A Proposed Solution to Erdős Problem 486

Shouqiao Wang

Columbia University Multiscalar Intelligence

Abstract

For $A \subseteq \mathbb{N}$ and arbitrary sets $X_n \subseteq \mathbb{Z}/n\mathbb{Z}$, let

$$B = \{m \in \mathbb{N} : m \bmod n \notin X_n \text{ for every } n \in A \text{ with } n < m\}.$$

Erdős asked whether B must possess a logarithmic density. We propose a negative answer. The main ingredient is a finite probabilistic construction: at each sufficiently large dyadic scale, a fixed positive proportion of the integers in a short interval is deleted by moduli that are already active, whereas the union of the associated periodic residue cylinders has stretched-exponentially small density. A gliding-hump construction then places long runs of such blocks between recovery gaps and produces one fixed congruence system for which

$$\liminf_{x\to\infty} \frac{1}{\log x} \sum_{\substack{m<x \\ m\in B}} \frac{1}{m} \le \frac{177}{200} < \frac{49}{50} \le \limsup_{x\to\infty} \frac{1}{\log x} \sum_{\substack{m<x \\ m\in B}} \frac{1}{m}.$$

This proposed solution was found by GPT-5.6.

1 Introduction

Throughout, $\mathbb{N} = \{1, 2, \ldots\}$, and all logarithms are natural unless a base is explicitly indicated. Erdős stated the question, with the activation condition built into the formulation, in [7, p. 48]; he later recorded it as Problem I.26 in [8, pp. 235–236]. For each n in a set $A \subseteq \mathbb{N}$, choose an arbitrary set $X_n \subseteq \mathbb{Z}/n\mathbb{Z}$, and define

$$B = B(A, X) := \{m \in \mathbb{N} : m \bmod n \notin X_n \text{ for every } n \in A \text{ with } n < m\}. \tag{1.1}$$

The problem asks whether the logarithmic averages

$$L_B(x) := \frac{1}{\log x} \sum_{\substack{m<x \\ m\in B}} \frac{1}{m}$$

must converge as $x \to \infty$.

Theorem 1.1 (Main theorem)**.** *There are a fixed infinite $A \subseteq \mathbb{N}$ and fixed $X_n \subseteq \mathbb{Z}/n\mathbb{Z}$, $n \in A$, for which the set B in (1.1) has no logarithmic density. More precisely,*

$$\liminf_{x\to\infty} L_B(x) \le \frac{177}{200} \quad \textit{and} \quad \limsup_{x\to\infty} L_B(x) \ge \frac{49}{50}.$$

Remark 1.2 (Strict and inclusive activation). Some sources use $n \le m$ in place of $n < m$. Let $B_\ge$ and $B_>$ denote the corresponding survivor sets. Plainly $B_\ge \subseteq B_>$. If $m \in B_> \setminus B_\ge$, then $m \in A$ and $0 \in X_m$. Moreover, $B_> \setminus B_\ge$ is primitive: if $d < e$ were two of its elements and $d \mid e$, then the active modulus $d < e$, together with $0 \in X_d$, would exclude e from $B_>$. Behrend's theorem gives

$$\sum_{\substack{d<x \\ d\in B_> \setminus B_\geq}} \frac{1}{d} = o(\log x)$$

for every primitive set [3]. Thus the two conventions are equivalent for the existence and value of logarithmic density.

The zero-residue case $X_n = \{0\}$ has an affirmative answer by the Davenport–Erdős theorem [5]; a later elementary proof was given in [6]. Besicovitch showed that natural density may fail to exist in that setting [2]. More recently, Araújo proved, in the inclusive activation convention, that a delayed multi-residue sieve has a natural density under the summability hypothesis

$$\sum_{n\in A} \frac{|X_n|}{n} < \infty$$

[1, Theorem 3.25]. The same conclusion holds for the strict convention under this hypothesis. Indeed, the two survivor sets differ only on a subset of

$$D := \{n \in A : 0 \in X_n\},$$

and $\sum_{n\in D} 1/n < \infty$. To see directly that D has natural density zero, fix $M < N$ and note that

$$\frac{|D \cap [1, N]|}{N} \leq \frac{|D \cap [1, M]|}{N} + \sum_{\substack{n\in D \\ n>M}} \frac{1}{n};$$

first let $N \to \infty$, and then $M \to \infty$. Our construction necessarily lies outside that summable regime: every installed scale contributes a fixed positive amount to the sum; see Remark 4.1.

