New Bounds for Automorphic L-functions

Gergely Harcos

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the
Department of Mathematics

June, 2003

Abstract

This dissertation contributes to the analytic theory of automorphic L-functions.

We prove an approximate functional equation for the central value of the L-series attached to an irreducible cuspidal automorphic representation π of GL_m over a number field. The approximation involves a smooth truncation of the Dirichlet series $L(s,\pi)$ and $L(s,\tilde{\pi})$ after about $\sqrt{C}$ terms, where C denotes the analytic conductor (of π and $\tilde{\pi}$ at the central point) introduced by Iwaniec and Sarnak. We investigate the decay rate of the cutoff function and its derivatives. We also see that the truncation can be made uniformly explicit at the cost of an error term. The results extend to products of central values.

We establish, via the Hardy–Littlewood circle method, a nontrivial bound on shifted convolution sums of Fourier coefficients coming from classical holomorphic or Maass cusp forms of arbitrary level and nebentypus. These sums are analogous to the binary additive divisor sum which has been studied extensively. We achieve polynomial uniformity in all the parameters of the cusp forms by carefully estimating the Bessel functions that enter the analysis. As an application we derive, extending work of Duke, Friedlander and Iwaniec, a subconvex estimate on the critical line for L-functions associated to character twists of these cusp forms.

We also study the shifted convolution sums via the Sarnak–Selberg spectral method. For holomorphic cusp forms this approach detects optimal cancellation over any totally real number field. For Maass cusp forms the method is burdened with complicated integral transforms. We succeed in inverting the simplest of these transforms whose kernel is built up of Gauss hypergeometric functions.

Acknowledgements

My gratitude to my advisor, Peter Sarnak, is due not only to his guidance in the preparation of the current work, but, most of all, to his permanent influence on my mathematical outlook. I feel fortunate to have glimpsed, at times, the mathematical purview provided by his wide perspective.

I should also manifest my debt to all my teachers, and to all companions who helped to form me as a mathematician. My friends, in general, are to be thanked for their valued support. What I owe to my family cannot be distinguished from who I am and who I am becoming; I can only ascertain that most of my worth comes from them. Before all, I thank my father, Péter Harcos, and Yvette Vajda, my bride.

Revised version (September, 2003). I am indebted to Florin Spinu for having pointed out an error in Chapter 2 of the original document. The error and several misprints have been corrected.

To My Father

Contents

Chapter 1

Introduction

1.1 Prologue

L-functions are among the most fundamental and most fascinating objects in number theory. An L-function can be attached to

(1) a smooth projective variety defined over a number field (Hasse, Weil),

(2) an irreducible complex or l-adic representation of the Galois group of a number field (Artin, Grothendieck), or

(3) a cusp form or irreducible cuspidal automorphic representation (Hecke, Langlands, Godement–Jacquet).

An L-function is defined in terms of local data. In each of the cases above, this local data consists of

(1) the number of points of the reduction of the projective variety to various finite fields,

(2) the eigenvalues of the Frobenius elements in the Galois group, or

(3) the Langlands parameters of the automorphic form or representation.

By definition, the L-function is given as an Euler product over the rational primes of the local data:

$$L(s) = \prod_p L_p(s).$$

Various results and conjectures relating these objects add up to the general philosophy that every L-function of arithmetic nature is a ratio of automorphic L-functions.

Besides their many combinatorial and algebraic properties, L-functions are very much analytic objects. Understanding their analytic behaviour is an important task, especially if it gives rise to arithmetic implications. A classical example is Chebotarev's density theorem on Frobenius elements in the Galois group. The analytic properties of an L-function are most accessible when the L-function is known to come from an automorphic form. Even in this case, our knowledge is surprisingly limited. It has been realized only recently how widely such knowledge could be applied to deep diophantine problems.

1.2 Size of an L-function

A foremost issue in such applications is that of the size of an L-function. Let us first fix our notation for a general discussion. We consider a number field F of degree d and an irreducible cuspidal automorphic representation π of GL_m over F with unitary central character. By Flath's theorem, π can be written uniquely as a restricted tensor product $\otimes_v \pi_v$, where π_v is an irreducible admissible representation of $\mathrm{GL}_m(F_v)$ for each place v of F. Accordingly, the complete L-function associated to π is defined as a product of local L-functions,

$$\Lambda(s, \pi) = \prod_v L(s, \pi_v).$$

It is convenient to collect the local factors for v underlying a given rational place w, and introduce the subproducts

$$L(s,\pi_w)=\prod_{v|w}L(s,\pi_v).$$

For the infinite place $w=\infty$ the subproduct takes the form

$$L(s,\pi_\infty)=\prod_{j=1}^{md}\pi^{\frac{\mu_j-s}{2}}\Gamma\left(\frac{s-\mu_j}{2}\right), \tag{1.1}$$

while for a finite rational prime $w=p$ we have

$$L(s,\pi_p)=\prod_{j=1}^{md}\frac{1}{1-\alpha_j(p)p^{-s}}. \tag{1.2}$$

(Note that π inside the first product refers to the positive constant, not the representation.) The numbers μ_j (resp. $\alpha_j(p)$) are called the Archimedean (resp. non-Archimedean) Langlands parameters and satify the following uniform bound by Theorem 1 of [Lu-Ru-Sa].

Theorem 1.1 (Luo–Rudnick–Sarnak).

$$\sup\{\Re\mu_j,\Re\log_p\alpha_j(p)\}\leq\frac{1}{2}-\frac{1}{m^2+1}. \tag{1.3}$$

The local factor $L(s,\pi_\infty)$ is distinguished in the sense that in vertical strips it decays exponentially while the other factors $L(s,\pi_p)$ remain bounded away from 0. This fact alone provides ample justification for isolating the finite part

$$L(s,\pi)=\prod_{p<\infty}L(s,\pi_p),\quad \Re s>\frac{3}{2}-\frac{1}{m^2+1}, \tag{1.4}$$

an absolutely convergent Euler product over the rational primes by (1.3). The resulting complete

$$\Lambda(s,\pi) = L(s,\pi_\infty)L(s,\pi)$$

extends to an entire function which is bounded in vertical strips (except for $\pi = |\det|^{it}$ when a simple pole occurs at $s = 1 - it$), and satisfies a functional equation of the form

$$N^{\frac{s}{2}}\Lambda(s,\pi) = \kappa N^{\frac{1-s}{2}}\Lambda(1-s,\tilde{\pi}). \tag{1.5}$$

N is the arithmetic conductor (a positive integer), κ is the root number (of modulus 1), and $\tilde{\pi}$ is the contragradient representation of π. The local L-functions of π and $\tilde{\pi}$ are connected by

$$\bar{L}(s,\pi_v) = L(\bar{s},\tilde{\pi}_v). \tag{1.6}$$

It is natural to expect that $L(s,\pi)$ has a moderate size in vertical strips, so that $\Lambda(s,\pi)$ inherits the exponential decay of the Archimedean factor $L(s,\pi_\infty)$. We shall formulate a more precise and more general statement using the analytic conductor introduced by Iwaniec and Sarnak [Iw-Sa]:

$$C(s,\pi) = \frac{N}{(2\pi)^{md}}\prod_{j=1}^{md}|s-\mu_j|.$$

In order to bound automorphic L-functions, it is essential to represent them as absolutely convergent Dirichlet series

$$L(s,\pi) = \sum_{n=1}^{\infty}\frac{\lambda_\pi(n)}{n^s}. \tag{1.7}$$

Certainly (1.2), (1.3) and (1.4) guarantee that $L(s,\pi)$ acquires this form in the half-plane $\Re s > \frac{3}{2} - \frac{1}{m^2+1}$. As a by-product, we also see that the coefficients satisfy

$$\lambda_\pi(n) \ll_{\epsilon,m,d} n^{\frac{1}{2}-\frac{1}{m^2+1}+\epsilon}$$

for any $\epsilon > 0$. Upon the Ramanujan–Selberg conjectures we could replace the occurrences of $\frac{1}{2} - \frac{1}{m^2+1}$ in (1.3) and in the previous inequality by 0. These improved bounds hold unconditionally in a certain average form by Theorem 4 of [Mol].

Theorem 1.2 (Molteni). *Uniformly in $\epsilon > 0$ and $x > 0$,*

$$\sum_{n \le x} |\lambda_\pi(n)| \ll_\epsilon x^{1+\epsilon} C\left(\tfrac{1}{2}, \pi\right)^\epsilon. \tag{1.8}$$

The implied constant depends only on ϵ, m and d.

It should be noted that Molteni assumes $\Re\mu_j \le 0$ for all j (cf. axiom (A4) in [Mol]), but his argument works equally well with the weaker bound (1.3). In particular, the Dirichlet series (1.7) is absolutely convergent in the larger half-plane $\Re s > 1$, and it satisfies

$$L(\sigma, \pi) \ll_{\sigma,\epsilon,m,d} C\left(\tfrac{1}{2}, \pi\right)^\epsilon, \quad \sigma > 1. \tag{1.9}$$

By replacing π with its twist $\pi \otimes |\det|^{it}$ this becomes

$$L(\sigma + it, \pi) \ll_{\sigma,\epsilon,m,d} C\left(\tfrac{1}{2} + it, \pi\right)^\epsilon, \quad \sigma > 1. \tag{1.10}$$

We can combine (1.9) with the functional equation (1.5) to deduce uniform bounds in the half-plane $\Re s < 0$. First,

$$L(\sigma, \pi) \ll_{\sigma,\epsilon,m,d} C\left(\tfrac{1}{2}, \pi\right)^{1/2-\sigma+\epsilon}, \quad \sigma < 0,$$

and then, by replacing π with $\pi \otimes |\det|^{it}$,

$$L(\sigma + it, \pi) \ll_{\sigma,\epsilon,m,d} C\left(\tfrac{1}{2} + it, \pi\right)^{1/2-\sigma+\epsilon}, \quad \sigma < 0. \tag{1.11}$$

Finally, we can interpolate between (1.10) and (1.11) by the Phragmén–Lindelöf convexity principle to obtain bounds inside the critical strip $0 < \Re s < 1$ (or on the boundaries away from the possible pole).

Convexity Bound. *For any* $0 < \sigma < 1$ *and any* $\epsilon > 0$*, there is a uniform bound*

$$L(\sigma + it, \pi) \ll_{\sigma,\epsilon} C\left(\tfrac{1}{2} + it, \pi\right)^{(1-\sigma)/2+\epsilon}. \tag{1.12}$$

The implied constant depends only on σ*,* ϵ*,* m *and* d*.*

The expontents given by (1.10) and (1.11) are sharp. We expect, however, that a much stronger inequality holds in place of the convexity bound.

Generalized Lindelöf Hypothesis. *For any* $0 < \sigma < 1$ *and any* $\epsilon > 0$*, there is a uniform bound*

$$L(\sigma + it, \pi) \ll_{\sigma,\epsilon} C\left(\tfrac{1}{2} + it, \pi\right)^{\max(0,1-2\sigma)/2+\epsilon}. \tag{1.13}$$

The implied constant depends only on σ*,* ϵ*,* m *and* d*.*

This very powerful statement is a consequence of the generalized Riemann hypothesis that all the roots of $\Lambda(s, \pi)$ lie on the critical line $\Re s = \frac{1}{2}$. In fact, the resolution of several deep equidistribution questions in number theory relies on a small but substantial improvement on the convexity bound in certain families of automorphic L-functions. For convenience and applicability we focus on the critical line $\Re s = \frac{1}{2}$.

Subconvexity Problem. *Show that there is a* $\delta = \delta(m, d) > 0$ *such that*

$$L(s, \pi) \ll_{m,d} C(s, \pi)^{1/4-\delta}, \quad \Re s = \frac{1}{2}. \tag{1.14}$$

Applications include equidistribution of lattice points on ellipsoids (Linnik's problem), characterization of integers represented by a given quadratic form over a number field (Hilbert's 11th problem), equidistribution of certain Galois orbits of CM-points on Shimura varieties (evidence toward the André–Oort conjecture), and equidistribution of mass in arithmetic quantum chaos.

1.3 Approximate functional equation

It is not obvious that the coefficients $\lambda_\pi(n)$ can be used to reveal the finer behaviour of $L(s,\pi)$ in the critical strip $0<\Re s<1$. This was originally realized for the Riemann zeta function

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

by Hardy and Littlewood in 1921 [Har-Lit]. They established an approximation to $\zeta(s)$, called an *approximate functional equation*, a special case of which reads as follows:

$$\zeta\left(\frac{1}{2}+it\right)=\sum_{n\leq\sqrt{\frac{|t|}{2\pi}}}\frac{1}{n^{\frac{1}{2}+it}}+\frac{\zeta\left(\frac{1}{2}+it\right)}{\zeta\left(\frac{1}{2}-it\right)}\sum_{n\leq\sqrt{\frac{|t|}{2\pi}}}\frac{1}{n^{\frac{1}{2}-it}}+O\big(|t|^{-\frac{1}{4}}\log|t|\big).$$

Note that the factor in front of the second sum is of modulus 1 and does not destroy the symmetry $t\leftrightarrow -t$. This formula was extended and studied by many researchers with focus generally restricted to small powers of Dirichlet L-functions or Dedekind L-functions. Among the few studies with a larger scope the most notable ones are by Chandrasekharan and Narasimhan [Ch-Na], Lavrik [La], and Ivić [Iv].

In Chapter 2 we shall present uniform variants of the approximate functional equation for all automorphic L-functions. We shall demonstrate that the values of $L(s,\pi)$ on the critical line $\Re s=\frac{1}{2}$ can be approximated as a sum of two Dirichlet series which have essentially $\sqrt{C(s,\pi)}$ terms. The relevance of the analytic conductor has not been displayed in this general context before. In fact, we had to do some "fine tuning" on the original analytic conductor of Iwaniec and Sarnak [Iw-Sa] in order to achieve our goal.

The result we obtain fits well into the philosophy that L-functions (or rather, L-values) should be considered in families [Iw-Sa]. We shall employ smooth cutoff functions as they are more natural for the problem and also yield better error terms. First we obtain an exact representation by an implicit cutoff function with uniform decay properties (Theorem 2.1). This formula is most useful for families whose Archimedean parameters remain bounded. The second representation (Theorem 2.2), inspired by the recent work of Ivić [Iv], has a more explicit main term at the cost of an error term. This formula works best in families where the Archimedean parameters grow large simultaneously. The proofs are based on standard Mellin transform techniques, and they make crucial use of the estimates of Luo–Rudnick–Sarnak (1.3) and Molteni (1.8). A variant of the method yields similar formulae for products of central values (e.g. for higher moments).

1.4 Amplification

The approximate functional equation reduces the subconvexity problem to cancellation in finite smooth sums

$$S(X,\pi)=\sum_{n=1}^{\infty}\lambda_\pi(n)w\left(\frac{n}{X}\right),$$

where $w:(0,\infty)\to\mathbb{C}$ is a fixed weight function of compact support on the positive axis. More precisely, by combining Corollary 2.1 with a smooth decomposition of unity, we can see that a variant of (1.14),

$$\forall\epsilon>0:\forall t\in\mathbb{R}:\quad L(\tfrac12+it,\pi)\ll_{\epsilon,m,d}C\left(\tfrac12+it,\pi\right)^{1/4-\delta+\epsilon},\tag{1.15}$$

follows from a uniform bound

$$S(X,\pi) \ll_{w,m,d} C\left(\tfrac{1}{2},\pi\right)^{1/4-\delta+\epsilon}\sqrt{X} \tag{1.16}$$

in the range $X \le C\left(\tfrac{1}{2},\pi\right)^{1/2+\epsilon}$. It should be observed that Molteni's bound (1.8) yields an even stronger estimate whenever $X \le C\left(\tfrac{1}{2},\pi\right)^{1/2-2\delta}$. The above inequality (with no restriction on X) is in fact equivalent to the subconvex bound (1.15), as can be seen from the representation

$$S(X,\pi) = \int_{1/2-i\infty}^{1/2+i\infty} L(s,\pi)X^s W(s)\,ds,$$

where

$$W(s) = \int_0^\infty w(x)x^s\frac{dx}{x}$$

denotes the Mellin transform of $w(x)$.

By this line of thought we also see that the generalized Lindelöf hypothesis (1.13) translates into strong square-root cancellation among the coefficients $\lambda_\pi(n)$:

$$S(X,\pi) \ll_{\epsilon,w,m,d} C\left(\tfrac{1}{2},\pi\right)^{\epsilon}\sqrt{X}.$$

In particular, we expect that in a family $\mathcal{F}$ of cusp forms π we have

$$\frac{1}{|\mathcal{F}|}\sum_{\pi\in\mathcal{F}}|S(X,\pi)|^2 \ll_{\epsilon,w,m,d} C^\epsilon X,$$

as long as the analytic conductors satisfy $C\left(\tfrac{1}{2},\pi\right) \asymp C$. It is often possible to apply ideas from harmonic analysis to establish the preceding square mean bound for certain families $\mathcal{F}$. As an immediate consequence, we obtain a pointwise bound

$$S(X,\pi) \ll_{\epsilon,w,m,d} C^\epsilon\sqrt{|\mathcal{F}|X}, \quad \pi\in\mathcal{F}.$$

If we can guarantee that $|\mathcal{F}| \ll C^{1/2-2\delta}$, then a subconvex bound for $L\left(\frac{1}{2}, \pi\right)$ is established in the form (1.16). In most cases, however, harmonic analysis just falls short of establishing subconvexity. This is not surprising in the light of the extensive deep applications of subconvex bounds in number theory. The roots of subconvexity lie in arithmetic.

Amplification is an arithmetic device to substitute for shortening the family $\mathcal{F}$. It appeared in the seminal work of Duke, Friedlander and Iwaniec [Fr-Iw, Du-Fr-Iw1]. The basic idea is to introduce nonnegative arithmetic weights $|a_\pi|^2$ so that

$$\frac{1}{|\mathcal{F}|}\sum_{\pi\in\mathcal{F}} |a_\pi|^2 |S(X,\pi)|^2 \ll_{\epsilon,w,m,d} C^\epsilon X,$$

while $|a_\pi|$ is larger than C^δ for a specific $\pi \in \mathcal{F}$ and some $\delta > 0$. Then we only need to guarantee that $|\mathcal{F}| \ll C^{1/2+\epsilon}$, and subconvexity follows. The details in carrying out this program can become very complicated. Much of this thesis is devoted to study shifted convolution sums of the coefficients $\lambda_\pi(n)$, the sums that lie at the heart of the amplification method in the cases where it is known to work.

1.5 Shifted convolution sums and the circle method

A particularly interesting (conjectural) family of automorphic representations consists of Rankin–Selberg products $\pi \otimes \rho$, where π is a fixed cusp form on GL_m and ρ varies over cusp forms on a fixed GL_n ($n \le m$). The L-functions $L(s, \pi\otimes\rho)$ can be defined intrinsically and the expected analytic properties have been established by the work of many authors. The approach of amplification to establish subconvexity for these L-functions naturally leads to shifted convolution sums for π:

$$D_f(a,b;h) = \sum_{am\pm bn=h} \lambda_\pi(m)\bar{\lambda}_\pi(n) f(am,bn). \tag{1.17}$$

Here a, b, h are positive integers and f is some nice weight function on $(0,\infty)\times(0,\infty)$, e.g. smooth and compactly supported on a box $[X,2X]\times[Y,2Y]$. If we have a uniform estimate

$$\sum_{m\le x}|\lambda_\pi(m)|^2 \ll_\pi x,$$

then the size of the sum (1.17) can be seen to be at most $O_{\epsilon,f,\pi}\big(\sqrt{XY}\big)$. In order to achieve subconvexity, we need to improve on this bound in the X and Y aspects with certain uniformity regarding the other parameters.

