

*Séminaire de Théorie des Nombres
Paris 1987-88*

**COMPORTEMENT STATISTIQUE DU NOMBRE
DE FACTEURS PREMIERS DES ENTIERS
M. BALAZARD**

I.— Introduction.

Si n est un entier positif, on note $\Omega(n)$ le nombre de facteurs premiers de n , comptés avec leurs multiplicités :

$$(1) \quad \Omega(n) = \sum_{p^\alpha | n} 1 = \sum_{p^\alpha || n} \alpha$$

où p désigne un nombre premier générique et α un entier positif générique. La fonction arithmétique Ω est complètement additive, c'est-à-dire que $\Omega(ab) = \Omega(a) + \Omega(b)$, quels que soient les entiers naturels a et b .

Dans cet exposé, nous étudions le comportement local de Ω . Posons :

$$(2) \quad \nu_x(\Omega(n) = k) = \frac{1}{x} \sum_{\substack{n \leq x \\ \Omega(n) = k}} 1$$

où x est un entier positif tendant vers l'infini, et k un entier positif. Dans (2), ν_x désigne la probabilité uniforme sur l'ensemble des entiers $1, 2, \dots, x$; comme $2^{\Omega(n)} \leq n$, on a $\nu_x(\Omega(n) = k) = 0$ pour $k > \frac{\log x}{\log 2}$ et il suffit d'étudier (2) pour $k \leq \frac{\log x}{\log 2}$.

Depuis le début de notre siècle, de nombreux auteurs ont donné des équivalents asymptotiques ou des majorations pour $\nu_x(\Omega(n) = k)$. Nous résumons ci-dessous ces travaux; pour ne pas alourdir cette présentation, nous omettons les termes d'erreurs effectifs connus pour les résultats (6) et (10).

$$(3) \quad \nu_x(\Omega(n) = k) \sim (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!} \quad \text{quand } x \rightarrow +\infty,$$

pour tout k fixé (Landau 1900, cf. [11]).

$$(4) \quad \nu_x(\Omega(n) = k) \leq c_0 \left(\frac{10}{9}\right)^{-k} (\log x)^{-1} S_{k-1} \left(\frac{10}{9}(\log \log x + c_1)\right)$$

uniformément pour $x \geq 3$ et $k \geq 1$, où c_0 et c_1 sont des constantes positives

absolues et $S_{k-1}(X) = \sum_{i=0}^{k-1} \frac{X^i}{i!}$ est la $(k-1)$ -ième somme partielle de la série exponentielle (Hardy et Ramanujan 1917, cf. [8]). Signalons qu'une inégalité fautive $(\nu_x(\Omega(n) = k) \leq c_0(\log x)^{-1}(\log \log x + c_1)^{k-1}/(k-1)!)$ a parfois été utilisée imprudemment à la place de (4).

$$(5) \quad \nu_x(\Omega(n) = k) \sim (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!} \quad \text{quand } x \rightarrow +\infty$$

et $|k - \log \log x| \leq B(\log \log x)^{\frac{1}{2}}$ où B est positif, arbitraire mais fixé (Erdős 1948, cf. [5]).

$$(6) \quad \nu_x(\Omega(n) = k) \sim F\left(\frac{k-1}{\log \log x}\right) (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!} \quad \text{quand } x \rightarrow +\infty,$$

uniformément pour $1 \leq k \leq (2-\epsilon)\log \log x$, où $\epsilon > 0$ est fixé et

$$F(z) = \frac{1}{\Gamma(z+1)} \prod_p \left(1 - \frac{1}{p}\right)^z \left(1 - \frac{z}{p}\right)^{-1} \quad (\text{Sathe-Selberg 1953-54, cf. [19]}).$$

$$(7) \quad \nu_x(\Omega(n) = k) \sim C(\log x) 2^{-k} \quad \text{quand } x \rightarrow +\infty,$$

uniformément pour $(2+\epsilon)\log \log x \leq k \leq B \log \log x$ où ϵ et B sont positifs et fixés et $C = \frac{1}{4} \prod_{p \geq 3} \left(1 + \frac{1}{p(p-2)}\right)$ (Selberg 1954, cf. [19]).

$$(8) \quad \nu_x(\Omega(n) = k) \leq c_2(\log x) k^4 2^{-k}$$

uniformément pour $x \geq 3$ et $k \geq 1$, où c_2 est une constante positive absolue (Erdős–Sárközy 1980, cf. [6]).

$$(9) \quad \nu_x(\Omega(n) = k) \leq c_3(\log x)(\log \log x)^{\frac{1}{2}-k}$$

uniformément pour $x \geq 3$ et $k \geq 1$, où c_3 est une constante positive absolue (Norton 1981, cf. [14]).

$$(10) \quad \nu_x(\Omega(n) = k) \sim C 2^{-k} \log(x 2^{-k}) \text{ quand } x 2^{-k} \rightarrow +\infty,$$

uniformément pour $k \geq (2+\epsilon)\log \log x$, $\epsilon > 0$ étant fixé (Nicolas 1984, cf. [13]). Signalons des travaux récents d'Azzouza, donnant des majorations explicites de $\nu_x(\Omega(n) = k)$, et utilisant la démonstration du théorème de Nicolas.

Le rapprochement entre les lois de répartition des fonctions arithmétiques et les lois probabilistes classiques est l'un des objectifs de la théorie probabiliste des nombres. Ainsi (6) montre que, pour $k \leq (2-\epsilon)\log \log x$, $\Omega(n)$ se comporte à peu près comme une variable de Poisson de paramètre $\log \log x$ et (10) indique une loi locale à peu près géométrique de raison $\frac{1}{2}$ pour $k \geq (2+\epsilon)\log \log x$.

