A two-class Erdős–Rankin lower bound from finite estimates

SolveAtHome collaboration

Chris Benjaminsen (`@Benjaminsen`, `@MoltkeBenjaminsen`)

`@nielsegberts` `@zemaj` `@MichaelRobartes`

solveathome.org

Documented project contributors; roles in Section 11

LaTeX exposition: attribution revision, 9 October 2026

Abstract

Let $P(y) = \prod_{p \le y} p$, with every prime at or below the real cutoff y included, and let $\mathcal{T}_y$ be the set of integers n for which $\gcd(n(n+2), P(y)) = 1$. We give a readable exposition of the accepted formal argument that the largest consecutive gap between members of $\mathcal{T}_y$ is at least a positive constant times

$$\frac{y(\log y)^3(\log\log\log y)^2}{(\log\log y)^4}$$

for all sufficiently large real y. A three-band residue cover reduces the problem to three finite counting budgets. A finite Selberg sieve bounds the auxiliary survivors, a shifted finite Euler product bounds the smooth exceptions, and an elementary factorial argument supplies enough reserved primes to complete the cover. The sufficient estimates use the fixed parameter $A = 12$ and require no additional prime-number-theorem or smooth-number-asymptotic hypothesis. The finite cyclic maximum is identified with an attained maximum over all integer starts, including negative starts and gaps crossing a period boundary. Seven additional original statements, some stronger, more general, or separate, remain open outside the three accepted mapped conclusions. Their exact domains are preserved in the linked obligation ledger. This edition presents the existing argument; it asserts neither novelty nor a new proof execution.

Edition and attribution. This paper faithfully reuses and reorganizes the mathematical exposition in the accepted project manuscript accompanying return #2597. SolveAtHome provides the collaborative research platform and public evidence record. Chris Benjaminsen supplied project direction and requested this attribution revision. The names beneath the title credit documented contributions; they are not an agreed formal coauthor list. Detailed roles, AI assistance, reviews, and external sources appear in Section 11.

The initial LaTeX draft was written by the selected `gpt-6-astra` AI assistant at high reasoning effort. Sol integrated and compiled it, condensed Section 9, and prepared this attribution revision and the presentation fixes identified in review #698. The accepted source exposition disclosed preparation by `gpt-6.1-sol` at high reasoning effort. This edition preserves the three mapped statements and the accepted mathematical argument; it adds no new proof execution, novelty claim, affiliation, or third-party endorsement.

1 The object and the three mapped conclusions

All logarithms are natural, and $\mathbb{N}$ includes 0. For a finite set S of natural primes put

$$P_S = \prod_{p \in S} p, \qquad \mathcal{T}_S = \{n \in \mathbb{Z} : \gcd(n(n+2), P_S) = 1\}.$$

The gcd is nonnegative also for negative integer arguments. Define $\mathrm{Consec}_S(s, d)$, for $s \in \mathbb{Z}$ and $d \in \mathbb{N}$, to mean

$$d > 0, \qquad s, s+d \in \mathcal{T}_S, \qquad \forall i \in \mathbb{N}\ (0 < i < d \implies s + i \notin \mathcal{T}_S).$$

Thus a distance d contains exactly $d-1$ excluded interior positions. The use of natural offsets loses no integer position strictly between the endpoints.

To define the finite object before asserting any maximum property, set

$$\mathcal{D}_S = \{d \in \mathbb{N} : 1 \le d \le P_S, \quad \exists s \in \mathbb{N}\ (s < P_S \text{ and } \mathrm{Consec}_S(s,d))\}, \qquad G_2(P_S) = \sup \mathcal{D}_S.$$

Here the finite natural supremum has default value 0 for an empty set. The proof below shows that this default never occurs for a prime family. Only the start s is restricted to a fundamental period; the endpoint $s+d$ may be at or beyond P_S.

For a real y, write

$$S_y = \{p \in \mathbb{N} : p \text{ is prime and } (p : \mathbb{R}) \le y\}, \qquad P(y) = P_{S_y}, \qquad F(y) = \frac{y(\log y)^3(\log\log\log y)^2}{(\log\log y)^4}.$$

On the eventual positive domain used throughout, the finite implementation of S_y is exactly the primes in $\{0, \dots, \lfloor y \rfloor\}$. Every prime up to y belongs to this family. The later auxiliary sieve with cutoff $\sqrt{y}$ does not redefine it.

Theorem 1 (Finite gap bound; mapped claim `main_eq1`). *There exist $c \in \mathbb{R}$ with $c > 0$ and $Y \in \mathbb{R}$ such that, for every real $y \ge Y$,*

$$cF(y) \le \big(G_2(P(y)) : \mathbb{R}\big). \tag{1}$$

The exact declaration is `PrimaryMain.manuscript_main_goal` in `PrimaryMain.lean`. Its target definition is `SieveParameters.ManuscriptMainGoal`. The same positive constant and threshold work for all subsequent real y; they are not selected anew for each cutoff. The construction gives $c = 1/\,\mathrm{chooseB}(12)$. It proves the positivity of this coefficient and an eventual real threshold, without claiming an optimized numerical onset.

Theorem 2 (Attained integer maximum; mapped claim `integer_maximum`). *For every $S : \mathrm{Finset}\,\mathbb{N}$ such that every member of S is prime,*

$$\big(\exists s \in \mathbb{Z}\ \mathrm{Consec}_S(s, G_2(P_S))\big) \quad \textit{and} \quad \big(\forall s \in \mathbb{Z}\ \forall d \in \mathbb{N}\ [\mathrm{Consec}_S(s,d) \Longrightarrow d \le G_2(P_S)]\big).$$

The finite prime family S may be empty.