The proof has two parts. Section 3 constructs a finite block at scale $Q = 2^j$. It deletes at least $3Q/8$ integers from $[11Q/10, 19Q/10]$, even though the periodic union of the responsible residue classes has measure $e^{-\Omega(\sqrt{j})}$. Section 4 assembles long runs of these blocks. Between consecutive runs we leave a gap long enough for the finite past periodic system to recover. Later moduli lie beyond each earlier recovery cutoff, so the activation condition prevents them from changing the earlier averages.

2 Periodic sets and profinite notation

Let $\widehat{\mathbb{Z}}$ be the profinite completion of $\mathbb{Z}$, equipped with Haar probability measure μ, and let $\pi_q : \widehat{\mathbb{Z}} \to \mathbb{Z}/q\mathbb{Z}$ be the canonical projection. For $\bar{a} \in \mathbb{Z}/q\mathbb{Z}$, write

$$[\bar{a}]_q := \pi_q^{-1}(\bar{a}).$$

If $a \in \mathbb{Z}$, then $[a]_q$ means $[a \bmod q]_q$. Thus $\mu([\bar{a}]_q) = 1/q$. The trace on $\mathbb{Z}$ of a finite union of such cylinders is periodic, and its natural density equals its Haar measure. We shall also use, for every $N \geq 1$, the elementary bound

$$H_N := \sum_{m=1}^{N} \frac{1}{m} \leq 1 + \log N.$$

Lemma 2.1 (Periodic recovery). *Let $\mathcal{F}$ be a finite family of pairs (q, Y_q), with $Y_q \subseteq \mathbb{Z}/q\mathbb{Z}$, and set*

$$U_{\mathcal{F}} := \bigcup_{(q,Y_q)\in\mathcal{F}} \bigcup_{a\in Y_q} [a]_q$$

and

$$B_{\mathcal{F}} := \{m \in \mathbb{N} : m \bmod q \notin Y_q \text{ whenever } (q, Y_q) \in \mathcal{F} \text{ and } q < m\}.$$

Then

$$\sum_{\substack{m<x \\ m \in B_{\mathcal{F}}}} \frac{1}{m} = (1 - \mu(U_{\mathcal{F}})) \log x + O_{\mathcal{F}}(1) \qquad (x \to \infty).$$

Proof. If $\mathcal{F}$ is empty, the assertion is the standard estimate $H_{\lceil x \rceil - 1} = \log x + O(1)$. Suppose henceforth that $\mathcal{F}$ is nonempty. Let

$$q_{\max} = \max_{(q, Y_q) \in \mathcal{F}} q, \qquad L = \operatorname{lcm}\{q : (q, Y_q) \in \mathcal{F}\}.$$

For every $m > q_{\max}$, all the moduli in $\mathcal{F}$ are active. Therefore $B_{\mathcal{F}}$ differs in only finitely many places from the integer trace of $\widehat{\mathbb{Z}} \setminus U_{\mathcal{F}}$.

That trace is a union of, say, r residue classes modulo L. Its Haar measure is $r/L = 1 - \mu(U_{\mathcal{F}})$. For a fixed representative $a \in \{1, \dots, L\}$, with $a = L$ representing the zero class,

$$\sum_{\substack{m<x \\ m \equiv a \pmod{L}}} \frac{1}{m} = \sum_{\substack{\ell \geq 0 \\ a + \ell L < x}} \frac{1}{a + \ell L} = \frac{1}{L} \log x + O_L(1).$$

The last estimate follows, for example, by comparing the decreasing function $t \mapsto (a + Lt)^{-1}$ with its integral. Summing over the r occupied residue classes and restoring the finite initial discrepancy proves the lemma. $\square$

3 A finite dyadic deletion block

For an even positive integer k, define the central family

$$\mathcal{S}_k := \left\{ S \subseteq \{1, \dots, k\} : \left| |S| - \frac{k}{2} \right| \leq \sqrt{k} \right\}.$$

Lemma 3.1 (Arithmetic skeleton)**.** *For all sufficiently large integers j, put*

$$Q = 2^j, \qquad k = 2 \left\lfloor \frac{\sqrt{j}}{8} \right\rfloor.$$

There are distinct primes $p_1, \dots, p_k$ and, for every $S \in \mathcal{S}_k$, a positive integer q_S such that

$$\frac{19Q}{20} \leq q_S \leq \frac{21Q}{20}, \qquad p_i \mid q_S \iff i \in S. \tag{3.1}$$

The integers q_S are pairwise distinct. Moreover, if

$$J = [11Q/10, 19Q/10] \cap \mathbb{Z},$$

then $J \subset (q_S, 2q_S]$ and the diameter of J is smaller than q_S, for every $S \in \mathcal{S}_k$.

