Erdős Problem #489:
An extended-limit theorem, a finiteness criterion, and broad positive results

Abstract

Let $A \subseteq \{2, 3, \dots\}$ and write

$$\mathcal{F}_A := \{n \geq 1 : \ a \nmid n \text{ for every } a \in A\} = \{b_1 < b_2 < \cdots\}.$$

In [6, p. 237], Erdős asked whether, under the sparseness hypothesis $|A \cap [1, x]| = o(x^{1/2})$, the limit

$$\lim_{x\to\infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2$$

exists and is finite. The model case $A = \{p^2 : p \text{ prime}\}$ is the sequence of squarefree numbers, for which Erdős had already proved existence of the second moment in [5].

This note proves the following general reduction. If

$$q_A(h) := \lim_{x\to\infty} \frac{1}{x} \#\{n \leq x : [n, n+h-1] \cap \mathcal{F}_A = \varnothing\} \qquad (h \geq 1),$$

then each $q_A(h)$ exists, and

$$\lim_{x\to\infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2 = 1 + 2 \sum_{h \geq 1} q_A(h) \in [0, +\infty].$$

Thus the normalized second moment always has an *extended real* limit, and the original finiteness question is equivalent to the summability of the empty-window frequencies $q_A(h)$.

We also record a broad finiteness theorem from the existing literature: if A is pairwise coprime and thin and the multiplicative semigroup $\langle A \rangle$ has index $\alpha \in (0, 1)$ in the sense of Gorodetsky–Mangerel–Rodgers, then $q_A(h) \ll h^{-(2-\alpha)}$, hence the second moment is finite. This covers, in particular, the squarefree example and many other classical $\mathcal{B}$-free systems. The note ends with a literature survey centered on Problem #489.

1 The problem and the main theorem

Erdős asked the following question in [6, p. 237]; it is also listed as Problem #489 on the Erdős problems site [3]. Let $A \subseteq \mathbb{N}$ satisfy

$$A(x) := |A \cap [1, x]| = o(x^{1/2}),$$

and let $\mathcal{F}_A = \{b_1 < b_2 < \cdots\}$ be the set of positive integers divisible by no element of A. Is it true that

$$\lim_{x\to\infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2$$

exists, and is finite?

Since $A \subseteq \{2, 3, \dots\}$, we always have $1 \in \mathcal{F}_A$, so $b_1 = 1$.

For $h \geq 1$, define the empty-window counting function

$$E_h(x) := \#\{n \leq x : [n, n+h-1] \cap \mathcal{F}_A = \varnothing\}.$$

Our first main result is the following.

Theorem 1. *Assume* $A(x) = o(x^{1/2})$*, and let* $\mathcal{F}_A = \{b_1 < b_2 < \cdots\}$.

1. *For every fixed* $h \geq 1$*, the limit*
$$q_A(h) := \lim_{x \to \infty} \frac{E_h(x)}{x}$$
exists.

2. *The limit*
$$L_A := \lim_{x \to \infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2$$
exists in the extended interval $[0, +\infty]$ *and is given by*
$$L_A = 1 + 2 \sum_{h \geq 1} q_A(h).$$
Consequently,
$$L_A < \infty \iff \sum_{h \geq 1} q_A(h) < \infty.$$

Thus the existence part of Problem #489 can be settled cleanly once one allows the value $+\infty$, and the remaining issue is to recognize when the empty-window tail is summable.

We also isolate a broad subclass in which finiteness follows from the current literature. Given a set $A \subseteq \{2, 3, \dots\}$, let $\langle A \rangle$ denote the multiplicative semigroup generated by A.

Corollary 2. *Assume that* A *is pairwise coprime and thin, and that*
$$N_{\langle A \rangle}(x) = x^{\alpha + o(1)} \qquad (x \to \infty)$$
for some $\alpha \in (0, 1)$*. Then*
$$q_A(h) \ll h^{-(2-\alpha)} \qquad (h \geq 1),$$
so in particular
$$\lim_{x \to \infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2 < \infty.$$

The proof is a one-line consequence of a theorem of Gorodetsky–Mangerel–Rodgers once Theorem 1 is available; see Section 5.

2 Thinness, density, and fixed local patterns

The sparseness condition immediately implies that A is thin.

