The sparse variance of squarefree numbers in arithmetic progressions to polynomial moduli: an elementary asymptotic formula

Write-up of an archived machine-generated proof run
(provenance 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 $t \geq 1$, consider

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

where $Q(x;q,l)$ counts squarefree $n \leq x$ with $n \equiv l \pmod q$ and $g(q,l)$ is the corresponding density. We show by elementary means that for each fixed $\delta \in (0, 1/9)$, $V_f(x,y) = C_f\, x^{1/2} f(y)^{3/2}\big(1 + O(x^{-\eta})\big)$ uniformly for $x^{7/9+\delta} \leq f(y) \leq x^{1-\delta}$, with an explicit Euler product $C_f > 0$ and $\eta = \min(\delta/12, 1/(6d))$; the implied constant depends only on f and δ. We also give explicit error terms valid throughout $f(y) \leq x$, and a version with an exact arithmetic main term. The proof rests on an exact identity for the variance with respect to a single modulus, a uniform count of pairs of squarefree numbers with a given difference, an Euler–Maclaurin analysis of sawtooth sums $\sum_s B(Z/s^2)$, and an average of a multiplicative function over polynomial values. We compare the result with exact numerical computations of V_f for $f = k^2 + 1$, $k^3 + 2$ and $2k + 1$ up to $x = 2.56 \cdot 10^8$.

Contents

1 Introduction **2**
1.1 The problem . . . 2
1.2 Background . . . 3
1.3 The main result . . . 3
1.4 Strategy . . . 4
1.5 Notation and conventions . . . 5

2 Preliminaries **5**
2.1 Counting . . . 5
2.2 Euler products . . . 5
2.3 Euler–Maclaurin summation . . . 6

3 Densities of squarefree numbers in progressions **7**

4 An exact decomposition of the single-modulus variance **10**

5 Pairs of squarefree numbers with a given difference **11**

6 Sawtooth sums and the main term for one modulus **13**
6.1 Sawtooth sums . . . 13
6.2 The main term for one modulus . . . 15

7 The variance for a single modulus **16**

8 Averages over polynomial moduli **17**
8.1 Preliminaries 17
8.2 Proof of Proposition 8.1(0), (b) and (c) 19
8.3 Proof of Proposition 8.1(a) 19

9 Proof of the main theorem **21**
9.1 Part (ii) 21
9.2 Part (i) 21
9.3 Part (iii) 22

10 Numerical evidence **22**
10.1 What was computed 22
10.2 The constants C_f 23
10.3 The data and the predicted constant 24
10.4 The exact-shape statement 25
10.5 Limitations 28

A Full tables of normalised values **29**

B Provenance **31**

1 Introduction

1.1 The problem

Let $\mu^2(n)$ denote the indicator function of the squarefree positive integers. For real $x \geq 1$, $q \in \mathbb{N}$ and $l \in \mathbb{Z}$ put

$$Q(x;q,l) = \#\{1 \leq n \leq x : \ n \equiv l \pmod q,\ \mu^2(n) = 1\}, \qquad Q(x) = Q(x;1,0).$$

The limit $g(q,l) = \lim_{x\to\infty} Q(x;q,l)/x$ exists (Proposition 3.1) and is the density of squarefree numbers in the class l modulo q. Its error term and the variance with respect to one modulus are

$$E(x;q,l) = Q(x;q,l) - g(q,l)\,x, \qquad V(x;q) = \sum_{l=1}^{q} E(x;q,l)^2.$$

The Montgomery–Hooley-type variance $V(x,Y) = \sum_{q\leq Y} V(x;q)$ over all moduli $q \leq Y$ was evaluated asymptotically by Vaughan [6]; we recall his result below.

This paper treats a *sparse* version, in which the moduli are restricted to the values of a polynomial. Throughout, $f \in \mathbb{Z}[t]$ is a polynomial of degree $d \geq 1$ such that

$$f(t) \geq 2 \quad \text{and} \quad f'(t) \geq 1 \qquad \text{for all real } t \geq 1. \tag{1.1}$$

In particular f has positive leading coefficient and is strictly increasing on $[1,\infty)$. We study

$$V_f(x,y) = \sum_{1\leq k\leq y} f'(k)\,V(x;f(k)) = \sum_{1\leq k\leq y} f'(k) \sum_{l=1}^{f(k)} E(x;f(k),l)^2, \qquad Y := f(y) \leq x, \tag{1.2}$$

where k runs over integers and $y \geq 1$ is real. The weight $f'(k)$ compensates for the sparseness of the moduli: by the mean value theorem, the number of integers in an interval $(f(k), f(k+1)]$ is about $f'(k)$, so $\sum_{k\leq y} f'(k)G(f(k))$ is a Riemann-sum approximation to $\int_{f(1)}^{Y} G(u)\,du$, and $V_f(x,y)$ is directly comparable with the full variance $V(x,Y)$.

The question is to determine the asymptotic behaviour of $V_f(x,y)$, with an explicit main term depending on f, in as large a range of $f(y) \leq x$ as possible.

1.2 Background

For the full variance, Vaughan [6, Theorem, (1.12)–(1.15)] proved, for k-free numbers in general and in the case $k = 2$ of squarefree numbers in particular, that whenever $Y \le x$,

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

with

$$c_2 = \frac{8C_1\,(-\zeta(-1/2))}{3\,\zeta(2)^2}, \qquad C_1 = \prod_p\Big(1 + \frac{(p-1)^2p^{3/2} - 2p^4 + 3p^3 - p}{p^2(p^2-1)^2}\Big). \tag{1.3}$$

(This is (1.14)–(1.15) of [6] with $k = 2$, written with Y in place of Vaughan's modulus bound Q, to avoid a clash with the counting function $Q(x)$; there x is assumed sufficiently large.) Vaughan's main term dominates when x/Y is large and Y is large compared with $x^{2/3}\exp(-\frac{2}{3}\tilde{c}(\log x)^{3/5}(\log\log x)^{-1/5})$, a quantity slightly smaller than $x^{2/3}$ [6, (1.16)]. His proof uses the circle method, with major-arc information on the exponential sum over k-free numbers from [1]. Earlier results, quoted in [6, (1.6)–(1.11)], have the shape $V(x,Y) = x^2\mathcal{F}(x/Y) + O(\cdots)$, with an error term of size at least about $x^{3/2}$.

For primes, Brüdern and Wooley [2] proved an analogue of the Montgomery–Hooley formula in which the moduli are restricted to the values $f(k)$ of a polynomial of degree at least 2, with the same weight $f'(k)$. They adapted the circle-method approach to the Montgomery–Hooley formula that they attribute to Goldston and Vaughan [3] and to Vaughan [5]. The problem considered here is the squarefree analogue of their result.

The argument presented below uses no circle method. It is completely elementary, and the price is a narrower range of validity of the asymptotic formula, namely $x^{7/9+\delta} \le f(y) \le x^{1-\delta}$; see Remark 1.2.

1.3 The main result

We need some notation. For a prime p let

$$a_p = 1 - \frac{3}{p^2} + \frac{2}{p^3} = \frac{(p-1)^2(p+2)}{p^3}, \qquad b_p = \frac{(p-1)(p^2-2+\sqrt{p})}{p^3}, \qquad c_p = \frac{(p-1)^2(p+1)}{p^3}. \tag{1.4}$$

Let $\rho(m) = \#\{a \bmod m :\ f(a) \equiv 0 \pmod m\}$, and put

$$\sigma_p = a_p\Big(1 - \frac{\rho(p)}{p}\Big) + b_p\Big(\frac{\rho(p)}{p} - \frac{\rho(p^2)}{p^2}\Big) + c_p\,\frac{\rho(p^2)}{p^2}, \qquad C_f = \frac{2}{3}\,I\prod_p \sigma_p, \tag{1.5}$$

where

$$I = \int_0^\infty B(u^{-2})\,du = -4\zeta(-\tfrac{1}{2}) = \frac{\zeta(3/2)}{\pi} = 0.831544899909\ldots, \qquad B(u) = \{u\}(1-\{u\}), \tag{1.6}$$

and $\{u\}$ is the fractional part of u. The three weights in σ_p are the proportions of residues k modulo p^2 with $p \nmid f(k)$, with $p \parallel f(k)$ and with $p^2 \mid f(k)$, respectively. We also need the explicit single-modulus main term

$$\mathrm{MT}(x;q) = A_0\, q\sum_{t\ge1}\eta(t)\,m_t(q)\,B\Big(\frac{x}{q\,m_t(q)}\Big), \tag{1.7}$$

where

$$A_0 = \prod_p\Big(1 - \frac{2}{p^2}\Big), \qquad \eta(t) = \mu^2(t)\prod_{p\mid t}\frac{1}{p^2-2}, \qquad m_t(q) = \frac{t^2}{\gcd(t^2,q)}. \tag{1.8}$$

The series (1.7) converges absolutely, since $0 \le B(u) \le u$ (Lemma 2.4).

Theorem 1.1. *Let $f \in \mathbb{Z}[t]$ have degree $d \geq 1$ and satisfy* (1.1). *The product $\prod_p \sigma_p$ converges absolutely, and $0 < C_f < \infty$. For every $\varepsilon > 0$ there is a constant $c(f, \varepsilon)$, depending only on f and ε, such that for all real $x \geq 2$ and $y \geq 1$ with $Y = f(y) \leq x$:*

(i) $|V_f(x,y) - C_f\, x^{1/2} Y^{3/2}| \leq c(f,\varepsilon)\Big(x^{1/3+\varepsilon} Y^{5/3} + x^{5/3+\varepsilon} + x^{1/2+\varepsilon} Y^{3/2-1/(2d)}\Big)$;

(ii) $\Big|V_f(x,y) - \sum_{1 \leq k \leq y} f'(k)\, \mathrm{MT}(x; f(k))\Big| \leq c(f,\varepsilon)(x^{5/3+\varepsilon} + x^{1/2} Y)$.

(iii) *Consequently, for each fixed $\delta \in (0, 1/9)$, uniformly for $x^{7/9+\delta} \leq f(y) \leq x^{1-\delta}$,*

$$V_f(x,y) = C_f\, x^{1/2} f(y)^{3/2}(1 + O_{f,\delta}(x^{-\eta})), \qquad \eta = \min\Big(\frac{\delta}{12}, \frac{1}{6d}\Big).$$

Remark 1.2 (The range). Each error term in (i) comes from one step of the proof. The term $x^{5/3+\varepsilon}$ comes from the pointwise bound $x^{2/3+\varepsilon}$ for the error in the count of pairs of squarefree numbers with a given difference (Proposition 5.1), summed trivially over about x/q differences. It is smaller than $x^{1/2}Y^{3/2}$ only when $Y \geq x^{7/9+\delta}$. The term $x^{1/3+\varepsilon}Y^{5/3}$ comes from the Euler–Maclaurin error $Z^{1/3}$ in the sawtooth sums of Proposition 6.1. It has relative size $x^{\varepsilon}(Y/x)^{1/6}$, which forces $Y \leq x^{1-\delta}$. The term $x^{1/2+\varepsilon}Y^{3/2-1/(2d)}$ comes from averaging over the polynomial values (Proposition 8.1). Statement (ii) avoids the sawtooth asymptotics altogether: its error term is of smaller order than $x^{1/2}Y^{3/2}$ throughout $x^{7/9+\delta} \leq Y \leq x$, including $Y \asymp x$, where the arithmetic main term $\sum_k f'(k)\mathrm{MT}(x; f(k))$ is kept as it is.

Remark 1.3 (Consistency with the full variance). Formally taking $f(t) = t$ in (1.5), so that $\rho \equiv 1$, gives $\sigma_p = C_{1,p}(1 - p^{-2})^2$ identically in p, where $C_{1,p}$ is the factor of C_1 in (1.3). Since $\prod_p (1 - p^{-2})^2 = \zeta(2)^{-2}$, the constant $\frac{2}{3} I \prod_p \sigma_p$ equals Vaughan's c_2 exactly. The identity of local factors was checked with computer algebra, and both products are evaluated numerically in Table 1. The polynomial $f(t) = t$ violates $f(1) \geq 2$, so this is a consistency check and not a case of Theorem 1.1.

1.4 Strategy

The proof has four parts.

An exact identity for one modulus (Section 4). Expanding the square, $\sum_l Q(x; q, l)^2$ counts pairs $n \equiv n' \pmod q$ of squarefree numbers up to x, that is, the diagonal together with pairs $(n, n+hq)$. Replace the number of pairs with difference m by its expected value $A(m)(x-m)$, where $A(m)$ is the density of pairs of squarefree numbers with difference m. Summing $A(hq)(x - hq)$ over h is an arithmetic-progression sum. The terms of order x^2/q and x in this sum cancel *exactly* against $x^2 \sum_l g(q,l)^2$ and the diagonal. The remainder is the lattice-point discrepancy $\mathrm{MT}(x; q)$, which arises from the identity $2\sum_{1 \leq j \leq T}(T - j) = T^2 - T + B(T)$. What is left is controlled by three explicit error terms.

Error terms (Sections 3 and 5). The pair counts are evaluated with a uniform error $O(x^{2/3+\varepsilon})$ by a Möbius truncation at $x^{1/3}$ (Proposition 5.1). The remaining terms are of size $O(x^{1/2})$ and $O(x^{3/2}/q)$ (Proposition 3.3).

The main term (Section 6). After a factorisation of t according to q, $\mathrm{MT}(x; q)$ is a weighted sum of sawtooth sums $\sum_s B(Z/s^2)$. By Euler–Maclaurin summation such a sum equals $I\sqrt{Z} + O(Z^{1/3})$, which gives $\mathrm{MT}(x; q) = h(q)\sqrt{xq} + O_\varepsilon(x^{1/3} q^{2/3+\varepsilon})$ with an explicit multiplicative-type function h (Proposition 6.2). This is the source of the order $x^{1/2}q^{1/2}$ of $V(x; q)$.

Averaging over polynomial values (Section 8). Finally, $\sum_{k \leq y} f'(k) f(k)^{1/2} h(f(k))$ is evaluated by expanding h as a divisor sum and counting k in the $\rho(e)$ residue classes modulo e on which $e \mid f(k)$. This gives $C_f Y^{3/2}$, where $\frac{2}{3}Y^{3/2}$ comes from $\int f' f^{1/2} = \frac{2}{3} f^{3/2}$.

Section 9 combines these steps and composes the error terms explicitly. Section 10 compares Theorem 1.1 with exact computations of $V_f(x, y)$ for three polynomials.

1.5 Notation and conventions

The letter p always denotes a prime. We write $p^j \parallel q$ if $p^j \mid q$ and $p^{j+1} \nmid q$, $v_p(q)$ for the exponent of p in q, $\omega(n)$ for the number of distinct prime factors of n, $\tau(n)$ for the number of divisors of n, and $\mathrm{rad}(n) = \prod_{p \mid n} p$. Further, $\mathbf{1}[\cdot]$ is the indicator of a statement, $(a,b) = \gcd(a,b)$ and $[a,b] = \mathrm{lcm}(a,b)$, and $\{u\} = u - \lfloor u \rfloor$. A quantity written θ (possibly with an index) is a real number with $|\theta| \le 1$, which may differ at each occurrence. Implied constants in O, $\ll$ and $\gg$ and constants $c(\cdot)$ depend only on the parameters shown. The symbols B, I, A_0, η, $m_t(q)$, MT, a_p, b_p, c_p, ρ, σ_p and C_f have the fixed meanings given above. In addition,

$$\kappa(t) = t^2\eta(t) = \mu^2(t)\prod_{p \mid t}\frac{p^2}{p^2-2}, \qquad S_2(q) = \sum_{l=1}^{q} g(q,l)^2, \qquad F(t) = f'(t)f(t)^{1/2}, \tag{1.9}$$

and the multiplicative-type function

$$h(q) = I\prod_{p \nmid q} a_p \prod_{p \parallel q} b_p \prod_{p^2 \mid q} c_p. \tag{1.10}$$

2 Preliminaries

This section collects the elementary tools used repeatedly later. The only external inputs are standard facts from Montgomery and Vaughan [4], cited precisely where they are used.

2.1 Counting

Lemma 2.1 (A tail bound)**.** *For every real $w > 0$, $\displaystyle\sum_{a > w} a^{-2} \le \frac{2}{w}$.*

Proof. If $w \ge 1$, let $k = \lfloor w \rfloor + 1 > w$. Then $\sum_{a \ge k} a^{-2} \le k^{-2} + \int_k^\infty u^{-2}\,du \le 2/k < 2/w$. If $0 < w < 1$, the sum is at most $\zeta(2) < 2 < 2/w$. □

Lemma 2.2 (Integers in a residue class)**.** *Let $N \ge 0$ be real, $L \in \mathbb{N}$ and $c \in \mathbb{Z}$. Then $\#\{1 \le n \le N : \ n \equiv c \pmod L\} = N/L + \theta$ with $|\theta| < 1$.*

Proof. Let $n_0 \in \{1, \dots, L\}$ be the least positive element of the class. If $n_0 > N$, the count is 0, and $0 \le N/L < n_0/L \le 1$. Otherwise the count is $\lfloor (N - n_0)/L \rfloor + 1$, which lies in $\big((N-n_0)/L, (N-n_0)/L + 1\big]$. This interval is contained in $(N/L - 1, N/L + 1)$, because $1 \le n_0 \le L$. □

Lemma 2.3 (Divisor bounds)**.** *For every $\varepsilon > 0$ there is $C(\varepsilon)$ with $\tau(n) \le C(\varepsilon) n^{\varepsilon}$ for all $n \in \mathbb{N}$ [4, (2.20)]. Moreover $2^{\omega(n)} \le \tau(n)$ and $\sum_{e \mid n} 4^{\omega(e)} \le \tau(n)^3$.*

Proof. The first statement is [4, (2.20)]. Each subset of the prime divisors of n gives a divisor, so $2^{\omega(n)} \le \tau(n)$. For $e \mid n$ we have $4^{\omega(e)} = (2^{\omega(e)})^2 \le \tau(e)^2 \le \tau(n)^2$, and there are $\tau(n)$ divisors e. □

2.2 Euler products

We use the following standard fact in the form given by Montgomery and Vaughan [4, Theorem 1.9] (with $s = 0$): *if G is multiplicative and $\sum_{n \ge 1} |G(n)| < \infty$, then*

$$\sum_{n \ge 1} G(n) = \prod_{p}\Big(1 + \sum_{j \ge 1} G(p^j)\Big). \tag{2.1}$$

We also use that if $\sum_p |u_p| < \infty$ and $1 + u_p \ne 0$ for all p, then $\prod_p (1 + u_p)$ converges absolutely to a nonzero limit, and such products may be multiplied and divided factor by factor. Absolute convergence of $\sum |G(n)|$ for a multiplicative G follows from $\prod_p (1 + \sum_j |G(p^j)|) < \infty$, because every partial sum of $\sum |G(n)|$ is bounded by a finite subproduct.

