1. Introduction

In [4, p. 80], Erdős wrote:

I just want to state without proof a special result in this direction, namely

$$\frac{\log\log n}{\log n}n + c_9\frac{n}{(\log n)^2} \le g_3(n) \le \frac{\log\log n}{\log n}n + c_{10}\frac{n}{(\log n)^2}. \tag{1.1}$$

It is not clear whether (1.1) *can be sharpened.*

This leads to the following problem, which also appears as Problem #796 on Bloom's Erdős Problems website [1].

Problem 1.1. *Let $k \ge 2$ and let $g_k(n)$ be the largest possible size of $A \subseteq \{1, \dots, n\}$ such that every m has $< k$ solutions to $m = a_1a_2$ with $a_1 < a_2 \in A$.*

Is it true that

$$g_3(n) = \frac{\log\log n}{\log n}n + (c + o(1))\frac{n}{(\log n)^2}$$

for some constant c?

In this note, we show that (1.1) is false, thereby providing a counterexample to Problem 1.1. We also give a corrected version of the problem.

2. A lower bound

Fix $n \ge 2$. Then $g_3(n)$ denote the largest cardinality of a set $A \subseteq \{1, 2, \dots, n\}$ with the property that for every integer m the equation

$$m = a_1a_2, \qquad a_1 < a_2,\ a_1, a_2 \in A,$$

has at most 2 solutions.

Let M be the Meissel–Mertens constant. Then we have the following lower bound.

Proposition 2.1. *For all sufficiently large n,*

$$g_3(n) \ge \frac{n\log\log n}{\log n} + (M + 1 + o(1))\frac{n}{\log n}.$$

Proof. Let

$$P := \{p \le n :\ p \text{ is prime}\}, \qquad S := \{pq \le n :\ p, q \text{ primes and } q > \sqrt{n}\},$$

and define

$$A := P \cup S \subseteq \{1, 2, \dots, n\}.$$

1. A has at most two multiplicative representations. Call a prime *large* if it exceeds $\sqrt{n}$.

Claim 1. *Every element $a \in A$ has at most one large prime divisor (counted without multiplicity).*

Proof of Claim 1. If $a \in P$ then a is prime, so the claim is trivial. If $a \in S$, then $a = pq$ with $q > \sqrt{n}$ prime and $p \le n/q < \sqrt{n}$, so q is the unique large prime divisor of a. □

Claim 2. *For every integer m, the number of solutions of $m = a_1a_2$ with $a_1 < a_2$ and $a_1, a_2 \in A$ is at most 2.*

Proof of Claim 2. Fix m and suppose $m = a_1a_2$ with $a_1 < a_2$ and $a_1, a_2 \in A$. By Claim 1, each a_i contributes at most one large prime divisor, so m has at most two large prime divisors (counted with multiplicity).

Case 0: m has no large prime divisor. Then neither a_1 nor a_2 can lie in S, hence $a_1, a_2 \in P$ and $a_1, a_2 \le \sqrt{n}$. Thus m is the product of two primes and the representation is unique up to order; with the constraint $a_1 < a_2$, there is at most one solution.

Case 1: m has exactly one large prime divisor $q > \sqrt{n}$. Then exactly one of a_1, a_2 is divisible by q. The factor divisible by q is either $q \in P$ or $pq \in S$ with $p \le \sqrt{n}$ prime. The other factor has no large prime divisor, hence must be a prime $\le \sqrt{n}$ (an element of P). Consequently, m has the form

$$m = qpr$$

where $p, r \le \sqrt{n}$ are primes (not necessarily distinct), and any solution corresponds to deciding whether q is paired with p or with r. This yields at most 2 solutions (with fewer if $p = r$).

Case 2: m has exactly two large prime divisors (counted with multiplicity). Let the large prime divisors of m (with multiplicity) be $q_1, q_2 > \sqrt{n}$. By Claim 1, each $a \in A$ contains at most one large prime divisor, hence in any representation $m = a_1a_2$ with $a_1 < a_2$ and $a_1, a_2 \in A$, the two large primes q_1, q_2 must be split between a_1 and a_2, i.e. each of a_1, a_2 is divisible by exactly one of q_1, q_2. Moreover, by the definition of $A = P \cup S$, every element of A is either a prime $\le n$, or a product pq with $q > \sqrt{n}$ prime and p prime (necessarily $p < \sqrt{n}$). In particular, if $a \in A$ is divisible by a large prime $q > \sqrt{n}$, then

$$a = q \cdot u \qquad \text{with } u \in \{1\} \cup \{\text{primes } \le \sqrt{n}\}.$$

Therefore any solution $m = a_1a_2$ forces

$$a_1 = q_1u_1, \qquad a_2 = q_2u_2$$

(up to swapping q_1, q_2), where $u_1, u_2 \in \{1\} \cup \{\text{primes } \le \sqrt{n}\}$.

If $q_1 \ne q_2$, then the only possible ambiguity comes from swapping which small factor (u_1 or u_2) is paired with q_1 or q_2. Since each u_i is either 1 or a single prime, there are at most two such pairings:

$$(q_1u_1)(q_2u_2) \quad \text{or} \quad (q_1u_2)(q_2u_1),$$

and after imposing $a_1 < a_2$ this gives at most 2 solutions (and fewer if $u_1 = u_2$).

If $q_1 = q_2 = q$, then every admissible factor must contain exactly one copy of q, hence $a_1 = qu_1$, $a_2 = qu_2$ with u_1, u_2 as above. Swapping yields the same unordered pair, so with the constraint $a_1 < a_2$ there is at most one solution.

