hexagon:2611.00754v1

THE SHARP ASYMPTOTIC FOR DISTINCT MULTIPLES IN AN INTERVAL: A LEAN FORMALIZATION

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

ABSTRACT. Let $f(n)$ be the least nonnegative integer h for which the open interval $(n, n+h)$ contains distinct integers $a_1, \dots, a_n$ satisfying $i \mid a_i$. We give a complete Lean 4 formalization of the asymptotic formula

$$f(n) \sim \frac{\sqrt{2}}{e}\, n\sqrt{\log n}.$$

This formula was previously announced by Turturean. The lower bound uses completely multiplicative weights, an exact convolution over powers of 2 and 3, and two inverse-exponential profiles. For the upper bound, we construct fractional matching weights from finite histograms of prime valuations. Explicit oscillation and stability estimates transfer the continuous transport identities to the integer divisibility graph, including the smallest sources. The companion development proves both bounds for all sufficiently large integers and checks their correspondence with the original open-interval minimum. Its complete theorem has been replayed, together with its imported declarations, by Lean's kernel from an empty environment.

1. INTRODUCTION

Choosing a multiple of each positive integer just beyond n is individually easy. Requiring all of the chosen multiples to be distinct produces a global matching problem. Erdős and Pomerance studied this phenomenon in [3]; the diagonal asymptotic is recorded as Erdős problem 710 [2, 1].

For $n \in \mathbb{N}$, define $f(n)$ to be the least $h \in \mathbb{N}$ such that there is an injection $a : \{1, \dots, n\} \to \mathbb{N}$ with

$$n < a(i) < n + h, \qquad i \mid a(i) \quad (1 \le i \le n). \tag{1.1}$$

The minimum exists: $a(i) = (n+1)i$ is admissible when $h = n^2 + 1$. We take $f(0) = 0$, which agrees with the empty matching convention.

Theorem 1.1. *For the minimum in (1.1),*

$$\lim_{n\to\infty} \frac{f(n)}{n\sqrt{\log n}} = c_*, \qquad c_* := \frac{\sqrt{2}}{e}. \tag{1.2}$$

The limit is along all natural numbers.

The formula in Theorem 1.1 was announced earlier by David Turturean [8]. We do not claim priority for its mathematical discovery. The present contribution is a complete formalization and a finite-histogram implementation of the upper-bound transport. In particular, the proof does not posit an unproved approximation principle for atomic prime-factor distributions. The lower bound follows the elementary two-prime convolution method described in the earlier manuscript.

Date: November 2, 2026.

2020 *Mathematics Subject Classification.* 11N37, 05D15, 68V15.

Key words and phrases. distinct multiples, divisibility matching, prime number theorem, weighted counting, Lean, formal verification.

Email addresses: howew@purdue.edu (Will Howe), huffma97@purdue.edu (Grant Huffman), jenwin@purdue.edu (Jensen Kohlmeyer), kruerl@purdue.edu (Liam Kruer).

Throughout the proof we put

$$L = \log n, \qquad S = n\sqrt{L}. \tag{1.3}$$

All assertions involving these quantities are made for sufficiently large n, so that $L > 0$. The argument proves a leading asymptotic and does not assert an effective threshold or a second-order term.

AI assistance. OpenAI Codex assisted with mathematical proof development, Lean code, and the preparation of this manuscript. The verification claims below refer to the recorded compiler and kernel checks. They are not a claim of independent human refereeing or of acceptance by a journal or prize committee.

2. The finite matching reductions

Let $M(n)$ be the least right endpoint T for which distinct multiples of $1, \dots, n$ can be selected from $(n, T] \cap \mathbb{N}$. The two conventions satisfy the exact identity

$$f(n) = M(n) - n + 1 \qquad (n \geq 1). \tag{2.1}$$

Indeed, the integers in $(n, T]$ are precisely those in $(n, n + (T - n + 1))$. Consequently, a matching with right endpoint $T = \lfloor cS \rfloor$ gives $\limsup f(n)/S \leq c$.

We use two familiar consequences of finite matching theory. They are included to specify the precise inequalities transferred to Lean.