Lemma 3. *If* $A(x) = o(x^{1/2})$*, then*
$$\sum_{a \in A} \frac{1}{a} < \infty.$$

Proof. Write $A = \{a_1 < a_2 < \cdots\}$. From $A(a_n) = n = o(a_n^{1/2})$ we obtain $a_n/n^2 \to \infty$. In particular, $a_n \geq n^2$ for all sufficiently large n, and hence
$$\sum_{a \in A} \frac{1}{a} = \sum_{n \geq 1} \frac{1}{a_n} \ll \sum_{n \geq 1} \frac{1}{n^2} < \infty. \qquad \square$$

Let
$$\mathcal{M}(A) := \{n \geq 1 : \exists a \in A \text{ with } a \mid n\}$$
be the set of multiples of A. Then $\mathcal{F}_A = \mathbb{N} \setminus \mathcal{M}(A)$. The following classical fact is exactly what is needed to guarantee that $\mathcal{F}_A$ has positive density.

Proposition 4 (classical)**.** *If A is thin, then $\mathcal{M}(A)$ has an asymptotic density. If moreover $1 \notin A$, then*

$$d(\mathcal{M}(A)) < 1,$$

so $\mathcal{F}_A$ has positive asymptotic density.

Proof. This is the theorem of Erdős on thin sets of multiples together with the Heilbronn–Rohrbach inequality; see [14, Lemma 5] and [10, Chapter 0]. ☐

Corollary 5. *Under the hypothesis $A(x) = o(x^{1/2})$, the density*

$$\rho_A := d(\mathcal{F}_A)$$

exists and satisfies $\rho_A > 0$.

We now prove existence of all fixed local pattern frequencies by periodic truncation. For $Y \geq 1$, let

$$A_{\leq Y} := A \cap [1, Y], \qquad \mathcal{F}_{A,\leq Y} := \{n \geq 1 : \ a \nmid n \text{ for every } a \in A_{\leq Y}\}.$$

Since $A_{\leq Y}$ is finite, the indicator $\mathbf{1}_{\mathcal{F}_{A,\leq Y}}$ is periodic.

Lemma 6. *Fix $m \geq 1$ and a binary word $w = (w_0, \dots, w_{m-1}) \in \{0,1\}^m$. Then the density*

$$\delta_A(w) := \lim_{x\to\infty} \frac{1}{x} \#\{n \leq x : (\mathbf{1}_{\mathcal{F}_A}(n), \dots, \mathbf{1}_{\mathcal{F}_A}(n+m-1)) = w\}$$

exists.

Proof. For each Y, periodicity gives existence of the density

$$\delta_Y(w) := \lim_{x\to\infty} \frac{1}{x} \#\{n \leq x : (\mathbf{1}_{\mathcal{F}_{A,\leq Y}}(n), \dots, \mathbf{1}_{\mathcal{F}_{A,\leq Y}}(n+m-1)) = w\}.$$

If the length-m words attached to n differ for $\mathcal{F}_A$ and $\mathcal{F}_{A,\leq Y}$, then for some $0 \leq j < m$ the number $n + j$ is divisible by an element $a \in A$ with $a > Y$. Therefore

$$\begin{aligned}
\limsup_{x\to\infty} \frac{1}{x} \#\Big\{n \leq x : (\mathbf{1}_{\mathcal{F}_A}(n), \dots, \mathbf{1}_{\mathcal{F}_A}(n+m-1)) \neq (\mathbf{1}_{\mathcal{F}_{A,\leq Y}}(n), \dots, \mathbf{1}_{\mathcal{F}_{A,\leq Y}}(n+m-1))\Big\} \\
\leq \sum_{j=0}^{m-1} \sum_{\substack{a\in A\\ a>Y}} \limsup_{x\to\infty} \frac{1}{x} \#\{n \leq x : a \mid n+j\} \\
\leq m \sum_{\substack{a\in A\\ a>Y}} \frac{1}{a}.
\end{aligned}$$

By Lemma 3, the tail on the right tends to 0 as $Y \to \infty$. Hence $\delta_Y(w)$ is a Cauchy net as $Y \to \infty$, so it converges to some value $\delta_A(w)$. The same estimate shows that the word frequency for $\mathcal{F}_A$ differs from $\delta_Y(w)$ by at most $m \sum_{a>Y,\, a\in A} 1/a + o(1)$, whence the required density exists and equals $\delta_A(w)$. ☐

Corollary 7. *For each $h \geq 1$, the empty-window density*

$$q_A(h) = \lim_{x\to\infty} \frac{E_h(x)}{x}$$

exists.