Historically, the first examples of shifted convolution sums were generalized binary additive divisor sums, whose coefficients are given in terms of the divisor function:

$$D_f^\tau(a,b;h)=\sum_{am\pm bn=h}\tau(m)\tau(n)f(am,bn).$$

Note that the $\tau(n)$'s generate $\zeta^2(s)$, and they also appear as Fourier coefficients of the modular form $\frac{\partial}{\partial s}E(z,s)\big|_{s=1/2}$, where $E(z,s)$ is the Eisenstein series for $\mathrm{SL}_2(\mathbb{Z})$. These sums have been studied extensiviely since 1926, when Kloosterman published his famous refinement of the circle method [Kl]. A short summary of subsequent developments can be found in [Du-Fr-Iw2].

The crucial insight of Kloosterman was to make use of the very regular distribution of Farey fractions on the unit interval. By applying Voronoï-type summation formulae for the relevant exponential generating functions (which in turn reflect modular transformation properties), the binary additive sum in question decomposes to a main term and an error term in a natural fashion. The main term arises, because $E(z,s)$ is not cuspidal, and the error term is expressed in terms of Kloosterman sums

$$S(m,n;q)=\sideset{}{^*}\sum_{d\ (\mathrm{mod}\ q)} e_q\big(dm+\bar{d}n\big),$$

for which a nontrivial bound is needed. Kloosterman [Kl] did provide a nontrivial bound, and later Weil [We] and Esterman [Es] proved the optimal estimate.

This classical approach was revived recently by Duke, Friedlander and Iwaniec [Du-Fr-Iw2]. The Farey dissection being disguised as the δ-method, the Voronoï-type summation formula is still utilized at all frequencies so as to yield the following general result.

Theorem 1.3 (Duke–Friedlander–Iwaniec). *Let a, b coprime and assume that the partial derivatives of the weight function f satisfy the estimate*

$$x^k y^l f^{(k,l)}(x,y) \ll_{k,l} \left(1+\frac{x}{X}\right)^{-1}\left(1+\frac{y}{Y}\right)^{-1} P^{k+l} \tag{1.18}$$

with some $P, X, Y \geq 1$ for all $k, l \geq 0$. Then

$$D_f^{\tau}(a,b;h) = \int_0^{\infty} g(x, \mp x \pm h)\,dx + O\big(P^{5/4}(X+Y)^{1/4}(XY)^{1/4+\epsilon}\big),$$

where the implied constant depends only on ϵ,

$$g(x,y) = f(x,y)\sum_{q=1}^{\infty}\frac{(ab,q)}{abq^2}c_q(h)(\log x - \lambda_{aq})(\log y - \lambda_{bq}),$$

$c_q(h) = S(h,0;q)$ denotes Ramanujan's sum, and λ_{aq}, λ_{bq} are constants given by

$$\lambda_{aq} = 2\gamma + \log\frac{aq^2}{(a,q)^2}.$$

As was pointed out in [Du-Fr-Iw2], the error term is smaller than the main term whenever

$$P^{5/4}ab \ll (X+Y)^{-5/4}(XY)^{3/4-\epsilon}.$$

In Chapter 3 we shall extend the above ideas to exhibit nontrivial cancellation in the shifted convolution sums (1.17) for cuspidal automorphic representations π of

GL_2 over $\mathbb{Q}$. In fact, we shall estimate the more general sums

$$D_f(a,b;h) = \sum_{am \pm bn = h} \lambda_\phi(m)\lambda_\psi(n) f(am, bn), \tag{1.19}$$

where $\lambda_\phi(m)$ (resp. $\lambda_\psi(n)$) are the normalized Fourier coefficients of a classical holomorphic or weight zero Maass cusp form ϕ (resp. ψ) of arbitrary level and nebentypus. The conclusion is recorded in Theorem 3.1. In Chapter 4 we shall apply the result about shifted convolution sums to obtain a subconvex bound for the values $L(s, \phi \otimes \chi)$, where ϕ is a primitive form in the sense of Atkin–Lehner theory [At-Le, Li, At-Li], s is a fixed point on the critical line, and χ runs through primitive Dirichlet characters of conductor prime to the level of ϕ (Theorem 4.1). A specialization to the central point $s = 1/2$ yields, via Waldspurger's theorem and its generalization [Wal, Sh], nontrivial bounds for the Fourier-coefficients of holomorphic or Maass cusp forms of half-integral weight. These bounds in turn can be applied to resolve Linnik's problem [Du, Du-SP].

1.6 Shifted convolution sums and spectral theory

A different spectral approach was developed by Sarnak for all levels. The method can be traced back to the discovery of Rankin and Selberg, that for a holomorphic cusp form

$$\phi(z) = \sum_{n=1}^{\infty} \rho_\phi(n) e(nz)$$

of weight k, level N and arbitrary nebentypus, there is an integral representation

$$\sum_{n=1}^{\infty} \frac{|\rho_\phi(n)|^2}{n^{s+k-1}} = \frac{(4\pi)^{s+k-1}}{\Gamma(s+k-1)} \int_{\Gamma\backslash\mathcal{H}} y^k |\phi(z)|^2 E(z,s) \frac{dx\,dy}{y^2}, \tag{1.20}$$

where $\Gamma\backslash\mathcal{H}$ is a fundamental domain for the action of the Hecke congruence subgroup $\Gamma = \Gamma_0(N)$ on the upper half-plane $\mathcal{H} = \{x + iy : y > 0\}$, and

$$E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} y^s(\gamma z)$$

denotes Eisenstein's series. The above identity can be proved by a simple unfolding technique, and it shows that the summatory function of the coefficients $|\rho_\phi(n)|^2$ depends largely on the analytic properties of $E(z,s)$. The Eisenstein series $E(z,s)$ is a meromorphic function in the s-plane with the only pole at $s = 1$ in the half-plane $\Re s \geq 1/2$. The pole at $s = 1$ is simple with residue explicitly given by

$$\operatorname*{res}_{s=1} E(z,s) = \frac{1}{\operatorname{vol}\bigl(\Gamma\backslash\mathcal{H}\bigr)}.$$

The connection with the shifted convolution sums (1.17) becomes apparent if we specify $\Gamma = \Gamma_0(Nab)$, replace $E(z,s)$ by the Poincaré series

$$P_h(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} y^s(\gamma z) e(-hx(\gamma z)),$$

and the $\Gamma_0(N)$-invariant product $y^k|\phi(z)|^2$ by the Γ-invariant product $y^k\phi(az)\bar{\phi}(bz)$. We obtain, by the same unfolding technique,

$$\sum_{am-bn=h} \frac{\rho_\phi(m)\bar{\rho}_\phi(n)}{(am+bn)^{s+k-1}} = \frac{(2\pi)^{s+k-1}}{\Gamma(s+k-1)} \int_{\Gamma\backslash\mathcal{H}} y^k\phi(az)\bar{\phi}(bz)P_h(z,s)\frac{dx\,dy}{y^2}. \tag{1.21}$$

The integral equals, by definition, the Petersson inner product of the Γ-invariant functions $U(z) = y^k\phi(az)\bar{\phi}(bz)$ and $\bar{P}_h(z,s)$, and it can be decomposed according to the spectrum of $L^2(\Gamma\backslash\mathcal{H})$. The discrete part of the spectrum corresponds to an orthonormal basis of Maass cusp forms

$$\phi_0(x+iy) = \frac{1}{\mathrm{vol}^{1/2}(\Gamma\backslash\mathcal{H})},$$

$$\phi_j(x+iy) = \sqrt{y}\sum_{n\neq 0}\lambda_j(n)K_{i\tau_j}\big(2\pi|n|y\big)e(nx), \quad j=1,2,\ldots,$$

while the continuous spectrum is provided by the Eisenstein series

$$E_{\mathfrak{c}}(\cdot,\tfrac{1}{2}+i\tau) = \delta_{\mathfrak{c}}y^s + \eta_{\mathfrak{c}}(s)y^{1-s} + \sqrt{y}\sum_{n\neq 0}\lambda_{\mathfrak{c},\tau}(n)K_{i\tau}\big(2\pi|n|y\big)e(nx), \quad \tau\in\mathbb{R},$$

where $\mathfrak{c}$ is a singular cusp of $\Gamma\backslash\mathcal{H}$. The decomposition reads, at least formally, as

$$\begin{aligned} I(s) = \langle U, \bar{P}_h(.,s)\rangle &= \sum_{j=0}^{\infty}\langle U,\phi_j\rangle\langle\phi_j,\bar{P}_h(\cdot,s)\rangle \\ &\quad + \sum_{\mathfrak{c}}\frac{1}{4\pi}\int_{-\infty}^{\infty}\langle U, E_{\mathfrak{c}}(\cdot,\tfrac{1}{2}+i\tau)\rangle\langle E_{\mathfrak{c}}(\cdot,\tfrac{1}{2}+i\tau),\bar{P}_h(\cdot,s)\rangle\, d\tau. \end{aligned}$$

We observe that the inner products $\langle\phi_j,\bar{P}_h(.,s)\rangle$ and $\langle E_{\mathfrak{c}}(\cdot,\frac{1}{2}+i\tau),\bar{P}_h(\cdot,s)\rangle$ can be unfolded to

$$\langle\phi_j,\bar{P}_h(.,s)\rangle = \frac{\lambda_j(h)}{4(\pi h)^{s-\frac{1}{2}}}\Gamma\left(\frac{s-\frac{1}{2}+i\tau_j}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-i\tau_j}{2}\right),$$

$$\langle E_{\mathfrak{c}}(\cdot,\tfrac{1}{2}+i\tau),\bar{P}_h(\cdot,s)\rangle = \frac{\lambda_{\mathfrak{c},\tau}(h)}{4(\pi h)^{s-\frac{1}{2}}}\Gamma\left(\frac{s-\frac{1}{2}+i\tau}{2}\right)\Gamma\left(\frac{s-\frac{1}{2}-i\tau}{2}\right),$$

where $\frac{1}{4}+\tau_j^2$ (resp. $\frac{1}{4}+\tau^2$) denotes the Laplacian eigenvalue of ϕ_j (resp. of $E_{\mathfrak{c}}(\cdot,\frac{1}{2}+i\tau)$).

It follows that the size of $I(s)$ (including the location of its poles) are determined by the exceptional spectrum of $\Gamma\backslash\mathcal{H}$, the size of the Fourier coefficients $\lambda_j(h)$ and $\lambda_{\mathfrak{c},\tau}(h)$, and the size of the triple products $\langle U,\phi_j\rangle$ and $\langle U, E_{\mathfrak{c}}(\cdot,\frac{1}{2}+i\tau)\rangle$. We know that $\lambda_{\mathfrak{c},\tau}(h)$ is of size at most h^{ϵ}, and the Ramanujan conjecture predicts the same for

$\lambda_j(h)$. In addition, the Selberg conjecture predicts that the exceptional spectrum is empty, that is, $\tau_j \in \mathbb{R}$. As a substitute for these conjectures, we shall only assume the following statement which is known for many nontrivial values $\theta < 1/2$ (cf. (1.3)):

Hypothesis. For any cusp form π on GL_2 over $\mathbb{Q}$, the local Langlands parameters $\mu_{j,\pi}$ and $\alpha_{j,\pi}(p)$ $(j = 1, 2)$ satisfy

$$\begin{aligned} |\Re \mu_{j,\pi}| &\leq \theta, \quad \text{if } \pi_\infty \text{ is unramified;} \\ \left|\Re \log_p \alpha_{j,\pi}(p)\right| &\leq \theta, \quad \text{if } \pi_p \text{ is unramified } (p < \infty). \end{aligned}$$

The behaviour of the triple products $\langle U, \phi_j\rangle$ and $\langle U, E_{\mathfrak{c}}(\cdot, \frac{1}{2}+i\tau)\rangle$ was only understood recently by Sarnak [Sa1, Sa2]. He showed that

$$\langle U, \phi_j\rangle \ll_\phi \big(1 + |\tau_j|\big)^{k+1} e^{-\frac{\pi}{2}|\tau_j|},$$

and similarly for $\langle U, E_{\mathfrak{c}}(\cdot, \frac{1}{2}+i\tau)\rangle$. Note that the exponential decay in the eigenvalue parameter τ_j (resp. τ) exactly compensates the exponential decay of the coefficient $\Gamma(s+k-1)$ in (1.21).

If $\phi_1, \phi_2, \ldots$ are suitably chosen Maass–Hecke cuspidal eigenforms, then this argument leads to the powerful estimate

$$J(s) = \sum_{am-bn=h} \frac{\rho_\phi(m)\bar{\rho}_\phi(n)}{(am+bn)^{s+k-1}} \ll_{\phi,\epsilon} (ab)^{1-\frac{k}{2}} h^{\frac{1}{2}+\theta-\sigma+\epsilon}|s|^3, \quad \Re s \geq \frac{1}{2}+\theta+\epsilon. \tag{1.22}$$

Note that $\theta = 7/64$ is eligible by the recent work of Kim and Sarnak [Ki]. The strength of this result comes from the fact that it can be combined with the technique of Mellin transforms to yield a nontrivial bound for any shifted convolution sum

$$\sum_{am-bn=h} \lambda_\phi(m)\bar{\lambda}_\phi(n) W\left(\frac{am+bn}{h}\right),$$

where W is an arbitrary smooth function $(1,\infty)\to\mathbb{C}$ of compact support, and

$$\lambda_\phi(m) = m^{\frac{1-k}{2}}\rho_\phi(m)$$

denotes the normalized Fourier coefficients of ϕ. To see this connection, we introduce for convenience the variable

$$u = \frac{am+bn}{h},$$

as well the function

$$V(u) = (1-u^{-2})^{\frac{1-k}{2}}W(u),$$

then for any $\sigma > 1$ we get

$$\sum_{am-bn=h}\lambda_\phi(m)\bar{\lambda}_\phi(n)W(u) = (4ab)^{\frac{k-1}{2}}\sum_{am-bn=h}\frac{\rho_\phi(m)\bar{\rho}_\phi(n)}{(am+bn)^{k-1}}V(u)$$

$$= \frac{1}{2\pi i}\int_{(\sigma)}(4ab)^{\frac{k-1}{2}}h^s J(s)\hat{V}(s)\,ds.$$

We can rewrite (1.22) as

$$(4ab)^{\frac{k-1}{2}}h^s J(s) \ll_{\phi,\epsilon} (ab)^{\frac{1}{2}}h^{\frac{1}{2}+\theta+\epsilon}|s|^3, \quad \Re s \geq \frac{1}{2}+\theta+\epsilon,$$

therefore by shifting $\sigma > 1$ to any $\sigma > \frac{1}{2}+\theta$ we can conclude that

$$\sum_{am-bn=h}\lambda_\phi(m)\bar{\lambda}_\phi(n)W(u) \ll_{\phi,\epsilon} (ab)^{\frac{1}{2}}h^{\frac{1}{2}+\theta+\epsilon}\sup_{\sigma+i\mathbb{R}}\left|s^3\hat{V}(s)\right|.$$

In particular, if W is supported on $(X,2X)$, then we obtain

$$\sum_{am-bn=h}\lambda_\phi(m)\bar{\lambda}_\phi(n)W(u) \ll_{\phi,\sigma,\epsilon} (ab)^{\frac{1}{2}}h^{\frac{1}{2}+\theta+\epsilon}X^\sigma\max_{j=0,1,2,3}\left\|V^{(j)}\right\|_\infty, \quad \sigma \geq \frac{1}{2}+\theta+\epsilon.$$

For a Maass cusp form of weight κ and level N the analogous argument leads to complicated integral transforms. For such a form ϕ the Fourier expansion reads

$$\phi(x+iy) = \sum_{n \neq 0} \rho_\phi(n) \tilde{W}_{\frac{n}{|n|}\frac{\kappa}{2}, i\mu}\big(4\pi|n|y\big) e(nx),$$

where

$$\tilde{W}_{\alpha,\beta}(y) = \left\{ \frac{\Gamma\left(\frac{1}{2}+\beta-\alpha\right)}{\Gamma\left(\frac{1}{2}+\beta+\alpha\right)} \right\}^{1/2} W_{\alpha,\beta}(y),$$

$$W_{\alpha,\beta}(y) = \frac{e^{y/2}}{2\pi i} \int_{(\sigma)} \frac{\Gamma(w-\beta)\Gamma(w+\beta)}{\Gamma\left(\frac{1}{2}+w-\alpha\right)} \, y^{\frac{1}{2}-w} \, dw, \quad \sigma > |\Re\beta|,$$

is the (normalized) Whittaker function. The normalization is introduced in order to retain the coefficients $\rho_\phi(n)$ after the Maass operators have been applied. More precisely, if k is an integer of the same parity as κ, then

$$\phi_k(x+iy) = \sum_{n \neq 0} \rho_\phi(n) \tilde{W}_{\frac{n}{|n|}\frac{k}{2}, i\mu}\big(4\pi|n|y\big) e(nx)$$

is a Maass form of weight k and the same Petersson norm as ϕ:

$$\langle \phi_k, \phi_k \rangle = \langle \phi, \phi \rangle.$$

See Section 4 of [Du-Fr-Iw3] for details.

The unfolding technique yields an identity

$$(2\pi h)^{s-1} \int_{\Gamma\backslash\mathcal{H}} \phi_k(az)\bar{\phi}_k(bz) P_h(z,s) \frac{dx\,dy}{y^2} = \sum_{am-bn=h} \rho_\phi(m)\bar{\rho}_\phi(n) H_{s,k,i\mu}\left(\frac{am+bn}{h}\right),$$

where

$$H_{s,k,i\mu}(u) = \int_0^\infty \tilde{W}_{\frac{u+1}{|u+1|}\frac{k}{2},i\mu}\big(|u+1|y\big)\bar{\tilde{W}}_{\frac{u-1}{|u-1|}\frac{k}{2},i\mu}\big(|u-1|y\big)y^{s-2}\,dy, \quad u \neq \pm 1.$$

The main question that arises in the light of the above discussion is which weight functions $W : \mathbb{R} \to \mathbb{C}$ can be obtained by an averaging device from the $H_{s,k,i\mu}$ corresponding to values s on a vertical line $\sigma + i\mathbb{R}$ ($\sigma > 1$) and all even (resp. odd) integers k. In Chapter 5 we shall make the first step in answering these questions by obtaining a fairly precise description of the span of the functions $H_{s,0,i\mu}$ (Theorem 5.1).

Chapter 2

Approximate functional equation

2.1 Overview

We shall approximate the values of a principal L-function $L(s,\pi)$ on the critical line $\Re s = \frac{1}{2}$ as a sum of two truncated Dirichlet series which have about $\sqrt{C(s,\pi)}$ terms. We borrow notation from Section 1.2, and we also refer the reader to Section 1.3 for an introduction. The results of this chapter were published in [Ha1].

In order to keep the argument as clean as possible, we shall only display our formulae for the central value $L\left(\frac{1}{2},\pi\right)$. This results in no loss of generality, as $L\left(\frac{1}{2}+it,\pi\right)$ can be interpreted as the central value corresponding to the twisted representation $\pi \otimes |\det|^{it}$. For convenient reference we record the change of parameters in the formulae as we twist π by a 1-dimensional representation.

$$\pi \rightsquigarrow \pi \otimes |\det|^{it}; \quad L\left(\tfrac{1}{2},\pi\right) \rightsquigarrow L\left(\tfrac{1}{2}+it,\pi\right); \quad C\left(\tfrac{1}{2},\pi\right) \rightsquigarrow C\left(\tfrac{1}{2}+it,\pi\right);$$

$$\lambda_\pi(n) \rightsquigarrow n^{-it}\lambda_\pi(n); \qquad \mu_j \rightsquigarrow \mu_j - it; \qquad N \rightsquigarrow N; \qquad \kappa \rightsquigarrow N^{-it}\kappa.$$

For the rest of this chapter π will be a fixed cusp form on GL_m over a number field F, and C will abbreviate

$$C = C\left(\frac{1}{2}, \pi\right) = \frac{N}{(2\pi)^{md}} \prod_{j=1}^{md} \left|\frac{1}{2} - \mu_j\right|. \tag{2.1}$$

Theorem 2.1. *There is a smooth function $f : (0, \infty) \to \mathbb{C}$ and a complex number λ of modulus 1 depending only on the Archimedean parameters μ_j $(j = 1, \ldots, md)$ such that*

$$L\left(\frac{1}{2}, \pi\right) = \sum_{n=1}^{\infty} \frac{\lambda_\pi(n)}{\sqrt{n}} f\left(\frac{n}{\sqrt{C}}\right) + \kappa\lambda \sum_{n=1}^{\infty} \frac{\bar{\lambda}_\pi(n)}{\sqrt{n}} \bar{f}\left(\frac{n}{\sqrt{C}}\right). \tag{2.2}$$

The function f and its partial derivatives $f^{(k)}$ $(k = 1, 2, \ldots.)$ satisfy the following uniform growth estimates at 0 and infinity:

$$f(x) = \begin{cases} 1 + O_\sigma(x^\sigma), & 0 < \sigma < \frac{1}{m^2+1}; \\ O_\sigma(x^{-\sigma}), & \sigma > 0; \end{cases} \tag{2.3}$$

$$f^{(k)}(x) = O_{\sigma,k}(x^{-\sigma}), \quad \sigma > k - \tfrac{1}{m^2+1}. \tag{2.4}$$

The implied constants depend only on σ, k, m and d.

