hexagon:2611.00765v1

THE THREE PARTS OF ERDŐS PROBLEM 12: COUNTING FUNCTIONS AND RECIPROCAL SUMS

WILL HOWE, GRANT HUFFMAN, JENSEN KOHLMEYER, AND LIAM KRUER

Abstract. We give a unified account of the three numbered questions in Erdős Problem 12 for sets in which no element divides the sum of two distinct larger elements. The positive answer to the square-root counting question and the negative answer to a universal power-saving bound are prior AlphaProof Nexus results of Google DeepMind. We include attributed proofs following its released Lean constructions. For the reciprocal-sum question, we retain the proof that one absolute constant bounds the sum for every such set, finite or infinite. The distinction between parameterwise dense examples and the uniform reciprocal bound is made explicit. A companion single-file Lean formalization combines the two adapted prior proofs with the reciprocal-sum proof and checks all three statements together.

1. The three questions and their attribution

A set $A \subseteq \mathbb{N} = \{1, 2, \ldots\}$ has *Property P* if there are no pairwise distinct elements $a, b, c \in A$ such that

$$a < b, \qquad a < c, \qquad a \mid b + c.$$

In particular, equality of the two larger summands is permitted by the formulation and does not constitute a forbidden triple. Write $A(N) = |A \cap \{1, \ldots, N\}|$. Erdős Problem 12 asks three separate questions about the counting function and reciprocal sum of such an infinite set.

Theorem 1.1 (Part (i), AlphaProof Nexus)**.** *There is an infinite set A with Property P such that*

$$\liminf_{N\to\infty} \frac{A(N)}{\sqrt{N}} > 0.$$

Theorem 1.2 (Part (ii), AlphaProof Nexus)**.** *There is no absolute $c > 0$ such that every infinite set A with Property P satisfies $A(N) < N^{1-c}$ for infinitely many N. More precisely, for each $0 < c < 1$ there are an infinite set A_c with Property P and an integer N_c such that*

$$A_c(N) \geq N^{1-c} \qquad (N \geq N_c). \tag{1.1}$$

The set and the threshold may depend on c.

Theorem 1.3 (Part (iii), with a uniform bound)**.** *There is an absolute constant C such that every set $A \subseteq \mathbb{N}$ with Property P satisfies*

$$\mathcal{H}(A) := \sum_{a\in A} \frac{1}{a} \leq C. \tag{1.2}$$

The same constant applies to finite and infinite sets.

Date: November 5, 2026.

2020 *Mathematics Subject Classification.* 11B05, 11N35, 11N36.

Key words and phrases. Erdős problems, reciprocal sums, divisibility, sieve methods, primitive sets.

E-mail. howew@purdue.edu (W. Howe), huffma97@purdue.edu (G. Huffman), jenwin@purdue.edu (J. Kohlmeyer), kruerl@purdue.edu (L. Kruer).

The answers to the three numbered questions are therefore, respectively, *yes*, *no*, and *yes*. The first two answers are prior results of the Google DeepMind AlphaProof Nexus work of Tsoukalas, Kovsharov, Shirobokov et al. [11]. They are included here with attribution, exposition, and adaptations of the released Lean proofs [4, 5]; they are not contributions first obtained in this manuscript. The reciprocal-sum argument of Theorem 1.3 is the argument developed in the original manuscript by the four named authors. The present version brings the three statements into one account and one checked Lean source file.

The quantifier order in (1.1) matters: $\forall c\, \exists A_c\, \exists N_c$. We use precisely this parameterwise assertion. The construction below does not assert that one fixed set works for every c. The lower bounds in part (ii) are compatible with the uniform reciprocal bound in part (iii): the onset N_c and the set itself vary with the exponent.

Erdős and Sárközy [3] proved that every infinite set with Property P has density zero, and formulated the reciprocal-sum problem and its stronger uniform version. For the counting questions, Section 2 follows the congruence-block constructions in the released AlphaProof Nexus Lean files. Those are the concrete proof instances incorporated in the companion formalization. The constructions are presented here directly, rather than by relying on admitted statements in the Formal Conjectures catalog.

The proof of part (iii) has three main steps. A quantitative incompatibility between separated blocks of rough elements is combined with sieve estimates to control prime fibres carrying inherited congruence restrictions. We then represent reciprocal sums by random prime-factor prefix weights. An adapted rule for selecting the restrictions gives a uniform total expected weight. The restriction estimate is a special case of Bera–Viswanadham [2]; Section 3 includes a direct proof in the form needed here. The Mertens prefix weights are classical; see Alexeev et al. [1]. Earlier work of Plasma AI [8] contains the reduction to odd primitive sets, restricted reciprocal-sum estimates, and the problem of controlling accumulated congruence restrictions. These attributions apply to the reciprocal-sum argument, independently of the provenance of parts (i) and (ii).

Notation and conventions. All logarithms are natural unless a base is indicated. We write $\mathrm{e}(t) = \exp(2\pi i t)$, and $\|t\|_{\mathbb{R}/\mathbb{Z}}$ for distance from a real number to the nearest integer. The constants implicit in $\ll$ and $\asymp$ are absolute unless a dependence is indicated. An integer is *D-rough* if no prime factor is at most D, and *D-smooth* if every prime factor is at most D. The integer 1 satisfies both conventions. We write $P^+(g)$ for the largest prime factor of $g > 1$, with $P^+(1) = 1$. A set is *primitive* if no two distinct members divide one another. Property P is inherited by subsets and by division by a common divisor. The paper uses positive integers; the Lean statements use natural numbers, with $1/0 = 0$. The constructions below consist of positive integers.

2. The counting constructions: parts (i) and (ii)

This section is an attributed exposition of the released proofs [4, 5]. We isolate the common congruence argument and then make the two choices of scale. Omitting finitely many initial blocks in the exposition simplifies the estimates and does not affect the asserted asymptotic statements. The Lean implementations verify explicit initial bounds as well.

Lemma 2.1 (Congruence separation)**.** *Let B_j be finite sets of positive integers, with every element of B_i smaller than every element of B_j when $i < j$. Suppose that there are pairwise distinct primes $q_j > 2$ such that elements of B_j are $0 \pmod{q_j}$ and $1 \pmod{q_i}$ for every $i < j$. If each B_j lies in an interval $[L_j, U_j]$ with $2U_j < 3L_j$, then $\bigcup_j B_j$ has Property P.*

Proof. Suppose $a \in B_i$, $a < b, c$, and $a \mid b + c$. If either larger element lies in a later block, then $b + c$ is 1 or $2 \pmod{q_i}$, whereas $a \equiv 0 \pmod{q_i}$. This is impossible because $q_i > 2$. Otherwise both larger elements lie in B_i, and $2a < b + c \le 2U_i < 3L_i \le 3a$, which precludes divisibility. □

2.1. **A positive square-root lower limit.** Choose increasing primes $q_0, q_1, \dots$, starting with $q_0 = 5$, so that $q_{j+1} \le 2q_j + 2$. Bertrand's postulate gives such a choice, and induction yields $q_j + 2 \le 7 \cdot 2^j$. Put

$$K_j = 729 \prod_{i<j} q_i, \qquad P_j = K_j q_j, \qquad M_j = P_j^8, \qquad T_j = \left\lfloor \frac{P_{j+1}^4}{2} \right\rfloor.$$

The Chinese remainder theorem supplies $r_j \in [0, P_j)$ such that $r_j \equiv 1 \pmod{K_j}$ and $r_j \equiv 0 \pmod{q_j}$. Define

$$B_j = \{M_j + r_j + tP_j : 0 \le t < T_j\}.$$

Then $|B_j| = T_j$. Also $P_{j+1} = P_j q_{j+1}$, while $P_j \ge 3645 \cdot 5^j$. Consequently

$$\frac{q_{j+1}^4}{P_j^3} \longrightarrow 0, \qquad \frac{T_j P_j}{M_j} \longrightarrow 0.$$

Thus, after removing finitely many initial blocks, each B_j lies in $[M_j, (7/5)M_j]$. The scales are strictly separated because $M_{j+1} = M_j q_{j+1}^8$. The residues of B_j have the form in Lemma 2.1, since $q_i \mid K_j$ for $i < j$. Their union A therefore has Property P.