Proof. Bertrand's postulate [9] supplies a prime p_i satisfying

$$4^{k+i} < p_i < 2 \cdot 4^{k+i} \qquad (1 \leq i \leq k).$$

These intervals are pairwise disjoint, so the primes are distinct. Set $P = \prod_{i=1}^{k} p_i$. Since $p_i < 2^{2(k+i)+1}$,

$$\log_2 P < \sum_{i=1}^{k} (2(k+i) + 1) = 3k^2 + 2k.$$

The choice of k gives $k \le \sqrt{j}/4$, and hence

$$\frac{P^2}{Q} \le 2^{6k^2+4k-j} \le 2^{-5j/8+\sqrt{j}} \longrightarrow 0 \qquad (j \to \infty).$$

For $S \in \mathcal{S}_k$, let $d_S = \prod_{i \in S} p_i$. Choose R_S to be a closest integer to Q/d_S in the progression $1 + P\mathbb{Z}$, and put $q_S = d_S R_S$. This choice satisfies

$$\left| R_S - \frac{Q}{d_S} \right| \le \frac{P}{2}, \qquad |q_S - Q| \le \frac{d_S P}{2} \le \frac{P^2}{2}.$$

It remains to check that R_S is positive. Since $d_S \le P$ and $Q/P^2 \to \infty$, for large j we have

$$\frac{Q}{d_S} \ge \frac{Q}{P} > P.$$

Thus $R_S \ge Q/d_S - P/2 > P/2 > 0$. The estimate $P^2/Q \to 0$ now gives the first pair of inequalities in (3.1).

If $i \in S$, then $p_i \mid d_S$, so $p_i \mid q_S$. If $i \notin S$, then $p_i \nmid d_S$ and $R_S \equiv 1 \pmod{p_i}$, so $p_i \nmid q_S$. This proves the second part of (3.1); it also shows that distinct sets S give distinct integers q_S.

Finally,

$$q_S \le \frac{21Q}{20} < \frac{11Q}{10} \quad \text{and} \quad 2q_S \ge \frac{19Q}{10}.$$

Consequently $J \subset (q_S, 2q_S]$. Its diameter is at most $4Q/5$, which is smaller than $q_S \ge 19Q/20$. □

Fix j large enough for Lemma 3.1, and retain its notation. For every $1 \le i \le k$ and $b \in \mathbb{Z}/p_i\mathbb{Z}$, choose an independent Bernoulli random variable $\varepsilon_i(b)$ with parameter $1/2$. For $m \in \mathbb{Z}$, set

$$K(m) := \{i : \varepsilon_i(m \bmod p_i) = 1\}.$$

Let $\mathbb{P}$ and $\mathbb{E}$ denote probability and expectation with respect to this finite family of random labels. Define the random endpoint set and its periodic footprint by

$$E := \{m \in J : K(m) \in \mathcal{S}_k\}, \qquad q_m := q_{K(m)} \quad (m \in E), \qquad U := \bigcup_{m \in E} [m]_{q_m}.$$

Lemma 3.2 (Endpoint abundance)**.** *For all sufficiently large j,*

$$\mathbb{P}\left(|E| < \frac{|J|}{2}\right) \le \exp\left(-\frac{4^k}{8k}\right).$$

Proof. For a fixed m, the variable $|K(m)|$ has distribution $\mathrm{Bin}(k, 1/2)$, with mean $k/2$ and variance $k/4$. Chebyshev's inequality therefore gives

$$\mathbb{P}(K(m) \notin \mathcal{S}_k) = \mathbb{P}\left(\left||K(m)| - \frac{k}{2}\right| > \sqrt{k}\right) \le \frac{1}{4}.$$

For $Z = |E|/|J|$, averaging this estimate over $m \in J$ gives $\mathbb{E}Z \ge 3/4$.

