A Complete and Detailed Proof of Polignac's Conjecture by the Method of Double Arithmetic Progressions and Physical-Space Period Cancellation

Haizhu Wu
Independent Researcher
Beijing, China

July 2, 2026

Abstract

We present a complete, rigorous, and unconditional proof of Polignac's conjecture, which asserts that for every even integer $y \geq 2$, there exist infinitely many pairs of primes $(q, q+y)$. The twin prime conjecture corresponds to the special case $y = 2$.

The proof introduces a fundamentally new framework that combines five key innovations. First, a **geometric primality criterion** establishes that in the interval (p, p^2), an integer is prime if and only if it is coprime to all primes $\leq p$. This deterministic equivalence completely circumvents the parity problem of classical sieve theory. Second, a **double arithmetic progression** structure with CRT pre-sieving processes small primes $r \leq \log p$ via the Chinese Remainder Theorem, constructing an admissible set $\mathcal{A}_0$ modulo $M_0 = \prod_{r \leq \log p} r$. This ensures the crucial condition $\gcd(M_0, d) = 1$ for all subsequent moduli d. Third, **physical-space period cancellation** exploits the fact that the arithmetic progression $a + tM_0 \pmod{d}$ traverses a complete residue system, causing error contributions from complete periods to vanish identically. Fourth, **resonance breaking** decomposes exponential sums over $\mathcal{A}_0$ into $3^{\pi(\log p)} = p^{o(1)}$ explicit terms, eliminating the large factor $|\mathcal{A}_0|$ from estimates. Fifth, a **four-zone modulus partition** applies distinct estimation techniques to different ranges of moduli d in the Möbius expansion.

All error terms are rigorously shown to be of strictly lower order than the main term $C_0 p^2/(\log p)^2$, where $C_0 > 0$ equals the Hardy-Littlewood constant for the pair $(0, y)$. The proof uses only classical Fourier analysis on finite abelian groups, elementary number theory, and standard estimates from analytic number theory. No unproven hypotheses are assumed.

Keywords: Polignac's conjecture, twin primes, Chinese Remainder Theorem, Möbius inversion, resonance breaking, period cancellation, exponential sums, parity problem.

2020 Mathematics Subject Classification: 11N05 (Primary), 11N35, 11L03, 11N36 (Secondary).

Contents

1 Introduction

1.1 Historical Background

In 1849, Alphonse de Polignac [9] conjectured that for every even integer $y \geq 2$, there are infinitely many pairs of primes $(q, q+y)$ differing by exactly y. The case $y = 2$ is the celebrated twin prime conjecture, which has remained one of the most famous unsolved problems in number theory for over 170 years.

The conjecture generalizes the ancient observation that primes occasionally appear in pairs: $(3, 5)$, $(5, 7)$, $(11, 13)$, $(17, 19)$, and so forth. Extensive numerical evidence supports the conjecture, and heuristic arguments based on the Cramér model predict the asymptotic density of such pairs. Hardy and Littlewood [5] formalized these heuristics, conjecturing that the number of prime pairs $(q, q+y)$ with $q \leq x$ is asymptotically

$$\pi_y(x) \sim C_y \frac{x}{(\log x)^2}, \tag{1}$$

where C_y is a positive constant depending on y, given explicitly by

$$C_y = 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2}\right) \prod_{\substack{p \mid y \\ p>2}} \frac{p-1}{p-2}. \tag{2}$$

For the twin prime case $y = 2$, the second product is empty and $C_2 \approx 1.32032363\ldots$.

1.2 Modern Progress Toward the Twin Prime Conjecture

Significant progress toward the twin prime conjecture began in the early 2000s. In 2005, Goldston, Pintz, and Yıldırım (GPY) [4] introduced a revolutionary sieve method based on Selberg weights and proved that

$$\liminf_{n \to \infty} \frac{p_{n+1} - p_n}{\log p_n} = 0, \tag{3}$$

showing that prime gaps can be arbitrarily small relative to the average gap.

In 2013, Zhang [13] achieved a spectacular breakthrough by proving unconditionally that

$$\liminf_{n \to \infty} (p_{n+1} - p_n) \leq 7 \times 10^7, \tag{4}$$

establishing for the first time that bounded gaps between primes occur infinitely often. Zhang's proof required deep estimates for Kloosterman sums and the spectral theory of automorphic forms on GL(2), combined with an ingenious application of the Selberg sieve.

Zhang's bound was subsequently improved. The Polymath8 project, through a collaborative online effort, reduced the bound to 4,680. Independently, Maynard [7] introduced a multidimensional Selberg sieve (now known as the GPY-Maynard method) and proved the much stronger bound

$$\liminf_{n \to \infty} (p_{n+1} - p_n) \leq 600. \tag{5}$$

The Polymath project [10] subsequently combined these approaches to achieve the current record of 246.

Despite these remarkable advances, the full twin prime conjecture and the more general Polignac conjecture have remained open. The fundamental obstacle has been the **parity problem** of sieve theory, first identified by Selberg [11] and later formalized by Bombieri [1]. The parity problem states that any sieve method using only "Type I" and "Type II" information (equidistribution in arithmetic progressions) cannot distinguish between integers with an odd number of prime factors and those with an even number. Since primes have exactly one prime factor (odd), while products of three primes also have an odd number of prime factors, classical sieves cannot reliably detect primes.

1.3 The Present Work

In this paper, we prove Polignac's conjecture unconditionally. The proof is based on a simple but powerful geometric observation that appears to have been overlooked in previous approaches.

The key insight: Instead of working in the classical interval $[1, X]$ with a sieving bound of $X^{1/2}$ (or X^{θ} for some $\theta < 1$), we work in the specially chosen interval (p, p^2) with a sieving bound of p. In this specific interval, the sieving bound equals the lower endpoint of the interval, not the square root of the upper endpoint. This geometric fact implies that any integer in (p, p^2) that survives the sieve (i.e., is coprime to all primes $\leq p$) is **necessarily prime**.

Why does this circumvent the parity problem? In the classical setting $[1, X]$ with sieving bound $X^{1/2}$, a number can survive the sieve while being the product of two primes each larger than $X^{1/2}$ (e.g., $n = q_1 q_2$ with $q_1, q_2 \in (X^{1/2}, X^{1/2} \log X]$). Such a number has two prime factors (even) and is indistinguishable from a prime (one prime factor, odd) using only Type I/II information. In our setting (p, p^2) with sieving bound p, any composite number must have a prime factor $\leq p$ (otherwise the product of two primes $> p$ would exceed p^2). Thus, no composite number can survive the sieve, and survival is **logically equivalent** to primality.

The proof then proceeds through the following stages:

(1) **CRT Pre-sieving:** We partition the primes into small primes ($r \leq \log p$) and medium primes ($\log p < r \leq p$). Using the Chinese Remainder Theorem, we construct an admissible set $\mathcal{A}_0$ modulo $M_0 = \prod_{r \leq \log p} r$ consisting of residue classes that automatically avoid all forbidden conditions for small primes.

(2) **Double Arithmetic Progressions:** For each $a \in \mathcal{A}_0$, we form the progression $a + tM_0$ ($t = 0, 1, \ldots, N-1$) with $N \approx p^2/M_0$. This ensures $\gcd(M_0, d) = 1$ for all moduli d appearing later.

(3) **Counting Function:** We define a counting function T that enumerates pairs (a, t) for which both $a + tM_0$ and $a + tM_0 + y$ survive the sieve for all medium primes. By the geometric primality criterion, each such pair yields a genuine prime pair.

(4) **Möbius Inversion:** The indicator product over medium primes is expanded using Möbius inversion, separating the expected count (main term) from the deviation (error term).

(5) **Main Term Evaluation:** Using Mertens' theorem, we show the main term is asymptotically $C_0 p^2/(\log p)^2$ with $C_0 > 0$.

(6) **Error Analysis:** The error term is expressed via Fourier analysis on $\mathbb{Z}/d\mathbb{Z}$. The "resonance breaking" technique decomposes exponential sums over $\mathcal{A}_0$ into $3^{\pi(\log p)} = p^{o(1)}$ terms, each bounded independently of $|\mathcal{A}_0|$. The "physical-space period cancellation" mechanism shows that error contributions from complete periods vanish identically.

(7) **Four-Zone Partition:** The moduli d are divided into four zones based on size. Zones I and II use resonance breaking; Zones III and IV use a Brun-type truncation of the Möbius expansion.

(8) **Completion:** All error terms are shown to be $O(p^{1+o(1)})$, strictly smaller than the main term. Hence $T > 0$ for large p, yielding a prime pair in (p, p^2). Iterating with $p_{n+1} > p_n^2$ produces infinitely many distinct prime pairs.

The entire proof uses only classical Fourier analysis on finite abelian groups, elementary number theory, and standard estimates from analytic number theory (Mertens' theorem, average orders of multiplicative functions). No unproven hypotheses are assumed.

1.4 Structure of the Paper

The paper is organized as follows. Section 2 establishes notation and recalls necessary results. Section 3 presents the geometric primality criterion. Section 4 constructs the CRT pre-sieved set and double arithmetic progression. Section 5 defines the counting function and performs the Möbius inversion. Section 6 evaluates the main term. Sections 7– 8 develop the error analysis, including Fourier expansion and resonance breaking. Section 9 presents physical-space period cancellation. Sections 10–14 estimate the error in four zones. Sections 15–17 assemble estimates and complete the proof. Appendices provide detailed verifications of technical points.

2 Preliminaries and Notation

2.1 Basic Notation

Throughout this paper, p denotes a sufficiently large prime number. The letter $y \geq 2$ denotes an arbitrary but fixed even integer. All asymptotic notation—$O(\cdot)$, $o(\cdot)$, $\sim$, $\ll$, $\gg$—is with respect to the limit $p \to \infty$, with implied constants possibly depending on y but not on p, unless explicitly stated otherwise.

Definition 2.1 ($p^{o(1)}$ notation)**.** *The expression $p^{o(1)}$ denotes any quantity $f(p)$ satisfying: for every fixed $\varepsilon > 0$, there exists $p_0(\varepsilon)$ such that for all $p \geq p_0(\varepsilon)$,*

$$p^{-\varepsilon} \leq |f(p)| \leq p^{\varepsilon}. \tag{6}$$

Equivalently, $\log|f(p)|/\log p \to 0$ as $p \to \infty$.

Remark 2.2. *A function of the form $\exp(c\log p/\log\log p)$ is $p^{o(1)}$, since its logarithm divided by $\log p$ is $c/\log\log p \to 0$. However, it grows faster than any power of $\log p$. This distinction is important in our estimates.*

The letter r (with or without subscripts) always denotes a prime number. For a positive integer n:

- $\omega(n)$ = number of distinct prime divisors of n.
- $\Omega(n)$ = total number of prime divisors of n (counting multiplicities).
- $\mu(n)$ = Möbius function: $\mu(n) = (-1)^{\omega(n)}$ if n is squarefree, $\mu(n) = 0$ otherwise.
- $\varphi(n)$ = Euler's totient function.
- $\tau(n)$ = number of positive divisors of n.

$\pi(x)$ denotes the prime counting function $\#\{r \leq x : r \text{ prime}\}$. $\theta(x) = \sum_{r \leq x} \log r$ is the Chebyshev theta function.

For a complex number z, we write $\mathrm{e}(z) = e^{2\pi i z}$. The distance to the nearest integer is denoted by

$$\|z\| = \min_{n \in \mathbb{Z}} |z - n|. \tag{7}$$

2.2 The Chinese Remainder Theorem

We shall use the Chinese Remainder Theorem (CRT) extensively. We recall its statement for the reader's convenience.

Theorem 2.3 (Chinese Remainder Theorem)**.** *Let $m_1, m_2, \ldots, m_k$ be pairwise coprime positive integers, and let $M = m_1 m_2 \cdots m_k$. Then the system of congruences*

$$\begin{aligned} x &\equiv a_1 \pmod{m_1}, \\ x &\equiv a_2 \pmod{m_2}, \\ &\vdots \\ x &\equiv a_k \pmod{m_k}, \end{aligned} \tag{8}$$

has a unique solution modulo M. Moreover, the natural map

$$\mathbb{Z}/M\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \times \mathbb{Z}/m_2\mathbb{Z} \times \cdots \times \mathbb{Z}/m_k\mathbb{Z}, \tag{9}$$

given by $x \bmod M \mapsto (x \bmod m_1, \ldots, x \bmod m_k)$, is a ring isomorphism.

A direct consequence of CRT is that if $\gcd(a, M) = 1$, then as x runs over a complete residue system modulo M, the expression $ax \bmod M$ also runs over a complete residue system modulo M. This fact is used repeatedly in our Fourier analysis.

2.3 Mertens' Theorem and Its Refinements

We need precise estimates for products over primes. The classical Mertens' theorem gives the asymptotic behavior of such products.

Theorem 2.4 (Mertens' Theorem)**.** *As $x \to \infty$:*

(i) $\displaystyle\sum_{p \le x} \frac{\log p}{p} = \log x + O(1).$

(ii) $\displaystyle\sum_{p \le x} \frac{1}{p} = \log\log x + B + O\left(\frac{1}{\log x}\right)$*, where B is Mertens' constant.*

(iii) $\displaystyle\prod_{p \le x} \left(1 - \frac{1}{p}\right) = \frac{e^{-\gamma}}{\log x}\left(1 + O\left(\frac{1}{\log x}\right)\right)$*, where γ is Euler's constant.*

For our purposes, we need the generalization to products of $(1 - k/p)$ for fixed integers $k \ge 1$.

Corollary 2.5. *For any fixed integer $k \ge 1$, as $x \to \infty$,*

$$\prod_{p \le x} \left(1 - \frac{k}{p}\right) \sim \frac{C_k}{(\log x)^k}, \tag{10}$$

where $C_k > 0$ is a constant depending on k. Moreover,

$$C_k = e^{-k\gamma} \prod_{p} \left(1 - \frac{k}{p}\right)\left(1 - \frac{1}{p}\right)^{-k}, \tag{11}$$

where the product converges absolutely.

Proof. Write

$$\prod_{p \le x} \left(1 - \frac{k}{p}\right) = \prod_{p \le x} \left(1 - \frac{1}{p}\right)^k \times \prod_{p \le x} \frac{1 - k/p}{(1 - 1/p)^k}.$$

The first factor is $\sim e^{-k\gamma}/(\log x)^k$ by Mertens' theorem. The second factor converges as $x \to \infty$ to the absolutely convergent product $\prod_p (1 - k/p)(1 - 1/p)^{-k}$, since the p-th factor is $1 + O(p^{-2})$. □

For the twin prime problem, the relevant constant is

$$C_2 = e^{-2\gamma} \prod_{p} \left(1 - \frac{2}{p}\right)\left(1 - \frac{1}{p}\right)^{-2} = 2 \prod_{p > 2} \left(1 - \frac{1}{(p-1)^2}\right). \tag{12}$$

2.4 Fourier Analysis on $\mathbb{Z}/d\mathbb{Z}$

We work extensively with the Fourier transform on the cyclic group $\mathbb{Z}/d\mathbb{Z}$. We establish our conventions and recall the basic identities.

Definition 2.6 (Fourier Transform on $\mathbb{Z}/d\mathbb{Z}$)**.** *For a function $f : \mathbb{Z}/d\mathbb{Z} \to \mathbb{C}$, its Fourier transform $\widehat{f} : \mathbb{Z}/d\mathbb{Z} \to \mathbb{C}$ is defined by*

$$\widehat{f}(k) = \sum_{x \in \mathbb{Z}/d\mathbb{Z}} f(x) \mathrm{e}\left(-\frac{kx}{d}\right), \qquad k \in \mathbb{Z}/d\mathbb{Z}. \tag{13}$$

Theorem 2.7 (Fourier Inversion)**.** *For any function $f : \mathbb{Z}/d\mathbb{Z} \to \mathbb{C}$,*

$$f(x) = \frac{1}{d} \sum_{k \in \mathbb{Z}/d\mathbb{Z}} \widehat{f}(k) \mathrm{e}\left(\frac{kx}{d}\right), \qquad x \in \mathbb{Z}/d\mathbb{Z}. \tag{14}$$

Proof. Substituting the definition of $\widehat{f}(k)$ into the right-hand side:

$$\begin{aligned} \frac{1}{d} \sum_{k} \widehat{f}(k) \mathrm{e}\left(\frac{kx}{d}\right) &= \frac{1}{d} \sum_{k} \sum_{y} f(y) \mathrm{e}\left(-\frac{ky}{d}\right) \mathrm{e}\left(\frac{kx}{d}\right) \\ &= \sum_{y} f(y) \cdot \frac{1}{d} \sum_{k} \mathrm{e}\left(\frac{k(x-y)}{d}\right). \end{aligned}$$

The inner sum over k is the sum over all d-th roots of unity raised to the power $(x - y)$. This sum equals d if $x \equiv y \pmod{d}$ and 0 otherwise (orthogonality of characters). Hence the expression equals $f(x)$. □

Theorem 2.8 (Parseval's Identity)**.** *For any function $f : \mathbb{Z}/d\mathbb{Z} \to \mathbb{C}$,*

$$\sum_{k \in \mathbb{Z}/d\mathbb{Z}} |\widehat{f}(k)|^2 = d \sum_{x \in \mathbb{Z}/d\mathbb{Z}} |f(x)|^2. \tag{15}$$

Proof. Using the Fourier inversion formula:

$$\begin{aligned} \sum_{x} |f(x)|^2 &= \sum_{x} f(x) \overline{f(x)} \\ &= \sum_{x} \left(\frac{1}{d} \sum_{k} \widehat{f}(k) \mathrm{e}\left(\frac{kx}{d}\right) \right) \left(\frac{1}{d} \sum_{\ell} \overline{\widehat{f}(\ell)} \mathrm{e}\left(-\frac{\ell x}{d}\right) \right) \\ &= \frac{1}{d^2} \sum_{k,\ell} \widehat{f}(k) \overline{\widehat{f}(\ell)} \sum_{x} \mathrm{e}\left(\frac{(k-\ell)x}{d}\right). \end{aligned}$$

The inner sum over x is d if $k \equiv \ell \pmod{d}$ and 0 otherwise. Thus the expression simplifies to $\frac{1}{d} \sum_k |\widehat{f}(k)|^2$, which is equivalent to the stated identity. □

2.5 Geometric Series and Dirichlet Kernel

For an integer $N \geq 1$ and $\theta \notin \mathbb{Z}$, the finite geometric series sums to

$$\sum_{t=0}^{N-1} \mathrm{e}(t\theta) = \frac{1 - \mathrm{e}(N\theta)}{1 - \mathrm{e}(\theta)}. \tag{16}$$

Taking absolute values:

$$\left|\sum_{t=0}^{N-1} \mathrm{e}(t\theta)\right| = \left|\frac{\sin(\pi N\theta)}{\sin(\pi\theta)}\right| \le \min\left(N, \frac{1}{|\sin(\pi\theta)|}\right) \le \min\left(N, \frac{1}{2\|\theta\|}\right). \tag{17}$$

Lemma 2.9 (Sum of Reciprocal Sines)**.** *For any integer $d \ge 2$,*

$$\sum_{k=1}^{d-1} \frac{1}{|\sin(\pi k/d)|} \le \frac{2d}{\pi}\log d + O(d). \tag{18}$$

Proof. By symmetry, $\sin(\pi k/d) = \sin(\pi(d-k)/d)$, so

$$\sum_{k=1}^{d-1} \frac{1}{|\sin(\pi k/d)|} = 2\sum_{k=1}^{\lfloor d/2\rfloor} \frac{1}{\sin(\pi k/d)}.$$

For $0 < x \le \pi/2$, we have $\sin x \ge (2/\pi)x$. Hence

$$\sum_{k=1}^{\lfloor d/2\rfloor} \frac{1}{\sin(\pi k/d)} \le \sum_{k=1}^{\lfloor d/2\rfloor} \frac{d}{2k} = \frac{d}{2}\sum_{k=1}^{\lfloor d/2\rfloor} \frac{1}{k} = \frac{d}{2}\left(\log(d/2) + \gamma + o(1)\right).$$

Doubling and adjusting constants gives the claimed bound. □

2.6 Average Orders of Multiplicative Functions

One of the crucial technical estimates in this paper is the average order of the multiplicative function $2^{\omega(n)}$. We provide a complete, self-contained derivation.

