la notation systématique

$$(12) \quad u := \frac{\log x}{\log y} \quad (x \geq y \geq 2)$$

et nous posons, lorsque  $\alpha$  est défini par (8),  $Z = Z(x, y) = \alpha \sqrt{\varphi_1''(\alpha, y)}$ .

**Théorème 2 [1].** *On a*

$$(13) \quad \Psi_1(x, y) = x^\alpha \zeta_1(\alpha, y) G(Z) \left\{ 1 + O\left(\frac{1}{u}\right) \right\}$$

uniformément pour  $x \geq y \geq 2$  tels que  $\vartheta(y) > 2 \log x$ .

L'apparition de la fonction  $G$  dans une formule asymptotique issue de la méthode du col est inhabituelle dans un cadre arithmétique. Dans le cas de la fonction  $\Psi(x, y)$ , l'analogie de la quantité  $Z$  est toujours  $\gg \min(y, \log x) / \log y \gg 1$  et l'erreur commise en remplaçant  $G(Z)$  par  $1/(\sqrt{2\pi}Z)$  dans la formule finale est englobée par le terme d'erreur. Ici, au contraire, le facteur impliquant  $G$  est nécessaire dès que l'on a, pour une constante absolue convenable  $c$  dans  $]0, \frac{1}{2}[$ ,

$$N_y^{(1/2)-c} \leq x < N_y^{1/2}.$$

En insérant dans (13) des estimations explicites pour  $\alpha$  et  $\varphi_1''(\alpha, y)$ , on peut obtenir des évaluations explicites de  $\Psi_1(x, y)$ . Des formules asymptotiques de ce type ont été établies par Naïmi dans [9]<sup>(2)</sup> pour  $(\log x)^{1+\varepsilon} < y \leq x^{1/(\log_2 x)^2}$  avec  $\varepsilon > 0$  arbitraire et fixé, et, par une méthode différente, par Ivic et Tenenbaum dans [8] pour  $(\log x)^{2+\varepsilon} \leq y \leq x$ . Les mêmes travaux contiennent également des estimations uniformes du rapport  $\Psi_1(x, y)/\Psi(x, y)$ , dont une forme affaiblie est

$$(14) \quad \frac{\Psi_1(x, y)}{\Psi(x, y)} = \frac{1}{\zeta(\beta_1)} + o(1) \quad (x \geq y \geq 2, x \rightarrow \infty),$$

avec  $\beta_1 := \max\{1, 2 \log(1 + y/\log x) / \log y\}$ . (Voir [14], exercice corrigé III.5.7, pour une preuve directe de ce résultat simplifié.)

Nous donnons aux corollaires 'sPetrovplus' et 'sPetrovlarge' *infra* des conséquences explicites de (13) valables pour les « petites » valeurs de  $y$ , notamment celles pour lesquelles on ne disposait jusqu'à présent que de majorations de  $\Psi_1(x, y)$ . Cela permet, en particulier, de décrire quantitativement un changement de phase du comportement asymptotique de  $\Psi_1(x, y)$  autour de  $y = 2 \log x$ . Ces résultats, pourraient d'ailleurs être encore précisés ou étendus au prix de certaines complications des énoncés.

Le premier de nos deux corollaires est un théorème de grandes déviations qui mesure le caractère gaussien de la répartition des diviseurs de  $N_y$  et fournit ainsi une description probabiliste simple du changement de phase mentionné plus haut.

---

2. Voir [10] pour une présentation détaillée.

**Corollaire 3 [1].** Soient  $y \geq 2$ ,  $N_y = e^{\vartheta(y)}$ ,  $D(y)^2 = \frac{1}{4} \sum_{p \leq y} (\log p)^2$ . On a uniformément pour  $0 \leq z \ll (y/\log y)^{1/4}$  et  $x = \sqrt{N_y} e^{-zD(y)}$ ,

$$(15) \quad \Psi_1(x, y) = \tau(N_y, x) = 2^{\pi(y)} \Phi^*(z) \left\{ 1 + O\left(\frac{z^4 + 1}{u}\right) \right\}.$$

*Remarque.* Une estimation de même nature résulte d'un théorème effectif sur les grandes déviations dans le théorème de la limite centrale dû à Petrov et généralisant un résultat de Cramér. En appliquant le théorème VIII.2 de [11] à la variable aléatoire  $S_y$  définie en (7), on obtient pour  $z \geq 0$ ,  $k := \pi(y)$ ,  $z = o(\sqrt{k})$ ,  $x = \sqrt{N_y} e^{-zD(y)}$

$$(16) \quad \Psi_1(x, y) = 2^{\pi(y)} \Phi^*(z) e^{(z^3/\sqrt{k})g_y(z/\sqrt{k})} \left\{ 1 + O\left(\frac{1+z}{\sqrt{k}}\right) \right\}.$$

Ici,  $g_y(v) := \sum_{j=1}^{\infty} a_{j,y} v^j$  où  $\{a_{j,y}\}_{j=1}^{\infty}$  est une suite de nombres complexes satisfaisant à  $|a_{j,y}| \leq A^j$  ( $j \geq 1$ ) pour une constante convenable  $A$  indépendante de  $y$ , et  $a_{1,y} \asymp 1$ . La formule (13), qui est valable sans restriction sur  $z \geq 0$ , est toujours plus précise que (16). Dans le domaine considéré au Corollaire 3, on a  $k = \pi(y) \asymp u$ . Le théorème de Petrov fournit donc lorsque  $z \ll u^{1/4}$

$$(17) \quad \Psi_1(x, y) = 2^{\pi(y)} \Phi^*(z) \left\{ 1 + O\left(\frac{1+z}{\sqrt{u}} + \frac{z^4}{u}\right) \right\}.$$

Cette évaluation est donc moins précise que (15) dans le domaine considéré.

Le second corollaire est une formule asymptotique pour

$$\log \left\{ \tau(N_y, x) / \tau(N_y) \right\} = \log \left\{ \Psi_1(x, y) / 2^{\pi(y)} \right\}$$

dans un domaine plus large. Nous posons

$$\gamma(v) := \frac{1}{\log 2} \left\{ \frac{v}{e^v + 1} + \log(1 + e^{-v}) \right\}.$$

**Corollaire 4 [1].** Soient  $y \geq 2$ ,  $N_y = e^{\vartheta(y)}$ ,  $x$  satisfaisant à  $2 \log x < \vartheta(y) \leq (\log x)^3$  et  $v := \log(\{\vartheta(y)/\log x\} - 1) > 0$ . On a

$$\Psi_1(x, y) = 2^{\pi(y)\gamma(v)} \{1 + O((v+1)/\log y)\}.$$

Pour la commodité du lecteur, nous mentionnons que, dans le domaine non couvert par cet énoncé, disons  $y \gg (\log x)^3$ , une formule asymptotique pour  $\log \Psi_1(x, y)$  résulte immédiatement de (14) et d'évaluations uniformes classiques pour  $\log \Psi(x, y)$ , par exemple [13], théorème III.5.2.

