Erdős Problem #1054: a consolidated write-up of the forum proofs

Compiled from comments by
jif, Terence Tao, Vjekoslav Kovač, Thomas Bloom,
Liam Price/GPT-5.5 Pro, and principia_math

June 24, 2026

Abstract

This note consolidates the main proof ideas appearing in the discussion thread for Erdős Problem #1054. The problem asks about the least integer m for which N is the sum of the k smallest divisors of m, for some $k \geq 1$. The write-up has three purposes. First, it records jif's proof/certification that the function is defined for every positive integer except 2 and 5. Second, it gives a clean version of the Tao–Kovač density argument, culminating in Kovač's strengthened bound

$$\#\{N \leq X : f(N) \leq \delta N\} \ll \exp\{-\exp((1/\delta)^c)\}X$$

for some $c > 0$ and sufficiently small δ. Third, it records, but does not reproduce, the June 2026 principia_math/Principia Math claim resolving the remaining limsup question, because the forum page gives only a summary and links to an external Overleaf/Lean proof. All attributions are made to the forum names or author names used in the thread.

1 Problem, notation, and attributions

Let

$$1 = d_1(m) < d_2(m) < \cdots < d_{\tau(m)}(m) = m$$

be the positive divisors of m in increasing order. We say that N is represented by m if

$$N = d_1(m) + \cdots + d_k(m)$$

for some $1 \leq k \leq \tau(m)$. Erdős Problem #1054 asks about

$$f(N) = \min\{m \geq 1 : N \text{ is represented by } m\},$$

when this minimum exists. The question, as stated on the Erdős Problems site and attributed there to Guy's collection, is whether $f(N) = o(N)$, perhaps for almost all N, and whether

$$\limsup_{N\to\infty} \frac{f(N)}{N} = \infty.$$

The formulation is closely related to Erdős Problem #468, with an index shift coming from whether 1 is counted among the divisors.

For density estimates it is convenient to put $f(2) = f(5) = \infty$, since these two values are not represented; this convention does not affect any asymptotic statement.

For $j \geq 0$ define

$$\sigma_j(m) = \sum_{i=1}^{\tau(m)-j} d_i(m),$$

with the convention that the sum is 0 if $j \geq \tau(m)$. Thus $\sigma_j(m)$ is the sum of all divisors of m except the j largest ones. The condition "N is the sum of the k smallest divisors of m" is the same as

$$N = \sigma_j(m) \qquad (j = \tau(m) - k).$$

The principal contributions used below are as follows.

- **jif** gave a proof/certificate that the represented positive integers are exactly $\mathbb{N} \setminus \{2, 5\}$, using finite verification, prime-window subset-sum intervals, Rosser–Schoenfeld bounds, and an explicit consequence of Helfgott's ternary Goldbach theorem.
- **Terence Tao** gave the original density obstruction to $f(N) = o(N)$, via estimates for the sums $\sigma_j(m)$ and a Markov/union-bound argument.
- **Vjekoslav Kovač (Vjeko_Kovac)** wrote a short self-contained note, based on Tao's argument and comments by Thomas Bloom, proving a much stronger uniform upper bound for the density of $\{N : f(N) \leq \delta N\}$.
- **Thomas Bloom** pointed out the dependence on the moment parameter and how to optimize it.
- **Liam Price** posted links to GPT-5.5 Pro proofs giving intermediate fixed-power improvements. These are superseded by Kovač's shorter argument, but are part of the thread history.
- **principia_math / Principia Math** posted a claimed resolution of the remaining limsup question. Since the thread text gives only the statement and links to an Overleaf/Lean proof, this note records that claim separately rather than treating it as independently checked here.

2 Representability and well-definedness

We first isolate the elementary obstruction.

