IMPROVED LONG GAPS BETWEEN PRIMES

OPENAI

Abstract. We prove that, for all sufficiently large X,

$$G(X) \gg \frac{\log X (\log_2 X)^2 \log_4 X}{(\log_3 X)^2},$$

where $G(X)$ is the largest gap between consecutive primes not exceeding X and $\log_j$ denotes the j-fold iterated logarithm. The proof is due to GPT 6 Astra.

1. Introduction

Let p_n denote the nth prime and let $G(X) = \max_{p_{n+1} \le X} (p_{n+1} - p_n)$. We write $\log_j$ for the j-fold iterated logarithm. Westzynthius [15] first proved that $G(X)/\log X \to \infty$. Erdős [3] subsequently obtained

$$G(X) \gg \frac{\log X \log_2 X}{(\log_3 X)^2},$$

and Rankin [14] improved this to

$$G(X) \gg \frac{\log X \log_2 X \log_4 X}{(\log_3 X)^2}. \tag{1.1}$$

Erdős discussed the difficulty of improving Rankin's estimate in [4], and later asked whether its constant could be made arbitrarily large [5], originally offering a \$10,000 prize for this improvement [1]. This was proved independently by Maynard [11] and by Ford, Green, Konyagin and Tao [8]. Their subsequent joint work [9] gave the quantitative bound

$$G(X) \gg \frac{\log X \log_2 X \log_4 X}{\log_3 X}.$$

More recently, an argument by GPT 5.6 Sol [13], posted on the Erdős Problems website [1], gives the further bound

$$G(X) \gg \frac{\log X \log_2 X}{\log_4 X},$$

improving the work of Ford, Green, Konyagin and Tao and the subsequent joint work with Maynard. We prove the following bound.

Theorem 1.1. *There is an absolute constant $c > 0$ such that, for all sufficiently large X,*

$$G(X) \ge c \frac{\log X (\log_2 X)^2 \log_4 X}{(\log_3 X)^2}.$$

This improves (1.1) by a factor $\log_2 X$. The main input is the following statement about translates of a set of integers. Put $Q(x) = \prod_{p \le x} p$.

Proposition 1.2 (Short translates)**.** *There is an absolute constant $0 < \delta < 1/2$ such that the following holds for all sufficiently large x. Let $x < H \le x(\log x)^2$, let $Q = Q(x)$, and let $S \subseteq [1, H] \cap \mathbb{Z}$ have cardinality $k \le \delta x$. For every integer $0 \le b < Q$, there is an integer $1 \le t \le e^x$ such that $b + Qt + s$ is composite for every $s \in S$.*

The estimates are uniform in H, b, and S. Section 2 outlines the choice of weight and its truncation. Section 3 deduces Theorem 1.1 from Proposition 1.2; Section 4 proves the proposition. All implicit constants are absolute unless a dependence is indicated.

In the accompanying GitHub repository `openai/LongGapsBetweenPrimes`, we have included a complete Lean proof of Theorem 1.1 in `LongGapsBetweenPrimes.lean`.

2. Proof outline

The main task is Proposition 1.2. Thus $Q = \prod_{p \le x} p$, $S = \{s_1, \dots, s_k\} \subseteq [1, H]$, $H \le x(\log x)^2$, and $k \le \delta x$, with $0 \le b < Q$ prescribed. Write $L_i(t) = b + Qt + s_i$ for the integers we want to make composite. To describe the scales, take a large fixed constant C and a small fixed constant $0 < \kappa < 1/3$, and put

$$z = x^C, \qquad y = \exp\Big(\frac{x}{(\log x)^C}\Big), \qquad D = e^{\kappa x}, \qquad T = \lfloor e^x \rfloor.$$

We choose δ sufficiently small in terms of these constants. The auxiliary primes lie in $(z, y]$, and $P = \prod_{z < p \le y} p$ is their product. For large x, we have $H < z < y < Q$. The lower cutoff z makes interactions between different forms small; the upper cutoff y leaves room to truncate a product of divisor sums at D and still average it over $1 \le t \le T$.

