Sparse variance for squarefree numbers in arithmetic progressions to polynomial moduli

A write-up of an archived, model-generated proof

Archived run `zt-thesis-plan`, commit `867fc9b`
(models and workflow are recorded in Appendix B)

Abstract

For a polynomial $f \in \mathbb{Z}[t]$ of degree $d \geq 1$ with $f(t) \geq 2$ and $f'(t) \geq 1$ for $t \geq 1$, we study the weighted variance

$$V_f(x,y) = \sum_{k \leq y} f'(k) \sum_{l=1}^{f(k)} \Big(Q(x; f(k), l) - g(f(k), l)\, x \Big)^2$$

of the number $Q(x;q,l)$ of squarefree $n \leq x$ in the residue class $l \bmod q$, where the moduli are restricted to values of f and every residue class is included. We prove that, uniformly for $f(y) \leq x$,

$$V_f(x,y) = x^2 F_f(x/f(y)) + O_{f,\varepsilon}(x^{2-\eta_d+\varepsilon}), \qquad \eta_d = \min\Big\{\frac{1}{5}, \frac{1}{4d-1}\Big\},$$

with an explicit profile F_f satisfying $F_f(H) = K_f H^{-3/2} + O(H^{-13/8})$, and hence $V_f(x,y) \sim K_f x^{1/2} f(y)^{3/2}$ for $x^{1-2\eta_d/3+\delta} \leq f(y) \leq x^{1-\delta}$. The constant $K_f > 0$ is $-\frac{8}{3}\zeta(-\frac{1}{2})$ times an absolutely convergent Euler product built from the numbers of roots of f modulo p and p^2; no irreducibility or squarefreeness assumption on f is needed. The proof combines an elementary uniform estimate for correlations of squarefree numbers, an average of the resulting singular series over polynomial moduli, and a Mellin-transform analysis of a Dirichlet series whose pole at $s = -\frac{1}{2}$ produces the main term. We compare the predictions with exact computations of V_f up to $x = 2.56 \cdot 10^8$ for $f(t) = t^2+1$, t^3+2 and $2t+1$.

Contents

1 Introduction **3**
- 1.1 Squarefree numbers in arithmetic progressions . . . 3
- 1.2 Sparse moduli . . . 4
- 1.3 The main result . . . 4
- 1.4 Numerical illustration . . . 6
- 1.5 Strategy of the proof . . . 6
- 1.6 Organisation and conventions of this document . . . 7

2 Notation and elementary tools **8**
- 2.1 Conventions . . . 8
- 2.2 Elementary estimates . . . 8
- 2.3 Infinite products and Euler products . . . 9
- 2.4 Ramanujan sums . . . 9
- 2.5 The polynomial and its roots . . . 10

3 Local densities and the singular series **12**
- 3.1 The density of squarefree numbers in a residue class . . . 12
- 3.2 The mean square of the densities . . . 13
- 3.3 The singular series . . . 14

4 Expanding the variance **16**
- 4.1 Notation . . . 16
- 4.2 Counting pairs in a residue class . . . 16
- 4.3 The variance on an interval of moduli . . . 17
- 4.4 The initial moduli . . . 18

5 Correlations of squarefree numbers **19**

6 Averaging over polynomial moduli **21**
- 6.1 The averaged singular series . . . 21
- 6.2 A partial summation lemma . . . 22
- 6.3 The averaging proposition . . . 23

7 The Dirichlet series of the averaged singular series **25**
- 7.1 Facts about the zeta function . . . 25
- 7.2 The local factors . . . 26
- 7.3 The factorisation . . . 26

8 The smoothed sum Θ_f and the profile F_f **29**
8.1 Growth in a vertical strip . . . 29
8.2 The Mellin representation . . . 30
8.3 The integral identity . . . 31

9 Proof of the main theorem **32**
9.1 The variance on an interval of moduli . . . 32
9.2 Choice of the cutoffs and the error budget . . . 33
9.3 Proof of Corollary 1.2 . . . 34
9.4 Where the exponents come from . . . 34

10 Numerical evidence **35**
10.1 What is computed . . . 35
10.2 The constants . . . 36
10.3 The data against the profile . . . 37
10.4 Extrapolation to small ratios . . . 38
10.5 Limitations . . . 40

A Full numerical tables **41**

B Provenance of this document **43**

Chapter 1

Introduction

1.1 Squarefree numbers in arithmetic progressions

Let μ be the Möbius function, so that $\mu^2(n)$ is the indicator function of the squarefree positive integers. For a modulus $q \geq 1$, an integer l and real $x \geq 1$ put

$$Q(x;q,l) = \#\{n \leq x : \ n \equiv l \pmod q, \ n \text{ squarefree}\}.$$

The limit $g(q,l) = \lim_{x\to\infty} Q(x;q,l)/x$ exists; its closed form is recalled in Proposition 3.1. Unlike the primes, the squarefree numbers populate residue classes that are *not* coprime to the modulus, and their density $g(q,l)$ depends on l through the common divisors of l and q. For instance, $g(q,l) = 0$ whenever p^2 divides both q and l for some prime p. The error term

$$E(x;q,l) = Q(x;q,l) - g(q,l)\,x$$

measures the irregularity of the distribution, and its mean square over all residue classes and all moduli up to Y,

$$V(x,Y) = \sum_{q\leq Y} \sum_{l=1}^{q} E(x;q,l)^2,$$

is the squarefree analogue of the Barban–Davenport–Halberstam and Montgomery–Hooley variance for primes [8, 5].

For $Y \leq x$ the bound $V(x,Y) \ll_\varepsilon x^{1+\varepsilon}Y$ is Lemma 2.2 of Brüdern, Granville, Perelli, Vaughan and Wooley [1] (with their parameter $k = 2$); it is quoted precisely in quoted result 4.4 below. Asymptotic formulae for $V(x,Y)$ go back to work of Orr, Warlimont and Croft, surveyed in the introduction of Vaughan's paper [11]; according to that survey, Warlimont [12] proved $V(x,Y) = x^2\mathcal{F}(x/Y) + O(x^{3/2}(\log x)^{7/2})$ for $1 \leq Y \leq x$, with $\mathcal{F}(u) \sim c_2\, u^{-3/2}$ as $u \to \infty$, where c_2 is the constant in (1.2) below; and the argument of Croft [3] contains an error affecting the constant. Vaughan [11, Theorem] broke the barrier $x^{3/2}$ inherent in the earlier methods. For squarefree numbers (his $k = 2$) his theorem states that, for $Y \leq x$,

$$V(x,Y) = c_2 x^{1/2}Y^{3/2} + O\Big(x^{1/4}Y^{7/4}\exp\Big(-c^*\frac{(\log 2x/Y)^{3/5}}{(\log\log 3x/Y)^{1/5}}\Big)\Big) + O\Big(x^{3/2}\exp\Big(-\tilde{c}\frac{(\log x)^{3/5}}{(\log\log x)^{1/5}}\Big)\Big), \quad (1.1)$$

for certain positive constants $c^*, \tilde{c}$, where

$$c_2 = \frac{8C_1\,(-\zeta(-\tfrac12))}{3\,\zeta(2)^2}, \qquad C_1 = \prod_p\Big(1 + \frac{(p-1)^2p^{3/2} - 2p^4 + 3p^3 - p}{p^2(p^2-1)^2}\Big) \quad (1.2)$$

[11, (1.13)–(1.15)]; in the numerator of the local factor of C_1, the term $3p^3$ is the sum $p^{2k-1}+2p^{k+1}$ of Vaughan's general formula at $k=2$. The main term is thus of order $x^{1/2}Y^{3/2}$, much smaller than the "diagonal" size xY one might naively expect; this reflects strong cancellation, which will also be the central feature of the present work.

1.2 Sparse moduli

Brüdern and Wooley [2] asked the analogous question for primes when the moduli are restricted to the values $f(k)$ of a polynomial, weighting the modulus $f(k)$ by $f'(k)$. The weight makes the sparse sum behave like the full sum over moduli $q \le f(y)$, since $f'(k)\,\mathrm{d}k = \mathrm{d}f(k)$. They established an asymptotic formula of Montgomery–Hooley type for this sparse variance, improving the upper bound of Mikawa and Peneva [6]. Their method follows the approach of Goldston and Vaughan [4] and Vaughan [10]: the variance is linked to the number of solutions of $p_1 - p_2 = hf(k)$, which is then counted by the circle method.

This document treats the analogous problem for squarefree numbers. Throughout, $f \in \mathbb{Z}[t]$ is a polynomial of degree $d \ge 1$ such that

$$f(t) \ge 2 \quad \text{and} \quad f'(t) \ge 1 \qquad \text{for all real } t \ge 1, \tag{1.3}$$

and

$$V_f(x,y) = \sum_{k \le y} f'(k) \sum_{l=1}^{f(k)} E\big(x; f(k), l\big)^2 \qquad (f(y) \le x). \tag{1.4}$$

Here k runs over positive integers, $y \ge 1$ is real, and all residue classes modulo $f(k)$ are included. We write

$$\rho(m) = \#\{a \bmod m : \ f(a) \equiv 0 \pmod m\}, \qquad \rho(1) = 1,$$

and $Y = f(y)$. No irreducibility, squarefreeness or discriminant condition is imposed on f, and f may have fixed prime divisors.

1.3 The main result

To state the result we need an arithmetic weight. For a positive integer h and a prime p let $v_p(h)$ be the exponent of p in h, and define

$$W_f(h) = \prod_p w_p(h), \qquad w_p(h) = \begin{cases} 1 - \dfrac{2}{p^2} + \dfrac{\rho(p^2)}{p^4}, & p \nmid h, \\ 1 - \dfrac{2}{p^2} + \dfrac{\rho(p)}{p^3}, & p \parallel h, \\ 1 - \dfrac{1}{p^2}, & p^2 \mid h. \end{cases} \tag{1.5}$$

Equivalently, $w_p(h) = (1-2p^{-2})(1+\rho(p^j)/(p^j(p^2-2)))$ with $j = \max\{2 - v_p(h), 0\}$. In Chapter 6 we shall see that $W_f(h)$ is the average over k of the singular series $\mathfrak{S}(hf(k))$ that governs the number of pairs of squarefree integers with difference $hf(k)$. The product converges absolutely and $0 < W_f(h) \le 1$ (Lemma 6.1). Let

$$D_f(s) = \sum_{h \ge 1} W_f(h) h^{-s} \qquad (\operatorname{Re} s > 1).$$

By Proposition 7.2, D_f continues meromorphically to $\operatorname{Re} s > -\frac{3}{4}$, where its only poles are simple poles at $s = 1$ and $s = -\frac{1}{2}$. Define

$$R_f = \operatorname*{Res}_{s=1} D_f(s), \qquad B_f = \lim_{s\to 1}\Big(D_f(s) - \frac{R_f}{s-1}\Big) - \frac{3}{2}R_f, \qquad K_f = -\frac{16}{3}\operatorname*{Res}_{s=-1/2} D_f(s), \tag{1.6}$$

and, for real $H \geq 1$,

$$\Theta_f(H) = \sum_{1\leq h\leq H} W_f(h)\frac{(H-h)^2}{h}, \qquad F_f(H) = R_f \log H + B_f + \frac{1}{\zeta(2)H} - \frac{\Theta_f(H)}{H^2}. \tag{1.7}$$

Theorem 1.1 (Main theorem). *Let f satisfy (1.3), let $d = \deg f$, and put*

$$\eta_d = \min\Big\{\frac{1}{5}, \frac{1}{4d-1}\Big\}.$$

For every ε with $0 < \varepsilon < \eta_d/2$ there are constants $C(f,\varepsilon) > 0$ and $x_0(f,\varepsilon)$ such that, uniformly for real $x \geq x_0(f,\varepsilon)$ and real $y \geq 1$ with $Y = f(y) \leq x$,

$$|V_f(x,y) - x^2 F_f(x/Y)| \leq C(f,\varepsilon)\, x^{2-\eta_d+\varepsilon}, \tag{1.8}$$

$$|V_f(x,y) - K_f x^{1/2} Y^{3/2}| \leq C(f,\varepsilon)(x^{3/8}Y^{13/8} + x^{2-\eta_d+\varepsilon}). \tag{1.9}$$

The constant K_f is positive and is given by the absolutely convergent Euler product

$$K_f = -\frac{8}{3}\,\zeta(-\tfrac{1}{2})\prod_p\Big(1-\frac{1}{p}\Big)\Big(1+\frac{1}{p}-\frac{2+\rho(p)}{p^2}+\frac{\rho(p)}{p^{5/2}}-\frac{\rho(p^2)}{p^{7/2}}+\frac{\rho(p^2)}{p^4}\Big). \tag{1.10}$$

Moreover

$$R_f = \frac{1}{\zeta(2)^2}\prod_p\Big(1+\frac{(1-1/p)(\rho(p)+\rho(p^2))}{(p^2-1)^2}\Big). \tag{1.11}$$

The function F_f is explicit, and Proposition 8.5 shows that $|F_f(H) - K_f H^{-3/2}| \leq C(f)H^{-13/8}$ for all $H \geq 1$; (1.9) is (1.8) combined with this expansion at $H = x/Y$. Dividing (1.9) by its main term gives the following.

Corollary 1.2. *Let $0 < \delta < \eta_d/3$. Uniformly for $x \geq x_0(f,\delta)$ and $x^{1-2\eta_d/3+\delta} \leq f(y) \leq x^{1-\delta}$,*

$$V_f(x,y) = K_f\, x^{1/2} f(y)^{3/2}(1 + O_{f,\delta}(x^{-\delta/8})).$$

The proof of Corollary 1.2 is given in Section 9.3. The exponents for small degrees are as follows.

d	η_d	error in (1.8)	range of Corollary 1.2
1	$1/5$	$x^{9/5+\varepsilon}$	$x^{13/15+\delta} \leq f(y) \leq x^{1-\delta}$
2	$1/7$	$x^{13/7+\varepsilon}$	$x^{19/21+\delta} \leq f(y) \leq x^{1-\delta}$
3	$1/11$	$x^{21/11+\varepsilon}$	$x^{31/33+\delta} \leq f(y) \leq x^{1-\delta}$
$d \geq 2$	$1/(4d-1)$	$x^{2-1/(4d-1)+\varepsilon}$	$x^{1-2/(3(4d-1))+\delta} \leq f(y) \leq x^{1-\delta}$

Remark 1.3 (Two regimes). Theorem 1.1 distinguishes two descriptions of the variance. Suppose that Y/x tends to zero and that, in addition, the additive error $x^{2-\eta_d+\varepsilon}$ in (1.9) is $o(x^{1/2}Y^{3/2})$; this holds, for instance, throughout the range of Corollary 1.2. Then (1.9) shows that the variance is asymptotic to the power law $K_f x^{1/2}Y^{3/2}$, as in Vaughan's theorem (1.1). For smaller Y the theorem gives only an upper bound for the difference, and no asymptotic formula. When the ratio $r = Y/x$ is fixed, (1.8) gives $V_f(x,y) = x^2 F_f(1/r) + O(x^{2-\eta_d+\varepsilon})$. Here $x^2 F_f(1/r) =$

$K_f x^{1/2}Y^{3/2} \cdot F_f(1/r)r^{-3/2}/K_f$, and by (8.3) the factor $F_f(1/r)r^{-3/2}/K_f = 1 + O(r^{1/8})$ is independent of x. The theorem does not say that this factor equals one, and the numerical values in Chapter 10 show that it does not. The normalised quantity $V_f(x,y)/(x^{1/2}Y^{3/2})$ therefore approaches the ratio-dependent profile $F_f(1/r)r^{-3/2}$ rather than the constant K_f. This distinction is essential for interpreting the numerical data.

Remark 1.4 (Consistency with Vaughan's constant). For $f(t) = t + 1$ one has $\rho(m) = 1$ for all m, and $V_f(x,y) = V(x,Y) - E(x;1,1)^2$ with $Y = f(y) = y + 1$, because the moduli $f(k) = k + 1$, $k \le y$, are exactly the integers $2 \le q \le Y$; by Proposition 3.1 the subtracted term is $O(x)$. A direct computation shows that for $\rho(p) = \rho(p^2) = 1$ each local factor in (1.10) equals $(1-p^{-2})^2$ times the local factor of C_1 in (1.2). Since $\prod_p (1-p^{-2})^2 = \zeta(2)^{-2}$, it follows that $K_{t+1} = c_2$. Thus the constant of Theorem 1.1 agrees with Vaughan's constant in the one case where both results apply; numerically $K_{t+1} = 0.128479$ (see Table 10.1). For $f(t) = t + 1$, Theorem 1.1 is of course much weaker than (1.1) as far as the range and the error term are concerned. This comparison was made while preparing this document and is not part of the archived proof (see Appendix B).

Remark 1.5 (Limitations). The asymptotic formula is established only for $f(y) \ge x^{1-2\eta_d/3+\delta}$, not throughout $f(y) \le x$. Two features of the argument fix the exponent η_d (see Section 9.2). The cap $1/5$ comes from summing the error $O(x^{4/5})$ of the correlation estimate (Proposition 5.1) over about $x \log x$ pairs (h,k). The exponent $1/(4d-1)$ comes from balancing the bound for small moduli, which is obtained by inserting the full variance bound of [1], against the error of the averaging over polynomial values in Proposition 6.4.

1.4 Numerical illustration

Chapter 10 compares the theorem with an archived numerical snapshot of exact computations of $V_f(x,y)$ for $f(t) = t^2+1$, t^3+2 and $2t+1$, for $2.5\cdot 10^5 \le x \le 2.56\cdot 10^8$ and $f(y)/x \in [0.005, 1]$; the computations are repeated by the script `reproduce.py` accompanying this document. The constants are

$$K_{t^2+1} = 0.138083, \qquad K_{t^3+2} = 0.135608, \qquad K_{2t+1} = 0.142613,$$

computed from (1.10) with primes up to 10^7. At the largest x the data agree with the profile $x^2 F_f(x/Y)$ to within a fraction of one per cent for t^2+1 and $2t+1$, and to within a few per cent for t^3+2, where y is only a few hundred. The snapshot's empirical extrapolation to $r \to 0$ is consistent with K_f within its heuristic uncertainty envelope. Section 10.4 splits its offset from K_f into a part due to the fitting model itself and a part due to the difference between data and profile at finite x. The numbers are evidence, not proof; Chapter 10 explains the method and its limitations.

1.5 Strategy of the proof

The proof follows a classical route: expand the square, count pairs of squarefree numbers in the same residue class, and evaluate the resulting averages. It has seven steps, corresponding to Chapters 3 to 9.

(1) *Local densities* (Chapter 3). We compute $g(q,l)$, the second moment $q\sum_l g(q,l)^2 = \sum_{r\mid q}\lambda(r)$, and the Ramanujan expansion $\mathfrak{S}(n) = \sum_r \gamma(r)^2 c_r(n)$ of the singular series

$$\mathfrak{S}(n) = \prod_p \Big(1 - \frac{2}{p^2}\Big) \prod_{p^2\mid n} \frac{p^2-1}{p^2-2}.$$

(2) *Expanding the variance* (Chapter 4). Opening the square and counting ordered pairs of squarefree numbers in a common residue class gives

$$V_f(x;z,y) \approx cx(Y-Z)+2\sum_{z<k\le y} f'(k)\sum_{h\le x/f(k)} C_x(hf(k))-x^2\sum_{z<k\le y} f'(k)\sum_{l} g(f(k),l)^2,$$

where $c=\zeta(2)^{-1}$, $Z=f(z)$ and $C_x(m)=\sum_{n\le x-m}\mu^2(n)\mu^2(n+m)$. The moduli $f(k)$ with $k\le z$ are disposed of by the bound of [1].

(3) *Correlations* (Chapter 5). An elementary argument gives $C_x(m)=(x-m)\mathfrak{S}(m)+O(x^{4/5})$ uniformly in $1\le m\le x$.

