Distribution Agreement

Signature:

______________________________ __________

Robert J. Lemke Oliver Date

Topics in Analytic Number Theory

By

Robert J. Lemke Oliver
Doctor of Philosophy

Mathematics

Ken Ono
Advisor

David Borthwick
Committee Member

David Zureick-Brown
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Topics in Analytic Number Theory

By

Robert J. Lemke Oliver
M.A., University of Wisconsin-Madison, 2010

Advisor: Ken Ono, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Mathematics
2013

Abstract

Topics in Analytic Number Theory
By Robert J. Lemke Oliver

In this thesis, the author proves results using the circle method, sieve theory and the distribution of primes, character sums, modular forms and Maass forms, and the Granville-Soundararajan theory of *pretentiousness*. In particular, he proves theorems about partitions and q-series, *almost-prime* values of polynomials, Gauss sums, modular forms, quadratic forms, and multiplicative functions exhibiting extreme cancellation. This includes a proof of the Alder-Andrews conjecture, generalizations of theorems of Iwaniec and Ono and Soundararajan, and answers to questions of Zagier and Serre, as well as questions of the author in the Granville-Soundararajan theory of pretentiousness.

Topics in Analytic Number Theory

By

Robert J. Lemke Oliver
M.A., University of Wisconsin-Madison, 2010

Advisor: Ken Ono, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Mathematics
2013

Contents

Chapter 1

Introduction

Arguably the first work in analytic number theory was done by Euler, who considered the function $\zeta(x)$, $x \in \mathbb{R}$, defined via the series

$$\zeta(x) := \sum_{n=1}^{\infty} \frac{1}{n^x}.$$

While Euler's principal motivation may have been the resolution of the so-called *Basel problem* – the evaluation $\zeta(2) = \pi^2/6$ – from our own, modern perspective, probably the most far-reaching and important aspect of his work relates to the value $\zeta(1)$. Of course, this is the harmonic series which was known, even in Euler's time, to diverge to infinity. However, Euler observed that $\zeta(x)$ also possesses the product representation

$$\zeta(x) = \prod_p \left(1 - \frac{1}{p^x}\right)^{-1},$$

valid for any x such that the series converges (where, here and throughout, the index p runs only over primes); such a product representation for a series is known today as an *Euler product*. From this product, the behavior at $x = 1$ of $\zeta(x)$ implies something about the behavior of the primes: namely, that the series $\sum_p \frac{1}{p}$ diverges. Moreover, using more precise information about the divergence of the harmonic series (which Euler himself studied), it was possible to deduce the rate of divergence. Since the harmonic series diverges logarithmically – that is,

$$\sum_{n<X} \frac{1}{n} \sim \log X,$$

as X goes to infinity – and the sum of the reciprocals of the primes is almost the logarithm of $\zeta(1)$, Euler deduced that

$$\sum_{p<X} \frac{1}{p} \sim \log \log X. \tag{1.1}$$

Although Euler did not think in distributional terms, this was nonetheless essentially the first distributional statement about the primes post-Euclid: being prime may be rare, but it cannot be so rare that the sum of the reciprocals of the primes converges.

Probably the first person to think seriously about the distribution of the primes was Gauss, who, in 1792 or 1793, when he was 15 years old, noticed that the number of primes in chiliads – intervals of length 1000 – decreased logarithmically. This led him to conjecture that, if we let $\pi(X) := \#\{p < X\}$, then

$$\pi(X) \sim \frac{X}{\log X}.$$

This conjecture (of course, now proved, and known as the prime number theorem) is consistent with Euler's observation (1.1). In particular, (1.1) is a sort of average version of the prime number theorem, so that it is implied by the prime number theorem, but it is unable to prove the prime number theorem (for example, it can't rule out conspiracies of the sort, e.g., that there are almost no primes starting with the digit 9).

Moving past beautiful work of Chebychev – who proved, for example, Bertrand's Postulate that there are always primes between X and $2X$, but whose techniques are fundamentally unable to prove the prime number theorem – it wasn't until 1859, more than sixty years after Gauss's conjecture was made, that an attack on the prime number theorem was proposed. In his seminal memoir, which has been at the heart of the past 150 years of analytic number theory, Riemann outlines a program whose central object of study is the zeta function

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s},$$

now thought of as a function of a complex variable. Riemann showed that $\zeta(s)$ satisfies the functional equation $Z(s) = Z(1-s)$, where

$$Z(s) := \pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

is the so-called "completed ζ-function", whence $\zeta(s)$ possesses an analytic continuation to the entire complex plane except for a simple pole at $s = 1$. Here, we must remark that, in fact, Euler had a primitive understanding of the functional equation for integral arguments, but that his arguments are only made meaningful by Riemann (e.g., the evaluation $\zeta(-1) = -1/12$ is meaningless without a notion of analytic continuation). In this memoir, Riemann shows that the prime number theorem would follow if one could show that $\zeta(s) \neq 0$ for $\Re(s) = 1$, and in a quantitative form by bounding zeros away from the 1-line. This program was completed, independently, by Hadamard and de la Vallée Poussin, in 1896. Although subsequent progress has been made, and, indeed, we will return to this theme later, it is not at the heart of our story, so we postpone discussion and move back in time.

In work published in 1837, Dirichlet considered the following question: in an arithmetic progression $a \pmod{q}$, with $(a, q) = 1$, how many primes are there? Legendre conjectured that there were infinitely many, and while some progressions (e.g., $p \equiv 3 \pmod{4}$) fall easily to analogues of Euclid's proof, many more (e.g., $p \equiv 1 \pmod{4}$) do not, and, we know now, cannot. But if Euclid's proof cannot generalize, what about Euler's? That is, what can be said about the divergence of the series

$$\sum_{p \equiv a (\mathrm{mod}\, q)} \frac{1}{p}?$$

What Dirichlet observed is that the above sum can be decomposed naturally in terms of what we now call *Dirichlet characters modulo* q: homomorphisms $\chi\colon (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}^\times$ extended to all of $\mathbb{Z}$ by reduction modulo q. For $(a, q) = 1$, these satisfy the orthogonality relation

$$\frac{1}{\phi(q)} \sum_{\chi (\mathrm{mod}\, q)} \chi(n)\overline{\chi(a)} = \begin{cases} 1 & \text{if } n \equiv a \, (\mathrm{mod}\, q) \\ 0 & \text{otherwise} \end{cases},$$

where $\phi(q) := \#\{n \leq q : (n, q) = 1\}$, whence

$$\sum_{p \equiv a (\mathrm{mod}\, q)} \frac{1}{p} = \frac{1}{\phi(q)} \sum_{\chi (\mathrm{mod}\, q)} \overline{\chi(a)} \sum_{p} \frac{\chi(p)}{p}.$$

It turns out that, analogous to Euler's work, the inner summation can be analyzed by studying the function $L(s, \chi)$ defined via

$$L(s, \chi) := \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \prod_{p} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}$$

at the point $s = 1$. In particular, because each χ is periodic, if χ is non-trivial, we have that

$$\sum_{n<X} \chi(n) = O(1)$$

as $X \to \infty$, whence partial summation easily shows that $L(s, \chi)$ is analytic in the region $\Re(s) > 0$. If $\chi = \chi_0$ is trivial, then $L(s, \chi_0) \doteq \zeta(s)$, where $\doteq$ means the equality is valid up to a finite product over primes, and so $L(s, \chi_0)$ possesses a pole at $s = 1$. However, because the connection between $L(s, \chi)$ and $\sum_p \chi(p)/p^s$ is given by taking logarithms (actually, for technical reasons, typically logarithmic derivatives), it is also necessary to establish that $L(1,\chi) \neq 0$. This is typically done via two formulae: first, one shows that

$$\lim_{\sigma\to 1^+} \prod_{\chi (\mathrm{mod}\, q)} L(\sigma,\chi) \geq 1,$$

necessarily including the trivial character χ_0, so that, bearing in mind the pole of $L(s,\chi_0)$ at $s=1$, at most one $L(s,\chi)$ can be zero (such χ is necessarily real, as $L(1,\chi) = L(1,\bar{\chi})$), and then one shows that, if χ is real, i.e. quadratic, then we have Dirchlet's beautiful class number formula

$$L(1,\chi) \doteq h(d)$$

where $h(d)$ is the class number of the quadratic field associated to χ, and the equality is valid up to explicit non-zero terms depending (non-trivially!) on d. This is enough, after some small effort, to show that the series

$$\sum_{p\equiv a (\mathrm{mod}\, q)} \frac{1}{p}$$

diverges; in fact, we even get the weak equidistribution statement that

$$\sum_{\substack{p<X \\ p\equiv a (\mathrm{mod}\, q)}} \frac{1}{p} \sim \frac{1}{\phi(q)} \log\log X.$$

It is also possible to carry out Riemann's program for the primes $p \equiv a \,(\mathrm{mod}\, q)$, and prove the prime number theorem for primes in arithmetic progressions, which requires the fact that the functions $L(s,\chi)$ behave very similarly to $\zeta(s)$. Let $\tau(\chi)$ denote the *Gauss sum* associated to χ,

$$\tau(\chi) := \sum_{a (\mathrm{mod}\, q)} \chi(a) e^{2\pi i a/q},$$

which arises naturally in the decomposition of χ into the (additive) Fourier basis $\{e^{2\pi i a x/q}\}_{a (\mathrm{mod}\, q)}$, and satisfies $|\tau(\chi)| = q^{1/2}$. Then each $L(s,\chi)$ satisfies a functional equation of the form

$$\Lambda(s,\chi) = \frac{\tau(\chi)}{q^{1/2}} \Lambda(1-s,\bar{\chi}),$$

where

$$\Lambda(s,\chi) := q^{s/2}\pi^{-s/2}\Gamma\left(\frac{s+\delta}{2}\right)L(s,\chi)$$

and $\delta = 0$ or 1 according to the parity of χ; we call $\Lambda(s,\chi)$ the completion of $L(s,\chi)$ (the fact that $\tau(\chi)$ arises is not surprising, as Poisson summation is the main tool used to prove the functional equation). Thus, $\zeta(s)$ and the functions $L(s,\chi)$ form the prototype for a class of functions for which we can carry out Riemann's program, and it is an overriding philosophy of number theory that all such functions are of arithmetic interest. Such functions are called *L-functions*, and, together, $\zeta(s)$ and all $L(s,\chi)$ (which we call Dirichlet L-functions) form the complete set of *degree one* L-functions, the degree being a fundamental measure of the complexity of an L-function.

Not all questions about primes can be attacked via L-functions, however. For example, the famous twin prime conjecture, asserting that there are infinitely many primes p such that $p+2$ is also prime, seems to be a fundamentally different beast. Nevertheless, there is a somewhat natural attack on it (which, of course, has fundamental problems – we currently have no idea how to actually prove the twin prime conjecture). Recall the sieve of Eratosthenes, which would predict that the proportion of primes in the interval $(X^{1/2}, X]$ should be dictated by a product over primes up to $X^{1/2}$. Namely, we expect that

$$\pi(X) - \pi(X^{1/2}) \approx X \prod_{p \le X^{1/2}} \left(1 - \frac{1}{p}\right),$$

which we think of as, for each prime p, discarding those integers that are divisible by p. It turns out that this formula is incorrect: it yields the right order of magnitude, $X/\log X$, but it is off by a constant factor.

Proceeding nonetheless, we might expect that the number of twin primes up to X, $\pi_2(X)$, might be given by a formula of the sort

$$\begin{aligned}\pi_2(X) - \pi_2(X^{1/2}) &\approx X \prod_{p \le X^{1/2}} \frac{\#\{n \,(\mathrm{mod}\, p) : n, n+2 \text{ not divisible by } p\}}{p} \\ &= X \cdot \frac{1}{2} \cdot \prod_{3 \le p \le X^{1/2}} \left(1 - \frac{2}{p}\right) \\ &\asymp \frac{X}{\log^2 X}.\end{aligned}$$

But, given the problems with even the sieve of Eratosthenes, could such a formula possibly be true? Heuristically, we would expect so, but we are currently unable to prove it. To see why, consider the sieve of Eratosthenes in more detail. For any $z \ge 2$, let

$$P(z) := \prod_{p \le z} p,$$

$\omega(d) := \#\{p : p|d\}$, $\mu(d) := (-1)^{\omega(d)}$ if d is squarefree, and $\mu(d) = 0$ otherwise. By inclusion-exclusion, we have that

$$\begin{aligned}\pi(X) - \pi(X^{1/2}) &= \sum_{d|P(X^{1/2})} \mu(d) \cdot \#\{n \le X : d \mid n\} \\ &= \sum_{d|P(X^{1/2})} \mu(d) \left\lfloor \frac{X}{d} \right\rfloor \\ &= X \sum_{d|P(X^{1/2})} \frac{\mu(d)}{d} + \sum_{d|P(X^{1/2})} \mu(d) \left(\frac{X}{d} - \left\lfloor \frac{X}{d} \right\rfloor \right) \\ &= X \prod_{p \le X^{1/2}} \left(1 - \frac{1}{p}\right) + O\left(\sum_{d|P(X^{1/2})} 1 \right).\end{aligned}$$

But this error term is

$$O\left(2^{\pi(X^{1/2})}\right) = O\left(2^{X^{1/2}}\right),$$

so it is remarkable that the product is even of the same order of magnitude as $\pi(X)$! The ingenious idea of Brun was that, if we're willing to sacrifice the strength of our conclusion (e.g., to obtain only an upper bound, or a result on numbers with few prime factors – so called *almost-primes*), then the method can be salvaged and be made extremely general. Proceeding along these lines, Brun was able to show that

$$\pi_2(X) \ll \frac{X}{\log^2 X},$$

from which he derived the compelling corollary that the series

$$\sum_{\substack{p: \\ p+2 \text{ is prime}}} \frac{1}{p}$$

converges. He was also able to show that there are infinitely many integers n such that, together, n and $n+2$ have at most nine prime factors. Following along these lines, "approximate" versions of many classical conjectures have been proved: the results on twin primes have been stengthened, there are almost-prime versions of Goldbach's conjecture, and it is known that polynomials represent almost-primes infinitely often. One result we must mention, due to Goldston, Pintz, and Yildirim [28], is the best unconditional result toward the problem of bounded gaps between primes and was published in 2009. In particular, they prove that

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

where p_n denotes the n-th prime. In view of the prime number theorem, the average gap between p_n and p_{n+1} is $\log n$, so this says that, compared to the average gap, gaps between primes can be arbitrarily small.

Despite the above discussion, it would be a mistake to assert that all of analytic number theory has been motivated by the study of the primes. Indeed, throughout the 19th century, when much of the formative work on the distribution of primes was being done, another, disjoint, theory was being developed by the likes of Dedekind, Fricke, Gauss, Jacobi, Kronecker, and Weierstrass, to be continued by the likes of Deuring, Hardy, Hecke, Hilbert, Maass, Ramanujan, and Siegel in the early 20th century (and, of course, many others since). This is the theory of *modular forms*, which, broadly speaking, are functions on the upper half plane $\mathbb{H}$ satisfying remarkable symmetry properties. Examples of modular forms include the partition generating function $P(q)$, given by

$$P(q) := \sum_{n=0}^{\infty} p(n)q^n = \prod_{n=1}^{\infty} \frac{1}{1-q^n},$$

a formula first proved by Euler, where $p(n)$ denotes the number of partitions of n (i.e., unordered tuples of integers summing to n), Jacobi's theta function $\theta(q)$ given by

$$\theta(q) := \sum_{n=-\infty}^{\infty} q^{n^2},$$

and the elliptic j-function, which plays a remarkable role in explicit class field theory over imaginary quadratic fields (this is the theory of complex multiplication, which we will not touch upon further, but which Hilbert declared to be the most beautiful theory in all of science). One remarkable result we must mention, which birthed the ubiquitous *circle method*, is a theorem of Hardy and Ramanujan, which asserts that

$$p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}}.$$

Essentially, the proof of this relied upon a simple observation using the Cauchy integral formula and an extremely clever exploitation of the symmetry properties of the function $P(q)$.

In recent years, the theory of modular forms has blossomed into one of central arithmetic interest. Because of their symmetry properties, modular forms naturally give rise to L-functions of degree two, and, in a certain sense, seem to give rise to all such L-functions. This is of extreme importance: to any elliptic curve over $\mathbb{Q}$ (the solutions to the Diophantine equation $y^2 = x^3 + Ax + B$), one can construct what we call the *Hasse-Weil L-function*, but it turns out that actually proving that it is an L-function in a legitimate way is extremely difficult and is done by showing that it agrees with the L-function associated to a modular form. This was the crucial step in the resolution of Fermat's Last Theorem by Wiles and Taylor, with the full modularity result for all elliptic curves over $\mathbb{Q}$ being due to Breuil, Conrad, Diamond, and Taylor. More generally, all L-functions (of any degree) are expected to come from generalizations of modular forms known as automorphic forms, and any L-function is expected to be of arithmetic interest; this is the celebrated Langlands program, which is the focus of much research in number theory, but about which we shall say no more.

The theory of L-functions has been at the heart of analytic number theory ever since Riemann's 1859 memoir and is an area of active research (cf., the Langlands program mentioned above). Nevertheless, there is a real sense in which analytic number theory is stuck. The zeros of any L-function are expected to be of extreme importance (cf. Riemann's program to prove the prime number theorem), but actually saying anything non-trivial is extremely hard. Given any L-function, we expect that all zeros should lie on the line of symmetry of the functional equation, $\Re(s) = 1/2$, that is, we expect all zeros to be as far as possible from the 1-line (which, e.g., would yield best possible quantitative versions of the prime number theorem); this is the infamous Riemann hypothesis, which is one of the Clay prize problems (strictly speaking, the Riemann hypothesis only concerns $\zeta(s)$, with the generalized Riemann hypothesis, GRH, concerning any L-function). For the Riemann zeta function $\zeta(s)$, any zero $s = \sigma + it$ is known only to satisfy

$$\sigma < 1 - \frac{c}{\log^{2/3}|t| \log\log^{1/3}|t|},$$

whereas for a general L-function, typically, if we are able to say anything, we are only able to say that

$$\sigma < 1 - \frac{c}{\log |t|}.$$

In fact, for general L-functions, the situation is often even worse: we sometimes cannot rule out the existence of an exceptional zero lying inside the above region incredibly close to $s = 1$. Such a zero is called a *Siegel zero*. Siegel zeros cause many problems in analytic number theory, and are one of the principal sources of ineffectivity of many otherwise beautiful theorems. Moreover, the above "zero-free regions" are classical: the bounds for general L-functions (although hard to establish in general) essentially follow from the classical techniques of Hadamard and de la Vallée Poussin, whereas the bounds for $\zeta(s)$ are from 1958. That is, when it comes to the zeros of L-functions, we have been stuck for more than 50 years.

If analytic number theory is able to say shockingly little about the zeros of L-functions, is there another approach to the classical questions, one which fundamentally avoids this issue? That is, is it possible to do analytic number theory without L-functions? In 1948, Erdős and Selberg found an elementary proof (meaning it avoided L-functions, not that it was simple – his contributions were part of what led to Selberg earning the Fields Medal) of the prime number theorem, and it was expected that this would open the floodgates, and that a new era of analytic number theory would be ushered in. This new era never materialized, and the elementary proof, while beautiful, was essentially relegated to being a curiosity of mathematics. In recent years, however, Granville and Soundararajan have written a series of papers which introduces ideas that finally establish a context for the elementary proof of Erdős and Selberg. This is the *pretentious* view of analytic number theory.

How does pretentiousness work? There is an overriding philosophy in analytic number theory that the poles of a Dirichlet series correspond to the summatory function of its coefficients, with zeros providing information about the multiplicative structure of the coefficients (e.g., if they're multiplicative, their correlation with the primes). Post-Riemann, analytic number theory has almost exclusively focused on the analytic side of this philosophy: to study a multiplicative function, one should look at its Dirichlet series and determine all salient information about its analytic properties. What Granville and Soundararajan propose is a systematic, general study of multiplicative functions themselves, completely bypassing the need to find analytic information about the Dirichlet series (which, in every case of interest, requires looking at the series in a region where it's defined only by analytic continuation). Where there is ground to be gained is by proving deep theorems about the structure of multiplicative functions. One can imagine taking a function of interest, plugging it into this machinery that's been developed, and, almost for free, receiving interesting arithmetic information about the function. Put another way, pretentiousness works by considering extremely general multiplicative functions simultaneously, while the standard L-function approach (which cannot touch certain, more combinatorial, multiplicative functions that naturally fit into the pretentious view) works with *individual* functions. Pretentiousness is a very young theory, so while, as yet, pretentious techniques can only match the results produced by more classical approaches, it is nevertheless an exciting time in analytic number theory: we finally have a new idea.

1.1 Gauss sums

Recall that, to a Dirichlet character $\chi \pmod q$, one associates the Gauss sum

$$\tau(\chi) := \sum_{a \pmod q} \chi(a) e^{2\pi i a/q},$$

and that $\tau(\chi)/q^{1/2}$ arises naturally as the root number of the functional equation satisfied by the Dirichlet L-function $L(s, \chi)$. Zagier [18] asked, in general, when root numbers are roots of unity.

There is another flavor of Gauss sum: given a finite field $\mathbb{F}_q$ and a character χ of the multiplicative group $\mathbb{F}_q^\times$, one associates the Gauss sum

$$g(\chi) := \sum_{a \in \mathbb{F}_q^\times} \chi(a) e^{2\pi i \mathrm{Tr}(a)/p},$$

where $q = p^f$, and, for any $x \in \mathbb{F}_q$, $\mathrm{Tr}(x) := x + x^p + \cdots + x^{p^{f-1}}$ denotes the trace map $\mathrm{Tr}\colon \mathbb{F}_q \to \mathbb{F}_p$. This quantity is still of critical interest, as it connects the multiplicative structure of χ to the additive structure of $\mathbb{F}_q$. Motivated by the question of Zagier, we consider when $\varepsilon(\chi) := q^{-1/2} g(\chi)$ is a root of unity (it is classical that $|g(\chi)| = q^{1/2}$). In Chapter 2, using p-adic techniques, we prove the following classification of characters χ with $\varepsilon(\chi)$ a root of unity, which depends only on the order m of χ (this result was also proved by Evans, with a longer, fundamentally different, proof).

Theorem 1.1. *Suppose that χ is a multiplicative character of order m, and let r be the order of p modulo m. Then $\varepsilon(\chi)$ is a root of unity if and only if, for each $t \in \mathbb{Z}$, we have that*

$$\sum_{i=0}^{m-1} \overline{tp^i} = \frac{rm}{2},$$

where $\overline{tp^i}$ denotes the reduction of tp^i modulo m in $\{0, \ldots, m-1\}$.

Essentially the idea of the proof is to use the Gross-Koblitz formula [36], which is the beautiful fact that $g(\chi)$ is, up to an explicit power of p, given by the product of values of the p-adic gamma function $\Gamma_p(z)$, which is defined via p-adic interpolation of factorials. The necessity of the condition is established by considering the p-adic valuation of $g(\chi)$ and its conjugates, and sufficiency is established via a result of Gross and Koblitz on when the product of values of $\Gamma_p(z)$ is a root of unity.

1.2 Sieve theory and the distribution of primes

Recall that sieve techniques, developed by Brun and expanded by others, consider "approximations" to classical questions about the distribution of primes (e.g., the twin prime conjecture and Goldbach's conjecture): letting $\mathcal{P}_2$ denote the set of integers that are either prime or the product of two primes, it is known that there are infinitely many primes p such that $p+2$ is in $\mathcal{P}_2$, and it is also known that every sufficiently large even integer is the sum of a prime and an element of $\mathcal{P}_2$. Another classical problem concerns primes represented by polynomials. That is, given an irreducible polynomial $F(x)$, is $F(n)$ prime infinitely often, where $n \in \mathbb{Z}$? It turns out that there can be *local obstructions* (e.g., the polynomial $F(x) = x^2 + x + 2$ is always even), but for polynomials without local obstructions, we can follow similar heuristic reasoning as we did with the twin primes, and we conjecture that $F(n)$ is prime infinitely often. Dirichlet's theorem on primes in arithmetic progressions yields this conjecture in the case $\deg F(x) = 1$, but for no higher degree polynomials is it known. In Chapter 3, we prove the following theorem.

Theorem 1.2. *Let $F(x) \not\equiv x^2 + x \pmod 2$ be an irreducible quadratic polynomial. Then there are infinitely many n such that $F(n)$ is in $\mathcal{P}_2$.*

This theorem is the best possible result obtainable via sieve theory, as the parity problem dictates that it is impossible to detect primes with a linear sieve (which we are necessarily forced to use: on average, we discard one residue class per prime). Theorem 1.2 is a generalization of an Inventiones paper of Iwaniec [44], where he proves the same result for $F(x) = x^2 + 1$. Among other new inputs, Iwaniec's proof relied upon the equidistribution of roots to the quadratic congruence $F(x) \equiv 0 \pmod m$, which is established using the arithmetic of the underlying field, $\mathbb{Q}(i)$. There are also deep theorems of Friedlander and Iwaniec [25] and of Heath-Brown [41], who proved that the polynomials $x^2 + y^4$ and $x^3 + 2y^3$ are prime infinitely often (here, we bypass the restriction that sieves cannot detect primes because of the bivariate nature of the polynomials), but, again, we see that the underlying fields, $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt[3]{2})$, are simple. When proving Theorem 1.2, to establish equidistribution, we must consider arbitrary quadratic fields, both of positive and negative discriminant, and of class number greater than 1. Both of these features present serious technical difficulties.

1.3 The analytic theory of modular forms

The theory of modular forms is, naturally, very broad. In this thesis, we consider several questions: the Alder-Andrews and Andrews conjectures in the theory of partitions, a question concerning two types of special modular forms known as eta-quotients and theta functions, and a question concerning the representation of integers by ternary quadratic forms.

1.3.1 The Alder-Andrews and Andrews conjectures

Recall that the *circle method* was developed by Hardy and Ramanujan to attack the problem of determining the asymptotic behavior of the partition function, $p(n)$. The method easily generalizes to the general problem of estimating the coefficients of a modular form with a pole, but it has also proved fruitful even when the generating function is not modular. Indeed, the best known results toward Waring's problem and the Goldbach conjecture are proved using the circle method, where estimates for exponential sums are used as substitutes for modularity.

A classical identity from the theory of partitions, due to Euler, is that the number of partitions of an integer n into odd parts is the same as the number of partitions into distinct parts. Let $Q_d(n)$ denote the number of partitions of n into parts congruent to $\pm 1 \pmod{d+3}$, and let $q_d(n)$ denote the number of partitions of n such that the difference between any two parts is at least d. Euler's identity is the statement that $Q_1(n) = q_1(n)$, and the second Rogers-Ramanujan identity implies that $Q_2(n) = q_2(n)$. This exact trend does not continue, however, as an identity of Schur shows that $Q_3(n) \leq q_3(n)$, with equality holding only for finitely many n. Motivated by these classical results, Alder conjectured in 1953 that $Q_d(n) \leq q_d(n)$ for all values of n and d. With C. Alfes and M. Jameson, we proved the following theorem (presented in Section 4.1).

Theorem 1.3. *Alder's conjecture is true.*

Andrews [2] and Yee [80] proved Alder's conjecture for $d \geq 31$ via combinatorial techniques, where there is enough wiggle room in the generating functions to construct an injection, thus leaving those cases of small d where analytic number theory intervenes. To prove the conjecture, we obtained explicit asymptotic formulae using the circle method, both for modular and non-modular generating functions, and verified the conjecture for all "small" n. However, proceeding naively, one first obtains a bound on n of the order 10^{200}, so that considerable effort is needed to reduce the problem to something computationally feasible. In response to our proof of Alder's conjecture, Andrews made a related conjecture, which, with Jameson, we prove, in Section 4.2, holds asymptotically.

1.3.2 Eta-quotients and theta functions

In the theory of modular forms, there is a stark dichotomy between two classes of arithmetic interest: there are modular forms of integer weight, and there are those of half-integer weight. Although many deep questions abound, the integer weight theory is much more developed than that of half-integer weight. This is no accident: the spaces of half-integral weight modular forms enjoy less rigid structure than integral weight spaces, which opens the door for yet deeper theory. Indeed, the coefficients of half-integral weight forms frequently encode values of L-functions, and this is both the source of much delight and much frustration.

There are two natural examples of half-integer weight forms: theta functions, either associated to a Dirichlet character or to a quadratic form of odd dimension, and certain η-quotients, which are modular forms generated by infinite products. Zagier has asked for a classification of those η-quotients which are holomorphic, and his student, Mersmann [58], showed that, for any given weight, there are essentially only finitely many. Since, by a theorem of Serre and Stark, all weight $1/2$ forms are linear combinations of theta functions associated to even Dirichlet characters, this shows that only finitely many η-quotients can also be equal to a theta function associated to an even Dirichlet character. Two questions immediately jump out: which η-quotients are they, and what can be said about theta functions associated to odd Dirichlet characters? We answer both of these questions in Section 4.3.

Theorem 1.4. *There are exactly seven η-quotients that are theta functions for an even Dirichlet character, and there are exactly five that are theta functions for an odd Dirichlet character.*

In fact, this theorem also applies to certain linear spaces generated by theta functions, but not to the full linear span. Since any theta function associated to a Dirichlet character is necessarily lacunary, meaning almost all of its Fourier coefficients are 0, this also provides a partial answer to a problem of Serre [73] to classify the lacunary η-quotients. Although there are other classifications of η-quotients known, this is some of the first progress toward Serre's problem in half-integral weight.

1.3.3 Representation by ternary quadratic forms

It is a classical problem to classify the integers represented by a positive definite, integral quadratic form. In particular, one may ask when such a form represents all positive integers. This is now known due to recent theorems of Conway-Schneeberger-Bhargava [9] and Bhargava-Hanke [8], a pair of results known as the 15 and 290 theorems. If a form does not represent all positive integers, perhaps the next best thing would be for it to represent all integers which are locally represented (represented modulo m for all integers m). Such forms are known as *regular* quadratic forms, and Jagy, Kaplansky, and Schiemann [47] proved that there are at most 913 regular ternary quadratic forms, of which 899 are known to be regular. Adapting techniques of Ono and Soundararajan [65], we prove the following theorem in Section 4.4.

Theorem 1.5. *Assume the GRH for all Dirichlet and modular L-functions. Then each of the remaining 14 forms is regular.*

The subtlety of this question, and the need for the GRH, is that the theta function associated to any ternary quadratic form can be canonically decomposed into two parts, one of which is, essentially, the class number of an imaginary quadratic field, the other of which is, essentially, the central value of a modular L-function. It is known, ineffectively, due to Duke and Schulze-Pillot [22], that the class number part eventually dominates, proving that any sufficiently large and locally represented integer is globally represented. However, similar in spirit to work of Granville and Stark [35] who show that a failure of a form of the *abc*-conjecture implies the existence of a Siegel zero, we exploit the ineffectivity to prove the following theorem, also in Section 4.4.

Theorem 1.6. *Assume the GRH for modular L-functions. There is an explicitly computable constant $C(Q)$ such that if $n \geq C(Q)$ is locally represented by Q but not globally, then some Dirichlet L-function has a Siegel zero.*

1.4 The pretentious view of analytic number theory

In the pretentious view of analytic number theory, one looks for deep structure theorems concerning multiplicative functions. One beautiful theorem among many in the subject – and the first foundational result – is due to Halász, which classifies those functions with large partial sums, i.e., multiplcative $f(n)$ such that both $|f(n)| \leq 1$, $f(n) \in \mathbb{C}$, and

$$S_f(X) := \sum_{n<X} f(n) \gg X.$$

Granville and Soundararajan define a distance $\mathbb{D}(f,g)$ between two multiplicative functions as a sort of average of the difference on primes. Generically, we expect that $\mathbb{D}(f,g)$ should be infinite, but in the event that it is not, we say that $f(n)$ is $g(n)$-pretentious. Halász's theorem states that if $S_f(X) \gg X$ as $X \to \infty$, then $f(n)$ is n^{it}-pretentious for some $t \in \mathbb{R}$ – that is, it "comes from" a natural example of a function with large sums.

The principal focus of Chapter 5 of this thesis is the following complementary question: suppose that $S_f(x)$ exhibits more cancellation than it has a right to – that is, suppose that $S_f(X) \ll X^{1/2-\delta}$ as $X \to \infty$ for some $\delta > 0$ – must $f(n)$ be pretentious to a Dirichlet character? While this question is likely intractable in general, in Section 5.1, we provide an affirmative answer for a natural class of functions $\mathcal{S}_K$ defined via the arithmetic of any number field K, where one can see Dirichlet characters naturally arising.

Theorem 1.7. *If $f \in \mathcal{S}_K$ satisfies $S_f(X) \ll X^{1/2-\delta}$ for some $\delta > 0$, then f coincides with a Dirichlet character.*

It turns out, somewhat surprisingly, that pretentiousness does not "detect" power cancellation – it is possible for $S_f(x) \ll 1$, say, and for $g(n)$ to be $f(n)$-pretentious, and yet for $S_g(x) \not\ll x^{1-\epsilon}$ for any $\epsilon > 0$. In work with J. Jung, we rectify this situation by establishing new definitions of pretentiousness that do detect power cancellation. As an example of our work, which is presented in Section 5.2, we have the following theorem, which should be thought of as establishing the right pretentious framework to ask the question of extreme cancellation.

Theorem 1.8. *If $f(n)$ is $g(n)$-strongly pretentious and $S_f(x) \ll x^\alpha$ for some $\alpha > 0$, then $S_g(x) \ll x^\alpha$.*

Chapter 2

Gauss sums over finite fields and roots of unity

Let $p > 2$ be a prime, and let $q = p^f$ for some $f \geq 1$. Let $\psi\colon \mathbb{F}_p \to \mathbb{C}^\times$ be a non-trivial additive character, and let $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ be a non-trivial multiplicative character. The Gauss sum $g(\chi) = g(\chi, \psi)$ associated to χ is given by

$$g(\chi) := \sum_{x \in \mathbb{F}_q^\times} \chi(x)\psi(\mathrm{tr}(x)), \tag{2.1}$$

where $\mathrm{tr}(x) := x + x^p + \ldots + x^{p^{f-1}}$. The determination of $g(\chi)$ is of central importance in analytic number theory as it reflects both the multiplicative and additive structure of $\mathbb{F}_q$. Classical arguments show that $|g(\chi)| = \sqrt{q}$. On the other hand, the quantity $\varepsilon(\chi) := g(\chi)/\sqrt{q}$ has only been determined for χ of certain orders (see [7] for a comprehensive treatment of recent results). Motivated by private communications with Zagier, we determine when $\varepsilon(\chi)$ is a root of unity.

Theorem 2.1. *Let $\chi\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ be a multiplicative character of order m and let r be the order of p modulo m. The quantity $\varepsilon(\chi)$ is a root of unity if and only if for every integer t coprime to m we have that*

$$\sum_{i=0}^{r-1} \overline{tp^i} = \frac{rm}{2}, \tag{2.2}$$

where $\overline{tp^i}$ denotes the canonical representative of tp^i modulo m in $[0,\ldots,m-1]$.

Remark: After this work was done, the author learned that Theorem 2.1 was first obtained by Evans [24]. Evans's proof used Stickelberger's relation on the decomposition of $g(\chi)$ into prime ideals (see [43]). An equivalent condition, essentially (2.7) below, was later obtained by Yang and Zheng [79], again using Stickelberger's relation. We give a different proof of Theorem 2.1, one based on a deep theorem of Gross and Koblitz [36] relating Gauss sums to the p-adic gamma function.

2.1 The Gross-Koblitz formula

Let $p > 2$ be a prime and $q = p^f$ for some $f \geq 1$. The p-adic gamma function $\Gamma_p(z)\colon \mathbb{Z}_p \to \mathbb{Z}_p^\times$ is defined by

$$\Gamma_p(z) := \lim_{\substack{m\to z\\ m\in\mathbb{Z}}} (-1)^m \prod_{\substack{j<m\\ (j,p)=1}} j. \tag{2.3}$$

Let $\omega_f\colon \mathbb{F}_q^\times \to \mathbb{C}^\times$ be the Teichmüller character of $\mathbb{F}_q$, $\psi\colon \mathbb{F}_p \to \mathbb{C}^\times$ be a non-trivial additive character, and $\zeta_p = \psi(1)$. Let $\pi \in \mathbb{Q}_p(\zeta_p)$ be the unique element satisfying both $\pi^{p-1} = -p$ and $\zeta_p \equiv 1 + \pi \pmod{\pi^2}$. For integers $0 \leq a < q-1$, the Gauss sum $g(\omega_f^{-a})$ is defined by

$$g(\omega_f^{-a}) := -\sum_{x\in\mathbb{F}_q^\times} \omega_f^{-a}(x)\psi(\mathrm{tr}(x)), \tag{2.4}$$

where $\mathrm{tr}(x) := x + x^p + \ldots + x^{p^{f-1}}$. The Gross-Koblitz formula [36] states that

$$g(\omega_f^{-a}) = \pi^{S(a)} \prod_{j=0}^{f-1} \Gamma_p\left(\left\{\frac{ap^j}{q-1}\right\}\right), \tag{2.5}$$

where $S(a)$ denotes the sum of digits in the base p expansion of a and, for any $x \in \mathbb{R}$, $\{x\} := x - \lfloor x \rfloor$ denotes the fractional part of x.

2.2 Proof of Theorem 2.1

Let χ be a multiplicative character of $\mathbb{F}_q^\times$ of order m. There is a unique a such that $0 \leq a < q-1$ and $\chi = \omega_f^{-a}$. Since $g(\chi) \in \mathbb{Q}(\zeta_p, \zeta_{q-1})$, $\varepsilon(\chi)$ is a root of unity if and only if $g(\chi)^{2p(q-1)} = q^{p(q-1)}$. The Gross-Koblitz formula (2.5) yields that

$$g(\chi)^{2p(q-1)} = p^{2p(q-1)S(a)/(p-1)} \left(\prod_{j=0}^{f-1} \Gamma_p \left(\left\{ \frac{ap^j}{q-1} \right\} \right) \right)^{2p(q-1)}, \tag{2.6}$$

and by comparing the p-adic valuation of both sides, we see that a necessary condition for $\varepsilon(\chi)$ to be a root of unity is $S(a) = \frac{f(p-1)}{2}$. In fact, if χ' is another character of $\mathbb{F}_q^\times$ of order m, then there is an element of $\mathrm{Gal}(\mathbb{Q}(\zeta_p, \zeta_m))$ taking $g(\chi)$ to $g(\chi')$. Hence, $\varepsilon(\chi)$ is a root of unity if and only if $\varepsilon(\chi')$ is. Thus, if $\varepsilon(\chi)$ is a root of unity, for all t coprime to m we have that

$$S(\overline{ta}^{(q-1)}) = \frac{f(p-1)}{2}, \tag{2.7}$$

where $\overline{ta}^{(q-1)}$ is the canonical reduction of ta modulo $q-1$. This condition will prove to be sufficient to guarantee that $\varepsilon(\chi)$ is a root of unity. To see this, we begin by reinterpreting the sum of digits function $S(a)$.

Lemma 2.2. *For any* $0 \leq b < q-1$*, we have that*

$$\sum_{j=0}^{f-1} \left\{ \frac{bp^j}{q-1} \right\} = \frac{S(b)}{p-1}.$$

Proof. Write $b = \sum_{i=0}^{f-1} b_i p^i$. For any $0 \leq j \leq f-1$, we observe that $bp^j \equiv b^{(j)} \pmod{q-1}$ where $0 \leq b^{(j)} < q-1$ is the j-th iterate of the cyclic permutation on the base p digits of b. Hence, we have that

$$\begin{aligned} \sum_{j=0}^{f-1} \left\{ \frac{bp^j}{q-1} \right\} &= \frac{1}{q-1} \sum_{j=0}^{f-1} b^{(j)} \\ &= \frac{S(b)}{p-1}. \end{aligned}$$

$\square$

Write $a = t_0 \cdot (a, q-1)$ for some t_0 coprime to m. Since $m = \frac{q-1}{(a,q-1)}$, we have that

$$\left\{\frac{ap^j}{q-1}\right\} = \left\{\frac{t_0 p^j}{m}\right\} = \frac{\overline{t_0 p^j}}{m},$$

whence

$$\sum_{j=0}^{f-1}\left\{\frac{ap^j}{q-1}\right\} = \frac{f}{r}\sum_{j=0}^{r-1}\frac{\overline{t_0 p^j}}{m}, \tag{2.8}$$

where $\overline{tp^j}$ is the reduction of tp^j modulo m and r is the multiplicative order of p modulo m. Hence, by Lemma 2.2, (2.7) holds for t coprime to m if and only if we have that

$$\sum_{j=0}^{r-1}\overline{tp^j} = \frac{rm}{2}. \tag{2.9}$$

This establishes the necessity of (2.2) in the statement of Theorem 2.1. Sufficiency follows immediately from a result of Gross and Koblitz [36]: If $\{a_1, \ldots, a_k, n_1, \ldots, n_k\}$ is a set of integers such that, for all u coprime to m, $\sum_{i=1}^{k} n_i \cdot \overline{ua_i}$ is an integer independent of u, then the product

$$\prod_{i=1}^{k}\prod_{j=0}^{f-1}\Gamma_p\left(\frac{\overline{a_i p^j}}{m}\right)^{n_i}$$

is a root of unity. We apply this result with $k = r$, $a_i = p^i$, and $n_i = 2$, showing that if (2.2) is satisfied, then $\varepsilon(\chi)$ is a root of unity.

Chapter 3

Almost-primes represented by quadratic polynomials

Let $G(x) = c_g x^g + c_{g-1}x^{g-1} + \ldots + c_1 x + c_0 \in \mathbb{Z}[x]$ be an irreducible polynomial of degree g and discriminant D, and let $\rho(m) = \rho_G(m)$ denote the number of incongruent solutions to the congruence $G(n) \equiv 0 \pmod{m}$. Throughout, we assume that $c_g > 0$ and $\rho(p) \neq p$ for all primes p. The question of how often $G(x)$ represents primes is the content of a conjecture by Bouniakowsky [12], and, more generally, by Schinzel [70] and Bateman and Horn [6]:

Conjecture 3.1. Assuming the notation and hypotheses above, we have that

$$\#\{1 \leq n \leq x : G(n) \text{ is prime}\} \sim \Gamma_G \cdot \frac{x}{\log x},$$

where

$$\Gamma_G := \frac{1}{g} \prod_{p \text{ prime}} \left(1 - \frac{\rho(p)}{p}\right)\left(1 - \frac{1}{p}\right)^{-1}.$$

The prime number theorem for primes in arithmetic progressions implies that this conjecture is true when $g = 1$. Very little is known if $g \geq 2$.

Remark. There have been fantastic recent results on the related problem for polynomials in two variables, such as $x^2 + y^4$ and $x^3 + 2y^3$, which Friedlander and Iwaniec [25] and Heath-Brown [41] have shown represent primes infinitely often; in fact, they have obtained the asymptotic orders of the sets of such primes.