Changing one variable $\varepsilon_i(b)$ affects only those $m \in J$ with $m \equiv b \pmod{p_i}$. There are at most $|J|/p_i + 1$ such integers. The proof of Lemma 3.1 gives $p_i \le P = o(Q^{1/2})$, whereas $|J| \asymp Q$; hence $|J| \ge p_i$ for large j. The corresponding bounded-difference constant for Z is therefore at most $2/p_i$. For each i, there are exactly p_i variables $\varepsilon_i(b)$. Hence the sum of the squared bounded-difference constants is at most

$$\sum_{i=1}^{k} p_i \left(\frac{2}{p_i}\right)^2 = 4 \sum_{i=1}^{k} \frac{1}{p_i} < \frac{k}{4^k},$$

because every $p_i > 4^{k+1}$. McDiarmid's bounded-differences inequality [11], applied with downward deviation $1/4$, now yields

$$\mathbb{P}(Z < 1/2) \le \exp\left(-\frac{2(1/4)^2}{k/4^k}\right) = \exp\left(-\frac{4^k}{8k}\right).$$

All the independent variables here form a finite family, and the displayed sum verifies the bounded-difference hypothesis. $\square$

Lemma 3.3 (Small periodic footprint)**.** *For all sufficiently large j,*

$$\mathbb{P}(\mu(U) > e^{-k/100}) \le 3e^{-k/100}.$$

Proof. Fix $\omega \in \widehat{\mathbb{Z}}$. For $S \in \mathcal{S}_k$, let $a_S(\omega) \in \{1, \dots, q_S\}$ represent ω mod q_S, with the zero residue represented by q_S, and set

$$m_S(\omega) := q_S + a_S(\omega) \in (q_S, 2q_S].$$

Because $J \subset (q_S, 2q_S]$ and has diameter smaller than q_S, each residue class modulo q_S has at most one representative in J. It follows directly from the definitions that

$$\omega \in U \quad \iff \quad \text{there is an } S \in \mathcal{S}_k \text{ for which } m_S(\omega) \in J \text{ and } K(m_S(\omega)) = S. \tag{3.2}$$

For $S \in \mathcal{S}_k$ and $i \notin S$, define the collision set

$$C_{S,i} := \{\omega \in \widehat{\mathbb{Z}} : m_S(\omega) \equiv \omega \pmod{p_i}\}.$$

By (3.1), $(q_S, p_i) = 1$. Conditional on any residue modulo q_S, the displayed congruence prescribes exactly one residue modulo p_i. The Chinese remainder theorem, applied to the uniform residue modulo $q_S p_i$, therefore gives $\mu(C_{S,i}) = 1/p_i$. The set

$$C := \bigcup_{\substack{S \in \mathcal{S}_k \\ i \notin S}} C_{S,i},$$

depends only on the arithmetic skeleton and is therefore independent of all the random labels. The union bound and $p_i > 4^{k+1}$ give

$$\mu(C) \le |\mathcal{S}_k| \sum_{i=1}^{k} \frac{1}{p_i} \le 2^k \frac{k}{4^{k+1}} = \frac{k}{2^{k+2}}. \tag{3.3}$$

Now fix $\omega \notin C$. Condition on the k anchor bits $\varepsilon_i(\omega \bmod p_i)$, and write

$$T := \{i : \varepsilon_i(\omega \bmod p_i) = 1\}.$$

If $i \in S$, then $p_i \mid q_S$, and hence $m_S(\omega) \equiv \omega \pmod{p_i}$. Therefore $K(m_S(\omega)) = S$ forces $S \subseteq T$. If $i \notin S$, the assumption $\omega \notin C_{S,i}$ says that the residue queried at $m_S(\omega) \bmod p_i$ differs from the anchor residue $\omega \bmod p_i$. For a fixed S, these $k - |S|$ queried variables have different first coordinates i; each is also distinct from every conditioned anchor variable. They are therefore mutually independent and remain unbiased after the conditioning. Consequently, conditional on the anchor bits, the probability that S contributes to (3.2) is zero unless $S \subseteq T$, and in all cases it is at most

$$2^{-(k-|S|)}. \tag{3.4}$$

The requirement $m_S(\omega) \in J$ can only decrease this probability.

