Prime-Gap Dyadic Series: Perturbations, Exact Criteria and Certificates

Proofs, counterexamples and computations for Erdős Problem #251

Will Cook*

September 2026

Abstract

A convergent dyadic series with nonnegative integer coefficients can be made rational by adding nonnegative integer corrections on a sparse set. Given the original sequence, a finite prefix to retain, and any allowance $f(n)$ tending to infinity, we fix a set of permitted correction indices of upper Banach density zero and a nondegenerate interval above the original sum. Every target in this interval is attained by corrections supported in that set and eventually bounded by the prescribed function. Each fixed modulus eventually divides both the corrections and their cumulative sums, with a cutoff independent of the target. For the choice $f(n) = (\log(n+3))^{\varepsilon}$, $\varepsilon > 0$, we can also make the total variation distance between the empirical unnormalised block distributions, sampled at integer starts in $[X, 2X)$, tend to zero for lengths $o(\log\log X)$. The construction varies the split of a fixed total between adjacent coordinates and corrects the cumulative residue immediately before them. For prime gaps the cumulative positions remain asymptotic to $n\log n$, and the construction preserves the property that each fixed nonzero polynomial with integer coefficients in a fixed number of consecutive gaps vanishes only on a density-zero set. No primality conclusion is asserted for the cumulative positions.

For primes $p_0 = 2, p_1 = 3, \ldots$ and gaps $g_n = p_{n+1} - p_n$, summation by parts reduces the prime series to the gap series. The complete tails $T_N = \sum_{j\geq 1} g_{N+j}2^{-j}$ satisfy $T_{N+1} = 2T_N - g_{N+1}$; we use this recurrence to characterise rationality and to derive finite tests with explicit remainder bounds. Counterexamples show why several weaker coefficient conditions do not suffice. An exact finite certificate shows that a rational value of the prime series would have denominator at least $2^{39998} > 10^{12040}$. This denominator bound does not establish irrationality.

1 Introduction

Throughout, the primes are indexed from zero: $p_0 = 2$, $p_1 = 3$, $p_2 = 5$, and $g_n = p_{n+1} - p_n$. Erdős Problem #251 concerns

$$\Pi = \sum_{n\geq 0} \frac{p_n}{2^{n+1}} = 2/2 + 3/4 + 5/8 + 7/16 + \cdots,$$

Authorship and AI use. Will Cook built and directed the research infrastructure and reviewed the claims when he could. AI agents did most of the research and drafting. Cook did not independently verify every claim.

with the question recorded in [1, p. 94], [2, p. 62] and [4, p. 103]. The catalogue lists this as Problem #251 [22]. We construct rational gap series that retain specified congruences and block statistics of the actual prime gaps. No condition in the construction guarantees that the cumulative positions are prime. We also give exact tail criteria for irrationality of Π, but do not establish them for the actual gaps. The first values are

i	0	1	2	3	4	5	6	7	8	9
p_i	2	3	5	7	11	13	17	19	23	29
g_i	1	2	2	4	2	4	2	4	6	2

Thus $g_0 = 1$ and every later gap is even, since all primes after 2 are odd. The parity will matter in the criterion involving two consecutive tail differences. A second normalisation with denominator 2^i also occurs, and at every finite horizon it is exactly twice this one:

$$\sum_{i=0}^{n-1} \frac{p_i}{2^i} = 2 \sum_{i=0}^{n-1} \frac{p_i}{2^{i+1}}.$$

This factor-of-two identity preserves rationality. We use Π as normalised above throughout.

The perturbation theorem in Section 2 applies to arbitrary nonnegative integer coefficients. The independent actual-prime argument starts with summation by parts, identifies the complete tails and characterises their integral differences. Its finite certificates use an explicit prime bound, not the perturbation construction. The counterexamples in Section 8 identify coefficient properties that cannot supply the missing prime-gap estimate.

Notation. We use $\mathbb{N} = \{0, 1, 2, \ldots\}$, write φ for Euler's totient function, and write $\operatorname{den} q$ for the positive denominator of $q \in \mathbb{Q}$ in lowest terms. The notation $\operatorname{dist}(x, \mathbb{Z})$ means the distance from x to the nearest integer. A rational or real number is *integral* when it belongs to $\mathbb{Z}$. A sequence (a_n) is *eventually periodic with period* $h \geq 1$ when $a_{n+h} = a_n$ for every sufficiently large n. A set $A \subseteq \mathbb{N}$ has *density zero* when $|A \cap [1, x]| = o(x)$, and a property holds for *almost all* indices when its exceptional set has density zero. A sum over an empty range of indices is zero. A property holds *cofinally* when it holds at arbitrarily large indices; this does not imply that those indices have positive density.

2 Sparse congruence-preserving perturbations

Two adjacent corrections can have a fixed ordinary sum while their dyadic contribution varies. For example, the pairs $(0, 6), (2, 4), (4, 2), (6, 0)$ at $n, n+1$ have ordinary sum 6 and dyadic contributions $6, 8, 10, 12$ divided by 2^{n+2}. If the cumulative correction before them is 1, adding 1 at $n-1$ makes it even; changing the pair then preserves that residue.

Changing finitely many integer coefficients adds a dyadic rational and cannot alter rationality. To rationalise an irrational sum, we must therefore allow infinitely many corrections. The theorem repeats the paired operation with moduli divisible by every fixed integer eventually. Widely separated triples give sparse support. The range of choices at each triple is large enough to compensate for the intervening powers of two,

so the attainable sums fill an interval. The permitted index set and the interval are chosen before the target.

Density and block conventions. All intervals of indices contain integers, and log is natural unless a base is displayed. Upper Banach density zero means

$$\lim_{H\to\infty} \sup_{u\in\mathbb{N}} \frac{|S \cap [u, u+H)|}{H} = 0.$$

For integers $X \geq 1$ and $m \geq 1$, let $\mu_{a,X,m}$ be the empirical probability measure obtained by choosing an integer $n \in [X, 2X)$ uniformly and observing $(a_n, \dots, a_{n+m-1})$. Equal blocks are counted with multiplicity. We use $d_{\mathrm{TV}}(\mu, \nu) = \sup_B |\mu(B) - \nu(B)|$. For the same function of the block, bounded in absolute value by B_0, its means under the two measures differ by at most $2B_0 d_{\mathrm{TV}}(\mu, \nu)$. No rescaling of the coefficients is implicit.

Theorem 2.1 (sparse changes preserving congruences)**.** *Let $a_n \in \mathbb{N}$ and $A = \sum_{n\geq 0} a_n 2^{-n-1} < \infty$. For every cutoff K and every $f : \mathbb{N} \to \mathbb{R}$ tending to infinity, there are a set $S \subseteq [K, \infty)$ of upper Banach density zero and a nondegenerate interval $I \subset (A, \infty)$ with the following property. Every $r \in I$ has the form*

$$r = \sum_{n\geq 0} (a_n + e_n) 2^{-n-1},$$

where $e_n \in \mathbb{N}$ is supported on S, $e_n \leq f(n)$ eventually, and, for every integer $q \geq 1$, there is an N_q, chosen independently of r, such that

$$q \mid e_n \quad \text{and} \quad q \mid \sum_{i<n} e_i \qquad (n \geq N_q).$$

For every $\varepsilon > 0$, the construction can instead be chosen with $e_n \leq (\log(n+3))^{\varepsilon}$ eventually and $|S \cap [X, 2X)| = O_\varepsilon(X/\log\log X)$. In that case the empirical distributions of original and corrected unnormalised blocks of length $m(X) = o(\log\log X)$, sampled at the same integer starts in $[X, 2X)$, have total variation distance tending to zero, uniformly over r. Precisely the coupling error is at most $m|S \cap [X, 2X + m)|/X$; with $\|\Phi\|_\infty \leq B_0$ the bounded-test error is at most twice this quantity times B_0, also for tests depending on the starting index.

The hypotheses allow any nonnegative integer sequence of polynomial growth, but not $a_n = 2^n$, whose dyadic series diverges. The correction allowance can grow as slowly as $\log\log(n+3)$. A fixed bound is not possible here: eventual boundedness together with eventual divisibility by every fixed modulus would make the corrections eventually zero; their fixed total would then have to be divisible by every modulus and hence be zero. Nonnegative corrections could not change the sum. The set S contains every permitted correction index; the nonzero corrections may occupy a smaller, target-dependent subset. Requiring upper Banach density zero is stronger than requiring ordinary density zero. For example, the union of the integer intervals $[k!, k! + k)$, $k \geq 3$, has ordinary density zero but upper Banach density one. If $k! \leq x < (k+1)!$, at most $k(k+1)/2$ indices have occurred, so their proportion is at most $k(k+1)/(2k!) \to 0$. Nevertheless each interval is filled and their lengths are unbounded. The hypothesis rules out this concentration in long intervals, however far apart those intervals lie.

Proof. We first arrange the congruences, then choose sizes and spacings that respect f, and finally prove that every target in an interval is attained. Choose centres n_j, indexed by $j \geq 0$, and an initial $n_{-1} \geq K$, with spacings $s_j = n_j - n_{j-1} \geq 4$. Let M_j be positive moduli with $M_j \mid M_{j+1}$, and put $D_j = 2^{s_j} - 1$. The value of D_j is chosen so that the weighted range of one pair becomes a difference of two successive powers of $1/2$. Those differences will telescope in the interval argument. All these parameters will be chosen before the target.

Preserving the congruences. Let C_j denote the cumulative correction before the triple at n_j. Starting with $C_0 = 0$, define

$$c_j = (-C_j) \bmod M_j, \qquad C_{j+1} = C_j + c_j + M_j D_j.$$

For any choice $d_j \in \{0, \dots, D_j\}$, put

$$e_{n_j-1} = c_j, \qquad e_{n_j} = M_j d_j, \qquad e_{n_j+1} = M_j(D_j - d_j), \tag{1}$$

and put $e_n = 0$ off these triples. The spacing makes the triples disjoint. Their total is $c_j + M_j D_j$, independent of d_j, so C_j really is the cumulative correction $C_j = \sum_{i<n_j-1} e_i$ before its residue correction, for every choice of the d_j. The pair alone could not repair that cumulative residue: its total is already a multiple of M_j. The residue correction does, since $M_j \mid C_j + c_j$.

The preceding correction contributes $c_j 2^{-n_j}$ to the dyadic sum. It therefore changes the location of the attainable interval, although it does not change the range obtained by varying d_j.

Assume that each fixed q divides M_j eventually, and choose J with $q \mid M_J$. At index n_J, just after its residue correction, the cumulative sum is divisible by q. Both entries of the pair are also divisible by q, so every later cumulative sum through that stage remains divisible by q. Inductively $q \mid C_j$ and $q \mid M_j$ imply $q \mid c_j$, since $M_j \mid C_j + c_j$. Thus every subsequent residue correction, pair entry and zero entry is divisible by q, as is every intermediate cumulative sum. We may take $N_q = n_J$. The first residue correction need not itself be divisible by q, and its coordinate $n_J - 1$ is deliberately outside this eventual range.

Keeping the corrections small and sparse. The factorial modulus must grow to include every divisor, and the spacing must grow to make the support sparse. Since f need not be monotone, we first replace it by a nondecreasing lower bound. Define

$$h(n) = \inf_{m \geq n} \min(f(m), m).$$

This finite real number is nondecreasing in n, satisfies $h(n) \leq n$, and tends to infinity. Choose $n_{-1} \geq K$ with $h(n_{-1}) \geq 32$ and recursively set

$$\begin{aligned} k_j &= \max\{k \in \mathbb{N} : k \geq 2,\ k!2^{k+2} \leq h(n_{j-1})\}, \\ M_j &= k_j!, \qquad s_j = k_j + 2, \qquad n_j = n_{j-1} + s_j. \end{aligned} \tag{2}$$

The maximum exists: $k = 2$ is eligible and $k!2^{k+2}$ tends to infinity. Since h is nondecreasing, the integers k_j are nondecreasing; since $n_j \to \infty$, they tend to infinity. The

factorials form a divisibility chain and eventually contain every fixed divisor, as do the distinguished terms in [20, Proposition 8]. At each of the three coordinates n in stage j,

$$0 \le e_n \le M_j 2^{s_j} \le h(n_{j-1}) \le f(n), \qquad e_n \le n_{j-1} < n.$$

Here the residue correction is smaller than M_j and each pair entry is at most $M_j D_j$. Thus every correction series converges by comparison with $\sum_n n2^{-n-1}$.

Let $S = \bigcup_j \{n_j - 1, n_j, n_j + 1\}$. It is independent of the choices d_j. To see upper Banach density zero, fix $H > 0$. Outside a finite initial segment, consecutive centres are at least H apart, since $s_j \to \infty$. An interval of length L therefore meets at most $3(L/H + 2)$ late permitted correction indices, plus the fixed finite number of early indices. Divide by L, take the supremum over interval positions, then let $L \to \infty$ and $H \to \infty$.

Filling an interval. The preceding choices satisfy the size and congruence requirements. We now show that their weighted sums have no gaps. Put $w_j = M_j 2^{-n_j - 2}$ and

$$F_j = \sum_{i \ge j} D_i w_i, \qquad \beta = \sum_{j \ge 0} (c_j 2^{-n_j} + D_j w_j).$$

The size estimates just proved give convergence of both series, $F_j \to 0$, and $F_0 > 0$. The constant β is positive and independent of the choices d_j. The complete weighted correction, including every residue correction, is exactly

$$\sum_n e_n 2^{-n-1} = \beta + \sum_j d_j w_j. \tag{3}$$

For $i > j$, we have $M_i \ge M_j$ and

$$D_i w_i = M_i (2^{-n_{i-1}-2} - 2^{-n_i - 2}) \ge M_j (2^{-n_{i-1}-2} - 2^{-n_i - 2}).$$

For $J > j$, the finite sum is therefore at least $M_j (2^{-n_j - 2} - 2^{-n_J - 2})$. Since $n_J \to \infty$, letting $J \to \infty$ yields $F_{j+1} \ge w_j$. Adjacent intervals $[dw_j, dw_j + F_{j+1}]$, for $0 \le d \le D_j$, therefore overlap or touch. Their union is $[0, F_j]$; no strict inequality is needed. For any $y \in [0, F_0]$, choose successive integers d_j leaving each remainder in $[0, F_{j+1}]$. Since these capacities tend to zero, the resulting series has sum y. The finite choice set $\{0, w_j, \dots, D_j w_j\}$ has maximum and diameter $D_j w_j$ and largest successive gap w_j. The size estimates give $\sum_j D_j w_j < \infty$; together with $w_j \le F_{j+1}$, this verifies the hypotheses of the covering lemma of Crmarić–Kovač [27, Lemma 4]. Fridy's generalised-base lemma [9, Lemma, p. 194] and the reciprocal-choice argument in Kovač–Tao [11, Lemma 5.1] are antecedents. The finite-choice argument needs no monotonicity of the weights: increasing n_j alone would not ensure that $w_j = M_j 2^{-n_j - 2}$ decreases, since the moduli also grow. Only the covering inequality $w_j \le F_{j+1}$ is used. Equation (3) therefore realises every r in the fixed interval $I = (A + \beta, A + \beta + F_0) \subset (A, \infty)$.

Preserving growing blocks. The spacing controls how many coordinates are changed, whereas the product $M_j 2^{s_j}$ controls their size. For a prescribed $\varepsilon > 0$, allow exponents $\varepsilon/4$ for M_j and $\varepsilon/2$ for 2^{s_j}. Their sum is less than ε, leaving room in the correction bound. Replace (2) by

$$\begin{aligned} s_j &= \left\lfloor \frac{\varepsilon}{2} \log_2 \log(n_{j-1} + 3) \right\rfloor, \qquad n_j = n_{j-1} + s_j, \\ k_j &= \max\{k \in \mathbb{N} : k \ge 2,\ k! \le (\log(n_{j-1} + 3))^{\varepsilon/4}\}, \qquad M_j = k_j!. \end{aligned} \tag{4}$$

Take the initial cutoff so large that $s_j \geq 4$ and the maximum is nonempty from the first step. Again k_j is nondecreasing and tends to infinity. Keep $D_j = 2^{s_j} - 1$ and the same recursion for c_j. All triple entries obey

$$e_n \leq M_j 2^{s_j} \leq (\log(n_{j-1} + 3))^{3\varepsilon/4} \leq (\log(n + 3))^{\varepsilon}.$$

Enlarge the initial cutoff, if necessary, so that $(\log(n_{j-1} + 3))^{3\varepsilon/4} \leq n_{j-1}$ from the first step. Then every correction again satisfies $e_n \leq n$, and the common bound $\sum_n n2^{-n-1} < \infty$ gives convergence of F_j and β. The congruence and interval arguments therefore apply to this schedule. Also s_j is comparable, with constants depending on ε, to $\log\log(n_{j-1} + 3)$, and $s_j = o(n_{j-1})$. Hence consecutive centres near X are separated by at least a constant multiple of $\log\log X$, giving $|S \cap [X, 2X)| = O_\varepsilon(X/\log\log X)$. The estimate on $[X, 3X)$ follows by covering it with $[X, 2X)$ and $[2X, 4X)$.

Finally sample both length-m blocks at the same uniform integer start in $[X, 2X)$. They can differ only if that block meets S. Each changed coordinate belongs to at most m such blocks, so the total variation distance is at most $m|S \cap [X, 2X + m)|/X$. If a test bounded by B_0 also depends on the starting index, its two values still agree on every unchanged block at that same index and differ by at most $2B_0$ elsewhere. Its mean difference is therefore bounded by $2B_0 m|S \cap [X, 2X + m)|/X$ as well. This conclusion uses the coupling, not just the marginal distributions of the blocks. If $m \leq X$, then $[X, 2X + m) \subseteq [X, 3X)$, so the preceding support estimate gives

$$\frac{m\,|S \cap [X, 2X + m)|}{X} = O_\varepsilon\left(\frac{m}{\log\log X}\right).$$

For integer $m = m(X) = o(\log\log X)$, the condition $m \leq X$ holds for all large X and the bound tends to zero. This comparison uses unnormalised blocks of coefficient values. It requires neither a limiting distribution nor any randomness or independence in the original sequence. □

Identical marginal distributions alone would not control an index-dependent test. For example, let $a_n = n \bmod 2$ and $b_n = 1 - a_n$. For every even X, their one-coordinate empirical distributions on $[X, 2X)$ are equal, but the test $\mathbf{1}_{\{x = n \bmod 2\}}$ has mean 1 for (n, a_n) and 0 for (n, b_n). This example is not a sparse perturbation; it isolates why the proof couples blocks at the same index.

