Quart. J. Math **51** (2000), 385–403.

A rate estimate in Billingsley's theorem for the size distribution of large prime factors *

by Gérald Tenenbaum

(*Institut Élie Cartan, Université Henri-Poincaré, BP 239, 54506 Vandœuvre Cedex, France*)

Abstract. Let $\{P_j(m) : 1 \leqslant j \leqslant \omega(m)\}$ denote the decreasing sequence of distinct prime factors of a positive integer m. We provide an asymptotic expansion for the distribution function

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) := \nu_n \{m : P_j(m) > n^{\alpha_j} \ (1 \leqslant j \leqslant k)\}$$

which is valid uniformly in a large range for $\overrightarrow{\boldsymbol{\alpha_k}} := (\alpha_1, \dots, \alpha_k)$. When $k \geqslant 2$, we give an asymptotic formula for the same quantity which holds with no restriction at all on $\overrightarrow{\boldsymbol{\alpha_k}}$. A sample consequence of this second result is that, given any fixed $k \geqslant 2$, the formula

$$\nu_n\{P_k(m) \leqslant y\} = r_k(u)\{1 + O(1/\log y)\}$$

holds uniformly for $2 \leqslant y \leqslant n$, where u is defined by $n = y^u$ and r_k is a suitable distribution function.

1. Introduction and statement of results

For integer $m \geqslant 1$, let $\omega(m)$ denote the total number of prime factors of m, counted without multiplicity, and write $P_j(m)$ $(1 \leqslant j \leqslant \omega(m))$ for the jth largest distinct prime factor of m. Billingsley proved in 1972 that the joint size distribution of the vector

$$\Big(\frac{\log P_1(m)}{\log n}, \dots, \frac{\log P_{\omega(m)}(m)}{\log n}\Big),$$

regarded as a random vector on $\Omega_n := \{m : 1 \leqslant m \leqslant n\}$ with uniform probability ν_n, converges in law to the Poisson-Dirichlet distribution with parameter 1, which we shall denote by $PD(1)$.

The Poisson-Dirichlet process has many equivalent definitions : see in particular Kingman [10], and Arratia [1]. The classical definition is the following. If

$$X_1 \geqslant X_2 \geqslant \cdots$$

are the points of a Poisson process with rate, or intensity, $x \mapsto \mathrm{e}^{-x}/x$,(1) then

$$Z := \sum_{j \geqslant 1} X_j$$

* We include here some corrections with respect to the published version.

1. In other words, for $0 < a < b$, the number $N_{a,b}$ of points X_j which fall into $]a, b]$ is a Poisson random variable with parameter $\int_a^b \mathrm{e}^{-x}\,\mathrm{d}x/x$.

is almost surely finite and $V_j = X_j/Z$ defines a sequence having distribution $PD(1)$. In particular, if $(V_1, V_2, \ldots)$ has distribution $PD(1)$, then $V_1 \geqslant V_2 \geqslant \cdots$ and $\sum_{j\geqslant 1} V_j = 1$. Another construction of the Poisson-Dirichlet distribution is obtained by considering an infinite sequence $U_1, U_2, \ldots$ of independent uniform random variables on $]0,1[$ and defining

$$X_1 = U_1,\ X_2 := (1-U_1)U_2,\ X_3 := (1-U_1)(1-U_2)U_3, \ldots$$

so that $\sum_{j\geqslant 1} X_j = 1$ almost surely. Then the permutation $(X_{\sigma(1)}, X_{\sigma(2)}, \ldots)$ obtained by rearranging the X_j in non-increasing order also has distribution $PD(1)$. A third realisation has recently been obtained by Arratia, Barbour and Tavaré [2] : the U_j being as above, put $Y_j := \prod_{1\leqslant i\leqslant j} U_i$ and $T := \sum_{j\geqslant 1} Y_j$; then the conditional law $(Y_1, Y_2, \ldots | T = 1)$ is distributed according to $PD(1)$.

Write $\overrightarrow{\boldsymbol{\alpha_k}} := (\alpha_1, \ldots, \alpha_k)$ and

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) := \nu_n\big\{m : P_j(m) > n^{\alpha_j}\ (1 \leqslant j \leqslant k)\big\}.$$

Billingsley's theorem is plainly equivalent to the statement that, for fixed k and $n \to \infty$,

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = \mathbb{P}(V_1 > \alpha_1, \ldots, V_k > \alpha_k) + o(1). \tag{1·1}$$

A simple proof of this, completely different from Billingsley's, has recently been given by Donnelly and Grimmett [6]. Their approach is based on a size-biased random permutation of the $P_j(m)$ which fits the order of the X_j defined above.

In this article, we investigate the rate of convergence in (1·1). The most precise result known so far in this direction is the formula of Knuth and Trabb Pardo [11] which states that

$$F_n(0, \ldots, 0, \alpha) = \mathbb{P}(V_k > \alpha) + \frac{\sigma_k(\alpha)}{\log n} + O_\alpha\Big(\frac{1}{(\log n)^2}\Big)$$

for any fixed $\alpha > 0$, $k \geqslant 1$, and a suitable function $\sigma_k(\alpha)$. We provide an asymptotic development for the left-hand side of (1·1) according to negative powers of $\log n$.

Our statement involves a small parameter $\varepsilon \in]0, \frac{1}{3}[$ which will be fixed throughout the paper. We introduce the notation

$$L_\varepsilon(y) := \mathrm{e}^{(\log y)^{3(1-\varepsilon)/5}} \quad (y \geqslant 2), \quad \kappa(\varepsilon, x) := \frac{(\log_2 x)^{5(1+2\varepsilon)/3}}{\log x} \quad (x > 3).$$

Here and subsequently, $\log_k$ denotes the k-fold iterated logarithm.

Theorem. *Let k be a positive integer. There exists a sequence of real functions $\{\varphi_h\}_{h=0}^\infty$ defined on $[0,1]^k$, an increasing sequence of integers $\{R_h\}_{h=0}^\infty$ with $R_0 = 0$, and a sequence of affine linear forms $\{\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})\}_{r=1}^\infty$ having the following properties :*

(i) φ_0 is continuous and, for $h \geqslant 1$, φ_h is continuous except perhaps on $\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}}) = 0$ for $R_{h-1} < r \leqslant R_h$. Moreover, for all fixed $h \geqslant 1$, we have

$$\varphi_h(\overrightarrow{\boldsymbol{\alpha_k}}) \ll_h 1/\alpha_k^h \qquad (\alpha_k > 0).$$

(ii) For each r, we have

$$\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}}) = c_r - \sum_{1\leqslant j\leqslant k} m_{jr}\alpha_j \tag{1·2}$$

where the m_{jr} are non negative integers.

(iii) For arbitrary but fixed $H \geqslant 0$ and $\varepsilon \in]0, \frac{1}{3}[$, we have

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = \sum_{0\leqslant h\leqslant H} \frac{\varphi_h(\overrightarrow{\boldsymbol{\alpha_k}})}{(\log n)^h} + O\Big(\frac{1}{(\alpha_k \log n)^{H+1}}\Big), \tag{1·3}$$

uniformly in the range

$$\begin{cases} \overrightarrow{\boldsymbol{\alpha_k}} \in [0,1]^k,\ \alpha_k > \kappa(\varepsilon, n), \\ \displaystyle\min_{\substack{1\leqslant r\leqslant R_H\\ \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})>0}} \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}}) > K_H \frac{\log_2 n}{\log n}, \end{cases}$$

where K_H is a suitable constant depending only on H. The implied constant in (1·3) may depend on H and ε.

(iv) For $k \geqslant 2$, we have

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = \varphi_0(\overrightarrow{\boldsymbol{\alpha_k}}) + O\bigg(\frac{(\log 1/\alpha_k)^{k-2}}{\log n}\bigg) \tag{1·4}$$

uniformly for $\overrightarrow{\boldsymbol{\alpha_k}} \in]0,1]^k$, $n \geqslant 2$.

We note that it is always possible to define $\varphi_h(\overrightarrow{\boldsymbol{\alpha_k}})$ on the critical hyperplanes in such a way that (1·3) holds for any fixed $\overrightarrow{\boldsymbol{\alpha_k}}$ with $\alpha_k > 0$: indeed any $\overrightarrow{\boldsymbol{\alpha_k}}$ with $\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}}) = 0$ for some value(s) of r is a limit point of vectors $\overrightarrow{\boldsymbol{\alpha_k^*}}$ with $\Lambda_r\big(\overrightarrow{\boldsymbol{\alpha_k^*}}\big) < 0$ for the same value(s) of r.

