NOTES ON ERDŐS PROBLEM #451: DENSITY ESTIMATES AND HEURISTIC EVIDENCE

QUANYU TANG

Abstract. We consider Erdős Problem #451 from the Erdős Problems website: estimate n_k, the smallest integer $> 2k$ such that

$$\prod_{1\le i\le k}(n_k - i)$$

has no prime factor in the interval $(k, 2k)$. For this problem it is natural to introduce the quantity

$$D_k := \prod_{\substack{k<p<2k \\ p \text{ prime}}} \left(1 - \frac{k}{p}\right),$$

which represents the density of admissible integers n (those for which the block $[n-k, n-1]$ contains no multiple of a prime p with $k < p < 2k$). We establish the asymptotic behavior

$$\log \frac{1}{D_k} = (\log 4 + o(1))\,\frac{k}{\log k},$$

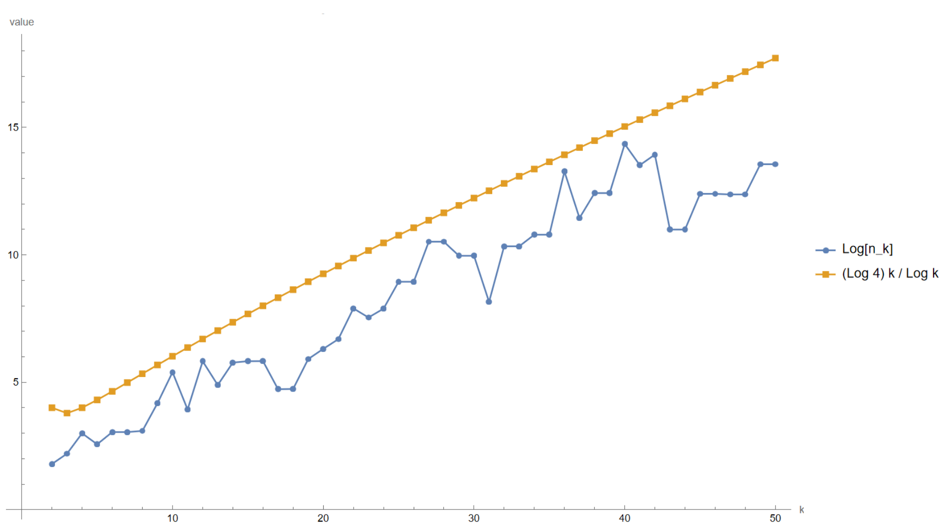
and conjecture that this density asymptotic also governs the growth of the minimal solution n_k.

1. Introduction and Motivation

On Bloom's Erdős Problems website, Problem #451 asks:

Problem 1.1. *Estimate n_k, the smallest integer $> 2k$ such that $\prod_{1\le i\le k}(n_k - i)$ has no prime factor in the open interval $(k, 2k)$.*

Numerical experiments for $1 \le k \le 50$ show that $\log n_k$ grows roughly like $k/\log k$. In particular, the plot of $\log n_k$ against $(\log 4)\,k/\log k$ exhibits very similar growth. Moreover, if we replace $\log 4$ by $\log 3$, then the two curves intersect:



These data suggest the asymptotic order

Conjecture 1.2. $\log n_k = \Theta\left(\frac{k}{\log k}\right)$,

and more precisely make it natural to conjecture the upper bound

Conjecture 1.3. $\log n_k \;\le\; (\log 4 + o(1))\,\frac{k}{\log k}$.

Why does the constant $\log 4$ appear? The point is that one can pass from the "no prime factor in $(k, 2k)$" constraint to a clean density calculation. Let

$$\mathcal{P}_k \;:=\; \{\, p : p \text{ is prime and } k < p < 2k \,\}, \qquad M_k \;:=\; \prod_{\substack{k<p<2k \\ p \text{ prime}}} p.$$