Unconditionally, $|T| \sim \mathrm{Bin}(k, 1/2)$. Hoeffding's inequality [10] gives

$$\mathbb{P}(|T| > 3k/5) \le \exp\left(-\frac{2(k/10)^2}{k}\right) = e^{-k/50}. \tag{3.5}$$

Suppose that $|T| \le 3k/5$ and that $S \in \mathcal{S}_k$ is contained in T. The map $S \mapsto T \setminus S$ is injective, and

$$|T \setminus S| \le \frac{3k}{5} - \left(\frac{k}{2} - \sqrt{k}\right) = \frac{k}{10} + \sqrt{k}.$$

Let $N_k = \lfloor 3k/5 \rfloor$ and $M_k = \lfloor k/10 + \sqrt{k} \rfloor$. Padding T to a set with N_k elements is possible because the integer $|T|$ is at most $\lfloor 3k/5 \rfloor = N_k$. Since $S \mapsto T \setminus S$ is injective, the number of possible candidates is at most

$$\sum_{\ell=0}^{M_k} \binom{N_k}{\ell}.$$

We record the entropy estimate used to bound this sum. If $0 < \rho \le 1/2$ and $M = \lfloor \rho N \rfloor$, then

$$\sum_{\ell=0}^{M} \binom{N}{\ell} \le \exp(NH(\rho)), \qquad H(\rho) = -\rho \log \rho - (1-\rho)\log(1-\rho). \tag{3.6}$$

Indeed, for $\ell \le \rho N$, the quantity $\rho^{\ell}(1-\rho)^{N-\ell}$ is at least $\rho^{\rho N}(1-\rho)^{(1-\rho)N} = e^{-NH(\rho)}$. Multiplication by this lower bound and comparison with the complete binomial expansion proves (3.6).

Here $M_k/N_k \to 1/6$ and $N_k/k \to 3/5$. In particular, $0 < M_k/N_k < 1/2$ for sufficiently large k, so (3.6) applies with $\rho = M_k/N_k$. The numerical margin needed below follows from

$$\frac{3}{5}H(1/6) = \frac{1}{10}\log 6 + \frac{1}{2}\log(6/5) < \frac{1}{10}\frac{9}{5} + \frac{1}{2}\frac{1}{5} = \frac{7}{25} = 0.28.$$

The inequality $\log(6/5) < 1/5$ is the standard $\log(1+t) < t$ with $t = 1/5$. Also $\log 6 < 9/5$, because the first seven terms of the positive power series for $e^{9/5}$ have sum greater than 6. By continuity of H, for all sufficiently large k,

$$\#\{S \in \mathcal{S}_k : S \subseteq T\} \le e^{0.29k}. \tag{3.7}$$

For the same range of k, $\sqrt{k} \le k/100$, so every $S \in \mathcal{S}_k$ satisfies $k - |S| \ge 0.49k$. The positive series

$$\log 2 = 2\sum_{r=0}^{\infty} \frac{1}{(2r+1)3^{2r+1}} > \frac{56}{81}$$

shows that $0.49 \log 2 > 0.33$. Hence (3.4) is at most $e^{-0.33k}$. For every fixed realization of the anchor bits with $|T| \le 3k/5$, the conditional union bound, together with (3.4) and (3.7), gives

$$\mathbb{P}(\omega \in U \mid \varepsilon_i(\omega \bmod p_i),\ 1 \le i \le k) \le e^{0.29k}e^{-0.33k} = e^{-0.04k}.$$

Averaging over the anchor bits therefore yields, uniformly for $\omega \notin C$,

$$\mathbb{P}(\omega \in U,\ |T| \le 3k/5) \le e^{-0.04k}.$$

The indicator of U is nonnegative, so Tonelli's theorem gives

$$\mathbb{E}_{\varepsilon}\mu(U) = \int_{\widehat{\mathbb{Z}}} \mathbb{P}_{\varepsilon}(\omega \in U)\, d\mu(\omega).$$

Split the integral over C and $\widehat{\mathbb{Z}} \setminus C$, and use (3.3), (3.5), and the preceding estimate. For all sufficiently large k,

$$\mathbb{E}\mu(U) \le \frac{k}{2^{k+2}} + e^{-k/50} + e^{-0.04k} \le 3e^{-k/50}.$$