À l'instar de la plupart des estimations issues de la méthode du col, le Théorème 2 fournit comme application spécifique une formule semi-asymptotique pour la fonction  $x \mapsto \Psi_1(x, y)$ .

**Corollaire 5 [1].** On a, uniformément pour  $x \geq y \geq 2$  tels que  $\vartheta(y) > 2 \log(cx)$  et  $1 \leq c \leq y$ ,

$$(18) \quad \Psi_1(cx, y) = c^\alpha \Psi_1(x, y) \left\{ 1 + O\left(\frac{1}{\sqrt{u}}\right) \right\}.$$

La méthode fournit également de bonnes estimations dans le cas des intervalles courts. Le corollaire suivant possède une application arithmétique remarquable, que nous développerons dans le prochain exposé. La symétrie des diviseurs de  $N_y := e^{\vartheta(y)}$  permet d'obtenir une estimation valable sans condition sur le couple  $(x, y)$ .

**Corollaire 6 [1].** Soit  $\kappa \geq 1$ . On a uniformément sous les conditions  $x \geq 2$ ,  $y \geq 2$ ,  $1 \leq z \leq \min\{x, y^\kappa\}$ ,

$$(19) \quad \Psi_1(x + x/z, y) - \Psi_1(x, y) \ll \Psi_1(x, y)/z.$$

## 4. Crible d'Ératosthène

Sans entrer dans tous les détails, nous mentionnons ici que la méthode du col fournit également des renseignements précis dans le problème du crible d'Ératosthène, i.e. de l'évaluation de la fonction

$$\Phi(x, y) := \sum_{\substack{n \leq x \\ P^-(n) > y}} 1.$$

Nous avons vu dans l'exposé précédent une majoration de l'écart entre  $\Phi(x, y)$  et son approximation probabiliste

$$x \prod_{p \leq y} \left(1 - \frac{1}{p}\right) = \frac{x}{\zeta(1, y)}.$$

La formule de Perron permet d'écrire, pour tout réel  $\kappa > 1$ ,

$$(20) \quad \Phi(x, y) = \frac{1}{2\pi i} \int_{\kappa-i\infty}^{\kappa+i\infty} \frac{\zeta(s)x^s}{\zeta(s, y)s} ds \quad (x \notin \mathbb{Z}^+).$$

Le résidu en  $s = 1$  vaut  $x/\zeta(1, y)$ . L'intégrande est une fonction holomorphe de  $s$  pour  $s \neq 0$  ou  $1$ , et tend vers 0 lorsque  $|\tau| \rightarrow +\infty$  dans toute bande verticale  $0 < \sigma_0 \leq \sigma \leq 1$ . On peut donc déplacer l'abscisse d'intégration vers la gauche et évaluer la nouvelle intégrale par la méthode du col. On constate alors, contrairement aux cas considérés jusqu'ici, qu'il n'existe pas de point critique réel. L'équation au point-selle possède en fait une infinité de solutions complexes proches d'une droite verticale et le terme principal provient des contributions des deux solutions conjuguées les plus proches de l'axe réel.

La formule finale fait intervenir la fonction  $\omega$  de Buchstab, qui est essentiellement l'inverse de convolution de celle de Dickman. Elle est définie comme la solution continue à droite de l'équation différentielle aux différences

$$\{t\omega(t)\}' = \omega(t-1) \quad (t > 2)$$

avec la condition initiale  $t\omega(t) = 1$  ( $1 \leq t \leq 2$ ). On prolonge cette fonction à  $\mathbb{R}$  tout entier en posant  $\omega(t) = 0$  ( $t < 1$ ). L'approximation de  $\Phi(x, y)$  produite par la méthode du col s'exprime en fonction des deux quantités

$$(21) \quad \mu_y(u) := \int_0^{+\infty} \omega(u-v)y^{-v} dv, \quad W(x, y) := x\mu_y(u) \frac{e^\gamma \log y}{\zeta(1, y)},$$

où  $\gamma$  désigne la constante d'Euler, et d'une fonction positive  $R(u)$  issue de la théorie des équations différentielle aux différences (voir [7]) et dont, pour simplifier l'exposition, nous omettons ici la définition précise, nous contentant d'indiquer la formule asymptotique

$$(22) \quad R(u) = \varrho(u) \exp \left\{ \frac{-\pi^2 u + o(u)}{2(\log u)^2} \right\} \quad (u \rightarrow \infty).$$

**Théorème 7 [15].** Soit  $\varepsilon > 0$ . Avec les notations  $L_\varepsilon := e^{(\log y)^{(3/5)-\varepsilon}}$ , et  $Y_\varepsilon := e^{(\log y)^{(3/2)-\varepsilon}$ , on a uniformément pour  $x \geq y \geq 2$

$$(23) \quad \Phi(x, y) = W(x, y) + O(E(x, y)),$$

avec

$$(24) \quad E(x, y) := \begin{cases} xR(u)/L_\varepsilon + x\varrho(u)/Y_\varepsilon, & \text{si } (x, y) \in (H_\varepsilon), \\ \Psi(x, y), & \text{dans le cas contraire.} \end{cases}$$

On déduit en particulier de ce résultat les estimations explicites suivantes.

**Corollaire 8.** Soit  $\varepsilon > 0$ . On a, uniformément pour  $(x, y) \in (H_\varepsilon)$ ,

$$\Phi(x, y) = \frac{x}{\zeta(1, y)} + O\left(\frac{xR(u)}{\log y} + \frac{x\varrho(u)}{Y_\varepsilon}\right), \quad (25)$$

$$\Phi(x, y) = \frac{e^\gamma \{x\omega(u) - y\}}{\zeta(1, y)} + O\left(\frac{xR(u) \log(u+1)}{(\log y)^2} + \frac{x\varrho(u)}{Y_\varepsilon}\right). \quad (26)$$

Comme dans le cas de  $\Psi(x, y)$ , le terme principal abstrait se révèle très maniable et propice aux applications. À titre d'exemple, mentionnons la formule duale de (6), obtenue dans [15] : posant

$$A := e^{-\gamma} \{1 + \gamma\} + \int_1^\infty \{\omega(t) - e^{-\gamma}\} \frac{dt}{t} + \sum_p \left\{ \frac{\log p}{(p-1)\zeta(1, p)} + e^{-\gamma} \log\left(1 - \frac{1}{p}\right) \right\},$$

on a, pour tout  $\varepsilon > 0$  et uniformément pour  $x \geq 2$ ,

$$(27) \quad \sum_{1 < n \leq x} \log P^-(n) = e^{-\gamma} x \log_2 x + Ax + O(xe^{-(\log x)^{3/8-\varepsilon}}).$$

## 5. Loi du $k$ -ième facteur premier

Il existe beaucoup d'autres applications récentes de la méthode du col en théorie analytique des nombres, notamment à des problèmes requérant une extension à plusieurs variables complexes. Il nous entraînerait largement hors du cadre de cet exposé de décrire ces développements ici. Contentons-nous de mentionner une application à un problème de nature probabiliste dans lequel l'intégrale à évaluer est une intégrale de Cauchy.

