ON INTEGERS n FOR WHICH $n \nmid P(n)!$

KEVIN FORD

ABSTRACT. Answering a question of Paul Erdős, we determine asymptotically how many positive integers $n \leqslant x$ satisfy $n \nmid P(n)!$, where $P(n)$ is the largest prime factor of n.

1. INTRODUCTION

Let $S(n)$ be the smallest integer k so that $n|k!$. This function was first considered by Lucas in 1883 [**8**] and rediscovered many times since, and sometimes bears the name "the Kempner function" or the "Smarandache function". If $P(n)$ denotes the largest prime factor of n, then usually $S(n) = P(n)$, see sequence A057109 in the On-Line Enclyclopedia of Integer Sequences [**9**]. Let $N(x)$ be the number of $n \leqslant x$ for which $S(n) \neq P(n)$. In this paper we give an asymptotic formula for $N(x)$. Before that, we reveal the history of the problem and attempts to solve it.

In 1991, Paul Erdős [**3**] proposed the problem of showing that $N(x) = o(x)$ to the American Mathematical Monthly. The problem proposal did not contain a solution, however. At the request of Paul Batemen, the Problems Editor of the Monthly at that time, a solution was prepared by the author, who showed that $N(x) \ll xe^{-c\sqrt{\log x}}$ for some positive c. Only one other solution was received, that of I. Katsanas, which was published in the Monthly in February 1994 [**7**].

In 1998, Steven Finch asked the author to provide him with details of his argument (hinted at in [**7**]), and in October 1998 he sent a draft to Finch containing a proof of a full asymptotic formula for $N(x)$. A few months later, the draft appeared in the "Smarandache Notions Journal" (rather bizarrely, the minus signs in all exponents were replaced with blank space). The "paper" was neither submitted to, nor refereed by, this publication. However, it was reviewed by Mathematical Reviews [**4**].

Meanwhile, S. Akbik [**1**] proved the upper bound $N(x) \ll xe^{-(1/4)\sqrt{\log x}}$. In 2003, J.-M. De Koninck and N. Doyon [**2**] claimed the rough bound $N(x) = x\exp\{-(2+o(1))\sqrt{\log x \log_2 x}\}$, which is incorrect (the constant 2 should be $\sqrt{2}$). Two years later, A. Ivić [**6**] published an asymptotic formula, which corrected a mistake in the first author's manuscript and differed by a multiplicative constant. In this article we present the author's 1998 argument, including the correction of Ivić. The author thanks Aleksandar Ivić for helpful conversations.

We begin by introducing the Dickman-de Bruijn function $\rho(u)$, defined recursively by

$$\rho(u) = 1 \quad (0 \leqslant u \leqslant 1), \qquad \rho(u) = 1 - \int_1^u \frac{\rho(v-1)}{v}\, dv \quad (u > 1).$$

For $u > 1$ let $\xi = \xi(u)$ be defined by

$$u = \frac{e^{\xi} - 1}{\xi}.$$

It can be easily shown that $\xi(u)$ is increasing and that

$$\xi(u) = \log u + \log_2 u + O\left(\frac{\log_2 u}{\log u}\right). \tag{1.1}$$

Finally, let $u_0 = u_0(x)$ be defined by the equation

$$\log x = u_0^2 \xi(u_0). \tag{1.2}$$

The function $u_0(x)$ may also be defined directly by

$$\log x = u_0 \left(x^{1/u_0^2} - 1\right).$$

It is straightforward to show that

$$u_0 = \left(\frac{2\log x}{\log_2 x}\right)^{\frac{1}{2}} \left(1 - \frac{\log_3 x}{2\log_2 x} + \frac{\log 2}{2\log_2 x} + O\left(\left(\frac{\log_3 x}{\log_2 x}\right)^2\right)\right). \tag{1.3}$$

We can now state our main result.

Theorem 1. *We have*

$$N(x) \sim 2^{1/4}\pi^{1/2}(\log x \log_2 x)^{3/4} x^{1-1/u_0} \rho(u_0) \qquad (x \to \infty).$$

One cannot write the asymptotic formula in terms of "simple" functions, but there is an easily digestible corollary.

Corollary 2. *We have*

$$N(x) = x \exp\left\{ -\sqrt{2\log x \log_2 x} + O\Big(\Big(\frac{\log x}{\log_2 x}\Big)^{1/2} \log_3 x\Big)\right\}.$$

The asymptotic formula can be made a bit simpler, without reference to the function ρ and written only in terms of u_0.