Il est naturel de s'interroger sur ce brusque changement de nature des formules asymptotiques pour $\nu_x(\Omega(n) = k)$. La solution du problème est donnée par la considération d'une nouvelle loi probabiliste simple.

II.— La loi Poisson–géométrique.

Considérons une loi de Poisson de paramètre $\lambda \geq 0$, définie par la formule

$$(11) \quad p_k = e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \quad k = 1, 2, \dots,$$

et une loi géométrique de raison r :

$$(12) \quad g_k = (1-r)r^k \quad k = 0, 1, 2, \dots$$

Nous appelons loi Poisson–géométrique de paramètre λ et de raison r le produit de convolution :

$$(13) \quad (p \star g)_k = \sum_{j=1}^k p_j g_{k-j} = e^{-\lambda(1-r)r^{k-1}} S_{k-1}(\lambda/r)$$

où, comme pour (4), S_{k-1} est la $(k-1)$ -ième somme partielle de la série exponentielle.

Afin d'obtenir l'ordre de grandeur de $(p \star g)_k$, rappelons les résultats classiques suivants :

- i) si $k \leq (1-\epsilon)X$, $S_k(X) \sim \frac{X^k}{k!}$ pour $X \rightarrow +\infty$, $\epsilon > 0$ fixé
- ii) si $k \geq (1+\epsilon)X$, $S_k(X) \sim e^X$ pour $X \rightarrow +\infty$, $\epsilon > 0$ fixé
- iii) si $k = X + t\sqrt{X}$, $S_k(X) \sim \frac{e^X}{\sqrt{2\pi}} \int_{-\infty}^t \exp(-u^2/2) du$ uniformément pour $|t| \leq B$, $B > 0$ fixé.

Le lecteur intéressé trouvera des résultats beaucoup plus précis dans [15].

Ainsi, si r est fixé et λ grand, on a :

$$(14) \quad (p \star g)_k \sim e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} (1-r) \text{ si } k \leq (1-\epsilon)\frac{\lambda}{r}$$

$$(15) \quad (p \star g)_k \sim e^{\lambda(\frac{1}{r}-1)} (1-r)r^{k-1} \text{ si } k > (1+\epsilon)\frac{\lambda}{r},$$

avec une transition gaussienne, donnée par iii), entre ces deux zones.

La loi Poisson-géométrique a donc un comportement double : (14) évoque plutôt une loi de Poisson de paramètre λ et (15) une loi géométrique de raison r . En rapprochant (14) et le résultat de Sathe-Selberg (6), (15) et le résultat de Nicolas (10), et compte-tenu également de l'inégalité de Hardy et Ramanujan (4), le théorème suivant semble naturel.

THEOREME 1 (Balazard, Delange, Nicolas 1988). *Uniformément par rapport à k , on a :*

$$(16) \quad \nu_x(\Omega(n) = k) \sim f(\min(2, \frac{k-1}{\log \log y})) \frac{2^{-k}}{\log y} S_{k-1}(2 \log \log y)$$

quand $y = x 2^{-k} \rightarrow +\infty$, où $f(z) = \frac{2^{1-z}}{\Gamma(z+1)} \prod_{p \geq 3} (1 - \frac{1}{p})^z (1 - \frac{z}{p})^{-1}$.

Ce théorème est annoncé, sous une forme plus précise, dans [4]. Sa démonstration complète se trouve dans [2] ; le paragraphe suivant en contient les principales idées. Il montre que la loi locale de $\Omega(n)$ pour $1 \leq n \leq x$ est en gros une loi Poisson-géométrique de "paramètre" $\log \log y$ et de raison $\frac{1}{2}$ ($\log \log y$ dépend de k , mais assez faiblement : on a $0 \leq \log \log x - \log \log y \leq \log \frac{1}{\epsilon}$ si $k \leq (1-\epsilon) \frac{\log x}{\log 2}$). Observons que la fonction f est continue et strictement positive sur $[0, 2]$, et que $f(2) = C$.

En utilisant pour $S_k(X)$ des estimations plus précises que i) et ii) ci-dessus, on peut retrouver le résultat de Sathe-Selberg (6) et celui de Nicolas (10) à partir du théorème 1. Combiné avec iii), il donne pour la *zone critique* $k \sim 2 \log \log x$ un nouveau résultat, observé pour la première fois par Delange :

COROLLAIRE. *Pour tout $B > 0$ fixé, on a :*

$$(17) \quad \nu_x(\Omega(n) = k) \sim C 2^{-k} \log x \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp(-u^2/2) du \quad \text{quand } x \rightarrow +\infty,$$

uniformément pour $-B \leq t = \frac{k-2 \log \log x}{\sqrt{2 \log \log x}} \leq B$.

III.- Interprétation probabiliste et plan de la démonstration du théorème 1.

La loi locale d'une somme de deux variables aléatoires indépendantes est le produit de convolution de leurs lois locales respectives. La signification du théorème 1 est que dans l'écriture

$$(18) \quad \Omega(n) = \nu_2(n) + \Omega'(n),$$

les fonctions arithmétiques $v_2(n)$ (= valuation 2-adique de n) et $\Omega'(n)$ (= nombre de facteurs premiers impaires de n) se comportent pour $n \leq x$ comme deux variables aléatoires indépendantes. La loi locale de v_2 est approximativement géométrique de raison $\frac{1}{2}$ et celle de Ω' approximativement de Poisson de paramètre $\log \log x$. Bien entendu, la même analyse montre que la loi de Ω' a une composante géométrique prépondérante de raison $\frac{1}{3}$ mais celle-ci s'avère négligeable par rapport à la loi de v_2 .