Désignons par  $p_1(n) < \dots < p_{\omega(n)}(n)$  la suite des facteurs premiers d'un entier générique  $n$ . Erdős a montré que, si  $\xi(n) \rightarrow \infty$ , alors

$$(28) \quad \exp \exp\{(1 - \varepsilon)j\} \leq p_j(n) \leq \exp \exp\{(1 + \varepsilon)j\} \quad (\xi(n) \leq j \leq \omega(n))$$

pour presque tout entier  $n$ .<sup>(3)</sup>

Erdős définit

$$\lambda_k(p) := \text{dens}\{n \geq 1 : p_k(n) = p\} \quad (k \geq 1, p \geq 2),$$

et interprète (28) comme l'assertion que la somme

$$\sum_p \lambda_k(p) = 1$$

est dominée par la contribution des nombres premiers  $p$  tels que

$$(1 - \varepsilon)k \leq \log_2 p \leq (1 + \varepsilon)k.$$

Il souligne ensuite comme un « paradoxe » le fait que  $\lambda_k(p)$  est maximal for  $p = e^{k+o(k)}$  et donne comme explication qu'« il y a beaucoup plus de valeurs de  $p$  autour de  $e^{e^k}$  qu'autour de  $e^k$ . »

---

3. C'est-à-dire lorsque  $n$  parcourt une suite de densité naturelle 1.

La méthode du col est susceptible de fournir des évaluations précises de  $\lambda_k(p)$  et, grâce aux formules semi-asymptotiques, des estimations du mode  $p_k^*$  défini comme le plus petit nombre premier  $p$  tel que  $\lambda_k(p_k^*) = \max_p \lambda_k(p)$ . On a

$$\lambda_k(p) = \frac{1}{p} \prod_{\substack{q < p \\ q \text{ prime}}} \left(1 - \frac{1}{q}\right) s_{k-1}(p)$$

où

$$s_j(p) := \sum_{\substack{P^+(m) < p \\ \omega(m) = j}} \frac{1}{m} \quad (j \geq 0),$$

de sorte que

$$F_p(z) := \prod_{q < p} \left(1 + \frac{z}{q-1}\right) = \sum_{j=0}^{\infty} s_j(p) z^j.$$

La méthode du col a été employée par Erdős–Tenenbaum dans [4] pour évaluer  $\lambda_k(p)$  et  $p_k^*$ . On déduit par exemple des résultats obtenus que, désignant par  $p_k$  le  $k$ -ième nombre premier et posant  $\varepsilon_p := \log_3 p / \log_2 p$ , on a

$$\lambda_k(p) = \frac{\exp \left\{ (1 + O(\varepsilon_p)) k \log_2(\log p / \log k) + O(k) \right\}}{k! p \log p}$$

uniformément pour  $p \geq p_k$  — donc en fait sans restriction !

Comme promis, les formules semi-asymptotiques fournissent des évaluations précises du maximum et du mode. Posant  $\ell_1 := \log k$ ,  $\ell_2 := \log_2 k$ , on a

$$\begin{aligned} \max_p \lambda_k(p) &= \exp \left\{ -k \left( \ell_1 - \ell_2 - 1 + \frac{2\ell_2 + 1}{\ell_1} + \frac{2\ell_2^2 - \ell_2 + O(1)}{\ell_1^2} \right) \right\}, \\ \log p_k^* &= \frac{k}{\ell_1} \left\{ 1 + \frac{2\ell_2}{\ell_1} + \frac{2\ell_2^2 - 3\ell_2 + O(1)}{\ell_1^2} \right\}. \end{aligned}$$

On voit ainsi que l’assertion initiale d’Erdős  $\log p_k^* \sim k$  est fautive, mais que son explication heuristique et sa conception sont pertinentes.

La méthode du col peut également être adaptée à la détermination de la loi globale de  $p_k(n)$ . Posons

$$\lambda_k^*(y) := \text{dens} \{n : p_k(n) > y\} = \sum_{p > y} \lambda_k(p).$$

Erdős a montré en 1969 [3] que, si l’on définit implicitement  $z = z(k, y)$  par  $y = \exp \exp \{k + z\sqrt{k}\}$ , alors

$$\lambda_k^*(y) = \Phi^*(z) + o(1) \quad (k \rightarrow \infty),$$

où  $\Phi^*(z)$  est la fonction de répartition décroissante de la loi de Gauss, définie en (9). En faisant appel aux résultats de [4] issus de la méthode du col, De Koninck et Tenenbaum [2] montrent que l’on a, uniformément pour  $y > 0$ ,

$$(29) \quad \lambda_k^*(y) = \Phi^*(z) - \frac{\Phi_0(z)}{\sqrt{2\pi k}} + O\left(\frac{1}{k}\right),$$

avec  $\Phi_0(z) = e^{-z^2/2} \left\{ \frac{1}{3} + A - \frac{1}{3} z^2 \right\}$ , et

$$A := \gamma - \sum_p \left\{ \log \left( \frac{1}{1 - 1/p} \right) - \frac{1}{p} \right\} \approx 0.26150.$$

La formule (29) atteste des deux aspects fondamentaux de la loi de répartition de  $p_k(n)$  en mettant en évidence, à côté d’un terme principal probabiliste, un terme correctif où intervient une constante arithmétique.

Nous pouvons également déduire de ces résultats d'autres estimations fines concernant la loi de répartition du  $k$ -ième facteur premier d'un entier. Un exemple intéressant est celui de la valeur médiane de cette loi, que nous définissons comme le plus petit nombre premier  $\hat{p} = \hat{p}_k$  tel que

$$\lambda_k^*(\hat{p}) \leq \frac{1}{2}.$$

Il est clair que  $\hat{p}_1 = 2$  et l'on peut montrer que  $\hat{p}_2 = 37$ ,  $\hat{p}_3 = 42719$ . La notion de valeur médiane constitue un modèle heuristique quantitatif pour la taille usuelle du  $k$ -ième facteur premier d'un entier.