Here we consider how frequently $G(x)$ represents numbers that are "almost prime." To this end, let P_r denote the set of squarefree positive integers with at most r distinct prime factors. The best general result along the lines of the above conjecture asserts that a degree g polynomial $G(x)$ represents P_{g+1} infinitely often. For $g \leq 7$, this is due to Kuhn [52], Wang [78], and Levin [54], and for general g this follows from work of Buhštab [16] and Richert [68]. In the special case of $G(x) = x^2 + 1$, a deep theorem of Iwaniec [44] states that $G(x)$ represents P_2 infinitely often. To prove this, Iwaniec obtained a new form of the error in the linear sieve, and he proved an equidistribution result about the roots of the quadratic congruence $x^2 + 1 \equiv 0 \pmod{m}$. By generalizing Iwaniec's result, we are able to obtain the following theorem.

Theorem 3.2. *If $G(x) = c_2x^2 + c_1x + c_0 \in \mathbb{Z}[x]$ is irreducible, with $c_2 > 0$ and $\Gamma_G \neq 0$, then there are infinitely many positive integers n such that $G(n)$ is in P_2.*

Remark. 1) If $G(x) = c_2x^2 + c_1x + c_0 \in \mathbb{Z}[x]$ is irreducible, with $c_2 > 0$ and $\Gamma_G = 0$, then, since $\rho_G(p) \leq 2$ for all primes p, we must have that $\rho_G(2) = 2$. The polynomials $G_0(x) := G(2x)/2$ and $G_1(x) := G(2x+1)/2$ are irreducible, have integer coefficients, and satisfy $\rho_{G_0}(2) = \rho_{G_1}(2) = 1$. Theorem 3.2 then shows that $G(n)$ is $2P_2$ infinitely often.

2) The author, in unpublished work, has obtained conditions on higher degree $G(x)$ which would allow one to conclude that $G(x)$ represents P_g infinitely often. Unfortunately, these conditions are rather technical, and there are no higher degree polynomials yet known to satisfy them.

To prove Theorem 3.2, we use the method employed by Iwaniec [44] to consider arbitrary quadratic polynomials. In Section 3.1, we transform the original problem into a sifting problem to which we can apply Iwaniec's linear sieve inequality. To obtain non-trivial cancellation in the resulting error terms and deduce Theorem 3.2, we need a result on the distribution of roots of $G(x)$ to various moduli, which we prove in Section 3.2. To prove this result for $G(x) = x^2+1$, Iwaniec made use of the fact that $\mathrm{disc}(x^2+1) = -4$ is negative, which allowed him to use the theory of positive definite quadratic forms. It is here, therefore, that most of the additional work in handling arbitary quadratic polynomials is necessary, to account for the fact that the discriminant may be positive and also that $G(x)$ may not be monic. This equidistribution problem also provides the obstruction for establishing the analogue of Theorem 3.2 for higher degree polynomials.

3.1 Proof of Theorem 3.2

We assume from here on out that $G(x)$ is a fixed irreducible quadratic polynomial with positive leading coefficent such that $\rho(2) \neq 2$. We apply the method of Iwaniec [44] to obtain an estimate for

$$\#\{1 \leq n < x : G(n) \in P_2\}.$$

We will introduce a weighted sum in Section 3.1.1 which will change the problem into one of establishing estimates of sifting functions, which we study by using the linear sieve in Section 3.1.2. In Section 3.1.3, we then use these estimates to complete the proof of Theorem 3.2.

3.1.1 A weighted sum

If we let

$$\mathcal{A} = \mathcal{A}_x := \{G(n) : 1 \leq n < x\}, \tag{3.1}$$

we wish to estimate the sum

$$\sum_{a \in \mathcal{A} \cap P_2} 1.$$

To do so, we introduce a weight function $w(n)$ and instead sum $w(a)$. Let λ be a real number such that $2 \le \lambda < 3$, and assume x is sufficiently large so that $G(n) \le x^\lambda$ for all $n \le x$. If n is a positive integer, let p_n and $\omega(n)$ denote the smallest prime divisor of n and the number of distinct prime divisors of n, respectively. For a prime $p < x^{\lambda/2}$ such that $p|n$, let

$$\omega_p(n) := \begin{cases} 1 - \frac{\log p}{\lambda/2 \log x} & \text{if } p = p_n \\ \frac{\log p_n}{\lambda/2 \log x} & \text{if } p > p_n \text{ and } p < x^{\lambda/4} \\ 1 - \frac{\log p}{\lambda/2 \log x} & \text{if } p > p_n \text{ and } p \ge x^{\lambda/4}, \end{cases}$$

then let

$$w(n) := 1 - \frac{\lambda/2}{3-\lambda} \sum_{p|n, p<x^{\lambda/2}} \omega_p(n). \tag{3.2}$$

Remark. The weights $w(n)$ are the same weights that Iwaniec used, which are due to Richert (unpublished, see [44]). Laborde [53] developed weights which would yield a slightly better implied constant for the asymptotic $\#(\mathcal{A} \cap P_2) \gg \frac{x}{\log x}$, but since we have suppressed the constant, we choose to use Richert's weights to maintain continuity with Iwaniec.

We require a lemma due to Iwaniec [44, Lemma 1], which asserts that the weight function $w(n)$ detects P_2 for squarefree n.

Lemma 3.3 (Iwaniec)**.** *If $n \le x^\lambda$ and $w(n) > 0$, then n has at most 2 distinct prime factors.*

By Lemma 3.3, for any $z \le x^{\lambda/4}$ we have that

$$\#\{a \in \mathcal{A} : a \in P_2\} \ge \sum_{\substack{a \in \mathcal{A} \\ (a, P(z))=1 \\ a \text{ squarefree}}} w(a),$$

where $P(z) = \prod_{p<z} p$. If $z = x^\gamma$ for some $\gamma > 0$, there are few non-squarefree $a \in \mathcal{A}$ such that $(a, P(z)) = 1$, as

$$\sum_{\substack{n<x \\ (G(n),P(z))=1 \\ G(n) \text{ not squarefree}}} 1 \ll x^{\lambda/2} z^{-1/2} + x^{2/3} \log^{4/3} x,$$

which we obtain by Iwaniec's argument for x^2+1 and an application of the square sieve [19, Theorem 2.3.5]. Hence, we consider the sum

$$W(\mathcal{A},z) = \sum_{\substack{a\in\mathcal{A}\\(a,P(z))=1}} w(a), \tag{3.3}$$

with the goal of showing that $W(\mathcal{A},z) \gg \frac{x}{\log x}$. For any positive integer q, let

$$\mathcal{A}_q := \{a \in \mathcal{A} : a \equiv 0 \,(\mathrm{mod}\ q)\}.$$

Following Iwaniec, we can write $W(\mathcal{A},z)$ in terms of the sifting functions

$$S(\mathcal{A}_q,u) := \#\{a \in \mathcal{A}_q : (a,P(u)) = 1\}, \tag{3.4}$$

namely we have that

$$\begin{aligned} W(\mathcal{A},z) \;&=\; S(\mathcal{A},z) + \frac{\lambda/2}{3-\lambda}\Bigg[\sum_{z\le p<x^{\lambda/4}}\ \sum_{z\le p_1<p} \frac{\log p/p_1}{\lambda/2\log x} S(\mathcal{A}_{pp_1},p_1) \\ &\quad - \sum_{z\le p<x^{\lambda/4}}\left(\left(1-2\frac{\log p}{\lambda/2\log x}\right)S(\mathcal{A}_p,p) + \frac{\log p}{\lambda/2\log x}S(\mathcal{A}_p,z)\right) \\ &\quad - \sum_{x^{\lambda/4}<p<x^{\lambda/2}}\left(1-\frac{\log p}{\lambda/2\log x}\right)S(\mathcal{A}_p,z)\Bigg]. \end{aligned} \tag{3.5}$$

3.1.2 The linear sieve

We have reduced the problem to that of obtaining a lower bound for the function $W(\mathcal{A},z)$ defined by (3.3), and by (3.4) and (3.5) this reduces to the problem of obtaining good estimates for the sifting functions $S(\mathcal{A}_q,u)$. We recall the following linear sieve inequality [44, Lemma 2].

Lemma 3.4 (Iwaniec). *Let $q \geq 1$, $u \geq 2$, $M \geq 2$, and $N \geq 2$. For any $\eta > 0$ we have*

$$\begin{aligned} S(\mathcal{A}_q, u) &\leq V(u)x\frac{\rho(q)}{q}(F(s)+E) + 2^{\eta^{-7}}R(\mathcal{A}_q; M, N), \\ S(\mathcal{A}_q, u) &\geq V(u)x\frac{\rho(q)}{q}(f(s)-E) - 2^{\eta^{-7}}R(\mathcal{A}_q; M, N), \end{aligned}$$

where $s = \log MN/\log u$, $E \ll \eta s^2 + \eta^{-8}e^{-s}(\log MN)^{-1/3}$, and

$$V(u) = \prod_{p<u}\left(1 - \frac{\rho(p)}{p}\right).$$

The functions $F(s)$ and $f(s)$ are the continuous solutions of the system of differential-difference equations

$$\begin{aligned} sf(s) &= 0 && \text{if} \quad 0 < s \leq 2, \\ sF(s) &= 2e^C && \text{if} \quad 0 < s \leq 3, \\ (sf(s))' &= F(s-1) && \text{if} \quad s > 2, \\ (sF(s))' &= f(s-1) && \text{if} \quad s > 3, \end{aligned}$$

where C is Euler's constant. The error term $R(\mathcal{A}_q; M, N)$ has the form

$$R(\mathcal{A}_q; M, N) = \sum_{m<M, n<N, mn|P(u)} a_m b_n r(\mathcal{A}_q; mn), \tag{3.6}$$

where

$$r(\mathcal{A}_q; d) := |\mathcal{A}_{[q,d]}| - \frac{\rho([q,d])}{[q,d]}x,$$

and the coefficients a_m and b_n are real numbers, bounded by 1 in absolute value, and supported on squarefree values of m and n.

The functions $F(s)$ and $f(s)$ both tend to 1 monotonically as $s \to \infty$, $F(s)$ from above and $f(s)$ from below. Thus, we wish to choose M and N so that s is large, but we do so at the expense of increasing the size of the error term $R(\mathcal{A}_q; M, N)$. Consequently, we are mainly concerned with bounding $R(\mathcal{A}_q; M, N)$ for large values of M and N.

Lemma 3.5. *With notation as in Lemma 3.4, for any $\epsilon > 0$ we have*

$$\sum_{m<x^{1-8\epsilon}} \left| \sum_{\substack{n<x^{\gamma_0-\gamma_1\epsilon} \\ (n,m)=1}} b_n r(\mathcal{A}; mn) \right| \ll x^{1-\epsilon},$$

where $\gamma_0 := \frac{1-\alpha_0}{2(1+\beta_0)}$ and $\gamma_1 := \frac{4\alpha_0}{1+\beta_0}$, where α_0 and β_0 are defined in Lemma 3.6.

Before we prove Lemma 3.5, we state a result whose proof we postpone until Section 3.2 (see Lemma 3.10).

Lemma 3.6. *Let q be a squarefree number, d an odd divisor of q, μ an integer prime to d, and ω a root of $G(x)$ modulo d. Furthermore, let $M < M_1 < 2M$ and $0 \le \alpha < \beta < 1$. Let $P(M_1, M; q, d, \mu, \omega, \alpha, \beta)$ denote the number of pairs of integers m, Ω such that $M < m < M_1$, $(m, q) = 1$, $m \equiv \mu \,(mod\ d)$, $\alpha \le \frac{\Omega}{mq} < \beta$, $G(\Omega) \equiv 0 \,(mod\ mq)$, and $\Omega \equiv \omega \,(mod\ d)$. Then there are constants $A(q) > 0$, $\alpha_0 < 1$ and β_0 such that, for every $\epsilon > 0$,*

$$P(M_1, M; q, d, \mu, \omega, \alpha, \beta) = (\beta - \alpha)(M_1 - M)\rho\left(\frac{q}{d}\right)\frac{A(q)}{\phi(d)} + O\left(M^{\alpha_0+\epsilon} q^{\beta_0+\epsilon}\right).$$

Proof of Lemma 3.5. Let

$$B(x; m, N) := \sum_{n<N,(n,m)=1} b_n r(\mathcal{A}; mn).$$

Our initial task will be to bound $B(x; m, N)$ by using Lemma 3.6. By the Cauchy-Schwarz inequality, we get

$$\sum_{M<m<2M} |B(x; m, N)| \le M^{\frac{1}{2}} \left(\sum_{M<m<2M} B(x; m, N)^2 \right)^{\frac{1}{2}}. \tag{3.7}$$

Since we have that

$$B(x; m, N) = \sum_{\substack{0\le v<m \\ G(v)\equiv 0 (\text{mod } m)}} \sum_{\substack{n<N \\ (n,m)=1}} b_n \left(\sum_{\substack{k<x \\ k\equiv v (\text{mod } m) \\ G(k)\equiv 0 (\text{mod } n)}} 1 - \frac{x}{m}\frac{\rho(n)}{n} \right),$$

the Cauchy-Schwarz inequality implies that

$$\begin{aligned} B(x;m,N)^2 &\leq \rho(m) \sum_{\substack{0\leq v<m \\ G(v)\equiv 0 (\mathrm{mod}\ m)}} \left[\sum_{\substack{n<N \\ (n,m)=1}} b_n \left(\sum_{\substack{k<x \\ k\equiv v (\mathrm{mod}\ m) \\ G(k)\equiv 0 (\mathrm{mod}\ n)}} 1 - \frac{x}{m}\frac{\rho(n)}{n} \right) \right]^2 \\ &\ll M^\epsilon \sum_{\substack{0\leq v<m \\ G(v)\equiv 0 (\mathrm{mod}\ m)}} \left[\sum_{\substack{n<N \\ (n,m)=1}} b_n \left(\sum_{\substack{k<x \\ k\equiv v (\mathrm{mod}\ m) \\ G(k)\equiv 0 (\mathrm{mod}\ n)}} 1 - \frac{x}{m}\frac{\rho(n)}{n} \right) \right]^2 . \end{aligned}$$

Expanding the square on the right-hand side and reintroducing the sum over m, we get that

$$\sum_{M<m<2M} B(x;m,N)^2 \ll M^\epsilon \left(W(x;M,N) - 2xV(x;M,N) + x^2 U(M,N) \right), \tag{3.8}$$

where

$$W(x;M,N) := \sum_{M<m<2M} \sum_{\substack{0\leq v<m \\ G(v)\equiv 0 (\mathrm{mod}\ m)}} \sum_{\substack{n_1,n_2<N \\ (n_1n_2,m)=1}} b_{n_1} b_{n_2} \sum_{\substack{k_1,k_2<x \\ k_1\equiv k_2\equiv v (\mathrm{mod}\ m) \\ G(k_1)\equiv G(k_2)\equiv 0 (\mathrm{mod}\ n)}} 1, \tag{3.9}$$

$$V(x;M,N) := \sum_{M<m<2M} \sum_{\substack{0\leq v<m \\ G(v)\equiv 0 (\mathrm{mod}\ m)}} \frac{1}{m} \sum_{\substack{n_1,n_2<N \\ (n_1n_2,m)=1}} b_{n_1} b_{n_2} \frac{\rho(n_2)}{n_2} \sum_{\substack{k<x \\ k\equiv v (\mathrm{mod}\ m) \\ G(k)\equiv 0 (\mathrm{mod}\ n_1)}} 1, \tag{3.10}$$

and

$$U(M,N) := \sum_{M<m<2M} \sum_{\substack{0\leq v<m \\ G(v)\equiv 0 (\mathrm{mod}\ m)}} \frac{1}{m^2} \sum_{\substack{n_1,n_2<N \\ (n_1n_2,m)=1}} b_{n_1} b_{n_2} \frac{\rho(n_1)\rho(n_2)}{n_1 n_2}. \tag{3.11}$$

We will estimate $W(x;M,N)$, $V(x;M,N)$, and $U(M,N)$ separately with the goal of showing that their main terms cancel in the expression (3.8). Our main tools to this end are Lemma 3.6 and partial summation. We follow the method of Iwaniec [44, Proof of Proposition 1] closely, with more effort being necessary only in the estimation of $W(x; M, N)$. Consequently, we state only the results for $U(M, N)$ and $V(x; M, N)$, noting that they follow in the same fashion as the estimate of $W(x; M, N)$ we provide below. In particular, the required estimate for $U(M, N)$ is

$$U(M,N) = \frac{1}{2M} \sum_{n_1,n_2<N} b_{n_1} b_{n_2} \frac{\rho(n_1)\rho(n_2)}{n_1 n_2} A([n_1,n_2]) + O\left(M^{\alpha_0-2+\epsilon} N^{2\beta_0+\epsilon}\right), \tag{3.12}$$

and the required estimate for $V(x; M, N)$ is

$$V(x;M,N) = \frac{x}{2M} \sum_{n_1,n_2<N} b_{n_1} b_{n_2} \frac{\rho(n_1)\rho(n_2)}{n_1 n_2} A([n_1,n_2]) + O\left(x^{\epsilon} + x M^{\alpha_0-2+\epsilon} N^{2\beta_0+\epsilon}\right). \tag{3.13}$$

Follwing Iwaniec's method for $W(x; M, N)$ as far as we can, we obtain

$$W(x;M,N) = \sum_{n_1,n_2<N} b_{n_1} b_{n_2} T^*(n_1,n_2;x,M) + O\left(x^{1+\epsilon}\right),$$

where to define $T^*(n_1, n_2; x, M)$ we need to first define the integers c and d. For integers $l_1, l_2 < \frac{x}{M}$, let $0 \le c < [n_1, n_2]$ be the solution to

$$\begin{aligned} c &\equiv l_1 \left(\bmod \frac{n_1}{(n_1,n_2)}\right) \\ c &\equiv l_2 \left(\bmod \frac{n_2}{(n_1,n_2)}\right) \\ c &\equiv l_1 \left(\bmod (n_1,n_2)\right), \end{aligned}$$

and let

$$d := \frac{(n_1,n_2)}{(n_1,n_2,l_1-l_2)}.$$

With the above definitions, we have

$$T^*(n_1,n_2;x,M) := \sum_{\substack{l_1,l_2<\frac{x}{M} \\ l_1\equiv l_2 (\bmod (2,n_1,n_2))}} \sum_{\substack{0\le\mu<d \\ (\mu,d)=1}} \sum_{\substack{0\le v<d \\ G(\mu l_1+v)\equiv 0 (\bmod d) \\ G(\mu l_2+v)\equiv 0 (\bmod d)}} \Sigma_1, \tag{3.14}$$

where

$$\Sigma_1 := \sum_{\substack{M<m<M_1,(m,n_1n_2)=1 \\ m\equiv\mu (\mathrm{mod}\ d), cm\leq\Omega<(c+1)m \\ \Omega\equiv\mu l_1+v (\mathrm{mod}\ d), G(\Omega)\equiv 0 (\mathrm{mod}\ m[n_1,n_2])}} 1,$$

and $M_1 = \min\left(2M, \frac{x}{l_1}, \frac{x}{l_2}\right)$. In fact, Σ_1 is precisely

$$P\left(M_1, M; [n_1,n_2], d, \mu, \mu l_1 + v, \frac{c}{[n_1,n_2]}, \frac{c+1}{[n_1,n_2]}\right),$$

so Lemma 3.6 implies that

$$\begin{aligned} T^*(n_1,n_2;x,M) &= \frac{A([n_1,n_2])\rho([n_1,n_2])}{[n_1,n_2]} \sum_{\substack{l_1,l_2<\frac{x}{M} \\ l_1\equiv l_2 (\mathrm{mod}\ (2,n_1,n_2))}} \frac{M_1-M}{\rho(d)\phi(d)} \sum_{\mu,v} 1 \\ &\quad + O\left(x^2 M^{\alpha_0-2+\epsilon} N^{2\beta_0+\epsilon}\right). \end{aligned} \tag{3.15}$$

The sum $\sum_{\mu,v} 1$ is counting the number of integers μ and v modulo d such that $(\mu,d)=1$ and $G(\mu l_1+v)\equiv G(\mu l_2+v)\equiv 0\ (\mathrm{mod}\ d)$. This is the same as the number of choices of μl_1+v and μl_2+v such that $G(\mu l_1+v)\equiv G(\mu l_2+v)\equiv 0\ (\mathrm{mod}\ d)$ and their difference, $\mu(l_1-l_2)$, is invertible modulo d. Since d is squarefree and the number of solutions is multiplicative in d, there are exactly $\rho(d)\psi(d)$ ways of doing this, where $\psi(d)$ is the multiplicative function defined by $\psi(p) := \rho(p)-1$ for each prime p. Hence, the sum in (3.15) is equal to

$$\phi((n_1,n_2))^{-1} \sum_{\substack{l_1,l_2<\frac{x}{M} \\ l_1\equiv l_2 (\mathrm{mod}\ (2,n_1,n_2))}} \phi((n_1,n_2,l_1-l_2))\psi\left(\frac{(n_1,n_2)}{(n_1,n_2,l_1-l_2)}\right)(M_1-M).$$

Since $\rho(p)=0,1,$ or 2, we must have that $\psi(p)=0,\pm 1$. We first note that if $\psi(p)=-1$ for some $p \mid [n_1,n_2]$, then $\rho(p)=0$ and so $T^*(n_1,n_2;x,M)$ would then be 0. We therefore assume that $\psi(p)\neq -1$ and evaluate $T^*(n_1,n_2;x,M)$.

Let $n \mid (n_1,n_2)$ be maximal such that $\psi(n)=1$, and let $n_0 = \frac{(n_1,n_2)}{n}$. Since we have

$$\psi\left(\frac{(n_1,n_2)}{(n_1,n_2,l_1-l_2)}\right) = \psi\left(\frac{n_0}{(n_0,l_1-l_2)}\right),$$

it follows that $\psi\left(\frac{(n_1,n_2)}{(n_1,n_2,l_1-l_2)}\right)=0$ unless $n_0 \mid (l_1-l_2)$. Hence, we consider

$$\frac{1}{\phi((n_1,n_2))}\sum_{\substack{l_1,l_2<\frac{x}{M}\\ l_1\equiv l_2 (\mathrm{mod}\ n_0)}}\frac{\phi((n_1,n_2,l_1-l_2))}{\psi\left(\frac{(n_1,n_2,l_1-l_2)}{n_0}\right)}(M_1-M),$$

which, by using the fact that $(n_1,n_2,l_1-l_2)=n_0(n,l_1-l_2)$, is given by

$$\frac{1}{\phi(n)}\sum_{\substack{l_1,l_2<\frac{x}{M}\\ l_1\equiv l_2 (\mathrm{mod}\ n_0)}}\phi((n,l_1-l_2))(M_1-M).$$

We now have that, letting $\xi := \mu * \phi$,

$$\begin{aligned}\sum_{\substack{0<l_1<l_2\\ l_1\equiv l_2 (\mathrm{mod}\ n_0)}}\phi((n,l_1-l_2)) &= \sum_{\substack{0<l_1<l_2\\ l_1\equiv l_2 (\mathrm{mod}\ n_0)}}\sum_{t\mid(n,l_1-l_2)}\xi(t)\\ = \frac{l_2}{n_0}\sum_{t\mid n}\frac{\xi(t)}{t}+O\left(\phi(n)\right) &= \frac{l_2\phi(n)\rho(n)}{n_0 n}+O\left(\phi(n)\right),\end{aligned}$$

where the last equality follows from the evaluation of $\sum_{t\mid n}\frac{\xi(t)}{t}$ on primes. We are thus led to consider

$$\sum_{l_2<\frac{x}{M}} l_2\left(\min\left(2M,\frac{x}{l_2}\right)-M\right)=\frac{x^2}{4M}+O(x).$$

Inserting these estimates into (3.15), we now see that

$$\begin{aligned}T^*(n_1,n_2;x,M) &= \frac{x^2}{2M}\left(A([n_1,n_2])\frac{\rho([n_1,n_2])}{n_1n_2}\rho(n)\right)\\ &\quad +O\left(xN^{\epsilon}\frac{\rho(n_1)\rho(n_2)}{n_1n_2}+x^2M^{\alpha_0-2+\epsilon}N^{2\beta_0+\epsilon}\right).\end{aligned}$$

Hence, we have

$$\begin{aligned}W(x;M,N) &= \frac{x^2}{2M}\sum_{n_1,n_2<N} b_{n_1}b_{n_2}\frac{\rho(n_1)\rho(n_2)}{n_1n_2}\frac{\rho(n)}{\rho((n_1,n_2))}A([n_1,n_2])\\ &\quad +O\left(x^{1+\epsilon}+x^2M^{\alpha_0-2+\epsilon}N^{2+2\beta_0+\epsilon}\right).\end{aligned}$$

Since primes $p \mid n_0$ satisfy $\psi(p) = 0$ and hence $\rho(p) = 1$, we have that $\rho((n_1, n_2)) = \rho(n)$. This implies the required estimate, that

$$\begin{aligned} W(x; M, N) &= \frac{x^2}{2M} \sum_{n_1, n_2 < N} b_{n_1} b_{n_2} \frac{\rho(n_1)\rho(n_2)}{n_1 n_2} A([n_1, n_2]) \\ &\quad + O\left(x^{1+\epsilon} + x^2 M^{\alpha_0 - 2 + \epsilon} N^{2 + 2\beta_0 + \epsilon}\right). \end{aligned} \tag{3.16}$$

Inserting the estimates (3.12), (3.13), and (3.16) into (3.8), we see that the main terms cancel, and we obtain that

$$\sum_{M < m < 2M} B(x; m, N)^2 \ll \left(x + x^2 M^{\alpha_0 - 2} N^{2 + 2\beta_0}\right) x^\epsilon M^\epsilon N^\epsilon. \tag{3.17}$$

Returning to the statement of the lemma, let $N = x^{\gamma_0 - \gamma_1 \epsilon}$. With this choice of N, it suffices to show for any $M < x^{1 - 8\epsilon}$ that

$$\sum_{M < m < 2M} |B(x; m, N)| \ll x^{1 - 3\epsilon/2}.$$

If $M < x^{1 - \gamma_0 - \epsilon}$, the trivial estimate

$$|B(x; m, N)| \leq \rho(m) \sum_{n < N} \rho(n) \ll \rho(m) N$$

yields the desired result.

If $M > x^{1 - \gamma_0 - \epsilon}$, we use the estimate (3.17) in equation (3.7), and obtain

$$\begin{aligned} \sum_{M < m < 2M} |B(x; m, N)| &\ll \left((Mx)^{1/2} + x M^{\frac{\alpha_0 - 1}{2}} N^{1 + \beta_0}\right) x^\epsilon M^\epsilon N^\epsilon \\ &\ll x^{1 - 3\epsilon/2} \end{aligned}$$

by our choice of $M < x^{1 - 8\epsilon}$ and $N = x^{\gamma_0 - \gamma_1 \epsilon}$. □

Armed with Lemma 3.5, we are now able to acquire the desired estimate for the sifting functions $S(\mathcal{A}_q, u)$.

Lemma 3.7. *If $z < x^{\lambda/2r}$, then for any $\epsilon > 0$ and x sufficiently large, we have*

$$\sum_{\substack{q<x^{1-\epsilon}\\(q,P(z_q))=1}} c_q S(\mathcal{A}_q, z_q) < V(z)x\cdot$$

$$\cdot\left(\sum_{\substack{q<x^{1-\epsilon}\\(q,P(z_q))=1}} c_q \frac{\rho(q)}{q} F\left(\frac{(1+\gamma_0)\log x - \log q}{\log z_q}\right)\frac{\log z}{\log z_q} + O_{\log z}(\epsilon)\right),$$

with γ_0 as defined in Lemma 3.5, provided that for each q, $z \leq z_q < x^{\lambda/2r}$ and $0 \leq c_q \leq 1$.

This lemma is essentially the same as Proposition 2 in [44], so we present it without proof. We obtain a lower bound for the sum in Lemma 3.7 by replacing F with f.

3.1.3 Proof of Theorem 3.2

With Lemma 3.7 at our disposal, we obtain a lower bound for the size of the set

$$\{1 \leq n < x : G(n) \in P_2\}.$$

We wish to apply Lemma 3.4 and Lemma 3.7 to equation (3.5) to obtain a lower bound for $W(\mathcal{A}, z)$. We may do this for each term in (3.5) but the short sum

$$\sum_{x^{1-\epsilon}\leq p<x}\left(1 - \frac{\log p}{\lambda/2\log x}\right) S(\mathcal{A}_p, z).$$

However, in this case, we make the estimate

$$S(\mathcal{A}_p, z) \ll \frac{x}{p\log(x/p)},$$

yielding the bound $O\left(\frac{\epsilon x}{\log x}\right)$. For notational convenience, set

$$\alpha := 1 + \gamma_0 \text{ and } \gamma := \frac{\log z}{\log x}.$$

By partial summation, we obtain

$$\begin{aligned}
W(\mathcal{A},z) &> V(z)x\left(f\left(\frac{\alpha}{\gamma}\right)+\left[\int_\gamma^{\frac{1}{2}}\int_\gamma^u \frac{u-t}{1}\frac{\gamma}{t}f\left(\frac{\alpha-u-t}{t}\right)\frac{dt}{t}\frac{du}{u}\right.\right.\\
&\quad -\int_\gamma^{\frac{1}{2}}\left((1-2u)\frac{\gamma}{u}F\left(\frac{\alpha-u}{u}\right)+uF\left(\frac{\alpha-u}{\gamma}\right)\right)\frac{du}{u}\\
&\quad \left.\left.-\int_{\frac{1}{2}}^1(1-u)F\left(\frac{\alpha-u}{\gamma}\right)\frac{du}{u}\right]-\epsilon\right)\\
&=: V(z)x(W-\epsilon),
\end{aligned}$$

where we have let λ tend to 2, which is permitted by continuity. Since $\Gamma_G\neq 0$, we have that $V(z)\asymp\log^{-1}x$ by Mertens' Theorem and we wish to show that $W>0$.

We observe that W decreases monotonically as α increases from 1, so we wish to find $\gamma<\frac{1}{2}$ such that $W|_{\alpha=1}>0$. However, we will not immediately substitute $\alpha=1$ into the above formula. Instead, we will choose $\gamma=\frac{\alpha}{6}$ and take the limit as α tends to 1 from the right. Using that

$$sF(s)=2e^C\left(1+\int_2^{s-1}\log(u-1)\frac{du}{u}\right)$$

if $3\leq s\leq 5$, and

$$sf(s)=2e^C\left(\log(s-1)+\int_3^{s-1}\int_2^{t-1}\log(u-1)\frac{du}{u}\frac{dt}{t}\right)$$

if $4\leq s\leq 6$, we obtain

$$\begin{aligned}
W &= \frac{\alpha e^C}{3}\left(\log\left(\frac{5}{6}\alpha\right)-\frac{\alpha-1}{\alpha}\log(\alpha-1)\right.\\
&\quad\left.-\int_2^4\left[t\log\left(\frac{6(t+1)}{5(t+2)}\right)+(t+1)\log\left(1-\frac{t}{5}\right)\right]\frac{\log(t-1)}{t(t+1)}dt\right).
\end{aligned}$$

Upon taking the limit $\alpha\to 1^+$, we see that

$$W_1=\frac{e^C}{3}\left(\log\left(\frac{5}{6}\right)-\int_2^4\left[t\log\left(\frac{6(t+1)}{5(t+2)}\right)+(t+1)\log\left(1-\frac{t}{5}\right)\right]\frac{\log(t-1)}{t(t+1)}dt\right),$$

which a numerical computation reveals to be positive.

3.2 An equidistribution result for the congruence $G(x) \equiv 0 \pmod{m}$

Here we prove Lemma 3.6, an equidistribution result for the roots of the congruence $G(x) \equiv 0 \pmod{m}$, where $G(x)$ is any irreducible quadratic polynomial. The proof of Theorem 3.2 is complete once this lemma is proved. Before we can do this, however, we need a result concerning the Dirichlet series $L(s, \psi) := \sum_{m=1}^{\infty} \frac{\psi(m)}{m^s}$, where $\psi = \rho * \mu$ and $\rho(m)$ is the number of incongruent solutions to $G(x) \equiv 0 \pmod{m}$. Although we only need this result for $\deg(G(x)) = 2$, we prove the following result for any irreducible polynomial $G(x)$.

Lemma 3.8. *The series $L(s, \psi)$ converges to a positive real number at $s = 1$.*

Proof. One, admittedly easier, way to establish this result would be to observe that

$$L(s, \psi) = \frac{\zeta_K(s)}{\zeta(s)} \cdot A(s),$$

where $\zeta_K(s)$ denotes the Dedekind zeta function of the splitting field of $G(x)$, and $A(s)$ is given by an absolutely convergent product in the region $\Re(s) > 1/2$. However, we eschew this attack to present what we believe to be a more aesthetically pleasing proof.

If D is the discriminant of $G(x)$, then, by Hensel's Lemma, we can express the Euler product for $L(s, \psi)$ as

$$L(s, \psi) \quad = \quad \lambda_D(s) \prod_{p \nmid D} \left(1 + \frac{\psi(p)}{p^s}\right) =: \lambda_D(s) L_0(s, \psi),$$

where $\lambda_D(s)$ is the product arising from primes $p \mid D$. Since it is a finite product, it will have no bearing on the convergence of $L(1, \psi)$. Thus, we are only concerned with the convergence of $L_0(1, \psi)$. Assuming that s is tending to 1 in the half-plane $\Re(s) > 1$, we have that

$$\begin{aligned}
\log\left(L_0(s,\psi)\right) &= \sum_{p\nmid D} \log\left(1 + \frac{\psi(p)}{p^s}\right) \\
&= \sum_{p\nmid D} \frac{\psi(p)}{p^s} + O\left(\sum_{p\nmid D} \frac{1}{p^{2\Re(s)-\epsilon}}\right) \\
&= \sum_{p\nmid D} \frac{\psi(p)}{p^s} + O(1).
\end{aligned}$$

Since $\rho(p)$ can be interpreted Galois theoretically and depends only on the conjugacy class C of Frob_p in $\mathrm{Gal}(G)$, we have, letting $\mathrm{Gal}(G)^{\#}$ denote the set of conjugacy classes of $\mathrm{Gal}(G)$ and recalling that $\psi(p) = \rho(p) - 1$,

$$\begin{aligned}
\sum_{p\nmid D} \frac{\psi(p)}{p^s} &= \sum_{C\in \mathrm{Gal}(G)^{\#}} (\rho(C) - 1) \sum_{\mathrm{Frob}_p \in C} p^{-s} \\
&= \sum_{C\in \mathrm{Gal}(G)^{\#}} (\rho(C) - 1) \frac{\#C}{\#\mathrm{Gal}(G)} \log\left(\frac{1}{s-1}\right) + \theta(s),
\end{aligned}$$

where $\theta(s)$ is holomorphic for $\Re(s) \geq 1$. The last equality follows from the Chebotarev Density Theorem (for example, see Proposition 1.5 of [72]). The value of $\rho(C)$ is the number of roots of $G(x)$ in $\mathbb{C}$ fixed by elements of C, so letting $\mathrm{Fix}(C)$ (resp. $\mathrm{Fix}(\sigma)$, for $\sigma \in \mathrm{Gal}(G)$) be the number of fixed points of an element of C (resp. the number of fixed points of σ), we have that

$$\begin{aligned}
\sum_{C\in \mathrm{Gal}(G)^{\#}} \#C \cdot (\rho(C) - 1) &= \sum_{C\in \mathrm{Gal}(G)^{\#}} \#C \cdot \mathrm{Fix}(C) - \#\mathrm{Gal}(G) \\
&= \sum_{\sigma\in \mathrm{Gal}(G)} \mathrm{Fix}(\sigma) - \#\mathrm{Gal}(G) = 0,
\end{aligned}$$

by Burnside's Lemma. Hence, we see that $\log\left(L_0(s,\psi)\right) = O(1)$ as s tends to 1. Thus, the infinite product converges and $L_0(1,\psi)$ exists, whence $L(1,\psi)$ does as well. The fact that $L(1,\psi)$ is positive and real comes immediately from its Euler product and the definition of $\psi(m)$. $\square$

We will also need a lemma of Iwaniec [44, Lemma 7] on the approximation of the characteristic function $\chi_I(t)$ of the interval $I := [\alpha, \beta) \subseteq [0, 1)$ by Fourier series.

Lemma 3.9 (Iwaniec)**.** *Let $2\Delta < \beta - \alpha < 1 - 2\Delta$. There exist two functions $A(t)$ and $B(t)$ such that*

$$|\chi_I(t) - A(t)| = B(t)$$

and

$$\begin{aligned} A(t) &= \beta - \alpha + \sum_{h \neq 0} A_h e(ht) \\ B(t) &= \Delta + \sum_{h \neq 0} B_h e(ht), \end{aligned}$$

with Fourier coefficients A_h and B_h satisfying

$$|A_h|, |B_h| \leq \min\left(\frac{1}{|h|}, \frac{\Delta^{-2}}{|h|^3}\right) =: C_h. \tag{3.18}$$

Armed with Lemmas 3.8 and 3.9, we now prove the main result of this section, which is a generalization of Iwaniec's Lemma 4 [44], and is the precise statement of our Lemma 3.6. For a squarefree integer q we define

$$A(q) := \frac{\phi(q)}{q} \frac{L(1, \psi)}{L_q(1, \psi)}, \tag{3.19}$$

where $\phi(n)$ is Euler's totient function,

$$L_q(1, \psi) := \prod_{p|q} \left(1 + \frac{\psi(p)}{p} + \ldots + \frac{\psi(p^{r_p})}{p^{r_p}}\right), \tag{3.20}$$

and r_p is the smallest integer such that $\psi(p^k) = 0$ for all $k > r_p$. We note that r_p exists as a consequence of Hensel's Lemma because $G(x)$ is irreducible.

Lemma 3.10. *Let q be a squarefree number, d an odd divisor of q, μ an integer prime to d, and ω a root of $G(x)$ modulo d. Furthermore, let $M < M_1 < 2M$ and $0 \le \alpha < \beta < 1$. Let $P(M_1, M; q, d, \mu, \omega, \alpha, \beta)$ denote the number of pairs of integers m, Ω such that $M < m < M_1$, $(m, q) = 1$, $m \equiv \mu \,(mod\ d)$, $\alpha \le \frac{\Omega}{mq} < \beta$, $G(\Omega) \equiv 0 \,(mod\ mq)$, and $\Omega \equiv \omega \,(mod\ d)$. Then there are constants $\alpha_0 < 1$ and β_0 such that, for every $\epsilon > 0$,*

$$P(M_1, M; q, d, \mu, \omega, \alpha, \beta) = (\beta - \alpha)(M_1 - M)\rho\left(\frac{q}{d}\right)\frac{A(q)}{\phi(d)} + O\left(M^{\alpha_0+\epsilon} q^{\beta_0+\epsilon}\right).$$

Proof. By Lemma 3.9, we have that

$$P(M_1, M; q, d, \mu, \omega, \alpha, \beta) = (\beta - \alpha) \sum_{\substack{M<m<M_1,(m,q)=1,m\equiv\mu(\text{mod } d) \\ 0\le\Omega<mq,G(\Omega)\equiv 0(\text{mod } mq),\Omega\equiv\omega(\text{mod } d))}} 1$$
$$+O\left(\rho(q)\Delta M + \sum_{h\neq 0} C_h \left| \sum_{\substack{M<m<M_1,(m,q)=1,m\equiv\mu(\text{mod } d) \\ 0\le\Omega<mq,G(\Omega)\equiv 0(\text{mod } mq),\Omega\equiv\omega(\text{mod } d)}} e\left(\frac{h\Omega}{mq}\right)\right|\right). \tag{3.21}$$

By the Chinese Remainder Theorem, the sum in the main term above is given by

$$\rho\left(\frac{q}{d}\right) \sum_{\substack{M<m<M_1 \\ (m,q)=1 \\ m\equiv\mu(\text{mod } d)}} \rho(m) = \rho\left(\frac{q}{d}\right) \sum_{\substack{a\le T \\ (a,q)=1}} \psi(a) \sum_{\substack{\frac{M}{a}<b<\frac{M_1}{a},(b,q/d)=1 \\ b\equiv\mu\bar{a}(\text{mod } d)}} 1$$
$$+\rho\left(\frac{q}{d}\right) \sum_{\substack{b<2M^{1/2} \\ (b,q)=1}} \sum_{\substack{\max\left(\frac{M}{b},T\right)<a<\frac{M_1}{b} \\ a\equiv\mu\bar{b}(\text{mod } d),(a,q/d)=1}} \psi(a).$$

If $(a, D) = 1$, then $\psi(a) = \left(\frac{D}{a}\right)\mu(a)^2$. Hence, we have that

$$\rho\left(\frac{q}{d}\right)\sum_{\substack{M<m<M_1\\(m,q)=1\\m\equiv\mu(\text{mod } d)}}\rho(m) = \rho\left(\frac{q}{d}\right)\sum_{a\le T,(a,q)=1}\psi(a)\left(\phi\left(\frac{q}{d}\right)\frac{M_1-M}{aq}+O\left(\phi\left(\frac{q}{d}\right)\right)\right)$$

$$+\,O\left(\rho(q)\phi(q)M^{\frac{1}{2}+\epsilon}\right)$$

$$=\rho\left(\frac{q}{d}\right)\phi\left(\frac{q}{d}\right)\frac{M_1-M}{q}\sum_{\substack{a\le T\\(a,q)=1}}\frac{\psi(a)}{a}+O\left(\rho(q)\phi(q)T^{1+\epsilon}+\rho(q)\phi(q)M^{\frac{1}{2}+\epsilon}\right)$$

$$=\rho\left(\frac{q}{d}\right)\phi\left(\frac{q}{d}\right)\frac{M_1-M}{q}\frac{L(1,\psi)}{L_q(1,\psi)}$$

$$+\,O\left(\rho(q)\phi(q)\left(M\frac{\phi(q)}{q}T^{-1+\epsilon}+T^{1+\epsilon}+M^{\frac{1}{2}+\epsilon}\right)\right).$$

By choosing $T = M^{\frac{1}{2}}$, we see that the error above is $O\left(M^{\frac{1}{2}+\epsilon}q^{1+\epsilon}\right)$.

We now estimate the error term in (3.21), which is

$$O\left(\rho(q)\Delta M+\sum_{h\neq 0}C_h\left|\sum_{\substack{M<m<M_1,(m,q)=1,m\equiv\mu(\text{mod } d)\\0\le\Omega<mq,G(\Omega)\equiv 0(\text{mod } mq),\Omega\equiv\omega(\text{mod } d)}}e\left(\frac{h\Omega}{mq}\right)\right|\right).$$

We will bound the above sum by an estimate of the form

$$\sum_{\substack{M<m<M_1,(m,q)=1,\\m\equiv\mu(\text{mod } d),0\le\Omega<mq,\\G(\Omega)\equiv 0(\text{mod } mq),\Omega\equiv\omega(\text{mod } d)}}e\left(\frac{h\Omega}{mq}\right)\ll M^{\alpha_2+\epsilon}q^{\beta_2+\epsilon}\sum_{h\neq 0}C_h\left(1+hM^{\alpha_3+\epsilon}q^{\beta_3+\epsilon}\right)\tau(h),\tag{3.22}$$

and, upon summing over h, we find that the error is

$$O\left(M^{\alpha_2+\epsilon}q^{\beta_2+\epsilon}\left(1+\frac{M^{\alpha_3+\epsilon}q^{\beta_3+\epsilon}}{\Delta}\right)(\log\Delta)^2\right),\tag{3.23}$$

where $\alpha_2 < 1$, $\alpha_3 < 1-\alpha_2$, and β_2 and β_3 are real numbers, and the last equality has come from (3.18).

