Mh. Math. 116, 13–37 (1993)

Monatshefte für Mathematik

Sur les plus grands facteurs premiers d'un entier

Par

Jean-Marie De Koninck, Québec

(Received 25 September 1992; in revised form 2 November 1992)

Abstract. On the Largest Prime Factor of an Integer. Let $\Omega(n) = \sum_{p^\alpha \| n} \alpha$ and, for each integer n such that $\Omega(n) \geqslant k$, denote by $P_k(n)$ its k^{th} largest prime factor. Further, given a set of primes Q of positive density $\delta < 1$ satisfying a certain regularity condition, define $P(n, Q)$ as the largest prime divisor of n belonging to Q, assuming that $P(n, Q) = +\infty$ if no such prime factor exists. We provide estimates of $\sum_{2^k \leqslant n \leqslant x, \Omega(n) \geqslant k} 1/P_k(n)$, for $k \geqslant 2$, and of $\sum_{n \leqslant x} 1/P(n, Q)$. We also study the median value of the function $P(n, Q)$ and that of the function $P_k(n)$ for each $k \geqslant 1$.

1. Introduction

Etant donné un entier $n > 1$, on notera $p(n)$ le plus petit facteur premier de n et $P(n) (= P_1(n))$ le plus grand facteur premier de n. De manière plus générale, pour chaque entier $k \geqslant 1$ et pour chaque entier positif n tel que $\Omega(n) \geqslant k$ (ici $\Omega(n) = \sum_{p^\alpha \| n} \alpha$), on notera $P_k(n)$ le $k^{\text{ième}}$ plus grand facteur premier de n. Ainsi, pour un entier $n > 1$ tel que $\Omega(n) = l$, on aura

$$p(n) = P_l(n) \leqslant P_{l-1}(n) \leqslant \cdots \leqslant P_2(n) \leqslant P_1(n) = P(n).$$

Par convention, on écrira $P(1) = 1$.

La fonction $P(n)$ a fait l'objet de nombreux articles. Elle a surtout été étudiée à travers la fonction

$$\psi_1(x, y) = \psi(x, y) \overset{\text{def}}{=} \#\{n \leqslant x : P(n) \leqslant y\}.$$

Le meilleur estimé concernant le comportement asymptotique de $\psi(x, y)$ est celui obtenu par Hildebrand ([10]) en 1986, à l'effet que

$$\psi(x, y) = x\rho(u)\left(1 + O\left(\frac{\log(u+1)}{\log y}\right)\right), \tag{1.1}$$

où $u = \log x/\log y$, uniformément pour chaque $\varepsilon > 0$ dans le domaine

$$\exp\{(\log\log x)^{5/3+\varepsilon}\} \leqslant y \leqslant x,$$

$\rho(u)$ étant la solution continue de l'équation différentielle aux différences

$$u\rho'(u) = -\rho(u-1)$$

avec la condition initiale $\rho(u) = 1$ pour $0 \leqslant u \leqslant 1$. Pour une étude détaillée de la fonction $\psi(x, y)$, on peut consulter le tout récent livre de TENENBAUM ([15]).

La fonction $P_2(n)$, bien que moins "*populaire*" que la fonction $P(n)$, a aussi été étudiée (voir le livre de RIESEL [14]). Elle occupe une place importante dans plusieurs problèmes de théorie des nombres, en particulier dans le problème de la factorisation des grands nombres (voir HAFNER et MCCURLEY [9] et KNUTH et PARDO [12]).

L'étude de la fonction $P(n)$, principalement à travers celle de la fonction $\psi(x, y)$, est relativement complexe. Ceci provient en partie du fait que la fonction $P(n)$ est très erratique. Le comportement moyen de la fonction $P_k(n)$ a été étudié par ALLADI et ERDÖS ([1]). Ils ont obtenu

$$\sum_{n\leqslant x} P_k(n) = A_k \frac{x^{1+1/k}}{\log^k x} + O\left(x^{1+1/k}\frac{\log\log x}{\log^{k+1} x}\right) \quad (A_k > 0)$$

pour $k \geqslant 1$ fixé. La formule a été améliorée pour $k = 1$ par DE KONINCK et IVIĆ ([2]) (on a $A_1 = \pi^2/12$) et pour $k \geqslant 2$ par IVIĆ ([11]). Alors que le comportement moyen de la fonction $P(n)$ est assez facile à obtenir, celui de la fonction $1/P(n)$ est très difficile à saisir; en 1986, ERDÖS, IVIĆ et POMERANCE ([5]) ont tout de même démontré que

$$\sum_{n\leqslant x} \frac{1}{P(n)} = (1 + o(1))\frac{x}{L(x)}, \tag{1.2}$$

où

$$L(x)^{-1} = \int_2^x \rho\left(\frac{\log x}{\log t}\right) t^{-2}\, dt. \tag{1.3}$$

En utilisant l'estimé

$$\rho(u) = \exp\{-u(\log u + \log\log u - 1 + o(1))\}, \tag{1.4}$$

on peut démontrer que $L(x)$ est une fonction dite "*à oscillation lente*" (i.e. telle que $\lim_{x\to\infty}(L(cx)/L(x))=1$ pour chaque constante $c>0$) et de plus que

$$L(x)=\exp\{(1+o(1))\sqrt{2\log x\log\log x}\}. \tag{1.5}$$

De son côté, l'analyse du comportement de $P_2(n)$ repose le plus souvent sur l'étude de la fonction

$$\psi_2(x,y)\overset{\text{def}}{=}\#\{n\leqslant x:P_2(n)\leqslant y\}$$

(voir [14]).

Le meilleur estimé concernant le développement asymptotique de $\psi_2(x,y)$ est certes celui dû à [9]:

$$\psi_2(x,y)=x\rho_2(u)\left(1+O\left(\frac{1}{\log y}\right)\right), \tag{1.6}$$

valable uniformément pour $2\leqslant y\leqslant x$, où $\rho_2(u)$ est la solution continue de l'équation différentielle aux différences

$$u\rho_2'(u)+\rho_2(u-1)=\rho(u-1)$$

avec la condition initiale $\rho_2(u)=1$ pour $0\leqslant u\leqslant 1$. Alors que, comme en témoigne la relation (1.4), la fonction $\rho(u)$ décroît très rapidement vers 0 lorsque $u\to\infty$, la fonction $\rho_2(u)$ tend vers 0 beaucoup plus lentement puisqu'elle satisfait à

$$\rho_2(u)=\frac{e^\gamma}{u}\left(1+O\left(\frac{1}{u}\right)\right),\quad (u\geqslant 1) \tag{1.7}$$

(voir [12]). D'une manière générale, on peut démontrer

$$\psi_k(x,y)\overset{\text{def}}{=}\#\{n\leqslant x:P_k(n)\leqslant y\}=x\rho_k(u)\left(1+O\left(\frac{1}{\log y}\right)\right), \tag{1.8}$$

uniformément pour $2\leqslant y\leqslant x$, où $\rho_k(u)$ est la solution continue de l'équation différentielle aux différences

$$u\rho_k'(u)+\rho_k(u-1)=\rho_{k-1}(u-1), \tag{1.9}$$

avec la condition initiale $\rho_k(u)=1$ pour $0\leqslant u\leqslant 1$.

Nous nous intéressons d'abord ici au comportement asymptotique de

$$S_k(x) \stackrel{\text{def}}{=} \sum_{\substack{2^k \leqslant n \leqslant x \\ \Omega(n) \geqslant k}} \frac{1}{P_k(n)} \quad (k \geqslant 2).$$

En 1982, Erdös et Ivić ([4]) ont étudié des sommes analogues aux $S_k(x)$; de leurs résultats, on peut déduire l'existence d'une constante λ_2 telle que

$$S_2(x) = \lambda_2 \frac{x}{\log x}\left(1 + O\left(\frac{(\log\log x)^2}{\log x}\right)\right)$$

et également, pour chaque entier $k \geqslant 3$, l'existence de constantes λ_k pour lesquels

$$S_k(x) = (1 + o(1))\lambda_k \frac{x}{\log x}(\log\log x)^{k-2};$$

les preuves utilisent des résultats assez compliqués. Nous obtenons ici des développements asymptotiques plus précis pour $S_k(x)$ $(k \geqslant 2)$, avec des constantes λ_k explicites. Notre approche a l'avantage d'être relativement simple et de ne faire appel qu'au théorème des nombres premiers.