We also observe that the remainder term in (1·4) may always be chosen to be $o(1)$[2] since we may assume with no loss of generality that

$$\alpha_k > 1/(2\log n).$$

As will transpire from the proof, the limit

$$\varphi_0(\overrightarrow{\boldsymbol{\alpha_k}}) = \mathbb{P}(V_1 > \alpha_1, \ldots, V_k > \alpha_k)$$

has a simple expression in terms of the Dickman function—see (3·13) below. The Dickman function is the density function of the law of $1/V_1$ and is defined as the unique continuous solution on $\mathbb{R}^+$ of the difference-differential equation

$$u\varrho'(u) + \varrho(u-1) = 0$$

with initial condition $\varrho(u) = 1$ $(0 \leqslant u \leqslant 1)$. As usual, we set $\varrho(u) = 0$ for $u < 0$. It is immediate that the kth derivative $\varrho^{(k)}(u)$ is defined for $u \in \mathbb{R} \smallsetminus \{0, 1, \ldots, k\}$. We define $\varrho^{(k)}(j)$ by right-continuity for $0 \leqslant j \leqslant k$.

2. And actually $\ll (\log_2 n)^{k-2}/\log n$.

Formula (1·4) extends and makes more precise a result of Hafner and McCurley [8]. For instance, using the fact (shown in Lemma 4 below) that, for fixed $k \geqslant 2$,

$$r_k(v) := 1 - \varphi_0(0, \dots, 0, 1/v) \asymp \frac{\big(\log 2v\big)^{k-2}}{v} \qquad (v \geqslant 1), \tag{1·5}$$

we deduce as a special case of (1·4) that

$$\frac{1}{n} \sum_{\substack{m \leqslant n \\ P_k(m) \leqslant y}} 1 = r_k(u)\Big\{1 + O\Big(\frac{1}{\log y}\Big)\Big\} \tag{1·6}$$

for any fixed $k \geqslant 2$ and uniformly for $2 \leqslant y \leqslant n$, with $u := (\log n)/\log y$.

Our method essentially rests on Saias's extension [12] of de Bruijn's approximation formula [5][(3)]

$$\Psi(x,y) := \sum_{\substack{m \leqslant x \\ P_1(m) \leqslant y}} 1 = \Big\{1 + O\Big(\frac{1}{L_\varepsilon(y)}\Big)\Big\} x \int_{1-}^{\infty} \varrho\Big(\frac{\log x/t}{\log y}\Big) \,\mathrm{d}\Big(\frac{[t]}{t}\Big), \tag{1·7}$$

valid uniformly for $x \geqslant 2$, $y \geqslant x^{\kappa(\varepsilon,x)}$. The integral on the right is best defined as a Lebesgue–Stieltjes integral, since the jump of $\varrho(u)$ at $u = 0$ causes a formal difficulty in the frame of the Riemann-Stieltjes integral when x is a positive integer. Alternatively, one can, as did de Bruijn, first restrict the definition to non-integer values of x, so that it makes sense as a Riemann–Stieltjes integral, and then extend the definition to $\mathbb{R}^+$ by right-continuity. We note that, when $y > x$, the integral is equal to $[x]/x$, so the error term actually vanishes.

The discontinuities of the functions φ_h cannot be avoided, as shown by the case $k = 1$—see [12] or [14], pp. 390–391.

It may be observed that our theorem is equally valid, with no change in the statement, if the prime factors $P_j(m)$ are counted with multiplicities.

2. Technical preparation

Let $P^-(m)$ stand for the smallest prime factor of an integer $m > 1$, with the convention that $P^-(1) = \infty$. Denote by $\mathcal{E}_k(x,y)$ the set of all integers $m \leqslant x$ with $\omega(m) = k$, $P^-(m) > y$, and by $\pi_k(x,y)$ the cardinality of $\mathcal{E}_k(x,y)$. We also set $\mathcal{E}_k(x) := \mathcal{E}_k(x,1)$ and $\pi_k(x) := \pi_k(x,1)$.

We need a uniform upper bound for $\pi_k(x,y)$. This will be provided in Lemma 2 below as a straightforward consequence of the following result, which, with other applications in mind, we state in a much more general context.

Lemma 1. *Let $A > 0$, $B > 0$ and assume f is a non-negative, multiplicative arithmetic function satisfying the conditions*

$$\sum_{p^\nu \leqslant x} f(p^\nu) \log p^\nu \leqslant Ax \qquad (x \geqslant 2), \qquad \sum_p \sum_{\nu \geqslant 2} \frac{f(p^\nu)}{p^\nu} \leqslant B.$$

3. De Bruijn proved the validity of (1·7) in the range $y \geqslant \exp(\log x)^{5/8+\varepsilon}$ for any fixed $\varepsilon > 0$. For a proof of Saias' result, see also [14], theorem III.5.9.

Then we have uniformly for $k \geqslant 1$, $x \geqslant 2$,

$$\sum_{m \in \mathcal{E}_k(x)} f(m) \ll \frac{Ax}{\log x} \frac{\left(\sum_{p \leqslant x} f(p)/p + B\right)^{k-1}}{(k-1)!}, \tag{2·1}$$

where the implicit constant is absolute. In particular, for any constant $R > 0$, we have, uniformly under the condition $1 \leqslant k \leqslant R \log_2 x$,

$$\sum_{m \in \mathcal{E}_k(x)} f(m) \ll \pi_k(x) \exp\Big\{\varrho^* \sum_{p \leqslant x} \frac{f(p) - 1}{p}\Big\} \tag{2·2}$$

with $\varrho^* := (k-1)/\log_2 x$ and where the implicit constant depends at most on A, B, R.

Proof. Write $S_k(x)$ for the left-hand side of (2·1) and put

$$T_k(x) := \sum_{m \in \mathcal{E}_k(x)} f(m) \log m.$$

Our first step consists in finding an upper bound for $T_k(x)$. We have

$$\begin{aligned} T_k(x) &= \sum_{\substack{mp^\nu \leqslant x,\, p \nmid m \\ \omega(m) = k-1}} f(p^\nu) f(m) \log p^\nu \leqslant Ax \sum_{\substack{m \leqslant x \\ \omega(m) = k-1}} \frac{f(m)}{m} \\ &\leqslant \frac{Ax}{(k-1)!} \Big(\sum_{p^\nu \leqslant x} \frac{f(p^\nu)}{p^\nu} \Big)^{k-1} \leqslant \frac{Ax}{(k-1)!} \Big(\sum_{p \leqslant x} \frac{f(p)}{p} + B \Big)^{k-1}. \end{aligned}$$

A simple integration by parts now suffices to establish (2·1). Setting

$$L(x) := \sum_{p \leqslant x} f(p)/p,$$

we have

$$\begin{aligned} S_k(x) &= \int_{3/2}^x \frac{\mathrm{d}T_k(u)}{\log u} = \frac{T_k(x)}{\log x} + \int_{3/2}^x \frac{T_k(u)}{u(\log u)^2} \,\mathrm{d}u \\ &\leqslant \frac{Ax}{\log x} \frac{\{L(x) + B\}^{k-1}}{(k-1)!} + A \frac{\{L(x) + B\}^{k-1}}{(k-1)!} \int_{3/2}^x \frac{\mathrm{d}u}{(\log u)^2} \\ &\ll \frac{Ax}{\log x} \frac{\{L(x) + B\}^{k-1}}{(k-1)!}. \end{aligned}$$

The bound (2·2) follows immediately from (2·1) and classical estimates for $\pi_k(x)$[4] since, in the required range, we have

$$\begin{aligned} \frac{x}{\log x} \frac{\{L(x) + B\}^{k-1}}{(k-1)!} &\leqslant \frac{x}{\log x} \frac{(\log_2 x)^{k-1}}{(k-1)!} \left\{ 1 + \frac{\sum_{p \leqslant x} \{f(p) - 1\}/p + B}{\log_2 x} \right\}^{k-1} \\ &\ll \pi_k(x) \exp\Big\{ \varrho^* \sum_{p \leqslant x} \frac{f(p) - 1}{p} \Big\}. \end{aligned}$$

4. See, e.g., [14], theorem II.6.4.

This completes the proof.