If $\alpha_3 < 0$, we take $\Delta = M^{\alpha_3} q^{\beta_3}$, yielding that the error in equation (3.21) is

$$O\left(M^{1+\alpha_3+\epsilon} q^{\beta_3+\epsilon} + M^{\alpha_2+\epsilon} q^{\beta_2+\epsilon}\right),$$

in which case we may take $\alpha_0 = \max\left(\frac{1}{2}, \alpha_2, 1+\alpha_3\right)$ and $\beta_0 = \max\left(1, \beta_2, \beta_3\right)$. If $\alpha_3 \geq 0$, we take $\Delta = M^{\frac{\alpha_2+\alpha_3-1}{2}} q^{\beta_3}$, yielding that the error in equation (3.21) is

$$O\left(M^{\frac{1+\alpha_2+\alpha_3}{2}+\epsilon} q^{\beta_3+\epsilon} + M^{\frac{1+\alpha_2+\alpha_3}{2}+\epsilon} q^{\beta_2+\epsilon}\right),$$

and we may take $\alpha_0 = \max\left(\frac{1}{2}, \frac{1+\alpha_2+\alpha_3}{2}\right)$ and $\beta_0 = \max\left(1, \beta_2, \beta_3\right)$. Thus, it only remains to establish (3.22).

We begin by removing the condition that $(m,q)=1$ by Möbius inversion:

$$\sum_{\substack{M<m<M_1,(m,q)=1,m\equiv\mu(\text{mod } d)\\ 0\leq\Omega<qm,G(\Omega)\equiv 0(\text{mod } mq),\Omega\equiv\omega(\text{mod } d)}} e\left(\frac{h\Omega}{mq}\right)$$
$$= \sum_{l|\frac{q}{d}} \mu(l) \sum_{\substack{qM<E<qM_1,E\equiv\mu q(\text{mod } dq),E\equiv 0(\text{mod } lq)\\ 0\leq\Omega<E,G(\Omega)\equiv 0(\text{mod } E),\Omega\equiv\omega(\text{mod } d)}} e\left(\frac{h\Omega}{E}\right).$$

We will estimate the inner sum by using the theory of quadratic forms, a method originally due to Hooley [42]. If c_2 and E are relatively prime, there is a bijection between roots $G(\Omega)\equiv 0\,(\text{mod } E)$ and quadratic forms $[E,y,z]$ of discriminant D, given explicitly by $\Omega = \frac{y-c_1}{2}\overline{c_2}$, where $0\leq\overline{c_2}<E$ is the inverse of c_2 modulo E. To apply this correspondence, therefore, we first take out the part of E not relatively prime to c_2, getting

$$\sum_{\substack{qM<E<qM_1,E\equiv\mu q(\text{mod } dq)\\ E\equiv 0(\text{mod } lq),0\leq\Omega<E\\ G(\Omega)\equiv 0(\text{mod } E),\Omega\equiv\omega(\text{mod } d)}} e\left(\frac{h\Omega}{E}\right)$$
$$= \sum_{\substack{f\leq T\\ (f,c_1)=1}}^{*} \sum_{\substack{0\leq u<fc_2\\ (u,c_2)=1}} \sum_{\substack{0\leq v<f\\ G(v)\equiv 0(\text{mod } f)\\ v\equiv\omega(\text{mod } (d,f))}} e\left(\frac{hv\bar{u}}{f}\right) \sum_{E,\Omega}^{*} e\left(\frac{h\Omega\bar{f}}{E}\right)$$
$$+O\left((qM)^{1+\epsilon}T^{-1+\epsilon}\right),$$

where the star on the first summation indicates that f is composed only of primes dividing c_2, $\bar{u}$ is the inverse of u modulo fc_2, $\bar{f}$ is the inverse of f modulo E, T is a parameter to be specified later, and the star on the innermost summation indicates that E and Ω satisfy $\frac{qM}{f} < E < \frac{qM_1}{f}$, $fE \equiv 0 \,(\mathrm{mod}\ lq)$, $fE \equiv \mu q \,(\mathrm{mod}\ dq)$, $E \equiv u \,(\mathrm{mod}\ fc_2)$, $0 \leq \Omega < E$, $\Omega \equiv \omega \left(\mathrm{mod}\ \frac{d}{(d,f)}\right)$, and $G(\Omega) \equiv 0 \,(\mathrm{mod}\ E)$.

We are now able to use the bijection between roots of quadratic congruences and quadratic forms. From the explicit construction described above, we have that

$$\begin{aligned} \sum_{E,\Omega}^{*} e\left(\frac{h\Omega\bar{f}}{E}\right) &= \sum_{[E,y,z]}^{*} e\left(\frac{h\overline{fc_2}(y-c_1)}{2E}\right) \\ &= \sum_{[E,y,z]}^{*} e\left(\frac{h(y-c_1)}{2fc_2E} - \frac{h\bar{u}(y-c_1)}{2fc_2}\right), \end{aligned}$$

where we have transferred the congruence conditions on Ω to conditions on y. Now, suppose the form $[E,y,z]$ is equivalent to $[a, 2b+c_1, c]$ under the action of $\Gamma^0(fc_2)$. In other words, there is an $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \Gamma^0(fc_2)$ such that

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} a & \frac{2b+c_1}{2} \\ \frac{2b+c_1}{2} & c \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} E & \frac{y}{2} \\ \frac{y}{2} & z \end{pmatrix},$$

where

$$\Gamma^0(fc_2) = \left\{ \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL_2(\mathbb{Z}) : \beta \equiv 0 \,(\mathrm{mod}\ fc_2) \right\}.$$

Then we have that

$$E = a\alpha^2 + (2b+c_1)\alpha\gamma + c\gamma^2 =: E_{\alpha,\gamma}, \tag{3.24}$$

and

$$y = 2a\alpha\beta + (2b+c_1)(\alpha\delta + \beta\gamma) + 2c\gamma\delta. \tag{3.25}$$

Hence, we see that

$$\frac{y-c_1}{2} = a\alpha\beta + c\gamma\delta + b(\alpha\delta+\beta\gamma) + c_1\beta\gamma,$$

from which it follows that

$$\alpha\frac{y-c_1}{2} = \beta E_{\alpha,\gamma} + c\gamma + b\alpha.$$

Thus, we have that

$$\begin{aligned}
&\frac{h(y-c_1)}{2fc_2E} - \frac{h\bar{u}(y-c_1)}{2fc_2} \\
&\qquad = \frac{h\beta}{fc_2\alpha} + \frac{h(c\gamma+b\alpha)}{fc_2\alpha E_{\alpha,\gamma}} - \frac{h\bar{u}(\beta E_{\alpha,\gamma}+c\gamma+b\alpha)}{fc_2\alpha} \\
&\qquad \equiv \frac{h\left(\left(\overline{fc_2}fc_2-1\right)c\bar{u}\gamma - \overline{fc_2}fc_2\bar{\gamma}\right)}{fc_2\alpha} + \frac{h(c\gamma+b\alpha)}{fc_2\alpha E_{\alpha,\gamma}} - \frac{hb\bar{u}}{fc_2} \pmod 1 \\
&\qquad =: \frac{h\left(\left(\overline{fc_2}fc_2-1\right)c\bar{u}\gamma - \overline{fc_2}fc_2\bar{\gamma}\right)}{fc_2\alpha} + h\phi_{\alpha,\gamma},
\end{aligned}$$

where $\bar{\gamma}$ and $\overline{fc_2}$ are the inverses of γ and fc_2 modulo α, respectively. To simplify notation, we denote by $\theta_{\alpha,\gamma}$ the quantity on the right hand side of the final equation. We note that we may obtain a similar expression for $\theta_{\alpha,\gamma}$ with γ in the denominator. With this notation, we have that

$$\sum_{E,\Omega}{}^{*} e\left(\frac{h\Omega\bar{f}}{E}\right) = \sum_{Q=[a,2b+c_1,c]}{}^{\prime} \sum_{\alpha,\gamma}{}^{*} e\left(\theta_{\alpha,\gamma}\right), \tag{3.26}$$

where the outer sum runs over a set of representatives of quadratic forms $Q = [a, 2b+c_1, c]$ of discriminant D under the action of $\Gamma^0(fc_2)$, and the inner sum runs over coprime integers α and γ such that $\frac{qM}{f} < a\alpha^2 + (2b + c_1)\alpha\gamma + c\gamma^2 < \frac{qM_1}{f}$, restricted to one representation of the form (3.24) and (3.25), and satisfying

$$\begin{aligned}
&fE_{\alpha,\gamma} \equiv 0 \pmod{lq}, \\
&fE_{\alpha,\gamma} \equiv \mu q \pmod{dq}, \\
&E_{\alpha,\gamma} \equiv u \pmod{fc_2}, \text{ and} \\
&\left(\tfrac{1-\bar{u}E_{\alpha,\gamma}}{c_2}\right)(c\gamma+b\alpha) - \alpha\omega \equiv 0 \left(\bmod \tfrac{d}{(d,f)}\right).
\end{aligned} \tag{3.27}$$

If either α or γ is fixed, the number of simultaneous soltuions to these congruences, c_G, is bounded by $(q,c)\tau(q)(fc_2)^{\frac{1}{2}}$. Since $c = O(1)$ if $G(x)$ is monic, we have that c_G is $O(q^\epsilon)$ if $G(x)$ is monic and $O(q^{1+\epsilon}f^{\frac{1}{2}})$ otherwise.

Returning to (3.26), we now break into two cases, depending on the sign of D. If D is negative, then the forms $[a, 2b+c_1, c]$ are positive definite, and we may write

$$\sum_{E,\Omega}{}^{*} e\left(\frac{h\Omega\bar{f}}{E}\right) = \sum_{Q=[a,2b+c_1,c]}{}' \frac{1}{|\Gamma_Q|} \sum_{\alpha,\gamma}{}^{*} e\left(\theta_{\alpha,\gamma}\right), \tag{3.28}$$

where the summation over α and γ is no longer restricted to one representation of (3.24) and (3.25) and Γ_Q is the isotropy subgroup of Q in $\Gamma^0(fc_2)$. We consider this case completely before handling the indefinite case, $D > 0$.

Since the number of reduced forms is finite, we are primarily concerned with estimating

$$\sum_{\alpha,\gamma}{}^{*} e\left(\theta_{\alpha,\gamma}\right) \;=\; \sum_{|\gamma|<|\alpha|}{}^{*} e\left(\theta_{\alpha,\gamma}\right) + \sum_{|\alpha|<|\gamma|}{}^{*} e\left(\theta_{\alpha,\gamma}\right).$$

These two sums can be handled in the same way, so we will only provide details for the first. In this case, we have that

$$\begin{aligned}&\left|\sum_{|\gamma|<|\alpha|}{}^{*} e\left(\theta_{\alpha,\gamma}\right)\right| \\ &\qquad \ll c_G \sum_{\alpha} \sup_{\lambda,\Lambda} \left| \sum_{\gamma\equiv\lambda (\mathrm{mod}\ \Lambda)}{}^{*} e\left(\frac{h\left(\left(\overline{fc_2}fc_2 - 1\right)c\bar{u}\gamma - \overline{fc_2}fc_2\bar{\gamma}\right)}{fc_2\alpha} + h\phi_{\alpha,\gamma}\right)\right|.\end{aligned} \tag{3.29}$$

We will use partial summation to handle this inner sum. To do so, we note that

$$\phi_{\alpha,\gamma} - \phi_{\alpha,\gamma+1} \ll \frac{\max(|a|,|b|,|c|)}{|\alpha|qM}. \tag{3.30}$$

We will also need the following estimate for incomplete Kloosterman sums, which can be derived from Weil's bound via the method of completion.

Lemma 3.11. *If u, v, and s are integers and if $0 < r_2 - r_1 < 2s$, then, for any integers λ and Λ, we have that*

$$\sum_{\substack{r_1<r<r_2,(r,s)=1\\ r\equiv\lambda(mod\ \Lambda)}} e\left(\frac{ur+v\bar{r}}{s}\right) \ll s^{\frac{1}{2}+\epsilon}(u,v,s)^{\frac{1}{2}}.$$

Now, by using Lemma 3.11 and (3.30) with partial summation in (3.29), we get that

$$\begin{aligned}\left|\sideset{}{^*}\sum_{|\gamma|<|\alpha|} e\left(\theta_{\alpha,\gamma}\right)\right| &\ll c_G q^{\frac{1}{4}+\epsilon} M^{\frac{1}{4}+\epsilon} f^{\frac{1}{4}} \left(1+\frac{h\max(|a|,|b|,|c|)}{qM}\right)\sum_{\alpha}(\alpha,h)^{\frac{1}{2}}\\ &\ll c_G q^{\frac{3}{4}+\epsilon} M^{\frac{3}{4}+\epsilon} f^{-\frac{1}{4}+\epsilon}\left(1+\frac{h\max(|a|,|b|,|c|)}{qM}\right)\tau(h).\end{aligned}$$

We obtain the same estimate for $\sum^*_{|\alpha|<|\gamma|}$.

If $G(x)$ is monic, then $\max(|a|,|b|,|c|) \ll |D|^{\frac{1}{2}} = O(1)$ by the theory of reduced forms for $SL_2(\mathbb{Z})(=\Gamma^0(1))$. Since the number of reduced forms is finite and depends only on the discriminant, we then have that

$$\sideset{}{^*}\sum_{E,\Omega} e\left(\frac{h\Omega}{E}\right) = O\left(q^{\frac{3}{4}+\epsilon}M^{\frac{3}{4}+\epsilon}\left(1+\frac{h}{qM}\right)\tau(h)\right).$$

The same estimate holds for $\sum_{m,\Omega} e\left(\frac{h\Omega}{mq}\right)$, establishing (3.22).

If $G(x)$ is not monic, by considering the coset representatives of $\Gamma^0(fc_2)$ in $SL_2(\mathbb{Z})$, which can be taken modulo fc_2, we obtain $\max(|a|,|b|,|c|) = O(f^2)$, from which it follows that

$$\sideset{}{^*}\sum_{E,\Omega} e\left(\frac{h\Omega}{E}\right) \ll q^{\frac{7}{4}+\epsilon}M^{\frac{3}{4}+\epsilon}f^{\frac{1}{4}+\epsilon}H_D(fc_2)\left(1+\frac{hf^2}{qM}\right)\tau(h),$$

where $H_D(fc_2)$ denotes the number of reduced forms of discriminant D with respect to the action of $\Gamma^0(fc_2)$. By again considering the coset representatives of $\Gamma^0(fc_2)$ in $SL_2(\mathbb{Z})$, we see that

$$H_D(fc_2) \leq H_D(1)[SL_2(\mathbb{Z}):\Gamma^0(fc_2)] \ll f^{1+\epsilon}.$$

Hence, we have that

$$\begin{aligned}\sum_{m,\Omega} e\left(\frac{h\Omega}{mq}\right) &\ll (qM)^{1+\epsilon}T^{-1+\epsilon} \\ &+ q^{\frac{7}{4}+\epsilon}M^{\frac{3}{4}+\epsilon}\tau(h)\sum_{f\le T}{}^{*}\sum_{(u,fc_2)=1}\rho(f)H_D(fc_2)f^{\frac{1}{4}+\epsilon}\left(1+\frac{hf^2}{qM}\right) \\ &\ll (qM)^{1+\epsilon}T^{-1+\epsilon} + q^{\frac{7}{4}+\epsilon}M^{\frac{3}{4}+\epsilon}T^{\frac{9}{4}+\epsilon}\tau(h)\left(1+\frac{hT^2}{qM}\right)\sum_{f\le T}{}^{*}1 \\ &\ll (qM)^{1+\epsilon}T^{-1+\epsilon} + q^{\frac{7}{4}+\epsilon}M^{\frac{3}{4}+\epsilon}T^{\frac{9}{4}+\epsilon}\tau(h)\left(1+\frac{hT^2}{qM}\right),\end{aligned}$$

where, on the last line, we have used that there are $O(T^{\epsilon})$ values of $f \le T$ whose prime divisors all divide c_2. Upon choosing $T = q^{-\frac{3}{13}}M^{\frac{1}{13}}$, we see that (3.22) holds, with

$$\sum_{m,\Omega} e\left(\frac{h\Omega}{mq}\right) \ll q^{\frac{16}{13}+\epsilon}M^{\frac{12}{13}+\epsilon}\left(1+hq^{-\frac{19}{13}}M^{-\frac{11}{13}}\right)\tau(h).$$

We now consider the indefinite case (i.e. when $D > 0$). To deduce (3.22) from the sum in (3.26), we apply the theory of Pell-type equations. If $D \equiv 0 \pmod 4$, let

$$u^2 - \frac{D}{4}v^2 = 1$$

be chosen such that $\tau := u + v\sqrt{\frac{D}{4}}$ is minimal with $\tau > 1$. If $\tau^m = u_m + v_m\sqrt{\frac{D}{4}}$, let $k = k_{fc_2}$ be the smallest positive integer such that $v_k \equiv 0 \pmod{fc_2}$. If $D \equiv 1 \pmod 4$, let

$$u^2 + uv - \frac{D-1}{4}v^2 = 1$$

be chosen such that $\tau := u + v\left(\frac{1+\sqrt{D}}{2}\right)$ is minimal with $\tau > 1$. If $\tau^m = u_m + v_m\left(\frac{1+\sqrt{D}}{2}\right)$, we again let k be the smallest positive integer such that $v_k \equiv 0 \pmod{fc_2}$.

With this notation, since we may take $a > 0$, there is a unique representative of (3.24) and (3.25) satisfying $\alpha > 0$ and

$$-\frac{2a(\tau^k - 1)}{b + (\tau^k + 1)\sqrt{D}}\alpha < \gamma \leq \frac{2a(\tau^k - 1)}{(\tau^k + 1)\sqrt{D} - b}\alpha.$$

We apply the same techniques as in the positive definite case and find that

$$\sum_{E,\Omega}{}^{*} e\left(\frac{h\Omega}{E}\right) \ll c_G q^{\frac{3}{4}+\epsilon} M^{\frac{3}{4}+\epsilon} f^{\frac{9}{4}+\epsilon} H_D(fc_2)\left(1 + \frac{hf^2}{qM}\right)\tau(h),$$

from which we derive that

$$\sum_{m,\Omega} e\left(\frac{h\Omega}{mq}\right) \ll q^{\frac{3}{4}+\epsilon} M^{\frac{3}{4}+\epsilon}\left(1 + \frac{h}{qM}\right)\tau(h)$$

if $G(x)$ is monic, and

$$\sum_{m,\Omega} e\left(\frac{h\Omega}{mq}\right) \ll q^{\frac{8}{7}+\epsilon} M^{\frac{20}{21}+\epsilon}\left(1 + hq^{-\frac{9}{7}} M^{-\frac{18}{21}}\right)\tau(h)$$

if $G(x)$ is not monic. This establishes (3.22).

□

Chapter 4

The analytic theory of modular forms

4.1 The Alder-Andrews conjecture

(The results in this section are joint with Marie Jameson and Claudia Alfes.)

A famous identity of Euler states that the number of partitions into odd parts equals the number of partitions into distinct parts, and the first Rogers-Ramanujan identity tells us that the number of partitions into parts which are $\pm 1 \pmod 5$ equals the number of partitions into parts which are 2-distinct (a d-distinct partition is one where the difference between any two parts is at least d). Another related identity is a theorem of Schur which states that the partitions of n into parts which are $\pm 1 \pmod 6$ are in bijection with the partitions of n into 3-distinct parts where no consecutive multiples of 3 appear. In 1956, these three facts encouraged H.L. Alder to consider the partition functions $q_d(n) := p(n|d\text{-distinct parts})$ and $Q_d(n) := p(n|\text{parts} \pm 1 \pmod{d+3})$.

Conjecture (Alder). If $\Delta_d(n) = q_d(n) - Q_d(n)$, then, for any $d, n \geq 1$, we have that $\Delta_d(n) \geq 0$.

By the above discussion, the conjecture is true for $d \leq 3$, and the inequality can be replaced by an equality for $d = 1$ and 2. Large tables of values seem to suggest, however, that $q_d(n)$ and $Q_d(n)$ are rarely equal. Andrews [2] refined Alder's conjecture (see [4] for more information on this conjecture):

Conjecture (Alder-Andrews). For $4 \leq d \leq 7$ and $n \geq 2d + 9$, or $d \geq 8$ and $n \geq d + 6$, $\Delta_d(n) > 0$.

Remark. For any given d, there are only finitely many n not covered by the Alder-Andrews conjecture, and a simple argument shows that $\Delta_d(n) \geq 0$ for these n.

In essence, Alder's conjecture asks us to relate the coefficients of

$$\sum_{n=0}^{\infty} Q_d(n)q^n = \prod_{n=1}^{\infty} \frac{1}{(1 - q^{n(d+3)-(d+2)})(1 - q^{n(d+3)-1})}$$

and

$$\sum_{n=0}^{\infty} q_d(n)q^n = \sum_{n=0}^{\infty} \frac{q^{d\binom{n}{2}+n}}{(1 - q)(1 - q^2)\cdots(1 - q^n)}.$$

Although the first generating function is essentially a weight 0 modular form, the second is generally not modular (except in the cases $d = 1$ and 2). This is the root of the difficulty in proving Alder's conjecture, since the task is to relate Fourier coefficients of two functions which have different analytic properties.

Nonetheless, there have been several significant advances toward proving Alder's conjecture. Using combinatorial methods, Andrews [2] proved that Alder's conjecture holds for all values of d which are of the form $2^s - 1$, $s \geq 4$. In addition, Yee ([80], [81]) proved that the conjecture holds for $d = 7$ and for all $d \geq 32$. These results are of great importance because they resolve the conjecture except for $4 \leq d \leq 30, d \neq 7, 15$.

In addition, Andrews [2] deduced that $\lim_{n\to\infty} \Delta_d(n) = +\infty$ using powerful results of Meinardus ([56], [57]) which give asymptotic expressions for the coefficients $q_d(n)$ and $Q_d(n)$. Unfortunately, a mistake in [57] implies that one must argue further to establish this limit. We correct the proof of Meinardus's main theorem (see the discussion after (4.21)) and show that the statement of the theorem remains unchanged. We first prove the following result, which can be made explicit:

Theorem 4.1. *Let $d \geq 4$ and let $\alpha \in [0,1]$ be the root of $\alpha^d + \alpha - 1 = 0$. If $A := \frac{d}{2}\log^2\alpha + \sum_{r=1}^{\infty}\frac{\alpha^{rd}}{r^2}$, then for every positive integer n we have*

$$\Delta_d(n) = \frac{A^{1/4}}{2\sqrt{\pi\alpha^{d-1}(d\alpha^{d-1}+1)}} n^{-3/4}\exp(2\sqrt{nA}) + \mathcal{E}_d(n),$$

where $\mathcal{E}_d(n) = O\left(n^{-\frac{5}{6}}\exp(2\sqrt{nA})\right)$.

Remark. The main term of $\Delta_d(n)$ is the same as the main term for $q_d(n)$ (cf. Theorem 4.6).

In the course of proving Theorem 4.1, we derive explicit approximations for $Q_d(n)$ and $q_d(n)$ (see Theorems 4.3 and 4.6, respectively). Using these results, we obtain the following:

Theorem 4.2. *The Alder-Andrews Conjecture is true.*

In order to prove Theorems 4.1 and 4.2, we consider $q_d(n)$ and $Q_d(n)$ independently and then compare the resulting effective estimates. Accordingly, in Section 4.1.1, we give explicit asymptotics for $Q_d(n)$, culminating in Theorem 4.3. Next, in Section 4.1.2, we laboriously make Meinardus's argument effective (and correct) in order to give an explicit asymptotic formula for $q_d(n)$ in Theorem 4.6. In Section 4.1.3 we use the results from Sections 4.1.1 and 4.1.2 to prove Theorems 4.1 and 4.2.

4.1.1 Estimate of $Q_d(n)$ with explicit error bound

As before, let $Q_d(n)$ denote the number of partitions of n whose parts are $\pm 1 \pmod{d+3}$. From the work of Meinardus, we have that

$$Q_d(n) \sim \frac{(3d+9)^{-\frac{1}{4}}}{4\sin\left(\frac{\pi}{d+3}\right)} n^{-\frac{3}{4}} \exp\left(n^{\frac{1}{2}} \frac{2\pi}{\sqrt{3(d+3)}}\right).$$

In this formula, only the order of the error is known. We will bound the error explicitly, following closely the method of Meinardus [56] as it is presented by Andrews in Chapter 6 of [3]. This allows us to prove the following theorem:

Theorem 4.3. *If $d \geq 4$ and n is a positive integer, then*

$$Q_d(n) = \frac{(3d+9)^{-\frac{1}{4}}}{4\sin\left(\frac{\pi}{d+3}\right)} n^{-\frac{3}{4}} \exp\left(n^{\frac{1}{2}} \frac{2\pi}{\sqrt{3(d+3)}}\right) + R(n),$$

where $R(n)$ is an explicitly bounded function (see (4.10) *at the end of this section).*

Remark. An exact formula for $Q_d(n)$ is known due to the work of Subrahmanyasastri [75]. In addition, by using Maass-Poincaré series, Bringmann and Ono [13] obtained exact formulas in a much more general setting. However, we do not employ these results since the formulas are extremely complicated, and Theorems 4.1 and 4.2 do not require this level of precision.

Preliminary Facts

Consider the generating function f associated to $Q_d(n)$,

$$f(\tau) := \prod_{\substack{n \equiv \pm 1 (d+3) \\ n \geq 0}} (1-q^n)^{-1} = 1 + \sum_{n=1}^{\infty} Q_d(n) q^n,$$

where $q = e^{-\tau}$ and $\Re(\tau) > 0$. Let $\tau = y + 2\pi i x$. We can then obtain a formula for $Q_d(n)$ by integrating $f(\tau)$ against $e^{n\tau}$. Consequently, we require an approximation of $f(\tau)$ so that we may make use of this integral formula. To do this, we need an additional function,

$$g(\tau) := \sum_{\substack{n\equiv\pm1(d+3)\\ n\geq 0}} q^n.$$

Lemma 4.4. *If* $\arg\tau > \frac{\pi}{4}$ *and* $|x| \leq \frac{1}{2}$*, then* $\Re(g(\tau)) - g(y) \leq -c_2 y^{-1}$*, where* c_2 *is an explicitly given constant depending only on* d*.*

Proof. For notational convenience, we will consider the expression

$$-y\left(\Re(g(\tau)) - g(y)\right).$$

Expanding, we find that

$$-y\left(\Re(g(\tau)) - g(y)\right) = S_1 + S_2 + S_3,$$

where

$$S_1 := \frac{\big(1-\cos(2\pi x)\big)\left(e^{(3d+8)y} - e^{(2d+5)y} - e^{(d+4)y} + e^{y}\right)}{\left(\frac{e^{(d+3)y}-1}{y}\right)\left(e^{(2d+6)y} - 2e^{(d+3)y}\cos(2\pi(d+3)x) + 1\right)},$$

$$S_2 := \frac{\big(1-\cos(2\pi(d+2)x)\big)\left(e^{(2d+7)y} - e^{(2d+5)y} - e^{(d+4)y} + e^{(d+2)y}\right)}{\left(\frac{e^{(d+3)y}-1}{y}\right)\left(e^{(2d+6)y} - 2e^{(d+3)y}\cos(2\pi(d+3)x) + 1\right)},$$

and

$$S_3 := \frac{\big(1-\cos(2\pi(d+3)x)\big)\left(2e^{(2d+5)y} + 2e^{(d+4)y}\right)}{\left(\frac{e^{(d+3)y}-1}{y}\right)\left(e^{(2d+6)y} - 2e^{(d+3)y}\cos(2\pi(d+3)x) + 1\right)}.$$

When $y = 0$, each of S_1, S_2, and S_3 is defined. Namely, $S_1 = 0$, $S_2 = 0$, and $S_3 = \frac{2}{d+3}$.

Since these functions are even in x, we may assume $x \geq 0$. Further, the condition $\arg\tau > \frac{\pi}{4}$ implies that $y < 2\pi x$. To find c_2, we note that each $S_i \geq 0$ and so it suffices to bound one away from 0. We do this in three different cases.

Case 1: Suppose that $y \geq \frac{1}{2}$. Since $\frac{1}{2} > x > \frac{1}{2\pi}y$, it follows that $1 - \cos(2\pi x) > 1 - \cos\frac{1}{2}$ and that S_1 is bounded away from 0. In particular,

$$S_1 \geq \frac{\pi\left(1-\cos\frac{1}{2}\right)\left(e^{\frac{3d+8}{2}} - e^{\frac{2d+5}{2}} - e^{\frac{d+4}{2}} + e^{\frac{1}{2}}\right)}{\left(e^{\pi(d+3)}-1\right)\left(e^{\pi(d+3)}+1\right)^2}. \tag{4.1}$$

Case 2: Suppose that $y < \frac{1}{2}$ and $\left|x - \frac{k}{d+3}\right| < \frac{y}{d+3}$ for some positive integer k. Although less obvious than in Case 1, S_1 will again be bounded away from 0:

$$S_1 \geq \frac{\pi\left(1-\cos\frac{\pi}{d+3}\right)}{e^{\pi(d+3)}-1}\frac{e^{(3d+8)y} - e^{(2d+5)y} - e^{(d+4)y} + e^{y}}{\left(e^{(d+3)y}-1\right)^2 + 8\pi^2 y^2 e^{(d+3)y}},$$

and so

$$S_1 \geq \frac{2\pi^3\left(1-\cos\frac{\pi}{d+3}\right)(d+2)(d+3)}{\left(e^{\pi(d+3)}-1\right)\left(\left(e^{(d+3)\pi}+1\right)^2 + 8\pi^4 e^{(d+3)\pi}\right)}. \tag{4.2}$$

Case 3: Suppose that $y < \frac{1}{2}$ and $\left|x - \frac{k}{d+3}\right| \geq \frac{y}{d+3}$ for some non-negative integer k. We additionally assume $\left|x - \frac{k}{d+3}\right| \leq \frac{1}{2(d+3)}$. This is permitted since every x is covered as we vary k. It will be S_3 that is bounded away from 0.

Let $u := 2\pi(d+3)\left|x - \frac{k}{d+3}\right|$ and note that $0 \leq u \leq \pi$, $y \leq \frac{u}{2\pi}$, and $\cos u = \cos 2\pi(d+3)x$. Now, we have that

$$S_3 \geq \frac{4\pi}{e^{(d+3)\pi}-1}\frac{1-\cos u}{\left(e^{\frac{(d+3)u}{2\pi}}-1\right)^2 + 2e^{\frac{(d+3)u}{2\pi}}(1-\cos u)},$$

and a tedious analysis of the derivative of this function implies for $d \geq 4$ that

$$S_3 \geq \frac{8\pi}{\left(e^{(d+3)\pi}-1\right)\left(\left(e^{\frac{d+3}{2}}-1\right)^2 + 4e^{\frac{d+3}{2}}\right)}. \tag{4.3}$$

Obviously, we may take c_2 to be the minimum of the bounds (4.1), (4.2), and (4.3). □

Using Lemma 4.4, we now obtain an approximation for $f(\tau)$.

Lemma 4.5. *If* $|\arg\tau| \le \frac{\pi}{4}$ *and* $|x| \le \frac{1}{2}$*, then*

$$f(\tau) = \exp\left(\frac{\pi^2}{3(d+3)}\tau^{-1} + \log\left(\frac{1}{2\sin\frac{\pi}{d+3}}\right) + f_2(\tau)\right),$$

where $f_2(\tau) = O\left(y^{\frac{1}{2}}\right)$ *is an explicitly bounded function. Furthermore, if* $y \le y_{\max}$ *is sufficiently small,* $0 < \delta < \frac{2}{3}$*,* $0 < \epsilon_1 < \frac{\delta}{2}$*,* $\beta := \frac{3}{2} - \frac{\delta}{4}$*, and* $y^\beta \le |x| \le \frac{1}{2}$*, then there is a constant* c_3 *depending on* d*,* ϵ_1 *and* δ *such that*

$$f(y + 2\pi i x) \le \exp\left(\frac{\pi^2}{3(d+3)}y^{-1} - c_3 y^{-\epsilon_1}\right).$$

Remark. The discussion of the size of $y_{\max}$ will follow (4.4).

Proof. From page 91 of Andrews [3], we have that

$$\log f(\tau) = \tau^{-1}\frac{\pi^2}{3(d+3)} + \log\left(\frac{1}{2\sin\frac{\pi}{d+3}}\right) + \frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \tau^{-s}\Gamma(s)\zeta(s+1)D(s)ds,$$

where $D(s)$ is the Dirichlet series

$$D(s) := \sum_{\substack{n\equiv\pm1(d+3)\\ n\ge0}} \frac{1}{n^s}$$

which converges for $\Re(s) > 1$. Writing

$$D(s) = (d+3)^{-s}\left(\zeta\left(s, \frac{1}{d+3}\right) + \zeta\left(s, \frac{d+2}{d+3}\right)\right),$$

where $\zeta(s,a)$ is the Hurwitz zeta function, we see that $D(s)$ can be analytically continued to the entire complex plane except for a pole of order 1 and residue $\frac{2}{d+3}$ at $s = 1$ (see, for example, page 255 of Apostol's book [5]).

We bound the integral by noting that $|D(s)| \le |\zeta(s)|$, obtaining

$$\left|\frac{1}{2\pi i}\int_{-\frac{1}{2}-i\infty}^{-\frac{1}{2}+i\infty} \tau^{-s}\Gamma(s)\zeta(s+1)D(s)ds\right| \le \xi\sqrt{y},$$

where

$$\xi := \frac{\sqrt{2}}{2\pi}\int_{-\infty}^{\infty}\left|\zeta\left(\frac{1}{2}+it\right)\zeta\left(-\frac{1}{2}+it\right)\Gamma\left(-\frac{1}{2}+it\right)\right| dt.$$

The first statement of the lemma follows.

Remark. Numerical estimates show that $\xi < .224$.

To prove the second statement, we again follow the method of Andrews [3]. We consider two cases: (1) $y^{\beta} \le |x| \le \frac{y}{2\pi}$ and (2) $\frac{y}{2\pi} \le |x| \le \frac{1}{2}$. In the first, we see that $|\arg \tau| \le \frac{\pi}{4}$, so we apply the first statement of the lemma, getting

$$\begin{aligned}
&\log|f(y+2\pi ix)| \\
&\le \frac{\pi^2 y^{-1}}{3(d+3)} + \frac{\pi^2 y^{-1}}{3(d+3)}\left(\left(1+4\pi^2x^2y^{-2}\right)^{-\frac{1}{2}}-1\right) + \log\left(\frac{1}{2\sin\frac{\pi}{d+3}}\right) + \xi\sqrt{y} \\
&\le \frac{\pi^2}{3(d+3)}y^{-1} - c_4 y^{-\frac{\delta}{2}},
\end{aligned}$$

where

$$c_4 := \frac{\pi^4}{3(d+3)}\left(2-\frac{3}{2}y_{\max}^{1-\frac{\delta}{2}}\right) - \log\left(\frac{1}{2\sin\frac{\pi}{d+3}}\right)y_{\max}^{\frac{\delta}{2}} - \xi y_{\max}^{\frac{1+\delta}{2}}.$$

In the second case, we have that

$$\log|f(y+2\pi ix)| = \log f(y) + \Re\left(g(\tau)\right) - g(y),$$

and using Lemma 4.4, we obtain

$$\log|f(y+2\pi ix)| \;\le\; \frac{\pi^2}{3(d+3)}y^{-1} - c_5 y^{-1}$$

where

$$c_5 := c_2 - y_{\max}\log\left(\frac{1}{2\sin\frac{\pi}{d+3}}\right) - \xi y_{\max}^{\frac{3}{2}}. \tag{4.4}$$

We let $c_3{:=}\min\left(c_4(y_{\max})^{\epsilon_1-\frac{\delta}{2}}, c_5(y_{\max})^{\epsilon_1-1}\right)$ and take $y_{\max}$ to be small enough so that $c_3{>}0$. $\square$

Remark. In the proof of Theorem 4.1, we need only bound $Q_d(n)$ since it is of lower order than $q_d(n)$. We shall ignore the restriction on $y_{\max}$ for convenience.

Proof of Theorem 4.3

From the Cauchy integral theorem, we have

$$\begin{aligned} Q_d(n) &= \frac{1}{2\pi i}\int_{\tau_0}^{\tau_0+2\pi i} f(\tau)\exp(n\tau)\,d\tau \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} f(y+2\pi ix)\exp(ny+2n\pi ix)\,dx. \end{aligned}$$

Applying the saddle point method, we take $y = n^{-\frac{1}{2}}\pi/\sqrt{3(d+3)}$ and we let $m := ny$ for notational simplicity. Assuming the notation in Lemma 4.5, for $n \geq 6$, we have $y_{\max} \leq \left(\frac{1}{2\pi}\right)^{\frac{1}{\beta-1}}$, so that both cases in the proof of the second statement of Lemma 4.5 are nonvacuous. We have that

$$Q_d(n) = e^m \int_{-y^\beta}^{y^\beta} f(y+2\pi ix)\exp(2\pi inx)\,dx + e^m R_1,$$

where

$$R_1 := \left(\int_{-\frac{1}{2}}^{-y^\beta} + \int_{y^\beta}^{\frac{1}{2}}\right) f(y+2\pi ix)\exp(2\pi inx)\,dx.$$

By Lemma 4.5,

$$|R_1| \leq \exp\left[\frac{\pi^2}{3(d+3)}\left(\frac{m}{n}\right)^{-1} - c_3\left(\frac{m}{n}\right)^{-\epsilon_1}\right],$$

so

$$|e^m R_1| \leq \exp\left(2m - c_3 m^{\epsilon_1}\left(\frac{\pi^2}{3(d+3)}\right)^{-\epsilon_1}\right). \tag{4.5}$$

Using Lemma 4.5, write

$$Q_d(n) = \exp\left(2m - \log\left(2\sin\frac{\pi}{d+3}\right)\right)\int_{-(m/n)^\beta}^{(m/n)^\beta} \exp(\phi_1(x))\,dx + \exp(m)R_1,$$

where

$$\phi_1(x) := m\left[\left(1+\frac{2\pi i x n}{m}\right)^{-1}-1\right]+2\pi i n x+g_1(x)$$

and $|g_1(x)| \le \xi\sqrt{\frac{\pi^2}{3m(d+3)}}$.

After making the change of variables $2\pi x = (m/n)\omega$, we obtain

$$Q_d(n) = \exp\left(2m+\log\frac{m}{n}+\log\left(\frac{1}{2\sin\frac{\pi}{d+3}}\right)-\log 2\pi\right)I+\exp(m)R_1,$$

where

$$I \;:=\; \int_{-c_{10}m^{1-\beta}}^{c_{10}m^{1-\beta}} \exp(\phi_2(\omega))\,d\omega, \qquad c_{10} := 2\pi\left(\frac{\pi^2}{3(d+3)}\right)^{\beta-1},$$

and

$$\phi_2(\omega) := m\left(\frac{1}{1+i\omega}-1+i\omega\right)+g_1(\omega).$$

We must now find an asymptotic expression for I. Write

$$I = \int_{-c_{10}m^{1-\beta}}^{c_{10}m^{1-\beta}} \exp\left(-m\omega^2\right)\,d\omega+R_2, \tag{4.6}$$

where

$$R_2 := \int_{-c_{10}m^{1-\beta}}^{c_{10}m^{1-\beta}} \exp\left(-m\omega^2\right)\left(\exp(\phi_3(\omega))-1\right)d\omega,$$

with

$$\phi_3(\omega) := m\left(\frac{1}{1+i\omega}-1+i\omega+\omega^2\right)+g_1(\omega).$$

Simplifying, we find that

$$|\phi_3(\omega)| \le c_{10}^3 m^{\frac{3\delta-2}{4}}+\xi\sqrt{\frac{\pi^2}{3m(d+3)}}. \tag{4.7}$$

Substituting $m_{\min} = 2^{\frac{2-\delta}{4}}\pi^{\frac{10-\delta}{4}}(3(d+3))^{-1}$ for m in (4.7), it follows that

$$|\phi_3(\omega)| \;\le\; \frac{2^{\frac{44+8\delta-3\delta^2}{16}}\pi^{\frac{76+8\delta-3\delta^2}{16}}}{3(d+3)}+\xi(2\pi)^{\frac{\delta-2}{8}} =: \phi_{3,\max}.$$

Thus, letting $c_6 := \frac{\exp(\phi_{3,\max})-1}{\phi_{3,\max}}$, we have

$$|\exp(\phi_3(\omega)) - 1| \leq m^{-\frac{1}{2}+\frac{3\delta}{4}} \left(c_6 c_{10}^3 + \xi c_6 m_{\min}^{-\frac{3\delta}{4}} \sqrt{\frac{\pi^2}{3(d+3)}} \right) =: m^{-\frac{1}{2}+\frac{3\delta}{4}} c_7.$$

Hence, we conclude that $|R_2| \leq 2c_{10}c_7 m^{\delta-1}$.

Computing the integral in (4.6), we see that

$$\int_{-c_{10}m^{1-\beta}}^{-c_{10}m^{1-\beta}} \exp\left(-m\omega^2\right) d\omega = \left(\frac{\pi}{m}\right)^{\frac{1}{2}} + g_2(m), \tag{4.8}$$

where $|g_2(m)| \leq 2m^{-\frac{1}{2}} \exp\left(-c_{10} m^{\frac{\delta}{4}}\right)$. Thus, we find that $I = \left(\frac{\pi}{m}\right)^{\frac{1}{2}} + g_2(m) + R_2$. Combining these results, we obtain the desired expression for $Q_d(n)$,

$$Q_d(n) = \left(\frac{n^{-\frac{3}{4}} (3(d+3))^{-\frac{1}{4}}}{4 \sin\left(\frac{\pi}{d+3}\right)} \right) \exp\left(n^{\frac{1}{2}} \frac{2\pi}{\sqrt{3(d+3)}} \right) + R(n), \tag{4.9}$$

where

$$\begin{aligned} |R(n)| \leq \\ & n^{-\frac{1}{4}} \left(\frac{\pi^{\frac{1}{2}} (3(d+3))^{-\frac{3}{4}}}{2\sin\left(\frac{\pi}{d+3}\right)} \right) \exp\left(n^{\frac{1}{2}} \frac{2\pi}{\sqrt{3(d+3)}} - n^{\frac{\delta}{8}} 2\pi^{2-\frac{\delta}{4}} (3(d+3))^{-2+\frac{3\delta}{8}} \right) \\ & + n^{-1+\frac{\delta}{2}} \left(\frac{c_7 \pi^{1+\frac{\delta}{2}}}{(3(d+3))^2 \sin\left(\frac{\pi}{d+3}\right)} \right) \exp\left(n^{\frac{1}{2}} \frac{2\pi}{\sqrt{3(d+3)}} \right) \\ & + \exp\left(n^{\frac{1}{2}} \frac{2\pi}{\sqrt{3(d+3)}} - c_3 n^{\frac{\epsilon_1}{2}} \left(\frac{\pi^2}{3(d+3)} \right)^{-\frac{3\epsilon_1}{2}} \right). \end{aligned} \tag{4.10}$$

4.1.2 Estimate of $q_d(n)$ with explicit error bound

Theorem 2 of [57] (with $k = m = 1$ and $\ell = d$) gives

$$q_d(n) \sim \frac{A^{1/4}}{2\sqrt{\pi\alpha^{d-1}(d\alpha^{d-1}+1)}} n^{-3/4} \exp(2\sqrt{nA}),$$

where α and A depend only on d (see their definitions below in Theorem 4.6). We will bound the error explicitly, following closely the paper of Meinardus [57]. We make his calculations effective, and we obtain the following theorem.