This is `IntegerMainBinding.integer_maximum_binding` in `IntegerMainBinding.lean`. Its explicit Lean parameters are `S : Finset Nat` and `hS : KKFiniteCoreDraft.PrimeFamily S`. The latter means $\forall p \in S$, p is prime. It is a defining domain condition and is retained in this statement.

Theorem 3 (Literal integer bound; mapped claim `integer_main_eq1`). *There exist $c \in \mathbb{R}$, $c > 0$, and $Y \in \mathbb{R}$ such that for every real $y \ge Y$ there is $d \in \mathbb{N}$ satisfying all three conditions*

$$\exists s \in \mathbb{Z}\ \mathrm{Consec}_{S_y}(s,d), \qquad \forall s \in \mathbb{Z}\ \forall e \in \mathbb{N}\ [\mathrm{Consec}_{S_y}(s,e) \Longrightarrow e \le d], \qquad cF(y) \le (d : \mathbb{R}).$$

This is `IntegerMainBinding.manuscript_integer_main_goal` in `IntegerMainBinding.lean`. It follows by taking $d = G_2(P(y))$, using the proved prime-family property of S_y, and applying Theorems 1 and 2. In the accepted inventory, `assumptions: []` means that the final claims have no additional analytic input or custom axiom. It does not erase explicit parameters, domain conditions, or the hypotheses of intermediate lemmas.

Why the maximum is an integer maximum. The polynomial $n(n+2)$ has the same residue at congruent integer arguments modulo P_S. Hence survivor membership is periodic even at negative arguments. The residue $P_S - 1$ survives, since its two factors are congruent to -1 and 1. Every interval of P_S consecutive positions contains a survivor; every consecutive distance is therefore at most P_S. There is a next survivor after any given survivor, so $\mathcal{D}_S$ is nonempty and its finite supremum is attained.

For an arbitrary integer start s, reduce it to its nonnegative remainder modulo P_S. For each offset, adding it before or after this normalization preserves the survivor predicate. An integer consecutive gap consequently has the same distance as a gap starting in $\{0, \dots, P_S - 1\}$. Conversely, every finite natural start is an integer start. This proves both assertions of Theorem 2; it also explains why omitting the crossing gap would change the object. For $S = \varnothing$, $P_S = 1$, every integer survives, and the attained largest distance is 1.

Plan of the mathematical argument. A covering phase turns a block of m excluded positions into a survivor gap of distance at least $m+1$. The early prime choices almost cover that block. The uncovered indices split into auxiliary sieve survivors, a short initial interval, and two smooth-number sets. The three estimates are arranged to fit within a reserve of distinct primes, one for each remaining index. Sections 2–8 give the finite details and the threshold assembly. They are mathematical explanations of the accepted argument, not additional independently mapped or separately reviewed claims of this edition. Section 9 summarizes the additional formalization scope and links the exact preserved original obligation ledger.

2 Covering an interval and producing a gap

At each prime $p \in S$, choose one residue a_p. Its admissible pair is

$$\{a_p, a_p - 2\} \pmod{p}.$$

A cover of $1, \dots, m$ means that for each index i some prime satisfies $i \equiv a_p \pmod{p}$ or $i + 2 \equiv a_p \pmod{p}$. The Chinese remainder theorem supplies a common phase $s \equiv -a_p \pmod{p}$. At that phase these alternatives become $p \mid s + i$ or $p \mid s + i + 2$. Every covered index therefore gives a nonsurvivor $s + i$.

For every $m \in \mathbb{N}$, it follows that

$$\exists (a_p) \text{ covering } 1, \dots, m \quad \Longleftrightarrow \quad m + 1 \le G_2(P_S). \tag{2}$$

For the forward implication, a covered block need not begin immediately after a survivor. One may translate it by a period and choose the greatest preceding survivor; its next survivor lies beyond the block. Their distance is at least $m+1$. For the reverse implication, take the first m interior positions of an attained gap of distance at least $m+1$, and choose the negative of its left endpoint at every prime. This proves (2) with the required endpoints.

The case $m = 0$ is included: the cover is vacuous, while the attained positive gap gives $1 \le G_2(P_S)$. For the empty prime family, $P_S = 1$ and $G_2(P_S) = 1$, so the only coverable length is 0. No gap of distance zero is introduced.

The completion step is finite. Suppose an early set of primes has left R indices uncovered and a disjoint reserved set has at least R primes. An injection assigns a different reserved prime p_i to each uncovered i. Set $a_{p_i} = i \pmod{p_i}$, retaining the old choices at early primes. Every remaining index is now covered. The resulting cover uses the union of the early and reserved prime families, and (2) gives a gap of distance at least $m+1$. No prescribed location of this gap is asserted.

The relevant proofs are the CRT and phase bridges in `Arithmetic` and `Bridges`, the gap arguments in `Normalization` and `Gaps`, and the injection and completion in `Reserved` and `ReserveBudget`.

3 The parameters and the two different residue systems

Write

$$L = \log y, \qquad L_2 = \log L, \qquad L_3 = \log L_2.$$

The construction uses

$$z_0 = L^{12}, \qquad z_1 = \exp\left(\frac{LL_3}{12L_2}\right), \qquad m = \left\lfloor \frac{y}{B} \frac{L^3 L_3^2}{L_2^4} \right\rfloor. \tag{3}$$

The fixed positive constant B will be given explicitly in Section 5. On the proved sufficiently large domain all displayed logarithmic denominators are positive, the mass inside the floor is nonnegative, and

$$3 < z_0 < z_1 < \sqrt{y} < y/2, \qquad \frac{2(m+2)}{y} \le z_0. \tag{4}$$

These inequalities are proved from a real threshold, rather than being assumptions in Theorem 1. For example, the geometry theorem works for $A \ge 4, B \ge 1$, with

$$Y_{\mathrm{geom}}(A) = \max\{\exp(4096 \log 4),\ \exp(\exp(16A))\}.$$

It gives $L \geq 3072$, $L_2 \geq 16A$, and $L_3 \geq 1/2$. Elementary exponential and logarithmic inequalities then prove the ordering and the cofactor bound. The additive 2 in $m+2$ is retained throughout: the second linear form is $i+2$, which can exceed m.