Remark. The upper bound in Lemma 1 is, in a certain sense, optimal. For instance, when f fulfils Wirsing's conditions

$$0 \leqslant f(p^\nu) \leqslant \lambda_1 \lambda_2^{\nu-1} \qquad (p \text{ prime}, \ \nu \geqslant 1),$$

for suitable constants λ_1, λ_2 with $\lambda_1 > 0$, $0 < \lambda_2 < 2$, and furthermore satisfies

$$f(p^\nu) = 1 \qquad (p > y)$$

with $y = y(x)$ such that $\log_2 y = o\big(\sqrt{\log_2 x}\big)$ as $x \to \infty$, it can be shown as a straightforward application of the Selberg–Delange method,[5] that the asymptotic formula

$$\frac{1}{\pi_k(x)} \sum_{m \in \mathcal{E}_k(x)} f(m) = \{1 + o(1)\} \prod_{p \leqslant x} \frac{1 + \varrho^* \sum_{\nu=1}^\infty f(p^\nu)/p^\nu}{1 + \varrho^*/(p-1)},$$

is uniformly valid for $1 \leqslant k \leqslant R \log_2 x$.

Lemma 2. *There is an absolute constant $c_0 > 0$ such that, uniformly for $k \geqslant 1$ and $x \geqslant y \geqslant 2$, we have*

$$\pi_k(x, y) \ll \frac{x}{\log x} \frac{(\log u + c_0)^{k-1}}{(k-1)!} \tag{2·3}$$

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

Proof. We apply (2·1) with $f(p^\nu) = 1$ for $p > y$ and $f(p^\nu) = 0$ otherwise.

Lemma 3. *For $k \geqslant 1$ and $x \geqslant y \geqslant 2$, let $M_k(x, y)$ denote the number of integers $m \leqslant x$ such that $P_k(m)^2 | m$ and $P_k(m) > y$. Then, we have*

$$M_1(x, y) \ll_H x \frac{\mathrm{e}^{-Hu} + \mathrm{e}^{-H\sqrt{\log x}}}{y \log y}, \tag{2·4}$$

for any fixed integer $H > 0$ and uniformly in x, y, and

$$M_k(x, y) \ll \frac{x}{y \log x} \frac{(\log u + c_0)^{k-2}}{(k-2)!} \qquad (k \geqslant 2), \tag{2·5}$$

uniformly in k, x, y and with $u := (\log x)/\log y$.

Proof. We note that $M_k(x, y) = 0$ when $y > \sqrt{x}$. Our main ingredient is the bound

$$\Psi(x, y) \ll_c x\mathrm{e}^{-cu} + x^{1/3}, \tag{2·6}$$

valid for any $c > 0$, and proved in [15], exercise III.5.6 with solution.[6]

5. See [14], chapter II.5
6. Note that the exponent $1/3$ in (2·6) could be replaced by any positive constant. We do not need such precision here.

First consider the case $k=1$. We have by (2·6)

$$M_1(x,y)=\sum_{y<p\leqslant\sqrt{x}}\Psi(x/p^2,p)\ll_c\sum_{y<p\leqslant\sqrt{x}}\frac{x^{1-c/\log p}}{p^2}+\sum_{p\leqslant\sqrt{x}}\frac{x^{1/3}}{p^{2/3}}$$
$$\ll_c\int_y^{\sqrt{x}}\frac{x^{1-c/\log t}\,\mathrm{d}t}{t^2\log t}+\frac{\sqrt{x}}{\log x},$$

by partial summation. To bound the last integral, we observe that the function

$$t\mapsto c(\log x)/(\log t)+\tfrac{1}{2}\log t$$

is unimodal and attains its minimum at $t=t_c(x):=\mathrm{e}^{\sqrt{2c\log x}}$. Hence, for $2\leqslant y\leqslant t_c(x)$, the integral is

$$\ll\int_y^\infty\frac{x\mathrm{e}^{-\sqrt{2c\log x}}\,\mathrm{d}t}{t^{3/2}\log t}\ll x\frac{\mathrm{e}^{-\sqrt{2c\log x}}}{\sqrt{y}\log y}\leqslant x\frac{\mathrm{e}^{-d\sqrt{\log x}}}{y\log y},$$

with $d:=\sqrt{\tfrac{1}{2}c}$, while, for $t_c(x)<y\leqslant\sqrt{x}$, it is

$$\ll\int_y^\infty\frac{x\mathrm{e}^{-cu-\frac{1}{2}\log y}\,\mathrm{d}t}{t^{3/2}\log t}\ll x\frac{\mathrm{e}^{-cu}}{y\log y}.$$

Selecting, for instance, $c:=2H^2$ yields the required bound.

When $k\geqslant 2$, we observe that any integer m counted by $M_k(x,y)$ can be written as a product $m=ab$ with $P^-(b)>P_1(a)>y$, $\omega(b)=k-1$ and $P_1(a)^2|a$. Therefore, we plainly have

$$M_k(x,y)\leqslant\sum_{b\in\mathcal{E}_{k-1}(x,y)}M_1(x/b,y).$$

From (2·4) with, say, $H=1$, we deduce that

$$M_k(x,y)\ll\frac{x}{y\log y}\sum_{b\in\mathcal{E}_{k-1}(x,y)}\bigg(\frac{\mathrm{e}^{-u}}{b^{1-1/\log y}}+\frac{\mathrm{e}^{-\sqrt{\log x/b}}}{b}\bigg).\tag{2·7}$$

We estimate the b-sum above using (2·3) and splitting the range into intervals of the form $]y\mathrm{e}^j,y\mathrm{e}^{j+1}]$. We find that it is

$$\ll\frac{(\log u+c_0)^{k-2}}{(k-2)!}\sum_{0\leqslant j\leqslant J}\frac{\mathrm{e}^{-u+j/\log y}+\mathrm{e}^{-\sqrt{J-j}}}{j+\log y}$$

with $J:=1+[\log(x/y)]$. A simple calculation shows that the sum over j above is $\ll 1/u$. Inserting into (2·7) yields the stated bound.

Our next preliminary result is a rough estimate of the natural density $r_k(v)$ of those integers m such that $P_k(m)^v\leqslant m$. It is clear that $r_k(v)$ is the function given by (1·5).

Lemma 4. *Let $\delta \in]0,1[$. There exist positive constants c_1, c_2, c_3, possibly depending on δ, such that, for all real numbers $v \geqslant 1$ and all integers k with $2 \leqslant k \leqslant (1-\delta)\log v - c_3$, we have*

$$c_1 \frac{(\log v)^{k-2}}{(k-2)!\,v} \leqslant r_k(v) \leqslant c_2 \frac{(\log v)^{k-2}}{(k-2)!\,v}. \tag{2·8}$$

In particular, we have

$$r_k(v) \asymp_k (\log 2v)^{k-2}/v \qquad (v \geqslant 1) \tag{2·9}$$

for all fixed $k \geqslant 2$.

Proof. Let $R_k(n,v)$ denote the number of integers $m \leqslant n$ such that $P_k(m) \leqslant n^{1/v}$, so that

$$R_k(n,v) = r_k(v)n + o(n) \qquad (n \to \infty).$$

Any m counted by $R_k(n,v)$ may be written uniquely as a product $m = ab$ with $P_1(a) \leqslant n^{1/v}$ and $P^-(b) > n^{1/v}$, $0 \leqslant \omega(b) \leqslant k-1$. Let $R_j^*(n,v)$ denote the subsum corresponding to the condition $\omega(b) = j$. We plainly have

$$R_0^*(n,v) = \Psi(n, n^{1/v}) = n\varrho(v) + o(n) \qquad (n \to \infty),$$

so we may turn our attention to the case $1 \leqslant j \leqslant k-1$. Clearly

$$R_j^*(n,v) = \sum_{b \in \mathcal{E}_j(n, n^{1/v})} \Psi\big(n/b, n^{1/v}\big).$$

Using the bound (see [14], theorem III.5.1)

$$\Psi(x,y) \ll x\mathrm{e}^{-u/2} \qquad (x \geqslant 2,\ y \geqslant 2), \tag{2·10}$$

and applying Lemma 2 to estimate the b-sum via partial summation yields

$$R_j^*(n,v) \ll n \frac{(\log v + c_0)^{j-1}}{(j-1)!\,v}$$

uniformly for $n \geqslant 2$ and $v \leqslant (\log n)/\log 2$. We omit the details of this derivation, which are standard. Since this last upper bound is an increasing function of j for $j \leqslant \log v + c_0$, this proves the upper bound in (2·8).

