Ramanujan Journal
51 (2020), 153–161.

On local laws for the number of small prime factors*

Gérald Tenenbaum

Abstract. Investigating a question of Alladi, we describe the local distribution of small prime factors of integers, with emphasis on the transition phase occurring for certain values of the parameters.

Keywords: local distribution of prime factors, restricted prime factors, small prime factors, saddle-point method.

2010 Mathematics Subject Classification: 11N25, 11N37, 11N60.

1. Introduction and statement of results

Let $\nu(n,y) := \sum_{p|n,\, p\leqslant y} 1$ denote the number of prime factors not exceeding y, counted without multiplicity, of a natural integer n. K. Alladi proposed to estimate

$$N_k(x,y) := \sum_{\substack{n\leqslant x\\ \nu(n,y)=k}} 1 \qquad (x\geqslant y\geqslant 2,\ k\ll \log_2 y)$$

with the purpose of describing the phase transition between the case of small y, when we expect

$$N_k(x,y)\asymp \frac{x(\log_2 y)^k}{k!\log y} \tag{1·1}$$

and the case of large y, when the parameter k, in the above estimate, should be replaced by $k-1$.(1) Of course, as stated, this phenomenon can only be brought to light when $r:=k/\log_2 y$ is small.

As we shall see, this question is quite appealing and a complete answer still eludes us. We propose here, as a benchmark, a first, concise study resting principally on results derived through the saddle-point method. In parallel to the present work, Alladi and Molnar [1] have tackled the problem using mainly Buchstab's iteration method.

We observe incidentally that the situation changes in nature when $k/\log_2 y$ becomes large. For instance, we have from corollary 1 in [5],

$$\pi_k(x) := N_k(x,x)\sim \frac{F(r)}{\Gamma(r+1)}\frac{\mathrm{e}^{-kh/2}x(\log_2 x)^{k-1}}{(k-1)!\log x} = o\Big(\frac{x(\log_2 x)^{k-1}}{(k-1)!\log x}\Big)$$

if $r:=k/\log_2 x\to\infty$, $k=o\big((\log_2 x)^2\big)$, with

$$F(z):=\prod_p\Big(1+\frac{z}{p-1}\Big)\Big(1-\frac1p\Big)^z,\quad h=\Big(\frac{\log(2+\mathrm{e}^\gamma y\log y)}{\log_2 x}\Big)^2. \tag{1·2}$$

In the case of relatively small values of y, an asymptotic formula with remainder easily follows from estimates established in [3]. Here and in the sequel, γ denotes Euler's constant.

---

* We include here the correction of a misprint appearing in the statement of Theorem 1.5 of the published version.

1. Here and in the sequel we let $\log_j$ denote the j-th iterated logarithm.

Theorem 1.1. *There exists an absolute constant $c > 0$ such that, uniformly under the conditions $3 \leqslant y \leqslant x^{c(\log_3 x)/\log_2 x}$, $r := k/\log_2 y \ll 1$, we have*

$$N_k(x,y) = F(r)\mathrm{e}^{\gamma(r-1)}\frac{x(\log_2 y)^k}{k!\log y}\Big\{1+O\Big(\frac{k}{(\log_2 y)^2}\Big)\Big\}. \tag{1·3}$$

Next, we investigate the case of larger values of y. The following result is an easy consequence of a special case of corollary 2.4 in [8]. We write

$$S_z(x,y) := \sum_{n\leqslant x} z^{\nu(n,y)} \qquad (x \geqslant y \geqslant 1,\ z\in\mathbb{C}).$$

Theorem 1.2. *Let $\kappa > 0$. Uniformly for $3 \leqslant y \leqslant x$, $1 \leqslant k \leqslant (\log_2 y)/\kappa$, we have*

$$N_k(x,y) \ll x\frac{(\log_2 y)^k}{k!\,\log y}\cdot \tag{1·4}$$

Moreover, for any fixed $\kappa \in]0,1[$ and uniformly for $3 \leqslant y \leqslant x$, $\kappa \leqslant r := k/\log_2 y \leqslant 1/\kappa$, we have

$$N_k(x,y) = S_r(x,y)\frac{(\log_2 y)^k}{k!\mathrm{e}^k}\Big\{1+O\Big(\frac{1}{\sqrt{\log_2 y}}\Big)\Big\}. \tag{1·5}$$