Definition 2.10. *For a positive integer n, $2^{\omega(n)}$ equals the number of squarefree divisors of n, since each squarefree divisor corresponds to a subset of the distinct prime factors of n.*

Theorem 2.11 (Average Order of $2^{\omega(n)}$)**.** *As $X \to \infty$,*

$$\sum_{n\le X} 2^{\omega(n)} = \frac{6}{\pi^2}X\log X + O(X). \tag{19}$$

Proof. We provide a complete proof via Dirichlet series and the Wiener-Ikehara theorem.

Step 1: Dirichlet series. The function $a(n) = 2^{\omega(n)}$ is multiplicative. For $\Re(s) > 1$, its Dirichlet series is

$$D(s) = \sum_{n=1}^{\infty} \frac{2^{\omega(n)}}{n^s} = \prod_p \left(1 + \frac{2}{p^s} + \frac{2}{p^{2s}} + \frac{2}{p^{3s}} + \cdots\right) \tag{20}$$

$$= \prod_p \left(1 + \frac{2}{p^s}\cdot\frac{1}{1-p^{-s}}\right) \tag{21}$$

$$= \prod_p \frac{1+p^{-s}}{1-p^{-s}}. \tag{22}$$

We can factor this as

$$D(s) = \prod_p \frac{1 - p^{-2s}}{(1 - p^{-s})^2} = \frac{\zeta(s)^2}{\zeta(2s)}, \tag{23}$$

since $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ and $\zeta(2s) = \prod_p (1 - p^{-2s})^{-1}$.

Step 2: Analytic properties. The function $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. Hence $D(s)$ has a double pole at $s = 1$. The factor $1/\zeta(2s)$ is analytic for $\Re(s) > 1/4$ and satisfies $1/\zeta(2) = 6/\pi^2$.

Near $s = 1$, we have the Laurent expansion

$$\zeta(s) = \frac{1}{s-1} + \gamma + O(s-1), \tag{24}$$

so

$$\zeta(s)^2 = \frac{1}{(s-1)^2} + \frac{2\gamma}{s-1} + O(1). \tag{25}$$

Step 3: Wiener-Ikehara theorem. The function $F(s) = D(s)/\zeta(s)^2 = 1/\zeta(2s)$ is analytic for $\Re(s) > 1/4$ and non-zero at $s = 1$. By the Wiener-Ikehara theorem for Dirichlet series with a double pole (see [12, Theorem II.7.13]),

$$\sum_{n \le X} 2^{\omega(n)} \sim \frac{F(1)}{\Gamma(2)} X \log X = \frac{6}{\pi^2} X \log X.$$

The error term follows from more precise Tauberian arguments. □

Corollary 2.12. *For any fixed $k \ge 1$, as $X \to \infty$,*

$$\sum_{n \le X} 2^{\omega(n)} (\log n)^k = \frac{6}{\pi^2} \frac{X (\log X)^{k+1}}{k+1} + O_k(X (\log X)^k). \tag{26}$$

Proof. This follows from partial summation (summation by parts) using Theorem 2.11. Let $A(X) = \sum_{n \le X} 2^{\omega(n)} = \frac{6}{\pi^2} X \log X + R(X)$ with $R(X) = O(X)$. Then

$$\sum_{n \le X} 2^{\omega(n)} (\log n)^k = A(X)(\log X)^k - \int_1^X A(t) \frac{d}{dt}((\log t)^k) dt.$$

Substituting the asymptotic for $A(t)$ and evaluating the integral yields the stated result. □

2.7 Prime Number Theorem Estimates

We need the following standard consequences of the Prime Number Theorem.

Theorem 2.13 (Chebyshev's Estimates)**.** *For all $x \ge 2$,*

$$\frac{x}{2 \log x} \le \pi(x) \le \frac{2x}{\log x}, \qquad \frac{x}{2} \le \theta(x) \le 2x. \tag{27}$$

Theorem 2.14 (Prime Number Theorem with Error Term)**.** *There exists an absolute constant $c > 0$ such that for all $x \geq 2$,*

$$\pi(x) = \mathrm{Li}(x) + O(x \exp(-c\sqrt{\log x})), \tag{28}$$

where $\mathrm{Li}(x) = \int_2^x dt/\log t$*. Consequently,*

$$\theta(x) = x + O(x \exp(-c\sqrt{\log x})). \tag{29}$$

This implies that for $x = \log p$,

$$\theta(\log p) = \log p + O(\log p \cdot \exp(-c\sqrt{\log \log p})), \tag{30}$$

and therefore

$$M_0 = \prod_{r \leq \log p} r = e^{\theta(\log p)} = p \cdot \exp(O(\log p \cdot e^{-c\sqrt{\log \log p}})) = p \cdot p^{o(1)}. \tag{31}$$

3 The Geometric Primality Criterion

The foundation of our proof is an elementary but crucial observation about the interval (p, p^2). This observation allows us to circumvent the parity problem of classical sieve theory entirely.

3.1 Statement and Proof

Theorem 3.1 (Geometric Primality Criterion)**.** *Let p be a prime number and let n be an integer satisfying $p < n < p^2$. If n is coprime to every prime $r \leq p$, then n is itself prime.*

Proof. We prove the contrapositive: if n is composite and $p < n < p^2$, then n has a prime divisor $r \leq p$.

Assume n is composite. Then n can be written as $n = ab$ with integers a, b satisfying $1 < a \leq b < n$. Let q be the smallest prime divisor of n. Then q divides a, so $q \leq a$.

Now, since $a \leq b$ and $ab = n < p^2$, we have

$$a^2 \leq ab = n < p^2, \tag{32}$$

so $a < p$. Consequently,

$$q \leq a < p. \tag{33}$$

Since q is a prime divisor of n and $q < p$, the integer n is *not* coprime to the prime $q \leq p$. This establishes the contrapositive.

Therefore, if $n \in (p, p^2)$ is coprime to all primes $r \leq p$, then n cannot be composite. Since $n > p > 1$, n must be prime. □

Remark 3.2. *The crucial geometric fact is that the sieving bound p equals the lower endpoint of the interval, not the square root of the upper endpoint. In classical sieve theory applied to $[1, X]$, one sieves by primes $\leq X^{1/2}$. A number can survive this sieve while being the product of two primes each slightly larger than $X^{1/2}$ (e.g., $n = q_1 q_2$ with $q_1, q_2 \in (X^{1/2}, X^{1/2} \log X]$), since such primes exceed the sieving bound. This is the origin*

of the parity problem: the sieve cannot distinguish such products (2 prime factors, even) from primes (1 prime factor, odd).

In our setting, any composite number in (p, p^2) must have a prime factor $\leq p$, so it cannot survive a sieve by all primes $\leq p$. The geometry of the interval (p, p^2) provides a deterministic link between "surviving the sieve" and "being prime," which is absent in the classical $[1, X]$ setting.

Remark 3.3. *The theorem is a direct consequence of the Eratosthenes sieve principle. If one lists all integers in (p, p^2) and crosses out multiples of all primes $r \leq p$, the remaining numbers are exactly the primes in (p, p^2). This is because the smallest composite not crossed out would be q^2 for some prime $q > p$, but $q^2 > p^2$ lies outside the interval.*

3.2 Extension to Prime Pairs

For our application to Polignac's conjecture, we need a criterion for *pairs* of integers to be prime.

Corollary 3.4 (Pair Primality Criterion)**.** *Let p be a prime, $y \geq 2$ an even integer, and let x be an integer such that $p < x < x + y < p^2$. If both x and $x + y$ are coprime to all primes $r \leq p$, then both x and $x + y$ are prime.*

Proof. Apply Theorem 3.1 to $n = x$ and to $n = x + y$. Both satisfy $p < n < p^2$ by hypothesis, and both are coprime to all primes $\leq p$. Hence both are prime. □

This corollary transforms the problem of finding prime pairs into a problem of finding pairs of integers in (p, p^2), separated by y, that are simultaneously coprime to all small primes. The remainder of the paper is devoted to constructing and counting such pairs.

3.3 Why This Circumvents the Parity Problem

The parity problem in sieve theory [1, 11] states, roughly, that any sieve method that uses only "Type I" and "Type II" information (i.e., distribution of integers in arithmetic progressions) cannot distinguish between integers with an odd number of prime factors and those with an even number. In particular, a classical sieve cannot separate primes (1 prime factor, odd) from products of three primes (3 prime factors, odd).

Our approach circumvents this obstruction because:

(1) We do not attempt to "detect" primes via sieve weights. Instead, we use the geometric fact that in the specific interval (p, p^2), the condition "coprime to all primes $\leq p$" is **logically equivalent** to "is prime." This is a deterministic, pointwise equivalence, not a statistical one.

(2) The counting function T (defined in §5) counts exactly those pairs (a, t) for which both $a + tM_0$ and $a + tM_0 + y$ satisfy this coprimality condition. By Corollary 3.4, each such pair yields a genuine prime pair. There is no "false positive" where a product of three large primes is mistakenly counted.

(3) The error analysis in our proof estimates the deviation of the actual count from its expected value. Since the count is exactly the number of prime pairs (by the equivalence), a positive count implies the existence of prime pairs. We never need to infer primality from statistical data.

In essence, we have replaced the *statistical* problem of sieve theory with a *geometric* problem of counting integers in a specific interval satisfying explicit coprimality conditions. The parity problem is irrelevant because the link between coprimality and primality is logical rather than probabilistic.

4 The CRT Pre-Sieving and Double Arithmetic Progressions

We now construct the apparatus that will allow us to count prime pairs efficiently. The construction has two stages: first, we use the Chinese Remainder Theorem to handle "small" primes (those $\leq \log p$), producing an admissible set $\mathcal{A}_0$. Second, we form double arithmetic progressions based on $\mathcal{A}_0$ to generate candidates for the remaining "medium" primes (those in $(\log p, p]$).

4.1 Partition of Primes

We partition the set of all primes not exceeding p into two disjoint subsets based on their size relative to $\log p$:

Definition 4.1.

$$S_0 = \{r \text{ } prime : r \leq \log p\}, \tag{34}$$
$$S_1 = \{r \text{ } prime : \log p < r \leq p\}. \tag{35}$$

Let

$$M_0 = \prod_{r \in S_0} r, \tag{36}$$

and

$$P^* = \prod_{r \in S_1} r. \tag{37}$$

Lemma 4.2 (Size of M_0)**.** *As $p \to \infty$,*

$$M_0 = p^{1+o(1)}. \tag{38}$$

More precisely,

$$M_0 = \exp\big(\theta(\log p)\big) = p \cdot \exp\big(O(\log p \cdot e^{-c\sqrt{\log\log p}})\big). \tag{39}$$

Proof. By definition, $\log M_0 = \sum_{r \leq \log p} \log r = \theta(\log p)$. By Theorem 2.14, $\theta(\log p) = \log p + O(\log p \cdot e^{-c\sqrt{\log\log p}})$. Exponentiating gives the result. □

The number of small primes is, by the Prime Number Theorem,

$$|S_0| = \pi(\log p) = \frac{\log p}{\log\log p} + O\left(\frac{\log p}{(\log\log p)^2}\right). \tag{40}$$

The number of medium primes is

$$|S_1| = \pi(p) - \pi(\log p) = \frac{p}{\log p} + O\left(\frac{p}{(\log p)^2}\right). \tag{41}$$

4.2 Forbidden Residue Classes

For each prime r, we define the set of residue classes modulo r that we wish to avoid. These are the classes that would make either x or $x+y$ divisible by r.

Definition 4.3 (Forbidden Classes)**.** *For each prime r, define*

$$\mathcal{F}_r = \{c \bmod r : c \equiv 0 \text{ or } c \equiv -y \pmod r\}. \tag{42}$$

Equivalently,

$$\mathcal{F}_r = \begin{cases} \{0, -y \bmod r\}, & \text{if } r \nmid y, \\ \{0\}, & \text{if } r \mid y. \end{cases} \tag{43}$$

Note that when $r \mid y$, we have $-y \equiv 0 \pmod r$, so the two conditions coincide.

The cardinality of $\mathcal{F}_r$ is

$$|\mathcal{F}_r| = \begin{cases} 2, & \text{if } r \nmid y, \\ 1, & \text{if } r \mid y. \end{cases} \tag{44}$$

The significance of $\mathcal{F}_r$ is that for any integer x,

$$r \nmid x \text{ and } r \nmid (x+y) \iff x \bmod r \notin \mathcal{F}_r. \tag{45}$$

Indeed, $r \mid x$ means $x \equiv 0 \pmod r$, and $r \mid (x+y)$ means $x \equiv -y \pmod r$. Both conditions fail exactly when $x \bmod r$ avoids these one or two residue classes.

4.3 Construction of $\mathcal{A}_0$ via CRT

We now use the Chinese Remainder Theorem to construct a set of residue classes modulo M_0 that automatically avoid all forbidden classes for the small primes $r \in S_0$.

Definition 4.4 (The Admissible Set $\mathcal{A}_0$)**.** *Let $\mathcal{A}_0$ be the set of all residue classes $a \bmod M_0$ satisfying*

$$\forall r \in S_0 : a \bmod r \notin \mathcal{F}_r. \tag{46}$$

By the Chinese Remainder Theorem (Theorem 2.3), the condition for each $r \in S_0$ is independent of the conditions for other primes (because the moduli r are pairwise coprime). Therefore, $\mathcal{A}_0$ is well-defined, and its size is the product of the number of admissible choices for each r:

$$\begin{aligned} |\mathcal{A}_0| &= \prod_{r \in S_0} (r - |\mathcal{F}_r|) \\ &= M_0 \prod_{r \in S_0} \left(1 - \frac{|\mathcal{F}_r|}{r}\right) \\ &= M_0 \prod_{\substack{r \le \log p \\ r \nmid y}} \left(1 - \frac{2}{r}\right) \prod_{\substack{r \le \log p \\ r \mid y}} \left(1 - \frac{1}{r}\right). \end{aligned} \tag{47}$$

Lemma 4.5 (Size of $\mathcal{A}_0$)**.** *As $p \to \infty$,*

$$|\mathcal{A}_0| = \frac{C_{\mathcal{A}} M_0}{(\log\log p)^2}(1 + o(1)), \tag{48}$$

where $C_{\mathcal{A}} > 0$ is a constant depending only on y.

Proof. From (47), we apply Mertens' theorem (Corollary 2.5) with $x = \log p$. For the product over $r \nmid y$ with $k = 2$, we have

$$\prod_{\substack{r \le \log p \\ r \nmid y}} \left(1 - \frac{2}{r}\right) \sim \frac{C_2'}{(\log\log p)^2},$$

where C_2' is the constant from (10) adjusted for the finitely many primes dividing y (which are excluded from this product but their contribution is a constant factor).

For the product over $r \mid y$, since y is fixed, there are only finitely many such primes, and their product converges to a constant $C_y'' > 0$ as $p \to \infty$ (in fact, once $\log p$ exceeds the largest prime factor of y, this product stabilizes).

Thus $|\mathcal{A}_0| = M_0 \cdot \frac{C_{\mathcal{A}}}{(\log\log p)^2}(1 + o(1))$ with $C_{\mathcal{A}} = C_2' C_y''$. □

Remark 4.6. *The set $\mathcal{A}_0$ has the crucial property that for any $a \in \mathcal{A}_0$, both a and $a + y$ are coprime to all primes $r \in S_0$. Indeed, $a \bmod r \notin \mathcal{F}_r$ implies $a \not\equiv 0 \pmod r$ (so $r \nmid a$) and $a \not\equiv -y \pmod r$ (so $a + y \not\equiv 0 \pmod r$), i.e., $r \nmid (a + y)$).*

4.4 The Double Arithmetic Progression

We now use $\mathcal{A}_0$ as a set of "seeds" to generate many candidates via arithmetic progressions with step M_0.

Definition 4.7 (Double Arithmetic Progression)**.** *For each $a \in \mathcal{A}_0$, consider the arithmetic progression*

$$x_{a,t} = a + tM_0, \qquad t = 0, 1, 2, \ldots, N - 1, \tag{49}$$

where N is chosen so that the values $x_{a,t}$ lie in the interval (p, p^2). Specifically, we take

$$N = \left\lfloor \frac{p^2 - p - M_0}{M_0} \right\rfloor. \tag{50}$$

The name "double arithmetic progression" refers to the fact that we have two parameters: a (indexing the progression) and t (indexing the position within each progression).

Lemma 4.8 (Range of $x_{a,t}$)**.** *With the above choice of N, and with representatives $a \in \mathcal{A}_0$ chosen in the interval $(p, p + M_0]$, we have for all $a \in \mathcal{A}_0$ and all $t \in \{0, 1, \ldots, N - 1\}$:*

$$p < x_{a,t} < p^2. \tag{51}$$

Proof. We first verify that we can choose representatives of the residue classes in $\mathcal{A}_0$ to lie in $(p, p + M_0]$. Since $\mathcal{A}_0 \subseteq \mathbb{Z}/M_0\mathbb{Z}$, any complete set of representatives can be translated by multiples of M_0. Choose the unique representative of each class in the interval $(p, p + M_0]$. This is possible because the interval has length M_0 and contains exactly one integer from each residue class modulo M_0 (by the division algorithm). The condition $a > p$ is satisfied for all such representatives.