Theorem 4.6. *Let α be the unique real number in $[0,1]$ satisfying $\alpha^d+\alpha-1=0$, and let $A := \frac{d}{2}\log^2\alpha + \sum_{r=1}^{\infty}\frac{\alpha^{rd}}{r^2}$. If n is a positive integer, then*

$$q_d(n) = \frac{A^{1/4}}{2\sqrt{\pi\alpha^{d-1}(d\alpha^{d-1}+1)}} n^{-3/4}\exp(2\sqrt{nA}) + r_d(n)$$

where $|r_d(n)|$ can be bounded explicitly (see the end of this section).

Preliminary Facts

For fixed $d \geq 4$, we have the generating function

$$f(z) := \sum_{n=0}^{\infty} q_d(n)e^{-nz} \tag{4.11}$$

with $z = x + iy$. Hence, we obviously have that

$$q_d(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(z)e^{nz}\,dy. \tag{4.12}$$

Therefore to estimate $q_d(n)$ we require strong approximations for $f(z)$.

Lemma 4.7. *Assuming the notation above, for $|y| \leq x^{1+\epsilon}$ and $x < \beta$, where*

$$\beta := \min\left(-\frac{\pi}{\log\rho}\xi, \frac{2\alpha^{2-d}}{\pi d}, \frac{1}{2d} + \rho\left(\frac{1}{2} - \frac{\pi^2}{24}\right)\right)^{\frac{1}{\epsilon}}$$

and $0 < \xi < 1$ is a constant, we have that

$$f(z) = \left(\alpha^{d-1}\left(d\alpha^{d-1}+1\right)\right)^{-\frac{1}{2}} e^{\frac{A}{z}}(1 + f_{err}(z)),$$

where $f_{err}(z) = o(1)$ is an explicitly bounded function.

Lemma 4.8. *Assuming the notation above, for $x < \beta$ and $x^{1+\epsilon} < |y| \leq \pi$, we have that*

$$|f(x+iy)| \leq \sqrt{\frac{2\pi}{dx}} e^{-\eta\rho x^{2\epsilon-1}}(1+f_2(\rho,x)) \exp\left(\frac{A}{x} + \frac{1-d}{2}\log\alpha + f_1(\rho,x)\right),$$

where f_1, f_2 are functions given in Lemma 4.9, and η is an explicitly given constant.

Remark. Although $\epsilon = \frac{11}{24}$ in [57], we will benefit by varying ϵ in our work.

To prove Lemmas 4.7 and 4.8, we follow [57] and write, by the Cauchy Integral Theorem,

$$f(z) = \frac{1}{2\pi i}\int_{\mathcal{C}} H(w,z)\Theta(w,z)\,\frac{dw}{w} \tag{4.13}$$

where $\mathcal{C}$ is a circle of radius $\rho := 1-\alpha$ centered at the origin,

$$H(w,z) := \prod_{n=1}^{\infty}(1-we^{-nz})^{-1}, \quad \text{and} \quad \Theta(w,z) := \sum_{n=-\infty}^{\infty} e^{-\frac{d}{2}n(n-1)z}w^{-n}. \tag{4.14}$$

Therefore, we estimate $H(w,z)$ and $\Theta(w,z)$.

Lemma 4.9. *Let $\rho = \alpha^d = 1-\alpha$ and suppose $w = \rho e^{i\phi}$ with $-\pi \leq \phi < \pi$. Then for $|y| \leq x^{1+\epsilon}$ and $x < \beta$,*

$$H(w,z) = \exp\left(\frac{1}{z}\sum_{r=1}^{\infty}\frac{w^r}{r^2} + \frac{1}{2}\log(1-w) + f_1(w,z)\right) \tag{4.15}$$

and

$$\Theta(w,z) = \sqrt{\frac{2\pi}{dz}}\exp\left(\frac{\log^2 w}{2dz} - \frac{1}{2}\log w\right)(1+f_2(w,z)), \tag{4.16}$$

where, as $x \to 0$, $f_1(w,z) = O\left(x^{\frac{1}{2}}\right)$ and

$$f_2(w,z) = O\left(x + \exp\left[-\frac{c_0}{x}(\pi - |\phi|) + c_1 x^{\epsilon-1}\right]\right)$$

are explicitly bounded functions.

Proof. First, (4.14) and the inverse Mellin transform yield

$$\log H(w,z) = \frac{1}{2\pi i}\int_{2-i\infty}^{2+i\infty} z^{-s}\Gamma(s)\zeta(s)D(s+1,w)\,ds, \tag{4.17}$$

where $\zeta(s)$ is the Riemann zeta function, $\Gamma(s)$ is the Gamma function, and $D(s,w) := \sum_{r\geq 1}\frac{w^r}{r^s}$, which is defined as a function of s for all fixed w with $|w|<1$.

Note that if $\theta_0 := \arctan x^\epsilon$, then

$$|z^{1/2-it}| \leq |z|^{1/2}e^{\theta_0|t|} \leq \left(1+x^{2\epsilon}\right)^{\frac{1}{4}} x^{1/2}e^{\theta_0|t|}.$$

By changing the curve of integration and accounting for the poles at $s=0$ and 1, we have

$$\log H(w,z) = \frac{1}{z}\sum_{r\geq 1}\frac{w^r}{r^2} + \frac{\log(1-w)}{2} + f_1(w,z),$$

where

$$\begin{aligned} |f_1(w,z)| &= \left|\frac{1}{2\pi i}\int_{-\infty}^{\infty} z^{1/2-it}\Gamma\left(-\frac{1}{2}+it\right)\zeta\left(-\frac{1}{2}+it\right)D\left(\frac{1}{2}+it,w\right)dt\right| \\ &\leq \left(1+x^{2\epsilon}\right)^{\frac{1}{4}} 2^{-\frac{5}{2}}\pi^{-\frac{3}{2}}\zeta\left(\frac{3}{2}\right)\frac{\rho}{1-\rho}\frac{4}{\frac{\pi}{2}-\theta_0}x^{\frac{1}{2}} =: f_1(x) \end{aligned}$$

This proves the first statement as $x^{2\epsilon}$ and $\theta_0 = \arctan x^\epsilon$ both tend toward 0 as $x\to 0$.

The transformation properties of theta functions functions give

$$\Theta(w,z) = \sqrt{\frac{2\pi}{dz}}e^{\frac{(\log w - dz/2)^2}{2dz}}\sum_{\mu=-\infty}^{\infty} e^{\frac{-2\pi^2\mu^2}{dz}-\frac{2\pi i\mu}{dz}(\log w - dz/2)}. \tag{4.18}$$

The argument on page 295 of [57] completes the proof of the lemma, with

$$\begin{aligned} |f_2(w,z)| \leq e^{\frac{d|z|}{8}}&\left[e^{\frac{dx\sqrt{1+x^{2\epsilon}}}{8}} - 1 + 2\frac{\exp\left(-\frac{4\pi^2(1-\xi)}{dx(1+x^{2\epsilon})}\right)}{1-\exp\left(-\frac{2\pi^2(1-\xi)}{dx(1+x^{2\epsilon})}\right)}\right] \\ &+2\exp\left(\frac{2\pi(|\phi|-\pi)}{dx\left(1+x^{2\epsilon}\right)} - \frac{2\pi\log\rho}{d}x^{\epsilon-1} + \frac{d|z|}{8}\right) \\ &=: f_2(\phi,z) = f_2^0(z) + f_2^\phi(z)\exp\left(\frac{2\pi|\phi|}{dx\left(1+x^{2\epsilon}\right)}\right). \end{aligned} \tag{4.19}$$

$\square$

We now prove Lemmas 4.7 and 4.8 using Lemma 4.9.

Proof of Lemma 4.7. Recall from (4.13) that

$$f(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} H(w,z)\Theta(w,z)\,\frac{dw}{w}.$$

Let $\phi_0 = x^c$ with $\frac{3}{8} < c < \frac{1}{2}$. Then

$$f(z) = \frac{1}{2\pi i} \int_{\rho e^{-i\phi_0}}^{\rho e^{i\phi_0}} H(w,z)\Theta(w,z)\,\frac{dw}{w} + \frac{1}{2\pi i} \int_{\mathcal{B}} H(w,z)\Theta(w,z)\,\frac{dw}{w}, \tag{4.20}$$

where $\mathcal{B}$ is the circle $\mathcal{C}$ without the arc $\rho e^{-i\phi_0}$ to $\rho e^{i\phi_0}$.

We first estimate the second integral in (4.20). We note the error of Meinardus [57] in the bound of $\Theta(w,z)$ provided between (25) and (26). From Lemma 4.9, we have that

$$|\Theta(w,z)| \leq \sqrt{\frac{2\pi}{d|z|}}\rho^{-\frac{1}{2}} \exp\left(\frac{x\log^2\rho}{2d(x^2+y^2)} - \frac{\phi^2 x}{2d(x^2+y^2)} + \frac{y\phi\log\rho}{d(x^2+y^2)}\right)(1+|f_2(w,z)|)\,. \tag{4.21}$$

The term $\frac{y\phi\log\rho}{d(x^2+y^2)}$ does not appear in [57] and tends to infinity if $y\phi < 0$. This term arises from the main term of $\Theta(w,z)$, so its contribution cannot be ignored. Furthermore, it is $O\left(x^{\epsilon-1}\right)$, and hence cannot be combined into the negative $O\left(x^{2c-1}\right)$ term arising from $\frac{\phi^2 x}{2d(x^2+y^2)}$. However, the bound Meinardus claims on the product $|H(w,z)\Theta(w,z)|$ is correct. To see this, we need more than the bound $|H(w,z)| \leq H(\rho,x)$ that was originally thought to be sufficient.

From Lemma 4.9, we have that

$$|H(w,z)| \leq \exp\left(|f_1(w,z)|\right)(1+\rho)^{\frac{1}{2}} \exp\left(\Re\left(\frac{1}{z}\sum_{r=1}^{\infty}\frac{w^r}{r^2}\right)\right), \tag{4.22}$$

and

$$\Re\left(\frac{1}{z}\sum_{r=1}^{\infty}\frac{w^r}{r^2}\right) = \frac{x}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r\cos(r\phi)}{r^2}+\frac{y}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r\sin(r\phi)}{r^2}$$
$$= \frac{x}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}+\frac{x}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(\cos(r\phi)-1)$$
$$-\frac{y\phi\log(1-\rho)}{x^2+y^2}+\frac{y}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(\sin(r\phi)-r\phi).$$

Since $\rho=\alpha^d$ and $1-\rho=\alpha$, combining this with (4.22) and (4.21), we see that

$$|H(w,z)\Theta(w,z)|\le\sqrt{\frac{2\pi(1+\rho)}{d|z|\rho}}(1+|f_2(w,z)|)\exp\left(|f_1(w,z)|+\frac{Ax}{x^2+y^2}\right.$$
$$\left.-\frac{\phi^2x}{x^2+y^2}\cdot\left[\frac{1}{2d}+\frac{1}{\phi^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(1-\cos(r\phi))-\frac{y}{\phi^2x}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(\sin(r\phi)-r\phi)\right]\right)$$

Hence, as $x\to 0$ we recover Meinardus's bound on $|H(w,z)\Theta(w,z)|$.

Using the notation of Lemma 4.9,

$$\left|\int_{\mathcal{B}}H(w,z)\Theta(w,z)\frac{dw}{w}\right|\le\sqrt{\frac{2\pi}{d|z|}}\left(\frac{1+\rho}{\rho}\right)^{\frac{1}{2}}\exp\left(f_1(x)+\frac{Ax}{x^2+y^2}\right)\cdot$$
$$\left[\left(1+f_2^0(z)\right)\cdot\int_{\mathcal{B}}e^{-\psi(\phi,z)}d\phi+f_2^{\phi}(z)\int_{\mathcal{B}}\exp\left(-\psi(\phi,z)+\frac{2\pi|\phi|}{dx\,(1+x^{2\epsilon})}\right)d\phi\right],\tag{4.23}$$

where

$$\psi(\phi,z):=\frac{\phi^2x}{2d(x^2+y^2)}+\frac{x}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(1-\cos(r\phi))$$
$$-\frac{y}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(\sin(r\phi)-r\phi).\tag{4.24}$$

We evaluate the two integrals of (4.23) separately.

For the integral $\int_{\mathcal{B}}\exp\left(-\psi(\phi,z)\right)d\,\phi$, we consider two parts based on the sign of $y\phi$. We assume that $y>0$. When $\phi>0$, $\sin(r\phi)-r\phi<0$ for all r, and so $\psi(\phi, z) > 0$. Then

$$\begin{aligned}\int_{\phi_0}^{\pi} \exp\left(-\psi(\phi, z)\right) d\,\phi \quad &\leq \quad \frac{1}{\psi_\phi(\phi_0, z)}\left[\exp\left(-\psi(\phi_0, z)\right) - \exp\left(-\psi(\nu\phi_0, z)\right)\right] \\ &+ \quad \frac{1}{\psi_\phi(\nu\phi_0, z)}\left[\exp\left(-\psi(\nu\phi_0, z)\right) - \exp\left(-\psi(\pi, z)\right)\right],\end{aligned} \tag{4.25}$$

where $\nu > 1$ is a constant.

When $\phi < 0$, we note that $\sin(r\phi) - r\phi > 0$, and so

$$\begin{aligned}&\psi(\phi, z)\\ &\quad\geq \frac{\phi^2 x}{2d(x^2+y^2)} + \frac{x}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(1-\cos(r\phi)) - \frac{x^{1+\epsilon}}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}\left(\frac{r^3\phi^3}{6}\right)\\ &\quad= \frac{\phi^2 x}{2d(x^2+y^2)} + \frac{x}{x^2+y^2}\sum_{r=1}^{\infty}\frac{\rho^r}{r^2}(1-\cos(r\phi)) + \frac{\phi^3 x^{1+\epsilon}\alpha^{d-2}}{6(x^2+y^2)}\\ &\quad=: \hat{\psi}(\phi, z),\end{aligned} \tag{4.26}$$

whence, using that $\frac{\pi^2}{2}\alpha^{d-2}x^\epsilon \leq \frac{\pi}{d}$,

$$\begin{aligned}&\int_{\phi_0}^{\pi} \exp(-\psi(-\phi, z))d\phi\\ &\qquad\leq \frac{1}{\hat{\psi}_\phi(-\phi_0, z)}\left[\exp\left(-\hat{\psi}(-\phi_0, z)\right) - \exp\left(-\hat{\psi}(-\nu\phi_0, z)\right)\right] +\\ &\qquad\quad+\frac{1}{\hat{\psi}_\phi(-\nu\phi_0, z)}\left[\exp\left(-\hat{\psi}(-\nu\phi_0, z)\right) - \exp\left(-\hat{\psi}(-\pi, z)\right)\right].\end{aligned} \tag{4.27}$$

We now consider the second integral in (4.23). A weaker bound on $\psi(\phi, z)$ suffices. In particular, we have $\psi(\phi, z) \geq k\phi^2$, where

$$k := \frac{x}{x^2+y^2}\left(\frac{1}{2d} - \frac{\pi\alpha^{d-2}}{6}\left|\frac{y}{x}\right| + \rho\left(\frac{1}{2} - \frac{\pi^2}{24}\right)\right),$$

which is positive since $x < \beta$.

Hence, we have that

$$\begin{aligned}\int_{\mathcal{B}} \exp\left(-\psi(\phi, z) + \frac{2\pi|\phi|}{dx\,(1+x^{2\epsilon})}\right) d\phi \leq \frac{dx\,(1+x^{2\epsilon})}{\pi - kdx\,(1+x^{2\epsilon})}\cdot\\ \left[\exp\left(-k\pi^2 + \tfrac{2\pi^2}{dx(1+x^{2\epsilon})}\right) - \exp\left(-k\phi_0^2 + \tfrac{2\pi\phi_0}{dx(1+x^{2\epsilon})}\right)\right].\end{aligned} \tag{4.28}$$

Using (4.25), (4.27), and (4.28) in (4.23) gives an explicit bound for the second integral of (4.20), say $E_{\mathcal{B}}(z)$.

Following page 297 of [57], the first integral of (4.20) is given by

$$\begin{aligned} I &:= \frac{1}{2\pi i}\int_{\rho e^{-i\phi_0}}^{\rho e^{i\phi_0}} H(w,z)\Theta(w,z)\,\frac{dw}{w} \\ &= \frac{1}{\sqrt{2\pi d z}}\exp\left(\frac{A}{z}+\frac{1-d}{2}\log\alpha\right)\left(I_{\text{main}}+I_{\text{error}}\right), \end{aligned}$$

where

$$I_{\text{main}} := \int_{-\phi_0}^{\phi_0}\exp\left(-\frac{\phi^2}{2dz}\left(d\alpha^{d-1}+1\right)\right)\,d\phi \tag{4.29}$$

and

$$\begin{aligned} I_{\text{error}} :=\int_{-\phi_0}^{\phi_0}&\left(\exp\left(\log\left(\frac{1-\rho e^{i\phi}}{1-\rho}\right)+f_3(w,z)+f_1(w,z)\right)\left(1+f_2(w,z)\right)-1\right) \\ &\cdot\exp\left(-\frac{\phi^2}{2dz}\left(d\alpha^{d-1}+1\right)\right)\,d\phi, \end{aligned} \tag{4.30}$$

where $|f_3(w,z)|\le \frac{\rho e}{6(1-\rho e)^2}\phi^3$. Then we have

$$|I_{\text{error}}|\le 2\phi_0\left(\frac{1-\rho\cos\phi_0}{1-\rho}\exp\left(f_1(x)+\frac{\rho e}{6(1-\rho e)^2}\phi_0^3\right)\left(1+f_2(\phi_0,z)\right)-1\right), \tag{4.31}$$

and

$$I_{\text{main}} = \sqrt{\frac{\pi z d}{d\alpha^{d-1}+1}} - 2\int_{\phi_0}^{\infty}\exp\left(-\frac{\phi^2}{2dz}\left(d\alpha^{d-1}+1\right)\right)\,d\phi. \tag{4.32}$$

Hence, it follows that

$$I = \frac{\alpha^{\frac{1-d}{2}}}{\sqrt{d\alpha^{d-1}+1}}\exp\left(\frac{A}{z}\right)+\hat{E}_{\phi_0}(w,z), \tag{4.33}$$

where

$$
\begin{aligned}
&|\hat{E}_{\phi_0}(w,z)| \\
&\le \frac{\alpha^{\frac{d-1}{2}}}{\sqrt{2\pi d|z|}} \exp\left(\frac{Ax}{x^2+y^2}\right) \left[\frac{(2d|z|)^{\frac{1}{2}}}{\phi_0\left(d\alpha^{d-1}+1\right)^{\frac{1}{2}}} \exp\left(-\frac{\phi_0^2 x\left(d\alpha^{d-1}+1\right)}{2d\left(x^2+y^2\right)}\right)\right. \\
&\left. + 2\phi_0\left(\frac{1-\rho\cos\phi_0}{1-\rho}\exp\left(f_1(x)+\frac{\rho e}{6(1-\rho e)^2}\phi_0^3\right)(1+f_2(\phi_0,z))-1\right)\right] \\
&=: E_{\phi_0}(z).
\end{aligned}
\tag{4.34}
$$

Hence, we finally see that

$$
|f_{err}(z)| \le \left(E_{\phi_0}(z)+E_{\mathcal{B}}(z)\right)\left(\alpha^{d-1}\left(d\alpha^{d-1}+1\right)\right)^{\frac{1}{2}} \exp\left(\frac{-Ax}{x^2+y^2}\right).
$$

$\square$

Proof of Lemma 4.8. In order to bound f for $x^{1+\epsilon} < |y| \le \pi$, note that $|\Theta(w,z)| \le \Theta(\rho,x)$ by (4.14), which also yields that

$$
\log|H(w,z)| = \Re\{\log H(w,z)\} \le \log H(\rho,x) + \Re\left\{w\sum_{n\ge1}e^{-nz}\right\} - \rho\sum_{n\ge1}e^{-nx}.
$$

On the other hand, we have that

$$
\begin{aligned}
&\Re\left\{w\sum_{n\ge1}e^{-nz}\right\} - \rho\sum_{n\ge1}e^{-nx} \\
&\qquad \le -\rho x^{2\epsilon-1}\left(\beta^{1-2\epsilon}e^{-\beta}\right)\left(\frac{1}{1-e^{-\beta}} - \frac{1}{\sqrt{1-2e^{-\beta}\cos\beta^{1+\epsilon}+e^{-2\beta}}}\right).
\end{aligned}
$$

To see this, note that

$$
\begin{aligned}
&\Re\left\{w\sum_{n\ge1}e^{-nz}\right\} - \rho\sum_{n\ge1}e^{-nx} \\
&\qquad \le -\rho e^{-x}\left(\frac{1}{1-e^{-x}} - \frac{1}{\sqrt{1-2e^{-x}\cos x^{1+\epsilon}+e^{-2x}}}\right).
\end{aligned}
$$

This then gives

$$\begin{aligned}
&\frac{\Re\left\{w\sum_{n\geq 1}e^{-nz}\right\}-\rho\sum_{n\geq 1}e^{-nx}}{-\rho x^{2\epsilon-1}}\\
&\qquad\qquad\geq x^{1-2\epsilon}e^{-x}\left(\frac{1}{1-e^{-x}}-\frac{1}{\sqrt{1-2e^{-x}\cos x^{1+\epsilon}+e^{-2x}}}\right)\\
&\qquad\qquad\geq \beta^{1-2\epsilon}e^{-\beta}\left(\frac{1}{1-e^{-\beta}}-\frac{1}{\sqrt{1-2e^{-\beta}\cos \beta^{1+\epsilon}+e^{-2\beta}}}\right)\\
&\qquad\qquad=: \eta.
\end{aligned}$$

The statement of Lemma 4.8 now follows from (4.13) and Lemma 4.9. $\square$

Proof of Theorem 4.6

From (4.12), it follows that $q_d(n) = I_1 + I_2$, where

$$I_1 := \frac{1}{2\pi}\int_{-x^{1+\epsilon}}^{x^{1+\epsilon}} f(z)e^{nz}\,dy \quad \text{and} \quad I_2 := \frac{1}{2\pi}\left(\int_{-\pi}^{-x^{1+\epsilon}} + \int_{x^{1+\epsilon}}^{\pi}\right) f(z)e^{nz}\,dy.$$

In this proof, we let $x = \sqrt{\frac{A}{n}}$. Following the idea of page 291 of [57], we split I_1 as

$$I_1 = \gamma e^{\frac{2A}{x}} \int_{-x^{1+\epsilon}}^{x^{1+\epsilon}} e^{-\frac{y^2 A}{x^3}}\,dy + E_2 + E_3, \tag{4.35}$$

where γ is given by

$$\gamma := \frac{1}{2\pi\sqrt{\alpha^{d-1}\left(d\alpha^{d-1}+1\right)}},$$

E_2 by

$$E_2 := \gamma e^{\frac{2A}{x}} \int_{-x^{1+\epsilon}}^{x^{1+\epsilon}} e^{-\frac{y^2 A}{x^3}} \left(e^{A\left(\frac{y^4+ixy^3}{x^3(x^2+y^2)}\right)} - 1\right) dy,$$

and

$$E_3 := \gamma e^{\frac{2A}{x}} \int_{-x^{1+\epsilon}}^{x^{1+\epsilon}} e^{\frac{-xy^2+iy^3}{x^2(x^2+y^2)}} f_{err}(z)\,dy.$$

The first integral in (4.35) can be written

$$\gamma e^{\frac{2A}{x}} \int_{-x^{1+\epsilon}}^{x^{1+\epsilon}} e^{-\frac{y^2 A}{x^3}}\,dy = \gamma e^{\frac{2A}{x}}\sqrt{\frac{\pi x^3}{A}} + E_1 \tag{4.36}$$

where

$$|E_1| \leq \frac{\gamma}{A\sqrt{2}} x^{2-\epsilon} e^{\frac{2A}{x} - Ax^{2\epsilon-1}}. \tag{4.37}$$

For E_2, we further split the integral:

$$\begin{aligned} E_2 &= \gamma e^{\frac{2A}{x}} \int_{|y|\leq x^{1+\epsilon_2}} e^{-\frac{y^2 A}{x^3}} \left(e^{A\left(\frac{y^4+ixy^3}{x^3(x^2+y^2)}\right)} - 1 \right) dy \\ &\quad + \gamma e^{\frac{2A}{x}} \int_{x^{1+\epsilon_2}\leq |y| \leq x^{1+\epsilon}} e^{-\frac{y^2 A}{x^3}} \left(e^{A\left(\frac{y^4+ixy^3}{x^3(x^2+y^2)}\right)} - 1 \right) dy, \end{aligned}$$

with $\epsilon_2 > \epsilon$, $\epsilon_2 > \frac{1}{3}$. Then

$$\begin{aligned} &\left| \gamma e^{\frac{2A}{x}} \int_{|y|\leq x^{1+\epsilon_2}} e^{-\frac{y^2 A}{x^3}} \left(e^{A\left(\frac{y^4+ixy^3}{x^3(x^2+y^2)}\right)} - 1 \right) dy \right| \leq \\ &\qquad \gamma e^{\frac{2A}{x}} \left(\exp\left(Ax^{3\epsilon_2-1}\right) - 1 \right) \sqrt{\frac{\pi x^3}{A}} \end{aligned} \tag{4.38}$$

and

$$\begin{aligned} &\left| \gamma e^{\frac{2A}{x}} \int_{x^{1+\epsilon_2}\leq |y| \leq x^{1+\epsilon}} e^{-\frac{y^2 A}{x^3}} \left(e^{A\left(\frac{y^4+ixy^3}{x^3(x^2+y^2)}\right)} - 1 \right) dy \right| \leq \\ &\qquad \gamma \exp\left(\frac{2A}{x} - \frac{Ax^{\epsilon_2-2}}{1+x^{2\epsilon}} \right) x^3 \left(1 + x^{2\epsilon}\right) + \frac{\gamma x^3}{A} \exp\left(\frac{2A}{x} - Ax^{\epsilon_2-2} \right). \end{aligned} \tag{4.39}$$

Finally, for E_3, we have

$$|E_3| \leq \gamma e^{\frac{2A}{x}} \left| f_{err}^{\max} \right| \left(\pi x^3 \left(1 + x^{2\epsilon}\right) \right)^{\frac{1}{2}}. \tag{4.40}$$

To bound I_2, we apply Lemma 4.8 to find that

$$|I_2| \leq \sqrt{\frac{2\pi}{dx}} e^{-\eta\rho x^{2\epsilon-1}} (1 + f_2(\rho, x)) \exp\left(nx + \frac{A}{x} + \frac{1-d}{2} \log\alpha + f_1(\rho, x) \right). \tag{4.41}$$

Finally, we obtain

$$q_d(n) = \frac{A^{1/4}}{2\sqrt{\pi\alpha^{d-1}(d\alpha^{d-1}+1)}} n^{-3/4} \exp(2\sqrt{nA}) + E_1 + E_2 + E_3 + I_2,$$

where $|E_1 + E_2 + E_3 + I_2|$ is bounded using the expressions in (4.37) - (4.41). The result follows with $|r_d(n)| \leq |E_1| + |E_2| + |E_3| + |I_2|$.

4.1.3 Proof of Alder's Conjecture

Using Theorems 4.3 and 4.6, we are now able to prove our main results.

Proof of Theorem 4.1. Applying the results of Sections 4.1.1 and 4.1.2, we have that

$$\Delta_d(n) \quad = \quad q_d(n) - Q_d(n) = \frac{A^{1/4}}{2\sqrt{\pi\alpha^{d-1}(d\alpha^{d-1}+1)}} n^{-3/4}\exp(2\sqrt{nA}) + \mathcal{E}_d(n),$$

where $\mathcal{E}_d(n) = r_d(n) - Q_d(n)$. In the proof of Lemma 4.5, we relax the restriction on $y_{\max}$. Thus, Theorem 4.3 implies

$$Q_d(n) = O\left(\exp\left(\frac{2\pi}{\sqrt{3d+9}} n^{1/2} + c_0 n^{\frac{1}{6}}\right)\right),$$

where c_0 is some positive constant.

By Theorem 4.6, $|r_d(n)| \leq |E_1| + |E_2| + |E_3| + |I_2|$, and a careful examination of each of these terms shows that $E_1 = O\left(n^{-\frac{5}{6}} e^{2\sqrt{An}}\right)$, $E_2 = O\left(n^{-\frac{3}{2}\epsilon_2 - \frac{1}{4}} e^{2\sqrt{An}}\right)$, $E_3 = O\left(n^{-\frac{15}{16}} e^{2\sqrt{An}}\right)$, and $I_2 = O\left(n^{\frac{1}{4}} e^{2\sqrt{An} - \eta\rho x^{2\epsilon-1}}\right)$. Hence, by choosing $\epsilon_2 \geq \frac{7}{18}$, the result follows. □

Proof of Theorem 4.2. The works of Yee ([80],[81]) and Andrews [2] show that $\Delta_d(n) \geq 0$ when $d \geq 31$ and can be easily modified to show that the inequality is strict when $n \geq d+6$. For each remaining $4 \leq d \leq 30$, we use Theorems 4.3 and 4.6 to compute the smallest n such that our bounds imply $\Delta_d(n) > 0$. We denote this n by $\Omega(d)$, and a C++ program computed the values of $\Delta_d(n) \leq \Omega_d(n)$, which then confirmed the remaining cases of the Alder-Andrews Conjecture. As an example, we find that when $d = 30$, $\Omega(30) \leq 9.77 \cdot 10^6$. To get this, we take $\delta = 10^{-10}$ and $\epsilon_1 = 5 \cdot 10^{-11}$ in Theorem 4.3 and, in Theorem 4.6, $\epsilon = .16906$, $\epsilon_2 = .499999$, $\xi = .99$, $c = .375000001$, and $\nu = 1$. Other d are similar, and all satisfy $\Omega(d) \leq \Omega(30)$. □

4.2 A conjecture of Andrews

(The results in this section are joint with Marie Jameson.)

The first Göllnitz-Gordon identity states that the number of partitions of n into 2-distinct parts, with difference at least 4 between even parts, equals the number of partitions of n into parts congruent to $\pm 1, 4 \pmod 8$. Here, a d-distinct partition is defined to be a partition in which the difference between any two parts is at least d. In addition, an identity of Schur states that the number of partitions of n into 3-distinct parts, with difference at least 6 between multiples of 3, equals the number of partitions of n into parts congruent to $\pm 1 \pmod 6$.

It is natural to investigate whether this phenomenon has a generalization to further $d \geq 3$, and in this direction H.L. Alder [1] showed that if $d > 3$, the number of partitions of n into d-distinct parts where parts divisible by d differ by at least $2d$ is not equal to the number of partitions of n into parts taken from any set of integers whatsoever. G.E. Andrews considered a different generalization by considering the functions

$$q_d^*(n) := p(n|d\text{-distinct parts, no consecutive multiples of } d) \tag{4.42}$$

$$Q_d^*(n) := p(n|\text{parts} \equiv \pm 1, \pm(d+2) \pmod{4d}). \tag{4.43}$$

At a 2009 conference in Ottawa, he made the following conjecture[1] to accompany Alder's Conjecture (for more information on Alder's Conjecture, see Section 4.1, as well as [2], [4], [80], and [81]).

Conjecture (Andrews). For $d > 1$ and $n \geq 1$, we have that

$$q_d^*(n) - Q_d^*(n) \geq 0.$$

[1] A few days after the conference, in a private communication, he modified the conjecture. We are concerned with this modification.

Clearly, the conjecture holds for $d = 2$ and $d = 3$ by the Göllnitz-Gordon and Schur identities. Although the truth of this conjecture remains open, we show that the conjecture holds for sufficiently large n.

Theorem 4.10. *For fixed $d > 3$,*

$$\lim_{n\to\infty} (q_d^*(n) - Q_d^*(n)) = +\infty.$$

In addition to making the above conjecture, Andrews defined

$$Q_d^{**}(n) := p(n|\text{parts} \equiv \pm 1, \pm(d+2), \pm(d+6), \dots, \pm(d+4j+2) \pmod{4d}),$$

where $j = \lfloor (d-2)/4 \rfloor$, and wondered which values of d would yield

$$q_d^*(n) - Q_d^{**}(n) \geq 0. \tag{4.44}$$

Clearly, (4.44) implies the truth of Andrews's conjecture. Unfortunately, it is not true for all values of d (it fails, for example, when $d = 14$ and $n = 644$). We find which values of d cause (4.44) to hold (or fail) asymptotically.

Theorem 4.11. *Assuming the notation above, the following are true:*

1. *If $4 \leq d \leq 13$ or $d = 17$, and $d \neq 6$ or 10, then*

$$\lim_{n\to\infty} (q_d^*(n) - Q_d^{**}(n)) = +\infty.$$

2. *If $d = 14$ or 15, or $d \geq 18$, then*

$$\lim_{n\to\infty} (Q_d^{**}(n) - q_d^*(n)) = +\infty.$$

Remark. To establish Theorem 4.11, we show that the orders of $q_d^*(n)$ and $Q_d^{**}(n)$ are different.

Remark. Theorem 4.11 does not apply when $d = 6, 10,$ or 16. In these cases, we expect $q_d^*(n)$ to be asymptotically larger than $Q_d^{**}(n)$.

In the next section, to establish these results, we find an asymptotic expression for $q_d^*(n)$ by relating it to $q_d(n)$, where

$$q_d(n) := p(n|d\text{-distinct parts}).$$

Asymptotics for $q_d(n)$ are known, and they have been very helpful in proving Alder's Conjecture and its refinement by Andrews. We use these formulae to prove Theorem 4.10 in Section 4.2.1 and Theorem 4.11 in Section 4.2.1.

4.2.1 Proof of Andrews's conjecture in the limit

As stated above, rather than find an asymptotic formula for $q_d^*(n)$ directly, we instead compare it to the function $q_d(n)$ without the added difference condition between multiples of d. We speculate that relatively few d-distinct partitions of n have consecutive multiples of d, so we expect that $q_d^*(n) \asymp q_d(n)$. At present, this sort of relation is difficult to obtain. However, we note that

$$q_d^*(n) \geq q_{d+1}(n) \tag{4.45}$$

for all $d \geq 2$ and $n \geq 1$. From G. Meinardus [57] or from Theorem 4.6 (which corrects the proof), we have the asymptotic formula

$$q_d(n) \sim c_0 n^{-\frac{3}{4}} \exp\left(2\sqrt{A_d n}\right), \tag{4.46}$$

where c_0 is an explicit constant depending only on d,

$$A_d := \frac{d \log^2 \rho}{2} + \sum_{r=1}^{\infty} \frac{\rho^{rd}}{r^2}, \tag{4.47}$$

and $\rho = \rho_d$ is the unique root of $x^d + x - 1 = 0$ in the interval $[0, 1]$. Hence, (4.45) and (4.46) imply that

$$q_d^*(n) \gg n^{-\frac{3}{4}} \exp\left(2\sqrt{A_{d+1} n}\right). \tag{4.48}$$

Proof of Theorem 4.10

We must consider $Q_d^*(n)$. A result of V.V. Subrahmanyasastri (Theorem 10, [75]) yields the following asymptotic formula for $Q_d^*(n)$:

$$Q_d^*(n) \sim c_1 n^{-\frac{3}{4}} \exp\left(\pi\sqrt{\frac{2n}{3d}}\right),$$

where c_1 is an explicit constant. Recalling (4.48), our task is now to show that for all $d \geq 4$,

$$\pi\sqrt{\frac{2}{3d}} < 2\sqrt{A_{d+1}},$$

or, equivalently, that

$$\sqrt{2dA_{d+1}} > \frac{\pi}{\sqrt{3}}. \tag{4.49}$$

From (4.47), we have that

$$2dA_{d+1} \quad \geq \quad d^2 \log^2 \rho_{d+1},$$

and so we consider when

$$|d \log \rho_{d+1}| > \frac{\pi}{\sqrt{3}}. \tag{4.50}$$

Since $|d \log \rho_{d+1}|$ is increasing in d, one can verify that (4.50) holds for $d \geq 13$. A numerical computation verifies (4.49) in the remaining $4 \leq d \leq 12$.

Proof of Theorem 4.11

As in Section 4.2.1, Theorem 10 of [75] applies. In particular, we have that

$$Q_d^{**}(n) \sim c_2 n^{-\frac{3}{4}} \exp\left(\pi\sqrt{\frac{n\left\lfloor 2+\frac{d-2}{4}\right\rfloor}{3d}}\right)$$

for an explicit constant c_2. A numerical computation now shows that

$$\pi\sqrt{\frac{\left\lfloor 2+\frac{d-2}{4}\right\rfloor}{3d}} < 2\sqrt{A_{d+1}}$$

for $4 \leq d \leq 13$, $d \neq 6, 10$, and for $d = 17$. Hence, for these d, (4.48) implies that

$$\lim_{n\to\infty} (q_d^*(n) - Q_d^{**}(n)) = +\infty.$$

For the other values of $d \neq 6, 10$, or 16, instead of showing that

$$\lim_{n\to\infty} (Q_d^{**}(n) - q_d^*(n)) = +\infty,$$

we show the stronger statement that

$$\lim_{n\to\infty} (Q_d^{**}(n) - q_d(n)) = +\infty.$$

Hence, we must show that for these d,

$$\pi\sqrt{\frac{\lfloor 2 + \frac{d-2}{4} \rfloor}{3d}} > 2\sqrt{A_d}. \tag{4.51}$$

But

$$\pi\sqrt{\frac{\lfloor 2 + \frac{d-2}{4} \rfloor}{3d}} > \frac{\pi}{\sqrt{12}}$$

for all d and

$$2\sqrt{A_d} < \frac{\pi}{\sqrt{12}}$$

for $d \geq 26$. A numerical computation verifies (4.51) in the remaining cases.

Remark. One can check that (4.51) fails to hold when $d = 6, 10$, or 16. Our above speculation that $q_d^*(n) \asymp q_d(n)$ suggests that, for these values of d,

$$\lim_{n\to\infty} (q_d^*(n) - Q_d^{**}(n)) = +\infty,$$

although a stronger result along the lines of (4.48) would be needed to prove this.

4.3 Eta-quotients and theta functions

Jacobi's Triple Product Identity states that

$$\prod_{n=1}^{\infty}(1-x^{2n})(1+x^{2n-1}z^2)(1+x^{2n-1}z^{-2}) = \sum_{n=-\infty}^{\infty} z^{2m}x^{m^2}, \tag{4.52}$$

which is surprising because it gives a striking closed form expression for an infinite product. Using (4.52), one can derive many elegant q-series identities. For example, one has Euler's identity

$$q\prod_{n=1}^{\infty}(1-q^{24n}) = \sum_{k=-\infty}^{\infty}(-1)^k q^{(6k+1)^2} \tag{4.53}$$

and Jacobi's identity

$$\prod_{n=1}^{\infty}\frac{(1-q^{2n})^5}{(1-q^n)^2(1-q^{4n})^2} = \sum_{k=-\infty}^{\infty} q^{k^2}. \tag{4.54}$$

Both (4.53) and (4.54) can be viewed as identities involving Dedekind's eta-function $\eta(z)$, which is defined by

$$\eta(z) := q^{1/24}\prod_{n=1}^{\infty}(1-q^n), \tag{4.55}$$

where $q := e^{2\pi i z}$. It is well known that $\eta(z)$ is essentially a half-integral weight modular form, a fact which Dummit, Kisilevsky, and McKay [23] exploited to classify all the eta-products (functions of the form $\prod_{i=1}^{s}\eta(n_i z)^{t_i}$, where each n_i and each t_i is a positive integer) whose q-series have multiplicative coefficients. Martin [55] later obtained the complete list of integer weight eta-quotients (permitting the t_i to be negative) with multiplicative coefficients.

The right hand sides of both (4.53) and (4.54) also have an interpretation in terms of half-integral weight modular forms: they are examples of *theta functions*. Given a Dirichlet character ψ, the theta function $\theta_\psi(z)$ of ψ is given by

$$\theta_\psi(z) := \sum_{n}\psi(n)n^{\delta}q^{n^2}, \tag{4.56}$$

where $\delta = 0$ or 1 according to whether ψ is even or odd. The summation over n in (4.56) is over the positive integers, unless ψ is the trivial character, in which case the summation is over all integers. With this language, (4.53) becomes

$$\eta(24z) = \theta_{\chi_{12}}(z),$$

where $\chi_{12}(n) = \left(\frac{12}{n}\right)$ and $\left(\frac{\cdot}{\cdot}\right)$ is the Jacobi symbol. This fact is subsumed into the theorem of Dummit, Kisilevsky, and McKay, as $\eta(24z)$ is an eta-product and any theta function necessarily has multiplicative coefficients. However, we note that (4.54) is equivalent to

$$\frac{\eta(2z)^5}{\eta(z)^2\eta(4z)^2} = \theta_1(z),$$

which is covered neither by the theorem of Dummit, Kisilevsky, and McKay (as the left-hand side is a quotient of eta-functions, not merely a product), nor is it covered by the theorem of Martin (as the modular forms involved are of half-integral weight). It is therefore natural to ask which eta-quotients are theta functions.

Theorem 4.12. *1. The following eta-quotients are the only ones which are* *theta functions for an even character:*

$$\begin{aligned}
\frac{\eta(2z)^5}{\eta(z)^2\eta(4z)^2} &= \sum_{n=-\infty}^{\infty} q^{n^2}, \\
\frac{\eta(8z)\eta(32z)}{\eta(16z)} &= \sum_{n=1}^{\infty} \left(\frac{2}{n}\right) q^{n^2}, \\
\frac{\eta(16z)^2}{\eta(8z)} &= \sum_{n=1}^{\infty} \left(\frac{n}{2}\right)^2 q^{n^2}, \\
\frac{\eta(6z)^2\eta(9z)\eta(36z)}{\eta(3z)\eta(12z)\eta(18z)} &= \sum_{n=1}^{\infty} \left(\frac{n}{3}\right)^2 q^{n^2}, \\
\eta(24z) &= \sum_{n=1}^{\infty} \left(\frac{12}{n}\right) q^{n^2}, \\
\frac{\eta(48z)^3}{\eta(24z)\eta(96z)} &= \sum_{n=1}^{\infty} \left(\frac{24}{n}\right) q^{n^2}, \\
\frac{\eta(48z)\eta(72z)^2}{\eta(24z)\eta(144z)} &= \sum_{n=1}^{\infty} \left(\frac{n}{6}\right)^2 q^{n^2}, \\
\frac{\eta(24z)\eta(96z)\eta(144z)^5}{\eta(48z)^2\eta(72z)^2\eta(288z)^2} &= \sum_{n=1}^{\infty} \left(\frac{18}{n}\right) q^{n^2}.
\end{aligned}$$

2. The following eta-quotients are the only ones which are theta functions *for an odd character:*

$$\eta(8z)^3 = \sum_{n=1}^{\infty} \left(\frac{-4}{n}\right) nq^{n^2},$$
$$\frac{\eta(16z)^9}{\eta(8z)^3\eta(32z)^3} = \sum_{n=1}^{\infty} \left(\frac{-2}{n}\right) nq^{n^2},$$
$$\frac{\eta(3z)^2\eta(12z)^2}{\eta(6z)} = \sum_{n=1}^{\infty} \left(\frac{n}{3}\right) nq^{n^2},$$
$$\frac{\eta(48z)^{13}}{\eta(24z)^5\eta(96z)^5} = \sum_{n=1}^{\infty} \left(\frac{-6}{n}\right) nq^{n^2},$$
$$\frac{\eta(24z)^5}{\eta(48z)^2} = \sum_{n=1}^{\infty} \left(\frac{n}{12}\right) nq^{n^2}.$$

In fact, we establish a broader classification theorem. Given a positive integer m, let Θ_m^0 denote the linear span of the set of all theta functions associated to an even character ψ whose modulus is m together with its 'twists' by $\chi_{2,0}$, $\chi_{3,0}$, and $\chi_{6,0}$, where $\chi_{r,0}$ denotes the principal character modulo r, and let Θ_m^1 denote the analogous space associated to odd characters of modulus m. Here, the twist of a theta function associated to ψ by another character χ is the theta function associated to $\psi\chi$. For convenience, let Θ_m denote the union of Θ_m^0 and Θ_m^1. We call an element of Θ_m monic if its q-expansion has the form $1 + O(q)$ or $q + O(q^4)$.