Proposition 2.1 (The exceptional values 2 and 5)**.** *The positive integers* 2 *and* 5 *are not represented. The integers* $1, 3, 4$ *are represented.*

Proof. Every nonempty prefix sum of divisors begins with 1. Hence a two-term prefix sum is at least $1 + 2 = 3$, so 2 cannot occur. If 5 were a two-term prefix sum, it would have to be $1 + 4$, but 4 cannot be the smallest divisor after 1, since a number whose least nontrivial divisor is 4 would have no divisor 2 and hence would be odd, impossible if it is divisible by 4. With three or more terms the prefix sum is at least $1 + 2 + 3 = 6$ when it exists; in any case it is never 5. Finally 1, 3, and 4 are represented by $m = 1, 2, 3$, respectively. □

The positive result rests on two standard gadgets plus explicit finite certificates.

Lemma 2.2 (Prime-window prefix gadget, jif)**.** *Let $M < p_1 < \cdots < p_t \leq X < M^2$ be primes, and let*

$$m = p_1 p_2 \cdots p_t.$$

Then the first divisors of m *are*

$$1, p_1, p_2, \dots, p_t.$$

Consequently, if $N = p_{i_1} + \cdots + p_{i_s}$ *is a subset sum of primes in* $(M, X]$*, then* $N + 1$ *is represented by* $m = p_{i_1} \cdots p_{i_s}$*.*

Proof. Every composite divisor of m is a product of at least two primes, hence is $> M^2$. Since $X < M^2$, all primes $p_i \le X$ occur before any composite divisor. Taking the product over a selected subset gives initial divisors 1 followed by exactly the selected primes in increasing order, and the relevant prefix sum is $1 + N$. ☐

Lemma 2.3 (Interval-extension trick, jif)**.** *Suppose subset sums of already chosen positive integers cover every integer in an interval* $[C, U]$*. If the next available integer* p *satisfies*

$$p \le U - C + 1,$$

then after adjoining p *the subset sums cover* $[C, U + p]$*.*

Proof. The old sums cover $[C, U]$ and the old sums plus p cover $[C + p, U + p]$. The hypothesis $p \le U - C + 1$ says $C + p \le U + 1$, so these two integer intervals touch or overlap. Their union is therefore $[C, U + p]$. ☐

The following theorem is jif's well-definedness result as recorded in the forum and in the accompanying GitHub repository. The hand proof below identifies the elementary structure; the finite and numerical inequalities are those certified in the repository's verifier output.

Theorem 2.4 (jif's representability theorem)**.** *Assuming the finite certificates and numerical inequalities verified in jif's repository, every positive integer other than* 2 *and* 5 *is represented. Equivalently,* $f(N)$ *is finite exactly for*

$$N \in \mathbb{N} \setminus \{2, 5\}.$$

Proof. The values $1, 3, 4$ are represented and $2, 5$ are not, by Proposition 2.1. It remains to cover $n \ge 6$.

The verifier records four overlapping ranges.

(i) A direct divisor-prefix search covers every $6 \le n \le 10^7$, except that the value 7 is separately represented as $7 = 1 + 2 + 4$ by $m = 4$.

(ii) A prime-window certificate with $M = 10^4$ and $X = 99{,}000{,}000 < M^2$ gives representations for every

$$469{,}616 \le n \le 273{,}803{,}744{,}799{,}154.$$

(iii) A larger seed proves that every integer in $[105{,}000{,}000, 156{,}000{,}000]$ is a sum of distinct primes in $(20{,}000{,}000, 40{,}000{,}000)$. Since the seed length is $51{,}000{,}001$ and the next prime after $40{,}000{,}000$ is $40{,}000{,}003$, the interval-extension lemma starts. Bertrand's postulate continues the induction, and Rosser–Schoenfeld bounds for $\pi(x)$ lower-bound the available prime mass up to $399{,}000{,}000{,}000{,}000 < 20{,}000{,}000^2$. This yields representations for every

$$105{,}000{,}001 \le n \le 1{,}111{,}351{,}202{,}532{,}220{,}892{,}436{,}000{,}001.$$

(iv) For the tail one uses the explicit Helfgott-type input certified by jif: every odd $T \geq 10^{27}$ is a sum of three distinct odd primes, each larger than $T/(30000 \log T)$. The numerical margins excluding small or repeated coordinates are checked using the explicit constants quoted from Helfgott's Section 7 and Rosser–Schoenfeld's estimate for sums of Λ.