For large j,

$$T_j = \tfrac{1}{2}\sqrt{M_{j+1}} + O(1).$$

If N falls inside the short interval containing B_j, the preceding complete block contributes $T_{j-1} = \frac{1}{2}\sqrt{M_j} + O(1)$ elements. If N falls in the following gap, block B_j contributes $T_j = \frac{1}{2}\sqrt{M_{j+1}} + O(1)$ elements. These two cases give

$$A(N) \ge \tfrac{1}{3}\sqrt{N}$$

for every sufficiently large N. To verify that the lower limit is finite as well, note that $r_j > 0$, so $A(M_j) = \sum_{i<j} |B_i|$, up to the omitted finite prefix. Since $P_{i+1} \le P_j 5^{-(j-i-1)}$,

$$A(M_j) \le \frac{1}{2} P_j^4 \sum_{k \ge 0} 5^{-4k} < P_j^4 = \sqrt{M_j}.$$

This proves Theorem 1.1. In particular, the lower limit is bounded away from zero and bounded above; this agrees with the real-valued lower-limit convention used by the Lean statement.

2.2. **Failure of a universal power saving.** Fix $0 < c < 1$. Choose distinct odd primes p_j with $p_j \le 2^{j+2}$, and put $Q_j = \prod_{i \le j} p_i$. Then $Q_j \le 2^{(j+2)^2}$. Choose an integer $a \ge 4$ with $ac \ge 3$, and define

$$M_j = 2^{a(j+2a)^2}.$$

Choose a residue $s_j \pmod{Q_j}$ that is $0 \pmod{p_j}$ and $1 \pmod{p_i}$ for $i < j$, and set

$$B_j = \{n \in [10M_j, 14M_j] : n \equiv s_j \pmod{Q_j}\}.$$

The intervals are disjoint and ordered because $14M_j < 10M_{j+1}$. They satisfy the strict width inequality in Lemma 2.1. Once again their union A_c has Property P, and it is infinite because the blocks are nonempty for all sufficiently large j.

Counting a residue class in an interval gives

$$|B_j| \ge \frac{4M_j}{Q_j} - 1.$$

The logarithm to base 2 of $4M_j/Q_j$ is at least $2 + a(j+2a)^2 - (j+2)^2$, whereas

$$\log_2\big((14M_{j+1})^{1-c}\big) = (1-c)\big(\log_2 14 + a(j+1+2a)^2\big).$$

The difference has leading term $(ac-1)j^2$, whose coefficient is positive. Therefore, for all sufficiently large j,

$$|B_j| \ge (14M_{j+1})^{1-c}.$$

For $14M_j \le N < 14M_{j+1}$, the complete block B_j is already counted in $A_c(N)$. This proves (1.1). If a universal saving $d > 0$ existed, apply the construction with $c = \min(d, 1/2)$. For sufficiently large $N \ge 1$, $A_c(N) \ge N^{1-c} \ge N^{1-d}$, contradicting the required infinitely many strict inequalities. This proves Theorem 1.2.

3. A restriction estimate for sifted integers

Let $\mathcal{P}$ be any set of primes. Define a multiplicative function supported on squarefree products of primes in $\mathcal{P}$ by

$$h_{\mathcal{P}}(1) = 1, \qquad h_{\mathcal{P}}(d) = \prod_{p \mid d} \frac{1}{p-1},$$

and put $h_{\mathcal{P}}(d) = 0$ for all other positive integers d. For $z \ge 2$, set

$$G_{\mathcal{P}}(z) = \sum_{d<z} h_{\mathcal{P}}(d), \qquad \mathcal{S}(N, \mathcal{P}, z) = \{1 \le n \le N : p \nmid n \text{ for every } p \in \mathcal{P},\ p < z\}.$$

Using strict inequalities at the sieve cutoff removes endpoint ambiguities.

Lemma 3.1. *Let $2 \le z \le N^{1/4}$, and let $\mathcal{B} \subseteq \mathbb{R}/\mathbb{Z}$ be a finite δ-separated set. For every fixed $q > 2$ and every complex sequence (a_n) supported on $\mathcal{S}(N, \mathcal{P}, z)$,*

$$\left(\sum_{\xi \in \mathcal{B}} \left| \sum_{n \le N} a_n \mathrm{e}(n\xi) \right|^q \right)^{1/q} \ll_q \left(\frac{N + \delta^{-1}}{G_{\mathcal{P}}(z)} \sum_{n \le N} |a_n|^2 \right)^{1/2}. \tag{3.1}$$

The constant is uniform in $\mathcal{P}$.

Proof. We give the enveloping-sieve argument in the form needed here. For a parameter $z_0 \ge 2$, let $\mathcal{P}' = \mathcal{P} \cap [z_0, z)$, write $h' = h_{\mathcal{P}'}$, and put $G' = \sum_{d<z} h'(d)$. For $p \in \mathcal{P}'$, define

$$u_p(n) = \frac{1 - p\mathbf{1}_{p \mid n}}{p-1}, \qquad u_d(n) = \prod_{p \mid d} u_p(n).$$

The mean of u_p modulo p is zero. Its nonzero additive Fourier coefficients all equal $-1/(p-1)$, and

$$u_p(n)^2 = h'(p) + (h'(p) - 1)u_p(n).$$

Consequently the nonnegative periodic function

$$\beta(n) = \frac{1}{(G')^2} \left(\sum_{\substack{d<z \\ h'(d)>0}} u_d(n) \right)^2 \tag{3.2}$$

equals 1 on $\mathcal{S}(N, \mathcal{P}, z)$. Orthogonality and the Chinese remainder theorem show that its constant Fourier coefficient is $1/G'$.

A nonconstant Fourier coefficient can occur only at a reduced fraction a/ℓ, where $\ell < z^2$ is squarefree and supported on $\mathcal{P}'$. Its absolute value is at most

$$\frac{3^{\omega(\ell)} h'(\ell)}{G'} \ll \frac{1}{G'\sqrt{\ell}}, \tag{3.3}$$

where $\omega(\ell)$ is the number of distinct prime factors. Indeed, in a term $u_d u_e$, a prime dividing ℓ has at most three membership patterns in d, e, and contributes a coefficient of magnitude at most $h'(p)$. A prime outside ℓ must occur in both factors to contribute a constant term, which is $h'(p)$. Summing those common factors costs at most G'. Finally, $3/(p-1) \le p^{-1/2}$ for all sufficiently large primes, and the finitely many remaining primes contribute an absolute factor.