(4) *Averaging over polynomial moduli* (Chapter 6). Using the Ramanujan expansion, the dependence of $\mathfrak{S}(hf(k))$ and of $\sum_l g(f(k),l)^2$ on k is periodic modulo each r. Partial summation replaces the sums over k by integrals over $t=f(k)$. The singular series average to $W_f(h)$, and the density term becomes $R_f\log(Y/Z)$, with the same constant R_f.

(5) *The Dirichlet series* (Chapter 7). The factorisation $D_f(s)=\zeta(s)\zeta(2s+2)H_f(s)$ with H_f an Euler product holomorphic for $\operatorname{Re}s>-\frac34$ gives the residues and the value $D_f(0)=-\frac12\zeta(2)^{-1}$.

(6) *The smoothed sum* (Chapter 8). A Mellin integral with kernel $2/(s(s+1)(s+2))$ and a contour shift to $\operatorname{Re}s=-\frac58$ give

$$\Theta_f(H)=R_fH^2\log H+B_fH^2+cH-K_fH^{1/2}+O(H^{3/8}).$$

The term cH comes from the pole of the kernel at $s=0$ and the value $D_f(0)$.

(7) *Assembly* (Chapter 9). The integral produced in step (4) equals $\frac12(Z^2\Theta_f(x/Z)-Y^2\Theta_f(x/Y))$ exactly. In the combination of steps (2)–(6), the terms $x^2R_f\log(Y/Z)$, the B_f-terms and the diagonal term $cx(Y-Z)$ cancel identically, which leaves $x^2F_f(x/Y)-x^2F_f(x/Z)$. Choosing $Z=x^{2d/(4d-1)}$ and estimating every error term proves Theorem 1.1.

The cancellation in step (7) is the reason for the shape of the profile F_f. All three large terms, of sizes $x^2\log x$, x^2 and xY, disappear, and what remains is $x^2F_f(x/Y)\approx K_fx^{1/2}Y^{3/2}$, produced by the pole of $\zeta(2s+2)$ at $s=-\frac12$.

1.6 Organisation and conventions of this document

Chapter 2 collects elementary facts used repeatedly: tail estimates, the divisor bound, infinite products, Ramanujan sums and the growth of f. Each of Chapters 3 to 9 proves one of the seven propositions of the archived proof; their statements are kept with the hypotheses, ranges and uniformity of the archive. Chapter 10 describes the numerical evidence, Appendix A lists the full numerical tables, and Appendix B records how this document was produced. The file `changes.md` maps every statement to its archived source.

Only two results are quoted from the literature: Lemma 2.2 of [1], used for the small moduli in Proposition 4.5, and Corollary 10.5 of [9] (a consequence of the functional equation of ζ), used to bound ζ in a vertical strip in Lemma 8.2. Everything else is proved here.

Chapter 2

Notation and elementary tools

2.1 Conventions

The letter p always denotes a prime, and products $\prod_p$ run over all primes. Unless a range says otherwise, summation variables $h, k, l, m, n, q, r, a, b, d, e, j$ run over positive integers. Logarithms are natural. We write $c = \zeta(2)^{-1} = 6/\pi^2$, $\tau(n)$ for the number of divisors of n, φ for Euler's totient, $v_p(n)$ for the exponent of p in n, $p \parallel n$ when $v_p(n) = 1$, (a, b) and $[a, b]$ for the greatest common divisor and least common multiple, $\mathbf{1}_{\mathcal{P}}$ for the indicator of a property $\mathcal{P}$, and $e(t) = e^{2\pi i t}$. For real t, $\lfloor t \rfloor$ and $\{t\}$ are the integer and fractional parts. The polynomial f always satisfies (1.3), $d = \deg f$, and $Y = f(y)$, $Z = f(z)$. "Constants depending only on f" may depend on the coefficients of f. The notation $A \ll B$ means $|A| \le CB$ for a constant C whose dependence is indicated by subscripts.

2.2 Elementary estimates

Lemma 2.1. *(a) $\sum_{j\ge1} j^{-2} \le 2$. For an integer $N \ge 1$, $\sum_{j>N} j^{-2} \le 1/N$. For real $T > 0$, $\sum_{j>T} j^{-2} \le 2/T$.*

(b) For real $y \ge 1$, $\sum_{j\le y} 1/j \le 1 + \log y$.

(c) For real $u > 1$ and an integer $N \ge 1$, $\sum_{p>N} p^{-u} \le \sum_{n>N} n^{-u} \le N^{1-u}/(u-1)$.

(d) For $\eta > 0$ and $x \ge 2$, $\log(2x) \le J_\eta x^\eta$ with $J_\eta = \log 2 + 1/(e\eta)$. For $x \ge 1$, $\log x \le 5x^{1/5}$. For fixed $j \ge 1$ and $u > 0$ there is $C_{j,u}$ with $(\log 2x)^j \le C_{j,u} x^u$ for $x \ge 2$.

(e) For integers $d \ge 1$ and $H \ge 1$, $\sum_{h\le H} h^{-1+1/d} \le dH^{1/d}$.

Proof. (a) Compare j^{-2} with $\int_{j-1}^{j} t^{-2}\,\mathrm{d}t$ for $j \ge 2$: $\sum_{j\ge1} j^{-2} \le 1 + \int_1^\infty t^{-2}\,\mathrm{d}t = 2$ and $\sum_{j>N} j^{-2} \le \int_N^\infty t^{-2}\,\mathrm{d}t = 1/N$. For $T \ge 1$ put $N = \lfloor T \rfloor \ge T/2$; then $\sum_{j>T} j^{-2} = \sum_{j>N} j^{-2} \le 1/N \le 2/T$. For $0 < T < 1$ the sum is at most $2 \le 2/T$. (b) and (c) follow in the same way by comparison with $\int t^{-1}\,\mathrm{d}t$ and $\int_N^\infty t^{-u}\,\mathrm{d}t$. (d) The function $u \mapsto u^{-\eta}\log u$ on $[1,\infty)$ has maximum $1/(e\eta)$, so $\log x \le x^\eta/(e\eta)$, and $\log 2 \le (\log 2)x^\eta$. Also $\log x = \int_1^x t^{-1}\,\mathrm{d}t \le \int_1^x t^{-4/5}\,\mathrm{d}t \le 5x^{1/5}$. The last claim follows from the first with $\eta = u/j$. (e) Since $t^{-1+1/d}$ is nonincreasing and integrable on $(0, H]$, the sum is at most $\int_0^H t^{-1+1/d}\,\mathrm{d}t = dH^{1/d}$. □

Lemma 2.2 (Divisor bound)**.** *For every $\eta > 0$ put $D_\eta = \prod_{p<2^{1/\eta}} (1 - p^{-\eta})^{-2}$. Then $\tau(n) \le D_\eta n^\eta$ for all $n \ge 1$.*

Proof. By multiplicativity it suffices to bound $(a+1)p^{-\eta a}$ for each prime power $p^a \parallel n$. If $p \ge 2^{1/\eta}$, then $a+1 \le 2^a \le p^{\eta a}$. If $p < 2^{1/\eta}$, then $(a+1)p^{-\eta a} \le \sum_{j\ge0}(j+1)p^{-\eta j} = (1-p^{-\eta})^{-2}$. Factors of D_η for primes not dividing n are at least one. □

2.3 Infinite products and Euler products

Lemma 2.3. *Let (e_p) be complex numbers indexed by the primes, with $\sum_p |e_p| < \infty$.*

(a) The partial products $P_N = \prod_{p \le N}(1+e_p)$ converge to a limit P, and $|P| \le \exp(\sum_p |e_p|)$.

(b) For every $N \ge 1$, $|\prod_{p>N}(1+e_p) - 1| \le \exp(\sum_{p>N} |e_p|) - 1$.

(c) If $1 + e_p \ne 0$ for all p, then $P \ne 0$; if all $1 + e_p$ are positive reals, then $P > 0$.

(d) If $e_p = e_p(s)$ are holomorphic on an open set $\Omega \subset \mathbb{C}$ and $\sum_p \sup_{s \in K} |e_p(s)| < \infty$ for every compact $K \subset \Omega$, then $P(s)$ is holomorphic on Ω.

Proof. For a finite set S of primes, expanding the product and using the triangle inequality gives $|\prod_{p\in S}(1+e_p) - 1| \le \prod_{p\in S}(1+|e_p|) - 1 \le \exp(\sum_{p\in S}|e_p|) - 1$. Hence, for $M > N$,

$$|P_M - P_N| = |P_N| \Big| \prod_{N<p\le M} (1+e_p) - 1 \Big| \le e^{\sum_p |e_p|} \Big(e^{\sum_{p>N} |e_p|} - 1 \Big),$$

which tends to zero as $N \to \infty$. This proves (a), and letting $M \to \infty$ in the finite version gives (b). For (c), choose N_0 with $|e_p| \le \frac{1}{2}$ for $p > N_0$. For $|w| \le \frac{1}{2}$ the principal logarithm satisfies $|\log(1+w)| \le 2|w|$, so $\prod_{p>N_0}(1+e_p) = \exp(\sum_{p>N_0} \log(1+e_p))$ is nonzero, and it is positive if the factors are positive. The finitely many remaining factors are nonzero (positive). For (d), the estimate in (a) holds uniformly on each compact K, so the holomorphic functions P_N converge locally uniformly and the limit is holomorphic. □

Lemma 2.4. *Let a be a multiplicative function with $\sum_n |a(n)| < \infty$. Then $\sum_{n\ge1} a(n) = \prod_p \sum_{j\ge0} a(p^j)$, and the product converges absolutely in the sense of Lemma 2.3.*

Proof. The local series converge absolutely. For a bound N, multiplying the finitely many local series with $p \le N$ gives $\sum a(n)$ over the integers n all of whose prime factors are at most N, by unique factorisation and multiplicativity. Every omitted n exceeds N, so the difference from $\sum_n a(n)$ is at most $\sum_{n>N} |a(n)|$, which tends to zero. Finally $\sum_p |\sum_{j\ge1} a(p^j)| \le \sum_{n\ge2} |a(n)| < \infty$. □

Applied to $a(n) = n^{-s}$ and $a(n) = \mu(n)n^{-2}$, Lemma 2.4 gives $\zeta(s) = \prod_p (1-p^{-s})^{-1}$ for $\operatorname{Re} s > 1$ and

$$c = \zeta(2)^{-1} = \sum_{n\ge1} \frac{\mu(n)}{n^2} = \prod_p \Big(1 - \frac{1}{p^2}\Big). \tag{2.1}$$

Since $\sum_p 2p^{-2} < \infty$ and each factor is positive, Lemma 2.3 also shows that $\prod_p (1 - 2p^{-2})$ converges to a positive number.

2.4 Ramanujan sums

For $r \ge 1$ and $n \in \mathbb{Z}$ let $c_r(n) = \sum_{1\le a\le r,\,(a,r)=1} e(an/r)$.

Lemma 2.5. *(a) $c_r(n) = \sum_{s|(r,n)} s\,\mu(r/s)$; in particular $c_r(n)$ is an integer depending only on n modulo r.*

(b) $|c_r(n)| \le \varphi(r)$, and $c_{rs}(n) = c_r(n)c_s(n)$ whenever $(r,s) = 1$.

(c) $c_p(n) = p\mathbf{1}_{p|n} - 1$, $c_{p^2}(n) = p^2\mathbf{1}_{p^2|n} - p\mathbf{1}_{p|n}$, and hence

$$c_p(n) + c_{p^2}(n) = p^2\mathbf{1}_{p^2|n} - 1. \tag{2.2}$$

Proof. Insert $\mathbf{1}_{(a,r)=1} = \sum_{b\mid(a,r)} \mu(b)$ and write $a = ba'$:

$$c_r(n) = \sum_{b\mid r} \mu(b) \sum_{a'=1}^{r/b} e\Big(\frac{a'n}{r/b}\Big) = \sum_{b\mid r} \mu(b)\frac{r}{b}\mathbf{1}_{(r/b)\mid n},$$

since a complete geometric sum of $e(a'n/m)$ over $a' \bmod m$ equals $m\mathbf{1}_{m\mid n}$. Writing $s = r/b$ gives (a). The bound $|c_r(n)| \le \varphi(r)$ is trivial, since $c_r(n)$ is a sum of $\varphi(r)$ numbers of modulus one. By (a), $r \mapsto c_r(n)$ is the Dirichlet convolution of the multiplicative functions $s \mapsto s\mathbf{1}_{s\mid n}$ and μ, hence multiplicative. Part (c) is (a) for $r = p, p^2$. □

2.5 The polynomial and its roots

Lemma 2.6. *There are constants $0 < \kappa_1 \le 1 \le \kappa_2$, depending only on f, such that for all real $t \ge 1$*

$$\kappa_1 t^d \le f(t) \le \kappa_2 t^d, \qquad 1 \le f'(t) \le \kappa_2 t^{d-1}, \qquad |f''(t)| \le \kappa_2 t^{d-2}. \tag{2.3}$$

(For $d = 1$, $f'' = 0$.) Put $U = \kappa_2/\kappa_1$ and $V = U + U^2$. Then:

(a) f is continuous, strictly increasing and unbounded on $[1,\infty)$; for every $Z \ge f(1)$ there is a unique real $z \ge 1$ with $f(z) = Z$. Moreover $f(k+1) - f(k) \ge 1$, $f(t) \ge t + 1$, and $f(y) \le x$ implies $y < x$ and $y \le (x/\kappa_1)^{1/d}$.

(b) $w(t) = f'(t)/f(t)$ satisfies $0 < w(t) \le U/t$ and $|w'(t)| \le V/t^2$ for $t \ge 1$.

(c) If $x \ge 2$, $1 \le z \le y$ and $f(y) \le x$, then

$$\sum_{z<k\le y} \frac{f'(k)}{f(k)} \le U \sum_{k\le y} \frac{1}{k} \le U(1 + \log y) \le 2U\log(2x). \tag{2.4}$$

(d) Put $\alpha = 1 - 1/d$ and $K_1 = 3\kappa_2\kappa_1^{-\alpha}$. For $1 \le z \le y$, the weighted count $\mathcal{W} = \sum_{z<k\le y} f'(k)$ satisfies

$$|\mathcal{W} - (Y - Z)| \le K_1 Y^\alpha, \qquad 0 \le \mathcal{W} \le (1 + K_1)Y. \tag{2.5}$$

(e) For $k \le z$ we have $f'(k) \le \kappa_2\kappa_1^{-\alpha} Z^\alpha$.

Proof. The leading coefficient A_d of f is nonzero, and it is positive because $f(t) \ge 2$ for all large t. The function $f(t)/t^d$ is positive and continuous on $[1,\infty)$ with the positive limit A_d, so it is bounded above and below by positive constants. Bounding each coefficient of f' and f'' gives the upper bounds for f' and f'', and $f' \ge 1$ is a hypothesis. We may take $\kappa_1 \le 1 \le \kappa_2$.

(a) Monotonicity and $f(k+1) - f(k) = \int_k^{k+1} f' \ge 1$ follow from $f' \ge 1$, and $f(t) \ge f(1) + (t-1) \ge t + 1$. The existence and uniqueness of z follow from continuity and strict monotonicity, and the last claims follow from $x \ge f(y) \ge y + 1$ and $x \ge f(y) \ge \kappa_1 y^d$.

(b) We have $w(t) \le \kappa_2 t^{d-1}/(\kappa_1 t^d) = U/t$ and $w' = f''/f - (f'/f)^2$, so $|w'| \le U/t^2 + U^2/t^2$.

(c) The first inequality is (b), the second is Lemma 2.1(b), and the third uses $y < x$ and $1 + \log x \le 2\log(2x)$, valid because $\log x + 2\log 2 > 1$.

(d) Let $n = \lfloor z \rfloor$ and $m = \lfloor y \rfloor$. For each k with $n < k \le m$, $|f'(k) - \int_{k-1}^k f'| \le \int_{k-1}^k |f''|$, and $[k-1, k] \subset [1,\infty)$ because $k \ge 2$. Summing gives $|\mathcal{W} - \int_n^m f'| \le \int_1^y |f''| \le \kappa_2 y^{d-1}$; for $d \ge 2$ this uses $\int_1^y t^{d-2}\,\mathrm{d}t \le y^{d-1}/(d-1)$, and for $d = 1$ the integral vanishes. Replacing $[n, m]$ by $[z, y]$ changes the integral by at most $2\sup_{1\le t\le y} f'(t) \le 2\kappa_2 y^{d-1}$. Since $\int_z^y f' = Y - Z$ and $y^{d-1} \le (Y/\kappa_1)^\alpha$, the first inequality follows. The second follows from the first, because $Y \ge 2$ gives $Y^\alpha \le Y$.

(e) $f'(k) \le \kappa_2 k^{d-1} \le \kappa_2 z^{d-1} \le \kappa_2 (Z/\kappa_1)^\alpha$. □

Lemma 2.7. *Let A_d be the leading coefficient of f, put $C_\rho = \max\{d, |A_d|\}$, and let $\mathcal{B}$ be the finite set of primes dividing every coefficient of f.*

(a) $\rho(mn) = \rho(m)\rho(n)$ *whenever* $(m, n) = 1$.

(b) $\rho(p^2) \le p\rho(p)$ *and* $\rho(p^j) \le p^j$ *for* $j \ge 0$.

(c) If $p \notin \mathcal{B}$, *then* $\rho(p) \le d$ *and* $\rho(p^2) \le dp$.

(d) $\rho(p) \le \min\{C_\rho, p\}$ *and* $\rho(p^2) \le C_\rho p$ *for every prime* p.

No assumption on the content, discriminant or factorisation of f is needed.

Proof. (a) The Chinese remainder theorem identifies residues modulo mn with pairs of residues modulo m and n, and $f(a) \equiv 0 \bmod mn$ if and only if both components are roots. (b) Each root modulo p^2 reduces to a root modulo p, and each residue modulo p has p lifts modulo p^2. (c) For $p \notin \mathcal{B}$ the reduction of f modulo p is a nonzero polynomial of degree at most d over the field $\mathbb{F}_p$, so it has at most d roots; then use (b). (d) If $p \nmid A_d$, (c) applies. If $p \mid A_d$, then $\rho(p) \le p \le |A_d|$. The bound for $\rho(p^2)$ follows from (b). □

Chapter 3

Local densities and the singular series

This chapter collects the arithmetic of a single modulus: the density $g(q,l)$, its mean square over residue classes, and the expansion of the singular series $\mathfrak{S}$ in Ramanujan sums. The same coefficient $\lambda(r)$ governs both the mean square of the densities and the singular series. This shared coefficient is the reason why, in Chapter 9, the density term and the off-diagonal term produce the same constant R_f and cancel. (Archived source: proposition P1.)

3.1 The density of squarefree numbers in a residue class

The starting point is the identity

$$\mu^2(n) = \sum_{m^2 \mid n} \mu(m) \qquad (n \geq 1), \tag{3.1}$$

which follows by multiplying local factors: a prime p with $p^2 \mid n$ contributes the factor $1 + \mu(p) = 0$, while all other primes contribute 1.

Proposition 3.1. *For every integer $q \geq 1$, every integer l and real $x \geq 1$,*

$$\bigl|Q(x;q,l) - x\,g(q,l)\bigr| \leq 3\sqrt{x}, \qquad g(q,l) = \frac{1}{q} \sum_{\substack{m \geq 1 \\ (m^2,q) \mid l}} \frac{\mu(m)(m^2,q)}{m^2}, \tag{3.2}$$

where the series converges absolutely; in particular $g(q,l) = \lim_{x\to\infty} Q(x;q,l)/x$. Equivalently, $g(q,l) = 0$ if some p^2 divides both q and l, and otherwise

$$g(q,l) = \frac{c}{q} \prod_{p \mid q} \Bigl(1 - \frac{1}{p^2}\Bigr)^{-1} \prod_{\substack{p \| q \\ p \mid l}} \Bigl(1 - \frac{1}{p}\Bigr). \tag{3.3}$$

In particular $0 \leq g(q,l) \leq 1/q$. For $l = 0$ the series reads $g(q,0) = \sum_{m \geq 1} \mu(m)/[q,m^2]$.