Call a form uncovered if it has no prime factor in P. We seek a nonnegative weight for which the weighted expected number of uncovered forms is less than 1. This supplies a translate where every form has a proper prime factor, since $L_i(t) \ge Q > y$. We first consider the complete average $\mathbb{E}$ over a uniformly chosen residue $t \bmod P$, and then pass to $1 \le t \le T$. Under this complete average, residue choices at distinct primes are independent, while a prime $p > H$ can divide at most one form.

For each auxiliary prime p and each divisor $d \mid P$, define

$$\psi_{i,p}(t) = \begin{cases} -1, & p \mid L_i(t), \\ 1/(p-1), & p \nmid L_i(t), \end{cases} \qquad \psi_{i,d}(t) = \prod_{p \mid d} \psi_{i,p}(t), \qquad \psi_{i,1}(t) = 1.$$

Each $\psi_{i,p}$ has mean zero, since p divides $L_i(t)$ in one residue class modulo p. For a fixed form these divisor functions are orthogonal, with squared mean $1/\varphi(d)$, where φ is Euler's totient function. If the form is uncovered, then $\psi_{i,d} = 1/\varphi(d)$ for every $d \mid P$. We therefore seek coefficients $a(d)$ satisfying

$$a(1) = 1, \qquad \sum_{\substack{d \mid P \\ d > 1}} \frac{a(d)}{\varphi(d)} = -1.$$

The resulting factor

$$F_i(t) = \sum_{d \mid P} a(d) \psi_{i,d}(t)$$

has mean 1 and vanishes on uncovered forms.

How should one choose $a(d)$? A divisor d contributes squared mass $a(d)^2/\varphi(d)$ and uses $\log d$ of a logarithmic modulus budget. We therefore minimize $\sum_{d \mid P,\, d > 1} a(d)^2 \log d/\varphi(d)$ subject to the cancellation above. The minimum is attained by

$$B = \sum_{\substack{d \mid P \\ d > 1}} \frac{1}{\varphi(d) \log d}, \qquad a(d) = -\frac{1}{B \log d} \quad (d > 1).$$

This explains the reciprocal logarithm. Put $A = \sum_{d \mid P} a(d)^2/\varphi(d) = \mathbb{E} F_i^2$. For the auxiliary range above, $B \asymp 1$ and $A = 1 + O((\log x)^{-2})$. Thus a single factor is usually close to 1, while vanishing exactly on the rare uncovered forms.

The product $\prod_i F_i$ vanishes whenever any form is uncovered, but its expansion involves moduli too large to average over $t \le T$. We therefore retain only terms whose divisor indices $r_1, \dots, r_k$ are pairwise coprime and satisfy $r_1 \cdots r_k \le D$. Call the resulting sum $R(t)$ and use $w(t) = R(t)^2$. Pairwise coprimality assigns each selected prime to at most one form; since all auxiliary primes exceed z, the loss from this restriction is small. The choice $D = e^{\kappa x}$ with $\kappa < 1/3$ makes the error in averaging the weight over the shorter interval small as well. Squaring a truncated divisor sum to obtain a nonnegative weight is also central to the enveloping sieve of Green and Tao [10, Appendices 7–8].

To understand the cutoff, ignore the small correlations between different forms and view the squared mass through independent divisor indices with distribution

$$\Pr(r_i = d) = \frac{a(d)^2}{A\varphi(d)} \qquad (d \mid P).$$

Each index has mean logarithmic size $1/(AB) \asymp 1$. The total expected logarithmic cost is therefore $O(k)$, while the available budget is $\log D = \kappa x$. Since $k \le \delta x$, choosing δ sufficiently small leaves ample room. Most indices equal 1: these expansion indices, rather than all the prime factors of a typical $L_i(t)$, determine the cost. A concentration estimate gives

$$\Pr(r_1 \cdots r_k > D) \le \exp\Big(-\frac{\log D}{2 \log y}\Big).$$

Our choice of y gives $\log D / \log y = \kappa (\log x)^C$, which is much larger than $\log k$. Thus the discarded squared mass is negligible even after summing the errors over all k forms.