The comparison concerns absolute, not relative, error in event frequencies. For example, it applies to fixed block lengths and to $m(X) = \lfloor\sqrt{\log\log X}\rfloor$ for large X, but makes no vanishing-error assertion for lengths comparable to $\log\log X$. An $o(1)$ change may erase an event whose probability tends to zero; relative preservation requires an error small compared with that probability. The comparison is of finite coefficient blocks, not complete weighted tails: coefficients beyond an unchanged block can still change its tail.

The printed construction and the formal construction use different corrections. Their statements and quantifier order are compared in Appendix C.

Corollary 2.2 (target intervals arbitrarily close to the original sum)**.** *For the sparse theorem for an arbitrary sequence, the target interval can additionally be required to lie in $(A, A + \eta)$ for any prescribed $\eta > 0$.*

Proof. In the general schedule, every supported entry satisfies $e_n \le n$. Choose an integer $K' \ge K$ sufficiently large that $\sum_{n\ge K'} n2^{-(n+1)} = (K'+1)2^{-K'} < \eta$, and start the construction beyond K'. This still leaves the originally prescribed prefix unchanged. The complete interval of corrections has positive lower endpoint and upper endpoint at most this sum. Thus the interval may be chosen as close to A as desired; this does not say that one fixed allowance permits every target above A. □

3 Context and mathematical dependencies

Relation to prior work. The passage from the primes to their gaps is already public. Tao posted on the problem's forum thread on 7 October 2025 that summation by parts makes the question equivalent to irrationality of $\sum_n (p_{n+1} - p_n)2^{-n}$, and named the shape of the missing input as a sufficiently quantitative and uniform prime-tuples hypothesis giving statistical control of the binary expansion of about $\log\log n$ consecutive gaps [23, comment of 7 October 2025]. Theorem 4.3 below is that reduction in exact form, with the endpoint retained, with convergence supplied by an elementary bound, and with the statement checked by the Lean kernel. The elementary bound replaces the prime number theorem in this reduction. The prime number theorem is still used in the different argument establishing $P_n \sim n\log n$ in Corollary 8.6; that application also uses Schlage-Puchta's fixed-polynomial theorem.

A sufficiently uniform Hardy–Littlewood hypothesis gives conditional results of a different kind. Land's draft states conditional irrationality [13, Conjecture 1, Theorem 2]; Ringer's draft states conditional normality in each integer base [26]. For a fixed base $b \ge 2$, the number in the latter statement is $\sum_{n\ge 0} p_n b^{-(n+1)}$, and normality is to base b. Changing b changes the number; this is not absolute normality of a single constant. The authors supply formalisation material, which has not been independently rebuilt for this revision. The uniform hypothesis in [12, Conjecture 1.3] counts prime translates of every nonempty admissible tuple of $k \le (\log\log x)^3$ distinct integer shifts in $[0, (\log x)^2]$. Its main term is the tuple's singular series times $\int_2^x (\log t)^{-k}\,dt$, with absolute error at most $Cx^{1-\delta}$. The logarithmic integral cannot simply be replaced by its leading asymptotic while retaining that error bound. Here admissible means that the shifts omit a residue class modulo each prime. The positive constants C, δ and the starting threshold are chosen before x and the tuple. Qualitative asymptotics for each fixed tuple do not provide this uniformity.

In the version rechecked on 18 September 2026, Ringer also gives weaker assumptions for the joint positions of the first L primes after a prime basepoint, with L growing on the order of $\log\log X$. These are not single-gap marginal estimates. Here X measures prime values: the basepoints are primes in $(X, 2X]$, sampled uniformly. One assumption controls a weighted sum of adverse tuple-count errors; another permits bounded domination by a comparison law rather than total-variation approximation. In the weighted error condition, the adverse direction alternates with the order in inclusion–exclusion: undercounts matter at some orders and overcounts at others. An upper sieve estimate alone does not supply these assumptions, and none is used below.

Our index band $[X, 2X)$ corresponds to primes in $[p_X, p_{2X})$. The prime number theorem gives $\pi(2p_X) = 2X + o(X)$, so this set and $(p_X, 2p_X]$ have symmetric difference of

size $o(X)$ and each has $X + o(X)$ elements. Their uniform measures share mass equal to the intersection size divided by the larger sample size. This tends to 1, so their total variation distance is $o(1)$. The comparison uses the same uniformly bounded observable, with the same normalisation at common basepoints. It does not by itself transfer the stronger weighted, relative, or quantitative estimates for growing prime patterns.

Erdős stated on p. 93 of the 1958 article that $\sum_{n\geq 1} p_{n-1}^k/n!$ is irrational for every $k \geq 1$, and wrote there that the proof for $k > 1$ is complicated enough that only the case $k = 1$ is printed [1, pp. 93–95]. A proof of the full family appears in Schlage-Puchta's Theorem 3, which gives the stronger statement that $1, S_0, S_1, S_2, \ldots$ are linearly independent over $\mathbb{Q}$, where $S_k = \sum_{n\geq 1} p_{n-1}^k/n!$ in our zero-based indexing [18, Theorem 3].

On page 103 of his 1988 problem paper Erdős separately stated the fixed-denominator problem: he could not prove that $\sum_{n\geq 1} p_{n-1}^k/2^n$ is irrational for every $k \geq 1$, and wrote that the case $k = 1$ was probably already very difficult. He also stated the variable-denominator expectation that $\sum_{n\geq 1} p_{n-1}/(d_1 \cdots d_n)$ is irrational whenever $d_n \geq 2$ and $d_n = o(p_{n-1})$ [4, p. 103]. That expectation is false: on 15 April 2026 Kovač posted a note, with the printed attribution ChatGPT 5.4 Pro (orchestrated by Vjeko Kovač), constructing such a sequence (d_n) for which the sum is exactly 1 [24, Theorem 1 and proof, pp. 1–2].

Section 3 of the 1958 article concerns variable product denominators, not the dyadic denominator sequence [1, pp. 96–97]. Its printed growth condition (5) uses nonstandard asymptotic typography; we neither use that condition as a hypothesis nor assign it a modern quantified interpretation. The simple endpoint example is unambiguous: if $G_n = \prod_{j=1}^{n}(p_{j-1} + 1)$ and $G_0 = 1$, then

$$\frac{p_{n-1}}{G_n} = \frac{1}{G_{n-1}} - \frac{1}{G_n}, \qquad \sum_{n\geq 1} \frac{p_{n-1}}{G_n} = 1.$$

Neither this variable-denominator example nor the counterexample in [24] addresses fixed dyadic denominators.

There is a related theorem with dyadic denominators and bounded coefficients. If $P(m)$ denotes the largest prime factor of m, Erdős and Pomerance proved that

$$\sum_{m\geq 2} \frac{\mathbf{1}_{\{P(m)>P(m+1)\}}}{2^m}$$

is irrational [3, §7, p. 320]. Erdős and Graham record the complementary indicator on p. 62: equality of the two largest prime factors is impossible for consecutive integers, so its series is $1/2$ minus the displayed one and is irrational as well. That theorem concerns a sequence of zeros and ones comparing largest prime factors, whereas the numerators in Π are unbounded.

Schlage-Puchta gives a necessary condition for rationality of a number formed by concatenating integer blocks in a base $b \geq 2$ that is not a proper power, under the stated monotonicity and growth assumptions [18, Theorem 2, p. 1]. The positions of those blocks depend on their lengths; they are not the fixed positions of the coefficients in our series. This concatenation theorem is not used here. The input we use is his Lemma 4. For each fixed nonzero polynomial $F \in \mathbb{Z}[x_0, \ldots, x_k]$, the actual prime gaps

satisfy $F(g_n, \dots, g_{n+k}) \neq 0$ outside a set of indices of natural density zero [18, Lemma 4, pp. 5–6]. Its proof uses Selberg's sieve. Section 8 draws two consequences from it.

Theorems 8.1 and 2.1 fill intervals of attainable values. The first is a binary expansion. For the second, the direct reference is Crmarić–Kovač's finite-choice covering lemma [27, Lemma 4]. For finite nonnegative choice sets, each with at least two elements and with summable maxima, their lemma gives a single interval when each stage's largest successive gap is at most the sum of the later diameters. They apply it to product-denominator series. In the notation of the proof of Theorem 2.1, the choices are $\{0, w_j, \dots, D_j w_j\}$, and the required inequality is $w_j \leq \sum_{i>j} D_i w_i$. Fridy's generalised-base lemma [9, Lemma, p. 194] is an antecedent with nonincreasing weights; the finite-choice form avoids that additional assumption. Kovač–Tao use sets of reciprocal choices [11, Lemma 5.1, Theorem 2.5]. Van Doorn–Kovač use finite subsums and a distinguished subsequence of denominators: each distinguished denominator is divisible by every earlier denominator, and each positive integer divides some denominator [20, Lemma 7, Proposition 8]. Hence every fixed positive integer divides all sufficiently late distinguished denominators. It is this eventual divisibility, not a condition on every denominator in their sequence, that our factorial moduli share. At a distinguished cutoff, their finite sums are all multiples of the reciprocal of the last denominator. Summing the covering inequalities between successive distinguished cutoffs gives the hypothesis of their Lemma 7 for the whole finite prefix. That lemma leaves no gaps in this grid between zero and the full finite sum. Once the target's denominator divides that last denominator and the target lies in this range, a finite representation follows. Their proposition thus represents rational targets in a half-open interval by finite subsums, whereas our infinite choices attain every real target in a fixed interval.

A further comparison is van Doorn's exchange $\{21d, 28d\} \leftrightarrow \{20d, 30d\}$: both reciprocal sums are $1/(12d)$, while the ordinary sums are $49d$ and $50d$ [29, proof of Theorem 3]. Our paired corrections preserve the ordinary sum and vary the dyadic contribution instead. This is an analogy between conserved quantities, not an application of his partition theorem. We claim no new interval-covering principle.

For binary subsums, each term is either retained or omitted; the set of all resulting sums is called the achievement set. The powers 2^{-n}, $n \geq 1$, fill $[0, 1]$ by binary expansion. In contrast, the powers 3^{-n} omit $(1/6, 1/3)$, since all terms after the first sum to $1/6$. More complicated subsum sets can contain intervals even when the simple covering inequality fails at some indices. Bartoszewicz–Filipczak–Szymonik use a central run of attainable integer block sums to obtain interior in a multigeometric family [32, Theorem 2]. Prus-Wiśniowski–Ptak construct Cantorvals for which the overlap indices $\{n : a_n \leq \sum_{i>n} a_i\}$ have density zero [31, Theorem]. These compact sets equal the closure of their interiors, and both endpoints of every nondegenerate interval component are accumulation points of one-point components. Thus a finite-stage gap between possible subsums does not by itself rule out an interval elsewhere in the full set of subsums. Nor is it enough merely to know that neither comparison between a term and its remaining tail holds eventually. The survey [30, §1] records the surrounding classification. Our variable-digit proof verifies covering at every stage and does not depend on a classification theorem for binary subsums. Nitecki's preprint [36, Theorem 14] gives an ex-

position explicitly crediting the Guthrie–Nymann classification; its title and numbering differ from the published Monthly article.

The inequality pattern alone is insufficient even when its index set is specified exactly: Marchwicki–Miska's construction [38, Theorem 2.1], with the repaired proof in [37], can realise any infinite strict term-dominating index set with a Cantor achievement set. The repair preserves the original uniqueness conclusion; the latter paper also gives a simpler proof of a weaker statement without uniqueness. Nowakowski's Star Procedure [39, Definition 2, Theorem 3.1] is a different sufficient route to a Cantorval, requiring every stage of a positivity recursion. These comparisons explain why our proof checks overlap at every stage instead of relying only on the set of indices at which overlap occurs.

Automatic sequences and the limit of the analogy. Adamczewski–Drmota–Müllner compute logarithmic densities for each fixed automatic sequence along primes. Their Theorem 1.2 also gives a sufficient condition for natural densities to exist and be rational; Theorem 1.4 reduces the logarithmic-density calculation to integers coprime to a suitable fixed modulus [28, Theorems 1.2, 1.4]. The general logarithmic densities are not all rational. The fixed finite-alphabet automaton is essential: the result is not a theorem about unbounded consecutive prime gaps, their complete dyadic tails, or a growing family of residue or carry automata. Any such application would have to specify the automaton, prove that it computes the required observable, and supply uniform errors as its state space grows. We do not infer any of those steps from the density theorem.

The uniformity required from prime statistics. Kuperberg's large-set singular-series estimates allow a specified growing cardinality, not arbitrary dimension [12, Theorem 1.1]. Her arithmetic-progression and smooth-weight estimates are a useful fixed-parameter comparison [33]; they do not by themselves give a law for ordered consecutive-gap blocks. Jha's revised Poisson-tail preprint [34] concerns growing short-interval prime counts under a strong Hardy–Littlewood hypothesis. Counting primes in an interval does not by itself control the weighted differences of consecutive gaps used by our tail criteria. We do not infer the required estimates for growing blocks from these results or from a fixed-dimensional limit theorem.

At the cited revision, Problem #251 appears as an unproved statement in the *Formal Conjectures* repository [25]. Its zero-based Lean sum starts with the zeroth prime over 2^0, so it is twice the displayed normalisation and has equivalent irrationality status; its proof is `sorry`.

The strategy. An ordinary binary expansion uses digits in $\{0, 1\}$ and is rational exactly when those digits are eventually periodic. The primes p_n are coefficients, not binary digits; carrying them changes the sequence to which this criterion applies. Even bounded integer coefficients can be nonperiodic and have a rational dyadic sum, as Proposition D.6 illustrates. A different classical approach uses the growth of denominators in series of unit fractions. Erdős and Straus named a sum $\sum_k 1/a_k$ over a strictly increasing sequence of positive integers an *Ahmes series* [6]; for such a series the condition $a_k^{1/2^k} \to \infty$ is sufficient for irrationality. Its growth scale is sharp: replacing divergence by a sufficiently large fixed lower threshold does not suffice, since shifted Sylvester se-

quences grow like C^{2^k} for arbitrarily large C and have rational reciprocal sum. This is not a necessary condition for irrationality. Both statements and their attribution are recorded in the introduction of Kovač and Tao [11, §1]. Splitting each term $p_i/2^{i+1}$ into p_i copies of $2^{-(i+1)}$ writes Π as a sum of unit fractions with repetitions, and the denominators occurring in it are exactly the powers of two. Even ignoring the repetitions the growth hypothesis fails, since $(2^n)^{1/2^n} \to 1$, and the growth criterion does not apply.

For the dyadic series below, we instead examine the reduced denominators of the complete tails under a rationality assumption. Rationality of the sum is equivalent to an eventual integrality condition on differences of those tails, and that equivalence uses nothing about the numerators beyond the fact that they are integers. The condition does not by itself force the numerators to repeat: Proposition 8.10, applied to $K_n = n$, produces the integer sequence $\kappa_n = n-1$, which is unbounded and hence not eventually periodic, and whose dyadic sum is zero.

Proof dependencies. The sparse theorem is independent of primes: residue correction and finite-choice covering give the interval, and counting changed starts gives its statistical consequences. Applying this construction to prime gaps then uses Schlage-Puchta's lemma and the prime number theorem. The recurrence classification needs integer coefficients and Euler's congruence, not any prime-tuple conjecture. The signed certificate bounds one tail difference using an explicit majorant and exact finite prime data. To prove irrationality by this local criterion, it would suffice to establish the simultaneous inequalities of Section 10 arbitrarily late for every positive shift. We do not prove that condition or assert that it is necessary. The appendices give the finite computations, further equivalent criteria, complete counterexamples and formal references. The short paper gives the main construction and the prime-tail criterion without these extensions.

4 Summation by parts, with the endpoint retained

Summation by parts expresses a weighted sum of a sequence in terms of its first value, its consecutive differences and one final term. We use finite sums first, so the identity needs neither growth nor convergence assumptions. The linked definitions are the dyadic partial sum and the weighted sum of consecutive differences. The formula below writes both sums explicitly.

Proposition 4.1 (finite summation by parts)**.** *For every rational sequence P and every $n \geq 0$,*

$$\sum_{i=0}^{n} \frac{P(i)}{2^{i+1}} = P(0) + \sum_{i=0}^{n-1} \frac{P(i+1) - P(i)}{2^{i+1}} - \frac{P(n)}{2^{n+1}}.$$

Proof. At $n = 0$ both sides equal $P(0)/2$. When n is replaced by $n+1$, the change on the right is

$$\frac{P(n+1) - P(n)}{2^{n+1}} - \frac{P(n+1)}{2^{n+2}} + \frac{P(n)}{2^{n+1}} = \frac{P(n+1)}{2^{n+2}}.$$

This is the next term on the left, which proves the identity by induction. □

Specialising to $P(i) = p_i$, whose first value is $p_0 = 2$, and writing $g_i = p_{i+1} - p_i$ for the zero-based gaps, gives the reformulation.

Theorem 4.2 (prime-gap reformulation)**.** *Let $p_0 = 2, p_1 = 3, \ldots$ be the primes in increasing order and $g_i = p_{i+1} - p_i$. For every $n \geq 0$,*

$$\sum_{i=0}^{n} \frac{p_i}{2^{i+1}} = 2 + \sum_{i=0}^{n-1} \frac{g_i}{2^{i+1}} - \frac{p_n}{2^{n+1}}.$$

The leading 2 is the first prime, not a normalising constant. At $n = 2$, for instance, the left side is $2/2 + 3/4 + 5/8 = 19/8$ and the right side is $2 + (1/2 + 2/4) - 5/8 = 19/8$.