Dans [2], une formule approchée pour  $\hat{p}_k$  est déduite de (29). On a

$$(30) \quad \log_2 \hat{p}_k = k - b + O(1/\sqrt{k}) \quad (k \rightarrow \infty),$$

avec  $b := \frac{1}{3} + A \approx 0,59483$ . Comme  $b > 0$ , on obtient que, pour  $k$  assez grand,

$$\text{dens}\{n : p_k(n) > \exp \exp k\} < \frac{1}{2},$$

une conséquence arithmétique relativement inattendue.

## Bibliographie

- [1] R. de la Bretèche & G. Tenenbaum, Sur les lois locales de la répartition du  $k$ -ième diviseur d'un entier, *Proc. London Math. Soc.* (3) **84** (2002), 289–323.
- [2] J.-M. De Koninck & G. Tenenbaum, Sur la loi de répartition du  $k$ -ième facteur premier d'un entier, *Math. Proc. Camb. Phil. Soc.* **133** (2002), 191–204.
- [3] P. Erdős, On the distribution of prime divisors, *Aequationes Math.* **2**, (1969), 177–183.
- [4] P. Erdős & G. Tenenbaum, Sur les densités de certaines suites d'entiers, *Proc. London Math. Soc.* (3) **59** (1989), 417–438.
- [5] A. Hildebrand, On the number of positive integers  $\leq x$  and free of prime factors  $> y$ . *J. Number Theory* **22** (1986), 289–307.
- [6] A. Hildebrand & G. Tenenbaum, On integers free of large prime factors, *Trans. Amer. Math. Soc.* **296** (1986), 265–290.
- [7] A. Hildebrand & G. Tenenbaum, On a class of difference differential equations arising in number theory, *J. d'Analyse* **61** (1993), 145–179.
- [8] A. Ivić & G. Tenenbaum, Local densities over integers free of large prime factors, *Quart. J. Math. (Oxford)*, (2) **37** (1986), 401–417.
- [9] M. Naïmi, Les entiers sans facteur carré et certaines généralisations, Thèse de troisième cycle, Université de Tunis, 1987.
- [10] M. Naïmi, Les entiers sans facteur carré  $\leq x$  dont leurs facteurs premiers  $\leq y$ , Groupe de travail en théorie analytique et élémentaire des nombres, 1986–87, 69–76, *Publ. Math. Orsay* 88-01, Univ. Paris XI, Orsay, 1988.
- [11] V. V. Petrov, *Sums of independent random variables*, Springer Verlag, Berlin, Heidelberg, New York, 1975.
- [12] É. Saias, Sur le nombre des entiers sans grand facteur premier, *J. Number Theory* **32**, no. 1 (1989), 78–99.
- [13] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, Cours spécialisés, no. 1, Société Mathématique de France (1995), xv + 457 pp.
- [14] G. Tenenbaum, en collaboration avec Jie Wu, *Exercices corrigés de théorie analytique et probabiliste des nombres*, Cours spécialisés, no. 2, Société Mathématique de France (1996), xiv + 251 pp.
- [15] G. Tenenbaum, Crible d'Ératosthène et modèle de Kubilius, in : K. Györy, H. Iwaniec, J. Urbanowicz (eds.), *Number Theory in Progress*, Proceedings of the conference in honor of Andrzej Schinzel, Zakopane, Poland 1997, 1099–1129, Walter de Gruyter, Berlin, New York, 1999.

Gérald Tenenbaum  
 Institut Élie Cartan  
 Université Henri Poincaré–Nancy 1  
 BP 239  
 54506 Vandœuvre Cedex  
 France

# Entiers friables : trois applications arithmétiques

Gérald Tenenbaum

## 1. Introduction : modèle de Kubilius

Ces dernières années, l'importance des entiers friables en théorie des nombres n'a cessé de croître, tant en raison des méthodes théoriques nouvelles qui leur font jouer un rôle essentiel que par les applications fondamentales aux problèmes de factorisation.

Les applications algorithmiques sont liées au fait que les entiers friables sont précisément ceux pour lesquels une factorisation peut être facilement obtenue : il n'est nécessaire d'essayer que les petits nombres premiers.

L'incidence des nombres friables dans les problèmes d'arithmétique provient notamment du développement des techniques de crible. Nous avons vu dans la preuve élémentaire de Daboussi pour le théorème des nombres premiers comment la décomposition  $n = ab$  avec

$$P^+(a) \leq y < P^-(b)$$

avec  $y$  fixé peut être exploitée non trivialement.

Les développements récents de la méthode du cercle, dûs à Vaughan et Wooley (voir par exemple le chapitre 12 de [17]) reposent largement sur les propriétés des entiers friables, qui sont choisis comme ensembles de sommations de sommes d'exponentielles. Les bonnes propriétés de répartition des diviseurs des entiers friables (par exemple  $d_{j+1} \leq yd_j$  si  $d_j$  et  $d_{j+1}$  sont deux diviseurs consécutifs d'un entier  $y$ -friable  $n$ ) permettent de contrôler le nombre de solutions de certaines équations diophantiennes, qui sont à leur tour interprétées comme des normes- $L^p$  de sommes d'exponentielles.

Les résultats obtenus à l'aide de la méthode du col sur les fonctions de crible  $\Psi(x, y)$  et  $\Phi(x, y)$  ont également des applications en théorie probabiliste des nombres.

Soit  $\Omega_n := \{1, 2, \dots, n\}$  et  $\mathcal{T}_y$  la tribu des événements de  $\Omega_n$  définis par des conditions de divisibilité relatives aux puissances de nombres premiers  $p^j$  avec  $p \leq y$ . Les événements  $\mathcal{T}_y$ -mesurables de  $\Omega_n$  sont donc des réunions d'ensembles de la forme

$$E_a := \{m \leq n : m = ab, P^-(b) > y\} \quad (P^+(a) \leq y).$$

Pour la mesure empirique uniforme  $\nu_n$  sur  $\Omega_n$ , on a donc

$$\nu_n(E_a) = \frac{1}{n} \Phi\left(\frac{n}{a}, y\right) \quad (P^+(a) \leq y).$$