This shared cutoff has a close analogue in the Maynard–Tao sieve. Maynard [12, §7] chooses a one-variable function g and restricts the product $\prod_i g(kt_i)$ to nonnegative variables $t_1, \dots, t_k$ with $\sum_i t_i \le 1$. Concentration under the density proportional to g^2 controls the mass lost at the cutoff. He also leaves enough room in the other coordinates for one coordinate's integral to run over its full support. Here $\log r_i / \log D$ plays the role of t_i, and the spare room keeps a one-coordinate divisor sum close to zero when that form is uncovered. The resulting errors are small in weighted average, making the weighted expected number of uncovered forms less than 1.

3. The Erdős–Rankin construction

The deduction is a straightforward modification of the Erdős–Rankin argument, as presented in [6, §2.2]. We use Proposition 1.2 in place of the final assignment of a distinct unused prime to each remaining position.

Proof of Theorem 1.1 assuming Proposition 1.2. For a sufficiently small absolute constant $\eta > 0$, put

$$H = \left\lfloor \eta \frac{x (\log x)^2 \log_3 x}{(\log_2 x)^2} \right\rfloor, \qquad w = (\log x)^4, \qquad z = H^{\log_3 x / (5 \log_2 x)}.$$

Then $2H/x < w < z < x/2$ and $\log H = (1 + o(1)) \log x$. Write $Q = Q(x)$. For each prime $p \le x$, we choose a residue $a_p \pmod p$; it covers the positions $n \in [1, H] \cap \mathbb{Z}$ with $n \equiv a_p \pmod p$. Choose $a_p = 0$ for $p \le w$ and for $z < p \le x/2$. A composite survivor in $[1, H]$ cannot have a prime factor greater than $x/2$, since the complementary factor would be between 2 and $2H/x < w$. Thus the surviving set S_0 consists of primes and z-smooth integers, whose prime factors are all at most z.

Let $\Psi(H, z)$ count all positive z-smooth integers at most H, including 1. Rankin's argument [14, Lemma II] gives $\Psi(H, z) \le H^\sigma \prod_{p \le z} (1 - p^{-\sigma})^{-1}$ for $0 < \sigma < 1$. Take $u = \log H / \log z = 5 \log_2 x / \log_3 x$ and $\sigma = 1 - \log u / \log z$. Mertens' estimates, together with $p^{1-\sigma} - 1 \le (u - 1) \log p / \log z$, give, for an absolute constant C,

$$\Psi(H, z) \ll H \log z \, \exp(-u \log u + Cu).$$

The right-hand side is $H(\log x)^{-4+o(1)}$, so $\Psi(H,z) \le H/(\log x)^3$ for large x. The sharper estimates of de Bruijn [2] are not needed here. By the prime number theorem, $|S_0| \le \pi(H) + \Psi(H,z) \ll H/\log x$.

Choose the classes a_p for $w < p \le z$ successively and greedily. At prime p this removes at least a $1/p$ fraction of the remaining positions. Let S be the set left after these choices. Mertens' product estimate gives

$$|S| \le |S_0| \prod_{w<p\le z} (1-1/p) \ll \frac{H}{\log x}\frac{\log w}{\log z} \ll \frac{H(\log_2 x)^2}{(\log x)^2 \log_3 x} \ll \eta x. \tag{3.1}$$

Choose arbitrary classes for $x/2 < p \le x$ and discard any further positions covered from S. For sufficiently small η, the remaining set satisfies $|S| \le \delta x$.