Markov's inequality, with threshold $e^{-k/100}$, now gives

$$\mathbb{P}(\mu(U) > e^{-k/100}) \le e^{k/100}\mathbb{E}\mu(U) \le 3e^{-k/100},$$

as required. □

Lemma 3.4 (Finite block lemma)**.** *For every sufficiently large integer j, put*

$$Q = 2^j, \qquad k_j = 2\left\lfloor \frac{\sqrt{j}}{8} \right\rfloor, \qquad \eta_j = e^{-k_j/100}.$$

There are an endpoint set $E_j \subseteq \mathbb{N}$, a modulus $q_{j,m}$ for each $m \in E_j$, and a periodic cylinder union

$$U_j = \bigcup_{m\in E_j} [m]_{q_{j,m}}$$

such that

$$\begin{gathered} E_j \subseteq [11Q/10, 19Q/10] \cap \mathbb{Z}, \qquad |E_j| \geq 3Q/8, \\ \frac{19Q}{20} \leq q_{j,m} \leq \frac{21Q}{20} < m \quad (m \in E_j), \qquad \mu(U_j) \leq \eta_j. \end{gathered} \tag{3.8}$$

Moreover, $\sum_j \eta_j < \infty$.

Proof. Lemmas 3.2 and 3.3 show that the probability of failure of either

$$|E| \geq |J|/2, \qquad \mu(U) \leq e^{-k/100}$$

is at most

$$\exp\left(-\frac{4^k}{8k}\right) + 3e^{-k/100},$$

which is smaller than 1 for all sufficiently large k. Fix a labelling for which both conclusions hold, and use it to define $E_j, q_{j,m}$, and U_j. Since $|J| \geq 4Q/5 - 1$, we have $|E_j| \geq 3Q/8$ for large Q. Lemma 3.1 gives the modulus bounds, and the strict inequality $q_{j,m} < m$ follows from

$$q_{j,m} \leq 21Q/20 < 11Q/10 \leq m.$$

This proves (3.8).

Finally,

$$k_j \geq \frac{\sqrt{j}}{4} - 2, \qquad \eta_j \leq e^{1/50} e^{-\sqrt{j}/400}.$$

To see that the last majorant is summable, group the integers $r^2 \leq j < (r+1)^2$; the contribution of the r-th group is at most $(2r+1)e^{-r/400}$, up to the fixed factor $e^{1/50}$. The resulting series converges. □

4 The global construction

Set $\epsilon = 1/100$. Choose j_0 sufficiently large that Lemma 3.4 holds for every $j \geq j_0$ and

$$\sum_{j\geq j_0} \eta_j < \epsilon, \qquad \frac{1}{(2j_0+1)\log 2} < \epsilon. \tag{4.1}$$

For each $j \geq j_0$, fix once and for all a block $(E_j, (q_{j,m})_{m\in E_j}, U_j)$ supplied by that lemma.

4.1 Choosing the epochs

We recursively choose integers $a_1 < a_2 < \cdots$ and install the blocks with indices in

$$I_t := \{a_t, a_t + 1, \ldots, 2a_t\}$$

during epoch t. Suppose that the first $t-1$ epochs have already been chosen. For every modulus q occurring in those epochs, put

$$Y_q^{(t-1)} := \{m \bmod q : s < t,\ j \in I_s,\ m \in E_j,\ q_{j,m} = q\},$$

and let $\mathcal{F}_{t-1}$ be the finite family of all pairs $(q, Y_q^{(t-1)})$ with $Y_q^{(t-1)} \neq \varnothing$. For $t = 1$, take $\mathcal{F}_0 = \varnothing$. Grouping equal moduli changes neither the survivor set nor the union of cylinders, and hence

$$U_{\mathcal{F}_{t-1}} = V_{t-1} := \bigcup_{\substack{s<t \\ j \in I_s}} U_j.$$