Lemma 2.1 (Weighted obstruction)**.** *Suppose $U_i \geq 0$ and $V_m \geq 0$ are weights on $1 \leq i \leq n$ and $n < m \leq T$, respectively. If $U_i \leq V_m$ whenever $i \mid m$, then a matching into $(n, T]$ implies*

$$\sum_{i=1}^{n} U_i \leq \sum_{n < m \leq T} V_m.$$

Proof. Sum $U_i \leq V_{a(i)}$ along the matching and use injectivity and nonnegativity of the target weights. □

Lemma 2.2 (Fractional matching)**.** *Let $w(i, m) \geq 0$ vanish unless $i \mid m$ and $n < m \leq T$. If*

$$\sum_{m} w(i, m) \geq 1 \quad (1 \leq i \leq n), \qquad \sum_{i=1}^{n} w(i, m) \leq 1 \quad (n < m \leq T), \tag{2.2}$$

then there is a matching into $(n, T]$.

Proof. For a set A of sources, summing the row inequalities over A and then the column inequalities over its set $N(A)$ of neighbors gives $|A| \leq |N(A)|$. Hall's theorem [5] gives the matching. □

3. Multiplicative weights and scalar profiles

For $y > 0$ and a prime $p > 3$, put

$$b_y(p) = \max\left\{0, \frac{1 - \sqrt{p/y}}{1 + \sqrt{p/y}}\right\}. \tag{3.1}$$

Assign weight 1 to the primes 2 and 3, and extend completely multiplicatively to positive integers:

$$\rho_y(m) = \prod_{\substack{p^a \| m \\ p > 3}} b_y(p)^a.$$

In particular, prime powers retain their full multiplicities.

Lemma 3.1. *For every positive integer $q < y$,*

$$\rho_y(q) \geq \frac{1-\sqrt{q/y}}{1+\sqrt{q/y}}. \tag{3.2}$$

Proof. Write $h(t) = (1-t)/(1+t)$. For $a, b \geq 0$, clearing positive denominators gives

$$h(a+b) \leq h(a)h(b),$$

since the difference of the cross-products is $2ab(a+b) \geq 0$. Let $p_1, \ldots, p_r$ list, with multiplicity, the prime factors of q that exceed 3. Each $\sqrt{p_j} \geq 2$, and $x + y \leq xy$ for $x, y \geq 2$. Induction, including the empty-list case, therefore gives

$$\sum_{j=1}^{r} \sqrt{p_j} \leq \prod_{j=1}^{r} \sqrt{p_j} \leq \sqrt{q}.$$

Put $t_j = \sqrt{p_j/y}$. Then $\sum_j t_j \leq \sqrt{q/y} < 1$. The function h is decreasing and nonnegative on $[0,1]$. Induction using the displayed two-variable inequality yields

$$h\left(\sqrt{q/y}\right) \leq h\left(\sum_j t_j\right) \leq \prod_j h(t_j) = \rho_y(q).$$

For $r = 0$, the last product is 1, and the conclusion still holds. □

For $t \in [0,1]$, let $u(t), v(t) \in [0,1]$ be the unique solutions of

$$ue^{u-1} = t, \qquad ve^{1-v} = t, \tag{3.3}$$

and set $\Phi(t) = 1 - u(t)$ and $H(t) = 1 + v(t)$. The defining maps are continuous and strictly increasing on $[0,1]$.

Lemma 3.2. *For $r, t \in [0,1]$,*

$$\Phi(t) \leq H(r)\frac{1-\sqrt{rt}}{1+\sqrt{rt}}. \tag{3.4}$$

Proof. Write $x = u(t)$ and $z = v(r)$. The exponential inequality $(1-d)e^d \leq 1$, with $d = (x-z)/2$, gives $(2+z-x)e^{(x-z)/2} \leq 2$. Multiplication by $\sqrt{xz}$ and the arithmetic-geometric mean inequality give

$$(2+z-x)\sqrt{xz}\, e^{(x-z)/2} \leq x + z.$$

Since $\sqrt{rt} = \sqrt{xz}\, e^{(x-z)/2}$, rearrangement is exactly (3.4). □

For a fixed endpoint T, define

$$U_i = \rho_y(i)\Phi\left(\frac{T}{yi}\right)\mathbf{1}_{i>T/y}, \qquad V_m = \rho_y(m)H(m/T) \quad (m \leq T). \tag{3.5}$$

If $i \mid m \leq T$ and $i > T/y$, write $m = iq$. Then $q < y$ and $(m/T)(T/(yi)) = q/y$. Lemmas 3.1 and 3.2, together with multiplicativity, imply $U_i \leq V_m$. For $i \leq T/y$ the inequality follows from $U_i = 0$. Thus (3.5) supplies a weighted obstruction on every divisibility edge, not merely on average.