Theorem 4.13. *1. The only eta-quotients which are monic elements of* Θ_m^0

for some m are those in Theorem 4.12 together with

$$\begin{aligned}
\frac{\eta(z)^2}{\eta(2z)} &= \sum_{n=-\infty}^{\infty} \left(1 - 2\left(\frac{n}{2}\right)^2\right) q^{n^2} = \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2},\\
\frac{\eta(z)\eta(4z)\eta(6z)^2}{\eta(2z)\eta(3z)\eta(12z)} &= \sum_{n=-\infty}^{\infty} \left(1 - \frac{3}{2}\left(\frac{n}{3}\right)^2\right) q^{n^2},\\
\frac{\eta(2z)^2\eta(3z)}{\eta(z)\eta(6z)} &= \sum_{n=-\infty}^{\infty} \left(1 - 2\left(\frac{n}{2}\right)^2 - \frac{3}{2}\left(\frac{n}{3}\right)^2 + 3\left(\frac{n}{6}\right)^2\right) q^{n^2},\\
\frac{\eta(8z)^5}{\eta(4z)^2\eta(16z)^2} &= \sum_{n=-\infty}^{\infty} \left(1 - \left(\frac{n}{2}\right)^2\right) q^{n^2},\\
\frac{\eta(9z)^2}{\eta(18z)} &= \sum_{n=-\infty}^{\infty} \left(1 - 2\left(\frac{n}{2}\right) - \left(\frac{n}{3}\right)^2 + 2\left(\frac{n}{6}\right)^2\right) q^{n^2},\\
\frac{\eta(18z)^5}{\eta(9z)^2\eta(36z)^2} &= \sum_{n=-\infty}^{\infty} \left(1 - \left(\frac{n}{3}\right)^2\right) q^{n^2},\\
\frac{\eta(4z)\eta(16z)\eta(24z)^2}{\eta(8z)\eta(12z)\eta(48z)} &= \sum_{n=-\infty}^{\infty} \left(1 - \left(\frac{n}{2}\right)^2 - \frac{3}{2}\left(\frac{n}{3}\right)^2 + \frac{3}{2}\left(\frac{n}{6}\right)^2\right) q^{n^2},\\
\frac{\eta(72z)^5}{\eta(36z)^2\eta(144z)^2} &= \sum_{n=-\infty}^{\infty} \left(1 - \left(\frac{n}{2}\right)^2 - \left(\frac{n}{3}\right)^2 + \left(\frac{n}{6}\right)^2\right) q^{n^2},\\
\frac{\eta(3z)\eta(18z)^2}{\eta(6z)\eta(9z)} &= \sum_{n=1}^{\infty} \left(2\left(\frac{n}{6}\right)^2 - \left(\frac{n}{3}\right)^2\right) q^{n^2},\\
\frac{\eta(8z)^2\eta(48z)}{\eta(16z)\eta(24z)} &= \sum_{n=1}^{\infty} \left(3\left(\frac{n}{6}\right)^2 - 2\left(\frac{n}{2}\right)^2\right) q^{n^2}.
\end{aligned}$$

2. The only eta-quotients which are monic elements of Θ^1_m for some m are those in Theorem 4.12 together with

$$\frac{\eta(6z)^5}{\eta(3z)^2} = \sum_{n=1}^{\infty} \left(2\left(\frac{n}{12}\right) - \left(\frac{n}{3}\right)\right) n q^{n^2}.$$

Remark. A theorem of Mersmann [58] classifying holomorphic eta-quotients implies that there are essentially only finitely many eta-quotients which could be in any Θ_m, even allowing non-monic elements. Unfortunately, while Mersmann's result can be made effective, the computations necessary to prove either case of Theorem 4.13 in this way would be prohibitively large. Consequently, our proof proceeds along fundamentally different lines. We note that Mersmann's result is slightly misquoted in [15] - the theorem credited to Mersmann on Page 30 of [15] is stronger than what he proves in his thesis.

Our proof proceeds as follows. Instead of using the method employed by Mersmann [58] - essentially a careful study of the order of vanishing of eta-quotients - we make use of the combinatorial properties of eta-quotients and the constraints on the q-series of theta functions. Combined with the theory of modular forms, in particular the Fricke involution $W_{k,M}$, asymptotic formulae, Eisenstein series, and Shimura's correspondence, the classification in Theorems 4.12 and 4.13 reduces to a case by case analysis. In this analysis, we make great use of the simple observation that if $a > b$, then $(1+q^a)(1+q^b) = 1 + q^b + O(q^a)$. In this regard, we also need the solution to a classical Diophantine problem.

4.3.1 Preliminary Facts

We begin by recalling some basic facts about modular forms. A holomorphic function $f\colon \mathbb{H} \to \mathbb{C}$ is a weakly holomorphic modular form of weight $k \in \frac{1}{2}\mathbb{Z}$ for the subgroup $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ if $f(\gamma z) = \epsilon_\gamma (cz+d)^k f(z)$ for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$, acting in the usual way by fractional linear transformation, where ϵ_γ is a suitable fourth root of unity. Moreover, we require for each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ that $(f \mid_k \gamma)(z) := (cz+d)^{-k} f(\gamma z)$ is represented by a Fourier series of the form

$$(f \mid_k \gamma)(z) = \sum_{n \geq n_0} a_\gamma(n) q_N^n,$$

where $q_N := e^{2\pi i/N}$. In fact, there are only finitely many such series required, one for each "cusp" of $\Gamma \setminus \mathbb{H}$, that is, an element $\rho \in \Gamma \setminus \mathbb{Q}$. In this case, we let $q_\rho := q_N$. The space of all weakly holomorphic modular forms of weight k on Γ is denoted by $M_k^!(\Gamma)$; its subspace consisting of all forms which are holomorphic (resp. vanishing) at the cusps is denoted by $M_k(\Gamma)$ (resp. $S_k(\Gamma)$) (these are the spaces of modular forms and cusp forms, respectively). For the subgroups we are concerned with, namely the congruence subgroups of level N,

$$\Gamma_1(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{N} \right\},$$

weakly holomorphic modular forms are fixed under the substitution $z \mapsto z+1$, and so they have a Fourier series at infinity with respect to the variable $q := e^{2\pi i z}$. Although we will briefly need the Fourier expansions at other cusps, it is this Fourier series (also called a q-expansion) that is of the most interest to us.

If $f(z) \in M_k^!(\Gamma_1(N))$, then $f(z)$ is said to be modular of level N. If $f(z)$ is holomorphic, then necessarily $k \geq 0$ and the space $M_k(\Gamma_1(N))$ decomposes naturally into two pieces: the previously mentioned cusp space $S_k(\Gamma_1(N))$ and the so-called Eisenstein space $\mathcal{E}_k(\Gamma_1(N))$. If $k \in \mathbb{Z}$, exploiting this decomposition, the size of the Fourier coefficients $a_f(n)$ of $f(z)$ is well-understood. In particular, letting $a_f(n) = a_{\text{cusp}}(n) + a_{\text{Eis}}(n)$, then we have that both

$$a_{\text{Eis}}(n) \ll_{f,\epsilon} n^{k-1+\epsilon},$$

due to the explicit nature of the coefficients (see (4.61) below), and

$$a_{\text{cusp}}(n) \ll_{f,\epsilon} n^{(k-1)/2+\epsilon},$$

which is the celebrated bound of Deligne. If $k \in 1/2 + \mathbb{Z}$, then the coefficients of both the Eisenstein series and the cuspidal part are not understood nearly as well, as both frequently encode values of L-functions. Nevertheless, polynomial bounds are known for each. In particular, we have the "trivial" bound that

$$a_{\mathrm{cusp}}(n) \ll_f n^{k/2},$$

valid for all $k \geq 1/2$ and

$$a_{\mathrm{Eis}}(n) \ll_{f,\epsilon} n^{k-1+\epsilon}$$

for $k \geq 3/2$, and $a_{\mathrm{Eis}}(n) \ll n^{\epsilon}$ if $k = 1/2$. Although stronger bounds are known (most recently, due to Blomer and Harcos [11]), it is only the fact that each is polynomially bounded that will be relevant to us. This is because if $f(z)$ is weakly holomorphic, but not holomorphic, then the coefficients of $f(z)$ are of a fundamentally different size. Namely, for n in certain arithmetic progressions depending on the level, they satsify

$$\log |a_f(n)| \gg n^{1/2}.$$

This is due, in various settings, to Rademacher and Zuckerman ([66], [67], [82], [83]), and, more recently, to Bringmann and Ono [14]. We will find this vast difference in size useful later on.

We recall that, given a Dirichlet character ψ of modulus r, the function $\theta_\psi(z)$ is defined by $\theta_\psi(z) := \sum_n \psi(n) n^\delta q^{n^2}$, where $\delta = 0$ or 1 according to whether ψ is even or odd. In both cases, the summation over n is assumed to be over the positive integers unless $r = 1$, in which case the sum is over all integers. It is classical that $\theta_\psi(z)$ is a modular form of weight $1/2$ if ψ is even and of weight $3/2$ if ψ is odd. Each $\theta_\psi(z)$ is of level $4r^2$ and, moreover, if $r \neq 1$, then it is a cusp form. Regardless of the parity of ψ, we refer to $\theta_\psi(z)$ as a theta function of modulus r.

The twist of a theta function $\theta_\psi(z)$ by a character χ is the theta function associated to $\psi\chi$. Given a positive integer m, we let Θ_m^0 denote the linear span of the set of all weight $1/2$ theta functions whose moduli are m and their twists by $\chi_{2,0}$, $\chi_{3,0}$, and $\chi_{6,0}$, and we let Θ_m^1 denote the analogous space for weight 3/2 theta functions. Let Θ_m be the union of Θ_m^0 and Θ_m^1.

Dedekind's eta-function $\eta(z)$ is defined by

$$\eta(z) := q^{1/24} \prod_n (1 - q^n).$$

It is almost a modular form of weight 1/2 on $SL_2(\mathbb{Z})$, in the sense that

$$\eta\left(\frac{-1}{z}\right) = (-iz)^{1/2}\eta(z), \tag{4.57}$$

but it fails to transform suitably under $z \mapsto z+1$. However, since $\eta(z)$ is non-vanishing away from the cusps, a function of the form

$$f(z) = \prod_{d|N} \eta(dz)^{r_d} \tag{4.58}$$

will be a weakly holomorphic modular form on $\Gamma_1(24N)$ if $\sum_{d|N} dr_d \equiv 0 \pmod{24}$. This level may not be sharp, in the sense that $f(z)$ may be a weakly holomorphic modular form on $\Gamma_1(M)$ for some proper divisor M of $24N$, but what is important for our purposes is that the only primes dividing the level of $f(z)$ are those dividing N together with 2 and 3. We call a function of the form (4.58) satisfying this condition an eta-quotient. The order of vanishing of $f(z)$ at the cusp $\rho := \frac{\alpha}{\delta}$ is given by [64, Theorem 1.65]

$$\mathrm{ord}_{z=\rho} f(z) = \frac{N}{24} \sum_{d|N} \frac{(d,\delta)^2 r_d}{(\delta, \frac{N}{\delta}) d\delta}. \tag{4.59}$$

Lemma 4.14. *Suppose that $f(z) = \prod_{d|N} \eta(dz)^{r_d}$ is an element of Θ_m. Set $a := 1 + \max(1, \nu_2(m))$ and $b := \max(1, \nu_3(m))$, where $\nu_p(\cdot)$ is the standard p-adic valuation, and let m_0 be the maximal divisor of m coprime to 6. Then $r_d = 0$ for $d \nmid 2^{2a}3^{2b}m_0^2$.*

Proof. Given a theta function $\theta(z)$ of modulus r and weight k, it is well known that $\theta(z) \mid W_{k,4r^2} := (-2rz)^{-k}\theta\left(\frac{-1}{4r^2 z}\right)$ is again a modular form of weight k whose Fourier series has integral exponents, where $W_{k,4r^2}$ is the usual Fricke involution (see [64, Proposition 3.8]). This property also holds for $\theta(z)|W_{k,4r^2t}$ for any t, and so we see that the operator $W_{k,2^{2a}3^{2b}m_0^2}$ sends Θ_m to the union of two spaces of modular forms whose Fourier series have integral exponents. If $f(z)$ is in Θ_m and has weight k, therefore, we must have that $f(z)|W_{k,2^{2a}3^{2b}m_0^2}$ is a modular form of weight k with only integral exponents.

On the other hand, we compute using (4.57) that

$$\begin{aligned} f(z)|W_{k,2^{2a}3^{2b}m_0^2} &= (-2^a3^bm_0z)^{-k}\prod_{d|N}\eta\left(\frac{-d}{2^{2a}3^{2b}m_0^2z}\right)^{r_d} \\ &= C\prod_{d|N}\eta\left(\frac{2^{2a}3^{2b}m_0^2z}{d}\right)^{r_d} =: C\tilde{f}(z) \end{aligned}$$

for some constant C. But the only way for $\tilde{f}(z)$ to have a Fourier series with integral exponents is if for each $d \nmid 2^{2a}3^{2b}m_0^2$ we have that $r_d = 0$. $\square$

The above lemma limits the eta-quotients which are in Θ_m for a fixed m, but we still need a way to control the possible values of m. The following proposition permits us to do that. First, though, we fix notation. Given a weakly holomorphic modular form $f(z) = \sum_{n\gg-\infty} a(n)q^n$ of level N, the U_p operator for a prime p is defined by

$$f(z)|U_p := \sum_{n\gg-\infty} a(pn)q^n.$$

It is well known that $f(z)|U_p$ is again a weakly holomorphic modular form of level N if $p \mid N$ and level pN if $p \nmid N$.

Proposition 4.15. *If $f \in M_k^!(\Gamma_1(N))$ and $p \nmid N$ is prime, then $f(z)|U_p = 0$ if and only if $f(z) = 0$.*

Before proving Proposition 4.15, we deduce its application to the problem at hand.

Corollary 4.16. *If $f(z) = \prod_{d|N} \eta(dz)^{r_d}$ is in Θ_m for some m, then the only primes dividing m are 2 and 3.*

Proof. We first show that any eta-quotient $f(z)$ is not annihilated by the U_p operator for any $p \geq 5$. Fix such a prime and write $f(z) = f_1(z)f_2(pz)$, with

$$\begin{aligned} f_1(z) &:= \prod_{d|N, p\nmid d} \eta(dz)^{r_d}, \text{ and} \\ f_2(z) &:= \prod_{d|N, p|d} \eta\left(\frac{dz}{p}\right)^{r_d}, \end{aligned}$$

where an empty product has value 1. We now have that $f(24z)|U_p = f_1(24z)|U_p \cdot f_2(24z)$. Since $f_1(24z)$ is a weakly holomorphic modular form of level indivisible by p, we see by Proposition 4.15 that $f_1(24z)|U_p$ is non-zero, and since $f_2(24z) \neq 0$, we also have that $f(24z)|U_p$ is non-zero. Since $p \nmid 24$ by assumption, it follows that that $f(z)|U_p \neq 0$.

On the other hand, functions in Θ_m have the property that their coefficients are supported on exponents which are coprime to m. Hence, for any prime divisor p of m, we must have that U_p annihilates Θ_m. Thus, if $f(z)$ is in Θ_m, the only way it can have this property is if the only primes dividing m are 2 and 3. □

The proof of Proposition 4.15 relies upon a lemma on sums of almost-everywhere multiplicative functions, which we define to be functions satisfying $f(mn) = f(m)f(n)$ for any coprime m and n, neither of which is divisible by any of a finite set of primes called the *bad primes*. As an example, if $f(n)$ is a multiplicative function such that $f(t) \neq 0$ for some t, then $f(tn)/f(t)$ is not generically multiplicative. It is, however, almost-everywhere multiplicative away from the primes dividing t.

Lemma 4.17. *Suppose that $f_1, \cdots, f_s$ are almost-everywhere multiplicative functions which are each non-zero for an infinite set of primes. Moreover,* *assume that, for each $i \neq j$, $f_i(p) \neq f_j(p)$ for an infinite number of primes p. If $c_1 f_1(n) + \ldots + c_s f_s(n) = 0$ for all n indivisible by every bad prime, then each $c_i = 0$.*

Proof. We proceed by induction, noting that the result is obviously true if $s = 1$.

If $s \geq 2$, we may assume by way of contradiction that each $c_i \neq 0$, so we have that

$$f_s(n) = -\sum_{i=1}^{s-1} \frac{c_i}{c_s} f_i(n) \tag{4.60}$$

for every n not divisible by any bad prime. Let m and n be coprime integers not divisible by any bad prime. We then have that both

$$f_s(mn) = -\sum_{i=1}^{s-1} \frac{c_i}{c_s} f_i(m) f_i(n)$$

and

$$\begin{aligned} f_s(mn) &= \left(-\sum_{i=1}^{s-1} \frac{c_i}{c_s} f_i(m)\right)\left(-\sum_{i=1}^{s-1} \frac{c_i}{c_s} f_i(n)\right) \\ &= \sum_{i=1}^{s-1} \left(\frac{c_i}{c_s} \sum_{j=1}^{s-1} \frac{c_j}{c_s} f_j(m)\right) f_i(n). \end{aligned}$$

Equating these two expressions for $f_s(mn)$, we obtain that

$$\sum_{i=1}^{s-1} \left(\frac{c_i}{c_s} f_i(m) + \frac{c_i}{c_s} \sum_{j=1}^{s-1} \frac{c_j}{c_s} f_j(m)\right) f_i(n) = 0,$$

which, by our induction hypothesis, can only happen if for each i and m, we have that

$$\frac{c_i}{c_s}\left(f_i(m) + \sum_{j=1}^{s-1} \frac{c_j}{c_s} f_j(m)\right) = 0.$$

Since $c_i \neq 0$ for each i, we then get a linear combination of $f_j(m)$ equaling 0, and again using the induction hypothesis, we find that $c_i = -c_s$ and all other $c_j = 0$. Since we assumed that each $c_j \neq 0$, this can only happen if $s = 2$, and in that case, (4.60) yields that $f_1(n) = f_2(n)$ for all n away from the set of bad primes. But since these functions were assumed to be distinct, this cannot happen. $\square$

Proof of Proposition 4.15. If the Fourier expansion of $f(z)$ at the cusp ρ is given by

$$f(z) = \sum_{n \gg -\infty} a_\rho(n) q_\rho^{n+\kappa_\rho},$$

then the principal part of $f(z)$ at ρ is

$$f_\rho^-(z) = \sum_{n+\kappa_\rho<0} a_\rho(n) q_\rho^{n+\kappa_\rho}.$$

Following either the classical work of Rademacher and Zuckerman ([66], [67], [82], [83]) or the recent work of Bringmann and Ono [14], we can write $f(z) = f^-(z) + f_{\mathrm{hol}}(z)$, where $f^-(z)$ is a linear combination of so-called Maass-Poincaré series which matches the principal part of $f(z)$ at each cusp and $f_{\mathrm{hol}}(z)$ is a holomorphic modular form. Since the coefficients of the Maass-Poincaré series grow superpolynomially along certain arithmetic progressions modulo N and the coefficients of $f_{\mathrm{hol}}(z)$ are polynomially bounded (see the above discussion), in order for $f(z)|U_p$ to be 0, we must have that $f^-(z) = 0$ and $f(z) = f_{\mathrm{hol}}(z)$. In the case that $k < 0$, we are now done, as there are no holomorphic modular forms of negative weight. If $k = 0$, the only holomorphic modular forms are constant and are preserved under U_p. Hence, in this case too, we must have that $f(z) = 0$.

If $k = 1/2$, a deep theorem of Serre and Stark [71] states that $f(z)$ must be a linear combination of weight 1/2 theta functions $\theta_\chi(z)$ of level dividing N and their dilates $\theta_\chi(tz)$, with $t \cdot \mathrm{cond}(\chi) \mid N$. Thus, if $(n, N) = 1$, we have that

$$0 = a_f(p^2n^2) = \sum_\chi c_\chi \chi(p) \chi(n),$$

where the sum runs over the characters of conductor dividing N, and each c_χ is a constant. But by the linear independence of characters, we must have that each $c_\chi \chi(p) = 0$, whence $c_\chi = 0$ since $(p, \mathrm{cond}(\chi)) = 1$. Considering iteratively $a_f(tp^2n^2)$ in the same way, the result follows if $k = 1/2$.

We may now suppose that $k \geq 1$. In the case that k is an integer, following [21], a basis for the Eisenstein space of $M_k(\Gamma_1(N))$ is given by

$$\{E_k^{\varepsilon,\psi,t}(z) : (\varepsilon, \psi, t) \in A_{k,N}\},$$

where we define $E_k^{\varepsilon,\psi,t}(z)$ using the series

$$E_k^{\varepsilon,\psi}(z) = c_{k,\varepsilon,\psi} + \sum_{n=1}^{\infty} \left(\sum_{d|n} \varepsilon(n/d)\psi(d)d^{k-1} \right) q^n \tag{4.61}$$

by $E_k^{\varepsilon,\psi,t}(z) = E_k^{\varepsilon,\psi}(tz)$ for $k \neq 2$ or $(\varepsilon, \psi) \neq (1,1)$, and $E_2^{1,1,t}(z) = E_2^{1,1}(z) - tE_2^{1,1}(tz)$ for $t \neq 1$. In the above, $c_{k,\varepsilon,\psi}$ is a constant and $A_{k,N}$ is the set of triples (ε, ψ, t) where ε and ψ are primitive Dirichlet characters of conductor u and v, respectively, with $(\varepsilon\psi)(-1) = (-1)^k$, and t is a positive integer such that $tuv \mid N$. In the case $k = 2$ we exclude the triple $(1,1,1)$, and in the case $k = 1$, we require the first two elements of a triple to be unordered. We note that the Fourier coefficients of the Eisenstein series $E_k^{\varepsilon,\psi}(z)$ are multiplicative; we denote these coefficients by $\sigma_{k-1}^{\varepsilon,\psi}(n)$.

For the cusp space $S_k(\Gamma_1(N))$, we may choose a basis of Hecke eigenforms, so that any holomorphic modular form is a linear combination of forms $g_1(z), \cdots, g_s(z)$, each with Fourier coefficients $a_1(n), \cdots, a_s(n)$ that are "essentially" multiplicative: in the case that $a_i(n)$ arises from a newform (that is, a form not coming from some level $M \mid N$), it is legitimately multiplicative, but if $a_i(n)$ arises from a non-newform, then it is of the form $a_i(tn) = a_j(n)$ for some $t \mid N$ and the coefficients $a_j(n)$ of some newform (the role of t will be handled easily in the proof below). In particular, if

$f(z) = \sum a(n)q^n$, we may, for $(n, N) = 1$, write

$$a(n) = c_1 f_1(n) + \cdots c_r f_r(n)$$

for some constants $c_1, \cdots, c_r$ and multiplicative $f_i(n)$, coming either from an Eisenstein series or an eigenform. Since $f(z)|U_p = 0$, we must have that $a(pn) = 0$ for all n. In particular, for $(n, pN) = 1$, we must have that

$$0 = c_1 f_1(pn) + \cdots + c_r f_r(pn) = c_1 f_1(p) f_1(n) + \cdots + c_r f_s(p) f_r(n),$$

and we may omit any $f_i(n)$ arising from an Eisenstein series $E_k^{\varepsilon,\psi}(tz)$ with $t > 1$ or from a non-newform (since $(n, N) = 1$ and $t \mid N$, the omitted coefficients are 0 and will have no effect on $a(pn)$). By Lemma 4.17, we must have that each $c_i f_i(p) = 0$. Now, $f_i(p)$ may be zero for some i, but in that case we necessarily have that $f_i(p^2) \neq 0$ (this follows from the Euler product expansion in the case of a cusp form and by a direct computation in the case of Eisenstein series, which can only have this property if $k = 1$). Thus, by also considering $a(p^2 n)$, we see that each c_i not arising from an $E_k^{\varepsilon,\psi}(tz)$ or a non-newform must be 0. Iteratively letting t be the smallest divisor of N not yet considered and repeating the argument above for $a(tpn)$ and $a(tp^2n)$, we see that all $c_i = 0$, whence $f(z) = 0$ identically.

In the case that the weight is half-integral, we may still choose as a basis of the cusp space a sequence of Hecke eigenforms, but we no longer know that there is a basis of the Eisenstein space with multiplicative Fourier coefficients. We proceed, therefore, to show that $f(z)$ is a cusp form. The argument in the integer weight case then applies, showing that $f(z) = 0$.

Consider the image of $f(z)$ under the Shimura map $S_{\lambda,\tau} : M_{\lambda+\frac{1}{2}}(\Gamma_1(N)) \to M_{2\lambda}(\Gamma_1(N))$, where $\lambda := k - 1/2$ and τ is any squarefree positive integer (often, this map is only defined for the cusp-space; see, for example, work of Jagathesan and Manickam [46] on extensions of this). If $f(z) = \sum a(n)q^n$, the non-constant terms of $S_{\lambda,\tau}(f(z)) = \sum b(n)q^n =: F(z)$ are given by the

Dirichlet series formula [64, Theorem 3.14]

$$\sum_{n=1}^{\infty} b(n)n^{-s} = L(s+1-\lambda, \chi_\tau) \sum_{n=1}^{\infty} a(\tau n^2)n^{-s},$$

where χ_τ is a Dirichlet character. Hence, we have that

$$b(n) = \sum_{d|n} d^{\lambda-1}\chi_\tau(d)a\left(\frac{\tau n^2}{d^2}\right).$$

In particular if $(p,n)=1$, then

$$\begin{aligned} b(p^m n) &= \sum_{d|p^m n} d^{\lambda-1}\chi_\tau(d)a\left(\frac{\tau p^{2m}n^2}{d^2}\right) \\ &= \sum_{d|n} (p^m d)^{\lambda-1}\chi_\tau(p^m d)a\left(\frac{\tau n^2}{d^2}\right) \\ &= p^{m(\lambda-1)}\chi_\tau(p^m)b(n), \end{aligned}$$

since $a(pr)=0$ for any r. Let $F_0(z)$ denote the projection of $F(z)$ into the Eisenstein space of $M_{2\lambda}(\Gamma_1(N))$. We can express $F_0(z)$ as

$$F_0(z) = \sum_{(\varepsilon,\psi,t)} a_{\varepsilon,\psi,t} E_{2\lambda}^{\varepsilon,\psi}(tz), \tag{4.62}$$

where we are summing over all triples (ε,ψ,t) such that ε and ψ are primitive characters of conductor dividing N and t is any divisor of N (thus, we include triples (ε,ψ,t) which do not arise in $A_{2\lambda,N}$). We require that $a_{\varepsilon,\psi,t}=0$ when $(\varepsilon,\psi,t)\notin A_{2\lambda,N}$, unless $\lambda=1$, where we let $a_{1,1,1}$ absorb the coefficients of $E_2^{1,1}(z)$ arising from the terms $E_2^{1,1,t}(z)$ in the basis expansion of the Eisenstein space of $M_2(\Gamma_1(N))$. We then have that, if $(n,pN)=1$,

$$b(p^m n) = \sum_{\varepsilon,\psi} a_{\varepsilon,\psi,1}\sigma_{2\lambda-1}^{\varepsilon,\psi}(p^m n) + O\left((p^m n)^{\lambda-1/2+\epsilon}\right)$$

by Deligne's bound. Similarly, we also have for such n that

$$b(n) = \sum_{\varepsilon,\psi} a_{\varepsilon,\psi,1}\sigma_{2\lambda-1}^{\varepsilon,\psi}(n) + O\left(n^{\lambda-1/2+\epsilon}\right).$$

Since $b(p^m n) = p^{m(\lambda-1)}\chi_\tau(p^m)b(n)$, we must have that

$$\sum_{\varepsilon,\psi} \tilde{a}_{\varepsilon,\psi}\sigma_{2\lambda-1}^{\varepsilon,\psi}(n) = O\left((p^m n)^{\lambda-1/2+\epsilon}\right), \tag{4.63}$$

where

$$\tilde{a}_{\varepsilon,\psi} = a_{\varepsilon,\psi,1}\cdot(\sigma_{2\lambda-1}^{\varepsilon,\psi}(p^m) - p^{m(\lambda-1)}\chi_\tau(p^m)).$$

Let $n = \ell_1\ell_2$, where ℓ_1 and ℓ_2 are large primes such that $\ell_2 \asymp \ell_1^{1/2}$. Then (4.61) implies that

$$\begin{aligned}
\sigma_{2\lambda-1}^{\varepsilon,\psi}(\ell_1\ell_2) &= (\psi(\ell_1)\ell_1^{2\lambda-1} + \varepsilon(\ell_1))\cdot(\psi(\ell_2)\ell_2^{2\lambda-1} + \varepsilon(\ell_2)) \\
&= \psi(\ell_1)\psi(\ell_2)(\ell_1\ell_2)^{2\lambda-1} + \psi(\ell_1)\varepsilon(\ell_2)\ell_1^{2\lambda-1} + O\left(\ell_1^{\lambda-1/2}\right).
\end{aligned}$$

Using this in (4.63) and dividing by $(\ell_1\ell_2)^{2\lambda-1}$, we see that

$$\sum_{\varepsilon,\psi} \tilde{a}_{\varepsilon,\psi}\psi(\ell_1)\psi(\ell_2) = O\left(\ell_1^{-\lambda+1/2} + p^{m(\lambda-1/2)+\epsilon}\ell_1^{-\lambda+1/2+\epsilon}\right),$$

and letting ℓ_1 and ℓ_2 tend to infinity along fixed arithmetic progressions modulo N, we see that, in fact,

$$\sum_{\varepsilon,\psi} \tilde{a}_{\varepsilon,\psi}\psi(\ell_1)\psi(\ell_2) = 0.$$

Hence, (4.63) and the expansion of $\sigma_{2\lambda-1}(\ell_1\ell_2)$ now yield that

$$\sum_{\varepsilon,\psi} \tilde{a}_{\varepsilon,\psi}\psi(\ell_1)\varepsilon(\ell_2) = O\left(\ell_1^{-\lambda+1/2} + p^{m(\lambda-1)+\epsilon}\ell_1^{-\frac{1}{2}(\lambda-1/2)+\epsilon}\right),$$

and again letting ℓ_1 and ℓ_2 tend to infinity along fixed arithmetic progressions, we see that

$$\sum_{\varepsilon,\psi} \tilde{a}_{\varepsilon,\psi}\psi(\ell_1)\varepsilon(\ell_2) = 0. \tag{4.64}$$

Since ℓ_1 and ℓ_2 were chosen to be in arbitrary arithmetic progressions modulo N, this can be viewed as an equation in terms of the matrix $K_N \otimes K_N$, where

K_N is the $\phi(N) \times \phi(N)$ matrix whose components are $\chi(a)$ as a runs over elements of $(\mathbb{Z}/N\mathbb{Z})^\times$ and χ runs over its characters. Since K_N is invertible (it is the tensor product of Vandermonde matrices arising from the cyclic factors of $(\mathbb{Z}/N\mathbb{Z})^\times$), $K_N \otimes K_N$ is as well. Hence, the only way for (4.64) to hold is if each $\tilde{a}_{\varepsilon,\psi} = 0$. Recall that

$$\tilde{a}_{\varepsilon,\psi} = a_{\varepsilon,\psi,1} \cdot (\sigma_{2\lambda-1}^{\varepsilon,\psi}(p^m) - p^{m(\lambda-1)}\chi_\tau(p^m)),$$

and since $\sigma_{2\lambda-1}^{\varepsilon,\psi}(p^m) \asymp p^{m(2\lambda-1)}$, by considering large enough m, we conclude that each $a_{\varepsilon,\psi,1} = 0$. By iteratively letting t be the smallest divisor of N not yet considered and looking at $b(tp^m n)$, the above argument shows that each $a_{\varepsilon,\psi,t} = 0$. Consequently, we must have that $F_0(z) = 0$ and $F(z)$ is a cusp form. But since this is true independent of the choice of τ in the map $S_{\lambda,\tau}$, we also have that $f(z)$ is itself a cusp form. We now proceed as in the integer-weight case. Strictly speaking, the coefficients of half-integer weight eigenforms are only almost-everywhere multiplicative in square classes (that is, for t squarefree, $a(tn^2)/a(t)$ is multiplicative in n, provided that $(n, N) = 1$), but by considering each square class separately, the result follows. $\square$

Before we can prove Theorem 4.13, we need one further lemma.

Lemma 4.18. *The only $a = 2^i 3^j$ which are one less than a square are $a = 3, 8, 24, 48,$ and 288.*

Proof. Suppose $a = n^2 - 1 = (n+1)(n-1)$. If a is odd, then we must have that both $n-1$ and $n+1$ are powers of 3, and so $a = 3$. If a is not divisible by 3, then both $n-1$ and $n+1$ must be powers of 2, and so $a = 8$. Lastly, suppose that a is divisible by both 2 and 3. Since both $n-1$ and $n+1$ are then required to be even, the pair $\left(\frac{n-1}{2}, \frac{n+1}{2}\right)$ must be either $(2^{i-2}, 3^j)$ or $(3^j, 2^{i-2})$. It is a classical result due to Levi ben Gerson that the only powers of 2 and 3 which differ by one are $(2,3)$, $(3,4)$, and $(8,9)$. These lead to $a = 24$, 48, and 288, respectively. $\square$

Remark. Of course, by Mihăilescu's resolution of Catalan's conjecture [59] it is known that 8 and 9 are the only consecutive perfect powers.

4.3.2 Proof of Theorem 4.13

We begin by considering a few concrete cases of Theorem 4.12. Although we could prove Theorem 4.13 without doing so, this allows us to illustrate the constructive approach we shall take.

Lemma 4.14 and Corollary 4.16 together tell us that if $f(z) = \prod_{d|N} \eta(dz)^{r_d}$ is an eta-quotient which is also a theta function $\theta_\psi(z)$ of modulus r, then r must be divisible only by the primes 2 and 3, and we may take $N = 4r^2$. Since the coefficients of any eta-quotient are real, we must also have that ψ is a quadratic character. The key observation which leads to Theorem 4.12 is that if we know the modulus of a quadratic theta function, then we know the first few terms in its Fourier series, at least up to a sign.

First, we consider whether $\theta_1(z) = 1 + 2\sum_{n=1}^{\infty} q^{n^2}$ is an eta-quotient. Let $\eta_0(z) := \prod_{n=1}^{\infty}(1-q^n)$, so that $f(z) = q^{a_f} \prod_{d|N} \eta_0(dz)^{r_d}$, where $a_f = \sum_{d|N} \frac{dr_d}{24}$. In order for $f(z)$ to be equal to $\theta_1(z)$, we must have that $a_f = 0$, as $\eta_0(dz) = 1 + O(q^d)$. In fact, we know more:

$$\eta_0(dz)^{r_d} = 1 - r_d q^d + O(q^{2d}). \tag{4.65}$$

Consequently, in order for the Fourier expansion of $f(z)$ to match that of

$$\theta_1(z) = 1 + 2q + 2q^4 + 2q^9 + O(q^{16}),$$

we must have that $r_1 = -2$ or, equivalently, that -2 is the exact power of $\eta(z)$ dividing $f(z)$ (we say that $\eta(z)^{-2}$ divides $f(z)$). The Fourier series of $\eta_0(z)^{-2}$ is given by

$$\eta_0(z)^{-2} = 1 + 2q + 5q^2 + O(q^3),$$

and since there is no q^2 term in the Fourier expansion of $\theta_1(z)$, we see that $\eta(2z)^5$ must also divide $f(z)$ in order to cancel the $5q^2$ term in the Fourier series for $\eta_0(z)^{-2}$. This leads us to consider

$$\frac{\eta_0(2z)^5}{\eta_0(z)^2} = 1 + 2q - 4q^5 + O(q^6).$$

Since we now need to add $2q^4$ to this Fourier expansion, we see that $\eta(4z)^{-2}$ must also divide $f(z)$. We compute that

$$\frac{\eta_0(2z)^5}{\eta_0(z)^2\eta_0(4z)^2} = 1 + 2q + 2q^4 + 2q^9 + O(q^{16}),$$

which matches perfectly the Fourier expansion of $\theta_1(z)$! Since for $f(z) = \frac{\eta(2z)^5}{\eta(z)^2\eta(4z)^2}$, we have that $a_f = 0$, $f(z)$ is a candidate for an eta-quotient representation of $\theta_1(z)$. One easily verifies via (4.59) that $f(z)$ is holomorphic, and then the Sturm bound [64, Theorem 2.58] implies that $f(z) = \theta_1(z)$.

We now suppose that ψ is a character whose modulus is a positive power of 2 and that the weight of $\theta_\psi(z)$ is 1/2. The Fourier series of $\theta_\psi(z)$ must start as

$$\theta_\psi(z) = q \pm q^9 \pm q^{25} \pm q^{49} + O(q^{81}).$$

Consequently, with the notation from before, we must have that $a_f = 1$, and from (4.65), that the term $\eta(dz)^{r_d}$ dividing $f(z)$ with smallest d must be either $\eta(8z)$ or $\eta(8z)^{-1}$. We consider only the first case before turning to the general situation of Theorem 4.13.

The Fourier expansion of $\eta_0(8z)$ is given by

$$\eta_0(8z) = 1 - q^8 - q^{16} + O(q^{40}).$$

As before, we now see that $\eta(16z)^{-1}$ must divide $f(z)$, and

$$\frac{\eta_0(8z)}{\eta_0(16z)} = 1 - q^8 - q^{24} + q^{32} + O(q^{40}).$$

Since the modulus of ψ is a power of 2, the proof of Lemma 4.14 implies that we cannot change the coefficient of q^{24}, as it would require the level to be divisible by 3. Of course, this is acceptable, as the coefficient of q^{25} in the expansion of $\theta_\psi(z)$ is permitted to be -1. Continuing, we see that $\eta(32z)$ must divide $f(z)$, which leads us to

$$\frac{\eta_0(8z)\eta_0(32z)}{\eta_0(16z)} = 1 - q^8 - q^{24} + q^{48} + q^{80} - q^{120} - q^{168} + O(q^{224}),$$

a very promising Fourier series. We compute that for $f(z) = \frac{\eta(8z)\eta(32z)}{\eta(16z)}$, we have that $a_f = 1$, $f(z)$ is holomorphic, and $f(z) = \theta_{\chi_8}(z)$ where $\chi_8(\cdot) = \left(\frac{2}{\cdot}\right)$ is a primitive character of conductor 8.

A priori it is conceivable that there are other theta functions expressible as eta-quotients divisible by $\eta(8z)$. However, since we have now reached a Fourier series whose exponents are supported on the squares, if $f(z)$ were additionally divisible by some $\eta(dz)^{r_d}$ with $d > 32$ chosen minimally, then we must have both that d is a power of 2, based on the level, and that d is one less than a square, based on its effect on the Fourier series. But Lemma 4.18 implies that there are no such d, so we have produced the only eta-quotient divisible by $\eta(8z)$ which is a theta function.

We now turn to the proof of Theorem 4.13, assuming that $f(z) = \prod \eta(dz)^{r_d}$ is in Θ_m for some m whose only prime factors are 2 and 3. Under the assumption that $f(z)$ is monic, we must consider two cases: either the Fourier series of $f(z)$ has the form $1 + O(q)$ or it has the form $q + O(q^4)$.

In the first case, the Fourier expansion has the form $1 + O(q)$ and we must have that $m = 1$ since $f(z)$ is a linear combination of theta functions and the only theta function whose Fourier expansions begins in this fashion is $\theta_1(z)$. Lemma 4.14 then tells us that we are limited to factors $\eta(dz)$ with $d \mid 144$. Unfortunately, we must now split into further cases, depending on the location of the first non-zero coefficient after the constant term.

If the Fourier series begins $1 - aq + O(q^2)$ with $a \neq 0$, then we must have that $\eta(z)^a$ is a factor of $f(z)$. We have that

$$\eta_0(z)^a = 1 - aq + \frac{a(a-3)}{2}q^2 + O(q^3),$$

and consequently, we must have that $\eta(2z)^{a_2}$ divides $f(z)$, where $a_2 = \frac{a(a-3)}{2}$. We then see that $\eta(3z)^{a_3}$ must also divide $f(z)$, where $a_3 = \frac{a^3-4a}{3}$. The coefficient of q^4 can be arbitrary, however, so we let $\eta(4z)^b$ denote the power of $\eta(4z)$ dividing $f(z)$. We then compute the coefficient of q^5 in $\eta_0(z)^a\eta_0(2z)^{a_2}\eta_0(3z)^{a_3}\eta_0(4z)^b$ to be

$$a\left(b - \frac{1}{20}a^4 + \frac{3}{4}a^2 - \frac{1}{2}a - \frac{6}{5}\right).$$

This coefficient must be 0, since we cannot 'fix' it with some $\eta(dz)$ with $d \mid 144$. Since a was assumed to be non-zero, we see that b is in fact not allowed to be arbitrary. We let a_4 be the required value, namely $\frac{1}{20}a^4 - \frac{3}{4}a^2 + \frac{1}{2}a + \frac{6}{5}$. We then continue as before, seeing that we must have a factor of $\eta(6z)^{a_6}$, where $a_6 = -\frac{1}{30}a^6 - \frac{1}{2}a^3 + \frac{8}{15}a^2 + 2a$. The coefficient of q^7 is then

$$-\frac{2}{35}a(a-1)(a+1)(a-2)(a+2)(a^2+5),$$

and, again since $a \neq 0$, we must have that $a = \pm 1, \pm 2$.