To prove the lower bound, we apply a special case of a result of Sárközy [13] concerning the distribution of the function $\Omega(m,y)$, which counts *with* multiplicity the number of prime factors of m that exceed y. Write $N_k(x,y)$ for the number of integers $m \leqslant x$ with $\Omega(m,y) = k$. Sárközy's result, in the form given by Balazard [3], states that

$$N_k(x,y) \gg x \frac{(\log u)^{k-1}}{(k-1)!\,u} \tag{2·11}$$

uniformly for $x \geqslant y \geqslant 2$, $k \leqslant (2-\delta)\log u - c_3$, with $u := (\log x)/\log y$. We have, however,

$$R_{k-1}^*(n,v) \geqslant N_{k-1}\big(n, n^{1/v}\big) - \sum_{1 \leqslant j \leqslant k-1} M_j\big(n, n^{1/v}\big)$$

because any integer m counted by $N_{k-1}(n, n^{1/v})$ and not by the sum over j above has $k-1$ distinct prime factors greater than $n^{1/v}$. The lower bound estimate in (2·8) hence follows from (2·11), (2·4) and (2·5).

We note that (2·9) is an immediate consequence of (2·8) when v is sufficiently large in terms of k. Since $r_k(v)$ is obviously a non-increasing function of v, this is all that is needed.

3. Proof of the Theorem

Put $y_k := n^{\alpha_k}$. We may assume with no loss in generality that $y_k > \frac{3}{2}$. We plainly have

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = \frac{1}{n} \sum_{\substack{p_k<\cdots<p_1\leqslant n\\ p_j>n^{\alpha_j}\ (1\leqslant j\leqslant k)}} \Psi\Big(\frac{n}{p_1\cdots p_k}, p_k\Big) + O\Big(\frac{1}{n}\sum_{1\leqslant j\leqslant k-1} M_j(n, y_k)\Big). \tag{3·1}$$

The error term is zero for $k = 1$ and by (2·5) it is

$$\ll \frac{\{\log(1/\alpha_k) + c_0\}^{k-2}}{y_k(\log n)(k-2)!} \tag{3·2}$$

when $k \geqslant 2$.[7]

Let $D = D(\alpha_1, \dots, \alpha_k; n)$ denote the subset of $\mathbb{R}^k$ defined by the inequalities

$$(D)\begin{cases} \frac{3}{2} < t_k < \cdots < t_1 \leqslant n \\ t_j > n^{\alpha_j}\ (1 \leqslant j \leqslant k). \end{cases} \tag{3·3}$$

Then the main term in (3·1) may be re-written as

$$F_n^*(\overrightarrow{\boldsymbol{\alpha_k}}) := \frac{1}{n}\int_D \Psi\Big(\frac{n}{t_1\cdots t_k}, t_k\Big)\prod_{j=1}^k \mathrm{d}\pi(t_j).$$

We first show that

$$F_n^*(\overrightarrow{\boldsymbol{\alpha_k}}) = \frac{1}{n}\int_D \Psi\Big(\frac{n}{t_1\cdots t_k}, t_k\Big)\prod_{j=1}^k \frac{\mathrm{d}t_j}{\log t_j} + O\Big(\frac{1+\{\log(1/\alpha_k)\}^{k-2}}{L_\varepsilon(y_k)\log n}\Big). \tag{3·4}$$

This is obtained by inserting the formula

$$\prod_{j=1}^k \mathrm{d}\pi(t_j) - \prod_{j=1}^k \frac{\mathrm{d}t_j}{\log t_j} = \sum_{\ell=1}^k \mathrm{d}\mu_\ell(t_1, \dots, t_k),$$

with[8]

$$\mathrm{d}\mu_\ell(t_1,\dots,t_k) := \prod_{1\leqslant j\leqslant \ell}\mathrm{d}\pi(t_j)\prod_{\ell<j\leqslant k}\frac{\mathrm{d}t_j}{\log t_j} - \prod_{1\leqslant j<\ell}\mathrm{d}\pi(t_j)\prod_{\ell\leqslant j\leqslant k}\frac{\mathrm{d}t_j}{\log t_j},$$

and expanding the Ψ-term as a sum over integers. We thus see that the error term in (3·4) may be rewritten as

$$\frac{1}{n}\sum_{1\leqslant\ell\leqslant k}\sum_{m\leqslant n}\mu_\ell(D_m) = \sum_{1\leqslant\ell\leqslant k} Z_\ell,$$

7. We could actually obtain a slightly better bound, but such improvement would not be useful for our purpose.

8. We use the standard convention that an empty product is equal to 1.

say, where D_m is the intersection of D with the subset of $\mathbb{R}^k$ defined by the inequalities

$$P_1(m) \leqslant t_k, \quad t_1 \cdots t_k \leqslant n/m.$$

Using Fubini's theorem, we write $\mu_\ell(D_m)$ as a k-tuple integral with innermost term corresponding to the variable t_ℓ. By the prime number theorem, this is

$$\ll \frac{1}{m\prod_{j\neq\ell} t_j} L_{\varepsilon/3}\Big(\frac{n}{m\prod_{j\neq\ell} t_j}\Big)^{-1},$$

when $m\prod_{j\neq\ell} t_j \leqslant n/\max\{y_k, P_1(m)\}$, and zero otherwise.

If $k=1$, we sum trivially over m and obtain

$$\begin{aligned} Z_1 &\ll \sum_{\substack{P_1(m)\leqslant y_k\\ m\leqslant n/y_k}} \frac{1}{mL_{\varepsilon/3}(n/m)} + \sum_{\substack{P_1(m)>y_k\\ m\leqslant n/P_1(m)}} \frac{1}{mL_{\varepsilon/3}(n/m)} \\ &\ll \int_1^{n/y_k} \frac{\Psi(t,y_k)}{t^2 L_{\varepsilon/3}(n/t)}\,\mathrm{d}t + \sum_{y_k<p\leqslant\sqrt{n}} \frac{1}{p} \sum_{\substack{r\leqslant n/p^2\\ P_1(r)\leqslant p}} \frac{1}{rL_{\varepsilon/3}(n/pr)} \\ &\ll \int_1^{n/y_k} \frac{\Psi(t,y_k)}{t^2 L_{\varepsilon/3}(n/t)}\,\mathrm{d}t + \sum_{y_k<p\leqslant\sqrt{n}} \frac{1}{p} \int_1^{n/p^2} \frac{\Psi(t,p)}{t^2 L_{\varepsilon/3}(n/pt)}\,\mathrm{d}t. \end{aligned}$$

Inserting the bound

$$\begin{aligned} \int_1^{z/y} \frac{\Psi(t,y)}{t^2 L_{\varepsilon/3}(z/t)}\,\mathrm{d}t &\ll \int_{\log y}^{\log z} \exp\Big\{-\frac{\log z - v}{2\log y} - v^{(3-\varepsilon)/5}\Big\}\,\mathrm{d}v \\ &\ll \frac{z^{-1/(4\log y)}}{L_{\varepsilon/2}(y)} + \frac{1}{L_{\varepsilon/2}(z)}, \end{aligned}$$

which follows from (2·10) for $z \geqslant y > \frac{3}{2}$, we get

$$\begin{aligned} Z_1 &\ll \frac{1}{L_\varepsilon(y_k)\log n} + \sum_{y_k<p\leqslant\sqrt{n}} \frac{1}{p}\Big\{\frac{\mathrm{e}^{-(\log n)/8\log y_k}}{L_{\varepsilon/2}(y_k)} + \frac{1}{L_{2\varepsilon/3}(n)}\Big\} \\ &\ll \frac{1}{L_\varepsilon(y_k)\log n}. \end{aligned}$$

When $k>1$, we observe that, for $z \geqslant y^2 \geqslant 2$,

$$\int_y^{z/y} \frac{\mathrm{d}t}{tL_{\varepsilon/3}(z/t)\log t} + \sum_{y<p\leqslant z/y} \frac{1}{L_{\varepsilon/3}(z/p)p} \ll \frac{1}{L_{2\varepsilon/5}(y)\log z},$$

and, writing $T_{\ell,s} := \prod_{j\neq \ell,s} t_j$, we apply this with $y = \max\{P_1(m), y_k\}$, $z = n/T_{\ell,s}$, to the integral relative to t_s for some $s \in [1,k] \smallsetminus \{\ell\}$. This yields

$$Z_\ell \ll \int_{[y_k,n]^{k-2}} H\Big(\frac{n}{T_{\ell,s}}\Big) \prod_{\substack{1\leqslant j<\ell\\ j\neq s}} \frac{\mathrm{d}\pi(t_j)}{t_j} \prod_{\substack{\ell<j\leqslant k\\ j\neq s}} \frac{\mathrm{d}t_j}{t_j \log t_j} \tag{3·5}$$

with

$$H(z) := \sum_{m\leqslant z/\max\{y_k,P_1(m)\}} \frac{1}{m\log(z/m) L_{\varepsilon/2}\big(P_1(m)+y_k\big)}.$$