4. Counting powers of two and three

Put $A = \log 2$, $D = \log 3$, and

$$\Psi(t) = \#\{(a,b) \in \mathbb{N}^2 : Aa + Db \le t\},$$

with $\Psi(t) = 0$ when $t < 0$.

Lemma 4.1. *As $t \to \infty$,*

$$\Psi(t) = \frac{t^2}{2AD} + \frac{A+D}{2AD}t + o(t). \tag{4.1}$$

Consequently, for every fixed $0 < c_1 < c_2$ and $C > 0$,

$$\Psi(t+s) - \Psi(t) = \frac{ts}{AD} + \frac{s^2 + (A+D)s}{2AD} + o(L) \tag{4.2}$$

uniformly for $c_1 L \le t \le c_2 L$ and $|s| \le C \log L$.

Proof. Let $N = \lfloor t/A \rfloor$. Exact summation gives

$$\Psi(t) = (N+1)\left(\frac{t}{D} + 1\right) - \frac{A}{2D}N(N+1) - \sum_{a=0}^{N}\left\{\frac{t}{D} - a\frac{A}{D}\right\}.$$

The number A/D is irrational, since a rational relation would give an equality between nontrivial powers of 2 and 3. For an irrational rotation, the average of the fractional-part function tends uniformly in the initial phase to $1/2$. One proves this first for Fourier characters by the finite geometric-series formula, then for continuous periodic functions by uniform approximation, and finally for fractional parts by continuous upper and lower approximations near the jump. Hence the last sum is $(N+1)/2 + o(N)$ uniformly in the phase. Expanding $N = t/A + O(1)$ proves (4.1).

The remainder divided by t is uniformly small once t exceeds a fixed threshold. Both t and $t+s$ lie in a fixed positive multiple of L for large L. Subtracting the two expansions therefore gives (4.2), uniformly on the stated range. □

We also use the elementary bound

$$0 \le \Psi(t) \le (1 + t/A)(1 + t/D) \qquad (t \ge 0). \tag{4.3}$$

5. An exact probabilistic convolution

Every positive integer has a unique representation $2^a 3^b r$ with $(r, 6) = 1$. Consequently, for $X > 0$,

$$F_y(X) := \sum_{m \le X} \rho_y(m) = \sum_{\substack{r \le X \\ (r,6)=1}} \rho_y(r)\Psi(\log(X/r)). \tag{5.1}$$

This is an exact counting identity.

For each prime $3 < p \le y$, let E_p be an independent geometric random variable with

$$\mathbb{P}(E_p = k) = (1 - b_y(p))b_y(p)^k \quad (k \ge 0),$$

and put

$$U = \sum_{3<p\le y} E_p \log p, \qquad Z_y = \prod_{3<p\le y} (1 - b_y(p))^{-1}. \tag{5.2}$$

The random integer $R = \prod p^{E_p}$ has probability $\rho_y(r)/Z_y$ at every supported integer r. Unique factorization therefore turns (5.1) into

$$F_y(e^t) = Z_y \mathbb{E}\Psi(t - U). \tag{5.3}$$

Lemma 5.1. *We have $\mathbb{E}U \sim y/2$ as $y \to \infty$. If $y = \beta L$, where $\beta > 0$, and $\theta > \beta/2$, then there is $c > 0$ such that*

$$\mathbb{P}(U > \theta L) \le e^{-c\sqrt{L}} \tag{5.4}$$

for all sufficiently large L.

Proof. For $p \le y$,

$$\frac{b_y(p)}{1 - b_y(p)} = \frac{\sqrt{y/p} - 1}{2}.$$

The prime number theorem, with partial summation, gives

$$\sum_{p \le y} \frac{\log p}{\sqrt{p}} \sim 2\sqrt{y}, \qquad \sum_{p \le y} \log p \sim y.$$

Removing 2 and 3 changes the resulting expression by $O(\sqrt{y})$, so

$$\mathbb{E}U = \frac{\sqrt{y}}{2} \sum_{3 < p \le y} \frac{\log p}{\sqrt{p}} - \frac{1}{2} \sum_{3 < p \le y} \log p \sim \frac{y}{2}.$$

