#A81 INTEGERS 13 (2013)

THE AVERAGE LARGEST PRIME FACTOR

Eric Naslund
Department of Mathematics, Princeton University, Princeton, New Jersey
naslund@math.princeton.edu

Received: 2/18/13, Revised: 7/17/13, Accepted:12/5/13 , Published: 12/13/13

Abstract

Let $P(n)$ denote the largest prime factor of an integer n. In this paper we look at the average of $P(n)$, and prove that

$$\frac{1}{x}\sum_{n\leq x} P(n) = \mathrm{li}_g(x) + O_\epsilon\left(xe^{-c(\log x)^{3/5-\epsilon}}\right),$$

where $\mathrm{li}_g(x) = \int_2^x \frac{t}{x}\frac{[x/t]}{\log t}dt$. This improves a result of De Koninck and Ivíc and allows us to deduce their asymptotic expansion

$$\frac{1}{x}\sum_{n\leq x} P(n) = c_0\frac{x}{\log x} + c_1\frac{1!x}{(\log x)^2} + \cdots + c_{k-1}\frac{(k-1)!x}{(\log x)^k} + O\left(\frac{x}{(\log x)^{k+1}}\right)$$

as a corollary, with the advantage that we can give the constants explicitly as

$$c_n = \frac{1}{2^{n+1}}\sum_{j=0}^{n}\frac{2^j(-1)^j\zeta^{(j)}(2)}{j!}.$$

1. Introduction

Let $P(n)$ denote the largest prime factor of an integer n. The moments of $P(n)$ were first looked at in 1976 by Knuth and Pardon [4], and concerning the average, they proved that

$$\sum_{n\leq x} P(n) \sim \frac{\pi^2}{12}\frac{x^2}{\log x}. \tag{1}$$

The above asymptotic also appeared in Erdős and Alladi's 1977 paper [1] on additive arithmetic functions. Later, De Koninck and Ivić [2] proved that $\sum_{n\leq x} P(n)$ has an asymptotic expansion of the form

$$\sum_{n\leq x} P(n) = x^2\left(\frac{d_1}{\log x} + \frac{d_2}{\log^2 x} + \cdots + \frac{d_m}{\log^m x} + O\left(\frac{1}{\log^{m+1} x}\right)\right)$$

where the constants d_i are computable, but not given explicitely. This expansion appears again in [3], where Ivić finds a similar formula for the k^{th} largest prime factor. In this paper, we calculate the average of $P(n)$ up to an error of the form $O_\epsilon\left(xe^{-c(\log x)^{3/5-\epsilon}}\right)$, like that of the prime number theorem. This allows us to deduce De Koninck and Ivić's expansion as a corollary, as well as give the constants d_i explicitely. Our main result is:

Theorem 1. *Letting $li_g(x) = \int_2^x \frac{t}{x}\frac{[x/t]}{\log t}dt$, we have that*

$$\frac{1}{x}\sum_{n\leq x} P(n) = li_g(x) + O_\epsilon\left(xe^{-c(\log x)^{3/5-\epsilon}}\right). \tag{2}$$

For any integer k, $li_g(x)$ has the asymptotic expansion

$$li_g(x) = c_0\frac{x}{\log x} + c_1\frac{1!x}{(\log x)^2} + \cdots + c_{k-1}\frac{(k-1)!x}{(\log x)^k} + O\left(\frac{x}{(\log x)^{k+1}}\right) \tag{3}$$

where

$$c_n = \frac{1}{2^{n+1}}\sum_{j=0}^{n}\frac{2^j(-1)^j\zeta^{(j)}(2)}{j!}. \tag{4}$$

These constants c_n satisfy $c_n = 1 + O\left(\frac{n}{2^n}\right)$.

Notice in particular that $c_0 = \frac{\pi^2}{12}$, so we deduce that

$$\sum_{n\leq x} P(n) = \frac{\pi^2}{12}\frac{x^2}{\log x} + O\left(\frac{x^2}{(\log x)^2}\right).$$

Although this expansion has previously been presented in [2], and again in [3], there are some advantages to the result above. We are able to give the constants c_n explicitely in terms of the zeta function without much additional work, and our proof is basic requiring only an application of the hyperbola method. Furthermore, we give the explicit integral form which has an error term like that of the prime number theorem, paralleling how $\pi(x)$ is approximated by the function $\mathrm{li}(x)$.