Lemma 2.4 (The functions η and κ)**.** *The functions η and κ of* (1.8) *and* (1.9) *are multiplicative, nonnegative, and supported on squarefree integers. For all t,*

$$0 \le \eta(t) \le 2^{\omega(t)} t^{-2}, \qquad 0 \le \kappa(t) \le 2^{\omega(t)}. \tag{2.2}$$

Further, $0 < A_0 \le 1$ and

$$\sum_{t \ge 1} \eta(t) = \prod_p \Big(1 + \frac{1}{p^2 - 2}\Big) = \frac{6}{\pi^2 A_0}. \tag{2.3}$$

For all real $x > 0$ and $q \in \mathbb{N}$, the terms of the series (1.7) *satisfy $0 \le \eta(t) m_t(q) B\big(x/(q m_t(q))\big) \le \eta(t)\, x/q$. Hence $\mathrm{MT}(x;q)$ converges absolutely and $0 \le \mathrm{MT}(x;q) \le 6x/\pi^2$.*

Proof. Multiplicativity and the support are clear from the definitions. Since $p^2 - 2 \ge p^2/2$ for every prime, $\eta(p) \le 2/p^2$ and $\kappa(p) \le 2$, which gives (2.2). Each factor $1 - 2/p^2$ lies in $[\frac{1}{2}, 1)$ and $\sum_p 2/p^2 < \infty$, so $0 < A_0 \le 1$. By (2.1), applied to η (whose series converges since $\sum_t 2^{\omega(t)} t^{-2} \le \sum_t \tau(t) t^{-2} = \zeta(2)^2$ by [4, Theorem 1.8]),

$$\sum_t \eta(t) = \prod_p \Big(1 + \frac{1}{p^2 - 2}\Big) = \prod_p \frac{1 - p^{-2}}{1 - 2p^{-2}} = \frac{\prod_p (1 - p^{-2})}{A_0} = \frac{6}{\pi^2 A_0},$$

using $\prod_p (1 - p^{-2}) = 1/\zeta(2) = 6/\pi^2$. Finally $0 \le B(u) \le \min(u, \frac{1}{4})$ for $u \ge 0$, so $m\, B(x/(qm)) \le x/q$, and $\mathrm{MT}(x;q) \le A_0 q \cdot (x/q) \sum_t \eta(t) = 6x/\pi^2$. □

The next evaluation occurs twice, in the second moment of g (Proposition 3.1(iv)) and in the density of pairs of squarefree numbers (Proposition 5.1).

Lemma 2.5 (Coprime pairs of squarefree numbers)**.** *For every $e \in \mathbb{N}$,*

$$\sum_{\substack{a,b \ge 1,\ \mu^2(ab) = 1 \\ (ab, e) = 1}} \frac{\mu(a)\mu(b)}{a^2 b^2} = \prod_{p \nmid e} \Big(1 - \frac{2}{p^2}\Big) = A_0 \prod_{p \mid e} \frac{p^2}{p^2 - 2},$$

and the series converges absolutely. Here $\mu^2(ab) = 1$ means exactly that a and b are squarefree and coprime.

Proof. Group the pairs according to $n = ab$. For a squarefree n there are $2^{\omega(n)}$ pairs (a, b) with $ab = n$, each with $\mu(a)\mu(b) = \mu(n)$. The sum is therefore $\sum_n G(n)$ with $G(n) = \mu(n) 2^{\omega(n)} n^{-2} \mathbf{1}[(n, e) = 1]$, which is multiplicative with $\sum |G(n)| \le \zeta(2)^2$. By (2.1) the sum equals $\prod_{p \nmid e} (1 - 2/p^2)$. The last expression follows by dividing A_0 by the finitely many factors $1 - 2/p^2 = (p^2 - 2)/p^2 \ne 0$ with $p \mid e$. □

2.3 Euler–Maclaurin summation

Put $\psi(u) = \{u\} - \frac{1}{2}$ and

$$\psi_2(u) = \tfrac{1}{2}\big(\{u\}^2 - \{u\}\big) + \tfrac{1}{12} = \tfrac{1}{12} - \tfrac{1}{2} B(u). \tag{2.4}$$

Then ψ_2 is continuous and 1-periodic, $\psi_2' = \psi$ off the integers, $|\psi_2| \le \frac{1}{12}$ and $|\psi| \le \frac{1}{2}$, and $\psi_2(n) = \frac{1}{12}$, $\psi(n) = -\frac{1}{2}$ for $n \in \mathbb{Z}$.

Lemma 2.6 (Euler–Maclaurin)**.** *Let $a < b$ be real and let $G \in C^2[a, b]$ be real- or complex-valued. Then*

$$\sum_{a < n \le b} G(n) = \int_a^b G(u)\, du + \psi(a)G(a) - \psi(b)G(b) - \psi_2(a)G'(a) + \psi_2(b)G'(b) - \int_a^b \psi_2(u) G''(u)\, du.$$

Suppose now that G is C^2 on $[a,\infty)$, that $G(u)\to 0$ and $G'(u)\to 0$ as $u\to\infty$, and that $G''\in L^1[a,\infty)$. Then

$$\lim_{b\to\infty}\Big(\sum_{a<n\le b} G(n) - \int_a^b G(u)\,du\Big) = \psi(a)G(a) - \psi_2(a)G'(a) - \int_a^\infty \psi_2(u)G''(u)\,du.$$

In particular, if one of $\sum_{n>a} G(n)$ and $\int_a^\infty G(u)\,du$ converges, then so does the other, and the formula holds with $b=\infty$ and the terms at b omitted.

Proof. Since $\lfloor u\rfloor = u - \frac12 - \psi(u)$, we have $\sum_{a<n\le b} G(n) = \int_{(a,b]} G\,d\lfloor u\rfloor = \int_a^b G\,du - \int_{(a,b]} G\,d\psi$. As ψ is right-continuous and of bounded variation and G is continuous, integration by parts for Riemann–Stieltjes integrals gives $\int_{(a,b]} G\,d\psi = G(b)\psi(b) - G(a)\psi(a) - \int_a^b \psi G'\,du$. Finally $\int_a^b \psi G'\,du = [\psi_2 G']_a^b - \int_a^b \psi_2 G''\,du$. (As a check, $G\equiv 1$ gives $\lfloor b\rfloor - \lfloor a\rfloor$.) For complex-valued G, apply this to the real and imaginary parts.

For $b\to\infty$, move $\int_a^b G\,du$ to the left-hand side of the finite formula. On the right, the terms at b tend to 0, because ψ and ψ_2 are bounded and $G(b), G'(b)\to 0$. The last integral converges to $\int_a^\infty \psi_2 G''\,du$, because ψ_2 is bounded and $G''\in L^1$. This proves the limit formula. The partial sums $\sum_{a<n\le b} G(n)$, for real b, are those of the series $\sum_{n>a} G(n)$. So if the series converges, then $\int_a^b G\,du$ converges as $b\to\infty$, and conversely; in either case the limit formula becomes the stated formula with $b=\infty$. (The series and the integral need not converge separately: $G(u)=\cos(\log u)/u$ on $[1,\infty)$ satisfies the hypotheses, but $\int_1^X G\,du = \sin(\log X)$ has no limit as $X\to\infty$, and hence neither have the partial sums of $\sum_{n\ge 2} G(n)$.) □

In the applications below (Proposition 6.1), G is u^{-s} with $\Re s>1$ or a linear combination of u^{-2} and u^{-4}. In both cases the series and the integral converge absolutely, so the last statement of the lemma applies.

Lemma 2.6 is the Euler–Maclaurin formula of the first and second order; compare [4, §1.3], where the continuation of $\zeta(s)$ to $\Re s>0$ is obtained in the same way [4, Theorem 1.12].

3 Densities of squarefree numbers in progressions

We first determine $g(q,l)$ and its first two moments over l. We then show that two of the four terms in the decomposition of Section 4 are small. Throughout this section $q\in\mathbb{N}$ is fixed, and for $d\in\mathbb{N}$ we write

$$g_d = \gcd(d^2, q).$$

Proposition 3.1 (Closed form and moments of g)**.** *For every $q\in\mathbb{N}$ and $l\in\mathbb{Z}$:*

(i) *The limit $g(q,l) = \lim_{x\to\infty} Q(x;q,l)/x$ exists, and*

$$g(q,l) = \sum_{d\ge 1} \mu(d)\,\frac{g_d}{d^2 q}\,\mathbf{1}[g_d \mid l], \tag{3.1}$$

where the series converges absolutely. Moreover $0\le g(q,l)\le 1/q$, and $g(q,l)$ depends only on l mod q.

(ii) $\sum_{l=1}^{q} g(q,l) = 6/\pi^2$.

(iii) *For every squarefree $n\in\mathbb{N}$,*

$$g(q,n) = \frac{6}{\pi^2}\,\frac{1}{q}\prod_{p\mid q}\Big(1-\frac{1}{p^2}\Big)^{-1}\prod_{\substack{p\mid(n,q)\\ p\,\|\,q}}\Big(1-\frac{1}{p}\Big).$$

(iv) $S_2(q) = \sum_{l=1}^{q} g(q,l)^2 = \dfrac{A_0}{q}\sum_{t\ge 1}\eta(t)\dfrac{\gcd(t^2,q)}{t^2} = \dfrac{A_0}{q}\sum_{t\ge 1}\dfrac{\eta(t)}{m_t(q)}$.

(v) $\sum_{t\ge 1}\eta(t) = 6/(\pi^2 A_0)$.

Proof. (i) We use $\mu^2(n) = \sum_{d^2\mid n}\mu(d)$. For $d > \sqrt{x}$ no $n \le x$ is divisible by d^2, so

$$Q(x;q,l) = \sum_{d\le\sqrt{x}}\mu(d)N_d(x), \qquad N_d(x) = \#\{n \le x :\ n \equiv l\ (q),\ d^2 \mid n\}.$$

Write $n = d^2 m$ with $1 \le m \le x/d^2$. The congruence $d^2 m \equiv l \pmod q$ is solvable in m if and only if $g_d \mid l$. If $g_d \nmid l$, then $N_d(x) = 0$. If $g_d \mid l$, the congruence is equivalent to $(d^2/g_d)m \equiv l/g_d \pmod{q/g_d}$. As $(d^2/g_d, q/g_d) = 1$, this is a single residue class modulo q/g_d, and Lemma 2.2 gives $N_d(x) = \frac{g_d}{d^2 q}x + \theta_d$. Hence

$$Q(x;q,l) = x\sum_{d\le\sqrt{x}}\mu(d)\frac{g_d}{d^2 q}\mathbf{1}[g_d \mid l] + \theta\sqrt{x}.$$

Since $g_d \le q$, the d-th term of (3.1) has modulus at most d^{-2}. So the series converges absolutely, and by Lemma 2.1 its tail beyond $\sqrt{x}$ is at most $2/\sqrt{x}$ for $x \ge 1$. Therefore

$$|Q(x;q,l) - x\,g(q,l)| \le 3\sqrt{x} \qquad (x \ge 1), \tag{3.2}$$

with $g(q,l)$ defined by (3.1), and $Q(x;q,l)/x \to g(q,l)$. From $0 \le Q(x;q,l) \le x/q + 1$ we get $0 \le g(q,l) \le 1/q$. Since $g_d \mid q$, the condition $g_d \mid l$ depends only on $l \bmod q$.

(ii) For each d, exactly q/g_d of the integers $1 \le l \le q$ are divisible by g_d. Summing the absolutely convergent series (3.1) over the q values of l and interchanging the sums gives $\sum_l g(q,l) = \sum_d \mu(d)d^{-2} = 6/\pi^2$.

(iii) Let n be squarefree. Only squarefree d contribute to (3.1). For squarefree d, $g_d = \prod_{p\mid d}\gcd(p^2,q)$ is a product of coprime prime powers, so $g_d \mid n$ if and only if $\gcd(p^2,q) \mid n$ for every $p \mid d$. Hence, for squarefree d,

$$\mu(d)\frac{g_d}{d^2}\mathbf{1}[g_d \mid n] = \prod_{p\mid d}u(p), \qquad u(p) = -\frac{\gcd(p^2,q)}{p^2}\mathbf{1}[\gcd(p^2,q) \mid n],$$

and $q\,g(q,n) = \sum_d \mu^2(d)\prod_{p\mid d}u(p)$. We have $|u(p)| = p^{-2}$ for $p \nmid q$, and $|u(p)| \le 1$ for the finitely many $p \mid q$. So $\sum_p |u(p)| < \infty$ and, by (2.1), $q\,g(q,n) = \prod_p(1+u(p))$. Now

- if $p \nmid q$, then $\gcd(p^2,q) = 1$ and $1 + u(p) = 1 - p^{-2}$;
- if $p \,\|\, q$, then $\gcd(p^2,q) = p$ and $1 + u(p) = 1 - p^{-1}\mathbf{1}[p \mid n]$;
- if $p^2 \mid q$, then $\gcd(p^2,q) = p^2 \nmid n$, since n is squarefree, and $1 + u(p) = 1$.

Hence $q\,g(q,n) = \prod_{p\nmid q}(1-p^{-2})\prod_{p\|q,\ p\mid n}(1-p^{-1})$. The claim follows from $\prod_{p\nmid q}(1-p^{-2}) = \frac{6}{\pi^2}\prod_{p\mid q}(1-p^{-2})^{-1}$.

(iv) Squaring (3.1) (a product of two absolutely convergent series) and summing over $1 \le l \le q$ gives

$$S_2(q) = \frac{1}{q^2}\sum_{d,e\ge 1}\mu(d)\mu(e)\frac{g_d g_e}{d^2 e^2}\,\#\{1 \le l \le q :\ g_d \mid l,\ g_e \mid l\}.$$

Both g_d and g_e divide q, so the count is $q/[g_d,g_e]$, and $g_d g_e/[g_d,g_e] = (g_d,g_e) = \gcd\big((d,e)^2,q\big)$. Thus, with $G(t) = \gcd(t^2,q) \le q$,

$$S_2(q) = \frac{1}{q}\sum_{d,e\ \text{squarefree}}\mu(d)\mu(e)\frac{G((d,e))}{d^2 e^2},$$

which converges absolutely. Write $t=(d,e)$, $d=td'$ and $e=te'$. This is a bijection from pairs of squarefree d,e onto triples (t,d',e') with t squarefree, $d'e'$ squarefree and $(d'e',t)=1$, and $\mu(d)\mu(e)=\mu(d')\mu(e')$. Summing over t last and using Lemma 2.5 with $e=t$,

$$S_2(q)=\frac{1}{q}\sum_{t\ \text{squarefree}}\frac{G(t)}{t^4}A_0\prod_{p\mid t}\frac{p^2}{p^2-2}=\frac{A_0}{q}\sum_{t\ge 1}\eta(t)\frac{\gcd(t^2,q)}{t^2},$$

because $t^{-4}\prod_{p\mid t}p^2/(p^2-2)=t^{-2}\eta(t)$ for squarefree t. Finally $\gcd(t^2,q)/t^2=1/m_t(q)$.

(v) This is (2.3). □

For $q=1$, (iv) reads $A_0\sum_t\eta(t)t^{-2}=\prod_p(1-p^{-2})^2=36/\pi^4=g(1,1)^2$, as it should.

The second ingredient is an elementary count of squarefree numbers coprime to a given squarefree number.

Lemma 3.2. *Let r be squarefree and put $T_r(y)=\#\{1\le m\le y:\ \mu^2(m)=1,\ (m,r)=1\}$ for real $y>0$. Then*

$$|T_r(y)-\nu_r y|\le 3\cdot 2^{\omega(r)}\sqrt{y},\qquad \nu_r=\sum_{d\ge 1}\sum_{e\mid r}\frac{\mu(d)\mu(e)}{[d^2,e]},$$

where the double series converges absolutely and $|\nu_r|\le 2\cdot 2^{\omega(r)}$. In particular $\nu_1=6/\pi^2$, so $|Q(y)-6y/\pi^2|\le 3\sqrt{y}$ for all $y>0$.