4.1 The infinite identity and the irrationality equivalence

Write

$$u_n = \frac{p_n}{2^{n+1}}, \qquad v_n = \frac{g_n}{2^{n+1}}$$

for the terms of the prime series and of the gap series, so that $\sum_{n\geq 0} u_n = \Pi$. The termwise identity $v_n = 2u_{n+1} - u_n$ is the dyadic discrete derivative. It expresses each gap term as an integer combination of two consecutive prime terms, so once (u_n) is summable the gap series can be summed by rearranging two copies of the prime series.

Theorem 4.3 (infinite prime-gap identity)**.** *Both series converge and*

$$\sum_{n\geq 0} \frac{p_n}{2^{n+1}} \;=\; 2 + \sum_{n\geq 0} \frac{g_n}{2^{n+1}}.$$

Proof. The bound $p_n \leq 1250(n+1)^4$ of Appendix A makes (u_n) summable, and $0 < g_n \leq p_{n+1}$ then makes (v_n) summable. The shifted sequence (u_{n+1}) is summable, so summing $v_n = 2u_{n+1} - u_n$ and using $u_0 = 1$ gives

$$\sum_{n\geq 0} v_n = 2\left(\sum_{n\geq 0} u_n - u_0\right) - \sum_{n\geq 0} u_n = \sum_{n\geq 0} u_n - 2. \qquad \square$$

The prime number theorem is needed neither for this convergence nor for the identity. Its asymptotic $p_n \sim n \log n$ is used in Corollary 8.4 and in the argument counting repeated tail values [17, Chapter 6].

Corollary 4.4 (exact irrationality reformulation)**.** Π *is irrational if and only if* $S = \sum_{n\geq 0} g_n 2^{-(n+1)}$ *is irrational. The corresponding zero-based series with denominator* 2^n *equals* $4 + 2S$ *and has the same irrationality status.*

Proof. The identity just proved gives $\Pi = 2 + S$. Adding a rational constant or multiplying by a nonzero rational number preserves rationality and irrationality. $\square$

Concretely, $\Pi = 3.674643966\ldots$ is irrational if and only if $S = 1.674643966\ldots$ is. Neither irrationality statement is established here.

5 The tail recurrence and the exact criteria

Rescaled tails turn the series question into a recurrence. Suppose $\sum_{i\ge 0} a_i 2^{-(i+1)}$ converges with every a_i an integer, and rescale its tails by putting

$$T_N = 2^{N+1} \sum_{i>N} \frac{a_i}{2^{i+1}} = \sum_{j\ge 1} \frac{a_{N+j}}{2^j}.$$

Two facts follow immediately. First $T_{N+1} = 2T_N - a_{N+1}$, so advancing one index doubles the rescaled tail and subtracts a single coefficient. Second $T_0 = 2\sum_{i\ge 0} a_i 2^{-(i+1)} - a_0$, so the sum is rational exactly when T_0 is. This section studies that recurrence, assuming nothing about the coefficients beyond the fact that they are integers, and resumes the prime-gap instance at the end.

Definition 5.1. Let $g : \mathbb{N} \to \mathbb{Z}$ and $T : \mathbb{N} \to \mathbb{Q}$ or $T : \mathbb{N} \to \mathbb{R}$. Say T satisfies the *dyadic tail recurrence* with *coefficients* g when $T_{N+1} = 2T_N - g_{N+1}$ for every N. Write $\sigma_h(N) = T_{N+h} - T_N$ for the *shift* of length h at N, and call a real number *integral* when it belongs to $\mathbb{Z}$.

For a complete-tail example, take $g_n = 2$ for odd $n \ge 1$ and $g_n = 4$ for even $n \ge 1$. Its complete tail at index 0 is $T_0 = (2/2 + 4/4)/(1 - 1/4) = 8/3$, and the tails alternate between $8/3$ and $10/3$. Thus $\sigma_1(N)$ is always $2/3$ or $-2/3$, whereas $\sigma_2(N) = 0$ for every N. Nonintegrality at one prescribed shift length does not imply irrationality, even for positive even coefficients and genuine tails.

Modulo integers, the recurrence is repeated doubling: the coefficient term does not affect the fractional part. An integral shift means that the fractional part has returned to an earlier value. Iterating the recurrence h times makes this observation explicit: it multiplies T_N by 2^h and accumulates an integer. Define $B_{0,N} = 0$ and $B_{h+1,N} = 2B_{h,N} + g_{N+h+1}$, so that $B_{h,N} = g_{N+1}2^{h-1} + \cdots + g_{N+h}$ is the weighted integer sum accumulated in those h steps.

Theorem 5.2 (block identity)**.** *For every N and h,*

$$T_{N+h} = 2^h T_N - B_{h,N}, \qquad \text{hence} \qquad \sigma_h(N) = (2^h - 1)\, T_N - B_{h,N}.$$

Proof. The case $h = 0$ follows from $B_{0,N} = 0$. If the first identity holds at h, then

$$T_{N+h+1} = 2(2^h T_N - B_{h,N}) - g_{N+h+1} = 2^{h+1} T_N - B_{h+1,N}.$$

Subtracting T_N gives the second identity. □

Subtracting the tail recurrences at $N + h$ and N gives

$$\sigma_h(N+1) = 2\sigma_h(N) - (g_{N+h+1} - g_{N+1}), \tag{5}$$

which is the recurrence for the tail differences. Since $B_{h,N}$ is an integer, Theorem 5.2 also gives the integrality criterion: $\sigma_h(N)$ is integral if and only if $(2^h - 1)T_N$ is integral. For fixed h in this general recurrence, T_N therefore determines whether the shift is integral, regardless of the later integer coefficients. For a complete tail, however, changing those

coefficients also changes T_N; the observation does not permit arbitrary changes to prime gaps.

For rational $T_N = u/v$ in lowest terms, the recurrence gives $T_{N+1} = (2u - g_{N+1}v)/v$. The numerator has greatest common divisor $\gcd(2, v)$ with v, because $\gcd(u, v) = 1$. Thus

$$\operatorname{den}(T_{N+1}) = \frac{\operatorname{den}(T_N)}{\gcd(2, \operatorname{den}(T_N))}, \qquad \sigma_h(N) \in \mathbb{Z} \iff \operatorname{den}(T_N) \mid 2^h - 1. \tag{6}$$

Each even denominator loses exactly one factor of 2 and an odd denominator is unchanged, so after finitely many steps the denominator is an odd integer d; Euler's congruence then gives $d \mid 2^{\varphi(d)} - 1$, and the multiplicative order of 2 modulo d is the least positive such exponent, with the convention that this order is 1 when $d = 1$. The hypothesis that d is odd cannot be dropped: if $T_N = 1/2$ then $2^h - 1$ is odd for every $h \geq 1$, so no shift at N is integral. The resulting periodicity is modulo $\mathbb{Z}$: the fractional parts eventually repeat, not necessarily the full tail values or the integer coefficients.

Theorem 5.3 (exact rationality classification)**.** *Let $T : \mathbb{N} \to \mathbb{R}$ satisfy $T_{N+1} = 2T_N - g_{N+1}$ with integer coefficients g. The following are equivalent:*

(i) *T_0 is rational;*

(ii) *$\sigma_h(N)$ is an integer for some $h \geq 1$ and some N;*

(iii) *for some fixed $h \geq 1$, $\sigma_h(N)$ is an integer at every sufficiently large N.*

Consequently T_0 is irrational if and only if every positive-length shift is nonintegral at every index, equivalently if and only if for every $h \geq 1$ and every cutoff some later N has $\sigma_h(N) \notin \mathbb{Z}$.

Proof. For (i)⇒(iii), the block identity makes every T_N rational. By (6), its denominator is a fixed odd integer d for all sufficiently large N. Choose $h = \varphi(d)$; Euler's congruence and the same formula show that every subsequent $\sigma_h(N)$ is integral. The implication (iii)⇒(ii) is immediate. For (ii)⇒(i), the block identity gives $\sigma_h(N) = (2^h - 1)T_N - B_{h,N}$ with $B_{h,N}$ and $\sigma_h(N)$ integers and $2^h - 1 \neq 0$, so T_N is rational, and $T_N = 2^N T_0 - B_{N,0}$ then makes T_0 rational. Negating the three conditions gives the two irrationality formulations. □

Lemma 5.4 (the boundary condition identifying a true tail)**.** *Let $a_1, a_2, \ldots$ be real numbers with $\sum_{j \geq 1} |a_j| 2^{-j} < \infty$, and let $U_{N+1} = 2U_N - a_{N+1}$. Then*

$$U_N = \sum_{j \geq 1} a_{N+j} 2^{-j} \quad \textit{for every } N \quad \iff \quad 2^{-N} U_N \longrightarrow 0.$$

Proof. Iteration gives $2^{-N} U_N = U_0 - \sum_{j=1}^{N} a_j 2^{-j}$. The limit is zero exactly when U_0 equals the full series; subtracting its first N terms then gives the claimed tail formula. Without the boundary condition, one may add $C 2^N$ to every U_N without changing the recurrence. □

For instance, zero coefficients and initial value $U_0 = 1/12$ give $\frac{1}{12}, \frac{1}{6}, \frac{1}{3}, \frac{2}{3}, \frac{4}{3}, \ldots$, with $U_{N+h} - U_N = (2^h - 1)2^N/12$. Shifts of length 2 become integral at index 2; shifts of length

1 never do. These are not complete tails of the zero sequence, since $2^{-N}U_N = 1/12$ does not tend to zero. The denominator calculation remains valid, but the boundary condition excludes this extra homogeneous term.

The actual prime-gap tails. Put

$$T_N = \sum_{j\geq 1} \frac{g_{N+j}}{2^j},$$

which converges by Theorem 4.3. Reindexing identifies this sum with the rescaled complete tail of S, as in the shifted-gap series identity. Splitting off the first term gives $T_{N+1} = 2T_N - g_{N+1}$ without a rationality hypothesis, as formalised in the real tail recurrence. Since $T_0 = 2S - 1 = 2.349287932\ldots$, the numbers Π, S and T_0 have the same rationality status, and Theorem 5.3 specialises to the actual series: irrationality of Π is equivalent to nonintegrality arbitrarily far out for each positive shift of T. The formal statement is the irrationality equivalence for the actual prime tails. Thus the passage from the series to its complete tails is proved. What remains is to establish the required nonintegrality for these prime-gap tails.

5.1 Pairs of congruent indices and least-common-multiple shifts

The previous criterion fixes the distance between the two indices. An equivalent version allows that distance to vary, but requires the indices to be congruent modulo each prescribed positive integer. It is not a weaker hypothesis on the tail: the next theorem proves that it is equivalent to irrationality.

Theorem 5.5 (pairs of congruent indices)**.** *S is irrational if and only if for every $t \geq 1$ and every N_0 there are $N, M \geq N_0$ with $M \equiv N \pmod{t}$ and $T_M - T_N \notin \mathbb{Z}$.*

Proof. If S is irrational, take $M = N + t$ for the N supplied by Theorem 5.3 at shift length t. Conversely, suppose S is rational. By (6) the orbit reaches an index N_0 beyond which the reduced denominator is a fixed odd d; let t be the multiplicative order of 2 modulo d, taking $t = 1$ when $d = 1$. For $N, M \geq N_0$ with $M \equiv N$ modulo t, the difference $T_M - T_N$ is then an integer, contradicting the stated property at this t and this cutoff. □

Modulo one, the tail recurrence is multiplication by two. Once the reduced denominator is the fixed odd integer d, a difference $T_M - T_N$ is integral precisely when $M - N$ is divisible by the order of 2 modulo d. Theorem 5.5 therefore allows the distance between the indices to vary, but requires a nonintegral pair beyond every cutoff for every prescribed modulus.

Equal tail values do not help this criterion, since their difference is zero. Section D.4 shows that rationality and the prime number theorem would in fact force many equal tail values. It also explains why, beyond the rational denominator cutoff, an interval condition excluding integers already contradicts rationality at indices congruent modulo the corresponding multiplicative order. These are ordinary deductions, not additional formalised statements.

A second reformulation tests one prescribed sequence of pairs of indices. Let $L_0 = 1$ and $L_j = \operatorname{lcm}(1, \ldots, j)$.

Theorem 5.6 (criterion using least common multiples)**.** *Let $g : \mathbb{N} \to \mathbb{Z}$ and $T : \mathbb{N} \to \mathbb{R}$ satisfy $T_{N+1} = 2T_N - g_{N+1}$. Then T_0 is irrational if and only if $T_{2L_j} - T_{L_j} \notin \mathbb{Z}$ for every $j \geq 0$.*

Proof. Theorem 5.3 gives the forward implication. Conversely, if T_0 is rational then some positive shift length h is integral at every basepoint beyond an index N_0. Choose j so large that $N_0 \leq L_j$ and $h \mid L_j$. The telescoping identity $\sigma_{a+b}(N) = \sigma_a(N) + \sigma_b(N + a)$ shows by induction that every positive multiple of h is integral at every such basepoint, so $T_{2L_j} - T_{L_j}$ is an integer. □

One could use factorials instead of least common multiples, since they have the same eventual divisibility property. The least common multiples are no larger at each index. Theorems 5.5 and 5.6 change which tail differences one must test, but remain exact equivalences for integer-coefficient recurrences. Neither proves the needed nonintegrality for the actual prime gaps.

6 A local certificate, and one actual pair

An integer in $(-1, 1)$ must be zero. Thus, if two consecutive tail differences lie in that interval and are both integral, both vanish. The recurrence then forces $g_{N+h+1} = g_{N+1}$. An unequal pair of gaps therefore gives a finite way to rule out simultaneous integrality, provided that both complete tail differences have been bounded.

Theorem 6.1 (adjacent small-shift obstruction)**.** *Let T satisfy the dyadic tail recurrence with integer coefficients g, and fix h and N. If*

$$-1 < \sigma_h(N) < 1, \qquad -1 < \sigma_h(N+1) < 1, \qquad g_{N+h+1} \neq g_{N+1},$$

then $\sigma_h(N)$ and $\sigma_h(N+1)$ cannot both be integers. Consequently, if such a pair occurs beyond every threshold, the h-shift is not eventually integral.

Proof. An integer strictly between -1 and 1 is zero. If both shifts were integers, both would vanish, and (5) would give $g_{N+h+1} = g_{N+1}$. The final assertion chooses one such adjacent pair after the alleged onset of integrality. □

Corollary 6.2 (real form and the sufficient condition)**.** *The same statement holds for a real orbit, with the same proof. If for every $h \geq 1$ and every cutoff some later N satisfies the three displayed conditions for the actual prime gaps, then Π is irrational.*

The unequal-gap condition alone is available: at $h = 1$, $N = 2$ it reads $g_4 = 2 \neq 4 = g_3$, and Proposition 6.3 gives such inequalities at arbitrarily large indices for each fixed h. The missing assertion is that both small-tail inequalities and the unequal-gap condition hold at the same arbitrarily late indices. Separate infinite sets of witnesses for the three requirements need not intersect.

Proposition 6.3 (prime gaps do not become periodic)**.** *For every positive h, the actual consecutive-prime-gap sequence is not eventually periodic with period h.*

Proof. For $m \geq 2$ the integers $(m+1)! + 2, \ldots, (m+1)! + m + 1$ are composite, so the gaps are unbounded. An eventually periodic sequence of natural numbers has finite range after its preperiod and a bounded initial segment, hence is bounded. □

Lean checks the factorial argument as unboundedness of the actual gaps and the conclusion as non-eventual periodicity. Far stronger lower bounds for large gaps are known [8, Theorem 1]; unboundedness is all that is needed.

With even gap differences, the three conditions in Theorem 6.1 reduce to two possible signed intervals, each accompanied by a prescribed gap difference. For fixed h write $D_N = \sigma_h(N)$ and $\delta_N = g_{N+h+1} - g_{N+1}$, so that $D_{N+1} = 2D_N - \delta_N$.

Theorem 6.4 (an equivalent signed interval test)**.** *Assume δ_N is even. The conjunction $-1 < D_N < 1$, $-1 < D_{N+1} < 1$, $\delta_N \neq 0$ is equivalent to*

$$(\delta_N = 2 \text{ and } \tfrac{1}{2} < D_N < 1) \quad \text{or} \quad (\delta_N = -2 \text{ and } -1 < D_N < -\tfrac{1}{2}).$$

Proof. Since $\delta_N = 2D_N - D_{N+1}$, the bounds $|D_N| < 1$ and $|D_{N+1}| < 1$ give $-3 < \delta_N < 3$. A nonzero even integer in this interval is 2 or -2. If $\delta_N = 2$, the inequality $-1 < 2D_N - 2 < 1$, together with $|D_N| < 1$, is equivalent to $1/2 < D_N < 1$. Negating both differences gives the case $\delta_N = -2$. Each substitution is reversible. □

For the actual gaps, both indices $N + 1$ and $N + h + 1$ are positive, so δ_N is even for every $N \geq 0$. Theorem 6.4 therefore applies to the actual tails. Its proof uses order, the recurrence and the evenness of δ_N; it does not require rational tail values. This is the real version used in Theorem 8.7.

6.1 An explicit remainder and a certified actual pair

Both window conditions involve complete infinite tails. Bounding the omitted terms turns each test into a finite integer comparison.

Proposition 6.5 (explicit remainder)**.** *For integers $h, N \geq 0$ and $L \geq 1$ put*

$$F_{h,N,L} = \sum_{j=1}^{L} \frac{g_{N+h+j} - g_{N+j}}{2^j}, \qquad P(x) = x^4 + 8x^3 + 36x^2 + 104x + 150,$$

$$E_{h,N,L} = \frac{1250}{2^L}\bigl(P(N + h + L + 2) + P(N + L + 2)\bigr).$$

Then $|\sigma_h(N) - F_{h,N,L}| \leq E_{h,N,L}$. In particular $|F_{h,N,L}| + E_{h,N,L} < 1$ certifies $|\sigma_h(N)| < 1$, and $\operatorname{dist}(F_{h,N,L}, \mathbb{Z}) > E_{h,N,L}$ certifies $\sigma_h(N) \notin \mathbb{Z}$.