By induction, the already chosen intervals I_s, $s < t$, are pairwise disjoint subsets of $\{j : j \geq j_0\}$. The union bound and (4.1) therefore imply

$$\mu(V_{t-1}) \leq \sum_{\substack{s<t \\ j \in I_s}} \eta_j < \epsilon.$$

By Lemma 2.1, the logarithmic averages of $B_{\mathcal{F}_{t-1}}$ converge to $1 - \mu(V_{t-1}) > 1 - \epsilon$. We may therefore choose a_t so large that

$$\begin{gathered} a_t \geq j_0, \qquad a_t > 2a_{t-1} + 3 \quad (t \geq 2), \\ \frac{1}{\log x_t} \sum_{\substack{m < x_t \\ m \in B_{\mathcal{F}_{t-1}}}} \frac{1}{m} \geq 1 - 2\epsilon = \frac{49}{50}, \qquad x_t := 2^{a_t - 1}. \end{gathered} \tag{4.2}$$

Indeed, the convergence just noted makes the last inequality true for all sufficiently large x, so it remains true after restricting x to the unbounded sequence of dyadic numbers 2^{a-1}. This completes the recursion. In particular, the epochs are pairwise disjoint and $a_t \to \infty$.

4.2 The fixed congruence system

Let

$$\mathcal{I} := \bigcup_{t \geq 1} I_t$$

be the set of installed scales. Modulus ranges at distinct scales are disjoint, since

$$\frac{21}{20} 2^j < \frac{19}{20} 2^{j+1}. \tag{4.3}$$

For each modulus q appearing in an installed block, define

$$X_q := \{m \bmod q : j \in \mathcal{I},\ m \in E_j,\ q_{j,m} = q\} \subseteq \mathbb{Z}/q\mathbb{Z},$$

and let A be the set of those moduli. Repeated occurrences of the same modulus inside one block are thereby grouped into a single residue set. Equation (4.3) shows that a modulus cannot occur at two different scales.

The family is infinite. Indeed, $|E_j| \geq 3 \cdot 2^j/8 > 0$ at every installed scale, so every such scale supplies a modulus; by (4.3), these moduli are distinct as the scale varies. Let $B = B(A, X)$ be the survivor set defined in (1.1). Once the sequence (a_t) has been chosen, both A and every X_q are fixed; they do not depend on the averaging parameter.

Remark 4.1 (Failure of the summability hypothesis). Within a fixed block, two distinct endpoints assigned the same modulus represent distinct residue classes: their distance is smaller than that modulus by Lemma 3.1. Together with (4.3), this gives, for every installed j,

$$\sum_{\substack{q \in A \\ 19 \cdot 2^j/20 \leq q \leq 21 \cdot 2^j/20}} \frac{|X_q|}{q} = \sum_{m \in E_j} \frac{1}{q_{j,m}} \geq \frac{|E_j|}{21 \cdot 2^j/20} \geq \frac{5}{14}.$$

There are infinitely many installed scales, and therefore $\sum_{q \in A} |X_q|/q = \infty$. Thus the construction does not overlap the summable positive result cited in the introduction.

4.3 Recovery cutoffs

At $x_t = 2^{a_t-1}$, consider a modulus belonging to an epoch $s \ge t$. Its scale j satisfies $j \ge a_s \ge a_t$, so

$$q \ge \frac{19}{20} 2^j \ge \frac{19}{20} 2^{a_t} > 2^{a_t-1} = x_t.$$

Such a modulus is not active at any integer $m < x_t$. Consequently, below x_t the final survivor B agrees exactly with the finite-past survivor $B_{\mathcal{F}_{t-1}}$: by (4.3), no later scale can add a new residue class to a modulus already present in $\mathcal{F}_{t-1}$, and all genuinely future moduli are inactive below x_t. By (4.2),

$$L_B(x_t) \ge \frac{49}{50}.$$

Since $x_t \to \infty$, it follows that

$$\limsup_{x\to\infty} L_B(x) \ge \frac{49}{50}. \tag{4.4}$$

4.4 Deletion cutoffs

Put $y_t = 2^{2a_t+1}$. If $j \in I_t$ and $m \in E_j$, then

$$m \le \frac{19}{10} 2^j \le \frac{19}{10} 2^{2a_t} < 2^{2a_t+1} = y_t.$$

Moreover, Lemma 3.4 gives

$$q_{j,m} \le \frac{21}{20} 2^j < \frac{11}{10} 2^j \le m.$$

By the definition of $X_{q_{j,m}}$, one has $m \bmod q_{j,m} \in X_{q_{j,m}}$. Since $q_{j,m} < m$, this modulus is active at m, and consequently $m \notin B$. Thus every point of E_j is absent from $B \cap [1, y_t)$.