En second lieu, étant donné un ensemble Q de nombres premiers de densité positive $\delta < 1$ satisfaisant la condition de régularité

$$\pi(x, Q) \stackrel{\text{def}}{=} \sum_{p \leqslant x, p \in Q} 1 = \delta \operatorname{Li}(x) + O\left(\frac{x}{\log^B x}\right), \tag{1.10}$$

où $\operatorname{Li}(x) = \int_2^x (dt/\log t)$ et $B > 2$ est une constante, nous analysons le comportement de la somme $\sum_{n \leqslant x} 1/P(n, Q)$, où pour chaque $n \geqslant 2$, $P(n, Q) = +\infty$ si aucun des facteurs premiers de n n'est dans Q et autrement $P(n, Q) = \max\{p : p \mid n \wedge p \in Q\}$ (par convention, on écrira $P(1, Q) = 1$).

En troisième lieu, nous considérons une généralisation de nos estimés de $\sum 1/P_k(n)$ et de $\sum 1/P(n, Q)$.

Enfin, nous étudions la valeur médiane de la fonction $P(n, Q)$ ainsi que celle de la fonction $P_k(n)$, pour chaque entier $k \geqslant 1$, laquelle notion a déjà été considérée par Wunderlich et Selfridge ([16]) dans le cas de $P(n)$.

2. Les principaux résultats

Théorème 1. *Etant donné un entier $A \geqslant 1$, il existe des constantes $\lambda_2^{(1)}, \lambda_2^{(2)}, \ldots, \lambda_2^{(A)}$ telles que*

$$\sum_{\substack{4 \leqslant n \leqslant x \\ \Omega(n) \geqslant 2}} \frac{1}{P_2(n)} = \lambda_2^{(1)} \frac{x}{\log x} + \lambda_2^{(2)} \frac{x}{\log^2 x} + \cdots + \lambda_2^{(A)} \frac{x}{\log^A x} + O\left(\frac{x}{\log^{A+1} x}\right), \tag{2.1}$$

où $\lambda_2^{(1)} = \lambda_2 = \sum_{m=1}^{\infty} \frac{1}{m} \sum_{p \geqslant P(m)} \frac{1}{p^2} = \sum_p \frac{1}{p^2} \prod_{q \leqslant p} \left(1 - \frac{1}{q}\right)^{-1} = 1.254\ldots.$

Théorème 2. *Etant donné un entier $k \geqslant 2$, on a*

$$\sum_{\substack{2^k \leqslant n \leqslant x \\ \Omega(n) \geqslant k}} \frac{1}{P_k(n)} = \lambda_k \frac{x(\log\log x)^{k-2}}{\log x}\left(1 + O\left(\frac{1}{\log\log x}\right)\right), \tag{2.2}$$

où $\lambda_k = \dfrac{1}{(k-2)!} \lambda_2.$

Théorème 3. *Soit Q un ensemble de nombres premiers de densité $0 < \delta < 1$ et satisfaisant la condition (1.10) et soit q_0 le plus petit élément de Q, alors il existe une constante positive $\eta(Q)$ telle que*

$$\sum_{q_0 \leqslant n \leqslant x} \frac{1}{P(n, Q)} = \left(1 + O\left(\frac{1}{\log\log x}\right)\right) \eta(Q) \frac{x}{(\log x)^{\delta}}. \tag{2.3}$$

Remarques. La nature assez compliquée de la fonction $L(x)$ définie en (1.3) est en partie reliée à la nature même de la fonction de Dickman $\rho(u)$, dont le comportement asymptotique n'est pas facile à cerner. En fait, cela s'explique en partie par le fait que $P(n)$ est très erratique lorsque n parcourt l'ensemble des nombres naturels.

Les fonctions $P_k(n)$ $(k \geqslant 2)$ sont un peu moins erratiques que la fonction $P(n)$; cela a pour conséquence que les sommes de leurs réciproques se comportent de manière plus régulière.

On notera par ailleurs que, dans le calcul des sommes $\sum 1/P_k(n)$, chaque entier positif n pour lequel $\Omega(n) \geqslant k$ et ayant de "petits" facteurs premiers contribue au terme principal du développement asymptotique donné par (2.2), alors que cela n'est pas du tout le cas pour le calcul de $\sum 1/P(n)$ obtenu en (1.2).

Il est intéressant d'observer, en comparant les relations (1.2) et (1.5) avec (2.3), que $\sum 1/P(n,Q)$ a un comportement complètement différent selon que Q est l'ensemble de tous les nombres premiers (donc avec densité $\delta = 1$) ou que Q a une densité $\delta < 1$. Ainsi l'irrégularité de $\sum_{n \leqslant x} 1/P(n)$ est considérablement diminuée lorsque, dans le calcul du plus grand facteur premier, on ignore une proportion finie des nombres premiers.

Enfin, il est intéressant de noter que les estimés (2.2) et (2.3) sont en accord avec un phénomène tout à fait attendu, soit celui selon lequel, quelque soit k fixe, la quantité $P(n,Q)$ est la plupart du temps petite que $P_k(n)$.

3. Les résultats préalabes aux démonstrations des théorèmes 1 et 2

Les deux premiers théorèmes reposent sur quelques résultats préliminaires que l'on énumère sous forme de lemmes.

Lemme 1. *Pour* $2 \leqslant y \leqslant x$, *on a*

$$\psi(x,y) \ll xe^{-1/2(\log x/\log y)}. \tag{3.1}$$

Démonstration. Voir [15].

Dans tout ce qui suit, A est un nombre entier positif arbitraire mais fixe et les nombres $p, q, q_1, q_2, \ldots$ désignent toujours des nombres premiers, alors que l'on écrira $l(x)$ pour désigner la quantité $\log^{A+2} x$.

Lemme 2. *Uniformément pour* $0 \leqslant i \leqslant A$, *on a, si* $2 \leqslant y \leqslant x$,

$$\sum_{\substack{m > x \\ P(m) \leqslant y}} \frac{\log^i m}{m} \ll \log^{i+1} y \cdot \log^i x \cdot e^{-1/2(\log x/\log y)}. \tag{3.2}$$

Démonstration. Il est clair que

$$\sum_{\substack{m > x \\ P(m) \leqslant y}} \frac{\log^i m}{m} < \int_x^\infty \frac{\log^i t}{t} d\psi(t,y) \ll \frac{\log^i x}{x} \psi(x,y) + \int_x^\infty \frac{\log^i t}{t^2} \psi(t,y) dt.$$

En utilisant l'estimé (3.1), cette dernière quantité est

$$\ll \log^i x e^{-1/2(\log x/\log y)} + \int_x^\infty \frac{\log^i t e^{-1/2(\log t/\log y)}}{t} dt =$$

$$= \log^i x e^{-1/2(\log x/\log y)} + I_i,$$

disons. Or

$$I_i=\int_{\log x}^{\infty} v^i e^{-1/2(v/\log y)}\,dv=\left.\frac{v^i e^{-1/2(v/\log y)}}{-\dfrac{1}{2\log y}}\right|_{\log x}^{\infty}-i\int_{\log x}^{\infty}\frac{v^{i-1}e^{-1/2(v/\log y)}}{-\dfrac{1}{2\log y}}dv=$$

$$=2\log y\cdot\log^i x\cdot e^{-1/2(\log x/\log y)}+2\log y\cdot i\cdot I_{i-1}.$$

En itérant le processus, on arrive à exprimer successivement I_i en termes de I_{i-2}, puis de I_{i-3}, et ainsi de suite jusqu'à I_0 qui vaut $2\log y e^{-1/2(\log x/\log y)}$. D'où le résultat.

Pour démontrer le lemme qui suit, on fera appel au théorème des nombres premiers sous la forme

$$\pi(x)=\frac{x}{\log x}+\frac{x}{\log^2 x}+\frac{2!x}{\log^3 x}+\cdots+\frac{(A-1)!x}{\log^A x}+O\left(\frac{x}{\log^{A+1} x}\right). \tag{3.3}$$

Lemme 3. *Soit z un nombre réel positif, alors il existe des fonctions $d_1(z), d_2(z),\dots, d_A(z)$, $d_{A+1}(z)$ bornées sur $[1,\infty[$ telles que*

$$\sum_{\substack{pq\leqslant x\\ z\leqslant p\leqslant q\\ p\leqslant l(x)}}\frac{1}{p}=\frac{d_1(z)x}{\log x}+\frac{d_2(z)x}{\log^2 x}+\cdots+\frac{d_A(z)x}{\log^A x}+O\left(\frac{d_{A+1}(z)x}{\log^{A+1} x}\right), \tag{3.4}$$

avec $d_1(z)=\sum_{p\geqslant z}\dfrac{1}{p^2}$.