For primes strictly below $y/2$, choose $a_p = 1$ when $z_0 < p \leq z_1$, and choose $a_p = 0$ otherwise. Reserve every prime with $y/2 \leq p \leq y$. Thus the early covering pairs are $\{1, -1\}$ in the middle band and $\{0, -2\}$ elsewhere. The early and reserved sets form an exact partition of the full prime cutoff, including its endpoints.

For counting only, introduce a larger forbidden set at each prime $p \leq \sqrt{y}$:

$$\Omega_p = \begin{cases} \{0, -2, 1, -1\} \pmod p, & z_0 < p \leq z_1, \\ \{0, -2\} \pmod p, & \text{otherwise.} \end{cases} \tag{5}$$

The actual cover still has one pair at each prime. Formula (5) describes an auxiliary sieve, whose benefit is an additional factor from the middle band. Its local cardinality $g(p)$ is 1 at $p = 2$, 2 at $p = 3$, 4 at middle-band primes, and 2 at other primes at least 5. A collision of the four middle-band residues would force a prime divisor of 1, 2 or 3; the band lies above 3, so no collision occurs. All local densities satisfy $0 < g(p)/p < 1$.

The code supplies canonical natural representatives and proves their equality with these sets in $\mathbb{Z}/p\mathbb{Z}$. In particular, the single class at $p = 2$ is accounted for explicitly.

4 The exact count of indices still needing a prime

Let R count the positive indices $1 \leq i \leq m$ missed by all early covering pairs, and let $S(m, \Omega)$ count indices in that same interval avoiding all auxiliary sets (5). Let

$$\Psi(X, Z) = \#\{1 \leq n \leq X : \ p \mid n, \ p \text{ prime} \Rightarrow p \leq Z\}.$$

This count includes 1, excludes 0, and uses an inclusive real prime cutoff. In the construction $X = m + 2$ is a natural number and $Z = z_1$ is real.

The proved finite inequality is

$$R \leq S(m, \Omega) + \lfloor \sqrt{y} + 2 \rfloor + 2\Psi(m + 2, z_1). \tag{6}$$

To see the dichotomy behind it, suppose an early-uncovered index fails the auxiliary avoidance condition. A hit in the middle-band pair $\{1, -1\}$ would already cover it. Therefore a prime $p \leq \sqrt{y}$ divides $k = i$ or $k = i + 2$. If k is prime, it is small. In the remaining case consider its largest prime factor q. Early noncoverage implies every prime factor of k lies in $(z_0, z_1]$ or in $[y/2, \infty)$. If $q \leq z_1$, the number is smooth. Otherwise $q \geq y/2$, while

$$1 < k/q \leq 2(m + 2)/y \leq z_0.$$

The cofactor has a prime factor at most z_0, contradicting the earlier exclusion. Hence each uncovered index belongs to the auxiliary survivor set, the short initial interval, or one of the two smooth sets. Cardinalities of unions give (6); overlaps only improve the upper bound.

At the same large threshold, `ReserveBudget.short_floor_budget` proves

$$\lfloor \sqrt{y} + 2 \rfloor \leq \frac{y}{24L}. \tag{7}$$

The other two terms will be bounded by $y/(4L)$ and $y/(24L)$, respectively. These exact budgets sum to $y/(3L)$.

5 A finite sieve and its explicit sufficient constant

We describe the proved sieve instance, including where its constants come from. For $1 \leq n \leq X$, define the incidence label

$$D(n) = \prod_{\substack{p \leq \sqrt{y} \\ n \bmod p \in \Omega_p}} p, \qquad w_t = \#\{1 \leq n \leq X : D(n) = t\}.$$

The labels form a finite set and the weights are nonnegative cardinalities with total mass X. Every label divides the auxiliary squarefree prime product. For every divisor d of that product, simultaneous local hits specify exactly $g(d) = \prod_{p|d} g(p)$ residue classes modulo d, by CRT. Each class has interval count differing from X/d by at most 1, even if $d > X$. Thus the actual weights satisfy

$$\sum_{d|t} w_t = \frac{Xg(d)}{d} + r_d, \qquad |r_d| \le g(d). \tag{8}$$

The sifted weight sum is exactly $S(X, \Omega)$. These identities are established for the counting weights, not supplied as unspecified sieve hypotheses.

Put $\nu(p) = g(p)/p$, extending multiplicatively on squarefree divisors, and $h(d) = \prod_{p|d} \nu(p)/(1-\nu(p))$. At a real level $H > 1$, the Selberg denominator is

$$J(H) = \sum_{\substack{d|P_{\text{aux}} \\ d^2 \le H}} h(d).$$

The explicit coefficients are

$$\lambda_d = \nu(d)^{-1} J(H)^{-1} \mu(d) \sum_{\substack{k|P_{\text{aux}} \\ d|k,\ k^2 \le H}} h(k).$$

They have $\lambda_1 = 1$, vanish outside $d^2 \le H$, and their quadratic main term is exactly $1/J(H)$, by finite Möbius inversion and diagonalization. Expanding the square against (8) gives the bound by $X/J(H)$ plus the actual remainder sum. The construction uses the squared-level convention above throughout.

The remainder is controlled coarsely but explicitly. One has $|\lambda_d| \le d$ on supported divisors. A squarefree n has $3^{\omega(n)}$ ordered divisor pairs with least common multiple n. Their aggregate coefficient therefore has absolute value at most $n^2 3^{\omega(n)}$, while $g(n) \le 4^{\omega(n)}$. Only $n \le H$ contributes. The factors at 2 and 3 give

$$12^{\omega(n)} \le 4n^2,$$

so the entire remainder is at most $4H^5$.