Proof. Let $M = \lfloor \sqrt{x} \rfloor$ and $d_m = (m^2, q)$. By (3.1), $Q(x;q,l) = \sum_{m \leq M} \mu(m) N_m$, where N_m counts $n \leq x$ with $n \equiv l \bmod q$ and $m^2 \mid n$. These two congruences are soluble exactly when $d_m \mid l$: writing $n = m^2 u$ turns them into $m^2 u \equiv l \bmod q$, and after division by d_m the coefficient m^2/d_m is coprime to q/d_m. When soluble, the solutions form one residue class modulo qm^2/d_m, so N_m differs from $x d_m/(qm^2)$ by at most one. Summing over $m \leq M$,

$$Q(x;q,l) = \frac{x}{q} \sum_{\substack{m \leq M \\ d_m \mid l}} \frac{\mu(m) d_m}{m^2} + E_1, \qquad |E_1| \leq M.$$

Each term of the series in (3.2) has modulus at most m^{-2} because $d_m \le q$, so the series converges absolutely, and by Lemma 2.1(a) the omitted tail contributes at most $x\sum_{m>M} m^{-2} \le x/M$. Since $\sqrt{x}/2 \le M \le \sqrt{x}$, the total error is at most $M + x/M \le 3\sqrt{x}$. Dividing by x and letting $x \to \infty$ identifies the limit.

For the closed form, only squarefree m contribute. By absolute convergence (Lemma 2.4) the series factors over primes. If $t = v_p(q)$, including p in m contributes $-p^{\min(2,t)-2}$, and is allowed exactly when $p^{\min(2,t)} \mid l$. Hence the local factor of $q\,g(q,l)$ is

$$1 - p^{-2}\ (t=0), \qquad \begin{cases} 1 & (p \nmid l) \\ 1-p^{-1} & (p \mid l) \end{cases}\ (t=1), \qquad \begin{cases} 1 & (p^2 \nmid l) \\ 0 & (p^2 \mid l) \end{cases}\ (t \ge 2),$$

these conditions being meaningful also for $l = 0$. Combining with (2.1) gives (3.3) and the zero case. All local factors lie in $[0,1]$, so $0 \le q\,g(q,l) \le 1$. For $l = 0$ every m is admissible and $d_m/(qm^2) = 1/[q,m^2]$. □

3.2 The mean square of the densities

Define the multiplicative function G by $G(1) = 1$, $G(p) = G(p^2) = -1/(p^2-1)$ and $G(p^j) = 0$ for $j \ge 3$, and put

$$\gamma(r) = c\,G(r), \qquad \lambda(r) = \varphi(r)\gamma(r)^2. \tag{3.4}$$

Thus $\lambda(r) = c^2\varphi(r)\prod_{p\mid r}(p^2-1)^{-2}$ if no prime cube divides r, and $\lambda(r) = 0$ otherwise. Its prime-power values are

$$\lambda(p) = c^2\frac{p-1}{(p^2-1)^2}, \qquad \lambda(p^2) = c^2\frac{p(p-1)}{(p^2-1)^2}. \tag{3.5}$$

Proposition 3.2. *(a) For every $q \ge 1$, $q\sum_{l=1}^{q} g(q,l)^2 = \sum_{r\mid q}\lambda(r)$.*

(b) $\sum_{r\ge1}\lambda(r) = c$.

(c) Every r in the support of λ can be written uniquely as $r = ab^2$ with a, b squarefree and coprime, and $0 \le \lambda(ab^2) \le c^2a^{-3}b^{-2}$.

(d) For every real $R > 0$,

$$\mathcal{T}(R) := \sum_{r>R}\lambda(r) \le \frac{10}{3}R^{-1/2}, \qquad \mathcal{N}(R) := \sum_{r\le R} r\lambda(r) \le \frac{5}{3}R^{1/2}.$$

In particular there is an absolute constant C_0 with $\mathcal{T}(R) \le C_0R^{-1/2}(\log 2R)^2$ and $\mathcal{N}(R) \le C_0R^{1/2}(\log 2R)^2$ for $R \ge 2$.

Proof. (a) For $t \ge 1$ define $b_{p,t}(l) = 1 - p^{-1}$ if $t = 1$ and $p \mid l$; $b_{p,t}(l) = 0$ if $t \ge 2$ and $p^2 \mid l$; and $b_{p,t}(l) = 1$ otherwise. By Proposition 3.1, including its zero case,

$$q\,g(q,l) = c\prod_{p^t\,\|\,q}\Big(1-\frac{1}{p^2}\Big)^{-1} b_{p,t}(l).$$

For $t = 1$ there are $p-1$ residues with $b_{p,1} = 1$ and one with $b_{p,1} = 1 - p^{-1}$, so

$$\Big(1-\frac{1}{p^2}\Big)^{-2}\frac{1}{p}\sum_{l \bmod p} b_{p,1}(l)^2 = \frac{p^4-2p^2+p}{(p^2-1)^2} = 1 + \frac{p-1}{(p^2-1)^2} = 1 + \frac{\lambda(p)}{c^2}.$$

For $t \ge 2$, exactly p^{t-2} of the p^t residues have $b_{p,t} = 0$ and the others have $b_{p,t} = 1$, so

$$\Big(1 - \frac{1}{p^2}\Big)^{-2} \frac{1}{p^t} \sum_{l \bmod p^t} b_{p,t}(l)^2 = \Big(1 - \frac{1}{p^2}\Big)^{-1} = 1 + \frac{1}{p^2 - 1} = 1 + \frac{\lambda(p) + \lambda(p^2)}{c^2},$$

using $(p-1) + p(p-1) = p^2 - 1$. By the Chinese remainder theorem the residues modulo q correspond to tuples of residues modulo the prime powers $p^t \parallel q$, so the average of $(q\, g(q,l))^2$ over $l \bmod q$ factors into these local averages. Hence

$$q \sum_{l=1}^{q} g(q,l)^2 = \frac{1}{q} \sum_{l=1}^{q} (q\, g(q,l))^2 = c^2 \prod_{p^t \parallel q} \Big(1 + \sum_{j=1}^{t} \frac{\lambda(p^j)}{c^2}\Big) = \sum_{r \mid q} \lambda(r).$$

(c) The representation is standard. With $A_p = p^3/((p-1)(p+1)^2)$, (3.5) reads $\lambda(p) = c^2 A_p / p^3$ and $\lambda(p^2) = c^2 A_p / p^2$. Since $(p-1)(p+1)^2 - p^3 = p^2 - p - 1 \ge 1$, we have $0 < A_p \le 1$, and multiplying over the prime factors gives the bound.

(b) By (c), $\sum_r \lambda(r) \le c^2 \sum_a a^{-3} \sum_b b^{-2} < \infty$. The function λ / c^2 is multiplicative (note $\lambda(1) = c^2$), and $1 + (\lambda(p) + \lambda(p^2))/c^2 = (1 - p^{-2})^{-1}$. Hence Lemma 2.4 and (2.1) give $\sum_r \lambda(r) = c^2 \prod_p (1 - p^{-2})^{-1} = c^2 \zeta(2) = c$.

(d) Dropping the conditions on a, b in (c) only enlarges the nonnegative sums. By Lemma 2.1(a) with $T = \sqrt{R/a}$,

$$\mathcal{T}(R) \le c^2 \sum_{a \ge 1} a^{-3} \sum_{b > \sqrt{R/a}} b^{-2} \le 2c^2 R^{-1/2} \sum_a a^{-5/2} \le \frac{10}{3} R^{-1/2},$$

using $c \le 1$ and $\sum_a a^{-5/2} \le 1 + \int_1^\infty t^{-5/2}\, \mathrm{d}t = \frac{5}{3}$. Similarly $r\lambda(r) \le c^2 a^{-2}$ for $r = ab^2$, and for each $a \le R$ there are at most $\sqrt{R/a}$ values of b, so $\mathcal{N}(R) \le c^2 R^{1/2} \sum_{a \le R} a^{-5/2} \le \frac{5}{3} R^{1/2}$. For $R \ge 2$, $\log 2R > 1$, which gives the weaker forms with $C_0 = 10/3$. □

3.3 The singular series

For $n \ge 1$ define

$$\mathfrak{S}(n) = \prod_p \Big(1 - \frac{2}{p^2}\Big) \prod_{p^2 \mid n} \frac{p^2 - 1}{p^2 - 2} = \prod_p \Big(1 - \frac{2}{p^2} + \frac{\mathbf{1}_{p^2 \mid n}}{p^2}\Big). \tag{3.6}$$

The second form shows that every local factor differs from 1 by at most $2/p^2$, so the product converges absolutely (Lemma 2.3). This is the singular series $S_2(n)$ that appears in the asymptotic formula for representations of n as a sum of two squarefree numbers in [1, Corollary 3]; here it governs pairs with a given difference (Chapter 5).

Proposition 3.3. *For every integer $n \ge 1$ the series $\sum_{r \ge 1} \gamma(r)^2 c_r(n)$ converges absolutely, with $\sum_r |\gamma(r)^2 c_r(n)| \le c$, and*

$$\mathfrak{S}(n) = \sum_{r \ge 1} \gamma(r)^2 c_r(n) = \sum_{r \ge 1} \lambda(r) \frac{c_r(n)}{\varphi(r)}. \tag{3.7}$$

Proof. By Lemma 2.5(b), $|\gamma(r)^2 c_r(n)| \le \lambda(r)$, so the series converges absolutely with the stated bound by Proposition 3.2(b). The summand is c^2 times the function $G(r)^2 c_r(n)$, which is multiplicative in r (Lemma 2.5(b)) and vanishes unless r is cube-free. Hence Lemma 2.4 and (2.2) give

$$\sum_{r \ge 1} \gamma(r)^2 c_r(n) = c^2 \prod_p \Big(1 + \frac{c_p(n) + c_{p^2}(n)}{(p^2 - 1)^2}\Big) = \prod_p \Big(1 - \frac{1}{p^2}\Big)^2 \Big(1 + \frac{p^2 \mathbf{1}_{p^2 \mid n} - 1}{(p^2 - 1)^2}\Big),$$

where the factor c^2 was distributed over the primes using (2.1). If $p^2 \nmid n$, the local factor equals $((p^2-1)^2-1)/p^4 = 1-2/p^2$; if $p^2 \mid n$, it equals $(1-p^{-2})^2(1+(p^2-1)^{-1}) = 1-p^{-2}$. This is the second form of (3.6). □

Chapter 4

Expanding the variance

We open the square in the definition of V_f and separate three kinds of terms: a diagonal term, a mixed term involving the densities, and an off-diagonal term that counts pairs of squarefree numbers whose difference is a multiple of the modulus. The moduli $f(k)$ with k below a cutoff z are removed by a known upper bound. (Archived source: proposition P2.)

4.1 Notation

For real $x \geq 2$ and integers $1 \leq m \leq x$ put

$$C_x(m) = \sum_{1 \leq n \leq x-m} \mu^2(n)\mu^2(n+m). \tag{4.1}$$

For a modulus q write $H(q) = \sum_{l=1}^{q} g(q,l)^2$. For real $1 \leq z \leq y$ with $Z = f(z) \leq Y = f(y) \leq x$ define

$$V_f(x;z,y) = \sum_{z<k\leq y} f'(k) \sum_{l=1}^{f(k)} E(x; f(k), l)^2, \tag{4.2}$$

$$T_f(z,y) = \sum_{z<k\leq y} f'(k) \sum_{l=1}^{f(k)} g(f(k),l)^2 = \sum_{z<k\leq y} f'(k) H(f(k)), \tag{4.3}$$

$$S_f(x;z,y) = \sum_{z<k\leq y} f'(k) \sum_{1 \leq h \leq x/f(k)} C_x(hf(k)). \tag{4.4}$$

Since f is increasing, $f(k) \leq Y \leq x$ for $k \leq y$, and every argument $hf(k)$ of C_x in (4.4) lies in $[1,x]$. Note that $V_f(x,y) = V_f(x,z) + V_f(x;z,y)$ for $1 \leq z \leq y$.

4.2 Counting pairs in a residue class

Lemma 4.1. *For real $x \geq 2$ and every integer $1 \leq q \leq x$,*

$$\sum_{l=1}^{q} Q(x;q,l)^2 = \sum_{n \leq x} \mu^2(n) + 2 \sum_{1 \leq h \leq x/q} C_x(hq).$$

Proof. The left side counts ordered pairs (n_1, n_2) of squarefree integers in $[1,x]$ with $n_1 \equiv n_2 \bmod q$. Pairs with $n_1 = n_2$ give the first term. For the others, the positive difference is hq for exactly one integer $1 \leq h \leq x/q$; the pairs with $n_2 - n_1 = hq$ are counted by $C_x(hq)$, and reversing the order gives the factor two. □

Lemma 4.2. *For every integer $q \ge 1$ and real $x \ge 1$, the mixed sum*

$$\Delta(x;q) = \sum_{l=1}^{q} g(q,l)E(x;q,l)$$

satisfies $|\Delta(x;q)| \le 3\sqrt{x}\,\tau(q)/q$.

Proof. Put $q_1 = \prod_{p\|q} p$ and $A_q = \prod_{p\mid q}(1-p^{-2})^{-1}$. By (2.1), $cA_q = \prod_{p\nmid q}(1-p^{-2}) \le 1$. For squarefree n the zero case of Proposition 3.1 cannot occur, and expanding $\prod_{p\mid(q_1,n)}(1-p^{-1})$ in (3.3) gives

$$\mu^2(n)g(q,n) = \frac{cA_q}{q}\mu^2(n)\sum_{a\mid q_1,\ a\mid n}\frac{\mu(a)}{a}.$$

The squarefree $n \le x$ divisible by a are counted by $Q(x;a,0)$, and Proposition 3.1 gives $|Q(x;a,0) - x\,g(a,0)| \le 3\sqrt{x}$. Hence, with $B(q) = \frac{cA_q}{q}\sum_{a\mid q_1}\frac{\mu(a)}{a}\,g(a,0)$,

$$\Big|\sum_{n\le x}\mu^2(n)g(q,n) - xB(q)\Big| \le \frac{3\sqrt{x}}{q}\sum_{a\mid q_1}\frac{1}{a} \le \frac{3\sqrt{x}\,\tau(q)}{q}. \tag{4.5}$$

Since $g(q,n)$ depends only on $n \bmod q$, $\sum_{n\le t}\mu^2(n)g(q,n) = \sum_{l=1}^{q} g(q,l)Q(t;q,l)$. Dividing by t and letting $t\to\infty$, the right side tends to $H(q)$ and, by (4.5), the left side tends to $B(q)$. Thus $B(q) = H(q)$, and since $\Delta(x;q) = \sum_l g(q,l)Q(x;q,l) - xH(q) = \sum_{n\le x}\mu^2(n)g(q,n) - xH(q)$, the claim follows from (4.5). □

4.3 The variance on an interval of moduli

Proposition 4.3. *For every $\varepsilon > 0$ there is $C(f,\varepsilon) > 0$ such that, uniformly for real $x \ge 2$ and $1 \le z \le y$ with $Z = f(z) \le Y = f(y) \le x$,*

$$\Big|V_f(x;z,y) - \big[cx(Y-Z) + 2S_f(x;z,y) - x^2T_f(z,y)\big]\Big| \le C(f,\varepsilon)(x^{3/2+\varepsilon} + xY^{1-1/d}).$$

Proof. Fix $q = f(k) \le x$. Since $Q = E + xg$, expanding $\sum_l Q^2$ and using Lemma 4.1 and the definition of Δ gives the exact identity

$$\sum_{l=1}^{q}E(x;q,l)^2 = \sum_{n\le x}\mu^2(n) + 2\sum_{1\le h\le x/q}C_x(hq) - x^2H(q) - 2x\Delta(x;q).$$

Multiply by $f'(k)$ and sum over $z < k \le y$. With $\mathcal{W} = \sum_{z<k\le y}f'(k)$, the difference to be estimated is exactly

$$\mathcal{R} = \Big(\sum_{n\le x}\mu^2(n) - cx\Big)\mathcal{W} + cx(\mathcal{W} - (Y-Z)) - 2x\sum_{z<k\le y}f'(k)\Delta(x;f(k)).$$

Diagonal error. Proposition 3.1 with $q=1$ gives $|\sum_{n\le x}\mu^2(n) - cx| \le 3\sqrt{x}$, since $g(1,0) = \sum_m \mu(m)m^{-2} = c$. By (2.5) and $Y \le x$, the first term is at most $3(1+K_1)\sqrt{x}Y \le 3(1+K_1)x^{3/2}$.

Endpoint error. By (2.5), the second term is at most $cK_1xY^{1-1/d}$.

Mixed error. Let $\eta = \varepsilon/2$. By Lemma 4.2, Lemma 2.2 and $f(k) \le x$, each term satisfies $|f'(k)\Delta(x;f(k))| \le 3D_\eta x^{1/2+\eta}f'(k)/f(k)$. Summing with (2.4) and using $\log 2x \le J_\eta x^\eta$ (Lemma 2.1(d)), we find that the third term is at most

$$2x\cdot 3D_\eta x^{1/2+\eta}\cdot 2U\log(2x) \le 12D_\eta U J_\eta\, x^{3/2+\varepsilon}.$$

Adding the three bounds proves the proposition with $C(f,\varepsilon) = 3(1+K_1) + 12D_{\varepsilon/2}UJ_{\varepsilon/2} + cK_1$. The argument also covers $z = y$ and intervals containing no integer. □

4.4 The initial moduli

For the moduli $f(k)$ with $k \le z$ we use the full-variance bound of Brüdern, Granville, Perelli, Vaughan and Wooley. In their notation μ_k is the indicator of the k-free numbers, so $\mu_2 = \mu^2$, and their parameter k is unrelated to our summation variable.

Quoted result 4.4 ([1, Lemma 2.2 and (2.6)–(2.7)])**.** *For positive integers a, q define $E_k(x;q,a)$ by*

$$\sum_{\substack{1\le n\le x\\ n\equiv a \pmod q}} \mu_k(n) = \frac{x}{\zeta(k)} g_k(q,(q,a)) + E_k(x;q,a), \qquad \Upsilon_k(x,Q) = \sum_{1\le q\le Q}\sum_{a=1}^{q} |E_k(x;q,a)|^2.$$

Suppose that Q and x are positive real numbers and that $k \ge 2$ is an integer. Then for each $\varepsilon > 0$,

$$\Upsilon_k(x,Q) \ll x^{2/k+\varepsilon} Q^{2-2/k} \quad (1 \le Q \le x), \qquad \Upsilon_k(x,Q) \ll Q^2 \log(2Q) \quad (Q > x).$$

The inner sum in Υ_k runs over all residue classes, coprime to q or not. We check that the normalisation agrees with ours without using the formula for g_k. Fix $1 \le a \le q$ and apply the lemma with $k = 2$, $Q = q$, $\varepsilon = \frac{1}{2}$ and $x = t \ge \max\{2, q\}$. By positivity, $|E_2(t;q,a)|^2 \le \Upsilon_2(t,q) \ll t^{3/2} q$, so $E_2(t;q,a)/t \to 0$ as $t \to \infty$. Dividing the definition by t therefore shows that $\zeta(2)^{-1} g_2(q,(q,a)) = \lim_t Q(t;q,a)/t = g(q,a)$. Hence $E_2 = E$, and with $k = 2$ the lemma states that

$$V(x,Q) = \sum_{q\le Q}\sum_{l=1}^{q} E(x;q,l)^2 \le B_\varepsilon x^{1+\varepsilon} Q \qquad (1 \le Q \le x) \tag{4.6}$$

for every $\varepsilon > 0$, where B_ε depends only on ε. The same bound is recalled as (1.5) in [11].