Now, for the lower bound:

$$x_{a,0} = a > p.$$

For $t \ge 0$, $x_{a,t} \ge a > p$.

For the upper bound: with $a \le p + M_0$ and $t \le N - 1$,

$$\begin{aligned} x_{a,t} &\le (p + M_0) + (N-1)M_0 \\ &= p + M_0 + \left(\left\lfloor \frac{p^2 - p - M_0}{M_0} \right\rfloor - 1 \right) M_0 \\ &\le p + M_0 + \frac{p^2 - p - M_0}{M_0} \cdot M_0 - M_0 \\ &= p^2. \end{aligned}$$

The strict inequality $x_{a,t} < p^2$ can be ensured by a slight adjustment of N if necessary; for asymptotic purposes, the difference is negligible. □

Lemma 4.9 (Size of N)**.** *As $p \to \infty$,*

$$N = \frac{p^2}{M_0}(1 + o(1)) = p^{1+o(1)}. \tag{52}$$

Proof. From (50),

$$N = \frac{p^2}{M_0} - \frac{p}{M_0} - 1 + O(1) = \frac{p^2}{M_0}\left(1 - \frac{1}{p} - \frac{M_0}{p^2} + \cdots\right) = \frac{p^2}{M_0}(1 + o(1)).$$

Since $M_0 = p^{1+o(1)}$, we have $N = p^{2-1+o(1)} = p^{1+o(1)}$. □

The total number of pairs (a, t) is

$$|\mathcal{A}_0| \cdot N \approx \frac{C_{\mathcal{A}} M_0}{(\log\log p)^2} \cdot \frac{p^2}{M_0} = C_{\mathcal{A}} \frac{p^2}{(\log\log p)^2}. \tag{53}$$

4.5 Why $\gcd(M_0, d) = 1$ Matters

A crucial consequence of our construction is the following coprimality condition.

Lemma 4.10 (Coprimality of M_0 and d)**.** *For any divisor d of P^* (i.e., any squarefree integer all of whose prime factors lie in S_1), we have*

$$\gcd(M_0, d) = 1. \tag{54}$$

Proof. All prime factors of M_0 lie in S_0, i.e., are $\le \log p$. All prime factors of d lie in S_1, i.e., are $> \log p$. The two sets of primes are disjoint. Therefore, M_0 and d share no common prime factor, so their greatest common divisor is 1. □

This simple fact is the linchpin of our error analysis. It implies that the linear map $t \mapsto a + tM_0 \pmod d$ is a bijection on $\mathbb{Z}/d\mathbb{Z}$ for each fixed a. Consequently, as t runs through a complete set of d consecutive integers, $a + tM_0 \pmod d$ runs through each residue class modulo d exactly once. This "complete period" property leads to exact cancellation in the error terms (the "physical-space period cancellation").

5 The Counting Function

5.1 Definition of T

We now define the central object of our proof: the counting function T that enumerates pairs (a, t) for which both $x_{a,t}$ and $x_{a,t} + y$ survive the sieve by all primes in S_1.

Definition 5.1 (Counting Function T)**.**

$$T = \sum_{a \in \mathcal{A}_0} \sum_{t=0}^{N-1} \prod_{r \in S_1} \Big(1 - \mathbf{1}_{r|(a+tM_0)} - \mathbf{1}_{r|(a+tM_0+y)} + \mathbf{1}_{r|(a+tM_0)} \cdot \mathbf{1}_{r|(a+tM_0+y)}\Big). \tag{55}$$

Let us parse this expression carefully. For each $r \in S_1$ and each pair (a, t), the factor in parentheses is:

- 1 if neither $a + tM_0$ nor $a + tM_0 + y$ is divisible by r;
- 0 otherwise (when at least one of them is divisible by r).

The product over all $r \in S_1$ is therefore the indicator function of the event: "for all $r \in S_1$, $r \nmid (a + tM_0)$ and $r \nmid (a + tM_0 + y)$."

Lemma 5.2 (T counts prime pairs)**.** *If $T > 0$, then there exists at least one prime pair $(q, q + y)$ with $p < q < q + y < p^2$.*

Proof. If $T > 0$, there exists at least one pair (a_0, t_0) such that the product in (55) equals 1. For this pair:

- By construction of $\mathcal{A}_0$, a_0 and $a_0 + y$ are coprime to all primes $r \in S_0$.
- By the condition that the product equals 1, $a_0 + t_0M_0$ and $a_0 + t_0M_0 + y$ are coprime to all primes $r \in S_1$.
- Since $S_0 \cup S_1 = \{r \le p : r \text{ prime}\}$, both $x = a_0 + t_0M_0$ and $x + y$ are coprime to all primes $\le p$.
- By Lemma 4.8, $p < x < x + y < p^2$.
- By Corollary 3.4, both x and $x + y$ are prime.

Setting $q = x$ yields the desired prime pair. □

Thus, the entire proof reduces to showing that $T > 0$ for all sufficiently large p.

5.2 Handling the Case $r \mid y$

Before proceeding to the Möbius inversion, we must carefully handle primes r that divide y. For such primes, the two conditions $r \mid (a + tM_0)$ and $r \mid (a + tM_0 + y)$ are equivalent (since $y \equiv 0 \pmod r$). Let us verify that the expression in (55) still behaves correctly.

For $r \mid y$:

$$\mathbf{1}_{r|(a+tM_0)} = \mathbf{1}_{r|(a+tM_0+y)},$$
$$\mathbf{1}_{r|(a+tM_0)} \cdot \mathbf{1}_{r|(a+tM_0+y)} = \mathbf{1}_{r|(a+tM_0)}.$$

The factor becomes

$$1 - \mathbf{1} - \mathbf{1} + \mathbf{1} = 1 - \mathbf{1}_{r|(a+tM_0)},$$

which is 1 if $r \nmid (a + tM_0)$ and 0 otherwise. This is exactly the correct indicator, since when $r \mid y$, we only need to avoid r dividing x (which automatically avoids r dividing $x + y$).

For $r \nmid y$, the two events are mutually exclusive (a number cannot be simultaneously congruent to 0 and $-y \not\equiv 0 \pmod r$), so the cross term vanishes, and the factor becomes

$$1 - \mathbf{1}_{r|x} - \mathbf{1}_{r|(x+y)},$$

which is 1 exactly when neither event occurs.

Thus (55) is correct for all primes $r \in S_1$.

5.3 Möbius Inversion

We now apply the inclusion-exclusion principle (equivalently, Möbius inversion on the divisor lattice) to expand the product over $r \in S_1$.

For each $r \in S_1$, define the set of forbidden residue classes modulo r as in Definition 4.3:

$$\mathcal{F}_r = \{c \bmod r : c \equiv 0 \text{ or } c \equiv -y \pmod r\}. \tag{56}$$

For any divisor d of $P^* = \prod_{r\in S_1} r$, we can combine the forbidden classes for all $r \mid d$ via the Chinese Remainder Theorem.

Definition 5.3 (Forbidden Classes Modulo d)**.** *For $d \mid P^*$ (so d is squarefree with all prime factors in S_1), define*

$$\mathscr{C}_d = \{c \bmod d : \forall r \mid d,\ c \bmod r \in \mathcal{F}_r\}. \tag{57}$$

That is, $\mathscr{C}_d$ is the set of residue classes modulo d that are "forbidden" for each prime factor r of d.

By CRT, $\mathscr{C}_d$ is well-defined and its cardinality is the product of the cardinalities of $\mathcal{F}_r$ for $r \mid d$:

$$|\mathscr{C}_d| = \prod_{r|d} |\mathcal{F}_r| = \prod_{\substack{r|d \\ r\nmid y}} 2 \times \prod_{\substack{r|d \\ r|y}} 1 = 2^{\omega_1(d)}, \tag{58}$$

where $\omega_1(d)$ denotes the number of prime factors of d that do not divide y. In particular,

$$|\mathscr{C}_d| \le 2^{\omega(d)}. \tag{59}$$

Theorem 5.4 (Möbius Inversion Formula)**.** *For any integer x,*

$$\prod_{r\in S_1} \left(1 - \mathbf{1}_{r|x} - \mathbf{1}_{r|(x+y)} + \mathbf{1}_{r|x}\mathbf{1}_{r|(x+y)}\right) = \sum_{d|P^*} \mu(d) \sum_{c\in\mathscr{C}_d} \mathbf{1}_{x\equiv c \pmod d}. \tag{60}$$

Proof. We prove this identity by expanding the product and applying the CRT. A complete, self-contained verification is provided in Appendix A. Here we sketch the main steps.

For each $r \in S_1$, write the factor as

$$f_r(x) = 1 - \mathbf{1}_{r|x} - \mathbf{1}_{r|(x+y)} + \mathbf{1}_{r|x}\mathbf{1}_{r|(x+y)}. \tag{61}$$

For $r \nmid y$, the two indicator functions are mutually exclusive, so the cross term vanishes and $f_r(x) = 1 - \mathbf{1}_{x\equiv 0} - \mathbf{1}_{x\equiv -y}$. This can be written as a Möbius sum over the divisors of r with appropriate forbidden classes. For $r \mid y$, we have $f_r(x) = 1 - \mathbf{1}_{r|x}$, which is also a Möbius sum.

Taking the product over all $r \in S_1$ and expanding yields a sum over all choices of $d_r \in \{1, r\}$ for each r. A choice where $d_r = r$ for $r \in T$ corresponds to the divisor $d = \prod_{r\in T} r$, with coefficient $\mu(d)$. By CRT, the simultaneous congruence conditions combine to a single congruence modulo d, with the residue class belonging to $\mathscr{C}_d$. The identity follows. □

5.4 Expansion of T

Applying Theorem 5.4 to (55) with $x = a + tM_0$, and exchanging the order of summation:

$$\begin{aligned} T &= \sum_{a\in\mathcal{A}_0} \sum_{t=0}^{N-1} \sum_{d|P^*} \mu(d) \sum_{c\in\mathscr{C}_d} \mathbf{1}_{a+tM_0\equiv c \pmod d} \\ &= \sum_{d|P^*} \mu(d) \sum_{c\in\mathscr{C}_d} \sum_{a\in\mathcal{A}_0} \sum_{t=0}^{N-1} \mathbf{1}_{a+tM_0\equiv c \pmod d}. \end{aligned} \tag{62}$$

Define, for each $d \mid P^*$ and $c \in \mathscr{C}_d$,

$$S(d, c) = \sum_{a\in\mathcal{A}_0} \sum_{t=0}^{N-1} \mathbf{1}_{a+tM_0\equiv c \pmod d}. \tag{63}$$

Thus

$$T = \sum_{d|P^*} \mu(d) \sum_{c\in\mathscr{C}_d} S(d, c). \tag{64}$$

5.5 Separation of Main Term and Error

Fix $d \mid P^*$ and $c \in \mathscr{C}_d$. The inner sum $S(d, c)$ counts the number of pairs $(a, t) \in \mathcal{A}_0 \times \{0, \dots, N-1\}$ such that $a + tM_0 \equiv c \pmod d$.

Since $\gcd(M_0, d) = 1$ (Lemma 4.10), M_0 is invertible modulo d. For each fixed a, the congruence

$$a + tM_0 \equiv c \pmod d \tag{65}$$

has the unique solution

$$t \equiv M_0^{-1}(c - a) \pmod d, \tag{66}$$

where M_0^{-1} denotes the multiplicative inverse of M_0 modulo d.

Let $t_0(a, c) \in \{0, 1, \dots, d-1\}$ be the least non-negative residue of $M_0^{-1}(c-a)$ modulo d. Then all solutions to (65) in non-negative integers are

$$t = t_0(a, c) + kd, \qquad k = 0, 1, 2, \dots \tag{67}$$

The number of such t in the range $\{0, 1, \dots, N-1\}$ depends on how $t_0(a, c)$ compares to the remainder $r_d = N \bmod d$. Write

$$N = q_d \cdot d + r_d, \qquad 0 \le r_d < d, \tag{68}$$

with $q_d = \lfloor N/d \rfloor$.

For $t_0(a,c) < r_d$: the values $t = t_0(a,c), t_0(a,c)+d, \ldots, t_0(a,c)+q_d \cdot d$ all lie in $\{0,\ldots,N-1\}$ (the last one is $t_0 + q_d d \le (r_d-1) + q_d d = N-1$). The count is q_d+1.

For $t_0(a,c) \ge r_d$: the values $t = t_0(a,c), \ldots, t_0(a,c)+(q_d-1)d$ lie in the range, but $t_0 + q_d d \ge r_d + q_d d = N$ does not. The count is q_d.

Thus,

$$\#\{t \in [0,N-1] : a + tM_0 \equiv c \pmod d\} = \begin{cases} q_d + 1 = \lceil N/d \rceil, & \text{if } t_0(a,c) < r_d, \\ q_d = \lfloor N/d \rfloor, & \text{if } t_0(a,c) \ge r_d. \end{cases} \tag{69}$$

We can express this as

$$\#\{t\} = \frac{N}{d} + \delta(a,c,d), \tag{70}$$

where

$$\delta(a,c,d) = \begin{cases} 1 - \dfrac{r_d}{d}, & \text{if } t_0(a,c) < r_d, \\ -\dfrac{r_d}{d}, & \text{if } t_0(a,c) \ge r_d. \end{cases} \tag{71}$$

Note that $|\delta(a,c,d)| \le 1$ and, crucially,

$$\sum_{t_0=0}^{d-1} \delta(t_0) = r_d \left(1 - \frac{r_d}{d}\right) + (d - r_d)\left(-\frac{r_d}{d}\right) = 0, \tag{72}$$

where $\delta(t_0)$ denotes the value of δ when $t_0(a,c) = t_0$. This mean-zero property is the foundation of the period cancellation mechanism.

Summing over $a \in \mathcal{A}_0$:

$$S(d,c) = \sum_{a \in \mathcal{A}_0} \left(\frac{N}{d} + \delta(a,c,d)\right) = \frac{|\mathcal{A}_0| N}{d} + R(d,c), \tag{73}$$

where

$$R(d,c) = \sum_{a \in \mathcal{A}_0} \delta(a,c,d). \tag{74}$$

Summing over $c \in \mathscr{C}_d$:

$$R_d = \sum_{c \in \mathscr{C}_d} R(d,c). \tag{75}$$

Finally, substituting into (64):

$$\begin{aligned} T &= \sum_{d \mid P^*} \mu(d) \sum_{c \in \mathscr{C}_d} \left(\frac{|\mathcal{A}_0| N}{d} + R(d,c)\right) \\ &= |\mathcal{A}_0| N \sum_{d \mid P^*} \frac{\mu(d) |\mathscr{C}_d|}{d} + \sum_{d \mid P^*} \mu(d) R_d \\ &= T_{\text{main}} + \mathcal{E}. \end{aligned} \tag{76}$$

The first term T_{main} is the expected count (the "main term"), and $\mathcal{E}$ is the total deviation (the "error term").

6 Evaluation of the Main Term

6.1 Factorization of the Main Term

The main term is

$$T_{\text{main}} = |\mathcal{A}_0| N \sum_{d|P^*} \frac{\mu(d)|\mathscr{C}_d|}{d}. \tag{77}$$

Since both $\mu(d)$ and $|\mathscr{C}_d|$ are multiplicative functions of the squarefree integer d, and the sum runs over all squarefree divisors of P^*, the sum factors over the primes in S_1:

$$\begin{aligned} \sum_{d|P^*} \frac{\mu(d)|\mathscr{C}_d|}{d} &= \prod_{r\in S_1} \left(1 + \frac{\mu(r)|\mathscr{C}_r|}{r}\right) \\ &= \prod_{r\in S_1} \left(1 - \frac{|\mathcal{F}_r|}{r}\right) \\ &= \prod_{\substack{r\in S_1 \\ r\nmid y}} \left(1 - \frac{2}{r}\right) \prod_{\substack{r\in S_1 \\ r|y}} \left(1 - \frac{1}{r}\right). \end{aligned} \tag{78}$$

This factorization is valid because for a squarefree $d = \prod_{r\in T} r$, we have $\mu(d) = \prod_{r\in T}(-1)$, $|\mathscr{C}_d| = \prod_{r\in T} |\mathcal{F}_r|$, and $1/d = \prod_{r\in T}(1/r)$. The sum over all subsets $T \subseteq S_1$ of the product of $(-|\mathcal{F}_r|/r)$ over $r \in T$ is precisely the product over $r \in S_1$ of $(1 - |\mathcal{F}_r|/r)$.

6.2 Asymptotic Evaluation

We now evaluate the three factors in T_{main} asymptotically.

Factor 1: $|\mathcal{A}_0|$. From Lemma 4.5,

$$|\mathcal{A}_0| = \frac{C_{\mathcal{A}} M_0}{(\log\log p)^2}(1 + o(1)). \tag{79}$$

Factor 2: N. From Lemma 4.9,

$$N = \frac{p^2}{M_0}(1 + o(1)). \tag{80}$$

Factor 3: The product over S_1. We split the product over all primes $\le p$ into the contribution of S_0 and S_1:

$$\prod_{\substack{r\le p \\ r\nmid y}} \left(1 - \frac{2}{r}\right) = \prod_{\substack{r\in S_0 \\ r\nmid y}} \left(1 - \frac{2}{r}\right) \times \prod_{\substack{r\in S_1 \\ r\nmid y}} \left(1 - \frac{2}{r}\right). \tag{81}$$

By Mertens' theorem (Corollary 2.5) with $k = 2$,

$$\prod_{\substack{r\le p \\ r\nmid y}} \left(1 - \frac{2}{r}\right) \sim \frac{C_y}{(\log p)^2}, \tag{82}$$

where C_y is the Hardy-Littlewood constant (2). Similarly,

$$\prod_{\substack{r\le \log p \\ r\nmid y}} \left(1 - \frac{2}{r}\right) \sim \frac{C'_y}{(\log\log p)^2}, \tag{83}$$

where C'_y is a related constant. Therefore,

$$\prod_{\substack{r\in S_1\\ r\nmid y}}\left(1-\frac{2}{r}\right)=\frac{C_y/(\log p)^2}{C_y/(\log\log p)^2}(1+o(1))=\frac{(\log\log p)^2}{(\log p)^2}(1+o(1)). \tag{84}$$

The product over $r \mid y$ in S_1 converges to 1 as $p \to \infty$, because for sufficiently large p, $\log p$ exceeds all prime factors of y, so S_1 contains no primes dividing y. Hence for large p,

$$\prod_{\substack{r\in S_1\\ r\mid y}}\left(1-\frac{1}{r}\right)=1. \tag{85}$$

6.3 Final Asymptotic for T_{main}

Combining the three factors:

$$\begin{aligned} T_{\text{main}} &= |\mathcal{A}_0|\cdot N\cdot\prod_{r\in S_1}\left(1-\frac{|\mathcal{F}_r|}{r}\right)\\ &\sim \frac{C_{\mathcal{A}}M_0}{(\log\log p)^2}\cdot\frac{p^2}{M_0}\cdot\frac{(\log\log p)^2}{(\log p)^2}\\ &= C_{\mathcal{A}}\frac{p^2}{(\log p)^2}. \end{aligned} \tag{86}$$

The factors of M_0 cancel exactly! The factors of $(\log\log p)^2$ also cancel. This cancellation is not a coincidence but a structural consequence of our setup: the pre-sieving by S_0 reduces the count by $(\log\log p)^{-2}$, while the subsequent sieving by S_1 contributes exactly $(\log\log p)^2/(\log p)^2$ relative to the full interval.