For the tail, choose a small $0 < \delta < 1/2$ and let $\lambda = \delta/(2\sqrt{y})$. The elementary inequality $\log p \le 2\sqrt{p}$ implies $\lambda \log p \le \delta(1 - b_y(p))$. Using $e^t - 1 \le (1 + \delta)t$ for $0 \le t \le \delta$, the geometric generating function and $-\log(1 - x) \le x/(1 - x)$ give

$$\log \mathbb{E}e^{\lambda U} \le K_\delta \lambda \mathbb{E}U, \qquad K_\delta = \frac{1 + \delta}{1 - \delta(1 + \delta)}. \tag{5.5}$$

Here δ is chosen so the denominator is positive, and $K_\delta \to 1$ as $\delta \downarrow 0$. Choose it so that $K_\delta \beta/2 < \theta$. Markov's inequality then bounds the probability in (5.4) by $\exp(-\lambda(\theta L - K_\delta \mathbb{E}U)) \le e^{-c\sqrt{L}}$. □

Proposition 5.2 (Uniform weighted counting)**.** *Fix $0 < \beta < 2$, $\beta/2 < \theta < 1$, and $C > 0$. For $y = \beta L$, define*

$$W = W_{\beta,\theta}(L) := \frac{Z_y}{AD} \mathbb{E}\big[(L - U)\mathbf{1}_{U \le \theta L}\big]. \tag{5.6}$$

Then $W > 0$ for large L, $W \asymp Z_y L$, and

$$F_y(e^{L+s}) - F_y(e^L) = sW + o(W) \qquad (|s| \le C \log L) \tag{5.7}$$

uniformly in s.

Proof. On $U \le \theta L$, the variable $t = L - U$ lies in $[(1 - \theta)L, L]$. Apply (4.2) inside (5.3). The linear term has expectation sW/Z_y; the quadratic and remaining linear terms are $O((\log L)^2) = o(L)$. The uniform remainder is $o(L)$. On the exceptional event, the two lattice counts are at most $O(L^2)$ by (4.3), while (5.4) makes their expected contribution $o(L)$. Multiplication by Z_y proves the asserted error. Finally, $(1 - \theta)L \le L - U \le L$ on the main event and that event has probability tending to one. Thus W is bounded above and below by positive constant multiples of $Z_y L$. □

6. The lower bound

The two profile losses can be evaluated exactly:

$$\int_0^\infty u(e^{-s})\,\mathrm{d}s = \frac{3}{2}, \qquad \int_0^\infty v(e^{-s})\,\mathrm{d}s = \frac{1}{2}. \tag{6.1}$$

Indeed, (3.3) gives $s = -\log u + 1 - u$ on the left and $s = -\log v - 1 + v$ on the right, reducing the integrals to $\int_0^1 (1 + u)\,\mathrm{d}u$ and $\int_0^1 (1 - v)\,\mathrm{d}v$.

Fix $0 < \kappa < c_*$ and choose

$$\kappa^2 e^2 < \beta < 2, \qquad \beta/2 < \theta < 1. \tag{6.2}$$

Let $T = \lfloor \kappa S \rfloor$, $y = \beta L$, and

$$a = \log(T/n), \qquad b = \log(ny/T). \tag{6.3}$$

Both windows are positive and $O(\log L)$, and

$$b - a \longrightarrow \log(\beta/\kappa^2) > 2. \tag{6.4}$$

Lemma 6.1. *With the weights in (3.5),*

$$\sum_{i \le n} U_i \ge W\left(b - \frac{3}{2} - o(1)\right), \qquad \sum_{n < m \le T} V_m \le W\left(a + \frac{1}{2} + o(1)\right). \tag{6.5}$$

Proof. Place mass $\rho_y(m)$ at $\log(m/n)$. Proposition 5.2 says that the cumulative mass of this measure differs from that of $W\,\mathrm{d}s$ by $o(W)$, uniformly on $[-b, a]$. For a bounded monotone test function, integration by layers bounds the integral discrepancy by the cumulative discrepancy times a constant depending only on its height. This statement applies to atomic measures as well as continuous measures.

For sources, the test function on $[-b, 0]$ is $\Phi(e^{-b-s})$, of height at most 1; for targets it is $H(e^{s-a})$, of height at most 2. Their integrals against $\mathrm{d}s$ are

$$b - \frac{3}{2} + u(e^{-b}) + \frac{u(e^{-b})^2}{2}, \qquad a + \frac{1}{2} - v(e^{-a}) + \frac{v(e^{-a})^2}{2},$$

respectively. The first is at least $b - 3/2$, and the second at most $a + 1/2$. These inequalities prove (6.5). □

By (6.4), the source total in (6.5) is eventually strictly greater than the target total. Lemma 2.1 rules out a matching into $(n, T]$. Using (2.1) and then letting $\kappa \uparrow c_*$ gives

$$\liminf_{n \to \infty} \frac{f(n)}{S} \ge c_*. \tag{6.6}$$

7. A sharp variational inequality

We turn to the upper bound. The analytic estimate that determines the constant is elementary.