In order to exploit (1·5) we need to estimate $S_r(x,y)$. This depends on the quantity

$$\sigma_r(u) := \int_0^{u-1}\omega(u-t)\varrho_r(t)\,\mathrm{d}t + \varrho_r(u) \quad (u \geqslant 1),$$

where ϱ_r is the r-th fractional convolution power of the Dickman function $\varrho = \varrho_1$—it satisfies $\varrho_r(u) = \varrho(u)r^{u+o(u)}$ as $u\to\infty$, see [9] for further details—and ω is Buchstab's function—see, e.g., [6], §III.6.3. Appealing to the estimate $\omega(t) = \mathrm{e}^\gamma + O(t^{-t})$, proved for instance in [6], th. III.6.8, we easily obtain that for fixed $r > 0$,

$$\sigma_r(u) > 0 \quad (u > 1), \qquad \sigma_r(u) = \mathrm{e}^{\gamma(r-1)} + O(u^{-u}) \qquad (u\to\infty). \tag{1·6}$$

For fixed $\varepsilon > 0$, we define the domain

$$\exp\big\{(\log_2 x)^{5/3+\varepsilon}\big\} \leqslant y \leqslant x. \tag{H_ε}$$

Theorem 1.3. *Let $\varepsilon > 0$. Uniformly for $(x,y)\in(H_\varepsilon)$, $3 \leqslant y \leqslant \sqrt{x}$, $r := k/\log_2 y \ll 1$, and with $u := (\log x)/\log y$, we have*

$$S_r(x,y) = F(r)\sigma_r(u)x(\log y)^{r-1}\Big\{1+O\Big(\frac{1}{\log y}\Big)\Big\}. \tag{1·7}$$

Consequently, for any fixed $\kappa\in]0,1[$ and uniformly for $\sqrt{x} \geqslant y \geqslant 3$, $\kappa \leqslant r \leqslant 1/\kappa$, we have

$$N_k(x,y) = \frac{F(r)\sigma_r(u)x(\log_2 y)^k}{k!\log y}\Big\{1+O\Big(\frac{1}{\sqrt{\log_2 y}}\Big)\Big\}. \tag{1·8}$$

Note that, in view of (1·6), formulae (1·8) and (1·3) coincide for large u.

At this point, we are left with two cases :

(a) $x^{c(\log_3 x)/\log_2 x} < y \leqslant \sqrt{x}$ and $k \leqslant \kappa\log_2 y$ with fixed, arbitrary small $\kappa > 0$;

(b) $\sqrt{x} < y \leqslant x$, $1 \leqslant k \leqslant (\log_2 y)/\kappa$.

We first show that behaviour (1·1) holds throughout range (a), and somewhat beyond.

Theorem 1.4. *Let* $c \in]0,1[$ *and* $\kappa > 0$ *be fixed. Then, estimate* (1·1) *holds uniformly for* $2 \leqslant y \leqslant x^{1-c}$, $1 \leqslant k \leqslant (\log_2 y)/\kappa$.

From the above, we now know that the desired threshold must occur in range (b). Our final, general statement takes into account the case when $w := (\log x)/\log(3x/y)$ is large.

Theorem 1.5. *Let* $\kappa > 0$. *Uniformly for* $x \geqslant y \geqslant 3$, $1 \leqslant k \leqslant (\log_2 y)/\kappa$, *we have*

$$N_k(x,y) \asymp \frac{x(\log_2 y)^{k-1}}{(k-1)!\log y} + \frac{x(\log_2 \min(3x/y,y))^k}{k!\log y}. \tag{1·9}$$

It follows in particular from this statement that, when $w \geqslant 3$, the estimate (1·1) holds if $1 \leqslant k \ll (\log_2 y)/\log w$, while we have

$$N_k(x,y) \asymp \frac{x(\log_2 y)^{k-1}}{(k-1)!\log y} \tag{1·10}$$

when $1 + (\log_2 y)(\log_2 w)/\log w \leqslant k \ll \log_2 y$. In the middle range, formula (1·9) describes the transition.

We have thus determined the true order of magnitude of $N_k(x,y)$ for large y and all $k \ll \log_2 y$, and, in a number of cases, provided an asymptotic formula with remainder. It is a challenging problem, with interesting methodological issues, to obtain asymptotic estimates in the remaining ranges.