Proof. The checked bound $p_n \leq 1250(n + 1)^4$ gives $0 \leq g_m \leq p_{m+1} \leq 1250(m + 2)^4$, so the omitted absolute tail is at most

$$1250 \sum_{j>L} \frac{(N + h + j + 2)^4 + (N + j + 2)^4}{2^j}.$$

The polynomial identity $2P(x) = (x + 1)^4 + P(x + 1)$ telescopes to $\sum_{k=1}^{J} (m + k)^4 2^{-k} = P(m) - P(m + J)2^{-J}$, whose last term tends to zero. Apply it with $m = N + h + L + 2$ and $m = N + L + 2$ after writing $j = L + k$. The two tests follow from the triangle inequality and the definition of distance to $\mathbb{Z}$. □

Proposition 6.6 (a certified adjacent pair)**.** *For the actual prime gaps, $h = 1$ and $N = 2$ satisfy the three hypotheses of Theorem 6.1: both $\sigma_1(2)$ and $\sigma_1(3)$ lie in $(-1, 1)$ and are nonintegral, and $g_4 = 2 \neq 4 = g_3$.*

Proof. Take $L = 40$ and $Q = 2^{40} = 1099511627776$. Exact integer arithmetic over the first 46 primes gives

N	$QF_{1,N,40}$	$QE_{1,N,40}$
2	−662838684750	11764181250
3	873345886050	12805761250

In each row $|QF| + QE < Q$, which puts the whole certified interval inside $(-1, 1)$, and QE is smaller than the distance from QF to the nearest multiple of Q, which excludes every integer. The margins are 424908761776 and 213359980476 respectively. These are strict integer comparisons rather than inferences from rounded numerical tails. Appendix B replays them. □

This computation proves the local condition at one pair of indices. The quantifiers over every h and every cutoff remain.

Proposition 6.7 (a one-tail signed certificate)**.** *Let $D_{N+1} = 2D_N - \delta_N$ with real D_N, and suppose $\delta_N = 2s$ for $s \in \{-1, 1\}$. If integers A, B, Q satisfy $Q > 0$, $B \geq 0$, $|QD_N - A| \leq B$, and*

$$2sA - Q > 2B, \qquad Q - sA > B,$$

then $1/2 < sD_N < 1$, $|D_{N+1}| < 1$, and both D_N, D_{N+1} are nonintegral. In particular the adjacent small-shift obstruction is certified from just one tail enclosure.

Proof. The inequalities place the complete interval $[(sA - B)/Q, (sA + B)/Q]$ inside $(1/2, 1)$. Thus $sD_{N+1} = 2sD_N - 2 \in (-1, 0)$, proving every assertion. □

For $h = 1, N = 2$, the exact data already used above give $Q = 2^{40}$, $A = -662838684750$, $B = 11764181250$ and $s = -1$. The two positive integer margins are respectively 202637379224 and 424908761776. The second enclosure in Proposition 6.6 remains a useful independent cross-check, not a logical necessity.

7 Two lower bounds on a possible rational denominator

An exact rational enclosure of Π can exclude all rational values whose denominator lies below an explicit bound. The two results here use the same polynomial estimate for the omitted prime terms, but have different supplied verification records. Both bounds are finite; neither proves irrationality.

Theorem 7.1 (a denominator bound checked by exact integer comparisons)**.** *Let $a \in \mathbb{Z}$ and let b be a positive integer. If $\Pi = a/b$ then $b \geq 2^{589} > 10^{177}$, and the same floor holds for every rational equal to S.*

Proof. Put $c = 1229$, the number of primes below 10^4, and $A = \sum_{i=0}^{c-1} p_i 2^{c-i-1}$, so that $A/2^c$ is the cth partial sum of Π. All terms are positive, so $A/2^c \le \Pi$. For the upper bound, $p_{c+j} \le 1250(c+j+1)^4$. Since $c \ge 9$, each successive ratio of $(c+1+j)^4$ is at most $(11/10)^4 < 3/2$. Hence $(c+1+j)^4 \le (c+1)^4(3/2)^j$ for all $j \ge 0$. The omitted dyadic terms are therefore bounded by a geometric series with ratio $3/4$, whose sum gives

$$\frac{A}{2^c} \le \Pi \le \frac{2A + 5000(c+1)^4}{2^{c+1}}.$$

The four positive integers u, v, u', v' printed in Appendix B satisfy

$$u'v - uv' = 1, \qquad \frac{u}{v} < \frac{A}{2^c}, \qquad \frac{2A + 5000(c+1)^4}{2^{c+1}} < \frac{u'}{v'}, \qquad v + v' \ge 2^{589},$$

all four being integer comparisons after clearing the positive denominators. If $\Pi = a/b$ then $av - bu \ge 1$ and $bu' - av' \ge 1$, and the determinant identity gives

$$b = b(u'v - uv') = (av - bu)v' + (bu' - av')v \ge v + v' \ge 2^{589}.$$

The only properties needed of the bracketing fractions are their order, positive denominators and determinant one. Their decimal expansions and the way they were found play no part in this denominator bound. Since $589 \log_{10} 2 > 177$, the floor exceeds 10^{177}. If $S = a/b$, apply the same argument to $\Pi = (a + 2b)/b$. □

The Lean kernel decides the four inequalities on the same literals by `decide +kernel` with no `native_decide`, over a trial-division sieve it re-runs on $[2, 10^4)$: the certificate, the floor for the prime series, and the floor for the gap series. Its analytic input is the same enclosure, over the polynomial bound at line 360. The four Farey literals have 177 and 178 digits.

Theorem 7.2 (certified continued-fraction exclusion)**.** *Every rational equal to* Π*, and hence every rational equal to* S*, has reduced denominator* $q \ge 2^{39997}$*, and therefore* $q > 10^{12040}$*.*

Proof. The accompanying integer-arithmetic verifier constructs an 80000-bit enclosure with 23369 common partial quotients and checks the following certificate. It uses neither floating-point arithmetic nor a Lean build.

Let $c = 80128$ and $A = \sum_{i=0}^{c-1} p_i 2^{c-i-1}$, with the primes obtained by a sieve below 1200000. The polynomial bound and the quartic identity of Proposition 6.5 give

$$\frac{A}{2^c} \le \Pi \le \frac{A + 1250P(c)}{2^c}.$$

Round the lower endpoint down and the upper endpoint up to multiples of 2^{-80000}. The resulting integers L, U satisfy $U - L = 1$ and $\Pi \in [L2^{-80000}, U2^{-80000}]$. This is the width reported as 1 in the scaled receipt. The verifier computes the common continued-fraction prefix of these two rational endpoints by integer division and inversion.

Its last two convergents, ordered as $u/v < u'/v'$, satisfy

$$u'v - uv' = 1, \qquad u\,2^{80000} < Lv, \qquad Uv' < u'\,2^{80000}, \qquad v + v' \ge 2^{39998}.$$

Thus the entire enclosure lies strictly between two Farey neighbours. The same determinant calculation as in Theorem 7.1 gives $q \geq v + v'$ for every rational value of Π. It proves the stated bound and, in fact, the slightly stronger dyadic bound $q \geq 2^{39998}$; the strict decimal bound remains 10^{12040}. The continued fractions used in the computation are those of the rational enclosure endpoints, not an assumed infinite continued fraction for Π. Once the neighbours are found, only the four displayed integer comparisons are needed. □

The original receipt uses the best-approximation property of convergents [10, §6, Theorems 16 and 17]; Short gives an accessible statement and proof [19, Theorem 1.1]. The receipt is the continued-fraction receipt, which records the generating program and its digest, the power-of-two exponent 39997, the bit length 39998, the strict decimal power 12040, the decimal digit count 12041, the largest partial quotient 973919 at index 16442, and an arithmetic self-check of the same routine against the known expansions of e and π. Those bit and digit counts describe the original routine's last convergent, not an arbitrary rational in the enclosure. Bit length 39998 implies a denominator at least 2^{39997}, and 12041 decimal digits imply only a non-strict lower bound 10^{12040}. The bound for q instead uses the checked sum $v + v' \geq 2^{39998}$. The strict decimal inequality follows from the separate integer comparison $2^{39997} > 10^{12040}$.

A rational number may have a denominator beyond either bound. Extending the computation can exclude a larger finite range, but a longer prefix need not improve the bound at every step and cannot by itself exclude all denominators. These statements are denominator exclusions, not estimates for an irrationality exponent.

8 What cannot supply the missing input

The counterexamples in this section test which information about the gaps could force irrationality. They preserve selected growth, residue and nonconcentration properties while changing the dyadic value. The sparsity result has a different role: it rules out a positive-proportion lower bound for the two-window event, while leaving sparse-witness arguments available.

8.1 Bounded residue-preserving perturbations

Here a single positive modulus M is fixed in advance. Adding either zero or M to each sufficiently late coefficient preserves that modulus and gives a binary choice at every index. Unlike the sparse theorem, this construction need not preserve all moduli eventually and need not have a support of density zero.

Theorem 8.1 (bounded-perturbation obstruction)**.** *Let a_n be natural numbers with $\sum_{n\geq 0} a_n 2^{-(n+1)}$ convergent. For every integer $M \geq 1$ and every cutoff K there are digits $\varepsilon_n \in \{0,1\}$, zero for $n < K$, such that $\sum_{n\geq 0}(a_n + M\varepsilon_n)2^{-(n+1)}$ is rational.*

Proof. Write $A = \sum_{n\geq 0} a_n 2^{-(n+1)}$ and choose a rational $r \in (A, A + M2^{-K})$. A binary expansion of $(r - A)/M$ has the form $\sum_{n\geq K} \varepsilon_n 2^{-(n+1)}$ with $\varepsilon_n \in \{0,1\}$; set $\varepsilon_n = 0$ for $n < K$. The perturbed series converges and equals r. □

The construction is covered by the supplied formal verification record. It follows that a property shared by every allowed perturbation cannot by itself force irrationality, since one such perturbation has a rational sum. This is an ordinary consequence of the construction, not a separate formal theorem. The construction is also checked at the actual consecutive prime gaps for every $M \geq 1$ and every K, the rational perturbed prime-gap series, with the perturbed gap lying in $[g_n, g_n + M]$ and congruent to g_n modulo M, the gap bounds. For each fixed $k \geq 1$, the sum of k consecutive gaps increases by at most kM. Thus infinitely many bounded gaps and bounded clusters of each fixed size survive, but their original numerical bounds need not. Nonnegative corrections also preserve lower bounds for large gaps. Taking $M = 2q$ preserves parity and the cumulative positions modulo a prescribed q, hence any limiting residue frequencies of those positions. These conclusions concern the size and residue information in the results discussed in Section 10; they establish neither primality of the new positions nor preservation of every quantitative assertion of those theorems.

8.2 Algebraic nonconcentration survives the rationalising perturbation

The bounded perturbation also preserves fixed-block polynomial nonconcentration. The relevant input is Schlage-Puchta's Lemma 4: for every polynomial $F \in \mathbb{Z}[x_0, \dots, x_k]$ which does not vanish identically, $F(g_n, \dots, g_{n+k}) \neq 0$ for almost all n [18, Lemma 4, pp. 5–6]. Say that an integer sequence a has *fixed-block nonconcentration* when it satisfies that conclusion for every $k \geq 0$ and every nonzero $F \in \mathbb{Z}[x_0, \dots, x_k]$. The property transfers across bounded perturbations with no independence or randomness assumption.

This nonconcentration condition excludes every bounded integer sequence: if its values lie in a finite set E, the nonzero polynomial $\prod_{c \in E}(x_0 - c)$ vanishes at every index. It also excludes a sequence satisfying a fixed nonzero polynomial relation on every short block, such as $a_n = n$, for which $x_1 - x_0 - 1$ vanishes. The density-zero exceptional set depends on the fixed polynomial. There cannot be one such set for all polynomials: $x_0 - a_n$ vanishes on the block beginning at n, so a common exceptional set would contain every index. No uniform estimate for growing families of polynomials or block lengths is asserted.

In the next proposition, N, H are nonnegative integers, h is an integer, and all counts are over integer indices and integer triples.

Proposition 8.2 (finite counting for shifted gap differences)**.** *Write p_n for the primes indexed from $p_0 = 2$ and $g_n = p_{n+1} - p_n$. For $h \geq 2$ and $r \in \mathbb{Z}$, let $M_{h,r}(N)$ count $n < N$ with $g_{n+h} - g_n = r$. Let $Q_{N,H,r}$ count triples (x, d, s) with $x < p_N$, $0 < d < s \leq H$, $d + r > 0$, and all four integers $x, x + d, x + s, x + s + d + r$ prime. Then, for every $N, H \geq 0$,*

$$(H + 1)M_{h,r}(N) \leq (h + 1)p_{N+h+1} + (H + 1)Q_{N,H,r}.$$

Proof. For a counted index n, put $x = p_n$, $d = g_n$ and $s = p_{n+h} - p_n$. If the total span $p_{n+h+1} - p_n$ is at most H, then $0 < d < s \leq H$, $d + r = g_{n+h} > 0$, and the four primes are precisely $x, x + d, x + s, x + s + d + r$. The map is injective because $x = p_n$ determines n. For the remaining indices the integer span is at least $H + 1$. Each gap appears in at most $h + 1$ of the spans, so their total is at most $(h + 1)\sum_{i<N+h} g_i < (h + 1)p_{N+h+1}$.

Multiply the resulting bound on the exceptional count by $H+1$ and add the small-span contribution. □

The displayed finite shifted-count bound is unconditional by the ordinary proof above. Its historical source locator is retained; Appendix C identifies the relocated proof in the supplied packet and distinguishes its recorded release status from a checked build. To obtain zero density from this particular finite bound, one would need, for every $\varepsilon > 0$ and all large N, a choice of H making its right side less than $\varepsilon N(H+1)$. For $h = 1$, the same definitions give $s = d$, so $x + d = x + s$. The strict condition $d < s$ in the four-prime count fails; that case requires a three-prime count instead. Zero density itself is already known for every fixed $h \geq 1$ and $r \in \mathbb{Z}$: apply Schlage-Puchta's lemma above to the nonzero polynomial $x_h - x_0 - r$. Thus the finite reduction should not be read as leaving that upper bound open. Neither upper bound supplies the late occurrences or weighted-tail separation required for irrationality.

For the perturbation claim, first suppose that a has fixed-block nonconcentration and every correction is either 0 or a fixed integer $M \geq 1$. For a fixed $r \in \mathbb{Z}$, an equality $b_{n+1} - b_n = r$ then requires $a_{n+1} - a_n \in \{r - M, r, r + M\}$. Each of these three equalities holds on a set of density zero. The general argument uses the same finite-union step for translated polynomials.

Theorem 8.3 (nonconcentration is perturbation-stable)**.** *Let $a : \mathbb{N} \to \mathbb{Z}$ have fixed-block nonconcentration, let $E \subset \mathbb{Z}$ be finite, and let $b_n = a_n + e_n$ with $e_n \in E$ for every n. Then b has fixed-block nonconcentration.*

Proof. Fix $k \geq 0$ and a nonzero $F \in \mathbb{Z}[x_0, \dots, x_k]$. For a tuple $\mathbf{e} = (e^{(0)}, \dots, e^{(k)}) \in E^{k+1}$ put $F_{\mathbf{e}}(x_0, \dots, x_k) = F(x_0 + e^{(0)}, \dots, x_k + e^{(k)})$. The substitution $x_i \mapsto x_i + e^{(i)}$ is a ring automorphism of $\mathbb{Z}[x_0, \dots, x_k]$ with inverse $x_i \mapsto x_i - e^{(i)}$, so $F_{\mathbf{e}} \neq 0$. If $F(b_n, \dots, b_{n+k}) = 0$ then $F_{\mathbf{e}}(a_n, \dots, a_{n+k}) = 0$ for the tuple $\mathbf{e} = (e_n, \dots, e_{n+k})$, whence

$$\{\, n : F(b_n, \dots, b_{n+k}) = 0 \,\} \subseteq \bigcup_{\mathbf{e} \in E^{k+1}} \{\, n : F_{\mathbf{e}}(a_n, \dots, a_{n+k}) = 0 \,\}.$$

Each set on the right has density zero by hypothesis, and there are only finitely many such sets because E^{k+1} is finite. Their union therefore has density zero. □

Finiteness of the correction set is essential to this argument; no independence between the corrections and the sequence is needed. Without a restriction on the corrections, the choice $e_n = -a_n$ would give $b_n = 0$, which fails nonconcentration even for the polynomial x_0.

Corollary 8.4 (nonconcentration does not force irrationality)**.** *Fix $M \geq 1$ and $K \geq 0$, and let b be the perturbed sequence supplied by Theorem 8.1 at the actual prime gaps. Then $\sum_{n \geq 0} b_n 2^{-(n+1)}$ is rational, $b_n = g_n$ for $n < K$, $b_n - g_n \in \{0, M\}$ and $b_n \equiv g_n \pmod{M}$ for every n, b has fixed-block nonconcentration, and the cumulative sequence $P_n = 2 + \sum_{i<n} b_i$ satisfies $p_n \leq P_n \leq p_n + Mn$ and hence $P_n \sim n \log n$.*

Proof. Theorem 8.1 gives the rational sum, prefix and congruence assertions. Summing $0 \leq b_i - g_i \leq M$ and using $\sum_{i<n} g_i = p_n - 2$ gives $p_n \leq P_n \leq p_n + Mn$. Now $p_n \sim$

$n \log n$ [17, Chapter 6] implies $P_n \sim n \log n$. The perturbation takes values in the fixed two-element set $E = \{0, M\}$, so Theorem 8.3 applies to the actual gaps, which have fixed-block nonconcentration by Schlage-Puchta's lemma [18, Lemma 4]. □

Fixed-block polynomial nonconcentration, taken together with a prescribed finite prefix of actual gaps, a pointwise bound $b_n \le g_n + M$, every residue modulo M, positivity and cumulative growth at the prime-number-theorem scale, is therefore compatible with a rational dyadic value. Taking M even also preserves eventual evenness. To preserve a prescribed modulus q together with parity, take $M = 2q$, with the corresponding pointwise bound $b_n \le g_n + 2q$. The terms P_n are not asserted to be prime. The nonconcentration conclusion is for each fixed polynomial in a fixed block; it gives no uniform bound for polynomial families or block lengths growing with the index.