Lemma 7.1. *For $m \ge 0$ and $v > 0$,*

$$\frac{m}{(1 + 2m)v} \le \frac{m}{2} + P(v), \qquad P(v) := \frac{\max(\sqrt{2} - \sqrt{v}, 0)^2}{4v}. \tag{7.1}$$

For $0 < \ell \le 2 \le V$,

$$\int_\ell^V P(v)\,\mathrm{d}v = \frac{1}{2}\log(2/\ell) - \frac{3}{2} + \sqrt{2\ell} - \frac{\ell}{4}. \tag{7.2}$$

Proof. For $v \ge 2$, the left side of (7.1) is at most $m/2$. For $v < 2$, maximizing $m/((1+2m)v) - m/2$ over $m \ge 0$ gives $1 + 2m = \sqrt{2/v}$ and maximum $P(v)$; equivalently, clearing denominators proves the inequality by a square. On $(0, 2)$,

$$P(v) = \frac{1}{2v} - \frac{\sqrt{2}}{2\sqrt{v}} + \frac{1}{4}.$$

Integrating and using $P = 0$ above 2 gives (7.2). □

Fix $c > c_*$. Choose $b > 0$ small, $B > 1$, $0 < \tau < 1$, and $w > 0$ so that

$$\tau^{-1}\left(\frac{B}{2} + \frac{1}{2}\log\frac{2}{b^2} - \frac{3}{2} + \sqrt{2}b - \frac{b^2}{4}\right) < \log(c/b) - 4w, \tag{7.3}$$

and $b < \min(1, c/2)$. Such choices exist: at $B = \tau = 1$ the expression in parentheses is $\log(c_*/b) + \sqrt{2}b - b^2/4$, which is strictly less than $\log(c/b)$ for small b. The strict inequality persists after a small increase of B, a small decrease of τ, and a sufficiently small positive choice of w.

8. Transport on a finite prime histogram

Choose $s_0 > 0$ small enough that $V := c/s_0 > 2$. Let

$$v_k = b^2 + kh \quad (0 \le k \le N), \qquad Nh = V - b^2. \tag{8.1}$$

The mesh $h > 0$ will be fixed sufficiently small before n is allowed to grow. Write $\nu_p(j)$ for the exponent of the prime p in j. For an integer $j \ge 1$, define its logarithmic prime masses and histogram densities by

$$\mu_k(j) = \sum_{v_k L < p \le v_{k+1} L} \nu_p(j) \frac{\log p}{L}, \qquad m_k(j) = \frac{\mu_k(j)}{h}. \tag{8.2}$$

Unique factorization gives the exact budget

$$\sum_{k=0}^{N-1} m_k(j) h \le \frac{\log j}{L}. \tag{8.3}$$

For any nonnegative histogram (m_k) of total mass at most B, put

$$C_k = \frac{\tau}{m_k + \tau(1 + m_k)}, \quad D_k = C_k(1 + m_k), \quad S_k = \sum_{r=k}^{N-1} D_r \log(v_{r+1}/v_r). \tag{8.4}$$

Then $0 < C_k, D_k \le 1$, $S_N = 0$, and

$$m_k C_k = \tau(1 - D_k), \qquad 1 - D_k \le \tau^{-1} \frac{m_k}{1 + 2m_k}. \tag{8.5}$$

The total deficit $\mathcal{D}$ satisfies

$$\begin{aligned} \mathcal{D} &:= \sum_k (1 - D_k) \log(v_{k+1}/v_k) \\ &= \log(V/b^2) - S_0 \\ &\le \tau^{-1} \left(\frac{B}{2} + \int_{b^2}^{V} P(v)\, \mathrm{d}v \right) < \log(c/b) - 4w. \end{aligned} \tag{8.6}$$

The middle inequality follows by integrating (7.1) on each constant-density bin. In particular,

$$S_0 > -\log s_0 - \log b + 4w. \tag{8.7}$$

Let $K_w(t) = \max(w - |t|, 0)/w^2$, so that $K_w \ge 0$, its support is $[-w, w]$, and its integral is 1. Set $\sigma_0 = \log s_0 - 2w$, and for $v \in (v_k, v_{k+1}]$ define

$$G_m(v, \sigma) = \frac{C_k}{v} K_w\left(\sigma - \sigma_0 - S_k + D_k \log(v/v_k)\right). \tag{8.8}$$

Values at bin endpoints are assigned consistently to the half-open bins. When estimating a single bin, we use the continuous extension of its displayed formula to that bin's two endpoints; changing endpoint values does not change any of the integrals.