Proof. Insert $\mu^2(m)=\sum_{d^2\mid m}\mu(d)$ and $\mathbf{1}[(m,r)=1]=\sum_{e\mid(m,r)}\mu(e)$. This gives $T_r(y)=\sum_{d\ge1}\sum_{e\mid r}\mu(d)\mu(e)\lfloor y/[d^2,e]\rfloor$, where only $d\le\sqrt{y}$ contribute because $[d^2,e]\ge d^2$. The same inequality gives $|\nu_r|\le 2^{\omega(r)}\zeta(2)\le 2\cdot 2^{\omega(r)}$. If $y\ge 1$, there are at most $2^{\omega(r)}\sqrt{y}$ pairs (d,e) with $d\le\sqrt{y}$, $e\mid r$. Replacing each floor by $y/[d^2,e]$ costs at most 1 per pair. Completing the d-sum costs at most $y\,2^{\omega(r)}\sum_{d>\sqrt{y}}d^{-2}\le 2\cdot 2^{\omega(r)}\sqrt{y}$, by Lemma 2.1. If $0<y<1$, then $T_r(y)=0$ and $|\nu_r y|\le 2\cdot 2^{\omega(r)}y\le 2\cdot 2^{\omega(r)}\sqrt{y}$. For $r=1$ only $e=1$ occurs, and $\nu_1=\sum_d\mu(d)d^{-2}=6/\pi^2$. □

Proposition 3.3 (Two small terms)**.** *Let $c_0=3e^{2\zeta(3/2)}$. For all real $x\ge 1$ and all integers $1\le q\le x$:*

(i) $|Q(x)-6x/\pi^2|\le c_0\sqrt{x}$;

(ii) $\left|\sum_{l=1}^{q}g(q,l)E(x;q,l)\right|\le c_0\sqrt{x}/q$.

The hypothesis $q\le x$ is not used in the proof of (ii).

Proof. (i) This is the case $r=1$ of Lemma 3.2, since $3\le c_0$.

(ii) Let $q_1=\prod_{p\parallel q}p$. Since $g(q,\cdot)$ is q-periodic,

$$S(x):=\sum_{l=1}^{q}g(q,l)Q(x;q,l)=\sum_{n\le x}\mu^2(n)g(q,n),\qquad \sum_{l=1}^{q}g(q,l)E(x;q,l)=S(x)-x\,S_2(q).\quad(3.3)$$

For squarefree n the last product in Proposition 3.1(iii) runs over $p\mid(n,q_1)$, and $\prod_{p\mid(n,q_1)}(1-1/p)=\sum_{r\mid(n,q_1)}\mu(r)/r$. Put $c_q=\frac{6}{\pi^2}\prod_{p\mid q}(1-p^{-2})^{-1}$; then $0<c_q\le\frac{6}{\pi^2}\zeta(2)=1$ and $g(q,n)=\frac{c_q}{q}\sum_{r\mid q_1,\ r\mid n}\mu(r)/r$. Interchanging the finite sum over r with the sum over n,

$$S(x)=\frac{c_q}{q}\sum_{r\mid q_1}\frac{\mu(r)}{r}N_r(x),\qquad N_r(x)=\#\{n\le x:\ r\mid n,\ \mu^2(n)=1\}.$$

For squarefree r, $n = rm$ is squarefree if and only if m is squarefree and $(m, r) = 1$, so $N_r(x) = T_r(x/r)$. By Lemma 3.2,

$$S(x) = M_q\, x + R_q(x), \qquad M_q = \frac{c_q}{q} \sum_{r \mid q_1} \frac{\mu(r)\nu_r}{r^2}, \qquad |R_q(x)| \le \frac{3\sqrt{x}}{q} \sum_{r \text{ squarefree}} \frac{2^{\omega(r)}}{r^{3/2}}.$$

The last sum equals $\prod_p (1 + 2p^{-3/2}) \le \exp\bigl(2\sum_p p^{-3/2}\bigr) \le e^{2\zeta(3/2)}$, so $|R_q(x)| \le c_0\sqrt{x}/q$ for all $x > 0$.

It remains to show $M_q = S_2(q)$. As $x \to \infty$, $Q(x; q, l)/x \to g(q, l)$ for each of the finitely many l, so $S(x)/x \to S_2(q)$. On the other hand $S(x)/x = M_q + O(x^{-1/2})$. Hence $M_q = S_2(q)$, and by (3.3) the sum $\sum_l g(q, l)E(x; q, l)$ equals $R_q(x)$. □

4 An exact decomposition of the single-modulus variance

For real $x \ge 1$ and an integer $m \ge 1$ let

$$\begin{aligned} C_x(m) &= \#\{n \ge 1 :\ n + m \le x,\ \mu^2(n) = \mu^2(n + m) = 1\}, \\ A(m) &= A_0 \sum_{t^2 \mid m} \eta(t), \qquad D_x(m) = C_x(m) - A(m)(x - m). \end{aligned} \tag{4.1}$$

Here $A(m)$ is the density of integers n with n and $n + m$ both squarefree (Proposition 5.1), so $A(m)(x - m)$ is the expected value of $C_x(m)$ and $D_x(m)$ is its error.

Proposition 4.1 (Exact decomposition)**.** *For every real $x \ge 1$ and every $q \in \mathbb{N}$,*

$$V(x; q) = \mathrm{MT}(x; q) + 2 \sum_{1 \le h \le x/q} D_x(hq) + \Bigl(Q(x) - \frac{6x}{\pi^2}\Bigr) - 2x \sum_{l=1}^{q} g(q, l)E(x; q, l).$$

The sum over h is empty if $q > x$.

All four terms have a transparent meaning. The second term collects the errors in the counts of pairs, and the last two are of size $O(\sqrt{x})$ and $O(x^{3/2}/q)$ by Proposition 3.3. The main term $\mathrm{MT}(x; q)$ is what remains of the *expected* pair counts after an exact cancellation of the terms of order x^2/q and x.

Proof. Write $Q_l = Q(x; q, l)$, $g_l = g(q, l)$ and $E_l = Q_l - g_l x$. Expanding the square and using $\sum_l g_l Q_l = \sum_l g_l E_l + xS_2(q)$,

$$V(x; q) = \sum_l Q_l^2 - 2x \sum_l g_l Q_l + x^2 S_2(q) = \sum_{l=1}^{q} Q_l^2 - x^2 S_2(q) - 2x \sum_{l=1}^{q} g_l E_l. \tag{4.2}$$

Counting pairs. The sum $\sum_l Q_l^2$ counts ordered pairs (n, n') of squarefree integers in $[1, x]$ with $n \equiv n' \pmod q$. The diagonal contributes $Q(x)$. The pairs with $n < n'$ are the pairs $(n, n + hq)$ with $h \ge 1$ and $n + hq \le x$, and the pairs with $n > n'$ are their mirror images. If $hq \ge x$ there is no $n \ge 1$ with $n + hq \le x$. Hence

$$\sum_l Q_l^2 = Q(x) + 2 \sum_{1 \le h \le x/q} C_x(hq) = Q(x) + 2 \sum_{1 \le h \le x/q} D_x(hq) + 2\Sigma_A, \tag{4.3}$$

where $\Sigma_A = \sum_{1 \le h \le x/q} A(hq)(x - hq)$.

The expected pair counts. Only squarefree t contribute to $A(hq)$, because η vanishes elsewhere. For squarefree t and $h \in \mathbb{N}$ we have

$$t^2 \mid hq \iff m_t(q) \mid h. \tag{4.4}$$

Indeed, with $\gamma = \gcd(t^2, q)$, $t^2 \mid hq$ holds if and only if $(t^2/\gamma) \mid h(q/\gamma)$, and since $(t^2/\gamma, q/\gamma) = 1$ this holds if and only if $t^2/\gamma = m_t(q)$ divides h. Both sums in Σ_A are finite ($h \le x/q$ and $t^2 \le hq \le x$), so they may be interchanged:

$$\Sigma_A = A_0 \sum_{t \ge 1} \eta(t) \sum_{\substack{1 \le h \le x/q \\ m_t(q) \mid h}} (x - hq).$$

Fix t and write $m = m_t(q)$, $M = qm$ and $X = x/M$. Put $h = jm$; then $x - hq = M(X - j)$ and $1 \le j \le \lfloor X \rfloor$. With $N = \lfloor X \rfloor$ and $\vartheta = \{X\}$,

$$\sum_{j=1}^{N} (X - j) = NX - \frac{N(N+1)}{2} = \frac{1}{2}(X - \vartheta)(X + \vartheta - 1) = \frac{1}{2}(X^2 - X + B(X)).$$

This also holds when $N = 0$: then $X < 1$, $B(X) = X - X^2$, and both sides vanish. Therefore

$$2\Sigma_A = A_0 \sum_{t \ge 1} \eta(t) \Big[\frac{x^2}{q\, m_t(q)} - x + q\, m_t(q)\, B\Big(\frac{x}{q\, m_t(q)} \Big) \Big]. \tag{4.5}$$

The bracket vanishes whenever $q\, m_t(q) > x$, since it is twice an empty sum. We now split the bracket into its three terms. Each of the three resulting series converges absolutely: $\eta \ge 0$ and $\sum_t \eta(t) < \infty$ by Lemma 2.4, $m_t(q) \ge 1$, and $0 \le q\, m\, B(x/(qm)) \le x$. So the splitting is legitimate. By Proposition 3.1(iv) and (v), and the definition (1.7) of MT,

$$2\Sigma_A = x^2 S_2(q) - \frac{6x}{\pi^2} + \mathrm{MT}(x; q). \tag{4.6}$$

Conclusion. Inserting (4.6) into (4.3), and then (4.3) into (4.2), the terms $x^2 S_2(q)$ cancel. This leaves exactly the stated identity. No estimate has been used. □

Remark 4.2. Proposition 4.1 says that the size of $V(x; q)$, about $\sqrt{xq}$, comes from the lattice-point discrepancy $B(X)$ in $\sum_{j \le X}(X - j) = \frac{1}{2}(X^2 - X) + \frac{1}{2}B(X)$. It does not come from any irregularity in the distribution of squarefree numbers. Heuristically, for t coprime to q the term of $\mathrm{MT}(x; q)$ is $A_0 q \kappa(t) B(x/(qt^2))$, and the sum of $B(x/(qt^2))$ over t is of order $\sqrt{x/q}$, so $\mathrm{MT}(x; q)$ has order $q\sqrt{x/q} = \sqrt{xq}$; see Section 6. In the computations of Section 10 the identity was also checked numerically: with every other term computed by brute force, it holds at $x = 3000$ for all $q \le 40$ to a relative accuracy of $4.8 \cdot 10^{-10}$ (`reproduce.py`, function `identity_selftest`).

5 Pairs of squarefree numbers with a given difference

The following estimate counts pairs of squarefree numbers with a given difference m, uniformly in $m \le x$. It is the bottleneck for the lower end of the range in Theorem 1.1. The exponent $\frac{2}{3}$ is the one familiar from the closely related additive problem: as reported in [1, §1], an elementary argument of Evelyn and Linfoot gives the number of representations of n as a sum of two k-free numbers with an error $O_{\varepsilon,k}(n^{2/(k+1)+\varepsilon})$, which is $O_\varepsilon(n^{2/3+\varepsilon})$ for squarefree numbers. We give a complete proof of the version needed here.

Proposition 5.1 (Uniform pair count)**.** *For every $\varepsilon > 0$ there is a constant $c_3(\varepsilon)$ such that*

$$|D_x(m)| \le c_3(\varepsilon)\, x^{2/3+\varepsilon}$$

for all real $x \ge 2$ and all integers $1 \le m \le x$, where $D_x(m)$ is defined in (4.1)*.*

Proof. Fix $x \geq 2$ and $1 \leq m \leq x$, and put $N = x - m \geq 0$ and $z = x^{1/3} \geq 1$. Then $C_x(m) = \sum_{1 \leq n \leq N} \mu^2(n)\mu^2(n+m)$, with n running over positive integers. Put $\tau^* = \max_{k \leq x} \tau(k)$. By Lemma 2.3, $\tau^* \leq C(\varepsilon)x^{\varepsilon}$, and τ^* also bounds $\tau(m)$.

Step 1: truncation. For $k \in \mathbb{N}$ let

$$S(k) = \sum_{d \leq z,\ d^2 \mid k} \mu(d), \qquad R(k) = \mu^2(k) - S(k) = \sum_{d > z,\ d^2 \mid k} \mu(d).$$

Then $|S(k)| \leq \tau(k)$ and $|R(k)| \leq \#\{d > z : \ d^2 \mid k\}$. With $\alpha = \mu^2(n)$ and $\beta = \mu^2(n+m)$ we have $\alpha\beta - S(n)S(n+m) = R(n)\beta + S(n)R(n+m)$, and $|\beta| \leq 1$. Hence

$$C_x(m) = \sum_{n \leq N} S(n)S(n+m) + \mathcal{E}_1, \qquad |\mathcal{E}_1| \leq \sum_{n \leq N} |R(n)| + \tau^* \sum_{n \leq N} |R(n+m)|.$$

For each $d > z$ the number of $n \leq N$ with $d^2 \mid n$ is at most $N/d^2 \leq x/d^2$. The number of $n \leq N$ with $d^2 \mid n+m$ is at most the number of multiples of d^2 up to x, that is, at most x/d^2. By Lemma 2.1, both sums are at most $x \sum_{d > z} d^{-2} \leq 2x/z = 2x^{2/3}$. So $|\mathcal{E}_1| \leq 2(1+\tau^*)x^{2/3}$.

Step 2: congruences. Expanding S,

$$\sum_{n \leq N} S(n)S(n+m) = \sum_{d_1, d_2 \leq z} \mu(d_1)\mu(d_2)\, \#\{1 \leq n \leq N : \ d_1^2 \mid n,\ d_2^2 \mid n+m\}.$$

The system $n \equiv 0\ (d_1^2)$, $n \equiv -m\ (d_2^2)$ is solvable if and only if $(d_1^2, d_2^2) = (d_1, d_2)^2$ divides m. In that case its solutions form one residue class modulo $L = [d_1^2, d_2^2]$, and Lemma 2.2 gives the count $N/L + \theta$. There are at most $\lfloor z \rfloor^2 \leq x^{2/3}$ pairs (d_1, d_2). So

$$\sum_{n \leq N} S(n)S(n+m) = N \sum_{\substack{d_1, d_2 \leq z \\ (d_1,d_2)^2 \mid m}} \frac{\mu(d_1)\mu(d_2)}{[d_1^2, d_2^2]} + \mathcal{E}_2, \qquad |\mathcal{E}_2| \leq x^{2/3}.$$

Step 3: completion. Let Σ be the same sum over all $d_1, d_2 \geq 1$ with $(d_1, d_2)^2 \mid m$. Write $d_1 = ea$, $d_2 = eb$ with $e = (d_1, d_2)$ and $(a, b) = 1$; then $[d_1^2, d_2^2] = e^2a^2b^2$, and the condition reads $e^2 \mid m$. By symmetry, the pairs with $\max(d_1, d_2) > z$ contribute in absolute value at most twice the contribution of the pairs with $d_1 > z$, that is, at most

$$2 \sum_{e^2 \mid m} \frac{1}{e^2} \sum_{a > z/e} \frac{1}{a^2} \sum_{b \geq 1} \frac{1}{b^2} \leq 2 \cdot \frac{\pi^2}{6} \sum_{e^2 \mid m} \frac{1}{e^2} \cdot \frac{2e}{z} = \frac{2\pi^2}{3z} \sum_{e^2 \mid m} \frac{1}{e} \leq \frac{7\,\tau(m)}{z},$$

by Lemma 2.1 with $w = z/e$. This also shows that Σ converges absolutely. Hence

$$N \sum_{\substack{d_1, d_2 \leq z \\ (d_1,d_2)^2 \mid m}} \frac{\mu(d_1)\mu(d_2)}{[d_1^2, d_2^2]} = N\Sigma + \mathcal{E}_3, \qquad |\mathcal{E}_3| \leq \frac{7x\tau^*}{z} = 7\tau^* x^{2/3}.$$

Step 4: evaluation of Σ. A nonzero term needs $d_1 = ea$ and $d_2 = eb$ squarefree. Then e, a, b are squarefree and pairwise coprime, and $\mu(ea)\mu(eb) = \mu(a)\mu(b)$. By absolute convergence and Lemma 2.5,

$$\Sigma = \sum_{\substack{e \text{ squarefree} \\ e^2 \mid m}} \frac{1}{e^2} \sum_{\substack{\mu^2(ab) = 1 \\ (ab, e) = 1}} \frac{\mu(a)\mu(b)}{a^2b^2} = A_0 \sum_{\substack{e \text{ squarefree} \\ e^2 \mid m}} \frac{1}{e^2} \prod_{p \mid e} \frac{p^2}{p^2 - 2} = A_0 \sum_{t^2 \mid m} \eta(t) = A(m).$$

Conclusion. Since $N = x - m$, we get $C_x(m) = A(m)(x - m) + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3$, so

$$|D_x(m)| \leq \bigl(2(1+\tau^*) + 1 + 7\tau^*\bigr)x^{2/3} \leq \bigl(3 + 9C(\varepsilon)\bigr)\, x^{2/3+\varepsilon}. \qquad \square$$

6 Sawtooth sums and the main term for one modulus

6.1 Sawtooth sums

The function $B(u) = \{u\}(1-\{u\})$ is continuous, 1-periodic, takes values in $[0, \frac14]$, and satisfies $B(u) = u - u^2$ for $0 \le u \le 1$. The sums we need are $\sum_{s\ge1} B(Z/s^2)$ and weighted versions of them. Their size $\asymp \sqrt{Z}$ has two sources: the terms with $s > \sqrt{Z}$, where $B(Z/s^2) = Z/s^2 - Z^2/s^4$, and the terms with $s \le \sqrt{Z}$, where B oscillates. The natural approximation is the integral $\int_0^\infty B(Z/u^2)\,du = I\sqrt{Z}$ (substitute $u = \sqrt{Z}\,w$). For $q \in \mathbb{N}$ put

$$\kappa_q(t) = \kappa(t)\mathbf{1}[(t,q)=1], \qquad \Delta_q = \prod_p \Big(1-\frac{1}{p}\Big)\Big(1+\mathbf{1}[p\nmid q]\frac{p}{p^2-2}\Big).$$

For $p \nmid q$ the factor of Δ_q is $(p-1)^2(p+2)/(p(p^2-2)) = 1 + O(p^{-2})$; there are finitely many other factors, so the product converges absolutely.