Proposition 8.5 (nonconcentration under sparse changes)**.** *Let $a, b : \mathbb{N} \to \mathbb{Z}$ agree off a set S of ordinary density zero. If a has fixed-block nonconcentration, then so does b. No boundedness assumption on $a - b$ is needed.*

Proof. For fixed k and a nonzero $F \in \mathbb{Z}[x_0, \dots, x_k]$, a zero of $F(b_n, \dots, b_{n+k})$ either is already a zero at the a-block or has $n + i \in S$ for some $0 \le i \le k$. Each fixed translate of a density-zero set has density zero, and a finite union retains that property. This proves the assertion for each fixed F and k. It gives no uniform estimate for polynomial families or lengths growing with n. □

Corollary 8.6 (simultaneous prime-gap countermodel)**.** *For every prescribed finite prime-gap prefix and $0 < \varepsilon \le 1$, there is an altered sequence $b = g + e$, with $P_n = 2 + \sum_{i<n} b_i$, for which one can simultaneously impose a rational dyadic value, nonnegative integer corrections eventually at most $(\log(n+3))^{\varepsilon}$, all fixed eventual coefficient and cumulative congruences, fixed-block polynomial nonconcentration, and vanishing total variation distance between unnormalised block distributions for lengths $o(\log \log X)$. The cumulative positions satisfy*

$$0 \le P_n - p_n = O_\varepsilon\left(\frac{n(\log(n+3))^{\varepsilon}}{\log \log n}\right), \qquad P_n \sim n \log n.$$

Proof. Apply Theorem 2.1 at a rational target. Proposition 8.5 transfers the nonconcentration supplied by Schlage-Puchta's Lemma 4 [18] to the corrected sequence. To obtain the cumulative support bound, split $[\sqrt{n}, n)$ into dyadic intervals, use $\log \log x \asymp \log \log n$ there, and bound the first $\sqrt{n}$ indices trivially. Thus $|S \cap [0, n)| = O(n/\log \log n)$. The pointwise correction bound gives the displayed estimate, including a fixed constant for the exceptional prefix. Since $(\log n)^{\varepsilon - 1}/\log \log n \to 0$ for $0 < \varepsilon \le 1$, the prime number theorem [17, Chapter 6] gives the last conclusion. The position series is rational as well: reversing the order of summation of nonnegative terms gives

$$\sum_{n \ge 0} \frac{P_n}{2^{n+1}} = 2 + \sum_{i \ge 0} b_i \sum_{n > i} 2^{-n-1} = 2 + \sum_{i \ge 0} \frac{b_i}{2^{i+1}} \in \mathbb{Q}.$$

The finite value on the right also proves convergence on the left. □

These are ordinary deductions from the sparse theorem and the stated external analytic input. In fact, all conclusions also hold for every $\varepsilon > 1$: apply the corollary with exponent 1. Its correction is smaller than the prescribed allowance, and its cumulative estimate $O(n \log(n+3)/\log\log n)$ is stronger than the displayed bound and still gives $P_n \sim n \log n$. Thus the stated range $0 < \varepsilon \leq 1$ suffices to obtain every positive allowance exponent; no larger correction is required.

The congruences imply that for every fixed integer $y \geq 2$, all sufficiently late P_n have no prime divisor at most y: use modulus $y!$ and $p_n > y$ to get $\gcd(P_n, y!) = 1$. This is stronger than eventual oddness, but it does not establish primality. Excluding divisors up to $\sqrt{P_n}$ would require a bound growing with n, whereas the congruence cutoff depends on the fixed y. Even primality of every P_n would not identify the sequence with the consecutive primes, since a prime-valued sequence may skip primes.

The two preservation arguments are different. The bounded construction uses a finite set of translated polynomials to preserve nonconcentration; it need not preserve empirical block frequencies. The sparse construction instead compares blocks at the same indices, most of which are unchanged. Its total variation bound is absolute, so it need not preserve relative frequencies of rare events. Moreover, its vanishing-error range is $m = o(\log\log X)$, not m comparable to $\log\log X$. Neither comparison establishes preservation of the quantitative prime-pattern hypotheses of the conditional irrationality and normality results discussed in Section 3.

8.3 The two-window event has density zero

The same nonconcentration lemma shows that the three conditions of the small-pair test can hold only on a set of density zero. This does not prevent that set from containing arbitrarily large indices.

Theorem 8.7 (sparsity of the two-window event)**.** *Fix $h \geq 1$. The set of $N \geq 1$ at which the three hypotheses of Theorem 6.1 hold for the actual prime gaps has density zero. For the same h, the set of N with $g_{N+h+1} = g_{N+1}$ also has density zero.*

Proof. For $N \geq 1$ every gap involved is even, so $\delta_N = g_{N+h+1} - g_{N+1}$ is even. If $|\sigma_h(N)| < 1$, $|\sigma_h(N+1)| < 1$ and $\delta_N \neq 0$, then $\delta_N = 2\sigma_h(N) - \sigma_h(N+1)$ lies in $(-3, 3)$, so $\delta_N = \pm 2$. Apply Schlage-Puchta's lemma [18, Lemma 4] at index $n = N+1$ with $k = h$ to the two polynomials $x_h - x_0 - 2$ and $x_h - x_0 + 2$, neither of which vanishes identically: each of $\{N : \delta_N = 2\}$ and $\{N : \delta_N = -2\}$ has density zero, and the event is contained in their union. The second assertion is the same lemma applied to $x_h - x_0$. □

Unequal gaps occur at almost every index, a stronger conclusion than Proposition 6.3. But combining unequal gaps with the two small tail differences forces their difference to be exactly 2 or -2, which occurs only on a density-zero set. Thus a proof of Problem 10.2 must find arbitrarily large successful indices in that sparse set. A positive-density lower bound for the same event is impossible. An averaging argument that detects a sparse set is not ruled out.

The proof of Schlage-Puchta's lemma gives a quantitative upper bound as well. For fixed h, it bounds the number of eligible indices in $[X, 2X)$ by $O_h(X/\log\log X)$. To see the two contributions, discard starts whose block of $h+1$ gaps contains a gap exceeding

$\log X \log\log X$. The sum of these block sums is $O_h(X \log X)$ by the prime number theorem, so there are at most $O_h(X/\log\log X)$ discarded starts. For the remaining starts, the sieve calculation in the cited proof bounds the zeros of each of $x_h - x_0 - 2$ and $x_h - x_0 + 2$ by

$$O_h\left(\frac{X(\log\log X)^{2h+2}}{\log X}\right) = o_h\left(\frac{X}{\log\log X}\right).$$

The shift from $n = N + 1$ changes only the band endpoints. Thus the same upper bound applies to the two-window event. This is an ordinary consequence of the cited sieve argument, not a new lower bound or an asymptotic formula for the number of successful indices.

The observed proportions in Section 9 decrease across the sampled bands. These finite observations establish neither an asymptotic rate nor the required existence beyond every cutoff. Counting by prime size and counting by gap index also give different cutoffs and must not be compared without converting between them.

8.4 Recurring gap values differing by two do not suffice

Theorem 6.4 requires both a gap difference of 2 or -2 at a prescribed distance h and a tail difference in the corresponding interval. Infinite recurrence of two values differing by 2 in one residue class guarantees neither their occurrence at that distance nor the weighted-tail bound. The next example has the stated recurrence and growth properties, yet its dyadic sum is rational and every tail difference is integral. It is a sequence of integers, not of actual prime gaps.

The mechanism is to choose integer tails first, then recover the coefficients from $a_n = 2U_{n-1} - U_n$. Constant tails $U_n = 4$ give constant coefficients $a_n = 4$. Raising one isolated tail to 6 changes the two affected coefficients from $4, 4$ to $2, 8$, with the same combined dyadic contribution. A slowly growing even background will supply cumulative growth $n \log n$; factorial indices will impose the recurrence in residue classes while keeping these changes sparse.

Theorem 8.8 (recurring values are not enough)**.** *There is a sequence $(a_n)_{n\ge1}$ of positive even integers with the following properties. The values 2 and 4 each occur infinitely often at indices divisible by every fixed $t \ge 1$. The sequence is unbounded, not eventually periodic, and satisfies $a_n = O(\log n)$. The series $\sum_{n\ge1} a_n 2^{-n}$ equals 6, and every scaled tail $\sum_{j\ge1} a_{N+j}2^{-j}$ is an integer, so every tail shift is integral. The increasing odd sequence $P_n = 3 + \sum_{j\le n} a_j$ satisfies $P_n \sim n \log n$.*

Proof. Put $b_n = 2\lceil \frac{1}{2}\log(n+64)\rceil$, so that b_n is even, $b_n \ge 6$, and $b_n = \log(n+64) + O(1)$. Since $\frac{1}{2}\log(n+65) - \frac{1}{2}\log(n+63) < 1$ for every $n \ge 1$, consecutive increments of b lie in $\{0, 2\}$ and $b_{n+1} - b_{n-1} \le 2$. Call an index n *special* when $n = k!$ or $n = 2\,k!$ for some $k \ge 5$, and define

$$U_n = \begin{cases} 2b_{n-1} - 2, & n = k!,\ k \ge 5, \\ 2b_{n-1} - 4, & n = 2\,k!,\ k \ge 5, \\ b_n, & \text{otherwise}, \end{cases} \qquad a_n = 2U_{n-1} - U_n \quad (n \ge 1).$$

For $k \geq 5$ one has $k! < 2\,k! < (k+1)!$, so all special indices are distinct. They are even and therefore never adjacent. Each U_n is an even integer by construction, so each a_n is an even integer.

At an ordinary index whose predecessor is also ordinary, $a_n = 2b_{n-1} - b_n = b_{n-1} - (b_n - b_{n-1}) \geq 6 - 2 = 4$. At a special index the predecessor is ordinary and $a_n = 2b_{n-1} - (2b_{n-1} - r) = r$, which is 2 at $n = k!$ and 4 at $n = 2\,k!$. Immediately after a special index carrying r,

$$a_{n+1} = 2(2b_{n-1} - r) - b_{n+1} \geq 4b_{n-1} - 2r - (b_{n-1} + 2) = 3b_{n-1} - 2r - 2 \geq 8.$$

Every a_n is therefore a positive even integer, and $U_n = O(\log n)$ gives $a_n = O(\log n)$. For $k \geq 5$, the index $n = k! + 2$ and its predecessor are ordinary: both lie strictly between $k!$ and $2\,k!$, where there are no special indices. Along these indices $a_n \geq b_{n-1} - 2 \to \infty$. Thus (a_n) is unbounded and hence not eventually periodic.

The definition $a_n = 2U_{n-1} - U_n$ is exactly $a_n 2^{-n} = U_{n-1}2^{-(n-1)} - U_n 2^{-n}$, so for every $N \geq 0$ and $m \geq 1$

$$\sum_{j=1}^{m} \frac{a_{N+j}}{2^j} = U_N - \frac{U_{N+m}}{2^m}.$$

Since $U_n = O(\log n)$ the endpoint tends to zero, so the tail at N equals the integer U_N; at $N = 0$ the series equals $U_0 = b_0 = 6$. Every difference $U_{N+h} - U_N$ is an integer, so every tail shift is integral.

For each t and every sufficiently large k we have $t \mid k!$, hence $t \mid 2\,k!$, and $a_{k!} = 2$ while $a_{2k!} = 4$: both values recur infinitely often at indices congruent to 0 modulo t.

Finally, summing $a_j = 2U_{j-1} - U_j$ gives $\sum_{j=1}^{n} a_j = 2U_0 + \sum_{j=1}^{n-1} U_j - U_n$. The baseline $\sum_{j<n} b_j = n \log n + O(n)$. The special indices up to n number $O(\log n / \log\log n)$ and each changes the summand by $O(\log n)$, so their total contribution is $O((\log n)^2) = o(n)$. Hence $P_n = 3 + \sum_{j \leq n} a_j \sim n \log n$, and P_n is odd and increasing. □

The factorial schedule above proves recurrence in the zero residue class. The following separate construction strengthens this to every residue class.

8.5 A rational series with recurring values in every residue class

There is a synthetic integer sequence $(a_n)_{n \geq 1}$ with all of the following properties. Each a_n is positive and divisible by 2, and

$$a_n \leq 4 \log(n+1) + 24.$$

For every $t \geq 1$, every $0 \leq r < t$, and every cutoff N, there are $i, j \geq N$ such that

$$i \equiv j \equiv r \pmod{t}, \qquad a_i = 2, \quad a_j = 4.$$

For every integer B and cutoff N, some $n \geq N$ satisfies $a_n > B$. For every $h \geq 1$ there is no cutoff after which $a_{n+h} = a_n$ for all n. Nevertheless, all the complete tails

$$U_N = \sum_{j \geq 1} \frac{a_{N+j}}{2^j} \quad (N \geq 0)$$

are integers, $U_0 = 6$, and $U_{N+h} - U_N \in \mathbb{Z}$ for every $N, h \geq 0$. The sequence

$$P_N = 3 + \sum_{j=1}^{N} a_j$$

is strictly increasing and odd, with $P_N/(N \log N) \to 1$. These are synthetic positions; no P_N is asserted to be prime.

Enumerate all triples (t, r, m) with $t \geq 1$, $0 \leq r < t$ and $m \geq 0$, each once. For the kth triple choose $c_k \equiv r \pmod{t}$, with $c_0 \geq 100$, $c_{k+1} \geq c_k + 3$ and $c_k \geq 2^{k^2}$ for $k \geq 1$. Each arithmetic progression is unbounded, so these lower bounds can be met recursively. Set $v_k = 2$ for even m and $v_k = 4$ for odd m. For each (t, r) both choices then occur infinitely often. Use the even baseline

$$b_n = 6 + 2\left\lfloor \frac{\log(n+1)}{2} \right\rfloor, \qquad U_n = \begin{cases} 2b_{n-1} - v_k, & n = c_k, \\ b_n, & n \notin \{c_k : k \geq 0\}, \end{cases}$$

and define $a_n = 2U_{n-1} - U_n$. The separation of the centres will ensure that $a_{c_k} = v_k$.

The baseline is even and at least 6. For $n \geq 1$, the increase of $\log(n+1)$ over two consecutive steps is less than 2, so the floor in the definition gives $b_n - b_{n-1} \in \{0, 2\}$ and $b_{n+1} - b_{n-1} \leq 2$. Away from a centre and its successor, $a_n = 2b_{n-1} - b_n \geq 4$. At a centre, $a_{c_k} = v_k$. Immediately afterwards,

$$a_{c_k+1} = 4b_{c_k-1} - 2v_k - b_{c_k+1} \geq 3b_{c_k-1} - 2v_k - 2 \geq 8.$$

Since the centres are at least three apart, these cases exhaust all indices. Also $0 < U_n \leq 2b_n$, so $a_n \leq 4b_{n-1} \leq 4\log(n+1) + 24$. The indices $n = c_k + 2$ lie outside the centres and their successors, because consecutive centres are at least three apart. Along these indices $a_n \geq b_{n-1} - 2 \to \infty$. Thus the sequence is unbounded and cannot be eventually periodic.

Finite telescoping gives

$$\sum_{j=1}^{m} a_{N+j} 2^{-j} = U_N - U_{N+m} 2^{-m}.$$

The bound $U_n = O(\log(n+1))$ makes the last term tend to zero. Hence the complete tail equals the integer U_N, and its initial value is 6. For the cumulative growth, $b_n = \log(n+1) + O(1)$ and there are at most $1 + \sqrt{\log_2 N}$ centres up to N. Replacing b_n by U_n changes its partial sum by $O((\log N)^{3/2})$. Therefore

$$\sum_{j=1}^{N} a_j = 2U_0 + \sum_{j=1}^{N-1} U_j - U_N = N \log N + O(N),$$

which gives the stated asymptotic for P_N.

The full statement is kernel-checked in Lean.

Thus positivity, evenness, unboundedness, nonperiodicity, logarithmic size, cumulative growth $N \log N$, and recurrence of the values 2 and 4 in every residue class are jointly compatible with integral tails. These properties alone cannot prove that a tail difference is nonintegral. The sequence is not asserted to consist of prime gaps: its logarithmic bound in particular excludes the extreme large gaps of the primes. The construction therefore does not address hypotheses that use those extreme gaps.

8.6 Polynomial coefficients and telescoping sums

Proposition 8.9 (quadratic polynomial-shift countermodel)**.** *Put* $c_n = 2(n^2+4n+2)$ *and* $U_n = 2(n+4)^2$*. Then* c_n *is positive, even and strictly increasing,* $U_{n+1} = 2U_n - c_{n+1}$*, every shift* $U_{N+h} - U_N$ *is integral,* $c_{n+1} - c_n = 4n+10$ *is never* ± 2*, and*

$$\sum_{j\geq 1} \frac{c_j}{2^j} = 32.$$

Proof. Direct expansion gives the recurrence, and the finite telescope is $\sum_{j=1}^{n} c_j 2^{-j} = 32 - 2(n+4)^2 2^{-n}$, whose last term tends to zero. The remaining assertions follow from the integer values of U_n and from $c_{n+1} - c_n = 4n+10 \geq 10$. □

The series starts at $j = 1$ and equals the initial rescaled tail $U_0 = 32$. Under the opening convention its sum is $\sum_{n\geq 0} c_n 2^{-(n+1)} = (c_0+32)/2 = 18$. The formal construction checks the recurrence, strict growth, integrality of every shift, exclusion of adjacent differences of size two, and the value 32 for the shifted series, whose summand is $c_{n+1}/2^{n+1}$; the rational value is recorded at line 109. Positivity, parity, strict growth, unboundedness and nonperiodicity are therefore jointly compatible with rationality, and the adjacent-gap condition of Theorem 6.4 fails at every index. The telescoping mechanism is the one used in the note of ChatGPT 5.4 Pro (orchestrated by Vjeko Kovač) against the variable-denominator expectation of Erdős [24, Theorem 1 and proof, pp. 1–2]; the sequence is not a result about Problem #251.