Proposition 8.1. *The kernel (8.8) has the following properties:*

(i) *If* $\log s_0 \le \sigma \le -\log b + w$, *then*

$$\sum_k \int_{v_k}^{v_{k+1}} (1+m_k) G_m(v,\sigma)\,\mathrm{d}v = 1.$$

(ii) *For every* $\vartheta \in \mathbb{R}$,

$$\sum_k \int_{v_k}^{v_{k+1}} m_k G_m(v,\vartheta - \log v)\,\mathrm{d}v \le \tau.$$

(iii) *If* $G_m(v,\sigma) \ne 0$, *then*

$$\log b + w < \sigma + \log v < \log c - w. \tag{8.9}$$

Proof. Set $x = \log v$. In the row integral the coefficient becomes $C_k(1+m_k) = D_k$, while the kernel argument has slope D_k in x. The integral on a bin is therefore the integral of K_w from $\sigma - \sigma_0 - S_k$ to $\sigma - \sigma_0 - S_{k+1}$. These intervals telescope. By (8.7), their union covers $[-w, w]$ on the indicated source range.

For a column, the argument has slope $D_k - 1$. The identity $m_k C_k = \tau(1 - D_k)$ makes the bin integral τ times an integral of K_w. The endpoint map

$$x \longmapsto x + S_k - D_k(x - \log v_k)$$

has slope $1 - D_k \ge 0$ and joins continuously across the bins. The column integrals consequently telescope to τ times the integral of K_w on an interval, which is at most τ. This also handles bins with $m_k = 0$, where the column contribution vanishes.

The same monotone endpoint map takes values between $\log(b^2) + S_0$ and $\log V$. A nonzero kernel value differs from the corresponding map value, after adding σ_0, by less than w. Using (8.7) and $V = c/s_0$ proves (8.9). □

9. Uniform discretization and one-prime stability

Set $\ell = b^2$. On every bin, uniformly over all nonnegative histograms, the kernel is bounded by $1/(\ell w)$. Both functions $v \mapsto G_m(v,\sigma)$ and $v \mapsto G_m(v,\vartheta - \log v)$ have oscillation at most

$$Kh, \qquad K = \frac{1}{\ell^2 w^2} + \frac{1}{\ell^2 w}. \tag{9.1}$$

To see this, use the $1/w^2$ Lipschitz bound for K_w, $|\log v - \log u| \le |v-u|/\ell$, and $0 \le C_k, D_k \le 1$. For the column argument, $|D_k - 1| \le 1$ provides the same estimate.

Lemma 9.1 (Finite quadrature)**.** *Let g have oscillation at most ε about its left-endpoint value on an interval of length h. Let nonnegative atomic weights in that interval have total mass M. For $q \ge 0$,*

$$\left| \sum_r a_r g(x_r) - q \int g(v)\,\mathrm{d}v \right| \le \varepsilon(M + qh) + |M - qh|\,|g(v_{\mathrm{left}})|. \tag{9.2}$$

Proof. Subtract $g(v_{\mathrm{left}})$ from both the finite sum and the integral. Their two errors are bounded by εM and $\varepsilon q h$; the constant terms differ by $(M - qh)g(v_{\mathrm{left}})$. □

Let

$$P_k(L) = \sum_{v_k L < p \le v_{k+1} L} \frac{\log p}{L}.$$

The prime number theorem gives $P_k(L) \to h$ for every fixed bin. Applying Lemma 9.1 to the row weights $(1+\nu_p(j))\log p/L$ and density $1+m_k(j)$ yields a total row error bounded by

$$E_L = Kh\left(\sum_k P_k(L) + Nh + 2B\right) + \frac{1}{\ell w}\sum_k |P_k(L) - h|. \tag{9.3}$$

Thus $E_L \to 2Kh(Nh+B)$. For columns with the fixed histogram of j, the atomic mass is exactly $m_k(j)h$; the corresponding error is at most $2KhB$.