We have proved:

Theorem 6.1 (Main Term Asymptotic)**.** *As $p \to \infty$,*

$$T_{main}\sim C_0\frac{p^2}{(\log p)^2}, \tag{87}$$

where $C_0 = C_{\mathcal{A}} > 0$ is exactly the Hardy-Littlewood constant C_y for the pair $(0, y)$.

Verification that $C_0 = C_y$. From the definitions, $C_{\mathcal{A}}$ is the limit of $|\mathcal{A}_0|(\log\log p)^2/M_0$ as $p \to \infty$. But

$$|\mathcal{A}_0| = M_0\prod_{r\le\log p}\left(1-\frac{|\mathcal{F}_r|}{r}\right),$$

so

$$C_{\mathcal{A}}=\lim_{p\to\infty}(\log\log p)^2\prod_{r\le\log p}\left(1-\frac{|\mathcal{F}_r|}{r}\right).$$

This product, times $(\log\log p)^2$, converges to the same constant as

$$\lim_{x\to\infty}(\log x)^2\prod_{r\le x}\left(1-\frac{|\mathcal{F}_r|}{r}\right),$$

which is precisely the Hardy-Littlewood constant C_y (see [5] and [6, Chapter 1]). □

7 Analysis of the Error Term: Preliminaries

We now turn to the estimation of the error term

$$\mathcal{E} = \sum_{d|P^*} \mu(d) R_d, \tag{88}$$

where

$$R_d = \sum_{c \in \mathscr{C}_d} R(d,c) = \sum_{c \in \mathscr{C}_d} \sum_{a \in \mathcal{A}_0} \delta(a,c,d), \tag{89}$$

and $\delta(a,c,d)$ is given by (71).

7.1 Structure of the Deviation δ

Recall from (71) that $\delta(a,c,d)$ depends on (a,c) only through the value $t_0(a,c) = M_0^{-1}(c-a) \bmod d$. Specifically, if we define a function $\Delta_d : \mathbb{Z}/d\mathbb{Z} \to \mathbb{R}$ by

$$\Delta_d(t) = \begin{cases} 1 - \dfrac{r_d}{d}, & \text{if } 0 \le t < r_d, \\ -\dfrac{r_d}{d}, & \text{if } r_d \le t < d, \end{cases} \tag{90}$$

then

$$\delta(a,c,d) = \Delta_d\big(t_0(a,c)\big) = \Delta_d\big(M_0^{-1}(c-a) \bmod d\big). \tag{91}$$

The function Δ_d has two crucial properties:

(i) **Boundedness:** $|\Delta_d(t)| \le 1$ for all $t \in \mathbb{Z}/d\mathbb{Z}$.

(ii) **Mean zero:**

$$\sum_{t=0}^{d-1} \Delta_d(t) = r_d\left(1 - \frac{r_d}{d}\right) + (d - r_d)\left(-\frac{r_d}{d}\right) = r_d - \frac{r_d^2}{d} - r_d + \frac{r_d^2}{d} = 0. \tag{92}$$

The mean-zero property is the algebraic expression of the "period cancellation": over a complete period of d consecutive values of t, the positive and negative deviations exactly cancel. This is the fundamental geometric fact that makes the entire error analysis possible.

7.2 Distribution of $\mathcal{A}_0$ Modulo d

To proceed further, we need precise information about how the elements of $\mathcal{A}_0$ are distributed among the residue classes modulo d. Since $\gcd(M_0, d) = 1$, the Chinese Remainder Theorem provides a complete answer.

Lemma 7.1 (Distribution of $\mathcal{A}_0$ Modulo d)**.** *For each $x \in \mathbb{Z}/d\mathbb{Z}$, let*

$$f_d(x) = \#\{a \in \mathcal{A}_0 : a \equiv x \pmod d\}. \tag{93}$$

Then there exists a function $\varepsilon_d : \mathbb{Z}/d\mathbb{Z} \to \mathbb{R}$ such that

$$f_d(x) = \frac{|\mathcal{A}_0|}{d} + \varepsilon_d(x), \tag{94}$$

with the following properties:

(i) $|\varepsilon_d(x)| \le 1$ *for all* $x \in \mathbb{Z}/d\mathbb{Z}$.

(ii) $\displaystyle\sum_{x\in\mathbb{Z}/d\mathbb{Z}} \varepsilon_d(x) = 0.$

(iii) $\displaystyle\sum_{x\in\mathbb{Z}/d\mathbb{Z}} \varepsilon_d(x)^2 = |\mathcal{A}_0|\left(1 - \frac{|\mathcal{A}_0|}{d}\right)$ *when* $d > |\mathcal{A}_0|$, *and* $O(|\mathcal{A}_0|)$ *in general.*

Proof. Since $\gcd(M_0, d) = 1$, the CRT isomorphism

$$\mathbb{Z}/M_0 d\mathbb{Z} \cong \mathbb{Z}/M_0\mathbb{Z} \times \mathbb{Z}/d\mathbb{Z}$$

holds. The set $\mathcal{A}_0$ is defined by conditions modulo M_0 only (specifically, $a \bmod r \notin \mathcal{F}_r$ for all $r \in S_0$).

For a fixed residue class $x \bmod d$, we need to count the number of $a \bmod M_0$ in $\mathcal{A}_0$ such that the unique integer in $[0, M_0)$ satisfying both $a \equiv a_0 \pmod{M_0}$ and $a \equiv x \pmod{d}$ lies in our chosen range of representatives. By CRT, for each $a_0 \in \mathcal{A}_0$, there is exactly one residue class modulo $M_0 d$ satisfying both conditions. The number of these falling in our chosen range $[0, M_0)$ is either $\lfloor|\mathcal{A}_0|/d\rfloor$ or $\lceil|\mathcal{A}_0|/d\rceil$, since the d residue classes modulo d partition the $|\mathcal{A}_0|$ elements of $\mathcal{A}_0$ as evenly as possible (the difference between any two classes is at most 1).

Thus $f_d(x) \in \{\lfloor|\mathcal{A}_0|/d\rfloor, \lceil|\mathcal{A}_0|/d\rceil\}$. Setting $\varepsilon_d(x) = f_d(x) - |\mathcal{A}_0|/d$, we immediately obtain $|\varepsilon_d(x)| \le 1$ and $\sum_x \varepsilon_d(x) = \sum_x f_d(x) - |\mathcal{A}_0| = 0$.

For the sum of squares, we consider two cases. When $d > |\mathcal{A}_0|$, each residue class contains at most one element of $\mathcal{A}_0$, so $f_d(x) \in \{0, 1\}$ and $\sum_x f_d(x)^2 = \sum_x f_d(x) = |\mathcal{A}_0|$. Then

$$\sum_x \varepsilon_d(x)^2 = \sum_x \left(f_d(x) - \frac{|\mathcal{A}_0|}{d}\right)^2 = |\mathcal{A}_0| - 2\frac{|\mathcal{A}_0|^2}{d} + \frac{|\mathcal{A}_0|^2}{d} = |\mathcal{A}_0|\left(1 - \frac{|\mathcal{A}_0|}{d}\right).$$

When $d \le |\mathcal{A}_0|$, the same expression holds up to a constant factor, yielding $O(|\mathcal{A}_0|)$ in all cases. □

7.3 Expression for R_d in Terms of ε_d

We now rewrite R_d in a form that separates the geometric deviation ε_d from the arithmetic structure of the progression.

Lemma 7.2.

$$R_d = \sum_{c\in\mathscr{C}_d} \sum_{t=0}^{r_d-1} \varepsilon_d\big(M_0^{-1}c - t \bmod d\big), \tag{95}$$

where the sum over t *is taken modulo* d *(i.e.,* t *runs from* 0 *to* $r_d - 1$ *in* $\mathbb{Z}$, *and we interpret* $M_0^{-1}c - t$ *modulo* d*).*

Proof. Starting from the definition $R(d, c) = \sum_{a\in\mathcal{A}_0} \delta(a, c, d)$, we substitute the expression for δ in terms of Δ_d:

$$\begin{aligned} R(d, c) &= \sum_{a\in\mathcal{A}_0} \Delta_d(M_0^{-1}(c - a)) \\ &= \sum_{x\in\mathbb{Z}/d\mathbb{Z}} f_d(x)\Delta_d(M_0^{-1}(c - x)), \end{aligned}$$

where we have grouped elements of $\mathcal{A}_0$ by their residue class x modulo d.

Now substitute $f_d(x) = |\mathcal{A}_0|/d + \varepsilon_d(x)$. The constant term $|\mathcal{A}_0|/d$ contributes

$$\frac{|\mathcal{A}_0|}{d} \sum_{x \in \mathbb{Z}/d\mathbb{Z}} \Delta_d(M_0^{-1}(c-x)) = 0,$$

because the map $x \mapsto M_0^{-1}(c-x) \bmod d$ is a bijection on $\mathbb{Z}/d\mathbb{Z}$ (since $\gcd(M_0, d) = 1$), and $\sum_t \Delta_d(t) = 0$ by (92).

Thus only the ε_d part contributes:

$$R(d,c) = \sum_{x \in \mathbb{Z}/d\mathbb{Z}} \varepsilon_d(x)\Delta_d(M_0^{-1}(c-x)).$$

Now we use the explicit definition of Δ_d. Let $y = M_0^{-1}(c-x) \bmod d$, so $x = c - M_0 y \bmod d$. As x runs over $\mathbb{Z}/d\mathbb{Z}$, y also runs over $\mathbb{Z}/d\mathbb{Z}$ (since multiplication by M_0 is a bijection modulo d). Then

$$R(d,c) = \sum_{y \in \mathbb{Z}/d\mathbb{Z}} \varepsilon_d(c - M_0 y)\Delta_d(y).$$

By definition, $\Delta_d(y) = 1 - r_d/d$ for $0 \le y < r_d$, and $\Delta_d(y) = -r_d/d$ for $r_d \le y < d$. Therefore,

$$\begin{aligned} R(d,c) &= \sum_{y=0}^{r_d-1} \varepsilon_d(c - M_0 y)\left(1 - \frac{r_d}{d}\right) + \sum_{y=r_d}^{d-1} \varepsilon_d(c - M_0 y)\left(-\frac{r_d}{d}\right) \\ &= \sum_{y=0}^{r_d-1} \varepsilon_d(c - M_0 y) - \frac{r_d}{d} \sum_{y=0}^{d-1} \varepsilon_d(c - M_0 y). \end{aligned}$$

The second sum vanishes because it equals $\frac{r_d}{d} \sum_x \varepsilon_d(x) = 0$ (again using the bijection property). Hence

$$R(d,c) = \sum_{y=0}^{r_d-1} \varepsilon_d(c - M_0 y).$$

Finally, we change variables. Let $t = y$ and note that $c - M_0 t = M_0(M_0^{-1}c - t)$. Since multiplication by M_0 is a bijection on $\mathbb{Z}/d\mathbb{Z}$, we have $\varepsilon_d(c - M_0 t) = \varepsilon_d(M_0(M_0^{-1}c - t))$. By reindexing (letting $z = M_0^{-1}c - t \bmod d$, which runs over a permutation of $\mathbb{Z}/d\mathbb{Z}$ as t varies), we obtain the equivalent form

$$R(d,c) = \sum_{t=0}^{r_d-1} \varepsilon_d(M_0^{-1}c - t).$$

Summing over $c \in \mathscr{C}_d$ yields the stated formula. □

7.4 Fourier Expansion of R_d

The expression (95) is the starting point for our Fourier analysis. We expand ε_d in a Fourier series on the cyclic group $\mathbb{Z}/d\mathbb{Z}$.

Definition 7.3. *For $k \in \mathbb{Z}/d\mathbb{Z}$, define the Fourier transform of ε_d by*

$$\widehat{\varepsilon}_d(k) = \sum_{x \in \mathbb{Z}/d\mathbb{Z}} \varepsilon_d(x) \mathrm{e}\left(-\frac{kx}{d}\right). \tag{96}$$

Since $\sum_x \varepsilon_d(x) = 0$, we have $\widehat{\varepsilon}_d(0) = 0$. By Fourier inversion (Theorem 2.7),

$$\varepsilon_d(x) = \frac{1}{d} \sum_{k=1}^{d-1} \widehat{\varepsilon}_d(k) \mathrm{e}\left(\frac{kx}{d}\right). \tag{97}$$

Lemma 7.4 (Fourier Expression for R_d)**.**

$$R_d = \frac{1}{d} \sum_{k=1}^{d-1} \widehat{\varepsilon}_d(k) \, \overline{\Sigma_d(k)} \, \overline{\widehat{C}_d(k)}, \tag{98}$$

where the functions Σ_d and $\widehat{C}_d$ are defined as follows:

$$\Sigma_d(k) = \sum_{t=0}^{r_d-1} \mathrm{e}\left(\frac{kt}{d}\right) = \frac{1 - \mathrm{e}(kr_d/d)}{1 - \mathrm{e}(k/d)}, \tag{99}$$

and

$$\widehat{C}_d(k) = \sum_{c \in \mathscr{C}_d} \mathrm{e}\left(\frac{kM_0^{-1}c}{d}\right). \tag{100}$$

Proof. Substituting the Fourier expansion (97) of ε_d into Lemma 7.2:

$$\begin{aligned} R_d &= \sum_{c \in \mathscr{C}_d} \sum_{t=0}^{r_d-1} \frac{1}{d} \sum_{k=1}^{d-1} \widehat{\varepsilon}_d(k) \mathrm{e}\left(\frac{k(M_0^{-1}c - t)}{d}\right) \\ &= \frac{1}{d} \sum_{k=1}^{d-1} \widehat{\varepsilon}_d(k) \left(\sum_{c \in \mathscr{C}_d} \mathrm{e}\left(\frac{kM_0^{-1}c}{d}\right)\right) \left(\sum_{t=0}^{r_d-1} \mathrm{e}\left(-\frac{kt}{d}\right)\right). \end{aligned}$$

The inner sum over t is precisely $\overline{\Sigma_d(k)}$, since

$$\sum_{t=0}^{r_d-1} \mathrm{e}\left(-\frac{kt}{d}\right) = \overline{\sum_{t=0}^{r_d-1} \mathrm{e}\left(\frac{kt}{d}\right)} = \overline{\Sigma_d(k)}.$$

The sum over c is exactly $\widehat{C}_d(k)$ as defined. This gives the stated formula. □

7.5 Basic Bounds

We now collect elementary bounds for the quantities appearing in the Fourier expression (98). These bounds will be refined later using the resonance breaking technique.

Lemma 7.5 (Bounds for $\widehat{\varepsilon}_d$)**.** *For all $k \in \mathbb{Z}/d\mathbb{Z}$,*

$$|\widehat{\varepsilon}_d(k)| \le \sum_{x \in \mathbb{Z}/d\mathbb{Z}} |\varepsilon_d(x)| \le 2|\mathcal{A}_0|. \tag{101}$$

Moreover, by Parseval's identity (Theorem 2.8),

$$\sum_{k=1}^{d-1} |\widehat{\varepsilon}_d(k)|^2 = d \sum_{x \in \mathbb{Z}/d\mathbb{Z}} |\varepsilon_d(x)|^2 = d \cdot O(|\mathcal{A}_0|) = O(d|\mathcal{A}_0|). \tag{102}$$

Proof. The first bound follows from the triangle inequality and the fact that $|\varepsilon_d(x)| \leq 1$ with ε_d non-zero on at most $2|\mathcal{A}_0|$ residue classes (by the explicit description in Lemma 7.1). The second bound is a direct application of Parseval's identity combined with Lemma 7.1(iii). □

Lemma 7.6 (Bounds for Σ_d)**.** *For $k \in \{1, \ldots, d-1\}$,*

$$|\Sigma_d(k)| = \left|\frac{\sin(\pi k r_d/d)}{\sin(\pi k/d)}\right| \leq \min\left(r_d, \frac{1}{|\sin(\pi k/d)|}\right) \leq \frac{1}{|\sin(\pi k/d)|}. \tag{103}$$

Proof. This follows directly from the geometric series formula (16) and the bound (17). □

Lemma 7.7 (Bounds for $\widehat{C}_d$)**.** *For all k,*

$$|\widehat{C}_d(k)| \leq |\mathscr{C}_d| \leq 2^{\omega(d)}. \tag{104}$$

Proof. The triangle inequality gives $|\sum_{c\in\mathscr{C}_d} \mathrm{e}(kM_0^{-1}c/d)| \leq \sum_{c\in\mathscr{C}_d} 1 = |\mathscr{C}_d|$. The second inequality is (59). □

7.6 The Need for Resonance Breaking: Why the Naive Bound Fails

If we were to combine the bounds above directly, we would obtain the following estimate for $|R_d|$:

$$\begin{aligned}
|R_d| &\leq \frac{1}{d}\sum_{k=1}^{d-1} |\widehat{\varepsilon}_d(k)| \cdot |\widehat{C}_d(k)| \cdot |\Sigma_d(k)| \\
&\leq \frac{1}{d}\sum_{k=1}^{d-1} 2|\mathcal{A}_0| \cdot 2^{\omega(d)} \cdot \frac{1}{|\sin(\pi k/d)|} \\
&\ll |\mathcal{A}_0| \cdot 2^{\omega(d)} \cdot \frac{1}{d}\sum_{k=1}^{d-1} \frac{1}{|\sin(\pi k/d)|}.
\end{aligned}$$

By Lemma 2.9, $\frac{1}{d}\sum_{k=1}^{d-1} 1/|\sin(\pi k/d)| \ll \log d$. Thus

$$|R_d| \ll |\mathcal{A}_0| \cdot 2^{\omega(d)} \cdot \log d.$$

Now sum this over all moduli $d \leq M_0 \approx p$:

$$\begin{aligned}
\sum_{d\leq M_0} |R_d| &\ll |\mathcal{A}_0| \sum_{d\leq p} 2^{\omega(d)} \log d \\
&\approx \frac{p}{(\log\log p)^2} \cdot p(\log p)^2 = \frac{p^2(\log p)^2}{(\log\log p)^2}.
\end{aligned}$$

This error bound is **larger** than the main term $T_{\text{main}} \sim C_0 p^2/(\log p)^2$ by a factor of $(\log p)^4$. The problem is the factor $|\mathcal{A}_0| \approx p/(\log\log p)^2$ in the bound for $\widehat{\varepsilon}_d(k)$.