Rationality also fails to force the coefficients to repeat. Let $K : \mathbb{N} \to \mathbb{Q}$ be arbitrary and put $\kappa_n = 2K_n - K_{n+1}$: the coefficient of the telescoping series.

Proposition 8.10 (exact telescoping)**.** *For every* $n \geq 0$*,* $\sum_{i=0}^{n-1} \kappa_i 2^{-(i+1)} = K_0 - K_n 2^{-n}$.

Proof. The sum is empty when $n = 0$. Passing from n to $n+1$ adds $(2K_n - K_{n+1})2^{-(n+1)} = K_n 2^{-n} - K_{n+1}2^{-(n+1)}$, which replaces the terminal term by the required one. □

The identity $\kappa_n = 2K_n - K_{n+1}$ has the algebraic form used in the rationality criterion of Erdős and Straus when $a_n = 2$. That criterion has additional hypotheses: for integers b_n and positive integers a_n with $a_n > 1$ for all large n and $|b_n|/(a_{n-1}a_n) \to 0$, the series $\sum_n b_n/(a_1 \cdots a_n)$ is rational exactly when some positive integer B and integers c_n satisfy $Bb_n = c_n a_n - c_{n+1}$ and $|c_{n+1}| < a_n/2$ for all large n [7, Theorem 2.1, pp. 85–86]. With $a_n = 2$, the smallness hypothesis becomes $|b_n|/4 \to 0$. Since the b_n are integers, this forces $b_n = 0$ eventually. Thus the criterion in this specialisation does not apply even to a bounded integer sequence that is nonzero infinitely often, let alone to the positive prime gaps. Proposition 8.10, by contrast, imposes no growth condition on K. The infinite series converges precisely when $2^{-n}K_n$ has a finite limit, and then its sum is $K_0 - \lim_{n\to\infty} 2^{-n}K_n$. The sum equals K_0 precisely when that limit is zero; polynomial growth suffices. For $K_n = 2^n$, every κ_n is zero, so the series converges to 0, not to $K_0 = 1$.

The finite telescoping identity is also formalised. Taking $K_0 = \frac{5}{2}$ and $K_n = 2n+2$ for $n \geq 1$ gives $\kappa_0 = 1$ and $\kappa_n = 2n$, so the coefficients $1, 2, 4, 6, 8, \ldots$ are positive, even after the first term, unbounded and not eventually periodic, while their dyadic sum is $\frac{5}{2}$.

Rationality alone therefore cannot imply eventual periodicity even for a positive, parity-correct integer coefficient sequence, and rationality cannot be contradicted merely by Proposition 6.3.

The arithmetic-progression formulations of nonintegrality do not create new prime-distribution information. Their full definitions and exact hypotheses are preserved in Appendix E.

9 Finite numerical searches

The two searches described below concern finite ranges of the prime-gap tails. Their receipts record floating-point observations, not exact certificates or proofs of occurrence beyond every cutoff. The exact integer checks are given separately in Appendix B. The searches have not been rerun for this edition.

Over the 6 841 648 primes below 1.2×10^8 and offsets $h = 1, \dots, 16$, the scan records hits for the reduced two-window event of Theorem 6.4 at every tested offset, with observed proportions between 0.00418 and 0.008248 and nearly balanced signs. At $h = 1$ it records 56 427 hits among 6 841 564 candidate indices, the last at the prime 119 995 753. Multiplying the observed proportion in each band by that band's mean of $\log p$ gives 0.17671 in the first band and 0.13148 in the last. This finite observation does not establish an asymptotic counting law or cofinality. The tails in this scan use double-precision arithmetic. The receipt gives no certified enclosure for the accumulated rounding error, so neither a count nor an individual hit is an exact certificate. A bound for the omitted infinite tail would not, by itself, bound that rounding error. Proposition 6.6 gives a separate pair certified by exact integer arithmetic. The receipt is the adjacent-pair search receipt.

A second scan uses the 1 270 607 primes below 2×10^7 and pairs of indices in the half-open range $1\,143\,545 \le n, m < 1\,270\,540$. It searches each residue class modulo each $1 \le t \le 20$ for the condition of Theorem 5.5, selecting pairs by a next-gap difference of 2 and a shifted-tail window. At $t = 20$, every residue class has at least 90 150 counted pairs; pairs may share indices. For each class, the program also records the larger index of a selected late witness. The minimum of those recorded indices is 1 270 520; it need not come from the class with the fewest pairs. The saved sample pairs pass the resulting two-window check in floating-point arithmetic. These observations are not exact certificates, and the scan supplies neither arbitrary moduli nor witnesses beyond every prescribed cutoff. The receipt is the congruent-pair search receipt. The public computation guide provides the programs, dependencies and exact replay commands for all three recorded runs, including the continued-fraction calculation above.

10 Remaining prime-gap estimates

Theorem 5.3 reduces irrationality of the actual prime series to nonintegrality of its tail differences. The following two problems distinguish that equivalent condition from the stronger, local condition used by the finite certificates.

Problem 10.1 (nonintegral tail differences of each positive length)**.** For every $h \geq 1$ and every N_0, prove that some $N \geq N_0$ satisfies

$$\sum_{j\geq 1} \frac{g_{N+h+j} - g_{N+j}}{2^j} \notin \mathbb{Z}. \tag{7}$$

Problem 10.2 (two small tail differences at arbitrarily large indices)**.** For every $h \geq 1$ and every N_0, prove that some $N \geq N_0$ satisfies

$$|\sigma_h(N)| < 1, \qquad |\sigma_h(N+1)| < 1, \qquad g_{N+h+1} \neq g_{N+1}. \tag{8}$$

Problem 10.1 and the variable-offset criterion of Theorem 5.5 are each equivalent to irrationality. By Corollary 6.2, Problem 10.2 is sufficient; no converse is asserted. These logical reformulations do not supply the required prime-gap estimate.

The counterexamples retain a fixed prefix and modulus (Theorem 8.1), polynomial nonconcentration and cumulative growth $n \log n$ (Corollary 8.4), or all the congruences and block statistics of the sparse construction. The distinct construction in Section 8.5 makes two values differing by 2 recur in every residue class. Each has a rational sum; none guarantees prime cumulative positions.

For the actual gaps, Theorem 8.7 shows that the event in (8) has density zero for each fixed h. A proof of its occurrence must therefore find arbitrarily late witnesses inside a sparse set, rather than prove that they occupy a positive proportion. A finite prefix alone cannot do this: it can be retained while the dyadic sum is made rational. A finite block together with a proved bound on the omitted tail can still certify one pair, as in Proposition 6.6. The distinction is between a finite certified instance and an occurrence theorem beyond every cutoff.

Proposition 6.5 gives sufficient integer inequalities for window membership at a chosen truncation length L. Failure of such a test need not mean that the complete tail lies outside the window: its error interval may still cross an endpoint. For fixed h and $\varepsilon > 0$, the choice $L = \lceil (4+\varepsilon)\log_2(N+2) \rceil$ gives $E_{h,N,L} = O_{h,\varepsilon}(N^{-\varepsilon})$. Indeed, $L = O_\varepsilon(\log N)$, both arguments of the quartic P are $O_{h,\varepsilon}(N)$, and $2^{-L} \leq (N+2)^{-4-\varepsilon}$. Requiring a fixed positive margin from both endpoints is stronger than strict window membership in (8). At this depth, the test bounds the signed integer $2^L F_{h,N,L}$, not just its residue modulo 2^L, and also requires $g_{N+h+1} - g_{N+1} = \pm 2$. For fixed h and ε, checking one certificate uses $O_\varepsilon(\log N)$ gap values and the explicit remainder bound. This counts data, not bit operations. The unproved input is the existence of certified witnesses beyond every cutoff for each fixed h; constants uniform in h are not required.

The following published results control gap sizes and clusters, not the signed weighted sums required by this criterion. Zhang's bounded-gap theorem [21, Theorem 1, p. 1122] produces infinitely many bounded consecutive-prime gaps. Maynard bounds $\liminf_n (p_{n+m} - p_n)$ for every fixed m, giving bounded clusters of every fixed size [16, Theorem 1.1, p. 384], with the explicit unconditional bound $\liminf_n (p_{n+1} - p_n) \leq 600$ [16, Theorem 1.3, p. 385]. Polymath subsequently improved this to 246 [35, Theorem 1.4(i) of arXiv v4]; neither numerical bound controls the signed weighted tails required here. The large-gap theorem of Ford, Green, Konyagin, Maynard and Tao gives

an effective lower bound for the largest single consecutive-prime gap below X [8, Theorem 1]. A theorem giving bounded clusters at each fixed cluster size does not by itself supply a weighted condition on windows whose length grows with the basepoint, and none of these results is used as a proof input here.

Conditionally the picture is different. Tao's comment names uniform quantitative prime-tuples control of about $\log\log n$ consecutive gaps as the plausible route [23, comment of 7 October 2025], and Land's draft carries that out under Kuperberg's uniform prime-tuples conjecture [13]. The countermodels in Section 8 show that the listed size, congruence and block-statistical properties alone do not force irrationality. These constructions are not shown to enumerate the consecutive primes, so they do not rule out arguments that combine those properties with further information about primes.

Verification boundary. The smaller denominator bound has a recorded kernel-decided integer certificate. The larger bound is supported by the archived computation and the independent integer-arithmetic replay described above. Neither is an irrationality theorem, and the replay is not a Lean kernel check. The older 4096-bit calculation and the separate justification of its tail bound are discussed in Section D.7. Declaration-level build status is recorded below, separately from the ordinary proofs and numerical receipts.

Statements and declarations

Artefact and data availability. Declaration links identify immutable revisions, not a live branch. The main source links in this manuscript retain revision `3d6d938d696f`; release-only declarations are linked to `52f29ad173b0`. The supplied declaration index distinguishes `ci_checked`, `outside_checked_build`, and `release_only`; presence of a file or link alone is not evidence that its proof was checked. The recorded CI receipt is run 35073961520. The declaration index records the status of each result. The build record reports `lake build ErdosProblems Erdos249257` with Lean 4.29.1. No new execution of that build is claimed here.

Lean 4 [14] and mathlib [15] provide the proof-checking environment. A checked proof statement, its hypotheses and its transitive axiom report must be distinguished from a challenge containing `sorry`, an executable certificate, or an ordinary proof. Formal checking does not establish novelty, attribution, or the faithfulness of an unexamined mathematical interpretation.

The finite-computation receipts and generating programs remain separately identified in the pinned computation guide. For this edition the self-contained verifier in Appendix B was rerun; it checks the adjacent pair and the bound 2^{589}. A separate verifier, supplied as `checks/verify_independent_cf.py`, regenerates the 80000-bit enclosure and checks the Farey certificate for the larger exclusion. It reproduces the common-prefix length and the receipt's denominator threshold, and checks the stronger neighbouring-denominator sum stated above. It does not rerun the original program's floating-point statistics or its self-tests on e and π. The prime-gap frequency scans were not rerun, and neither verifier is a new Lean build.

Funding and competing interests. This work received no external funding. The author declares no competing interests.

Acknowledgements. I thank Wouter van Doorn for advice on explaining unfamiliar hypotheses, removing unnecessary terminology, and using notation only when it helps the reader. His comments concerned an earlier note on Problem #243; this acknowledgement does not imply that he reviewed or endorsed the mathematics of the present paper. The problem numbering follows the Erdős Problems catalogue maintained by Thomas Bloom [22]. The logarithmic-scale construction and the transfer of nonconcentration are retained from the earlier review materials accompanying this project; that provenance is not evidence of independent mathematical review.

A An elementary polynomial bound for the primes

We prove $p_n \le 1250(n+1)^4$ for every $n \ge 0$ by the central binomial coefficient method of Erdős's proof of Chebyshev's theorem [5, §1, pp. 194–196]. Let $\pi(x)$ count the primes at most x. For an integer $m \ge 4$,

$$4^m < m\binom{2m}{m} \le m(2m)^{\pi(2m)}. \tag{9}$$

For the first inequality, $m\binom{2m}{m}/4^m$ equals $70/64$ at $m = 4$ and its ratio at successive indices is $(2m+1)/(2m) > 1$. For the second, the exponent of a prime ℓ in $\binom{2m}{m}$ is $\sum_{k\ge1}(\lfloor 2m/\ell^k\rfloor - 2\lfloor m/\ell^k\rfloor)$, each summand is 0 or 1, and every summand with $\ell^k > 2m$ vanishes; the exponent is therefore at most $\lfloor \log_\ell(2m)\rfloor$. Thus the full power of each prime appearing in $\binom{2m}{m}$ is at most $2m$. Multiplying over at most $\pi(2m)$ distinct primes gives the second inequality.

Now fix $n \ge 0$, put $x = n + 5$ and $m = x^4 \ge 625$, and suppose $\pi(2m) \le n$. Since $x \le 2^x$ and $n + 4(n+1)x = 4x^2 - 15x - 5 \le 2x^4$,

$$m(2m)^n = 2^n x^{4(n+1)} \le 2^{\,n+4(n+1)x} \le 2^{\,2x^4} = 4^m,$$

contradicting (9). Hence $\pi(2m) > n$, so at least $n+1$ primes lie below $2m$ and therefore

$$p_n \le 2(n+5)^4 \le 1250(n+1)^4,$$

the last step because $(n+5) \le 5(n+1)$ for $n \ge 0$. The coarser bound $p_n \le 1250(n+1)^4$ is checked by the kernel at line 360.

B Integer certificates

The following calculation uses trial division and integer arithmetic only. It reproduces the two rows of Proposition 6.6 and every inequality used in Theorem 7.1. The four literals are the certificate the Lean kernel decides in `KernelDenominatorFloor.lean`; their use here requires no floating-point approximation. Adjacent quoted strings are concatenated by Python.

```
from math import isqrt

primes = [n for n in range(2, 10000)
          if all(n % d for d in range(2, isqrt(n) + 1))]
gaps = [q - p for p, q in zip(primes, primes[1:])]
P = lambda x: x**4 + 8*x**3 + 36*x*x + 104*x + 150
Q = 2**40
expected = [(-662838684750, 11764181250),
            (873345886050, 12805761250)]
for N, target in zip((2, 3), expected):
    D = sum((gaps[N+1+j] - gaps[N+j])*2**(40-j)
            for j in range(1, 41))
    B = 1250*(P(N+43) + P(N+42))
    distance = min(D % Q, Q - D % Q)
    if (D, B) != target or not (abs(D)+B < Q and B < distance):
        raise ArithmeticError("tail certificate failed")
if gaps[4] == gaps[3]:
    raise ArithmeticError("gap mismatch failed")

u = int(
    "80656418571526529321760196321868980032711628291714663348 27"
    "30836077944152787174450335094078559889033699885255507461 59"
    "73558897922500842023448210201391609566636587897181681526 62"
    "0217"
)

v = int(
    "21949451244136632321439709245412633124220695246356153605 18"
    "42470773519582218168307201892899048316629550843926902486 83"
    "12162917239885377332351730406072544968385302138677814423 35"
    "1745"
)

up = int(
    "65394371014981626268824118924709052221082600085685554459 75"
    "30261633155217846686899097127598765624846200590981384174 69"
    "52328398881853163742723332773896114831173340004938675849 23"
    "912"
)

vp = int(
    "17796110757898665511984272416217096301201363232815981766 37"
    "84906052492297355019687649040361529707294916566508739385 80"
    "62030254663133898102912437860054991645374993836120818051 60"
    "767"
)

c = len(primes)
A = sum(p*2**(c-1-i) for i, p in enumerate(primes))
checks = (c == 1229,
          u > 0 and v > 0 and up > 0 and vp > 0,
          up*v - u*vp == 1,
          u*2**c < A*v,
          (2*A + 5000*(c+1)**4)*vp < up*2**(c+1),
          v+vp >= 2**589,
```

```
          2**589 > 10**177)
if not all(checks):
    raise ArithmeticError("denominator certificate failed")
print("Both tail rows and the denominator floor verified.")
```

C Guide to the formal sources

The links in this appendix and after selected proofs identify formal definitions and statements at fixed revisions and line numbers. The declaration index records their build status; the ordinary arguments are printed in the paper itself.

The summation-by-parts identity and polynomial prime bound concern the actual primes. The recurrence results apply to arbitrary integer coefficients and, as stated, rational or real sequences. Substituting the actual prime-gap tail requires its convergence and boundary condition. The finite certificates use the same polynomial bound to control the omitted terms. The counterexamples concern other coefficient sequences, not a construction of new primes. The complete link inventory is also retained in the TeX source.

Sparse construction. The source index records the interval theorem for an arbitrary sequence, its rational-target consequence, the polylogarithmic interval theorem, the block total-variation bound and the uniform bounded-test theorem as `ci_checked`. The formal construction chooses one coordinate at a time, rather than the triples used here. Its polylogarithmic statement fixes the permitted set, interval, and support-count constant and cutoff before the target; the eventual congruences and convergence are quantified after it. The common support-count bound and same-index block comparison therefore give a target-independent statistical error bound. A common congruence cutoff instead requires the construction, which chooses it from the schedule and modulus. The printed schedule independently supplies the common cutoffs and bounds in Theorem 2.1. The recorded build status is described under Statements and declarations; no new Lean build is claimed.

Finite shifted-gap count. The supplied release contains a relocated proof of the finite shifted-gap count in Proposition 8.2. Its natural-number triples agree with the printed integer triples: x, d, s and $x + s + d + r$ are positive under the stated conditions. The finite inequality has no sieve premise; the separate sieve-dependent density-zero implication is not used here. The declaration index records this file as `release_only`, not `ci_checked`; source availability does not establish a successful build.