Remark 2.1. The range $0 < \sigma < \frac{1}{m^2+1}$ in (2.3) can be widened to $0 < \sigma < \frac{1}{2}$ for all representations π which are tempered at ∞, that is, conjecturally for all π. Similarly, upon the Ramanujan–Selberg conjecture the range of σ in (2.4) can be extended to $\sigma > k - \frac{1}{2}$.

Combining the theorem with Molteni's bound (1.8) we obtain that the size of the central value $L\left(\frac{1}{2}, \pi\right)$ can be very well approximated with the first $C^{1/2+\epsilon}$ Dirichlet coefficients.

Corollary 2.1. *For any positive numbers ϵ and A,*

$$L\left(\frac{1}{2}, \pi\right) = \sum_{n \le C^{1/2+\epsilon}} \frac{\lambda_\pi(n)}{\sqrt{n}} f\left(\frac{n}{\sqrt{C}}\right) + \kappa\lambda \sum_{n \le C^{1/2+\epsilon}} \frac{\bar{\lambda}_\pi(n)}{\sqrt{n}} \bar{f}\left(\frac{n}{\sqrt{C}}\right) + O_{\epsilon,A}(C^{-A}).$$

The implied constant depends only on ϵ, A, m and d.

In particular, by applying (1.8) again, we can reconstruct the convexity bound (1.12) for the central value (in fact for all values on the critical line).

In a family of representations π, it is often desirable to see that the weight functions f do not vary too much. In fact, assuming that the Archimedean parameters are not too small, one can replace f by an explicit function g (independent of π) and derive an approximate functional equation with a nontrivial error term, that is, an error substantially smaller than the convexity bound furnished by the above corollary. To state the result, we introduce

$$\eta = \min_{j=1,\ldots,md} \left| \frac{1}{2} - \mu_j \right|. \tag{2.5}$$

Theorem 2.2. *Let $g : (0, \infty) \to \mathbb{R}$ be a smooth function with the functional equation $g(x) + g(1/x) = 1$ and derivatives decaying faster than any negative power of x as $x \to \infty$. Then, for any $\epsilon > 0$,*

$$L\left(\frac{1}{2}, \pi\right) = \sum_{n=1}^{\infty} \frac{\lambda_\pi(n)}{\sqrt{n}} g\left(\frac{n}{\sqrt{C}}\right) + \kappa\lambda \sum_{n=1}^{\infty} \frac{\bar{\lambda}_\pi(n)}{\sqrt{n}} g\left(\frac{n}{\sqrt{C}}\right) + O_{\epsilon,g}(\eta^{-1} C^{1/4+\epsilon}),$$

where λ (of modulus 1) is given by (2.8), and the implied constant depends only on ϵ, g, m and d.

Remark 2.2. The formula is really of value when the family under consideration satisfies $\eta \gg C^\delta$ with some fixed $\delta > 0$.

2.2 The implicit form

In this section we prove Theorem 2.1. We introduce the auxiliary function

$$F(s,\pi_\infty)=\frac{1}{2}C^{-s/2}N^s\frac{L\left(\frac{1}{2}+s,\pi_\infty\right)L\left(\frac{1}{2},\tilde{\pi}_\infty\right)}{L\left(\frac{1}{2}-s,\tilde{\pi}_\infty\right)L\left(\frac{1}{2},\pi_\infty\right)}+\frac{1}{2}C^{s/2}, \tag{2.6}$$

which is holomorphic in the half plane $\Re s>-\frac{1}{m^2+1}$ by (1.1) and (1.3). With this notation we can rewrite the functional equation (1.5) as

$$F(s,\pi_\infty)L\left(\tfrac{1}{2}+s,\pi\right)=\kappa\lambda F(-s,\tilde{\pi}_\infty)L\left(\tfrac{1}{2}-s,\tilde{\pi}\right), \tag{2.7}$$

where

$$\lambda=\frac{L\left(\frac{1}{2},\tilde{\pi}_\infty\right)}{L\left(\frac{1}{2},\pi_\infty\right)}. \tag{2.8}$$

It follows from (1.6) that $|\lambda|=1$, $F(0,\pi_\infty)=1$, and

$$\bar{F}(s,\pi_\infty)=F(\bar{s},\tilde{\pi}_\infty). \tag{2.9}$$

We also fix an entire function $H(s)$ which satisfies the growth estimate

$$H(s)\ll_{\sigma,A}\big(1+|s|\big)^{-A},\quad \Re s=\sigma; \tag{2.10}$$

on vertical lines. In addition, we shall assume that $H(0)=1$ and that $H(s)$ is symmetric with respect to both axes:

$$H(s)=H(-s)=\bar{H}(\bar{s}). \tag{2.11}$$

Such a function can be obtained as the Mellin transform of a smooth function $h:\ (0,\infty)\to\mathbb{R}$ which has total mass 1 with respect to the measure dx/x, functional equation $h(1/x)=h(x)$, and derivatives decaying faster than any negative power of x as $x \to \infty$:

$$H(s) = \int_0^\infty h(x) x^s \frac{dx}{x}.$$

Using these two auxiliary functions and taking an arbitrary $0 < \sigma < \frac{1}{m^2+1}$, we can express the central value $L\left(\frac{1}{2}, \pi\right)$ via the residue theorem as

$$\begin{aligned} L\left(\frac{1}{2}, \pi\right) &= \frac{1}{2\pi i} \int_{(\sigma)} L\left(\frac{1}{2} + s, \pi\right) F(s, \pi_\infty) H(s) \frac{ds}{s} \\ &\quad - \frac{1}{2\pi i} \int_{(-\sigma)} L\left(\frac{1}{2} + s, \pi\right) F(s, \pi_\infty) H(s) \frac{ds}{s}. \end{aligned}$$

This step is justified by the convexity bound (1.12), inequality (2.10) and Lemma 2.1 below. Applying a change of variable $s \mapsto -s$ in the second integral we get, by the functional equations (2.7) and (2.11),

$$\begin{aligned} L\left(\frac{1}{2}, \pi\right) &= \frac{1}{2\pi i} \int_{(\sigma)} L\left(\frac{1}{2} + s, \pi\right) F(s, \pi_\infty) H(s) \frac{ds}{s} \\ &\quad + \frac{\kappa\lambda}{2\pi i} \int_{(\sigma)} L\left(\frac{1}{2} + s, \tilde{\pi}\right) F(s, \tilde{\pi}_\infty) H(s) \frac{ds}{s}. \end{aligned}$$

The second integral is minus the complex conjugate of the first one, as can be seen by another change of variable $s \mapsto \bar{s}$ combined with the functional equations (1.6), (2.9) and (2.11). Therefore we obtain the representation (2.2) of Theorem 2.1 by defining

$$f\left(\frac{x}{\sqrt{C}}\right) = \frac{1}{2\pi i} \int_{(\sigma)} x^{-s} F(s, \pi_\infty) H(s) \frac{ds}{s}. \tag{2.12}$$

For any nonnegative integer k we also have

$$f^{(k)}(x) = \frac{(-1)^k}{2\pi i} \int_{(\sigma)} x^{-s-k} C^{-s/2} F(s, \pi_\infty) H(s) s(s+1) \dots (s+k-1) \frac{ds}{s}. \tag{2.13}$$

When $k = 0$, the integrand in this expression is holomorphic for $\Re s > -\frac{1}{m^2+1}$ with the exception of a simple pole at $s = 0$ with residue 1. So in this case we are free to move the line of integration to any nonzero $\sigma > -\frac{1}{m^2+1}$, but negative σ's will pick up an additional value 1 from the pole at $s = 0$. When $k > 0$, the integrand is holomorphic in the entire half plane $\Re s > -\frac{1}{m^2+1}$, so the line of integration can be shifted to any $\sigma > -\frac{1}{m^2+1}$ without changing the value of the integral. Henceforth, by (2.10) and (2.13), the truth of inequalities (2.3) and (2.4) is reduced to the following:

Lemma 2.1. *For any $\sigma > -\frac{1}{m^2+1}$, there is a uniform bound*

$$2C^{-s/2}F(s,\pi_\infty) - 1 \ll_\sigma \big(1+|s|\big)^{md\sigma}, \quad \Re s = \sigma. \tag{2.14}$$

The implied constant depends only on σ, m and d.

We start with the following simple estimate.

Lemma 2.2. *For any $\alpha > -\sigma$, there is a uniform bound*

$$\frac{\Gamma(z+\sigma)}{\Gamma(z)} \ll_{\alpha,\sigma} |z+\sigma|^\sigma, \quad \Re z \geq \alpha.$$

Proof of Lemma 2.2. The function $\Gamma(z+\sigma)/\Gamma(z)$ is holomorphic in a neighborhood of $\Re z \geq \alpha$. For $|z| > 2|\sigma|$ we get, using Stirling's formula,

$$\frac{\Gamma(z+\sigma)}{\Gamma(z)} \ll_\sigma \left|\frac{(z+\sigma)^{z+\sigma-1/2}}{z^{z-1/2}}\right| \ll_\sigma |z+\sigma|^\sigma.$$

The rest of the values of z (those with $\Re z \geq \alpha$ and $|z| \leq 2|\sigma|$) form a compact set, so for these we simply have

$$\frac{\Gamma(z+\sigma)}{\Gamma(z)} \ll_{\alpha,\sigma} 1 \ll_{\alpha,\sigma} |z+\sigma|^\sigma. \quad \square$$

Proof of Lemma 2.1. Let $s = \sigma + it$. For any $j = 1, \dots, md$, we apply Lemma 2.2 with

$$\alpha = \frac{1}{2(m^2+1)} - \frac{\sigma}{2}, \quad z = \frac{1}{4} - \frac{\mu_j}{2} - \frac{\sigma}{2} + \frac{it}{2}$$

to see that

$$\frac{\Gamma\left(\frac{1}{4}-\frac{\mu_j}{2}+\frac{\sigma}{2}+\frac{it}{2}\right)}{\Gamma\left(\frac{1}{4}-\frac{\mu_j}{2}-\frac{\sigma}{2}+\frac{it}{2}\right)} \ll_{\sigma,m} \left|\frac{1}{4}-\frac{\mu_j}{2}+\frac{\sigma}{2}+\frac{it}{2}\right|^{\sigma}.$$

This is the same as

$$\frac{\Gamma\left(\frac{1}{4}-\frac{\mu_j}{2}+\frac{s}{2}\right)}{\Gamma\left(\frac{1}{4}-\frac{\bar{\mu}_j}{2}-\frac{s}{2}\right)} \ll_{\sigma,m} \left|\frac{1}{2}-\mu_j+s\right|^{\sigma}.$$

It follows from (1.3) that

$$\left|\frac{1}{2}-\mu_j+s\right| \leq \left|\frac{1}{2}-\mu_j\right| + |s| \ll_m \left|\frac{1}{2}-\mu_j\right| \big(1+|s|\big),$$

therefore we have

$$\frac{\Gamma\left(\frac{1}{4}-\frac{\mu_j}{2}+\frac{s}{2}\right)}{\Gamma\left(\frac{1}{4}-\frac{\bar{\mu}_j}{2}-\frac{s}{2}\right)} \ll_{\sigma,m} \left|\frac{1}{2}-\mu_j\right|^{\sigma} \big(1+|s|\big)^{\sigma}.$$

Taking the product of these inequalities over all $j = 1, \ldots, md$, and using (1.1), (1.6) and (2.1), we get

$$\frac{L\left(\frac{1}{2}+s,\pi_\infty\right)}{L\left(\frac{1}{2}-s,\tilde{\pi}_\infty\right)} \ll_{\sigma,m,d} \left(\frac{C}{N}\right)^{\sigma} \big(1+|s|\big)^{md\sigma}, \quad \Re s = \sigma.$$

By (2.6), this is equivalent to (2.14), completing the proof of Lemma 2.1 and Theorem 2.1. □

2.3 The explicit form

Our aim is to deduce Theorem 2.2. We can assume that $H(s)$ is the Mellin transform of $h(x) = -xg'(x)$. Indeed, $h : (0,\infty) \to \mathbb{R}$ is a smooth function with the functional equation $h(1/x) = h(x)$ and derivatives decaying faster than any negative power of x

as $x \to \infty$, therefore $H(s)$ is entire and satisfies (2.10) and (2.11). Also,

$$H(0) = -\int_0^\infty g'(x) = g(0+) = 1.$$

Equivalently, $H(s)/s$ is the Mellin transform of $g(x)$, because by partial integration it follows that

$$-\int_0^\infty g'(x)x^s dx = s\int_0^\infty g(x)x^s \frac{dx}{x}.$$

In any case, $g(x)$ can be expressed as an inverse Mellin transform

$$g(x) = \frac{1}{2\pi i}\int_{(\sigma)} x^{-s}H(s)\frac{ds}{s}.$$

The idea is to compare $g(x)$ with the function $f(x)$ given by (2.12). We have, for any $\sigma > 0$,

$$f(x) - g(x) = \frac{1}{2\pi i}\int_{(\sigma)} x^{-s}\big\{C^{-s/2}F(s,\pi_\infty) - 1\big\}H(s)\frac{ds}{s}.$$

In fact, the integrand is holomorphic in the entire half plane $\Re s > -\frac{1}{m^2+1}$, so the line of integration can be shifted to any $\sigma > -\frac{1}{m^2+1}$ without changing the value of the integral. In particular, the choice $\sigma = 0$ is permissible, that is,

$$f(x) - g(x) = \frac{1}{2\pi i}\int_{-\infty}^\infty x^{-it}\big\{C^{-it/2}F(it,\pi_\infty) - 1\big\}H(it)\frac{dt}{t}. \tag{2.15}$$

Note that x^{-it} and $2C^{-it/2}F(it,\pi_\infty)-1$ are of modulus 1. For any $\epsilon > 0$, the values of t with $|t| \geq \min(\eta/2, C^\epsilon)$ contribute $O_{\epsilon,g,m,d}(\eta^{-1})$ to the integral. This follows from (2.10) and $\eta \ll C^{1/md}$. We estimate the remaining contribution via the following lemma.

Lemma 2.3. *For any $\epsilon > 0$, there is a uniform bound*

$$2C^{-it/2}F(it,\pi_\infty) - 2 \ll_\epsilon |t|\eta^{-1}C^\epsilon, \quad |t| < \min(\eta/2, C^\epsilon).$$

The implied constant depends only on ϵ, m and d.

Proof. As $2C^{-it/2}F(it,\pi_\infty) - 1$ lies on the unit circle, it suffices to show that

$$\log\bigl\{2C^{-it/2}F(it,\pi_\infty) - 1\bigr\} \ll_{\epsilon,m,d} |t|\eta^{-1}C^\epsilon, \quad |t| < \min(\eta/2, C^\epsilon).$$

Here the left hand side is understood as a continuous function defined via the principal branch of the logarithm near $t = 0$. Using (2.1), (2.6), (1.1) and (1.6) we can see that the derivative (with respect to t) of the left hand side is given by

$$i\Re\sum_{j=1}^{md}\left\{\frac{\Gamma'}{\Gamma}\left(\frac{1}{4} - \frac{\mu_j}{2} + \frac{it}{2}\right) - \log\left(\frac{1}{4} - \frac{\mu_j}{2}\right)\right\},$$

so we can further reduce the lemma to

$$\frac{\Gamma'}{\Gamma}\left(\frac{1}{4} - \frac{\mu_j}{2} + \frac{it}{2}\right) - \log\left(\frac{1}{4} - \frac{\mu_j}{2}\right) \ll_{\epsilon,m,d} \eta^{-1}C^\epsilon, \quad |t| < \min(\eta/2, C^\epsilon). \tag{2.16}$$

Here $\frac{1}{4} - \frac{\mu_j}{2} + \frac{it}{2}$ has real part at least $\frac{1}{2(m^2+1)}$ by (1.3) and absolute value at least $\eta/4$ by (2.5). Therefore, a standard bound yields

$$\frac{\Gamma'}{\Gamma}\left(\frac{1}{4} - \frac{\mu_j}{2} + \frac{it}{2}\right) = \log\left(\frac{1}{4} - \frac{\mu_j}{2} + \frac{it}{2}\right) + O_m(\eta^{-1}).$$

For $|t| < \min(\eta/2, C^\epsilon)$ we can also see that

$$\log\left(\frac{1}{4} - \frac{\mu_j}{2} + \frac{it}{2}\right) = \log\left(\frac{1}{4} - \frac{\mu_j}{2}\right) + O(\eta^{-1}C^\epsilon).$$

It follows from (1.3) that $C \gg_{m,d} 1$, therefore the last two estimates add up to (2.16) as required. $\square$

Returning to the integral (2.15), it follows from Lemma 2.3 that the values of t with $|t| < \min(\eta/2, C^{\epsilon})$ contribute at most $O_{\epsilon,g,m,d}(\eta^{-1}C^{2\epsilon})$. Altogether we have, by $C \gg_{m,d} 1$,

$$f(x) - g(x) = O_{\epsilon,g,m,d}(\eta^{-1}C^{2\epsilon}).$$

We conclude Theorem 2.2 by combining this estimate with Corollary 2.1 and Molteni's bound (1.8).

Chapter 3

Shifted convolution sums and the circle method

3.1 Overview

We shall establish, in the spirit of Duke, Friedlander and Iwaniec, a nontrivial bound for the shifted convolution sums (1.17) arising from classical holomorphic or Maass cusp forms for the Hecke congruence subgroups. We refer the reader to Section 1.5 for an introduction. The notions in the following theorem will be defined in the next section. The result, in less explicit form, will also appear in [Ha2].

Theorem 3.1. *Let $\lambda_\phi(m)$ (resp. $\lambda_\psi(n)$) be the normalized Fourier coefficients of a holomorphic or Maass cusp form ϕ (resp. ψ) of level N and arbitrary nebentypus character modulo N. Let $|\tilde{\mu}|$ (resp. $|\tilde{\nu}|$) denote the Archimedean size of ϕ (resp. ψ), and suppose that f satisfies (1.18). Then for coprime a and b we have*

$$D_f(a, b; h) \ll P^{11/10} N^{9/5} |\tilde{\mu}\tilde{\nu}|^{9/5+\epsilon} (ab)^{-1/10} (X+Y)^{1/10} (XY)^{2/5+\epsilon},$$

where the implied constant depends only on ϵ.