Separating the prime factors below z_0 gives

$$G_{\mathcal{P}}(z) \le G' \prod_{p<z_0} (1-1/p)^{-1} \ll G' \log(2z_0). \tag{3.4}$$

For a nonzero vector $c = (c_\xi)_{\xi\in\mathcal{B}}$, put

$$R = \frac{\|c\|_1^2}{\|c\|_2^2} \ge 1, \qquad z_0 = (2R)^2, \qquad \eta = \min\{\delta/4, 1/(4z^4)\}.$$

Choose a nonnegative integrable function ψ majorizing 1 on $[1,N]$, with Fourier transform supported on $[-\eta,\eta]$, integral $O(N+\eta^{-1})$, and decay $O((1+t^2)^{-1})$. One construction is an absolute multiple of the convolution of $\mathbf{1}_{[-1/\eta,N+1+1/\eta]}$ with

$$\eta\left(\frac{\sin(\pi\eta t)}{\pi\eta t}\right)^2.$$

Poisson summation implies

$$\sum_{n\in\mathbb{Z}} \psi(n)\mathrm{e}(n\theta) = 0 \quad \text{if } \|\theta\|_{\mathbb{R}/\mathbb{Z}} > \eta, \qquad \left|\sum_{n\in\mathbb{Z}} \psi(n)\mathrm{e}(n\theta)\right| \ll N+\eta^{-1}.$$

Use $\psi\beta$ to majorize the dual square sum. The constant coefficient of β contributes only when $\xi=\xi'$. For a fixed ordered pair (ξ,ξ'), at most one reduced nonzero fraction a/ℓ, $\ell<z^2$, can satisfy $\|\xi-\xi'+a/\ell\|_{\mathbb{R}/\mathbb{Z}} \le \eta$, since distinct such fractions have distance at least z^{-4}. Every nonzero denominator in (3.3) is at least z_0. We obtain

$$\sum_{n\in\mathcal{S}(N,\mathcal{P},z)} \left|\sum_{\xi\in\mathcal{B}} c_\xi \mathrm{e}(n\xi)\right|^2 \ll \frac{N+\eta^{-1}}{G'}\left(\|c\|_2^2 + z_0^{-1/2}\|c\|_1^2\right).$$

If $z_0 \ge z$, then $\beta=1$ and the second term is absent. Since $z^4 \le N$, $z_0^{1/2} = 2R$, and (3.4) holds, this gives

$$\sum_{n\in\mathcal{S}(N,\mathcal{P},z)} \left|\sum_{\xi\in\mathcal{B}} c_\xi \mathrm{e}(n\xi)\right|^2 \ll \frac{N+\delta^{-1}}{G_{\mathcal{P}}(z)} \log(2R)\,\|c\|_2^2. \tag{3.5}$$

Write

$$U^2 = \frac{N+\delta^{-1}}{G_{\mathcal{P}}(z)} \sum_n |a_n|^2, \qquad F(\xi) = \sum_n a_n \mathrm{e}(n\xi).$$

The one-frequency case of (3.5) gives $\max_{\xi\in\mathcal{B}} |F(\xi)| \ll U$. For $\mathcal{E}_\lambda = \{\xi\in\mathcal{B} : |F(\xi)| \ge \lambda U\}$, choose unit complex phases realizing the absolute values in the pairing with F. Cauchy–Schwarz and (3.5) imply

$$\lambda^2 |\mathcal{E}_\lambda| \ll \log(2|\mathcal{E}_\lambda|) \quad \text{whenever } \mathcal{E}_\lambda \ne \varnothing.$$

Thus $|\mathcal{E}_\lambda| \ll \lambda^{-2}\log(2+\lambda^{-1})$ on the bounded range where the level set is nonempty. Integrating this estimate against $q\lambda^{q-1}$ proves (3.1). □

Lemma 3.2. *There are absolute constants $c_0>0$ and D_0 with the following properties. If $D \ge D_0$, $N \ge D$, and $z = \frac{1}{2}\min\{D, N^{1/4}\}$, then*

$$G_{\{\text{all primes}\}}(z) \gg \log D. \tag{3.6}$$

If $16 \le K \le D^{c_0}$, then

$$G_{\{p>4K\}}(z) \gg \frac{\log D}{\log(2K)}. \tag{3.7}$$

Proof. Put $z_1 = z^c$, where $c > 0$ is a sufficiently small absolute constant. In the Euler product $\prod_{4K<p\le z_1}(1+1/(p-1))$, normalize the weight of each squarefree product to a probability distribution. A prime p is then selected with probability $1/p$. The expected logarithm of the product is

$$\sum_{4K<p\le z_1} \frac{\log p}{p} \ll c \log z.$$

Markov's inequality shows that at least half the total weight is carried by products smaller than z, after c is fixed sufficiently small. Mertens' estimates give

$$\prod_{4K<p\le z_1} (1-1/p)^{-1} \asymp \frac{\log z_1}{\log(2K)}.$$

Choose c_0 small enough that $z_1 > 8K$ for all sufficiently large D. Since $\log z_1 \gg \log D$, this proves (3.7). The same argument with all primes at most z_1 proves (3.6). □

4. Interaction between separated rough blocks

Lemma 4.1. *Let $D \ge D_0$, let $X < Y$ be powers of two with $X \ge D$, and put $K = Y/X$. Assume $16 \le K \le D^{c_0}$. Suppose that $E \subseteq [X, 2X)$ and $T \subseteq [Y, 2Y)$ are D-rough subsets of one set with Property P. With*

$$U = \log D, \qquad u = \frac{U|E|}{X}, \qquad v = \frac{U|T|}{Y},$$

one has

$$\sqrt{u}\, v \ll K^{-1/8}. \tag{4.1}$$

Proof. The one-frequency case of Lemma 3.1, together with (3.6), gives $u, v \ll 1$. We may assume both sets are nonempty.

Work on $\mathbb{Z}/M\mathbb{Z}$, where $M = 32Y$, with normalized Fourier transform

$$\widehat{w}(\xi) = \frac{1}{M} \sum_{n \bmod M} w(n) \mathrm{e}(-n\xi/M).$$

Function-space norms are normalized by M^{-1}, while Fourier-space ℓ^q-norms use counting measure. Set

$$f = U\mathbf{1}_E, \qquad t = U\mathbf{1}_T, \qquad \alpha = \frac{1}{M}\sum_n f(n) = \frac{u}{32K}.$$

Let $k_0 = K/4$, $k_1 = 4K$, and $J = [k_0, k_1] \cap \mathbb{Z}$. Define

$$b(n) = \frac{1}{|J|} \sum_{\substack{a\in E,\ k\in J \\ 2ka=n}} f(a).$$

All the products in this definition are smaller than M. Its Fourier transform and mean satisfy

$$B(\xi) := \widehat{b}(\xi) = \frac{1}{|J|}\sum_{k\in J} \widehat{f}(2k\xi), \qquad \frac{1}{M}\sum_n b(n) = \alpha.$$

Every prime factor of $2k$ is at most $4K < D$. The D-rough part of $n = 2ka$ is therefore exactly a. In particular the representation determines a uniquely, and

$$b(n) \ll U/K. \tag{4.2}$$

The support of b avoids all primes in $(4K, D]$.