La démonstration du théorème 1 utilise cette même idée : séparer dans $\Omega(n)$ les deux quantités $v_2(n)$ et $\Omega'(n)$. Si n est entier ≥ 1 , écrivons $n = 2^\alpha m$ avec $\alpha = v_2(n)$ et m impair. Nous avons :

$$\begin{aligned} n \leq x \text{ et } \Omega(n) = k &\Leftrightarrow 2^\alpha m \leq x \text{ et } \alpha + \Omega(m) = k \\ &\Leftrightarrow 2^{k-\Omega(m)} m \leq x, \Omega(m) \leq k \text{ et } \alpha = k - \Omega(m). \end{aligned}$$

Par conséquent, $x \nu_x(\Omega(n) = k)$ est exactement le nombre d'entiers m impairs tels que $m 2^{-\Omega(m)} \leq x 2^{-k}$ et $\Omega(m) \leq k$. En notant $\psi(m) = m 2^{-\Omega(m)}$, $y = x 2^{-k}$, la lettre m désignant dans toute la suite un entier générique impair, nous avons :

$$(19) \quad x \nu_x(\Omega(n) = k) = \sum_{\substack{\psi(m) \leq y \\ \Omega(m) \leq k}} 1.$$

Cette formule est due à Halász et a déjà été utilisée par Nicolas pour démontrer (10). Elle peut aussi servir au calcul numérique rapide de $x \nu_x(\Omega(n) = k)$ pour de grandes valeurs de x .

La démonstration de Nicolas pour (10) est élémentaire mais nous pouvons utiliser à partir de (19) des méthodes d'analyse complexe. La formule de Cauchy nous donne en effet :

$$(20) \quad x \nu_x(\Omega(n) = k) = \frac{1}{2\pi i} \int_{|z|=r} \sum_{\psi(m) \leq y} z^{\Omega(m)} \frac{z^{-k-1}}{1-z} dz$$

où $r < 1$ est arbitraire, la circonférence $|z| = r$ étant parcourue une fois dans le sens positif.

On est ainsi amené à chercher une estimation précise de la somme $\sum_{\psi(m) \leq y} z^{\Omega(m)}$. Cette estimation est obtenue par la méthode d'intégration

complexe exposée par Selberg dans [19]. La fonction génératrice $\sum \frac{z^{\Omega(m)}}{m \psi(m)^s}$ se

factorise en $\zeta(s) z^{2s} G(s, z)$, la fonction $G(s, z)$ étant holomorphe pour $|z| < \frac{3}{2}$, $\text{Re } s > \max(\frac{1}{2}, \frac{\log |z|}{\log \frac{3}{2}})$. La nouveauté est ici la dépendance en la

variable s de l'exposant de la fonction ζ (cf. également [1]). On obtient finalement :

$$(21) \quad \sum_{\psi(m) \leq y} z^{\Omega(m)} = z f(2z) y (\log y)^{2z-1} + O_R(y (\log y)^{2\text{Re}z-2} (\log \log y))$$

uniformément pour $|z| \leq R < \frac{3}{2}$ et $y \geq 3$, où f est définie dans l'énoncé du théorème 1.

D'après (20) et (21), $x \nu_x(\Omega(n) = k)$ est somme de deux termes $\frac{y}{\log y} T_1$ et $\frac{y}{\log^2 y} (\log \log y) T_2$ avec :

$$T_1 = \frac{1}{2\pi i} \int_{|z|=r} \frac{f(2z)}{1-z} z^{-k} (\log y)^{2z} dz$$

$$T_2 = \frac{1}{2\pi i} \int_{|z|=r} O_R\left(\left|\frac{z^{-k-1}}{1-z}\right| (\log y)^{2\text{Re}z}\right) dz.$$

Les méthodes connues d'évaluation asymptotique d'intégrales complexes (méthodes de Laplace, du col, des résidus) peuvent alors être employées avec succès pour estimer T_1 et T_2 . Au prix de calculs assez fins (voir [2], pages 103–109) on peut montrer que :

$$T_1 \sim f(2r) S_{k-1}(2 \log \log y) \text{ et } T_2 = o(T_1)$$

quand $y \rightarrow +\infty$, uniformément en k , où $r = \frac{S_{k-2}(2 \log \log y)}{S_{k-1}(2 \log \log y)}$ (on peut supposer $k \geq 2$).

Pour terminer la démonstration du théorème 1, on montre que

$$(22) \quad 2r \sim \min \left(2, \frac{k-1}{\log \log y} \right), \quad y \rightarrow +\infty$$

uniformément pour $k \geq 2$ (voir le paragraphe VI ci-dessous).

Bien entendu, les méthodes utilisées permettent d'évaluer les termes d'erreurs précisant les équivalences asymptotiques ci-dessus.

IV.— Cas d'autres fonctions.

La fonction $\Omega(n)$ est un exemple de fonction complètement additive valant 0 ou 1 en chaque nombre premier. De façon générale, une telle fonction s'écrit :

$$(23) \quad \Omega_E(n) = \sum_{\substack{p^\nu \parallel n \\ p \in E}} \nu,$$

E désignant un ensemble quelconque de nombres premiers. Le problème du comportement statistique de $\Omega_E(n)$ pour $1 \leq n \leq x$ a été étudié notamment par Halász, Sárközy et Norton ; il intervient dans de nombreuses questions de théorie analytique des nombres, notamment dans des situations où l'ensemble E peut dépendre de x (cf. [7]).