Proposition 4.5. *For every $\varepsilon > 0$ there is $C(f,\varepsilon)$ such that, uniformly for real $x \ge 2$ and real $z \ge 1$ with $f(z) \le x$,*

$$0 \le V_f(x,z) \le C(f,\varepsilon)\, x^{1+\varepsilon} f(z)^{2-1/d}.$$

Proof. Let $Z = f(z)$. By Lemma 2.6(a) the moduli $f(k)$, $k \le z$, are distinct integers in $[2, Z]$, and by Lemma 2.6(e) their weights satisfy $0 < f'(k) \le \kappa_2 \kappa_1^{-\alpha} Z^\alpha$ with $\alpha = 1 - 1/d$. Each inner sum $\sum_l E^2$ is nonnegative. We may therefore first bound each weight by this maximum and then enlarge the set of moduli to all $q \le Z$:

$$0 \le V_f(x,z) \le \kappa_2 \kappa_1^{-\alpha} Z^\alpha \sum_{q\le Z}\sum_{l=1}^{q} E(x;q,l)^2 \le \kappa_2 \kappa_1^{-\alpha} B_\varepsilon\, x^{1+\varepsilon} Z^{1+\alpha}.$$

Here (4.6) is applicable because $1 \le 2 \le Z \le x$, and $1 + \alpha = 2 - 1/d$. □

Remark 4.6. It is essential to use positivity before replacing the weights by their maximum. Bounding a sparse sum by the full sum over all $q \le Z$ costs the factor $Z^{1-1/d}$. This loss is one of the two sources of the exponent $1/(4d-1)$ in Theorem 1.1.

Chapter 5

Correlations of squarefree numbers

The off-diagonal term S_f is a weighted sum of the correlations $C_x(m)$ of (4.1). We need an asymptotic formula for $C_x(m)$ that is uniform in the whole range $1 \le m \le x$, including differences m close to x, where the number $x - m$ of available pairs is small. Counts of pairs of squarefree numbers with a given difference have a long history (see for example [7], and the discussion of the related binary additive problem in [1, §1]). The archived proof gives a self-contained elementary argument with an explicit absolute constant. The archived plan anticipated a circle-method argument based on [1], but the proof does not need it. (Archived source: proposition P3.)

Proposition 5.1. *There is an absolute constant K_{c} such that for all real $x \ge 2$ and all integers $1 \le m \le x$,*

$$\bigl|C_x(m) - (x-m)\mathfrak{S}(m)\bigr| \le K_{\mathrm{c}}\, x^{4/5}.$$

One may take $K_{\mathrm{c}} = 67 + 2D_{1/10}$, with $D_{1/10}$ as in Lemma 2.2. In particular, for every $\varepsilon > 0$ the bound $C(\varepsilon)x^{4/5+\varepsilon}$ holds, which is the form recorded in the archived plan.

Proof. Put $L = x - m \ge 0$ and $H = \lfloor x^{2/5} \rfloor$, so that $\frac{1}{2}x^{2/5} \le H \le x^{2/5}$ and $H \ge 1$. All bounds below hold for every real $L \ge 0$, including $L < 1$, when $C_x(m) = 0$.

Truncation. For $k \ge 1$ split (3.1) as $\mu^2(k) = T_H(k) + R_H(k)$, where

$$T_H(k) = \sum_{\substack{e \le H \\ e^2 \mid k}} \mu(e), \qquad R_H(k) = \sum_{\substack{e > H \\ e^2 \mid k}} \mu(e).$$

Then $|T_H(k)| \le \tau(k)$ and $|R_H(k)| \le \#\{e > H : e^2 \mid k\}$. Each of the sets $\{n : 1 \le n \le L\}$ and $\{n + m : 1 \le n \le L\}$ consists of integers in $[1, x]$, because $m + L = x$. Each contains at most x/e^2 multiples of e^2. Therefore, by Lemma 2.1(a),

$$\sum_{1 \le n \le L} |R_H(n)| \le \sum_{e > H} \frac{x}{e^2} \le \frac{x}{H}, \qquad \sum_{1 \le n \le L} |R_H(n+m)| \le \frac{x}{H}. \tag{5.1}$$

The exact identity

$$\mu^2(n)\mu^2(n+m) - T_H(n)T_H(n+m) = R_H(n)\mu^2(n+m) + T_H(n)R_H(n+m)$$

together with $0 \le \mu^2 \le 1$ and $|T_H(n)| \le \tau(n) \le D_{1/10}x^{1/10}$ (Lemma 2.2) bounds each summand by $|R_H(n)| + D_{1/10}x^{1/10}|R_H(n+m)|$. Summing with (5.1),

$$\Bigl|C_x(m) - \sum_{1 \le n \le L} T_H(n)T_H(n+m)\Bigr| \le (1 + D_{1/10}x^{1/10})\frac{x}{H} \le 2x^{3/5} + 2D_{1/10}x^{7/10}. \tag{5.2}$$

Counting the truncated congruences. Let $e_1, e_2 \le H$ and $g = (e_1, e_2)$. The congruences $n \equiv 0 \bmod e_1^2$ and $n \equiv -m \bmod e_2^2$ are soluble exactly when $g^2 \mid m$. Necessity follows by reducing modulo g^2. For sufficiency write $e_1 = ga$, $e_2 = gb$ with $(a,b) = 1$ and $n = e_1^2 u$; the second congruence becomes $a^2 u \equiv -m/g^2 \bmod b^2$, which determines u modulo b^2. Hence the solutions form one residue class modulo $e_1^2 e_2^2/g^2$. For any residue class modulo h with representative r, the number of its members in $[1, L]$ is $\lfloor (L-r)/h \rfloor - \lfloor -r/h \rfloor = L/h + \theta$ with $|\theta| \le 1$. Expanding both truncated sums therefore gives

$$\sum_{1 \le n \le L} T_H(n) T_H(n+m) = L\, \mathcal{A}_H(m) + E_H, \qquad |E_H| \le H^2 \le x^{4/5}, \tag{5.3}$$

where

$$\mathcal{A}_H(m) = \sum_{\substack{e_1, e_2 \le H \\ (e_1,e_2)^2 \mid m}} \frac{\mu(e_1)\mu(e_2)(e_1,e_2)^2}{e_1^2 e_2^2}.$$

Completing the series. Let $\mathcal{A}(m)$ denote the same sum over all $e_1, e_2 \ge 1$. Writing $e_1 = ga$, $e_2 = gb$ and dropping the coprimality of a, b shows absolute convergence: $\sum_{e_1,e_2} (e_1,e_2)^2/(e_1 e_2)^2 \le \sum_g g^{-2} \sum_a a^{-2} \sum_b b^{-2} \le 8$. For the tail with $e_1 > H$, the same substitution gives the bound $\sum_g g^{-2} (\sum_{a > H/g} a^{-2})(\sum_b b^{-2})$. For $g \le H$ the summand is at most $g^{-2}(2g/H) \cdot 2 = 4/(gH)$ by Lemma 2.1(a), and summing over $g \le H$ gives at most $4(1 + \log H)/H$. For $g > H$ the summand is at most $4/g^2$, and these sum to at most $4/H$. By symmetry the tail with $e_2 > H$ has the same bound, so

$$|\mathcal{A}(m) - \mathcal{A}_H(m)| \le \frac{8(2 + \log H)}{H}, \tag{5.4}$$

uniformly in m, because the condition $(e_1, e_2)^2 \mid m$ was dropped.

The Euler product. By absolute convergence, $\mathcal{A}(m)$ factors over primes, the argument of Lemma 2.4 applying verbatim to the double sum. Only the exponents 0 and 1 occur in e_1, e_2. At a prime p the four choices contribute 1, $-p^{-2}$, $-p^{-2}$, and $p^{-2}\mathbf{1}_{p^2 \mid m}$, so by (3.6)

$$\mathcal{A}(m) = \prod_p \Big(1 - \frac{2}{p^2} + \frac{\mathbf{1}_{p^2 \mid m}}{p^2}\Big) = \mathfrak{S}(m).$$

Conclusion. Since $0 \le L \le x$, $H \ge \frac{1}{2} x^{2/5}$ and $\log H \le \frac{2}{5} \log x \le 2x^{1/5}$ (Lemma 2.1(d)), (5.4) gives

$$L\, |\mathcal{A}_H(m) - \mathfrak{S}(m)| \le 16 x^{3/5} (2 + \tfrac{2}{5} \log x) \le 32 x^{3/5} + 32 x^{4/5} \le 64 x^{4/5}.$$

Combining with (5.2) and (5.3) gives $|C_x(m) - L\mathfrak{S}(m)| \le (2 + 2D_{1/10} + 1 + 64) x^{4/5}$. □

Remark 5.2. The constant $D_{1/10}$ is astronomically large, but only its existence matters. For the application in Chapter 9 only the exponent $4/5$ is relevant: summing the error over the roughly $x \log x$ weighted pairs (h, k) costs $x^{9/5+\varepsilon}$, and this caps the saving in Theorem 1.1 at $\eta_d \le 1/5$.

Chapter 6

Averaging over polynomial moduli

After Chapters 4 and 5 the variance is expressed through two sums over k: the density term $T_f(z,y)$ and the main part of the off-diagonal term,

$$M_f(x;z,y)=\sum_{z<k\le y} f'(k)\sum_{1\le h\le x/f(k)}(x-hf(k))\mathfrak{S}(hf(k)). \tag{6.1}$$

Both involve arithmetic functions of $f(k)$. Expanding them over r with the coefficients $\lambda(r)$ of Chapter 3 makes the dependence on k periodic modulo r. Partial summation then replaces the sum over k by an integral in the variable $t=f(k)$, with $f'(k)$ acting as the Jacobian. This is the role of the weight $f'(k)$. (Archived source: proposition P4, a crux of the plan.)

6.1 The averaged singular series

For $h\ge 1$ and a prime p put

$$\vartheta_p(h)=\frac{1}{p^2}\#\{a \bmod p^2:\ p^2\mid hf(a)\}.$$

If $j=\max\{2-v_p(h),0\}$, then $p^2\mid hf(a)$ is equivalent to $p^j\mid f(a)$. Counting lifts from residues modulo p^j to residues modulo p^2 gives

$$\vartheta_p(h)=\frac{\rho(p^j)}{p^j}\in[0,1],\qquad\text{so}\qquad 1-\frac{2}{p^2}+\frac{\vartheta_p(h)}{p^2}=w_p(h), \tag{6.2}$$

where $w_p(h)$ is the local factor in (1.5). Comparing with the second form of (3.6), $W_f(h)=\prod_p w_p(h)$ is what one obtains by replacing the indicator $\mathbf{1}_{p^2\mid hf(k)}$ in $\mathfrak{S}(hf(k))$ by its average $\vartheta_p(h)$ over k. Lemma 6.2 below makes this precise.

Lemma 6.1. *For every $h\ge 1$ and prime p, $1-2p^{-2}\le w_p(h)\le 1-p^{-2}$. Hence the product $W_f(h)$ converges absolutely, uniformly in h, and $0<W_f(h)\le 1$. The series*

$$\widetilde{R}_f=\sum_{r\ge 1}\lambda(r)\frac{\rho(r)}{r}$$

converges absolutely and $0\le\widetilde{R}_f\le c$.

Proof. The bounds for w_p follow from $\vartheta_p(h)\in[0,1]$, so $|w_p(h)-1|\le 2p^{-2}$, and Lemma 2.3 applies. The factors are positive, so $W_f(h)>0$, and they are at most one. Finally $0\le\rho(r)/r\le 1$ and $\sum_r\lambda(r)=c$. □

For $r, h \geq 1$ define

$$\beta_r(h) = \frac{1}{r} \sum_{a=1}^{r} \frac{c_r(hf(a))}{\varphi(r)}, \tag{6.3}$$

the mean value of the r-periodic sequence

$$k \longmapsto \frac{c_r(hf(k))}{\varphi(r)}.$$

Note $|\beta_r(h)| \leq 1$ by Lemma 2.5(b).

Lemma 6.2. *For every $h \geq 1$, $\sum_{r \geq 1} \lambda(r)\beta_r(h) = W_f(h)$, the series converging absolutely and uniformly in h.*

Proof. Absolute and uniform convergence follow from $|\beta_r(h)| \leq 1$ and $\sum_r \lambda(r) = c$. If $(r, s) = 1$, the Chinese remainder theorem identifies $a \bmod rs$ with pairs $(a \bmod r,\ a \bmod s)$. Moreover $c_{rs} = c_r c_s$, $c_r(hf(a))$ depends only on $a \bmod r$, and $\varphi(rs) = \varphi(r)\varphi(s)$. Hence $\beta_{rs}(h) = \beta_r(h)\beta_s(h)$. Thus $\lambda(r)\beta_r(h)$ is c^2 times a multiplicative function supported on cube-free r, and Lemma 2.4 gives

$$\sum_r \lambda(r)\beta_r(h) = c^2 \prod_p \Big(1 + \frac{(p-1)\beta_p(h) + (p^2-p)\beta_{p^2}(h)}{(p^2-1)^2}\Big)$$

by (3.5). The average defining $\beta_p(h)$ may equally be taken over $a \bmod p^2$, since its summand has period p. Therefore, by (2.2),

$$(p-1)\beta_p(h) + (p^2-p)\beta_{p^2}(h) = \frac{1}{p^2} \sum_{a \bmod p^2} (c_p + c_{p^2})(hf(a)) = p^2\vartheta_p(h) - 1.$$

Distributing $c^2 = \prod_p (1-p^{-2})^2$ over the primes, the local factor becomes

$$\Big(1 - \frac{1}{p^2}\Big)^2 \Big(1 + \frac{p^2\vartheta_p(h) - 1}{(p^2-1)^2}\Big) = \frac{(p^2-1)^2 + p^2\vartheta_p(h) - 1}{p^4} = 1 - \frac{2}{p^2} + \frac{\vartheta_p(h)}{p^2} = w_p(h),$$

by (6.2). No restriction on f modulo p has been used. □

6.2 A partial summation lemma

Lemma 6.3. *Let b be a periodic sequence with integer period $r \geq 1$, $|b(k)| \leq 1$, and mean $m_b = r^{-1}\sum_{a=1}^{r} b(a)$. Let $1 \leq u \leq v$ be real and let F be continuously differentiable on $[u, v]$. Then*

$$\Big|\sum_{u < k \leq v} b(k)F(k) - m_b \int_u^v F(t)\,\mathrm{d}t\Big| \leq 4r\Big(|F(v)| + \int_u^v |F'(t)|\,\mathrm{d}t\Big).$$

Proof. Let $P(t) = \sum_{k \leq t} b(k) - m_b t$. Writing $t = qr + s$ with $q = \lfloor t/r \rfloor$ and $0 \leq s < r$, the q complete periods cancel, the remaining at most r terms have modulus at most 1, and $|m_b s| \leq r$; so $|P(t)| \leq 2r$. Hence $D_u(t) = \sum_{u < k \leq t} b(k) - m_b(t-u) = P(t) - P(u)$ satisfies $|D_u(t)| \leq 4r$ and $D_u(u) = 0$. Riemann–Stieltjes integration by parts gives the identity

$$\sum_{u < k \leq v} b(k)F(k) - m_b \int_u^v F(t)\,\mathrm{d}t = \int_u^v F(t)\,\mathrm{d}D_u(t) = F(v)D_u(v) - \int_u^v F'(t)D_u(t)\,\mathrm{d}t,$$

from which the bound follows. □

With $F = w = f'/f$ on $[z, y]$, Lemma 2.6(b) gives $|w(y)| \leq U/y \leq U/z$ and $\int_z^y |w'| \leq V/z$. Since $\int_z^y w = \log(Y/Z)$, we obtain for every r-periodic b as in Lemma 6.3 and $1 \leq z \leq y$:

$$\Big|\sum_{z < k \leq y} b(k)\frac{f'(k)}{f(k)} - m_b \log\frac{Y}{Z}\Big| \leq \frac{4D_w r}{z}, \qquad D_w = U + V. \tag{6.4}$$

6.3 The averaging proposition

Proposition 6.4. *There is $C(f) > 0$ such that the following holds uniformly for real $x \geq 2$, $1 \leq z \leq y$ with $Z = f(z) \leq Y = f(y) \leq x$, and real $2 \leq R \leq x$. With*

$$\mathcal{A} = \mathcal{A}(x, Z, R) = Z^{-1/d} R^{1/2} (\log 2R)^2 + R^{-1/2} (\log 2R)^2 \log(2x)$$

and

$$I_f(x; Z, Y) = \int_Z^Y \sum_{1 \leq h \leq x/t} (x - ht) W_f(h) \, \mathrm{d}t, \tag{6.5}$$

we have

$$|T_f(z, y) - \widetilde{R}_f \log(Y/Z)| \leq C(f) \, \mathcal{A}, \tag{6.6}$$

$$|M_f(x; z, y) - I_f(x; Z, Y)| \leq C(f) \, x^2 \mathcal{A}. \tag{6.7}$$

No condition on the discriminant, content or factorisation of f is imposed.

Throughout the proof, $\mathcal{T}(R)$ and $\mathcal{N}(R)$ are as in Proposition 3.2(d), so that $\mathcal{T}(R) \leq C_0 R^{-1/2} (\log 2R)^2$ and $\mathcal{N}(R) \leq C_0 R^{1/2} (\log 2R)^2$. If $z = y$ both assertions are trivial, so we assume $z < y$.

Proof of (6.6). By Proposition 3.2(a) applied to $q = f(k)$,

$$T_f(z, y) = \sum_{z < k \leq y} \frac{f'(k)}{f(k)} \sum_{r \mid f(k)} \lambda(r).$$

The part with $r > R$ is nonnegative and at most $\mathcal{T}(R) \sum_{z<k\leq y} f'(k)/f(k) \leq 2U\mathcal{T}(R) \log(2x)$ by (2.4). For each $r \leq R$, the sequence $b(k) = \mathbf{1}_{r \mid f(k)}$ has period r and mean $\rho(r)/r$, so (6.4) gives

$$\sum_{z < k \leq y} \frac{f'(k)}{f(k)} \mathbf{1}_{r \mid f(k)} = \frac{\rho(r)}{r} \log \frac{Y}{Z} + e_r, \qquad |e_r| \leq \frac{4 D_w r}{z}.$$

Multiplying by $\lambda(r)$ and summing over $r \leq R$ produces an error at most $4 D_w \mathcal{N}(R)/z$. Completing $\sum_{r \leq R} \lambda(r)\rho(r)/r$ to $\widetilde{R}_f$ costs at most $\mathcal{T}(R) \log(Y/Z) \leq \mathcal{T}(R) \log(2x)$, because $\rho(r)/r \leq 1$ and $2 \leq Z \leq Y \leq x$. Since $Z \leq \kappa_2 z^d$, we have $z^{-1} \leq \kappa_2^{1/d} Z^{-1/d}$, and therefore

$$|T_f(z, y) - \widetilde{R}_f \log(Y/Z)| \leq C_0 \Big(4 D_w \kappa_2^{1/d} Z^{-1/d} R^{1/2} (\log 2R)^2 + (2U+1) R^{-1/2} (\log 2R)^2 \log(2x) \Big).$$

□

For the second assertion we first need a bound for the weight functions. Let $H_0 = \lfloor x/Z \rfloor$. For $1 \leq h \leq H_0$ define

$$v_h = \min\{y, \, f^{-1}(x/h)\}, \qquad F_h(t) = f'(t)(x - h f(t)) \quad (z \leq t \leq v_h),$$

which is legitimate because $x/h \geq Z = f(z)$ (Lemma 2.6(a)). Then $z \leq v_h \leq (x/(\kappa_1 h))^{1/d}$ and $0 \leq x - h f(t) \leq x$ on $[z, v_h]$. Put

$$K_2 = \begin{cases} \kappa_2 + \kappa_2^2/\kappa_1, & d = 1, \\ \kappa_2 + \kappa_2/(d-1) + \kappa_2^2/((2d-1)\kappa_1), & d \geq 2. \end{cases}$$

Lemma 6.5. *For $1 \leq h \leq H_0$, $|F_h(v_h)| + \displaystyle\int_z^{v_h} |F_h'(t)| \, \mathrm{d}t \leq K_2 \, x \, v_h^{d-1}$.*

Proof. We have $|F_h(v_h)| \le \kappa_2 x v_h^{d-1}$ and $|F_h'| \le |f''|(x - hf) + hf'^2$. For $d \ge 2$,

$$\int_z^{v_h} |F_h'| \le \kappa_2 x \int_z^{v_h} t^{d-2}\,\mathrm{d}t + h\kappa_2^2 \int_z^{v_h} t^{2d-2}\,\mathrm{d}t \le \frac{\kappa_2}{d-1} x v_h^{d-1} + \frac{h\kappa_2^2}{2d-1} v_h^{2d-1},$$

and $hv_h^d \le x/\kappa_1$ bounds the last term by $\kappa_2^2 x v_h^{d-1}/((2d-1)\kappa_1)$. For $d = 1$, $f'' = 0$ and $\int_z^{v_h} |F_h'| \le h\kappa_2^2 v_h \le \kappa_2^2 x/\kappa_1$. □

Proof of (6.7). *Truncation in r.* Put

$$\mathfrak{S}_R(n) = \sum_{r \le R} \lambda(r) \frac{c_r(n)}{\varphi(r)}, \qquad W_R(h) = \sum_{r \le R} \lambda(r)\beta_r(h).$$

By Proposition 3.3 and Lemma 6.2 and the bounds $|c_r(n)| \le \varphi(r)$ and $|\beta_r(h)| \le 1$,

$$|\mathfrak{S}(n) - \mathfrak{S}_R(n)| \le \mathcal{T}(R), \qquad |W_f(h) - W_R(h)| \le \mathcal{T}(R)$$

uniformly in n, h. Let M_R and I_R be M_f and I_f with $\mathfrak{S}$ and W_f replaced by $\mathfrak{S}_R$ and W_R. For fixed k there are at most $x/f(k)$ values of h, each with $0 \le x - hf(k) \le x$, so by (2.4)

$$|M_f - M_R| \le \mathcal{T}(R) x^2 \sum_{z < k \le y} \frac{f'(k)}{f(k)} \le 2U\mathcal{T}(R) x^2 \log(2x).$$

Similarly, for each $t \in [Z, Y]$ there are at most x/t values of h, so $|I_f - I_R| \le \mathcal{T}(R)x^2 \int_Z^Y \mathrm{d}t/t \le \mathcal{T}(R)x^2 \log(2x)$.