For $a \in E$, put

$$I_a = [k_0 a, k_1 a] \cap \mathbb{Z}, \qquad b_0(n) = \sum_{a\in E} \frac{f(a)}{|I_a|} \mathbf{1}_{\{n=2\ell : \ell \in I_a\}}.$$

Then b_0 has mean α, is $O(\alpha)$ everywhere, and is at least $c\alpha$ on the even integers in $[2Y, 4Y]$. Indeed, $|I_a| \asymp Y$, and this latter interval is contained in the doubled interval $2I_a$ for every $a \in E$. Write $B_0 = \widehat{b_0}$.

A uniform Fourier comparison. We claim that

$$\|B - B_0\|_{\ell^\infty} \ll \alpha u^{-1/2} K^{-2/9}. \tag{4.3}$$

Fix a frequency ξ, and put $\lambda = 2\xi/M \pmod 1$. Throughout this comparison take $q = 9/4$.

First suppose that $\|h\lambda\|_{\mathbb{R}/\mathbb{Z}} \geq 1/(100X)$ for every $1 \leq h \leq 4K$. The frequencies $k\lambda$, $k \in J$, are $1/(100X)$-separated. Lemma 3.1, applied on $[1, 2X]$ and normalized by U/M, gives an ℓ^q-norm $O(\sqrt{u}/K)$ on these frequencies. Hölder's inequality then yields

$$|B(\xi)| \ll \frac{\sqrt{u}}{K} K^{-1/q}.$$

The geometric average over I_a also gives $|B_0(\xi)| \ll \alpha/K$.

Otherwise, for some $h \leq 4K$, reduce the nearby rational approximation to obtain

$$\lambda = \frac{r}{d} + \beta \pmod 1, \qquad (r, d) = 1, \qquad d \leq 4K, \qquad |\beta| \leq \frac{1}{100dX}.$$

If $2 \leq d \leq K^{1/2}$, every $a \in E$ is a unit modulo d, so

$$\|a\lambda\|_{\mathbb{R}/\mathbb{Z}} \geq \frac{1}{d} - 2X|\beta| \geq \frac{1}{2d}.$$

The geometric average over J therefore gives $|B(\xi)| \ll \alpha d/K \ll \alpha K^{-1/2}$.

If $d > K^{1/2}$, divide J into consecutive chunks of length at most d. Within a chunk the frequencies $k\lambda$ are separated by at least $1/(2d)$, since $d \leq 4K \ll X$. Apply Lemma 3.1 to each chunk and then average over J. This gives

$$|B(\xi)| \ll \frac{\sqrt{u}}{K} d^{-1/q} \ll \frac{\sqrt{u}}{K} K^{-1/(2q)}.$$

In both cases with $d \geq 2$, the interval average defining B_0 is bounded by

$$|B_0(\xi)| \ll \frac{\alpha d}{Y} \ll \frac{\alpha K}{Y}.$$

This term is smaller than the right side of (4.3), after decreasing c_0 if needed.

It remains to consider $d = 1$. Use the lift of λ with $|\lambda| \leq 1/(100X)$, and put $\rho = M|\lambda|$. If $\rho \geq K^{1/4}$, geometric summation gives

$$|B(\xi)| + |B_0(\xi)| \ll \alpha/\rho \ll \alpha K^{-1/4}.$$

If $\rho < K^{1/4}$, couple the uniform measure on I_a to the grid $\{ka : k \in J\}$ by intervals $[ka, (k+1)a)$. The endpoint adjustment has total variation $O(1/K)$, and the phase changes by $O(a|\lambda|) = O(\rho/K)$ inside each interval. Hence

$$|B(\xi) - B_0(\xi)| \ll \alpha(1 + \rho)/K.$$

These bounds prove (4.3), since $1/(2q) = 2/9$ and $u \ll 1$.

The forbidden correlation. By (4.2),

$$\|b\|_{L^2(\mathbb{Z}/M\mathbb{Z})} \ll \alpha\sqrt{U/u}.$$

Apply Lemma 3.1 to b, using the prime band $(4K, D]$, the full $1/M$-spaced frequency grid, and (3.7). The result is

$$\|B\|_{\ell^q} \ll \alpha u^{-1/2}\sqrt{\log(2K)}.$$

Since $b_0 \ll \alpha$, Parseval gives $\|B_0\|_{\ell^2} \ll \alpha$, and thus $\|B_0\|_{\ell^q} \ll \alpha$. Interpolating with (4.3), with $q/9 = 1/4$, gives

$$\|B - B_0\|_{\ell^9} \ll \alpha u^{-1/2} K^{-1/6} \log(2K)^{1/8} \ll \alpha u^{-1/2} K^{-1/8}. \tag{4.4}$$

The full-prime restriction estimate likewise gives $\left\|\widehat{t}\right\|_{\ell^q} \ll \sqrt{v}$. Fourier inversion and Hölder's inequality, with exponents 9 and 9/8, now show that replacing b by b_0 in

$$\mathbb{E}_{x,y \bmod M} t(x)t(y)b(x+y)$$

changes the expression by at most $C\alpha v u^{-1/2} K^{-1/8}$.

Every element of T is odd. Thus $x + y$ is an even integer in $[2Y, 4Y]$ whenever $x, y \in T$, and the expression with b_0 is at least $c\alpha v^2$. The expression with b vanishes. A nonzero term would give the integer equality $x + y = 2ka$ with $x, y > a$ and a, x, y in the underlying set. Distinct x, y contradict Property P. If $x = y$, then $x = ka$, which is impossible because $k \geq 2$ is D-smooth and x is D-rough. Therefore $c\alpha v^2 \leq C\alpha v u^{-1/2} K^{-1/8}$, which proves (4.1). □

5. Prime sets with inherited congruence restrictions

We use the analytic large sieve and the Brun–Titchmarsh inequality in their classical forms [6, 7]. In particular, if a set of primes occupies at most $\vartheta\varphi(m)$ reduced residue classes modulo m, then, for a fixed $a > 1$, $x \geq m^a$, $x \geq 2$, and $y \geq x$, partial summation gives

$$\sum_{\substack{p \in Q \\ x \leq p < y}} \frac{1}{p} \ll_a \vartheta \left(1 + \log \frac{\log y}{\log x}\right). \tag{5.1}$$

Indeed, Brun–Titchmarsh bounds the counting function by $O_a(\vartheta t/\log t)$ for $t \geq x$.

Lemma 5.1. *Every set Q of primes with Property P satisfies $\mathcal{H}(Q) \ll 1$, with one absolute constant.*

Proof. Remove the possible prime 2, which costs at most 1/2. If the remaining set is nonempty, let q be its least element. The reciprocal sum over $[q, q^2]$ is $O(1)$ by the ordinary upper bound for primes and partial summation.