2. Proof of Theorem 1.1

Let $z \in \mathbb{C}$. By a slight modification of the proof of theorem 02 of [4] (i.e. appealing to classical upper bounds for the number $\Psi(x,y)$ of y-friable integers not exceeding x instead of the simple Rankin bound used in [4]) we get, uniformly for $2 \leqslant y \leqslant x^{c(\log_3 x)/\log_2 x}$, $z \ll 1$,

$$S_z(x,y) = x\prod_{p\leqslant y}\Big(1+\frac{z-1}{p}\Big) + O\Big(\frac{x}{(\log x)^2}\Big).$$

Since the main term of (1·3) is of order $\gg x/\log y$, we see that $N_k(x,y)/x$ is, to the stated accuracy, approximated by the coefficient of z^k in the product over primes. However, this has been studied in [3]. For $k \ll \log_2 y$, let us write

$$L := \log\Big(\frac{\log y}{\log(k+1)}\Big), \quad M := \log_2 y - \log\big(1+\log^+(k/L)\big) = \log_2 y + O(1), \quad \varrho := k/M.$$

Then, by corollary 1 of [3],

$$N_k(x,y) = x\prod_{p\leqslant y}\Big(1+\frac{\varrho-1}{p}\Big)\frac{M^k}{k!\,\mathrm{e}^k}\Big\{1+O\Big(\frac{k}{(\log_2 y)^2}\Big)\Big\}. \tag{2·1}$$

Formula (1·3) follows by writing $\varrho = r + O\big(r/\log_2 y\big)$ and noticing that the main terms of (1·3) and (2·1) coincide to within the stated accuracy.

3. Proof of Theorem 1.2

The upper bound(1·4) follows immediately from corollary 2.4(iii) of [8]. The same statement provides (1·5) with $E(y) := \sum_{p\leqslant y} 1/p$ instead of $\log_2 y$ and $k/E(y)$ instead of r. However, shifting $E(y)$ by a bounded amount does not perturb the asymptotic formula as stated. This can be seen in either of two ways. The first is by directly inspecting the proof displayed in [8], where a quantity $\mathrm{e}^{(z-r)E(y)}\{1+O(\vartheta)\}$ gets integrated over the circle $z = r\mathrm{e}^{i\vartheta}$; the second is by using the simple estimate

$$S_r(x,y) \asymp x(\log y)^{r-1} \qquad (r \asymp 1), \tag{3·1}$$

which readily follows, for instance, from theorem 1.1 of [7]. We only outline this approach. Applying (3·1) with $(1 \pm v)r$ in place of r and choosing v optimally, we obtain, for bounded, positive v,

$$\frac{1}{S_r(x,y)} \sum_{\substack{n\leqslant x\\ |\nu(n,y)-r\log_2 y|>vr\log_2 y}} r^{\nu(n,y)} \ll (\log y)^{-rQ(v)}$$

with $Q(v) \asymp v^2$ if $0 \leqslant v \leqslant 1/2$, and $Q(v) \asymp (1+v)\log(1+v)$ if $v > 1/2$. Integrating over v, we obtain

$$\sum_{n\leqslant x} r^{\nu(n,y)}\{\nu(n,y) - r\log_2 y\}^2 \ll S_r(x,y)\log_2 y. \tag{3·2}$$

From this, we see that $S_r(x,y)r^{-k} = \sum_{n\leqslant x} r^{\nu(n,y)-k}$ varies by at most a factor

$$1 + O(1/\sqrt{\log_2 y})$$

when $r - k/\log_2 y \ll 1/\log_2 y$.

We note in passing that (3·2) also follows from the weighted version of the Turán-Kubilius inequality proved in [2]. However, verifying the hypotheses then turns out to be more complicated.

4. Proof of Theorem 1.3

Using the classical notation $\Phi(x,y)$ for the number of uncancelled integers in the sieve by prime factors $\leqslant y$ and letting $P^+(n)$ denote the largest prime factor of an integer n with the convention that $P^+(1) = 1$, we have

$$S_r(x,y) = T_r(x,y) + U_r(x,y) - U_r(x/y,y), \tag{4·1}$$

with

$$T_r(x,y) := \sum_{\substack{m\leqslant x/y\\ P^+(m)\leqslant y}} r^{\nu(m)}\Phi\Big(\frac{x}{m},y\Big), \qquad U_r(x,y) := \sum_{\substack{m\leqslant x\\ P^+(m)\leqslant y}} r^{\nu(m)}.$$

By a result in [9], writing $u := (\log x)/\log y$, we have

$$U_r(x,y) = \Big\{1 + O\Big(\frac{\log(u+1)}{\log y}\Big)\Big\}F(r)\varrho_r(u)x(\log y)^{r-1} \tag{4·2}$$

in (H_ε), with notation (1·2).