The finite approximations. All remaining sums are finite and may be rearranged. For integers $k \in (z, y]$, the condition $h \le x/f(k)$ is equivalent to $k \le v_h$, and $k > z$ forces $h < x/Z$. Hence

$$M_R = \sum_{r \le R} \lambda(r) \sum_{h \le H_0} \sum_{z < k \le v_h} F_h(k) \frac{c_r(hf(k))}{\varphi(r)}.$$

The sequence $k \mapsto c_r(hf(k))/\varphi(r)$ has period r, because f has integer coefficients, and its mean is $\beta_r(h)$. The substitution $t = f(u)$ gives

$$\int_z^{v_h} F_h(u)\,\mathrm{d}u = \int_Z^{\min\{Y, x/h\}} (x - ht)\,\mathrm{d}t,$$

and therefore $I_R = \sum_{r \le R} \lambda(r) \sum_{h \le H_0} \beta_r(h) \int_z^{v_h} F_h(u)\,\mathrm{d}u$. By Lemmas 6.3 and 6.5, the inner difference for fixed r and h is at most $4rK_2 x v_h^{d-1} \le 4rK_2 \kappa_1^{-(d-1)/d} x^{2-1/d} h^{-1+1/d}$. If $v_h = z$ the difference is zero. By Lemma 2.1(e), $\sum_{h \le H_0} h^{-1+1/d} \le d(x/Z)^{1/d}$, so for fixed r the total difference is at most $K_3 r x^2 Z^{-1/d}$ with $K_3 = 4dK_2\kappa_1^{-(d-1)/d}$. No logarithm arises here. Multiplying by $\lambda(r)$ and summing over $r \le R$ gives

$$|M_R - I_R| \le K_3 x^2 Z^{-1/d} \mathcal{N}(R).$$

Combining the three estimates proves

$$|M_f - I_f| \le C_0 x^2 \Big(K_3 Z^{-1/d} R^{1/2} (\log 2R)^2 + (2U+1) R^{-1/2} (\log 2R)^2 \log(2x) \Big).$$

Together with the first part, Proposition 6.4 holds with $C(f) = C_0 \max\{4D_w \kappa_2^{1/d}, K_3, 2U + 1\}$. □

Remark 6.6. The error $x^2 Z^{-1/d} R^{1/2}$ comes from the k-sampling of a weight of total variation about $x v_h^{d-1}$, summed over h with the h-dependent endpoint v_h. The archived plan points out that using the endpoint y for every h, instead of v_h, would lose the claimed range. This error is small only when $Z^{1/d}$, which is comparable to z, is large compared with $R^{1/2}(\log 2R)^2$. The tail term $R^{-1/2}(\log 2R)^2 \log(2x)$, on the other hand, needs $R^{1/2}$ to be large compared with $(\log 2R)^2 \log(2x)$. The choice $R = Z^{1/d}$ in Chapter 9 balances the two powers: both terms are then $Z^{-1/(2d)}$ up to logarithmic factors.

Chapter 7

The Dirichlet series of the averaged singular series

The averaged singular series $W_f(h)$ is a product of local factors that depend only on whether $p \nmid h$, $p \parallel h$ or $p^2 \mid h$. Its Dirichlet series therefore factors as $\zeta(s)$ times an Euler product, and a second factorisation exhibits a pole of $\zeta(2s+2)$ at $s = -\frac{1}{2}$. This pole, which is invisible in the region of absolute convergence, produces the main term of Theorem 1.1. (Archived source: proposition P5, a crux of the plan.)

7.1 Facts about the zeta function

Lemma 7.1. *(a) For every integer $N \geq 1$ and $\operatorname{Re} w > 0$, $w \neq 1$,*

$$\zeta(w) = \sum_{n \leq N} n^{-w} + \frac{N^{1-w}}{w-1} - w \int_N^\infty \{u\} u^{-w-1}\,\mathrm{d}u. \tag{7.1}$$

In $\operatorname{Re} w > 0$ the only pole of ζ is a simple pole at $w = 1$ with residue 1, and $\lim_{w\to 1}\bigl(\zeta(w) - \frac{1}{w-1}\bigr) = \gamma_E$, Euler's constant.

(b) Let $\mathcal{B}_2(u) = \{u\}^2 - \{u\} + \frac{1}{6}$. The formula

$$\zeta(w) = \frac{1}{w-1} + \frac{1}{2} + \frac{w}{12} - \frac{w(w+1)}{2} \int_1^\infty \mathcal{B}_2(u) u^{-w-2}\,\mathrm{d}u \tag{7.2}$$

continues ζ meromorphically to $\operatorname{Re} w > -1$, with the single pole at $w = 1$. In particular $\zeta(0) = -\frac{1}{2}$ and $-\frac{1}{4} \leq \zeta(-\frac{1}{2}) \leq -\frac{1}{6}$.

Proof. (a) For $\operatorname{Re} w > 1$ write $\sum_{n>N} n^{-w} = \int_N^\infty u^{-w}\,\mathrm{d}\lfloor u \rfloor$, replace $\lfloor u \rfloor = u - \{u\}$, and integrate the second part by parts; the boundary term at N vanishes because $\{N\} = 0$. This gives (7.1). The integral converges absolutely and locally uniformly for $\operatorname{Re} w > 0$ because $0 \leq \{u\} < 1$, which gives the continuation and the description of the pole. Subtracting $1/(w-1)$ and letting $w \to 1$ gives $\sum_{n \leq N} n^{-1} - \log N - \int_N^\infty \{u\} u^{-2}\,\mathrm{d}u$. The integral lies in $[0, 1/N]$, and letting $N \to \infty$ gives γ_E by its definition as $\lim(\sum_{n \leq N} n^{-1} - \log N)$.

(b) With $N = 1$, (7.1) becomes $\zeta(w) = \frac{1}{w-1} + \frac{1}{2} - w\int_1^\infty (\{u\} - \frac{1}{2}) u^{-w-1}\,\mathrm{d}u$. The function $\mathcal{B}_2$ is continuous, satisfies $|\mathcal{B}_2| \leq \frac{1}{6}$, and has derivative $2(\{u\} - \frac{1}{2})$ between integers. Integration by parts over $[1, T]$ for integers $T \geq 2$ gives

$$\int_1^T \Bigl(\{u\} - \frac{1}{2}\Bigr) u^{-w-1}\,\mathrm{d}u = \frac{T^{-w-1} - 1}{12} + \frac{w+1}{2} \int_1^T \mathcal{B}_2(u) u^{-w-2}\,\mathrm{d}u,$$

and letting $T \to \infty$ for $\operatorname{Re} w > 1$ yields (7.2). The last integral and its w-derivatives are dominated on $\operatorname{Re} w \geq \sigma_0 > -1$ by $(\log u)^m u^{-\sigma_0-2}/6$, so the integral is holomorphic for $\operatorname{Re} w > -1$. At $w = 0$ the formula gives $-1 + \frac{1}{2} = -\frac{1}{2}$. At $w = -\frac{1}{2}$ it gives $\zeta(-\frac{1}{2}) = -\frac{5}{24} + \frac{1}{8}\int_1^\infty \mathcal{B}_2(u) u^{-3/2}\,du$, and the integral has modulus at most $\frac{1}{6}\cdot 2 = \frac{1}{3}$. □

7.2 The local factors

For each prime p put

$$\alpha_p = 1 - \frac{2}{p^2} + \frac{\rho(p^2)}{p^4}, \qquad \beta_p = \frac{\rho(p)}{p^3} - \frac{\rho(p^2)}{p^4}, \qquad \delta_p = \frac{1}{p^2} - \frac{\rho(p)}{p^3}. \tag{7.3}$$

By (1.5), $w_p(h)$ equals α_p, $\alpha_p + \beta_p$ or $\alpha_p + \beta_p + \delta_p$ according as $p \nmid h$, $p \parallel h$ or $p^2 \mid h$. By Lemma 2.7 (with $C_\rho = \max\{d, |A_d|\}$),

$$1 - \frac{2}{p^2} \leq \alpha_p \leq 1, \quad 0 \leq \beta_p \leq \frac{\rho(p)}{p^3} \leq \frac{C_\rho}{p^3}, \quad 0 \leq \delta_p \leq \frac{1}{p^2}, \quad \alpha_p + \beta_p + \delta_p = 1 - \frac{1}{p^2}. \tag{7.4}$$

Here $\beta_p \geq 0$ because $\rho(p^2) \leq p\rho(p)$, $\delta_p \geq 0$ because $\rho(p) \leq p$, and $\alpha_p \leq 1$ because $\rho(p^2) \leq p^2$. Define

$$L_p(s) = \alpha_p + \beta_p p^{-s} + \delta_p p^{-2s}, \qquad M_p(s) = (1 - p^{-2s-2}) L_p(s), \tag{7.5}$$

and $P_f(s) = \prod_p L_p(s)$, $H_f(s) = \prod_p M_p(s)$. In the notation of the archived proposition, $L_p(s) = (1-2p^{-2})B_p(s)$ with $B_p(s) = a_p + (b_p - a_p)p^{-s} + (c_p - b_p)p^{-2s}$, $a_p = 1 + \rho(p^2)/(p^2(p^2-2))$, $b_p = 1 + \rho(p)/(p(p^2-2))$, $c_p = 1 + 1/(p^2-2)$. Expanding the product, with $z = p^{-s}$,

$$M_p(s) = \alpha_p + \beta_p z + (\delta_p - \alpha_p p^{-2}) z^2 - \beta_p p^{-2} z^3 - \delta_p p^{-2} z^4, \qquad \delta_p - \alpha_p p^{-2} = -\frac{\rho(p)}{p^3} + \frac{2}{p^4} - \frac{\rho(p^2)}{p^6}. \tag{7.6}$$

The cancellation of the terms $\pm p^{-2}$ in the coefficient of z^2 is what allows H_f to converge beyond $\operatorname{Re} s = -\frac{1}{2}$.

7.3 The factorisation

Proposition 7.2. *(a) The series $D_f(s) = \sum_h W_f(h) h^{-s}$ converges absolutely for $\operatorname{Re} s > 1$.*

(b) The products P_f and H_f converge absolutely and locally uniformly, and define holomorphic functions, on $\operatorname{Re} s > -\frac{1}{2}$ and $\operatorname{Re} s > -\frac{3}{4}$ respectively. For every fixed $\sigma > -\frac{3}{4}$ there is $C(f,\sigma)$ with $|H_f(\sigma + it)| \leq C(f,\sigma)$ for all real t.

(c) For $\operatorname{Re} s > 1$, $D_f(s) = \zeta(s)P_f(s) = \zeta(s)\zeta(2s+2)H_f(s)$. The function $\zeta(s)\zeta(2s+2)H_f(s)$ continues D_f meromorphically to $\operatorname{Re} s > -\frac{3}{4}$, and its only possible poles there are simple poles at $s = 1$ and $s = -\frac{1}{2}$.

(d) $R_f = \operatorname{Res}_{s=1} D_f(s) = P_f(1)$. Moreover $R_f = \widetilde{R}_f$, where $\widetilde{R}_f$ is the series of Lemma 6.1, and (1.11) holds. Moreover $B_f = (\gamma_E - \frac{3}{2})R_f + P_f'(1)$, and $R_f > 0$.

(e) $D_f(0) = -\frac{1}{2\zeta(2)} = -\frac{c}{2}$.

(f) We have $\operatorname{Res}_{s=-1/2} D_f(s) = \frac{1}{2}\zeta(-\frac{1}{2})H_f(-\frac{1}{2})$, and hence

$$K_f = -\tfrac{8}{3}\,\zeta(-\tfrac{1}{2})\,H_f(-\tfrac{1}{2}).$$

Moreover $H_f(-\frac{1}{2}) > 0$ and $K_f > 0$, and K_f is given by the absolutely convergent product (1.10).

Proof. (a) By Lemma 6.1, $0 < W_f(h) \le 1$, so $\sum_h |W_f(h)h^{-s}| \le \sum_h h^{-\operatorname{Re} s}$.

(b) For $\operatorname{Re} s \ge \sigma$, (7.4) gives $|L_p(s) - 1| \le 2p^{-2} + C_\rho p^{-3-\sigma} + p^{-2-2\sigma}$. If $\sigma > -\frac{1}{2}$ all exponents exceed one, so by Lemma 2.1(c) the sum over p converges uniformly on $\operatorname{Re} s \ge \sigma$, and Lemma 2.3 applies. For M_p, (7.6) and (7.4) give, for $\operatorname{Re} s \ge \sigma$,

$$|M_p(s) - 1| \le \frac{2}{p^2} + \frac{C_\rho}{p^{3+\sigma}} + \Big(\frac{C_\rho}{p^3} + \frac{2}{p^4} + \frac{C_\rho}{p^5}\Big)p^{-2\sigma} + \frac{C_\rho}{p^{5+3\sigma}} + \frac{1}{p^{4+4\sigma}}.$$

For $\sigma > -\frac{3}{4}$ the exponents 2, $3+\sigma$, $3+2\sigma$, $4+2\sigma$, $5+2\sigma$, $5+3\sigma$ and $4+4\sigma$ all exceed one. Hence $\sum_p \sup_{\operatorname{Re} s \ge \sigma} |M_p(s) - 1| =: E_H(\sigma) < \infty$, and Lemma 2.3 gives absolute and locally uniform convergence, holomorphy, and $|H_f(\sigma + it)| \le \exp(E_H(\sigma))$ for all t.

(c) By Lemma 6.1, $W_f(1) = \prod_p \alpha_p > 0$, and $u(h) = W_f(h)/W_f(1)$ is multiplicative with $u(p) = (\alpha_p + \beta_p)/\alpha_p$ and $u(p^e) = (\alpha_p + \beta_p + \delta_p)/\alpha_p$ for $e \ge 2$. For $\operatorname{Re} s > 1$ and $z = p^{-s}$ its local series is

$$1 + \frac{\alpha_p + \beta_p}{\alpha_p} z + \frac{\alpha_p + \beta_p + \delta_p}{\alpha_p} \cdot \frac{z^2}{1 - z} = \frac{\alpha_p + \beta_p z + \delta_p z^2}{\alpha_p(1 - z)} = \frac{L_p(s)}{\alpha_p(1 - p^{-s})}.$$

Since $\sum_h |u(h)h^{-s}| < \infty$, Lemma 2.4 gives $\sum_h u(h)h^{-s} = \prod_p L_p(s)\alpha_p^{-1}(1-p^{-s})^{-1}$. Multiplying by the convergent product $W_f(1) = \prod_p \alpha_p$ and using the Euler product of ζ gives $D_f(s) = \zeta(s)P_f(s)$; the limits of partial products multiply. For $\operatorname{Re} s > -\frac{1}{2}$ we have $\operatorname{Re}(2s+2) > 1$, so $\prod_p (1 - p^{-2s-2})^{-1} = \zeta(2s+2)$ converges absolutely, and multiplying the products gives $P_f(s) = \zeta(2s+2)H_f(s)$. By Lemma 7.1(b), $\zeta(s)$ and $\zeta(2s+2)$ are meromorphic for $\operatorname{Re} s > -\frac{3}{4}$, since then $\operatorname{Re} s > -1$ and $\operatorname{Re}(2s+2) > \frac{1}{2}$. Their only poles there are at $s = 1$ and at $2s+2 = 1$, both simple. Since H_f is holomorphic, $\zeta(s)\zeta(2s+2)H_f(s)$ is the required continuation.

(d) Since P_f is holomorphic near 1 and $\zeta(s) = \frac{1}{s-1} + \gamma_E + O(s-1)$ by Lemma 7.1(a), we have $D_f(s) = \frac{P_f(1)}{s-1} + \gamma_E P_f(1) + P_f'(1) + O(s-1)$. This gives $R_f = P_f(1)$ and the formula for B_f from (1.6). Substituting (7.3),

$$L_p(1) = \alpha_p + \frac{\beta_p}{p} + \frac{\delta_p}{p^2} = 1 - \frac{2}{p^2} + \frac{1}{p^4} + \frac{(1 - 1/p)(\rho(p) + \rho(p^2))}{p^4} = \Big(1 - \frac{1}{p^2}\Big)^2 (1 + T_p),$$

with $T_p = (1 - 1/p)(\rho(p) + \rho(p^2))/(p^2 - 1)^2 \ge 0$. By Lemma 2.7(d) and $p^2 - 1 \ge \frac{3}{4}p^2$, $T_p \le \frac{16C_\rho}{9}(p^{-3} + p^{-4})$, which is summable. Using (2.1), $R_f = \zeta(2)^{-2} \prod_p (1 + T_p) > 0$, which is (1.11). For the identification with $\widetilde{R}_f$, note that $\lambda(r)\rho(r)/r = c^2 v(r)$, where v is the nonnegative multiplicative function supported on cube-free integers with

$$v(p) = \frac{(1 - 1/p)\rho(p)}{(p^2 - 1)^2}, \qquad v(p^2) = \frac{\varphi(p^2)\rho(p^2)}{p^2(p^2 - 1)^2} = \frac{(1 - 1/p)\rho(p^2)}{(p^2 - 1)^2},$$

using Lemma 2.7(a) for multiplicativity of ρ. Thus $1 + v(p) + v(p^2) = 1 + T_p$. Since $\sum_r v(r) < \infty$ by Lemma 6.1, Lemma 2.4 gives $\widetilde{R}_f = c^2 \prod_p (1 + T_p) = R_f$.

(e) $L_p(0) = \alpha_p + \beta_p + \delta_p = 1 - p^{-2}$, so $P_f(0) = c$ by (2.1). The identity $D_f = \zeta P_f$ persists on $\operatorname{Re} s > -\frac{1}{2}$ by (c), and both factors are regular at 0; by Lemma 7.1(b), $D_f(0) = \zeta(0)c = -c/2$.