Enumerate the primes of Q greater than q^2 as $p_1 < p_2 < \cdots$, and set

$$L_0 = q, \qquad L_i = qp_1 \cdots p_i, \qquad x_i = \max\{L_i^2, p_{i+1}\}, \qquad y_i = L_{i+1}^2.$$

For every earlier prime modulus $r \in \{q, p_1, \ldots, p_i\}$, the later primes occupy at most one residue in each pair $\{a, -a\}$ of nonzero residues modulo r. Otherwise two distinct later primes have sum divisible by r. The Chinese remainder theorem therefore gives a fraction at most 2^{-i-1} of the reduced classes modulo L_i. Moreover $\log y_i \leq 3 \log x_i$. Equation (5.1) bounds the mass in $[x_i, y_i)$ by $O(2^{-i})$. These intervals cover all remaining primes of Q. For a finite list, its last prime is smaller than the final product squared, so the same argument applies. The resulting geometric series proves the claim. □

Lemma 5.2. *There are absolute constants $A_0 \geq 3$ and C with the following property. Let $Q \subseteq [Y, \infty)$ be a set of primes with Property P. Let $q_1, \ldots, q_k < Y$ be distinct odd primes, and set*

$$M = q_1 \cdots q_k, \qquad \delta = 2^{-k},$$

with $M = 1$ if $k = 0$. Suppose $Y \geq M^{A_0}$ and, for each q_i, the residues of Q occupy at most one member of every pair of opposite nonzero classes modulo q_i. Then

$$\mathcal{H}(Q) \leq C \frac{\delta(1+k)}{\log(2Y)}. \tag{5.2}$$

Proof. For bounded Y, the condition $M^{A_0} \le Y$ bounds M and k. Lemma 5.1 then suffices. We henceforth take Y larger than a fixed absolute threshold. The choice of A_0, made below in terms of c_0, is fixed once and for all.

The Chinese remainder theorem puts all primes of Q in at most $\delta\varphi(M)$ reduced residue classes modulo M. For any dyadic $V \ge Y/2$, Brun–Titchmarsh gives

$$|Q \cap [V, 2V)| \ll \frac{\delta V}{\log V}. \tag{5.3}$$

Here $M \le Y^{1/A_0}$, and $A_0 \ge 3$. For $M = 1$, use the ordinary prime upper bound. We will also use

$$\delta \ge M^{-1} \ge Y^{-1/A_0}. \tag{5.4}$$

Call the dyadic bin at $V = 2^j$ *light* if its cardinality is at most $\delta V/j^2$. All light bins together contribute $O(\delta/\log Y)$ to $\mathcal{H}(Q)$. If there is no other bin, the proof is complete.

Let $[U, 2U)$, $U = 2^J$, be the first bin that is not light, and put $n = |Q \cap [U, 2U)| > \delta U/J^2$. Choose

$$s = \lceil 8 \log_2 J \rceil$$

distinct primes from that bin. This is possible for all sufficiently large Y, by (5.4). Let L be M times their product, and set

$$V_0 = UJ^8\delta^{-8}.$$

Choose A_0 sufficiently large that, uniformly for $U \ge Y/2$,

$$4V_0/U \le (U/2)^{c_0} \tag{5.5}$$

for all sufficiently large Y. This follows from (5.4). The same range gives $L > V_0$.

The near range. Take $D = U/2$. In dyadic bins meeting $[U, V_0]$, every prime is D-rough. The normalized count

$$w_V = \frac{\log D}{V}|Q \cap [V, 2V)|$$

is $O(\delta)$, by (5.3). If two such counts are at least $w > 0$, then Lemma 4.1 bounds their ratio of scales by $O(w^{-12})$, unless the scales differ by a factor less than 16. The hypotheses of that lemma hold throughout the range by (5.5). It follows that

$$\#\{V : w_V \ge w\} \ll 1 + \log(C/w).$$

Integrating this bound over $0 < w \ll \delta$ gives

$$\mathcal{H}(Q \cap [U, V_0]) \ll \frac{\delta(1 + \log(1/\delta))}{\log U}. \tag{5.6}$$

The middle range. Let $T = Q \cap [V, 2V)$, where $V > 2U$. For every prime $p \in Q \cap [U, 2U)$, the set T occupies at most $(p-1)/2$ classes modulo p. Orthogonality and Cauchy–Schwarz give

$$\sum_{j=1}^{p-1} \left| \sum_{t \in T} \mathrm{e}(jt/p) \right|^2 \ge |T|^2.$$

The reduced fractions j/p belonging to different such primes are distinct and have spacing at least $1/(4U^2)$. The analytic large sieve therefore gives

$$n|T|^2 \ll (V + U^2)|T|, \qquad \mathcal{H}(T) \ll \frac{1}{n} + \frac{U^2}{Vn}.$$

Summing over bins meeting $[V_0, L]$, including the at most two boundary bins, yields

$$\mathcal{H}(Q \cap [V_0, L]) \ll \frac{sJ^3}{\delta U} + \frac{\delta^7}{J^6} \ll \frac{\delta}{J}. \tag{5.7}$$

For the first inequality use $\log L = O(sJ)$ and $n > \delta U/J^2$. For the second, divide by δ/J and use

$$\frac{sJ^4}{\delta^2 U} + \frac{\delta^6}{J^5} = o(1),$$

uniformly in the permitted range, by (5.4).

The far range. On $(L, L^2]$, use the external q_i and the $\lfloor s/2 \rfloor$ smallest selected internal primes as moduli. Their product is at most $\sqrt{ML}$, hence at most $L^{2/3}$ for sufficiently large Y. The occupied fraction of reduced classes is at most $\delta 2^{-\lfloor s/2 \rfloor}$. Equation (5.1) therefore bounds the mass in this range by $O(\delta 2^{-\lfloor s/2 \rfloor})$.

Enumerate the primes of Q above L^2 as $p_1 < p_2 < \cdots$, and now put $L_i = Lp_1 \cdots p_i$, with $L_0 = L$. The primes in $[L_i^2, L_{i+1}^2)$ are at least p_{i+1}, and occupy at most the fraction $\delta 2^{-s-i}$ of the reduced classes modulo L_i. With

$$x_i = \max\{L_i^2, p_{i+1}\}, \qquad y_i = L_{i+1}^2,$$

one has $\log y_i \le 3 \log x_i$. Another application of (5.1) gives mass $O(\delta 2^{-s-i})$. As in Lemma 5.1, these intervals also cover every term of a finite list. Altogether the far range contributes

$$O(\delta 2^{-s/2}) = O(\delta/J).$$

Combine this with the light-bin bound, (5.6), and (5.7). Since $\log(1/\delta) = k \log 2$ and $\log U \gg \log Y$, we obtain (5.2). □

6. Prime-factor prefixes

Let A be an odd primitive set with Property P. If $1 \in A$, then $A = \{1\}$, so suppose $1 \notin A$. For each positive integer g, define

$$Q_g = \{p \text{ prime} : p \ge P^+(g),\ gp \in A\}.$$

Each Q_g has Property P, and all its primes are odd. Every element of A has a unique representation gp in which p is its largest prime factor. Hence

$$\mathcal{H}(A) = \sum_{g \ge 1} \frac{\mathcal{H}(Q_g)}{g}. \tag{6.1}$$

The term at $g = 1$ is $O(1)$, by Lemma 5.1.

Independently for all primes p, let V_p have geometric distribution

$$\mathbb{P}(V_p = j) = (1 - 1/p)p^{-j}, \qquad j = 0, 1, 2, \ldots.$$

Form the ordered list consisting of V_p copies of p, with primes in increasing order. A *prefix product* is the product of an initial finite segment of this list.