If $a = -2$, the above yields that $f(z)$ is divisible by $\frac{\eta(2z)^5}{\eta(z)^2\eta(4z)^2}$, which is exactly the eta-quotient representation of $\theta_1(z)$ we found earlier. In particular, the exponents in its Fourier series are supported on the squares, and any additional factor $\eta(dz)^{r_d}$ of $f(z)$ must then have the property that d is a square. But because of the $2q$ term in the Fourier series of $f(z)$, this would introduce a term of order q^{d+1} in the Fourier series which we cannot change. Since $d+1$ is not a perfect square, we have found the only possible $f(z)$ with $a = -2$. Similarly, we find that there is exactly one form in Θ_1 for each of $a = -1$, $a = 1$, and $a = 2$:

$$\frac{\eta(2z)^2\eta(3z)}{\eta(z)\eta(6z)}, \quad \frac{\eta(z)\eta(4z)\eta(6z)^2}{\eta(2z)\eta(3z)\eta(12z)}, \quad \text{and} \quad \frac{\eta(z)^2}{\eta(2z)}.$$

If the first non-zero coefficient of $f(z)$ after the constant term is $-aq^4$, the argument above translates almost exactly to this case, with the modification that q is replaced by q^4. One must argue that there can be no $\eta(9z)^{r_9}$ factor, but this is clear, as it would introduce an irreparable q^{13} in the Fourier expansion. We thus obtain in this case the previous eta-quotients with z replaced by $4z$. Two of these are in Θ_1, namely

$$\frac{\eta(8z)^5}{\eta(4z)^2\eta(16z)^2} \text{ and } \frac{\eta(4z)\eta(16z)\eta(24z)^2}{\eta(8z)\eta(12z)\eta(48z)}.$$

If the Fourier series starts $1-aq^9+O(q^{16})$, we proceed similarly, obtaining two elements of Θ_1^0, namely $\frac{\eta(9z)^2}{\eta(18z)}$ and $\frac{\eta(18z)^5}{\eta(9z)^2\eta(36z)^2}$. Throughout all of the remaining cases, corresponding to the first non-constant term being q^{16}, q^{36}, and q^{144}, we obtain exactly two forms whose coefficients are supported on the squares: $\frac{\eta(36z)^2}{\eta(72z)}$, which is not in Θ_1, and $\frac{\eta(72z)^5}{\eta(36z)^2\eta(144z)^2}$, which is in Θ_1.

We now consider the case when the Fourier expansion of $f(z)$ has the form $q+O(q^4)$. By Lemma 4.18, the factor $\eta(dz)^{r_d}$ of $f(z)$ with $r_d \neq 0$ and d minimal has $d \in \{3, 8, 24, 48, 288\}$. We consider each of these cases in turn.

If $a := r_3 \neq 0$, in order for the Fourier coefficients to cancel properly, we must also have that $\eta(6z)^{a_2}\eta(9z)^{a_3}\eta(12z)^{a_4}\eta(18z)^{a_5}$ divides $f(z)$, where $a_2 = \frac{a(a-3)}{2}$, $a_3 = \frac{a(a-2)(a+2)}{3}$, $a_4 = \frac{a(a+2)(a-1)^2}{4}$, and $a_5 = -\frac{a(a-2)(a+2)(a^3+4a+15)}{30}$. The coefficient of q^{21} in the Fourier expansion of the corresponding $f(z)$ is

$$-\frac{2}{35}a(a-1)(a+1)(a-2)(a+2)(a^2+5),$$

from which we see that a must be one of $\pm1, \pm2$. These give rise to

$$\frac{\eta(3z)^2\eta(12z)^2}{\eta(6z)}, \frac{\eta(3z)\eta(18z)^2}{\eta(6z)\eta(9z)}, \frac{\eta(6z)^2\eta(9z)\eta(36z)}{\eta(3z)\eta(12z)\eta(18z)}, \text{ and } \frac{\eta(6z)^5}{\eta(3z)^2}.$$

As before, one verifies that each of these is in some Θ_m, and no additional factors $\eta(dz)^{r_d}$ can be added to each of these eta-quotients while maintaining the property that the Fourier coefficients are supported on the squares.

If $r_3 = 0$ and $a := r_8 \neq 0$, then we must also have that

$$\eta(16z)^{a_2}\eta(24z)^b\eta(32z)^{a_4}$$

divides $f(z)$, where $a_2 = \frac{a(a-3)}{2}$, b is arbitrary, and $a_4 = ab - \frac{1}{12}a^4 + \frac{7}{12}a^2 + \frac{1}{2}a$. Requiring the coefficient of q^{40} to be 0 (as we are not permitted to change it), we must have that

$$b = \frac{2a^4 - 20a^2 + 18}{15a}.$$

Since b is an integer, we have that $a \mid 18$, and one observes that the above is an integer only for $a \mid 6$. If $a = \pm 1$, we find the two forms

$$\frac{\eta(8z)\eta(32z)}{\eta(16z)} \text{ and } \frac{\eta(16z)^2}{\eta(8z)},$$

both of which are theta functions. Any factors of $\eta(48z)$ or $\eta(288z)$ would introduce irreparable coefficients, so these are all the forms arising from $a = \pm 1$. If $a = \pm 2$, we find

$$\frac{\eta(8z)^2\eta(48z)}{\eta(16z)\eta(24z)} \text{ and } \frac{\eta(16z)^5\eta(24z)\eta(96z)}{\eta(8z)^2\eta(32z)^2\eta(48z)^2},$$

the first of which is in Θ_2^0, the second of which is also a form of some interest, namely $\theta_\chi(z) + 3\theta_\chi(9z)$ where $\chi(n) = \left(\frac{2}{n}\right)$, but the Serre-Stark basis theorem [71] implies that it cannot lie in any Θ_m. If $a = \pm 3$, we obtain the forms

$$\eta(8z)^3 \text{ and } \frac{\eta(16z)^9}{\eta(8z)^3\eta(32z)^3},$$

both of which are theta functions. Lastly, if $a = \pm 6$, we obtain no forms, eventually running into a non-zero coefficient of q^{88}.

We now suppose that $\eta(24z)^a\eta(48z)^b$ is the smallest divisor of $f(z)$. We consider these two variables, since either coefficient can be arbitrary. Consequently, we permit one, but not both, of a and b to be 0. We proceed as before, eventually finding potentially problematic coefficients of q^{240} and q^{264}, say A_1 and A_2, respectively. Both A_1 and A_2 are polynomials in a and b, whose degrees in a are 10 and 11, respectively, and whose degrees in b are both 5. A_1 is irreducible, whereas A_2 has a factor of a and is otherwise irreducible. Substituting $a = 0$ into A_1, we find that we must have $b(b^4 - 6) = 0$, which cannot hold since we assumed that not both a and b are 0. Now, we compute the resultant in b of A_1 and A_2/a. This yields a degree 50 polynomial in a whose only rational roots are $a = 1, a = -1, a = 5$, and $a = -5$, which yield that $b = 0$ or -2, $b = 1$ or 3, $b = -2$, and $b = 13$, respectively. These yield the forms $\eta(24z)$, $\frac{\eta(24z)\eta(96z)\eta(144z)^5}{\eta(48z)^2\eta(72z)^2\eta(288z)^2}$, $\frac{\eta(48z)\eta(72z)^2}{\eta(24z)\eta(144z)}$, $\frac{\eta(48z)^3}{\eta(24z)\eta(96z)}$, $\frac{\eta(24z)^5}{\eta(48z)^2}$, and $\frac{\eta(48z)^{13}}{\eta(24z)^5\eta(96z)^5}$.

Finally, we suppose that $\eta(288z)^a$ is the smallest divisor of $f(z)$. We then must also have that

$$\eta(576z)^{a_2}\eta(864z)^{a_3}\eta(1152z)^{a_4}$$

divides $f(z)$, where $a_2 = \frac{a(a-3)}{2}$, $a_3 = \frac{a(a-2)(a+2)}{3}$, and $a_4 = \frac{a(a+2)(a-1)^2}{4}$. The coefficient of q^{1440} is $\frac{1}{5}a(a^4 - 6)$, which is non-zero, and so there are no such $f(z)$. This finishes the proof of Theorem 4.13.

4.4 Representation by ternary quadratic forms

The problem of determining, given a positive definite integral quadratic form, the integers represented by the quadratic form has motivated, and indeed encodes, a great deal of modern number theory. The problem of determining which forms are *universal* – forms that represent every positive integer – originates with Lagrange's four squares theorem, but it is only recently that a complete characterization has been found; this is the so called "15 Theorem" of Bhargava, Conway, and Schneeberger [9] and the "290 Theorem" of Bhargava and Hanke [8]. In addition, there is very recent work of Rouse [69] proving a "451 Theorem" for representation of odd integers.

When dealing with such problems, arguably the deepest case is that of ternary quadratic forms, bearing in mind that there are always local obstructions, so that the interesting problem becomes to determine the locally represented integers which are globally represented. The reason for the depth in this case is that the number of representations of an integer can be canonically decomposed into a "large" part and a "small" part, neither of which is well-understood. These notions are only valid asymptotically, and a theorem on representations follows by determining the point after which the large part truly is larger than the small part, a method first explicitly employed by Ono and Soundararajan [65] in their study of Ramanujan's ternary quadratic form; these techniques have subsequently been improved upon by Kane [50], Jetchev-Kane [48], and Chandee [17]. Both the large and small parts have an alternative arithmetic interpretation – the large part corresponds to the class number of an imaginary quadratic field (and hence to the value of a Dirichlet L-function), and the small part corresponds to the central critical value of a modular L-function. Thus, the general problem of determining when the large part dominates requires a great deal of understanding of the behavior of L-functions, much of which is beyond our current technology.

Here, we are concerned principally with classifying *regular* positive definite integral ternary quadratic forms. A quadratic form is regular if the only obstructions to representation are local obstructions, which, as mentioned above, is the natural generalization of universal forms to the ternary setting. Jagy, Kaplansky, and Schiemann [47] proved that there are at most 913 regular ternary quadratic forms, and they proved that 891 of these forms are indeed regular. In subsequent work of Oh [63], eight more forms in the list of 913 were proved to be regular. The purpose of this paper is to establish that the remaining 14 forms are regular, albeit conditionally; see Sections 4.4.1 and 4.4.1 for the list of these forms.

Theorem 4.19. *Assume the GRH for all Dirichlet L-functions and all mod-*

ular L-functions. Then each of the remaining 14 ternary quadratic forms mentioned above is regular.

Remark. As the proof of Theorem 4.19 will show, we do not actually need the GRH for all modular L-functions. Rather, we need it for the set of quadratic twists of certain weight two newforms.

While it is obviously unfortunate that we are not able to provide an unconditional proof of this result, the fact that the GRH plays a role should not be surprising. The only general method to obtain results on representation depends on the decomposition into the large and small parts mentioned above, both of which encode values of L-functions. We understand both objects very well assuming the GRH, but for neither do we currently possess unconditional bounds of sufficient quality.

Motivated by work of Granville and Stark [35], who established that a form of the *abc*-conjecture implies that there are no Siegel zeros of Dirichlet L-functions, we consider what exceptional arithmetic consequences would arise from the failure of a large locally represented integer to be globally represented.

Theorem 4.20. *Let Q be a ternary quadratic form of discriminant Δ, and assume the GRH for the family of L-functions associated to quadratic twists of newforms of conductor dividing Δ. Moreover, given any integer n, let $\overline{n}$ denote the image of n in the finite set $\prod_{p|\Delta} \mathbb{Q}_p^\times / \left(\mathbb{Q}_p^\times\right)^2$. Then there is an explicitly computable constant $N(Q, \mathbf{a})$ such that if $n \geq N(Q, \mathbf{a})$, $\overline{n} = \mathbf{a}$, n is squarefree, and n is locally represented but is not globally represented, then there is a Siegel zero of some Dirichlet L-function.*

Remark. By a Siegel zero of a Dirichlet L-function, we mean a real zero $\sigma < 1$ of some $L(s, \chi)$, where χ is a primitive real Dirichlet character to the modulus q and

$$\sigma > 1 - \frac{c}{\log 3q},$$

where c is some positive real number. Of course, we allow the quantity $N(Q, \mathbf{a})$ in Theorem 4.20 to depend upon the choice of c.

The constant $N(Q, \mathbf{a})$ in Theorem 4.20 is especially nice in the case that the cuspidal part of the theta function associated to Q is a Hecke eigenform. As an example of this, we have the following application to Ramanujan's ternary quadratic form, $Q = x^2 + y^2 + 10z^2$, which, in their pioneering paper, Ono and Soundararajan [65] proved represents all odd integers greater than 2719 under the assumption of the GRH.

Corollary 4.21. *Assume the GRH for the L-functions of all quadratic twists of the elliptic curve $y^2 = x^3 + x^2 + 4x + 4$. If the quadratic form $Q = x^2 + y^2 + 10z^2$ does not represent an odd integer $n \geq 2.8 \cdot 10^{25}$, then some Dirichlet L-function has a Siegel zero with*

$$c = 342395 \cdot n^{-0.392} \log^2 n.$$

Moreover, if c is fixed, if Q does not represent a locally represented integer $n \geq 8.179 \cdot 10^{24} \cdot c^{-2.793}$, then some Dirichlet L-function possesses a Siegel zero.

Lastly, for completeness, we consider the complementary question of, assuming the GRH for Dirichlet L-functions, deducing from the failure of a locally represented integer to be globally represented exceptional behavior for the arithmetic of modular forms. While we are able to state a more general result (see the third remark following the theorem), we focus on a case of more arithmetic interest. We say that a ternary quadratic form Q is *associated to an elliptic curve $E/\mathbb{Q}$* if the cuspidal part of its theta function is a Hecke eigenform which lifts, under the Shimura correspondence, to the cusp form associated to E. Also, given any elliptic curve E, let E_d denote its quadratic twist by d, and let $Ш(E_d)$ denote the Tate-Shafarevich group of E_d.

Theorem 4.22. *Assume the GRH for Dirichlet L-functions and the Birch and Swinnerton-Dyer conjecture for rank 0 elliptic curves, and suppose that Q is associated to the elliptic curve E. If n is locally represented but not globally, then there is a positive integer $d \gg n$ for which*

$$|\text{Ш}(E_{-d})| \gg_E \frac{d}{\log^4 d},$$

where the implied constant can be made explicit.

Three remarks:

1) Ramanujan's quadratic form $x^2 + y^2 + 10z^2$ is an example of a quadratic form associated to an elliptic curve; namely, it is associated to the curve $y^2 = x^3 + x^2 + 4x + 4$ given in Corollary 4.21. In this case, we would have $d = 10n$.

2) While the lower bound on the size of $\text{Ш}(E_d)$ does not contradict the Goldfeld-Szpiro conjecture [26] that, for any $E/\mathbb{Q}$ with conductor N, $|\text{Ш}(E)|$ is $O_\epsilon\left(N^{1/2+\epsilon}\right)$ uniformly in E, in fact a stronger statement is expected for the family of quadratic twists. In particular, the Ramanujan conjecture for half-integral weight modular forms would imply that $|\text{Ш}(E_d)| \ll_{E,\epsilon} d^{1/2+\epsilon}$.

3) In the event that Q is not associated to an elliptic curve, it is still possible to deduce similar sorts of arithmetic implications. To any packet of Galois representations, and in particular to a newform, one can associate a Tate-Shafarevich group, and it is possible, under the appropriate conjectures, to deduce that some quadratic twist of a newform associated to Q would have an usually large Tate-Shafarevich group in this sense. See, for example, work of Bloch and Kato [10] for more information on such objects.

This paper is organized as follows. In Section 4.4.1, we go over the necessary background in more detail, and we prove Theorem 4.19. In Sections 4.4.2 and 4.4.3, we prove Theorems 4.20 and 4.22, respectively.

4.4.1 Representation by ternary quadratic forms

We begin this section by going into more detail on the decomposition alluded to in the introduction, and we discuss the general approach to be taken to prove Theorem 4.19; this comprises Section 4.4.1. This approach proves to be technically slightly easier in the case that the form in question is in a genus of size two. This is the case for 11 of the 14 forms, and we prove that each is regular in Section 4.4.1. The remaining three forms are each in a genus of size three, and we dispatch of these in Section 4.4.1.

Eisenstein series and cusp forms

We begin with a brief review of the theory of quadratic forms as it relates to the theory of modular forms. Since we are only concerned with the case of positive definite integral ternary quadratic forms, it is to be understood that when we talk about a quadratic form, it is assumed to be such. Now, given two quadratic forms Q_1 and Q_2, we say that Q_1 and Q_2 are (globally) equivalent if there is some matrix $\gamma \in GL_3(\mathbb{Z})$ such that the variable substition $(x, y, z) \mapsto \gamma \cdot (x, y, z)$ takes Q_1 to Q_2; we say that they are locally equivalent if for each prime p, there is some matrix $\gamma_p \in GL_3(\mathbb{Z}_p)$ taking Q_1 to Q_2. The genus of a form Q, denoted by $\mathcal{G}(Q)$, is the set of forms locally equivalent to Q modulo global equivalence.

We can express Q in the form

$$Q(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\tau} A\mathbf{x},$$

where A is a symmetric 3×3 matrix with integer entries and even diagonal entries. The discriminant Δ of Q is the determinant of A, and the level of Q is the least integer N for which NA^{-1} has integer entries and even diagonal entries. The theta function associated to Q is given by

$$\theta_Q(z) := \sum_{\mathbf{x} \in \mathbb{Z}^3} q^{Q(\mathbf{x})}, \qquad q := e^{2\pi i z},$$

and it is a classical fact that $\theta_Q(z)$ is a modular form of weight $3/2$, level N, and nebentypus $\left(\frac{-4\Delta}{\cdot}\right)$. As such, it can be decomposed as

$$\theta_Q(z) = E(z) + C(z), \tag{4.66}$$

where $E(z)$ is an Eisentein series and $C(z)$ is cusp form. In fact, $E(z)$ can always be found from the theta functions of the forms in the genus of Q by the formula

$$E(z) = \frac{\sum_{Q' \in \mathcal{G}(Q)} \frac{1}{|\mathrm{Aut}(Q')|} \theta_{Q'}(z)}{\sum_{Q' \in \mathcal{G}(Q)} \frac{1}{|\mathrm{Aut}(Q')|}},$$

where $\mathrm{Aut}(Q')$ denotes the (finite) automorphism group of Q'. In addition, there is the Siegel mass formula, which asserts that the Fourier coefficients of $E(z)$ are essentially class numbers of imaginary quadratic fields multiplied by certain local densities. In particular, we have, if $E(z) = \sum a_E(n) q^n$, that

$$a_E(n) = \frac{24 h(-Mn)}{M w(-Mn)} \prod_{p \mid 2N} \beta_p(n) \cdot \frac{1 - \chi(p) \left(\frac{n}{p}\right) p^{-1}}{1 - p^{-2}}, \tag{4.67}$$

where $h(-d)$ denotes the class number of the field $\mathbb{Q}(\sqrt{-d})$, $w(-d)$ denotes the number of roots of unity in $\mathbb{Q}(\sqrt{-d})$, M is a rational number depending on $n \pmod{8N^2}$ such that $-nM$ is a fundamental discriminant, and the quantities $\beta_p(n)$ are certain local densities depending on the image of n in the finite set

$$\prod_{p \mid \Delta} \mathbb{Q}_p^\times / \left(\mathbb{Q}_p^\times\right)^2.$$

See work of Hanke [40] or Rouse [69] for more information on these densities. Hence, if n is locally represented, each $\beta_p(n) \neq 0$, and we have that

$$a_E(n) \gg_Q h(-Mn) \gg_\epsilon n^{1/2-\epsilon},$$

where the last inequality is the notoriously ineffective theorem of Siegel. Currently, the best effective lower bound on class numbers is due to Goldfeld

[27]. This relies on the deep work of Gross and Zagier [37], and would yield $a_E(n) \gg_\epsilon \log^{1-\epsilon} n$. As we will see, this bound is far too small to be of use, and it is for this reason that we must assume the GRH for Dirichlet L-functions. Assuming the GRH, the best explicit lower bounds on class numbers are due to Chandee [17], and are, of course, of the same quality as Siegel's lower bound.

To handle the cuspidal part $C(z)$ of $\theta_Q(z)$, we note that $C(z)$ can be decomposed as a linear combination of eigenforms. Each of these is either a unary theta function, which necessarily has coefficients supported on a single square class, or is subject to a theorem of Waldspurger [77], which says that the coefficients in certain square classes are essentially the square roots of central L-values of quadratic twists of the integral weight cusp form associated to the eigenform via Shimura's correspondence. Precisely, we have the following.

Theorem (Waldspurger). Suppose that $f(z) \in S_{\lambda+1/2}(\Gamma_0(N), \chi)$ is a half-integer weight eigenform of each of the Hecke operators $T(p^2)$, $p \nmid N$, and eigenvalues $\lambda(p)$. Moreover, assume that $F(z) \in S_{2\lambda}(\Gamma_0(N), \chi^2)$ has the same system of eigenvalues under each $T(p)$. If n_1 and n_2 are two positive squarefree integers with $n_1/n_2 \in (\mathbb{Q}_p^\times)^2$ for each $p \mid N$ and $f(z) = \sum_{n=1}^\infty a(n)q^n$, then

$$a(n_1)^2 \left(\frac{n_2}{n_1}\right)^{\lambda-1/2} \chi\left(\frac{n_2}{n_1}\right) L\left(1, F \otimes \chi^{-1}\chi_{n_2}\right) = a(n_2)^2 L\left(1, F \otimes \chi^{-1}\chi_{n_1}\right),$$

where $n_i^* = (-1)^\lambda n_i$.

Now, assuming either the Ramanujan conjecture for half-integral weight cusp forms or the GRH for the family of quadratic twists of weight two newforms, it is possible to show that the Fourier coefficients $a_C(n)$ satisfy $a_C(n) \ll_\epsilon n^{1/4+\epsilon}$. Unconditionally, the convexity bound for the family of quadratic twists yields that $a_C(n) \ll_\epsilon n^{1/2+\epsilon}$, which is not sufficient to establish any asymptotic result (recall that $a_E(n) \gg_\epsilon n^{1/2-\epsilon}$). The best known subconvexity bound, due to Blomer and Harcos [11], yields that $a_C(n) \ll_\epsilon n^{7/16+\epsilon}$, which, combined with Siegel's theorem, is enough to establish an asymptotic result. However, as is the case for the coefficients of the Eisenstein series, this upper bound is ineffective, and so we must assume the GRH. Doing so, the best explicit results are again due to Chandee [17].

At this stage, assuming the GRH for Dirichlet L-functions and for the family of L-functions associated to quadratic twists of certain weight two newforms, we are able to obtain that if n is locally represented but is not globally represented, then $a_E(n) + a_C(n) = 0$, and from Chandee's bounds, we are able to rule this out for large values of n. A finite computation then suffices to establish the result. Unsurprisingly, if the cuspidal part of $\theta_Q(z)$ is an eigenform, the bounds are marginally easier to assemble, and we are left with computations that are shorter. Since this is generically the case if the genus of Q is of size two, we consider those forms first before considering the forms in a genus of size three.

Proof of Theorem 4.19: Genera of size two

Of the 14 quadratic forms whose regularity remains unproven, there are 11 that are in a genus of size two; see Table 2.1 for the list of these forms. As mentioned above, for each of these forms, the cuspidal part of the theta function is an eigenform, whose system of eigenvalues necessarily comes from a weight two newform. In fact, each of these newforms is associated to a rational elliptic curve, and we have indicated the Cremona label of each in Table 2.1. With this information and Chandee's bounds, it is now possible to put into action the approach described in the previous section.

The form with the smallest discriminant is $Q := 3x^2 + 6y^2 + 14z^2 + 4yz + 2xz + 2xy$, with discriminant $224 = 2^5 \cdot 7$, and it is associated to the elliptic curve E with Cremona label 32a, given by the Weierstrass equation $E : y^2 =$

Form	**Disc.**	**Curve**	**Req.** n	**Sec.**
$3x^2+6y^2+14z^2+4yz+2xz+2xy$	224	32a	$6.1\cdot 10^6$	75
$x^2+5y^2+13z^2+2yz+xz+xy$	240	48a	$6.7\cdot 10^6$	318
$x^2+6y^2+13z^2+3yz+xz$	297	99b	$2.7\cdot 10^8$	3103
$2x^2+5y^2+11z^2+2yz+2xz+xy$	405	27a	$1.1\cdot 10^4$	0.2
$3x^2+5y^2+15z^2+3yz+3xz+3xy$	720	48a	$2.2\cdot 10^7$	381
$x^2+10y^2+29z^2+5yz+xz$	1125	225b	$3.8\cdot 10^8$	3508
$5x^2+8y^2+11z^2-4yz+xz+2xy$	1620	27a	$8.5\cdot 10^8$	23703
$2x^2+15y^2+32z^2+15yz+xz$	3375	225c	$8.3\cdot 10^8$	6386
$5x^2+13y^2+33z^2-6yz+3xz+xy$	8232	1176h	$7.2\cdot 10^5$	47
$9x^2+11y^2+29z^2-4yz+3xz+6xy$	10125	225b	$9.4\cdot 10^4$	3
$11x^2+15y^2+39z^2-3yz+6xz+3xy$	24696	1176h	$2.4\cdot 10^6$	217

Table 4.1: The 11 forms in a genus of size two.

x^3-x. For each of the 32 square classes $\mathbf{a}$ in $\mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2\times\mathbb{Q}_7^\times/(\mathbb{Q}_7^\times)^2$, we can find constants a, b, and d such that

$$r_Q(n)=a_E(n)+a_C(n)=ah(-bn)\pm dn^{1/4}L(1,E\otimes\chi_{-56n})^{1/2}$$

whenever $\overline{n}=\mathbf{a}$; if $\mathbf{a}$ is not represented, then $a=d=0$. As an example, if $\mathbf{a}=(3,1)$, we find that $a=1/4$, $b=56$, and $d=0.422\ldots$. Using Dirichlet's class number formula, we have that

$$h(-56n)=\frac{1}{\pi}\sqrt{56n}L(1,\chi_{-56n}),$$

and so if n is not represented, we have that

$$\frac{L(1,E\otimes\chi_{-56n})^{1/2}}{L(1,\chi_{-56n})}\geq 1.409\cdot n^{1/4}.$$

On the other hand, using Chandee's theorems, we find that

$$\frac{L(1,E\otimes\chi_{-56n})^{1/2}}{L(1,\chi_{-56n})}\leq 10.091\cdot n^{0.124},$$

which implies that $n \leq 6.108 \cdot 10^6$. Similar computations for the other square classes yield either the same or smaller bounds on n, so it suffices to check that Q is regular for $n \leq 6.108 \cdot 10^6$. For convenience of computation, we note that elementary arguments imply that Q represents n if and only if $Q' := w^2 + 17y^2 + 10yz + 41z^2$ represents $3n$. We check the regularity of Q' up to $20 \cdot 10^6$ for integers divisible by 3, which takes 75 seconds.

The other forms of genus two are handled in exactly the same fashion. The required bounds on n are recorded in Table 2.1, along with the time required for the computation.

Proof of Theorem 4.19: Genera of size three

We now turn to the remaining three forms; see Table 2.2 for the list. For forms in a genus of size greater than two, we no longer expect the cuspidal part of the theta function to be an eigenform. Nonetheless, we are in the lucky situation that the cuspidal part of the first form, $Q := 5x^2+9y^2+15z^2+9yz+3xz+3xy$, happens to be an eigenform, so it can be dispatched as in Section 4.4.1; we have recorded the relevant data in Table 2.2. Moreover, while the cuspidal parts of the theta functions of the remaining two forms are not eigenforms under all Hecke operators $T(p^2)$ for $p \nmid N$, each is an eigenform under some $T(p^2)$. We exploit this fact to more easily compute the decomposition of $C(z)$ into eigenforms.

Form	**Disc.**	**Curves**	**Req.** n	**Time**
$5x^2 + 9y^2 + 15z^2 + 9yz + 3xz + 3xy$	2160	48a	$6.7 \cdot 10^6$	320
$5x^2 + 9y^2 + 17z^2 + 6yz + 5xz + 3xy$	2592	32a, 288e	$2.4 \cdot 10^7$	1974
$5x^2 + 9y^2 + 27z^2 + 3xz + 3xy$	4536	56a, 504d	$7.0 \cdot 10^8$	30161

Table 4.2: The three forms in a genus of size three.

For the third form in Table 2.2, $Q := 5x^2 + 9y^2 + 27z^2 + 3xz + 3xy$,

$C(z)$ is an eigenform under the Hecke operators $T(p^2)$ for $p = 5, 7$, and 13, with eigenvalues $2, -1$, and 2, respectively. There are only two newforms with these eigenvalues in the appropriate weight two spaces, and they are associated with the elliptic curves 56a and 504d. For convenience, we denote these two curves by E_1 and E_2, respectively. Following the same approach as above, for each of the 128 classes $\mathbf{a} \in \mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2 \times \mathbb{Q}_3^\times/(\mathbb{Q}_3^\times)^2 \times \mathbb{Q}_7^\times/(\mathbb{Q}_7^\times)^2$, we have, if $\overline{n} = \mathbf{a}$, that

$$r_Q(n) = ah(-bn) \pm d_1 n^{1/4} L(1, E_1 \otimes \chi_{-14n})^{1/2} \pm d_2 n^{1/4} L(1, E_2 \otimes \chi_{-14n})^{1/2},$$

where each of a, b, d_1, and d_2 can be computed explicitly. We obtain for the squareclass $\mathbf{a} = (3, 2, 3)$, that

$$a = 1/4,\ \ b = 56,\ \ d_1 = 0.851\ldots,\ \ d_2 = 0.0801\ldots,$$

which yields a bound of the form

$$\frac{d_1\sqrt{L(1, E_1 \otimes \chi_{-14n})} + d_2\sqrt{L(1, E_2 \otimes \chi_{-14n})}}{L(1, \chi_{-56n})} \geq 0.595 \cdot n^{1/4},$$

and, following Chandee, we have that

$$\frac{d_1\sqrt{L(1, E_1 \otimes \chi_{-14n})} + d_2\sqrt{L(1, E_2 \otimes \chi_{-14n})}}{L(1, \chi_{-56n})} \leq 7.743 \cdot n^{0.124}.$$

This yields a bound on n of $6.918 \cdot 10^8$. Similar computations reveal only smaller bounds.

For the second form in Table 2.2, $Q := 5x^2 + 9y^2 + 17z^2 + 6yz + 5xz + 3xy$, $C(z)$ is an eigenform for every $p \equiv 3 \pmod 4$ with eigenvalue 0, indicating it is associated to weight two CM cusp forms. There are eight such forms of possible level, each, again, associated to an elliptic curve. However, we find that of these eight systems of eigenvalues, only two play a role in $C(z)$; we have listed the Cremona data for each in Table 2.2. We follow the same technique as above, and have listed the relevant information in Table 2.2.

4.4.2 Siegel zeros: Proof of Theorem 4.20

In this section, we consider the arithmetic consequences resulting from a large locally represented integer failing to be globally represented. The essential idea of the proof of Theorem 4.20 comes from equation (4.66), which states that

$$\theta_Q(z) = E(z) + C(z).$$

Similar to what we did in Section 4.4.1, by assuming the GRH for the family of modular L-functions arising from quadratic twists of newforms of conductor dividing $\Delta(Q)$, we are able to use Chandee's theorems [17] to obtain explicit upper bounds on the Fourier coefficients $a_C(n)$ of $C(z)$. Consequently, if n is locally represented, so that $a_E(n)$ is non-zero, and is not globally represented, so that $a_E(n) + a_C(n) = 0$, we obtain an explicit upper bound on $a_E(n)$, which would seemingly contradict the ineffective lower bound $a_E(n) \gg_\epsilon n^{1/2-\epsilon}$. Of course, the rectification of this comes from the fact that the implied constant depends upon a possible Siegel zero of some Dirichlet L-function. In particular, if there are no Siegel zeros $\sigma < 1$ satisfying

$$\sigma > 1 - \frac{c}{\log 3q},$$

then, following standard techniques (see Davenport [20, §21], for example) a lower bound of the form

$$h(d) \geq \alpha \cdot c e^{-8c} \frac{d^{1/2}}{\log^2 d} \tag{4.68}$$

can be obtained, where $\alpha = 1.288 \ldots \cdot 10^{-4}$. Thus, if we are able to use the above ideas to contradict the bound (4.68), we will have produced a Siegel zero. This is obviously now our goal.

Following the techniques developed in Section 4.4.1, by applying Chandee's theorem, we obtain a bound of the form

$$a_C(n) \leq r \cdot n^s,$$

for some explicit constants r, s depending only on the class of n in

$$\prod_{p \mid \Delta} \mathbb{Q}_p^\times/(\mathbb{Q}_p^\times)^2.$$

In fact, by varying the parameter x in Chandee's bound [17, Equation (19)], we can obtain many different values of (r, s), a fact which we will exploit for the purposes of optimization whenever dealing with a specific form – see the proof of Corollary 4.21 – but in general we only require x to be chosen so that $s < 1/2$. Provided that there is no contribution to $C(z)$ from a unary theta function, we are guaranteed to be able to make this choice (see equation (4.69) below), and if we restrict n to be squarefree and greater than the level of Q, we bypass this issue entirely. We could also require, if we write $n = n_0 n_1^2$ with n_0 squarefree, that n_0 is greater than the level, but we have chosen the statement we did for aesthetic purposes. At this point, Theorem 4.20 follows immediately. We now prove Corollary 4.21 to make this more explicit.

Proof of Corollary 4.21. We begin by noting that $Q = x^2 + y^2 + 10z^2$ has discriminant 40 and is associated to the elliptic curve $E : y^2 = x^3+x^2+4x+4$ with Cremona label 20a. For each square class $\mathbf{a} \in \mathbb{Q}_2^\times/(\mathbb{Q}_2^\times)^2 \times \mathbb{Q}_5^\times/(\mathbb{Q}_5^\times)^2$, and in particular each odd square class, we can find a, b, and d such that

$$r_Q(n) = ah(-bn) \pm d\sqrt{L(1, E \otimes \chi_{-10n})},$$

so that if n is not represented by Q, the bound

$$h(-bn) \leq \frac{d}{a}\sqrt{L(1, E \otimes \chi_{-10n})}$$

must hold. On the other hand, if there are no Siegel zeros for some $c < 1/8$, we also have the bound (4.68), so if the inequality

$$c \geq \frac{4ed}{a\alpha b^{1/2}} n^{-1/2} \log^2 n \sqrt{L(1, E \otimes \chi_{-10n})}$$

holds, we will have arrived at a contradiction. The bound for $L(1, E\otimes\chi_{-10n})$ obtained from Chandee's theorem is independent of the square class $\mathbf{a}$, and we compute that the constant is largest for the class $(1,1)$, with $a = 2/3$, $b = 40$, and $d = 1.572\ldots$. This yields that

$$c \geq 31480 \cdot n^{-1/2} \log^2 n \sqrt{L(1, E \otimes \chi_{-10n})}$$

will be problematic. As mentioned above, using Chandee's theorem, it is possible to bound the L-value by rn^s, where each of r and s depend upon a parameter x. In particular, we have that

$$s = \frac{1+\lambda}{\log x} \tag{4.69}$$

and

$$\log r = \Re \sum_{m\leq x} \frac{a(m)}{m^{\frac{1}{2}+\frac{\lambda}{\log x}} \log m} \frac{\log \frac{x}{m}}{\log x} + \frac{2(\lambda^2+\lambda)}{\log^2 x} + \frac{8e^{-\lambda}}{x^{\frac{1}{2}} \log^2 x} + \frac{1+\lambda}{2\log x} \log\left(\frac{800}{\pi^2}\right), \tag{4.70}$$

where $\frac{\log x}{2} \geq \lambda \geq \lambda_0$, $\lambda_0 = 0.4912\ldots$ is the unique positive solution to $e^{-\lambda_0} = \lambda_0 + \lambda_0^2/2$, and the $a(m)$'s are the coefficients of the Dirichlet series

$$-\frac{L'}{L}(s, E \otimes \chi_{-10n}).$$

By taking $\lambda = \lambda_0$ and $x = 1000$, we obtain that $s = 0.215\ldots$ and $r = 118.285\ldots$. Thus, if $c \geq 342395 \cdot n^{-0.392} \log^2 n$, this yields a contradiction. However, we have assumed that $c \leq 1/8$, and we note that this bound is only below that if $n \geq 2.8 \cdot 10^{25}$. This is the stated result.

Moreover, the bounds on a, b, and d are all worst when $\mathbf{a} = (1,1)$. Using the same values of λ and x, and that $\log n \leq e^{1/\epsilon} n^\epsilon$ for every $\epsilon > 0$, we find that if

$$n \geq 8.179 \cdot 10^{24} \cdot c^{-2.793},$$

then some Dirichlet L-function must have a Siegel zero. This establishes Theorem 4.20. $\square$

4.4.3 Tate-Shafarevich groups: Proof of Theorem 4.22

We now turn our attention to the proof of Theorem 4.22. The starting point is again the inequality

$$h(-bn) \leq \frac{d}{a}\sqrt{L(1, E \otimes \chi_{-Dn})},$$

which must hold if n is in the locally represented square class $\mathbf{a}$ but n is not globally represented. Here, we have assumed that Q is associated to the elliptic curve E. Instead of proceeding as we did in the proof of Theorem 4.20, however, we now assume GRH for Dirichlet L-functions, from which we obtain a lower bound for the central critical value of the modular L-function of the form

$$L(1, E \otimes \chi_{-Dn}) \geq \frac{a^2\alpha^2 b}{64 d^2 e^2}\frac{n}{\log^4(bn)} \gg \frac{n}{\log^4 n},$$

which should be compared to the Ramanujan bound $L(1, E \otimes \chi_{-d}) \ll_{E,\epsilon} d^{\epsilon}$. This inequality immediately guarantees that the analytic rank of $E \otimes \chi_{-Dn}$ is 0, which in turn yields that the arithmetic rank is 0 and the Tate-Shafarevich group $Ш(E \otimes \chi_{-Dn})$ is finite. However, if we want more control over the size of the Tate-Shafarevich group, we must assume the strong form of the Birch and Swinnerton-Dyer conjecture for rank 0 curves, which asserts, for any elliptic curve E' of rank 0, that

$$L(1, E') = \frac{\#Ш(E') \cdot \mathrm{Tam}(E') \cdot \Omega(E')}{\#E'_{\mathrm{tors}}(\mathbb{Q})^2},$$

where $Ш(E')$ denotes the Tate-Shafarevich group of $E'/\mathbb{Q}$, $\Omega(E')$ is the real period of E', $\mathrm{Tam}(E')$ is the Tamagawa number of E', and $E'_{\mathrm{tors}}(\mathbb{Q})$ denotes the rational torsion subgroup of E'. As E' varies over the family of quadratic twists of E, the torsion subgroup $E'_{\mathrm{tors}}(\mathbb{Q})$ is bounded by Mazur's theorem. In fact, a stronger bound can be obtained – apart from possible 2-torsion, there are only finitely many twists with non-trivial torsion subgroup – but this is essentially irrelevant for our theorem. Moreover, the real period varies in a predictable manner; namely, we have that

$$\frac{\Omega(E \otimes \chi_{-d})}{\Omega(E \otimes \chi_{-4})} = d^{-1/2}.$$

While the Tamagawa numbers are harder to control, the general bound

$$\mathrm{Tam}(E \otimes \chi_{-d}) \ll d^{1/2}$$

holds uniformly in E (see, e.g., [26]). The net effect of this is that, if n is locally represented but not globally, the inequality

$$\#Ш(E \otimes \chi_{-Dn}) \gg_E \frac{n}{\log^4 n}$$

must hold, where the implied constant can be made explicit. As mentioned in the introduction, this contradicts standard conjectures about the size of the Tate-Shafarevich group in the family of quadratic twists.

Chapter 5

The pretentious view of analytic number theory

In a recent series of papers, Granville and Soundararajan ([29], [31], [32], [33], [34] as a few examples) introduced the notion of *pretentiousness* in the study of multiplicative functions taking values in the complex unit disc, essentially the idea that if two functions are "close" in some sense, they should exhibit the same behavior. One striking example of this philosophy is a theorem of Halász [38], which can be interpreted as saying that given a multiplicative function $f(n)$ with $|f(n)| \leq 1$ for all n, the partial sums

$$S_f(x) := \sum_{n \leq x} f(n)$$

are large if and only if $f(n)$ "pretends" to be n^{it} for some $t \in \mathbb{R}$ (possibly 0). To make this precise, define the distance between two multiplicative functions $f(n)$ and $g(n)$ taking values in the complex unit disc to be

$$\mathbb{D}(f,g)^2 := \sum_p \frac{1 - \Re(f(p)\bar{g}(p))}{p},$$

where here and throughout, the summation over p is taken to be over primes. This distance is typically infinite, but in the event that it is finite, we follow Granville and Soundararajan and say that $f(n)$ and $g(n)$ are pretentious to each other, or that $f(n)$ is $g(n)$-pretentious. Halász's theorem then says that if $S_f(x) \gg x$, then $f(n)$ must be n^{it}-pretentious for some t. In other words, Halász's theorem classifies those $f(n)$ for which $S_f(x)$ is as large as possible.

Halász's theorem, in essence, tells a beautiful story about those functions exhibiting exceptionally little cancellation (read: essentially no cancellation at all) in their partial sums, so what about those functions which exhibit exceptionally large amounts? Since for generic $f(n)$ taking values in the complex unit disc, the best we can typically hope for is $S_f(x) \ll_\epsilon x^{1/2+\epsilon}$, we are interested in when $S_f(x)$ exhibits more than square root cancellation. In particular, we ask the following question.

Question 5.1. If $f(n)$ is a completely multiplicative function, bounded by 1 in absolute value, such that both $\sum_{n \le x} |f(n)|^2 \gg x$ and $S_f(x) \ll x^{\frac{1}{2}-\delta}$ hold for some fixed $\delta > 0$, must $f(n)$ be $\chi(n)n^{it}$-pretentious for some Dirichlet character χ and some $t \in \mathbb{R}$?

The reason for the condition that

$$\sum_{n \le x} |f(n)|^2 \gg x$$

is twofold. First, we wish to exclude functions like $f(n) = n^{-a}$ for some $a > 0$, and second, this condition is necessary for $\mathbb{D}(f, f)$ to be finite, and therefore for $f(n)$ to be pretentious to any function. In other words, this condition is necessary for $f(n)$ to fit into the context of pretentiousness.

5.1 Multiplicative functions dictated by Artin symbols

Question 5.1 appears to be intractable at present. Not all is lost, however, as we are able to provide an answer for a certain, natural class of functions, which moreover seems like a not unreasonable place to look for conspiracies akin to the periodicity of Dirichlet characters.

This class of functions will be defined via the arithmetic of number fields, with Dirichlet characters arising from cyclotomic extensions. Thus, let $K/\mathbb{Q}$ be a finite Galois extension, not necessarily abelian, and let $\left(\frac{K/\mathbb{Q}}{\cdot}\right)$ denote the Artin symbol, so that for each rational prime p unramified in K, $\left(\frac{K/\mathbb{Q}}{p}\right)$ is the conjugacy class in $\mathrm{Gal}(K/\mathbb{Q})$ of elements acting like Frobenius modulo $\mathfrak{p}$ for some prime $\mathfrak{p}$ of K dividing p. We let $\mathcal{S}_K$ denote the class of complex-valued completely multiplicative functions $f(n)$ satisfying the following two properties.