(f) Near $s = -\frac{1}{2}$, $\zeta(2s+2) = \frac{1/2}{s+1/2} + O(1)$, while $\zeta(s)$ and $H_f(s)$ are holomorphic. This gives the residue, and the formula for K_f follows from (1.6). At $s = -\frac{1}{2}$,

$$M_p(-\tfrac{1}{2}) = \Big(1 - \frac{1}{p}\Big)(\alpha_p + \beta_p\sqrt{p} + \delta_p p) = \Big(1 - \frac{1}{p}\Big)\Big(1 + \frac{1}{p} - \frac{2 + \rho(p)}{p^2} + \frac{\rho(p)}{p^{5/2}} - \frac{\rho(p^2)}{p^{7/2}} + \frac{\rho(p^2)}{p^4}\Big),$$

by direct substitution of (7.3). This is positive because $\alpha_p > 0$ and $\beta_p, \delta_p \ge 0$. The deviations $|M_p(-\frac{1}{2}) - 1|$ are summable by (b), so Lemma 2.3(c) gives $H_f(-\frac{1}{2}) > 0$, and Lemma 7.1(b) gives $\zeta(-\frac{1}{2}) < 0$. Hence $K_f > 0$, and (1.10) is the product $H_f(-\frac{1}{2}) = \prod_p M_p(-\frac{1}{2})$, with the factors at each prime kept together. □

Remark 7.3. Both poles in (c) are genuine, since $R_f > 0$ and $K_f > 0$. The value $D_f(0) = -c/2$ is essential: in Chapter 8 it produces the term $+cH$ in the expansion of $\Theta_f(H)$, which cancels the diagonal term $cx(Y - Z)$ of Proposition 4.3. The identity $R_f = \widetilde{R}_f$ makes the logarithmic terms cancel as well.

Chapter 8

The smoothed sum Θ_f and the profile F_f

The integral $I_f(x;Z,Y)$ of Proposition 6.4 turns out to be an exact combination of values of the smoothed partial sum

$$\Theta_f(H)=\sum_{1\le h\le H} W_f(h)\frac{(H-h)^2}{h}.$$

We evaluate $\Theta_f(H)$ by a Mellin integral of D_f and a contour shift past the poles at $s=1$, $s=0$ (from the kernel) and $s=-\tfrac12$. The residues give the four terms $R_fH^2\log H$, B_fH^2, cH and $-K_fH^{1/2}$. (Archived source: proposition P6.)

8.1 Growth in a vertical strip

We need one result from the literature. It is a consequence of the functional equation $\zeta(s)=\zeta(1-s)2^s\pi^{s-1}\Gamma(1-s)\sin(\pi s/2)$ [9, Corollary 10.4] and Stirling's formula.

Quoted result 8.1 ([9, Corollary 10.5])**.** *Suppose that $A>0$ is fixed. Then $|\zeta(s)|\asymp\tau^{1/2-\sigma}|\zeta(1-s)|$ uniformly for $|\sigma|\le A$ and $|t|\ge1$. Here $s=\sigma+it$ and $\tau=|t|+4$.*

We use it with $A=2$, only for $-\frac58\le\sigma\le\frac14$ and $|t|\ge2$. Let c_{MV} be a constant for the upper inequality there, so that $|\zeta(s)|\le c_{\mathrm{MV}}(|t|+4)^{1/2-\sigma}|\zeta(1-s)|$. No conditional estimate for ζ is used. (The holomorphy of ζ in the strip away from $s=1$ is already contained in Lemma 7.1(b).)

Lemma 8.2. *(a) Let $\delta\in\{\frac14,\frac34\}$ and $C_\delta=(\frac32)^{1-\delta}(\frac{1}{1-\delta}+1)+\frac{4}{\delta}$. Then $|\zeta(\alpha+iv)|\le C_\delta|v|^{1-\delta}$ for $\delta\le\alpha\le6$ and $|v|\ge2$.*

(b) For $s=\sigma+it$ with $-\frac58\le\sigma\le2$ and $T=|t|\ge2$,

$$|\zeta(s)|\le C_\zeta T^{11/8},\qquad |\zeta(2s+2)|\le C_{3/4}(2T)^{1/4},\qquad C_\zeta=\max\{C_{1/4},\,3^{9/8}c_{\mathrm{MV}}C_{3/4}\}.$$

Proof. (a) Let $V=|v|$ and use (7.1) with $N=\lceil V\rceil$, so $V\le N\le\frac32V$. The finite sum has modulus at most $\sum_{n\le N}n^{-\delta}\le N^{1-\delta}/(1-\delta)\le(\frac32)^{1-\delta}V^{1-\delta}/(1-\delta)$. Next, $|N^{1-s}/(s-1)|\le N^{1-\delta}/V\le(\frac32)^{1-\delta}V^{1-\delta}$. Finally $|s\int_N^\infty\{u\}u^{-s-1}\,du|\le|s|N^{-\alpha}/\alpha\le(V+6)V^{-\delta}/\delta\le4V^{1-\delta}/\delta$.

(b) For $\frac14\le\sigma\le2$ apply (a) with $\delta=\frac14$: $|\zeta(s)|\le C_{1/4}T^{3/4}\le C_{1/4}T^{11/8}$. For $-\frac58\le\sigma\le\frac14$, $\mathrm{Re}(1-s)\in[\frac34,\frac{13}{8}]$, so (a) with $\delta=\frac34$ and quoted result 8.1 give $|\zeta(s)|\le c_{\mathrm{MV}}(3T)^{1/2-\sigma}C_{3/4}T^{1/4}\le3^{9/8}c_{\mathrm{MV}}C_{3/4}T^{11/8}$, using $T+4\le3T$ and $\frac14\le\frac12-\sigma\le\frac98$. Finally, $\mathrm{Re}(2s+2)\in[\frac34,6]$ and $|\mathrm{Im}(2s+2)|=2T\ge4$, so (a) with $\delta=\frac34$ applies. □

The bound of Proposition 7.2(b) for H_f holds on each vertical line, but the contour shift needs a bound that is uniform on a strip. The archived proof supplies it directly from the Euler product.

Lemma 8.3. *Let $\mathcal{B}$ be the finite set of primes dividing every coefficient of f. For $-\frac{5}{8} \le \operatorname{Re} s \le 2$,*

$$|H_f(s)| \le M_H := e^{8(1+d)} \prod_{p \in \mathcal{B}} (1 + p^{-3/4})(1 + p^{5/8} + p^{5/4}).$$

Consequently, with $C_D = 2^{1/4} M_H C_\zeta C_{3/4}$,

$$|D_f(\sigma + it)| \le C_D |t|^{13/8} \qquad (-\tfrac{5}{8} \le \sigma \le 2,\ |t| \ge 2). \tag{8.1}$$

Proof. Let $z = p^{-s}$, so $|z| \le p^{5/8}$. For $p \notin \mathcal{B}$, Lemma 2.7(c) gives $\rho(p) \le d$ and $\rho(p^2) \le dp$. Together with (7.4) this yields $|\alpha_p - 1| \le 2p^{-2}$, $\beta_p \le dp^{-3}$, $\delta_p \le p^{-2}$ and $|\delta_p - \alpha_p p^{-2}| \le dp^{-3} + 2p^{-4} + dp^{-5} \le (2d+1)p^{-3}$. By (7.6),

$$|M_p(s) - 1| \le 2p^{-2} + dp^{-19/8} + (2d+1)p^{-7/4} + dp^{-25/8} + p^{-3/2} \le (4 + 4d)p^{-3/2},$$

and $\sum_p (4+4d)p^{-3/2} \le (4+4d)\int_1^\infty u^{-3/2}\,\mathrm{d}u = 8(1+d)$. Hence $|\prod_{p \notin \mathcal{B}} M_p(s)| \le \prod (1 + |M_p(s) - 1|) \le e^{8(1+d)}$. For the finitely many $p \in \mathcal{B}$ use $0 \le \alpha_p, \beta_p, \delta_p \le 1$ to get $|M_p(s)| \le (1 + p^{-2}|z|^2)(1 + |z| + |z|^2) \le (1 + p^{-3/4})(1 + p^{5/8} + p^{5/4})$. The bound (8.1) then follows from $D_f(s) = \zeta(s)\zeta(2s+2)H_f(s)$ and Lemma 8.2(b). □

8.2 The Mellin representation

Lemma 8.4. *For real $v > 0$,*

$$\frac{1}{2\pi i} \int_{1-i\infty}^{1+i\infty} \frac{2v^z}{z(z+1)(z+2)}\,\mathrm{d}z = \begin{cases} (1 - v^{-1})^2, & v \ge 1, \\ 0, & 0 < v < 1. \end{cases}$$

Proof. The integral converges absolutely. For $v \ge 1$, close the contour by a rectangle to the left, with horizontal sides at heights $\pm T$ and a left side on $\operatorname{Re} z = -N - \frac{1}{2}$ with an integer $N \ge 3$. On the horizontal sides the integrand is $O_v(T^{-3})$, so their contribution tends to zero as $T \to \infty$. On the left side each of $|z|, |z+1|, |z+2|$ is at least $\sqrt{(N - 3/2)^2 + (\operatorname{Im} z)^2}$, so the normalised integral is at most $\frac{v^{-N-1/2}}{\pi} \int_{\mathbb{R}} ((N - \frac{3}{2})^2 + t^2)^{-3/2}\,\mathrm{d}t = \frac{2v^{-N-1/2}}{\pi(N-3/2)^2}$, which tends to zero also when $v = 1$. The residues at $z = 0, -1, -2$ are 1, $-2v^{-1}$ and v^{-2}, with sum $(1 - v^{-1})^2$. For $v < 1$, close to the right instead; the right side $\operatorname{Re} z = N$ contributes at most $2v^N/(\pi N^2) \to 0$, and there are no poles. □

Proposition 8.5. *There is $C(f) > 0$ such that, uniformly for real $H \ge 1$,*

$$\left|\Theta_f(H) - \left(R_f H^2 \log H + B_f H^2 + cH - K_f H^{1/2}\right)\right| \le C(f) H^{3/8}, \tag{8.2}$$

and consequently

$$|F_f(H) - K_f H^{-3/2}| \le C(f) H^{-13/8}, \qquad |F_f(H)| \le (K_f + C(f)) H^{-3/2}. \tag{8.3}$$

Proof. Mellin representation. Apply Lemma 8.4 with $v = H/h$, multiply by $H^2 W_f(h)/h$ and sum over $h \ge 1$; note $H^2(1 - h/H)^2 = (H - h)^2$, and a term with $h = H$ vanishes on both sides. With $s = z + 1$ the integrand becomes $2W_f(h) H^{s+1} h^{-s}/((s-1)s(s+1))$ on $\operatorname{Re} s = 2$. The sum over h of its absolute values is at most $2H^3\zeta(2)/|(s-1)s(s+1)|$, which is integrable

on the line, because $|(s-1)s(s+1)| \geq |\operatorname{Im} s|^3$ there and the denominator does not vanish. Hence summation and integration may be interchanged, and

$$\Theta_f(H) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} G_H(s)\,\mathrm{d}s, \qquad G_H(s) = \frac{2D_f(s)H^{s+1}}{(s-1)s(s+1)}.$$

Contour shift. In the strip $-\frac{5}{8} \leq \operatorname{Re} s \leq 2$ the poles of G_H are at $s=1$ and $s=-\frac{1}{2}$ (from D_f, Proposition 7.2(c)) and at $s=0$ (from the kernel); the pole of the kernel at -1 lies outside. On the horizontal segments at height $\pm T$, $T \geq 2$, each factor of the denominator has modulus at least T, and $|H^{s+1}| \leq H^3$. So (8.1) bounds the integrand by $2C_D H^3 T^{-11/8}$. The segments have length $\frac{21}{8}$, so their total contribution is at most $\frac{21}{2}C_D H^3 T^{-11/8} \to 0$ as $T \to \infty$, for each fixed H. On the line $\operatorname{Re} s = -\frac{5}{8}$ we have $|(s-1)s(s+1)| \geq \frac{13}{8}\cdot\frac{5}{8}\cdot\frac{3}{8} = \frac{195}{512}$. With $M_0 = \max_{|t|\leq 2}|D_f(-\frac{5}{8}+it)|$, finite because the line contains no pole, the part $|t| \leq 2$ contributes at most $4\cdot\frac{512}{195}\cdot 2M_0 H^{3/8}$ to the integral of $|G_H|$. For $|t| \geq 2$, (8.1) gives $|G_H| \leq 2C_D H^{3/8}|t|^{-11/8}$, and these tails contribute at most $4C_D H^{3/8}\int_2^\infty t^{-11/8}\,\mathrm{d}t = \frac{32}{3}C_D 2^{-3/8}H^{3/8}$. Hence the residue theorem gives

$$\Theta_f(H) = \sum_{s_0 \in \{1,0,-1/2\}} \operatorname*{Res}_{s=s_0} G_H(s) + E_f(H), \qquad |E_f(H)| \leq C_* H^{3/8},$$

where $C_* = \frac{1}{2\pi}(\frac{4096}{195}M_0 + \frac{32}{3}C_D 2^{-3/8})$ depends only on f.

Residues. Write $D_f(s) = \frac{R_f}{s-1} + L_f + O(s-1)$ near $s=1$, so that $B_f = L_f - \frac{3}{2}R_f$ by (1.6), and put $q(s) = 2/(s(s+1))$. Then $q(1)=1$ and $q'(1) = -\frac{3}{2}$, and $G_H(s) = q(s)H^{s+1}D_f(s)/(s-1)$ has the residue

$$\begin{aligned}\operatorname*{Res}_{s=1} G_H &= R_f \frac{\mathrm{d}}{\mathrm{d}s}\big(q(s)H^{s+1}\big)\Big|_{s=1} + L_f q(1) H^2 \\ &= H^2(R_f \log H - \tfrac{3}{2}R_f + L_f) = R_f H^2 \log H + B_f H^2.\end{aligned}$$

At $s=0$, Proposition 7.2(e) gives

$$\operatorname*{Res}_{s=0} G_H = \frac{2D_f(0)H}{(-1)\cdot 1} = -2D_f(0)H = cH.$$

At $s=-\frac{1}{2}$ the denominator equals $(-\frac{3}{2})(-\frac{1}{2})(\frac{1}{2}) = \frac{3}{8}$, so by (1.6)

$$\operatorname*{Res}_{s=-1/2} G_H = \frac{16}{3}H^{1/2}\operatorname*{Res}_{s=-1/2} D_f = -K_f H^{1/2}.$$

This proves (8.2) with $C(f) = 1 + C_*$, for every real $H \geq 1$, integral or not. Substituting into (1.7) gives the exact identity $F_f(H) = K_f H^{-3/2} - E_f(H)/H^2$, which implies (8.3) since $H^{-13/8} \leq H^{-3/2}$. □

8.3 The integral identity

Lemma 8.6. *For real $x \geq 2$ and $2 \leq Z \leq Y \leq x$,*

$$2\int_Z^Y \sum_{1\leq h\leq x/t} (x-ht)W_f(h)\,\mathrm{d}t = Z^2\Theta_f(x/Z) - Y^2\Theta_f(x/Y).$$

Proof. Write $u_+ = \max\{u,0\}$. For $t \in [Z,Y]$ only $h \leq x/Z$ contribute, so the left side equals $\sum_{h\leq x/Z} W_f(h)\cdot 2\int_Z^Y (x-ht)_+\,\mathrm{d}t$. The function $t \mapsto (x-ht)_+^2$ is continuously differentiable with derivative $-2h(x-ht)_+$, hence $2\int_Z^Y (x-ht)_+\,\mathrm{d}t = \big((x-hZ)_+^2 - (x-hY)_+^2\big)/h$. On the other hand, by the definition of Θ_f, $Z^2\Theta_f(x/Z) = \sum_{h\leq x/Z} W_f(h)(x-hZ)_+^2/h$ and $Y^2\Theta_f(x/Y) = \sum_{h\leq x/Z} W_f(h)(x-hY)_+^2/h$; in the second sum the terms with $h > x/Y$ vanish. Subtracting gives the identity, with no error term. Note that $x/Z, x/Y \geq 1$. □

Chapter 9

Proof of the main theorem

We now combine Propositions 4.3, 4.5, 5.1, 6.4, 7.2 and 8.5 and Lemma 8.6. Throughout, $d = \deg f$,

$$a = 4d - 1, \qquad \theta = \frac{2d}{a}, \qquad \eta = \eta_d = \min\Big\{\frac{1}{5}, \frac{1}{a}\Big\},$$

and ε with $0 < \varepsilon < \eta/2$ is fixed. We put $\varepsilon_1 = \varepsilon/2$ and apply Propositions 4.3 and 4.5 with ε_1 in place of ε. Constants written C_f depend only on f, constants written $C_{f,\varepsilon}$ depend only on f and ε, and both may change from line to line. (Archived source: proposition P7.)

9.1 The variance on an interval of moduli

Lemma 9.1. *Let $x \ge 2$, $1 \le z \le y$, $2 \le Z = f(z) \le Y = f(y) \le x$ and $2 \le R \le x$, and let $\mathcal{A} = \mathcal{A}(x, Z, R)$ be as in Proposition 6.4. Then*

$$\Big|V_f(x; z, y) - x^2\big(F_f(x/Y) - F_f(x/Z)\big)\Big| \le C_{f,\varepsilon}(x^{3/2+\varepsilon_1} + xY^{1-1/d} + x^{9/5}\log(2x) + x^2\mathcal{A}).$$

Proof. Correlations. For $z < k \le y$ there are at most $x/f(k)$ integers h with $1 \le h \le x/f(k)$, and each $m = hf(k)$ is an integer in $[1, x]$. By Proposition 5.1, replacing $C_x(m)$ by $(x-m)\mathfrak{S}(m)$ costs at most $K_{\mathrm{c}}x^{4/5}$ for each pair (h, k). After multiplying by $f'(k)$ and summing over h, the cost is at most $K_{\mathrm{c}}x^{9/5}f'(k)/f(k)$, and by (2.4)

$$|S_f(x; z, y) - M_f(x; z, y)| \le 2UK_{\mathrm{c}}\, x^{9/5}\log(2x). \tag{9.1}$$

Averaging. Proposition 6.4 applies under exactly the present hypotheses and gives

$$|M_f(x; z, y) - I_f(x; Z, Y)| \le C_f x^2\mathcal{A}, \qquad |T_f(z, y) - R_f\log(Y/Z)| \le C_f\mathcal{A},$$

where we used $\widetilde{R}_f = R_f$ (Proposition 7.2(d)).

Combination. By Proposition 4.3, the factor 2 in front of S_f and the factor x^2 in front of T_f give

$$\Big|V_f(x; z, y) - \big[cx(Y - Z) + 2I_f(x; Z, Y) - x^2R_f\log(Y/Z)\big]\Big| \le C_{f,\varepsilon}(x^{3/2+\varepsilon_1} + xY^{1-1/d}) + 4UK_{\mathrm{c}}x^{9/5}\log(2x) + 3C_f x^2\mathcal{A}.$$

Exact cancellation. By Lemma 8.6, $2I_f(x; Z, Y) = Z^2\Theta_f(x/Z) - Y^2\Theta_f(x/Y)$. By the definition (1.7) of F_f, with $c = \zeta(2)^{-1}$, direct subtraction gives

$$x^2\big(F_f(x/Y) - F_f(x/Z)\big) = -x^2R_f\log(Y/Z) + cx(Y - Z) + Z^2\Theta_f(x/Z) - Y^2\Theta_f(x/Y),$$

because $x^2 \cdot c/(x/Y) = cxY$ and the B_f-terms cancel. Hence the bracket equals $x^2(F_f(x/Y) - F_f(x/Z))$. □

9.2 Choice of the cutoffs and the error budget

Proof of (1.8). Take

$$Z = x^{\theta} = x^{2d/a}, \qquad R = Z^{1/d} = x^{2/a}.$$

Both exponents lie in $(0,1)$. Choose $x_0 = x_0(f) \geq 2$ so that $Z \geq \max\{2, f(1)\}$ and $R \geq 2$ for $x \geq x_0$; then also $Z \leq x$ and $R \leq x$. By Lemma 2.6(a) there is a unique real $z \geq 1$ with $f(z) = Z$. Then

$$\begin{gathered} Z^{-1/d} R^{1/2} = R^{-1/2} = x^{-1/a}, \qquad 1 + \theta\Big(2 - \frac{1}{d}\Big) = \frac{8d-3}{a} = 2 - \frac{1}{a}, \\ \Big(2 - \frac{1}{a}\Big) - \Big(\frac{1}{2} + \frac{3d}{a}\Big) = \frac{6d-5}{2a} > 0. \end{gathered} \tag{9.2}$$

Case $Y \geq Z$. Then $z \leq y$, and $V_f(x,y) = V_f(x,z) + V_f(x;z,y)$. Hence

$$|V_f(x,y) - x^2 F_f(x/Y)| \leq V_f(x,z) + x^2 |F_f(x/Z)| + \Big| V_f(x;z,y) - x^2 \big(F_f(x/Y) - F_f(x/Z)\big) \Big|.$$

We bound every term; the table summarises the result.