The histogram used by an actual incoming edge is that of j/p, rather than that of j. The following estimate controls this change.

Lemma 9.2. *For fixed ℓ, h, w, τ, N, there is a constant C_0 such that multiplying an integer j by one prime p changes its kernel by at most*

$$C_0\frac{\log p}{Lh} \tag{9.4}$$

at every point of every bin, uniformly in the coordinate σ. One admissible choice is

$$C_0 = \frac{1+\tau}{\tau\ell w} + \frac{(2N+1)h}{\tau\ell^2 w^2}.$$

Proof. Multiplication by p increases only its bin mass, by $\log p/L$ if it is in the selected prime range. Hence every density changes by at most $\delta = \log p/(Lh)$. Direct subtraction in (8.4) gives

$$|C(m) - C(m')| \le \frac{1+\tau}{\tau}|m-m'|, \qquad |D(m) - D(m')| \le \tau^{-1}|m-m'|.$$

Each logarithmic bin length is at most h/ℓ. Bounding the differences of the two prefix sums that define S_k gives $|S_k - S_k'| \le 2N\delta h/(\tau\ell)$. Apply the Lipschitz bound for K_w and the bound $|\log(v/v_k)| \le h/\ell$ in (8.8). This gives (9.4) with the displayed constant. □

Since $p \le VL$, the right side of (9.4) is bounded by $C_0\log(VL)/(Lh) = o(1)$. Combining this with the column quadrature and (8.3), the total regular column weight, before a final amplification, is at most

$$\tau + \left(2Kh + C_0\frac{\log(VL)}{Lh}\right)B. \tag{9.5}$$

In reindexing incoming edges, only primes that divide the target are included. For a nondivisor, its valuation is zero; no artificial edge from the integer quotient is added.

10. Completion of the upper matching

Choose an amplification factor $\gamma > 1$ with $\gamma\tau < 1$. Next choose $s_0 > 0$ small enough, and then the fixed mesh h small enough, that

$$\gamma(1 - 2Kh(Nh+B)) > 1, \qquad \gamma(\tau + 2KhB) + \frac{2\gamma s_0}{c} < 1. \tag{10.1}$$

This is possible because $Nh = V - b^2$ remains fixed while $h \downarrow 0$. All parameters are now fixed independently of n.

Put

$$T = \lfloor cS \rfloor, \quad q = \lfloor b\sqrt{L} \rfloor, \quad I = \frac{s_0 S}{L}. \tag{10.2}$$

For large n, $q \ge 1$, $VL \le T$, $2nq < T$, and $\log T/L \le B$.

Sources $i > n/q$ are assigned the distinct targets qi, which lie in $(n, nq]$. It remains to match $i \le n/q$ into $(nq, T]$.

For the regular sources $I < i \le n/q$, use edges $i \mapsto ip$ with weight

$$\gamma \frac{\log p}{L} \nu_p(ip) G_{m(i)}\left(\frac{p}{L}, \log \frac{iL}{S}\right) \tag{10.3}$$

when p belongs to one of the selected bins. The identity $\nu_p(ip) = \nu_p(i) + 1$ makes the row sum exactly the discretized row integral. Since

$$\log s_0 < \log(iL/S) \le -\log b + w$$

for all such sources once n is large, (9.3) and (10.1) make every regular row sum at least 1. Moreover,

$$\log(iL/S) + \log(p/L) = \log(ip/S),$$

so (8.9) puts every nonzero target strictly between $be^w S$ and $ce^{-w} S$. These bounds place it in $(nq, T]$ for large n.

For the small sources $i \le I$, use primes $T/(2i) < p \le T/i$ with weights

$$\gamma \frac{2i}{T} \log p \, \nu_p(ip). \tag{10.4}$$

The row sum is at least

$$\gamma \frac{2i}{T} \left(\vartheta(T/i) - \vartheta(T/(2i))\right),$$

where $\vartheta(x) = \sum_{p \le x} \log p$. It tends uniformly to γ, because $T/i \ge Tq/n \to \infty$ and $\vartheta(X) - \vartheta(X/2) \sim X/2$. Thus each small row also has weight at least 1. Its targets lie in $(T/2, T] \subset (nq, T]$.