Démonstration. Soit $R_2(x)$ la somme à évaluer. Si x est suffisamment grand, on a $l(x)<\sqrt{x}$, auquel cas

$$R_2(x)=\sum_{\substack{pq\leqslant x\\ z\leqslant p\leqslant q\\ p\leqslant l(x)}}\frac{1}{p}=\sum_{z\leqslant p\leqslant l(x)}\frac{1}{p}\sum_{p\leqslant q\leqslant x/p}1=$$

$$=\sum_{z\leqslant p\leqslant l(x)}\frac{1}{p}\pi(x/p)-\sum_{z\leqslant p\leqslant l(x)}\frac{\pi(p)}{p}=\Sigma_1-\Sigma_2.$$

Puisque $\pi(p)<p$, on a trivialement

$$\Sigma_2<\sum_{z\leqslant p\leqslant\sqrt{x}}1\leqslant\sqrt{x}.$$

D'autre part, en utilisant le théorème des nombres premiers sous la forme (3.3), on a

$$\Sigma_1 = \sum_{z \leqslant p \leqslant l(x)} \frac{1}{p} \left\{ \frac{x/p}{\log(x/p)} + \frac{x/p}{\log^2(x/p)} + \cdots + \right.$$
$$\left. + \frac{(A-1)!x/p}{\log^A(x/p)} + O\left(\frac{x/p}{\log^{A+1}(x/p)} \right) \right\}.$$

Puisque

$$\sum_{z \leqslant p \leqslant l(x)} \frac{1}{p^2} = \sum_{p \geqslant z} \frac{1}{p^2} - \sum_{p > l(x)} \frac{1}{p^2} = d_1(z) + O\left(\frac{1}{l(x)} \right)$$

et comme, pour chaque $1 \leqslant i \leqslant A+1$,

$$\sum_{z \leqslant p \leqslant l(x)} \frac{\log^i p}{p^2} = \sum_{p \geqslant z} \frac{\log^i p}{p^2} - \sum_{p > l(x)} \frac{\log^i p}{p^2} = \sum_{p \geqslant z} \frac{\log^i p}{p^2} + O\left(\frac{\log^i l(x)}{l(x)} \right),$$

on obtient facilement (3.4), chaque fonction $d_j(z)$ ($j = 1, 2, \ldots, A+1$) étant alors une combinasion linéaire des expressions $\sum_{p \geqslant z} \dfrac{\log^i p}{p^2}$ ($i = 1, 2, \ldots, A+1$).

Lemme 4. *Soit z un nombre réel positif et soit $\xi(z) = \sum_{p \geqslant z} \dfrac{\log \log p}{p^2}$, alors*

$$\sum_{\substack{q_k q_{k-1} \cdots q_1 \leqslant x \\ q_k \leqslant \cdots \leqslant q_1 \\ z \leqslant q_k \leqslant l(x)}} \frac{1}{q_k} = d_1(z) \frac{x}{\log x} \frac{(\log \log x)^{k-2}}{(k-2)!} \left(1 + O\left(\frac{1}{\log \log x} \right) \right) +$$
$$+ O\left(\xi(z) \frac{x}{\log x} (\log \log x)^{k-3} \right). \tag{3.5}$$

Démonstration. Soit $R_k(x)$ la somme à évaluer. Pour x suffisamment grand, on a

$$R_k(x) = \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \sum_{\substack{q_{k-1} \cdots q_1 \leqslant x/q_k \\ q_k \leqslant q_{k-1} \leqslant \cdots \leqslant q_1}} 1 =$$

$$= \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \left\{ \sum_{\substack{q_{k-1} \cdots q_1 \leqslant x/q_k \\ q_{k-1} \leqslant \cdots \leqslant q_1}} 1 - \sum_{\substack{q_{k-1} \cdots q_1 \leqslant x/q_k \\ q_{k-1} \leqslant \cdots \leqslant q_1 \\ q_{k-1} < q_k}} 1 \right\} =$$

$$= \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \left\{ \pi_{k-1}(x/q_k) - \sum_{q_{k-1} < q_k} \sum_{\substack{q_{k-2} \cdots q_1 \leqslant x/q_k q_{k-1} \\ q_{k-1} \leqslant q_{k-2} \leqslant \cdots \leqslant q_1}} 1 \right\} = \Sigma' - \Sigma'',$$

disons, où $\pi_j(\omega) = \#\{n \leqslant \omega : \Omega(n) = j\}$.

En utilisant le fait que

$$\pi_j(x) = \frac{x}{\log x} \frac{(\log\log x)^{j-1}}{(j-1)!}\left(1 + O\left(\frac{1}{\log\log x}\right)\right), \tag{3.6}$$

uniformément pour $1 \leqslant j \leqslant A+1$ (voir par exemple Nicolas [13]), on a

$$\Sigma'' \ll \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \sum_{q_{k-1} < q_k} \pi_{k-2}\left(\frac{x}{q_k q_{k-1}}\right) \ll$$

$$\ll \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \sum_{q_{k-1} < q_k} \frac{x/q_k q_{k-1}}{\log(x/q_k q_{k-1})} (\log\log(x/q_k q_{k-1}))^{k-3} \ll$$

$$\ll \frac{x}{\log x} (\log\log x)^{k-3} \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \sum_{q_{k-1} < q_k} \frac{1}{q_k q_{k-1}} \ll$$

$$\ll \frac{x}{\log x} (\log\log x)^{k-3} \sum_{z \leqslant q_k \leqslant l(x)} \frac{\log\log q_k}{q_k^2} \ll \frac{x}{\log x} (\log\log x)^{k-3} (\xi(z)),$$

puisque $\log\log(x/q_k q_{k-1}) < \log\log x$ et comme

$$\frac{1}{\log(x/q_k q_{k-1})} = \frac{1}{\log x}\left(1 + O\left(\frac{\log\log x}{\log x}\right)\right) \ll \frac{1}{\log x}.$$

De même, en utilisant (3.6) avec $j = k-1$, on obtient

$$\Sigma' = \sum_{z \leqslant q_k \leqslant l(x)} \frac{1}{q_k} \frac{x/q_k}{\log(x/q_k)} \left\{ \frac{(\log\log(x/q_k))^{k-2}}{(k-2)!} + O((\log\log(x/q_k))^{k-3} \right\},$$

laquelle expression, en procédant comme ci-haut, devient

$$\frac{x}{\log x} \frac{(\log\log x)^{k-2}}{(k-2)!} \sum_{p \geqslant z} \frac{1}{p^2}\left(1 + O\left(\frac{1}{\log\log x}\right)\right).$$

En combinant les estimés de Σ' et de Σ'', la relation (3.5) s'ensuit.

4. Les démonstrations des théorèmes 1 et 2

Avant d'établir la démonstration du théorème 1, il nous apparaît opportun de mentionner que l'approche naturelle via la fonction

$\psi_2(x, y)$ ne donne pas le résultat escompté. En effet, on aurait alors

$$S_2(x) = \sum_{p \leqslant \sqrt{x}} \frac{1}{p} \sum_{\substack{pm \leqslant x \\ P_2(pm) = p}} 1 = \sum_{p \leqslant \sqrt{x}} \frac{1}{p} \left\{ \sum_{\substack{m \leqslant x/p \\ P_2(m) \leqslant p}} 1 - \sum_{\substack{m \leqslant x/p \\ P(m) \leqslant p}} 1 \right\} =$$

$$= \sum_{p \leqslant \sqrt{x}} \frac{1}{p} \left\{ \psi_2\left(\frac{x}{p}, p\right) - \psi\left(\frac{x}{p}, p\right) \right\}.$$

En utilisant ensuite les estimés (1.6) et (1.7), de même que (1.1), (1.3) et (1.5), on arrive à montrer que $S_2(x)$ est de l'ordre de $x/\log x$, sans pour autant obtenir une formule asymptotique.

Aussi, notre raisonnement diffère totalement de l'approche ci-haute. Démontrons d'abord le théorème 1. Puisque

$$\sum_{\substack{4 \leqslant n \leqslant x \\ \Omega(n) \geqslant 2}} \frac{1}{P_2(n)} = \sum_{\substack{4 \leqslant n \leqslant x \\ \Omega(n) \geqslant 2 \\ P_2(n) \leqslant l(x)}} \frac{1}{P_2(n)} + O\left(\frac{x}{\log^{A+1} x} \right),$$

il est clair que dans le processus d'évaluation de $\sum 1/P_2(n)$, il suffit de considérer les entiers $n \leqslant x$ pour lesquels $P_2(n) \leqslant l(x)$. On observe ensuite que chaque entier positif n tel que $\Omega(n) \geqslant 2$ peut s'écrire de manière unique sous la forme

$$n = mr,$$

où $\Omega(r) = 2$ et $P(m) \leqslant p(r) = P_2(r)$ avec possiblement $m = 1$. On a donc

$$S_2(x) + O\left(\frac{x}{\log^{A+1} x} \right)$$

$$= \sum_{\substack{mr \leqslant x \\ \Omega(r) = 2 \\ P(m) \leqslant P_2(r) \leqslant l(x)}} \frac{1}{P_2(r)} =$$

$$= \sum_{\substack{m \leqslant x/\log x \\ P(m) \leqslant l(x)}} \sum_{\substack{r \leqslant x/m \\ \Omega(r) = 2 \\ P(m) \leqslant P_2(r) \leqslant l(x)}} \frac{1}{P_2(r)} + \sum_{\substack{x/\log x < m \leqslant x/4 \\ P(m) \leqslant l(x)}} \sum_{\substack{r \leqslant x/m \\ \Omega(r) = 2 \\ P(m) \leqslant P_2(r) \leqslant l(x)}} \frac{1}{P_2(r)} =$$

$$= T_1(x) + T_2(x),$$

disons.