Requiring that the whole block $[n-k, n-1]$ contain no multiple of any $p \in \mathcal{P}_k$ is equivalent to requiring that the k integers $n-1, n-2, \ldots, n-k$ are all coprime to M_k. Fix a prime $p \in \mathcal{P}_k$. Modulo p, the interval $[n-k, n-1]$ contains a multiple of p if and only if

$$n \equiv 1, 2, \ldots, k \pmod{p}.$$

Hence the *allowed* residue classes modulo p are exactly the $p-k$ classes

$$\{\, 0,\ k+1,\ k+2,\ \ldots,\ p-1 \,\}.$$

Taking all primes $p \in \mathcal{P}_k$ together and applying the Chinese Remainder Theorem, the number of allowed residue classes modulo M_k is

$$A_k \;=\; \prod_{\substack{k<p<2k \\ p \text{ prime}}} (p-k),$$

so the "exact density" of admissible n in the integers is

$$D_k \;:=\; \frac{A_k}{M_k} \;=\; \prod_{\substack{k<p<2k \\ p \text{ prime}}} \left(1 - \frac{k}{p}\right).$$

Heuristically, the "typical gap" between admissible n_k is $\asymp D_k^{-1}$, so one expects the minimal solution n_k also to be of order D_k^{-1}. (However, I do not yet have a proof that n_k is indeed of order D_k^{-1}; this is precisely where things are currently stuck.)

In the next section we will explain how to obtain the asymptotic for D_k^{-1}, namely

$$\log \frac{1}{D_k} = (\log 4 + o(1))\,\frac{k}{\log k},$$

which is the key input toward the conjectured bound for $\log n_k$.

2. On the asymptotics of D_k^{-1}

Let

$$\log \frac{1}{D_k} = \sum_{\substack{k<p<2k \\ p \text{ prime}}} \log \frac{p}{p-k}.$$

We aim to show that this sum has the following asymptotic behavior.

Theorem 2.1. *As $k \to \infty$,*

$$\log \frac{1}{D_k} = (\log 4)\,\frac{k}{\log k} \;+\; O\Big(\frac{k}{(\log k)^2}\Big). \tag{2.1}$$

Before proving Theorem 2.1, we first establish some lemmas.

2.1. **Auxiliary lemmas.** Write $\pi(x)$ for the prime counting function and

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}.$$

The following lemma is standard, see [1, 2].

Lemma 2.2. *For any function F of bounded variation on $[k, 2k]$,*

$$\sum_{\substack{k<p\le 2k\\ p\ \text{prime}}} F(p) \ = \ \int_{(k,2k]} F(t)\, d\pi(t). \tag{2.2}$$

Lemma 2.3. *Let $F_k(t) := \log \frac{t}{t-k}$ for $t > k$ and*

$$I_k := \int_k^{2k} \frac{F_k(t)}{\log t}\, dt.$$

Then

$$I_k = (\log 4)\, \frac{k}{\log k} \ + \ O\left(\frac{k}{(\log k)^2}\right). \tag{2.3}$$

Proof. Substitute $t = k(1+u)$, $u \in (0, 1]$:

$$I_k = k \int_0^1 \frac{\log\left(\frac{1+u}{u}\right)}{\log k + \log(1+u)}\, du.$$

Since $0 \le \log(1+u) \le \log 2$, uniformly in u we have

$$\frac{1}{\log k + \log(1+u)} = \frac{1}{\log k} + O\left(\frac{1}{(\log k)^2}\right).$$

Moreover $\int_0^1 \left|\log\left(\frac{1+u}{u}\right)\right| du < \infty$, so we may integrate termwise to get

$$I_k = \frac{k}{\log k} \int_0^1 \log \frac{1+u}{u}\, du + O\left(\frac{k}{(\log k)^2}\right).$$

A direct computation gives $\int_0^1 \log\frac{1+u}{u}\, du = 2\log 2 = \log 4$. This completes the proof. $\square$

Lemma 2.4. *Let $\pi(x) = \mathrm{Li}(x) + E(x)$, where the prime number theorem holds with de la Vallée Poussin–type error*

$$E(x) \ = \ O\left(x e^{-c\sqrt{\log x}}\right) \qquad (x \to \infty)$$

for some constant $c > 0$. Then

$$\int_{(k,2k]} \log \frac{t}{t-k}\, dE(t) \ = \ O\left(\frac{k}{(\log k)^2}\right) \qquad (k \to \infty). \tag{2.4}$$

Proof. Write

$$F(t) := \log \frac{t}{t-k} \quad (t > k), \qquad I_k := \int_{(k,2k]} F(t)\, dE(t).$$

Split the interval:

$$I_k = \underbrace{\int_{(k,k+1]} F(t)\, dE(t)}_{=:I_0} + \underbrace{\int_{(k+1,2k]} F(t)\, dE(t)}_{=:I_1}.$$

Since $(k, k+1]$ contains at most one prime,

$$I_0 = \sum_{k<p\le k+1} F(p) - \int_k^{k+1} \frac{F(t)}{\log t}\, dt,$$

hence $|I_0| \le \log(k+1) + O(1) \ll k/(\log k)^2$.