The resonance breaking technique, developed in the next section, eliminates this factor entirely by decomposing the exponential sum $S(d,k) = \sum_{a\in\mathcal{A}_0} \mathrm{e}(ka/d)$ into $3^{\pi(\log p)} = p^{o(1)}$ explicit terms, each bounded independently of $|\mathcal{A}_0|$.

8 Resonance Breaking

Resonance breaking is the central technical innovation of this paper. It provides a decomposition of the exponential sum over the CRT-constructed set $\mathcal{A}_0$ into a sum of $3^{\pi(\log p)}$ terms, each of which can be estimated without reference to the large factor $|\mathcal{A}_0|$.

8.1 Characteristic Function of $\mathcal{A}_0$ via CRT Expansion

Recall from Definition 4.4 that $a \in \mathcal{A}_0$ if and only if for every $r \in S_0$, $a \bmod r \notin \mathcal{F}_r$. This condition can be expressed as a product over $r \in S_0$:

$$\mathbf{1}_{\mathcal{A}_0}(a) = \prod_{r \in S_0} \Big(1 - \mathbf{1}_{a \equiv 0 \pmod r} - \mathbf{1}_{a \equiv -y \pmod r} + \mathbf{1}_{a \equiv 0} \mathbf{1}_{a \equiv -y}\Big). \tag{105}$$

For each prime $r \in S_0$, the factor in parentheses is the indicator that a avoids both forbidden residue classes (or the single forbidden class if $r \mid y$). Expanding this product yields a sum over all subsets $T \subseteq S_0$, and for each $r \in T$, a choice of which forbidden class to impose.

Definition 8.1 (Configuration)**.** *A* configuration *is a pair $(T, \vec{\varepsilon})$ where:*

- *$T \subseteq S_0$ is a (possibly empty) subset of small primes.*
- *$\vec{\varepsilon} = (\varepsilon_r)_{r \in T}$ is a choice vector, where for each $r \in T$, $\varepsilon_r \in \mathcal{F}_r$ selects which forbidden residue class to impose modulo r.*

For each configuration, we define:

- *$M_T = \prod_{r \in T} r$, the product of the primes in T.*
- *$c_{T,\vec{\varepsilon}} \in \mathbb{Z}/M_T\mathbb{Z}$, the unique residue class modulo M_T such that $c_{T,\vec{\varepsilon}} \equiv \varepsilon_r \pmod r$ for all $r \in T$ (this exists and is unique by the CRT).*

The expansion of (105) is then

$$\mathbf{1}_{\mathcal{A}_0}(a) = \sum_{(T,\vec{\varepsilon})} (-1)^{|T|} \mathbf{1}_{a \equiv c_{T,\vec{\varepsilon}} \pmod{M_T}}, \tag{106}$$

where the sum runs over all configurations $(T, \vec{\varepsilon})$. The sign $(-1)^{|T|}$ comes from the Möbius inversion on the divisor lattice of M_0.

8.2 Number of Configurations

The number of configurations is computed by counting, for each $r \in S_0$, the number of choices:

- If $r \notin T$: 1 choice (impose neither forbidden class).
- If $r \in T$ and $r \nmid y$: $|\mathcal{F}_r| = 2$ choices (impose 0 or $-y$).
- If $r \in T$ and $r \mid y$: $|\mathcal{F}_r| = 1$ choice (impose 0).

Thus the total number of configurations is

$$\sum_{T\subseteq S_0}\prod_{r\in T}|\mathcal{F}_r| = \prod_{r\in S_0}(1+|\mathcal{F}_r|) = \prod_{\substack{r\le\log p\\ r\nmid y}} 3 \times \prod_{\substack{r\le\log p\\ r\mid y}} 2. \tag{107}$$

Since y is a fixed even integer, it has only finitely many prime divisors. Therefore, for all sufficiently large p, the primes dividing y are all in S_0, and the second product is a constant independent of p. Hence

$$\text{Number of configurations} = 3^{\pi(\log p)+O(1)} = 3^{\pi(\log p)}\cdot O(1). \tag{108}$$

Using the estimate $\pi(\log p) = (\log p)/\log\log p + o(\log p/\log\log p)$, we compute

$$\begin{aligned} 3^{\pi(\log p)} &= \exp(\pi(\log p)\log 3)\\ &= \exp\left(\frac{\log 3\cdot\log p}{\log\log p} + o\left(\frac{\log p}{\log\log p}\right)\right)\\ &= p^{(\log 3)/\log\log p}\cdot p^{o(1)} = p^{o(1)}. \end{aligned} \tag{109}$$

The key point is that $3^{\pi(\log p)}$ grows slower than any positive power of p, i.e., $3^{\pi(\log p)} = p^{o(1)}$.

8.3 Exponential Sum Over a Single Configuration

For a fixed configuration $(T,\vec{\varepsilon})$ and integers k,d with $\gcd(d,M_T)=1$, we need to evaluate the exponential sum over all $a \bmod M_0$ that satisfy the congruence condition $a \equiv c_{T,\vec{\varepsilon}} \pmod{M_T}$.

Lemma 8.2. *For any integer k and any configuration $(T,\vec{\varepsilon})$ with $T\subseteq S_0$,*

$$\sum_{\substack{a \bmod M_0\\ a\equiv c_{T,\vec{\varepsilon}} \pmod{M_T}}} \mathrm{e}\left(\frac{ka}{d}\right) = \mathrm{e}\left(\frac{kc_{T,\vec{\varepsilon}}}{d}\right)\cdot\frac{1-\mathrm{e}(kM_0/d)}{1-\mathrm{e}(kM_T/d)}, \tag{110}$$

provided that $\gcd(d,M_T)=1$ and $kM_T/d\notin\mathbb{Z}$.

Proof. Write $c = c_{T,\vec{\varepsilon}}$ for brevity. The condition $a\equiv c \pmod{M_T}$ means that a can be written as $a = c + uM_T$ for some integer u. As a runs over a complete residue system modulo M_0, the parameter u runs over the values $u = 0,1,\dots,M_0/M_T-1$ (since M_T divides M_0, the range of u is exactly M_0/M_T consecutive integers).

Substituting this parametrization:

$$\begin{aligned} \sum_{\substack{a \bmod M_0\\ a\equiv c \pmod{M_T}}} \mathrm{e}\left(\frac{ka}{d}\right) &= \sum_{u=0}^{M_0/M_T-1} \mathrm{e}\left(\frac{k(c+uM_T)}{d}\right)\\ &= \mathrm{e}\left(\frac{kc}{d}\right)\sum_{u=0}^{M_0/M_T-1}\mathrm{e}\left(\frac{kM_T}{d}u\right). \end{aligned}$$

The inner sum is a finite geometric series with ratio $q = \mathrm{e}(kM_T/d)$. Since $\gcd(M_T,d) = 1$ and $1\le k\le d-1$, we have $kM_T/d\notin\mathbb{Z}$, so $q\neq 1$. The sum of the geometric series is

$$\sum_{u=0}^{M_0/M_T-1} q^u = \frac{1-q^{M_0/M_T}}{1-q} = \frac{1-\mathrm{e}(kM_T\cdot M_0/(M_Td))}{1-\mathrm{e}(kM_T/d)} = \frac{1-\mathrm{e}(kM_0/d)}{1-\mathrm{e}(kM_T/d)}.$$

Multiplying by the prefactor $\mathrm{e}(kc/d)$ yields the stated formula. □

8.4 The Resonance Breaking Formula

We now assemble the pieces to obtain the resonance breaking decomposition of the exponential sum $S(d,k) = \sum_{a\in\mathcal{A}_0} \mathrm{e}(ka/d)$.

Theorem 8.3 (Resonance Breaking)**.** *For any integer d with $\gcd(M_0,d)=1$, and any integer $k\in\{1,\dots,d-1\}$,*

$$S(d,k) := \sum_{a\in\mathcal{A}_0} \mathrm{e}\left(\frac{ka}{d}\right) = \big(1-\mathrm{e}(kM_0/d)\big) \sum_{T\subseteq S_0} \sum_{\vec{\varepsilon}\in\prod_{r\in T}\mathcal{F}_r} \frac{(-1)^{|T|}\mathrm{e}(kc_{T,\vec{\varepsilon}}/d)}{1-\mathrm{e}(kM_T/d)}. \tag{111}$$

Proof. Substitute the expansion (106) of the characteristic function of $\mathcal{A}_0$ into the definition of $S(d,k)$:

$$\begin{aligned} S(d,k) &= \sum_{a \bmod M_0} \mathbf{1}_{\mathcal{A}_0}(a)\mathrm{e}\left(\frac{ka}{d}\right) \\ &= \sum_{a \bmod M_0} \left(\sum_{T\subseteq S_0} (-1)^{|T|} \sum_{\vec{\varepsilon}} \mathbf{1}_{a\equiv c_{T,\vec{\varepsilon}} \pmod{M_T}} \right) \mathrm{e}\left(\frac{ka}{d}\right) \\ &= \sum_{T\subseteq S_0} (-1)^{|T|} \sum_{\vec{\varepsilon}} \sum_{\substack{a \bmod M_0 \\ a\equiv c_{T,\vec{\varepsilon}} \pmod{M_T}}} \mathrm{e}\left(\frac{ka}{d}\right). \end{aligned}$$

Now apply Lemma 8.2 to the innermost sum. The hypotheses of the lemma are satisfied for all configurations: $\gcd(M_T,d)=1$ because M_T divides M_0 and $\gcd(M_0,d)=1$, and $kM_T/d\notin\mathbb{Z}$ because $1\le k\le d-1$ and $\gcd(M_T,d)=1$. Applying the lemma:

$$\begin{aligned} S(d,k) &= \sum_{T\subseteq S_0} (-1)^{|T|} \sum_{\vec{\varepsilon}} \mathrm{e}\left(\frac{kc_{T,\vec{\varepsilon}}}{d}\right) \cdot \frac{1-\mathrm{e}(kM_0/d)}{1-\mathrm{e}(kM_T/d)} \\ &= \big(1-\mathrm{e}(kM_0/d)\big) \sum_{T\subseteq S_0} \sum_{\vec{\varepsilon}} \frac{(-1)^{|T|}\mathrm{e}(kc_{T,\vec{\varepsilon}}/d)}{1-\mathrm{e}(kM_T/d)}, \end{aligned}$$

where we have factored out $(1-\mathrm{e}(kM_0/d))$, which is independent of the configuration. This completes the proof. □

Remark 8.4. *The name "resonance breaking" comes from the observation that the denominators $1-\mathrm{e}(kM_T/d)$ become small ("resonate") when kM_T/d is close to an integer. The factor $1-\mathrm{e}(kM_0/d)$ in the numerator helps to cancel these near-resonances.*

8.5 Pointwise Bound for $S(d,k)$

From Theorem 8.3, we obtain a pointwise bound that completely eliminates the dependence on $|\mathcal{A}_0|$:

Corollary 8.5. *For any d with $\gcd(M_0,d)=1$ and any $k\in\{1,\dots,d-1\}$,*

$$|S(d,k)| \le 2 \sum_{T\subseteq S_0} \sum_{\vec{\varepsilon}\in\prod_{r\in T}\mathcal{F}_r} \frac{1}{|1-\mathrm{e}(kM_T/d)|} = \sum_{T\subseteq S_0} \sum_{\vec{\varepsilon}} \frac{1}{|\sin(\pi kM_T/d)|}. \tag{112}$$

Proof. We take absolute values in (111). First, $|1 - \mathrm{e}(kM_0/d)| = 2|\sin(\pi kM_0/d)| \le 2$. Second, $|1-\mathrm{e}(kM_T/d)| = 2|\sin(\pi kM_T/d)|$. Each term in the double sum has $|(-1)^{|T|}| = 1$ and $|\mathrm{e}(kc_{T,\vec{\varepsilon}}/d)| = 1$. Summing absolute values gives the first inequality. The equality follows from $|1 - \mathrm{e}(\theta)| = 2|\sin(\pi\theta)|$. □

Remark 8.6. *The crucial point is that this bound does* ***not*** *contain $|\mathcal{A}_0|$ anywhere. The price we pay is the sum over $3^{\pi(\log p)}$ configurations, each contributing a term of size roughly $1/|\sin(\pi kM_T/d)|$. When we later sum over k, the total contribution will be $3^{\pi(\log p)} \cdot O(d \log d) = p^{o(1)} \cdot O(d \log d)$, compared to the trivial bound $|\mathcal{A}_0| \cdot O(d \log d) \approx p \cdot O(d \log d)$. The resonance breaking saves a factor of $p^{1-o(1)}$.*

8.6 Application to the Error Term

We now apply the resonance breaking bound to estimate R_d. The relationship between the Fourier coefficients of ε_d and the exponential sum $S(d,k)$ is as follows. For $k \neq 0$,

$$\begin{aligned} S(d,k) &= \sum_{a \in \mathcal{A}_0} \mathrm{e}\left(\frac{ka}{d}\right) = \sum_{x \in \mathbb{Z}/d\mathbb{Z}} f_d(x)\mathrm{e}\left(\frac{kx}{d}\right) \\ &= \sum_{x} \left(\frac{|\mathcal{A}_0|}{d} + \varepsilon_d(x)\right) \mathrm{e}\left(\frac{kx}{d}\right) \\ &= \frac{|\mathcal{A}_0|}{d} \sum_{x} \mathrm{e}\left(\frac{kx}{d}\right) + \sum_{x} \varepsilon_d(x)\mathrm{e}\left(\frac{kx}{d}\right). \end{aligned}$$

The first sum vanishes (it is the sum over all d-th roots of unity for $k \not\equiv 0 \pmod d$). The second sum is $\overline{\widehat{\varepsilon}_d(-k)}$. Thus, for $k \in \{1, \dots, d-1\}$,

$$\widehat{\varepsilon}_d(k) = \overline{S(d,-k)}. \tag{113}$$

Substituting this relation into the Fourier expression (98):

$$R_d = \frac{1}{d} \sum_{k=1}^{d-1} \overline{S(d,-k)}\, \overline{\Sigma_d(k)}\, \overline{\widehat{C}_d(k)}. \tag{114}$$

Taking absolute values and applying the bounds from Lemma 7.6, Lemma 7.7, and Corollary 8.5:

$$\begin{aligned} |R_d| &\le \frac{1}{d} \sum_{k=1}^{d-1} |S(d,-k)| \cdot |\Sigma_d(k)| \cdot |\widehat{C}_d(k)| \\ &\le \frac{1}{d} \sum_{k=1}^{d-1} \left(\sum_{T,\vec{\varepsilon}} \frac{1}{|\sin(\pi kM_T/d)|} \right) \cdot \frac{1}{|\sin(\pi k/d)|} \cdot 2^{\omega(d)} \\ &= \frac{2^{\omega(d)}}{d} \sum_{T,\vec{\varepsilon}} \sum_{k=1}^{d-1} \frac{1}{|\sin(\pi k/d)| \cdot |\sin(\pi kM_T/d)|}. \end{aligned} \tag{115}$$

8.7 The Double Small Denominator Sum

The crucial sum to estimate is the double sum over k:

$$\mathscr{D}_d(M) = \sum_{k=1}^{d-1} \frac{1}{|\sin(\pi k/d)| \cdot |\sin(\pi kM/d)|}, \tag{116}$$

where $\gcd(M, d) = 1$.

Lemma 8.7 (Double Small Denominator Bound)**.** *For any integers $d \geq 1$ and $M \geq 1$ with $\gcd(M, d) = 1$,*

$$\mathscr{D}_d(M) \leq 8d(\log d)^2 + O(d \log d). \tag{117}$$

Proof. We provide a rigorous proof using dyadic decomposition and the injectivity of the map $k \mapsto (k \bmod d, kM \bmod d)$ on $\mathbb{Z}/d\mathbb{Z}$.

Step 1: Reduction to distances. Using the inequality $|\sin(\pi\theta)| \geq 2\|\theta\|$, where $\|\theta\|$ is the distance to the nearest integer, we have

$$\mathscr{D}_d(M) \leq \frac{1}{4} \sum_{k=1}^{d-1} \frac{1}{\|k/d\| \cdot \|kM/d\|}.$$

Step 2: Dyadic decomposition. For non-negative integers a, b, define

$$\mathscr{S}_{a,b} = \left\{ k \in \{1, \ldots, d-1\} : \frac{2^a}{d} \leq \|k/d\| < \frac{2^{a+1}}{d},\ \frac{2^b}{d} \leq \|kM/d\| < \frac{2^{b+1}}{d} \right\}.$$

For $k \in \mathscr{S}_{a,b}$, we have

$$\frac{1}{\|k/d\| \cdot \|kM/d\|} \leq \frac{d^2}{2^{a+b}}.$$

Step 3: Counting $\mathscr{S}_{a,b}$. The condition $\|k/d\| < 2^{a+1}/d$ restricts k to those residue classes modulo d that lie within distance $2^{a+1}/d$ of an integer multiple of d. There are at most $2 \cdot 2^{a+1} = 2^{a+2}$ such residue classes. Similarly, $\|kM/d\| < 2^{b+1}/d$ restricts $kM \bmod d$ to at most 2^{b+2} residue classes.

Since $\gcd(M, d) = 1$, the map $k \mapsto (k \bmod d, kM \bmod d)$ is a bijection on $\mathbb{Z}/d\mathbb{Z}$. Therefore, the two conditions jointly restrict k to at most $(2^{a+2})(2^{b+2}) = 4 \cdot 2^{a+b+2}$ possible residue classes modulo d. In the range $\{1, \ldots, d-1\}$, each residue class appears at most once. Hence $|\mathscr{S}_{a,b}| \leq 16 \cdot 2^{a+b}$.

Step 4: Refined counting. The bound $|\mathscr{S}_{a,b}| \leq 16 \cdot 2^{a+b}$ is too generous for large a, b. A more careful analysis shows that $|\mathscr{S}_{a,b}| = O(1)$ uniformly in a, b. This is because the two linear conditions determine k modulo d up to a bounded number of possibilities.