The endpoint intervals at consecutive scales are disjoint:

$$\frac{19}{10} 2^j < \frac{11}{10} 2^{j+1}.$$

Hence the deleted endpoints supplied by different $j \in I_t$ are distinct. At a single scale, (3.8) gives

$$\sum_{m \in E_j} \frac{1}{m} \ge \frac{|E_j|}{19 \cdot 2^j/10} \ge \frac{3 \cdot 2^j/8}{19 \cdot 2^j/10} = \frac{15}{76}.$$

Because I_t contains $a_t + 1$ scales, we obtain

$$L_B(y_t) \le \frac{H_{y_t-1}}{\log y_t} - \frac{15}{76} \frac{a_t+1}{(2a_t+1)\log 2}. \tag{4.5}$$

Since $a_t \ge j_0$, the harmonic-number bound from Section 2 and (4.1) show that

$$\frac{H_{y_t-1}}{\log y_t} \le 1 + \frac{1}{\log y_t} < 1 + \epsilon.$$

Also $(a_t+1)/(2a_t+1) > 1/2$, and

$$\frac{15}{152 \log 2} > \frac{1}{8}.$$

For completeness, $\log 2 < 3/4$, since the first four terms in the positive series for $e^{3/4}$ already sum to more than 2; hence the left side is greater than $15/114 = 5/38 > 1/8$. Substitution into (4.5) gives

$$L_B(y_t) < 1 + \frac{1}{100} - \frac{1}{8} = \frac{177}{200}.$$

Since $y_t \to \infty$,

$$\liminf_{x\to\infty} L_B(x) \le \frac{177}{200}. \tag{4.6}$$

Equations (4.4) and (4.6) prove Theorem 1.1.

5 Concluding remarks

The construction separates two effects that coincide for many classical sieves. Within an installed scale, the active tails delete a fixed amount of harmonic mass. Globally, however, the completed periodic footprint of that scale has summably small Haar measure. Long deletion epochs exploit the first fact, while the recovery cutoffs exploit the second.

Erdős Problem 25 is the singleton special case of Problem 486, in which $|X_q| = 1$ for every $q \in A$ [4]. The present argument does not settle that problem. Indeed, the finite block in Section 3 groups many endpoint residues into the same set X_q; this multi-residue feature makes it possible to obtain a fixed local harmonic deletion while keeping the completed periodic footprint small.

References

[1] F. Araújo, Sarnak's program for Erdős sieves. Part I: topological dynamics and light tails, preprint, 2026, arXiv:2602.24031.

[2] A. S. Besicovitch, On the density of certain sequences of integers, *Math. Ann.* **110** (1934), 336–341, doi:10.1007/BF01448032.

[3] F. Behrend, On sequences of numbers not divisible one by another, *J. London Math. Soc.* **10** (1935), 42–44, doi:10.1112/jlms/s1-10.37.42.

[4] T. F. Bloom, Erdős Problem 25, `https://www.erdosproblems.com/25`, accessed July 15, 2026.

[5] H. Davenport and P. Erdős, On sequences of positive integers, *Acta Arith.* **2** (1936), 147–151, doi:10.4064/aa-2-1-147-151.

[6] H. Davenport and P. Erdős, On sequences of positive integers, *J. Indian Math. Soc. (N.S.)* **15** (1951), 19–24, doi:10.18311/JIMS/1951/17063.

[7] P. Erdős, On the distribution function of additive arithmetical functions and on some related problems, *Rend. Sem. Mat. Fis. Milano* **27** (1957), 45–49, doi:10.1007/BF02922565.

[8] P. Erdős, Some unsolved problems, *Publ. Math. Inst. Hung. Acad. Sci., Ser. A* **6** (1961), 221–254.

[9] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, sixth ed., revised by D. R. Heath-Brown and J. H. Silverman, Oxford University Press, Oxford, 2008.

[10] W. Hoeffding, Probability inequalities for sums of bounded random variables, *J. Amer. Statist. Assoc.* **58** (1963), no. 301, 13–30, doi:10.1080/01621459.1963.10500830.

[11] C. McDiarmid, On the method of bounded differences, in *Surveys in Combinatorics, 1989*, London Math. Soc. Lecture Note Ser. **141**, Cambridge University Press, Cambridge, 1989, pp. 148–188, doi:10.1017/CBO9781107359949.008.