For a prime subset Q, the subset expansion of $\prod_{p \in Q}(1 + h(p))$ has logarithmic first moment $\sum_{p \in Q} \nu(p) \log p$. If this is at most $\log H/4$, at least half the mass has $d^2 \le H$. Consequently $J(H) \ge \frac{1}{2} \prod_{p \in Q}(1 - \nu(p))^{-1}$. A finite prime-log estimate provides

$$\sum_{p \le N} \frac{\log p}{p} \le \log 4\,(2 + \log N), \qquad N \ge 2. \tag{9}$$

It follows from the elementary primorial bound $\theta(x) \le x \log 4$ and finite summation, and does not use the prime number theorem.

Choose $k = 512 \log 4$, $N = \lfloor \exp(L/k) \rfloor$, and $H = \exp(L/16)$. These satisfy the required moment condition on primes up to N. Passing from that restricted Euler product to the full auxiliary product costs a bounded factor. The proof bounds its logarithm by

$$E = 20 \log 4\, k + 80 \log 4.$$

Also $4H^5 \le y/(12L)$ on the stated domain. With $C_s = 2e^E$, the resulting finite estimate is

$$S(X, \Omega) \le C_s X V(\sqrt{y}) + \frac{y}{12L}, \qquad V(\sqrt{y}) = \prod_{p \le \sqrt{y}} (1 - g(p)/p). \tag{10}$$

The product is bounded by elementary finite estimates as well. Define $C = 4 \exp(10 \log 4)$. The ordinary inverse product obeys

$$\prod_{p \le N} (1 - 1/p)^{-1} \le C \log N \quad (N \ge 2),$$

and the harmonic sum gives the complementary upper bound for the ordinary product. Splitting a prime band therefore bounds its product by $C \log N_0 / \log N_1$. At primes at least 3, the basic two-class

factor is at most the square of the ordinary factor; middle-band primes supply two additional powers. The exceptional prime 2 costs at most a factor 2. Converting real cutoffs to floors, with the proved logarithmic floor bounds, gives

$$V(\sqrt{y}) \le 128C^2 \frac{(\log z_0/\log z_1)^2}{L^2}. \tag{11}$$

Set $K = C_s\, 128C^2$. Substitution of (3), using the natural floor bound, yields

$$S(m,\Omega) \le \left(\frac{K12^4}{B} + \frac{1}{12}\right)\frac{y}{L}.$$

The actual choice is

$$B = 1 + 6K12^4,$$

which is positive, at least 1, and gives $S(m,\Omega) \le y/(4L)$. These constants are deliberately generous; they are fixed expressions, not unspecified $O(1)$ factors. The proof establishes a sufficient one-sided estimate. It does not assert the original two-sided product asymptotic.

6 The smooth-number budget at $A = 12$

For any $\alpha > 0$, $X \ge 1$, and $Z \ge 2$, the finite Rankin inequality is

$$\Psi(X,Z) \le X^\alpha \prod_{p\le Z}(1-p^{-\alpha})^{-1}. \tag{12}$$

Indeed, each smooth $1 \le n \le X$ has $1 \le X^\alpha n^{-\alpha}$. Its weighted finite sum embeds into the sum over positive integers factored entirely over the finite set of primes $p \le Z$. The latter is a product of convergent geometric series because each $p^{-\alpha} < 1$. This is local convergence for a finite prime family; the all-natural series need not converge when $\alpha \le 1$. The identification with the actual smooth count uses `Nat.smoothNumbersUpTo` $X\,(\lfloor Z\rfloor_+ + 1)$, preserving the inclusive condition $p \le Z$ despite that library's strict bound convention.

Let $U = \log X/\log Z \ge 1$, suppose $\log Z \ge 1$ and $U \le \sqrt{Z}$, and set

$$\delta = \frac{\log U}{2\log Z}, \qquad \alpha = 1-\delta.$$

Then $0 \le \delta \le 1/4$, $\alpha > 0$, and $p^\delta \le \sqrt{U}$ for every prime $p \le Z$. The shifted local ratio satisfies

$$\frac{1-1/p}{1-p^{\delta-1}} = 1 + \frac{p^\delta - 1}{p - p^\delta} \le \exp\left(\frac{4(p^\delta-1)}{p}\right). \tag{13}$$

For this bound, $p^\delta \le \sqrt{p} \le 3p/4$ makes the denominator positive and at least $p/4$; then $1 + t \le e^t$ applies. The inequality $e^x - 1 \le xe^x$ gives

$$p^\delta - 1 \le \delta \log p\sqrt{U}.$$

Summing (13)'s exponent with (9), and using $2 + \log Z \le 3\log Z$, bounds the total shift cost by $6\log 4\sqrt{U}\log U$. Combining the product with the ordinary inverse bound and

$$X^{1-\delta} = X\exp(-U\log U/2)$$

proves

$$\Psi(X,Z) \le CX\log Z\exp\left(-\tfrac{1}{2}U\log U + 6\log 4\sqrt{U}\log U\right). \tag{14}$$

Now use the actual $X = m+2, Z = z_1, U = u$. The proved parameter estimates give $X \le 2yL^3$, $\log Z \le L$, and eventually $1 \le u \le \sqrt{Z}$, $Z \ge e$. They also give the quantified limits

$$\frac{\log X}{L} \longrightarrow 1, \qquad \frac{\log u}{L_3} \longrightarrow 1, \qquad \frac{u\log u}{L_2} \longrightarrow 12,$$

and $u \longrightarrow \infty$. For clarity, these limits can be read directly from quantitative identities. Put $r = \log X/L$. The floor and mass estimates give $1 \le r \le 1 + 7/\sqrt{L} \le 2$, and

$$u = r\frac{12L_2}{L_3}, \qquad \log u = L_3 + \log 12 - \log L_3 + \log r.$$

The bound $\log t \le 2\sqrt{t}$ makes both $r - 1$ and $(\log u)/L_3 - 1$ arbitrarily small at explicit sufficiently large thresholds. Their product gives $(u \log u)/L_2 \to 12$. Moreover $u \ge \sqrt{L_2}$ proves its growth. The same parameter estimates give $u \le 48L_2$ and $\log Z \ge \sqrt{L}/48$; at the chosen larger threshold, a quadratic exponential bound implies $48L_2 \le \sqrt{Z}$. Thus the range in (14) is proved for these exact parameters. These are parameter limits, not estimates for the number of smooth integers.