2. The Main Theorem

For each integer $n \leq x$, there is at most one prime $p > \sqrt{x}$ such that $p|n$, and for each prime p there will be exactly $\left[\frac{x}{p}\right]$ integers less than x which are divisible by p. Combining these two facts, we see that

$$\sum_{\sqrt{x}<p\leq x} p\left[\frac{x}{p}\right] \leq \sum_{n\leq x} P(n) \leq \sum_{p\leq x} p\left[\frac{x}{p}\right],$$

and since $\sum_{p\leq\sqrt{x}} p\left[\frac{x}{p}\right] \leq x\sum_{p\leq\sqrt{x}} 1 = O\left(\frac{x^{3/2}}{\log x}\right)$, it follows that

$$\sum_{n\leq x} P(n) = \sum_{p\leq x} p\left[\frac{x}{p}\right] + O\left(\frac{x^{3/2}}{\log x}\right). \tag{5}$$

By writing $\left[\frac{x}{p}\right] = \sum_{n\leq x:\ p|n} 1$, and rearranging the order of summation, we obtain

$$\sum_{p\leq x} p\left[\frac{x}{p}\right] = \sum_{p\leq x} p \sum_{n\leq x:\ p|n} 1 = \sum_{n\leq x}\sum_{p\leq \frac{x}{n}} p. \tag{6}$$

Letting $\mathcal{K}(x) = \sum_{p\leq x} p$, we have the estimate

$$\mathcal{K}(x) = \int_2^x \frac{t}{\log t}dt + O_\epsilon\left(x^2 e^{-c(\log x)^{3/5-\epsilon}}\right), \tag{7}$$

which follows from the prime number theorem (See [5]) along with partial summation. Using (7) along with equations (5), and (6), we have that

$$\begin{aligned}
\sum_{n\leq x} P(n) &= \sum_{n\leq x}\int_2^{\frac{x}{n}} \frac{t}{\log t}dt + \sum_{n\leq x} O_\epsilon\left(\frac{x^2}{n^2}e^{-c\left(\log\frac{x}{n}\right)^{3/5-\epsilon}}\right) + O\left(\frac{x^{\frac{3}{2}}}{\log x}\right)\\
&= \int_2^x \frac{t\left[\frac{x}{t}\right]}{\log t}dt + O_\epsilon\left(x^2 e^{-c(\log x)^{3/5-\epsilon}}\right) \qquad (8)\\
&= x\mathrm{li}_g(x) + O_\epsilon\left(x^2 e^{-c(\log x)^{3/5-\epsilon}}\right), \qquad (9)
\end{aligned}$$

which proves (2). To recover the asymptotic expansion in (3), we turn our attention to this integral function and make the substitution $t = \frac{x}{u}$ to obtain

$$\mathrm{li}_g(x) = \frac{1}{x}\int_2^x \frac{t\left[\frac{x}{t}\right]}{\log t}dt = x\int_1^{x/2} \frac{[u]}{u^3\log\left(\frac{x}{u}\right)}du. \tag{10}$$

We can rewrite the integral above as

$$\int_1^{x/2} \frac{[u]}{u^3\log\left(\frac{x}{u}\right)}du = \frac{1}{\log x}\int_1^{x/2}\frac{[u]}{u^3}\left(1-\frac{\log u}{\log x}\right)^{-1}du,$$

and for any integer $k\geq 1$, by the geometric series expansion

$$\left(1-\frac{\log u}{\log x}\right)^{-1} = 1+\frac{\log u}{\log x}+\cdots+\left(\frac{\log u}{\log x}\right)^{k-1}+\left(\frac{\log u}{\log x}\right)^k\left(1-\frac{\log u}{\log x}\right)^{-1},$$

we have that

$$\begin{aligned}
\int_1^{x/2}\frac{[u]}{u^3\log\left(\frac{x}{u}\right)}du &= \sum_{n=0}^{k-1}\frac{1}{(\log x)^{n+1}}\int_1^{x/2}\frac{[u]}{u^3}(\log u)^n du\\
&\quad+\frac{1}{(\log x)^{k+1}}\int_1^{x/2}\frac{[u]}{u^3}(\log u)^k\left(1-\frac{\log u}{\log x}\right)^{-1}du.
\end{aligned}$$