We spell out how the tail input yields divisor-prefix representations. Let n be even and large enough that $T = n - 1 \geq 10^{27}$. Write

$$n - 1 = p + q + r$$

with p, q, r distinct odd primes, each $> T/(30000 \log T)$. For $T \geq 10^{27}$ this lower bound gives $pq > T > r$. Hence for $m = pqr$ the initial divisors are $1, p, q, r$, and their sum is n.

For odd $n \geq 10^{27} + 10^8$, choose a prime

$$60000 \log n < \ell < 120000 \log n$$

using Bertrand's postulate. Put $T = n - 1 - \ell$, which is odd and at least 10^{27} in the range under consideration. Write $T = p + q + r$ as above. The verifier checks the elementary slack inequalities

$$\ell < p, \qquad \ell p > r.$$

Therefore the initial divisors of $m = \ell pqr$ are $1, \ell, p, q, r$, and their sum is n.

The four covered intervals

$$[6, 10^7], \quad [469, 616, 273, 803, 744, 799, 154], \quad [105, 000, 001, 1, 111, 351, 202, 532, 220, 892, 436, 000, 001],$$

$$[10^{27} + 10^8, \infty)$$

overlap. Thus every $n \geq 6$ is represented. Together with Proposition 2.1 this proves the claim, conditional only on the stated verifier outputs and explicit analytic inputs. □

Remark 2.5. This section is deliberately phrased as a certificate-based proof. The arithmetic for the large finite intervals is not reproduced line by line in this note; it is part of jif's Rust verifier and committed output.

3 The density obstruction of Tao and its sharpening by Kovač

The decisive observation is that representations with $m \leq \delta N$ force a large value of $\sigma_j(m)/m$.

Let

$$E_\delta(X) = \{N \leq X : f(N) \leq \delta N\}.$$

If $N \in E_\delta(X)$ and $m = f(N)$, then $m \leq \delta X$ and, for some $j \geq 0$,

$$N = \sigma_j(m), \qquad \frac{\sigma_j(m)}{m} = \frac{N}{m} \geq \frac{1}{\delta}.$$

Thus estimates for the moments of $\sigma_j(m)/m$ immediately give density bounds for $E_\delta(X)$.

3.1 Kovač's moment estimate

The following is the core estimate in Kovač's note. It is a cleaned-up version of his proof, based on Tao's original idea and Bloom's comments on the dependence on the moment parameter.

Lemma 3.1 (Kovač)**.** *There is an absolute constant $C > 0$ such that, for every integer $q \geq 3$ and every real $x \geq 1$,*

$$\sum_{m \leq x} \sum_{j \geq 0} \left(\frac{\sigma_j(m)}{m} \right)^q \ll \exp(Cq \log\log q) x.$$

The implicit constant is absolute.

Proof. Let

$$1 = r_1 < r_2 < \cdots < r_{\tau(m)} = m$$

be the divisors of m. Reflecting the divisor list gives

$$\frac{\sigma_j(m)}{m} = \sum_{i > j} \frac{1}{r_i}.$$

After expanding the q-th power and then summing over j, a fixed tuple $a_1, \dots, a_q$ of divisors is counted once for each divisor threshold not exceeding $\min a_i$. Hence

$$\sum_{j \geq 0} \left(\frac{\sigma_j(m)}{m} \right)^q = \sum_{\substack{r, a_1, \dots, a_q | m \\ r \leq \min(a_1, \dots, a_q)}} \frac{1}{a_1 \cdots a_q}.$$