Proof. Take in Lemma 6 the word $w = (0, \dots, 0)$ of length h. ☐

3 A gap lemma and an exact identity

Let

$$\mathcal{F}_A = \{b_1 < b_2 < \cdots\}, \qquad g_i := b_{i+1} - b_i \ (i \geq 1).$$

For integer $x \geq 1$, define

$$N(x) := \max\{i : b_i < x\}, \qquad r(x) := b_{N(x)+1} - x \in [0, \infty).$$

Thus $r(x)$ is the forward overshoot from x to the next A-free integer.

Proposition 8. *We have*

$$r(x) = o(x^{1/2}) \qquad (x \to \infty),$$

and therefore

$$\frac{r(x)^2}{x} \to 0.$$

Proof. By Corollary 5, $\rho_A > 0$. Choose Y so large that

$$\sum_{\substack{a \in A \\ a > Y}} \frac{1}{a} < \frac{\rho_A}{16}.$$

Let $\rho_Y := d(\mathcal{F}_{A,\leq Y})$. Since $\mathcal{F}_A \subseteq \mathcal{F}_{A,\leq Y}$ and $\mathcal{F}_{A,\leq Y} \setminus \mathcal{F}_A \subseteq \bigcup_{a>Y,\, a\in A} a\mathbb{N}$, we have

$$0 \leq \rho_Y - \rho_A \leq \sum_{\substack{a \in A \\ a > Y}} \frac{1}{a} < \frac{\rho_A}{16},$$

so $\rho_Y > \rho_A/2$.

Because $\mathbf{1}_{\mathcal{F}_{A,\leq Y}}$ is periodic, there are constants $c > 0$ and G_0 such that every interval of length $G \geq G_0$ contains at least cG elements of $\mathcal{F}_{A,\leq Y}$; for instance, we may take $c = \rho_A/4$.

Now let $(u, u+G)$ be a gap in $\mathcal{F}_A$, so that $u = b_i$ and $G = g_i$. Every point of $\mathcal{F}_{A,\leq Y} \cap (u, u+G)$ must be divisible by some $a \in A$ with $a > Y$. The divisors with $Y < a \leq G$ cover at most

$$\sum_{\substack{a \in A \\ Y < a \leq G}} \left(\left\lfloor \frac{G}{a} \right\rfloor + 1 \right) \leq G \sum_{\substack{a \in A \\ a > Y}} \frac{1}{a} + A(G)$$

points of the interval. Since $A(G) = o(G)$ and the tail sum is $< \rho_A/16$, this is at most $(c/2)G$ for all sufficiently large G.

Hence at least $(c/2)G$ points of $\mathcal{F}_{A,\leq Y} \cap (u, u+G)$ must be killed by divisors $a \in A$ with $a > G$. Such a divisor can hit at most one point of an interval of length G, so there are at least $(c/2)G$ distinct elements of $A \cap (G, u+G]$. Therefore

$$A(u+G) - A(G) \geq \frac{c}{2} G.$$

But $A(x) = o(x^{1/2})$, so the left-hand side is $o(\sqrt{u+G})$ and hence

$$G = o(\sqrt{u+G}).$$

Since $G \leq u + G$, this implies $G = o(\sqrt{u})$ as $u \to \infty$. Finally, if $x \in [b_i, b_{i+1})$, then $r(x) \leq g_i = o(\sqrt{b_i}) = o(\sqrt{x})$ because $b_i \sim x$. $\square$

The second ingredient is an exact combinatorial identity.

Lemma 9. *For every integer $x \geq 1$,*

$$\sum_{b_i < x} (b_{i+1} - b_i)^2 = x - 1 + r(x)^2 + 2 \sum_{h \geq 1} E_h(x).$$

Proof. Because $b_1 = 1$, there is no initial gap before the first point of $\mathcal{F}_A$. Each complete gap of length g_i with $b_i < x$ contributes exactly $\binom{g_i}{2}$ empty intervals. The only overcount occurs in the final partial gap ending at $b_{N(x)+1}$, and the intervals contained completely in $(x, b_{N(x)+1})$ contribute exactly $\binom{r(x)}{2}$. Hence

$$\sum_{h \geq 1} E_h(x) = \sum_{i \leq N(x)} \binom{g_i}{2} - \binom{r(x)}{2}.$$

Multiplying by 2 gives

$$2 \sum_{h \geq 1} E_h(x) = \sum_{i \leq N(x)} (g_i^2 - g_i) - r(x)^2 + r(x).$$

Also,

$$\sum_{i \leq N(x)} g_i = b_{N(x)+1} - b_1 = (x + r(x)) - 1.$$

Substituting this into the previous identity and rearranging yields the claim. □

4 Proof of Theorem 1

Proof of Theorem 1. Part (1) is Corollary 7.