Dans [3] nous avons comparé la loi locale $\nu_x(\Omega_E(n) = k)$ avec la loi Poisson-géométrique de raison $\frac{1}{p_1}$ et de paramètre $E_1(y)$, où :

$$p_1 = \min E; \quad E_1(u) = \sum_{\substack{p_1 < p \leq u \\ p \in E}} \frac{1}{p}; \quad y = x p_1^{-k},$$

et nous obtenons le résultat suivant :

THEOREME 2. *Posons $t = \min(p_1, \frac{k}{E_1(y)})$. Il existe des constantes positives absolues c_i ($4 \leq i \leq 9$) telles que*

$$\nu_x(\Omega_E(n) = k) \leq c_4 p_1^{-k} \exp(c_5 t - E_1(y)) S_k(p_1 E_1(y))$$

pour tout $x \geq 1$, tout E et tout entier naturel k ;

$$\nu_x(\Omega_E(n) = k) \geq c_6 p_1^{-k} \exp(-c_7 t \log(1+t) - E_1(y)) S_{k-1}(p_1 E_1(y))$$

pour tout $x \geq 1$, tout E et tout entier $k \geq 1$ tels que

$$E_1(y) \geq c_8 t \log(1+t) + c_9 .$$

On peut aussi compter une seule fois chaque diviseur premier de n et considérer la fonction

$$\omega(n) = \sum_{p|n} 1 .$$

Pour cette fonction, le résultat de Sathe et Selberg est

$$(24) \quad \nu_x(\omega(n) = k) \sim G\left(\frac{k-1}{\log \log x}\right) (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!}$$

quand $x \rightarrow +\infty$, uniformément pour $1 \leq k \leq B \log \log x$ où B est positif et fixé (mais quelconque) et $G(z) = \frac{1}{\Gamma(z+1)} \prod_p \left(1 - \frac{1}{p}\right)^z \left(1 + \frac{z}{p-1}\right)$ Hensley a montré (cf.

[9]) que (24) est vraie si et seulement si $k = o((\log \log x)^2 (\log \log \log x)^{-2})$. Récemment, Hildebrand et Tenenbaum ont donné une formule asymptotique pour $\nu_x(\omega(n) = k)$ valable pour $k \ll \log x (\log \log x)^{-2}$ cf. [10]. L'intervalle des valeurs de k pour lesquelles $\nu_x(\omega(n) = k) \neq 0$ est $0 \leq k \leq (1+o(1)) \log x (\log \log x)^{-1}$. Le traitement analytique de ce problème est nettement plus difficile que la démonstration de notre théorème 1 et nécessite l'emploi d'une méthode du col en deux variables complexes (cf. également [20]).

V.— Questions d'unimodalité.

Une suite (u_k) de réels ≥ 0 , définie pour k appartenant à un intervalle de \mathbb{Z} , est dite unimodale s'il existe $k_0 \in \mathbb{Z} \cup \{-\infty, +\infty\}$ tel que $u_k \leq u_{k+1}$ pour $k < k_0$, et $u_k \geq u_{k+1}$ pour $k \geq k_0$. L'article [12] de P. Medgyessy présente de façon agréable les faits fondamentaux concernant l'unimodalité des suites. Illustrons ces faits par les exemples suivants :

1) La loi de Poisson de paramètre λ , $p_k = e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!}$ définie pour $k \geq 1$, est unimodale. Son sommet est atteint en $k_0 = [\lambda] + 1$ (et aussi en $k_0 = \lambda$ si λ est entier).

2) La loi géométrique de raison r , $g_k = (1-r)r^k$, définie pour $k \geq 0$, est unimodale. Son sommet est atteint en $k = 0$.

3) Plus généralement, toute loi positive et log-concave (c'est-à-dire vérifiant $u_k^2 \geq u_{k-1} u_{k+1} > 0$ pour tout k) est unimodale.

4) Le produit de convolution de deux lois unimodales ne l'est pas forcément. Posons $u_0 = \frac{2}{3}$, $u_1 = u_2 = \frac{1}{6}$ et $u_k = 0$ si $k \neq 0, 1, 2$. Il est facile de vérifier que $u * u$ n'est pas unimodale. En revanche, le produit de convolution de deux lois positives et log-concaves est positif, log-concave et donc aussi unimodal. Ainsi la loi Poisson-géométrique est unimodale, ce qui n'est pas évident directement.

5) Soit x un réel. Pour tout entier k , soit u_k la somme des inverses des entiers positifs n vérifiant

$$p | n \Rightarrow p \leq x; \quad n \text{ sans facteur carré}; \quad \omega(n) = k.$$

La suite (u_k) est unimodale. En effet

$$P(z) = \sum_k u_k z^k = \prod_{p \leq x} \left(1 + \frac{z}{p}\right)$$

et

$$P(z)(1-z) = u_0 + (u_1 - u_0)z + (u_2 - u_1)z^2 + \dots$$

Or une version de la règle de Descartes affirme que si toutes les racines d'un polynôme à coefficients réels sont réelles, le nombre de racines positives est

égal au nombre de changements de signe dans la suite des coefficients du polynôme. Appliqué au polynôme $P(z)(1-z)$, cet énoncé montre qu'il y a un unique changement de signe dans la suite $u_0, u_1 - u_0, u_2 - u_1, \dots$ donc que la suite (u_k) est unimodale.

En 1948, Erdős a montré dans [5] que les suites

$$\sum_{\substack{n \leq x \\ \omega(n)=k}} \frac{1}{n}, \quad \sum_{\substack{n \leq x \\ \omega(n)=k}} \frac{\mu(n)^2}{n}, \quad \sum_{\substack{n \leq x \\ \Omega(n)=k}} \frac{1}{n}$$

sont unimodales si x est assez grand et conjecturé qu'il en est de même pour les suites

$$\sum_{\substack{n \leq x \\ \omega(n)=k}} 1, \quad \sum_{\substack{n \leq x \\ \omega(n)=k}} \mu(n)^2, \quad \sum_{\substack{n \leq x \\ \Omega(n)=k}} 1.$$

Nous confirmons cette conjecture pour cette dernière suite.