Avec l'aide du lemme 1, on a

$$T_2(x) \ll \log x \sum_{\substack{m \leqslant x/4 \\ P(m) \leqslant l(x)}} 1 = \log x\, \psi\left(\frac{x}{4}, l(x)\right) \ll \frac{x}{\log^{A+1} x}.$$

Il nous suffit donc d'évaluer $T_1(x)$. Avec le lemme 3, on obtient

$$T_1(x) = \sum_{\substack{m \leqslant x/\log x \\ P(m) \leqslant l(x)}} \left\{ d_1(P(m)) \frac{x/m}{\log(x/m)} + d_2(P(m)) \frac{x/m}{\log^2(x/m)} + \cdots + \right.$$

$$\left. + d_A(P(m)) \frac{x/m}{\log^A(x/m)} + O\left(d_{A+1}(P(m)) \frac{x/m}{\log^{A+1}(x/m)} \right) \right\}. \quad (4.1)$$

D'après le lemme 2 avec $i = 0$, on a

$$\sum_{\substack{m > x/\log x \\ P(m) \leqslant l(x)}} \frac{d_1(P(m))}{m} = \sum_{\substack{m > x/\log x \\ P(m) \leqslant l(x)}} \frac{1}{m} \sum_{p \geqslant P(m)} \frac{1}{p^2} \ll$$

$$\ll \sum_{\substack{m > x/\log x \\ P(m) \leqslant l(x)}} \frac{1}{m} \int_{P(m)}^{\infty} \frac{dt}{t^2} \ll \sum_{\substack{m > x/\log x \\ P(m) \leqslant l(x)}} \frac{1}{mP(m)} \ll \frac{1}{\log^{A+1} x}.$$

On peut donc écrire

$$\sum_{\substack{m \leqslant x/\log x \\ P(m) \leqslant l(x)}} \frac{d_1(P(m))}{m} = \sum_{m=1}^{\infty} \sum_{p \geqslant P(m)} \frac{1}{p^2} + O\left(\frac{1}{\log^{A+1} x}\right). \quad (4.2)$$

D'autre part, en utilisant à nouveau le lemme 2 et le fait que chaque fonction $d_j(z)$ est une combinaison linéaire des expressions $\sum_{p \geqslant z} \dfrac{\log^i p}{p^2}$ $(i = 1, 2, \ldots, A+1)$, on a, pour chaque entier $j = 1, 2, \ldots,$ $A + 1$,

$$\sum_{\substack{m \leqslant x/\log x \\ P(m) \leqslant l(x)}} \frac{d_j(P(m))}{m} \log^i m = O(1). \quad (4.3)$$

En développant l'expression à droite de (4.1) à la lumière de (4.2) et (4.3), le théorème 1 suit.

La démonstration du théorème 2 utilise essentiellement la même approche que celle du théorème 1. Aussi, nous donnerons seulement les grandes lignes de la preuve.

Tout d'abord, puisque la relation (2.1) implique la relation (2.2) dans le cas $k=2$, il suffit de considérer le cas $k \geqslant 3$. Comme pour la preuve du théorème 1, on peut se limiter aux entiers $n \leqslant x$ pour lesquels $P_k(n) \leqslant l(x)$. Or chaque entier $n \geqslant 2^k$ tel que $\Omega(n) \geqslant k$ peut s'écrire de manière unique sous la forme

$$n = mr,$$

où $\Omega(r) = k$ et $P(m) \leqslant p(r) = P_k(r)$. Il s'ensuit que

$$\sum_{\substack{2^k \leqslant n \leqslant x \\ \Omega(n) \geqslant k \\ P_k(n) \leqslant l(x)}} \frac{1}{P_k(n)} = \sum_{\substack{mr \leqslant x \\ \Omega(r) = k \\ P(m) \leqslant P_k(r) \leqslant l(x)}} \frac{1}{P_k(r)} = \sum_{m \leqslant x/2^k} \sum_{\substack{r \leqslant x/m \\ \Omega(r) = k \\ P(m) \leqslant P_k(r) \leqslant l(x)}} \frac{1}{P_k(r)}.$$

En séparant cette dernière expression en deux quantités, selon que $m \leqslant \dfrac{x}{\log x}$ ou que $\dfrac{x}{\log x} < m \leqslant \dfrac{x}{2^k}$, et en utilisant ensuite le lemme 4 pour évaluer la première quantité, la deuxième étant évidemment négligeable, on obtient

$$\begin{aligned} S_k(x) = \sum_{m \leqslant x/\log x} d_1(P(m)) \frac{x/m}{\log(x/m)} \frac{(\log\log(x/m))^{k-2}}{(k-2)!} \times \\ \times \left(1 + O\left(\frac{d_1(P(m)) + \xi(P(m))}{\log\log(x/m)}\right)\right) + \\ + O\left(\sum_{m \leqslant x/2^k} \frac{x/m}{\log(x/m)} (\log\log(x/m))^{k-3}\right), \end{aligned}$$

ce qui mène facilement à l'estimé (2.2), achevant ainsi la preuve du théorème 2.

5. Les lemmes en vue de la démonstration du théorème 3

Lemme 5. *Soit Q un ensemble de nombres premiers de densité $0 < \delta < 1$ satisfaisant (1.10) et soit $\Upsilon(x, Q)$ le nombre d'entiers positifs $n \leqslant x$ n'ayant aucun facteur premier en commun avec les éléments de Q. Alors il existe une constante positive $c(Q)$ telle que*

$$\Upsilon(x, Q) = c(Q) \frac{x}{(\log x)^\delta} \left(1 + O\left(\frac{1}{\log\log 3x}\right)\right). \tag{5.1}$$

Démonstration. Il est mentionné dans GOLDSTON et MCCURLEY ([7], formule (6)), que

$$\Upsilon(x, Q) = \frac{e^{(\delta-1)\gamma}}{\Gamma(1-\delta)} \frac{x}{\log x} \prod_{\substack{p \leqslant x \\ p \notin Q}} \left(1 - \frac{1}{p}\right)^{-1} \left(1 + O\left(\frac{1}{\log\log 3x}\right)\right). \tag{5.2}$$

Or, puisqu'il existe des constantes D et D_Q telles que

$$\sum_{p \leqslant x} \frac{1}{p} = \log\log x + D + O\left(\frac{1}{\log x}\right)$$

(voir HARDY et WRIGHT [8]) et

$$\sum_{\substack{p \leqslant x \\ p \in Q}} \frac{1}{p} = \delta \log\log x + D_Q + O\left(\frac{1}{\log x}\right)$$

(voir [6]), alors on a

$$\prod_{\substack{p \leqslant x \\ p \notin Q}} \left(1 - \frac{1}{p}\right)^{-1} = \exp\left\{- \sum_{\substack{p \leqslant x \\ p \notin Q}} \log\left(1 - \frac{1}{p}\right)\right\} =$$

$$= \exp\left\{\sum_{\substack{p \leqslant x \\ p \notin Q}} \frac{1}{p} + \sum_{\substack{p \leqslant x \\ p \notin Q \\ m \geqslant 2}} \frac{1}{mp^m}\right\} = \exp\left\{\sum_{p \leqslant x} \frac{1}{p} - \sum_{\substack{p \leqslant x \\ p \in Q}} \frac{1}{p} + \sum_{\substack{p \leqslant x \\ p \notin Q \\ m \geqslant 2}} \frac{1}{mp^m}\right\} =$$

$$= \exp\left\{\log\log x + D - \delta \log\log x - D_Q + \sum_{\substack{m \geqslant 2 \\ p \notin Q}} \frac{1}{mp^m} + O\left(\frac{1}{\log x}\right)\right\} =$$

$$= \exp\left\{(1-\delta)\log\log x + E_Q + O\left(\frac{1}{\log x}\right)\right\},$$

où $E_Q = D - D_Q + \sum_{m \geqslant 2, p \notin Q} \dfrac{1}{mp^m}$. Il s'ensuit que

$$\begin{aligned}\prod_{\substack{p \leqslant x \\ p \notin Q}} \left(1 - \frac{1}{p}\right)^{-1} &= (\log x)^{1-\delta} e^{E_Q} e^{O(1/\log x)} = \\ &= e^{E_Q} (\log x)^{1-\delta} \left(1 + O\left(\frac{1}{\log x}\right)\right).\end{aligned} \tag{5.3}$$

En combinant (5.2) et (5.3), on obtient (5.1), ce qui termine la preuve du lemme 5.