Proposition 6.1 (Sawtooth sums)**.**

(a) $I = \int_0^\infty B(u^{-2})\,du = \frac12\int_0^\infty B(v)v^{-3/2}\,dv = -4\zeta(-\frac12)$.

(b) *For every real $Z > 0$, $\big|\sum_{s\ge1} B(Z/s^2) - I\sqrt{Z}\big| \le c_1 Z^{1/3}$ with $c_1 = 60$.*

(c) *For every $\varepsilon > 0$ there is $c_5(\varepsilon)$ such that for all $q \in \mathbb{N}$ and real $Z > 0$,*

$$\Big|\sum_{t\ge1} \kappa_q(t) B(Z/t^2) - \Delta_q\, I\sqrt{Z}\Big| \le c_5(\varepsilon)\, q^{\varepsilon} Z^{1/3}.$$

Proof. (a) The substitution $v = u^{-2}$ gives the second expression. The integral converges, because $B(v) \le v$ near 0 and $B(v) \le \frac14$. Since B is continuous and piecewise C^1 with $B'(v) = 1 - 2\{v\}$, integration by parts on $[0, X]$ gives

$$\int_0^X B(v)v^{-3/2}\,dv = \big[-2B(v)v^{-1/2}\big]_0^X + 2\int_0^X B'(v)v^{-1/2}\,dv.$$

The boundary term tends to 0, because $B(v)v^{-1/2} \le v^{1/2}$ near 0 and $B(X)X^{-1/2} \le \frac14 X^{-1/2}$. Hence $\int_0^\infty Bv^{-3/2}\,dv = 2\lim_{N\to\infty}\int_0^N B'v^{-1/2}\,dv$ along integers N. On $(n, n+1)$ we have $B'(n^+) = 1$, $B'((n+1)^-) = -1$ and $B'' = -2$, so integrating by parts with the antiderivative $2v^{1/2}$ of $v^{-1/2}$,

$$\int_n^{n+1} B'(v)v^{-1/2}\,dv = -2(n+1)^{1/2} - 2n^{1/2} + 4\int_n^{n+1} v^{1/2}\,dv,$$

and summing over $0 \le n < N$,

$$\int_0^N B'(v)v^{-1/2}\,dv = -4\sum_{m=1}^{N} m^{1/2} + 2N^{1/2} + \tfrac83 N^{3/2}. \tag{6.1}$$

Now apply Lemma 2.6 on (N,∞) to $G(u) = u^{-s}$ with $\Re s > 1$, for which $\sum_{n>N} n^{-s}$ and $\int_N^\infty u^{-s}\,du = N^{1-s}/(s-1)$ converge absolutely, using $\psi(N) = -\frac12$ and $\psi_2(N) = \frac{1}{12}$:

$$\sum_{n>N} n^{-s} = \frac{N^{1-s}}{s-1} - \frac{N^{-s}}{2} + \frac{s}{12}N^{-s-1} - s(s+1)\int_N^\infty \psi_2(u)u^{-s-2}\,du.$$

The right-hand side is analytic in $\Re s > -1$, $s \ne 1$, because the integral converges locally uniformly there. By uniqueness of analytic continuation, $\zeta(s) = \sum_{n\le N} n^{-s}$ plus the right-hand side in that region. At $s = -\frac12$ we have $s(s+1) = -\frac14$ and $|\int_N^\infty \psi_2 u^{-3/2}\,du| \le \frac16 N^{-1/2}$, so

$$\sum_{m\le N} m^{1/2} = \zeta(-\tfrac12) + \tfrac23 N^{3/2} + \tfrac12 N^{1/2} + O(N^{-1/2}).$$

Inserting this into (6.1) shows that $\int_0^N B' v^{-1/2}\,dv \to -4\zeta(-\frac12)$. Hence $\int_0^\infty B v^{-3/2}\,dv = -8\zeta(-\frac12)$ and $I = -4\zeta(-\frac12)$. By the functional equation [4, Corollary 10.4] at $s=-\frac12$, $\zeta(-\frac12) = 2^{-1/2}\pi^{-3/2}\Gamma(\frac32)\sin(-\frac{\pi}{4})\zeta(\frac32) = -\zeta(\frac32)/(4\pi)$. So $I = \zeta(3/2)/\pi > 0$, the value stated in (1.6).

(b) *The case $Z<1$.* Here $0 \le B(Z/s^2) \le Z/s^2$, so $0 \le \sum_s B(Z/s^2) \le \zeta(2)Z < 2Z^{1/3}$. Also $0 < I\sqrt{Z} < Z^{1/3}$, since $I<1$. The difference of two numbers in $[0, 2Z^{1/3}]$ is at most $2Z^{1/3}$.

The case $Z \ge 1$. Put $\Lambda = Z^{1/3} \ge 1$ and $T = \lfloor \Lambda \rfloor \ge 1$. Then $\Lambda/2 < T \le \Lambda$, hence

$$\frac{Z}{T^3} < 8, \qquad K_0 := \frac{Z}{T^2} < 4\Lambda, \qquad \frac{Z^2}{T^5} < 32\Lambda. \tag{6.2}$$

Small s. Both $\sum_{s\le T} B(Z/s^2)$ and $\int_0^T B(Z/u^2)\,du$ lie in $[0, T/4] \subset [0, \Lambda/4]$, so their difference is at most $\Lambda/4$. It therefore suffices to show

$$\Big|\sum_{s>T} B(Z/s^2) - \int_T^\infty B(Z/u^2)\,du\Big| \le 57\Lambda, \tag{6.3}$$

which gives the claim with $57 + \frac14 < 60$.

A decomposition of B. For real $w \ge 0$ with $j = \lfloor w \rfloor$ we have $B(w) = (w-j)(1-w+j) = w - w^2 + 2jw - j - j^2$, and $\sum_{k=1}^{j} 2(w-k) = 2jw - j^2 - j$. Hence

$$B(w) = w - w^2 + 2\sum_{k\ge 1} (w-k)_+, \qquad (v)_+ = \max(v, 0).$$

Put $\Phi(u) = Z/u^2 - Z^2/u^4$ and $G_k(u) = (Z/u^2 - k)_+$, so that $B(Z/u^2) = \Phi(u) + 2\sum_{k\ge1} G_k(u)$. For $u \ge T$ only the $k < K_0$ contribute, and there are fewer than $K_0 < 4\Lambda$ of them. For such k, $G_k(u) = Z/u^2 - k$ on $[T, R_k]$ with $R_k = \sqrt{Z/k} > T$, and $G_k = 0$ on $[R_k, \infty)$.

Euler–Maclaurin. Apply Lemma 2.6 to Φ on (T, ∞) (where $\sum_{s>T}\Phi(s)$ and $\int_T^\infty \Phi$ converge absolutely, since $\Phi(u) = O(u^{-2})$), and to $Z/u^2 - k$ on $(T, R_k]$ for each $k < K_0$. Use $\psi(T) = -\frac12$, $\psi_2(T) = \frac{1}{12}$ and $G_k(R_k) = 0$, and add. Since $\int_T^\infty \Phi + 2\sum_k \int_T^{R_k} G_k = \int_T^\infty B(Z/u^2)\,du$ (all integrals converge absolutely), this gives

$$\sum_{s>T} B(Z/s^2) = \int_T^\infty B(Z/u^2)\,du + R_1 + R_2 + R_3 + R_4 + R_5,$$

where, writing G_k' and G_k'' for the derivatives of $Z/u^2 - k$:

- $R_1 = \psi(T)\big[\Phi(T) + 2\sum_{k<K_0} G_k(T)\big] = \psi(T) B(Z/T^2)$. The individual terms here have size about $Z^2/T^4 \asymp Z^{2/3}$, but their sum is $B(Z/T^2) \le \frac14$. So $|R_1| \le \frac18$. This is the one place where cancellation is essential.

- $R_2 = -\frac{1}{12}\big[\Phi'(T) + 2\sum_{k<K_0} G_k'(T)\big]$, with $\Phi'(T) = -2Z/T^3 + 4Z^2/T^5$ and $G_k'(T) = -2Z/T^3$. By (6.2), $|\Phi'(T)| < 16 + 128\Lambda$ and $|2\sum_k G_k'(T)| \le 4K_0 Z/T^3 < 128\Lambda$. So $|R_2| \le \frac{1}{12}(16 + 256\Lambda) \le \frac{272}{12}\Lambda$.

- $R_3 = -\int_T^\infty \psi_2 \Phi''\,du$ with $\Phi'' = 6Z/u^4 - 20Z^2/u^6$. So $|R_3| \le \frac{1}{12}(2Z/T^3 + 4Z^2/T^5) < \frac{1}{12}(16 + 128\Lambda) \le 12\Lambda$.

- $R_4 = -2\sum_{k<K_0} \int_T^{R_k} \psi_2 G_k''\,du$ with $G_k'' = 6Z/u^4$ and $\int_T^\infty 6Z u^{-4}\,du = 2Z/T^3 < 16$. So $|R_4| \le \frac{2}{12} K_0 \cdot 16 < \frac{32}{3}\Lambda$.

- $R_5 = 2\sum_{k<K_0} \psi_2(R_k) G_k'(R_k)$ with $G_k'(R_k) = -2Z/R_k^3 = -2k^{3/2}/\sqrt{Z}$. Since $k < 4\Lambda$, $\sum_{k<K_0} k^{3/2} < 4\Lambda \cdot 8\Lambda^{3/2}$, and with $\sqrt{Z} = \Lambda^{3/2}$ we get $|R_5| \le \frac{2}{12}\cdot 2 \cdot 32\Lambda = \frac{32}{3}\Lambda$. These are the kinks of the G_k at the points R_k. Summing over the $K_0 \asymp Z^{1/3}$ values of k is what produces the exponent $\frac13$.

Using $\Lambda \ge 1$, the total is at most $(\frac{1}{8}+\frac{272}{12}+12+\frac{32}{3}+\frac{32}{3})\Lambda = \frac{449}{8}\Lambda < 57\Lambda$, which proves (6.3).

(c) Fix q. The function κ_q is multiplicative with $\kappa_q(p) = p^2/(p^2-2)$ for $p \nmid q$, $\kappa_q(p) = 0$ for $p \mid q$ and $\kappa_q(p^j) = 0$ for $j \ge 2$. Let $\beta = \mu * \kappa_q$, so that $\kappa_q = 1 * \beta$. Then β is multiplicative, with $\beta(p^j) = \kappa_q(p^j) - \kappa_q(p^{j-1})$:

- for $p \mid q$: $\beta(p) = -1$ and $\beta(p^j) = 0$ for $j \ge 2$;
- for $p \nmid q$: $\beta(p) = \frac{2}{p^2-2}$, $\beta(p^2) = -\frac{p^2}{p^2-2}$ and $\beta(p^j) = 0$ for $j \ge 3$.

Absolute convergence. For $p \nmid q$, $\sum_j |\beta(p^j)| p^{-2j/3} \le \frac{2p^{-2/3}}{p^2-2} + 2p^{-4/3}$, because $p^2/(p^2-2) \le 2$. Hence

$$\Sigma_{2/3} := \sum_{n\ge1} \frac{|\beta(n)|}{n^{2/3}} \le \prod_{p\mid q}(1+p^{-2/3}) \prod_{p} \Big(1 + \frac{2p^{-2/3}}{p^2-2} + 2p^{-4/3}\Big) \le C_\beta\, 2^{\omega(q)} \le C_\beta C(\varepsilon) q^{\varepsilon},$$

with an absolute constant C_β, by Lemma 2.3. Similarly $\sum_n |\beta(n)|/n < \infty$ for each fixed q.

Rearrangement. For $W > 0$, $\sum_m B(W/m^2) \le \sum_m \min(\frac14, W/m^2) \le 3\sqrt{W}$. Indeed, for $W < 1$ the sum is at most $\zeta(2)W \le 2\sqrt{W}$, and for $W \ge 1$ it is at most $\frac14\sqrt{W} + W \cdot 2/\sqrt{W}$ by Lemma 2.1. Hence $\sum_{n,m} |\beta(n)| B(Z/(nm)^2) \le 3\sqrt{Z}\sum_n |\beta(n)|/n < \infty$, and we may rearrange:

$$\sum_{t\ge1} \kappa_q(t) B(Z/t^2) = \sum_{n,m\ge1} \beta(n) B\Big(\frac{Z/n^2}{m^2}\Big) = \sum_{n\ge1} \beta(n) \sum_{m\ge1} B\Big(\frac{Z/n^2}{m^2}\Big).$$

Main term. Part (b) holds for every positive argument. Applying it with Z/n^2 in place of Z, and using $I\sqrt{Z/n^2} = I\sqrt{Z}/n$,

$$\Big|\sum_t \kappa_q(t) B(Z/t^2) - I\sqrt{Z}\sum_n \frac{\beta(n)}{n}\Big| \le c_1 \sum_n |\beta(n)| \Big(\frac{Z}{n^2}\Big)^{1/3} = c_1 \Sigma_{2/3} Z^{1/3}.$$

The constant. By (2.1), $\sum_n \beta(n)/n = \prod_p (1 + \beta(p)/p + \beta(p^2)/p^2)$. For $p \mid q$ the factor is $1 - 1/p$. For $p \nmid q$ it is

$$1 + \frac{2}{p(p^2-2)} - \frac{1}{p^2-2} = \frac{p^3-3p+2}{p(p^2-2)} = \frac{(p-1)^2(p+2)}{p(p^2-2)} = \Big(1-\frac1p\Big)\Big(1+\frac{p}{p^2-2}\Big).$$

Hence $\sum_n \beta(n)/n = \Delta_q$, and (c) holds with $c_5(\varepsilon) = c_1 C_\beta C(\varepsilon)$. □

6.2 The main term for one modulus

Proposition 6.2 (Asymptotics of MT)**.** *For every $\varepsilon > 0$ there is $c_6(\varepsilon)$ such that for all real $x > 0$ and all $q \in \mathbb{N}$,*

$$\big|\mathrm{MT}(x;q) - h(q)\sqrt{xq}\big| \le c_6(\varepsilon)\, x^{1/3} q^{2/3+\varepsilon},$$

with h as in (1.10)*. In particular this holds for $x \ge 1$ and $1 \le q \le x$.*

Proof. Factorising t. Fix q. Every squarefree t factors uniquely as $t = t_0 t_1 t_2$, where t_0 is composed of primes $p \nmid q$, t_1 of primes $p \,\|\, q$, and t_2 of primes p with $p^2 \mid q$. Conversely, every such triple of squarefree numbers gives a squarefree t. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be the finite sets of admissible t_1 and t_2 (all of them divide q). For squarefree t, $\gcd(t^2, q) = \prod_{p\mid t} p^{\min(2, v_p(q))} = t_1 t_2^2$, so

$$m_t(q) = t_0^2 t_1, \qquad \frac{x}{q\, m_t(q)} = \frac{Z_1}{t_0^2} \quad \text{with} \quad Z_1 = \frac{x}{q\, t_1}, \qquad \eta(t) m_t(q) = \eta(t_1)\eta(t_2)\, t_1\, \kappa(t_0).$$

By the absolute convergence in Lemma 2.4 we may regroup:

$$\mathrm{MT}(x;q) = A_0\, q \sum_{t_1\in\mathcal{S}_1} \sum_{t_2\in\mathcal{S}_2} \eta(t_1)\eta(t_2)\, t_1\, \Sigma(t_1), \qquad \Sigma(t_1) = \sum_{t_0\ge1} \kappa_q(t_0) B\Big(\frac{Z_1}{t_0^2}\Big).$$

Applying Proposition 6.1(c). With $Z = Z_1 > 0$ we get $\Sigma(t_1) = \Delta_q I\sqrt{Z_1} + \theta\, c_5(\varepsilon) q^{\varepsilon} Z_1^{1/3}$. Since $t_1\sqrt{Z_1} = \sqrt{x/q}\, t_1^{1/2}$ and $t_1 Z_1^{1/3} = (x/q)^{1/3} t_1^{2/3}$,

$$\mathrm{MT}(x;q) = A_0 \Delta_q I \sqrt{xq} \sum_{t_1 \in \mathcal{S}_1} \eta(t_1) t_1^{1/2} \sum_{t_2 \in \mathcal{S}_2} \eta(t_2) + R,$$

$$|R| \le c_5(\varepsilon)\, q^{1+\varepsilon} \Big(\frac{x}{q}\Big)^{1/3} \sum_{t_1 \in \mathcal{S}_1} \eta(t_1) t_1^{2/3} \sum_{t_2 \in \mathcal{S}_2} \eta(t_2),$$

where we used $A_0 \le 1$. By (2.2), $\sum_{t_1} \eta(t_1) t_1^{2/3} \le \prod_p (1 + 2p^{-4/3})$ and $\sum_{t_2} \eta(t_2) \le \prod_p (1 + 2p^{-2})$, both absolute constants. So $|R| \le c_6(\varepsilon) x^{1/3} q^{2/3+\varepsilon}$.