THEOREME 3. *Si x est assez grand, la loi locale $\nu_x(\Omega(n) = k)$ est unimodale.*

Démonstration : Si $k \geq \frac{\log x}{\log 3}$ et si $n \leq x$ vérifie $\Omega(n) = k + 1$, alors n est pair sinon $n \geq 3^{\Omega(n)} > x$; l'entier $\frac{n}{2}$ est lui aussi $\leq x$ et vérifie $\Omega(n) = k$. Cela prouve que :

$$(25) \quad \text{si } k \geq \frac{\log x}{\log 3}, \nu_x(\Omega(n) = k + 1) \leq \nu_x(\Omega(n) = k).$$

Il nous suffit donc de montrer l'unimodalité de $\nu_x(\Omega(n) = k)$ pour $k < \frac{\log x}{\log 3}$. Avec cette condition, $y = x 2^{-k} > x^{1 - (\log 2)/(\log 3)}$ donc $y \rightarrow +\infty$, uniformément par rapport à k , et

$$(26) \quad 0 \leq \log \log x - \log \log y \leq \log \frac{\log 3}{\log(3/2)}.$$

La formule asymptotique (16) nous donne

$$(27) \quad \frac{\nu_x(\Omega(n)=k+1)}{\nu_x(\Omega(n)=k)} \sim \frac{f(t_2)}{f(t_1)} \frac{\log y}{2 \log(y/2)} \frac{S_k(2 \log \log(y/2))}{S_{k-1}(2 \log \log y)}$$

quand $y \rightarrow +\infty$, uniformément par rapport à k , où :

$$t_1 = \min(2, \frac{k-1}{\log \log y}); \quad t_2 = \min(2, \frac{k}{\log \log(y/2)}).$$

Comme $t_1 - t_2$ tend vers 0 quand $y \rightarrow +\infty$, uniformément par rapport à k , et comme f est continue et positive sur $[0,2]$, on peut supprimer le rapport $\frac{f(t_2)}{f(t_1)}$ dans (27).

D'autre part, le théorème des accroissements finis donne :

$$0 \leq S_k(2 \log \log y) - S_k(2 \log \log(y/2)) \leq S_{k-1}(2 \log \log y) \cdot 2 \log \frac{\log y}{\log(y/2)},$$

donc on peut réécrire (27) sous la forme :

$$(28) \quad \frac{\nu_x(\Omega(n)=k+1)}{\nu_x(\Omega(n)=k)} \sim \frac{1}{2} \frac{S_k(2 \log \log y)}{S_{k-1}(2 \log \log y)} \\ \sim \max\left(\frac{1}{2}, \frac{\log \log y}{k}\right) \text{ d'après (22),}$$

quand $y \rightarrow +\infty$, uniformément par rapport à $k \geq 1$.

Compte tenu de (26), on obtient donc :

$$(29) \quad \frac{\nu_x(\Omega(n)=k+1)}{\nu_x(\Omega(n)=k)} \sim \max\left(\frac{1}{2}, \frac{\log \log x}{k}\right)$$

quand $x \rightarrow +\infty$, uniformément par rapport à k , $1 \leq k < \frac{\log x}{\log 3}$.

Cela prouve que si $\delta \in]0,1[$ est fixé et $x \geq x_0(\delta)$, alors

$$\nu_x(\Omega(n) = k+1) > \nu_x(\Omega(n) = k) \quad \text{si } k \leq (1-\delta) \log \log x \\ \text{et } \nu_x(\Omega(n) = k+1) < \nu_x(\Omega(n) = k) \quad \text{si } k \geq (1+\delta) \log \log x.$$

Pour compléter la démonstration, il nous reste à prouver l'unimodalité de $\nu_x(\Omega(n)=k)$ dans le domaine

$$(1-\delta)\log\log x < k < (1+\delta)\log\log x.$$

Dans la formule de Sathe–Selberg (6), on sait que le rapport des deux termes de l'équivalence asymptotique est $1 + O_\epsilon\left(\frac{1}{\log\log x}\right)$. On obtient donc :

$$\frac{\nu_x(\Omega(n)=k+1)}{\nu_x(\Omega(n)=k)} = \frac{\log\log x}{k} \left(1 + O_\epsilon\left(\frac{1}{\log\log x}\right)\right)$$

pour $1 \leq k \leq (2-\epsilon)\log\log x$.

Cela prouve que

$$\nu_x(\Omega(n) = k+1) > \nu_x(\Omega(n) = k) \quad \text{si } k \leq \log\log x - C$$

$$\text{et } \nu_x(\Omega(n) = k+1) < \nu_x(\Omega(n) = k) \quad \text{si } k \geq \log\log x + C$$

où C est une constante absolue inconnue.

Finalement, il reste à démontrer l'unimodalité de $\nu_x(\Omega(n) = k)$ dans le domaine $|k - \log\log x| < C$. Pour cela, nous allons préciser le $O\left(\frac{1}{\log\log x}\right)$ de la formule de Sathe–Selberg à l'aide du lemme suivant.