- By Proposition 4.5 and (9.2), $V_f(x,z) \leq C_{f,\varepsilon} x^{1+\varepsilon_1} Z^{2-1/d} = C_{f,\varepsilon} x^{2-1/a+\varepsilon_1}$.
- By (8.3) with $H = x/Z \geq 1$, $x^2 |F_f(x/Z)| \leq C_f x^{1/2} Z^{3/2} = C_f x^{1/2+3d/a} \leq C_f x^{2-1/a}$, by (9.2).
- By Lemma 9.1 the remaining difference is at most $C_{f,\varepsilon}$ times $x^{3/2+\varepsilon_1} + xY^{1-1/d} + x^{9/5}\log(2x) + x^2\mathcal{A}$, and by (9.2) and $R \leq x$, $x^2\mathcal{A} \leq x^{2-1/a}((\log 2x)^2 + (\log 2x)^3)$.

origin	bound	comparison with $x^{2-\eta+\varepsilon}$
diagonal and mixed terms (Proposition 4.3)	$x^{3/2+\varepsilon_1}$	$\eta \leq \frac{1}{5} < \frac{1}{2}$
endpoint of $\sum f'(k)$ (Proposition 4.3)	$xY^{1-1/d} \leq x^{2-1/d}$	$\eta \leq \frac{1}{a} \leq \frac{1}{d}$
correlations (Proposition 5.1)	$x^{9/5}\log(2x)$	$\eta \leq \frac{1}{5}$, $\log 2x \ll_\varepsilon x^\varepsilon$
averaging (Proposition 6.4)	$x^{2-1/a}(\log 2x)^3$	$\eta \leq \frac{1}{a}$, $(\log 2x)^3 \ll_\varepsilon x^\varepsilon$
initial moduli (Proposition 4.5)	$x^{2-1/a+\varepsilon_1}$	$\eta \leq \frac{1}{a}$
lower boundary of the profile	$x^{1/2+3d/a} \leq x^{2-1/a}$	$\eta \leq \frac{1}{a}$

The logarithmic bounds come from Lemma 2.1(d). Every row is at most $C_{f,\varepsilon} x^{2-\eta+\varepsilon}$ for $x \geq 2$, which proves (1.8) when $Y \geq Z$.

Case $Y < Z$. Apply Proposition 4.5 with y in place of z; this is allowed since $y \geq 1$ and $f(y) = Y \leq x$. Because $2 - 1/d > 0$,

$$0 \leq V_f(x,y) \leq C_{f,\varepsilon} x^{1+\varepsilon_1} Y^{2-1/d} \leq C_{f,\varepsilon} x^{1+\varepsilon_1} Z^{2-1/d} = C_{f,\varepsilon} x^{2-1/a+\varepsilon_1}.$$

Also $x/Y \geq 1$, so (8.3) and (9.2) give $x^2|F_f(x/Y)| \leq C_f x^{1/2} Y^{3/2} \leq C_f x^{1/2+3d/a} \leq C_f x^{2-1/a}$. The triangle inequality and $\eta \leq 1/a$ give (1.8) in this case too.

All constants and x_0 depend only on f and ε, and real values of y and z need no separate treatment: the sums split exactly as $k \leq z$ and $z < k \leq y$, and the endpoint effects are contained in Proposition 4.3. □

Proof of (1.9). For every $Y \leq x$, (8.3) with $H = x/Y \geq 1$, multiplied by x^2, gives

$$\big|x^2 F_f(x/Y) - K_f x^{1/2} Y^{3/2}\big| \leq C_f x^2 (x/Y)^{-13/8} = C_f x^{3/8} Y^{13/8},$$

since $x^2(x/Y)^{-3/2} = x^{1/2}Y^{3/2}$. Combine this with (1.8). The formulas (1.10) and (1.11) and the positivity of K_f are Proposition 7.2(d),(f). □

Remark 9.2. The hypothesis $\varepsilon < \eta_d/2$ is not used in the argument above, which works for every $\varepsilon > 0$. We keep the hypothesis as stated.

9.3 Proof of Corollary 1.2

Let $0 < \delta < \eta/3$, apply (1.9) with $\varepsilon = \delta$ (admissible because $\delta < \eta/2$), and divide by $K_f x^{1/2} Y^{3/2}$, which is positive by Proposition 7.2(f). For $x^{1-2\eta/3+\delta} \le Y \le x^{1-\delta}$ the two relative errors are

$$\frac{x^{3/8} Y^{13/8}}{x^{1/2} Y^{3/2}} = \Big(\frac{Y}{x}\Big)^{1/8} \le x^{-\delta/8}, \qquad \frac{x^{2-\eta+\delta}}{x^{1/2} Y^{3/2}} \le x^{3/2-\eta+\delta-\frac{3}{2}(1-\frac{2\eta}{3}+\delta)} = x^{-\delta/2} \le x^{-\delta/8}.$$

Hence $|V_f(x,y)/(K_f x^{1/2} Y^{3/2}) - 1| \le C(f,\delta) x^{-\delta/8}$ for $x \ge x_0(f,\delta)$. The range is nonempty because $\delta < \eta/3$. □

9.4 Where the exponents come from

The cap $\eta_d \le \frac{1}{5}$ comes from summing the uniform correlation error $x^{4/5}$ of Proposition 5.1 over all pairs (h,k); the weighted number of these pairs is $\sum_k f'(k) x / f(k) \asymp x \log x$. The exponent $\frac{1}{4d-1}$ comes from balancing two terms. The first is the bound $x^{1+\varepsilon} Z^{2-1/d}$ for the initial moduli; it arises from multiplying the full-variance bound (4.6) by the largest weight $f'(k) \asymp Z^{1-1/d}$. The second is the averaging error $x^2 Z^{-1/(2d)}$ of Proposition 6.4 at $R = Z^{1/d}$. Equating $1 + \theta(2 - \frac{1}{d})$ with $2 - \frac{\theta}{2d}$ gives $\theta = \frac{2d}{4d-1}$. Improving either the treatment of small moduli or the correlation estimate would widen the range of Corollary 1.2. Neither improvement is attempted here.

Chapter 10

Numerical evidence

This chapter compares Theorem 1.1 with exact computations of $V_f(x, y)$. The computations come from the numerical snapshot accepted for the archived run (`source/numerics/snapshot`). Every number in the tables and figures below is recomputed by `reproduce.py` from that snapshot's own source code and from the formulas of Chapters 7 and 8; see the `README.md` accompanying this document. None of the numbers is part of the proof. The fits and their uncertainty envelopes are heuristic.

10.1 What is computed

The snapshot treats $f(t) = t^2 + 1$, $t^3 + 2$ and $2t + 1$, which satisfy (1.3), for

$$x \in \{2.5, 5\} \cdot 10^5 \cup \{1, 2, 4, 8, 16, 32, 64, 128, 256\} \cdot 10^6$$

and the target ratios $r \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1\}$; for each pair, y is the largest integer with $f(y) \le rx$. At the two largest values of x the extra ratio $r = 0.005$ is also computed, as a diagnostic. The normalised value

$$N_{\text{data}} = \frac{V_f(x, y)}{x^{1/2} f(y)^{3/2}}$$

is compared with the prediction

$$N_{\text{pred}} = F_f(H) H^{3/2} = \frac{x^2 F_f(x/Y)}{x^{1/2} Y^{3/2}}, \qquad H = \frac{x}{f(y)}.$$

By (8.3), $N_{\text{pred}} \to K_f$ as $H \to \infty$.

Algorithm of the snapshot. The computation follows the expansion of Chapter 4 exactly, with no approximation in the counts.

- The squarefree indicator up to x is sieved, and all correlations $C_x(m)$, $0 \le m \le x$, are obtained at once from one number-theoretic-transform autocorrelation modulo the prime 3221225473. Every correlation is at most x, which is below the prime, so no information is lost. Then $\sum_l Q(x; q, l)^2 = Q(x; 1, 1) + 2 \sum_{m \ge 1} C_x(mq)$ by Lemma 4.1.
- With $g(q, l) = c\, h(q, l)$, where h is the rational number given by (3.3), the coefficients $B_q = \sum_l Q(x; q, l) h(q, l)$ and $D_q = \sum_l h(q, l)^2$ are computed as exact rationals. B_q is computed from the counts $S_a = Q(x; a, 0)$ exactly as in the proof of Lemma 4.2. D_q is computed from the local factors of Proposition 3.2(a): $p(p^2 + p - 1)/((p - 1)(p + 1)^2)$ for $p \parallel q$ and $p^2/(p^2 - 1)$ for $p^2 \mid q$, divided by q.

- The weighted sums are combined as $V = \sum_k f'(k)(\sum_l Q^2 - 2cxB_{f(k)} + c^2x^2D_{f(k)})$ in 70-digit decimal arithmetic, with $c = 6/\pi^2$ evaluated to 85 digits. The integer parts are exact; V itself is irrational, so it is represented by a 70-digit approximation.

Validation in the snapshot code. The pipeline checks that the transform prime and its primitive root are valid. It checks the rational density formula against an independent enumeration of local residue conditions for all 1275 pairs (q, l) with $q \le 50$, and against brute-force frequencies at $x = 10^6$. It checks $\sum_l Q(x; q, l) = Q(x; 1, 1)$, and it compares, for every modulus at $x = 257$, the exact $\sum_l Q^2$ and both rational coefficients with direct residue counts. For every production value of x it also checks $C_x(0)$ and seven lags of the autocorrelation directly. All checks passed in the recomputation.

Reproduction. `reproduce.py` copies the snapshot's `compute.py` and `counts.cpp` unchanged, after checking their SHA-256 hashes against the acceptance record, compiles the C++ worker locally, and runs the snapshot pipeline on the full grid. In the run used to prepare this document, all 270 recomputed rows agreed exactly with the archived results: the integer counts were identical, and so were the 70-digit values of V. This comparison is recorded in `reproduction.json`.

10.2 The constants

Table 10.1 lists R_f, B_f, $H_f(-\frac{1}{2})$ and K_f, computed from the Euler products of Proposition 7.2 over the primes $p \le 10^7$. The root counts $\rho(p)$ and $\rho(p^2)$ are computed by direct enumeration for small p (for $\rho(p^2)$, $p \le 300$; for $\rho(p)$, $p \le 1000$). For larger p they follow from the classical rules: for t^2+1, $\rho(p) = 2$ if $p \equiv 1 \bmod 4$ and 0 otherwise; for t^3+2, $\rho(p) = 1$ if $p \equiv 2 \bmod 3$, and if $p \equiv 1 \bmod 3$, $\rho(p) = 3$ or 0 according as $(-2)^{(p-1)/3} \equiv 1 \bmod p$ or not; for $2t+1$ and $t+1$, $\rho(p) = 1$. For these p the roots are simple, so $\rho(p^2) = \rho(p)$ by Hensel's lemma. The rules are checked against the direct counts wherever both are available.

f	R_f	B_f	$H_f(-\frac{1}{2})$	K_f
t^2+1	0.392347679	−0.436842746	0.249083290	0.138082626
t^3+2	0.395332424	−0.441819624	0.244618515	0.135607519
$2t+1$	0.378696892	−0.424235662	0.257256057	0.142613308
$t+1$	0.420774325	−0.450308730	0.231758984	0.128478667

Table 10.1: Constants of Theorem 1.1, Euler products over $p \le 10^7$. The row $f(t) = t+1$ is the comparison of Remark 1.4.

Truncation. R_f and B_f are computed from products and sums whose factors differ from 1 by $O(p^{-3})$ or less, so their truncation errors are negligible at this precision. For K_f the factor $M_p(-\frac{1}{2})$ differs from 1 by about $(3+\rho(p))p^{-2}$. We therefore evaluate $\zeta(2)^{-3}\prod_p M_p(-\frac{1}{2})(1-p^{-2})^{-3}$, which is the same number. If $p \ge 5$ and $\rho(p), \rho(p^2) \le 3$, as is the case for $p > 3$ for all four polynomials, then multiplying out the formula in Proposition 7.2(f) gives

$$M_p(-\tfrac{1}{2}) - 1 = -\frac{3+\rho(p)}{p^2} + \frac{\rho(p)}{p^{5/2}} + \frac{2+\rho(p)}{p^3} - \frac{\rho(p)+\rho(p^2)}{p^{7/2}} + \rho(p^2)\Big(\frac{1}{p^4} + \frac{1}{p^{9/2}} - \frac{1}{p^5}\Big),$$

and hence $|M_p(-\frac{1}{2}) - 1| \le p^{-2}(6 + 3\cdot 5^{-1/2} + 1 + 6\cdot 5^{-3/2} + 0.2) \le 10p^{-2}$. Since $(1-p^{-2})^{-3} \le 1.131$ and $(1-p^{-2})^{-3} - 1 \le 3.6p^{-2}$, it follows that $|M_p(-\frac{1}{2})(1-p^{-2})^{-3} - 1| \le 15p^{-2}$. By

Lemma 2.3(b) and Lemma 2.1(c), the omitted primes change K_f by a relative amount of at most $e^{15/N} - 1 = 1.5 \cdot 10^{-6}$ for $N = 10^7$. This bound covers the truncation only; floating-point rounding is not certified by interval arithmetic. Table 10.2 shows the values at three cutoffs. They are consistent with an actual truncation error far below this bound.

prime cutoff	K_{t^2+1}	K_{t^3+2}	K_{2t+1}
10^5	0.1380827359	0.1356076276	0.1426134214
10^6	0.1380826350	0.1356075276	0.1426133169
10^7	0.1380826264	0.1356075192	0.1426133080

Table 10.2: K_f with Euler products truncated at different prime cutoffs.

Vaughan's constant. Evaluating (1.2) with the same cutoff gives $c_2 \approx 0.128479$, in agreement with $K_{t+1} = 0.128479$, as Remark 1.4 predicts.

Checks of Propositions 7.2 and 8.5. `reproduce.py` compares $\sum_{h \le 2 \cdot 10^5} W_f(h) h^{-3}$ with $\zeta(3) P_f(3)$ and finds agreement to within the truncation error. Table 10.3 shows the normalised remainder $(\Theta_f(H) - R_f H^2 \log H - B_f H^2 - cH)/H^{1/2}$, which by (8.2) equals $-K_f + O(H^{-1/8})$. The values are close to $-K_f$, consistent with this error term. Equation (8.2) does not imply a monotone approach, and for t^2+1 and $2t+1$ the distance to $-K_f$ grows slightly from $H = 10^4$ to $H = 10^5$.

H	$f = t^2 + 1$	$f = t^3 + 2$	$f = 2t + 1$
10^2	-0.140743	-0.138565	-0.144597
10^3	-0.138771	-0.136180	-0.142970
10^4	-0.138166	-0.135697	-0.142688
10^5	-0.138173	-0.135675	-0.142694
$-K_f$	-0.138083	-0.135608	-0.142613

Table 10.3: $(\Theta_f(H) - R_f H^2 \log H - B_f H^2 - cH)/H^{1/2}$ compared with $-K_f$.

10.3 The data against the profile

Table 10.4 compares the data at the largest $x = 256{,}000{,}000$ with the profile, and Fig. 10.1 plots both together with the limit K_f. Figure 10.2 shows how the relative deviation evolves with x; the full data are in Appendix A.

The observations are as follows. The bounds quoted in them are rounded outwards, so they may differ from Table 10.4 in the last digit.

- For $f(t) = 2t + 1$, where y is large (y between 639,999 and 127,999,999 at the largest x), data and profile agree to within 0.00054% in absolute value.
- For $f(t) = t^2 + 1$ (y between 1,131 and 15,999), the deviation lies between +0.0069% and +0.1018%. At every target ratio its absolute value has decreased by a factor of at least 18 from the smallest to the largest x. (The diagnostic ratio 0.005 is computed only at the two largest x.)

f	r	y	$f(y)/x$	N_{data}	N_{pred}	$N_{\mathrm{data}}/N_{\mathrm{pred}} - 1$ (%)
t^2+1	0.005	1,131	0.004997	0.140391	0.140249	+0.1017
	0.01	1,599	0.009988	0.140858	0.140773	+0.0601
	0.02	2,262	0.019987	0.142919	0.142819	+0.0701
	0.05	3,577	0.049980	0.144457	0.144406	+0.0352
	0.1	5,059	0.099975	0.150976	0.150943	+0.0214
	0.2	7,155	0.199977	0.156550	0.156515	+0.0222
	0.4	10,119	0.399977	0.162775	0.162752	+0.0141
	0.7	13,386	0.699941	0.169925	0.169900	+0.0146
	1	15,999	0.999875	0.171101	0.171089	+0.0069
t^3+2	0.005	108	0.004921	0.140761	0.137720	+2.2083
	0.01	136	0.009826	0.139990	0.138751	+0.8931
	0.02	172	0.019877	0.141202	0.141231	−0.0203
	0.05	233	0.049411	0.143912	0.142767	+0.8019
	0.1	294	0.099266	0.148477	0.147575	+0.6117
	0.2	371	0.199472	0.152339	0.151303	+0.6844
	0.4	467	0.397842	0.155499	0.154559	+0.6078
	0.7	563	0.697084	0.163990	0.163149	+0.5159
	1	634	0.995469	0.166706	0.166270	+0.2617
$2t+1$	0.005	639,999	0.005000	0.144845	0.144844	+0.0003
	0.01	1,279,999	0.010000	0.144598	0.144597	+0.0005
	0.02	2,559,999	0.020000	0.148368	0.148368	+0.0005
	0.05	6,399,999	0.050000	0.149343	0.149342	+0.0004
	0.1	12,799,999	0.100000	0.156703	0.156702	+0.0003
	0.2	25,599,999	0.200000	0.161710	0.161709	+0.0002
	0.4	51,199,999	0.400000	0.162451	0.162451	+0.0001
	0.7	89,599,999	0.700000	0.182390	0.182390	+0.0000
	1	127,999,999	1.000000	0.183691	0.183691	+0.0000

Table 10.4: Normalised variance at $x = 256{,}000{,}000$ against the predicted profile $N_{\mathrm{pred}} = F_f(H)H^{3/2}$, $H = x/f(y)$; r is the target ratio.

- For $f(t) = t^3 + 2$, y is only 108–634, and the deviation ranges from -0.0203% to $+2.2083\%$, the largest value occurring at the diagnostic ratio 0.005. Over the range of x it has decreased substantially, from up to 14.9% at the smallest x. Between the two largest x its absolute value decreased at 8 of the 9 ratios, the exception being $r = 0.005$; at $r = 0.02$ the deviation changed sign. This behaviour is compatible with large lower-order terms from the sampling of k, which the proof controls only through the $Z^{-1/d}$ terms of Proposition 6.4. That explanation is plausible but has not been proved.

The theorem's error bound is far from small in this range. Relative to the main term it is about $x^{-\eta_d} r^{-3/2}$, and at $x = 2.56 \cdot 10^8$ already $x^{-1/11} \approx 0.17$ for $d = 3$ and $x^{-1/7} \approx 0.06$ for $d = 2$. The data are therefore consistent with the theorem, and they agree with the predicted profile far better than the proved error term requires. However, they cannot test the error exponent.