The constant is $h(q)$. The two finite sums are Euler products over the primes dividing q:

$$\sum_{t_1 \in \mathcal{S}_1} \eta(t_1) t_1^{1/2} = \prod_{p \| q} \Big(1 + \frac{\sqrt{p}}{p^2 - 2}\Big), \qquad \sum_{t_2 \in \mathcal{S}_2} \eta(t_2) = \prod_{p^2 \mid q} \Big(1 + \frac{1}{p^2 - 2}\Big).$$

Multiplying the absolutely convergent products A_0 and Δ_q factor by factor, the local factor at p of the main coefficient is

$$\begin{aligned}
p \nmid q: &\quad \Big(1 - \tfrac{2}{p^2}\Big)\Big(1 - \tfrac{1}{p}\Big)\Big(1 + \tfrac{p}{p^2-2}\Big) = \frac{(p-1)(p^2+p-2)}{p^3} = \frac{(p-1)^2(p+2)}{p^3} = a_p, \\
p \| q: &\quad \Big(1 - \tfrac{2}{p^2}\Big)\Big(1 - \tfrac{1}{p}\Big)\Big(1 + \tfrac{\sqrt{p}}{p^2-2}\Big) = \frac{(p-1)(p^2-2+\sqrt{p})}{p^3} = b_p, \\
p^2 \mid q: &\quad \Big(1 - \tfrac{2}{p^2}\Big)\Big(1 - \tfrac{1}{p}\Big)\Big(1 + \tfrac{1}{p^2-2}\Big) = \frac{(p-1)(p^2-1)}{p^3} = c_p.
\end{aligned}$$

Together with the factor I, this is exactly $h(q)$. □

Remark 6.3. Since $b_p/a_p \ge 1 - 1/p$ and $c_p/a_p = (p+1)/(p+2) \ge 1 - 1/p$, we have $h(q) \ge I A_h\, \varphi(q)/q \gg_\varepsilon q^{-\varepsilon}$, where $A_h = \prod_p a_p > 0$ (Proposition 8.1(0)). So Proposition 6.2 gives, for $x \ge 1$ and $1 \le q \le x$,

$$\mathrm{MT}(x;q) = h(q)\sqrt{xq}\,\big(1 + O_\varepsilon(x^{2\varepsilon}(q/x)^{1/6})\big),$$

because the relative error is $\ll_\varepsilon q^{2\varepsilon}(q/x)^{1/6}$ and $q^{2\varepsilon} \le x^{2\varepsilon}$ for $q \le x$. (Without the restriction $q \le x$ the display can fail. Since $0 \le q\, m\, B(x/(qm)) \le x$ and $\sum_t \eta(t) < \infty$ (Lemma 2.4), (1.7) gives $0 \le \mathrm{MT}(x;q) \ll x$ for all $x > 0$. Hence $\mathrm{MT}(x;1)/(h(1)\sqrt{x}) \to 0$ as $x \to 0$, whereas for $\varepsilon > 1/12$ the displayed error term $x^{2\varepsilon} x^{-1/6}$ tends to 0.) The relative error is not small when $q \asymp x$. This is why the asymptotic formula of Theorem 1.1(iii) requires $Y \le x^{1-\delta}$, whereas the error term in (ii), which keeps MT unexpanded, has no such restriction. The numerical data in Section 10 suggest that, after summation over the moduli, the secondary term has relative size roughly proportional to $\sqrt{Y/x}$. That is consistent with, but sharper than, the proved bound.

7 The variance for a single modulus

Proposition 7.1 (Single-modulus variance)**.** *For every $\varepsilon > 0$ there is $c_7(\varepsilon)$ such that for all real $x \ge 2$ and all integers $1 \le q \le x$:*

(i) $|V(x;q) - \mathrm{MT}(x;q)| \le c_7(\varepsilon)\big(x^{5/3+\varepsilon}/q + \sqrt{x}\big)$;

(ii) $|V(x;q) - h(q)\sqrt{xq}| \le c_7(\varepsilon)\big(x^{1/3} q^{2/3+\varepsilon} + x^{5/3+\varepsilon}/q + \sqrt{x}\big)$.

Proof. (i) By Proposition 4.1, which holds for all $x \ge 1$ and $q \in \mathbb{N}$,

$$V(x;q) - \mathrm{MT}(x;q) = 2 \sum_{1 \le h \le x/q} D_x(hq) + \Big(Q(x) - \frac{6x}{\pi^2}\Big) - 2x \sum_{l=1}^{q} g(q,l) E(x;q,l).$$

For $1 \le h \le x/q$ the integer $m = hq$ satisfies $1 \le m \le x$. Since $x \ge 2$, Proposition 5.1 applies to each of the at most x/q terms. (This is the only place where $x \ge 2$ rather than $x \ge 1$ is needed.) It gives

$$\Big|2 \sum_{1 \le h \le x/q} D_x(hq)\Big| \le 2 \cdot \frac{x}{q} \cdot c_3(\varepsilon) x^{2/3+\varepsilon} = 2c_3(\varepsilon) \frac{x^{5/3+\varepsilon}}{q}.$$

Proposition 3.3(i) gives $|Q(x) - 6x/\pi^2| \le c_0\sqrt{x}$. Proposition 3.3(ii) gives

$$\Big|2x \sum_{l} g(q,l) E(x;q,l)\Big| \le 2c_0 \frac{x^{3/2}}{q} \le 2c_0 \frac{x^{5/3+\varepsilon}}{q},$$

since $x \ge 1$. By the triangle inequality, (i) holds with $c_7(\varepsilon) \ge 2c_3(\varepsilon) + 2c_0$.

(ii) Proposition 6.2 gives $|\mathrm{MT}(x;q) - h(q)\sqrt{xq}| \le c_6(\varepsilon) x^{1/3} q^{2/3+\varepsilon}$. Combining this with (i) by the triangle inequality proves (ii) with

$$c_7(\varepsilon) = \max\{2c_3(\varepsilon) + 2c_0,\ c_6(\varepsilon)\},$$

and (i) remains true with this larger constant. □

8 Averages over polynomial moduli

In this section f is fixed and satisfies (1.1). All constants may depend on f, and on ε where indicated. Write $f(t) = \alpha t^d + \dots$, with leading coefficient $\alpha \ge 1$, and let P_0 be the finite set of primes dividing the content of f (the gcd of its coefficients). Put $x_p = \rho(p)/p$ and $z_p = \rho(p^2)/p^2$.

Proposition 8.1 (Polynomial averages)**.**

(0) *$\sigma_p > 0$ for every p, and $\prod_p \sigma_p$ converges absolutely. Hence $0 < C_f < \infty$.*

Moreover, for every $\varepsilon > 0$ there are constants $c_8(f)$ and $c_8(f,\varepsilon)$ such that for all real $y \ge 1$, with $Y = f(y)$:

(a) $\Big|\sum_{1 \le k \le y} F(k)\, h(f(k)) - C_f Y^{3/2}\Big| \le c_8(f,\varepsilon)\, Y^{3/2+\varepsilon} y^{-1/2}$;

(b) $\displaystyle\sum_{1 \le k \le y} \frac{f'(k)}{f(k)} \le c_8(f)\,(1 + \log Y)$;

(c) $\displaystyle\sum_{1 \le k \le y} f'(k) \le c_8(f)\, Y$.

The main difficulty in (a) is that f may have repeated factors or a nontrivial content. Then $\rho(p^2)$ may be as large as $p\rho(p)$, and $\rho(p) = p$ is possible for $p \in P_0$. The proof uses only the bounds of Lemma 8.3 and never assumes $\rho(p^2) = \rho(p)$.

8.1 Preliminaries

Lemma 8.2 (Local factors)**.** *For every prime p, $a_p, b_p, c_p > 0$, and*

$$b_p - a_p = -\frac{1}{p} + p^{-3/2} + p^{-2} - p^{-5/2}, \qquad c_p - b_p = -p^{-3/2}(1 - p^{-1/2})(1 - p^{-1}).$$

Consequently $a_p \ge \frac{1}{2}$, $0 < 1 - a_p \le 3/p^2$, $|b_p - a_p| < 1/p$ and $-p^{-3/2} < c_p - b_p < 0$.

Proof. Positivity is clear from (1.4). Expanding (1.4) gives $b_p = 1 - \frac{1}{p} - \frac{2}{p^2} + \frac{2}{p^3} + p^{-3/2} - p^{-5/2}$ and $c_p = 1 - \frac{1}{p} - \frac{1}{p^2} + \frac{1}{p^3}$, from which the two identities follow by subtraction. (They were also confirmed with computer algebra; see Section 10.) Next, $a_2 = \frac{1}{2}$, and $a_p \geq 1 - 3/p^2 \geq \frac{2}{3}$ for $p \geq 3$. The difference $1 - a_p = 3/p^2 - 2/p^3$ lies in $(0, 3/p^2]$. In the first identity $-1/p + p^{-3/2} \in (-1/p, 0)$ and $p^{-2} - p^{-5/2} \in [0, p^{-2})$, so $|b_p - a_p| < 1/p$. In the second, $(1 - p^{-1/2})(1 - p^{-1}) \in (0, 1)$. □

Lemma 8.3 (Root counts)**.** *The function ρ is multiplicative. For every prime p, $\rho(p^2) \leq p\,\rho(p)$ and $\rho(p^j) \leq p^j$. For $p \notin P_0$, $\rho(p) \leq d$. Hence $0 \leq z_p \leq x_p \leq 1$ for every p, and $x_p, z_p \leq d/p$ for $p \notin P_0$.*

Proof. Multiplicativity follows from the Chinese remainder theorem. Every root of f modulo p^2 reduces to a root modulo p, and each residue modulo p has exactly p lifts modulo p^2, so $\rho(p^2) \leq p\rho(p)$. For $p \notin P_0$ the reduction of f modulo p is a nonzero polynomial of degree at most d over the field $\mathbb{Z}/p\mathbb{Z}$, so it has at most d roots. □

Lemma 8.4 (Shape of F)**.** *There are constants $0 < \gamma_1 \leq 1 \leq \gamma_2$, depending only on f, such that for all real $t \geq 1$*

$$\gamma_1 t^d \leq f(t) \leq \gamma_2 t^d, \qquad \gamma_1 t^{d-1} \leq f'(t) \leq \gamma_2 t^{d-1}, \qquad \gamma_1 t^{3d/2-1} \leq F(t) \leq \gamma_2 t^{3d/2-1}. \tag{8.1}$$

Consequently, for $y \geq 1$ and $Y = f(y)$: $\sup_{[1,y]} F \leq \gamma_3 F(y)$ with $\gamma_3 = \gamma_2/\gamma_1$; $F(y) \leq \gamma_4 Y^{3/2}/y$ with $\gamma_4 = \gamma_2 \gamma_1^{-3/2}$; $y \leq Y/\gamma_1$; and $\int_1^y F(t)\,dt = \frac{2}{3}(Y^{3/2} - f(1)^{3/2})$. Finally, $[1, \infty)$ is the union of at most $2d - 1$ closed intervals, on each of which F is monotone.

Proof. On $[1, \infty)$ the functions $f(t)/t^d$, $f'(t)/t^{d-1}$ and $F(t)/t^{3d/2-1}$ are continuous and positive, by (1.1). They have the positive limits α, αd and $\alpha^{3/2} d$ as $t \to \infty$, so they are bounded above and below by positive constants. Since $3d/2 - 1 > 0$, $\sup_{[1,y]} F \leq \gamma_2 y^{3d/2-1} \leq \gamma_3 F(y)$. Also $Y^{3/2}/y \geq \gamma_1^{3/2} y^{3d/2-1} \geq \gamma_1^{3/2} \gamma_2^{-1} F(y)$, and $Y \geq \gamma_1 y^d \geq \gamma_1 y$. The integral follows from $(f^{3/2})' = \frac{3}{2} f' f^{1/2}$. Finally $F' = P/(2\sqrt{f})$ with $P = 2ff'' + f'^2$. This is a polynomial of degree $2d - 2$ with leading coefficient $\alpha^2 d(3d - 2) > 0$ (for $d = 1$, $P = \alpha^2$), so it has at most $2d - 2$ real zeros, and between consecutive zeros F' has constant sign. □

Lemma 8.5 (Monotone weights in a residue class)**.** *Let $J \subset \mathbb{R}$ be a bounded interval (open, closed or half-open) and $G \geq 0$ monotone on J. Let $e \in \mathbb{N}$ and $r \in \mathbb{Z}$, and let $k_1 < \cdots < k_n$ be the integers $k \in J$ with $k \equiv r \pmod{e}$. Then*

$$\Big| \sum_{j=1}^{n} G(k_j) - \frac{1}{e} \int_J G(t)\,dt \Big| \leq 3 \sup_J G.$$

Proof. If $n = 0$, then J has length at most e and $\frac{1}{e}\int_J G \leq \sup_J G$. Let $n \geq 1$ and suppose G is nonincreasing. Then $G(k_{j+1}) \leq \frac{1}{e}\int_{k_j}^{k_{j+1}} G \leq G(k_j)$ for $j < n$. Summing over j, $\sum_{j \geq 2} G(k_j) \leq \frac{1}{e}\int_{k_1}^{k_n} G \leq \sum_{j<n} G(k_j)$, so $\big|\sum_j G(k_j) - \frac{1}{e}\int_{k_1}^{k_n} G\big| \leq \sup_J G$. Since $k_1 - e \notin J$ and $k_n + e \notin J$, the parts of J to the left of k_1 and to the right of k_n have length at most e. Each contributes at most $\sup_J G$ to $\frac{1}{e}\int_J G$. The nondecreasing case is symmetric. □

The lemma has content only when $\sup_J G < \infty$. In its application below, $G = F$ is continuous on the compact interval $[1, y]$, so this holds.

8.2 Proof of Proposition 8.1(0), (b) and (c)

(0) By Lemma 8.3 the weights $1 - x_p$, $x_p - z_p$ and z_p in (1.5) are nonnegative and sum to 1. So σ_p is a convex combination of $a_p, b_p, c_p > 0$, and $\sigma_p \geq \min(a_p, b_p, c_p) > 0$. Also $\sigma_p = a_p + (b_p - a_p)x_p + (c_p - b_p)z_p$. For $p \notin P_0$, Lemmas 8.2 and 8.3 give

$$|\sigma_p - 1| \leq (1 - a_p) + \frac{1}{p} \cdot \frac{d}{p} + p^{-3/2} \cdot \frac{d}{p} \leq \frac{3 + 2d}{p^2}. \tag{8.2}$$

Hence $\sum_p |\sigma_p - 1| < \infty$. Since every $\sigma_p > 0$, the product $\prod_p \sigma_p$ converges absolutely to a positive finite number. As $I > 0$ (Proposition 6.1(a)), $0 < C_f < \infty$. In the same way $A_h := \prod_p a_p$ converges absolutely, and $0 < A_h < 1$.

(b), (c) By (8.1), for $1 \leq k \leq y$ we have $f'(k)/f(k) \leq \gamma_3/k$ and $f'(k) \leq \gamma_2 k^{d-1}$. Hence

$$\sum_{k \leq y} \frac{f'(k)}{f(k)} \leq \gamma_3(1 + \log y) \leq \gamma_3(1 + \log Y + \log(1/\gamma_1)) \leq c_8(f)(1 + \log Y),$$

using $y \leq Y/\gamma_1$. Also $\sum_{k \leq y} f'(k) \leq \gamma_2 y \cdot y^{d-1} \leq \gamma_3 Y$.

8.3 Proof of Proposition 8.1(a)

A divisor-sum expansion of h. By (1.10), and since $\prod_p a_p$ converges absolutely,

$$h(q) = I\, A_h\, h^\natural(q), \qquad h^\natural(q) = \prod_{p \| q} \frac{b_p}{a_p} \prod_{p^2 | q} \frac{c_p}{a_p}.$$

The function $h^\natural$ is multiplicative, with $h^\natural(p) = b_p/a_p$ and $h^\natural(p^j) = c_p/a_p$ for $j \geq 2$. Let $w = \mu * h^\natural$, so that $h^\natural(q) = \sum_{e|q} w(e)$. Then w is multiplicative with

$$w(p) = \frac{b_p - a_p}{a_p}, \qquad w(p^2) = \frac{c_p - b_p}{a_p}, \qquad w(p^j) = 0 \quad (j \geq 3).$$

By Lemma 8.2 and $a_p \geq \frac{1}{2}$,

$$|w(p)| \leq \frac{2}{p}, \qquad |w(p^2)| \leq 2p^{-3/2}. \tag{8.3}$$

Thus $w(e) = 0$ unless e is cube-free, and every cube-free e is uniquely $e = ab^2$ with a, b squarefree and coprime. Since $2/p \leq 4p^{-3/4}$ and $2p^{-3/2} = 2(p^2)^{-3/4}$,

$$|w(e)| \leq 4^{\omega(e)} e^{-3/4}, \qquad |w(ab^2)| \leq 2^{\omega(ab)} a^{-1} b^{-3/2}. \tag{8.4}$$

For cube-free $e = ab^2$, Lemma 8.3 gives $\rho(p) \leq d$ for $p \mid a$, $p \notin P_0$; $\rho(p^2) \leq pd$ for $p \mid b$, $p \notin P_0$; and $\rho(p^{v_p(e)}) \leq p^2$ for $p \in P_0$. Hence $\rho(e) \leq C_P d^{\omega(e)} b$ with $C_P = \prod_{p \in P_0} p^2$, and

$$|w(e)|\rho(e) \leq C_P (2d)^{\omega(e)} a^{-1} b^{-1/2}, \qquad \frac{|w(e)|\rho(e)}{e} \leq C_P (2d)^{\omega(e)} a^{-2} b^{-5/2}. \tag{8.5}$$

Three estimates. Let $W \geq 1$.