Thus in all subcases, the number of solutions is at most 2.

In all cases, the number of solutions is at most 2. □

2. Size of A. Clearly

$$|A| = |P| + |S| = \pi(n) + |S|.$$

Thus by Proposition A.2, we know that

$$|A| = \pi(n) + |S| = \frac{n \log\log n}{\log n} + (M+1)\frac{n}{\log n} + o\left(\frac{n}{\log n}\right).$$

Claim 2 shows that A is admissible, hence $g_3(n) \ge |A|$. This completes the proof. □

3. Concluding Remarks

In [3, p. 261], Erdős stated that his [3, Theorem 3] could be sharpened to

$$g_3(n) \le \frac{\log\log n}{\log n} n + O\left(\frac{n}{(\log n)^{1+c}}\right),$$

where $c > 0$ is a suitable positive constant. In view of Proposition 2.1, we now conjecture that the second term in the upper bound is also of order $n/\log n$:

Conjecture 3.1. *There exists an absolute constant $c > 0$ such that, for all sufficiently large n,*

$$g_3(n) \le \frac{n\log\log n}{\log n} + c\frac{n}{\log n}.$$

More strongly, we also conjecture the following modified version of Problem 1.1:

Conjecture 3.2. *Is it true that*

$$g_3(n) = \frac{\log\log n}{\log n} n + (c + o(1))\frac{n}{\log n}$$

for some constant c?

References

[1] T. F. Bloom, Erdős Problem #796, `https://www.erdosproblems.com/796`, accessed 2025-12-30.

[2] Crisan, Dragos, and Radek Erban. "On the counting function of semiprimes." arXiv preprint arXiv:2006.16491 (2020).

[3] P. Erdős, On the multiplicative representation of integers. Israel Journal of Mathematics (1964), 251–261.

[4] P. Erdős, Some applications of graph theory to number theory. The Many Facets of Graph Theory (Proc. Conf., Western Mich. Univ., Kalamazoo, Mich., 1968) (1969), 77–82.

Appendix A. Preliminaries

Let $n \ge 3$. Define

$$S := \{pq \le n : \ p, q \text{ are primes and } q > \sqrt{n}\}.$$

Let $\pi(x)$ be the prime-counting function, and let $\pi_2(x)$ denote the *semiprime counting function* defined by

$$\pi_2(x) := \#\{(p,q) : \ p \le q, \ p, q \text{ prime}, \ pq \le x\}.$$

Lemma A.1. *For every $n \ge 3$,*

$$|S| = \pi_2(n) - \#\{(p,q) : \ p \le q \le \sqrt{n}, \ p, q \text{ prime}\} = \pi_2(n) - \frac{\pi(\sqrt{n})\big(\pi(\sqrt{n}) + 1\big)}{2}.$$

Proof. Every pair (p, q) of primes with $p \le q$ and $pq \le n$ falls into exactly one of the two disjoint classes:

$$\mathcal{A} := \{(p,q) : \ p \le q \le \sqrt{n}\}, \qquad \mathcal{B} := \{(p,q) : \ p \le q, \ pq \le n, \ q > \sqrt{n}\}.$$

Hence $\pi_2(n) = |\mathcal{A}| + |\mathcal{B}|$.

If $(p, q) \in \mathcal{A}$ then automatically $pq \le (\sqrt{n})^2 = n$, so $\mathcal{A}$ is simply the set of unordered pairs of primes $\le \sqrt{n}$ with repetition allowed. Therefore

$$|\mathcal{A}| = \binom{\pi(\sqrt{n})}{2} + \pi(\sqrt{n}) = \frac{\pi(\sqrt{n})\big(\pi(\sqrt{n}) + 1\big)}{2}.$$

If $(p, q) \in \mathcal{B}$ then $q > \sqrt{n}$ and $pq \le n$; this is equivalent to the condition that the integer pq lies in S. Moreover, because $q > \sqrt{n}$ implies $p < n/q < \sqrt{n} < q$, each element of S corresponds to a unique pair $(p, q) \in \mathcal{B}$ with $p \le q$. Thus $|\mathcal{B}| = |S|$, and the claimed identity follows. □

Clearly, we have the asymptotic formula (see, for example, [2, Theorem 2.3])

$$\pi_2(x) = \frac{x \log\log x}{\log x} + M\frac{x}{\log x} + o\left(\frac{x}{\log x}\right). \tag{A.1}$$

We also recall the prime number theorem $\pi(x) \sim x/\log x$.

Proposition A.2. *We have*

$$|S| = \frac{n \log\log n}{\log n} + M\frac{n}{\log n} + o\left(\frac{n}{\log n}\right).$$

Proof. By Lemma A.1,

$$|S| = \pi_2(n) - \frac{\pi(\sqrt{n})\big(\pi(\sqrt{n})+1\big)}{2}.$$

By the prime number theorem, $\pi(\sqrt{n}) = O(\sqrt{n}/\log n)$, hence

$$\pi(\sqrt{n})^2 = O\left(\frac{n}{(\log n)^2}\right) = o\left(\frac{n}{\log n}\right), \qquad \pi(\sqrt{n}) = o\left(\frac{n}{\log n}\right).$$

Therefore

$$\frac{\pi(\sqrt{n})\big(\pi(\sqrt{n})+1\big)}{2} = o\left(\frac{n}{\log n}\right).$$

Combining this with (A.1) at $x = n$ yields

$$|S| = \frac{n \log\log n}{\log n} + M\frac{n}{\log n} + o\left(\frac{n}{\log n}\right). \qquad \square$$