Finite summation by parts. Formalised as the summation-by-parts identity. The finite identity requires no positivity, monotonicity or convergence.

Prime-gap summation by parts. Formalised as the prime-gap summation by parts, using the gap partial sum and the agreement of the two gap formulas; the latter records that the natural-number difference $p_{n+1} - p_n$ agrees with the difference taken in $\mathbb{Q}$, which needs $p_n \le p_{n+1}$, the increase of consecutive primes.

Convergence and the infinite identity. The sources prove the polynomial prime bound, convergence of the prime series and the convergence transfer to the gap series. The identity in Theorem 4.3 is the unconditional prime-gap identity.

The irrationality equivalences. The source proves the irrationality equivalence between Π and S. For the second normalisation, it proves the identity $\sum_{n\geq 0} p_n 2^{-n} = 4+2S$ and the corresponding irrationality equivalence.

The general recurrence. The source defines the tail recurrence, the shift, and integrality.

Iteration of the recurrence. The source versions of Theorem 5.2 give the iterated block identity and the scaled shift identity.

Two consecutive small differences. The sources for Theorem 6.1 and Corollary 6.2 are the rational adjacent small-shift obstruction, its arbitrarily late consequence, the real small-pair implication, and the sufficient condition for irrationality.

Congruent indices. For Theorem 5.5, the source gives the criterion for congruent indices, the eventual congruence classification, and the irrationality criterion using congruent indices, using the actual real tail recurrence.

Denominators and integral shifts. The sources for (6) and its consequences give the denominator recurrence, its odd case and even case, the denominator classification, the totient shift, and the propagation of an integral shift to every later index.

Rationality of the initial value. For Theorem 5.3, the sources give two rationality criteria: one integral shift of positive length and eventual integrality for a fixed positive length. The corresponding irrationality criteria are nonintegrality at every index for every positive length and arbitrarily late nonintegrality for every positive length.

D Further criteria, examples and computational details

This appendix proves additional consequences of the recurrence, explains the finite truncation test, and gives a bounded counterexample and the full numerical table. These results are not needed for the sparse construction.

D.1 Totient-length shifts and persistence of integrality

Euler's congruence gives an explicit integral shift once the denominator is odd. The next proposition records that choice of length; the following one shows that integrality then persists at later indices.

Proposition D.1 (a shift of totient length)**.** *Let $T : \mathbb{N} \to \mathbb{Q}$ satisfy the dyadic tail recurrence with integer coefficients. If the reduced denominator d of T_N is odd, then $\sigma_{\varphi(d)}(N)$ is an integer.*

Proof. Since d is odd, 2 and d are coprime, so Euler's congruence gives $d \mid 2^{\varphi(d)} - 1$. Writing $2^{\varphi(d)} - 1 = dk$ and $T_N = u/d$ in lowest terms, $(2^{\varphi(d)} - 1)T_N = ku$ is an integer, and the integral-shift criterion transfers this to the shift. □

Proposition D.2 (propagation)**.** *Let $T : \mathbb{N} \to \mathbb{R}$ satisfy $T_{n+1} = 2T_n - a_{n+1}$ with integer coefficients, and define $\sigma_h(N) = T_{N+h} - T_N$. For fixed $h, N \geq 0$, if $\sigma_h(N)$ is an integer, then $\sigma_h(N + k)$ is an integer for every $k \geq 0$.*

Proof. The shift step identity gives $\sigma_h(N + 1) = 2\sigma_h(N) - (a_{N+h+1} - a_{N+1})$, an integer combination of an integer and two coefficients. Induct on k. □

These are the totient shift and the propagation theorems cited in the note. For example, if $\operatorname{den} T_N = 3$ then $\varphi(3) = 2$ and $3T_N$ is an integer, so $\sigma_2(N)$ is integral while $\sigma_1(N)$ is not; if $\operatorname{den} T_N = 5$ then $\sigma_4(N)$ is integral. In the orbit with every coefficient zero and $T_0 = 1/12$, the denominator is $12 = 2^2 \cdot 3$, the orbit reaches $T_2 = 1/3$ with odd denominator at $s = 2$, and $h = \varphi(3) = 2$ is exactly the shift length seen to be integral from index 2 onwards. The totient supplies one length, while the exact criterion is $\operatorname{den}(T_N) \mid 2^h - 1$. At an index with odd denominator $d > 1$, the least positive length is the multiplicative order of 2 modulo d. When $d = 1$, every positive length works; before the denominator becomes odd, no positive length works.

D.2 The finite truncation criterion

Here g_n denotes the actual prime gaps. We bound the omitted terms by a nonnegative majorant whose dyadic series converges. For an arbitrary real-valued majorant, convergence alone does not provide a computable remainder bound. The polynomial bound in Proposition 6.5 supplies an explicit rational bound, so the resulting certificate can be checked by integer arithmetic.

Proposition D.3 (finite truncation)**.** *Let $M : \mathbb{N} \to \mathbb{R}$ satisfy $M(n) \geq g_n$ for every n and $\sum_{n \geq 0} M(n)2^{-n} < \infty$, and put*

$$S_{h,N,L} = \sum_{j=1}^{L} \frac{g_{N+h+j} - g_{N+j}}{2^j}, \qquad R_{h,N,L}(M) = \sum_{j>L} \frac{M(N+h+j) + M(N+j)}{2^j}.$$

If for every $h \geq 1$ and every N_0 there are $N \geq N_0$ and $L \geq 1$ with $\operatorname{dist}(S_{h,N,L}, \mathbb{Z}) > R_{h,N,L}(M)$, then the nonintegrality condition in Problem 10.1 holds.

Proof. The part of $\sum_{j \geq 1}(g_{N+h+j} - g_{N+j})2^{-j}$ omitted from $S_{h,N,L}$ has absolute value at most $R_{h,N,L}(M)$, so under the displayed inequality the full sum lies at positive distance from every integer. □

Unlike the signed-window test, this criterion asks only for separation from the integers. The distance from the truncated sum to $\mathbb{Z}$ is determined by a finite integer calculation. Write the weighted block sum as

$$D_{h,N,L} = \sum_{j=1}^{L} 2^{L-j}(g_{N+h+j} - g_{N+j}), \qquad S_{h,N,L} = \frac{D_{h,N,L}}{2^L}.$$

Then

$$\operatorname{dist}(S_{h,N,L}, \mathbb{Z}) = 2^{-L} \min\{D_{h,N,L} \bmod 2^L,\ 2^L - (D_{h,N,L} \bmod 2^L)\},$$

where $D_{h,N,L} \bmod 2^L$ is the least nonnegative residue, including when $D_{h,N,L} < 0$. The criterion requires this residue to be more than $2^L R_{h,N,L}(M)$ from both 0 and 2^L. One successful comparison certifies nonintegrality at that index; a failed comparison is inconclusive. Irrationality requires such certificates arbitrarily late for every positive h. At prescribed logarithmic depth, the needed residue must be farther from both endpoints than the error bound. With the polynomial majorant in Proposition 6.5, the remainder bound is rational and explicit, so the whole test reduces to integer comparisons. The modular rewriting is an ordinary deduction, not a separately checked Lean declaration.

Proposition D.4 (completeness of finite separation)**.** *Suppose $D \in \mathbb{R}$, $S_L \in \mathbb{R}$ and $R_L \geq 0$ satisfy $|D - S_L| \leq R_L$ and $R_L \to 0$. Then*

$$D \notin \mathbb{Z} \quad \Longleftrightarrow \quad \text{there exists } L \text{ with } \operatorname{dist}(S_L, \mathbb{Z}) > R_L.$$

Proof. The reverse implication follows from the triangle inequality. For the forward implication put $\delta = \operatorname{dist}(D, \mathbb{Z}) > 0$. For all sufficiently large L, one has $R_L < \delta/2$, and the 1-Lipschitz property of distance gives $\operatorname{dist}(S_L, \mathbb{Z}) \geq \delta - R_L > R_L$. No monotonicity of the errors is required. □

For fixed h, N and the explicit polynomial remainder, testing $L = 1, 2, \ldots$ therefore stops with a certificate whenever $\sigma_h(N)$ is nonintegral. An unsuccessful comparison is inconclusive. A proof that every depth fails, unlike a finite unsuccessful run, would imply integrality by the same equivalence. None of this shows that suitable indices occur arbitrarily late. Quantitative estimates are still needed to guarantee separation at a prescribed truncation length as N varies. For a numerical example unrelated to primes, $[1 - 2^{-N}, 1 + 2^{-N}]$ contains both the noninteger $1 + 2^{-2N}$ and the integer 1 for every $N \geq 1$. Its radius tends to zero, yet it never certifies the changing value as nonintegral. This does not contradict completeness, which keeps the value fixed while refining its enclosure.

D.3 Nonintegral differences at multiples of each shift

Problem D.5 (multiples of each shift)**.** For every $r \geq 1$, is there an $m \geq 1$ such that the tail differences of length mr are nonintegral at arbitrarily large indices, as in (7)?

This is an exact reformulation, not a weaker mathematical target. If T_0 is irrational, every positive shift is nonintegral by the block identity, so the stated condition follows with $m = 1$. Conversely, rationality gives a shift $h \geq 1$ that is integral at every sufficiently late index. For each positive integer m, the identity

$$T_{N+mh} - T_N = \sum_{j=0}^{m-1} (T_{N+(j+1)h} - T_{N+jh})$$

makes every multiple mh eventually integral. Taking $r = h$ contradicts the stated condition. This ordinary deduction uses the checked rationality classification and a finite telescope; rearranging the quantifiers supplies no new estimate for the actual prime gaps.

D.4 Repeated tail values and congruent indices

In this subsection, $T_N = \sum_{j\ge1} g_{N+j}2^{-j}$ denotes the complete tail of the actual prime-gap series, rather than an arbitrary solution of the recurrence. The first argument uses the prime number theorem to find many equal tail values under rationality. The second distinguishes these repetitions from the nonintegral differences required at congruent indices. Both are ordinary deductions; the linked formal results concern the recurrence and the congruent-index criterion.

Repeated tail values. For an integer $X \ge 0$, summing $g_{N+1} = 2T_N - T_{N+1}$ over $0 \le N < X$ gives the exact identity

$$\sum_{N=0}^{X} T_N = p_{X+1} - p_1 - T_0 + 2T_X \le (p_{X+1} - p_0) + 2T_X.$$

The inequality uses $T_0 \ge 0$ and $p_1 \ge p_0$. The prime number theorem [17, Ch. 6] also gives

$$\frac{T_X}{p_X} = \sum_{j\ge1} \frac{g_{X+j}}{2^j p_X} \longrightarrow 0. \tag{10}$$

Indeed, for each fixed j, both p_{X+j+1}/p_X and p_{X+j}/p_X tend to 1, so their difference g_{X+j}/p_X tends to zero. For the uniform bound needed to pass the limit through the sum, the two-sided estimates $p_n \asymp n\log n$ give one constant K such that, for $X \ge 2$ and $j \ge 1$,

$$\begin{aligned}\frac{g_{X+j}}{p_X} &\le K\frac{(X+j+1)\log(X+j+1)}{X\log X}\\ &\le K(1+j)\left(1+\frac{\log(1+j)}{\log X}\right) \le K(1+j)^2.\end{aligned}$$

The second line uses $X + j + 1 \le X(1 + j)$ and $\log(1 + j) \le j\log 2 \le j\log X$. Since $\sum_{j\ge1}(1+j)^2 2^{-j} < \infty$, dominated convergence applies. Subtracting the exact summation identities at $2X$ and X now gives

$$\sum_{X<N\le2X} T_N = p_{2X+1} - p_{X+1} + 2(T_{2X} - T_X) \sim X\log X.$$

Indeed, the prime difference is asymptotic to $X\log X$, and (10) makes both tail terms $o(X\log X)$. Thus the mean complete tail in this band is asymptotic to $\log X$, without a rationality assumption. This is a mean, not a pointwise estimate. For the counting argument we only need the upper bound $\sum_{X<N\le2X} T_N \le K_0 X\log X$, valid for any fixed $K_0 > 1$ and all sufficiently large X. For each fixed $C > 1$, Markov's inequality then leaves at least $(1 - 1/C)X$ indices with $T_N \le K_0 C\log X$.

Under rationality, take X beyond the point where the reduced denominator is the fixed odd integer d. Among the indices just selected by Markov's inequality, the values lie in $d^{-1}\mathbb{Z}_{\ge0} \cap [0, K_0 C\log X]$. There are at most $dK_0 C\log X + 1$ such values. Pigeonhole therefore gives one value occurring at least

$$\frac{(1 - 1/C)X}{dK_0 C\log X + 1}$$

times. Taking, for example, $K_0 = 2$ gives a lower bound of order $\gg_{C,d} X/\log X$. The value may depend on X; no single value recurring in every band is established.

This count uses the growth of the actual primes. It is false for a general rational integer-coefficient recurrence: the quadratic countermodel of Proposition 8.9 has strictly increasing tails $T_N = 2(N+4)^2$. Even for the primes, equal-tail pairs have difference zero and hence do not supply the nonintegral difference required by the criterion using congruent indices.

The terminal term in the exact identity cannot simply be omitted. The factorial-gap argument in the proof of Proposition 6.3 shows that $T_X \geq g_{X+1}/2$ is unbounded, whereas $p_1 + T_0$ is fixed. The exact identity therefore gives $\sum_{N=0}^{X} T_N > p_{X+1}$ for arbitrarily large X. It does not establish that this strict inequality holds eventually.

The interval condition at congruent indices. Suppose the sum is rational, let d be the eventual odd reduced denominator, and let t be the order of 2 modulo d, with $t = 1$ when $d = 1$. Beyond the denominator cutoff, $M \equiv N \pmod{t}$ forces $T_M - T_N \in \mathbb{Z}$. Neither this difference nor its negative can then lie in $(\frac{1}{2}, 1)$. Thus a witness in that interval, at two such congruent indices, would contradict rationality without a further condition on the intervening gaps. This uses only the interval occurring in Theorem 6.4; it does not assert that its other hypotheses hold for arbitrary congruent pairs.

D.5 A bounded companion to the recurring-values countermodel

The logarithmic growth in Theorem 8.8 gives its cumulative sequence the same leading asymptotic $n \log n$ as the primes. The same mechanism is visible in the following bounded example.

Proposition D.6 (bounded recurring-values countermodel)**.** *Put $U_0 = 4$ and, for $n \geq 1$, $U_n = 6$ when $n = k!$ for some $k \geq 3$ and $U_n = 4$ otherwise, and set $a_n = 2U_{n-1} - U_n$ for $n \geq 1$. Then $a_n \in \{2, 4, 8\}$. For every $k \geq 3$, the value 2 occurs at index $k!$ and the value 4 at index $2\,k!$, so both recur infinitely often at indices divisible by any fixed $t \geq 1$. The series $\sum_{n\geq 1} a_n 2^{-n}$ equals 4 and every tail $\sum_{j\geq 1} a_{N+j}2^{-j}$ equals the integer U_N.*

Proof. For $k \geq 3$ the index $k!$ is even and at least 6, and $k!-1$ is odd, so the predecessor of a spike is an ordinary index; hence $a_{k!} = 8-6 = 2$, $a_{k!+1} = 12-4 = 8$, and $a_n = 8-4 = 4$ elsewhere. For $k \geq 2$ one has $k! < 2\,k! < (k+1)!$, so $2\,k!$ is not a factorial. Its odd predecessor $2\,k! - 1 \geq 3$ is not a factorial either; therefore $a_{2k!} = 4$. Since $t \mid k!$ for every large k, both values recur in the residue class 0 modulo t. The identity $a_n 2^{-n} = U_{n-1}2^{-(n-1)} - U_n 2^{-n}$ telescopes to $\sum_{j=1}^{m} a_{N+j}2^{-j} = U_N - U_{N+m}2^{-m}$, and U is bounded, so the endpoint vanishes. □

The indices with coefficient 2 are precisely the factorials $k!$, $k \geq 3$. They occur infinitely often with unbounded gaps, which is impossible for a value recurring in an eventually periodic sequence. Thus even this bounded coefficient sequence is nonperiodic, despite its rational dyadic sum. Its cumulative sum has linear growth. The logarithmic construction in Theorem 8.8 adds unboundedness and cumulative growth $N \log N$ without changing the telescoping mechanism. Neither construction guarantees prime cumulative positions.

D.6 Numerical counts and proportions

Section 9 describes the scan over the 6 841 648 primes under 1.2×10^8, with double-precision tails. The table gives its reported counts for $h = 1, \dots, 4$; the scan covered all offsets $h \le 16$. Each proportion divides the event count by $6\,841\,565 - h$ tested base-points, not by the number of primes.

h	events	proportion	$\delta = +2$	$\delta = -2$	last event at prime
1	56 427	0.008248	28 022	28 405	119 995 753
2	31 979	0.004674	15 742	16 237	119 993 807
3	30 233	0.004419	15 093	15 140	119 998 321
4	29 262	0.004277	14 606	14 656	119 993 473

The scan records hits at every tested offset $h \le 16$. The observed proportions range from 0.00418 to 0.008248, with nearly balanced signs at each offset. For $h = 1$, the first row below gives the proportions in the eight bands; the second multiplies each by its band's mean of $\log p$:

0.011285, 0.008951, 0.008303, 0.007936, 0.007651, 0.007507, 0.007253, 0.007094;

0.17671, 0.15058, 0.14420, 0.14067, 0.13766, 0.13667, 0.13333, 0.13148.

The ratio of the last rescaled proportion to the first is 0.744 at $h = 1$ and lies between 0.744 and 0.8121 across the measured offsets. Theorem 8.7 gives limiting density zero for each fixed offset. The bound $O_h(X/\log\log X)$ derived after its proof is only an upper bound; it predicts neither a monotone decline nor an asymptotic proportion. The finite-band ratios therefore do not establish an asymptotic rate.

D.7 Corrections to earlier records

The earlier tail-bound calculation. The 4096-bit program linked through Section 9 bounds T_s by summing the first 400 terms of $1250(s+j+2)^4 2^{-j}$ after rounding each down, then adding 1. This does not necessarily bound that polynomial series: at the recorded $s = 1\,270\,306$, the returned value is 40 below the exact sum $1250P(s+2)$, with P as in Proposition 6.5.