Un modèle probabiliste de  $(\Omega_n, \nu_n, \mathcal{T}_y)$  est obtenu, en se donnant sur un ensemble abstrait  $\Omega$ ,  $\pi(y)$  partitions

$$\Omega = \cup_{j \geq 0} \omega_{p,j} \quad (p \leq y)$$

et la tribu  $\mathcal{T}_y^*$  engendrée par les intersections finies d'ensembles  $\omega_{p,j}$ . On mime alors  $E_a$ , pour  $P^+(a) \leq y$ , par  $E_a^* := \cap_{p^j \parallel a} \omega_{p,j}$  et l'on compare  $\nu_n$  à la mesure de probabilité  $\mathbb{P}_y$  définie sur  $\Omega$  par

$$\mathbb{P}_y(\omega_{p,j}) = (1 - 1/p)p^{-j}$$

avec la condition que  $\omega_{p,j}$  et  $\omega_{q,k}$  sont indépendants si  $p \neq q$ . On peut donc associer à toute partie  $\mathcal{T}_y$ -mesurable  $E$  de  $\Omega_n$  une partie  $\mathcal{T}_y^*$ -mesurable  $E^*$  de  $\Omega$  par

$$E^* := \bigcup_{E_a \subset E} E_a^*.$$

La fonction de Kubilius  $K(n, y)$  est une jauge de l'écart entre les espaces probabilisés  $(\Omega_n, \nu_n, \mathcal{T}_y)$  et  $(\Omega, \mathbb{P}_y, \mathcal{T}_y^*)$ . Elle est définie par

$$K(n, y) := \sup_{E \in \mathcal{T}_y} |\nu_n(E) - \mathbb{P}_y(E^*)|.$$

Le « lemme fondamental » du modèle de Kubilius, dû à Kubilius [13], Barban–Vinogradov [2] et Elliott [6] énonce que, notant  $u := (\log n)/\log y$ , on a

$$(1) \quad K(n, y) \ll u^{-u/8} + n^{-1/15}.$$

Avec les approximations pour  $\Phi(x, y)$  et  $\Psi(x, y)$  issues de la méthode du col, on peut obtenir une estimation satisfaisante de  $K(n, y)$ . Le résultat s'exprime en partie à l'aide de la fonction

$$H(u) := \frac{1}{2} \int_{\mathbb{R}} |\omega(v) - e^{-\gamma}| \varrho(u - v) dv + \frac{1}{2} \varrho(u).$$

Cette fonction a été introduite par Arratia et Stark [1] qui ont montré que, pour chaque  $u$  fixé, on a  $\lim_{y \rightarrow \infty} K(y^u, y) = H(u)$ . On peut montrer que

$$H(u) = \varrho(u) 2^{u+o(u)} \quad (u \rightarrow \infty),^{(1)}$$

donc  $H(u)$  tend très vite vers 0 et en particulier

$$H(u) \ll u^{-u} \quad (u \geq 1).$$

Nous avons obtenu l'énoncé suivant.

**Théorème 1 [16].** Soient  $n \in \mathbb{N}$ ,  $y \in \mathbb{R}$  tels que  $n \geq y \geq 2$ . Pour chaque  $\varepsilon > 0$ , on a

$$(2) \quad K(n, y) \ll_{\varepsilon} \varrho(u) 2^{(1+\varepsilon)u} + n^{-1+\varepsilon}$$

où l'on a posé  $u := (\log n)/\log y$ . De plus, pour chaque  $\varepsilon > 0$ ,

$$(3) \quad K(n, y) = H(u) \left\{ 1 + O\left(\frac{\log(u+1) + 2^{-u} \log_2 y}{\log y}\right) \right\} + O\left(\frac{\varrho(u/2)^2}{\exp\{(\log y)^{3/2-\varepsilon}\}}\right)$$

uniformément dans le domaine

$$(H_{\varepsilon}) \quad n \geq 3, \quad \exp\{(\log_2 n)^{5/3+\varepsilon}\} \leq y \leq n.$$

On peut montrer que l'on ne peut remplacer  $\varepsilon$  par 0 dans aucune de ses deux occurrences au membre de droite de (2). Nous déduisons en particulier de (2) l'amélioration quasi-optimale de (1)

$$(4) \quad K(n, y) \ll_{\varepsilon} u^{-u} + n^{-1+\varepsilon} \quad (n \geq y \geq 2),$$

alors que (3) implique la validité, pour tout  $\varepsilon > 0$ , de la formule asymptotique

$$(5) \quad K(n, y) = H(u) \left\{ 1 + O\left(\frac{\log_2 y}{\log y}\right) \right\}$$

dans le domaine

$$(G_{\varepsilon}) \quad n \geq 3, \quad \exp\{(\log n)^{2/5+\varepsilon}\} \leq y \leq n.$$

---

1. Une évaluation plus précise est donnée dans [16].

## 2. Mode de la loi du $k$ -ième diviseur

Un autre problème arithmétique où les entiers friables se révèlent jouer un rôle capital et quelque peu inattendu est celui du mode de la loi du  $k$ -ième diviseur. Nous avons vu dans l'exposé précédent que si

$$\lambda_k(p) := \text{dens}\{n \geq 1 : p_k(n) = p\},$$

où  $p_k(n)$  désigne le  $k$ -ième facteur premier de  $n$ , atteint son maximum en  $p = p_k^*$ , alors on a

$$\log p_k^* \sim k/\log k \quad (k \rightarrow \infty).$$

On peut évidemment se poser la question analogue pour la loi du  $k$ -ème diviseur. Posons

$$\Lambda_k(d) := \text{dens}\{n \geq 1 : d_k(n) = d\},$$

où  $d_k(n)$  désigne le  $k$ -ième diviseur de  $n$ . Le problème est ici beaucoup plus complexe que dans le cas des facteurs premiers. La raison essentielle en est que des équations

$$p_k(n) = p \quad \text{et} \quad d_k(n) = d,$$

bien qu'apparemment semblables, sont de nature très différente. En effet, elles sont respectivement équivalentes à

$$\omega(n, p) = k, \quad \tau(n, d) = k,$$

où

$$\omega(n, t) := \sum_{\substack{p \leq t \\ p|n}} 1, \quad \tau(n, t) := \sum_{\substack{d \leq t \\ d|n}} 1.$$

Or, la fonction  $\omega(n, t)$  est, pour chaque  $t$  fixé une fonction additive de  $n$ , qui est donc déterminée par sa valeur sur les puissances de nombres premiers et pour laquelle on dispose de tout un arsenal de techniques effectives. La fonction  $n \mapsto \tau(n, t)$ , au contraire, ne possède aucune structure particulière, et l'étude de ses variations ne relève pas de méthodes générales de théorie analytique des nombres.