By the Chinese remainder theorem, choose $0 \le b < Q$ with $b \equiv -a_p \pmod p$ for every prime $p \le x$. For a covered position n, the integer $b + Qt + n$ is divisible by its covering prime for every $t \ge 1$, and is greater than that prime. Proposition 1.2 handles the remaining positions S. It supplies $1 \le t \le e^x$ such that $N+1, \dots, N+H$ are composite, where $N = b + Qt$. Since $\log Q(x) = (1+o(1))x$ by the prime number theorem, $N \le \exp((2+o(1))x)$. Taking $x = (\log X)/3$ gives

$$H = (1+o(1))\frac{\eta}{3}\frac{\log X (\log_2 X)^2 \log_4 X}{(\log_3 X)^2}.$$

The consecutive primes surrounding this interval both lie below X: the larger is at most $2(N+H) < X$ for large X, again by the prime number theorem. This proves the theorem. □

4. A truncated product that covers every form

4.1. **The weight.** Fix the data of Proposition 1.2, with $k \ge 1$, write $S = \{s_1, \dots, s_k\}$, and put $L_i(t) = b + Qt + s_i$ for $1 \le i \le k$ and $t \in \mathbb{Z}$. We first construct one divisor sum for each form, chosen to vanish when that form has no auxiliary prime factor. We then truncate their product to small moduli and square it.

Let P be a nonempty squarefree product of primes greater than H, and let $D > 1$ be a cutoff; their values will be chosen below. Recall the divisor functions introduced in Section 2. Writing $\mathbf{1}_E$ for the indicator of a condition E, we have, for $1 \le i \le k$, $p \mid P$ and $d \mid P$,

$$\psi_{i,p}(t) = \frac{1 - p\mathbf{1}_{\{p \mid L_i(t)\}}}{p-1}, \qquad \psi_{i,d}(t) = \prod_{p\mid d} \psi_{i,p}(t), \qquad \psi_{i,1}(t) = 1.$$

The factor $\psi_{i,p}$ is -1 when $p \mid L_i(t)$ and $1/(p-1)$ otherwise. Since $(p,Q) = 1$, the first case occurs in exactly one residue class modulo p, so this factor has mean zero. The roots for different forms are distinct, since $0 < |s_i - s_j| < p$. Recall also the coefficients

$$B = \sum_{\substack{d\mid P\\ d>1}} \frac{1}{\varphi(d)\log d}, \qquad a(1) = 1, \qquad a(d) = -\frac{1}{B\log d} \quad (d \mid P,\ d > 1). \tag{4.1}$$

By the Chinese remainder theorem, $\psi_{i,d}$ has mean zero modulo P for $d > 1$. Thus the sum $F_i(t) = \sum_{d\mid P} a(d)\psi_{i,d}(t)$ has mean 1. Call L_i uncovered at t if $(L_i(t), P) = 1$. In this case, every factor in $\psi_{i,d}$ equals $1/(p-1)$, and hence

$$F_i(t) = \sum_{d\mid P} \frac{a(d)}{\varphi(d)} = 1 - \frac{1}{B}\sum_{\substack{d\mid P\\ d>1}} \frac{1}{\varphi(d)\log d} = 0. \tag{4.2}$$

For $1 \le j \le k$, define the following set of ordered tuples:

$$\mathcal{R}_j = \{(r_1, \dots, r_j) : r_i \mid P,\ (r_i, r_\ell) = 1\ (i \ne \ell), r_1 \cdots r_j \le D\}.$$

Let $\mathcal{R}_0$ contain just the empty tuple; empty products are 1. All these regions use the same P and D. Our weight is

$$R(t) = \sum_{\mathbf{r}=(r_1,\dots,r_k)\in\mathcal{R}_k} \prod_{i=1}^{k} a(r_i)\psi_{i,r_i}(t), \qquad w(t) = R(t)^2. \tag{4.3}$$

Thus each selected prime is assigned to at most one form, and the product of all selected primes is at most D. The square ensures $w(t) \geq 0$. When a form is uncovered, its unrestricted coefficient sum is zero by (4.2). The restrictions in (4.3) leave only the excluded divisors; the estimates below bound their contribution.