On $[k+1, 2k]$ the function F is bounded and C^1, with

$$F(2k) = \log 2, \qquad F(k+1) = \log(k+1), \qquad F'(t) = \frac{1}{t} - \frac{1}{t-k} = -\frac{k}{t(t-k)}.$$

Integration by parts gives

$$I_1 = F(2k)E(2k) - F(k+1)E(k+1) - \int_{k+1}^{2k} E(t)\, F'(t)\, dt.$$

Put

$$H_k \;:=\; \sup_{t\in[k,2k]} t\, e^{-c\sqrt{\log t}} \;\asymp\; k\, e^{-c\sqrt{\log k}},$$

then $|E(t)| \ll H_k$ uniformly on $[k, 2k]$. Therefore

$$|F(2k)E(2k)| \ll H_k, \qquad |F(k+1)E(k+1)| \ll \log(k+1)\, H_k \ll H_k \log k,$$

and

$$\int_{k+1}^{2k} |E(t)F'(t)|\, dt \;\ll\; H_k \int_{k+1}^{2k} \frac{k}{t(t-k)}\, dt \;=\; H_k \left[\log \frac{t-k}{t}\right]_{k+1}^{2k} \;\asymp\; H_k \log k.$$

Hence

$$|I_1| \;\ll\; H_k \log k \;\asymp\; k\, e^{-c\sqrt{\log k}} \cdot \log k.$$

Hence $|I_1| \ll k/(\log k)^2$.

Combining the bounds for I_0 and I_1 yields (2.4). □

We are now ready to provide the following.

2.2. **Proof of Theorem 2.1.** By Lemma 2.2 with $F_k(t) = \log \frac{t}{t-k}$,

$$\log \frac{1}{D_k} = \sum_{\substack{k<p<2k \\ p \text{ prime}}} \log \frac{p}{p-k} = \int_{(k,2k]} \log\frac{t}{t-k}\, d\pi(t) = \int_k^{2k} \frac{\log\frac{t}{t-k}}{\log t}\, dt + \int_{(k,2k]} \log\frac{t}{t-k}\, dE(t).$$

The first integral equals $(\log 4)\frac{k}{\log k} + O\left(\frac{k}{(\log k)^2}\right)$ by Lemma 2.3, and the second is $O\left(\frac{k}{(\log k)^2}\right)$ by Lemma 2.4. This proves (2.1). □

3. Concluding remarks

We have now established the asymptotic density estimate

$$\log \frac{1}{D_k} \;=\; (\log 4 + o(1))\, \frac{k}{\log k}.$$

The remaining problem is to understand the connection between n_k and $\log\frac{1}{D_k}$. Intuitively, the average spacing of admissible integers n_k is $\asymp D_k^{-1}$, so one expects that the minimal solution n_k should also be of order D_k^{-1}. However, making this heuristic precise is delicate. At present, we do not know how to establish such a result rigorously. Bridging this gap between the density asymptotic and the actual minimal solution n_k is what remains to be done.

Based on the numerical figure in Section 1, we also computed n_k and D_k for $1 \le k \le 50$ and obtained the ratios

$$R_k := n_k\, D_k.$$

In this range the smallest and largest values of R_k occur at

$$\textbf{Min } R_k \text{ at } k = 46: \quad n_{46} = 241{,}251, \quad D_{46} = 5.5915609400 \times 10^{-7}, \quad R_{46} = 0.1348969668,$$

$$\textbf{Max } R_k \text{ at } k = 23: \quad n_{23} = 1884, \quad D_{23} = 0.0041253238, \quad R_{23} = 7.7721101000.$$

Acknowledgments

The author is grateful to Stijn Cambie for his helpful comments. The author also thanks Thomas Bloom for founding and maintaining the Erdős Problems website, which has been a valuable resource and inspiration for this work.

References

[1] H. Iwaniec and E. Kowalski, *Analytic Number Theory*.
[2] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*.

School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, P. R. China
Email address: tang_quanyu@163.com