Applying the bounds

$$\sum_{\substack{P_1(m)\leqslant y\\ m\leqslant z/y}} \frac{1}{m\log(z/m)} \ll \int_y^{z/y} \frac{\Psi(t,y)}{t^2\log(z/t)}\,\mathrm{d}t \ll \frac{\log y}{\log z}$$

and

$$\sum_{y<p\leqslant\sqrt{z}} \sum_{\substack{r\leqslant z/p^2\\ P_1(r)\leqslant p}} \frac{1}{pr\log(z/pr)L_{\varepsilon/2}(p)} \ll \sum_{y<p\leqslant\sqrt{z}} \frac{\log p}{p\log(z/p)L_{\varepsilon/2}(p)} \ll \frac{1}{L_\varepsilon(y)\log z}$$

which readily follow from (2·10) for $z \geqslant y^2 \geqslant 2$, we derive

$$H(z) \ll \frac{1}{L_\varepsilon(y_k)\log z}$$

and so we finally arrive at

$$Z_\ell \ll \frac{1}{L_\varepsilon(y_k)} \int_{\substack{[y_k,n]^{k-2}\\ T_{\ell,s}\leqslant n/y_k}} \frac{1}{\log(n/T_{\ell,s})} \prod_{\substack{1\leqslant j<\ell\\ j\neq s}} \frac{\mathrm{d}\pi(t_j)}{t_j} \prod_{\substack{\ell<j\leqslant k\\ j\neq s}} \frac{\mathrm{d}t_j}{t_j\log t_j}.$$

Now employing inductively the estimate

$$\int_y^{z/y} \frac{\mathrm{d}t}{t\log(z/t)\log t} + \sum_{y<p\leqslant z/y} \frac{1}{p\log(z/p)} \ll \frac{\log\{(\log z)/\log y\}}{\log z},$$

valid for $z \geqslant y^2 \geqslant 2$, yields (3·4).

We now embark on the proof of (i), (ii) and (iii), and therefore assume $\alpha_k > \kappa(\varepsilon, n)$. The error term in (3·2) is then obviously $\ll 1/n^{\kappa(\varepsilon,n)}$ and, since

$$1/L_\varepsilon(y_k) \ll \mathrm{e}^{-(\log_2 n)^{1+\varepsilon/3}},$$

we deduce from (3·4) that, for any fixed H,

$$F_n^*(\overrightarrow{\boldsymbol{\alpha_k}}) = \frac{1}{n}\int_D \Psi\Big(\frac{n}{t_1\cdots t_k}, t_k\Big)\prod_{j=1}^k \frac{\mathrm{d}t_j}{\log t_j} + O\Big(\frac{1}{(\log n)^{H+1}}\Big). \tag{3·6}$$

Next, we use (1·7) in the form

$$\frac{1}{n}\Psi\Big(\frac{n}{z}, y\Big) = \Big\{1 + O\left(\mathrm{e}^{-(\log_2 n)^{1+\varepsilon/3}}\right)\Big\}\frac{1}{z}\int_{1-}^{\infty}\varrho\Big(\frac{\log n/z - \log t}{\log y}\Big)\,\mathrm{d}\Big(\frac{[t]}{t}\Big),$$

valid uniformly for $1 \leqslant z \leqslant n$, $y > n^{\kappa(\varepsilon,n)}$. Inserting this into (3·6) yields

$$F_n^*(\overrightarrow{\boldsymbol{\alpha_k}}) = F_n^{**}(\overrightarrow{\boldsymbol{\alpha_k}}) + O\Big(\frac{1}{(\log n)^{H+1}}\Big),$$

with

$$\begin{aligned} F_n^{**}(\overrightarrow{\boldsymbol{\alpha_k}}) &= \int_D\int_{1-}^{\infty}\varrho\Big(\frac{\log n/tt_1\cdots t_k}{\log t_k}\Big)\,\mathrm{d}\Big(\frac{[t]}{t}\Big)\prod_{j=1}^k\frac{\mathrm{d}t_j}{\log t_j} \\ &= \int_{\Delta_k}\int_{0-}^{\infty}\varrho\Big(\frac{1-v-w_k}{v_k}\Big)\,\mathrm{d}\Big(\frac{[n^v]}{n^v}\Big)\prod_{j=1}^k\frac{\mathrm{d}v_j}{v_j}. \end{aligned} \tag{3·7}$$

Here and in what follows, we set

$$w_k := \sum_{1\leqslant j\leqslant k} v_j$$

and let $\Delta_k = \Delta_k(\alpha_1, \dots, \alpha_k)$ denote the sub-domain of $\mathbb{R}^k$ defined by the conditions

$$\begin{cases} v_k < \cdots < v_1 \\ v_j > \alpha_j \ (1 \leqslant j \leqslant k). \end{cases} \tag{Δ_k}$$

Our next step consists in reversing the order of integrations and approximating the inner k-tuple integral by a Taylor-type expansion analogous to that established in [7] (Lemma 4.2) for the ϱ-function.

Lemma 5. *For* $\overrightarrow{\boldsymbol{\alpha_k}} := (\alpha_1, \dots, \alpha_k) \in [0, 1[^k$, *define*

$$G(v; \overrightarrow{\boldsymbol{\alpha_k}}) := \int_{\Delta_k}\varrho\Big(\frac{1-v-w_k}{v_k}\Big)\prod_{j=1}^k\frac{\mathrm{d}v_j}{v_j} \qquad (v \in \mathbb{R}).$$

(i) G is continuous and, for each integer $H \geqslant 1$, G is of class $\mathcal{C}^H$ except perhaps at finitely many exceptional points v_r $(1 \leqslant r \leqslant R_H)$, at which the derivatives of G may have discontinuities of the first kind.

(ii) For each r, we have $v_r = \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})$, where Λ_r is a affine linear form in the α_j satisfying (1·2).

(iii) There exists two sequences of functions defined on $[0,1]^k$, $\{\gamma_h\}_{h=0}^{\infty}$ and $\{\vartheta_h\}_{h=1}^{\infty}$ such that :

(a) γ_0 is continuous and, for each $h \geqslant 1$, γ_h and ϑ_h are continuous except perhaps at $v = \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})$ $(1 \leqslant r \leqslant R_h)$;

(b) for all $H \geqslant 0$ and $v \geqslant 0$, we have

$$G(v,\overrightarrow{\boldsymbol{\alpha_k}}) = \sum_{0 \leqslant h \leqslant H} \frac{\gamma_h(\overrightarrow{\boldsymbol{\alpha_k}})}{h!} v^h + J_H(v) + Y_H(v), \tag{3·8}$$

with

$$\begin{aligned} Y_H(v) &:= \int_0^v \frac{(v-w)^H}{H!} G^{(H+1)}(w;\overrightarrow{\boldsymbol{\alpha_k}})\,\mathrm{d}w, \\ J_H(v) &:= \sum_{\substack{1 \leqslant r \leqslant R_H \\ 0 \leqslant \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}}) \leqslant v}} \vartheta_r(\overrightarrow{\boldsymbol{\alpha_k}})\{v - \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})\}^{d_r}, \end{aligned}$$

where $d_r := h$ for $R_{h-1} < r \leqslant R_h$ $(1 \leqslant h \leqslant H)$, with the convention that $R_0 := 0$. Moreover, for each $h \geqslant 1$, we have

$$\vartheta_r(\overrightarrow{\boldsymbol{\alpha_k}}) \ll 1/\alpha_k^h \qquad (\alpha_k > 0,\ R_{h-1} < r \leqslant R_h). \tag{3·9}$$

Proof. We note that

$$\Delta_k(\alpha_1,\dots,\alpha_k) = \Delta_k\big(\alpha_1^*,\dots,\alpha_k^*\big)$$

with $\alpha_j^* := \max_{j \leqslant r \leqslant k} \alpha_r$. So, we assume in what follows, without loss of generality, that

$$\alpha_k \leqslant \alpha_{k-1} \leqslant \cdots \leqslant \alpha_1. \tag{3·10}$$

We first prove assertions (i) and (ii) of Lemma 5. For $k=1$, we have

$$G(v;\overrightarrow{\boldsymbol{\alpha_1}}) = \int_{\alpha_1}^{\infty} \varrho\Big(\frac{1-v}{v_1} - 1\Big)\frac{\mathrm{d}v_1}{v_1} = 1 - \varrho\Big(\frac{(1-v)^+}{\alpha_1}\Big),^{(9)}$$

with an obvious interpretation when $\alpha_1 = 0$, so the required properties are satisfied.