Corollary 3. *We have*

$$N(x) \sim \frac{e^\gamma}{\sqrt{2}} (\log x)^{\frac{1}{2}} (\log_2 x) x^{1-2/u_0} \exp\left\{\int_0^{\frac{\log x}{u_0^2}} \frac{e^v - 1}{v}\, dv\right\} \qquad (x \to \infty),$$

where $\gamma = 0.5772\ldots$ is the Euler-Mascheroni constant.

2. The distribution of integers without large prime factors

We list standard estimates of the function $\Psi(x,y)$, which denotes the number of integers $n \leqslant x$ with $P(n) \leqslant y$. These may be found in [**5**] or [**10**, Ch. III.5].

Lemma 2.1 (Hildebrand)**.** *For every $\epsilon \in (0, 3/5)$,*

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

uniformly in $1 \leqslant u \leqslant \exp\{(\log y)^{3/5-\epsilon}\}$.

Lemma 2.2. *For $u \geqslant 1$,*

$$\begin{aligned}\rho(u) &= \left(1 + O\left(\frac{1}{u}\right)\right)\sqrt{\frac{\xi'(u)}{2\pi}}\exp\left\{\gamma - \int_1^u \xi(t)\,dt\right\} \\ &= \exp\left\{-u\left(\log u + \log_2 u - 1 + O\left(\frac{\log_2 u}{\log u}\right)\right)\right\}.\end{aligned}$$

Lemma 2.3. *If $u > 2$, $|v| \leqslant u/2$, then*

$$\rho(u-v) = \rho(u)\exp\{v\xi(u) + O((1+v^2)/u)\}.$$

3. THE PROOF OF THEOREM 1 AND COROLLARIES

We begin with a characterization of most of the n counted by $N(x)$. For brevity write

$$L(x) = \exp\left\{\sqrt{\log x \log_2 x}\right\}, \qquad E(x) = \frac{\log_3 x}{\log_2 x}.$$

Roughly speaking, we are aiming to show that $N(x) = xL(x)^{-\sqrt{2}+O(E(x))}$. The next Proposition shows that most of the integers counted by $N(x)$ have $P(n) \approx L(x)^{1/\sqrt{2}}$, and also n is divisible by the square of a prime very close to $P(n)$.

Proposition 1. *Define*

$$Y = L(x)^{1/\sqrt{2}-2E(x)^{1/3}}, \qquad Z = L(x)^{1/\sqrt{2}+2E(x)^{1/3}}.$$

Let $T(x)$ be the number of $n \leqslant x$ with $Y \leqslant P(n) \leqslant Z$ and such that there is a prime $q > P(n)/2$ with $q^2 \| n$. Then

$$N(x) = T(x) + O(xL(x)^{-\sqrt{2}-E(x)^{2/3}}).$$

Proof. Let

$$U = L(x)^{1/4}, \qquad V = L(x)^2.$$

Let N_1 be the number of n counted by $N(x)$ with $P(n) \leqslant U$, let N_2 be the number of n with $P(n) \geqslant V$, and let $N_3 = N(x) - N_1 - N_2$. By Lemmas 2.1 and 2.2,

$$N_1 \leqslant \Psi(x,U) = xL(x)^{-2+o(1)}.$$

Now consider $n \leqslant x$ with $P(n) > U$ and $n \nmid P(n)!$. Denoting $p = P(n)$, we have either $p^2|n$ or for some prime $q < p$ and $b \geqslant 2$ we have $q^b \| n$ and $q^b \nmid p!$. Since $p!$ is divisible by $q^{\lfloor p/q \rfloor}$ and $b \leqslant 2\log x$, it follows that $q > p/b > p^{1/2}$. In all cases n is divisible by the square of a prime $\geqslant p/(2\log x)$ and therefore

$$N_2 \leqslant \sum_{p \geqslant \frac{V}{2\log x}} \frac{x}{p^2} \ll \frac{x\log x}{V} \ll xL(x)^{-2+o(1)}.$$

Since $q > p^{1/2}$ it follows that $q^{\lfloor p/q \rfloor} \| p!$. If n is counted by N_3, there is a number $b \geqslant 2$ and prime $q \in (p/b, p]$ so that $q^b \| n$. For each $b \geqslant 2$, let $N_{3,b}$ be the number of n counted in N_3 such that $q^b \parallel n$ for some prime $q > p/b$. We have

$$\sum_{b \geqslant 9} N_{3,b} \ll \sum_{9 \leqslant b \leqslant 2\log x} \sum_{q \geqslant U/b} \frac{x}{q^b} \ll \frac{x}{U^8} = xL(x)^{-2}.$$