Fix an absolute constant $0 < \kappa < 1/3$; we will choose δ sufficiently small in terms of κ. Take $P = \prod_{z<p\leq y} p$, where

$$T = \lfloor e^x \rfloor, \qquad z = x^6, \qquad y = \exp\left(\frac{x}{(\log x)^5}\right), \qquad D = e^{\kappa x}.$$

For large x, we have $H < z < y$, as required. Put

$$\beta = \frac{1}{\log y}, \qquad u = \frac{\log D}{\log y} = \kappa(\log x)^5.$$

We will compare the average weight with the average weighted number of uncovered forms:

$$S_1 = \frac{1}{T}\sum_{1\leq t\leq T} w(t), \qquad S_2 = \frac{1}{T}\sum_{1\leq t\leq T} w(t)\sum_{i=1}^{k} \mathbf{1}_{\{(L_i(t),P)=1\}}. \tag{4.4}$$

The inequality $S_2 < S_1$ supplies a translate where every form has an auxiliary prime factor. For $1 \leq t \leq T$, these factors are proper, because $L_i(t) \geq Q > y$ for large x.

4.2. **The coefficient estimates.** The following bounds control the terms discarded at D.

Lemma 4.1. *Put $A = \sum_{d|P} a(d)^2/\varphi(d)$ and $A_\gamma = \sum_{d|P} a(d)^2 d^\gamma/\varphi(d)$. For sufficiently large x, we have $B \asymp 1$, and there is an absolute constant C such that*

$$0 \leq A - 1 \ll (\log x)^{-2}, \qquad |a(d)| \leq 1, \qquad \sum_{d|P} \frac{a(d)}{\varphi(d)} = 0. \tag{4.5}$$

For $0 \leq \gamma \leq 2\beta$, we have

$$\sum_{\substack{d|P\\ d>1}} \frac{|a(d)|d^\gamma}{\varphi(d)} \leq C, \qquad A_\gamma \leq A + C\gamma. \tag{4.6}$$

Moreover, for each $p \mid P$,

$$\sum_{\substack{d|P\\ p|d}} \frac{|a(d)|}{\varphi(d)} \leq \frac{C}{p}. \tag{4.7}$$

Proof. Put $V = \prod_{p\leq z}(1 - 1/p) \asymp 1/\log z$ and $M_z(v) = \sum_{d\leq v,\,(d,Q(z))=1} \mu(d)^2/\varphi(d)$, where μ is the Möbius function. The standard sieve summation formula [7, Theorem 4.4] gives, uniformly for $v \geq 2$,

$$|M_z(v) - V\log v| \ll 1.$$

Indeed, apply that theorem with $g(p) = \mathbf{1}_{\{p>z\}}/p$. Mertens' estimates verify its hypotheses with dimension 1 and loss parameter $L \ll \log z$; its singular series is V.

For $0 \leq \gamma \leq 2\beta$, put $U_\gamma = \sum_{d|P} d^\gamma/\varphi(d)$. Mertens' estimates give $U_0 \asymp \log y/\log z$, and $p^\gamma - 1 \ll \gamma \log p$ gives

$$\frac{U_\gamma}{U_0} = \prod_{z<p\leq y}\left(1 + \frac{p^\gamma - 1}{p}\right) \leq \exp(C\gamma\log y) \ll 1.$$

The part of B with $d > y$ is at most $U_0/\log y \ll 1/\log z$. A squarefree $d \le y$ is coprime to $Q(z)$ exactly when $d \mid P$. Partial summation of M_z therefore gives

$$\left| B - V \log \frac{\log y}{\log z} \right| \ll \frac{1}{\log z}.$$

In particular, $B \asymp 1$. A second partial summation, extended to all squarefree integers coprime to $Q(z)$, gives

$$A - 1 \le \frac{2}{B^2} \int_z^\infty \frac{M_z(v) - 1}{v(\log v)^3}\, dv \ll \frac{1}{(\log z)^2}.$$

Also $|a(d)| \le 1$ for large x, and the cancellation follows from the definition of B.