9. We use the classical notation $z^+ := \max(z,0)$.

For $k \geqslant 2$, we have

$$\begin{aligned} G(v;\overrightarrow{\boldsymbol{\alpha_k}}) &= \int_{\Delta_{k-1}} \prod_{j=1}^{k-1} \frac{\mathrm{d}v_j}{v_j} \int_{\alpha_k}^{v_{k-1}} \varrho\Big(\frac{1-v-w_{k-1}}{v_k}-1\Big)\frac{\mathrm{d}v_k}{v_k} \\ &= \int_{\Delta_{k-1}} \prod_{j=1}^{k-1} \frac{\mathrm{d}v_j}{v_j} \Big\{\varrho\Big(\frac{1-v-w_{k-1}}{v_{k-1}}\Big) - \varrho\Big(\frac{1-v-w_{k-1}}{\alpha_k}\Big)\Big\} \\ &= G(v;\overrightarrow{\boldsymbol{\alpha_{k-1}}}) - I_{k-1}(v;\overrightarrow{\boldsymbol{\alpha_k}}), \end{aligned}$$

say. Iterating, we obtain

$$\begin{aligned} G(v;\overrightarrow{\boldsymbol{\alpha_k}}) &= \int_{\alpha_1}^{\infty} \varrho\Big(\frac{1-v}{v_1}-1\Big)\frac{\mathrm{d}v_1}{v_1} - \sum_{2\leqslant j\leqslant k} I_{j-1}(v;\overrightarrow{\boldsymbol{\alpha_k}}) \\ &= 1-\varrho\Big(\frac{(1-v)^+}{\alpha_1}\Big) - \sum_{1\leqslant j\leqslant k-1} I_j(v;\overrightarrow{\boldsymbol{\alpha_k}}), \end{aligned}$$

with

$$I_j(v;\overrightarrow{\boldsymbol{\alpha_k}}) := \int_{\Delta_j} \varrho\Big(\frac{1-v-w_j}{\alpha_{j+1}}\Big)\prod_{r=1}^{j}\frac{\mathrm{d}v_r}{v_r}.$$

Let $D^hG(v;\overrightarrow{\boldsymbol{\alpha_k}})$ denote the hth derivative of $v\mapsto G(v;\overrightarrow{\boldsymbol{\alpha_k}})$ in the sense of distributions. For $v \geqslant 0$ and $h \geqslant 1$, we have

$$D^hG(v;\overrightarrow{\boldsymbol{\alpha_k}}) = G^{(h)}(v;\overrightarrow{\boldsymbol{\alpha_k}}) + \sum_{\substack{0\leqslant s\leqslant h-1\\ 0\leqslant m\leqslant s}} r_{m,s}\boldsymbol{\delta}_{m\alpha_1}^{(h-1-s)}(1-v)$$

with

$$\begin{aligned} G^{(h)}(v;\overrightarrow{\boldsymbol{\alpha_k}}) :=& \frac{(-1)^{h+1}}{\alpha_1^h}\varrho^{(h)}\Big(\frac{1-v}{\alpha_1}\Big) \\ &+ \sum_{1\leqslant j\leqslant k-1} \frac{(-1)^{h+1}}{\alpha_{j+1}^h}\int_{\Delta_j}\varrho^{(h)}\Big(\frac{1-v-w_j}{\alpha_{j+1}}\Big)\prod_{r=1}^{j}\frac{\mathrm{d}v_r}{v_r} \\ &+ \sum_{1\leqslant j\leqslant k-1}\sum_{\substack{0\leqslant s\leqslant h-1\\ 0\leqslant m\leqslant s}} r_{m,s}\int_{\Delta_j}\boldsymbol{\delta}_{m\alpha_j}^{(h-1-s)}(1-v-w_j)\prod_{r=1}^{j}\frac{\mathrm{d}v_r}{v_r} \end{aligned} \tag{3·11}$$

where

$$r_{m,s} := \varrho^{(s)}(m) - \varrho^{(s)}(m-) \qquad (0\leqslant m\leqslant s)$$

and $\boldsymbol{\delta}_x^{(t)}$ denotes the tth derivative of the Dirac measure at the point x.

This shows that $G(v;\overrightarrow{\boldsymbol{\alpha_k}})$ is h times differentiable except perhaps when $1-v=m\alpha_1$ with $0\leqslant m\leqslant h$, or when v is equal to an extremum of $1-m\alpha_{j+1}-w_j$ on Δ_j for some pair (j,m) with $1\leqslant j\leqslant k-1$, $0\leqslant m\leqslant h$. The right-hand side of (3·11) coincides with the hth derivative of $G(v;\overrightarrow{\boldsymbol{\alpha_k}})$, and defines a continuous function of v, outside the set of exceptional values. All possible discontinuities are of the form $\Lambda_r(\overrightarrow{\boldsymbol{\alpha}}_k)$ $(1\leqslant r\leqslant R_h)$, where Λ_r is, for each r, an affine linear form in k variables. The Λ_r are clearly of the form indicated in (1·2). Moreover, a simple computation suffices to show that each term arising on the left-hand side of (3·11) is $\ll 1/\alpha_k^h$. This implies, in particular, that the discontinuity $\vartheta_r^*(\overrightarrow{\boldsymbol{\alpha}}_k)$ of $G^{(h)}$ at $\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})$ satisfies

$$\vartheta_r^*(\overrightarrow{\boldsymbol{\alpha_k}})\ll 1/\alpha_k^h \qquad (1\leqslant r\leqslant R_h). \tag{3·12}$$

It remains to prove assertion (iii). We proceed by induction on H. When $H=0$, the formula follows by integrating the identity

$$\mathrm{d}G(w;\overrightarrow{\boldsymbol{\alpha_k}})=G'(w;\overrightarrow{\boldsymbol{\alpha_k}})\,\mathrm{d}w,$$

which holds because G is continuous, differentiable except perhaps at finitely many points and because G' is integrable.

Let us then assume (3·8) holds for a given $H\geqslant 0$. We have the equality between measures

$$\mathrm{d}G^{(H+1)}(w;\overrightarrow{\boldsymbol{\alpha_k}})=G^{(H+2)}(w;\overrightarrow{\boldsymbol{\alpha_k}})\,\mathrm{d}w+\sum_{R_H<r\leqslant R_{H+1}}\vartheta_r^*(\overrightarrow{\boldsymbol{\alpha_k}})\boldsymbol{\delta}_{\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})}(w)$$

where $\boldsymbol{\delta}_u(v)$ denotes the Dirac measure at $v=u$ and $\vartheta_r^*(\overrightarrow{\boldsymbol{\alpha_k}})$ is, for each r in the range $R_H<r\leqslant R_{H+1}$, the saltus of $G^{(H+1)}(v;\overrightarrow{\boldsymbol{\alpha_k}})$ at $\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})$. We deduce that

$$\begin{aligned}\int_0^v &(v-w)^H G^{(H+1)}(w;\overrightarrow{\boldsymbol{\alpha_k}})\,\mathrm{d}w\\ &=-\left[\frac{(v-w)^{H+1}G^{(H+1)}(w;\overrightarrow{\boldsymbol{\alpha_k}})}{H+1}\right]_{0+}^{v+}+\int_0^v\frac{(v-w)^{H+1}}{H+1}\,\mathrm{d}G^{(H+1)}(w;\overrightarrow{\boldsymbol{\alpha_k}})\\ &=\frac{G^{(H+1)}(0+;\overrightarrow{\boldsymbol{\alpha_k}})v^{H+1}}{H+1}+\sum_{\substack{R_H<r\leqslant R_{H+1}\\ 0\leqslant\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})\leqslant v}}\vartheta_r^*(\overrightarrow{\boldsymbol{\alpha_k}})\frac{\{v-\Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})\}^{H+1}}{H+1}\\ &\quad+\int_0^v\frac{(v-w)^{H+1}}{H+1}G^{(H+2)}(w;\overrightarrow{\boldsymbol{\alpha_k}})\,\mathrm{d}w.\end{aligned}$$

We set

$$\begin{aligned}\gamma_{H+1}(\overrightarrow{\boldsymbol{\alpha_k}})&:=G^{(H+1)}(0+;\overrightarrow{\boldsymbol{\alpha_k}}),\\ \vartheta_r(\overrightarrow{\boldsymbol{\alpha_k}})&:=\frac{\vartheta_r^*(\overrightarrow{\boldsymbol{\alpha_k}})}{(H+1)!}\quad (R_H<r\leqslant R_{H+1}).\end{aligned}$$

Dividing through by $H!$, applying the induction hypothesis, and taking (3·12) into account yields the required result.