Next, let $N_{3,b}(S)$ denote the number of integers counted in $N_{3,b}$ for which $p = P(n) \in S$. By Lemma 2.1 and the fact that ρ is decreasing, for $2 \leqslant b \leqslant 8$ and $S \subset (U, V)$,

$$\begin{aligned} N_{3,b}(S) &\leqslant \sum_{p \in S} \left(\Psi\left(\frac{x}{p^b}, p\right) + \sum_{p/b \leqslant q < p} \Psi\left(\frac{x}{pq^b}, p\right) \right) \\ &\ll x \sum_{p \in S} \left(\frac{1}{p^b} \rho\left(\frac{\log x}{\log p} - b\right) + \sum_{p/b \leqslant q < p} \frac{1}{pq^b} \rho\left(\frac{\log x - \log p - b \log q}{\log p}\right) \right) \\ &\ll x \sum_{p \in S} p^{-b} \rho\left(\frac{\log x}{\log p} - (b+1)\right) \\ &\ll x \sum_{m \in S} m^{-b} \exp\left\{ -\frac{\log x}{2 \log m} \log_2 x + O\left(\sqrt{\frac{\log x}{\log_2 x}} \log_3 x\right) \right\}, \end{aligned}$$

using Lemma 2.2 in the last step and replacing the sum over primes in S with a sum over all integers in S.

Now let $S = [L(x)^\alpha, L(x)^\beta]$, where $\frac{1}{4} \leqslant \alpha \leqslant \beta \leqslant 2$. Uniformly in α, β we have

$$\begin{aligned} N_{3,b}(S) &\ll x L(x)^{O(E(x))} \int_{L(x)^\alpha}^{L(x)^\beta} \frac{\exp\left\{-\frac{\log x \log_2 x}{2 \log t}\right\}}{t^b} \, dt \\ &= x L(x)^{O(E(x))} \int_\alpha^\beta L(x)^{-\frac{1}{2w} - w(b-1)} \, dw. \end{aligned} \tag{3.1}$$

Since the minimum of $\frac{1}{2w} + w(b-1)$ is $\sqrt{2(b-1)}$, occurring at $w = 1/\sqrt{2(b-1)}$,

$$N_{3,b} \ll x L(x)^{-\sqrt{2(b-1)} + o(1)} \ll x L(x)^{-2+o(1)} \qquad (3 \leqslant b \leqslant 8).$$

Combining this bound with our earlier bounds for N_1 and N_2 we conclude that

$$N(x) = N_{3,2} + O(x L(x)^{-2+o(1)}). \tag{3.2}$$

Now if $w \leqslant 1/\sqrt{2} - 2E(x)^{1/3}$ or $w \geqslant 1/\sqrt{2} + 2E(x)^{1/3}$ then $\frac{1}{2w} + w \geqslant \sqrt{2} + E(x)^{2/3}$. Using (3.1) we then have

$$N_{3,b}([U, Y) \cup (Z, V]) \ll x L(x)^{-\sqrt{2} - E(x)^{2/3}}.$$

Thus, by (3.2) we have

$$N(x) = N_{3,2}([Y, Z]) + O(x L(x)^{-\sqrt{2} - E(x)^{2/3}}). \tag{3.3}$$

Finally, if there is a prime $q > P(n)/2$ for which $q^2 \| n$, then $n \nmid P(n)!$ since $q \| P(n)!$. Thus, $N_{3,2}([Y, Z]) = T(x)$ and the proof is complete. □

The function $T(x)$ counts two types of integers, (i) those with $P(n)^2 \| n$ and (ii) those with $q^2 \| n$ for some $q \in (P(n)/2, P(n))$. Both papers [**4**] and [**6**] correctly estimate the count of integers in (i), but both contain mistakes in the count of integers of type (ii). In fact, as we shall show, the count of integers of type (ii) is asymptotically half the size of the count of integers of type (i)

Proof of Theorem 1. Again let $p = P(n)$. For n counted by $T(x)$, either $p^2 \| n$ or $q^2 \| n$ for some prime $q \in (p/2, p)$. In the latter case, $P(n/pq^2) < p$ and also $q \nmid n/pq^2$. Thus,

$$\begin{aligned} T(x) &= \sum_{Y<p<Z} \left(\Psi\left(\frac{x}{p^2}, p-1\right) + \sum_{\frac{p}{2} \leqslant q < p} \Psi\left(\frac{x}{pq^2}, p-1\right) + O\Big(\Psi\Big(\frac{x}{pq^3}, p\Big)\Big)\right) \\ &= S_1 + O(S_2), \end{aligned}$$

say. The sum S_2 is part of the expression used in the estimate for $N_{3,3}$ in Proposition 1 and hence

$$S_2 \ll xL(x)^{-2+o(1)}. \tag{3.4}$$

Using Lemma 2.1 we obtain

$$S_1 \sim x \sum_{Y<p<Z} \left(\frac{\rho\left(\frac{\log x}{\log p} - 2\right)}{p^2} + \sum_{p/2 \leqslant q < p} \frac{\rho\left(\frac{\log x - 2\log q}{\log p} - 1\right)}{pq^2} \right).$$