Each of these integrals is absolutely convergent on $[1,\infty)$, and so the last term contributes an error term of at most $O_k\left(\frac{1}{\log^{k+1} x}\right)$. Setting $c_k = \frac{1}{k!}\int_1^\infty \frac{[u]}{u^3}\left(\log u\right)^k du$, and applying the bound

$$\int_{x/2}^{\infty} \frac{[u]}{u^3}\left(\log u\right)^n du \leq \int_{x/2}^{\infty} \frac{\left(\log u\right)^n}{u^2} du = O\left(\frac{\left(\log x\right)^n}{x}\right),$$

it follows that

$$\mathrm{li}_g(x) = \frac{x}{\log x}\sum_{n=0}^{k-1}\frac{n! c_n}{\left(\log x\right)^n} + O_k\left(\frac{x}{\log^{k+1} x}\right). \tag{11}$$

To evaluate these constants c_n, we use exponential generating series. Integration by parts tells us that for $s>1$

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \int_1^\infty x^{-s} d\,[x] = s\int_1^\infty [x]\, x^{-s-1}dx,$$

and so

$$\sum_{k=0}^{\infty} c_k z^k = \int_1^\infty [x]\, x^{z-3}dx = \frac{\zeta(2-z)}{2-z}.$$

Multiplying the power series expansions for $\zeta(2-z)$ and $\frac{1}{2-z}$ yields

$$\left(\sum_{j=0}^{\infty}(-1)^j \zeta^{(j)}\left(2\right)\frac{z^j}{j!}\right)\left(\frac{1}{2}\sum_{k=0}^{\infty}\frac{z^k}{2^k}\right) = \frac{1}{2}\sum_{n=0}^{\infty} z^n\left(\sum_{k+j=n}(-1)^j\frac{\zeta^{(j)}\left(2\right)}{j!}\frac{1}{2^k}\right),$$

and hence

$$c_n = \frac{1}{2^{n+1}}\sum_{j=0}^{n}\frac{2^j(-1)^j\zeta^{(j)}\left(2\right)}{j!}.$$

To find the size of c_n we note that

$$\begin{aligned} \frac{(-1)^j\zeta^{(j)}\left(2\right)}{j!} &= \frac{1}{j!}\sum_{k=1}^{\infty}\frac{\left(\log k\right)^j}{k^2} \\ &= \frac{1}{j!}\int_1^\infty \frac{\left(\log x\right)^j}{x^2}dx + \frac{1}{j!}\int_1^\infty \frac{\left(j-2\log x\right)\log^{j-1} x}{x^3}\left\{x\right\}dx. \end{aligned}$$

Substituting $x=e^u$, the first integral becomes $\Gamma(j+1)$, and the second is $O\left(\frac{j!}{2^j}\right)$, which implies that

$$\frac{(-1)^j}{j!}\zeta^{(j)}\left(2\right) - 1 = O\left(\frac{1}{2^j}\right),$$

and so $c_n = 1 + O\left(\frac{n}{2^n}\right)$. This establishes all of Theorem 1.

Acknowledgment I am grateful to Greg Martin for his suggestions and encouragement.

References

[1] K. Alladi and P. Erdős. On an additive arithmetic function. *Pacific J. Math.* 71(2):275–294, 1977.

[2] Jean-Marie De Koninck and Aleksandar Ivić. The distribution of the average prime divisor of an integer. *Arch. Math. (Basel)* **43(1)** (1984), 37–43.

[3] Aleksandar Ivić. On the kth prime factor of an integer. *Zb. Rad. Prirod.-Mat. Fak. Ser. Mat.* **20(1)** (1990), 63–73.

[4] Donald E. Knuth and Luis Trabb Pardo. Analysis of a simple factorization algorithm. *Theoret. Comput. Sci.* **3(3)** (1976/77), 321–348.

[5] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.