Erdős [8] et Erdős–Tenenbaum [9] ont émis des hypothèses sur la localisation des  $d$  tels que

$$\Lambda_k(d) = \Lambda_k^* := \max_d \Lambda_k(d).$$

Pour énoncer ces conjectures, il convient de rappeler la valeur de l'ordre maximal de la fonction de diviseurs  $\tau(n) := \sum_{d|n} 1$ . On a classiquement

$$N_k := \min_{\tau(n) \geq k} n = K^{1+o(1)} \quad \text{avec} \quad K := k^{(\log_2 k)/\log 2}.$$

Comme  $d_k(n) = d$  implique  $\tau(n) \geq k$ , il est naturel de comparer  $d$  à une puissance de  $K$ . Erdős a énoncé sans démonstration dans [8] que

$$(6) \quad \Lambda_k(d) = \Lambda_k^* \Rightarrow d = K^{\log 2 + o(1)},$$

mais il a dû par la suite retirer cette assertion. On a trivialement  $\Lambda_k(d) \leq 1/d$  et

$$\Lambda_k^* \geq \Lambda_k(d_k) \geq \frac{1}{N_k} \prod_{p \leq N_k} (1 - 1/p) = 1/K^{1+o(1)}$$

si  $d_k$  est le  $k$ -ième diviseur de  $N_k$ . Ainsi

$$(7) \quad \Lambda_k(d) = \Lambda_k^* \Rightarrow d \leq K^{1+o(1)}.$$

De plus Erdős et Tenenbaum ont montré dans [9] que

$$\Lambda_k^* = 1/K^{1+o(1)}.$$

Cela les a conduit à proposer d'amender la formule (6) selon l'hypothèse que la majoration de (7) est en fait une égalité.

Il se trouve que cette conjecture est fautive : de la Bretèche et Tenenbaum ont établi dans [3] que l'on a

$$(8) \quad \Lambda_k(d) = \Lambda_k^* \Rightarrow d = K^{(1/2)+o(1)},$$

ce qui signifie qualitativement que, pour un entier  $d$  réalisant le maximum, on a  $\tau(d) = k^{(1/2)+o(1)}$  et que les entiers  $n = md$  tels que  $d_k(n) = d$  sont tels que  $\tau(m) = k^{(1/2)+o(1)}$ .

La preuve de (8) est très compliquée et nous nous bornons ici à indiquer comment les entiers friables, et plus spécifiquement les formules semi-asymptotiques issues de la méthode du col, interviennent dans le raisonnement.

Diverses considérations préliminaires permettent de réduire la démonstration de (8) à celle de la minoration asymptotique

$$\Lambda_k^* \geq K^{-1} K_1^{-\frac{1}{2}+o(1)}$$

où l'on a posé  $K_1 := k^{(\log_3 k)/\log 2}$ .

Soit  $\varepsilon > 0$ . On choisit  $d := \prod_{p \leq y} p$  avec  $y = (\frac{1}{2} + \varepsilon) \log(KK_1)$ , de sorte que

$$\tau(d) = 2^{\pi(y)} = k^{\frac{1}{2}+\varepsilon+o(1)}.$$

On a alors, pour tout ensemble  $\mathcal{M}$  d'entiers  $m$  tels que  $P^+(m) \leq d$  et  $d_k(md) = d$ ,

$$(9) \quad \Lambda_k^* \geq \Lambda_k(d) \geq \frac{1}{d} \prod_{p \leq d} (1 - 1/p) \sum_{m \in \mathcal{M}} \frac{1}{m} \approx \frac{1}{(KK_1)^{\frac{1}{2}+\varepsilon}} \sum_{m \in \mathcal{M}} \frac{1}{m},$$

où le symbole  $\approx$  signifie que l'on a égalité à un facteur  $K_1^{o(1)}$  près.

Nous devons donc minorer la somme en  $m$ . Soit  $\ell$  l'unique entier tel que  $\frac{1}{2}k/\tau(d) < 2^\ell \leq k/\tau(d)$ . Nous choisissons  $m$  de la forme  $m_1 r$  où  $m_1$  est sans facteur carré,  $\omega(m_1) = \ell$  et

$$p|m_1 \Rightarrow \ell < p \leq \ell^{1+\varepsilon}, \quad p|r \Rightarrow \ell^{1+\varepsilon} < p \leq d.$$

La dernière condition garantit qu'un entier  $m$  de cette forme est représentable de manière unique sous forme d'un produit  $m = m_1 r$ .

La somme  $\sum 1/m_1$  est le coefficient de  $z^\ell$  dans le polynôme

$$P(z) := \prod_{\ell < p \leq \ell^{1+\varepsilon}} (1 + z/p).$$

Or, pour  $z$  borné, on a  $P(z) \sim (1 + \varepsilon)^z$ . On s'attend donc à ce que

$$\sum \frac{1}{m_1} \approx \frac{\{\log(1 + \varepsilon)\}^\ell}{\ell!} \approx \frac{1}{K^{\frac{1}{2}-\varepsilon}},$$

et la méthode du col permet effectivement d'établir un tel résultat.

On a

$$(10) \quad \sum_{m \in \mathcal{M}} \frac{1}{m} = \sum_{m_1} \frac{1}{m_1} \sum_{r \in \mathcal{E}(m_1)} \frac{1}{r}$$

où  $\mathcal{E}(m_1) := \{r : p|r \Rightarrow \ell^{1+\varepsilon} < p \leq d, d_k(dm_1 r) = d\}$ . Au vu de (9) et (10), il nous suffit de montrer que

$$(11) \quad \min_{m_1} \sum_{r \in \mathcal{E}(m_1)} \frac{1}{r} \geq 1/K_1^{o(1)}.$$

Nous construisons les entiers  $r$  de  $\mathcal{E}(m_1)$  par un procédé inductif reposant sur l'identité

$$(12) \quad \tau(np, x) = \tau(n, x) + \tau(n, x/p) \quad (p \nmid n).$$

Posons  $r_0 = 1$ , et  $r_{j+1} = r_j p_j$  et  $k_j := k - \tau(dm_1 r_j, d)$ . On a

$$k_0 := k - \tau(dm_1, d) \leq 3k/4$$

par construction, et nous cherchons à construire des nombres premiers  $p_1, \dots, p_R$  de façon que

$$k_R = k - \tau(dm_1 p_1 \cdots p_R, d) = 0.$$

Supposons  $p_1, \dots, p_j$  déjà déterminés. Par (12), on a

$$(13) \quad k_{j+1} = k_j - \tau(dm_1 r_j, d/p_{j+1}).$$

Soient  $h := \log k$  et  $x_j, x_j^*$  des entiers tels que

$$\tau(dm_1 r_j, x_j) = [k_j(1 - 1/3h)], \quad \tau(dm_1 r_j, x_j^*) = [k_j(1 - 1/2h)].$$

On déduit donc de (13) que, pour tout  $p_{j+1} \in ]d/x_j^*, d/x_j]$ ,

$$k_{j+1} \leq k_j/h.$$

On note en particulier que le nombre  $R$  d'étapes nécessaires pour atteindre  $k_R = 0$  est  $\ll (\log k)/\log h$ . De plus, en supposant, pour simplifier l'exposition, que l'intervalle  $]d/x_j^*, d/x_j]$  soit suffisamment long, le théorème des nombres premiers garantit que

$$(14) \quad \sum_{d/x_j^* < p \leq d/x_j} \frac{1}{p} \gg \frac{x_j^* - x_j}{x_j^* \log d} \gg \frac{x_j^* - x_j}{x_j^* h^2}.$$