Summing over $m \leq x$ and bounding the number of multiples of $[r, a_1, \dots, a_q]$ by $x/[r, a_1, \dots, a_q]$, it is enough to prove

$$S := \sum_{\substack{r, a_1, \dots, a_q \geq 1 \\ r \leq \min a_i}} \frac{1}{a_1 \cdots a_q [r, a_1, \dots, a_q]} \ll \exp(Cq \log\log q).$$

For each i put

$$b_i = \frac{r}{(r, a_i)}, \qquad c_i = \frac{a_i}{(r, a_i)}.$$

Then $b_i \mid r$, $(b_i, c_i) = 1$, $a_i = rc_i/b_i$, and

$$[r, a_1, \dots, a_q] = r[c_1, \dots, c_q].$$

The condition $r \leq a_i$ becomes $b_i \leq c_i$. Dropping the coprimality condition only increases the sum, so

$$S \leq \sum_{r \geq 1} \frac{1}{r^{q+1}} \sum_{b_1, \dots, b_q | r} b_1 \cdots b_q \sum_{c_i \geq b_i} \frac{1}{c_1 \cdots c_q [c_1, \dots, c_q]}.$$

Fix $0 < \beta < 1$. Since $c_i \geq b_i$,

$$\frac{1}{c_i} \leq \frac{1}{b_i^{1-\beta} c_i^{\beta}}.$$

Therefore

$$S \leq S' S'',$$

where, with $\beta = (q-1)/q$,

$$S' = \sum_{r\geq 1} \frac{1}{r^2} \left(\sum_{d|r} \frac{1}{d^{(q-1)/q}} \right)^q$$

and

$$S'' = \sum_{c_1,\dots,c_q\geq 1} \frac{1}{c_1^{(q-1)/q}\cdots c_q^{(q-1)/q}[c_1,\dots,c_q]}.$$

It remains to estimate these two Euler products.

For S', multiplicativity gives

$$S' = \prod_p L_p,$$

where

$$L_p = \sum_{\ell\geq 0} p^{-2\ell} \left(\sum_{0\leq\gamma\leq\ell} p^{-((q-1)/q)\gamma} \right)^q .$$

For $p \leq q$,

$$L_p \leq (1-p^{-2})^{-1}(1-p^{-(q-1)/q})^{-q}.$$

Using

$$\sum_{p\leq q} p^{-(q-1)/q} \leq q^{1/q} \sum_{p\leq q} \frac{1}{p} \ll \log\log q,$$

we obtain

$$\prod_{p\leq q} L_p \ll \exp(Cq\log\log q).$$

For $p > q$, one has $qp^{-(q-1)/q} \ll 1$ and hence

$$L_p \leq 1 + O(p^{-2}).$$

Thus $\prod_{p>q} L_p \ll 1$, and so

$$S' \ll \exp(Cq\log\log q).$$

For S'', again by multiplicativity,

$$S'' = \prod_p M_p,$$

where

$$M_p = \sum_{\gamma_1,\dots,\gamma_q\geq 0} p^{-((q-1)/q)(\gamma_1+\cdots+\gamma_q)-\max_i\gamma_i}.$$