Take the eventual error tolerance $1/2$, so $u \log u \ge (23/2)L_2$, and enlarge the threshold until $\sqrt{u} \ge 288 \log 4$. Then the positive shift term in (14) is at most $u \log u/48$. Its total exponent is therefore at most

$$-\frac{23}{48}u \log u \le -\frac{529}{96}L_2 \le -\frac{11}{2}L_2.$$

The factor $X \log Z \le 2yL^4$ and $\exp(-11L_2/2) = L^{-11/2}$ give

$$\Psi(m + 2, z_1) \le \frac{2Cy}{L\sqrt{L}}.$$

Finally enlarge the threshold until $\sqrt{L} \ge 96C$. It follows that

$$2\Psi(m + 2, z_1) \le \frac{y}{24L}. \tag{15}$$

Every range and growth condition used here is discharged in `SmoothBudget.eventual_smooth_and_uncovered_budget`. Together with (6), (7), and the sieve quarter-budget, it proves the actual bound $R \le y/(3L)$. The loss of a factor roughly $1/2$ in the decay exponent is acceptable at $A = 12$; this argument does not establish the full original smooth-number asymptotic or its asserted consequence for every fixed $A > 4$.

7 Enough primes in the reserved interval

We next prove the needed literal supply

$$\#\{p \text{ prime} : y/2 \le p \le y\} \ge \frac{y}{3 \log y}. \tag{16}$$

This is derived by a factorial argument. To avoid confusing smooth counts with prime-power counts, write $\psi(N) = \sum_{n \le N} \Lambda(n)$ and $\theta(N) = \sum_{p \le N} \log p$, where $\Lambda(p^j) = \log p$ for prime p and $j \ge 1$, and Λ vanishes off positive prime powers.

Put $f(x) = x \log x - x$, with $f(0) = 0$. Logarithmic ratio inequalities show, for positive integers n,

$$\log n \le f(n + 1) - f(n) \le \log(n + 1).$$

Induction gives, for $N \ge 1$,

$$N \log N - N + 1 \le \log(N!) \le N \log N - N + 1 + \log N.$$

For $F(x) = \log(\lfloor x \rfloor!)$, this implies the uniform estimate

$$|F(x) - f(x)| \le 1 + \log(x + 1) \quad (x \ge 0). \tag{17}$$

For $x \ge 1$, let $n = \lfloor x \rfloor$. Logarithmic ratio inequalities give $(x - n) \log n \le f(x) - f(n) \le (x - n) \log x$. Since $0 \le x - n < 1$, this difference lies between 0 and $\log x$; combining it with the factorial bounds gives (17). The range $0 \le x < 1$ is handled separately, where the factorial is $0! = 1$ and $-1 \le f(x) \le 0$. Thus applying (17) at $x/30$ does not require $x \ge 30$.

Define the signed floor wheel and the factorial combination

$$w(n) = n - \lfloor n/2 \rfloor - \lfloor n/3 \rfloor - \lfloor n/5 \rfloor + \lfloor n/30 \rfloor,$$

$$T(x) = F(x) - F(x/2) - F(x/3) - F(x/5) + F(x/30), \qquad a = \frac{7}{15}\log 2 + \frac{3}{10}\log 3 + \frac{1}{6}\log 5.$$

Adding 30 changes the five terms in the wheel by $30, 15, 10, 6, 1$, whose signed combination is zero. Checking the thirty remainders therefore proves that the wheel takes only 0 and 1; it equals 1 for $1 \le n < 6$. These are signed integer identities; truncated natural subtraction is not used. The exact combination of the five f-terms is ax, so (17) gives

$$|T(x) - ax| \le 5(1 + \log(x+1)). \tag{18}$$

The divisor identity $\sum_{d|n} \Lambda(d) = \log n$, followed by finite summation, gives

$$\log(N!) = \sum_{n \le N} \Lambda(n)\lfloor N/n \rfloor.$$

Substituting it in the factorial combination yields the exact identity

$$T(N) = \sum_{n \le N} \Lambda(n) w(\lfloor N/n \rfloor).$$

Nonnegativity and the wheel properties now imply

$$0 \le T(N) \le \psi(N), \qquad \psi(N) - \psi(\lfloor N/6 \rfloor) \le T(N). \tag{19}$$

With $L_N = 1 + \log(N+1)$, strong induction proves

$$aN - 5L_N \le \psi(N) \le \frac{6}{5}aN + 5L_N^2. \tag{20}$$

For the upper bound at $N \ge 6$, take $M = \lfloor N/6 \rfloor$. The induction hypothesis and (19) bound $\psi(N)$ by $aN + 5L_N + (6/5)aM + 5L_M^2$. Since $M \le N/6$, the linear terms are at most $(6/5)aN$. Also $N + 1 \ge 3(M+1)$ and $\log 3 \ge 1$, so $L_M + 1 \le L_N$. Hence $L_N^2 - L_M^2 \ge L_N$, absorbing the new error. For $N < 6$, the elementary bound $\psi(N) \le N \log(N+1)$ handles the base case.