10.4 Extrapolation to small ratios

The snapshot also estimates the small-ratio constant empirically. Among the eight target ratios it selects the smallest one, 0.01 for all three polynomials, such that all rows at this or a larger ratio changed by less than 1% between the two largest x. It then fits $N(r) = C + a\sqrt{r} + br$ by unweighted least squares to the largest-x rows with $r \geq 0.01$, using the actual ratios. The

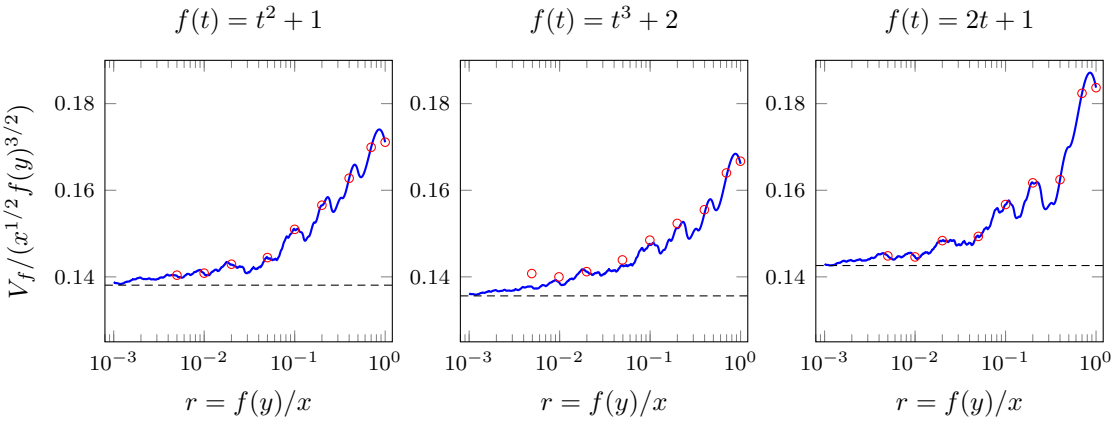


Figure 10.1: Predicted profile $F_f(1/r)\, r^{-3/2}$ (solid), data at $x = 256{,}000{,}000$ (circles) and the constant K_f (dashed), against $r = f(y)/x$ on a logarithmic axis.

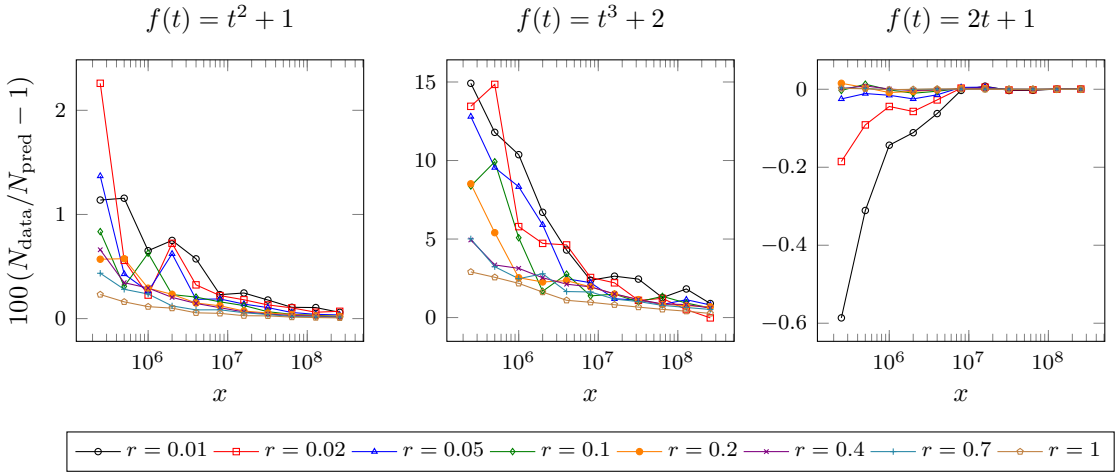


Figure 10.2: Relative deviation $100(N_{\text{data}}/N_{\text{pred}} - 1)$ against x, one line per target ratio r.

reported uncertainty is the maximum of twice the formal standard error and of the shifts under several variations: of the cutoff, of x, of the model, and of the fitted ratio window. It is a heuristic sensitivity envelope, not a confidence interval.

Table 10.5 shows that K_f lies within the heuristic envelope in all three cases ($K_f/C_{\text{data}} - 1 = +2.68\%$, -0.52%, $+0.65\%$). To see where the offset from K_f comes from, the same fit was applied to the *exact* predicted profile N_{pred} at the same actual ratios. Its intercept C_{prof} splits the offset exactly as

$$C_{\text{data}} - K_f = (C_{\text{data}} - C_{\text{prof}}) + (C_{\text{prof}} - K_f).$$

The second term is the bias of the three-parameter model: it is the error that the fit makes on the profile itself, which has no finite-x error. The first term is the difference between the data at $x = 256{,}000{,}000$ and the profile (Table 10.4), as seen through the fit. For $t^2 + 1$ and $2t + 1$ the model bias dominates (-0.003718 and -0.000928, against $+0.000110$ and $+0.000001$). For $t^3 + 2$ the data–profile term $+0.000528$ is larger than the model bias $+0.000186$, in line with the per-cent level deviations of the cubic data from the profile in Table 10.4. In this case most of the (small) offset from K_f comes from the finite-x behaviour of the data, not from the fitting model.

The model bias is not surprising. By (8.3) the profile $F_f(1/r)r^{-3/2}$ differs from K_f by a

	data		profile		decomposition		relative gap
f	C_{data}	envelope	C_{prof}	K_f	$C_{\text{data}} - C_{\text{prof}}$	$C_{\text{prof}} - K_f$	$K_f/C_{\text{data}} - 1$
t^2+1	0.134475	$\pm$0.007285	0.134364	0.138083	+0.000110	−0.003718	+2.68%
t^3+2	0.136321	$\pm$0.002404	0.135793	0.135608	+0.000528	+0.000186	−0.52%
$2t+1$	0.141686	$\pm$0.008218	0.141685	0.142613	+0.000001	−0.000928	+0.65%

Table 10.5: The snapshot's extrapolation C_{data} with its heuristic envelope; the same fit applied to the exact profile N_{pred} at the same actual ratios (C_{prof}); the constant K_f; and the decomposition $C_{\text{data}} - K_f = (C_{\text{data}} - C_{\text{prof}}) + (C_{\text{prof}} - K_f)$.

relative amount $O(r^{1/8})$. This is only an upper bound and does not determine the size or the shape of the correction. Empirically, Fig. 10.1 shows that the profile still varies noticeably below $r = 0.01$, and the model $C + a\sqrt{r} + br$ fitted on $r \geq 0.01$ does not reproduce its limit; the column $C_{\text{prof}} - K_f$ quantifies this. This is an observation about the computed profile, not a proved statement about F_f.

The ratio dependence is also consistent with the snapshot's log-log regressions (Table 10.6). At fixed ratio $V_f \approx x^2 F_f(1/r)$, so the slope in x should be close to 2, and the fits find $\alpha+\beta \approx 2$. Over the finite ratio range, however, the profile is not a pure power of r. So there is no reason for the fixed-x exponent β or the joint exponents to equal $\frac{3}{2}$ and $\frac{1}{2}$, and they do not.

f	fixed-x β	fixed-ratio $\alpha+\beta$	joint α	joint β	joint C
t^2+1	1.5454 (0.0039)	2.0008 (0.0006)	0.4543 (0.0013)	1.5445 (0.0011)	0.1731
t^3+2	1.5383 (0.0037)	1.9978 (0.0021)	0.4592 (0.0021)	1.5301 (0.0017)	0.1957
$2t+1$	1.5512 (0.0075)	2.0001 (0.0001)	0.4488 (0.0024)	1.5514 (0.0020)	0.1784

Table 10.6: Snapshot regressions of $\log V$ on $\log f(y)$ (fixed $x = 256{,}000{,}000$), on $\log x$ (fixed ratio: the mean of the eight ratio-specific slopes), and jointly on both. In parentheses: formal OLS standard errors for the fixed-x and joint fits; for the fixed-ratio column, the standard error of the mean of the eight slopes, that is, their sample standard deviation divided by $\sqrt{8}$, which is not a regression standard error. None of these is a rigorous uncertainty.

10.5 Limitations

The numerical evidence concerns three polynomials and $x \leq 2.56 \cdot 10^8$. In the cubic case, y is small and the agreement is only at the per-cent level. All uncertainties attached to fits are heuristic. The snapshot's validation tests the implementation on small cases and spot checks, not every production count. In the recomputation, every archived row was reproduced exactly; this confirms reproducibility, not the correctness of the shared code. The constants are computed in double precision from truncated Euler products, with a rigorous bound for the truncation of K_f but without certified rounding.

Appendix A

Full numerical tables

The tables list the normalised values $V_f(x,y)/(x^{1/2}f(y)^{3/2})$ for all x and all eight target ratios r (columns), as recomputed by `reproduce.py` with the snapshot code. The integer y is the largest with $f(y) \le rx$; the actual ratios $f(y)/x$ are recorded in `reproduction.json`.

x	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$2.5 \cdot 10^5$	0.14265	0.14605	0.14655	0.15222	0.15742	0.16378	0.17059	0.17164
$5 \cdot 10^5$	0.14256	0.14362	0.14505	0.15149	0.15742	0.16328	0.17031	0.17137
$1 \cdot 10^6$	0.14185	0.14316	0.14489	0.15191	0.15698	0.16318	0.17025	0.17136
$2 \cdot 10^6$	0.14189	0.14386	0.14533	0.15130	0.15688	0.16305	0.17010	0.17127
$4 \cdot 10^6$	0.14169	0.14330	0.14469	0.15128	0.15675	0.16294	0.17003	0.17122
$8 \cdot 10^6$	0.14116	0.14316	0.14470	0.15120	0.15671	0.16289	0.17003	0.17118
$1.6 \cdot 10^7$	0.14117	0.14309	0.14462	0.15115	0.15663	0.16283	0.16999	0.17115
$3.2 \cdot 10^7$	0.14105	0.14302	0.14458	0.15106	0.15660	0.16281	0.16996	0.17114
$6.4 \cdot 10^7$	0.14095	0.14297	0.14450	0.15101	0.15658	0.16279	0.16994	0.17112
$1.28 \cdot 10^8$	0.14091	0.14291	0.14447	0.15099	0.15656	0.16279	0.16993	0.17111
$2.56 \cdot 10^8$	0.14086	0.14292	0.14446	0.15098	0.15655	0.16278	0.16993	0.17110

Table A.1: Normalised variance $V_f(x,y)/(x^{1/2}f(y)^{3/2})$ for $f(t) = t^2 + 1$; rows: x, columns: target ratio r.

x	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$2.5 \cdot 10^5$	0.15991	0.16023	0.16097	0.15991	0.16317	0.16139	0.16939	0.17239
$5 \cdot 10^5$	0.15511	0.16267	0.15636	0.16132	0.15940	0.15903	0.16778	0.17086
$1 \cdot 10^6$	0.15356	0.14941	0.15418	0.15503	0.15509	0.15858	0.16628	0.17071
$2 \cdot 10^6$	0.14804	0.14791	0.15113	0.15002	0.15465	0.15769	0.16692	0.16960
$4 \cdot 10^6$	0.14472	0.14776	0.14625	0.15160	0.15478	0.15714	0.16521	0.16842
$8 \cdot 10^6$	0.14196	0.14483	0.14587	0.14956	0.15421	0.15751	0.16558	0.16824
$1.6 \cdot 10^7$	0.14240	0.14436	0.14442	0.14975	0.15358	0.15670	0.16484	0.16787
$3.2 \cdot 10^7$	0.14215	0.14278	0.14431	0.14905	0.15305	0.15610	0.16460	0.16738
$6.4 \cdot 10^7$	0.14043	0.14296	0.14392	0.14959	0.15260	0.15583	0.16453	0.16724
$1.28 \cdot 10^8$	0.14127	0.14195	0.14439	0.14896	0.15246	0.15594	0.16423	0.16699
$2.56 \cdot 10^8$	0.13999	0.14120	0.14391	0.14848	0.15234	0.15550	0.16399	0.16671

Table A.2: Normalised variance $V_f(x,y)/(x^{1/2}f(y)^{3/2})$ for $f(t) = t^3 + 2$; rows: x, columns: target ratio r.

x	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$2.5 \cdot 10^5$	0.14376	0.14809	0.14931	0.15670	0.16173	0.16245	0.18240	0.18370
$5 \cdot 10^5$	0.14415	0.14823	0.14933	0.15672	0.16172	0.16246	0.18239	0.18369
$1 \cdot 10^6$	0.14439	0.14830	0.14932	0.15670	0.16170	0.16245	0.18239	0.18369
$2 \cdot 10^6$	0.14444	0.14828	0.14931	0.15669	0.16170	0.16245	0.18239	0.18369
$4 \cdot 10^6$	0.14451	0.14833	0.14932	0.15669	0.16171	0.16245	0.18239	0.18369
$8 \cdot 10^6$	0.14459	0.14837	0.14935	0.15671	0.16171	0.16245	0.18239	0.18369
$1.6 \cdot 10^7$	0.14461	0.14838	0.14935	0.15671	0.16171	0.16245	0.18239	0.18369
$3.2 \cdot 10^7$	0.14459	0.14836	0.14934	0.15670	0.16171	0.16245	0.18239	0.18369
$6.4 \cdot 10^7$	0.14459	0.14837	0.14934	0.15670	0.16171	0.16245	0.18239	0.18369
$1.28 \cdot 10^8$	0.14460	0.14837	0.14934	0.15670	0.16171	0.16245	0.18239	0.18369
$2.56 \cdot 10^8$	0.14460	0.14837	0.14934	0.15670	0.16171	0.16245	0.18239	0.18369

Table A.3: Normalised variance $V_f(x, y)/(x^{1/2} f(y)^{3/2})$ for $f(t) = 2t + 1$; rows: x, columns: target ratio r.

Appendix B

Provenance of this document

Source. This document presents the proof recorded in an archived run of the project repository

https://gitlab.com/mshepherd/squarefree-numbers.

The run is labelled `zt-thesis-plan`. The provenance record lists three distinct revisions:

archive commit	`867fc9baf25edf4995748d91dd950e98cfd9debf`
harness commit	`552a8f01b9aa8504bbf03e82d149b1fdb3913980`
revision in every model record	`fc217220e6b3a0e570ab0acdb78e1a50f51a2e13`

The run followed the plan `thesis@gpt-6.1-sol`. Of the competing plans, this one was selected by Markus Schepke, the project owner. It decomposes the problem into seven propositions P1–P7, which correspond to Chapters 3 to 9. For this run the problem statement was accompanied by a short hint, written by the run operator with the project owner's authorisation, that pointed towards Vaughan's paper [11] and the work of Brüdern and Wooley [2] without describing a method. According to the intervention record, the owner authorised the operator to read the owner's master's thesis on this problem, which the project otherwise excludes, in order to write this hint; the run label `thesis-plan` reflects this. The numerical snapshot was reused, with the owner's authorisation, from the numerics stage of an earlier run of the same problem. Its acceptance record describes it as "operator selected; not a proof of numerical correctness".

Models. All mathematical content of the archive was produced by language models in the following roles; the recorded identifiers are model identifiers and command-line tool versions, not model weight versions, and no further version information is available.

role	model identifier	tool version
planner	`claude-opus-5-5`	Claude Code 2.1.311
planner	`codex:gpt-6.1-sol`	codex-cli 0.166.0
prover	`claude-sonnet-5-5`	Claude Code 2.1.311
prover	`codex:gpt-6.1-sol`	codex-cli 0.166.0
referee	`claude-opus-5-5`	Claude Code 2.1.311
referee	`codex:gpt-6.1-sol`	codex-cli 0.166.0
integrator	`claude-opus-5-5`	Claude Code 2.1.311
integrator	`codex:gpt-6.1-sol`	codex-cli 0.166.0

The acceptance record of the numerical snapshot names `codex-gpt-6.1-sol` as its candidate label. This label is not one of the model records above, and no tool version is recorded for the numerics stage. This write-up was produced by `claude-opus-5-5` (Claude Code 2.1.311) from the archived plan, statements, proofs, reviews, integration reports and numerical snapshot. The models designated to review the write-up are `claude-opus-5-5` and `codex:gpt-6.1-sol`.

Status of the evidence. Each of P1–P7 was approved in its first version by two independent model referees, one from each model family; their reports are archived next to the proofs. Both integration reports concluded that the propositions compose to the main theorem without loss. This agreement between models is evidence, not certification. Corollary 1.2 is the "in particular" clause of the plan's main theorem. It is not one of the approved propositions; its deduction from (1.9) (Section 9.3) is carried out in both integration reports, with $\varepsilon = \delta$ and $\varepsilon = \delta/2$ respectively. While preparing this document, every argument was rechecked step by step, and no mathematical gap was found. The changes made are editorial: notation, merging of repeated arguments into shared lemmas, a few explicit intermediate steps, and the fuller statement of some constants. Remark 1.4, the truncation bound, the consistency checks and the decomposition of the extrapolation offset in Chapter 10, and the corresponding parts of `reproduce.py` were added while preparing this document. The file `changes.md` lists every change and maps each result to its archived source.

Literature. The proof quotes only Lemma 2.2 of [1] and Corollary 10.5 of [9]; both statements were checked against the supplied copies. The other references serve as background. Bibliographic data for works not supplied as full text were taken from the reference lists of the supplied papers and from the project's bibliography candidate list, and were not checked against the originals.

Bibliography

[1] J. Brüdern, A. Granville, A. Perelli, R. C. Vaughan, and T. D. Wooley. On the exponential sum over k-free numbers. *Philos. Trans. Roy. Soc. London Ser. A*, 356:739–761, 1998. doi:10.1098/rsta.1998.0183.

[2] J. Brüdern and T. D. Wooley. Sparse variance for primes in arithmetic progression. *Quart. J. Math.*, 62:289–305, 2011. doi:10.1093/qmath/hap029.

[3] M. J. Croft. Square-free numbers in arithmetic progressions. *Proc. London Math. Soc. (3)*, 30:143–159, 1975. doi:10.1112/plms/s3-30.2.143.

[4] D. A. Goldston and R. C. Vaughan. On the Montgomery–Hooley asymptotic formula. In *Sieve Methods, Exponential Sums, and their Applications in Number Theory (Cardiff, 1995)*, volume 237 of *London Math. Soc. Lecture Note Ser.*, pages 117–142. Cambridge University Press, Cambridge, 1997. doi:10.1017/CBO9780511526091.010.

[5] C. Hooley. On the Barban–Davenport–Halberstam theorem. I. *J. Reine Angew. Math.*, 274/275:206–223, 1975. doi:10.1515/crll.1975.274-275.206.

[6] H. Mikawa and T. P. Peneva. Primes in arithmetic progressions to spaced moduli. *Arch. Math. (Basel)*, 84:239–248, 2005. doi:10.1007/s00013-004-1131-y.

[7] L. Mirsky. On the frequency of pairs of square-free numbers with a given difference. *Bull. Amer. Math. Soc.*, 55:936–939, 1949. doi:10.1090/S0002-9904-1949-09313-8.

[8] H. L. Montgomery. Primes in arithmetic progressions. *Michigan Math. J.*, 17:33–39, 1970. doi:10.1307/mmj/1029000373.

[9] H. L. Montgomery and R. C. Vaughan. *Multiplicative Number Theory I. Classical Theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007. doi:10.1017/CBO9780511618314.

[10] R. C. Vaughan. On a variance associated with the distribution of general sequences in arithmetic progressions. I. *Philos. Trans. Roy. Soc. London Ser. A*, 356:781–791, 1998. doi:10.1098/rsta.1998.0185.

[11] R. C. Vaughan. A variance for k-free numbers in arithmetic progressions. *Proc. London Math. Soc. (3)*, 91:573–597, 2005. doi:10.1112/S0024611505015352.

[12] R. Warlimont. Squarefree numbers in arithmetic progressions. *J. London Math. Soc. (2)*, 22:21–24, 1980. doi:10.1112/jlms/s2-22.1.21.