Let

$$A_h = \sum_{0\leq\gamma\leq h} p^{-((q-1)/q)\gamma}.$$

Grouping terms by $h = \max_i \gamma_i$ gives

$$M_p = \sum_{h\geq 0} p^{-h}(A_h^q - A_{h-1}^q), \qquad A_{-1} = 0.$$

For $p \leq q$ the rough bound

$$A_h \leq (1-p^{-(q-1)/q})^{-1}$$

yields

$$M_p \leq (1-p^{-1})^{-1}(1-p^{-(q-1)/q})^{-q},$$

and therefore

$$\prod_{p\leq q} M_p \ll \exp(Cq\log\log q).$$

For $p > q$, the mean value theorem gives

$$A_h^q - A_{h-1}^q \ll qp^{-h(q-1)/q},$$

so

$$M_p - 1 \ll q\sum_{h\geq 1} p^{-h(1+(q-1)/q)} \ll qp^{-2+1/q}.$$

Consequently

$$\sum_{p>q}(M_p-1) \ll q\sum_{n>q} n^{-2+1/q} \ll 1,$$

and $\prod_{p>q} M_p \ll 1$. Hence

$$S'' \ll \exp(Cq\log\log q).$$

Combining the estimates for S' and S'' proves the lemma, after adjusting C. □

3.2 The strengthened density theorem

Theorem 3.2 (Kovač; strengthening Tao's obstruction)**.** *There exists $c > 0$ such that, for all sufficiently small $\delta > 0$ and all $X \geq 1$,*

$$\#\{N \leq X : f(N) \leq \delta N\} \ll \exp\{-\exp((1/\delta)^c)\}X.$$

Consequently, for every fixed $M > 0$,

$$\#\{N \leq X : f(N) \leq \delta N\} \ll_M \delta^M X$$

uniformly in X.

Proof. Let $E_\delta(X)$ denote the set on the left. If $N \in E_\delta(X)$ and $m = f(N)$, then $m \leq \delta X$ and $N = \sigma_j(m)$ for some $j \geq 0$ with $\sigma_j(m)/m \geq 1/\delta$. Therefore, for every integer $q \geq 3$,

$$\begin{aligned}\#E_\delta(X) &\leq \#\left\{(m,j) : m \leq \delta X,\ \frac{\sigma_j(m)}{m} \geq \frac{1}{\delta}\right\}\\ &\leq \delta^q \sum_{m\leq\delta X}\sum_{j\geq 0}\left(\frac{\sigma_j(m)}{m}\right)^q\\ &\ll \delta^{q+1}\exp(Cq\log\log q)X,\end{aligned}$$

where Lemma 3.1 is used in the last line.

Choose $c = 1/(2C)$, after enlarging C if necessary, and take

$$q = \lfloor \exp((1/\delta)^c)\rfloor.$$

Then

$$\delta^{q+1}\exp(Cq\log\log q) = \exp\left(-(q+1)\log\frac{1}{\delta} + Cq\log\log q\right).$$

Because $\log\log q \sim c\log(1/\delta)$, the exponent is at most

$$-\frac{1}{2}q\log\frac{1}{\delta}$$

for sufficiently small δ. This is stronger than

$$-\exp((1/\delta)^c)$$

after slightly reducing c, and the theorem follows. The fixed-power bound $O_M(\delta^M X)$ is immediate from the double-exponential decay as $\delta \to 0$. $\square$

Corollary 3.3 (Failure of the "almost all" $o(N)$ statement). *For all sufficiently small fixed $\delta > 0$, the set*

$$\{N \in \mathbb{N} : f(N) > \delta N\}$$

has lower density arbitrarily close to 1. In particular, $f(N) = o(N)$ is false even as an almost-all statement.

Proof. Theorem 3.2 gives

$$\sup_{X\geq 1}\frac{\#\{N \leq X : f(N) \leq \delta N\}}{X} \to 0 \qquad (\delta \to 0).$$

Thus, for any prescribed $\eta > 0$, choosing δ small enough gives

$$\frac{\#\{N \leq X : f(N) > \delta N\}}{X} \geq 1-\eta$$

for all X up to the harmless exceptional convention at $2, 5$. $\square$

Remark 3.4 (Relation to Tao's original $O(\delta^2)$ bound). Tao's forum argument established the first decisive density obstruction, in the form

$$\#\{N \leq X : f(N) \leq \delta N\} \ll \delta^2 X.$$

One way to view Kovač's theorem is that it replaces Tao's first-moment/union-bound estimate by a flexible high-moment estimate. Liam Price's linked GPT-5.5 Pro write-ups supplied intermediate fixed-power improvements such as $O(\delta^3)$ and $O_k(\delta^k)$, but Kovač's note gives a shorter and quantitatively stronger version.