Il nous faut donc une minoration de  $(x_j^* - x_j)/x_j^*$ . Sous l'hypothèse que la fonction  $x \mapsto \tau(dm_1 r_j, x)$  est sous-linéaire en un sens faible, nous avons

$$\begin{aligned} \frac{k_j}{h} &\ll \tau(dm_1 r_j, x_j^*) - \tau(dm_1 r_j, x_j) \\ &\ll \frac{x_j^* - x_j}{x_j} \tau(dm_1 r_j, x_j) \ll \frac{x_j^* - x_j}{x_j} k_j \end{aligned}$$

donc

$$\frac{x_j^* - x_j}{x_j} \gg \frac{1}{h}$$

et, en reportant dans (14),

$$\sum 1/p_{j+1} \geq 1/h^3.$$

On en déduit que, pour tout  $m_1$ ,

$$\sum_{r \in \mathcal{E}(m_1)} \frac{1}{r} \gg 1/h^{4R} \gg 1/k^{O(1)}.$$

Cela implique bien (11).

Nous avons ainsi montré que la preuve de (8), c'est-à-dire la détermination du mode de la loi de répartition de  $d_k(n)$ , se ramène à établir la sous-linéarité faible de la fonction  $x \mapsto \tau(dm, x)$  pour certaines valeurs de  $m$ . Lorsque  $m = 1$ , on a, avec notre définition de  $d$ ,

$$\tau(dm, x) = \Psi_1(x, y).$$

Le résultat est alors une conséquence évidente de l'estimation

$$\Psi_1(x + x/z, y) - \Psi_1(x, y) \ll \Psi_1(x, y)/z \quad (1 \leq z \leq \min(x, y^c))$$

établie par la méthode du col, comme indiqué dans l'exposé précédent. En fait, on peut montrer par un argument relativement simple<sup>(2)</sup> que la sous-linéarité faible souhaitée dans le cas des valeurs de  $m$  considérées pour l'application envisagée est réductible à celle du cas  $m = 1$ . Cela achève notre description de la preuve de (8).

### 3. $P$ -sommation

Pour achever notre présentation des applications récentes des entiers friables à des problèmes de théorie analytique des nombres, nous mentionnons brièvement un sujet qui mériterait à lui seul un exposé complet, voire plusieurs.

Il s'agit du procédé de sommation introduite par Fouvry et Tenenbaum dans [10] et désigné sous le nom de  $P$ -sommation.

On dit qu'une série de nombres complexes  $\sum_{n=1}^{\infty} \alpha_n$ ,  $P$ -converge vers  $\alpha$  si la série  $\sum_{P(n) \leq y} \alpha_n$  converge pour tout  $y \geq 2$ , et si l'on a

$$\sum_{P(n) \leq y} \alpha_n = \alpha + o(1) \quad (y \rightarrow \infty).$$

On écrit alors

$$\sum_n \alpha_n = \alpha \quad (P).$$

On dit qu'une série convergente  $\sum_{n=1}^{\infty} \alpha_n$  est  $P$ -régulière si elle est  $P$ -convergente et si sa  $P$ -somme est égale à sa somme ordinaire, autrement dit si l'on a

$$\lim_{y \rightarrow \infty} \sum_{P(n) \leq y} \alpha_n = \sum_{n=1}^{\infty} \alpha_n.$$

Il est à noter qu'une série  $\sum_{P(n) \leq y} \alpha_n$  est convergente dès que  $\alpha_n \ll 1/n^c$  avec  $c > 0$ .

La  $P$ -sommation n'est pas un procédé régulier : pour  $(a, q) = 1$ , on a [10]

$$\sum_{n \geq 1} \frac{e(na/q)}{n} = \log \left( \frac{1}{1 - e(a/q)} \right) + \frac{\Lambda(q)}{\varphi(q)} \quad (P),$$

---

2. Voir le lemme 3.8 de [3].

où, conformément à l'usage, nous avons posé  $e(u) := e^{2\pi i u}$  ( $u \in \mathbb{R}$ ).

Une série peut être  $P$ -convergente sans être convergente : si  $\tau \in \mathbb{R}$ ,  $\tau \neq 0$ , on a

$$\sum_{n \geq 1} \frac{1}{n^{1+i\tau}} = \zeta(1+i\tau) \quad (P)$$

mais la série ne converge pas.

De nombreux problèmes de la théorie des nombres s'expriment simplement en termes de  $P$ -sommation. Par exemple, il découle du théorème de Mertens que

$$\sum_{P(n) \leq y} \frac{\mu(n)}{n} = \prod_{p \leq y} \left(1 - \frac{1}{p}\right) \rightarrow 0 \quad (y \rightarrow \infty),$$

c'est-à-dire

$$\sum_{n \geq 1} \frac{\mu(n)}{n} = 0 \quad (P).$$

Ainsi, le théorème des nombres premiers est élémentairement équivalent à la  $P$ -régularité de  $\sum_n \mu(n)/n$ .

Montgomery a montré (voir [10]) que la  $P$ -sommation est régulière lorsqu'on la restreint aux séries de terme général  $f(n)/n$  où  $f$  est une fonction multiplicative à valeurs dans le disque unité et satisfaisant à  $f(2^\nu) \neq -1$  pour au moins un entier  $\nu \geq 1$ .

Dans [4], de la Bretèche et Tenenbaum montrent que les  $P$ -séries de Fourier d'une large classe de fonctions évitent le phénomène de Gibbs. Soit  $VB_n(\mathbb{T})$  la classe des fonctions  $F$  qui sont 1-périodiques, à variation bornée sur  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ , et normalisées par  $F(\vartheta) = \frac{1}{2}\{F(\vartheta+) + F(\vartheta-)\}$  aux points de discontinuité. Toute fonction  $F$  de  $VB_n(\mathbb{T})$  est dérivable presque partout et la décomposition de Radon–Nikodym de  $dF$  s'écrit

$$(15) \quad dF(\vartheta) = F'(\vartheta) d\vartheta + d\sigma_F(\vartheta)$$

où la partie singulière  $\sigma_F$  est également à variation bornée sur  $\mathbb{T}$ . Désignons par  $BV_n^2(\mathbb{T})$  le sous-espace de  $BV_n(\mathbb{T})$  constitué des fonctions  $F$  dont la dérivée pp  $F'$  est dans  $L^2(\mathbb{T})$ . Nous notons  $a_n(F)$ ,  $b_n(F)$  les coefficients de Fourier d'une fonction  $F$  de  $L^1(\mathbb{T})$ , et nous définissons la  $P$ -série de Fourier de  $F$  par