We apply Lemma 2.3 with $u = \frac{\log x}{\log p} - 3$ and $v = 2\frac{\log(q/p)}{\log p}$, noting that by (1.1), we have

$$v\xi(u) \ll \sqrt{\frac{\log_2 x}{\log x}}.$$

It follows that

$$\rho\left(\frac{\log x - 2\log q}{\log p} - 1\right) \sim \rho\left(\frac{\log x}{\log p} - 3\right).$$

Thus, by the prime number theorem,

$$S_1 \sim x \sum_{Y<p<Z} \frac{1}{p^2} \left(\rho\left(\frac{\log x}{\log p} - 2\right) + \frac{1}{\log p} \rho\left(\frac{\log x}{\log p} - 3\right)\right).$$

By partial summation, the Prime Number Theorem with error term, and the change of variable $u = \log x/\log p$,

$$S_1 \sim x \int_{u_1}^{u_2} \left(\frac{\rho(u-2)}{u} + \frac{\rho(u-3)}{\log x}\right) x^{-1/u} du,$$

where

$$u_1 = \frac{\log x}{\log Z}, \qquad u_2 = \frac{\log x}{\log Y}.$$

By the definition of Y and Z, we have $u \sim u_0$ for $u_1 \leqslant u \leqslant u_2$. Hence, by Lemma 2.3,

$$\rho(u-2) \sim \rho(u)(u_0 \log u_0)^2, \qquad \rho(u-3) \sim \rho(u)(u_0 \log u_0)^3.$$

Thus, recalling (1.3), we obtain

$$S_1 \sim \frac{\sqrt{2}}{2} x (\log x)^{1/2} (\log_2 x)^{3/2} I, \quad I := \int_{u_1}^{u_2} \rho(u) x^{-1/u}\, du.$$

Combine this with our bound (3.4) on S_2 and Proposition 1 and we see that

$$N(x) = (1 + o(1)) \frac{\sqrt{2}}{2} x (\log x)^{1/2} (\log_2 x)^{3/2} I + O(xL(x)^{-\sqrt{2} - c(\log_2 x)^{-2/3}}). \tag{3.5}$$

Entending the range of integration to $[1,\infty)$ has a negligible effect on I by the above analysis, and after the change of variable $u=\frac{\log x}{\log t}$ we arrive at [**6**, Theorem 2]:

$$N(x)\sim 2x\int_2^x \rho\Big(\frac{\log x}{\log t}\Big)\frac{\log t}{t^2}\,dt. \tag{3.6}$$

To obtain Theorem 1, we need to analyze the integral I more carefully. Letting $u=u_0-v$, we will show that most of the contribution comes from very small v. In fact, u_0 was chosen in order to maximize $x^{-1/u}\rho(u)$. Using Lemma 2.3 and the relation $\xi(u_0)=\frac{\log x}{u_0^2}$, we see that the integrand is

$$\begin{aligned}\rho(u)x^{-1/u}&=\rho(u_0)x^{-1/u_0}\exp\left\{\frac{v\log x}{u_0^2}+\frac{\log x}{u_0}-\frac{\log x}{u_0-v}+O\Big(\frac{v^2+1}{u_0}\Big)\right\}\\&=\rho(u_0)x^{-1/u_0}\exp\left\{-\frac{v^2\log x}{u_0^2(u_0-v)}+O\Big(\frac{v^2+1}{u_0}\Big)\right\}.\end{aligned}$$