First, we require that $|f(p)| \leq 1$ for all primes p, with equality holding for all unramified primes, so that f both fits into the context of pretentiousness and is of the same size as a Dirichlet character. Secondly, generalizing the dependence of $\chi(p)$ only on the residue class of $p \pmod q$, we require $f(p)$ to depend only on the Artin symbol $\left(\frac{K/\mathbb{Q}}{p}\right)$. That is, if p_1 and p_2 are any two unramified primes such that

$$\left(\frac{K/\mathbb{Q}}{p_1}\right) = \left(\frac{K/\mathbb{Q}}{p_2}\right),$$

we must have that $f(p_1) = f(p_2)$. We note that if $K = \mathbb{Q}(\zeta_m)$, the m-th cyclotomic extension, $\mathcal{S}_K$ includes all Dirichlet characters modulo m, and by taking K to be a non-abelian extension, we can obtain other functions of arithmetic interest which are intrinsically different from Dirichlet characters; see the examples following Theorem 5.3. We are now interested in the following reformulation of Question 5.1 to the class of functions in $\mathcal{S}_K$.

Question 5.2. Suppose $f \in \mathcal{S}_K$ is such that $S_f(X) = O_f(X^{1/2-\delta})$ as $X \to \infty$ for some fixed $\delta > 0$. Must $f(n)$ coincide with a Dirichlet character? That is, must $f(p) = \chi(p)$ for all but finitely many primes?

Modifying techniques of Soundararajan [74] that were developed to show that degree 1 elements of the Selberg class arise from Dirichlet L-functions, we are able to answer this question in the affirmative.

Theorem 5.3. *If $f \in \mathcal{S}_K$ is such that $S_f(X) = O_f(X^{1/2-\delta})$, then $f(n)$ coincides with a Dirichlet character of modulus dividing the discriminant of K.*

Three remarks: First, while Question 5.1 fits nicely into the pretentious philosophy, the proof of Theorem 5.3 is highly non-pretentious, as it relies critically on L-function arguments. However, we still consider this proof to be of merit, as it highlights the interface between pretentious questions and techniques relying on L-functions.

Second, as the proof will show, the conditions on the class $\mathcal{S}_K$ are not optimal in two ways. In particular, first, the assumption that $f(n)$ is completely multiplicative is not necessary. What is required is that $f(p^2)$ is also determined by Artin symbols and that $|f(p^k)| \ll 1$ for all $k \geq 3$. Second, the assumption that $|f(p)| \leq 1$ is not necessary for primes which do not split completely in K, provided one has $|f(p^k)| \ll 1$. We have chosen the defintion of $\mathcal{S}_K$ as we did principally for aesthetic purposes.

Third, a restriction on the size of $f(p)$ is unlikely to be necessary for an interesting result to be obtained. In particular, the condition of constancy on primes with the same Frobenius implies that $S_f(X) = \tilde{O}(X)$ for all $f(n)$, and $S_f(X) = \tilde{O}(X^{1/2+\epsilon})$ for almost all $f(n)$ (where the notation $\tilde{O}$ indicates the statement is valid up to powers of $\log X$). Thus, the question of more than square root cancellation would still be of interest here. We would expect an analogue of Theorem 5.3 to hold, with Dirichlet characters replaced by the coefficients of any Artin L-function associated to $K/\mathbb{Q}$, but a proof would require significant progress toward understanding the nature of the Selberg class, among other things, and so is likely unobtainable.

We conclude this section with three examples of functions in some $\mathcal{S}_K$ which we believe to be of arithmetic interest.

Example 5.4. Let $F(x) = x^3 + x^2 - x + 1$, and let K be the splitting field of $F(x)$, which has Galois group $G \cong S_3$ and discriminant $-21296 = -2^4 \cdot 11^3$. Let $\rho(p)$ denote the number of inequivalent solutions to the congruence $F(x) \equiv 0 \pmod{p}$, and define the function $f \in \mathcal{S}_K$ by

$$f(p) = \begin{cases} -1 & \text{if } p \nmid 22 \text{ and } \rho(p) = 0, \\ 0 & \text{if } p \mid 22 \text{ or } \rho(p) = 1, \\ 1 & \text{if } p \nmid 22 \text{ and } \rho(p) = 3. \end{cases}$$

There is a unique Dirichlet character χ in $\mathcal{S}_K$, which corresponds to the alternating character of S_3, and is given by $\chi(p) = \left(\frac{-11}{p}\right)$. Alternatively, we can write $\chi(p)$ in terms of $\rho(p)$ by

$$\chi(p) = \begin{cases} -1 & \text{if } p \neq 11 \text{ and } \rho(p) = 1, \text{ or } p = 2, \\ 1 & \text{if } p \nmid 22 \text{ and } \rho(p) = 0 \text{ or } 3, \\ 0 & \text{if } p = 11. \end{cases}$$

Since $f(p) \neq \chi(p)$ for those primes p such that $\rho(p) = 0$ or 1 and since such primes occur a positive proportion of the time by the Chebotarev density theorem, we should not expect to see more than square root cancellation in the partial sums of $f(n)$ by Theorem 5.3, and indeed, we find the following.

X	$S_f(X)$	$\lvert S_f(X)\rvert/\sqrt{X}$	$\lvert S_f(X)\rvert/(X/\log^{7/6} X)$
10	0	0	0
10^2	-4	0.4	0.238
10^3	-12	0.379	0.114
10^4	-102	1.02	0.136
10^5	-736	2.327	0.127
10^6	-5757	5.757	0.123

In the next example, we discuss the apparent convergence in the fourth column.

Example 5.5. The astute reader may object to the above example by noting that $f(n)$ as constructed has mean $-1/6$ on the primes as a consequence of the Chebotarev density theorem. By using the Selberg-Delange method [76, Chapter II.5], we would expect that $S_f(X)$ should have order $X/(\log X)^{7/6}$, which indeed matches the data more closely (and explains the fourth column). The Selberg-Delange method breaks down when the mean on the primes is 0 or -1, with the latter possibility essentially corresponding to the Möbius function $\mu(n)$. Thus, the most interesting case occurs when the mean on the primes is 0. It is a simple exercise to see that any such function $g \in \mathcal{S}_K$ (where K is as in Example 5.4) arises as the "twist" of $\chi(n)$ – we must have that $g(p) = \omega\chi(p)$ for all primes $p \nmid 22$ and some ω satisfying $|\omega| = 1$. Taking $g(p) = i\chi(p)$ for all primes p, we compute the following.

X	$S_g(X)$	$\lvert S_g(X)\rvert/\sqrt{X}$
10	$1+i$	0.447
10^2	$2+i$	0.224
10^3	$6+2i$	0.2
10^4	$13+6i$	0.143
10^5	$36+50i$	0.195
10^6	$-260+215i$	0.337

Here, the fact that $S_g(X)$ is not $O(X^{1/2-\delta})$ for some $\delta > 0$ is less apparent than was the case in Example 1 (there is even more fluctuation than is visible in the limited information above – for example, $S_g(810000)/\sqrt{810000} \approx 0.059$), but nevertheless, since $g(n)$ does not coincide with a Dirichlet character, Theorem 5.3 guarantees that the partial sums are not $O(X^{1/2-\delta})$.

Example 5.6. Let $F(x) = x^4 + 3x + 3$, and let K be the splitting field of $F(x)$, which has Galois group $G \cong D_4$ and discriminant $1750329 = 3^6 \cdot 7^4$. There are five conjugacy classes of G, three of which, each of order two, can be determined by exploiting the quadratic subfields $\mathbb{Q}(\sqrt{-3})$ and $\mathbb{Q}(\sqrt{-7})$.

To distinguish the remaining two conjugacy classes, each of which consists of a single element, we exploit the factorization of $F(x)$ modulo p. As in the previous examples, let $\rho(p)$ denote the number of inequivalent solutions to the congruence $F(x) \equiv 0 \pmod{p}$, and additionally define $l(p)$ to be the pair $\left(\left(\frac{-3}{p}\right), \left(\frac{-7}{p}\right)\right)$. We now consider $f \in \mathcal{S}_K$ defined by

$$f(p) = \begin{cases} 1, & \text{if } l(p) = (1,1) \text{ and } \rho(p) = 4, \\ -1, & \text{if } l(p) = (1,1) \text{ and } \rho(p) = 0, \\ 1, & \text{if } l(p) = (-1,-1), \\ \zeta_3, & \text{if } l(p) = (-1,1), \\ \zeta_3^2, & \text{if } l(p) = (1,-1), \\ 0, & \text{if } p \mid 21. \end{cases}$$

We note that $f(n)$ is neither a Dirichlet character nor its twist – each of the three Dirichlet characters $\left(\frac{-3}{\cdot}\right), \left(\frac{-7}{\cdot}\right)$, and $\left(\frac{21}{\cdot}\right)$ has the same value on the singleton conjugacy classes, and these are the unique characters in $\mathcal{S}_K$ – yet it has mean 0 on the primes. We find the following.

X	$S_f(X)$	$\lvert S_f(X)\rvert/\sqrt{X}$
10	0	0
10^2	$4.5 - 2.598i$	0.520
10^3	$-11 + 6.928i$	0.411
10^4	$0.5 - 2.598i$	0.026
10^5	$-34 - 71.014i$	0.249
10^6	$-21 + 124.708i$	0.126

As with Example 5.5, we find the fact that $S_f(X)$ is not $O(X^{1/2-\delta})$ to be not entirely clear, yet it is guaranteed to be so. In this case, there is even more fluctuation in the values of $|S_f(X)|/\sqrt{X}$. For example, when $X = 7.61 \cdot 10^5$, we have that $|S_f(X)|/\sqrt{X} \approx 0.012$, yet when $X = 7.69 \cdot 10^5$, we have that

$|S_f(X)|/\sqrt{X} \approx 0.186$. Thus, without knowledge of Theorem 5.3, it would be difficult to guess the correct order of $S_f(X)$, although if one were forced to speculate, $O(X^{1/2})$ would probably be the most reasonable guess. In fact, under the generalized Riemann hypothesis, we have that $S_f(X) = O_\epsilon(X^{1/2+\epsilon})$ for all $\epsilon > 0$, so that by Theorem 5.3, square root cancellation is the truth in this case.

5.1.1 Proof of Theorem 5.3

We begin by recalling the setup in which we are working. $K/\mathbb{Q}$ is a finite Galois extension with Galois group $G := \mathrm{Gal}(K/\mathbb{Q})$, and $f \in \mathcal{S}_K$ if and only if $|f(p)| \leq 1$ for all primes p, with equality holding if p is unramified in K, and $f(p_1) = f(p_2)$ for all unramified primes p_1 and p_2 such that

$$\left(\frac{K/\mathbb{Q}}{p_1}\right) = \left(\frac{K/\mathbb{Q}}{p_2}\right).$$

We can therefore regard f as a class function of G, and as such, it can be decomposed [62] as

$$f = \sum_{\chi \in \mathrm{Irr}(G)} a_\chi \chi, \tag{5.1}$$

where $\mathrm{Irr}(G)$ denotes the set of characters associated to the irreducible representations of G and each $a_\chi \in \mathbb{C}$. To establish Theorem 5.3, we wish to show that $a_\chi = 1$ for some one-dimensional χ and $a_{\chi'} = 0$ for all other χ'. We do so incrementally, first establishing that each a_χ is, in fact, rational, then, using techniques due to Soundararajan [74] developed to study elements of the Selberg class of degree 1, we prove the result.

Let $L(s, f)$ denote the Dirichlet series associated to $f(n)$, so that we have

$$L(s, f) := \sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_p \left(1 - \frac{f(p)}{p^s}\right)^{-1},$$

recalling that $f(n)$ is completely multiplicative. By matching the coefficients of p^{-s} in each Euler factor, the decomposition (5.1) then guarantees the Euler product factorization

$$L(s,f) = \prod_{\chi} L(s,\chi)^{a_\chi} \prod_{p} \left(1 + O(p^{-2s})\right),$$

valid in the region of absolute convergence $\Re(s) > 1$, and where $L(s,\chi)$ is the Artin L-function associated to the representation attached to χ. In fact, we will need to go further with this factorization. The coefficient of p^{-2s} in the Euler product is again essentially a class function of G, so it can be decomposed as a linear combination of the characters χ, and we obtain

$$L(s,f) = \prod_{\chi} L(s,\chi)^{a_\chi} \prod_{\chi} L(2s,\chi)^{b_\chi} A(s), \tag{5.2}$$

where $A(s)$ is analytic and non-zero in the region $\Re(s) > 1/3$. We note that for each non-trivial χ, the L-function $L(2s,\chi)$ is analytic and non-zero in some neighborhood of the region $\Re(s) \geq 1/2$, and for the trivial character χ_0, we have that $L(2s,\chi_0) = \zeta(2s)$ is again analytic and non-zero in a neighborhood of $\Re(s) \geq 1/2$, except at $s = 1/2$. Thus, the product

$$\prod_{\chi} L(2s,\chi)^{b_\chi} A(s)$$

is analytic and non-zero in some neighborhood of $\Re(s) \geq 1/2$, except possibly at $s = 1/2$. Now, the condition that

$$\sum_{n \leq X} f(n) = O(X^{1/2-\delta})$$

guarantees that $L(s,f)$ has an analytic continuation to $\Re(s) > 1/2 - \delta$, so by the above, we must have that

$$\prod_{\chi} L(s,\chi)^{a_\chi}$$

is analytic in a neighborhood of $\Re(s) \geq 1/2$, except possibly at $s = 1/2$. In particular, we must have that

$$\mathrm{ord}_{s=s_0} L(s,f) = \sum_{\chi} a_\chi \mathrm{ord}_{s=s_0} L(s,\chi) \tag{5.3}$$

for any $s \neq 1/2$ with $\Re(s) \geq 1/2$.

Recall now that we wish to show that each a_χ is rational. This would follow from (5.3) above if there are $n := \#\mathrm{Irr}(G)$ choices $s_1, \ldots, s_n$ such that the vectors

$$(\mathrm{ord}_{s=s_i} L(s,\chi_1), \ldots, \mathrm{ord}_{s=s_i} L(s,\chi_n))$$

are linearly independent over $\mathbb{Q}$. If this is not the case, then there must be integers n_χ, not all zero, such that, for all $s_0 \neq 1/2$ with $\Re(s_0) \geq 1/2$, we have

$$\sum_{\chi} n_\chi \mathrm{ord}_{s=s_0} L(s,\chi) = 0.$$

Completing the L-functions and using the functional equation, this implies that the product,

$$\Lambda(s) := \prod_{\chi} \Lambda(s,\chi)^{n_\chi},$$

is entire and non-vanishing, except possibly at $s = 1/2$. We note that this is, essentially, an L-function of degree $d := \sum_\chi n_\chi \dim \chi$ and conductor $q := \prod_\chi q_\chi^{n_\chi}$, where each q_χ is the conductor of $L(s,\chi)$. Where this breaks from standard definitions of L-functions is in the gamma factor,

$$G(s) := \prod_{\chi} \Gamma(s,\chi)^{n_\chi},$$

where $\Gamma(s,\chi)$ is the gamma factor for $L(s,\chi)$, as typically one does not allow negative exponents. Nevertheless, much of the formalism carries through, and in particular, it is possible to show, for example by following Iwaniec and Kowalski [45, Theorem 5.8], that the number of zeros of height up to T is

$$N(T) = \frac{T}{\pi} \log \frac{qT^d}{(2\pi e)^d} + O(\log qT).$$

Since $\Lambda(s)$ is non-vanishing except possibly at $s = 1/2$, there can be no main term, and so we see that $d = 0$ and $q = 1$. Moreover, since $\Lambda(s)$ is entire and of order 1, we have the factorization

$$\Lambda(s) = (s - 1/2)^m e^{A+Bs},$$

where $m \in \mathbb{Z}$ and $A, B \in \mathbb{C}$. Since $d = 0$ and $q = 1$, by considering the functional equation, first as $s \to \infty$ with $s \in \mathbb{R}$, we see that $\Re(B) = 0$, and by $s = 1/2 + it, t \to \infty$, that $B = 0$. Writing $G(s)$ as

$$G(s) = \Gamma(s)^a \Gamma\left(\frac{s}{2}\right)^b \Gamma\left(\frac{s+1}{2}\right)^c$$

and applying Stirling's formula, we see that

$$\begin{aligned}\log G(s) \quad = \quad & \frac{2a+b+c}{2} s \log s - \left(a + (b+c)\left(\frac{1+\log 2}{2}\right)\right) s - \\ & \left(\frac{a+b}{2}\right) \log s + (a+b+c) \log \sqrt{2\pi} + b\frac{\log 2}{2} + O(|s|^{-1}).\end{aligned}$$

Because $\Lambda(s) = C \cdot (s - 1/2)^m$, the coefficients of both $s \log s$ and s must be zero. The fact that $2a+b+c = 0$ is a restatement of the fact that the degree is zero, but the condition that

$$a + (b+c)\left(\frac{1+\log 2}{2}\right) = 0$$

implies that $a = 0$ and $b = -c$, since each of a, b, and c is an integer. We thus have that

$$G(s) = (2s)^{b/2}(1 + O(|s|^{-1})),$$

whence

$$\prod_{\chi} L(s,\chi)^{n_\chi} = (s - 1/2)^m (2s)^{-b/2} (1 + O(|s|^{-1})).$$

Upon taking the limit as $s \to \infty$, we find that $b = m = 0$, so, in fact, $\prod_\chi L(s,\chi)^{n_\chi} = 1$. But this, of course, cannot happen, so no such n_χ exist, and each of the quantities a_χ must be rational.

At this stage, we are now able to prove the theorem. The advantage gained by knowing that each a_χ is rational, is that the function

$$F(s) := \prod_\chi L(s,\chi)^{a_\chi}$$

enjoys nice analytic properties. In particular, apart from a possible branch along the ray $(-\infty, 1/2]$, it will be entire. To see this, let k be the denominator of the a_χ, and note that we must have

$$\prod_\chi \Lambda(s,\chi)^{ka_\chi} = (s-1/2)^m h(s)^k$$

for some entire function $h(s)$. Ignoring the branch, $F(s)$ essentially behaves as an L-function of degree $\sum a_\chi \dim \chi$. However, we note that this is also the evaluation of $f(p)$ at a prime that splits completely in K – or, equivalently, at the identity of $\mathrm{Gal}(K/\mathbb{Q})$ – by (5.1). Thus, it is a rational number of absolute value 1, and so is either 1 or -1. However, there are no holomorphic L-functions of negative degree, as can be seen, for example, by a zero counting argument (which can be modified simply to account for the possible branch), and so the degree must be 1. Moreover, it is known that a degree 1 element of the Selberg class must come from a Dirichlet L-function, a fact which is originally due to Kaczorowksi and Perelli [49] and was reproved by Soundararajan [74]. However, as before, $F(s)$ does not satisfy the axioms of the Selberg class, as its gamma factor may have negative exponents, so we must modify Soundararajan's proof to our situation.

There are only two key components in Soundararajan's proof – an approximate functional equation for $F(s)$ and control of the gamma factors on the line $\Re(s) = 1/2$. The proof of the approximate functional equation naturally requires the analytic properties of $F(s)$, and as it may have a branch in our situation, we must modify the proof slightly; we do so in Lemma 5.7 below. On the other hand, the control over the gamma factors is provided from our assumption on the degree, so in particular, the same estimates hold. We state these estimates in (5.4) and we give the idea of Soundararajan's proof below, after the proof of Lemma 5.7.

Lemma 5.7. *For any $t \in \mathbb{R}$ such that $|t| \geq 2$ and any $X > 1$, we have that*

$$F(1/2+it) = \sum_{n=1}^{\infty} \frac{a(n)}{n^{1/2+it}} e^{-n/X} + O\left(X^{-1+\epsilon}(1+|t|)^{1+\epsilon} + e^{\frac{\log X}{\log |t|} - |t|}\right).$$

Proof. Consider the integral

$$I := \int_{(1)} F(1/2+it+w) X^w \Gamma(w) dw.$$

On one hand, replacing F by its Dirichlet series and directly computing, we find that I is given by

$$I = \sum_{n=1}^{\infty} \frac{a(n)}{n^{1/2+it}} e^{-n/X}.$$

On the other hand, moving the line of integration to the left, we find that

$$I = F(\frac{1}{2}+it) + \int_{(-1+\epsilon)} F(\frac{1}{2}+it+w) X^w \Gamma(w) dw + \int_{\mathcal{C}} F(\frac{1}{2}+it+w) X^w \Gamma(w) dw,$$

where the curve $\mathcal{C}$ is the union of the segments $(-1+\epsilon - it^+, -r - it^+)$ and $(-1+\epsilon - it^-, -r - it^-)$ – that is, above and below the branch – along with the circle of radius r centered at $w = -1/2$. The contribution from the line $\Re(w) = -1+\epsilon$ is handled as in Soundararajan's work via the functional equation, yielding a contribution of $O\left(X^{-1+\epsilon}(1+|t|)^{1+\epsilon}\right)$. Upon taking $r = \log^{-1}|t|$, we find that the contribution from the circle to $\mathcal{C}$ is $O\left(e^{\frac{\log X}{\log |t|}} e^{-|t|} \log^{a-1}|t|\right)$, where $a = -\mathrm{ord}_{s=1/2} F(s)$. Similarly, the contribution from the segments is $O\left(e^{-|t|}\log^b |t|\right)$, even accounting for a possible "trivial pole" of $F(s)$ at $s = 0$ (recall that we may be dividing by gamma factors), and we obtain the result. $\square$

As mentioned above, we must also have some control over the gamma factors on the line $\Re(s) = 1/2$. This is straightforward, as Stirling's formula yields, if $G(s) = \prod_\chi \Gamma(s,\chi)^{a_\chi}$, that there are constants $B, C \in \mathbb{R}$ such that

$$\frac{\overline{G}(1/2 - it)}{G(1/2 + it)} = e^{-it\log\frac{t}{2e} + iB + \frac{\pi i}{4}} C^{-it}\left(1 + O(t^{-1})\right). \tag{5.4}$$

This is not the most natural representation, but it proves to be convenient for the proof. Now, the idea of Soundararajan's proof is to consider, for any real $\alpha > 0$, the quantities

$$\mathcal{F}(\alpha, T) := \frac{1}{\sqrt{\alpha}} \int_{\alpha T}^{2\alpha T} F(1/2 + it) e^{it\log\frac{t}{2\pi e\alpha} - \frac{\pi i}{4}} dt$$

and

$$\mathcal{F}(\alpha) := \lim_{T\to\infty} \frac{1}{T}\mathcal{F}(\alpha, T).$$

Armed with Lemma 5.7, one can evalute $\mathcal{F}(\alpha)$ in two ways, either using the functional equation or not. The first method shows that $\mathcal{F}(\alpha) = 0$ unless $\alpha C q \in \mathbb{Z}$, where $q = \prod_\chi q_\chi^{a_\chi}$ and C is as in (5.4), in which case it is, essentially, the coefficient $a(\alpha C q)$. The second method, on the other hand, shows that $\mathcal{F}(\alpha)$ is periodic with period 1, whence $Cq \in \mathbb{Z}$ and the coefficients $a(n)$ are periodic modulo Cq. As remarked above, the proof of this follows Soundararajan's [74] exactly, with the only modification necessary being the replacement of his approximate functional equation with ours, Lemma 5.7. In fact, this argument extends to show that, if we modify the definition of the Selberg class to allow rational exponents on the gamma factors and for there to be finitely many lapses of holomorphicity, then still, the only degree 1 elements are those coming from the traditional Selberg class.

To conclude the proof of Theorem 5.3, we note that since the coefficients $F(s)$ are periodic modulo some q_0 and are also multiplicative, we must have that, away from primes dividing q_0, that they coincide with a Dirichlet character $\chi_{q_0} \pmod{q_0}$. Thus, we have that

$$\prod_{\chi} L(s,\chi)^{a_\chi} \doteq L(s,\chi_{q_0}),$$

where $\doteq$ means that equality holds up to a finite product over primes. The only way this can happen is if $\chi = \chi_{q_0}$ and $a_\chi = 1$ for some χ and $a_{\chi'} = 0$ for all others. This is exactly what we wished to show, so the result follows.

5.2 Pretentiously detecting power cancellation

(The results in this section are joint work with Junehyuk Jung.)

To study Question 5.1, we first ask that if $f(n)$ is $\chi(n)$-pretentious for some character χ, must $S_f(x)$ be small? This turns out to not be the case - by taking $f(p)$ to be 1 for primes lying in one of a suitably sparse set of dyadic intervals and to be $\chi(p)$ otherwise, one obtains a function which is $\chi(n)$-pretentious, but for which $S_f(x) \gg x/\log x$ for infinitely many x. Thus, we have a function, $f(n)$, which is pretentious to a function, $\chi(n)$, which exhibits as much cancellation as possible in its partial sums, and yet $S_f(x)$ is almost as large as possible. We therefore must ask whether there is a stronger notion of pretentiousness which preserves power savings.

To this end, given any two multiplicative functions $f(n)$ and $g(n)$, not necessarily bounded by 1, define the multiplicative function $h(n)$ by

$$g(n) = (f * h)(n),$$

where $(f * h)(n)$ represents the Dirichlet convolution of $f(n)$ and $h(n)$, and, for any $\beta > 0$, define the (possibly infinite) quantity $H_\beta(f,g)$ by

$$H_\beta(f,g) := \sum_p \sum_{k=1}^{\infty} \frac{|h(p^k)|}{p^{k\beta}}.$$

We caution that the convergence of this quantity is potentially asymmetric in $f(n)$ and $g(n)$. Motivated by the idea that if $f(n)$ and $g(n)$ are close, then each should need to be modified only slightly to obtain the other, we say that $f(n)$ and $g(n)$ are strongly β-pretentious to each other if both $H_\beta(f,g)$ and $H_\beta(g,f)$ are finite. If $f(n)$ and $g(n)$ are strongly β-pretentious for each $\beta > 0$, then we say that they are totally pretentious to each other.

Theorem 5.8. *Suppose that $f(n)$ and $g(n)$ are multiplicative functions, and that $S_f(x) \ll x^\alpha$ for some $\alpha > 0$. If $f(n)$ and $g(n)$ are totally pretentious to each other, then $S_g(x) \ll x^\alpha$. If, however, $f(n)$ and $g(n)$ are only strongly β-pretentious to each other, then we have that $S_g(x) \ll x^{max(\alpha,\beta)}$.*

Two remarks: First, it is apparent that the first statement of Theorem 5.8 regarding total pretentiousness is an immediate corollary to the second statement by taking $\beta < \alpha$. However, we consider its merit to be that it presupposes no knowledge of α to deduce that $S_g(x)$ and $S_f(x)$ exhibit the same level of cancellation.

Second, to obtain the conclusions of Theorem 5.8, it would suffice to suppose only that $H_\beta(f,g)$ is finite, with no hypothesis necessary on $H_\beta(g,f)$. We have chosen this formulation so that strong β-pretentiousness, and hence also total pretentiousness, is an equivalence relation. However, it is only the symmetry requirement that fails if we rely only on the finiteness of $H_\beta(f,g)$, in that if both $H_\beta(f,g)$ and $H_\beta(g,r)$ are finite, then so is $H_\beta(f,r)$.

Now, we wish to consider the extent to which strong and total pretentiousness relate to the traditional notion defined by $\mathbb{D}(f,g)$. We begin with the observation that, if $f(n)$ and $g(n)$ are bounded by 1 in absolute value, then we have that

$$\begin{aligned}\left(\sum_{p,k} \frac{\left|g(p^k) - f(p^k)\right|}{p^{k\gamma}}\right)^2 &\leq \left(\sum_{p,k} \frac{1}{p^{k(2\gamma-\beta)}}\right)\left(\sum_{p,k} \frac{\left|g(p^k) - f(p^k)\right|^2}{p^{k\beta}}\right) \\ &\ll \sum_{p,k} \frac{1 - \Re\left(f(p^k)\overline{g(p^k)}\right)}{p^{k\beta}},\end{aligned}$$

assuming that $\gamma > (1+\beta)/2$. This last quantity is a kind of generalized distance considered by Granville and Soundararajan in their book [30], and so the convergence of the initial series is, in this way, dictated by whether $f(n)$ and $g(n)$ are pretentious in a more traditional sense (although this observation is valid only if $\gamma > 1/2$). Moreover, since we have that $h(p) = g(p) - f(p)$, it is perhaps not unreasonable to hope that the convergence of this series is also related to the convergence of $H_\beta(f,g)$. Thus, define a distance $\widehat{\mathbb{D}}_{\beta,k}(f,g)$ by

$$\widehat{\mathbb{D}}_{\beta,k}(f,g) := \sum_p \sum_{j \le k} \frac{|g(p^j) - f(p^j)|}{p^{j\beta}},$$

and additionally define $\widehat{\mathbb{D}}_\beta := \widehat{\mathbb{D}}_{\beta,\infty}$. Our next theorem shows that, while $\widehat{\mathbb{D}}_\beta(f,g) < \infty$ does not imply that $H_\beta(f,g) < \infty$, it does imply the convergence for sufficiently large primes. We also consider what power cancellation can be deduced directly from assuming that $\widehat{\mathbb{D}}_{\beta,k}(f,g) < \infty$.

Theorem 5.9. *Let $f(n)$ and $g(n)$ be multiplicative functions satisfying $f(n)$, $g(n) = o(n^\delta)$ for some $\delta > 0$.*

1. If $\widehat{\mathbb{D}}_\beta(f,g) < \infty$, there is a $Y > 0$ such that if

$$H_\sigma(f,g;Y) := \sum_{p<Y} \sum_k \frac{|h(p^k)|}{p^{k\sigma}}$$

converges for some $\sigma \ge \beta$ and $\sigma > \delta$, then $H_\sigma(f,g) < \infty$.

2. Suppose that $S_f(x) \ll x^\alpha$ and that $\widehat{\mathbb{D}}_{\beta,k}(f,g) < \infty$. There is a $Y > 0$ such that if $H_\sigma(f,g;Y) < \infty$ for some $\sigma > 1/(k+1) + \delta$ also satisfying $\sigma \ge \max(\alpha,\beta)$, then $S_g(x) \ll x^\sigma$.

While it is unfortunate that we are unable to go from $\widehat{\mathbb{D}}_\beta(f,g) < \infty$ to $H_\beta(f,g) < \infty$ without checking the convergence of $H_\beta(f,g;Y)$, it is, in fact, generically necessary. If we let $f(n) = (-1)^{n+1}$, so that $f(2^k) = -1$ and $f(p^k) = 1$ for all $p \ne 2$ and all $k \ge 1$, and we let $g(n) = 1$, then we of course have that $S_f(x) \ll 1$ and that $S_g(x) \gg x$. However, since neither function is large and they differ only at the prime 2, we also have that $\widehat{\mathbb{D}}_\beta(f,g) < \infty$ for every $\beta > 0$, and we do not want to deduce any cancellation in $S_g(x)$. The reason that the theorem does not apply is that $|h(2^k)| = 2^k$, so that $H_\sigma(f,g;Y)$ diverges for every $\sigma \leq 1$.

Despite the above discussion, for certain classes of functions, we do not have to check the convergence of $H_\sigma(f,g;Y)$. We now present two such classes. The first class is motivated by the properties of the normalized coefficients of automorphic forms.

Definition 5.10. *Given a positive integer d, let $\mathcal{S}_d$ denote the set of "degree d" multiplicative functions, those functions $f(n)$ such that $f(n) = (f_1 * f_2 * \cdots * f_d)(n)$, where each $f_i(n)$ is a completely multiplicative function of modulus bounded by* 1.

As mentioned above, we are able to deduce a nice statement about pretentiousness in the context of degree d functions. Moreover, since the values at prime powers of a degree d function are determined by its values on the first d, it should stand to reason that the convergence of $\widehat{\mathbb{D}}_\beta(f,g)$ should be dictated by the convergence of $\widehat{\mathbb{D}}_{\beta,d}(f,g)$. We are able to show this as well. Thus, we have the following.

Theorem 5.11. *Let $f(n)$ and $g(n)$ be two degree d multiplicative functions such that $\widehat{\mathbb{D}}_{\beta,d}(f,g) < \infty$. We then have that both $\widehat{\mathbb{D}}_\beta(f,g)$ and $H_\beta(f,g)$ are finite. In particular, if we also know that $S_f(x) \ll x^\alpha$, then $S_g(x) \ll x^{\max(\alpha,\beta)}$.*

In the next class of functions, we return to the original setting of pretentiousness, functions of modulus bounded by 1.

Definition 5.12. *Let $f(n)$ be a mutiplicative function of modulus bounded by* 1*. We say that $f(n)$ is* good *at a prime p if there is no choice of $g(n)$*

with modulus bounded by 1 *for which the series*

$$\sum_k \frac{|h(p^k)|}{p^{k\sigma}}$$

fails to converge for some $\sigma > 0$. *We say that* $f(n)$ *is good if it is good at every prime* p.

This definition is, of course, exactly what we need to remove the condition on $H_\sigma(f, g; Y)$. However, we note two things: first, it is easy to give examples of good functions – any completely multiplicative function, say, since we have that $|h(p^k)| \leq 2$ – and, second, that it is possible to classify those functions which are good, which we do in Theorem 5.13. Also, we note that for any $f(n)$ and $g(n)$ bounded by 1, to check the convergence of $\widehat{\mathbb{D}}_\beta(f, g)$, it suffices to check the convergence of $\widehat{\mathbb{D}}_{\beta,\beta^{-1}}(f, g)$.

Theorem 5.13. *Let* $f(n)$ *and* $g(n)$ *be multiplicative functions of modulus bounded by* 1.

1. If $f(n)$ *is good and* $\widehat{\mathbb{D}}_\beta(f, g) < \infty$, *then* $H_\beta(f, g) < \infty$. *Thus, if* $S_f(x) \ll x^\alpha$, *we have that* $S_g(x) \ll x^{\max(\alpha,\beta)}$.

2. $f(n)$ *is good at* p *if and only if the function*

$$F_p(z) := \sum_{k=0}^{\infty} f\left(p^k\right) z^k$$

has no zeros in the open unit disc.

Finally, we return to the Granville-Soundararajan distance function, and we ask to what extent the natural modification

$$\mathbb{D}_\beta(f, g)^2 := \sum_p \frac{1 - \Re(f(p)\overline{g(p)})}{p^\beta}$$

allows one to detect power cancellation for multiplicative functions of modulus bounded by 1. For convenience, if $\mathbb{D}_\beta(f, g)$ is finite, we say that $f(n)$

and $g(n)$ are β-pretentious. As in Theorem 5.9, given $f(n)$ and $g(n)$, we will need a consideration of $h(n)$ at small primes, so we define

$$H_\sigma^2(f,g) := \sum_{p\le 13}\sum_{k=0}^{\infty}\frac{|h(p^k)|^2}{p^{k\sigma}},$$

although, strictly speaking, only the condition $p \le 4^{1/\sigma}$ is necessary (only $\sigma > 1/2$ will be used, so $p \le 13$ is, indeed, weaker). We have the following result, establishing both that β-pretentiousness is sufficient to detect some power cancellation, but that it is fundamentally unable to detect to the level we desire.

Theorem 5.14. *Let $f(n)$ and $g(n)$ be multiplicative functions bounded by 1 such that $S_f(x) \ll x^\alpha$ for some $\alpha < 1$, and suppose that $\mathbb{D}_\beta(f,g) < \infty$ for some $\beta \in (0,1]$.*

1. If $\sigma > 3/4$ is such that $\sigma \ge \max(\alpha, (1+\beta)/2)$ and $H_{2\sigma-1}^2(f,g) < \infty$, then $S_g(x) \ll x^\sigma$.

2. If $f(n)$ and $g(n)$ are both completely multiplicative, then $S_g(x) \ll x^{\max(\alpha,(1+\beta)/2)}$.

3. If $f(n)$ is completely multiplicative and $\beta \ge 2\alpha - 1$, there is a completely multiplicative function $f'(n)$ that is β-pretentious to $f(n)$ and is such that $S_{f'}(x)$ is not $O_\epsilon\left(x^{\frac{1+\beta}{2}-\epsilon}\right)$.

Three remarks: While it's perhaps unsatisfying that β-pretentiousness only detects power savings down to $O\left(x^{\frac{1+\beta}{2}}\right)$ even for completely multiplicative functions, the conclusion of the theorem can be strengthened if $f(n)$ and $g(n)$ are assumed to be real-valued. The reason for this is that our proof of optimality relies crucially on the fact that $1 - \Re(f(p)\bar{g}(p))$ can be much smaller than $|f(p) - g(p)|$, which is not the case if $f(n)$ and $g(n)$ take values in $[-1,1]$. Thus, if $f(n)$ and $g(n)$ are β-pretentious and real-valued, then we also have that $\widehat{\mathbb{D}}_{\beta,1}(f,g) < \infty$, and so Theorem 5.9 applies.

Second, as the proof of Theorem 5.14 will show, if we have the stronger condition that $S_f(x) = o(x^\alpha)$, then we may conclude that $S_g(x) = o\left(x^\sigma\right)$.

Third, there are quantitative versions of Halász's theorem, due to Halász [39], Montgomery [60], Tenenbaum [76], and Granville and Soundararajan [30] and [32], but all of these theorems are essentially unable to detect cancellation below $O\left(x\frac{\log\log x}{\log x}\right)$, and so are useless for the question of power cancellation. There is also very recent work of Koukoulopoulos [51], who establishes a variant of Halász's theorem allowing detection of cancellation down to the level of $O\left(x\exp\left(-c\sqrt{\log x}\right)\right)$, but, again, this is insufficient for our purposes.

In view of Theorem 5.14, which implies that β-pretentiousness is enough to detect power savings down to $O(x^{(1+\beta)/2})$, it's natural to ask what happens if $(1+\beta)/2 < \alpha$, so that we can detect below the order of magnitude of $S_f(x)$. That is, supposing we have precise information about $S_f(x)$, can we use β-pretentiousness to deduce precise information about $S_g(x)$? This is the content of our final theorem. For convenience, we state the necessary conditions on $f(n)$ and $g(n)$ here.

First, if $f(n)$ and $g(n)$ are both completely multiplicative, we only require that they are β-pretentious to each other for some $\beta > 0$. If, however, either is not completely multiplicative, we must also have that if $S_f(x) \ll_\epsilon x^{\alpha+\epsilon}$ for all $\epsilon > 0$, then $\alpha > 3/4$, and that both of the series $H^2_{2\sigma-1}(f,g)$ and $H^2_{2\sigma-1}(g,f)$ are convergent for some $\sigma < \alpha$.

Theorem 5.15. *Let $f(n)$ and $g(n)$ be as above.*

1. If $S_f(x) = x^\alpha\xi(x)$ for some function $\xi(x)$ satisfying $\xi(t) \ll_\epsilon t^\epsilon$, then $S_g(x) = x^\alpha\tilde{\xi}(x)$ for an explicitly given function $\tilde{\xi}(x)$ also satisfying $\tilde{\xi}(t) \ll_\epsilon t^\epsilon$.

2. If $\xi(t)$ satisfies the mean-square lower bound

$$\int_1^T |\xi(t)|^2 dt \gg_\epsilon T^{1-\epsilon},$$

then $\tilde{\xi}(t)$ does as well.

We have in mind the following two applications of Theorem 5.15: First, if $S_f(x)$ satisfies an asymptotic formula, then so does $S_g(x)$. For example, if the Dirichlet series associated to f, $L(s,f)$, has a finite number of poles on the line $\Re(s)=\alpha$ and is otherwise analytic on $\Re(s)>\alpha-\delta$ for some δ, then standard Tauberian theorems (for example, see [61]) show that

$$S_f(x)=\sum_{\substack{\rho:\Re(\rho)=\alpha\\ \mathrm{ord}_{s=\rho}L(s,f)<0}} x^{\rho}P_{\rho}(\log x)+O(x^{\alpha-\delta+\epsilon}),$$

where each $P_\rho(\log x)$ is a polynomial in $\log x$. Thus, with the notation of Theorem 5.15, we have that

$$\xi(x)=\sum_{\substack{\rho:\ \Re(\rho)=\alpha\\ \mathrm{ord}_{s=\rho}L(s,f)<0}} x^{\Im(\rho)}P_{\rho}(\log x)+O(x^{-\delta+\epsilon}),$$

and it is easy to see that $\xi(x)$ satisfies the required upper bound. Thus, we can apply Theorem 5.15, and it turns out that in this application, $\tilde{\xi}(x)$ works out to be

$$\tilde{\xi}(x)=\sum_{\substack{\rho:\ \Re(\rho)=\alpha\\ \mathrm{ord}_{s=\rho}L(s,f)<0}} x^{\Im(\rho)}Q_{\rho}(\log x)+O(x^{-\delta'})$$

for some suitably small $\delta'>0$, where $Q_\rho(\log x)$ is a polynomial in $\log x$ of the same degree as $P_\rho(\log x)$. Thus, the explicit nature of $\tilde{\xi}(t)$ is of use.

Second, if $S_f(x)$ exhibits a consistent level of cancellation, then so does $S_g(x)$. In the above situation, we made use of the explicit nature of $\tilde{\xi}(x)$ to deduce an asymptotic formula for $S_g(x)$, but in many cases, we would not be lucky enough to have an asymptotic formula for $S_f(x)$ with which to begin. However, it is often possible to deduce the weaker statement that $S_f(x) \not\ll_\epsilon x^{\alpha-\epsilon}$ for any $\epsilon>0$ – for example, $L(s,f)$ may have infinitely many poles on the line $\Re(s)=\alpha$. In this situation, the use of the mean-square lower bound becomes apparent - because $S_f(x)$ exhibits cancellation without satisfying an asymptotic formula, it is likely that $S_f(x)$ could be exceptionally small, perhaps even 0, for some values of x, but it also seems that this occurrence should be fairly rare. We can therefore deduce from Theorem 5.15 that if x^α is the right order of magnitude of $S_f(x)$ in this sense, then x^α is also the right order of magnitude for $S_g(x)$.

This paper is organized as follows: In Section 5.2.1, we consider strong pretentiousness and its relation to the Granville-Soundararajan distances, as discussed in the introduction. Thus, this is where Theorems 5.8-5.13 are proved. In Section 5.2.2, we consider the notion of β-pretentiousness, and establish Theorems 5.14 and 5.15

5.2.1 Strong pretentiousness

In this section, we consider the distances $H_\beta(f,g)$ and $\widehat{\mathbb{D}}_{\beta,k}(f,g)$ and their relation to each other. Thus, we prove Theorems 5.8-5.13, and we do so, in order, in Sections 5.2.1-5.2.1.

Detecting power cancellation

We now let $f(n)$, $g(n)$, and $h(n)$ be as in the hypotheses of Theorem 5.8. Thus, $f(n)$ and $g(n)$ are multiplicative and $h(n)$ is defined by $g(n) = (f * h)(n)$. We now prove Theorem 5.8.