4 The claimed limsup resolution

The thread also contains a June 22, 2026 comment by the user principia_math, claiming a proposed resolution of the remaining limsup question.

Claim 4.1 (principia_math / Principia Math, as stated on the forum). *For every fixed $A \geq 1$, the set of positive integers N satisfying*

$$f(N) > AN$$

has positive lower density. Consequently,

$$\limsup_{N\to\infty}\frac{f(N)}{N} = \infty.$$

The forum summary says that the proof was accompanied by an Overleaf LaTeX document and a Lean formalization. It also says that the Lean file isolates two classical analytic inputs as assumptions: a quantitative Mertens product estimate and an almost-all binary Goldbach theorem. The same comment notes that the authors later learned of jif's proof of well-definedness and that their Lean formalization instead used an almost-all binary Goldbach estimate.

Remark 4.2. This consolidated note does not reproduce the proof of Claim 4.1. The reason is not mathematical disagreement, but source availability: the thread text gives only a high-level summary, while the proof itself is in an external Overleaf/Lean link. Accordingly, Claim 4.1 is recorded as a claimed resolution with its stated dependencies, not as a proof checked in this document.

5 What is proved here and what remains source-dependent

Combining the thread contributions gives the following status.

(1) The values 2 and 5 are impossible, and jif's certificate-based proof covers all other positive integers. Therefore $f(N)$ is well-defined for all $N \notin \{2, 5\}$, subject to the finite verifier and explicit analytic inputs cited above.

(2) The assertion $f(N) = o(N)$ is false, not merely pointwise but in the almost-all sense: for small δ, the proportion of $N \le X$ with $f(N) \le \delta N$ is extremely small, uniformly in X.

(3) The best density upper bound written out in the thread is Kovač's

$$\sup_{X \ge 1} \frac{\#\{N \le X : f(N) \le \delta N\}}{X} \ll \exp\{-\exp((1/\delta)^c)\}.$$

(4) The remaining limsup question is claimed to have been resolved by principia_math / Principia Math, but its proof is not reproduced in the forum text and therefore is not incorporated as a checked proof here.

References

[1] Thomas F. Bloom, *Erdős Problem #1054*, Erdős Problems website, problem page and discussion thread, accessed 24 June 2026.
`https://www.erdosproblems.com/1054`
`https://www.erdosproblems.com/forum/thread/1054`

[2] Richard K. Guy, *Unsolved Problems in Number Theory*, third edition, Problem B2, Springer, 2004.

[3] Terence Tao (forum nickname `TerenceTao`), comment under Erdős Problem #1054, 1 November 2025, giving the density obstruction via the sums $\sigma_j(n)$.

[4] Vjekoslav Kovač (forum nickname `Vjeko_Kovac`), *An improved bound in Erdős Problem #1054*, short note dated 17 May 2026.
`https://web.math.pmf.unizg.hr/~vjekovac/files/Erdos_problem_1054.pdf`

[5] Thomas Bloom, comments under Erdős Problem #1054, 11 May 2026, on the dependence on the moment parameter and the optimization of the density estimate.

[6] Liam Price, comments under Erdős Problem #1054, 10 May 2026, linking GPT-5.5 Pro write-ups with intermediate fixed-power improvements.

[7] `jif`, comment under Erdős Problem #1054, 16 April 2026, and GitHub repository `jif-perso/erdos_1054`, giving the representability verifier.
`https://github.com/jif-perso/erdos_1054`

[8] Harald A. Helfgott, *The ternary Goldbach conjecture is true*, Annals of Mathematics Studies 195, Princeton University Press, 2015; see especially Section 7 for the explicit estimates used in jif's tail argument.

[9] J. Barkley Rosser and Lowell Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois Journal of Mathematics 6 (1962), 64–94.

[10] `principia_math` / Principia Math, comment under Erdős Problem #1054, 22 June 2026, claiming a proposed resolution of the limsup question and linking a LaTeX proof and Lean formalization.