Remark 3.1. We shall see in Section 3.6 that Cauchy's inequality implies

$$D_f(a,b;h) \ll N|\tilde{\mu}\tilde{\nu}|^{1/2}(ab)^{-1/2}(XY)^{1/2}. \tag{3.1}$$

The conclusion of the theorem supercedes this trivial bound whenever

$$P^{11}N^8|\tilde{\mu}\tilde{\nu}|^{13+\epsilon}(ab)^4 \ll \frac{(XY)^{1-\epsilon}}{X+Y}. \tag{3.2}$$

The proof of Theorem 3.1 is presented in Sections 3.2 through 3.7 and closely follows [Du-Fr-Iw2]. The heart of the argument is a Voronoï-type summation formula (see Section 3.3) for transforming certain exponential sums defined by the coefficients $\lambda_\phi(m)$ and $\lambda_\psi(n)$. As the level of the forms imposes some restriction on the frequencies in the formula, we replace (in Section 3.4) the classical Farey dissection (or the δ-method) with Jutila's variant of the circle method [Ju1]. The variant uses overlapping intervals, and hence provides great flexibility in the choice of frequencies. After transforming our exponential generating functions in Section 3.5, we encounter twisted Kloosterman sums

$$S_\chi(m,n;q) = \sideset{}{^*}\sum_{d \pmod q} \chi(d) e_q\big(dm + \bar{d}n\big),$$

where χ is a Dirichlet character mod q. We refer to the usual Weil–Estermann bound

$$\big|S_\chi(m,n;q)\big| \leq (m,n,q)^{1/2} q^{1/2} \tau(q), \tag{3.3}$$

for which the original proofs [We, Es] can be adapted. In Section 3.6 we apply a smooth dyadic decomposition, and conclude the theorem by optimizing the free parameters. In order to achieve polynomial uniformity in the Archimedean parameters of the cusp forms, we need to exhibit careful estimates for the Bessel functions involved in the summation formula. These estimates appear in Section 3.7 with detailed proofs.

3.2 Normalized Fourier coefficients

We define the normalized Fourier coefficients of cusp forms as follows. Let ϕ be a cusp form of level N and nebentypus χ, that is, a holomorphic cusp form of some integral weight k, or a real-analytic Maass cusp form of some nonnegative Laplacian eigenvalue $1/4 + \mu^2$. In the holomorphic case we write $k - 1 = 2i\mu$, in the real-analytic case we define $k = 0$, and in both cases we put $\tilde{\mu} = 1/2 + i\mu$ and call $|\tilde{\mu}|$ the *Archimedean size* of ϕ. This is in accordance with Section 1.2.

By definition, χ is a Dirichlet character mod N, and ϕ is a complex valued function on the upper half plane $\mathcal{H} = \{z : \Im z > 0\}$, which decays exponentially to zero at each cusp and satisfies a transformation rule with respect to the Hecke congruence subgroup $\Gamma_0(N)$:

$$\phi\left(\frac{az+b}{cz+d}\right) = \chi(d)(cz+d)^k\phi(z), \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N).$$

In particular, ϕ admits the Fourier expansion

$$\phi(x+iy) = \sum_{n\neq 0} \rho_\phi(n)W(ny)e(nx), \tag{3.4}$$

where

$$W(y) = \begin{cases} e^{-2\pi y} & \text{if } \phi \text{ is holomorphic,} \\ |y|^{1/2}K_{i\mu}\big(2\pi|y|\big) & \text{if } \phi \text{ is real-analytic.} \end{cases} \tag{3.5}$$

Here $e(x) = e^{2\pi i x}$, and $K_{i\mu}$ is the MacDonald-Bessel function. If ϕ is holomorphic,

$\rho_\phi(n)$ vanishes for $n < 0$. Writing

$$\langle \phi, \phi \rangle = \int_{\Gamma_0(N)\backslash\mathcal{H}} y^{k-2} |\phi(x+iy)|^2 \, dx \, dy,$$

we define the *normalized Fourier coefficients* of ϕ as

$$\lambda_\phi(n) = \begin{cases} \left(\frac{N(k-1)!}{\langle \phi, \phi \rangle (4\pi n)^{k-1}} \right)^{1/2} \rho_\phi(n) & \text{if } \phi \text{ is holomorphic,} \\ \left(\frac{N(4\pi|n|)}{\langle \phi, \phi \rangle \cosh \pi\mu} \right)^{1/2} \rho_\phi(n) & \text{if } \phi \text{ is real-analytic.} \end{cases} \tag{3.6}$$

This normalization corresponds to Rankin–Selberg theory which implies the following mean square estimate for the normalized Fourier coefficients (see Section 8.2 of [Iw1]):

$$c_N \sum_{1 \le |n| \le x} |\lambda_\phi(n)|^2 \sim x \qquad \text{as } x \to \infty,$$

$$1 \ll c_N \ll \log\log(3N).$$

More precisely,

$$c_N \asymp \frac{\operatorname{vol}\bigl(\Gamma_0(N)\backslash\mathcal{H}\bigr)}{N} = \frac{\pi}{3} \prod_{p|N} \left(1 + \frac{1}{p}\right).$$

We also have a good uniform upper bound for all $x > 0$ (see Theorem 3.2 and (8.7) and (9.34) in [Iw1]):

$$\sum_{1 \le |n| \le x} |\lambda_\phi(n)|^2 \ll x + N|\tilde{\mu}|, \tag{3.7}$$

where the implied constant is absolute.

Lemma 3.1. *For any $\epsilon > 0$ there is a uniform bound*

$$y^{k/2} \phi(x+iy) \ll \langle \phi, \phi \rangle^{1/2} |\tilde{\mu}|^{3/2+\epsilon} y^{-\epsilon}, \quad x \in \mathbb{R}, \; y > 1/2.$$

The implied constant depends only on ϵ.

Proof. We distungish between two cases.

Case 1. ϕ is holomorphic. By (3.4), (3.5) and (3.6), the statement is equivalent to

$$\left|\sum_{n=1}^{\infty}\lambda_\phi(n)(4\pi n)^{\frac{k-1}{2}}e^{-2\pi ny}e(nx)\right|^2 \ll_\epsilon (k-1)!k^{3+\epsilon}Ny^{-k-\epsilon}.$$

By the Cauchy-Schwartz inequality the left hand side can be estimated from above by

$$\left(\sum_{n=1}^{\infty}|\lambda_\phi(n)|^2(4\pi n)^{-1-\epsilon}\right)\left(\sum_{n=1}^{\infty}(4\pi n)^{k+\epsilon}e^{-4\pi ny}\right).$$

The first factor is $\ll_\epsilon Nk$ by the mean square bound (3.7), therefore it remains to show that

$$\sum_{n=1}^{\infty}(4\pi ny)^{k+\epsilon}e^{-4\pi ny} \ll_\epsilon (k-1)!k^{2+\epsilon}.$$

We accomplish this in stronger form by comparing the sum with the similar integral (note that $y \gg 1$):

$$\sum_{n=1}^{\infty}(4\pi ny)^{k+\epsilon}e^{-4\pi ny} \ll \sup_{y>0}\left\{(4\pi ny)^{k+\epsilon}e^{-4\pi ny}\right\} + \int_0^{\infty}(4\pi ny)^{k+\epsilon}e^{-4\pi ny}\,dy$$

$$= \left(\frac{k+\epsilon}{e}\right)^{k+\epsilon} + \Gamma(k+1+\epsilon) \ll_\epsilon (k-1)!k^{1+\epsilon}.$$

Case 2. ϕ is real-analytic. By (3.4), (3.5) and (3.6), the statement is equivalent to

$$\left|\sum_{n\neq 0}\lambda_\phi(n)K_{i\mu}\big(2\pi|n|y\big)e(nx)\right|^2 \ll_\epsilon e^{-\pi|\mu|}|\tilde{\mu}|^{3+\epsilon}Ny^{-1-\epsilon}.$$

By the Cauchy-Schwartz inequality the left hand side can be estimated from above by

$$\left(\sum_{n\neq 0}|\lambda_\phi(n)|^2|2\pi n|^{-1-\epsilon}\right)\left(\sum_{n\neq 0}|2\pi n|^{1+\epsilon}\left|K_{i\mu}\big(2\pi|n|y\big)\right|^2\right).$$

The first factor is $\ll_\epsilon N|\tilde{\mu}|$ by the mean square bound (3.7), therefore it remains to show that

$$\sum_{n\neq 0} |2\pi n y|^{1+\epsilon} e^{\pi|\mu|} \big| K_{i\mu}\big(2\pi|n|y\big)\big|^2 \ll_\epsilon |\tilde{\mu}|^{2+\epsilon}.$$

We accomplish this by employing Proposition 3.5 of Section 3.7, noting also that $y \gg 1$ and $|\Re(i\mu)| \leq 1/2$:

$$\begin{aligned}\sum_{n\neq 0} |2\pi n y|^{1+\epsilon} e^{\pi|\mu|} \big| K_{i\mu}\big(2\pi|n|y\big)\big|^2 = & \sum_{4|n|y<|\tilde{\mu}|} \dots + \sum_{|\tilde{\mu}|\leq 4|n|y<2|\tilde{\mu}|} \dots + \sum_{2|\tilde{\mu}|\leq 4|n|y} \dots \\ \ll_\epsilon & \ |\tilde{\mu}|^{2+\epsilon} \quad + \quad |\tilde{\mu}|^{1+\epsilon} \quad + \quad |\tilde{\mu}|^{\epsilon}. \quad \square\end{aligned}$$

Lemma 3.2. *For any $\epsilon > 0$ there is a uniform bound*

$$\big\| y^{k/2}\phi(x+iy)\big\|_\infty \ll \langle \phi, \phi\rangle^{1/2} |\tilde{\mu}|^{3/2+\epsilon}.$$

The implied constant depends only on ϵ.

Proof. It is known that any $z = x + iy$ can be represented as $z = \frac{aw+b}{cw+d}$, where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and w has imaginary part $\Im w > 1/2$. The proof of the previous lemma can be adapted almost verbatim to the cusp form $w \mapsto (cw+d)^{-k}\phi\left(\frac{aw+b}{cw+d}\right)$, so that we have, in particular,

$$y^{k/2}\phi(x+iy) = \frac{|\Im w|^{k/2}}{|cw+d|^k}\phi\left(\frac{aw+b}{cw+d}\right) \ll_\epsilon \langle \phi, \phi\rangle^{1/2} |\tilde{\mu}|^{3/2+\epsilon}. \quad \square$$

The proof of Theorem 3.1 is based on exponential sums of the form

$$T_{\phi,\alpha}(M) = \sum_{1\leq m\leq M} \lambda_\phi(m) e(\alpha m). \tag{3.8}$$

We shall use the following uniform variant of Wilton's classical estimate.

Proposition 3.1. *For any $\epsilon > 0$ there is a uniform bound*

$$T_{\phi,\alpha}(M) \ll N^{1/2}|\tilde{\mu}|^{2+\epsilon}M^{1/2+\epsilon}, \quad \alpha \in \mathbb{R},\ M > 0. \tag{3.9}$$

The implied constant depends only on ϵ.

Proof. We can clearly assume that M is a positive integer. As before, we distungish between two cases.

Case 1. ϕ is holomorphic. For any positive integer m we have, by (3.4), (3.5) and (3.6),

$$\lambda_\phi(m)e^{-2\pi my}e(m\alpha) = \left(\frac{N(k-1)!}{\langle\phi,\phi\rangle(4\pi m)^{k-1}}\right)^{1/2}\int_0^1 \phi(\alpha+\beta+iy)e(-m\beta)\,d\beta.$$

We multiply both sides by $(2\pi y)^{k/2+\epsilon}$, and integrate with respect to dy/y. We obtain

$$\lambda_\phi(m)m^{-1/2-\epsilon}e(m\alpha) = \int_0^1 \Phi_\alpha(\beta)e(-m\beta)\,d\beta,$$

where

$$\Phi_\alpha(\beta) = \frac{(2\pi)^{k/2+\epsilon}}{\Gamma(k/2+\epsilon)}\left(\frac{N(k-1)!}{\langle\phi,\phi\rangle(4\pi)^{k-1}}\right)^{1/2}\int_0^\infty y^{k/2+\epsilon}\phi(\alpha+\beta+iy)\frac{dy}{y}.$$

Note that the integral converges by Lemmata 3.1 and 3.2 and satisfies the uniform bound

$$\int_0^\infty y^{k/2+\epsilon}\phi(\alpha+\beta+iy)\frac{dy}{y} \ll_\epsilon \langle\phi,\phi\rangle^{1/2}k^{3/2+2\epsilon}.$$

It follows that

$$\Phi_\alpha(\beta) \ll_\epsilon N^{1/2}k^{7/4+\epsilon}.$$

By introducing the kernel

$$F_M(\beta) = \sum_{|m| \leq M} e(m\beta) = \frac{\sin \pi(2M+1)\beta}{\sin \pi\beta},$$

we can write

$$\sum_{m=1}^{M} \lambda_\phi(m) m^{-1/2-\epsilon} e(m\alpha) = \int_0^1 \Phi_\alpha(\beta) F_M(\beta)\, d\beta.$$

It is known that the $\mathcal{L}^1$-norm of F_M is $\ll \log(2M)$, therefore it follows that

$$\sum_{m=1}^{M} \lambda_\phi(m) m^{-1/2-\epsilon} e(m\alpha) \ll_\epsilon N^{1/2} k^{7/4+\epsilon} M^\epsilon.$$

From this bound (3.9) follows by partial summation.

Case 2. ϕ is real-analytic. For any nonzero integer m we have, by (3.4), (3.5) and (3.6),

$$\lambda_\phi(m) y^{1/2} K_{i\mu}(2\pi|m|y) e(m\alpha) = \left(\frac{4\pi N}{\langle \phi, \phi \rangle \cosh \pi\mu}\right)^{1/2} \int_0^1 \phi(\alpha + \beta + iy) e(-m\beta)\, d\beta.$$

We multiply both sides by $(2\pi)^{1/2+\epsilon} y^\epsilon$, and integrate with respect to dy/y. We obtain

$$\lambda_\phi(m) |m|^{-1/2-\epsilon} e(m\alpha) = \int_0^1 \Phi_\alpha(\beta) e(-m\beta)\, d\beta,$$

where

$$\Phi_\alpha(\beta) = \frac{8\pi^{1+\epsilon}}{\prod_\pm \Gamma\left(\frac{1}{4} + \frac{\epsilon}{2} \pm \frac{i\mu}{2}\right)} \left(\frac{N}{\langle \phi, \phi \rangle \cosh \pi\mu}\right)^{1/2} \int_0^\infty y^\epsilon \phi(\alpha + \beta + iy) \frac{dy}{y}.$$

Note that the integral converges by Lemmata 3.1 and 3.2 and satisfies the uniform bound

$$\int_0^\infty y^\epsilon \phi(\alpha + \beta + iy) \frac{dy}{y} \ll_\epsilon \langle \phi, \phi \rangle^{1/2} |\tilde{\mu}|^{3/2+2\epsilon}.$$

It follows that

$$\Phi_\alpha(\beta) \ll_\epsilon N^{1/2}|\tilde{\mu}|^{2+\epsilon},$$

and from this point we proceed exactly as in Case 1. $\square$

3.3 Summation formula

Various Voronoï-type summation formulas are fulfilled by the normalized Fourier coefficients. In the case of full level ($N = 1$) Duke and Iwaniec [Du-Iw] established such a formula for holomorphic cusp forms and Meurman [Me] for Maass cusp forms. These can be generalized to arbitrary level and nebentypus with obvious minor modifications as follows.

Proposition 3.2. *Let d and q be coprime integers such that $N \mid q$, and let g be a smooth, compactly supported function on $(0,\infty)$. If ϕ is a holomorphic cusp form of level N, nebentypus χ and integral weight k then*

$$\chi(d)\sum_{n=1}^{\infty}\lambda_\phi(n)e_q(dn)g(n) = \sum_{n=1}^{\infty}\lambda_\phi(n)e_q\bigl(-\bar{d}n\bigr)\hat{g}(n),$$

where

$$\hat{g}(y) = \frac{2\pi i^k}{q}\int_0^\infty g(x)J_{k-1}\left(\frac{4\pi\sqrt{xy}}{q}\right)\,dx.$$

If ϕ is a real-analytic Maass cusp form of level N, nebentypus χ and nonnegative Laplacian eigenvalue $1/4+\mu^2$ then

$$\chi(d)\sum_{n=1}^{\infty}\lambda_\phi(n)e_q(dn)g(n) = \sum_{\pm}\sum_{n=1}^{\infty}\lambda_\phi(\mp n)e_q\bigl(\pm\bar{d}n\bigr)g^{\pm}(n),$$

where

$$g^{-}(y) = -\frac{\pi}{q\cosh \pi\mu}\int_0^\infty g(x)\{Y_{2i\mu} + Y_{-2i\mu}\}\left(\frac{4\pi\sqrt{xy}}{q}\right)\, dx,$$

$$g^{+}(y) = \frac{4\cosh \pi\mu}{q}\int_0^\infty g(x)K_{2i\mu}\left(\frac{4\pi\sqrt{xy}}{q}\right)\, dx.$$

Here $\bar{d}$ *is a multiplicative inverse of* $d \bmod q$*,* $e_q(x) = e(x/q) = e^{2\pi ix/q}$ *and* J_{k-1}*,* $Y_{\pm 2i\mu}$*,* $K_{2i\mu}$ *are Bessel functions.*

The proof for the holomorphic case [Du-Iw] is a straightforward application of Laplace transforms. Meurman's proof for the real-analytic case [Me] is more involved, but only because he considers a wider class of test functions g and has to deal with delicate convergence issues. For smooth, compactly supported functions g as in our formulation these difficulties do not arise, and one can give a much simpler proof based on Mellin transformation, the functional equations of the L-series attached to additive twists of ϕ (see [Me]), and Barnes' formulae for the gamma function. Indeed, Lemma 5 in [St], a special case of Meurman's summation formula, has been proved by such an approach. We expressed the formula for the non-holomorphic case in terms of K- and Y-Bessel functions in order to emphasize the analogy with the Voronoï-type formula for the divisor function (where one has $\mu = 0$) as derived by Jutila [Ju4, Ju5].

Michel recently extended the above formula to all denominators [Mi1, Mi2]. The extension becomes quite involved when N is not square-free, and the proof relies heavily on Atkin–Lehner theory [At-Le, Li, At-Li]. We shall not use this general version.

3.4 Setting up the circle method

For sake of exposition we shall only present the case of Maass forms and the equation $am - bn = h$. The other cases follow along similar lines by changing Bessel functions and signs at relevant places of the argument. In our inequalities ϵ will always denote a small positive number whose actual value is allowed to change at each occurrence. Implied constants will always depend on ϵ. All other dependencies will be explicitly indicated.

Let ϕ (resp. ψ) be a Maass cusp form of level N, nebentypus χ (resp. ω) and Laplacian eigenvalue $1/4 + \mu^2 \geq 0$ (resp. $1/4 + \nu^2 \geq 0$) whose normalized Fourier coefficients are $\lambda_\phi(m)$ (resp. $\lambda_\psi(n)$). We shall first investigate $D_g(a, b; h)$ for smooth test functions $g(x, y)$ which are supported in a box $[A, 2A] \times [B, 2B]$ and have partial derivatives bounded by

$$g^{(k,l)} \ll_{k,l} A^{-k} B^{-l} P^{k+l}. \tag{3.10}$$

Our aim is to prove the estimate

$$D_g(a, b; h) \ll P^{11/10} N^{9/5} |\tilde{\mu}\tilde{\nu}|^{9/5+\epsilon} (ab)^{-1/10} (A+B)^{1/10} (AB)^{2/5+\epsilon}. \tag{3.11}$$

In Section 3.6 we shall deduce Theorem 3.1 from this bound by employing a partition of unity and decomposing appropriately any smooth test function $f(x, y)$ satisfying (1.18). In fact, (3.11) is a special case of Theorem 3.1, as can be seen upon setting $X = A$, $Y = B$, and $f(x, y) = g(x, y)$.

We shall assume that

$$P^{11} N^8 |\tilde{\mu}\tilde{\nu}|^{13+\epsilon} (ab)^4 \ll \frac{(AB)^{1-\epsilon}}{A+B}, \tag{3.12}$$

for otherwise (3.11) follows from the trivial upper bound

$$D_g(a,b;h) \ll N|\tilde{\mu}\tilde{\nu}|^{1/2}(ab)^{-1/2}(AB)^{1/2}. \tag{3.13}$$

The trivial bound itself is a consequence of $g \ll 1$, Cauchy's inequality, and the mean square estimate (3.7) applied to the forms ϕ and ψ.

