Asbjørn Christian Nordentoft

On arithmetic statistics and periods of automorphic forms

PhD thesis • Department of Mathematical Sciences • University of Copenhagen

Asbjørn Christian Nordentoft

Department of Mathematical Sciences, University of Copenhagen
Universitetsparken 5, 2100 København Ø, Denmark.

acnordentoft@outlook.com

This PhD thesis has been submitted to the PhD School of the Faculty of Science, University of Copenhagen, on September 30, 2020.

Adviser:	Morten Risager (University of Copenhagen)
Assessment committee:	Valentin Blomer (University of Bonn)
	Zeev Rudnick (Tel-Aviv University)
	Jasmin Matz (University of Copenhagen)

ISBN: 978-87-7125-037-4

Abstract

This thesis consists of seven papers contributing to the understanding of periods of automorphic forms and their connections to questions in arithmetic statistics.

In Paper A, we obtain a normal distribution result for central values of additive twists of L-functions of holomorphic cusp forms. This is a generalization to higher weight of a recent breakthrough due to Petridis and Risager, resolving a (-n average version of a) conjecture due to Mazur and Rubin. We furthermore present applications to certain "wide" families of automorphic L-functions.

In Paper B, we study the distribution of period polynomials associated to a fixed holomorphic cusp form. We determine the asymptotic joint distribution of the coefficients and obtain an asymptotic expression for the zeroes of the polynomials. This complements recent work of Jin, Ma, Ono and Soundararajan (and others).

In Paper C, we prove that additive twists associated to holomorphic cusp forms (with general level) define a quantum modular form in the sense of Zagier. We use this to obtain a reciprocity formula for a certain twisted first moment of L-functions, similar to reciprocity relations obtained by Conrey.

In Paper D, which is joint with Petru Constantinescu, we introduce an automorphic method for studying the residual distribution of modular symbols. We obtained a refinement of a result due to Lee and Sun (which resolved an average version of a conjecture of Mazur and Rubin), and furthermore generalize the results to quotients of general hyperbolic spaces. Finally, we resolve the conjecture of Mazur and Rubin in some very special cases using algebraic methods.

In Paper E, which is joint with Peter Humphries, we study sparse equidistribution of certain hyperbolic orbifolds associated to real quadratic fields introduced by Duke, Imamoḡlu and Tóth. Our main insight is that the Weyl sums that appear in the distribution problem can be related to automorphic periods, which in turn by work of Martin and Whitehouse can be related to central values of Rankin–Selberg L-functions.

In Paper F, we obtain a uniform sup norm bound for Eisenstein series using exponential sum methods, improving on a result due to Blomer. We use this to obtain a hybrid subconvexity bound for class group L-functions.

In Paper G, which is joint with Yiannis Petridis and Morten Risager, we study the mass distribution of holomorphic cusp forms on shrinking regions around infinity. In particular, we obtain an asymptotic formula for the quantum variance (extending results due to Luo and Sarnak), which exhibits a phase transition.

Resumé

Denne afhandling består af syv artikler, der bidrager til forståelsen af automorfe perioder og deres forbindelse til spørgsmål i aritmetisk statistik.

I Artikel A viser vi et normalfordelings-resultat for central værdier af additive twist af L-funktioner for holomorfe spidsformer. Dette er en generalisering af et nyligt gennembrud af Petridis og Risager, som løste (en gennemsnitlig version) af en formodning af Mazur og Rubin. Ydermere viser vi anvendelser til visse "brede" familier af automorfe L-funktioner.

I Artikel B undersøger vi fordelingen af periodepolynomierne tilknyttet til en fastholdt holomorf spidsform. Vi bestemmer den samlede asymptotiske fordelingen af koefficienterne og opnår et asymptotisk udtryk for nul-punkterne for polynomierne. Dette komplementerer nylige resultater af Jin, Ma, Ono og Soundararajan (blandt andre).

I Artikel C beviser vi at additive twist hørende til holomorfe spidsformer (med vilkårligt niveau) definerer en kvante-modulform i Zagiers terminologi. Vi benytter dette til at vise en reciprocitetsrelation for visse twistede første-momenter af L-funktioner, der minder om reciprocitetsrelationer opnået af Conrey.

I Artikel D, som er skrevet i samarbejde med Petru Constantinescu, introducerer vi en automorf metode til at studere den residuale fordeling af modulære symboler. Vi opnår en raffinering af et resultat af Lee og Sun (som løste en gennemsnitlig version af en formodning af Mazur og Rubin). Ydermere generaliserer vores resultater til generalle hyperbolske rum. Endeligt løser vi den fulde formodning af Mazur og Rubin i nogle specialtilfælde ved at bruge algebraiske metoder.