Splitting at $d = y$ and using $d^\gamma \le e^2$ below y gives

$$\sum_{\substack{d \mid P \\ d > 1}} \frac{|a(d)| d^\gamma}{\varphi(d)} \le e^2 + \frac{U_\gamma}{B \log y} \ll 1.$$

The inequality $d^\gamma - 1 \le \gamma d^\gamma \log d$ now gives

$$A_\gamma - A \le \frac{\gamma}{B} \sum_{\substack{d \mid P \\ d > 1}} \frac{|a(d)| d^\gamma}{\varphi(d)} \ll \gamma.$$

Finally, for $p \mid P$, write $d = pe$ with $e \mid P/p$. Separate $e = 1$ and use $\log(pe) \ge \log e$ for $e > 1$ to obtain

$$\sum_{\substack{d \mid P \\ p \mid d}} \frac{|a(d)|}{\varphi(d)} \le \frac{1}{B(p-1)} \left(\frac{1}{\log p} + B \right) \ll \frac{1}{p}.$$

□

4.3. **The average weight.** Expanding the factors ψ_{i,r_i} gives real coefficients $\lambda_{\mathbf{d}}$ such that

$$R(t) = \sum_{\mathbf{d} = (d_1, \dots, d_k) \in \mathcal{R}_k} \lambda_{\mathbf{d}} \prod_{i=1}^{k} \mathbf{1}_{\{d_i \mid L_i(t)\}}.$$

Their total absolute value is at most

$$\sum_{\mathbf{d} \in \mathcal{R}_k} |\lambda_{\mathbf{d}}| \le \sum_{\substack{n \mid P \\ n \le D}} k^{\omega(n)} \frac{\sigma_1(n)}{\varphi(n)} \le 2D^{1 + \log k / \log z} \le D^{3/2}, \tag{4.8}$$

where $\sigma_1(n) = \sum_{d \mid n} d$ and $\omega(n)$ counts distinct prime factors. We used $|a(d)| \le 1$, $\omega(n) \le \log D / \log z$, and $\sigma_1(n)/\varphi(n) \le 2$ for these n.

Lemma 4.2. *For sufficiently small $\delta > 0$, $S_1 = (1 + o(1))A^k$, uniformly under the hypotheses of Proposition 1.2.*

Proof. For a P-periodic function f, write $\mathbb{E}f = P^{-1} \sum_{t=0}^{P-1} f(t)$. Each compatible system of divisibility conditions in the expansion of R^2 specifies one residue class modulo a divisor of P. Its average over $1 \le t \le T$ differs from its complete average by at most $1/T$. Hence (4.8) gives $|S_1 - \mathbb{E}R^2| \ll D^3/T$. The local factors $\psi_{i,p}$ have mean zero, square mean $1/(p-1)$, and pairwise product mean $-1/(p-1)^2$ for distinct coordinates. Writing Δ for the diagonal sum, independence between primes gives

$$\Delta = \sum_{\mathbf{r} \in \mathcal{R}_k} \prod_{i=1}^{k} \frac{a(r_i)^2}{\varphi(r_i)}, \qquad |\mathbb{E}R^2 - \Delta| \ll \frac{k \log D}{z}\, \Delta. \tag{4.9}$$

Indeed, a prime occurring in only one of two tuples makes their product average zero, so only tuples with the same product can pair. The variance-one factors $\sqrt{p-1}\,\psi_{i,p}$ have product mean $-1/(p-1)$ for distinct coordinates. For a fixed tuple, the sum of absolute off-diagonal pairing factors is at most $\prod_{p\mid r_1\cdots r_k}(1+(k-1)/(p-1))-1 \ll k\log D/z$. The inequality $2|ab|\le a^2+b^2$ proves (4.9). This argument applies to arbitrary real coefficients on $\mathcal{R}_j$ for any $0\le j\le k$.