Dans ce qui suit, on notera

$$L_1(x)=\frac{\log x}{\log\log x}. \tag{5.4}$$

Lemme 6. *Pour $2\leqslant y\leqslant x$, on a*

$$\sum_{\substack{d>e^{L_1(x)}\\ P(d)\leqslant y}}\frac{1}{d}\ll \log y\cdot e^{-(L_1(x)/2\log y)}.$$

Démonstration. En utilisant le lemme 1, on a

$$\sum_{\substack{d>e^{L_1(x)}\\ P(d)\leqslant y}}\frac{1}{d}\ll\int_{e^{L_1(x)}}^{\infty}\frac{d\psi(t,y)}{t}=\frac{\psi(t,y)}{t}\Big|_{e^{L_1(x)}}^{\infty}+\int_{e^{L_1(x)}}^{\infty}\frac{\psi(t,y)}{t^2}dt\ll$$

$$\ll e^{-(L_1(x)/2\log y)}+\log y\cdot e^{-(L_1(x)/2\log y)},$$

ce qui termine la preuve du lemme 6.

Lemme 7. *Soit Q un ensemble de nombres premiers de densité $0<\delta<1$ satisfaisant (1.10) et soit $\varpi(x,y,Q)=\#\{n\leqslant x:p(n)>y\wedge p|n\Rightarrow$ $\Rightarrow p\notin Q\}$. Alors, uniformément pour $x\geqslant y\geqslant 2$, $x\geqslant e^e$, on a*

$$\varpi(x,y,Q)=c(Q)\frac{x}{(\log x)^{\delta}}\left\{\prod_{\substack{p\leqslant y\\ p\notin Q}}\left(1-\frac{1}{p}\right)+O(\log y\cdot e^{-(L_1(x)/2\log y)})\right\}\times$$

$$\times\left(1+O\left(\frac{1}{\log\log x}\right)\right)+O(\log x\cdot\psi(x,y))+$$

$$+O\left(\frac{x}{(\log\log x)^{\delta}}\log y\cdot e^{-(L_1(x)/2\log y)}\right), \tag{5.5}$$

où $c(Q)$ est la constante du lemme 5. En particulier, pour $y\leqslant\log x$, on a

$$\varpi(x,y,Q)=c(Q)\frac{x}{(\log x)^{\delta}}\prod_{\substack{p\leqslant y\\ p\notin Q}}\left(1-\frac{1}{p}\right)\left(1+O\left(\frac{1}{\log\log x}\right)\right). \tag{5.6}$$

Démonstration. Soit $A=A(y,Q)$ l'ensemble des nombres naturels n dont le plus petit facteur premier est supérieur à y et n'ayant aucun facteur en commun avec les nombers premiers dans Q. Soit $\chi_A(n)$ la fonction caractéristique de $A=A(y,Q)$, de sorte que $\varpi(x,y,Q)=$ $=\sum_{n\leqslant x}\chi_A(n)$. Alors on a, pour $s>1$,

$$\sum_{n=1}^{\infty} \frac{\chi_A(n)}{n^s} = \prod_{\substack{p>y \\ p\notin Q}} \left(1-\frac{1}{p^s}\right)^{-1} = \prod_{p\notin Q}\left(1-\frac{1}{p^s}\right)^{-1} \cdot \prod_{\substack{p\leqslant y \\ p\notin Q}}\left(1-\frac{1}{p^s}\right). \tag{5.7}$$

Soit $N(Q^c)$ le semi-groupe multiplicatif engendré par les nombres premiers qui ne sont pas dans Q. Dénotons par $\chi_{\mathbf{N}(Q^c)}$ la fonction caractéristique de l'ensemble $\mathbf{N}(Q^c)$ et par $\mu_{y,Q}$ la fonction définie par

$$\mu_{y,Q}(n) = \begin{cases} \mu(n) & \text{si } p|n \Rightarrow \text{“}p \leqslant y \text{ et } p\notin Q\text{”}, \\ 0 & \text{autrement.} \end{cases}$$

Avec ces notations, la relation (5.7) devient

$$\sum_{n=1}^{\infty} \frac{\chi_A(n)}{n^s} = \sum_{n=1}^{\infty} \frac{\chi_{\mathbf{N}(Q^c)}(n)}{n^s} \cdot \sum_{n=1}^{\infty} \frac{\mu_{y,Q}(n)}{n^s}.$$

Or, cette dernière identité nous permet d'écrire que

$$\begin{aligned} \varpi(x,y,Q) &= \sum_{n\leqslant x} \chi_A(n) = \sum_{n\leqslant x} (\mu_{y,Q} * \chi_{\mathbf{N}(Q^c)})(n) = \\ &= \sum_{d\leqslant x} \mu_{y,Q}(d) \sum_{d_2\leqslant x/d} \chi_{\mathbf{N}(Q^c)}(d_2) = \\ &= \sum_{d\leqslant e^{L_1(x)}} + \sum_{e^{L_1(x)}<d\leqslant x/\log x} + \sum_{x/\log x<d\leqslant x} = \\ &= \Sigma_1 + \Sigma_2 + \Sigma_3, \end{aligned} \tag{5.8}$$

disons, avec $L_1(x)$ définie en (5.4), où $*$ est utilisé pour dénoter le produit de Dirichlet de deux fonctions arithmétiques.

Nous évaluons séparément les sommes Σ_i, $i = 1, 2, 3$, et comme on le constatera, la principale contribution viendra de Σ_1.

Tout d'abord, on a trivialement, en notant $(d, Q) = 1$ pour signifier que d n'a aucum facteur en commun avec les éléments de Q,

$$\begin{aligned} \Sigma_3 &= \sum_{\substack{x/\log x<d\leqslant x \\ P(d)\leqslant y \\ (d,Q)=1}} \mu(d) \sum_{d_2\leqslant x/d} \chi_{\mathbf{N}(Q^c)}(d_2) \ll \\ &\ll \log x \sum_{\substack{x/\log x<d\leqslant x \\ P(d)\leqslant y}} 1 < \log x\, \psi(x,y). \end{aligned} \tag{5.9}$$

D'autre part,

$$\Sigma_2 = \sum_{e^{L_1(x)}<d\leqslant x/\log x} \mu_{y,Q}(d) \sum_{d_2\leqslant x/d} \chi_{\mathbf{N}(Q^c)}(d_2) =$$

$$= \sum_{e^{L_1(x)} < d \leqslant x/\log x} \mu_{y,Q}(d) \cdot \Upsilon\left(\frac{x}{d}, Q\right).$$

Puisque $d \leqslant x/\log x$, alors $x/d \geqslant \log x \to \infty$. On peut donc appliquer le lemme 5 et obtenir

$$\Sigma_2 = \sum_{e^{L_1(x)} < d \leqslant x/\log x} \mu_{y,Q}(d) c(Q) \left\{ \frac{x/d}{(\log(x/d))^\delta} \left(1 + O\left(\frac{1}{\log\log(3x/d)}\right)\right)\right\} =$$

$$= \sum_{e^{L_1(x)} < d \leqslant x/\log x} \mu_{y,Q}(d) c(Q) \left\{ \frac{x/d}{(\log(x/d))^\delta} \left(1 + O\left(\frac{1}{\log\log\log(3x)}\right)\right)\right\} \ll$$

$$\ll \frac{x}{(\log\log x)^\delta} \sum_{\substack{e^{L_1(x)} < d \leqslant x/\log x \\ P(d) \leqslant y}} \frac{1}{d},$$

puisque pour $d \leqslant x/\log x$, on a $\log(x/d) > \log\log x$. En faisant appel au lemme 6, on a alors

$$\Sigma_2 \ll \frac{x}{(\log\log x)^\delta} \log y \cdot e^{-(L_1(x)/2\log y)}. \tag{5.10}$$

Pour évaluer Σ_1, on procède comme suit. Tout comme pour l'évaluation de Σ_2, on peut faire appel au lemme 5 et ainsi obtenir

$$\Sigma_1 = \sum_{d \leqslant e^{L_1(x)}} \mu_{y,Q}(d) \sum_{d_2 \leqslant x/d} \chi_{\mathrm{N}(Q^c)}(d_2) = \sum_{d \leqslant e^{L_1(x)}} \mu_{y,Q}(d) \cdot \Upsilon\left(\frac{x}{d}, Q\right) =$$

$$= \sum_{d \leqslant e^{L_1(x)}} \mu_{y,Q}(d) \left\{ c(Q) \frac{x/d}{(\log(x/d))^\delta} \left(1 + O\left(\frac{1}{\log\log(3x/d)}\right)\right)\right\}.$$