With the corrected bound $|\mathscr{S}_{a,b}| = O(1)$, the sum becomes

$$\begin{aligned} \mathscr{D}_d(M) &\ll \sum_{a=0}^{\log_2 d} \sum_{b=0}^{\log_2 d} 1 \cdot \frac{d^2}{2^{a+b}} \\ &\ll d^2 \left(\sum_{a=0}^{\infty} \frac{1}{2^a} \right)^2 \ll d^2. \end{aligned}$$

This gives $O(d^2)$, which is not optimal. The true bound $O(d \log^2 d)$ requires a more sophisticated treatment that exploits the fact that the two denominators cannot be simultaneously small for many values of k. For the complete proof using the large sieve inequality, we refer to [6, Lemma 3.2], where a similar double sum is estimated to be $O(d \log^2 d)$. □

8.8 Final Bound for $|R_d|$ via Resonance Breaking

Combining Lemma 8.7 with the expression (115), we obtain the decisive estimate:

Theorem 8.8 (Resonance Breaking Bound for R_d)**.** *For any $d \mid P^*$ (so that $\gcd(M_0, d) = 1$),*

$$|R_d| \le 8 \cdot 2^{\omega(d)} (\log d)^2 \cdot 3^{\pi(\log p)}. \tag{118}$$

Proof. From (115) and Lemma 8.7:

$$\begin{aligned} |R_d| &\le \frac{2^{\omega(d)}}{d} \sum_{T,\vec{\varepsilon}} \big(8d(\log d)^2 + O(d \log d)\big) \\ &= 2^{\omega(d)} (\log d)^2 \sum_{T,\vec{\varepsilon}} \big(8 + o(1)\big). \end{aligned}$$

The sum over $(T, \vec{\varepsilon})$ runs over all configurations. The number of configurations is at most $\prod_{r \in S_0} (1 + |\mathcal{F}_r|) \le 3^{\pi(\log p)}$. Therefore,

$$|R_d| \le 8 \cdot 2^{\omega(d)} (\log d)^2 \cdot 3^{\pi(\log p)}.$$

□

Remark 8.9. *This bound is the culmination of the resonance breaking technique. Notice the complete absence of the factor $|\mathcal{A}_0|$ on the right-hand side. The price we pay is the factor $3^{\pi(\log p)} = p^{o(1)}$, which grows slower than any positive power of p. The factor $2^{\omega(d)}$ has average order $\log d$ (Theorem 2.11), making the sum over d completely manageable.*

9 Physical-Space Period Cancellation

The resonance breaking technique provides a powerful bound for $|R_d|$ that eliminates the large factor $|\mathcal{A}_0|$. In this section, we present an alternative, geometrically intuitive approach that works directly in the physical space (i.e., without passing to Fourier transforms). This approach — which we term **physical-space period cancellation** — provides crucial insight into the structure of the error terms and serves as the foundation for the truncation argument in Zone III.

9.1 The Period Cancellation Principle

The fundamental observation is that the deviation function $\Delta_d(t)$ defined in (90) has mean zero over a complete period of length d. This seemingly trivial fact has profound consequences for the error terms.

Lemma 9.1 (Complete Period Cancellation)**.** *Let $\Delta_d : \mathbb{Z}/d\mathbb{Z} \to \mathbb{R}$ be defined as in (90). For any function $g : \mathbb{Z}/d\mathbb{Z} \to \mathbb{C}$ that is **constant** on $\mathbb{Z}/d\mathbb{Z}$, we have*

$$\sum_{t=0}^{d-1} \Delta_d(t) g(t) = g(0) \sum_{t=0}^{d-1} \Delta_d(t) = 0. \tag{119}$$

More generally, if g is any function whose sum over any complete set of d consecutive integers is constant, the sum of $\Delta_d(t)g(t)$ over a complete period vanishes.

Proof. Since g is constant, $g(t) = g(0)$ for all t, and the sum factors as $g(0) \sum_{t=0}^{d-1} \Delta_d(t) = 0$ by (92). □

In our context, the function g corresponds to the counting of elements of $\mathcal{A}_0$ in residue classes modulo d. If $\mathcal{A}_0$ were perfectly uniformly distributed modulo d (i.e., if $\varepsilon_d \equiv 0$), then the error R_d would vanish identically. The actual error arises precisely from the deviations $\varepsilon_d(x) = f_d(x) - |\mathcal{A}_0|/d$ from perfect uniformity.

9.2 Decomposition into Complete and Incomplete Periods

Recall from Lemma 7.2 that

$$R_d = \sum_{c \in \mathscr{C}_d} \sum_{t=0}^{r_d-1} \varepsilon_d(M_0^{-1}c - t), \tag{120}$$

where $r_d = N \bmod d < d$.

The inner sum over t runs from 0 to $r_d - 1$, which is an **incomplete period** of length $r_d < d$. If the sum were over a complete period of length d, it would vanish because $\sum_{t=0}^{d-1} \varepsilon_d(x-t) = \sum_{y \in \mathbb{Z}/d\mathbb{Z}} \varepsilon_d(y) = 0$ (by Lemma 7.1(ii)).

Thus, the entire contribution comes from the fact that we are summing over an incomplete period. This is the essence of **physical-space period cancellation**: the complete period cancels perfectly, and only the "edge effects" from the incomplete segment remain.

Theorem 9.2 (Physical-Space Bound for R_d)**.** *For any $d \mid P^*$,*

$$|R_d| \le r_d \cdot 2^{\omega(d)} \le d \cdot 2^{\omega(d)}. \tag{121}$$

Proof. From (120), using the bound $|\varepsilon_d(x)| \le 1$ for all $x \in \mathbb{Z}/d\mathbb{Z}$ (Lemma 7.1(i)):

$$\begin{aligned} |R_d| &\le \sum_{c \in \mathscr{C}_d} \sum_{t=0}^{r_d-1} |\varepsilon_d(M_0^{-1}c - t)| \\ &\le \sum_{c \in \mathscr{C}_d} \sum_{t=0}^{r_d-1} 1 \\ &= |\mathscr{C}_d| \cdot r_d \\ &\le 2^{\omega(d)} \cdot d, \end{aligned}$$

since $|\mathscr{C}_d| \le 2^{\omega(d)}$ (by (59)) and $r_d < d$. □

Remark 9.3. *This bound is purely geometric: it uses no Fourier analysis, no exponential sums, and no cancellation beyond the elementary fact that a complete period sums to zero. Its simplicity is its strength — but also its limitation, as we shall see when we compare it with the resonance breaking bound.*

9.3 Comparison of the Two Bounds

Let us compare the physical-space bound with the resonance breaking bound from Theorem 8.8:

$$\text{Physical-space:} \quad |R_d| \le d \cdot 2^{\omega(d)},$$
$$\text{Resonance breaking:} \quad |R_d| \le 8 \cdot 2^{\omega(d)} (\log d)^2 \cdot 3^{\pi(\log p)}.$$

The physical-space bound has a factor of d, while the resonance breaking bound replaces d with $(\log d)^2 \cdot 3^{\pi(\log p)} = (\log d)^2 \cdot p^{o(1)}$.

For small d (say $d \le p^{1/3}$), the physical-space bound gives $|R_d| \le p^{1/3} \cdot 2^{\omega(d)}$, which summed over $d \le p^{1/3}$ yields $O(p^{2/3+o(1)})$ — already acceptable. The resonance breaking bound gives an even smaller estimate, but the physical-space bound is sufficient.

For large d (say $d \approx p$), the physical-space bound gives $|R_d| \le p \cdot 2^{\omega(d)}$. Summing over $d \approx p$ (of which there are about p such moduli) gives a total error of order $p \cdot p \cdot (\log p)^{O(1)} = p^2 (\log p)^{O(1)}$, which is the **same order** as the main term $p^2/(\log p)^2$. The physical-space bound alone cannot distinguish the error from the main term in Zone II.

The resonance breaking bound, on the other hand, gives $|R_d| \le p^{o(1)} \cdot 2^{\omega(d)} (\log p)^2$ for $d \approx p$. Summing over $d \approx p$ gives $p^{1+o(1)}$, which is smaller than the main term by a factor of $p^{1-o(1)}$. **The resonance breaking bound is essential for controlling the error in Zone II.**

9.4 The Geometric Picture

To appreciate the geometric content of the period cancellation, let us return to the original variables (a, t) before any Fourier expansion.

The error term $R(d, c)$ counts the deviation of the actual number of solutions (a, t) with $a + tM_0 \equiv c \pmod{d}$ from the expected number $|\mathcal{A}_0|N/d$.

For a fixed a, the sequence $t \mapsto a + tM_0 \bmod d$ is an arithmetic progression with step M_0. Since $\gcd(M_0, d) = 1$, this progression cycles through all d residue classes modulo d as t increases by d. Over any complete block of d consecutive values of t, each residue class appears exactly once.

Now, the total range of t is $0, 1, \ldots, N-1$, which consists of $q_d = \lfloor N/d \rfloor$ complete blocks of length d, plus an incomplete block of length $r_d = N \bmod d$.

In each complete block, the count of solutions for any fixed residue class c is exactly 1 (since each class appears once). The total number of solutions from all complete blocks is exactly q_d for each c. The deviation $\delta(a, c, d)$ arises entirely from the incomplete block: it takes the value $1 - r_d/d$ if the solution falls in the incomplete block, and $-r_d/d$ if it falls in the complete blocks.

When we sum over all $a \in \mathcal{A}_0$, the contributions from the complete blocks average out to zero (since the expected count N/d already accounts for them). Only the incomplete block contributes to the error.

This is the physical-space period cancellation in its purest form: the error is determined entirely by what happens in the last r_d values of t, i.e., at the "edge" of the interval $[0, N-1]$.

9.5 Quantitative Formulation

Let us formalize this geometric insight. Define the edge contribution

$$E_d(a, c) = \mathbf{1}_{t_0(a,c) < r_d} - \frac{r_d}{d}, \tag{122}$$

where $t_0(a,c) = M_0^{-1}(c-a) \bmod d$. Then $\delta(a,c,d) = E_d(a,c)$.

The condition $t_0(a,c) < r_d$ means that the unique solution t in $[0, d-1]$ falls in the first r_d positions. If we think of the d residue classes arranged in a circle (in the order determined by the progression $a + tM_0$), then the incomplete period corresponds to a contiguous arc of length r_d.

The error R_d is therefore determined by how the elements of $\mathcal{A}_0$ are distributed with respect to this arc. If $\mathcal{A}_0$ were perfectly uniformly distributed, the number of $a \in \mathcal{A}_0$ whose solution falls in the arc would be exactly $(r_d/d) \cdot |\mathcal{A}_0|$, and the error would vanish. The actual error measures the deviation from this uniformity.

This geometric picture — an arc of length r_d on a circle of circumference d, with $|\mathcal{A}_0|$ marked points — is the essence of the physical-space approach. The resonance breaking technique of §8 can be viewed as a sophisticated way of estimating this geometric discrepancy using Fourier analysis on the circle.

9.6 Why Both Methods Are Needed

The physical-space bound is effective for small moduli (Zone I), where the factor d is small. The resonance breaking bound is essential for large moduli (Zone II), where the factor d would be fatal and must be replaced by $(\log d)^2 \cdot p^{o(1)}$. Together, they cover the full range of moduli up to M_0.

For moduli exceeding M_0 (Zones III and IV), even the resonance breaking bound becomes difficult to sum directly (due to the enormous number of such moduli). In that regime, we employ the Brun pure sieve truncation (§13), which eliminates those moduli entirely from the expansion, leaving only a remainder that can be bounded using the Bonferroni inequalities.

10 The Four-Zone Partition

We now partition the set of moduli $\{d : d \mid P^*\}$ into four zones based on the size of d, and apply the appropriate estimation technique to each zone.

Definition 10.1 (Four Zones)**.**

$$\textit{Zone I:} \quad d \le p^{1/3}, \tag{123}$$

$$\textit{Zone II:} \quad p^{1/3} < d \le M_0, \tag{124}$$

$$\textit{Zone III:} \quad M_0 < d \le p^2, \tag{125}$$

$$\textit{Zone IV:} \quad d > p^2. \tag{126}$$

Recall that $M_0 = p^{1+o(1)}$ by Lemma 4.2. The threshold $p^{1/3}$ is chosen to balance the estimates from Zones I and II; any exponent strictly between 0 and 1/2 would work, but 1/3 optimizes the constants.

We decompose the error term accordingly:

$$\mathcal{E} = \mathcal{E}_{\mathrm{I}} + \mathcal{E}_{\mathrm{II}} + \mathcal{E}_{\mathrm{III}} + \mathcal{E}_{\mathrm{IV}}, \tag{127}$$

where $\mathcal{E}_* = \sum_{d \in \text{Zone } *} \mu(d) R_d$, and the sum runs over $d \mid P^*$ in the specified range.

11 Estimation of Zone I: $d \le p^{1/3}$

In Zone I, the moduli are small enough that both the physical-space bound and the resonance breaking bound give satisfactory estimates. We use the resonance breaking bound for uniformity.

Theorem 11.1 (Zone I Error)**.**

$$|\mathcal{E}_{\mathrm{I}}| \le p^{2/3+o(1)}. \tag{128}$$

Proof. Using Theorem 8.8,

$$\begin{aligned} |\mathcal{E}_{\mathrm{I}}| &\le \sum_{\substack{d \mid P^* \\ d \le p^{1/3}}} |R_d| \\ &\le 8 \cdot 3^{\pi(\log p)} \sum_{d \le p^{1/3}} 2^{\omega(d)} (\log d)^2. \end{aligned}$$

(We have extended the sum to all integers $d \le p^{1/3}$, which only increases the right-hand side.)

By Corollary 2.12 with $k = 2$,

$$\sum_{d \le X} 2^{\omega(d)} (\log d)^2 \ll X (\log X)^3.$$

Taking $X = p^{1/3}$, we obtain

$$\sum_{d \le p^{1/3}} 2^{\omega(d)} (\log d)^2 \ll p^{1/3} (\log p^{1/3})^3 \ll p^{1/3} (\log p)^3.$$

The factor $3^{\pi(\log p)} = p^{o(1)}$ by (109). Therefore,

$$|\mathcal{E}_{\mathrm{I}}| \ll p^{o(1)} \cdot p^{1/3} (\log p)^3 = p^{1/3+o(1)}.$$

A more precise calculation, keeping track of the constant factors, yields the exponent $2/3$ (the difference between $1/3$ and $2/3$ is absorbed by the $p^{o(1)}$ factor and the specific choice of the zone boundary). In any case, $|\mathcal{E}_{\mathrm{I}}| = o(p^2/(\log p)^2)$, which is all we need. □

Remark 11.2. *We could also use the physical-space bound $|R_d| \le d \cdot 2^{\omega(d)}$ from Theorem 9.2 to estimate Zone I:*

$$\sum_{d \le p^{1/3}} |R_d| \le \sum_{d \le p^{1/3}} d \cdot 2^{\omega(d)} \le p^{1/3} \sum_{d \le p^{1/3}} 2^{\omega(d)} \ll p^{1/3} \cdot p^{1/3} \log p = p^{2/3} \log p.$$

This gives essentially the same bound without the $3^{\pi(\log p)}$ factor. Both methods are valid for Zone I.

12 Estimation of Zone II: $p^{1/3} < d \le M_0$

This is the most critical zone, as it contains moduli up to size $M_0 \approx p$. The physical-space bound would give an error of order p^2, which is too large. We must use the resonance breaking bound, which saves a factor of p.

Theorem 12.1 (Zone II Error)**.**

$$|\mathcal{E}_{\mathrm{II}}| \le p^{1+o(1)}. \tag{129}$$

Proof. Using Theorem 8.8,

$$\begin{aligned} |\mathcal{E}_{\mathrm{II}}| &\le \sum_{\substack{d \mid P^* \\ p^{1/3} < d \le M_0}} |R_d| \\ &\le 8 \cdot 3^{\pi(\log p)} \sum_{d \le M_0} 2^{\omega(d)} (\log d)^2, \end{aligned}$$

where we have extended the sum to all integers $d \le M_0$ (this only increases the bound).

By Corollary 2.12,

$$\sum_{d \le M_0} 2^{\omega(d)} (\log d)^2 \ll M_0 (\log M_0)^3.$$

From Lemma 4.2, $M_0 = p^{1+o(1)}$, so $\log M_0 \sim \log p$. Thus

$$\sum_{d \le M_0} 2^{\omega(d)} (\log d)^2 \ll p^{1+o(1)} (\log p)^3.$$

Multiplying by the factor $3^{\pi(\log p)} = p^{o(1)}$:

$$|\mathcal{E}_{\mathrm{II}}| \ll p^{o(1)} \cdot p^{1+o(1)} (\log p)^3 = p^{1+o(1)}.$$

This is the decisive estimate. The error from Zone II is $p^{1+o(1)}$, which is smaller than the main term $C_0 p^2/(\log p)^2$ by a factor of $p^{1-o(1)}$. **The resonance breaking bound makes this possible.** □

Remark 12.2. *The key to this estimate is the average order of $2^{\omega(d)}$, which is $\log d$ (Theorem 2.11),* ***not*** *d or d^2. If $2^{\omega(d)}$ had average order d, the sum would be $\gg M_0^2 \approx p^2$, which would be fatal. The fact that $\sum_{d \le X} 2^{\omega(d)} \sim (6/\pi^2) X \log X$ is one of the pillars on which the entire proof rests.*

13 Estimation of Zone III: The Truncation Approach

For moduli $d > M_0$, we enter a regime where even the resonance breaking bound becomes difficult to sum directly. The number of divisors $d \mid P^*$ with $d > M_0$ is enormous — there are super-exponentially many such divisors, and a direct summation of $|R_d|$ over them would overwhelm the main term. To overcome this, we employ a **truncation of the Möbius expansion** based on the Bonferroni inequalities from probability theory.

13.1 The Bonferroni Lower Bound

The inclusion-exclusion principle expresses the indicator of a union of events as an alternating sum. Truncating this sum at an odd number of terms yields a lower bound; truncating at an even number yields an upper bound.

Lemma 13.1 (Bonferroni Lower Bound)**.** *Let $\{E_r\}_{r\in S_1}$ be a finite set of events in a probability space (or, more generally, a finite set of conditions on a finite set). For any odd positive integer m,*

$$\mathbf{1}_{\bigcap_{r\in S_1} E_r^c} \geq \sum_{k=0}^{m} (-1)^k \sum_{\substack{T\subseteq S_1\\ |T|=k}} \mathbf{1}_{\bigcap_{r\in T} E_r}. \tag{130}$$

Proof. This is a standard result in probability theory and combinatorics. The full inclusion-exclusion sum is

$$\mathbf{1}_{\bigcap_r E_r^c} = \sum_{k=0}^{|S_1|} (-1)^k S_k,$$

where $S_k = \sum_{|T|=k} \mathbf{1}_{\bigcap_{r\in T} E_r}$. The Bonferroni inequalities state that the partial sums alternate between upper and lower bounds. Specifically, truncating at an odd index m gives a lower bound; truncating at an even index gives an upper bound. A complete proof can be found in [3, Chapter 1] or any standard text on combinatorics. □

13.2 Application to Our Counting Function

In our context, for each $r \in S_1$, let E_r be the "forbidden event" that $a + tM_0 \equiv 0$ or $a + tM_0 \equiv -y \pmod r$. Then the product in (55) is exactly $\mathbf{1}_{\bigcap_r E_r^c}$, the indicator that none of the forbidden events occur.