In order to evaluate $T_r(x,y)$, we use the simple formula—see [6], corollary III.6.20—, valid in (H_ε),

$$\Phi(x,y) = \mathrm{e}^\gamma \frac{x\omega(u) - y}{\zeta(1,y)} + O\Big(\frac{x\varrho(u)}{(\log y)^2}\Big), \tag{4·3}$$

where ω is Buchstab's function, γ is Euler's constant, and $\zeta(s,y) := \prod_{p\leqslant y}(1-1/p^s)^{-1}$. This yields

$$T_r(x,y) = \frac{\mathrm{e}^\gamma x}{\zeta(1,y)}A_r(x,y) - \frac{\mathrm{e}^\gamma y}{\zeta(1,y)}U_r(x/y,y) + O\Big(\frac{x}{(\log y)^2}B_r(x,y)\Big)$$

where we have put

$$A_r(x,y) := \sum_{\substack{m\leqslant x/y\\ P^+(m)\leqslant y}} \frac{\omega(u-u_m)r^{\nu(m)}}{m}, \quad B_r(x,y) := \sum_{\substack{m\leqslant x/y\\ P^+(m)\leqslant y}} \frac{\varrho(u-u_m)r^{\nu(m)}}{m},$$

with $u_m := (\log m)/\log y$.

Appealing to (4·2), partial summation yields

$$\begin{aligned} A_r(x,y) &= \int_{0-}^{u-1} \frac{\omega(u-v)}{y^v}\,\mathrm{d}U_r(y^v,y)\\ &= \frac{U_r(x/y,y)}{x/y} + \int_0^{u-1} U_r(y^v,y)\Big\{\frac{\omega'(u-v)+\omega(u-v)\log y}{y^v}\Big\}\,\mathrm{d}v\\ &= F(r)(\log y)^r \int_0^{u-1} \varrho_r(v)\omega(u-v)\,\mathrm{d}v + O\Big((\log y)^{r-1}\Big) \end{aligned}$$

while we have trivially

$$B_r(x,y) \ll (\log y)^r$$

Assembling our estimates, we get (1·7).

We deduce (1·8) from (1·5) and (1·7) in (H_ε), and from (1·3) and (1·6) in the complementary domain.

5. Proof of Theorem 1.4

Put $\nu(n) := \nu(n,n)$. Note that, parallel to (4·1), we have the decomposition

$$N_k(x,y) = \sum_{\substack{m\leqslant x/y\\ \nu(m)=k\\ P^+(m)\leqslant y}} \Phi\Big(\frac{x}{m},y\Big) + \Theta_k(x,y), \tag{5·1}$$

with

$$\Theta_k(x,y) := \sum_{\substack{x/y<m\leqslant x\\ \nu(m)=k\\ P^+(m)\leqslant y}} 1.$$

In view of (1·4), we only have to establish the lower bound. For this, we retain the first term on the right-hand side of (5·1) and appeal to (4·3), using the fact that $\omega(u) \geqslant \frac12$ for $u \geqslant 1$. Writing $z := \frac13 y^c$, so that $x/z \geqslant 3y$, we get

$$N_k(x,y) \gg \frac{x}{\log y} \sum_{\substack{m\leqslant z\\ \nu(m)=k}} \frac{1}{m} \gg \frac{x(\log_2 z)^k}{k!\,\log y} \asymp \frac{x(\log_2 y)^k}{k!\,\log y}. \tag{5·2}$$

6. Proof of Theorem 1.5

We may plainly assume w to be arbitrarily large.

The case $k=1$ may be dealt with directly on observing that n is counted by $N_1(x,y)$ if, and only if, n is either a prime power p^j with $p\leqslant y$ or is of the form $p^j q$ where p and q are primes and $p^j \leqslant x/y$, $y<q\leqslant x/p^j$. We omit the details and assume $k\geqslant 2$ in the sequel.

First consider the case $k \ll (\log_2 y)/\log w$, so that (1·9) amounts to (1·1). In view of (1·4), we thus only have to show the lower bound. However, in the indicated range, this is provided by (5·2) with now $z := \frac13 y^{1/w}$ so that $x/z \geqslant 3y$ and $\log_2 z > \log_2 y - \log w + \log_2 2 \gg k$.

Next, we embark on proving (1·9) for the larger values of k. From (5·1), we have

$$N_k(x,y) \leqslant R_k(x,y) + \pi_k(x)$$

with

$$\pi_k(x) \asymp \frac{x}{\log y}\,\frac{(\log_2 y)^{k-1}}{(k-1)!}, \qquad R_k(x,y) \asymp \sum_{\substack{m\leqslant x/y\\ \nu(m)=k}} \frac{x}{m\log y} \ll \frac{x(\log_2 3x/y + c_1)^k}{k!\,\log y}.$$

Here and in the remainder of this proof, c_j $(j \geqslant 1)$ denotes an absolute constant. The expected upper bound follows on noticing that, on the right-hand side of (1·9), the order of magnitude of the second term can only exceed that of the first if $k \ll \log_2 3x/y$.