We are now in a position to complete the proof of our theorem.

The contribution to $F_n^{**}(\overrightarrow{\boldsymbol{\alpha_k}})$ from the main term in (3·8) may be computed using the formula (see e.g. [7])

$$\frac{(-1)^h}{h!}\int_{0-}^{\infty} v^h \,\mathrm{d}\Big(\frac{[n^v]}{n^v}\Big) = \frac{a_h}{(\log n)^h},$$

where the sequence $\{a_h\}_{h=1}^{\infty}$ is defined by the Taylor expansion

$$\frac{s\zeta(s+1)}{s+1} = \sum_{h=0}^{\infty} a_h s^h \qquad (|s|<1),$$

involving the Riemann zeta function. We thus obtain the contribution

$$\sum_{0\leqslant h\leqslant H} \frac{\varphi_h(\overrightarrow{\boldsymbol{\alpha_k}})}{(\log n)^h}$$

with $\varphi_h(\overrightarrow{\boldsymbol{\alpha_k}}) := (-1)^h a_h \gamma_h(\overrightarrow{\boldsymbol{\alpha_k}})$. It follows from the definition of the γ_h that, for fixed $h \geqslant 1$, we have

$$\varphi_h(\alpha_k) \ll_h 1/\alpha_k^h \qquad (\alpha_k > 0).$$

In particular,

$$\varphi_0(\overrightarrow{\boldsymbol{\alpha_k}}) = \gamma_0(\overrightarrow{\boldsymbol{\alpha_k}}) = \int_{\Delta_k} \varrho\Big(\frac{1-w_k}{v_k}\Big)\prod_{j=1}^{q}\frac{\mathrm{d}v_j}{v_j} \tag{3·13}$$

is bounded as a function of $\overrightarrow{\boldsymbol{\alpha_k}}$—actually $0 \leqslant \varphi_0(\overrightarrow{\boldsymbol{\alpha_k}}) \leqslant 1$.

We estimate the contributions to $F_n^{**}(\overrightarrow{\boldsymbol{\alpha_k}})$ from the remainder terms in (3·8) by partial summation as in [7]—equations (4·25) and (4·29)—to find that they are

$$\ll_H \frac{1}{(\alpha_k \log n)^{H+1}}.$$

We omit the technical details since they are identical to those appearing in [7]. This yields the formula

$$F_n^{**}(\overrightarrow{\boldsymbol{\alpha_k}}) = \sum_{0\leqslant h\leqslant H} \frac{\varphi_h(\overrightarrow{\boldsymbol{\alpha_k}})}{(\log n)^h} + O_H\Big(\frac{1}{(\alpha_k \log n)^{H+1}}\Big), \tag{3·14}$$

uniformly for $\overrightarrow{\boldsymbol{\alpha_k}} \in [0,1]^k$ and under the conditions

$$\alpha_k > 0, \qquad \min_{\substack{1\leqslant r\leqslant R_H\\ \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}})>0}} \Lambda_r(\overrightarrow{\boldsymbol{\alpha_k}}) > K_H \frac{\log_2 n}{\log n}.$$

The proof of assertions (i), (ii) and (iii) is therefore complete.

We now establish (iv). By (3·1) and (3·2), we have

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = \frac{1}{n} \sum_{\substack{p_k<\cdots<p_1\leqslant n\\ p_j>n^{\alpha_j}\ (1\leqslant j\leqslant k)}} \Psi\Big(\frac{n}{p_1\cdots p_k}, p_k\Big) + O\Big(\frac{\{\log(1/\alpha_k)+c_0\}^{k-2}}{(\log n)y_k(k-2)!}\Big). \tag{3·15}$$

for $k \geqslant 2$ and with no lower bound restriction upon α_k or y_k.[10] Instead of applying immediately (3·4), we first compute the inner p_k-sum by Buchstab's identity

$$\sum_{y<p\leqslant z} \Psi(x/p, p) = \Psi(x,z) - \Psi(x,y) \qquad (x \geqslant 1,\ z \geqslant y \geqslant 1).$$

To ease notation, we write $q := k-1 \geqslant 1$ throughout. The main term in (3·15) is hence equal to

$$\frac{1}{n} \sum_{\substack{p_q<\cdots<p_1\leqslant n\\ p_j>n^{\beta_j}\ (1\leqslant j\leqslant q)}} \Big\{\Psi\Big(\frac{n}{p_1\cdots p_q}, p_q\Big) - \Psi\Big(\frac{n}{p_1\cdots p_q}, n^{\alpha_k}\Big)\Big\}$$

with $\beta_j := \max(\alpha_j, \alpha_k)$ $(1 \leqslant j \leqslant q)$. By an argument similar to that which yields (3·4), we get from this and (3·15) that

$$\begin{aligned} F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = {} & \frac{1}{n}\int_{D^*} \Big\{\Psi\Big(\frac{n}{t_1\cdots t_q}, t_q\Big) - \Psi\Big(\frac{n}{t_1\cdots t_q}, n^{\alpha_k}\Big)\Big\} \prod_{j=1}^{q} \frac{\mathrm{d}t_j}{\log t_j} \\ & + O\Big(\frac{\{\log(1/\alpha_k)+c_0\}^{k-2}}{(\log n)L_\varepsilon(y_k)(k-2)!}\Big), \end{aligned}$$

where $D^* := D(\beta_1, \ldots, \beta_q; n)$ with the notation introduced in (3·3).

We want to approximate the main term by $\varphi_0(\overrightarrow{\boldsymbol{\alpha_k}})$ and insert Hildebrand's asymptotic formula [9] in the form

$$\Psi(x,y) = x\varrho(u)\Big\{1 + O\Big(\frac{\log(2+u)}{\log 2y}\Big)\Big\} + O(1), \tag{3·16}$$

valid uniformly for $x \geqslant 1$, $y \geqslant x^{\kappa(\varepsilon,x)}$.[11] The contribution from the term $O(1)$ in (3·16) is

$$\begin{aligned} & \ll \frac{1}{n}\int_{\substack{[y_k,n]^q\\ t_1\cdots t_q\leqslant n}} \prod_{j=1}^{q} \frac{\mathrm{d}t_j}{\log t_j} \\ & \ll \int_{\substack{[y_k,n]^{q-1}\\ t_1\cdots t_{q-1}\leqslant n/y_k}} \frac{1}{\log(2n/t_1\cdots t_{q-1})} \prod_{j=1}^{q-1} \frac{\mathrm{d}t_j}{t_j\log t_j} \ll \frac{\{\log(1/\alpha_k)\}^{q-1}}{\log n}. \end{aligned}$$

10. Of course, we still assume $y_k > \frac{3}{2}$.

11. The remainder term $O(1)$ takes care of very large values of y. It is, *a posteriori*, a simple matter to deduce (3·16) from (1·7).

Here and in the rest of this proof, all implicit constants may depend upon q. We obtain from the above bound :

$$F_n(\overrightarrow{\boldsymbol{\alpha_k}}) = \varphi_0(\overrightarrow{\boldsymbol{\alpha_k}}) + O\Big(\sum_{1\leqslant j\leqslant 3} W_j(n)\Big) + O\Big(\frac{\{\log(1/\alpha_k)\}^{q-1}}{\log n}\Big) \tag{3·17}$$

and

$$\begin{aligned} W_1(n) &:= \frac{1}{n}\int_{\overline{D^*}} \Big\{\Psi\Big(\frac{n}{t_1\cdots t_q}, t_q\Big) - \Psi\Big(\frac{n}{t_1\cdots t_q}, n^{\alpha_k}\Big)\Big\} \prod_{j=1}^{q} \frac{\mathrm{d}t_j}{\log t_j},\\ W_2(n) &:= \int_{\overline{\Delta}_q^*} \varrho\Big(\frac{1-w_q}{v_q}\Big)\prod_{j=1}^{q}\frac{\mathrm{d}v_j}{v_j},\\ W_3(n) &:= \frac{1}{\log n}\int_{\Delta_q^*} \varrho\Big(\frac{1-w_q}{v_q}\Big)\log\Big(2+\frac{1-w_q}{v_q}\Big)\prod_{j=1}^{q-1}\frac{\mathrm{d}v_j}{v_j}\frac{\mathrm{d}v_q}{v_q^2}, \end{aligned}$$

where $\overline{D^*}$ is the sub-domain of D^* corresponding to the extra condition

$$n/t_1\cdots t_q > L_q := \exp\exp\sqrt{\log t_q}, \tag{3·18}$$

Δ_q^* is the domain defined by the conditions

$$\begin{cases} v_q < \cdots < v_1 \\ v_j > \beta_j \ (1 \leqslant j \leqslant q), \end{cases}$$

and $\overline{\Delta}_q^*$ is the sub-domain of Δ_q^* corresponding to the supplementary conditions

$$1 - w_q = 1 - \sum_{1\leqslant j\leqslant q} v_j > z(v_q, n) := \frac{\mathrm{e}^{\sqrt{v_q \log n}}}{\log n}.$$