Lemma 6.1. *For $g > 1$, the probability that g is a prefix product is*

$$\mu_g = \frac{1}{g} \prod_{p < P^+(g)} (1 - 1/p) \asymp \frac{1}{g \log(2P^+(g))}. \tag{6.2}$$

Proof. Put $r = P^+(g)$. For primes $p < r$, the valuation V_p must equal $v_p(g)$. At r, the condition is $V_r \ge v_r(g)$, which has probability $r^{-v_r(g)}$. Independence gives the exact formula, and Mertens' product estimate gives the comparison. □

Assign to a prefix $g > 1$ the score

$$W_g = \log(2P^+(g))\, \mathcal{H}(Q_g).$$

Tonelli's theorem and (6.2) show that it is enough to prove

$$\mathbb{E} \sum_{\substack{g > 1 \\ g \text{ a prefix}}} W_g \ll 1. \tag{6.3}$$

Lemma 5.2 with $k = 0$ gives $W_g \ll 1$. More precisely, if $Q_g \neq \varnothing$ and $p_g = \min Q_g$, then

$$W_g \ll \frac{\log(2P^+(g))}{\log(2p_g)}. \tag{6.4}$$

We reserve p_g for this least terminal prime and use $r = P^+(g)$ for the prime coordinate of the prefix.

We may omit scores produced by the second or a later copy of a prime. Their total expectation is bounded by

$$C\sum_p \mathbb{E}(V_p - 1)_+ = C\sum_p \frac{1}{p(p-1)} < \infty.$$

The omitted copies remain in the list and in all later prefix products. Scores at primes below any fixed absolute threshold also have bounded total expectation. Thus only the first copy at each sufficiently large prime coordinate needs to be considered.

Lemma 6.2. *Suppose h and g are prefix products on one path, h occurs earlier, $Q_h, Q_g \neq \varnothing$, and $P^+(g) > p_h$. Then the residues of Q_g occupy at most one class from each opposite pair of nonzero classes modulo p_h. The terminal primes p_h from distinct nonempty prefixes on a path are distinct.*

Proof. We have $h \mid g$. If $p_h \mid g/h$, then $hp_h \in A$ divides every $gp \in A$, $p \in Q_g$. These are distinct elements because $P^+(g) > p_h$, contrary to primitivity. Thus $(g/h, p_h) = 1$. All primes in Q_g exceed p_h. For distinct $p, p' \in Q_g$, the congruence $p + p' \equiv 0 \pmod{p_h}$ would now imply

$$hp_h \mid gp + gp',$$

with both summands larger than hp_h, contradicting Property P. This proves the first assertion. If two distinct nested prefixes $h_1 \mid h_2$ had the same terminal prime p, then the distinct elements $h_1 p, h_2 p \in A$ would contradict primitivity. □

7. An adapted summation of prefix scores

We process every prime coordinate r in increasing order, including coordinates for which $V_r = 0$. At coordinate r, reveal V_r, retain all its copies in the prefix product, and score the first copy if $V_r \geq 1$. Let A_0 be the exponent in Lemma 5.2. The following scheduling rule uses selected terminal primes as congruence moduli, which we call *guards*.

After k guards have been activated, write M for their product and $\delta = 2^{-k}$. A *ready stage* starts at the first unprocessed prime at least R, where $R \geq M^{A_0}$ is determined by the previously revealed coordinates. Initially $k = 0$ and R is a fixed sufficiently large absolute threshold. A stage starts with no pending guard.

At a scored prefix g with $Q_g \neq \varnothing$, schedule p_g, retaining the smallest pending prime. If the newly scheduled minimum equals the current coordinate, activate it immediately after scoring. A guard already pending when its coordinate is reached is activated before revealing that coordinate, so that the coordinate belongs to the delay described below. The stage ends on activation.

Every finite pending prime is eventually reached, since all prime coordinates are processed. A ready stage that never activates therefore has no nonempty scored prefix. The guard selected in a completed stage arose from a previously visited prefix. Lemma 6.2 shows that the old guards impose the required opposite-class restrictions on every nonempty fibre in that ready stage. They are distinct and smaller than its current coordinate.

At a nonempty scored prefix g with $r = P^+(g)$, Lemma 5.2, applied with $Y = p_g \geq r$, gives

$$W_g \ll \delta(1 + k)\frac{\log r}{\log p_g}. \tag{7.1}$$

The harmless factors 2 in (6.4) are absorbed by the fixed lower threshold for r.

Lemma 7.1. *Conditional on the history at the start of a ready stage with k active guards, the expected total score in that stage is $O(2^{-k}(1 + k))$.*

Proof. Call a change of the pending minimum *large* if its logarithm decreases by a factor of at least 2. The first finite pending value is counted as large. At consecutive large changes, the ratio $\log r/\log p_g$ at least doubles, since r increases and the pending logarithm halves. Every such ratio is at most 1 before activation. Thus the sum of these ratios over all large changes is at most 2 on every path. Their total score is bounded by $C\delta(1+k)$, by (7.1).

For all other nonempty scored prefixes, let P_{old} be the pending prime known before revealing V_r. Then

$$\frac{\log r}{\log p_g} \le 2\frac{\log r}{\log P_{\mathrm{old}}}.$$

Interpret the right side as zero when no finite prime is pending. In that case the first nonempty prefix has already been assigned to the large-change class.

Both readiness at r and P_{old} are measurable before V_r is revealed. A first-copy score at r can occur only if $V_r \ge 1$, an event of conditional probability $1/r$. If P_* is the prime eventually activated, every ready coordinate is at most P_*, and every previous finite pending value is at least P_*. Consequently, pathwise,

$$\sum_{\substack{r \text{ ready}\\ P_{\mathrm{old}}<\infty}} \frac{1}{r}\frac{\log r}{\log P_{\mathrm{old}}} \le \frac{1}{\log P_*}\sum_{\substack{r\le P_*\\ r \text{ prime}}} \frac{\log r}{r} \ll 1. \tag{7.2}$$

For a stage without activation the sum is zero. Taking conditional expectation before each coordinate, and then applying Tonelli, proves the same $O(\delta(1+k))$ bound for all remaining scores. □

After activating P_*, put $M_{\mathrm{new}} = MP_*$. Process the remaining coordinates below $M_{\mathrm{new}}^{A_0}$ as a *delay*, then start the next ready stage at the first unprocessed prime at least that threshold. The old guards remain valid throughout the delay.

Lemma 7.2. *Conditional on the activation history, the expected total score in the delay is* $O(2^{-k}(1+k))$.

Proof. Equation (7.1) bounds every scored prefix during the delay by $C\delta(1+k)$, using $\log r/\log p_g \le 1$ and the old guards. Moreover $M^{A_0} \le R \le P_*$, so

$$M_{\mathrm{new}}^{A_0} = (MP_*)^{A_0} \le P_*^{A_0+1}.$$

Conditional on the activation history, all strictly later prime coordinates retain their independent geometric laws. Their expected number of first copies before the next ready stage is at most

$$\sum_{\substack{P_*<r\le P_*^{A_0+1}\\ r \text{ prime}}} \frac{1}{r} = O_{A_0}(1)$$

by Mertens' estimate. If P_* was pending before its coordinate, that coordinate is also in the delay and contributes at most one additional score. If the guard was scheduled at its own coordinate, that score was already counted in the ready stage. This proves the claim. □

Proposition 7.3. *Every odd primitive set A with Property P satisfies $\mathcal{H}(A) \ll 1$, with one absolute constant.*

Proof. Stage-start and activation histories are measurable at their respective prime coordinates. The conditional expectation statements in Lemmas 7.1 and 7.2 can therefore be applied on each event specifying these coordinates. All such coordinates are countable, and the rule never conditions on a future valuation.