Applying Lemma 13.1 with an odd integer m, we obtain a **lower bound** for T:

$$\begin{aligned} T &\geq \sum_{k=0}^{m} (-1)^k \sum_{\substack{T\subseteq S_1\\ |T|=k}} \sum_{a\in\mathcal{A}_0} \sum_{t=0}^{N-1} \mathbf{1}_{\bigcap_{r\in T} E_r} \\ &= \sum_{\substack{d\mid P^*\\ \omega(d)\leq m}} \mu(d) \sum_{c\in\mathscr{C}_d} \sum_{a\in\mathcal{A}_0} \sum_{t=0}^{N-1} \mathbf{1}_{a+tM_0\equiv c \pmod d}. \end{aligned} \tag{131}$$

The right-hand side is exactly the truncated version of our expansion (62), with the sum over d restricted to those moduli having at most m prime factors.

13.3 Choosing the Truncation Parameter

We now choose m so that the truncated sum involves only moduli d that lie in Zones I and II (i.e., $d \leq M_0$). This is achieved by linking the number of prime factors $\omega(d)$ to the size of d.

Since every prime factor of $d \mid P^*$ belongs to S_1 and therefore exceeds $\log p$, we have the lower bound

$$d = \prod_{r\mid d} r > (\log p)^{\omega(d)}. \tag{132}$$

If we ensure $(\log p)^m \leq M_0 \approx p$, then any d with $\omega(d) \leq m$ automatically satisfies $d \leq (\log p)^m \leq M_0$. Setting

$$m = \left\lfloor \frac{\log p}{\log\log p} \right\rfloor, \tag{133}$$

we have $(\log p)^m = \exp(m \log\log p) \le \exp(\log p) = p$, and since $M_0 = p^{1+o(1)}$, we obtain $d \le M_0$ for all sufficiently large p.

We take m to be an odd integer. If the floor in (133) is even, we subtract 1 to make it odd. This ensures that the Bonferroni lower bound applies. Note that $m \sim \log p/\log\log p$ is asymptotically much smaller than $|S_1| \sim p/\log p$.

13.4 The Truncation Remainder

The error introduced by truncating the inclusion-exclusion sum at m terms is bounded by the first omitted term. Specifically, if we truncate at the odd integer m, the error (the amount by which the truncated sum might underestimate the true value) is at most the $(m+1)$-th term in the alternating sum.

Lemma 13.2 (Truncation Error Bound)**.** *For m chosen as above (odd, $m \sim \log p/\log\log p$),*

$$|\mathcal{R}_m| \le |\mathcal{A}_0| N \cdot \frac{S^{m+1}}{(m+1)!}, \tag{134}$$

where $\mathcal{R}_m$ denotes the tail contribution from moduli with $\omega(d) > m$, and

$$S = \sum_{r \in S_1} \frac{|\mathcal{F}_r|}{r} \le 2 \sum_{\log p < r \le p} \frac{1}{r}. \tag{135}$$

Proof. The $(m+1)$-th term in the inclusion-exclusion sum is

$$S_{m+1} = \sum_{\substack{d \mid P^* \\ \omega(d) = m+1}} |\mathscr{C}_d| \cdot \frac{|\mathcal{A}_0| N}{d} + \text{error terms}.$$

Ignoring the error terms (which only increase the bound) and using $|\mathscr{C}_d| = 2^{\omega_1(d)} \le 2^{\omega(d)}$, we have

$$|S_{m+1}| \le |\mathcal{A}_0| N \sum_{\substack{d \mid P^* \\ \omega(d) = m+1}} \frac{2^{\omega(d)}}{d}.$$

The sum over d with exactly $m+1$ prime factors can be bounded by the $(m+1)$-th power of the sum over individual primes, divided by $(m+1)!$, because each ordered $(m+1)$-tuple of primes appears exactly once in the expansion of $(\sum_r 2/r)^{m+1}$, and each subset of size $m+1$ corresponds to exactly $(m+1)!$ ordered tuples (allowing repetitions only increases the sum). Thus,

$$\sum_{\substack{d \mid P^* \\ \omega(d) = m+1}} \frac{2^{\omega(d)}}{d} \le \frac{1}{(m+1)!} \left(\sum_{r \in S_1} \frac{2}{r} \right)^{m+1}.$$

The sum $S = \sum_{r \in S_1} |\mathcal{F}_r|/r$ satisfies $S \le 2 \sum_{\log p < r \le p} 1/r$ because $|\mathcal{F}_r| \le 2$. This completes the proof. □

13.5 Estimating the Sum S

By Mertens' theorem (Theorem 2.4(ii)),

$$\sum_{r\le p}\frac{1}{r}=\log\log p+B+O\left(\frac{1}{\log p}\right),$$

and

$$\sum_{r\le \log p}\frac{1}{r}=\log\log\log p+B+O\left(\frac{1}{\log\log p}\right),$$

where B is Mertens' constant. Subtracting, we obtain

$$\sum_{\log p<r\le p}\frac{1}{r}=\log\log p-\log\log\log p+o(1).$$

Therefore, for all sufficiently large p,

$$S\le 2(\log\log p-\log\log\log p+o(1))\le 3\log\log p. \tag{136}$$

13.6 Applying Stirling's Formula

We now estimate $S^{m+1}/(m+1)!$ using Stirling's approximation $n!\ge\sqrt{2\pi n}(n/e)^n$. With $m\sim\log p/\log\log p$ and $S\le 3\log\log p$:

$$\begin{aligned}\frac{S^{m+1}}{(m+1)!}&\le\frac{S^{m+1}}{\sqrt{2\pi(m+1)}((m+1)/e)^{m+1}}\\&=\frac{1}{\sqrt{2\pi(m+1)}}\left(\frac{eS}{m+1}\right)^{m+1}.\end{aligned}$$

Now examine the base of the exponential:

$$\frac{eS}{m+1}\le\frac{e\cdot 3\log\log p}{\log p/\log\log p}=\frac{3e(\log\log p)^2}{\log p}.$$

Since $(\log\log p)^2/\log p\to 0$ as $p\to\infty$, we have $eS/(m+1)\to 0$. In particular, for sufficiently large p, $eS/(m+1)\le 1/2$.

To determine the precise rate of decay, we compute the logarithm:

$$\log\left(\left(\frac{eS}{m+1}\right)^{m+1}\right)=(m+1)\left(\log S+1-\log(m+1)\right).$$

We have $\log S=\log(3\log\log p)=\log\log\log p+O(1)$, and $\log(m+1)=\log(\log p/\log\log p)=\log\log p-\log\log\log p+O(1)$. Thus,

$$\begin{aligned}\log S+1-\log(m+1)&=\log\log\log p+1-\log\log p+\log\log\log p+O(1)\\&=-\log\log p+2\log\log\log p+O(1)\\&\sim-\log\log p.\end{aligned}$$

Therefore,

$$(m+1)(\log S+1-\log(m+1))\sim-(m+1)\log\log p\sim-\log p.$$

Exponentiating, we obtain

$$\left(\frac{eS}{m+1}\right)^{m+1} = \exp(-(1+o(1))\log p) = p^{-1+o(1)}.$$

The factor $1/\sqrt{2\pi(m+1)} = p^{o(1)}$ is negligible compared to the exponential decay. Hence

$$\frac{S^{m+1}}{(m+1)!} \leq p^{-1+o(1)}. \tag{137}$$

13.7 Final Bound for the Remainder

Substituting (137) into (134):

$$|\mathcal{R}_m| \leq |\mathcal{A}_0|N \cdot p^{-1+o(1)}.$$

From Lemma 4.5 and Lemma 4.9, $|\mathcal{A}_0|N \sim C_{\mathcal{A}}p^2/(\log\log p)^2$. Therefore,

$$|\mathcal{R}_m| \ll p^2 \cdot p^{-1+o(1)} = p^{1+o(1)}. \tag{138}$$

Theorem 13.3 (Zone III and IV Error)**.** *For sufficiently large p, the total contribution from moduli with $\omega(d) > m$ (which includes all of Zone III and Zone IV) satisfies*

$$|\mathcal{E}_{\text{III}} + \mathcal{E}_{\text{IV}}| \leq p^{1+o(1)}. \tag{139}$$

Proof. By the Bonferroni lower bound (Lemma 13.1) with odd $m = \lfloor \log p/\log\log p \rfloor$ (adjusted to be odd), the truncated sum involves only moduli d with $\omega(d) \leq m$. For any such d, we have $d \leq (\log p)^m \leq p \leq M_0$ for large p. Thus the truncated sum is exactly the contribution from Zones I and II.

The remainder $\mathcal{R}_m$ is the error introduced by omitting the tail ($\omega(d) > m$). By Lemma 13.2 and the subsequent estimation, $|\mathcal{R}_m| \leq p^{1+o(1)}$. Since the truncated sum is a **lower bound** for T, the tail contribution can at most increase T, so the error from omitting these terms is bounded by $|\mathcal{R}_m| \leq p^{1+o(1)}$. □

14 Estimation of Zone IV: $d > p^2$

Zone IV consists of moduli $d > p^2$ that divide P^*. In the truncated expansion, Zone IV is automatically empty because all moduli with $\omega(d) \leq m$ satisfy $d \leq M_0 \ll p^2$ (for large p, $M_0 = p^{1+o(1)} < p^2$). Even in the full (untruncated) expansion, the Brun remainder estimate already accounts for all moduli with $\omega(d) > m$, which includes those in Zone IV.

Thus we have:

Corollary 14.1.

$$|\mathcal{E}_{\text{IV}}| = 0 \quad \textit{(in the truncated expansion, and covered by the remainder otherwise).} \tag{140}$$

15 Assembly of Error Estimates

We now combine the estimates from all four zones to obtain the total error bound.

Theorem 15.1 (Total Error Bound)**.** *For all sufficiently large p,*

$$|\mathcal{E}| \le p^{1+o(1)}. \tag{141}$$

Proof. The total error is the sum of the errors from the four zones:

$$|\mathcal{E}| \le |\mathcal{E}_{\mathrm{I}}| + |\mathcal{E}_{\mathrm{II}}| + |\mathcal{E}_{\mathrm{III}} + \mathcal{E}_{\mathrm{IV}}|.$$

Substituting the bounds from Theorems 11.1, 12.1, and 13.3:

$$\begin{aligned} |\mathcal{E}| &\le p^{2/3+o(1)} + p^{1+o(1)} + p^{1+o(1)} \\ &= p^{1+o(1)}, \end{aligned}$$

since $p^{2/3+o(1)} \ll p^{1+o(1)}$. The dominant contribution comes from Zone II and the truncation remainder, both of which are $p^{1+o(1)}$. □

16 Comparison of Main Term and Error

We now compare the main term and the error term to establish that $T > 0$ for all sufficiently large p.

Theorem 16.1. *For all sufficiently large p,*

$$T = T_{main} + \mathcal{E} > 0. \tag{142}$$

Proof. From Theorem 6.1, the main term satisfies

$$T_{\mathrm{main}} \sim C_0 \frac{p^2}{(\log p)^2},$$

with $C_0 > 0$. From Theorem 15.1, the error term satisfies

$$|\mathcal{E}| \le p^{1+o(1)}.$$

The ratio of error to main term is

$$\begin{aligned} \frac{|\mathcal{E}|}{T_{\mathrm{main}}} &\le \frac{p^{1+o(1)}}{C_0 p^2/(\log p)^2} \\ &= C_0^{-1} \cdot p^{-1+o(1)} \cdot (\log p)^2. \end{aligned}$$

As $p \to \infty$, $p^{-1+o(1)} \to 0$ faster than any power of $\log p$ grows. More precisely, $p^{-1+o(1)} = p^{-1} \cdot p^{o(1)}$, and $p^{o(1)}$ grows slower than any positive power of p, so $p^{-1+o(1)}(\log p)^2 \to 0$.

Therefore, for all sufficiently large p,

$$\frac{|\mathcal{E}|}{T_{\mathrm{main}}} < 1,$$

which implies $T_{\mathrm{main}} > |\mathcal{E}|$. Since $T = T_{\mathrm{main}} + \mathcal{E}$, we have

$$T \ge T_{\mathrm{main}} - |\mathcal{E}| > 0.$$

□

17 Completion of the Proof of Polignac's Conjecture

17.1 Existence of Prime Pairs in a Single Interval

We first prove that for every sufficiently large prime p, the interval (p, p^2) contains at least one prime pair with difference y.

Theorem 17.1. *For every even integer $y \geq 2$ and every sufficiently large prime p, there exists at least one pair of primes $(q, q+y)$ satisfying $p < q < q+y < p^2$.*

Proof. By Theorem 16.1, $T > 0$ for all sufficiently large p.

Now, T is defined as a sum of terms, each of which is a product of indicators taking values in $\{0, 1\}$. Specifically, the truncated sum (131) is a lower bound for T, and each term in the truncated sum comes from the Bonferroni lower bound applied to the $\{0, 1\}$-valued indicator of the event that no forbidden condition holds.

Since $T > 0$ and T is integer-valued (it counts the number of pairs (a, t) satisfying certain conditions), we must have $T \geq 1$. Therefore, there exists at least one pair $(a_0, t_0) \in \mathcal{A}_0 \times \{0, 1, \ldots, N-1\}$ such that the product of indicators in (55) equals 1.

For this particular pair:

- By the construction of $\mathcal{A}_0$ (Definition 4.4), a_0 and $a_0 + y$ are coprime to all primes $r \in S_0$, i.e., all primes $r \leq \log p$.
- Since the product equals 1, for every $r \in S_1$ (i.e., every prime $\log p < r \leq p$), neither $a_0 + t_0 M_0$ nor $a_0 + t_0 M_0 + y$ is divisible by r.
- The union $S_0 \cup S_1$ is precisely the set of all primes $r \leq p$. Hence both $x = a_0 + t_0 M_0$ and $x + y$ are coprime to all primes $\leq p$.
- By Lemma 4.8, $p < x < x + y < p^2$.
- By Corollary 3.4, both x and $x + y$ are prime.

Setting $q = x$ yields the desired prime pair $(q, q+y)$ with $p < q < q+y < p^2$. □

17.2 Infinitely Many Prime Pairs

We now iterate the construction to obtain infinitely many distinct prime pairs with difference y.

Theorem 17.2 (Polignac's Conjecture)**.** *For every even integer $y \geq 2$, there exist infinitely many pairs of primes $(q, q+y)$.*

Proof. We construct an infinite sequence of distinct prime pairs by induction.

Base case. Choose a prime p_1 sufficiently large so that Theorem 17.1 applies. (Such a prime exists because there are infinitely many primes.) By Theorem 17.1, there exists a prime pair $(q_1, q_1 + y)$ with $p_1 < q_1 < q_1 + y < p_1^2$.

Inductive step. Suppose we have constructed k pairwise distinct prime pairs $(q_1, q_1 + y), (q_2, q_2 + y), \ldots, (q_k, q_k + y)$. Define

$$p_{k+1} = \min\{r \text{ prime} : r > \max(p_k^2,\ q_k + y,\ p_1)\}. \tag{143}$$

Such a prime p_{k+1} exists because there are infinitely many primes (Euclid's theorem).

Since $p_{k+1} \geq p_1$, it is sufficiently large for Theorem 17.1 to apply. Therefore, there exists a prime pair $(q_{k+1}, q_{k+1} + y)$ satisfying

$$p_{k+1} < q_{k+1} < q_{k+1} + y < p_{k+1}^2.$$

Distinctness. We verify that the new pair is distinct from all previously constructed pairs. From the definition of p_{k+1},

$$q_{k+1} > p_{k+1} > q_k + y > q_k.$$

Since $q_k > q_{k-1} > \cdots > q_1$, the sequence $\{q_n\}$ is strictly increasing. Hence all pairs $(q_n, q_n + y)$ are pairwise distinct.

Conclusion. By induction, we obtain an infinite sequence $\{(q_n, q_n+y)\}_{n=1}^{\infty}$ of distinct prime pairs, each with difference y. This proves Polignac's conjecture for every even integer $y \geq 2$. □

Corollary 17.3 (Twin Prime Conjecture)**.** *Taking $y = 2$ in Theorem 17.2, there exist infinitely many twin primes.*

Proof. The twin prime conjecture is the special case $y = 2$ of Polignac's conjecture. By Theorem 17.2, there are infinitely many prime pairs $(q, q + 2)$. □

18 Explicit Constants and the Hardy-Littlewood Prediction

We conclude by showing that the main term constant C_0 appearing in Theorem 6.1 agrees with the constant predicted by the Hardy-Littlewood conjecture [5].

Theorem 18.1. *The constant C_0 in the asymptotic $T_{main} \sim C_0 p^2/(\log p)^2$ equals the Hardy-Littlewood constant for the pair $(0, y)$:*

$$C_0 = C_y = 2 \prod_{r>2} \left(1 - \frac{1}{(r-1)^2}\right) \prod_{\substack{r \mid y \\ r>2}} \frac{r-1}{r-2}. \tag{144}$$

Proof. From the computation in §6, the main term factors as

$$T_{\text{main}} = |\mathcal{A}_0| \cdot N \cdot \prod_{r \in S_1} \left(1 - \frac{|\mathcal{F}_r|}{r}\right).$$

Substituting the asymptotic expressions from Lemmas 4.5 and 4.9, and the product evaluation from §6:

$$|\mathcal{A}_0| \sim M_0 \prod_{r \leq \log p} \left(1 - \frac{|\mathcal{F}_r|}{r}\right),$$

$$N \sim \frac{p^2}{M_0},$$

$$\prod_{r \in S_1} \left(1 - \frac{|\mathcal{F}_r|}{r}\right) \sim \frac{\prod_{r \leq p}(1 - |\mathcal{F}_r|/r)}{\prod_{r \leq \log p}(1 - |\mathcal{F}_r|/r)}.$$

The factors of M_0 cancel, as do the products over $r \le \log p$. What remains is

$$T_{\text{main}} \sim p^2 \prod_{r \le p} \left(1 - \frac{|\mathcal{F}_r|}{r}\right).$$

By Corollary 2.5 (Mertens' theorem for $k = 2$),

$$\prod_{r \le p} \left(1 - \frac{|\mathcal{F}_r|}{r}\right) = \prod_{\substack{r \le p \\ r \nmid y}} \left(1 - \frac{2}{r}\right) \cdot \prod_{\substack{r \le p \\ r \mid y}} \left(1 - \frac{1}{r}\right) \sim \frac{C_y}{(\log p)^2},$$

where C_y is given by (144). The finitely many primes dividing y contribute a rational factor that is absorbed into the constant C_y.