As $g(x,y)$ is supported in $[A,2A]\times[B,2B]$, we can assume that $A,B \geq 1/2$, and also that

$$h \leq 2(A+B), \tag{3.14}$$

for otherwise $D_g(a,b;h)$ vanishes trivially. We shall attach, as in [Du-Fr-Iw2], a redundant factor $w(x-y-h)$ to $g(x,y)$, where $w(t)$ is a smooth function supported on $|t| \leq \delta^{-1}$ such that $w(0)=1$ and $w^{(i)} \ll_i \delta^i$. This, of course, does not alter $D_g(a,b;h)$. We choose

$$\delta = P\frac{A+B}{AB}, \tag{3.15}$$

so that, by (3.10), the new function

$$F(x,y) = g(x,y)w(x-y-h)$$

satisfies

$$|x-y-h| > \delta^{-1} \quad \Longrightarrow \quad F(x,y)=0, \tag{3.16}$$

and its partial derivatives are bounded by

$$F^{(k,l)} \ll_{k,l} \delta^{k+l}. \tag{3.17}$$

We apply the Hardy–Littlewood method to detect the equation $am-bn=h$, that is, we express $D_F(a,b;h)$ as the integral of a certain exponential sum over the unit interval $[0,1]$. We get

$$D_g(a,b;h) = D_F(a,b;h) = \int_0^1 G(\alpha)\, d\alpha, \tag{3.18}$$

where

$$G(\alpha) = \sum_{m,n} \lambda_\phi(m)\lambda_\psi(n)F(am,bn)e\big((am-bn-h)\alpha\big).$$

We shall approximate this integral by the following proposition of Jutila (a consequence of the main theorem in [Ju1]).

Proposition 3.3 (Jutila). *Let $\mathcal{Q}$ be a nonempty set of integers $Q \le q \le 2Q$, where $Q \ge 1$. Let $Q^{-2} \le \delta \le Q^{-1}$, and for each fraction d/q (in its lowest terms) denote by $I_{d/q}(\alpha)$ the characteristic function of the interval $[d/q-\delta, d/q+\delta]$. Write L for the number of such intervals, that is,*

$$L = \sum_{q\in\mathcal{Q}} \varphi(q),$$

and put

$$\tilde{I}(\alpha) = \frac{1}{2\delta L} \sum_{q\in\mathcal{Q}} \ \sideset{}{^*}\sum_{d \pmod q} I_{d/q}(\alpha).$$

If $I(\alpha)$ is the characteristic function of the unit interval $[0,1]$, then

$$\int_{-\infty}^{\infty} \big(I(\alpha) - \tilde{I}(\alpha)\big)^2\, d\alpha \ll \delta^{-1}L^{-2}Q^{2+\epsilon},$$

where the implied constant depends on ϵ only.

We shall choose some Q and apply the proposition with a set of denominators of the form

$$\mathcal{Q} = \big\{q \in [Q,2Q] : Nab \mid q \text{ and } (h,q) = (h,Nab)\big\}.$$

By a result of Jacobsthal [Ja], the largest gap between reduced residue classes mod

h is of size $\ll h^\epsilon$, whence (3.14) shows that

$$|\mathcal{Q}| \gg \frac{Q(AB)^{-\epsilon}}{Nab}, \tag{3.19}$$

assuming the right hand side exceeds some large positive constant $c = c(\epsilon) \geq 1$. Moreover, we shall assume that

$$Q^{-2} \leq \delta \leq Q^{-1}, \tag{3.20}$$

so that also

$$1 \leq Q \leq AB, \tag{3.21}$$

whence (3.19) yields

$$L \gg \frac{Q^2(AB)^{-\epsilon}}{Nab}. \tag{3.22}$$

We clearly have

$$|D_F(a,b;h) - \tilde{D}_F(a,b;h)| \leq \|G\|_\infty \|I - \tilde{I}\|_1, \tag{3.23}$$

where

$$\tilde{D}_F(a,b;h) = \int_{-\infty}^{\infty} G(\alpha)\tilde{I}(\alpha)\,d\alpha = \frac{1}{2\delta L}\sum_{q\in\mathcal{Q}}\ \sideset{}{^*}\sum_{d \pmod q} \int_{-\infty}^{\infty} G(\alpha) I_{d/q}(\alpha)\,d\alpha$$

$$= \frac{1}{2\delta L}\sum_{q\in\mathcal{Q}}\ \sideset{}{^*}\sum_{d \pmod q} \int_{-\delta}^{\delta} G(d/q+\beta)\,d\beta = \frac{1}{2\delta L}\sum_{q\in\mathcal{Q}}\ \sideset{}{^*}\sum_{d \pmod q} \mathfrak{I}_{d/q},$$

say. To derive an upper estimate for $G(\alpha)$, we express it as

$$G(\alpha) = \int_0^\infty \int_0^\infty F(x,y)e(-h\alpha)\,dT_{\phi,a\alpha}(x/a)\,dT_{\psi,-b\alpha}(y/b),$$

where the exponential sums $T_{\phi,a\alpha}$ and $T_{\psi,-b\alpha}$ are defined by (3.8). Using the uniform bound provided by Proposition 3.1, and also (3.16) and (3.17), it follows that

$$\|G\|_\infty \ll \frac{N|\tilde{\mu}\tilde{\nu}|^{2+\epsilon}}{(ab)^{1/2}}(AB)^{1/2+\epsilon}\|F^{(1,1)}\|_1 \ll \frac{N|\tilde{\mu}\tilde{\nu}|^{2+\epsilon}\delta}{(ab)^{1/2}} \cdot \frac{(AB)^{3/2+\epsilon}}{A+B}.$$

Also, by (3.22) and Proposition 3.3 we get

$$\|I-\tilde{I}\|_1 \leq 3\|I-\tilde{I}\|_2 \ll \frac{Nab}{\delta^{1/2}Q}(AB)^\epsilon,$$

so that (3.23) becomes

$$D_F(a,b;h) - \tilde{D}_F(a,b;h) \ll \frac{N^2|\tilde{\mu}\tilde{\nu}|^{2+\epsilon}(ab)^{1/2}\delta^{1/2}}{Q} \cdot \frac{(AB)^{3/2+\epsilon}}{A+B}. \tag{3.24}$$

3.5 Transforming exponential sums

The contribution of the interval $[d/q-\delta, d/q+\delta]$ can be expressed as

$$\mathfrak{I}_{d/q} = \int_{-\delta}^{\delta} G(d/q+\beta)\,d\beta = e_q(-dh)\sum_{m,n}\lambda_\phi(m)\lambda_\psi(n)e_q\big(d(am-bn)\big)E(m,n),$$

where

$$E(x,y) = F(ax,by)\int_{-\delta}^{\delta} e\big((ax-by-h)\beta\big)\,d\beta. \tag{3.25}$$

For further reference we record the following two simple consequences of (3.16) and (3.17):

$$E^{(k,l)} \ll_{k,l} \delta^{k+l+1}a^k b^l;$$

$$\|E^{(k,l)}\|_1 \ll_{k,l} \delta^{k+l}a^{k-1}b^{l-1}\frac{AB}{A+B}. \tag{3.26}$$

We assume that $q \in \mathcal{Q}$, hence $Nab \mid q$, and Proposition 3.2 yields

$$\mathfrak{I}_{d/q} = \overline{\chi\omega}(d)e_q(-dh)\sum_{\pm\pm}\sum_{m,n\geq 1}\lambda_\phi(\mp m)\lambda_\psi(\mp n)e_q\big(\bar{d}(\pm am \mp bn)\big)E^{\pm\pm}(m,n),$$

where the corresponding signs must be matched, and

$$E^{\pm\pm}(m,n) = \frac{ab}{q^2}\int_0^\infty\int_0^\infty E(x,y) M^{\pm}_{2i\mu}\left(\frac{4\pi a\sqrt{mx}}{q}\right) M^{\pm}_{2i\nu}\left(\frac{4\pi b\sqrt{ny}}{q}\right)\,dx\,dy,$$

$$M^{+}_{2ir} = (4\cosh\pi r)K_{2ir}, \quad M^{-}_{2ir} = -\frac{\pi}{\cosh\pi r}\{Y_{2ir}+Y_{-2ir}\}.$$

By summing over the residue classes we get

$$\sideset{}{^*}\sum_{d \pmod q} \mathfrak{I}_{d/q} = \sum_{\pm\pm}\sum_{m,n\geq 1}\lambda_\phi(\mp m)\lambda_\psi(\mp n) S_{\overline{\chi\omega}}(-h, \pm am \mp bn; q) E^{\pm\pm}(m,n). \tag{3.27}$$

In order to estimate the twisted Kloosterman sum, we observe that the greatest common divisor $(-h, \pm am \mp bn, q)$ divides $N(h,n,a)(h,m,b)$, as follows from the relations $(a,b)=1$ and $(h,q)=(h,Nab)$. Therefore (3.3) and (3.21) imply that

$$S_{\overline{\chi\omega}}(-h, \pm am \mp bn; q) \ll N^{1/2}(h,m)^{1/2}(h,n)^{1/2}Q^{1/2}(AB)^\epsilon. \tag{3.28}$$

We estimate $E^{\pm\pm}(m,n)$ by successive applications of integration by parts and the recurrence relations

$$\frac{d}{dz}\big(z^s K_s(z)\big) = -z^s K_{s-1}(z), \quad \frac{d}{dz}\big(z^s Y_s(z)\big) = z^s Y_{s-1}(z).$$

Using the first relation we can prove by induction on k that

$$K_s\big(\sqrt{z}\big) = \sum_{\kappa=0}^{k} c_{\kappa k} z^{\kappa-\frac{k}{2}}\left\{K_{s+k}\big(\sqrt{z}\big)\right\}^{(k)}$$

holds with appropriate constants satisfying

$$c_{\kappa k} = c_{\kappa k}(s) \ll_k \big(1+|s|\big)^{k-\kappa}, \quad 0\leq\kappa\leq k.$$

Clearly, for any $\eta > 0$ we also have

$$K_s\big(\eta\sqrt{z}\big) = \eta^{-k}\sum_{\kappa=0}^{k} c_{\kappa k} z^{\kappa-\frac{k}{2}} \left\{K_{s+k}\big(\eta\sqrt{z}\big)\right\}^{(\kappa)}.$$

Similarly, for any positive integer l there are constants

$$d_{\lambda l} = d_{\lambda l}(s) \ll_l \big(1+|s|\big)^{l-\lambda}, \quad 0 \le \lambda \le l,$$

such that for any $\theta > 0$ we have

$$Y_s\big(\theta\sqrt{z}\big) = \theta^{-l}\sum_{\lambda=0}^{l} d_{\lambda l} z^{\lambda-\frac{l}{2}} \left\{Y_{s+l}\big(\theta\sqrt{z}\big)\right\}^{(\lambda)}.$$

By specifying η and θ as

$$\eta = \frac{4\pi a\sqrt{m}}{q}, \qquad \theta = \frac{4\pi b\sqrt{n}}{q},$$

we obtain decompositions of $E^{\pm\pm}(m,n)$ accordingly. In particular, for each pair (k,l) it follows that

$$\begin{aligned} E^{\pm\pm}(m,n) \ll_{k,l} \frac{ab}{q^2}\frac{|\tilde{\mu}|^k|\tilde{\nu}|^l}{\eta^k\theta^l} \sup_{M_1,M_2} \sup_{\substack{0\le\kappa\le k\\ 0\le\lambda\le l}} \\ \int_0^\infty\int_0^\infty x^{\kappa-\frac{k}{2}} y^{\lambda-\frac{l}{2}} E(x,y) \left\{M_1\big(\eta\sqrt{x}\big)\right\}^{(\kappa)} \left\{M_2\big(\theta\sqrt{y}\big)\right\}^{(\lambda)} dx\, dy, \end{aligned} \quad (3.29)$$

where

$$M_1 \in \Big\{(\cosh\pi\mu)K_{k+2i\mu},\ (\cosh\pi\mu)^{-1}Y_{k+2i\mu},\ (\cosh\pi\mu)^{-1}Y_{k-2i\mu}\Big\},$$
$$M_2 \in \Big\{(\cosh\pi\nu)K_{l+2i\nu},\ (\cosh\pi\nu)^{-1}Y_{l+2i\nu},\ (\cosh\pi\nu)^{-1}Y_{l-2i\nu}\Big\}.$$

(3.25) shows that each integral above can be rewritten as

$$\int_{A/a}^{2A/a}\int_{B/b}^{2B/b}\left\{x^{\kappa-\frac{k}{2}}y^{\lambda-\frac{l}{2}}E(x,y)\right\}^{(\kappa,\lambda)}M_1\big(\eta\sqrt{x}\big)M_2\big(\theta\sqrt{y}\big)\,dx\,dy. \tag{3.30}$$

We shall pick a pair (k,l) for each (m,n) in such a way, that the following uniform estimates will hold:

$$\begin{aligned} M_1\big(\eta\sqrt{x}\big) &\ll_k |\tilde{\mu}|^{k+1+\epsilon}\big(\eta\sqrt{x}\big)^{-1/2}, && x\in[A/a,2A/a];\\ M_2\big(\theta\sqrt{y}\big) &\ll_l |\tilde{\nu}|^{l+1+\epsilon}\big(\theta\sqrt{y}\big)^{-1/2}, && y\in[B/b,2B/b]. \end{aligned} \tag{3.31}$$

As $|\Re(i\mu)|$ and $|\Re(i\nu)|$ are at most $1/4$, we can refer to the uniform estimates of Section 3.7 to see that (3.31) holds whenever the assignment $(m,n)\mapsto(k,l)$ is such that

$$\begin{aligned} k>0 \quad &\Longrightarrow \quad \eta\sqrt{A/a}>1,\\ l>0 \quad &\Longrightarrow \quad \theta\sqrt{B/b}>1. \end{aligned} \tag{3.32}$$

The integral (3.30) can be estimated by (3.26), (3.31), and the relation

$$\min(A\delta,B\delta)\geq 1,$$

which follows from (3.15). The resulting bound simplifies (3.29) to

$$E^{\pm\pm}(m,n)\ll_{k,l}\frac{AB}{q^2(A+B)}\frac{|\tilde{\mu}|^{2k+1+\epsilon}|\tilde{\nu}|^{2l+1+\epsilon}}{\eta^{k+\frac{1}{2}}\theta^{l+\frac{1}{2}}}\left(\frac{A}{a}\right)^{-\frac{k}{2}-\frac{1}{4}}\left(\frac{B}{b}\right)^{-\frac{l}{2}-\frac{1}{4}}(A\delta)^k(B\delta)^l.$$

Using that $Q\leq q\leq 2Q$ this can be rewritten as

$$E^{\pm\pm}(m,n)\ll_{k,l}\frac{|\tilde{\mu}\tilde{\nu}|^{\epsilon}(AB)^{1/2}}{\delta Q^2(A+B)}\left(\frac{A|\tilde{\mu}|^4(\delta Q)^2}{am}\right)^{\frac{k}{2}+\frac{1}{4}}\left(\frac{B|\tilde{\nu}|^4(\delta Q)^2}{bn}\right)^{\frac{l}{2}+\frac{1}{4}}. \tag{3.33}$$

This result is conditional under (3.31), but it suggests that in (3.27) we can neglect the contribution of those pairs (m,n) for which $am/A|\tilde{\mu}|^4$ or $bn/B|\tilde{\nu}|^4$ is greater than $(\delta Q)^2(AB)^\epsilon$.

Indeed, this will be the case if we specify $k=\lceil 200/\epsilon\rceil$ or $k=0$ (resp. $l=\lceil 200/\epsilon\rceil$ or $l=0$) depending on whether m (resp. n) is large or small in the above sense. We observe that

$$\begin{aligned} am > A|\tilde{\mu}|^4(\delta Q)^2(AB)^\epsilon \quad &\Longrightarrow \quad \eta\sqrt{A/a}>1,\\ bn > B|\tilde{\nu}|^4(\delta Q)^2(AB)^\epsilon \quad &\Longrightarrow \quad \theta\sqrt{B/b}>1,\end{aligned}$$

therefore our assignment $(m,n)\mapsto(k,l)$ satisfies (3.32), and with this choice (3.33) holds uniformly for all (m,n) with an implied constant depending only on ϵ.

It follows from (3.7) applied to ϕ and ψ, that

$$\begin{aligned} \sum_{1\le m\le x}|\lambda_\phi(\mp m)|(h,m)^{1/2} &\ll N^{1/2}|\tilde{\mu}|^{1/2}x\tau^{1/2}(h),\\ \sum_{1\le n\le y}|\lambda_\psi(\mp n)|(h,n)^{1/2} &\ll N^{1/2}|\tilde{\nu}|^{1/2}y\tau^{1/2}(h).\end{aligned} \tag{3.34}$$

Combining this bound with (3.28) and (3.33), we see that the total contribution to (3.27) of small pairs (m,n) is

$$\ll \frac{N^{3/2}|\tilde{\mu}\tilde{\nu}|^{3/2+\epsilon}\delta^3Q^{5/2}}{ab}\cdot\frac{(AB)^{3/2+\epsilon}}{A+B}.$$

On the other hand, (3.28), (3.33) and (3.34) similarly show that the remaining contribution from the pairs (m,n) with m or n large is

$$\ll \frac{N^{3/2}|\tilde{\mu}\tilde{\nu}|^{3/2+\epsilon}\delta^3Q^{5/2}}{ab}\cdot\frac{(AB)^{-50}}{A+B}.$$

To summarize, we have shown that

$$\sideset{}{^*}\sum_{d \pmod q} \mathfrak{I}_{d/q} \ll \frac{N^{3/2}|\tilde{\mu}\tilde{\nu}|^{3/2+\epsilon}\delta^3 Q^{5/2}}{ab} \cdot \frac{(AB)^{3/2+\epsilon}}{A+B}.$$

Hence, by (3.22),

$$\tilde{D}_F(a,b;h) = \frac{1}{2\delta L}\sum_{q\in\mathcal{Q}} \sideset{}{^*}\sum_{d \pmod q} \mathfrak{I}_{d/q} \ll \frac{N^{3/2}|\tilde{\mu}\tilde{\nu}|^{3/2+\epsilon}\delta^2 Q^{3/2}}{ab} \cdot \frac{(AB)^{3/2+\epsilon}}{A+B}. \qquad (3.35)$$

Inequalities (3.24) and (3.35) show that the optimal balance is achieved when

$$\delta^3 Q^5 \asymp N|\tilde{\mu}\tilde{\nu}|(ab)^3.$$

A natural choice is given by

$$\delta^3 Q^5 = N|\tilde{\mu}\tilde{\nu}|(cab)^3,$$

where c is the constant appearing in the remark after (3.19). Then, by (3.12), the conditions of Proposition 3.3 are satisfied, that is, both (3.20) and $Q \geq cNab(AB)^\epsilon$ hold. (3.24) and (3.35) add up to (3.11), using also (3.18).