The unrestricted diagonal sum is A^k. Rankin's bound and (4.6) show that the terms with $r_1\cdots r_k > D$ have total mass at most

$$D^{-\beta}A_\beta^k \le A^k\exp(-u+Ck\beta) \le A^k e^{-u/2}, \tag{4.10}$$

on choosing δ so that $Ck\le \frac12\log D$. The terms with a common prime in two coordinates have mass at most

$$\binom{k}{2}A^{k-2}\sum_{p\mid P}\left(\sum_{\substack{d\mid P\\ p\mid d}}\frac{a(d)^2}{\varphi(d)}\right)^2 \ll \frac{k^2}{z}A^k,$$

by $|a(d)|\le 1$ and (4.7). The discarded diagonal mass is $o(A^k)$, so $\Delta=(1+o(1))A^k$. Since $k\log D/z\to 0$ and $D^3=o(T)$, the lemma follows. □

4.4. **The uncovered forms.** When a form is uncovered, its full coefficient sum cancels. We use (4.6) again to bound the tail left by the cutoff.

Lemma 4.3. *For sufficiently small $\delta>0$, uniformly under the hypotheses of Proposition 1.2,*

$$S_2 \ll A^k k\left(e^{-u}+\frac{(\log D)^2}{z^2}\right)+\frac{kD^3}{T}.$$

Proof. It suffices to bound the contribution from L_k. For each divisor $m\mid P$ with $m\le D$, define

$$\varepsilon(m)=\sum_{\substack{d\mid P\\ d\le D/m\\ (d,m)=1}}\frac{a(d)}{\varphi(d)}.$$

Define the function $R_k:\mathbb{Z}\to\mathbb{R}$ by

$$R_k(t)=\sum_{\mathbf{r}=(r_1,\dots,r_{k-1})\in\mathcal{R}_{k-1}}\varepsilon(r_1\cdots r_{k-1})\prod_{j=1}^{k-1}a(r_j)\psi_{j,r_j}(t). \tag{4.11}$$

Here $\mathcal{R}_{k-1}$ consists of ordered, pairwise coprime $(k-1)$-tuples of divisors of P with product at most D. For a fixed such tuple with product m, the possible last coordinates $r_k=d$ in (4.3) satisfy precisely $d\mid P$, $(d,m)=1$, and $d\le D/m$. Thus R_k is obtained from R by replacing each factor $\psi_{k,d}$ by $1/\varphi(d)$ and summing over d. When $(L_k(t),P)=1$ this replacement changes no value, so

$$w(t)\mathbf{1}_{\{(L_k(t),P)=1\}}\le R_k(t)^2. \tag{4.12}$$

In the expansion with coefficients $\lambda_{\mathbf{d}}$, this replacement discards exactly the terms with $d_k>1$. Thus R_k has the same modulus bound D and coefficient bound $D^{3/2}$ as R. The Chinese remainder theorem gives

$$\left|\frac1T\sum_{1\le t\le T}R_k(t)^2-\mathbb{E}R_k^2\right|\ll\frac{D^3}{T}. \tag{4.13}$$

It remains to bound $\mathbb{E}R_k^2$ using (4.11). Since $\sum_{d|P} a(d)/\varphi(d) = 0$ and $a(d) < 0$ for $d > 1$, $\varepsilon(m)$ equals the sum of $|a(d)|/\varphi(d)$ over the excluded divisors. Hence

$$0 \le \varepsilon(m) \le \sum_{\substack{d|P \\ d>D/m}} \frac{|a(d)|}{\varphi(d)} + \sum_{p|m} \sum_{\substack{d|P \\ p|d}} \frac{|a(d)|}{\varphi(d)} \ll (m/D)^\beta + \frac{\log D}{z},$$

by (4.6) and (4.7).