$$(16) \quad F(\vartheta; y) := a_0(F) + \sum_{P(n) \leq y} \{a_n(F) \cos(2\pi n\vartheta) + b_n(F) \sin(2\pi n\vartheta)\}$$

pour toute valeur de  $\vartheta$  où cela a un sens. On peut montrer facilement que, lorsque  $F \in BV_n^2(\mathbb{T})$ , la série (16) est absolument convergente pour chaque  $y \geq 2$ . En effet, posant  $\tilde{F}(\vartheta) := \int_0^\vartheta F'(v) dv$ , de sorte que

$$a_n(\tilde{F}) = -\frac{b_n(F')}{2\pi n}, \quad b_n(\tilde{F}) = \frac{a_n(F')}{2\pi n} \quad (n \geq 1),$$

la formule de Parseval pour  $F' \in L^2(\mathbb{T})$  implique donc

$$(17) \quad \sum_{n=1}^{\infty} \{|a_n(\tilde{F})| + |b_n(\tilde{F})|\} \leq \left\{ \sum_{n=1}^{\infty} \frac{1}{2\pi^2 n^2} \sum_{n=1}^{\infty} \{|a_n(F')|^2 + |b_n(F')|^2\} \right\}^{1/2} < \infty.$$

**Théorème 2 [4].** Soit  $F \in BV_n^2(\mathbb{T})$ . On a

$$(18) \quad \lim_{y \rightarrow \infty} F(\vartheta; y) = F(\vartheta) \quad (\vartheta \in \mathbb{T})$$

et

$$(19) \quad \lim_{y \rightarrow \infty} \sup_{\vartheta \in \mathbb{T}} F(\vartheta; y) = \sup_{\vartheta \in \mathbb{T}} F(\vartheta).$$

Dans le même article, il est fait un usage crucial de la  $P$ -sommation pour étudier les conditions de validité de certaines identités obtenues formellement par interversion de sommations. Désignons par

$$B(\vartheta) := - \sum_{m \geq 1} \frac{\sin 2\pi m \vartheta}{\pi m}$$

la première fonction de Bernoulli, qui satisfait donc à

$$B(\vartheta) = \begin{cases} \langle \vartheta \rangle - \frac{1}{2} & \text{si } \vartheta \notin \mathbb{Z}, \\ 0 & \text{si } \vartheta \in \mathbb{Z}, \end{cases}$$

où  $\langle \vartheta \rangle$  désigne la partie fractionnaire du nombre réel  $\vartheta$ .

Le théorème suivant est représentatif des résultats obtenus dans [4].

**Théorème 3 [4].** On a

$$(20) \quad \sum_{m=1}^{\infty} \frac{\tau(m)}{\pi m} \sin(2\pi m \vartheta) + \sum_{n=1}^{\infty} \frac{B(n\vartheta)}{n} = 0$$

pour tout  $\vartheta \in \mathbb{Q}$ . Pour  $\vartheta \in \mathbb{R} \setminus \mathbb{Q}$ , la relation (20) est valide si, et seulement si, la série

$$(21) \quad \sum_{m=1}^{\infty} (-1)^m \frac{\log q_{m+1}}{q_m}$$

converge, où  $\{q_m\}_{m=1}^{\infty}$  désigne la suite des dénominateurs des réduites de  $\vartheta$ . Lorsque la série (21) diverge, il en va de même des deux séries de (20).

## Bibliographie

- [1] R. Arratia & D. Stark, A total variation distance invariance principle for primes, permutations and Poisson-Dirichlet. Manuscrit, 1997.
- [2] M.B. Barban & A.I. Vinogradov, On the number theoretic basis of probabilistic number theory. Dokl. Akad. Nauk SSSR 154 (1964), 495-496 ; = Soviet Math. Doklady 5 (1964), 96-98.
- [3] R. de la Bretèche & G. Tenenbaum, Sur les lois locales de la répartition du  $k$ -ième diviseur d'un entier, *Proc. London Math. Soc.* (3) **84** (2002), 289-323.
- [4] R. de la Bretèche & G. Tenenbaum, Sommes trigonométriques à coefficients arithmétiques, *J. Anal. Math.* **92** (2004), 1-79.
- [5] J.-M. De Koninck & G. Tenenbaum, Sur la loi de répartition du  $k$ -ième facteur premier d'un entier, *Math. Proc. Camb. Phil. Soc.* **133** (2002), 191-204.
- [6] P.D.T.A. Elliott, Probabilistic number theory : mean value theorems (Grundlehren der Math. Wiss. 239). Springer-Verlag, New York, Berlin, Heidelberg 1979.
- [7] — Probabilistic number theory : central limit theorems (Grundlehren der Math. Wiss. 240). Springer-Verlag, New York, Berlin, Heidelberg 1980.
- [8] P. Erdős, On some unconventional problems in number theory, *Astérisque* **61** (1979), 73-82.
- [9] P. Erdős & G. Tenenbaum, Sur les densités de certaines suites d'entiers, *Proc. London Math. Soc.* (3) **59** (1989), 417-438.
- [10] E. Fouvry & G. Tenenbaum, Entiers sans grand facteur premier en progressions arithmétiques, *Proc. London Math. Soc.* (3) **63** (1991), 449-494.

- [11] A. Hildebrand & G. Tenenbaum, On integers free of large prime factors, *Trans. Amer. Math. Soc.* **296** (1986), 265–290.
- [12] A. Hildebrand & G. Tenenbaum, On a class of difference differential equations arising in number theory, *J. d'Analyse* **61** (1993), 145–179.
- [13] J. Kubilius, Probabilistic methods in the theory of numbers. *Uspehi Mat. Nauk* (N.S.) 11 (1956), 2(68), 31–66 ; = Amer. Math. Soc. Translations, vol. 19 (1962), 47–85.
- [14] G. Tenenbaum, *Introduction à la théorie analytique et probabiliste des nombres*, Cours spécialisés, no. 1, Société Mathématique de France (1995), xv + 457 pp.
- [15] G. Tenenbaum, en collaboration avec Jie Wu, *Exercices corrigés de théorie analytique et probabiliste des nombres*, Cours spécialisés, no. 2, Société Mathématique de France (1996), xiv + 251 pp.
- [16] G. Tenenbaum, Crible d'Ératosthène et modèle de Kubilius, in : K. Győry, H. Iwaniec, J. Urbanowicz (eds.), *Number Theory in Progress*, Proceedings of the conference in honor of Andrzej Schinzel, Zakopane, Poland 1997, 1099–1129, Walter de Gruyter, Berlin, New York, 1999.
- [17] R.C. Vaughan, *The Hardy-Littlewood method*, second edition, Cambridge University Press, 1997.

Gérald Tenenbaum  
Institut Élie Cartan  
Université Henri Poincaré–Nancy 1  
BP 239  
54506 Vandœuvre Cedex  
France