The higher-prime-power tail satisfies, for $N \ge 2$,

$$0 \le \psi(N) - \theta(N) \le 2\sqrt{N}(\log N)^2. \tag{21}$$

Indeed, every nonprime prime power $n \le N$ is p^j with $p \le \lfloor \sqrt{N} \rfloor$ and $2 \le j \le \lfloor \log N / \log 2 \rfloor$. This finite set lies in the image of those base-exponent pairs. Its size is at most the number of pairs, without needing an injectivity assertion. Each summand is at most $\log N$, and $\log 2 \ge 2/3$ gives the stated loose coefficient 2.

The constants needed for a positive dyadic margin are proved rationally. The inequalities $(24/23)^{16} \le 2$, $(12/11)^{12} \le 3$, and $(8/7)^{12} \le 5$, together with $\log r \ge 1 - r^{-1}$, give

$$\log 2 \ge 2/3, \qquad \log 3 \ge 1, \qquad \log 5 \ge 3/2, \qquad a \ge 31/36.$$

Also $a \le 2$ follows from $\log t \le t - 1$ at $2, 3, 5$.

For real $y \ge 4$, use $N = \lfloor y \rfloor$, $M = \lfloor y/2 \rfloor$. Combining (20) and (21), with $N \ge y - 1$ and $M \le y/2$, proves

$$\theta(y) - \theta(y/2) \ge \frac{2a}{5}y - \mathcal{E}(y),$$

$$\mathcal{E}(y) = a + 5L_+ + 5L_+^2 + 2\sqrt{y}(\log y)^2, \qquad L_+ = 1 + \log(y+1).$$

There is a concrete error threshold. If $y \ge 300000^4$ and $\log y \ge 1$, put $s = \sqrt{\sqrt{y}} \ge 300000$. Then $y = s^4$, $\sqrt{y} = s^2$, and $\log y = 4 \log s \le 8\sqrt{s}$. Therefore $(\log y)^2 \le 64s$, $L_+ \le 10s$, and $L_+^2 \le 576s$. It follows that

$$\mathcal{E}(y) \le 128s^3 + 2932s \le 3060s^3 \le y/90.$$

The last inequality uses $90 \cdot 3060 = 275400 \le s$. Since $2a/5 \ge 31/90$, the difference of theta sums is at least $y/3$.

This difference sums $\log p$ over the strict interval $y/2 < p \le y$. Every such prime belongs to the reserved inclusive interval and contributes at most $\log y$. Hence it is at most the reserved cardinality times $\log y$, proving (16). `PrimeReserveBudget.actual_reserved_supply` uses the explicit threshold $\max\{300000^4, e\}$. No interval asymptotic is used or proved.

8 Completion, the coefficient, and the final quantifiers

Take a threshold large enough for the three uncovered-index budgets and (16). The actual uncovered cardinality satisfies

$$R \le y/(3L) \le \#\{p : y/2 \le p \le y\}.$$

Since both endpoint objects here are finite natural cardinalities, this real comparison gives exactly the cardinality inequality required for the injection. Section 2 then completes the cover using every prime at or below y, and proves

$$m + 1 \le G_2(P(y)).$$

Set $c = 1/B > 0$, with the fixed B of Section 5. On the positive large domain the floor mass in (3) equals $c \cdot yL^3L_3^2/L_2^4$. The elementary floor inequality gives

$$c\frac{yL^3L_3^2}{L_2^4} < m + 1 \le G_2(P(y)),$$

which proves (1), with its original logarithmic powers and full prime cutoff. The integer maximum characterization in Section 1 makes this literally a lower bound for the largest consecutive gap over all of $\mathbb{Z}$.

The threshold is a finite maximum of proved sufficient thresholds. In the smooth stage it includes the parameter threshold, the thresholds from the three quantified growth lemmas, and $\exp((96C)^2)$. The prime-supply stage has the explicit threshold above. Thus the existential eventual theorem has no remaining analytic premise. We do not claim an optimized or practical numerical onset for (1), nor a constant uniform in arbitrary choices of A, B.

A weaker endpoint also follows from the same proved main bound. On the geometry domain, $L_2^4 \le 64L$ and $LL_3^2 \ge 64$; hence $y \log y \le yL^3L_3^2/L_2^4$. `PrimaryMain.eventual_y_log_gap` therefore proves $G_2(P(y)) \gg y \log y$. This endpoint is an additional proved consequence, outside the three accepted mapped claims. It does not formalize the original separate short construction or its priced sieve constant.

A lower-bound construction supplies no upper bound for G_2, no conclusion about infinitely many twin primes, and no assertion about a specified interval containing a twin pair.

9 Additional formalization scope

The three main conclusions above have a complete accepted proof at their stated scope. None requires O1–O7 as an additional assumption: the finite argument uses the sufficient estimates in Sections 5–7 and the fixed choice $A = 12$. "Open" in the project ledger means that an additional original statement has not been closed by the accepted mapping. It does not assert that every listed statement is an unsolved mathematical problem; several are known external analytic results awaiting the specified formalization or correspondence check.

The exact original statements, domains, quantifiers, equation numbers, and historical source qualifications remain in the hash-pinned accepted source manuscript and its proof-to-exposition obligation ledger. The following summary links their existing follow-ups; it neither creates new jobs nor changes the status of an original obligation.

- **O1:** Mertens prime-harmonic asymptotic (original Eq. 8); follow-up #2599.
- **O2:** the uniform smooth-number asymptotic in its original range (Eq. 9); follow-up #2600.
- **O3:** the route for every fixed $A > 4$ (Eqs. 23–24). The parameter limits in Eq. 23 are already proved for $A \ge 4$; the smooth equality in Eq. 24 and its $o(y/L)$ consequence remain open in this mapping, beyond the sufficient $A = 12$ route proved here; follow-up #2601.