We establish the lower bound in three steps. First assume $k \geqslant 1 + (\log_2 w)(\log_2 y)/\log w$, so that, as stated in (1·10), the right-hand side of (1·9) is $\asymp \pi_k(x)$. Observe that, if an integer $n \leqslant x$ satisfies $\nu(n) = k$, then either n is counted by $N_k(x, y)$ or n is divisible by a prime $p \in]y, x]$. Therefore

$$\frac{x(\log_2 y)^{k-1}}{(k-1)!\log y} \asymp \pi_k(x) \leqslant N_k(x, y) + U_k(x, y),$$

with

$$\begin{aligned} U_k(x, y) &:= \sum_{\substack{y<p\leqslant x\\ mp\leqslant x\\ \nu(m)=k-1}} 1 \ll \sum_{\substack{m\leqslant x/y\\ \nu(m)=k-1}} \frac{x}{m\log y} \\ &\ll \frac{x(\log_2 3x/y + c_2)^{k-1}}{(k-1)!\log y} \ll \frac{x(\log_2 y - \log w + c_3)^{k-1}}{(k-1)!\log y}. \end{aligned} \tag{6·1}$$

By our assumption on k (and actually $k \geqslant 1 + C(\log_2 y)/\log w$ with a suitable large C suffices here), we thus have $U_k(x, y) \leqslant \frac{1}{2}\pi_k(x)$ and the required lower bound for $N_k(x, y)$ follows.

Next, we prove (1·9) for the range $1 + C(\log_2 y)/\log w < k \leqslant 1 + (\log_2 y)(\log_2 w)/\log w$ and under the extra assumption that $y \leqslant x/\log x$. We exploit this last hypothesis in the form that $k \leqslant (\log_2 w)(\log_2 y)/\log w$ implies $k \ll \log_2 3x/y$: this is obvious if, say, $w \leqslant \sqrt{\log y}$ and otherwise follows from the extra hypothesis since then $\log_2 3x/y \gg \log_3 y$. Consider (5·1). The first term on the right-hand side is plainly

$$\asymp \frac{x(\log_2 3x/y)^k}{k!\log y}\cdot$$

However, as previously noticed, since $\Theta_k(x, y) + U_k(x, y) = \pi_k(x)$, the upper bound (6·1) implies that $\Theta_k(x, y) \asymp \pi_k(x)$ provided C is large enough.

Finally, we consider the case $x/\log x < y \leqslant x$ and $k > 1 + C(\log y)/\log w$ with some sufficiently large constant C. Then the right-hand side of (1·9) is $\asymp \pi_k(x)$ because $k \geqslant 2$. Since $\Theta_k(x, y) \gg \pi_k(x)$, the required lower bound follows.

Acknowledgements. The author wishes to express warm thanks to Krishnaswami Alladi for bringing this problem to his attention and for subsequent interesting discussions on the matter.

References

[1] K. Alladi & T. Molnar, On the local distribution of the number of small prime factors, preprint.
[2] A. Biró & T. Szamuely, A Turán–Kubilius inequality with multiplicative weights, *Acta Math. Hung.* **70** (1996), n$^{\text{os}}$ 1-2, 39–56.
[3] P. Erdős & G. Tenenbaum, Sur les densités de certaines suites d'entiers, *Proc. London Math. Soc.* (3) **59** (1989), 417–438.
[4] R.R. Hall & G. Tenenbaum, *Divisors*, Cambridge tracts in mathematics, no 90, Cambridge University Press (1988).
[5] A. Hildebrand & G. Tenenbaum, On the number of prime factors of an integer, *Duke Math. J.* **56** no 3 (1988), 471–501.
[6] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, 3rd ed., Graduate Studies in Mathematics 163, Amer. Math. Soc. 2015.
[7] G. Tenenbaum, Fonctions multiplicatives, sommes d'exponentielles, et loi des grands nombres, *Indag. Math.* **27** (2016), 590–600.
[8] G. Tenenbaum, Valeurs moyennes effectives de fonctions multiplicatives complexes, *Ramanujan J.* **44**, n$^{\text{o}}$ 3 (2017), 641-701.
[9] G. Tenenbaum & J. Wu, Moyennes de certaines fonctions multiplicatives sur les entiers friables, *J. reine angew. Math.* **564** (2003), 119–166.

Gérald Tenenbaum
Institut Élie Cartan
Université de Lorraine
BP 70239
54506 Vandœuvre Cedex
France
internet: gerald.tenenbaum@univ-lorraine.fr