Proof of Theorem 5.8. Suppose that $S_f(x) \ll x^\alpha$ for some $\alpha > 0$. We first claim that the series $\sum_{n=1}^\infty |h(n)|/n^\beta$ is convergent. From this, we conclude that

$$\begin{aligned}\sum_{n\le x} g(n) &= \sum_{m\le x} h(m) \sum_{d\le x/m} f(d)\\ &\ll x^\alpha \sum_{m\le x} \frac{|h(m)|}{m^\alpha}\\ &\ll x^{\max(\alpha,\beta)},\end{aligned}$$

by partial summation. Thus, to establish the theorem, it just remains to show that the series above is convergent. However, this, too, is straightforward, as we have that

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{|h(n)|}{n^{\beta}} &= \prod_p \left(1 + \sum_{k=1}^{\infty} \frac{|h(p^k)|}{p^{k\beta}}\right) \\
&\leq \prod_p \exp\left(\sum_{k=1}^{\infty} \frac{|h(p^k)|}{p^{k\beta}}\right) \\
&= \exp\left(H_\beta(f,g)\right) < \infty.
\end{aligned}$$

Thus, we have proved Theorem 5.8. $\square$

Relation to Granville-Soundararajan distances: Proof of Theorem 5.9

We now wish to relate the finiteness of the distance $\widehat{\mathbb{D}}_\beta(f,g)$ to the finiteness of $H_\beta(f,g)$. For convenience, we recall that

$$H_\beta(f,g) := \sum_{p,k} \frac{|h(p^k)|}{p^{k\beta}}, \qquad \widehat{\mathbb{D}}_{\beta,k}(f,g) := \sum_p \sum_{j \leq k} \frac{|g(p^j) - f(p^j)|}{p^{j\beta}},$$

and that $\widehat{\mathbb{D}}_\beta(f,g) := \widehat{\mathbb{D}}_{\beta,\infty}(f,g)$.

From the definition of $h(n)$, we have that

$$g(p^k) - f(p^k) = \sum_{j=1}^{k} f(p^{k-j}) h(p^j),$$

which, by incorporating all the powers up to n, we may express in terms of the $n \times n$ matrix $A := \left(f(p^{i-j})\right)_{i,j \leq n}$, as

$$A \cdot (h(p), \cdots, h(p^n))^t = (g(p) - f(p), \cdots, g(p^n) - f(p^n))^t.$$

where we have set $f(p^j) = 0$ if $j < 0$. For any $k \geq 1$, define $D_f(k,p)$ to be the determinant of the $k \times k$ matrix (a_{ij}) given by

$$a_{ij} = \begin{cases} f(p^{i-j+1}) & \text{if } i - j + 1 \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

so that $(-1)^k D_f(k,p)$ is the $(n, n-k)$-th entry of the matrix A^{-1}. We now have that

$$h(p^n) = \sum_{k=0}^{n-1} (-1)^k (g(p^{n-k}) - f(p^{n-k})) D_f(k,p).$$

Therefore for $\sigma > 0$ sufficiently large, we have that

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{|h(p^n)|}{p^{n\sigma}} &\leq \sum_{n=1}^{\infty} \left(\sum_{k=1}^{n} |f(p^k) - g(p^k)| \cdot |D_f(n-k,p)| \right) p^{-n\sigma} \\ &= \left(\sum_{n=0}^{\infty} \frac{|D_f(n,p)|}{p^{n\sigma}} \right) \left(\sum_{m=1}^{\infty} \frac{|f(p^m) - g(p^m)|}{p^{m\sigma}} \right). \end{aligned}$$

Of course, at this stage, we would like to sum over p. Lemma 5.16 below states that the first quantity on the right hand side is uniformly bounded for p sufficiently large, say $p > Y_1$, provided that $f(n)$ is not too big, say $f(n) = o(n^\delta)$, and that $\sigma > \delta$. Thus, if we assume that $H_\sigma(f, g; Y_1)$ is finite, when we sum over p, the second summation on the right hand side will yield $\widehat{\mathbb{D}}_\sigma(f,g)$, and the first part of Theorem 5.9 follows.

Lemma 5.16. *If $f(n) = o(n^\delta)$ and $\sigma > \delta$, then for all but finitely many p, the series*

$$\sum_{n=0}^{\infty} \frac{|D_f(n,p)|}{p^{n\sigma}}$$

is convergent and uniformly bounded.

Proof. Let $M(k,p)$ be the maximum of the absolute value of the determinants of the $k \times k$ matrices (a_{ij}) which satisfy

$$|a_{ij}| \leq \begin{cases} p^{(i-j+1)\delta} & \text{if } i - j + 1 \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then, we observe that

$$M(k+1,p) \leq 2p^\delta M(k,p)$$

by cofactor expansion, and that $M(1,p) = p^\delta$. It therefore follows that

$$M(k,p) \le 2^{k-1} p^{k\delta},$$

which implies that the bound

$$|D_f(n,p)| < (2p^\delta)^n$$

holds for all but finitely many p. $\square$

Now, it remains to establish the second part of Theorem 5.9. To do so, we must be able to control the contribution of large prime powers to the sum

$$\sum_{m=1}^{\infty} \frac{|f(p^m) - g(p^m)|}{p^{m\sigma}}.$$

This control is provided by our assumption that $f(n), g(n) = o(n^\delta)$. In particular, it is straightforward to see that

$$\sum_{p>Y_2} \sum_{m=k+1}^{\infty} \frac{|f(p^m) - g(p^m)|}{p^{m\sigma}}$$

must converge for some Y_2, provided that $\sigma > \frac{1}{k+1} + \delta$. Thus, the second part of Theorem 5.9 is obtained with $Y = \max(Y_1, Y_2)$.

Degree d functions: Proof of Theorem 5.11

Suppose that $f(n)$ and $g(n)$ are multiplicative functions of degree d, and that $\widehat{\mathbb{D}}_{\beta,d}(f,g) < \infty$. We first show that $\widehat{\mathbb{D}}_\beta(f,g)$ is finite, and then we consider $H_\beta(f,g)$.

Lemma 5.17. *Let $f(n)$ and $g(n)$ be degree d multiplicative functions, and suppose that $\widehat{\mathbb{D}}_{\beta,d}(f,g) < \infty$. Then $\widehat{\mathbb{D}}_\beta(f,g) < \infty$.*

Proof. We begin with some general notation. For any given pair of integers $k, d \geq 0$, define the homogeneous symmetric polynomials r_k^d and q_k^d of degree k in d variables by

$$r_k^d(x_1, \cdots, x_d) := \begin{cases} 1, & \text{if } k = 0 \\ \displaystyle\sum_{1 \leq i_1 < \cdots < i_k \leq d} x_{i_1} x_{i_2} \cdots x_{i_k}, & \text{if } 1 \leq k \leq d \\ 0, & \text{if } k > d, \end{cases}$$

and

$$q_k^d(x_1, \cdots, x_d) := \sum_{j_1 + \cdots + j_d = k} x_1^{j_1} x_2^{j_2} \cdots x_d^{j_d}.$$

Then for an auxiliary variable X, we have that

$$\begin{aligned} \sum_{k=0}^{\infty} q_k^d X^k &= \prod_{j=1}^{d} \left(\sum_{k=0}^{\infty} x_j^k X^k \right) \\ &= \prod_{j=1}^{d} (1 - x_j X)^{-1} \\ &= \left(\sum_{k=0}^{d} (-1)^k r_k^d X^k \right)^{-1}, \end{aligned}$$

which implies that the identity

$$\sum_{j=0}^{k} (-1)^j r_{k-j}^d q_j^d = 0$$

holds for all $k \geq 1$.

Now, if $f(n)$ is a multiplicative function of degree d, so that $f = f_1 * \cdots * f_d$ where each f_i is completely multiplicative, we have that $f(p^k) = q_k^d(f_1(p), \ldots, f_d(p))$. Thus, if we set $\alpha_k(f, p) = r_k^d(f_1(p), \ldots, f_d(p))$ for $k = 0, \cdots, d$, we have that

$$\sum_{k=0}^{d} (-1)^k \alpha_k(f, p) f(p^{n-k}) = 0$$

for any $n \geq 0$, where, of course, we have set $f(p^r) = 0$ for $r < 0$. In particular, for any multiplicative functions $f(n)$ and $g(n)$ of degree d, since $\alpha_k(f,p) \ll_d 1$, we have that

$$|\alpha_k(f,p) - \alpha_k(g,p)| \ll_d |f(p) - g(p)| + |f(p^2) - g(p^2)| + \cdots + |f(p^d) - g(p^d)|$$

for any $k = 1, \cdots, d$. We are now ready to prove the lemma. Assume that $n \geq d+1$. Observing that $f(p^n) \ll_d n^{d-1}$ and $\alpha_k(f,p) \ll_d 1$, we have

$$\begin{aligned}
|f(p^n) - g(p^n)| &= \left| \sum_{k=1}^{d} (-1)^k \alpha_k(f,p) f(p^{n-k}) + (-1)^k \alpha_k(g,p) g(p^{n-k}) \right| \\
&\ll_d \sum_{k=1}^{d} \left| \alpha_k(f,p) f(p^{n-k}) + \alpha_k(g,p) g(p^{n-k}) \right| \\
&\ll_d \sum_{k=1}^{d} \left[\left| \alpha_k(f,p) \left(f(p^{n-k}) - g(p^{n-k}) \right) \right| \right. \\
&\qquad \left. + \left| g(p^{n-k}) \left(\alpha_k(f,p) - \alpha_k(g,p) \right) \right| \right] \\
&\ll_d \sum_{k=1}^{d} \left| f(p^{n-k}) - g(p^{n-k}) \right| \\
&\qquad + n^{d-1} \left(|f(p) - g(p)| + \cdots + |f(p^d) - g(p^d)| \right),
\end{aligned}$$

Since

$$\sum_{n=1}^{\infty} \frac{|f(p^n) - g(p^n)|}{p^{n\sigma}} \ll_d \sum_{n=1}^{\infty} \frac{n^{d-1}}{p^{n\sigma}}$$

is convergent, this inequality leads to

$$\begin{aligned}
\sum_{n=d+1}^{\infty} \frac{|f(p^n) - g(p^n)|}{p^{n\sigma}} &\ll_d \sum_{k=1}^{d} \sum_{n=d+1}^{\infty} \frac{|f(p^{n-k}) - g(p^{n-k})|}{p^{n\sigma}} \\
&\qquad + \sum_{n=d+1}^{\infty} \frac{n^{d-1}}{p^{n\sigma}} \left(|f(p) - g(p)| + \cdots + |f(p^d) - g(p^d)| \right) \\
&\ll_d \frac{1}{p^{\sigma}} \sum_{n=1}^{\infty} \frac{|f(p^n) - g(p^n)|}{p^{n\sigma}}.
\end{aligned}$$

Therefore for all sufficiently large p, we have

$$\sum_{n=d+1}^{\infty} \frac{|f(p^n) - g(p^n)|}{p^{n\sigma}} \ll_d \sum_{n=1}^{d} \frac{|f(p^n) - g(p^n)|}{p^{n\sigma}}.$$

By summing over p, we get the conclusion. □

It remains to show that $H_\beta(f, g)$ is finite. Recall for each prime p, that

$$\sum_{n=1}^{\infty} \frac{|h(p^n)|}{p^{n\beta}} \leq \left(\sum_{n=0}^{\infty} \frac{|D_f(n,p)|}{p^{n\beta}} \right) \left(\sum_{m=1}^{\infty} \frac{|f(p^m) - g(p^m)|}{p^{m\beta}} \right),$$

where $D_f(n,p)$ is as in Section 5.2.1. We will show that the first summation on the right hand side is uniformly bounded, so that by summing over p and using the finiteness of $\widehat{\mathbb{D}}_\beta(f,g)$, the result follows.

Lemma 5.18. *If $f(n)$ is a degree d multiplicative function and $\sigma > 0$, then, for all p, the series*

$$\sum_{n=0}^{\infty} \frac{|D_f(n,p)|}{p^{n\sigma}}$$

converges and is bounded independent of p.

Proof. Recall that we defined $D_f(k,p)$ so that the equation,

$$h(p^n) = \sum_{k=0}^{n-1} (-1)^k (g(p^{n-k}) - f(p^{n-k})) D_f(k,p),$$

holds. We may think of this as a linear polynomial in the variables $g(p^i)$ for $i = 1, \ldots, n$, and we note that the coefficient of $g(p^{n-j})$ is $D_f(j,p)$ for all j. On the other hand, from the definition of $h(n)$, we have the Euler product identity

$$\prod_p \left(\sum_{n=0}^{\infty} h(p^n) p^{-ns} \right) = \prod_p \left(\sum_{n=0}^{\infty} g(p^n) p^{-ns} \right) (1 - f_1(p)p^{-s}) \ldots (1 - f_d(p)p^{-s}),$$

where the $f_i(n)$ are the constituent completely multiplicative functions of $f(n)$. Thus, $h(p^n)$ can be expressed as a linear combination of the variables $g(p^i)$ for $i = n - d, \ldots, n$. Combining these two observations, we conclude that $D_f(k, p) = 0$ for $k \geq d + 1$. The result follows by noting that each of the $D_f(k, p)$ for $k \leq d$ can be bounded independent of p. $\square$

Good functions

Recall that a multiplicative function $f(n)$ of modulus bounded by 1 is good at p if there are no multiplicative functions $g(n)$, of modulus bounded by 1, such that the series

$$\sum_{k=0}^{\infty} \frac{|h(p^k)|}{p^{k\sigma}}$$

diverges for any $\sigma > 0$, and that $f(n)$ is good if it is good at each prime p. This condition ensures that $H_\sigma(f, g; Y)$ is finite for every $Y > 0$, so the first part of Theorem 5.13 is immediate, that the finiteness of $\widehat{\mathbb{D}}_\beta(f, g)$ implies the finiteness of $H_\beta(f, g)$. The second part, the classification of functions which are good at p, is proved along the following lines.

Recall that we defined

$$F_p(z) := \sum_{k=0}^{\infty} f\left(p^k\right) z^k,$$

and we wish to show that $f(n)$ is good at p if and only if $F_p(z)$ has no zeros in the open unit disc. To do this, we observe that $G_p(z) = F_p(z)H_p(z)$, where $G_p(z)$ and $H_p(z)$ are defined analogously to $F_p(z)$. Since $g(n)$ is bounded by 1, we must have that $G_p(z)$ is holomorphic in the disc. Now, the convergence of

$$\sum_{k=0}^{\infty} |h(p^k)| p^{-k\sigma}$$

is equivalent to the statement that $H_p(z)$ is holomorphic. Thus, the result follows.

5.2.2 β-pretentiousness

In this section, we consider the notion of β-pretentiousness in some detail. Recall that two multiplicative functions $f(n)$ and $g(n)$, both of modulus bounded by 1, are such that the series

$$\mathbb{D}_\beta(f,g) = \sum_p \frac{1 - \Re(f(p)\bar{g}(p))}{p^\beta}$$

converges, then they are said to be β-pretentious. In Section 5.2.2, we establish that if $f(n)$ and $g(n)$ are β-pretentious and if $S_f(x) \ll x^\alpha$, then we can detect power cancellation in $S_g(x)$. In Section 5.2.2, we construct a function $f'(n)$ which is β-pretentious to $f(n)$ and exhibits as little cancellation as possible in view of the estimates established in Section 5.2.2, thereby establishing their optimality. Thus, these two sections comprise the proof of Theorem 5.14. In Section 5.2.2, we establish Theorem 5.15 regarding what happens if we are permitted to detect more cancellation than exists.

Detecting power cancellation

The key result which we use to exhibit cancellation in Theorem 5.14 is the following proposition, which of course is reminiscent of the proof of Theorem 5.8.

Proposition 5.19. *Let $f(n)$ $g(n)$ be as above, and let $h(n)$ be defined by $g(n) = (f * h)(n)$. If the series*

$$\sum_{n=1}^{\infty} \frac{|h(n)|^2}{n^\sigma}$$

is convergent for some $\sigma > 0$, then $S_g(x) \ll x^{\max(\alpha,(1+\sigma)/2)}$. Moreover, if $S_f(x) = o(x^\alpha)$, then $S_g(x) = o(x^{\max(\alpha,(1+\sigma)/2)})$.

Proof. From the definition of $h(n)$, we have that

$$\begin{aligned}
\sum_{n\le x} g(n) &= \sum_{m\le x} h(m) \sum_{d\le x/m} f(d) \\
&\ll x^{\alpha} \sum_{m\le x} \frac{|h(m)|}{m^{\alpha}} \\
&\le x^{\alpha} \left(\sum_{m=1}^{\infty} \frac{|h(m)|^2}{m^{\sigma}} \right)^{1/2} \left(\sum_{m\le x} \frac{1}{m^{2\alpha-\sigma}} \right)^{1/2} \\
&\ll x^{\max(\alpha,(\sigma+1)/2)}.
\end{aligned}$$

If we have the stronger assumption that $S_f(x) = o(x^{\alpha})$, by splitting the sum over m on the first line according to whether m is large and proceeding in the same way, it is easily seen that $S_g(x) = o\left(x^{\max\left(\alpha, \frac{1+\sigma}{2}\right)}\right)$. □

In light of Proposition 5.19, to prove the first part of Theorem 5.14, it suffices to establish the following lemma.

Lemma 5.20. *If $f(n)$, $g(n)$, and $h(n)$ are as above, $|f(n)|, |g(n)| \le 1$ for all n, $f(n)$ and $g(n)$ are β-pretentious for some $\beta > 0$, and $\sigma > 1/2$ is such that $\sigma \ge \beta$, then the series*

$$\sum_{n=1}^{\infty} \frac{|h(n)|^2}{n^{\sigma}}$$

converges if the quantity

$$H(\sigma) = \sum_{p\le 4^{1/\sigma}} \sum_{k=0}^{\infty} \frac{|h(p^k)|^2}{p^{k\sigma}}$$

is finite.

Proof. Since $|g(n)| \le 1$ and $|f(n)| \le 1$, we have that

$$|h(p^k)| \le 2^k$$

for all p and all k. Therefore for $p > 4^{1/\sigma}$, one has

$$\sum_{k=1}^{\infty} \frac{|h(p^k)|^2}{p^{k\sigma}} \leq \frac{1 - \mathrm{Re}(f(p)\overline{g(p)})}{p^\sigma} + \frac{16}{p^{2\sigma}}(1 - 4/p^\sigma)^{-1}.$$

Thus, our assumption that $\sigma \geq \beta$ and that

$$\mathbb{D}_\beta(f, g) = \sum_p \frac{1 - \Re(f(p)\overline{g(p)})}{p^\beta}$$

is finite, together with the assumptions of the lemma, guarantee that the series

$$\sum_{n=1}^{\infty} \frac{|h(n)|^2}{n^\sigma} = \prod_p \left(\sum_{k=0}^{\infty} \frac{|h(p^k)|^2}{p^{k\sigma}} \right)$$

is absolutely convergent. □

To establish the cancellation for completely multiplicative functions claimed in the second part of Theorem 5.14, we have the following lemma.

Lemma 5.21. *If $f(n)$,$g(n)$, and $h(n)$ are as in Lemma 5.20 and $f(n)$ and $g(n)$ are completely multiplicative, then the series*

$$\sum_{n=1}^{\infty} \frac{|h(n)|^2}{n^\beta}$$

is convergent.

Proof. Since $h(p^k) = g(p^{k-1})(g(p) - f(p))$ for all primes p and all $k \geq 1$, we have that

$$|h(p^k)|^2 \leq |g(p) - f(p)|^2 \leq 2(1 - \Re(f(p)\bar{g}(p))).$$

Therefore, we have that

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{|h(n)|^2}{n^\beta} &= \prod_p \left(1 + \sum_{k=1}^{\infty} \frac{|h(p^k)|^2}{p^{k\beta}}\right) \\
&\leq \prod_p \left(1 + \frac{2(1-\Re(f(p)\bar{g}(p)))}{p^\beta}(1-p^{-\beta})^{-1}\right) \\
&\leq \exp\left(\sum_p \frac{2(1-\Re(f(p)\bar{g}(p)))}{p^\beta}(1-2^{-\beta})^{-1}\right) \\
&= \exp\left(2(1-2^{-\beta})^{-1}\mathbb{D}_\beta(f,g)\right) < \infty,
\end{aligned}$$

exactly as desired. □

To establish Theorem 5.14, it now remains to establish the optimality of the bound for completely multiplicative functions.

Optimality

It is worth noting at this point that there is another natural approach to proving the theorem, albeit one that is not entirely within the bounds of the pretentious philosophy. From the relation $g(n) = (f * h)(n)$, we have the Dirichlet series identity

$$L(s,g) = L(s,f)L(s,h).$$

The assumption that $S_f(x) \ll x^\alpha$ translates to $L(s,f)$ being analytic in the right half-plane $\Re(s) > \alpha$ and the assumption that $g(n)$ is β-pretentious to $f(n)$, in light of Lemma 5.20 and the Cauchy-Schwarz inequality, implies that $L(s,h)$ is analytic in the region $\Re(s) > \max\left(3/4, \frac{1+\beta}{2}\right)$. Standard arguments (e.g. Perron's formula) then imply the desired bound for $S_g(x)$. Our proof of optimality will proceed along similar lines. While it is somewhat unfortunate that we have to use this mildly non-pretentious argument, it is not entirely clear how to avoid its use.

Lemma 5.22. *Given any $\beta > 0$ and a completely multiplicative function $f(n)$ of modulus bounded by 1 such that $f(n)$ is 1-pretentious to itself, there is a completely multiplicative function $g(n)$ that is β-pretentious to $f(n)$, and which does not satisfy $S_g(x) \ll x^{(1+\beta)/2-\epsilon}$ for any $\epsilon > 0$.*

Proof. First, we may assume that $L(s, f)$ is analytic in the region $\Re(s) > (1+\beta)/2-\delta$ for some $\delta > 0$, otherwise we could simply take $g(n)$ to be $f(n)$. Let

$$g(p) := e\left(\frac{\omega_p}{p^{\frac{1-\beta}{2}} \log\log p}\right) f(p),$$

where $\omega_p = \pm 1$ is a system of signs to be specified later and, as is standard, $e(x) := e^{2\pi i x}$. It is easy to verify that $g(n)$ is β-pretentious to $f(n)$. Our goal is to force $L(s, h)$ to have a singularity at $s = \frac{1+\beta}{2}$. We compute the Euler product for $L(s, h)$ using the Taylor expansion of $e(x)$, getting that

$$\begin{aligned} L(s,h) &= \prod_p \left(1 + \frac{g(p) - f(p)}{p^s} + O\left(p^{-2s}\right)\right) \\ &= \prod_p \left(1 + \frac{2\pi i \omega_p f(p)}{p^{s+\frac{1-\beta}{2}} \log\log p} + O\left(p^{-2s} + p^{-s-1+\beta}\right)\right). \end{aligned}$$

The convergence of $L(s, h)$ at $s = \frac{1+\beta}{2}$ is thus dictated by the behavior of the series

$$P_f(\tau) := \sum_p \frac{i\omega_p f(p)}{p^\tau \log\log p}$$

as τ tends to 1 from the right. In particular, $L(s, h)$ will have a singularity at $s = \frac{1+\beta}{2}$ if we can force either the real part of $P_f(\tau)$ to tend to infinity, accounting for a (possibly fractional order) pole, or, failing that, to have the real part of $P_f(\tau)$ converge but the imaginary part diverge to infinity, accounting for an essential singularity. Obviously, we now choose ω_p to ensure one of these situations. If the series

$$\sum_p \frac{\Im(f(p))}{p \log\log p}$$

is not absolutely convergent, we choose $\omega_p = -\mathrm{sign}(\Im(f(p)))$, forcing $\Re(P_f(\tau))$ to diverge to infinity. If the series is absolutely convergent, we choose $\omega_p = \mathrm{sign}(\Re(f(p)))$, observing that

$$\begin{aligned} \sum_p \frac{|\Re(f(p))|}{p^\tau \log\log p} + \sum_p \frac{|\Im(f(p))|}{p^\tau \log\log p} &\geq \sum_p \frac{\Re(f(p))^2 + \Im(f(p))^2}{p^\tau \log\log p} \\ &= \sum_p \frac{|f(p)|^2}{p^\tau \log\log p} \\ &\geq \sum_p \frac{1}{p^\tau \log\log p} - \mathbb{D}_1(f,f), \end{aligned}$$

which tends to infinity as $\tau \to 1^+$. We thus have that

$$\Im\left(\sum_p \frac{i\omega_p f(p)}{p\log\log p}\right) = \sum_p \frac{|\Re(f(p))|}{p\log\log p} = \infty,$$

from which we conclude that $\Im(P_f(x))$ tends to infinity. We have thus constructed $g(n)$ so that $L(s,h)$ has a singularity at $s = \frac{1+\beta}{2}$, so provided that $L\left(\frac{1+\beta}{2}, f\right) \neq 0$, we obtain the result. If $L\left(\frac{1+\beta}{2}, f\right) = 0$, there is a $t \in \mathbb{R}$ such that $L\left(\frac{1+\beta}{2} + it, f\right) \neq 0$. We make the obvious modifications to the construction above to force $L(s,h)$ to have a singularity at $s = \frac{1+\beta}{2} + it$. $\square$

Asymptotic formulae

We now suppose we are in the situation of Theorem 5.15. That is, we assume that $f(n)$ is multiplicative, of modulus bounded by 1, and is such that

$$S_f(x) = x^\alpha \xi(x)$$

for some function $\xi(x)$ satisfying $\xi(t) \ll_\epsilon t^\epsilon$ for all $\epsilon > 0$, and we also assume that $\beta < 2\alpha - 1$. In addition, if $f(n)$ is not completely multiplicative, we assume that $\alpha > 3/4$ and that the series $H^2_{2\sigma-1}(f,g)$ and $H^2_{2\sigma-1}(g,f)$ are convergent. To establish a formula for $S_g(x)$, we note that

$$\begin{aligned}\sum_{n\le x} g(n) &= \sum_{m\le x} h(m) \sum_{d\le x/m} f(d)\\ &= x^{\alpha}\sum_{m\le x}\frac{h(m)}{m^{\alpha}}\xi(x/m),\end{aligned}$$

and so we naturally define $\tilde{\xi}(x)$ to be the convolution

$$\tilde{\xi}(x) := \sum_{m\le x}\frac{h(m)}{m^{\alpha}}\xi(x/m).$$

To see that $\tilde{\xi}(x) \ll x^{\epsilon}$, we merely note that

$$|\tilde{\xi}(x)| \le \sum_{m\le x}\frac{|h(m)|}{m^{\alpha}}|\xi(x/m)| \ll_{\epsilon} x^{\epsilon}\sum_{m\le x}\frac{|h(m)|}{m^{\alpha+\epsilon}}.$$

Our assumptions guarantee that the series on the right is convergent, whence the claimed bound. Now, suppose that

$$\int_1^T |\xi(t)|^2 dt \gg_{\epsilon} T^{1-\epsilon}.$$

Möbius inversion gives that

$$\xi(x) = \sum_{m\le x}\frac{\tilde{h}(m)}{m^{\alpha}}\tilde{\xi}(x/m),$$

where $\tilde{h}(n)$ is the Dirichlet inverse of $h(n)$ (i.e., $(h*\tilde{h})(1) = 1$ and $(h*\tilde{h})(n) = 0$ for $n > 1$). Using this and the Cauchy-Schwarz inequality in the above, we obtain that

$$\begin{aligned}T^{1-\epsilon} &\ll_{\epsilon} \int_1^T \left(\sum_{m\le t}\frac{|\tilde{h}(m)|^2}{m^{\beta}}\right)\left(\sum_{m\le t}\frac{|\tilde{\xi}(t/m)|^2}{m^{2\alpha-\beta}}\right)dt\\ &\le \sum_{m=1}^{\infty}\frac{|\tilde{h}(m)|^2}{m^{\beta}}\int_1^T\sum_{m\le t}\frac{|\tilde{\xi}(t/m)|^2}{m^{2\alpha-\beta}}dt\\ &= \sum_{m=1}^{\infty}\frac{|\tilde{h}(m)|^2}{m^{\beta}}\sum_{m\le T}\frac{1}{m^{2\alpha-\beta-1}}\int_1^{T/m}|\tilde{\xi}(t)|^2 dt\\ &\ll T^{2-2\alpha+\beta}\int_1^T|\tilde{\xi}(t)|^2\frac{dt}{t^{2-2\alpha+\beta}},\end{aligned}$$

where the infinite series is convergent by assumption, so we have absorbed it into the implied constant. Now, let

$$I := \int_1^T |\tilde{\xi}(t)|^2 dt,$$

and apply Hölder's inequality to get that

$$\begin{aligned}
\int_1^T |\tilde{\xi}(t)|^2 \frac{dt}{t^{2-2\alpha+\beta}} &\leq I^{\frac{2\alpha-\beta-1}{2}} \left(\int_1^T \frac{|\tilde{\xi}(t)|^2}{t^{\frac{2(2-2\alpha+\beta)}{3-2\alpha+\beta}}} dt \right)^{\frac{3-2\alpha+\beta}{2}} \\
&\ll_\epsilon I^{\frac{2\alpha-\beta-1}{2}} \left(\int_1^T t^{\frac{-2(2-2\alpha+\beta)}{3-2\alpha+\beta}+\epsilon} dt \right)^{\frac{3-2\alpha+\beta}{2}} \\
&\ll I^{\frac{2\alpha-\beta-1}{2}} T^{\frac{2\alpha-\beta-1}{2}+\epsilon}.
\end{aligned}$$

Using this in the above, we obtain that

$$I^{\frac{2\alpha-\beta-1}{2}} T^{\frac{3-2\alpha+\beta}{2}+\epsilon} \gg_\epsilon T^{1-\epsilon},$$

and so we have that

$$I^{\frac{2\alpha-\beta-1}{2}} \gg_\epsilon T^{\frac{2\alpha-\beta-1}{2}-\epsilon},$$

and the result follows, concluding the proof of Theorem 5.15.

Since the Dirichlet series $L(s,h)$ for $\Re(s) \geq \alpha$ plays a critical role in the definition of $\tilde{\xi}(x)$, it is useful to know whether it is 0. In particular, in applying Theorem 5.15 in the case when $S_f(x)$ satisfies an asymptotic formula, we might potentially lose a term in our formula if $L(\rho,h) = 0$ for some pole ρ of $L(s,f)$. However, we have the following simple observation.

Lemma 5.23. *If $f(n)$ and $g(n)$ are completely multiplicative and as above, then the Dirichlet series $L(s,h)$ associated to $h(n)$ is non-zero in the region $\Re(s) > (1+\beta)/2$.*

Proof. Since $h(n)$ is defined by the relation $g = f * h$, we have the Dirichlet series formula

$$L(s,h) = \frac{L(s,g)}{L(s,f)}.$$

By Lemma 5.21, this is absolutely convergent in the region $\Re(s) > (1+\beta)/2$. If we define $\tilde{h}(n)$ by $f = g * \tilde{h}$, the same argument applies to $L(s, \tilde{h})$. Since we also have that

$$L(s, \tilde{h}) = \frac{1}{L(s, h)},$$

this immediately yields the result. □

Of course, if $f(n)$ and $g(n)$ are not completely multiplicative, the analog of Lemma 5.23 can still be obtained with Lemma 5.20 replacing Lemma 5.21.

Bibliography

[1] Henry L. Alder. The nonexistence of certain identities in the theory of partitions and compositions. *Bull. Amer. Math. Soc.*, 54:712–722, 1948.

[2] George E. Andrews. On a partition problem of H. L. Alder. *Pacific J. Math.*, 36:279–284, 1971.

[3] George E. Andrews. *The theory of partitions*. Addison-Wesley Publishing Co., Reading, Mass.-London-Amsterdam, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2.

[4] George E. Andrews and Kimmo Eriksson. *Integer partitions*. Cambridge University Press, Cambridge, 2004.

[5] Tom M. Apostol. *Introduction to analytic number theory*. Springer-Verlag, New York, 1976. Undergraduate Texts in Mathematics.

[6] Paul T. Bateman and Roger A. Horn. A heuristic asymptotic formula concerning the distribution of prime numbers. *Math. Comp.*, 16:363–367, 1962.

[7] Bruce C. Berndt, Ronald J. Evans, and Kenneth S. Williams. *Gauss and Jacobi sums*. Canadian Mathematical Society Series of Monographs and Advanced Texts. John Wiley & Sons Inc., New York, 1998. A Wiley-Interscience Publication.

[8] M. Bhargava and J. Hanke. Universal quadratic forms and the 290-theorem. *Preprint.*

[9] Manjul Bhargava. On the Conway-Schneeberger fifteen theorem. In *Quadratic forms and their applications (Dublin, 1999)*, volume 272 of *Contemp. Math.*, pages 27–37. Amer. Math. Soc., Providence, RI, 2000.

[10] Spencer Bloch and Kazuya Kato. L-functions and Tamagawa numbers of motives. In *The Grothendieck Festschrift, Vol. I*, volume 86 of *Progr. Math.*, pages 333–400. Birkhäuser Boston, Boston, MA, 1990.

[11] Valentin Blomer and Gergely Harcos. Hybrid bounds for twisted L-functions. *J. Reine Angew. Math.*, 621:53–79, 2008.

[12] V Bouniakowsky. Nouveaux théorèmes relatifs à la distinction des nombres premiers et à la décomposition des entiers en facteurs. *Sc. Math. Phys.*, 6:305–329, 1857.

[13] Kathrin Bringmann and Ken Ono. Coefficients of harmonic maass forms. accepted for publication.

[14] Kathrin Bringmann and Ken Ono. Coefficients of harmonic weak maass forms. *Proc. of the 2008 Univ. of Flor. Conf.*, accepted for publication.

[15] Jan Hendrik Bruinier, Gerard van der Geer, Günter Harder, and Don Zagier. *The 1-2-3 of modular forms.* Universitext. Springer-Verlag, Berlin, 2008. Lectures from the Summer School on Modular Forms and their Applications held in Nordfjordeid, June 2004, Edited by Kristian Ranestad.

[16] A. A. Buhštab. Combinatorial strengthening of the sieve of Eratosthenes method. *Uspehi Mat. Nauk*, 22(3 (135)):199–226, 1967.

[17] Vorrapan Chandee. Explicit upper bounds for L-functions on the critical line. *Proc. Amer. Math. Soc.*, 137(12):4049–4063, 2009.

[18] H. Cohen and D. Zagier. Vanishing and non-vanishing theta values. *Ann. Sci. Math. Québec*, page to appear.

[19] Alina Carmen Cojocaru and M. Ram Murty. *An introduction to sieve methods and their applications*, volume 66 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2006.

[20] Harold Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2000. Revised and with a preface by Hugh L. Montgomery.

[21] Fred Diamond and Jerry Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[22] William Duke and Rainer Schulze-Pillot. Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids. *Invent. Math.*, 99(1):49–57, 1990.

[23] D. Dummit, H. Kisilevsky, and J. McKay. Multiplicative products of η-functions. In *Finite groups—coming of age (Montreal, Que., 1982)*, volume 45 of *Contemp. Math.*, pages 89–98. Amer. Math. Soc., Providence, RI, 1985.

[24] Ronald J. Evans. Generalizations of a theorem of Chowla on Gaussian sums. *Houston J. Math.*, 3(3):343–349, 1977.

[25] John Friedlander and Henryk Iwaniec. The polynomial X^2+Y^4 captures its primes. *Ann. of Math. (2)*, 148(3):945–1040, 1998.

[26] Dorian Goldfeld and Lucien Szpiro. Bounds for the order of the Tate-Shafarevich group. *Compositio Math.*, 97(1-2):71–87, 1995. Special issue in honour of Frans Oort.

[27] Dorian M. Goldfeld. The class number of quadratic fields and the conjectures of Birch and Swinnerton-Dyer. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 3(4):624–663, 1976.

[28] Daniel A. Goldston, János Pintz, and Cem Y. Yıldırım. Primes in tuples. I. *Ann. of Math. (2)*, 170(2):819–862, 2009.

[29] Andrew Granville. Pretentiousness in analytic number theory. *J. Théor. Nombres Bordeaux*, 21(1):159–173, 2009.

[30] Andrew Granville and K. Soundararajan. *Multiplicative number theory.* in preparation.

[31] Andrew Granville and K. Soundararajan. The spectrum of multiplicative functions. *Ann. of Math. (2)*, 153(2):407–470, 2001.

[32] Andrew Granville and K. Soundararajan. Decay of mean values of multiplicative functions. *Canad. J. Math.*, 55(6):1191–1230, 2003.

[33] Andrew Granville and K. Soundararajan. Large character sums: pretentious characters and the Pólya-Vinogradov theorem. *J. Amer. Math. Soc.*, 20(2):357–384 (electronic), 2007.

[34] Andrew Granville and Kannan Soundararajan. Pretentious multiplicative functions and an inequality for the zeta-function. In *Anatomy of integers*, volume 46 of *CRM Proc. Lecture Notes*, pages 191–197. Amer. Math. Soc., Providence, RI, 2008.

[35] Andrew Granville and H. M. Stark. *abc* implies no "Siegel zeros" for L-functions of characters with negative discriminant. *Invent. Math.*, 139(3):509–523, 2000.

[36] Benedict H. Gross and Neal Koblitz. Gauss sums and the p-adic Γ-function. *Ann. of Math. (2)*, 109(3):569–581, 1979.

[37] Benedict H. Gross and Don B. Zagier. Heegner points and derivatives of L-series. *Invent. Math.*, 84(2):225–320, 1986.

[38] G. Halász. Über die Mittelwerte multiplikativer zahlentheoretischer Funktionen. *Acta Math. Acad. Sci. Hungar.*, 19:365–403, 1968.

[39] G. Halász. On the distribution of additive and the mean values of multiplicative arithmetic functions. *Studia Sci. Math. Hungar.*, 6:211–233, 1971.

[40] Jonathan Hanke. Local densities and explicit bounds for representability by a quadratric form. *Duke Math. J.*, 124(2):351–388, 2004.

[41] D. R. Heath-Brown. Primes represented by $x^3 + 2y^3$. *Acta Math.*, 186(1):1–84, 2001.

[42] Christopher Hooley. On the greatest prime factor of a quadratic polynomial. *Acta Math.*, 117:281–299, 1967.

[43] Kenneth Ireland and Michael Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[44] Henryk Iwaniec. Almost-primes represented by quadratic polynomials. *Invent. Math.*, 47(2):171–188, 1978.

[45] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[46] T. Jagathesan and M. Manickam. On Shimura correspondence for non-cusp forms of half-integral weight. *J. Ramanujan Math. Soc.*, 23(3):211–222, 2008.

[47] William C. Jagy, Irving Kaplansky, and Alexander Schiemann. There are 913 regular ternary forms. *Mathematika*, 44(2):332–341, 1997.

[48] Dimitar Jetchev and Ben Kane. Equidistribution of Heegner points and ternary quadratic forms. *Math. Ann.*, 350(3):501–532, 2011.

[49] Jerzy Kaczorowski and Alberto Perelli. On the structure of the Selberg class. I. $0 \leq d \leq 1$. *Acta Math.*, 182(2):207–241, 1999.

[50] Ben Kane. Representations of integers by ternary quadratic forms. *Int. J. Number Theory*, 6(1):127–159, 2010.

[51] D. Koukoulopoulos. On multiplicative functions which are small on average. *Preprint.*

[52] P. Kuhn. Über die primteiler eines polynoms. proceedings of the International Congress of Mathematicians, Amsterdam. 2:35–37, 1954.

[53] Marc Laborde. Buchstab's sifting weights. *Mathematika*, 26(2):250–257 (1980), 1979.

[54] B. V. Levin. A one-dimensional sieve. *Acta Arith.*, 10:387–397, 1964/1965.

[55] Yves Martin. Multiplicative η-quotients. *Trans. Amer. Math. Soc.*, 348(12):4825–4856, 1996.

[56] Günter Meinardus. Asymptotische Aussagen über Partitionen. *Math. Z.*, 59:388–398, 1954.

[57] Günter Meinardus. Über Partitionen mit Differenzenbedingungen. *Math. Z.*, 61:289–302, 1954.

[58] Gerd Mersmann. Holomorphe η-produkte und nichtverschwindende ganze modulformen für $\gamma_0(n)$. *Master's Thesis, Univ. of Bonn*, 1991.

[59] Preda Mihăilescu. Primary cyclotomic units and a proof of Catalan's conjecture. *J. Reine Angew. Math.*, 572:167–195, 2004.

[60] H. Montgomery. A note on mean values of multiplicative functions. *Report No. 17, Institut Mittag-Leffler, Djursholm*, 1978.

[61] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative number theory. I. Classical theory*, volume 97 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2007.

[62] Jürgen Neukirch. *Algebraic number theory*, volume 322 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999. Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder.

[63] Byeong-Kweon Oh. Regular positive ternary quadratic forms. *Acta Arith.*, 147(3):233–243, 2011.

[64] Ken Ono. *The web of modularity: arithmetic of the coefficients of modular forms and q-series*, volume 102 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2004.

[65] Ken Ono and K. Soundararajan. Ramanujan's ternary quadratic form. *Invent. Math.*, 130(3):415–454, 1997.

[66] Hans Rademacher. On the expansion of the partition function in a series. *Ann. of Math. (2)*, 44:416–422, 1943.

[67] Hans Rademacher and Herbert S. Zuckerman. On the Fourier coefficients of certain modular forms of positive dimension. *Ann. of Math. (2)*, 39(2):433–462, 1938.

[68] H.-E. Richert. Selberg's sieve with weights. *Mathematika*, 16:1–22, 1969.

[69] Jeremy Rouse. Quadratic forms representing all odd positive integers. *Preprint.*

[70] A. Schinzel and W. Sierpiński. Sur certaines hypothèses concernant les nombres premiers. *Acta Arith. 4 (1958), 185–208; erratum*, 5:259, 1958.

[71] J.-P. Serre and H. M. Stark. Modular forms of weight 1/2. In *Modular functions of one variable, VI (Proc. Second Internat. Conf., Univ. Bonn, Bonn, 1976)*, pages 27–67. Lecture Notes in Math., Vol. 627. Springer, Berlin, 1977.

[72] Jean-Pierre Serre. Divisibilité de certaines fonctions arithmétiques. *Enseignement Math. (2)*, 22(3-4):227–260, 1976.

[73] Jean-Pierre Serre. Sur la lacunarité des puissances de η. *Glasgow Math. J.*, 27:203–221, 1985.

[74] K. Soundararajan. Degree 1 elements of the Selberg class. *Expo. Math.*, 23(1):65–70, 2005.

[75] V. V. Subrahmanyasastri. Partitions with congruence conditions. *J. Indian Math. Soc. (N.S.)*, 36:177–194 (1973), 1972.

[76] Gérald Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 46 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995.

[77] J.-L. Waldspurger. Sur les coefficients de Fourier des formes modulaires de poids demi-entier. *J. Math. Pures Appl. (9)*, 60(4):375–484, 1981.

[78] Yuan Wang. On sieve methods and some of their applications. *Sci. Sinica*, 11:1607–1624, 1962.

[79] Jin Yang and Weizhe Zheng. On a theorem of Chowla. *J. Number Theory*, 106(1):50–55, 2004.

[80] Ae Ja Yee. Partitions with difference conditions and Alder's conjecture. *Proc. Natl. Acad. Sci. USA*, 101(47):16417–16418 (electronic), 2004.

[81] Ae Ja Yee. Alder's conjecture. *J. Reine Angew. Math.*, 616:67–88, 2008.

[82] Herbert S. Zuckerman. On the coefficients of certain modular forms belonging to subgroups of the modular group. *Trans. Amer. Math. Soc.*, 45(2):298–321, 1939.

[83] Herbert S. Zuckerman. On the expansions of certain modular forms of positive dimension. *Amer. J. Math.*, 62:127–152, 1940.