- **O4:** the dyadic prime-count asymptotic (Eq. 10), stronger than the finite reserve bound used here; follow-up #2602.
- **O5:** the external one-class comparison and the general quoted sieve (Eqs. 5–7, 11), beyond the sufficient finite sieve bound; follow-up #2603.
- **O6:** the original harmonic-density expansion and two-sided Euler-product comparisons (Eqs. 20–21); follow-up #2604.
- **O7:** the separate short construction and its priced constant (Eq. 28); follow-up #2605.

Each obligation remains open outside the three mapped conclusions. Closing one would extend the formal record; it would not repair a missing hypothesis of the main theorem proved in this paper.

10 Exact source correspondence and assurance

The three theorem labels in Section 1 provide distinct locators for the three accepted claim IDs. Their correspondence is:

Claim ID	Exact accepted declaration
`main_eq1`	`PrimaryMain.manuscript_main_goal`
`integer_maximum`	`IntegerMainBinding.integer_maximum_binding`
`integer_main_eq1`	`IntegerMainBinding.manuscript_integer_main_goal`

The accepted `proof-to-exposition-map.json` records the frozen source correspondence for the mathematical sections. In particular, the finite core definitions in `FiniteCoreTargets.lean` are specifications; the target definitions themselves are not asserted as proofs. Their actual proofs are supplied by the arithmetic, bridge, normalization, and gap modules, and the final three declarations above supply the accepted endpoints.

The dependencies of the exposition are worth distinguishing. The covering step belongs to `Arithmetic`, `Bridges`, `Normalization`, `Gaps`, and `Reserved`. Actual local residue counts and incidence moments belong to `ManuscriptResidues`, `ResidueMoment`, `CRTMoments`, `IncidenceMoments`, and `DivisorMoments`. The finite sieve passes through `ActualBoundingSieve`, `SelbergFinite`, `SelbergError`, `SelbergEuler`, `SieveCoarse`, `SieveInterface`, `EulerRatio`, and `SieveParameters`. The ordinary-product estimates in `PrimeLog` and `MertensBand`, and their real-cutoff versions in `RealBand`, feed the explicit choice in `ManuscriptParameters`. The smooth budget uses `SmoothRankin`, `ShiftedRankin`, `SmoothParameters`, `SmoothLogLimit`, and `SmoothBudget`. The independent reserve-supply estimate uses `FactorialLogBounds`, `FactorialWheel`, `PrimeFactorial`, `PrimePsiBounds`, `PrimePowerTail`, and `PrimeReserveBudget`. Finally, `EarlyCover` and `ReserveBudget` assemble the cover before `PrimaryMain` and `IntegerMainBinding` yield the mapped claims. These source correspondences do not enlarge the inventory of accepted claims to the whole prose paper.

10.1 Public evidence and the edition boundary

The accepted source is return #2597. Its recorded execution is return #2598, receipt #31, under the named `lean-kernel-v1` assurance profile. The exact statements were addressed by statement review #695; independent review #696 accepted the three mapped conclusions. The latter records a `claude-opus-5-5` review at high effort, with spot inspection of the retained evidence and no Lean rerun. It expressly excludes O1–O7, the additional $y \log y$ consequence, and the expository prose of the intermediate sections from its mapped coverage.

These are records of the accepted source and proof package. The present LaTeX edition has different bytes and has not thereby inherited an independent fidelity review of its own. Its review status belongs to its ordinary submission record. No claim that this edition has already passed that review is made here.

For reproducible identification, the exact source manuscript has 41,786 bytes and SHA-256

`6c287c180b331318a19b3fb0a0e571147e01d785c4efa4341cf3bba5bd1c8e80`.

The accepted package fingerprint is

`91107e1d12018bd0b37bcf47471394133c5c1d3fc793900b990727acedc986e6`,

and the statement binding is

`cf2f5db08226d9151d245d8e103296d61d99de317df7e3af05eab0af6deda232`.

These identifiers refer to the accepted inputs; they are not a hash or a certification of this new LaTeX file.

The receipt reports a pass, matching target statements and proof bodies, standard-axiom auditing, and replay of 23,756 reachable declarations in an empty Lean kernel environment. Four negative controls—a mistaken statement, a mistyped proof, `sorryAx`, and a custom axiom—were rejected. The axiom policy permits only `propext`, `Classical.choice`, and `Quot.sound`; the three target reports list those standard axioms. There is no extra mathematical premise named Input S, M, H, or P in the three final claims.

This evidence has a precise trust boundary. The run used the released Lean 4.35.0-rc3 compiler and its kernel, together with retained build objects and their checked custody. No fresh clean source rebuild of the solution and dependencies was performed, and the compiler was not built from source during that check. The proof replay used one Lean kernel; there was no independent second kernel. The stronger comparator/export validation, primitive-export stage, and Nanoda check were not run. Retained exports were supplementary and unchecked under this profile. The execution was author-run; the independent review inspected its evidence rather than performing a new execution. No Lean build, replay, or proof check was run for this LaTeX drafting task. None of these records establishes a full-paper proof badge or the broader analytic obligations.

11 Mathematical lineage, contributions, and licenses

The construction follows the three-band architecture of Kalmynin and Konyagin [1] and the existing fixed-offset adaptation in the project record. Their polynomial-root extremal problem differs from the fixed-separation pair $n, n+2$ here. Their theorem is not transferred to this object merely by a change of notation. The present exposition makes no novelty or literature-completeness claim. Bibliographic metadata for [1, 2] was checked for this edition; the historical sieve citations and source-access qualifications in O1–O7 are preserved rather than represented as newly audited literature.

Platform and documented project contributions

SolveAtHome hosts the research assignments, manuscript versions, public source files, proof evidence, and independent review records that made this paper possible. The title-page collaboration credit belongs to that platform and the documented contributors below. Account handles identify submitted work and its operators; the accompanying model record identifies AI-performed work. A handle does not establish that its human operator personally derived every calculation. We use contribution-role attribution rather than infer formal coauthorship consent.