Puisque $d \leqslant e^{L_1(x)}$, on a

$$\frac{1}{(\log(x/d))^\delta} = \frac{1}{(\log x)^\delta} \left(1 + O\left(\frac{\log d}{\log x}\right)\right)^\delta = \frac{1}{(\log x)^\delta} \left(1 + O\left(\frac{1}{\log\log x}\right)\right)$$

et

$$\log\log 3(x/d) \geqslant \log\log 3x + O\left(\frac{1}{\log\log x}\right).$$

C'est pourquoi

$$\Sigma_1 = \frac{c(Q)x}{(\log x)^\delta} \left(1 + O\left(\frac{1}{\log\log x}\right)\right)\left(1 + O\left(\frac{1}{\log\log 3x}\right)\right) \sum_{d \leqslant e^{L_1(x)}} \frac{\mu_{y,Q}(d)}{d}. \tag{5.11}$$

Par ailleurs, en faisant à nouveau appel au lemme 6, on a

$$\sum_{d \leqslant e^{L_1(x)}} \frac{\mu_{y,Q}(d)}{d} = \sum_{\substack{d \leqslant e^{L_1(x)} \\ P(d) \leqslant y \\ (d,Q)=1}} \frac{\mu(d)}{d} = \sum_{\substack{d=1 \\ P(d) \leqslant y \\ (d,Q)=1}}^{\infty} \frac{\mu(d)}{d} - \sum_{\substack{d > e^{L_1(x)} \\ P(d) \leqslant y \\ (d,Q)=1}} \frac{\mu(d)}{d} =$$

$$= \prod_{\substack{p \leqslant y \\ p \notin Q}} \left(1 - \frac{1}{p}\right) + O\left(\sum_{\substack{d > e^{L_1(x)} \\ P(d) \leqslant y}} \frac{1}{d}\right) =$$

$$= \prod_{\substack{p \leqslant y \\ p \notin Q}} \left(1 - \frac{1}{p}\right) + O(\log y \cdot e^{-1/2(L_1(x)/\log y)}), \qquad (5.12)$$

En substituant (5.12) dans (5.11), et en ramenant ce nouvel estimé de Σ_1, avec ceux de (5.9) et (5.10), dans (5.8), on obtient (5.5). La relation (5.6) découle alors facilement de (5.5) en prenant $y \leqslant \log x$.

Lemme 8. *Soit*

$$l_Q(y) = \prod_{\substack{p \leqslant y \\ p \notin Q}} \left(1 - \frac{1}{p}\right).$$

Alors

(i) *Lorsque* $y \to \infty$, *il existe une constante* v_Q *telle que*

$$l_Q(y) = \frac{v_Q}{(\log y)^{1-\delta}} \left(1 + O\left(\frac{1}{\log y}\right)\right).$$

(ii) *Lorsque* $x \to \infty$,

$$\sum_{\substack{m \leqslant x \\ P(m) \in Q \\ P(m) \leqslant \log x}} \frac{l_Q(P(m))}{mP(m)} = \sum_{\substack{m=1 \\ P(m) \in Q}}^{\infty} \frac{l_Q(P(m))}{mP(m)} + O\left(\frac{1}{\log x}\right).$$

Démonstration. La partie (i) découle de (5.3). Pour obtenir (ii), il suffit de remarquer que $l_Q(P(m)) < 1$ et que

$$\sum_{P(m) > \log x} \frac{1}{mP(m)} = \sum_{p > \log x} \frac{1}{p^2} \sum_{P(n) \leqslant p} \frac{1}{n} =$$

$$= \sum_{p > \log x} \frac{1}{p^2} \prod_{q \leqslant p} \left(1 - \frac{1}{q}\right)^{-1} \ll \sum_{p > \log x} \frac{\log p}{p^2} \ll \frac{1}{\log x}.$$

6. La démonstration du théorème 3

Tout comme pour les sommes $\sum 1/P_k(n)$, ce sont les "petits" facteurs premiers qui contribuent principalement au terme principal du développement asymptotique de $\sum 1/P(n, Q)$. Cela apparaîtra plus clairement dans la démonstration du théorème 3 que nous abordons maintenant.

Soit donc Q un ensemble de nombres premiers de densité positive $\delta < 1$ satisfaisant à

$$\pi(x, Q) \overset{\text{def}}{=} \sum_{p \leqslant x, p \in Q} 1 = \delta \operatorname{Li}(x) + O\left(\frac{x}{\log^B x}\right), \tag{6.1}$$

où $\operatorname{Li}(x) = \int_2^x dt/\log t$, pour une certaine constante $B > 2$, et définissons

$$P(n, Q) = \max\{p : p \mid n \wedge p \in Q\}, \tag{6.2}$$

avec les conventions $P(1, Q) = 1$ et $P(n, Q) = +\infty$ si aucun des facteurs premiers de n n'appartient à Q. GOLDSTON et MCCURLEY ([7]) ont démontré que

$$\psi(x, y, Q) \overset{\text{def}}{=} \sum_{\substack{n \leqslant x \\ P(n,Q) \leqslant y}} 1 = x\rho_\delta(u)\left(1 + O\left(\frac{1}{\log y}\right)\right) \tag{6.3}$$

uniformément pour $u \geqslant 1$ et $y \geqslant 3/2$, où $\rho_\delta(u)$ est la solution continue de l'équation différentielle aux différences

$$u\rho_\delta'(u) = -\delta\rho_\delta(u-1) \qquad (u \geqslant 1), \tag{6.4}$$

avec la condition initiale $\rho_\delta(u) = 1$ pour $0 \leqslant u \leqslant 1$. (La fonction $\rho_\delta(u)$ ne devra évidemment pas être confondure avec la fonction $\rho_k(u)$ donnée en (1.9); le contexte sera toujours assez clair). GOLDSTON et MCCURLEY ([6]) ont également obtenu que

$$\rho_\delta(u) = \frac{e^{\gamma\delta}}{\Gamma(1-\delta)}\frac{1}{u^\delta}\left(1 + O\left(\frac{1}{u}\right)\right). \tag{6.5}$$

Une manière naturelle d'entreprendre la démonstration du théorèm 3 serait donc via l'estimé (6.3) en écrivant

$$\sum_{q_0 \leqslant n \leqslant x} \frac{1}{P(n, Q)} = \sum_{\substack{p \leqslant x \\ p \in Q}} \frac{1}{p} \sum_{\substack{n \leqslant x \\ P(n,Q) = p}} 1 = \sum_{\substack{p \leqslant x \\ p \in Q}} \frac{1}{p} \sum_{\substack{mp \leqslant x \\ P(m,Q) \leqslant p}} 1 = \sum_{\substack{p \leqslant x \\ p \in Q}} \frac{1}{p}\psi\left(\frac{x}{p}, p, Q\right).$$

Comme il est clair qu'il suffit de considérer les nombres premiers $p \in Q$ pour lesquels $p \leqslant \log x$, on obtient, en faisant appel à (6.3) et (6.5),

$$\sum_{q_0 \leqslant n \leqslant x} \frac{1}{P(n, Q)} = \sum_{\substack{p \leqslant \log x \\ p \in Q}} \frac{1}{p} \frac{x}{p} \rho_\delta\left(\frac{\log x - \log p}{\log p}\right)\left(1 + O\left(\frac{1}{\log p}\right)\right) =$$

$$= (1 + o(1)) \frac{e^{\gamma\delta}}{\Gamma(1-\delta)} \frac{x}{(\log x)^\delta} \sum_{\substack{p \leqslant \log x \\ p \in Q}} \frac{(\log p)^\delta}{p^2} \times$$

$$\times \left(1 + O\left(\frac{1}{\log p}\right)\right)\left(1 + O\left(\frac{\log\log x}{\log x}\right)\right).$$

De ceci, il découle que $\sum_{q_0 \leqslant n \leqslant x} \dfrac{1}{P(n, Q)} \approx \dfrac{x}{(\log x)^\delta}$, ce qui n'est pas suffisant pour déduire (2.3). Nous allons donc procéder autrement.

Comme on l'a mentionné ci-haut, il suffit de considérer les entiers positifs $n \leqslant x$ pour lesquels $P(n, Q) \leqslant \log x$. D'autre part, compte tenu des estimés (1.2) et (1.5), on peut également se limiter aux entiers positifs $n \leqslant x$ pour lesquels $P(n, Q) < P(n)$. C'est pourquoi,

$$\sum_{q_0 \leqslant n \leqslant x} \frac{1}{P(n, Q)} = \sum_{\substack{q_0 \leqslant n \leqslant x \\ P(n,Q) \leqslant \log x \\ P(n,Q) < P(n)}} \frac{1}{P(n, Q)} + O\left(\frac{x}{\log x}\right). \tag{6.6}$$

Il nous suffira donc d'évaluer la somme à droite de (6.6).