LEMME. Soit $H(z)$ une fonction holomorphe pour $|z| \leq R$. Posons

$$H(z)e^{tz} = \sum_{k=0}^{+\infty} a_k(t)z^k \quad \text{pour } |z| \leq R.$$

Nous avons alors :

$$(30) \quad a_k(t) = \frac{t^k}{k!} \left\{ H(\rho) - \frac{1}{2} \frac{\rho H''(\rho)}{t} + O\left(\frac{M}{t^2}\right) \right\}$$

uniformément pour $t > 0$ et $0 \leq k \leq Rt$, où $\rho = \frac{k}{t}$ et

$$M = \max_{|s| \leq R} (|s H'''(s)|) + \max_{|s| \leq R} (|s^2 H^{(4)}(s)|).$$

Démonstration du lemme : Par la formule de Cauchy, on a

$$a_k(t) = \frac{1}{2\pi i} \int_{|z|=\rho} H(z) e^{tz} z^{-k-1} dz$$

où la circonférence $|z| = \rho$ est parcourue une fois dans le sens positif, $\rho \leq R$ étant pour l'instant arbitraire.

Utilisons maintenant la formule de Taylor pour $H(z)$:

$$(31) \quad H(z) = H(\rho) + H'(\rho)(z-\rho) + H''(\rho) \frac{(z-\rho)^2}{2} + H'''(\rho) \frac{(z-\rho)^3}{6} + R(z, \rho)$$

$$\text{où } R(z, \rho) = \frac{1}{6} \int_{\rho}^z (z-s)^3 H^{(4)}(s) ds.$$

Les cinq termes du second membre de (31) donnent pour contribution à $a_k(t)$ respectivement :

$$\begin{aligned} & H(\rho) \frac{t^k}{k!}; \\ & H'(\rho) \left\{ \frac{t^{k-1}}{(k-1)!} - \rho \frac{t^k}{k!} \right\} = 0 \quad \text{si } \rho = \frac{k}{t}; \\ & \frac{1}{2} H''(\rho) \left\{ \frac{t^{k-2}}{(k-2)!} - 2\rho \frac{t^{k-1}}{(k-1)!} + \rho^2 \frac{t^k}{k!} \right\} = -\frac{1}{2} \rho H''(\rho) \frac{t^{k-1}}{k!} \quad \text{si } \rho = \frac{k}{t}; \\ & \frac{1}{6} H'''(\rho) \left\{ \frac{t^{k-3}}{(k-3)!} - 3\rho \frac{t^{k-2}}{(k-2)!} + 3\rho^2 \frac{t^{k-1}}{(k-1)!} - \rho^3 \frac{t^k}{k!} \right\} = \frac{1}{3} \rho H'''(\rho) \frac{t^{k-2}}{k!} \quad \text{si } \rho = \frac{k}{t}; \\ & \frac{1}{2\pi i} \int_{|z|=\rho} R(z, \rho) e^{tz} z^{-k-1} dz. \end{aligned}$$

Notons I cette dernière intégrale. Nous avons la majoration :

$$|R(z, \rho)| \leq A |z-\rho|^4$$

où $A = \frac{1}{24} \max_{|s| \leq R} |H^{(4)}(s)|$. Par conséquent :

$$\begin{aligned} |I| &\leq \frac{A}{2\pi} \int_{-\pi}^{\pi} (2\rho \sin \frac{\theta}{2})^4 e^{t\rho \cos \theta} \rho^{-k} d\theta \\ &\leq \frac{A}{2\pi} \rho^{4-k} e^{t\rho} \int_{-\infty}^{\infty} \theta^4 e^{-t\rho c\theta^2} d\theta \quad (c \text{ constante } > 0 \text{ absolue}) \\ &\ll A \rho^{\frac{3}{2}-k} \frac{e^{t\rho}}{t^{5/2}} \ll A \frac{t^k}{k!} \frac{\rho^2}{t^2} \quad \text{si } \rho = \frac{k}{t}, \end{aligned}$$

d'après la formule de Stirling.

En regroupant ces estimations on obtient bien l'énoncé du lemme.

Pour utiliser ce lemme, rappelons la formule servant de point de départ pour la démonstration du résultat de Sathé-Selberg (6) :

$$(32) \quad \sum_{n \leq x} z^{\Omega(n)} = z F(z) x(\log x)^{z-1} + Q(x, z)$$

où $|Q(x, z)| = O_R(x(\log x)^{\operatorname{Re} z - 2})$ pour $x \geq 3$ et $|z| \leq R$, R étant < 2 .

$$\text{Ecrivons } F(z)(\log x)^z = \sum_{k=0}^{+\infty} a_k(\log \log x) z^k \text{ pour } |z| < 2.$$

La formule de Cauchy et (32) nous donnent

$$(33) \quad \nu_x(\Omega(n) = k) = (\log x)^{-1} a_{k-1}(\log \log x) + \frac{1}{2\pi i} \int_{|z|=\rho} x^{-1} Q(x, z) z^{-k-1} dz$$

pour tout $\rho < 2$.

Choisissons $R = \frac{3}{2}$ et supposons $k-1 \leq \frac{3}{2} \log \log x$. En prenant $\rho = \frac{k-1}{\log \log x}$, la contribution de l'intégrale de (33) peut être majorée comme dans la démonstration du lemme :

$$\begin{aligned}
\left| \frac{1}{2\pi i} \int_{|z|=\rho} x^{-1} Q(x,z) z^{-k-1} dz \right| &\ll (\log x)^{-2} \rho^{-k} \int_{-\pi}^{\pi} e^{\rho \cos \theta} \log \log x d\theta \\
&\ll \rho^{-k} (\log x)^{\rho-2} (\rho \log \log x)^{-\frac{1}{2}} \\
&\ll \frac{(\log \log x)^{k-1}}{(k-1)!} (\log x)^{-2} \log \log x .
\end{aligned}$$

Par conséquent, (30) et (33) donnent

$$(34) \quad \nu_x(\Omega(n)=k) = (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!} \left\{ F(\rho) - \frac{1}{2} \frac{\rho F''(\rho)}{\log \log x} + O((\log \log x)^{-2}) \right\}$$

pour $x \geq 3$ et $0 \leq k-1 \leq \frac{3}{2} \log \log x$, où $\rho = \frac{k-1}{\log \log x}$.