Divide the identity of Lemma 9 by x. By Proposition 8, we obtain

$$\frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2 = 1 + 2 \sum_{h \geq 1} \nu_h(x) + o(1), \qquad \nu_h(x) := \frac{E_h(x)}{x}. \tag{1}$$

For each fixed x, the sequence $\nu_h(x)$ is decreasing in h; for each fixed h, it converges to $q_A(h)$.

Assume first that $\sum_{h \geq 1} q_A(h) = +\infty$. Then for every fixed H,

$$\liminf_{x \to \infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2 \geq 1 + 2 \sum_{1 \leq h \leq H} q_A(h)$$

by (1). Letting $H \to \infty$ yields

$$\lim_{x \to \infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2 = +\infty.$$

This is exactly the formula claimed in the theorem.

Now assume that $\sum_{h \geq 1} q_A(h) < \infty$. Since $\nu_h(x)$ is decreasing in h, dyadic decomposition gives, for each $H \geq 1$,

$$\sum_{h > H} \nu_h(x) \leq \sum_{m \geq 0} \sum_{2^m H < h \leq 2^{m+1} H} \nu_{2^m H}(x) \leq H \sum_{m \geq 0} 2^m \nu_{2^m H}(x).$$

Taking lim sup and using $\nu_{2^m H}(x) \to q_A(2^m H)$, we get

$$\limsup_{x \to \infty} \sum_{h > H} \nu_h(x) \leq H \sum_{m \geq 0} 2^m q_A(2^m H).$$

Since $q_A(h)$ is decreasing and summable, the dyadic tail on the right tends to 0 as $H \to \infty$; indeed,

$$H \sum_{m \geq 0} 2^m q_A(2^m H) \leq H q_A(H) + 2 \sum_{n \geq H} q_A(n) \to 0.$$

Therefore

$$\lim_{H \to \infty} \limsup_{x \to \infty} \sum_{h > H} \nu_h(x) = 0.$$

For each fixed H, we also have

$$\sum_{1 \leq h \leq H} \nu_h(x) \to \sum_{1 \leq h \leq H} q_A(h).$$

A standard truncation argument then yields

$$\sum_{h \geq 1} \nu_h(x) \to \sum_{h \geq 1} q_A(h).$$

Finally, returning to (1),

$$\lim_{x \to \infty} \frac{1}{x} \sum_{b_i < x} (b_{i+1} - b_i)^2 = 1 + 2 \sum_{h \geq 1} q_A(h),$$

as claimed. □

Remark 10. Lemma 6 proves more than Theorem 1: every fixed finite word in the indicator $\mathbf{1}_{\mathcal{F}_A}$ has an asymptotic frequency. In particular, the density of exact gaps of size g exists for every fixed g, since it is the frequency of the word $10^{g-1}1$.

5 Literature, broad finiteness classes, and what remains open

We now place Theorem 1 into the existing literature.

1. Squarefree and k-free gap moments

The squarefree case is $A = \{p^2 : p \text{ prime}\}$. In that model, Erdős proved already in 1951 that the second moment exists [5]. More generally, if $s_1 < s_2 < \cdots$ denotes the squarefree numbers, Chan proved in 2023 that

$$\sum_{s_{n+1} \leq x} (s_{n+1} - s_n)^\gamma \sim B(\gamma) x \qquad (0 \leq \gamma < 3.75),$$

improving a long chain of earlier work of Hooley, Filaseta, Trifonov, and Huxley [4, 7]. For k-free numbers with $k \geq 3$, Graham proved that

$$\sum_{s_{n+1} \leq x} (s_{n+1} - s_n)^\gamma \sim C_k(\gamma) x \qquad (0 \leq \gamma < 2k - 2 + 4/(k+1))$$

[9]. In particular, the second moment is finite in all classical k-free models.

2. General $\mathcal{B}$-free numbers and short intervals

For pairwise coprime thin forbidden sets, the $\mathcal{B}$-free literature is very large. Mirsky computed correlation functions in the k-free setting [13], and the modern survey of Avdeeva–Cellarosi–Sinai reviews the ergodic and statistical side of the subject [2]. For gaps and short intervals, important milestones include Kowalski–Robert–Wu [11], Matomäki [12], and Gorodetsky–Mangerel–Rodgers [8].