3.6 Dyadic decomposition

Our aim is to prove Theorem 3.1 for all test functions $f(x, y)$ satisfying (1.18). We fix an arbitrary smooth function

$$\rho : (0, \infty) \to \mathbb{R}$$

whose support lies in $[1,2]$ and which satisfies the following identity on the positive axis:

$$\sum_{i=-\infty}^{\infty} \rho\bigl(2^{-i/2}x\bigr) = 1.$$

To obtain such a function, we take an arbitrary smooth $\eta : (0,\infty) \to \mathbb{R}$ which is constant 0 on $(0,1)$ and constant 1 on $(\sqrt{2},\infty)$, and then define ρ as

$$\rho(x) = \begin{cases} \eta(x) & \text{if } 0 < x \leq \sqrt{2}, \\ 1 - \eta(x/\sqrt{2}) & \text{if } \sqrt{2} < x < \infty. \end{cases}$$

According to this partition of unity we decompose $f(x,y)$ as

$$f(x,y) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} f_{i,j}(x,y),$$

$$f_{i,j}(x,y) = f(x,y)\rho\left(\frac{x}{2^{i/2}X}\right)\rho\left(\frac{y}{2^{j/2}Y}\right).$$

Observe that

$$\operatorname{supp} f_{i,j} \subseteq [A_i, 2A_i] \times \bigl[B_j, 2B_j\bigr], \quad A_i = 2^{i/2}X, \quad B_j = 2^{j/2}Y, \tag{3.36}$$

whence (1.18) and $P \geq 1$ show that

$$\bigl(1 + 2^{i/2}\bigr)\bigl(1 + 2^{j/2}\bigr) f_{i,j}^{(k,l)} \ll_{k,l} A_i^{-k} B_j^{-l} P^{k+l}.$$

In other words, the bound (3.11) applies uniformly to each function

$$g_{i,j}(x,y) = \bigl(1 + 2^{i/2}\bigr)\bigl(1 + 2^{j/2}\bigr) f_{i,j}(x,y)$$

with the corresponding parameters $A = A_i$, $B = B_j$:

$$D_{g_{i,j}}(a, b; h) \ll P^{11/10} N^{9/5} |\tilde{\mu}\tilde{\nu}|^{9/5+\epsilon} (ab)^{-1/10} (A_i + B_j)^{1/10} (A_i B_j)^{2/5+\epsilon}.$$

This implies, for $\epsilon < 1/10$,

$$D_{f_{i,j}}(a, b; h) \ll 2^{-|i|/5} 2^{-|j|/5} P^{11/10} N^{9/5} |\tilde{\mu}\tilde{\nu}|^{9/5+\epsilon} (ab)^{-1/10} (X + Y)^{1/10} (XY)^{2/5+\epsilon}.$$

Finally,

$$D_f(a, b; h) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} D_{f_{i,j}}(a, b; h)$$

completes the proof of Theorem 3.1.

It should be noted that the trivial upper bound (3.1) mentioned in Section 3.1 follows by a similar reduction technique from the Cauchy bounds

$$D_{g_{i,j}}(a, b; h) \ll N |\tilde{\mu}\tilde{\nu}|^{1/2} (ab)^{-1/2} (A_i B_j)^{1/2}$$

of Section 3.4 (cf. (3.13)).

3.7 Bounds for Bessel functions

In this section we prove uniform bounds for Bessel functions of the first kind (Proposition 3.4) and of the second and third kinds (Proposition 3.5).

Proposition 3.4. *For any integer $k \geq 1$ the following uniform estimate holds:*

$$J_{k-1}(x) \ll \begin{cases} \frac{x^{k-1}}{2^{k-1}\Gamma\left(k-\frac{1}{2}\right)}, & 0 < x \leq 1; \\ k x^{-1/2}, & 1 < x. \end{cases}$$

The implied constant is absolute.

Proof. For $x > k^2$ the asymptotic expansion of J_{k-1} (see Section 7.13.1 of [Ol]) provides the stronger estimate $J_{k-1}(x) \ll x^{-1/2}$ with an absolute implied constant.

For $1 < x \leq k^2$ we use Bessel's original integral representation (see Section 2.2 of [Wat]),

$$J_{k-1}(x) = \frac{1}{2\pi} \int_0^{2\pi} \cos\big((k-1)\theta - x\sin\theta\big)\, d\theta,$$

to deduce that in this range

$$|J_{k-1}(x)| \leq 1 \leq kx^{-1/2}.$$

For the remaining range $0 < x \leq 1$ the required estimate follows from the Poisson-Lommel integral representation (see Section 3.3 of [Wat])

$$J_{k-1}(x) = \frac{x^{k-1}}{2^{k-1}\Gamma\left(k-\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right)} \int_0^{\pi} \cos(x\cos\theta)\sin^{2k-2}\theta\, d\theta. \quad \square$$

Proposition 3.5. *For any $\sigma > 0$ and $\epsilon > 0$ the following uniform estimates hold in the strip $|\Re s| \leq \sigma$:*

$$e^{-\pi|\Im s|/2} Y_s(x) \ll \begin{cases} \big(1+|\Im s|\big)^{\sigma+\epsilon} x^{-\sigma-\epsilon}, & 0 < x \leq 1+|\Im s|; \\ \big(1+|\Im s|\big)^{-\epsilon} x^{\epsilon}, & 1+|\Im s| < x \leq 1+|s|^2; \\ x^{-1/2}, & 1+|s|^2 < x. \end{cases}$$

$$e^{\pi|\Im s|/2} K_s(x) \ll \begin{cases} \big(1+|\Im s|\big)^{\sigma+\epsilon} x^{-\sigma-\epsilon}, & 0 < x \leq 1+\pi|\Im s|/2; \\ e^{-x+\pi|\Im s|/2} x^{-1/2}, & 1+\pi|\Im s|/2 < x. \end{cases}$$

The implied constants depend only on σ and ϵ.

Proof. The last estimate for Y_s follows from its asymptotic expansion (see Section 7.13.1 of [Ol]). The last estimate for K_s follows from Schläfli's integral representation (see Section 6.22 of [Wat]),

$$K_s(x) = \int_0^\infty e^{-x\cosh t} \cosh st.\, dt,$$

by noting that

$$\cosh t \geq 1 + t^2/2 \quad \text{and} \quad |\cosh st| \leq e^{\sigma t}.$$

We shall deduce the remaining uniform bounds from the integral representations

$$4K_s(x) = \frac{1}{2\pi i} \int_{\mathcal{C}} \Gamma\left(\frac{w-s}{2}\right) \Gamma\left(\frac{w+s}{2}\right) \left(\frac{x}{2}\right)^{-w} dw,$$

$$-2\pi Y_s(x) = \frac{1}{2\pi i} \int_{\mathcal{C}} \Gamma\left(\frac{w-s}{2}\right) \Gamma\left(\frac{w+s}{2}\right) \cos\left(\frac{\pi}{2}(w-s)\right) \left(\frac{x}{2}\right)^{-w} dw,$$

where the contour $\mathcal{C}$ is a broken line of 2 infinite and 3 finite segments joining the points

$$\begin{array}{ccc} -\epsilon - i\infty, & -\epsilon - i\big(2+2|\Im s|\big), & \sigma+\epsilon - i\big(2+2|\Im s|\big), \\ \sigma+\epsilon + i\big(2+2|\Im s|\big), & -\epsilon + i\big(2+2|\Im s|\big), & -\epsilon + i\infty. \end{array}$$

These formulae follow by analytic continuation from the well-known but more restrictive inverse Mellin transform representations of the K- and Y-Bessel functions, cf. formulae 6.8.17 and 6.8.26 in [Er].

If we write in the second formula

$$\cos\left(\frac{\pi}{2}(w-s)\right) = \cos\left(\frac{\pi}{2}w\right)\cos\left(\frac{\pi}{2}s\right) + \sin\left(\frac{\pi}{2}w\right)\sin\left(\frac{\pi}{2}s\right),$$

then it becomes apparent that the remaining inequalities of the lemma can be deduced from the uniform bound

$$\int_{\mathcal{C}} e^{\pi \max(|\Im s|,|\Im w|)/2} \left| \Gamma\left(\frac{w-s}{2}\right) \Gamma\left(\frac{w+s}{2}\right) \left(\frac{x}{2}\right)^{-w} dw \right| \\ \ll_{\sigma,\epsilon} \left(\frac{x}{1+|\Im s|}\right)^{-\sigma-\epsilon} + \left(\frac{x}{1+|\Im s|}\right)^{\epsilon}.$$

By introducing the notation

$$G(s) = e^{\pi|\Im s|/2}\Gamma(s),$$

$$M_s(x) = \int_{\mathcal{C}} \left| G\left(\frac{w-s}{2}\right) G\left(\frac{w+s}{2}\right) \left(\frac{x}{2}\right)^{-w} dw \right|,$$

the previous inequality can be rewritten as

$$M_s(x) \ll_{\sigma,\epsilon} \left(\frac{x}{1+|\Im s|}\right)^{-\sigma-\epsilon} + \left(\frac{x}{1+|\Im s|}\right)^{\epsilon}. \tag{3.37}$$

Case 1. $|\Im s| \le 1$.

If w lies on either horizontal segments of $\mathcal{C}$ or on the finite vertical segment joining $\sigma + \epsilon \pm i\big(2 + 2|\Im s|\big)$, then $w \pm s$ varies in a fixed compact set (depending only on σ and ϵ) disjoint from the negative axis $(-\infty, 0]$. It follows that for these values w we have

$$G\left(\frac{w-s}{2}\right) G\left(\frac{w+s}{2}\right) \ll_{\sigma,\epsilon} 1,$$

i.e.,

$$G\left(\frac{w-s}{2}\right) G\left(\frac{w+s}{2}\right) \left(\frac{x}{2}\right)^{-w} \ll_{\sigma,\epsilon} x^{-\sigma-\epsilon},$$

and the same bound holds for the contribution of these values to $M_s(x)$.

If w lies on either infinite vertical segments of $\mathcal{C}$, then

$$|\Im(w \pm s)| \asymp |\Im w| > 1,$$

whence Stirling's approximation yields

$$G\left(\frac{w-s}{2}\right)G\left(\frac{w+s}{2}\right) \asymp_\epsilon |\Im w|^{-\epsilon-1}.$$

It follows that the contribution of the infinite segments to $M_s(x)$ is $\ll_{\sigma,\epsilon} x^\epsilon$.

Altogether we infer that

$$M_s(x) \ll_{\sigma,\epsilon} x^{-\sigma-\epsilon} + x^\epsilon,$$

which is equivalent to (3.37).

Case 2. $|\Im s| > 1$.

If w lies on either horizontal segments of $\mathcal{C}$, then

$$|\Im(w \pm s)| \asymp |\Im s|,$$

whence Stirling's approximation yields

$$G\left(\frac{w-s}{2}\right)G\left(\frac{w+s}{2}\right) \asymp_{\sigma,\epsilon} |\Im s|^{\Re w-1},$$

i.e.,

$$G\left(\frac{w-s}{2}\right)G\left(\frac{w+s}{2}\right)\left(\frac{x}{2}\right)^{-w} \asymp_{\sigma,\epsilon} \frac{1}{|\Im s|}\left(\frac{|\Im s|}{x}\right)^{\Re w}.$$

It follows that the contribution of the horizontal segments to $M_s(x)$ is

$$\ll_{\sigma,\epsilon} |\Im s|^{-1+\sigma+\epsilon}x^{-\sigma-\epsilon} + |\Im s|^{-1-\epsilon}x^\epsilon.$$

If w lies on the finite vertical segment of $\mathcal{C}$ joining $\sigma+\epsilon \pm i\big(2+2|\Im s|\big)$, then

$$\Re(w \pm s) \geq \epsilon \qquad \text{and} \qquad \max|\Im(w \pm s)| \asymp |\Im s|,$$

whence Stirling's approximation implies

$$G\left(\frac{w-s}{2}\right)G\left(\frac{w+s}{2}\right) \ll_{\sigma,\epsilon} \begin{cases} |\Im s|^{\sigma+\epsilon/2-1/2} & \text{if } \min|\Im(w\pm s)| \le 1; \\ |\Im s|^{\sigma+\epsilon-1} & \text{if } \min|\Im(w\pm s)| > 1. \end{cases}$$

It follows that the contribution of the finite vertical segment to $M_s(x)$ is

$$\ll_{\sigma,\epsilon} |\Im s|^{\sigma+\epsilon}x^{-\sigma-\epsilon}.$$

If w lies on either infinite vertical segments of $\mathcal{C}$, then

$$|\Im(w\pm s)| \asymp |\Im w| > |\Im s|,$$

whence Stirling's approximation yields

$$G\left(\frac{w-s}{2}\right)G\left(\frac{w+s}{2}\right) \asymp_{\epsilon} |\Im w|^{-\epsilon-1}.$$

It follows that the contribution of the infinite vertical segments to $M_s(x)$ is

$$\ll_{\sigma,\epsilon} |\Im s|^{-\epsilon}x^{\epsilon}.$$

Altogether we infer that

$$M_s(x) \ll_{\sigma,\epsilon} |\Im s|^{\sigma+\epsilon}x^{-\sigma-\epsilon} + |\Im s|^{-\epsilon}x^{\epsilon},$$

which is equivalent to (3.37).

The proof of Proposition 3.5 is complete. $\square$

Chapter 4

Twists of Maass forms: a subconvex bound for L-functions

4.1 Overview

We shall prove a subconvex estimate on the critical line for L-functions associated to character twists of a fixed holomorphic or Maass cusp form ϕ of arbitrary level and nebentypus. We borrow notation from Section 3.2, and we also refer the reader to Section 1.2 for an introduction. The result, in less explicit form, will also appear in [Ha2].

We assume that ϕ is a *primitive form*, that is, a newform in the sense of [At-Le, Li, At-Li] normalized so that $\rho_\phi(1) = 1$. If we renormalize the Fourier coefficients of ϕ as

$$\lambda_\phi(n) = |n|^{\frac{1-k}{2}} \rho_\phi(n),$$

then $\lambda_\phi(n)$ $(n \geq 1)$ defines a character of the corresponding Hecke algebra, while $\lambda_\phi(-n) = \pm\lambda_\phi(n)$ (with a constant sign) when ϕ is a Maass form. In other words, ϕ defines a cuspidal automorphic representation of GL_2 over $\mathbb{Q}$ with arithmetic conductor N. The contragradient representation corresponds to the primitive cusp form

$\tilde{\phi}(z) = \bar{\phi}(-\bar{z})$ with renormalized Fourier coefficients $\lambda_{\tilde{\phi}}(n) = \bar{\lambda}_\phi(n)$. We note that by the powerful results of Iwaniec [Iw2] and Hoffstein–Lockhart [Ho-Lo], the old normalization (3.6) and the present one are essentially the same in that the scaling factor c between them satisfies

$$N^{-\epsilon}|\tilde{\mu}|^{-\epsilon} \ll_\epsilon c \ll_\epsilon N^\epsilon |\tilde{\mu}|^\epsilon.$$

We consider the twisted representations $\phi \otimes \chi$ as χ runs through the automorphic representations of GL_1 over $\mathbb{Q}$, that is, the primitive Dirichlet characters of the rational integers. In order to simplify our discussion, we shall assume that q, the conductor of χ, is prime to N. Then the analytic conductor of $\phi \otimes \chi$ satisfies

$$C(s, \phi \otimes \chi) \asymp q^2 N\big(|s|^2 + |\tilde{\mu}|^2\big), \qquad \Re s = \frac{1}{2}, \tag{4.1}$$

and for $\Re s > 1$ the associated L-function is given by

$$L(s, \phi \otimes \chi) = \sum_{n=1}^{\infty} \frac{\lambda_\phi(n)\chi(n)}{n^s}.$$

For a fixed point s on the critical line the convexity bound (1.12) implies that

$$L(s, \phi \otimes \chi) \ll_\epsilon |s|^{1/2+\epsilon} N^{1/4+\epsilon} |\tilde{\mu}|^{1/2+\epsilon} q^{1/2+\epsilon}.$$

Our aim is to decrease the exponent $1/2$ of q and still maintain polynomial control in the other parameters $|s|$, N, $|\tilde{\mu}|$.

Theorem 4.1. *Suppose that ϕ is a primitive holomorphic or Maass cusp form of Archimedean size $|\tilde{\mu}|$, level N and arbitrary nebentypus character mod N. Let $\Re s = 1/2$ and q be an integer prime to N. If χ is a primitive Dirichlet character modulo q, then*

$$L(s, \phi \otimes \chi) \ll |s|^{1+\epsilon} N^{9/8+\epsilon} |\tilde{\mu}|^{27/20+\epsilon} q^{1/2-1/54+\epsilon}, \tag{4.2}$$

where the implied constant depends only on ϵ.

A similar estimate with q-exponent $1/2 - 1/22$ was proved for holomorphic forms of full level in [Du-Fr-Iw1], and the improved exponent $1/2 - 7/130$ follows for holomorphic forms of arbitrary level as a special case of the main result in [Co-PS-Sa]. Duke, Friedlander and Iwaniec anticipated their method to be extendible to more general L-functions of rank two, and the present chapter is indeed an extension of their work. The very general Voronoï-formula of Michel enables one to establish Theorem 4.1 in slightly stronger form, e.g. with the original q-exponent $1/2 - 1/22$ of [Du-Fr-Iw1]. See [Mi2] for details.

Combining the estimate (4.2) at the central point $s = 1/2$ with Waldspurger's theorem [Wal] (see also [Koh, Sh]), we get the bound

$$c(q) \ll_\epsilon q^{1/4-1/108+\epsilon}, \quad q \text{ square-free}$$

for the normalized Fourier coefficients of half-integral weight forms of arbitrary level. Such a nontrivial bound is the key step in the solution of the general ternary Linnik problem given by Duke and Schulze-Pillot [Du, Du-SP].

The proof of Theorem 4.1 is presented in Sections 4.2 through 4.4. In Section 4.2 we reduce (4.2), via the approximate functional equation of Chapter 2, to an inequality about certain finite sums involving at most $C(s, \phi\otimes\chi)^{1/2+\epsilon}$ terms (cf. (4.1)). We prove this inequality in Section 4.3 by employing the amplification method. As discussed in Section 1.4, the idea is to consider a suitably weighted second moment of the finite sums arising from the family $\phi\otimes\chi$ of cusp forms (χ varies, ϕ is fixed). We choose the weights (called amplifiers) in such a way that one of the characters χ is emphasized, while the second moment average is still of moderate size. This forces, by positivity, $L(s, \phi\otimes\chi)$ to be small. In the course of evaluating the amplified second moment we encounter diagonal and off-diagonal terms. The off-diagonal terms decompose to shifted convolution sums, and at this point we apply Theorem 3.1.

4.2 Approximate functional equation

Using the approximate functional equation in the form Corollary 2.1, we can see that (4.2) is equivalent to

$$\sum_{n\leq C^{1/2+\epsilon}} \frac{\lambda_\phi(n)\chi(n)}{n^s} f\left(\frac{n}{\sqrt{C}}\right) \ll_\epsilon |s|^{1+\epsilon} N^{9/8+\epsilon} |\tilde{\mu}|^{27/20+\epsilon} q^{1/2-1/54+\epsilon},$$

where

$$C = C(s, \phi\otimes\chi) \ll |s|^2 N |\tilde{\mu}|^2 q^2,$$

and $f : (0,\infty) \to \mathbb{C}$ is a smooth function satisfying (2.3) and (2.4) with $m = 2$. In particular, we can write the left hand side as

$$\sum_{n\leq C^{1/2+\epsilon}} \frac{\lambda_\phi(n)\chi(n)g(n)}{\sqrt{n}},$$

where

$$g(x) = x^{1/2-s} f\left(\frac{x}{\sqrt{C}}\right)$$

satisfies the uniform bounds

$$g^{(k)}(x) \ll_k |s|^k x^{-k}.$$

Therefore, applying partial summation and a smooth dyadic decomposition, we can reduce Theorem 4.1 to the following

Proposition 4.1. *Let* $1 \leq T \leq \left(|s| N^{1/2} |\tilde{\mu}| q\right)^{1+\epsilon}$ *and* W *be a smooth complex valued function supported in* $[T, 2T]$ *such that* $W^{(k)} \ll_k |s|^k T^{-k}$. *Then*

$$\sum_{n=1}^{\infty} \lambda_\phi(n)\chi(n)W(n) \ll |s|^{5/6+\epsilon} N^{25/24+\epsilon} |\tilde{\mu}|^{71/60+\epsilon} q^{17/54+\epsilon} T^{2/3},$$

where the implied constant depends only on ϵ.