En supposant maintenant que $-C < \Delta_k = k - \log \log x < C$, on obtient :

$$\nu_x(\Omega(n)=k) = (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!} \left\{ 1 + \frac{F'(1)(\Delta_k - 1) - F''(1)/2}{\log \log x} + O((\log \log x)^{-2}) \right\}$$

et

$$\begin{aligned}
\nu_x(\Omega(n)=k+1) &= (\log x)^{-1} \frac{(\log \log x)^k}{k!} \left\{ 1 + \frac{F'(1)\Delta_k - F''(1)/2}{\log \log x} + O((\log \log x)^{-2}) \right\} \\
&= (\log x)^{-1} \frac{(\log \log x)^{k-1}}{(k-1)!} \left\{ 1 + \frac{\Delta_k(F'(1)-1) - F''(1)/2}{\log \log x} + O((\log \log x)^{-2}) \right\}
\end{aligned}$$

Par soustraction :

$$\begin{aligned}
\nu_x(\Omega(n)=k+1) - \nu_x(\Omega(n)=k) &= (\log x)^{-1} \frac{(\log \log x)^{k-2}}{(k-1)!} \left\{ F'(1) - \Delta_k + \right. \\
&\quad \left. + O((\log \log x)^{-1}) \right\} .
\end{aligned}$$

Ainsi, il existe une constante absolue $D > 0$, telle que

$$\nu_x(\Omega(n)=k+1) > \nu_x(\Omega(n)=k) \quad \text{si } k < \log \log x + F'(1) - \frac{D}{\log \log x} = A_x$$

et

$$\nu_x(\Omega(n)=k+1) < \nu_x(\Omega(n)=k) \quad \text{si } k > \log \log x + F'(1) - \frac{D}{\log \log x} = B_x .$$

Pour x assez grand, l'intervalle $[A_x, B_x]$ contient au plus un entier. S'il n'en contient pas, l'unimodalité est démontrée et $\nu_x(\Omega(n)=k)$ atteint son maximum pour $k = [A_x] + 1 = [B_x] + 1$. Si $[A_x, B_x]$ contient un entier k_0 , l'unimodalité reste acquise mais on ne connaît pas la position du maximum : suivant que $\nu_x(\Omega(n)=k_0+1) - \nu_x(\Omega(n)=k_0)$ est positif, négatif ou nul, ce maximum est atteint pour $k = k_0 + 1$, $k = k_0$ ou les deux.

Remarques.

1°) Notons $k(x)$ le plus petit entier k tel que $\nu_x(\Omega(n)=k)$ soit maximal. La démonstration du théorème 3 permet de conclure que $k(x)$ est une fonction non-décroissante de x , pour x assez grand. De plus on a $\nu_x(\Omega(n)=k(x)) = \nu_x(\Omega(n)=k(x)+1)$ pour une infinité d'entiers x . Ces observations sont dues à P. Erdős : nous laissons au lecteur les raisonnements très simples qui y mènent.

2°) Il serait intéressant de connaître le plus petit x_0 pour lequel la suite $(\nu_x(\Omega(n)=k))$ est unimodale si $x \geq x_0$. Il faudrait pour cela expliciter numériquement la formule de Selberg (32) et obtenir un algorithme efficace pour tester l'unimodalité de cette suite pour les premières valeurs de x .

3°) Concernant la fonction ω , la même démonstration que celle du théorème 3 et les résultats de Hildebrand et Tenenbaum prouvent que $\nu_x(\omega(n)=k)$ est unimodale dans le domaine $k \ll (\log x)(\log \log x)^{-2}$. Pour conclure, il manque un argument, peut-être élémentaire, permettant de montrer que :

$$\nu_x(\omega(n)=k+1) \leq \nu_x(\omega(n)=k) \text{ si } k \gg (\log x)(\log \log x)^{-2}.$$

4°) Concernant la fonction Ω_E , nous posons la question suivante : existe-t-il une constante absolue K telle que si $\sum_{\substack{p \leq x \\ p \in E}} \frac{1}{p} \geq K$, alors $\nu_x(\Omega(n)=k)$ est unimodale ?

Pour d'autres résultats d'unimodalité en théorie des nombres, le lecteur peut consulter les articles d'Odlyzko et Richmond [16], [17] et [18].

V.— Démonstration de (22).

Il s'agit de montrer que

$$(35) \quad \frac{S_{k-1}(X)}{S_k(X)} \sim \min\left(1, \frac{k}{X}\right),$$

quand $X \rightarrow +\infty$, uniformément pour $k \geq 1$.

Observons d'abord que $\frac{S_{k-1}(X)}{S_k(X)} \leq \min\left(1, \frac{k}{X}\right)$.

$$\begin{aligned} \text{Pour } k < X \text{ on a } S_k(X) &\leq \frac{X^k}{k!} \left(1 + \frac{k}{X} + \frac{k^2}{X^2} + \dots\right) \\ &= \frac{X^k}{k!} \frac{1}{1 - \frac{k}{X}} \end{aligned}$$

et d'autre part

$$S_k(X) = \frac{1}{2\pi i} \int_{|z|=\rho} e^{zX} \frac{z^{-k-1}}{1-z} dz \quad \text{pour tout } \rho < 1.$$

En écrivant $\frac{1}{1-z} = \frac{1}{1-\rho} + \frac{z-\rho}{(1-\rho)^2} + \frac{(z-\rho)^2}{(1-z)(1-\rho)^2}$, l'intégrale est somme de trois termes. Le premier vaut $\frac{1}{1-\rho} \frac{X^k}{k!}$; le deuxième est nul si $\rho = \frac{k}{X}$ et le troisième est

$$\ll \frac{X^k}{k!} \frac{\rho}{X(1-\rho)^3} \quad \text{si } \rho = \frac{k}{X},$$

par la méthode de Laplace et la formule de Stirling, comme dans la démonstration du théorème 3.