The latter paper gives a particularly clean route to finiteness of the second gap moment in a broad structured class. Recall their definition: a sequence $J \subseteq \mathbb{N}$ has *index* $\alpha \in [0,1]$ if $N_J(x) = x^{\alpha+o(1)}$.

Proof of Corollary 2. Let $G(X,H)$ denote the number of intervals $[n, n+H-1]$ with $n \leq X$ containing no A-free integers. Then $G(X,H) = E_H(X)$ up to the harmless endpoint convention already used throughout this note.

By Corollary 6.3 of Gorodetsky–Mangerel–Rodgers [8], if $\langle A \rangle$ has index $\alpha \in (0,1)$, then for every $\varepsilon > 0$ and every $k \geq 1$,

$$G(X,H) \ll_{A,\varepsilon,k} XH^{-(2-\alpha)k} \qquad (1 \leq H \leq X^{c_\alpha/k-\varepsilon}).$$

Fix H. Since the range condition is automatic for all sufficiently large X, dividing by X and letting $X \to \infty$ gives

$$q_A(H) \ll_A H^{-(2-\alpha)}$$

by taking $k = 1$. Because $2 - \alpha > 1$, the series $\sum_{H \geq 1} q_A(H)$ converges, and Theorem 1 implies finiteness of the second moment. □

3. Why the remaining case is still delicate

Even when the density of A-free integers is positive, very large gaps can occur. Alkan and Zaharescu constructed thin pairwise coprime forbidden sets with infinitely many unusually long consecutive gaps among the corresponding $\mathcal{B}$-free integers [1]. This does not contradict Theorem 1; it simply shows that the tail problem is nontrivial.

What remains genuinely open, at least from the standpoint of the present note, is the *unstructured* case of Problem #489:

> Assume only $A(x) = o(x^{1/2})$. Must the series $\sum_{h \geq 1} q_A(h)$ always converge? Equivalently, must the second gap moment of $\mathcal{F}_A$ always be finite?

Theorem 1 shows that this is the only remaining issue. The problem page itself still lists the original existence-and-finiteness question as open [3].

References

[1] E. Alkan and A. Zaharescu, *Consecutive large gaps in sequences defined by multiplicative constraints*, Canad. Math. Bull. **51** (2008), no. 2, 172–181.

[2] M. Avdeeva, F. Cellarosi, and Y. G. Sinai, *Ergodic and statistical properties of $\mathcal{B}$-free numbers*, Theory Probab. Appl. **61** (2017), no. 4, 618–642.

[3] T. F. Bloom, *Erdős Problem #489*, `https://www.erdosproblems.com/489`, accessed 2026-04-20.

[4] T. H. Chan, *On moments of gaps between consecutive squarefree numbers*, Moscow J. Combin. Number Theory **12** (2023), no. 4, 287–295.

[5] P. Erdős, *Some problems and results in elementary number theory*, Publ. Math. Debrecen **2** (1951), 103–109.

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

[7] K. F. Roth, *Starting with gaps between k-free numbers*, Int. J. Number Theory **11** (2015), no. 5, 1411–1435.

[8] O. Gorodetsky, A. P. Mangerel, and B. Rodgers, *Squarefrees are Gaussian in short intervals*, arXiv:2112.12234.

[9] S. W. Graham, *Moments of gaps between k-free numbers*, J. Number Theory **44** (1993), no. 1, 105–117.

[10] R. R. Hall, *Sets of Multiples*, Cambridge University Press, Cambridge, 1996.

[11] E. Kowalski, O. Robert, and J. Wu, *Small gaps in coefficients of L-functions and $\mathcal{B}$-free numbers in short intervals*, Rev. Mat. Iberoam. **23** (2007), no. 1, 281–326.

[12] K. Matomäki, *On the distribution of $\mathcal{B}$-free numbers and non-vanishing Fourier coefficients of cusp forms*, Glasg. Math. J. **54** (2012), no. 2, 381–397.

[13] L. Mirsky, *Arithmetical pattern problems relating to divisibility by r-th powers*, Proc. London Math. Soc. **50** (1949), 497–508.

[14] P. Pollack and C. Sanna, *Uncertainty principles connected with the Möbius inversion formula*, arXiv:1211.0189.