For a fixed target j, the incoming small-source weight is bounded by

$$\gamma \frac{2I}{T} \sum_{p \mid j} \nu_p(j) \log p \le \gamma \frac{2I}{T} \log j \le \gamma \frac{2I}{T} \log T \longrightarrow \frac{2\gamma s_0}{c}. \tag{10.5}$$

For regular sources, use (9.5) with the histogram of j and apply Lemma 9.2 to each actual incoming prime edge. The combined small and regular column sum is at most

$$\gamma \left(\tau + \left(2Kh + C_0 \frac{\log(VL)}{Lh}\right) B\right) + \gamma \frac{2I}{T} \log T < 1$$

for all sufficiently large n, by (10.1). The large sources use the disjoint target interval $(n, nq]$. Lemma 2.2 completes the matching.

We have proved that every $c > c_*$ admits a matching with right endpoint $\lfloor cS \rfloor$ for all sufficiently large n. Therefore

$$\limsup_{n \to \infty} \frac{f(n)}{S} \le c_*. \tag{10.6}$$

Combining (6.6) and (10.6) proves Theorem 1.1.

11. Formalization and verification

Theorem 1.1 is formalized in Lean 4 as the unconditional declaration `Erdos710.erdos_710`. Its statement uses the exact open-interval minimum in (1.1) and the limit along all natural numbers. Both bounds have complete proofs, with no `sorry`, `admit`, or additional mathematical assumptions. The development has passed a fresh kernel replay. Source revisions and reproducibility details are recorded in Appendix A.

Appendix A. Reproducibility of the Lean formalization

The companion development is identified by commit

`7844854432654d58bfbf6f39c4fdac3da2d736ed`.

It was rebuilt in isolation with Lean 4.35.0-rc3 and Mathlib [6] at revision

`c55e6e786f49471c72fbddbec5415808896aec1e`.

The prime number theorem is provided by the attributed source slice of [7], at revision

`c39a751132c88b6e8080b74c74023fd95b3d8be0`.

The finite matching definition indexes the divisors $1, \dots, n$ by `Fin n`, with divisor `i.val + 1`. The minimum is defined using `Nat.find` and its existence follows from the quadratic construction in Section 1. The development proves its equality with the `sInf` definition in the Formal Conjectures project [4]. The final statement is

```
Erdos710.erdos_710 :  Erdos710.SourceAsymptoticFormula.
```

It has no hypotheses and asserts the asymptotic equivalence of the original minimum with $(\sqrt{2}/\exp(1))n\sqrt{\log n}$.

The recorded verification rebuilt 86 Lean source modules and passed 493 explicit theorem axiom audits. A fresh kernel replay checked every local and imported declaration in the dependency closure of the final theorem from an empty environment. Its only axioms are `propext`, `Classical.choice`, and `Quot.sound`, the standard Lean foundations. The replay uses the kernel bundled with the pinned Lean version. The repository documentation supplies the build instructions and the correspondence between source modules and the proof above.

References

[1] T. F. Bloom, *Erdős problem 710*, Erdős Problems, https://www.erdosproblems.com/710, accessed November 2, 2026.

[2] P. Erdős, *Some of my forgotten problems in number theory*, Hardy–Ramanujan Journal **15** (1992), 34–50. https://doi.org/10.46298/hrj.1992.125.

[3] P. Erdős and C. Pomerance, *Matching the natural numbers up to n with distinct multiples in another interval*, Indagationes Mathematicae (Proceedings) **83** (1980), no. 2, 147–161. https://doi.org/10.1016/1385-7258(80)90018-9.

[4] The Formal Conjectures Authors, *Formal Conjectures*, definitions of Erdős Problems 710 and 711. Source blobs: `e65522825737943cf4de3b41fa00e101288af6f3` (710); `d0170ecd49a0621d4f3e951de1d0ec9690dab6a5` (711). https://github.com/google-deepmind/formal-conjectures.

[5] P. Hall, *On representatives of subsets*, Journal of the London Mathematical Society **s1-10** (1935), no. 1, 26–30. https://doi.org/10.1112/jlms/s1-10.37.26.

[6] The Mathlib Community, *Mathlib*, Lean mathematical library. The revision used is specified in Appendix A. https://github.com/leanprover-community/mathlib4.

[7] A. Kontorovich and contributors, *PrimeNumberTheoremAnd*. The revision used is specified in Appendix A. https://github.com/AlexKontorovich/PrimeNumberTheoremAnd.

[8] D. Turturean, *An asymptotic formula for distinct multiples in an interval*, manuscript, version compiled November 1, 2026 (25 pages). Public announcement, September 23, 2026; manuscript.