The quantities $W_1(n)$ and $W_2(n)$ can be treated in a very similar manner, and we only consider the first. The integration range is empty unless $y_k := n^{\alpha_k} \leqslant \lambda(n) := \exp(\log_2 n)^2$. When this is realised, we see by (2·10) that

$$\begin{aligned} W_1(n) &\ll \int_{y_k}^{\lambda(n)} \frac{\mathrm{d}t_q}{\log t_q} \int_{\substack{[y_k,n]^{q-1}\\ t_1\cdots t_{q-1}\leqslant n/t_qL_q}} \frac{\mathrm{e}^{-(\log n)/(2\log t_q)}}{(t_1\cdots t_q)^{1-1/(2\log t_q)}} \prod_{j=1}^{q-1}\frac{\mathrm{d}t_j}{\log t_j}\\ &\ll \int_{\alpha_k}^{(\log_2 n)^2/\log n} \frac{\mathrm{d}v_q}{v_q} \int_{\substack{[v_q,1]^{q-1}\\ w_{q-1}\leqslant 1-v_q-z(v_q,n)}} \mathrm{e}^{-(1-w_{q-1})/2v_q} \prod_{j=1}^{q-1}\frac{\mathrm{d}v_j}{v_j}. \end{aligned} \tag{3·19}$$

If $q = 1$, we trivially have

$$W_1(n) \ll \int_0^{(\log_2 n)^2/\log n} \mathrm{e}^{-1/2v_q}\frac{\mathrm{d}v_q}{v_q} \ll \frac{1}{\log n}.$$

When $q \geqslant 2$, we choose the v_{q-1}-integral as the innermost. It is

$$\begin{aligned}&\ll \mathrm{e}^{-(1-w_{q-2})/2v_q} \int_{v_q}^{1-w_{q-2}-z(v_q,n)/2} \mathrm{e}^{v_{q-1}/2v_q} \frac{\mathrm{d}v_{q-1}}{v_{q-1}} \\ &\ll \frac{\mathrm{e}^{-z(v_q,n)/4v_q} v_q}{1-w_{q-2}-z(v_q,n)/2} = \frac{\mathrm{e}^{-\nu(v_q \log n)} v_q}{1-w_{q-2}-z(v_q,n)/2},\end{aligned}$$

with $\nu(v) := \mathrm{e}^{\sqrt{v}}/4v$. We then observe that, since $z(v_q, n) \leqslant 1$,

$$\int_{\substack{[v_q,1]^{q-2} \\ w_{q-2} \leqslant 1-v_q-z(v_q,n)}} \frac{1}{1-w_{q-2}-z(v_q,n)/2} \prod_{1 \leqslant j \leqslant q-2} \frac{\mathrm{d}v_j}{v_j} \ll (1+\log 1/v_q)^{q-2},$$

as can be seen by a simple calculation. We thus obtain

$$\begin{aligned}W_1(n) &\ll (1+\log 1/\alpha_k)^{q-2} \int_{\alpha_k}^{(\log_2 n)^2/\log n} \mathrm{e}^{-\nu(v_q \log n)}\, \mathrm{d}v_q \\ &\ll \frac{(1+\log 1/\alpha_k)^{q-2}}{\log n} \int_{\log y_k}^{(\log_2 n)^2} \mathrm{e}^{-\nu(y)}\, \mathrm{d}y \ll \frac{(1+\log 1/\alpha_k)^{q-2}}{y_k \log n}.\end{aligned}$$

The same upper bound holds for $W_2(n)$.

It remains to bound $W_3(n)$. The inner v_q-integral is zero if $1 - w_{q-1} < \alpha_k$. Otherwise it does not exceed

$$\begin{aligned}&\int_{\beta_q}^{v_{q-1}} \varrho\Big(\frac{1-w_{q-1}}{v_q} - 1\Big) \log\Big(1 + \frac{1-w_{q-1}}{v_q}\Big) \frac{\mathrm{d}v_q}{v_q^2} \\ &\qquad \leqslant \frac{1}{1-w_{q-1}} \int_{(1-w_{q-1})/v_{q-1}}^{\infty} \varrho(w-1)\log(1+w)\,\mathrm{d}w \ll \frac{1}{1-w_{q-1}}.\end{aligned}$$

Inserting back in the multiple integral yields

$$\begin{aligned}W_3(n) &\ll \frac{1}{\log n} \int_{\substack{[\alpha_k,1]^{q-1} \\ 1-w_{q-1} \leqslant \alpha_k}} \frac{1}{1-w_{q-1}} \prod_{1 \leqslant j \leqslant q-1} \frac{\mathrm{d}v_j}{v_j} \\ &\ll \frac{\{1+\log(1/\alpha_k)\}^{q-1}}{\log n}.\end{aligned}$$

Inserting our estimates into (3·17), we obtain (iv), as stated.

Acknowledgements. The author takes pleasure in thanking here Richard Arratia for the friendly email conversation which motivated this study and Jie Wu for his help during the preparation of this article.

References

[1] Arratia, R. On the central role of scale invariant Poisson processes on $(0,\infty)$. *Microsurveys in Discrete Probability* (Princeton, NJ, 1997), 21–41, (edited by D. Aldous and J. Propp) DIMACS Ser. Discrete Math. Theoret. Comput. Sci. 41 (1998), Am. Math. Soc., Providence, RI.

[2] Arratia, R., Barbour, A. D., Tavaré, S., The Poisson Dirichlet distribution and the scale invariant Poisson process, *Comb., Prob. and Comp.*, to appear.

[3] Balazard, M., Remarques sur un théorème de G. Halász et A. Sárközy, *Bull. Soc. Math. France* **117** (1989), 389–413.

[4] Billingsley, P., On the distribution of large prime factors, *Period. Math. Hungar.* **2** (1972), 283–289.

[5] de Bruijn, N. G., On the number of positive integers $\leqslant x$ and free of prime factors $> y$, *Nederl. Akad. Wetensch. Proc. Ser. A* **54** (1951), 50–60.

[6] Donnelly, P., and Grimmett, G., On the asymptotic distribution of large prime factors, *J. London Math. Soc.* (2) **47** (1993), 395–404.

[7] Fouvry, É., and Tenenbaum, G., Entiers sans grand facteur premier en progressions arithmétiques, *Proc. London Math. Soc.* (3) **63** (1991), 449–494.

[8] Hafner, J. L., and McCurley, K. S., On the distribution of running time of certain integer factoring algorithm, *J. Algorithms* **10** *no. 4* (1989), 531–556.

[9] Hildebrand, A., On the number of positive integers $\leqslant x$ and free of prime factors $> y$, *J. Number Theory* **22** (1986), 289-307.

[10] Kingman, J. F. C., *Poisson Processes*, Oxford University Press, Oxford 1993.

[11] Knuth, D., and Trabb Pardo, L., Analysis of a simple factorization algorithm, *J. Theoret. Comput. Sci.* **3** (1976), 321–348.

[12] Saias, É., Sur le nombre des entiers sans grand facteur premier, *J. Number Theory* **32** (1989), 78–99.

[13] Sárközy, A., Remarks on a paper of G. Halász, *Period. Math. Hungar.* **8** (1977), 135–150.

[14] Tenenbaum, G., *Introduction to analytic and probabilistic number theory*, Cambridge studies in advanced mathematics, no. 46, Cambridge University Press, Cambridge 1995.

[15] Tenenbaum, G., in collaboration with J. Wu, *Exercices corrigés de théorie analytique et probabiliste des nombres*, Cours spécialisés, no. 2, Société Mathématique de France (1996), xiv + 251 pp.

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