The returned bound nevertheless has a valid justification. The stronger estimate $p_n \le 2(n+5)^4$ proved in Appendix A gives

$$T_s \le 2P(s+6) < 625(s+3)^4 \qquad (s \ge 0).$$

The last expression is already the first summand of the program, and all its remaining summands are nonnegative. The strict inequality follows from

$$625(s+3)^4 - 2P(s+6) = 623s^4 + 7436s^3 + 32958s^2 + 62972s + 40437 > 0.$$

Thus the rounding issue is in the stated majorant argument, not a failure of the returned enclosure. This argument justifies the analytic bound; it is not a rerun of the denominator-exclusion search.

Two different improvements to the exclusion record. An earlier internal record gave a denominator exclusion at 10^{602}. The kernel-decided floor $2^{589} > 10^{177}$ of Theorem 7.1 is numerically smaller but has a different verification status. It does not supersede 10^{602} in numerical strength. The continued-fraction exclusion $2^{39997} > 10^{12040}$ of Theorem 7.2 is the larger numerical bound. The bound has 12041 decimal digits; the corresponding strict power-of-ten lower bound is 10^{12040}, not 10^{12041}. The earlier manuscript reports a separate interval and Farey-neighbour replay. Independently of that report, this edition regenerates an 80000-bit enclosure and its Farey certificate, as described in Theorem 7.2. The original program was not executed, and the new verification is not a Lean check.

An insufficient proposed condition. An earlier record named Hardy-Littlewood k-tuple correlation of consecutive gaps at a fixed offset as the missing input. With the offset freed by Theorem 5.5 the route needs, for each modulus t, cofinally many congruent pairs with certified nonintegral tail difference. A further record proposed that two even values differing by 2, each occurring infinitely often inside a common index residue class, would supply the two-condition form. Theorem 8.8 shows that this implication is false for integer-coefficient recurrences. That entry is withdrawn.

The formal boundary. An earlier reading of the short note placed the relation between rational and real tails and the real small-pair implication among the results supported only by ordinary proofs. The supplied record instead identifies formal proofs at line 48 and line 99 of the real prime-gap tail module. A second reading placed the kernel-decided denominator floor and the criterion using congruent indices outside the source revision the note pins. Both modules are present at that revision, and both are linked from the note.

E Arithmetic-progression reformulations of integrality

One can ask whether an iterated tail difference lies in an arithmetic progression determined by the intervening coefficients. The next theorem shows exactly what this asks of the original tail difference. Under the stated growth bound, a second test using the fixed progression $2^r\mathbb{Z}$ also reduces to the same nonintegrality condition.

Here $D_N = \sigma_h(N)$ and $\delta_N = g_{N+h+1} - g_{N+1}$, with h fixed. For a longer block, set

$$B_{h,N,r} = \sum_{i=0}^{r-1} 2^{r-1-i}\delta_{N+i}; \qquad D_{N+r} = 2^r D_N - B_{h,N,r}.$$

The progression $-B_{h,N,r}+2^{r+1}\mathbb{Z}$ depends on the very coefficients used to compute D_{N+r}. Substituting the displayed recurrence will cancel that dependence; no new information is obtained by increasing r.

Theorem E.1 (equivalent arithmetic-progression tests)**.** *For every rational dyadic tail recurrence and all $h, N, r \geq 0$,*

$$D_{N+r} \in -B_{h,N,r} + 2^{r+1}\mathbb{Z} \quad \Longleftrightarrow \quad D_N \in 2\mathbb{Z}. \tag{11}$$

Consequently, if every δ_N is even, then

$$\begin{aligned}(\forall N_0\ \exists N, r:\ N_0 < N \text{ and } D_{N+r} \notin -B_{h,N,r} + 2^{r+1}\mathbb{Z}) \\ \iff \quad D_N \notin \mathbb{Z} \text{ for arbitrarily large } N.\end{aligned} \tag{12}$$

There is a second equivalence. Let $b : \mathbb{N} \to \mathbb{Q}$ satisfy $|D_N| \le b(N)$ for every N, and suppose that for every N and every positive integer q there is an r with

$$2b(N + r)q < 2^r. \tag{13}$$

Then

$$\begin{aligned}\forall N_0\ \exists N, r:\ N_0 < N \text{ and } \forall z \in \mathbb{Z}, \quad b(N + r) < |B_{h,N,r} - 2^r z| \\ \iff \quad D_N \notin \mathbb{Z} \text{ for arbitrarily large } N.\end{aligned} \tag{14}$$

Proof. Substituting $D_{N+r} = 2^r D_N - B_{h,N,r}$ into the first membership condition cancels the observed block from both sides and leaves $D_N = 2z$. This proves (11). When every δ_N is even, the recurrence also gives $D_{N+1} \in 2\mathbb{Z}$ if and only if $D_N \in \mathbb{Z}$. Thus eventual integrality would force eventual even integrality and contradict the left side of (12). This proves the forward implication. Conversely, given a cutoff, choose a nonintegral D_N beyond it and take $r = 0$ in (11).

For (14), eventual integrality and the block identity give some $z \in \mathbb{Z}$ with

$$B_{h,N,r} - 2^r z = -D_{N+r},$$

so the displayed strict separation contradicts $|D_{N+r}| \le b(N + r)$. Conversely, if D_N is nonintegral, choose a positive integer q with $1/q \le \operatorname{dist}(D_N, \mathbb{Z})$. Such a q exists because $\mathbb{Z}$ is closed and $D_N \notin \mathbb{Z}$. Choose r from (13). The block identity and the triangle inequality give, for every $z \in \mathbb{Z}$,

$$|B_{h,N,r} - 2^r z| \ge 2^r |D_N - z| - |D_{N+r}| \ge \frac{2^r}{q} - b(N + r) > b(N + r).$$

The same r works for all integers z, as required. □

The three displayed equivalences are assembled in one declaration, the three arithmetic-progression equivalences.

Evenness is essential in (12). Without it, the recurrence $T_N = N + 1$, $g_n = n - 1$ for $n \ge 1$, has $D_N = \sigma_1(N) = 1$ and $\delta_N = 1$ at every index. Thus D_N is never an even integer but is always an integer. By (11), the left side of (12) holds while the right side fails.

The growth hypothesis (13) is equivalent to

$$\liminf_{n\to\infty} 2^{-n} b(n) = 0.$$

Indeed, writing $n = N + r$, its inequality becomes $2^{-n} b(n) < 1/(2^{N+1} q)$. Since $b(n) \ge 0$, finding such an $n \ge N$ for every N and every positive integer q is exactly the stated liminf condition. In particular it holds for polynomial bounds and for Cc^n with fixed

$C > 0$ and $1 < c < 2$, but fails for $b(n) = 2^n$. A limit of zero is not necessary: the bound $b(n) = 1$ at odd indices and $b(n) = 2^n$ at even indices also satisfies the hypothesis. This example concerns the growth condition alone, not a bound for the prime tails. Only one suitable scale is needed to magnify the positive distance to $\mathbb{Z}$ beyond the error; no comparison at every sufficiently large scale is required.

The proof in fact uses no rationality of D_N: the integer q is chosen from its positive distance to $\mathbb{Z}$, not from a reduced denominator. Thus (14) holds for real integer-coefficient recurrences under the same bound and growth hypothesis. The same is true of (11), and of (12) when the coefficient differences are even. These extensions follow from the printed argument; the cited formal statement is the rational one.

The first test is exactly even integrality of D_N, regardless of r. The fixed-progression test compares $2^r \operatorname{dist}(D_N, \mathbb{Z})$ with the bound on the remaining tail. When D_N is nonintegral, the growth hypothesis makes the former larger than twice the latter at a suitable scale. Rewriting the test does not provide an estimate about the distribution of consecutive primes.

F Conclusions

The sparse construction attains every value in an interval while preserving the specified gap statistics and every fixed congruence eventually. Both rational and irrational values occur, so these properties do not determine rationality. The other counterexamples isolate further insufficient conditions. All these constructions alter integer sequences; they do not construct consecutive primes.

The exact tail criteria characterise rationality but do not produce actual-prime witnesses. One sufficient route to irrationality would establish the simultaneous small-difference conditions arbitrarily late for every positive shift. The finite certificates and denominator exclusions do not establish that conclusion.

References

[1] Paul Erdős, *Sur certaines séries à valeur irrationnelle*. L'Enseignement Mathématique **4** (1958), 93–100, doi:10.5169/seals-34629.

[2] Paul Erdős and Ronald L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*. Monographies de L'Enseignement Mathématique, L'Enseignement Mathématique, 1980.

[3] Paul Erdős and Carl Pomerance, *On the largest prime factors of n and n + 1*. Aequationes Mathematicae **17** (1978), 311–321, doi:10.1007/BF01818569.

[4] Paul Erdős, *On the irrationality of certain series: problems and results*. New Advances in Transcendence Theory, Cambridge University Press, 1988, pp. 102–109, doi:10.1017/CBO9780511897184.009.

[5] Paul Erdős, *Beweis eines Satzes von Tschebyschef*. Acta Litterarum ac Scientiarum Szeged **5** (1932), 194–198.

[6] Paul Erdős and Ernst G. Straus, *On the irrationality of certain Ahmes series*. Journal of the Indian Mathematical Society (N.S.) **27** (1964), 129–133. MR0175848.

[7] Paul Erdős and Ernst G. Straus, *On the irrationality of certain series*. Pacific Journal of Mathematics **55** (1974), no. 1, 85–92, doi:10.2140/pjm.1974.55.85.

[8] Kevin Ford, Ben Green, Sergei Konyagin, James Maynard and Terence Tao, *Long gaps between primes*. Journal of the American Mathematical Society **31** (2018), 65–105, doi:10.1090/jams/876; arXiv:1412.5029v3.

[9] John A. Fridy, *Generalized bases for the real numbers*. The Fibonacci Quarterly **4** (1966), no. 3, 193–201.

[10] A. Ya. Khinchin, *Continued Fractions*. University of Chicago Press, 1964.

[11] Vjekoslav Kovač and Terence Tao, *On several irrationality problems for Ahmes series*. Acta Mathematica Hungarica **175** (2025), 572–608, doi:10.1007/s10474-025-01528-0; arXiv:2406.17593v4. Statement and section numbers refer to arXiv version 4.

[12] Vivian Kuperberg, *Sums of singular series with large sets and the tail of the distribution of primes*. The Quarterly Journal of Mathematics **74** (2023), no. 4, 1457–1479, doi:10.1093/qmath/haad030; arXiv:2210.09775v2. Statement numbers refer to arXiv version 2, 15 June 2023.

[13] Johan Land, *A conditional proof of the irrationality of $\sum_{n\geq 1} p_n 2^{-n}$ under a uniform Hardy–Littlewood prime-tuples conjecture*. Research draft, 5 September 2026. Formalisation material accompanies the draft.

[14] Leonardo de Moura and Sebastian Ullrich, *The Lean 4 theorem prover and programming language*. Automated Deduction – CADE 28, Lecture Notes in Computer Science, Springer, 2021, pp. 625–635, doi:10.1007/978-3-030-79876-5_37.

[15] The mathlib Community, *The Lean mathematical library*. Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, ACM, 2020, pp. 367–381, doi:10.1145/3372885.3373824. Describes a Lean 3-era snapshot; the repository lock identifies the library used by the present Lean 4 sources.

[16] James Maynard, *Small gaps between primes*. Annals of Mathematics **181** (2015), 383–413, doi:10.4007/annals.2015.181.1.7.

[17] Hugh L. Montgomery and Robert C. Vaughan, *Multiplicative Number Theory I: Classical Theory*. Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2007, doi:10.1017/CBO9780511618314.

[18] Jan-Christoph Schlage-Puchta, *The irrationality of some number theoretical series*. Acta Arithmetica **126** (2007), no. 4, 295–303, doi:10.4064/aa126-4-1; arXiv:1105.1451v1. Published in 2007; arXiv upload is from 2011. Lemma and preprint page locators refer to arXiv version 1.

[19] Ian Short, *Ford Circles, Continued Fractions, and Rational Approximation*. The American Mathematical Monthly **118** (2011), no. 2, 130–135, doi:10.4169/amer.math.monthly.118.02.130; arXiv:0912.1997v1. Preprint title: Ford circles, continued fractions, and best approximation of the second kind. Preprint theorem locators refer to arXiv version 1.

[20] Wouter van Doorn and Vjekoslav Kovač, *Lacunary sequences whose reciprocal sums represent all rational numbers in an interval*. Acta Arithmetica **223** (2026), 275–295, doi:10.4064/aa251001-13-1; arXiv:2509.24971v3. Statement numbers refer to arXiv version 3.

[21] Yitang Zhang, *Bounded gaps between primes*. Annals of Mathematics **179** (2014), 1121–1174, doi:10.4007/annals.2014.179.3.7.

[22] Thomas F. Bloom, *Erdős Problem #251*. 2026. Catalogue snapshot accessed 6 September 2026.

[23] Erdős Problems contributors, *Erdős Problem #251 discussion thread*. 2026. Snapshot accessed 6 September 2026: Tao comment of 7 October 2025 and Land comments of 6 September 2026.

[24] ChatGPT 5.4 Pro (orchestrated by Vjeko Kovač), *On the Erdős problem #251*. Unpublished note, Department of Mathematics, University of Zagreb, 2026. Accessed 6 September 2026.

[25] The Formal Conjectures Authors, *FormalConjectures.ErdosProblems.251*. 2025. Pinned statement source, not a proof of Problem 251; accessed 28 July 2026.

[26] Stefan Ringer, *Local gap statistics, telescoping, and normality: a local-pattern approach to Erdős problem 251*. Preprint, 11 September 2026. Mutable main-branch TeX consulted 16 September 2026; the cited conditional statements were rechecked 18 September 2026. Formalisation material accompanies the draft.

[27] Tonći Crmarić and Vjekoslav Kovač, *On the irrationality of certain super-polynomially decaying series*. Colloquium Mathematicum **179** (2025), 55–68, doi:10.4064/cm9628-5-2025; arXiv:2504.18712v1. Lemma 4 is cited using arXiv version 1.

[28] Boris Adamczewski, Michael Drmota and Clemens Müllner, *(Logarithmic) densities for automatic sequences along primes and squares*. Transactions of the American Mathematical Society **375** (2022), no. 1, 455–499, doi:10.1090/tran/8476; arXiv:2009.14773v2. Theorems 1.2 and 1.4 refer to arXiv version 2, 13 April 2021; journal publication is 2022.

[29] Wouter van Doorn, *Partitions with prescribed sum of reciprocals: asymptotic bounds*. 2025; arXiv:2502.02200v2. Version 2, 23 July 2025.

[30] Szymon Głąb and Franciszek Prus-Wiśniowski, *Achievement sets – current results and open problems*. Real Analysis Exchange (2026), doi:10.14321/realanalexch.1766383782; arXiv:2512.17285v1. Advance publication, first available in Project Euclid 8 June 2026. Consulted text remains arXiv:2512.17285v1 (19 December 2025).

[31] Franciszek Prus-Wiśniowski and Jolanta Ptak, *Achievable Cantorvals almost without reversed Kakeya conditions*. 2024; arXiv:2412.08768v1. Version 1 submitted 11 December 2024. The sparse indices satisfy the overlap inequality, not its strict term-dominating reverse.

[32] Artur Bartoszewicz, Małgorzata Filipczak and Emilia Szymonik, *Multigeometric sequences and Cantorvals*. Central European Journal of Mathematics **12** (2014), no. 7, 1000–1007, doi:10.2478/s11533-013-0396-4; arXiv:1304.4218v2.

[33] Vivian Kuperberg, *Sums of singular series along arithmetic progressions and with smooth weights*. International Journal of Number Theory **21** (2025), no. 1, 53–74, doi:10.1142/S1793042125500046; arXiv:2301.06095v1. Preprint uploaded 15 January 2023; journal publication is 2025.

[34] Abhishek Jha, *The Poisson Tail Conjecture for primes in short intervals*. 2026; arXiv:2605.23014v2. Substantially revised version 2, 12 September 2026, 31 pages. Conditional statements must retain their strong Hardy–Littlewood hypotheses.

[35] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*. Research in the Mathematical Sciences **1** (2014), article 12, doi:10.1186/s40687-014-0012-7; arXiv:1407.4897v4. Theorem 1.4(i) refers to arXiv version 4 (22 December 2014); the journal numbers it Theorem 4(i). An erratum is recorded at doi:10.1186/s40687-015-0033-x.

[36] Zbigniew Nitecki, *Cantorvals and Subsum Sets of Null Sequences*. The American Mathematical Monthly **122** (2015), no. 9, 862–870, doi:10.4169/amer.math.monthly.122.9.862; arXiv:1106.3779v2. Consulted preprint: Subsum Sets: Intervals, Cantor Sets, and Cantorvals, version 2 (8 July 2013). Theorem 14 is attributed there to Guthrie–Nymann; its locator is not journal pagination.

[37] Piotr Miska, Franciszek Prus-Wiśniowski and Jolanta Ptak, *More on Kakeya Conditions for Achievement Sets*. Results in Mathematics **78** (2023), article 113, doi:10.1007/s00025-023-01890-x. Repairs an estimate in the 2021 proof, preserving its uniqueness conclusion, and gives a simpler proof of a weaker theorem without that conclusion.

[38] Jacek Marchwicki and Piotr Miska, *On Kakeya Conditions for Achievement Sets*. Results in Mathematics **76** (2021), article 181, doi:10.1007/s00025-021-01479-2. Theorem 2.1 is to be read with the proof repair in Miska–Prus-Wiśniowski–Ptak (2023).

[39] Piotr Nowakowski, *On a new condition implying that an achievement set is a Cantorval and its applications*. 2025; arXiv:2512.17761v1. Version 1, 19 December 2025. Theorem 3.1 requires the Star Procedure of Definition 2 never to break; no application to the present factorial weights is asserted.