Chris Benjaminsen. The public profiles `@Benjaminsen` and `@MoltkeBenjaminsen` identify Chris Benjaminsen. The project contribution record credits his research direction, correction of the attribution policy, project stewardship, and authorization to publish. The latter account's source-access report #3 and OCR-custody report #4 are inputs cited by the later source synthesis; their printed-page access limitations remain explicit. Under `@Benjaminsen`, return #1730 corrected the manuscript's analytic-source ranges, pinned-version citations, and short-construction constant. The disclosure repair in return #2222 preserved the earlier Astra preparation and separated it from the later corrections. This account also operated the accepted formalization, source-manuscript preparation, and check in source #2597 and check #2598, and operated the distinct-model reviewers listed below. These operational credits do not assign all mathematical or AI-performed work to Chris.

`@zemaj`. Return #49 records an early revision of this lower-bound manuscript; it was rejected pending corrections, so its numerical and source claims are not thereby accepted in this edition. Accepted return #158 contributed the historical cross-source sieve/OCR synthesis and priced-constant discussion. Its constant was subsequently corrected in #1730 and #2222. We credit the investigation and early drafting while preserving those limitations and later corrections.

`@nielsegberts`. Accepted return #1093 prepared the earlier three-band manuscript: the fixed-offset adaptation, finite incidence-weight explanation, analytic-input distinctions, and source-inspection

disclosures. Its AI worker was recorded as `gpt-6-astra`, with GitHub Copilot CLI named in the disclosure. This earlier contribution remains distinct from the new Astra LaTeX draft.

`@MichaelRobartes`. Review #73 critically reviewed the early #49 manuscript, identifying needed repairs to the sieve interface, finite ledger, example identification, normalizations, and claim calibration. The recorded worker was `gpt-6-astra`. This is historical review credit, not acceptance of a later edition.

AI assistance, formalization, and independent review

The public records support the following separate roles:

- `claude-fable-5-1` was the recorded AI worker on `@MoltkeBenjaminsen`'s source reports and `@zemaj`'s early manuscript/source synthesis. `gpt-6-astra` prepared the earlier `@nielsegberts` manuscript and performed the `@MichaelRobartes` review. `deepseek-v4-flash`, with Freebuff named in the historical disclosure, prepared the corrections in #1730.
- Sol (`gpt-6.1-sol`) prepared the accepted source exposition and formalization package #2597 and operated the authenticated worker check #2598, recorded in receipt #31. These are AI preparation and operator/execution roles under `@Benjaminsen`; the assurance limits are stated in Section 10.
- `claude-opus-5-5` supplied ordinary independent reviews under `@Benjaminsen`: the historical manuscript reviews #327 and #493, the accepted statement review #695, and the final source-and-receipt judgment #696. The model family differed from the author; the same account handle was disclosed. These reviewers assessed finished submissions, not coauthored the exposition or performed another Lean execution.
- The selected `gpt-6-astra`/high wrote the initial LaTeX draft for edition #2611. Sol integrated its layout, condensed the additional-obligation section, and compiled it. `claude-opus-5-5`/high accepted that exact edition's fidelity in review #698. Sol prepared this new attribution/presentation revision, reusing the same accepted mathematical evidence. Review #698 remains attached to the unchanged #2611 files; this revision requires its own fidelity verdict.

The earlier return and review histories remain the authority for their respective scopes. This list traces the records used by this paper; it is not a reconstruction of every classical fact's first inventor or a claim that third parties approved named coauthorship.

External mathematics, code, and licenses

The following adapted-source acknowledgments remain attached to the Lean files and are retained here for readers of the exposition:

- `SelbergFinite.lean` and `SelbergError.lean` adapt finite Möbius/diagonalization and lambda-square/lcm-pair arguments from Arend Mellendijk's PNT+, revision `2667e414c38e5a5dc9aa1946f16f13001e5cd3ed`. They preserve the 2023 copyright, author credit, and Apache 2.0 notice.
- `PrimeLog.lean`, `MertensBand.lean`, and `SelbergEuler.lean` adapt `openai/math`, revision `adc7f1241b42e322a6451854ab7e4b4c146bf78a`, under Apache 2.0. The headers identify the finite weighted prime-log, Mertens-factor/Rankin, finite Euler truncation, and denominator arguments and their local changes.
- `PrimeLog.lean` also credits Mathlib's `Chebyshev.lean`, copyright 2025 Alastair Irving, and names Alastair Irving, Terry Tao, and Ruben Van de Velde. The pinned Mathlib dependencies retain their contributors' notices and attribution.

Project prose is distributed under CC BY 4.0; this edition adapts the accepted exposition under those prose terms with attribution. The adapted Lean sources retain their separate Apache 2.0 notices and licenses. These acknowledgments do not grant new rights in external material or settle the distinct source-ledger repair described in job #3880. The seven open obligations already have follow-up returns #2599–#2605 and jobs #5411–#5417; recording them here creates no new follow-up or closure claim.

References

[1] A. B. Kalmynin and S. V. Konyagin, *A polynomial analogue of Jacobsthal function*, Izvestiya: Mathematics **88** (2) (2024), 225–235. doi:10.4213/im9467e. arXiv:2302.00459v2.

[2] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, Journal de Théorie des Nombres de Bordeaux **5** (2) (1993), 411–484. Primary journal record.

[3] Solve@Home twin-primes project, *A two-class Erdős–Rankin lower bound from finite estimates*, accepted source manuscript accompanying return #2597, with proof-to-exposition map, statement bundle, and frozen Lean sources; execution evidence in return #2598, statement review #695, and independent review #696. This is the primary source for the argument and the historical quotations in this edition.