The second-moment argument proving (4.9) applies to arbitrary real coefficients on $\mathcal{R}_{k-1}$. Applying it to R_k, using the bound for $\varepsilon(m)$, and then dropping the restrictions on the remaining coordinates, we obtain

$$\mathbb{E}R_k^2 \ll D^{-2\beta} A_{2\beta}^{k-1} + \frac{(\log D)^2}{z^2} A^{k-1} \ll A^k \left(e^{-u} + \frac{(\log D)^2}{z^2} \right).$$

The last inequality follows from (4.6), on choosing δ sufficiently small that $2Ck \le \log D$. Combine this with (4.12) and (4.13), and sum over the k forms. □

Proof of Proposition 1.2. The case $k = 0$ is immediate. Choose $\delta > 0$ sufficiently small for Lemmas 4.2 and 4.3. They give $S_1 > 0$ and

$$\frac{S_2}{S_1} \ll k e^{-u} + \frac{k(\log D)^2}{z^2} + \frac{kD^3}{T} \ll x e^{-\kappa(\log x)^5} + x^{-9} + e^{-(1-3\kappa+o(1))x}.$$

Since $\kappa < 1/3$, this tends to zero. If every $t \le T$ had an uncovered form, then $S_2 \ge S_1$. Hence some $1 \le t \le T$ has $(L_i(t), P) > 1$ for every i. Since every prime factor of P is smaller than every $L_i(t)$, all these values are composite. □

References

[1] T. F. Bloom, *Erdős Problem #4*. https://www.erdosproblems.com/4.

[2] N. G. de Bruijn, *On the number of positive integers $\le x$ and free of prime factors $> y$*, Nederl. Akad. Wetensch. Proc. Ser. A **54**, 50–60.

[3] P. Erdős, *On the difference of consecutive primes*, Quart. J. Math. Oxford Ser. **6**, 124–128.

[4] P. Erdős, *Some problems on the distribution of prime numbers*, in *Teoria dei numeri* (Varenna), C.I.M.E., 8 pp.

[5] P. Erdős, *Some of my favourite unsolved problems*, in *A Tribute to Paul Erdős* (A. Baker, B. Bollobás and A. Hajnal, eds.), Cambridge University Press, 467–478.

[6] K. Ford, *Anatomy of integers and random permutations*, lecture notes.

[7] K. Ford, *Sieve methods lecture notes*.

[8] K. Ford, B. Green, S. Konyagin and T. Tao, *Large gaps between consecutive prime numbers*, Ann. of Math. (2) **183**, 935–974.

[9] K. Ford, B. Green, S. Konyagin, J. Maynard and T. Tao, *Long gaps between primes*, J. Amer. Math. Soc. **31**, 65–105.

[10] B. Green and T. Tao, *Restriction theory of the Selberg sieve, with applications*, J. Théor. Nombres Bordeaux **18**, 147–182.

[11] J. Maynard, *Large gaps between primes*, Ann. of Math. (2) **183**, 915–933.

[12] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181**, 383–413.

[13] GPT 5.6 Sol, *A tilted residue-class construction for long prime-free intervals*, preprint. https://github.com/DottedCalculator/ai-math/blob/main/Erdos_4_GPT_5.6_Sol.pdf.

[14] R. A. Rankin, *The difference between consecutive prime numbers*, J. London Math. Soc. **13**, 242–247.

[15] E. Westzynthius, *Über die Verteilung der Zahlen, die zu den n ersten Primzahlen teilerfremd sind*, Comment. Phys.-Math. Soc. Sci. Fenn. **5**, no. 25, 1–37.