(i) *Tail.* For $e > W$ we have $1 \leq (e/W)^{1/2}$ and $e^{1/2} = a^{1/2} b$. By (8.5),

$$\sum_{e > W} \frac{|w(e)|\rho(e)}{e} \leq W^{-1/2} C_P \sum_{a,b \text{ squarefree}} \frac{(2d)^{\omega(a)+\omega(b)}}{a^{3/2} b^{3/2}} = c_4 W^{-1/2}, \tag{8.6}$$

with $c_4 = C_P \prod_p (1 + 2dp^{-3/2})^2$. In particular $\sum_e |w(e)|\rho(e)/e < \infty$.

(ii) *Short sum.* We have $(2d)^{\omega(n)} \le \tau(n)^{\log_2(2d)} \le C(f,\varepsilon) n^{\varepsilon}$ by Lemma 2.3, where the constant depends on d and ε. By (8.5), using $\sum_{b\le B} b^{-1/2} \le 2\sqrt{B}$,

$$\sum_{e\le W} |w(e)|\rho(e) \le C_P C(f,\varepsilon) W^{\varepsilon} \sum_{a\le W} \frac{1}{a} \sum_{b\le \sqrt{W}} b^{-1/2} \le 2C_P C(f,\varepsilon) W^{1/4+\varepsilon}(1+\log W). \tag{8.7}$$

(iii) *Euler product.* The function $e \mapsto w(e)\rho(e)/e$ is multiplicative, and its series converges absolutely by (i). By (2.1) and the second expression for σ_p in the proof of (0),

$$\sum_{e\ge 1} \frac{w(e)\rho(e)}{e} = \prod_p (1 + w(p)x_p + w(p^2)z_p) = \prod_p \frac{a_p + (b_p - a_p)x_p + (c_p - b_p)z_p}{a_p} = \frac{1}{A_h} \prod_p \sigma_p. \tag{8.8}$$

Truncation. Let $y \ge 1$ and $S = \sum_{k\le y} F(k) h^\natural(f(k))$. For $1 \le k \le y$ write $h^\natural(f(k)) = \sum_{e\mid f(k),\, e\le y} w(e) + T_k$ with $T_k = \sum_{e\mid f(k),\, e>y} w(e)$. By (8.4) and Lemma 2.3, and since $2 \le f(k) \le Y$ (as f is increasing on $[1,\infty)$),

$$|T_k| \le y^{-3/4} \sum_{e\mid f(k)} 4^{\omega(e)} \le y^{-3/4} \tau(f(k))^3 \le C(\varepsilon)\, y^{-3/4} Y^{\varepsilon}.$$

By Lemma 8.4, $\sum_{k\le y} F(k) \le y\gamma_3 F(y) \le \gamma_3\gamma_4 Y^{3/2}$, so

$$\Big|\sum_{k\le y} F(k) T_k\Big| \le C(\varepsilon)\gamma_3\gamma_4\, Y^{3/2+\varepsilon} y^{-3/4}. \tag{8.9}$$

Counting in residue classes. For $e \in \mathbb{N}$ let $N(e) = \sum_{1\le k\le y,\ e\mid f(k)} F(k)$. This is a sum over the $\rho(e)$ residue classes $r \bmod e$ with $f(r) \equiv 0 \pmod e$. Intersect the at most $2d-1$ closed intervals of Lemma 8.4 with $[1,y]$, and make them disjoint by removing from each its left endpoint whenever that point already belongs to the previous interval. This gives at most $2d-1$ disjoint intervals, closed or half-open, which cover $[1,y]$ and on each of which F is monotone. Applying Lemma 8.5 on each of them, and summing,

$$\Big|\sum_{\substack{1\le k\le y\\ k\equiv r\,(e)}} F(k) - \frac{1}{e}\int_1^y F(t)\,dt\Big| \le 3(2d-1) \sup_{[1,y]} F \le 6d\gamma_3 F(y).$$

(Stating Lemma 8.5 for arbitrary intervals covers the closed first interval $[1,\beta]$ directly; no value of F to the left of 1 is needed.) Summing over the $\rho(e)$ classes,

$$N(e) = \frac{\rho(e)}{e} \cdot \frac{2}{3}(Y^{3/2} - f(1)^{3/2}) + \theta_e\, 6d\gamma_3\, \rho(e) F(y). \tag{8.10}$$

Assembly. By the truncation, $S = \sum_{e\le y} w(e)N(e) + \sum_{k\le y} F(k)T_k$. By (8.6) and (8.8) with $W = y$, $\sum_{e\le y} w(e)\rho(e)/e = A_h^{-1}\prod_p \sigma_p + \theta c_4 y^{-1/2}$. Inserting (8.10) and using (8.7), (8.9) and $F(y) \le \gamma_4 Y^{3/2}/y$,

$$\begin{aligned} S &= \frac{2}{3}(Y^{3/2} - f(1)^{3/2})\Big(\frac{1}{A_h}\prod_p \sigma_p + \theta c_4 y^{-1/2}\Big) \\ &\quad + \theta\, 6d\gamma_3 F(y) \sum_{e\le y} |w(e)|\rho(e) + \theta C(\varepsilon)\gamma_3\gamma_4 Y^{3/2+\varepsilon} y^{-3/4} \\ &= \frac{2}{3A_h}\Big(\prod_p \sigma_p\Big) Y^{3/2} + O_{f,\varepsilon}\Big(1 + Y^{3/2}y^{-1/2} + Y^{3/2}y^{-3/4+\varepsilon}(1+\log y) + Y^{3/2+\varepsilon}y^{-3/4}\Big). \end{aligned}$$

The term $O_f(1)$ collects the contributions of $f(1)^{3/2}$. It is $\ll_f Y^{3/2}y^{-1/2}$, because $Y^{3/2}y^{-1/2} \ge \gamma_1^{3/2} y^{(3d-1)/2} \ge \gamma_1^{3/2}$. For $\varepsilon \le \frac{1}{8}$ we have $y^{-3/4+\varepsilon}(1+\log y) \ll_\varepsilon y^{-1/2}$, and since $Y \ge 2$ every error

term is then $\ll_{f,\varepsilon} Y^{3/2+\varepsilon} y^{-1/2}$. For $\varepsilon > \frac{1}{8}$ the claim follows from the case $\varepsilon = \frac{1}{8}$. Multiplying by IA_h gives

$$\sum_{k\le y} F(k)h(f(k)) = IA_h\, S = \frac{2}{3} I\Big(\prod_p \sigma_p\Big) Y^{3/2} + O_{f,\varepsilon}(Y^{3/2+\varepsilon} y^{-1/2}) = C_f Y^{3/2} + O_{f,\varepsilon}(Y^{3/2+\varepsilon} y^{-1/2}).$$

This proves (a). All implied constants depend only on f and ε, through d, $\gamma_1, \dots, \gamma_4$, C_P, $f(1)$ and the divisor-bound constants. □

9 Proof of the main theorem

Fix f as in Theorem 1.1. The finiteness and positivity of C_f are Proposition 8.1(0). Let $x \ge 2$ and $y \ge 1$ be real with $Y = f(y) \le x$, and fix $\varepsilon > 0$.

Range of the moduli. The hypothesis $f'(t) \ge 1$ for all *real* $t \ge 1$ makes f increasing on $[1, \infty)$. Hence every integer $1 \le k \le y$ satisfies

$$2 \le f(k) \le f(y) = Y \le x. \tag{9.1}$$

So Proposition 7.1 applies to $q = f(k)$ with the same x, for every $k \le y$, with a constant independent of k. By (1.2), $V_f(x,y) = \sum_{1\le k\le y} f'(k) V(x; f(k))$, and $f'(k) > 0$.

Elementary inequalities. Write $f(t) = \sum_{i\le d} \alpha_i t^i$ and $\gamma_0 = \sum_i |\alpha_i|$. For $y \ge 1$, $Y \le \gamma_0 y^d$, hence

$$y^{-1/2} \le \gamma_0^{1/(2d)}\, Y^{-1/(2d)}. \tag{9.2}$$

Since $Y \le x$, we have $Y^\varepsilon \le x^\varepsilon$ and $1 + \log Y \le 1 + \log x \le C'(\varepsilon) x^\varepsilon$. With Proposition 8.1(b),(c) this gives

$$\sum_{k\le y} \frac{f'(k)}{f(k)} \le c_8(f) C'(\varepsilon)\, x^\varepsilon, \qquad \sum_{k\le y} f'(k) \le c_8(f)\, Y. \tag{9.3}$$

9.1 Part (ii)

Apply Proposition 7.1(i) with $\varepsilon/2$ to each $q = f(k)$, multiply by $f'(k)$ and sum over $1 \le k \le y$:

$$\Big| V_f(x,y) - \sum_{k\le y} f'(k) \mathrm{MT}(x; f(k)) \Big| \le c_7(\tfrac{\varepsilon}{2}) \Big(x^{5/3+\varepsilon/2} \sum_{k\le y} \frac{f'(k)}{f(k)} + \sqrt{x} \sum_{k\le y} f'(k) \Big).$$

By (9.3) with $\varepsilon/2$, the right-hand side is at most

$$c_7(\tfrac{\varepsilon}{2})\, c_8(f) (C'(\tfrac{\varepsilon}{2})\, x^{5/3+\varepsilon} + x^{1/2} Y),$$

which is part (ii).

9.2 Part (i)

Apply Proposition 7.1(ii) with $\varepsilon/3$ to each $q = f(k)$ and use $f'(k)\sqrt{f(k)} = F(k)$:

$$\Big| V_f(x,y) - \sqrt{x} \sum_{k\le y} F(k) h(f(k)) \Big| \le c_7(\tfrac{\varepsilon}{3})\, (T_1 + T_2 + T_3),$$

where the three sums are bounded as follows.

- $T_1 = x^{1/3} \sum_{k\le y} f'(k) f(k)^{2/3+\varepsilon/3} \le x^{1/3} Y^{2/3+\varepsilon/3} \sum_{k\le y} f'(k) \le c_8(f)\, x^{1/3+\varepsilon/3} Y^{5/3}$, by $f(k) \le Y \le x$ and (9.3).
- $T_2 = x^{5/3+\varepsilon/3} \sum_{k\le y} f'(k)/f(k) \le c_8(f) C'(\tfrac{\varepsilon}{3})\, x^{5/3+2\varepsilon/3}$, by (9.3) with $\varepsilon/3$.

- $T_3 = \sqrt{x}\sum_{k\le y} f'(k) \le c_8(f)\, x^{1/2}Y \le c_8(f)\, x^{1/2}Y^{3/2-1/(2d)}$, because $Y \ge 2$ and $\frac{1}{2} - \frac{1}{2d} \ge 0$.

For the main term, Proposition 8.1(a) with $\varepsilon/3$ and (9.2) give

$$\Big|\sqrt{x}\sum_{k\le y} F(k)h(f(k)) - C_f x^{1/2}Y^{3/2}\Big| \le c_8(f,\tfrac{\varepsilon}{3})\gamma_0^{1/(2d)}\, x^{1/2}Y^{3/2+\varepsilon/3-1/(2d)}$$
$$\le c_8(f,\tfrac{\varepsilon}{3})\gamma_0^{1/(2d)}\, x^{1/2+\varepsilon/3}Y^{3/2-1/(2d)}.$$

By the triangle inequality,

$$|V_f(x,y) - C_f x^{1/2}Y^{3/2}| \le c(f,\varepsilon)\Big(x^{1/3+\varepsilon}Y^{5/3} + x^{5/3+\varepsilon} + x^{1/2+\varepsilon}Y^{3/2-1/(2d)}\Big),$$

with $c(f,\varepsilon) = c_7(\frac{\varepsilon}{3})c_8(f)(2 + C'(\frac{\varepsilon}{3})) + c_8(f,\frac{\varepsilon}{3})\gamma_0^{1/(2d)}$. This is part (i).

9.3 Part (iii)

Fix $\delta \in (0, 1/9)$. The interval $[x^{7/9+\delta}, x^{1-\delta}]$ is nonempty for $x > 1$ exactly because $7/9+\delta < 1-\delta$. Assume $x^{7/9+\delta} \le Y \le x^{1-\delta}$, put $\eta = \min(\delta/12, 1/(6d))$, and apply (i) with $\varepsilon = \min(\delta/12, 1/(9d))$. We divide each error term by $x^{1/2}Y^{3/2}$:

1. $x^{-1/6+\varepsilon}Y^{1/6} \le x^{-1/6+\varepsilon+(1-\delta)/6} = x^{\varepsilon-\delta/6} \le x^{-\delta/12} \le x^{-\eta}$, using $Y \le x^{1-\delta}$ and $\varepsilon \le \delta/12$;
2. $x^{7/6+\varepsilon}Y^{-3/2} \le x^{7/6+\varepsilon-\frac{3}{2}(7/9+\delta)} = x^{\varepsilon-3\delta/2} \le x^{-\delta/12} \le x^{-\eta}$, using $Y \ge x^{7/9+\delta}$ and $\frac{3}{2}\cdot\frac{7}{9} = \frac{7}{6}$;
3. $x^{\varepsilon}Y^{-1/(2d)} \le x^{\varepsilon-7/(18d)} \le x^{1/(9d)-7/(18d)} = x^{-5/(18d)} \le x^{-1/(6d)} \le x^{-\eta}$, using $Y \ge x^{7/9}$ and $\varepsilon \le 1/(9d)$.

Hence $|V_f(x,y) - C_f x^{1/2}Y^{3/2}| \le 3c(f,\varepsilon)\, x^{-\eta}x^{1/2}Y^{3/2}$, where ε depends only on δ and d. As $0 < C_f < \infty$, this is $V_f(x,y) = C_f x^{1/2}Y^{3/2}(1 + O_{f,\delta}(x^{-\eta}))$. □

Remark 9.1. The exponent η is conservative. The third ratio is in fact at most $x^{-5/(18d)}$, and the second at most $x^{-17\delta/12}$. We keep the exponent as it was stated in the archived proof.

10 Numerical evidence

This section compares Theorem 1.1 with exact computations of $V_f(x,y)$. Everything here is evidence and plays no part in the proofs. All tables and plots are regenerated by `manuscript/reproduce.py` (see `manuscript/README.md`). That script re-runs the accepted computation from its source code and recompiles its C++ part; it also computes the theoretical quantities. Numbers quoted in this section as bounds or ranges ("at most", "between", "lies in") are rounded outwards, so that the printed statement holds for the computed values; numbers quoted as approximate values ("about", "approximately", ≈) are rounded to the nearest digit.

10.1 What was computed

Grid. The three polynomials are $f(k) = k^2+1$, k^3+2 and $2k+1$, with weights $f'(k) = 2k$, $3k^2$ and 2. All three satisfy (1.1). The values of x are $x = 2.5\cdot 10^5 \cdot 2^j$ for $0 \le j \le 10$, so the largest is $x = 2.56\cdot 10^8$. For each x and each target ratio $r \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.4, 0.7, 1\}$, y is the largest integer with $f(y) \le rx$. At the two largest x the ratio $r = 0.005$ was added as a diagnostic. The quantity reported is the normalised variance

$$N(x,r) = \frac{V_f(x,y)}{x^{1/2}f(y)^{3/2}},$$

computed with the actual ratio $f(y)/x$. By Theorem 1.1(iii), $N(x,r) \to C_f$ when $x \to \infty$ and $f(y)/x \to 0$ suitably.

Method. The computation (archived numerical snapshot, files `compute.py` and `counts.cpp`) is exact in all integer quantities.

(1) The squarefree indicator $\mu^2(n)$, $n \le x$, is sieved. All autocorrelations

$$C(h) = \sum_{1 \le n \le x-h} \mu^2(n)\mu^2(n+h), \qquad 0 \le h \le x,$$

are obtained from one number-theoretic transform modulo the prime $p_0 = 3221225473$, with primitive root 5. Every $C(h)$ is at most $x < p_0$, so no information is lost.

(2) As in the proof of Proposition 4.1, $\sum_l Q(x;q,l)^2 = Q(x) + 2\sum_{m \ge 1} C(mq)$ for every modulus $q = f(k)$.

(3) The densities are $g(q,l) = \frac{6}{\pi^2}\tilde{g}(q,l)$ with an explicit rational $\tilde{g}(q,l)$ equivalent to Proposition 3.1(i),(iii). The coefficients $\sum_l Q(x;q,l)\tilde{g}(q,l)$ and $\sum_l \tilde{g}(q,l)^2$ are computed as exact rationals, using the counts of squarefree multiples of d as in the proof of Proposition 3.3.

(4) Finally $V(x;q) = \sum_l Q^2 - 2x\sum_l gQ + x^2\sum_l g^2$, as in (4.2), is accumulated with the weights $f'(k)$ in 70-digit decimal arithmetic.