Ainsi

$$(36) \quad S_k(X) = \frac{1}{1-\rho} \frac{X^k}{k!} \left(1 + O\left(\frac{\rho}{X(1-\rho)^2}\right)\right)$$

uniformément pour $X > 0$ et $\rho = \frac{k}{X} < 1$.

Soit maintenant $\epsilon > 0$. D'après (36) il existe une constante $B_1(\epsilon) > 0$ telle que

$$S_{k-1}(X) \geq \frac{1}{1 - \frac{k-1}{X}} \frac{X^{k-1}}{(k-1)!} (1 - \epsilon)$$

$$\text{si } k \leq X - B_1(\epsilon) \sqrt{X}.$$

Par conséquent

$$\begin{aligned} \frac{S_{k-1}(X)}{S_k(X)} &\geq \frac{k}{X} \frac{X-k}{X-k+1} (1-\epsilon) \\ &\leq \frac{k}{X} (1-2\epsilon) \quad \text{si } k \leq X - B_2(\epsilon) \sqrt{X}. \end{aligned}$$

D'autre part la formule de Stirling montre que $\frac{X^k}{k!} \ll \frac{e^X}{\sqrt{X}}$ pour tout k et

iii) (§ II) montre que

$$S_k(X) \gg_\epsilon e^X \quad \text{si } k > X - B_2(\epsilon) \sqrt{X}.$$

$$\begin{aligned} \text{Ainsi : } \frac{S_{k-1}(X)}{S_k(X)} &= 1 - \frac{X^k/k!}{S_k(X)} \geq 1 + O_\epsilon\left(\frac{1}{\sqrt{X}}\right) \\ &\geq (1-2\epsilon) \min\left(1, \frac{k}{X}\right) \end{aligned}$$

si $X \geq X_0(\epsilon)$ et $k > X - B_2(\epsilon) \sqrt{X}$.

$$\text{Finalement, } \frac{S_{k-1}(X)}{S_k(X)} \geq (1-2\epsilon) \min\left(1, \frac{k}{X}\right)$$

si $X \geq X_0(\epsilon)$ et pour tout $k \geq 1$: (35) est démontrée. Bien entendu, on peut rendre cette démonstration effective et fournir un terme d'erreur uniforme pour (35).

BIBLIOGRAPHIE

- [1] R. Balasubramanian et K. Ramachandra.— *On the number of integers n such that $nd(n) \leq x$* , Acta Arith. 49 (1988), 313–322.
- [2] M. Balazard.— *Sur la répartition des valeurs de certaines fonctions arithmétiques additives*, Thèse (Université de Limoges 1987).
- [3] M. Balazard.— *Remarques sur un théorème de G. Halasz et A. Sarközy*, Prépublication.
- [4] M. Balazard, H. Delange et J.-L. Nicolas.— *Sur le nombre de facteurs premiers des entiers*, C.R.A.S. 306 série I (1988), 511–514.
- [5] P. Erdős.— *On the integers having exactly k prime factors*, Ann. of Math. 49 (1948), 53–66.
- [6] P. Erdős et A. Sárközy.— *On the number of prime factors of integers*, Acta Sci. Math. 42 (1980), 237–246.
- [7] R.R. Hall et G. Tenenbaum.— *Divisors*, Cambridge University Press, 1988.
- [8] G.H. Hardy et S. Ramanujan.— *The normal number of prime factors of an integer*, Quarterly J. Math. 48 (1917), 76–92.
- [9] D. Hensley.— *The distribution of round numbers*, Proc. London Math. Soc. 54 (1987), 412–444.
- [10] A. Hildebrand et G. Tenenbaum.— *On the number of prime factors of an integer*, Duke Math. J. 56 (1988), 471–501.
- [11] E. Landau.— *Sur quelques problèmes relatifs à la distribution des nombres premiers*, Bull. Soc. Math. France 28 (1900), 25–38.
- [12] P. Medgyessy.— *On the unimodality of discrete distributions*, Period. Math. Hung. 2 (1972) 245–257.

- [13] J.-L. Nicolas.— *Sur la distribution des entiers ayant une quantité fixée de facteurs premiers*, Acta Arith. 44 (1984), 191–200.
- [14] K.K. Norton.— *On the number of restricted prime factors of an integers III*, L'Ens. Math. 28 (1982) 31–52.
- [15] K.K. Norton.— *Estimates for partial sums of the exponential series*, J. of Math. An. and Ap. 63 (1978), 265–296.
- [16] A.M. Odlyzko and L.B. Richmond.— *On the unimodality of high convolutions of discrete distributions*, Annals of Prob. 13 (1985), 299–306.
- [17] A.M. Odlyzko and L.B. Richmond.— *On the unimodality of some partition polynomials*, Europ. J. Comb. 3 (1982), 69–84.
- [18] A.M. Odlyzko and L.B. Richmond.— *On the compositions of an integer*, in : Combinatorial Mathematics VII, Springer Lecture Notes 829.
- [19] A. Selberg.— *Note on a paper by L.G. Sathe*, J. Indian Math. Soc. 18 (1954), 83–87.
- [20] G. Tenenbaum.— *La méthode du col en théorie analytique des nombres*, Sémin. Théorie des Nombres de Paris (1986–87), Birkhäuser (1988).

Michel Balazard
Département de Mathématiques
Faculté des Sciences
123, rue A. Thomas
87060 LIMOGES CEDEX