I Artikel E, som er skrevet i samarbejde med Peter Humphries, undersøger vi udtyndet ligefordeling for visse hyperbolske orbifolde associeret til reelle kvadratiske legemer, som er blevet defineret af Duke, Imamoḡlu og Tóth. Vores hovedindsigt er at Weyl-summerne som optræder i fordelingsproblemet kan relateres til automorfe perioder, som så igen kan relateres til Rankin–Selberg L-funktioner ved at bruge resultater af Martin of Whitehouse.

I Artikel F opnår vi uniforme sup-norms-begrænsninger for Eisenstein-rækker ved at bruge metoder i teorien for eksponentielle summer og forbedrer derved et resultat af Blomer. Vi anvender dette til at opnå en hybrid-sub-konveksitetsbegrænsning for klassegruppe-L-funktioner.

I Artikel G, som er skrevet i samarbejde med Yiannis Petridis og Morten Risager, undersøger vi masse-fordeling af holomorfe spidsformer på skrumpende områder omkring uendelig. Vi opnår blandt andet en asymptotisk formel for kvante-variansen (hvilket udbygger et resultat af Luo og Sarnak), som udviser en faseovergang.

Acknowledgements

First and foremost, I would like to thank Morten Risager. Ever since I took his *analytic number theory*-course, there has always been time for long and enlightening discussions. His way of approaching mathematics has learned me so much and I am very grateful for having had him as my advisor.

I would also like to express by gratitude to Dorian Goldfeld for his hospitality during my extended stay at Columbia University. This year was very formative for me, and I would like to thank him for his generosity and warm-heartedness.

It has been a pleasure to collaborate with Petru Constantinescu, Peter Humphries and Yiannis Petridis (as well as Morten Risager). You have all taught me many valuable lessons. Thank you for that.

Many other colleagues have played an important role in the past four years, in particular, Anders, Boris, Nicholas, Chung-Hang and my office mates Kaif, Riccardo, Malte, Robin and Philipp, all of whom I would like give special thanks.

I would also like to thank the non-mathematicians Sara, Oscar and Pernille for providing an outsiders perspective on my mathematical ideas.

Finally, I would like to thank the ones closest to me.

Asbjørn Christian Nordentoft
Copenhagen, September 2020

Contents

Part I

Introduction

CHAPTER 1

A GENTLE INTRODUCTION TO ARITHMETIC STATISTICS AND PERIODS

This thesis consists of the following seven papers within the field of analytic number theory.

Paper A: *Central values of additive twists of cuspidal L-functions* (submitted)

Paper B: *On the distribution of periods of holomorphic cusp forms and zeroes of period polynomials* (published in International Mathematics Research Notices)

Paper C: *A note on additive twists, reciprocity laws and quantum modular forms* (published in The Ramanujan Journal)

Paper D: *Residual equidistribution of modular symbols and cohomology classes for quotients of hyperbolic n-space* (preprint, joint with Petru Constantinescu)

Paper E: *Sparse equidistribution of hyperbolic orbifolds* (joint with Peter Humphries)

Paper F: *Hybrid subconvexity for class group L-functions and uniform sup norm bounds of Eisenstein series* (published in Forum Mathematicum)

Paper G: *Small scale equidistribution of Hecke eigenforms at infinity* (preprint, joint with Yiannis Petridis and Morten Risager)

The methods employed in the papers range from exponential sums, classical automorphic forms, spectral theory, to representation theory. Judging from the titles, it should be clear that the problems also range over a variety of subjects. There is however one common feature; they all touch upon questions related to *arithmetic statistics*. This should be interpreted in the widest possible sense to include distribution properties, extremal behavior and symmetry properties of arithmetic objects. If one takes a closer look at the approaches employed in the papers, one will notice that despite the differences in the techniques applied there is one common theme luring in the background. This is the notion of *periods of automorphic forms*. Since the birth of the theory of automorphic forms, periods of automorphic forms have played an important role and continue to be a central theme to this day. In the first part of this thesis, we give some general background on arithmetic statistics and automorphic forms with a special emphasis on periods. This focus will hopefully tie the seven papers together and shed some new light on some of the topics dealt with.