4.3 Amplification

Our purpose is to prove Proposition 4.1. As in [Du-Fr-Iw1], we shall estimate from both ways the amplified second moment

$$S = \sum_{\omega \bmod q}^{*} \left| \sum_{1 \leq l \leq L} \bar{\chi}(l)\omega(l) \right|^2 |S_\omega|^2,$$

where ω runs through the primitive characters modulo q, L is a parameter to be chosen later in terms of M and q, and

$$S_\omega = \sum_{n=1}^{\infty} \lambda_\phi(n)\omega(n)W(n).$$

Assuming $L \geq c(\epsilon)q^\epsilon$, it follows, using the result of Jacobsthal [Ja] that the largest gap between reduced residue classes mod q is of size $\ll q^\epsilon$, that

$$S \gg q^{-\epsilon} L^2 |S_\chi|^2. \tag{4.3}$$

Here and in the sequel implied contants may depend on ϵ.

On the other hand, expanding each primitive ω in S using Gauss sums and then extending the resulting summation to all characters mod q, we get by orthogonality,

$$S \leq \frac{\phi(q)}{q} \sum_{d \ (\mathrm{mod}\ q)}^{*} \left| \sum_m a(m) e_q(dm) \right|^2,$$

where

$$a(m) = \sum_{\substack{ln=m \\ 1 \leq l \leq L}} \bar{\chi}(l)\lambda_\phi(n)W(n).$$

It is clear that the coefficients $a(m)$ are supported in the interval $[1, M]$, where $M = 2LT$. Extending the summation to all residue classes d, the previous inequality becomes

$$S \leq \phi(q) \sum_{h \equiv 0 \pmod{q}} D(h), \tag{4.4}$$

where

$$D(h) = \sum_{m_1 - m_2 = h} a(m_1)\bar{a}(m_2).$$

We estimate the diagonal contribution $D(0)$ using the following Rankin–Selberg bound (Theorem 8.3 in [Iw1]):

$$\sum_{1 \leq n \leq x} |\lambda_\phi(n)|^2 \ll N^\epsilon |\tilde{\mu}|^\epsilon x.$$

Indeed, by $W \ll 1$ we get

$$D(0) = \sum_m |a(m)|^2 \ll \sum_{\substack{l_1 n_1 = l_2 n_2 \\ 1 \leq l_1, l_2 \leq L \\ T \leq n_1, n_2 \leq 2T}} \lambda_\phi(n_1)\bar{\lambda}_\phi(n_2)$$

$$\ll \sum_{\substack{1 \leq l \leq L \\ T \leq n \leq 2T}} |\lambda_\phi(n)|^2 \tau(nl) \ll N^\epsilon |\tilde{\mu}|^\epsilon M^\epsilon L \sum_{T \leq n \leq 2T} |\lambda_\phi(n)|^2,$$

whence

$$D(0) = \sum_m |a(m)|^2 \ll N^\epsilon |\tilde{\mu}|^\epsilon M^{1+\epsilon}. \tag{4.5}$$

We estimate the non-diagonal terms $D(h)$ $(h \neq 0)$ using Theorem 3.1. Clearly, we can rewrite each term as

$$D(h) = \sum_{1 \leq l_1, l_2 \leq L} \bar{\chi}(l_1)\chi(l_2) \sum_{l_1 n_1 - l_2 n_2 = h} \lambda_\phi(n_1)\bar{\lambda}_\phi(n_2) W(n_1)\bar{W}(n_2).$$

The inner sum is of type (1.19), because $\bar{\lambda}_\phi(n)$ is just the n-th renormalized Fourier coefficient of the contragradient cusp form $\tilde{\phi}(z) = \bar{\phi}(-\bar{z})$. For each pair (l_1, l_2) we apply Theorem 3.1 with $a = l_1/(l_1, l_2)$, $b = l_2/(l_1, l_2)$, $P = 2|s|$, $X = aT$ and $Y = bT$ to conclude that

$$\begin{aligned} D(h) &\ll L^2 |s|^{11/10} N^{9/5+\epsilon} |\tilde{\mu}|^{9/5+\epsilon} (a+b)^{1/10} (ab)^{3/10+\epsilon} T^{9/10+\epsilon} \\ &\ll |s|^{11/10} N^{9/5+\epsilon} |\tilde{\mu}|^{9/5+\epsilon} L^{27/10+\epsilon} T^{9/10+\epsilon}. \end{aligned} \tag{4.6}$$

4.4 Optimizing parameters

Inserting the bounds (4.5) and (4.6) into (4.4), it follows that

$$S \ll N^{\epsilon} |\tilde{\mu}|^{\epsilon} M^{\epsilon} \phi(q) \left(M + \frac{M}{q} |s|^{11/10} N^{9/5} |\tilde{\mu}|^{9/5} L^{27/10} T^{9/10} \right).$$

This shows that the optimal choice for L is provided by

$$q \asymp L^{27/10} T^{9/10}.$$

In order to maintain $L \geq c(\epsilon) q^{\epsilon}$, we choose

$$\left(|s| N^{1/2} |\tilde{\mu}| \right)^{9/10+\epsilon} q = L^{27/10} T^{9/10}. \tag{4.7}$$

This shows that

$$S \ll |s|^{2+\epsilon} N^{9/4+\epsilon} |\tilde{\mu}|^{27/10+\epsilon} q M^{1+\epsilon},$$

and then (4.3) yields

$$S_{\chi} \ll q^{\epsilon} L^{-1} |S|^{1/2} \ll \left(|s|^2 N^{9/4} |\tilde{\mu}|^{27/10} q T/L \right)^{1/2+\epsilon}.$$

Substituting (4.7) we get

$$S_\chi \ll \left\{|s|^2 N^{9/4} |\tilde{\mu}|^{27/10} q T \big(|s| N^{1/2} |\tilde{\mu}|\big)^{-1/3} q^{-10/27} T^{1/3}\right\}^{1/2+\epsilon},$$

which is precisely the conclusion of Proposition 4.1. □

The proof of Theorem 4.1 is complete.

Chapter 5

Shifted convolution sums and spectral theory

5.1 Overview

We shall obtain a fairly precise description of the continuous span of the functions $H_{s,0,i\mu}$ corresponding to values s on a vertical line $\sigma + i\mathbb{R}$, $\sigma > 1$. These functions play an important role in the Sarnak–Selberg spectral method applied to Maass forms. We refer the reader to Section 1.6 for an introduction. For convenience we shall assume that $\mu \in \mathbb{R}$.

By definition,

$$H_{s,0,i\mu}(u) = \int_0^\infty \tilde{W}_{0,i\mu}\big(|u+1|y\big)\bar{\tilde{W}}_{0,i\mu}\big(|u-1|y\big)y^{s-2}\,dy$$

$$= \frac{|u^2-1|^{\frac{1}{2}}}{\pi}\int_0^\infty K_{i\mu}\left(\frac{|u+1|y}{2}\right)K_{i\mu}\left(\frac{|u-1|y}{2}\right)y^s\frac{dy}{y}.$$

In particular, $H_{s,0,i\mu}(u)$ is an even function of u, therefore we can regard it as a function on the positive axis. Combining formulae 6.576.4 and 9.134.3 from [Gr-Ry], we can see that

$$H_{s,0,i\mu}(u) = M(s)u|1-u^{-2}|^{\frac{1}{2}+i\mu}G_s(u),$$

where

$$M(s) = \frac{2^{2s-3}\Gamma\left(\frac{s}{2}-i\mu\right)\Gamma^2\left(\frac{s}{2}\right)\Gamma\left(\frac{s}{2}+i\mu\right)}{\pi\Gamma(s)},$$

and

$$G_s(u) = \begin{cases} u^{2i\mu}F\left(\frac{s}{2}+i\mu, \frac{1}{2}+i\mu; \frac{s}{2}+\frac{1}{2}; u^2\right), & 0 \le u < 1; \\ u^{-s}F\left(\frac{s}{2}+i\mu, \frac{1}{2}+i\mu; \frac{s}{2}+\frac{1}{2}; u^{-2}\right), & 1 < u. \end{cases}$$

This explicit decomposition reduces our task to analyze the set of functions V on the positive axis that can be represented in the form

$$V(u) = \frac{1}{2\pi i}\int_{(\sigma)} V^{\bigstar}(s)G_s(u)\,ds. \tag{5.1}$$

In our formal definition we include an assumption on the growth rate of $V^{\bigstar}(s)$, which is natural and justified by the dependencies on s of the bounds in Lemmata 5.1 and 5.2.

Definition. Let V be an arbitrary complex valued function on the positive axis $(0,\infty)$, and $V^{\bigstar}(s)$ be a complex valued function on the vertical line $\sigma + i\mathbb{R}$ $(\sigma > 1)$, such that

$$\int_{(\sigma)} |s|^{3/2+\epsilon}|V^{\bigstar}(s)|\,ds < \infty \tag{5.2}$$

holds for some $\epsilon > 0$. Then $V^{\bigstar}$ is a $\bigstar$ transform of V if (5.1) is valid for all $u > 0$, $u \neq 1$.

Theorem 5.1. *Suppose that an arbitrary function $V : (0,\infty) \to \mathbb{C}$ has a $\bigstar$ transform on the vertical line $\sigma + i\mathbb{R}$ $(\sigma > 1)$. Then V is continuous at all points $u \neq 1$, the Mellin transform $V^*(z)$ of V is defined in $0 < \Re z < 1$, and $\frac{\Gamma\left(\frac{z}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{z}{2}+i\mu\right)}V^*(z)$ extends to a bounded holomorphic function in every half-plane $\Re z < \sigma_0 < \sigma$. Conversely,* *let $V:(0,\infty)\to\mathbb{C}$ be an arbitrary function which is continuous at all points $u\neq 1$ and has Mellin transform $V^*(z)$ defined in $0<\Re z<1$. If $K(z)=\frac{\Gamma\left(\frac{z}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{z}{2}+i\mu\right)}V^*(z)$ extends to a holomorphic function in some half-plane $\Re z<\sigma_0$ ($\sigma_0>1$) satisfying $K(z)\ll(1+|z|)^{-A}$ for some $A>2$, then V has a $\bigstar$ transform $V^\bigstar(s)$, which extends to a holomorphic function in $0<\Re s<\sigma_0$ satisfying $V^\bigstar(s)\ll_{\sigma,A}\big(1+|s|\big)^{-A-1/2}$.*

The theorem shows that the functions $H_{s,0,i\mu}$ form an incomplete system in the sense that some of the very natural functions V are excluded from their continuous span. For the Sarnak–Selberg method this negative conclusion has the message that the Maass operators must play a crucial role in a successful analysis.

Corollary 5.1. *Let $V:(0,\infty)\to\mathbb{C}$ be an arbitrary function compactly supported in $(0,1)\cup(1,\infty)$. If V has a $\bigstar$ transform, then it is identically zero.*

Proof. By Theorem 5.1, V is a continuous function of compact support whose Mellin transform vanishes at all negative odd integers. In other words, $V(u)$ is orthogonal to all functions u^{-2k} ($k=1,2,\dots$). It follows that $u^{-2}V(u)$ is orthogonal to all functions $p(u^{-2})$, where p is an arbitrary complex polynomial. These functions are dense among continuous functions on a compact interval by Weierstrass' approximation theorem, hence $V=0$. □

Corollary 5.2. *Let $V:(0,\infty)\to\mathbb{C}$ be an arbitrary function. If $V(u/c)$ has a $\bigstar$ transform for every $c>0$, then V is identically zero.*

Proof. By Theorem 5.1,

$$K(z)=\frac{\Gamma\left(\frac{z}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{z}{2}+i\mu\right)}V^*(z)$$

is defined in the strip $0<\Re z<1$ and extends to a holomorphic function in the half-plane $\Re z<1$. Moreover, for any $c>0$, $c^zK(z)$ is bounded. This forces $K(z)=0$ as follows. Let $f(u)$ be the inverse Mellin transform of $K(z)/(2-z)^2$, i.e.,

$$f(u)=\frac{1}{2\pi i}\int_{(\sigma)}u^{-z}\frac{K(z)}{(2-z)^2}\,dz$$

for any $\sigma < 1$. We can see that $f(u)$ is independendent of the particular line of integration. However, the assumption that $c^z K(z)$ is bounded for any $c > 0$ implies uniform bounds of the form

$$f(u) \ll_c (cu)^{-\sigma}, \quad u > 0, \quad \sigma < 1,$$

the implied constant depending on c only. By letting $\sigma \to -\infty$, we can conclude, for each $c > 0$, that $f(u)$ vanishes on $(1/c, \infty)$. Hence $f(u)$ is identically zero, and

$$K(z) = (2-z)^2 \int_0^\infty u^z f(u) \frac{du}{u} = 0, \quad \Re z < 1,$$

as claimed.

Therefore V^* must vanish in $0 < \Re z < 1$, which shows that $V(u) = 0$ as long as $u \neq 1$. We can repeat the argument with $V(2u)$ in place of $V(u)$ to see that $V(1) = 0$ must hold as well. $\square$

5.2 The integral transform

In this section we prove Theorem 5.1. To prove the first part, we shall assume that (5.1) holds for all $u > 0$, $u \neq 1$, where $\sigma > 1$, and $V^\star$ is a complex valued function on the vertical line $\sigma + i\mathbb{R}$ satisfying (5.2) for some $\epsilon > 0$. By applying formally the Mellin transform on both sides, we get

$$V^*(z) = \frac{1}{2\pi i} \int_{(\sigma)} V^\star(s) G_s^*(z)\, ds. \tag{5.3}$$

This step is justified by Fubini's theorem, if a sufficient uniform bound is provided for $G_s(u)$. We need to give a uniform estimate for the hypergeometric functions appearing in $G_s(u)$.

Lemma 5.1. *Let $\sigma > 1$ and $0 \leq u < 1$. Then for any $\epsilon > 0$ the following uniform bound holds on the vertical line $\sigma + i\mathbb{R}$:*

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; u\right) \ll |s|^{1/2+\epsilon}. \tag{5.4}$$

The implied constant depends only on σ and ϵ.

We postpone the proof of this lemma to the next section. It shows that the Mellin transforms $V^*(z)$ and $G_s^*(z)$ exist for $0 < \Re z < \sigma$, and that (5.3) is valid in this strip.

We can also see that V is continuous at $u \neq 1$, because the functions $G_s(u)$ are sufficiently uniformly continuous at these points. The relevant estimate reads as follows.

Lemma 5.2. *Let $\sigma > 1$ and $0 \leq v < u < 1$. Then for any $\epsilon > 0$ the following uniform bound holds on the vertical line $\sigma + i\mathbb{R}$:*

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; u\right) - F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; v\right) \ll (u - v)|s|^{3/2+\epsilon}.$$

We omit the proof of this result, as it is almost identical to that of Lemma 5.1.

It is essential that $G_s^*(z)$ can be determined explicitly. It is given by a special case of formula 2.21.1.3 from [Pr-Br-Ma]:

$$\tfrac{\Gamma(\alpha)\Gamma(a-\alpha)}{\Gamma(1-b+\alpha)\Gamma(c-\alpha)} = \tfrac{\Gamma(a)}{\Gamma(1-b)\Gamma(c)} \int_0^1 u^\alpha F(a,b;c;u) \tfrac{du}{u}$$

$$+ \tfrac{\Gamma(a)}{\Gamma(c-a)\Gamma(a-b+1)} \int_1^\infty u^{\alpha-a} F\left(a, a-c+1; a-b+1; \tfrac{1}{u}\right) \tfrac{du}{u}.$$

The formula is valid as long as $\Re(c - a - b) > -1$, $0 < \Re\alpha < \Re a$ and all the gamma values are finite on the right hand side. It can be verified formally by regarding the left hand side as a function of α and evaluating its inverse Mellin transform in u. By specializing the above formula to

$$\alpha = \tfrac{z}{2} + i\mu, \qquad a = \tfrac{s}{2} + i\mu, \qquad b = \tfrac{1}{2} + i\mu, \qquad c = \tfrac{s}{2} + \tfrac{1}{2},$$

and replacing u by u^2 in both integrals, we obtain the following result.

Lemma 5.3. *For $0 < \Re z < \sigma$ the Mellin transform of $G_s(u)$ is given by*

$$G_s^*(z) = c_{i\mu} \frac{\Gamma\left(\frac{s}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{s}{2} + i\mu\right)} \frac{\Gamma\left(\frac{s-z}{2}\right)}{\Gamma\left(\frac{s-z}{2} + \frac{1}{2} - i\mu\right)} \frac{\Gamma\left(\frac{z}{2} + i\mu\right)}{\Gamma\left(\frac{z}{2} + \frac{1}{2}\right)},$$

where $c_{i\mu}$ abbreviates the constant $\frac{1}{2}\Gamma\left(\frac{1}{2} - i\mu\right)$.

In particular, for $0 < \Re z < \sigma$, (5.3) can be rewritten as

$$\frac{\Gamma\left(\frac{z}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{z}{2} + i\mu\right)} V^*(z) = \frac{1}{2\pi i} \int_{(\sigma)} V^{\star}(s) \frac{\Gamma\left(\frac{s}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{s}{2} + i\mu\right)} \frac{c_{i\mu}\Gamma\left(\frac{s-z}{2}\right)}{\Gamma\left(\frac{s-z}{2} + \frac{1}{2} - i\mu\right)}\, ds. \tag{5.5}$$

The right hand side of this equation defines a bounded holomorphic function in every half-plane $\Re z < \sigma_0 < \sigma$, which concludes the proof of the first part of the theorem.

We turn to the second part of Theorem 5.1. Let $V : (0, \infty) \to \mathbb{C}$ be an arbitrary function which is continuous at all points $u \neq 1$. We assume that the Mellin transform $V^*(z)$ is defined for $0 < \Re z < 1$ such that $\frac{\Gamma\left(\frac{z}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{z}{2}+i\mu\right)} V^*(z)$ extends to a holomorphic function in some half-plane $\Re z < \sigma_0$ $(\sigma_0 > 1)$. If, in addition, we have a uniform bound

$$\frac{\Gamma\left(\frac{z}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{z}{2} + i\mu\right)} V^*(z) \ll (1 + |z|)^{-A}, \quad \Re z < \sigma_0 \tag{5.6}$$

for some $A > 2$, then the inverse Mellin transform of the left hand side is a continuous function $k : (0, \infty) \to \mathbb{C}$ vanishing on $(0, 1)$ such that

$$k^*(z) = \frac{\Gamma\left(\frac{z}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{z}{2} + i\mu\right)} V^*(z). \tag{5.7}$$

We claim that there is a continuous function $l : (0, \infty) \to \mathbb{C}$ vanishing on $(0, 1)$

such that (5.5) is solved by

$$V^\bigstar(s) = \frac{\Gamma\left(\frac{s}{2}+i\mu\right)}{\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)} l^*(s). \tag{5.8}$$

To see this, we also observe that for $\Re z < \sigma$

$$\frac{c_{i\mu}\Gamma\left(\frac{s-z}{2}\right)}{\Gamma\left(\frac{s-z}{2}+\frac{1}{2}-i\mu\right)} = j^*(z-s),$$

where

$$j(u) = (1-u^{-2})_+^{-1/2-i\mu}.$$

This is a special case of formula 3.251.1 from [Gr-Ry].

Notation. For $u \in \mathbb{R}$, u_+ abbreviates $\max(0,u)$.