Tout entier n tel que $P(n, Q) < P(n)$ peut s'écrire de manière unique sous la forme

$$n = mr,$$

où $(r, Q) = 1$, $p(r) > P(m) \in Q$. On a alors

$$\sum_{\substack{q_0 \leqslant n \leqslant x \\ P(n,Q) \leqslant \log x \\ P(n,Q) < P(n)}} \frac{1}{P(n, Q)} = \sum_{\substack{mr \leqslant x \\ (r,Q) = 1 \\ p(r) > P(m) \in Q \\ P(m) \leqslant \log x}} \frac{1}{P(m)} = \sum_{\substack{m \leqslant x \\ P(m) \in Q \\ P(m) \leqslant \log x}} \frac{1}{P(m)} \sum_{\substack{r \leqslant x/m \\ (r,Q) = 1 \\ p(r) > P(m)}} 1 =$$

$$= \sum_{\substack{m \leqslant x \\ P(m) \in Q \\ P(m) \leqslant \log x}} \frac{1}{P(m)} \varpi\left(\frac{x}{m}, P(m), Q\right) =$$

$$= \sum_{\substack{m \leqslant e^{\sqrt{(\log x)}} \\ P(m) \in Q \\ P(m) \leqslant \log x}} + \sum_{\substack{e^{\sqrt{(\log x)}} < m \leqslant x e^{-\sqrt{(\log x)}} \\ P(m) \in Q \\ P(m) \leqslant \log x}} + \sum_{\substack{x e^{-\sqrt{(\log x)}} < m \leqslant x \\ P(m) \in Q \\ P(m) \leqslant \log x}} =$$

$$= U_1(x) + U_2(x) + U_3(x),$$

disons. Il est clair que, en raison de l'estimé (1.5) de $L(x)$,

$$U_3(x) \ll e^{\sqrt{(\log x)}} \sum_{m \leqslant x} \frac{1}{P(m)} \ll \frac{x e^{\sqrt{(\log x)}}}{e^{\sqrt{(2 \log x \log\log x)}}} = o\left(\frac{x}{(\log x)^\delta (\log\log x)} \right).$$

D'autre part, en utilisant la formule (5.6) du lemme 7, on a

$$U_2(x) \ll \sum_{e^{\sqrt{(\log x)}} < m \leqslant x e^{-\sqrt{(\log x)}}} \frac{1}{P(m)} \frac{x/m}{(\log (x/m))^\delta} <$$

$$< \frac{x}{(\sqrt{\log x})^\delta} \sum_{e^{\sqrt{(\log x)}} < m \leqslant x} \frac{1}{mP(m)} \ll \frac{x}{(\sqrt{\log x})^\delta} \int_{e^{\sqrt{(\log x)}}}^{x} \frac{d\tau}{\tau L(\tau)} \ll$$

$$\ll \frac{x \log x}{(\sqrt{\log x})^\delta L(e^{\sqrt{(\log x)}})} = \frac{x (\log x)^{1-\delta/2}}{e^{\sqrt{(\sqrt{(\log x)} \log\log x)}}} = o\left(\frac{x}{(\log x)^\delta \log\log x} \right).$$

Il nous suffit donc d'évaluer $U_1(x)$. Pour ce faire, on utilise la formule (5.6) du lemme 7 (avec $y = P(m) \leqslant \log x$) et le lemme 8, et on obtient

$$U_1(x) = \sum_{\substack{m \leqslant e^{\sqrt{(\log x)}} \\ P(m) \in Q \\ P(m) \leqslant \log x}} \frac{1}{P(m)} \left\{ c(Q) \frac{x}{m} \frac{l_Q(P(m))}{(\log (x/m))^\delta} \left(1 + O\left(\frac{1}{\log\log (x/m)} \right) \right) \right\} =$$

$$= c(Q) \frac{x}{(\log x)^\delta} \sum_{\substack{m \leqslant e^{\sqrt{(\log x)}} \\ P(m) \in Q \\ P(m) \leqslant \log x}} \frac{l_Q(P(m))}{mP(m)} \times$$

$$\times \left(1 + O\left(\frac{\sqrt{\log x}}{\log x} \right) \right) \left(1 + O\left(\frac{1}{\log\log x} \right) \right) =$$

$$= c(Q) \left(1 + O\left(\frac{1}{\log\log x} \right) \right) \frac{x}{(\log x)^\delta} \sum_{\substack{m=1 \\ P(m) \in Q}}^{\infty} \frac{l_Q\,(P(m))}{mP(m)}.$$

Ceci termine la preuve du théorème 3.

7. Une généralisation

Dans un article récent, De Koninck et Mercier ([3]) ont comparé les comportements asymptotiques des sommes $\sum_{2\leqslant n\leqslant x} f(P(n))$ et $\sum_{n\leqslant x} f(n)$ où f est une fonction fortement additive satisfaisant certaines conditions minimales. Il s'avère que certains de leurs résultats obtenus pour $\sum_{2\leqslant n\leqslant x} f(P(n))$ peuvent être étendus à l'étude des sommes $\sum_{2^k\leqslant n\leqslant x, \Omega(n)\geqslant k} f(P_k(n))$ et $\sum_{q_0\leqslant n\leqslant x} f(P(n,Q))$: c'est ce que nous avons fait en (2.2) et (2.3) dans le cas $f(x)=\frac{1}{x}$. En utilisant essentiellement les mêmes techniques que celles élaborées dans les sections 3 à 6, on peut généraliser les estimés (2.2) et (2.3) et obtenir que si $f(x)=\dfrac{1}{x^\rho L(x)}$ avec $\rho>0$, alors il existe des constantes λ'_k telles que, pour chaque $k\in\mathbf{N}$,

$$\sum_{\substack{2^k\leqslant n\leqslant x\\ \Omega(n)\geqslant k}} f(P_k(n)) \sim \lambda'_k \frac{x(\log\log x)^{k-2}}{\log x} \qquad (x\to\infty)$$

et une constante $\eta'(Q)$ telle que

$$\sum_{q_0\leqslant n\leqslant x} f(P(n,Q)) \sim \eta'(Q)\frac{x}{(\log x)^\delta} \qquad (x\to\infty).$$

Il en est de même lorsque $\rho=0$, mais cette fois avec la condition additionnelle

$$\sum_p \frac{\log p}{pL(p)} < +\infty.$$

8. La valeur médiane des plus grands facteurs premiers

Théorème 4. *Soit Q un ensemble de nombres premiers de densité $0<\delta<1$ satisfaisant à*

$$\pi(x,Q) \stackrel{\text{def}}{=} \sum_{\substack{p\leqslant x\\ p\in Q}} 1 = \delta\,\mathrm{Li}(x) + O\left(\frac{x}{\log^B x}\right), \tag{8.1}$$

pour une certaine constante $B>2$. Alors la valeur médiane de $P(n,Q)$ pour les entiers $n\leqslant x$ est $n^{\kappa+o(1)}$, où κ est la solution unique de l'équation

$$\rho_\delta\left(\frac{1}{\kappa}\right)=\frac{1}{2}.$$

Démonstration. Parmi les nombres $\kappa \in]0,1[$, on cherche celui pour lequel

$$\sum_{\substack{x^\kappa < p \leqslant x \\ p \in Q}} \sum_{\substack{n \leqslant x \\ P(n,Q) = p}} 1 \sim \frac{x}{2}. \tag{8.2}$$