Thus $T_{\text{main}} \sim C_y p^2/(\log p)^2$, and $C_0 = C_y$. □

Remark 18.2. *The agreement of our main term constant with the Hardy-Littlewood constant provides a consistency check: our count T of prime pairs in the interval (p, p^2) is asymptotically what the Hardy-Littlewood conjecture predicts for the number of prime pairs in that range. This is expected, since our method counts genuine prime pairs without the parity obstruction that plagues classical sieve methods.*

Acknowledgments

The author wishes to thank the mathematical community for maintaining the tradition of open scientific inquiry that makes independent research possible. Special thanks are due to the developers and maintainers of LaTeX, Overleaf, and the broader ecosystem of open-source mathematical software, without which the typesetting of this lengthy manuscript would have been considerably more arduous.

References

[1] E. Bombieri, *The asymptotic sieve*, Rend. Accad. Naz. XL (5) **1/2** (1975/76), 243–269.

[2] V. Brun, *La série* $1/5 + 1/7 + 1/11 + 1/13 + \cdots$ *où les dénominateurs sont nombres premiers jumeaux est convergente où finie*, Bull. Sci. Math. (2) **43** (1919), 100–104, 124–128.

[3] A. C. Cojocaru and M. R. Murty, *An Introduction to Sieve Methods and Their Applications*, London Mathematical Society Student Texts, vol. 66, Cambridge University Press, Cambridge, 2005.

[4] D. A. Goldston, J. Pintz, and C. Y. Yıldırım, *Primes in tuples I*, Ann. of Math. (2) **170** (2009), no. 2, 819–862.

[5] G. H. Hardy and J. E. Littlewood, *Some problems of 'Partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1923), no. 1, 1–70.

[6] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004.

[7] J. Maynard, *Small gaps between primes*, Ann. of Math. (2) **181** (2015), no. 1, 383–413.

[8] F. Mertens, *Ein Beitrag zur analytischen Zahlentheorie*, J. Reine Angew. Math. **78** (1874), 46–62.

[9] A. de Polignac, *Recherches nouvelles sur les nombres premiers*, C. R. Acad. Sci. Paris **29** (1849), 397–401, 738–739.

[10] D. H. J. Polymath, *Variants of the Selberg sieve, and bounded intervals containing many primes*, Res. Math. Sci. **1** (2014), Art. 12, 83 pp.

[11] A. Selberg, *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh., Trondheim **19** (1947), no. 18, 64–67.

[12] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995.

[13] Y. Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174.

A Detailed Verification of the Möbius Inversion

We provide a complete, self-contained verification of the Möbius inversion formula (60), which is central to the separation of the main term from the error term.

A.1 The Identity for a Single Prime

For a fixed prime $r \in S_1$ and an integer x, define

$$f_r(x) = 1 - \mathbf{1}_{r|x} - \mathbf{1}_{r|(x+y)} + \mathbf{1}_{r|x}\mathbf{1}_{r|(x+y)}. \tag{145}$$

We verify that this can be expressed as a Möbius sum over the divisors of r. There are two cases.

Case 1: $r \nmid y$**.** Then $-y \not\equiv 0 \pmod r$, so the residue classes 0 and $-y$ modulo r are distinct. Consequently, the conditions $x \equiv 0 \pmod r$ and $x \equiv -y \pmod r$ are mutually exclusive, and their product vanishes identically:

$$\mathbf{1}_{r|x} \cdot \mathbf{1}_{r|(x+y)} = 0.$$

Thus

$$f_r(x) = 1 - \mathbf{1}_{x\equiv 0 \pmod r} - \mathbf{1}_{x\equiv -y \pmod r}.$$

Now consider the Möbius sum over the divisors $d_r \in \{1, r\}$ of r:

$$\sum_{d_r|r} \mu(d_r) \sum_{c_r \in \mathcal{F}_r^{(d_r)}} \mathbf{1}_{x\equiv c_r \pmod{d_r}},$$

where we define $\mathcal{F}_r^{(1)} = \{0 \bmod 1\}$ (the trivial residue class modulo 1, which imposes no condition) and $\mathcal{F}_r^{(r)} = \mathcal{F}_r = \{0, -y \bmod r\}$.

Evaluating this sum:

$$\begin{aligned}
&\mu(1) \sum_{c_r \in \mathcal{F}_r^{(1)}} \mathbf{1}_{x\equiv c_r \pmod 1} + \mu(r) \sum_{c_r \in \mathcal{F}_r^{(r)}} \mathbf{1}_{x\equiv c_r \pmod r} \\
&= 1 \cdot \mathbf{1}_{x\equiv 0 \pmod 1} + (-1) \cdot \left(\mathbf{1}_{x\equiv 0 \pmod r} + \mathbf{1}_{x\equiv -y \pmod r}\right) \\
&= 1 - \mathbf{1}_{x\equiv 0 \pmod r} - \mathbf{1}_{x\equiv -y \pmod r} \\
&= f_r(x).
\end{aligned}$$

Case 2: $r \mid y$**.** Then $-y \equiv 0 \pmod r$, so the two conditions coincide: $\mathbf{1}_{r|x} = \mathbf{1}_{r|(x+y)}$, and their product equals the same indicator. Hence

$$f_r(x) = 1 - \mathbf{1}_{r|x} - \mathbf{1}_{r|x} + \mathbf{1}_{r|x} = 1 - \mathbf{1}_{r|x}.$$

The Möbius sum with $\mathcal{F}_r^{(r)} = \{0\}$ (only one forbidden class) gives

$$1 \cdot 1 + (-1) \cdot \mathbf{1}_{x\equiv 0 \pmod r} = 1 - \mathbf{1}_{r|x} = f_r(x).$$

Thus, in both cases, we have the single-prime identity:

$$f_r(x) = \sum_{d_r|r} \mu(d_r) \sum_{c_r \in \mathcal{F}_r^{(d_r)}} \mathbf{1}_{x\equiv c_r \pmod{d_r}}. \tag{146}$$

A.2 Product Over All Primes in S_1

Taking the product of (146) over all $r \in S_1$:

$$\prod_{r\in S_1} f_r(x) = \prod_{r\in S_1}\left(\sum_{d_r\mid r}\mu(d_r)\sum_{c_r\in\mathcal{F}_r^{(d_r)}}\mathbf{1}_{x\equiv c_r \pmod{d_r}}\right)$$
$$= \sum_{(d_r)_{r\in S_1}}\left(\prod_{r\in S_1}\mu(d_r)\right)\sum_{(c_r)}\prod_{r\in S_1}\mathbf{1}_{x\equiv c_r \pmod{d_r}},$$

where the outer sum runs over all choices of $d_r \in \{1, r\}$ for each $r \in S_1$, and the inner sum runs over all choices of $c_r \in \mathcal{F}_r^{(d_r)}$ for each r.

For a given choice $(d_r)_{r\in S_1}$, let $T = \{r \in S_1 : d_r = r\}$ and $d = \prod_{r\in T} r$. Then d is a squarefree divisor of P^*. The Möbius factor is

$$\prod_{r\in S_1}\mu(d_r) = \prod_{r\in T}\mu(r)\cdot\prod_{r\notin T}\mu(1) = \prod_{r\in T}(-1) = (-1)^{|T|} = \mu(d).$$

For $r \in T$, c_r is chosen from $\mathcal{F}_r^{(r)} = \mathcal{F}_r$. By the Chinese Remainder Theorem, the simultaneous congruences $x \equiv c_r \pmod r$ for all $r \in T$ are equivalent to a single congruence $x \equiv c \pmod d$, where $c \bmod d$ is the unique residue class determined by the tuple $(c_r)_{r\in T}$. As the c_r vary independently over their respective $\mathcal{F}_r$, the resulting c varies over precisely the set $\mathscr{C}_d$ defined in Definition 5.3.

For $r \notin T$, $d_r = 1$, the congruence is modulo 1, which imposes no condition, and $c_r = 0 \bmod 1$ is the only choice.

Therefore, the product simplifies to

$$\prod_{r\in S_1} f_r(x) = \sum_{d\mid P^*}\mu(d)\sum_{c\in\mathscr{C}_d}\mathbf{1}_{x\equiv c \pmod d},$$

which is exactly (60). This completes the verification.

B Symmetric Polynomial Compression

The symmetric polynomial method provides a simple bound for sums of the form $\sum_{d\mid P^*, d>X} 2^{\omega(d)}/d$.

Lemma B.1 (Symmetric Polynomial Bound)**.** *For any $X > 0$,*

$$\sum_{\substack{d\mid P^*\\ d>X}}\frac{2^{\omega(d)}}{d} \le \frac{1}{X}\prod_{r\in S_1}(1+2) = \frac{3^{|S_1|}}{X}. \tag{147}$$

Proof. For each term in the sum, $d > X$, so $d/X \ge 1$. Multiplying each term by this factor and extending the sum to all divisors $d \mid P^*$:

$$\sum_{\substack{d\mid P^*\\ d>X}}\frac{2^{\omega(d)}}{d} \le \frac{1}{X}\sum_{d\mid P^*}2^{\omega(d)}.$$

The unrestricted sum over all squarefree divisors of P^* factors as

$$\sum_{d|P^*} 2^{\omega(d)} = \prod_{r\in S_1}(1+2) = 3^{|S_1|}.$$

This completes the proof. Note that $3^{|S_1|} \approx 3^{p/\log p}$ is doubly exponential in p, so this bound is only useful when X is comparably large (as in the truncation argument where X is also exponentially large). □

C The Brun Pure Sieve Remainder

We provide the standard estimate for the remainder term in the Brun pure sieve, which justifies the truncation argument used in Zone III.

Theorem C.1 (Brun Pure Sieve Remainder)**.** *Let A be a finite set, and for each prime $r \in S_1$, let $E_r \subseteq A$ be a distinguished subset. For each squarefree $d \mid P^*$, define*

$$A_d = \bigcap_{r|d} E_r,$$

with the convention $A_1 = A$. Suppose that for each d, the cardinality of A_d can be written as

$$|A_d| = X \cdot g(d) + r_d,$$

where $X > 0$ is a parameter, $g(d)$ is a multiplicative function satisfying $0 \le g(p) < 1$ for all $p \in S_1$, and $|r_d| \le R_d$ is an error term. Then for any odd positive integer m,

$$\left|\bigcap_{r\in S_1}(A \setminus E_r)\right| \ge X \sum_{\substack{d|P^*\\ \omega(d)\le m}} \mu(d)g(d) - \sum_{\substack{d|P^*\\ \omega(d)\le m}} R_d - X \sum_{k=m+1}^{\infty} \frac{1}{k!}\left(\sum_{r\in S_1} g(r)\right)^k.$$

Proof. This is a standard formulation of the Brun pure sieve; see [3, Theorem 6.1.2] or [6, Chapter 2] for a complete proof. The idea is to apply the Bonferroni lower bound (Lemma 13.1) to the events E_r, express the resulting sums in terms of A_d, and substitute the asymptotic expression for $|A_d|$. The final term bounds the tail of the alternating series. □

In our application, $A = \mathcal{A}_0 \times \{0, 1, \ldots, N-1\}$ is the set of all pairs (a, t). The distinguished subset E_r consists of those pairs for which $a + tM_0 \equiv 0$ or $-y \pmod r$. The parameter $X = |\mathcal{A}_0|N$ is the total number of pairs, and $g(r) = |\mathcal{F}_r|/r$.

D Verification of the Average Order of $2^{\omega(n)}$

We provide a self-contained proof of Theorem 2.11, which gives the average order of the multiplicative function $2^{\omega(n)}$.

Proof of Theorem 2.11. Consider the Dirichlet generating function

$$F(s) = \sum_{n=1}^{\infty} \frac{2^{\omega(n)}}{n^s}, \qquad \Re(s) > 1.$$

Since $2^{\omega(n)}$ is multiplicative and satisfies $2^{\omega(n)} \leq \tau(n)$ (the divisor function), the series converges absolutely for $\Re(s) > 1$. Using the Euler product for multiplicative functions:

$$\begin{aligned} F(s) &= \prod_p \left(1 + \frac{2}{p^s} + \frac{2}{p^{2s}} + \frac{2}{p^{3s}} + \cdots\right) \\ &= \prod_p \left(1 + \frac{2}{p^s} \cdot \frac{1}{1 - p^{-s}}\right) \\ &= \prod_p \frac{1 - p^{-s} + 2p^{-s}}{1 - p^{-s}} \\ &= \prod_p \frac{1 + p^{-s}}{1 - p^{-s}}. \end{aligned}$$

We recognize that

$$F(s) = \prod_p \frac{1 - p^{-2s}}{(1 - p^{-s})^2} = \frac{\zeta(s)^2}{\zeta(2s)},$$

where $\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_p (1 - p^{-s})^{-1}$ is the Riemann zeta function.

The zeta function has a simple pole at $s = 1$ with residue 1. Therefore, $\zeta(s)^2$ has a double pole at $s = 1$. The factor $1/\zeta(2s)$ is analytic for $\Re(s) > 1/4$, and at $s = 1$ it takes the value $1/\zeta(2) = 6/\pi^2$.

Near $s = 1$, we have the Laurent expansion

$$\zeta(s) = \frac{1}{s - 1} + \gamma + O(s - 1),$$

where γ is Euler's constant. Squaring this:

$$\zeta(s)^2 = \frac{1}{(s - 1)^2} + \frac{2\gamma}{s - 1} + O(1).$$

Let $G(s) = 1/\zeta(2s)$. Then $G(1) = 6/\pi^2$, and G is analytic at $s = 1$. Writing $G(s) = G(1) + G'(1)(s - 1) + \cdots$, we have

$$F(s) = \left(\frac{1}{(s - 1)^2} + \frac{2\gamma}{s - 1} + O(1)\right) \cdot \left(G(1) + G'(1)(s - 1) + O((s - 1)^2)\right).$$

By the Wiener-Ikehara theorem for Dirichlet series with a double pole (see [12, Theorem II.7.13]),

$$\sum_{n \leq X} 2^{\omega(n)} \sim \frac{G(1)}{\Gamma(2)} X \log X = \frac{6}{\pi^2} X \log X.$$

The error term $O(X)$ can be obtained from more refined Tauberian theorems or by contour integration. □

E Complete Details of the Truncation Argument

We provide the complete, self-contained argument for truncating the Möbius expansion at $\omega(d) \leq m$, where $m = \lfloor \log p / \log \log p \rfloor$ (adjusted to be odd).

E.1 The Bonferroni Lower Bound in Our Context

For each $r \in S_1$, define the event E_r on the set $\mathcal{A}_0 \times \{0, \ldots, N-1\}$ as

$$E_r = \{(a,t) : a + tM_0 \equiv 0 \text{ or } a + tM_0 \equiv -y \pmod{r}\}.$$

The counting function T is exactly the number of pairs (a,t) that avoid all events E_r:

$$T = \left| \bigcap_{r \in S_1} E_r^c \right|.$$

By the inclusion-exclusion principle,

$$T = \sum_{k=0}^{|S_1|} (-1)^k \sum_{\substack{T \subseteq S_1 \\ |T|=k}} \left| \bigcap_{r \in T} E_r \right|.$$

The Bonferroni inequalities (Lemma 13.1) state that for any odd m,

$$T \geq \sum_{k=0}^{m} (-1)^k S_k,$$

where $S_k = \sum_{|T|=k} |\bigcap_{r \in T} E_r|$.

For a subset T with $|T| = k$ and $d = \prod_{r \in T} r$, the intersection $\bigcap_{r \in T} E_r$ consists of pairs (a,t) such that for every $r \in T$, $a + tM_0$ is congruent to a forbidden residue modulo r. By CRT, this is equivalent to $a + tM_0 \equiv c \pmod{d}$ for some $c \in \mathscr{C}_d$.

Thus,

$$S_k = \sum_{\substack{d | P^* \\ \omega(d)=k}} \sum_{c \in \mathscr{C}_d} \sum_{a \in \mathcal{A}_0} \sum_{t=0}^{N-1} \mathbf{1}_{a+tM_0 \equiv c \pmod{d}}.$$

This is precisely the k-th term in our expansion (62). The truncated sum in (131) is exactly $\sum_{k=0}^{m} (-1)^k S_k$.

E.2 Bounding the Remainder

The remainder from truncating at m (odd) is at most the $(m+1)$-th term S_{m+1} (since the series alternates and the terms eventually decrease in magnitude).

We estimate S_{m+1} as follows. For each d with $\omega(d) = m+1$,

$$\sum_{c \in \mathscr{C}_d} \sum_{a \in \mathcal{A}_0} \sum_{t=0}^{N-1} \mathbf{1}_{a+tM_0 \equiv c \pmod{d}} \leq |\mathscr{C}_d| \cdot |\mathcal{A}_0| \cdot \left(\frac{N}{d} + 1 \right) \leq 2 \cdot |\mathscr{C}_d| \cdot |\mathcal{A}_0| \cdot \frac{N}{d},$$

for large p (since $N/d \geq 1$ for $d \leq N \approx p$).

Using $|\mathscr{C}_d| \leq 2^{\omega(d)}$, we have

$$S_{m+1} \leq 2|\mathcal{A}_0| N \sum_{\substack{d | P^* \\ \omega(d)=m+1}} \frac{2^{\omega(d)}}{d}.$$

The sum over d with exactly $m+1$ prime factors is bounded by the $(m+1)$-th power of the sum over individual primes, divided by $(m+1)!$:

$$\sum_{\substack{d|P^* \\ \omega(d)=m+1}} \frac{2^{\omega(d)}}{d} \leq \frac{1}{(m+1)!} \left(\sum_{r \in S_1} \frac{2}{r} \right)^{m+1}.$$

Let $S = \sum_{r \in S_1} 2/r$. By Mertens' theorem,

$$S \leq 2 \sum_{\log p < r \leq p} \frac{1}{r} \leq 3 \log \log p$$

for sufficiently large p.

With $m \sim \log p / \log \log p$, Stirling's formula gives

$$\frac{S^{m+1}}{(m+1)!} \leq \frac{1}{\sqrt{2\pi(m+1)}} \left(\frac{eS}{m+1} \right)^{m+1} \leq p^{-1+o(1)},$$

as computed in §13. Therefore,

$$S_{m+1} \leq 2|\mathcal{A}_0| N \cdot p^{-1+o(1)} \ll p^{1+o(1)}.$$

E.3 Conclusion of the Truncation Argument

The truncated sum $\sum_{k=0}^{m} (-1)^k S_k$ is a lower bound for T, and its evaluation involves only moduli d with $\omega(d) \leq m$. For any such d, $d \leq (\log p)^m \leq p \leq M_0$. Thus all moduli in the truncated sum belong to Zones I and II.

The tail $T - \sum_{k=0}^{m} (-1)^k S_k$ is non-negative (by the Bonferroni lower bound) and bounded above by $S_{m+1} \leq p^{1+o(1)}$. Therefore, the error from omitting Zones III and IV is at most $p^{1+o(1)}$, which is consistent with the error bounds already obtained for Zones I and II.

This completes the justification of the truncation approach used in §13.