The required identity (5.5) now reads

$$k^*(z) = \frac{1}{2\pi i}\int_{(\sigma)} j^*(z-s) l^*(s)\, ds, \quad 0 < \Re z < \sigma < \sigma_0.$$

If we also assume that

$$\int_{(\sigma)} |l^*(s)|\, ds < \infty$$

(which will be the case, cf. (5.10)), then a straightforward application of Fubini's theorem shows that the integral evaluates the Mellin transform of $j(u)l(u)$ at z. Indeed,

$$(jl)^*(z) = \int_0^\infty j(u)l(u)u^z \frac{du}{u} = \int_0^\infty j(u)\left\{\frac{1}{2\pi i}\int_{(\sigma)} l^*(s)u^{-s}\, ds\right\} u^z \frac{du}{u}$$

$$= \frac{1}{2\pi i}\int_{(\sigma)}\left\{\int_0^\infty j(u)u^{z-s}\frac{du}{u}\right\} l(s)\, ds = \frac{1}{2\pi i}\int_{(\sigma)} j^*(z-s)l^*(s)\, ds.$$

As $k(u)$ and $j(u)l(u)$ are continuous, (5.5) is now equivalent to

$$k(u) = j(u)l(u).$$

If we use the fact and assumption that both $k(u)$ and $l(u)$ vanish for $u < 1$, this becomes

$$l(u) = (1 - u^{-2})_+^{1/2+i\mu} k(u).$$

For $u > 1$ the first factor can be expanded according to the binomial theorem. The coefficients satisfy

$$\binom{\frac{1}{2} + i\mu}{j} \ll \frac{\Gamma\left(-\frac{1}{2} - i\mu + j\right)}{\Gamma(1 + j)} \ll (1 + j)^{-3/2},$$

therefore Fubini's theorem yields

$$l^*(s) = \sum_{j=0}^{\infty} \binom{\frac{1}{2} + i\mu}{j} (-1)^j k^*(s - 2j). \tag{5.9}$$

From this representation and (5.6)–(5.7) we can easily infer the bound

$$l^*(s) \ll_{\sigma,A} \big(1 + |s|\big)^{-A}, \quad 0 < \Re s < \sigma_0. \tag{5.10}$$

We found the recipe to construct a function $V^\star(s)$ satisfying (5.5). First we determine k^* according to (5.7). Then we define l^* by (5.9). Finally, $V^\star(s)$ is given by (5.8). By our assumptions on $V^*(z)$, it is clear that $V^\star(s)$ is analytic in $0 < \Re s < \sigma_0$, and from (5.10) it also follows that in this region it satisfies a uniform upper bound

$$V^\star(s) \ll_{\sigma,A} \big(1 + |s|\big)^{-A-1/2}.$$

Finally, (5.5) implies (5.1) for all points $u > 0$, $u \neq 1$, because V is continuous at all these points by assumption. This completes the proof of the theorem.

5.3 Bounds for hypergeometric functions

In this section we prove Lemma 5.1. For $0 \le u < \frac{1}{2}$ we use the representation

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; u\right) = \frac{\Gamma\left(\frac{s}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{1}{2} + i\mu\right)\Gamma\left(\frac{s}{2} - i\mu\right)} \int_0^1 t^{-\frac{1}{2}+i\mu}(1-t)^{\frac{s}{2}-i\mu-1}(1-tu)^{-\frac{s}{2}-i\mu}\, dt.$$

This identity is a special case of formula 9.111 from [Gr-Ry]. It follows that

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; u\right) \ll_\sigma \left|\frac{\Gamma\left(\frac{s}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{s}{2} - i\mu\right)}\right| \int_0^1 t^{-\frac{1}{2}}(1-t)^{-\frac{1}{2}}\, dt.$$

The integral on the right hand side is bounded, hence we have

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; u\right) \ll_\sigma |s|^{1/2}.$$

For the rest of this section we shall assume that $\frac{1}{2} \le u < 1$. We apply formula 9.131.1 from [Gr-Ry]:

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{1}{2} + i\mu; \tfrac{s}{2} + \tfrac{1}{2}; u\right) = (1-u)^{-\frac{s}{2}-i\mu} F\left(\tfrac{s}{2} + i\mu, \tfrac{s}{2} - i\mu; \tfrac{s}{2} + \tfrac{1}{2}; \tfrac{u}{u-1}\right). \tag{5.11}$$

Note that here $\frac{u}{u-1} \le -1$. We can express the hypergeometric function on the right hand side as a contour integral by formula 9.113 from [Gr-Ry]:

$$F\left(\tfrac{s}{2} + i\mu, \tfrac{s}{2} - i\mu; \tfrac{s}{2} + \tfrac{1}{2}; \tfrac{u}{u-1}\right) = \frac{1}{2\pi i}\int_{(-\epsilon)} \frac{\Gamma\left(\frac{s}{2} + i\mu + w\right)\Gamma\left(\frac{s}{2} - i\mu + w\right)\Gamma\left(\frac{s}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{s}{2} + i\mu\right)\Gamma\left(\frac{s}{2} - i\mu\right)\Gamma\left(\frac{s}{2} + \frac{1}{2} + w\right)}\Gamma(-w)\left(\frac{u}{1-u}\right)^w dw.$$

This formula is valid whenever

$$0 < \epsilon < \frac{\sigma}{2}.$$

In order to estimate the integral efficiently, we shift the contour to the line $\Re s = -\frac{\sigma}{2} - \epsilon$. This shift picks up the poles at

$$w = -\frac{s}{2} \pm i\mu.$$

To be precise, these are two simple poles when $\mu \neq 0$, and a double pole when $\mu = 0$. In both cases we can write the result as

$$\begin{aligned} F\left(\tfrac{s}{2}+i\mu, \tfrac{s}{2}-i\mu; \tfrac{s}{2}+\tfrac{1}{2}; \tfrac{u}{u-1}\right) &= \sum_{\pm} d_{\pm i\mu} \frac{\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{s}{2} \mp i\mu\right)} \left(\frac{u}{1-u}\right)^{-\frac{s}{2}\mp i\mu} \\ &+ \frac{1}{2\pi i} \int_{\left(-\frac{\sigma}{2}-\epsilon\right)} \frac{\Gamma\left(\frac{s}{2}+i\mu+w\right)\Gamma\left(\frac{s}{2}-i\mu+w\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{s}{2}+i\mu\right)\Gamma\left(\frac{s}{2}-i\mu\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}+w\right)} \Gamma(-w) \left(\frac{u}{1-u}\right)^{w} dw, \end{aligned}$$

where $d_{i\mu}$ and $d_{-i\mu}$ are suitable constants. It follows from (5.11) that

$$\begin{aligned} F\left(\tfrac{s}{2}+i\mu, \tfrac{1}{2}+i\mu; \tfrac{s}{2}+\tfrac{1}{2}; u\right) &\ll_\sigma |s|^{1/2} \\ &+ \int_{\left(-\frac{\sigma}{2}-\epsilon\right)} \left| \frac{\Gamma\left(\frac{s}{2}+i\mu+w\right)\Gamma\left(\frac{s}{2}-i\mu+w\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma\left(\frac{s}{2}+i\mu\right)\Gamma\left(\frac{s}{2}-i\mu\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}+w\right)} \Gamma(-w)\, dw \right|. \end{aligned} \tag{5.12}$$

It remains to estimate the last integral. In the light of the uniform estimate

$$\left| \frac{\Gamma\left(\frac{s}{2}+i\mu+w\right)\Gamma\left(\frac{s}{2}-i\mu+w\right)}{\Gamma\left(\frac{s}{2}+i\mu\right)\Gamma\left(\frac{s}{2}-i\mu\right)} \right| \ll_{\sigma,\epsilon} \left| \frac{\Gamma^2\left(\frac{s}{2}+w\right)}{\Gamma^2\left(\frac{s}{2}\right)} \right|, \quad \Re w = -\tfrac{\sigma}{2} - \epsilon,$$

we are left with estimating

$$\mathfrak{I} = \int_{\left(-\frac{\sigma}{2}-\epsilon\right)} \left| \frac{\Gamma^2\left(\frac{s}{2}+w\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma^2\left(\frac{s}{2}\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}+w\right)} \Gamma(-w)\, dw \right|. \tag{5.13}$$

The value of the integral does not change when s is replaced by $\bar{s}$, therefore we can assume that $\Im s > 0$. We split the integral into three parts.

Part 1. $\Im w > 0$. In this segment Stirling's formula implies

$$\frac{\Gamma^2\left(\frac{s}{2}+w\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma^2\left(\frac{s}{2}\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}+w\right)}\Gamma(-w) \ll_{\sigma,\epsilon} e^{-\pi\Im w}|w|^{\frac{\sigma}{2}-\frac{1}{2}+\epsilon}|s|^{1-\frac{\sigma}{2}}\left|\frac{s}{2}+w\right|^{-1-\epsilon}$$

$$\ll_{\sigma,\epsilon} |s|^{1-\frac{\sigma}{2}}\left|\frac{s}{2}+w\right|^{-1-\epsilon}.$$

It follows that the total contribution to the integral (5.13) is

$$\mathfrak{I}_1 \ll_{\sigma,\epsilon} |s|^{1-\frac{\sigma}{2}}|s|^{-\epsilon} \ll_{\sigma,\epsilon} |s|^{\frac{1}{2}}.$$

Part 2. $0 \geq \Im w \geq -\Im\frac{s}{2}$. In this segment Stirling's formula implies

$$\frac{\Gamma^2\left(\frac{s}{2}+w\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma^2\left(\frac{s}{2}\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}+w\right)}\Gamma(-w) \ll_{\sigma,\epsilon} |w|^{\frac{\sigma}{2}-\frac{1}{2}+\epsilon}|s|^{1-\frac{\sigma}{2}}\left|\frac{s}{2}+w\right|^{-1-\epsilon}$$

$$\ll_{\sigma,\epsilon} |s|^{\frac{1}{2}+\epsilon}\left|\frac{s}{2}+w\right|^{-1-\epsilon}.$$

It follows that the total contribution to the integral (5.13) is

$$\mathfrak{I}_2 \ll_{\sigma,\epsilon} |s|^{\frac{1}{2}+\epsilon}.$$

Part 3. $-\Im\frac{s}{2} > \Im w$. In this segment Stirling's formula implies

$$\frac{\Gamma^2\left(\frac{s}{2}+w\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}\right)}{\Gamma^2\left(\frac{s}{2}\right)\Gamma\left(\frac{s}{2}+\frac{1}{2}+w\right)}\Gamma(-w) \ll_{\sigma,\epsilon} e^{\pi\Im\left(\frac{s}{2}+w\right)}|w|^{\frac{\sigma}{2}-\frac{1}{2}+\epsilon}|s|^{1-\frac{\sigma}{2}}\left|\frac{s}{2}+w\right|^{-1-\epsilon}$$

$$\ll_{\sigma,\epsilon} |s|^{\frac{1}{2}+\epsilon}e^{\pi\Im\left(\frac{s}{2}+w\right)}\left|\frac{s}{2}+w\right|^{\frac{\sigma}{2}-\frac{3}{2}}.$$

It follows that the total contribution to the integral (5.13) is

$$\mathfrak{I}_3 \ll_{\sigma,\epsilon} |s|^{\frac{1}{2}+\epsilon}.$$

Altogether we can see that

$$\mathfrak{I} = \mathfrak{I}_1 + \mathfrak{I}_2 + \mathfrak{I}_3 \ll_{\sigma,\epsilon} |s|^{1/2+\epsilon},$$

therefore (5.12) and (5.13) imply the required bound (5.4).

The proof of Lemma 5.1 is complete. $\square$

Bibliography

[At-Le] A. O. L. Atkin, J. Lehner, Hecke operators on $\Gamma_0(m)$, *Math. Ann.* **185** (1970), 134–160.

[At-Li] A. O. L. Atkin, W.-C.W. Li, Twists of newsforms and pseudo-eigenvalues of W-operators, *Invent. Math.* **48** (1978), 221–243.

[Be-Re] J. Bernstein, A. Reznikov, Analytic continuation of representations and estimates of automorphic forms, *Ann. of Math.* **150** (1999), 329–352.

[Ch-Na] K. Chandrasekharan, R. Narasimhan, The approximate functional equation for a class of zeta-functions, *Math. Ann.* **152** (1963), 30–64.

[Co-PS-Sa] J. W. Cogdell, I. I. Piatetski-Shapiro, P. Sarnak, Estimates on the critical line for Hilbert modular L-functions and applications, in preparation

[Du] W. Duke, Hyperbolic distribution problems and half-integral weight Maass forms, *Invent. Math.* **92** (1988), 73–90.

[Du-Fr-Iw1] W. Duke, J. B. Friedlander, H. Iwaniec, Bounds for automorphic L-functions, *Invent. Math.* **112** (1993), 1–8.

[Du-Fr-Iw2] W. Duke, J. B. Friedlander, H. Iwaniec, A quadratic divisor problem, *Invent. Math.* **115** (1994), 209–217.

[Du-Fr-Iw3] W. Duke, J. B. Friedlander, H. Iwaniec, The subconvexity problem for Artin L-functions, *Invent. Math.* **149** (2002), 489–577.

[Du-Iw] W. Duke, H. Iwaniec, Bilinear forms in the Fourier coefficients of half-integral weight cusp forms and sums over primes, *Math. Ann.* **286** (1990), 783–802.

[Du-SP] W. Duke, R. Schulze-Pillot, Representations of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids, *Invent. Math.* **99** (1990), 49–57.

[Er] A. Erdélyi et al., Tables of integral transforms, Vol. I. [based on notes left by H. Bateman], McGraw-Hill, New York, 1954

[Es] T. Estermann, On Kloosterman's sum, *Mathematika* **8** (1961), 83–86.

[Fr-Iw] J. B. Friedlander, H. Iwaniec, A mean-value theorem for character sums, *Michigan Math. J.* **39** (1992), 153–159.

[Gr-Ry] I. S. Gradshteyn, I. M. Ryzhik, Tables of integrals, series, and products [corrected and enlarged by A. Jeffrey, incorporating the 4th edition edited by Yu. V. Geronimus and M. Yu. Tseytlin, translated from the Russian], Academic Press, New York, 1980

[Ha1] G. Harcos, Uniform approximate functional equation for principal L-functions, *Int. Math. Res. Not.* **2002**, 923–932.

[Ha2] G. Harcos, An additive problem in the Fourier coefficients of Maass forms, *Math. Ann.* **326** (2003), 347–365.

[Har-Lit] G. H. Hardy, J. E. Littlewood, The zeros of Riemann's zeta-function on the critical line, *Math. Z.* **10** (1921), 283–317.

[Ho-Lo] J. Hoffstein, P. Lockhart, Coefficients of Maass forms and the Siegel zero, With an appendix by D. Goldfeld, J. Hoffstein and D. Lieman, *Ann. of Math.* **140** (1994), 161–181.

[Iv] A. Ivić, An approximate functional equation for a class of Dirichlet series, *J. Anal.* **3** (1995), 241–252.

[Iw1] H. Iwaniec, Introduction to the spectral theory of automorphic forms, *Biblioteca de la Revista Matemática Iberoamericana*, Revista Matemática Iberoamericana, Madrid, 1995

[Iw2] H. Iwaniec, Small eigenvalues of Laplacian for $\Gamma_0(N)$, *Acta Arith.* **56** (1990), 65–82.

[Iw-Sa] H. Iwaniec, P. Sarnak, Perspectives on the analytic theory of L-functions, *Geom. Funct. Anal. Special Volume* (2000), 705–741.

[Ja] E. Jacobsthal, Über Sequenzen ganzer Zahlen, von denen keine zu n teilerfremd ist. I–III. (German), *Norke Vid. Selsk. Forh. Trondheim* **33** (1961), 117–139.

[Ju1] M. Jutila, Transformations of exponential sums, Proceedings of the Amalfi Conference on Analytic Number Theory (Maiori 1989), Univ. Salerno, Salerno, 1992, 263–270.

[Ju2] M. Jutila, The additive divisor problem and its analogs for Fourier coefficients of cusp forms. I., *Math. Z.* **223** (1996), 435-461.; II., ibid **225** (1997), 625–637.

[Ju3] M. Jutila, Convolutions of Fourier coefficients of cusp forms, *Publ. Inst. Math. (Beograd) (N.S.)* **65(79)** (1999), 31–51.

[Ju4] M. Jutila, On exponential sums involving the divisor function, *J. Reine Angew. Math.* **355** (1985), 173–190.

[Ju5] M. Jutila, A method in the theory of exponential sums, *Tata Lect. Notes Math.* **80**, Bombay, 1987

[Ki] H. H. Kim, Functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2 [with appendix 1 by D. Ramakrishnan and appendix 2 by H. H. Kim and P. Sarnak], *J. Amer. Math. Soc.* **16** (2003), 139–183.

[Kl] H. D. Kloosterman, On the representation of numbers in the form $ax^2 + by^2 + cz^2 + dt^2$, *Acta Math.* **49** (1926), 407–464.

[Koh] W. Kohnen, Fourier coefficients of modular forms of half-integral weight, *Math. Ann.* **271** (1985), 237–268.

[Ko-Mi-Va] E. Kowalski, P. Michel, J. VanderKam, Rankin–Selberg L-functions in the level aspect, *Duke Math. J.* **114** (2002), 123–191.

[La] A. F. Lavrik, Approximate functional equations of Dirichlet functions (Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* **32** (1968), 134–185.

[Li] W.-C.W. Li, Newforms and functional equations, *Math. Ann.* **212** (1975), 285–315.

[Lu-Ru-Sa] W. Luo, Z. Rudnick, P. Sarnak, On the generalized Ramanujan conjecture for $\mathrm{GL}(n)$, Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996), *Proc. Sympos. Pure Math.* **66(2)**, Amer. Math. Soc., Rhode Island, 1999, 301–310.

[Me] T. Meurman, On exponential sums involving the Fourier coefficients of Maass wave forms, *J. Reine Angew. Math.* **384** (1988), 192–207.

[Mi1] P. Michel, Complement to "Rankin–Selberg L-functions in the level aspect", unpublished notes (2000)

[Mi2] P. Michel, The subconvexity problem for Rankin–Selberg L-functions and equidistribution of Heegner points, *Ann. of Math.* **160** (2004), 185–236.

[Mol] G. Molteni, Upper and lower bounds at $s = 1$ for certain Dirichlet series with Euler product, *Duke Math. J.* **111** (2002), 133–158.

[Mot] Y. Motohashi, The binary additive divisor problem, *Ann. Sci. École Norm. Sup.* **27** (1994), 529–572.

[Ol] F. W. J. Olver, Asymptotics and special functions, *Computer Science and Applied Mathematics*, Academic Press, New York, 1974

[Pe] Y. Petridis, On squares of eigenfuncions for the hyperbolic plane and a new bound on certain L-series, *Int. Math. Res. Not.* **1995**, 111–127.

[Pe-Sa] Y. Petridis, P. Sarnak, Quantum unique ergodicity for $SL_2(\mathcal{O})\backslash\mathbf{H}^3$ and estimates for L-functions, *J. Evol. Equ.* **1** (2001), 277–290.

[Pr-Br-Ma] A. P. Prudnikov, Yu. A. Brychkov, O. I. Marichev, Integrals and series, Vol. 3 (More special functions), Gordon and Breach Science Publishers, New York, 1986

[Sa1] P. Sarnak, Integrals of products of eigenfunctions, *Int. Math. Res. Not.* **1994**, 251–260.

[Sa2] P. Sarnak, Estimates for Rankin–Selberg L-functions and quantum unique ergodicity, *J. Funct. Anal.* **184** (2001), 419–453.

[Se] A. Selberg, On the estimation of Fourier coefficients of modular forms, *Proc. Symp. Pure Math.* **8**, Amer. Math. Soc., Providence, R. I., 1965, 1–15.

[Sh] G. Shimura, On the Fourier coefficients of Hilbert modular forms of half-integral weight, *Duke Math. J.* **71** (1993), 501–557.

[St] T. Stefanicki, Non-vanishing of L-functions attached to automorphic representations of $GL(2)$ over Q, *J. Reine Angew. Math.* **474** (1996), 1–24.

[Wal] J.-L. Waldspurger, Sur les coefficients de Fourier des formes modulaires de poids demi-entier, *Math. Pures Appl.* **60** (1981), 375–484.

[Wat] G. N. Watson, A treatise on the theory of Bessel functions, Cambridge University Press, Cambridge, 1944

[We] A. Weil, On some exponential sums, *Proc. Nat. Acad. Sci. U. S. A.* **34** (1948), 204–207.