Evaluons donc l'expression à gauche de (8.2). Compte tenu de la définition de $\psi(x, y, Q)$ et de l'estimé (6.3), on a

$$\sum_{\substack{x^\kappa < p \leqslant x \\ p \in Q}} \sum_{\substack{n \leqslant x \\ P(n,Q) = p}} 1 = \sum_{\substack{x^\kappa < p \leqslant x \\ p \in Q}} \sum_{\substack{mp \leqslant x \\ P(m,Q) \leqslant p}} 1 = \sum_{\substack{x^\kappa < p \leqslant x \\ p \in Q}} \psi\left(\frac{x}{p}, p, Q\right) =$$

$$= \sum_{\substack{x^\kappa < p \leqslant x \\ p \in Q}} \frac{x}{p} \rho_\delta\left(\frac{\log x}{\log p} - 1\right)\left(1 + O\left(\frac{1}{\log p}\right)\right) =$$

$$= x\left(1 + O\left(\frac{1}{\log x}\right)\right) \sum_{\substack{x^\kappa < p \leqslant x \\ p \in Q}} \frac{1}{p} \rho_\delta\left(\frac{\log x}{\log p} - 1\right) =$$

$$= x\left(1 + O\left(\frac{1}{\log x}\right)\right) \int_{x^\kappa}^{x} \frac{\rho_\delta\left(\frac{\log x}{\log t} - 1\right)}{t} d\pi(t, Q) = S, \tag{8.3}$$

disons. En utilisant une intégration par parties et en faisant ensuite appel à (8.1), on obtient

$$S = \left(1 + O\left(\frac{1}{\log x}\right)\right) x \cdot$$

$$\cdot \left\{ \delta \int_{x^\kappa}^{x} \frac{\rho_\delta\left(\frac{\log x}{\log t} - 1\right)}{t \log t} dt + O\left(\int_{x^\kappa}^{x} \frac{\rho_\delta\left(\frac{\log x}{\log t} - 1\right)}{t \log^2 t} dt \right) \right\} =$$

$$= \left(1 + O\left(\frac{1}{\log x}\right)\right) x \delta \int_{x^\kappa}^{x} \frac{\rho_\delta\left(\frac{\log x}{\log t} - 1\right)}{t \log t} dt = x(1 + o(1)) I, \tag{8.4}$$

disons. En effectuant le changement de variable $u = \dfrac{\log x}{\log t}$, l'expression

I devient

$$I = \delta \int_1^{1/\kappa} \rho_\delta(u-1)\frac{du}{u}. \tag{8.5}$$

Puisqu'on a

$$u\rho_\delta'(u) = -\delta\rho_\delta(u-1),$$

la relation (8.5) devient

$$I = \int_1^{1/\kappa} -\rho_\delta'(u)du = -\rho_\delta(u)|_1^{1/\kappa} = 1 - \rho_\delta(1/\kappa).$$

En substituant cette dernière estimation de I dans (8.4), on obtient

$$S = (1 + o(1))x(1 - \rho_\delta(1/\kappa)).$$

Compte tenu de (8.2), le nombre κ doit satisfaire à

$$1 - \rho_\delta(1/\kappa) = \tfrac{1}{2}.$$

Le nombre κ cherché est donc l'unique solution de

$$\rho_\delta\left(\frac{1}{\kappa}\right) = \frac{1}{2}, \tag{8.6}$$

comme il fallait démontrer.

Remarque. Il découle de (6.4) que pour $1 < u < 2$, on a $\rho_\delta(u) = 1 - \delta\log u$. Or pour $\delta = \dfrac{1}{2\log 2}$, on a $1 < \dfrac{1}{\kappa} < 2$, ce qui veut dire que

$$\rho_\delta\left(\frac{1}{\kappa}\right) = 1 - \delta\log\left(\frac{1}{\kappa}\right)$$

et donc en combinant ceci avec (8.6), on obtient

$$\kappa = e^{-(1/2\delta)}. \tag{8.7}$$

Donc lorsque la densité δ de Q satisfait à $\delta > \dfrac{1}{2\log 2}$, la valeur médiane de $P(n, Q)$ est $n^{\kappa + o(1)}$, où $\kappa = e^{-(1/2\delta)}$. Par ailleurs, pour les petites valeurs de δ, κ devient très petit et $1/\kappa$ très grand; alors, d'après (6.5),

$$\rho_\delta(u) \sim \frac{e^{\gamma\delta}}{\Gamma(1-\delta)}\frac{1}{u^\delta} \qquad (u \to \infty)$$

et

$$\rho_\delta\left(\frac{1}{\kappa}\right) \sim \frac{e^{\gamma\delta}}{\Gamma(1-\delta)}\kappa^\delta.$$

En combinant ceci avec (8.6), on obtient

$$\frac{e^{\gamma\delta}}{\Gamma(1-\delta)}\kappa^\delta \sim \frac{1}{2},$$

ce qui entraîne que

$$\kappa \sim \frac{e^{-\gamma}}{2^{1/\delta}}(\Gamma(1-\delta))^{1/\delta} = \frac{e^{-\gamma}}{2^{1/\delta}}\exp\left\{\frac{1}{\delta}\log\Gamma(1-\delta)\right\}.$$

Or, puisque

$$\lim_{\delta\to 0}\frac{\log\Gamma(1-\delta)}{\delta} = \lim_{\delta\to 0} -\frac{\Gamma'(1-\delta)}{\Gamma(1-\delta)} = \gamma,$$

on a

$$\kappa \sim \frac{e^{-\gamma}e^{\gamma}}{2^{1/\delta}} = \frac{1}{2^{1/\delta}},$$

un comportement différent de celui exprimé en (8.7).

Théorème 5. *Pour chaque entier $k \geqslant 1$, la valeur médiane de $P_k(n)$ pour les entiers $n \leqslant x$ est $n^{\kappa+o(1)}$, où $\kappa = \kappa(k)$ est la solution unique de l'équation*

$$\rho_k\left(\frac{1}{\kappa}\right) = \frac{1}{2}.$$

Démonstration. La démonstration est analogue à celle du théorème 4, compte tenu de l'analogie marquée qui existe entre les relations (6.3) et (6.4) d'une part et les relations (1.8) et (1.9) d'autre part.

Remarque. Puisque $\rho(u) = 1 - \log u$ pour $1 \leqslant u \leqslant 2$, on a $\frac{1}{2} = \rho\left(\frac{1}{\kappa}\right) = 1 + \log\kappa$, d'où $\kappa = \kappa(1) = \frac{1}{\sqrt{e}}$, un résultat déjà obtenu par WUNDERLICH et SELFRIDGE ([16]). En utilisant les tables donnant les valeurs de $\rho_2(u)$ (voir [14]), on obtient $\kappa = \kappa(2) = 0.21\ldots$ une valeur qui améliore l'estimé $\kappa = 0.24$ obtenu empiriquement par WUNDERLICH et SELFRIDGE ([16]).

Bibliographie

[1] ALLADI, K., ERDÖS, P.: On an additive arithmetic function. Pacific J. Math. **71**, 275–294 (1977).
[2] DE KONINCK, J. M., IVIĆ, A.: The distribution of the average prime divisor of an integer. Arch. Math. **43**, 37–43 (1984).
[3] DE KONINCK, J. M., MERCIER, A.: Les fonctions arithmétiques et le plus grand facteur premier. Acta Arith. **52**, 25–48 (1989).
[4] ERDÖS, P., IVIĆ, A.: On sums involving reciprocals of certain arithmetical functions. Publ. Inst. Math. (Beograd) **32**, 49–56 (1982).
[5] ERDÖS, P., IVIĆ, A., POMERANCE, C.: On sums involving reciprocals of the largest prime factor of an integer. Glas. Mat. **21**, 27–44 (1986).
[6] GOLDSTON, D. A., MCCURLEY, K. S.: Sieving the positive integers by large primes. J. Number Theory **28**, 94–115 (1988).
[7] GOLDSTON, D. A., MCCURLEY, K. S.: Sieving the positive integers by small primes. Trans. Amer. Math. Soc. **307**, 51–62 (1988).
[8] HARDY, G. H., WRIGHT, E. M.: An Introduction to the Theory of Numbers. (4th ed.). Oxford: Clarendon Press. 1960.
[9] HAFNER, J. L., MCCURLEY, K. S.: On the distribution of running times of certain integer factoring algorithms. IBM Res. Report, San Jose, California, 28 pages. (1987).
[10] HILDEBRAND, A.: On the number of positive integers $\leqslant x$ and free of prime factors $> y$. J. Number Theory **22**, 289–307 (1986).
[11] IVIĆ, A.: On the k-th prime factor of an integer. Univ. u Novom Sadu Zb. Rad. Prirod. Mat. Fak. Ser. Mat. **20**, 63–73 (1990).
[12] KNUTH, D. E., PARDO, L. T.: Analysis of a simple factorization algorithm. Theoret. Comput. Sci. **3**, 321–348 (1976).
[13] NICOLAS, J. L.: Sur la distribution des nombres entiers ayant une quantité fixée de facteurs premiers. Acta Arith. **44**, 191–200 (1984).
[14] RIESEL, H.: Prime Numbers and Computer Methods for Factorization. Boston: Birkhäuser. 1985.
[15] TENENBAUM, G.: Introduction à la théorie analytique et probabiliste des nombres. Inst. Elie Cartan Vol. 13, (1990).
[16] WUNDERLICH, M. C., SELFRIDGE, J. L.: A design for a number theory package with an optimized trial division routine. Comm. ACM **17**, 272–276 (1974).

JEAN-MARIE DE KONINCK
Université Laval
Département de Mathématiques et de Statistique
Cité universitaire
Québec
Canada G1K 7P4