Each completed ready stage adds one new guard. Summing the two lemmas over $k \ge 0$, and restoring the omitted small-prime and repeated-copy scores, gives

$$\mathbb{E}\sum_{\substack{g>1\\ g \text{ a prefix}}} W_g \ll 1 + \sum_{k\ge 0} 2^{-k}(1+k) < \infty.$$

Equation (6.3), Lemma 6.1, and the decomposition (6.1) complete the proof. □

8. Completion of the proof

Proof of Theorem 1.3. It is enough to bound $\mathcal{H}(B)$ uniformly for finite sets B with Property P. Let $C \subseteq B$ consist of the elements minimal under divisibility. Every $b \in B \setminus C$ has a proper divisor in C. Assign b to one such divisor c. No c receives two distinct elements, since it would divide their sum and both would exceed it. Every assigned b is at least $2c$. Hence

$$\mathcal{H}(B) \le \frac{3}{2}\mathcal{H}(C). \tag{8.1}$$

The set C is primitive.

For $v \ge 0$, define

$$C_v = \{c/2^v : c \in C,\ v_2(c) = v\}.$$

Each C_v is an odd primitive set with Property P. Proposition 7.3 supplies one constant C_{odd} bounding all their reciprocal sums, including the possible singleton $\{1\}$. Therefore

$$\mathcal{H}(B) \le \frac{3}{2}\sum_{v\ge 0} 2^{-v}\mathcal{H}(C_v) \le 3C_{\mathrm{odd}}.$$

This bound is independent of B. Apply it to the finite truncations of an arbitrary A, and use monotone convergence to obtain (1.2). □

9. An explicit finite prefix bound

We give a finite formulation of the argument in Section 7, with explicit bounds for the total expected score. Fix an odd primitive set $B \subseteq [2, N]$ with Property P. In this section, the fibres Q_g and scores W_g are defined using B in place of A. Independently for each prime $p \le N$, choose $X_p \in \{0, 1, \dots, N+1\}$ with

$$\mathbb{P}(X_p = j) = (1 - 1/p)p^{-j} \quad (0 \le j < N+1), \qquad \mathbb{P}(X_p = N+1) = p^{-(N+1)}.$$

For $0 \le a \le N+1$, summing the geometric series gives $\mathbb{P}(X_p \ge a) = p^{-a}$. Thus, for every $2 \le g \le N$, the event that g is a prefix has exactly the probability in Lemma 6.1. Every expectation in this section is a finite sum.

Apply the scheduling rule of Section 7, processing integer times $r = 1, \dots, N$ and changing the state only at prime coordinates. Charge an activation coordinate using its incoming set of guards. Let k_r be the number of guards activated before time r, and put

$$w_k = 2^{-k}(1+k), \qquad R_k = 2^{-k}(2k+4), \qquad a_r = R_{k_r} - R_{k_{r+1}}.$$

Since $R_0 = 4$ and $R_k - R_{k+1} = w_k$, the quantity a_r is w_{k_r} at an activation and zero otherwise. In particular,

$$\sum_{r=1}^{N} a_r \le 4. \tag{9.1}$$

Choose C_{tail} uniformly in the prime-tail estimate so that $W_g \le C_{\mathrm{tail}}$ and, for a ready first-copy prefix with $r = P^+(g)$,

$$W_g \le C_{\mathrm{tail}} w_{k_r} \frac{\log(2r)}{\log(2p_g)} \le 2C_{\mathrm{tail}} w_{k_r} \frac{\log r}{\log p_g}.$$

Increase the fixed initial threshold if necessary.

9.1. **The ready coordinates.** If the pending terminal prime before time r is P, define

$$U_r = w_{k_r}\frac{\log r}{\log P}, \qquad T_r = \frac{w_{k_r}}{\log P}\sum_{\substack{p<r\\ p \text{ prime}}}\frac{\log p}{p}.$$

Set both quantities to zero when no prime is pending. A pending prime is at least the current coordinate. Between activations its value can only decrease.

Write $W_r^{(1)}$ for the first-copy score charged at a ready time, and set it to zero at other times. Then

$$W_r^{(1)} \le 4C_{\text{tail}}\left(\mathbf{1}_{\{X_r\ge 1\}}U_r + U_{r+1} - U_r + 2a_r\right), \tag{9.2}$$

where the indicator is zero at a composite time. To verify this inequality, first suppose that no activation occurs. If a nonempty prefix is scored, the new pending prime is at most its terminal prime p_g, and consequently

$$U_{r+1} \ge w_{k_r}\frac{\log r}{\log p_g}.$$

On this event the first three terms in parentheses in (9.2) sum to U_{r+1}. If no score occurs, $U_{r+1} \ge U_r$. At an activation, $U_{r+1} = 0$, $U_r \le w_{k_r}$, and $a_r = w_{k_r}$, so the activation term covers both the lost potential and the score. During a delay there is no pending prime and these quantities are zero.

The state before time r depends only on earlier coordinates. Independence therefore gives, for prime r,

$$\mathbb{E}\left[\mathbf{1}_{\{X_r\ge 1\}}U_r\right] = \frac{1}{r}\mathbb{E}U_r.$$

Choose $C_{\log} > 0$ such that $\sum_{p\le x}(\log p)/p \le C_{\log}\log x$ for $x \ge 2$. The weighted prime sum and the monotonicity of the pending minimum imply

$$\mathbf{1}_{\{r \text{ prime}\}}\frac{U_r}{r} \le T_{r+1} - T_r + C_{\log}a_r. \tag{9.3}$$

Indeed, without activation the increment of the numerator accounts for $(\log r)/r$, and a decrease of the denominator increases the potential. At an activation, the numerator through time r is at most $C_{\log}\log P$, so the lost potential together with the current summand is at most $C_{\log}w_{k_r}$.

Every terminal prime of a nonempty fibre is at most N. Hence no prime is pending after time N, and both potentials begin and end at zero. Summing (9.3) and using (9.1) gives

$$\sum_{r=1}^{N}\mathbb{E}\left[\mathbf{1}_{\{X_r\ge 1\}}U_r\right] \le 4C_{\log}.$$

Summation of (9.2) now bounds the expected ready contribution by

$$C_{\text{tail}}(32 + 16C_{\log}). \tag{9.4}$$

9.2. **The delayed coordinates.** Suppose that the latest activation was at a prime P, with j guards active before that activation and old product M. Put $H = (MP)^{A_0}$. The inequality $M^{A_0} \le P$ gives $H \le P^{A_0+1}$. Mertens' estimate supplies an absolute constant $C_{\text{delay}} > 0$, depending only on the fixed A_0, such that

$$\sum_{\substack{P\le p<P^{A_0+1}\\ p \text{ prime}}}\frac{1}{p} \le C_{\text{delay}}$$

for every possible activation prime P. Define the accumulated mass after this activation by

$$B_r = w_j\sum_{\substack{P\le p<\min(r,H)\\ p \text{ prime}}}\frac{1}{p}.$$

Then $0 \le B_r \le C_{\mathrm{delay}} w_j$. For $k_r \ge 1$, so that $j = k_r - 1$, set

$$D_r = C_{\mathrm{delay}} R_{k_r} + C_{\mathrm{delay}} w_{k_r - 1} - B_r.$$

Before the first activation set $D_r = C_{\mathrm{delay}} R_0$. These quantities are nonnegative. At a delayed prime coordinate the increment of B_r is w_j/r, so D_r decreases by that amount. At a ready coordinate without activation it is constant.