As $\frac{\log x}{u_0^2}\sim\frac{\log_2 x}{2}$, the main term in the exponential has larger order than the error term. Let I_1 be the part of the integral corresponding to $|v|\geqslant v_0:=(\log x)^{1/4}(\log_2 x)^{-1/2}$, and let $I_2=I-I_1$. Then

$$\begin{aligned}I_1&\ll\rho(u_0)x^{-1/u_0}\int_{v_0}^\infty\exp\left\{-0.3\frac{v^2(\log_2 x)^{3/2}}{(\log x)^{1/2}}\right\}dv\\&=\rho(u_0)x^{-1/u_0}(\log x)^{1/4}(\log_2 x)^{-3/4}\int_{(\log_2 x)^{1/4}}^\infty e^{-0.3w^2}\,dw\\&=o\Big(\rho(u_0)x^{-1/u_0}(\log x)^{1/4}(\log_2 x)^{-3/4}\Big)\end{aligned}$$

as $x\to\infty$. If $|v|\leqslant v_0$ then $\frac{v^2+1}{u_0}=o(1)$ and $\frac{v^2\log x}{u_0^2(u_0-v)}=\frac{v^2\log x}{u_0^3}+o(1)$. Hence

$$\begin{aligned}I_2&\sim\rho(u_0)x^{-1/u_0}\int_{-v_0}^{v_0}\exp\left\{-\frac{v^2\log x}{u_0^3}\right\}dv\\&=\frac{\rho(u_0)x^{-1/u_0}u_0^{3/2}}{(\log x)^{1/2}}\int_{|w|\leqslant v_0(\log x)^{1/2}u_0^{-3/2}}e^{-w^2}\,dw\\&\sim\frac{\rho(u_0)x^{-1/u_0}u_0^{3/2}}{(\log x)^{1/2}}\int_{-\infty}^\infty e^{-w^2}\,dw\\&\sim\frac{2^{3/4}\pi^{1/2}(\log x)^{1/4}}{(\log_2 x)^{3/4}}\rho(u_0)x^{-1/u_0},\end{aligned}$$

and it follows that

$$I\sim\frac{2^{3/4}\pi^{1/2}(\log x)^{1/4}}{(\log_2 x)^{3/4}}\rho(u_0)x^{-1/u_0}.$$

We insert this last bound into (3.5), and this provides the desired asymptotic for Theorem 1, provided that the error term is of smaller order than the main term. We use (1.3) and Lemma 2.2, together with the computation

$$\frac{\log x}{u_0}+u_0(\log u_0+\log_2 u_0)=\sqrt{2\log x\log_2 x}\Big(1+O(E(x))\Big)$$

and we see that

$$\rho(u_0)x^{-1/u_0} = L(x)^{-\sqrt{2}+O(E(x))}.$$

This completes the proof of Theorem 1, and gives Corollary 2 as well. □

Proof of Corollary 3. First observe that $\xi'(u) \sim u^{-1}$ and next use Lemma 2.2 to write

$$\rho(u_0) \sim \frac{e^\gamma}{\sqrt{2\pi u_0}} \exp\left\{-\int_1^{u_0} \xi(t)\,dt\right\}.$$

By the definitions of ξ and u_0, the change of variables $\xi(t) = v$, and the derivative formula

$$\xi(t)\,dt = v\,dt = e^v - \frac{e^v - 1}{v}\,dv,$$

we then obtain

$$\int_1^{u_0} \xi(t)\,dt = \int_0^{\xi(u_0)} e^v - \frac{e^v-1}{v}\,dv$$
$$= e^{\xi(u_0)} - 1 - \int_0^{\xi(u_0)} \frac{e^v-1}{v}\,dv.$$

Corollary 3 now follows from (1.2). In particular, $e^{\xi(u_0)} - 1 = u_0\xi(u_0) = \frac{\log x}{u_0}$.

References

[1] S. Akbik, On a density problem of Erdős, Int. J. Math. Sci **22** (1999), No. 3, 655–658.

[2] J.-M. De Koninck and N. Doyon, On a thin set of integers involving the largest prime factor function, Int. J. Math. Math. Sci. **19** (2003), 1185–1192.

[3] P. Erdős, Problem 6674, Amer. Math. Monthly **98** (1991), 965.

[4] K. Ford, The normal behavior of the Smarandache function, Smarandache Notions J. **10** (1999), 81–86. Unpublished. MR1682452

[5] A. Hildebrand and G. Tenenbaum, Integers without large prime factors, J. Théor. Nombres Bordeaux **5** (1993), 411–484.

[6] A. Ivić, A problem of Erdős involving the largest prime factor of n, Monatsh. Math. **145** (2005), 35–46.

[7] I. Katsanas, Solution to Problem 6674, Amer. Math. Monthly **101** (1994), 179.

[8] E. Lucas, Question Nr. 288. Mathesis **3** (1883), 232.

[9] N. A. Sloane, *The on-line encyclopedia of integer sequences.* `http://oeis.org`

[10] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory, 3rd ed.*, Amer. Math Soc., 2015.