The snapshot validates the implementation in several ways. It checks the transform against direct lag sums at every x. It compares $\tilde{g}$ with brute-force local densities for all $q \le 50$ and with empirical frequencies at $x = 10^6$. At $x = 257$ it compares every generated modulus with direct residue counts. During the original run, the operator and an integrator also recomputed some rows independently by direct counting (Appendix B).

Re-run. For this write-up the snapshot code was run again, unchanged, after its SHA-256 hashes had been checked against the acceptance record. The C++ part was compiled locally with `clang++ -O3`. All 270 rows (90 per polynomial, including the diagnostic rows) reproduce the archived integer counts exactly. The values of V_f agree to all stored digits (relative differences at most 0), and all fitted quantities are identical. The full grid needs about 5 GB of memory; the running times of the re-run that produced the tables are recorded in `reproduction.json` (fields `snapshot_rerun.seconds` and `snapshot_rerun.total_seconds_of_reproduce_py`).

10.2 The constants C_f

For the three polynomials the root counts are:

- k^2+1: $\rho(2) = 1$, $\rho(4) = 0$, and $\rho(p) = \rho(p^2) = 1 + \left(\frac{-1}{p}\right)$ for odd p;
- k^3+2: $\rho(2) = \rho(3) = 1$, $\rho(4) = \rho(9) = 0$; for $p \ge 5$, $\rho(p^2) = \rho(p)$, with $\rho(p) = 1$ if $p \equiv 2 \pmod 3$, and, if $p \equiv 1 \pmod 3$, $\rho(p) = 3$ or 0 according as $(-2)^{(p-1)/3} \equiv 1 \pmod p$ or not;
- $2k+1$: $\rho(2) = \rho(4) = 0$, and $\rho(p) = \rho(p^2) = 1$ for odd p.

For primes not dividing the discriminant all roots are simple, and each root modulo p lifts to exactly one root modulo p^2, which gives $\rho(p^2) = \rho(p)$. The formulas were checked by brute force modulo p and p^2 for all $p < 1000$. None of the three polynomials has a content prime, so $P_0 = \emptyset$, and (8.2) holds for every p with $K = 3 + 2d$. Hence, for $N \ge 30$,

$$\Big|\log \prod_{p>N} \sigma_p\Big| \le \frac{K}{1 - K/N^2} \sum_{p>N} \frac{1}{p^2} \le \frac{8K}{1 - K/N^2}\Big(\frac{1}{N^2} + \frac{1}{30N}\Big).$$

Here we used that every prime $p > N \ge 30$ lies in one of the 8 reduced residue classes modulo 30, and that each class contributes at most $N^{-2} + (30N)^{-1}$. Table 1 lists the partial values $\frac{2}{3} I \prod_{p \le N} \sigma_p$ and the resulting interval for C_f with $N = 10^7$. The partial values are rounded to nine decimals; the interval endpoints are rounded outwards, so that each printed interval contains the computed one. The intervals account rigorously for the omitted primes. They do not account for floating-point rounding in the partial products, which are computed in double

f	$N = 10^3$	$N = 10^4$	$N = 10^5$	$N = 10^6$	$N = 10^7$	interval for C_f
$k^2 + 1$	0.138152461	0.138088027	0.138083068	0.138082663	0.138082629	[0.138082603, 0.138082655]
$k^3 + 2$	0.135675596	0.135612820	0.135607954	0.135607555	0.135607522	[0.135607489, 0.135607555]
$2k + 1$	0.142685343	0.142618897	0.142613765	0.142613346	0.142613311	[0.142613291, 0.142613330]
t (formal)	0.128543562	0.128483702	0.128479079	0.128478702	0.128478670	[0.128478652, 0.128478687]
Vaughan c_2	0.128510919	0.128481180	0.128478873	0.128478684	0.128478668	

Table 1: Partial Euler products $\frac{2}{3} I \prod_{p \le N} \sigma_p$ and intervals for C_f. The row "t (formal)" uses $\rho \equiv 1$. The last row is Vaughan's constant (1.3) with C_1 truncated at $p \le N$; its limit equals that of the row above (Remark 1.3).

precision. As a check of that rounding (not a certified bound), the partial products for $N = 10^5$ were recomputed in 40-digit arithmetic; they agree with the double-precision values to a relative difference of at most $2.3 \cdot 10^{-16}$.

Apart from floating-point rounding, the intervals therefore fix the digits

$$C_{k^2+1} = 0.1380826\ldots, \qquad C_{k^3+2} = 0.135607\ldots, \qquad C_{2k+1} = 0.142613\ldots.$$

The partial products at $N = 10^7$ give the approximate values $C_{k^2+1} \approx 0.13808263$, $C_{k^3+2} \approx 0.13560752$ and $C_{2k+1} \approx 0.14261331$; the digits beyond those fixed above are only indicative. For $f(t) = t$ the formal value $\frac{2}{3} I \prod_{p \le 10^7} \sigma_p \approx 0.12847867$ agrees with Vaughan's $c_2 \approx 0.12847867$ (with C_1 truncated at the same point), as Remark 1.3 predicts. The local identity behind that remark, the identities of Lemma 8.2, and $\zeta(-\frac{1}{2}) = -\zeta(\frac{3}{2})/(4\pi)$ were also verified with computer algebra and 80-digit arithmetic (`reproduce.py`, function `symbolic_local_identity`).

10.3 The data and the predicted constant

Table 2 shows $N(x, r)$ at the largest x, together with the relative change from $x = 1.28 \cdot 10^8$. The full tables are in Appendix A. Two features are visible.

First, at fixed r the normalised values are stable in x. Between the two largest sizes they change by at most 0.91% for $k^3 + 2$, and by at most 0.04% for $k^2 + 1$ and $2k + 1$ (see also Fig. 2). Fits of $\log V_f$ against $\log x$ at fixed target ratio give $\alpha + \beta \approx 2$ (Table 3), consistent with the homogeneous main term $x^{1/2} Y^{3/2}$.

Second, $N(x, r)$ depends clearly on r. It rises from values between 0.140 and 0.145 at $r = 0.005$ to values between 0.166 and 0.184 at $r = 1$. A fit of $\log V_f$ against $\log f(y)$ at fixed x therefore gives exponents β between 1.538 and 1.552 instead of $\frac{3}{2}$ (Table 3). This does not contradict Theorem 1.1. Part (iii) concerns $f(y)/x \to 0$, and at a fixed ratio r the error term $x^{1/3+\varepsilon} Y^{5/3}$ of part (i) is as large as the main term. The ratio dependence is captured by the exact-shape statement (ii); see Section 10.4.

r	0.005	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$k^2 + 1$	0.14039	0.14086	0.14292	0.14446	0.15098	0.15655	0.16278	0.16993	0.17110
change (%)	−0.0233	−0.0363	+0.0051	−0.0102	−0.0118	−0.0043	−0.0083	−0.0003	−0.0030
$k^3 + 2$	0.14076	0.13999	0.14120	0.14391	0.14848	0.15234	0.15550	0.16399	0.16671
change (%)	+0.4079	−0.9050	−0.5298	−0.3279	−0.3250	−0.0824	−0.2807	−0.1486	−0.1688
$2k + 1$	0.14484	0.14460	0.14837	0.14934	0.15670	0.16171	0.16245	0.18239	0.18369
change (%)	+0.0003	−0.0000	+0.0002	+0.0005	+0.0002	+0.0001	+0.0001	−0.0000	−0.0000

Table 2: $N(x, r) = V_f(x, y)/(x^{1/2} f(y)^{3/2})$ at $x = 2.56 \cdot 10^8$ for the target ratios r (the column 0.005 is the diagnostic ratio). The "change" rows give the relative change, in percent, from $x = 1.28 \cdot 10^8$ to $x = 2.56 \cdot 10^8$.

Extrapolation to $r \to 0$. The snapshot estimates $\lim_{r \to 0} N$ empirically. At the largest x it fits the model $N(r) = C + a\sqrt{r} + br$ by unweighted least squares to the rows with r at or above

a "stable" cutoff. The cutoff is the smallest requested ratio such that every row at or above it changes by less than 1% between the two largest x; here it is 0.01 for all three polynomials. The reported uncertainty is a *heuristic sensitivity envelope*: the largest of twice the formal standard error of the intercept, the spread over neighbouring cutoffs, the shift to the previous x, the shifts under the alternative models $C + a\sqrt{r}$ and $C + a\sqrt{r} + br + cr^{3/2}$, and the shifts when the fitted window is cut at $r \le 0.1$, 0.2 or 0.4. It is not a confidence interval and not an error bound, and the model $C + a\sqrt{r} + br$ itself is not derived from the theorem. Table 3 compares the extrapolated constants with C_f. The relative differences are approximately -2.61%, $+0.53\%$ and -0.65% for k^2+1, k^3+2 and $2k+1$, all within the envelopes. Table 4 shows how strongly the extrapolated value depends on the model. For k^2+1 the alternatives range from about 0.134 to 0.142 and bracket C_f, so the difference of -2.61% reflects the choice of model rather than the constant. Figure 1 shows the data, the fitted curves and C_f.

f	$\alpha+\beta$ (fixed r)	β (fixed x)	$C_{\text{extr}} \pm$ envelope	C_f	$\frac{C_{\text{extr}}-C_f}{C_f}$	$\frac{C_{\text{extr}}-C_f}{\text{envelope}}$
k^2+1	2.0008 ± 0.0006	1.545 ± 0.004	0.1345 ± 0.0073	0.13808	-2.61%	-0.50
k^3+2	1.9978 ± 0.0021	1.538 ± 0.004	0.1363 ± 0.0024	0.13561	$+0.53\%$	$+0.30$
$2k+1$	2.0001 ± 0.0001	1.551 ± 0.007	0.1417 ± 0.0082	0.14261	-0.65%	-0.11

Table 3: Empirical exponents and extrapolated constants from the snapshot fits, compared with C_f from Table 1. "$\alpha+\beta$" is the mean of the eight fixed-ratio slopes of $\log V_f$ against $\log x$ (with the across-ratio standard error); "β" is the fixed-x slope against $\log f(y)$ at the largest x (with its regression standard error). The last column is $(C_{\text{extr}} - C_f)$ divided by the heuristic envelope; a value of absolute size at most 1 means that C_f lies within the envelope.

f	primary	$C + a\sqrt{r}$	$C + a\sqrt{r} + br + cr^{3/2}$	$r \le 0.1$	cutoff 0.005	previous x
k^2+1	0.13447	0.13837	0.13714	0.14176	0.13547	0.13451
k^3+2	0.13632	0.13759	0.13569	0.13856	0.13728	0.13763
$2k+1$	0.14169	0.14086	0.14145	0.14416	0.14188	0.14169

Table 4: Sensitivity of the extrapolated constant: the primary fit, the two alternative models, the window $r \le 0.1$, the lower cutoff 0.005, and the primary fit at the previous size $x = 1.28 \cdot 10^8$.

10.4 The exact-shape statement

Theorem 1.1(ii) replaces $C_f x^{1/2} Y^{3/2}$ by $\Sigma_{\text{MT}} = \sum_{k \le y} f'(k)\text{MT}(x; f(k))$. We evaluated Σ_{MT} for the rows listed in Table 5, together with $\sum_{k \le y} F(k)\overline{h}(f(k))$, the main term of Proposition 7.1(ii) summed over the moduli. We report

$$N_{\text{MT}} = \frac{\Sigma_{\text{MT}}}{x^{1/2}Y^{3/2}}, \qquad N_h = \frac{1}{Y^{3/2}} \sum_{k \le y} F(k)h(f(k)).$$

$\text{MT}(x;q)$ is computed from the factorisation $t = t_0 t_1 t_2$ of the proof of Proposition 6.2. For $t_0 \le \sqrt{Z_1}$ the terms are summed directly. For $t_0 > \sqrt{Z_1}$ we use $B(u) = u - u^2$ $(u < 1)$ and the closed forms

$$A_0 \sum_{(t,q)=1} \eta(t) = \frac{6}{\pi^2} \prod_{p \mid q} \frac{p^2-2}{p^2-1}, \qquad A_0 \sum_{(t,q)=1} \frac{\eta(t)}{t^2} = \frac{36}{\pi^4} \prod_{p \mid q} \frac{p^2(p^2-2)}{(p^2-1)^2},$$

which follow from (2.1) as in Lemma 2.4. The first 64 terms of the subtracted partial sums are evaluated in 80-digit arithmetic; A_0 and $\prod_p a_p$ are evaluated to high precision through the prime zeta function. The infinite series (1.7) is thus evaluated without truncation, up to floating-point

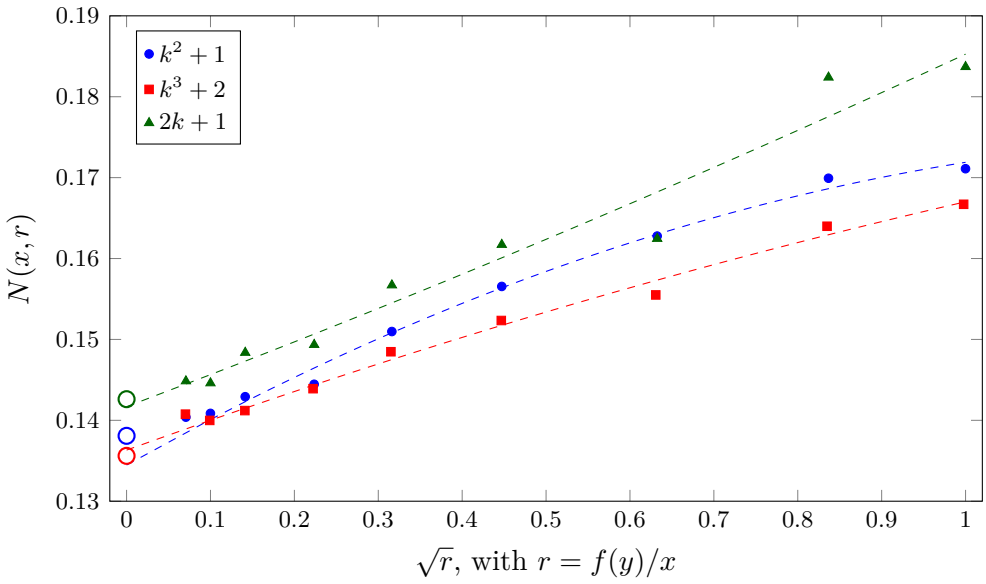


Figure 1: Normalised variance at $x = 2.56 \cdot 10^8$ against $\sqrt{r}$ (filled marks), the snapshot's empirical fits $C + a\sqrt{r} + br$ (dashed), and the proved constants C_f (open circles at $r = 0$).

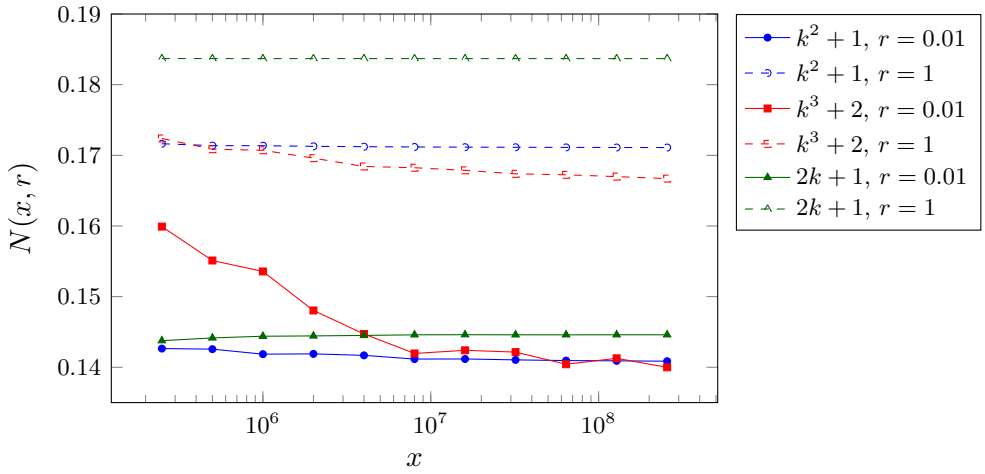


Figure 2: Stability in x: $N(x,r)$ against x for $r = 0.01$ and $r = 1$.

rounding. As an independent check, Proposition 4.1 was verified with this implementation of MT against brute-force values of every other term, for $x = 3000$ and all $q \le 40$ (relative discrepancy at most $4.8 \cdot 10^{-10}$). For $2k+1$ the number of moduli is about $x/2$, so only $x \le 10^6$ was treated.

f	x	r	y	N_{data}	N_{MT}	V_f/Σ_{MT}	N_h
k^2+1	10^6	0.01	99	0.14185	0.14171	1.00100	0.13966
		0.1	316	0.15191	0.15191	0.99999	0.13898
		1	999	0.17136	0.17136	1.00003	0.13825
	$1.6 \cdot 10^7$	0.01	399	0.14117	0.14118	0.99996	0.13853
		0.1	1264	0.15115	0.15115	1.00003	0.13825
		1	3999	0.17115	0.17115	1.00000	0.13812
	$2.56 \cdot 10^8$	0.005	1131	0.14039	0.14039	0.99998	0.13829
		0.01	1599	0.14086	0.14086	0.99999	0.13820
		0.1	5059	0.15098	0.15097	1.00001	0.13812
		1	15999	0.17110	0.17110	1.00000	0.13809
k^3+2	10^6	0.01	21	0.15356	0.15361	0.99968	0.15544
		0.1	46	0.15503	0.15499	1.00023	0.14231
		1	99	0.17071	0.17071	1.00003	0.13890
	$1.6 \cdot 10^7$	0.01	54	0.14240	0.14236	1.00024	0.14132
		0.1	116	0.14975	0.14974	1.00009	0.13844
		1	251	0.16787	0.16787	1.00000	0.13708
	$2.56 \cdot 10^8$	0.005	108	0.14076	0.14079	0.99982	0.13851
		0.01	136	0.13999	0.14000	0.99995	0.13776
		0.1	294	0.14848	0.14848	1.00000	0.13673
		1	634	0.16671	0.16671	1.00000	0.13602
$2k+1$	$2.5 \cdot 10^5$	0.01	1249	0.14376	0.14466	0.99381	0.14268
		0.1	12499	0.15670	0.15671	0.99992	0.14262
		1	124999	0.18370	0.18369	1.00002	0.14261
	10^6	0.01	4999	0.14439	0.14461	0.99848	0.14263
		0.1	49999	0.15670	0.15670	0.99997	0.14262
		1	499999	0.18369	0.18369	0.99999	0.14261

Table 5: Exact-shape comparison: the data $N_{\mathrm{data}} = N(x,r)$, the normalised sum N_{MT} of the exact main terms, the ratio V_f/Σ_{MT}, and the normalised h-average N_h. Rows with $r \in \{0.005, 0.01, 0.1, 1\}$ are shown; `reproduction.json` contains all 66 computed rows.