At a new activation, the previous remaining term $C_{\mathrm{delay}} w_{k_r - 1} - B_r$ is nonnegative. The new value satisfies

$$D_{r+1} \le C_{\mathrm{delay}}(R_{k_r+1} + w_{k_r}) = C_{\mathrm{delay}} R_{k_r} \le D_r.$$

If the activation coordinate $r = P$ belongs to the delay, the new B_{r+1} already includes w_{k_r}/r. In this case the first inequality above holds with the stronger left side $D_{r+1} + w_{k_r}/r$, so the potential pays this coordinate as well. Thus the total predictable delayed cost is at most $D_1 = 4C_{\mathrm{delay}}$. Each delayed first-copy score is at most $C_{\mathrm{tail}} w_j$, and its occurrence has conditional probability $1/r$. Its total expected contribution is therefore at most

$$4 C_{\mathrm{tail}} C_{\mathrm{delay}}. \tag{9.5}$$

The finitely many initial coordinates contribute an absolute constant. The repeated-copy contribution is bounded using

$$\sum_p \mathbb{E}(X_p - 1)_+ \le \sum_p \frac{1}{p(p-1)} < \infty.$$

Combining these estimates with (9.4) and (9.5) yields a bound for $\mathbb{E} \sum_{g \text{ a prefix}} W_g$ independent of B and N. Lemma 6.1 and (6.1) then give the uniform reciprocal bound for finite odd primitive sets. Section 8 extends it to all sets with Property P.

10. The single-file formalization and provenance

The companion file `Erdos12AllParts.lean` imports only Mathlib. It contains the local dependency closure of the reciprocal-sum argument and the two attributed AlphaProof Nexus proof developments. Its public interface is as follows.

Entry point	Mathematical statement
`part_i`	A good infinite set has positive lower limit at square-root scale.
`part_ii`	The universal power-saving assertion is false.
`part_iii`	Every good infinite set has a convergent reciprocal sum.
`all_parts`	The conjunction of the three preceding answers.
`parameterwise_dense`	The set and eventual counting threshold may depend on the exponent.
`uniform_reciprocal_bound`	One constant bounds the reciprocal sum of every set with Property P.

All entry points belong to the namespace `Erdos12AllParts`. The definition `IsGood` retains the quantifiers and strict inequalities of the original Formal Conjectures statement.

The check uses Lean `4.35.0-rc2` and Mathlib commit `065356127b1dc0016f66b7283ce0ce2c4055aa55`. The prior proof files were taken from Google DeepMind's `alphaproof-nexus-results` repository at commit `0647711a71183c1ea492ad60860776617ce1ea88`. They are distributed under the Apache License, Version 2.0, with copyright 2025 Google LLC. The companion file retains these notices and the license. Their original environment used Lean 4.27.0; the adaptation isolates their namespaces, removes redundant options and imports, updates the image-cardinality API, and expands the interval-residue counting lemma into a direct proof.

The analytic formalization also retains the Euler–Maclaurin and Mertens developments from `PrimeNumberTheoremAnd`, obtained via `chen_theorem`, and the vendored large-sieve development from `goldbach-lean` [9, 10]. Their recorded source revisions and attributions are preserved in the

Lean header and the validation/provenance files; these components are likewise recorded under Apache License 2.0 in the supplied project.

The final checks compile the combined source, check each public entry point, and inspect its axiom dependencies. The only axioms reported are `propext`, `Classical.choice`, and `Quot.sound`. The file introduces no admitted theorem and does not obtain the first two results by importing placeholder catalog declarations. A separate check compares all three endpoints with the pinned original catalog statements. The validation receipt records the source digest and commands.

Acknowledgments and contribution statement

Parts (i) and (ii), including the constructions and the released formal proofs adapted here, are credited to the AlphaProof Nexus work of Google DeepMind and Tsoukalas, Kovsharov, Shirobokov et al. [11]. The present manuscript supplies a unified exposition and the recorded adaptation to the pinned Lean environment. The four named authors retain responsibility for the reciprocal-sum argument, the integration, and the final mathematical claims.

ChatGPT and Codex, developed by OpenAI, assisted with the mathematical development of the reciprocal-sum argument, checking calculations, literature searches, source integration, and manuscript preparation. The earlier reductions and analytic ingredients used in part (iii) are credited in the introduction and body of the proof.

References

[1] B. Alexeev, K. Barreto, Y. Li, J. D. Lichtman, L. Price, J. I. Shah, Q. Tang and T. Tao, *Primitive sets and von Mangoldt chains: Erdős Problem #1196 and beyond*, preprint (2026), arXiv:2605.00301v1.

[2] T. Bera and G. K. Viswanadham, *Restriction estimates with sifted integers*, preprint (2026), arXiv:2512.21640v3.

[3] P. Erdős and A. Sárközy, *On the divisibility properties of sequences of integers*, Proc. London Math. Soc. (3) **21** (1970), 97–101. doi:10.1112/plms/s3-21.1.97.

[4] Google DeepMind, *AlphaProof Nexus results: Erdős Problem 12(i)*, Lean source (2026), Apache License, Version 2.0. Repository commit `0647711a71183c1ea492ad60860776617ce1ea88`. https://github.com/google-deepmind/alphaproof-nexus-results/blob/0647711a71183c1ea492ad60860776617ce1ea88/APNOutputs/ErdosProblems/erdos_12.parts.i.lean.

[5] Google DeepMind, *AlphaProof Nexus results: Erdős Problem 12(ii)*, Lean source (2026), Apache License, Version 2.0. Repository commit `0647711a71183c1ea492ad60860776617ce1ea88`. https://github.com/google-deepmind/alphaproof-nexus-results/blob/0647711a71183c1ea492ad60860776617ce1ea88/APNOutputs/ErdosProblems/erdos_12.parts.ii.lean.

[6] H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), 119–134.

[7] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I. Classical Theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007.

[8] Plasma AI, *Composite-modulus large sieve for Property P*, research notes, Erdős Problems (2026), online edition, accessed 4 November 2026.

[9] *PrimeNumberTheoremAnd*, formalization project, https://github.com/AlexKontorovich/PrimeNumberTheoremAnd, commit `c6c73610b406689c4b58325cbe1342ecca25d755`; retrieved through https://github.com/lixiang90/chen_theorem, commit `11dcca8fc3ab340358503278c46d72aef072a6e7`.

[10] *goldbach-lean*, large-sieve formalization, https://github.com/subfish-zhou/goldbach-lean, commit `9d4c3a4010d42aedd9e7231bf94386019cfffccc`.

[11] G. Tsoukalas, A. Kovsharov, S. Shirobokov et al., *Advancing mathematics research with AI-driven formal proof search*, preprint (2026), arXiv:2605.22763v2, Appendix B.4.