Over all 66 computed rows, V_f/Σ_{MT} lies in $[0.9938, 1.0011]$. At $x = 2.56 \cdot 10^8$ it differs from 1 by at most $1.8 \cdot 10^{-4}$. So the exact main term of Theorem 1.1(ii) accounts for the whole ratio dependence seen in Table 2. This agreement is much better than the proved error term $x^{5/3+\varepsilon} + x^{1/2}Y$. Relative to $x^{1/2}Y^{3/2}$, and ignoring the unknown constant and the factor x^{ε}, that error term is of size about $x^{-1/3}r^{-3/2}$. At $x = 2.56 \cdot 10^8$ this is about 1.6 for $r = 0.01$ and $1.6 \cdot 10^{-3}$ for $r = 1$. The pair-count errors $D_x(hq)$ evidently cancel substantially when summed over h, which the proof does not exploit. The quantity N_h approaches C_f as y grows, as Proposition 8.1(a) predicts. For k^3+2 the convergence is slower and less regular, because y is only a few hundred and the error is of order $y^{-1/2}$. The difference $N_{\mathrm{MT}} - N_h$ is the secondary term of the sawtooth sums. It is small for small r and grows with r. This is consistent with the relative bound $(Y/x)^{1/6}$ (up to x^{ε}) that Proposition 6.2 implies, but in the data the difference behaves roughly like a multiple of $\sqrt{r}$ (Fig. 3).

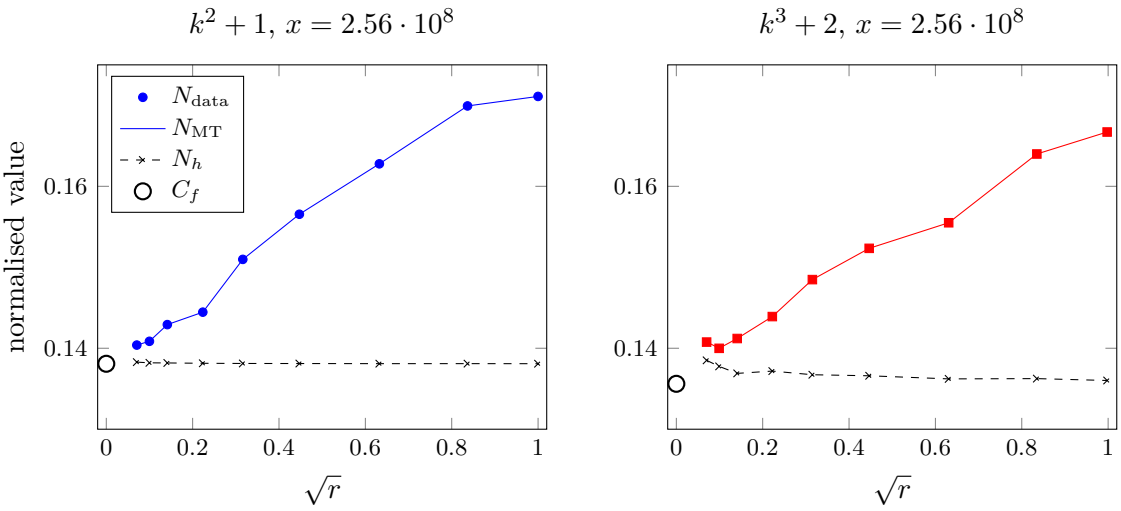


Figure 3: Data, exact main term N_{MT}, averaged asymptotic main term N_h and C_f at the largest x. Both panels use the legend of the left panel: filled marks are the data, the solid line is N_{MT}, the dashed line with crosses is N_h, and the open circle at $r=0$ is C_f.

10.5 Limitations

- The computations stop at $x=2.56\cdot10^8$. With the actual ratio $\varrho=f(y)/x$, a row lies in the range $x^{7/9+\delta}\le f(y)\le x^{1-\delta}$ of Theorem 1.1(iii) exactly when

$$\delta\le\delta_{\text{row}}:=\min\Big(\frac{\log\varrho}{\log x}+\frac{2}{9},\ -\frac{\log\varrho}{\log x}\Big).$$

  At $x=2.56\cdot10^8$ we have $x^{-2/9}\approx0.0135$. The rows with $r\le0.01$ have $\varrho\le r<x^{-2/9}$, so $\delta_{\text{row}}<0$ and they lie outside the range for every $\delta>0$. The rows with $r=1$ have ϱ close to 1 and $\delta_{\text{row}}\le2.4\cdot10^{-4}$, so they lie in the range only for very small δ; the implied constant in (iii) depends on δ and is not explicit. Over the whole grid at this x, δ_{row} is largest, about 0.103, at $r=0.1$. For each fixed $\delta>0$, the rows with a fixed target ratio r leave the range as x grows, since they need $x\le\varrho^{-1/\delta}$, where $\varrho\approx r$. The values of δ_{row} for all rows at the largest x are recorded in `reproduction.json`.

- The implied constants in Theorem 1.1 are not explicit, so no finite computation can confirm or refute it. The data can only be checked for consistency.

- The extrapolated constants depend on an empirical model, and their uncertainties are heuristic envelopes, not confidence intervals. The agreement in Table 3 is therefore qualitative evidence for the value of C_f, and the theorem does not rest on it.

- The values of C_f in Table 1 come with a rigorous bound for the omitted primes, but not with a certified bound for floating-point rounding; the 40-digit recomputation for $N=10^5$ is only a check.

- The exact-shape comparison uses double-precision arithmetic, with 80-digit evaluation only where cancellation occurs. Its agreement with the identity self-test is the main check of the implementation. For $2k+1$ it covers only $x\le10^6$.

References

[1] J. Brüdern, A. Granville, A. Perelli, R. C. Vaughan, and T. D. Wooley. On the exponential sum over k-free numbers. *Phil. Trans. R. Soc. Lond. A*, 356:739–761, 1998. No verified URL, DOI or arXiv identifier was available to this write-up; copy supplied with the project.

[2] J. Brüdern and T. D. Wooley. Sparse variance for primes in arithmetic progression. *Quart. J. Math.*, 62:289–305, 2011. No verified URL, DOI or arXiv identifier was available to this write-up; preprint copy supplied with the project.

[3] D. A. Goldston and R. C. Vaughan. On the Montgomery–Hooley asymptotic formula. In G. R. H. Greaves, G. Harman, and M. N. Huxley, editors, *Sieve Methods, Exponential Sums, and their Applications in Number Theory*, volume 237 of *London Math. Soc. Lecture Note Ser.*, pages 117–142. Cambridge University Press, 1997. doi:10.1017/CBO9780511526091.010.

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

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

[6] 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.

A Full tables of normalised values

The tables give $N(x,r) = V_f(x,y)/(x^{1/2}f(y)^{3/2})$ for all sizes and target ratios of the snapshot grid, regenerated by `reproduce.py`. The integer y is the largest with $f(y) \le rx$, and the normalisation uses the actual value $f(y)$.

x	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$2.5\cdot10^5$	0.142649	0.146052	0.146549	0.152224	0.157416	0.163777	0.170589	0.171642
$5\cdot10^5$	0.142565	0.143625	0.145049	0.151489	0.157419	0.163281	0.170311	0.171372
10^6	0.141855	0.143162	0.144886	0.151912	0.156975	0.163179	0.170250	0.171362
$2\cdot10^6$	0.141895	0.143860	0.145330	0.151304	0.156884	0.163051	0.170098	0.171269
$4\cdot10^6$	0.141686	0.143302	0.144687	0.151276	0.156752	0.162938	0.170032	0.171220
$8\cdot10^6$	0.141162	0.143155	0.144704	0.151199	0.156712	0.162890	0.170034	0.171183
$1.6\cdot10^7$	0.141175	0.143090	0.144622	0.151153	0.156632	0.162835	0.169986	0.171153
$3.2\cdot10^7$	0.141046	0.143021	0.144584	0.151056	0.156601	0.162812	0.169957	0.171141
$6.4\cdot10^7$	0.140949	0.142966	0.144505	0.151012	0.156578	0.162792	0.169935	0.171120
$1.28\cdot10^8$	0.140909	0.142912	0.144472	0.150993	0.156556	0.162789	0.169926	0.171106
$2.56\cdot10^8$	0.140858	0.142919	0.144457	0.150976	0.156550	0.162775	0.169925	0.171101

Table 6: $f(k) = k^2 + 1$.

x	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$2.5 \cdot 10^5$	0.159913	0.160232	0.160971	0.159910	0.163172	0.161392	0.169394	0.172387
$5 \cdot 10^5$	0.155112	0.162671	0.156355	0.161317	0.159404	0.159028	0.167785	0.170862
10^6	0.153562	0.149407	0.154182	0.155026	0.155087	0.158578	0.166278	0.170713
$2 \cdot 10^6$	0.148041	0.147909	0.151134	0.150024	0.154648	0.157691	0.166923	0.169599
$4 \cdot 10^6$	0.144716	0.147763	0.146250	0.151598	0.154780	0.157141	0.165214	0.168422
$8 \cdot 10^6$	0.141958	0.144833	0.145867	0.149564	0.154211	0.157510	0.165585	0.168245
$1.6 \cdot 10^7$	0.142399	0.144360	0.144416	0.149753	0.153584	0.156699	0.164844	0.167875
$3.2 \cdot 10^7$	0.142149	0.142778	0.144311	0.149048	0.153051	0.156096	0.164601	0.167379
$6.4 \cdot 10^7$	0.140427	0.142955	0.143925	0.149586	0.152600	0.155833	0.164525	0.167244
$1.28 \cdot 10^8$	0.141268	0.141954	0.144385	0.148961	0.152464	0.155936	0.164234	0.166987
$2.56 \cdot 10^8$	0.139990	0.141202	0.143912	0.148477	0.152339	0.155499	0.163990	0.166706

Table 7: $f(k) = k^3 + 2$.

x	0.01	0.02	0.05	0.1	0.2	0.4	0.7	1
$2.5 \cdot 10^5$	0.143761	0.148093	0.149307	0.156698	0.161734	0.162453	0.182396	0.183696
$5 \cdot 10^5$	0.144153	0.148232	0.149327	0.156723	0.161716	0.162465	0.182393	0.183694
10^6	0.144392	0.148302	0.149320	0.156700	0.161696	0.162452	0.182388	0.183690
$2 \cdot 10^6$	0.144437	0.148283	0.149305	0.156686	0.161701	0.162446	0.182389	0.183691
$4 \cdot 10^6$	0.144507	0.148327	0.149321	0.156694	0.161706	0.162449	0.182390	0.183693
$8 \cdot 10^6$	0.144593	0.148372	0.149349	0.156706	0.161713	0.162452	0.182390	0.183692
$1.6 \cdot 10^7$	0.144609	0.148375	0.149350	0.156706	0.161712	0.162452	0.182390	0.183691
$3.2 \cdot 10^7$	0.144592	0.148365	0.149341	0.156702	0.161709	0.162451	0.182390	0.183691
$6.4 \cdot 10^7$	0.144593	0.148366	0.149341	0.156702	0.161709	0.162451	0.182390	0.183692
$1.28 \cdot 10^8$	0.144598	0.148368	0.149342	0.156703	0.161709	0.162451	0.182390	0.183691
$2.56 \cdot 10^8$	0.144598	0.148368	0.149343	0.156703	0.161710	0.162451	0.182390	0.183691

Table 8: $f(k) = 2k + 1$.

B Provenance

The mathematics in Sections 2 to 9 is an edited presentation of an archived run of a multi-agent proof pipeline in which language models proposed, proved and refereed the results. This appendix records where the material comes from. A proposition-by-proposition map, with every editorial change and correction, is in `manuscript/changes.md`.

Archive.

- Source repository: `https://gitlab.com/mshepherd/squarefree-numbers`, archive label `zt-neutral`.
- Archive commit: `ed0aebaf461b602601856bf1b693fd76893d9aba`.
- Harness commit: `33aadc609359221ba5d81664438bce5f658f561c`.
- Revision recorded for all run stages: `8dfc53d01e56c9949c42061fdbe907a98a12a5d7`.

Models and roles (as recorded). The model identifiers are those recorded in the archive. The command-line tool versions are not model-weight versions, and model versions beyond these identifiers are unknown.

- Planning and probing: `claude-opus-5-5` and `codex:gpt-6.1-sol`; probing also `claude-sonnet-5-5`.
- Numerics: `claude-sonnet-5-5` and `codex:gpt-6.1-sol` produced candidate computations. The `codex:gpt-6.1-sol` snapshot was accepted.
- Proofs: `claude-sonnet-5-5` (prover).
- Refereeing: `claude-opus-5-5` and `codex:gpt-6.1-sol`. Each proposition was reviewed independently by both.
- Integration (assembly audit): `claude-opus-5-5` and `codex:gpt-6.1-sol`, independently.

Claude models ran under Claude Code 2.1.309 and Codex models under codex-cli 0.166.0.

Write-up stage. This text was written by `claude-opus-5-5` (Claude Code 2.1.311) from the archived run. It was refereed by `claude-opus-5-5` and `codex:gpt-6.1-sol` (the write-up stage records Claude Code 2.1.311 and codex-cli 0.166.0) in three rounds. Each round was followed by a revision, and the present version answers the third round. Some of the corrections described below were raised by these write-up referees.

Workflow. The problem statement restricts f to $\mathbb{Z}[t]$. It asks for the asymptotic behaviour of $V_f(x, y)$ in as large a range of $f(y) \leq x$ as can be justified, and it accepts a weaker range or error term provided it is correct. The weight $f'(k)$ and the class (1.1) of polynomials used here are those of the problem statement; neither was changed. The numerical candidates were computed before any plan existed, and one was accepted on computational grounds only, not because of its fitted constants. Eight proof plans were written, and the plan `simplest@claude-opus-5-5` was selected by the project owner. It splits the argument into nine propositions P1–P9. In this text, P1 is Proposition 3.1, P2 is Proposition 4.1, P3 is Proposition 5.1, P4 is Proposition 3.3, P5 is Proposition 6.1, P6 is Proposition 6.2, P7 is Proposition 7.1, P8 is Proposition 8.1, and P9 is Theorem 1.1 with its proof in Section 9. All nine proofs passed both referees in the first round, with only cosmetic remarks. Both integrators found that the propositions compose to the stated theorem without loss and without interface mismatches. The project owner's recorded decision reads: "Let's just assume the proof is correct and complete." A separate read-only operator audit found no substantive gap and left final mathematical acceptance to the project owner. The corrections made in this text are local and fall into three groups, listed individually in `changes.md`: (a) those recorded in the operator audit (the normalisation and convergence

wording in P1(iii), the endpoint $k = 1$ in P8, and unused numerical remarks in P6 and P8); (b) cosmetic points raised by the run's referees; and (c) points raised at the write-up stage. Group (c) contains one clarification of an archived auxiliary statement. For an infinite interval, the Euler–Maclaurin lemma in the archived P5 proof says that "the same holds" under the hypotheses $G, G' \to 0$, $G'' \in L^1$. Read as an identity between a convergent series and a convergent integral, this is false in that generality. Lemma 2.6 keeps exactly the archived hypotheses and states what letting $b \to \infty$ gives: the difference between the partial sums and the partial integrals converges, and the formula holds whenever the series or the integral converges. Both archived applications have absolutely convergent series and integrals, so no proposition and no step of the argument changes. No decision of the project owner on this point is recorded, and `changes.md` lists it for the owner's attention. The other items in group (c) concern only the exposition of this text, for example an arithmetic slip in an earlier version of the write-up, and do not touch the archived arguments. Agreement between models is evidence, not certification; the arguments are reproduced here in full so that they can be checked directly.

Independent checks during the run. Before acceptance, the operator compared both numerical candidates on overlapping rows and recomputed some rows by brute force. One integrator recomputed all 16 rows for $k^2 + 1$ and $k^3 + 2$ at $x = 250000$ by direct counting. The other computed $\sum f'(k)\mathrm{MT}(x; f(k))$ independently and compared it with the data. These checks are recorded in the archive. The computations in Section 10 were redone for this write-up as described there.

Literature. The run used the papers supplied with the problem and literature up to 2010, with the single exception of Brüdern–Wooley [2], which was supplied. The present text cites only works in the supplied corpus. See `README.md` for the limitations of the bibliography.