We will begin by giving a pictorial introduction to arithmetic statistics using a variety of examples from outside of mathematics. Then we will go on to explain a "toy example" of periods of automorphic forms. Both of these sections are intended for a general, non-mathematical audience and will be delivered in a rather informal style. After this, we will give a short account of the history of automorphic forms with a special emphasis on the role of periods. The exposition will be in the style of a survey. In particular, we will not give any of the required background but instead refer to some of the many excellent sources that already exist. This will lead to a more detailed discussion of the problems dealt with in this thesis. Finally, we will give summaries of the seven papers, which can be found in the second part of this thesis.

1.1 Statistics in number theory, and beyond

The overall topic of this thesis is *arithmetic statistics*. This is the field concerned with the statistical and distributional properties of arithmetic objects. We will begin by considering a very accessible example; the last digit of prime numbers. Recall that a prime number is a natural number (greater than 1) which is only divisible by 1 and the number itself. The list of primes begins as follows:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, \ldots$$

We notice that except for the primes 2 and 5 all primes end with one of the digits $1, 3, 7$ or 9 in their decimal representation. One can now wonder how the primes distribute among these digits. The following shows how the first 10.000 primes distribute among the four digits:

$$\mathbf{1} : 24.84\%, \quad \mathbf{3} : 25.15\%, \quad \mathbf{7} : 25.08\%, \quad \mathbf{9} : 24.91\%,$$

It looks like the primes distribute quite evenly among the four possible last digits! More precisely, we suspect that the primes *equidistribute* in the sense that if we count more and more primes, then the percentages as above should come closer and closer to 25% (formalized through the mathematical notion of *convergence*). Note that computations such as the above do *not* bring us any closer to concluding that the primes equidistribute; however many primes we check, it might just be that if we go a little further, then from some point on all primes have the same last digit, say. In this case, we however know that this cannot happen, since our suspicion has been proved mathematically! The result is known as the *Prime Number Theorem for Arithmetic Progressions* and was proved over a hundred years ago following the work of Dirichlet, Hadamard and de la Vallée Poussin. This is an example of a result in arithmetic statistics. It is useful to think of each prime number as representing a certain "experiment" and the last digit as being the "outcome" of this experiment. In this case the outcomes are limited to 4 options (excluding the primes 2 and 5). A

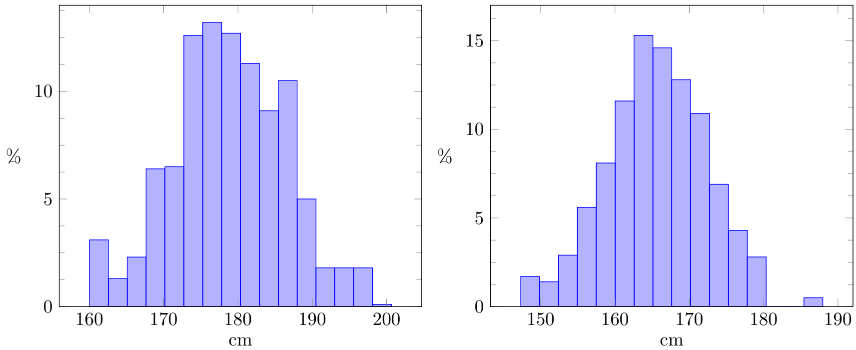

Figure 1.1: Histogram of heights of American males, resp. females age 30-39 (Data taken from U.S. National Center for Health Statistics, unpublished data, `https://www2.census.gov/library/publications/2010/compendia/statab/130ed/tables/11s0205.pdf`)

similar setting is encountered when looking at the biological sex of newborns, left and right handedness and many other examples from everyday life. It is an interesting exercise to consider, which kind of distribution one would expect in these more familiar cases.

This set-up is closely related to the subject of Paper D. As is the case with the last digit of primes, the main result in Paper D is that certain "experiments" equidistribute among a finite number of possible outcomes.

There are also many familiar situations where the outcome is *not* limited to a finite number of possibilities. Take for instance the sizes of snowflakes, heights of a population or the velocities of particles in a hot gas. Data for the two last examples are shown in Figures 1.1 and 1.2.

We notice something interesting; the data sets in these two examples are both following approximately the same "bell"-shape (up to translation and dilation). This shape arises all over the place in the natural and social sciences and is known as the *normal distribution* (or *Gauß curve*). It might seem like a miracle that this particular shape shows its face in such a variety of situations. Mathematics offers some kind of explanation to this phenomenon via the *Central Limit Theorem*. In loose terms this mathematical principle says that if your experiment depends on a large number of independent variables (or factors), then the outcome should be "bell"-shaped when plotted (although the validity of this explanation can be questioned [66]). This perspective is especially relevant in the case of Figure 1.1.

The normal distribution can also be derived from a set of simple axioms (or assumptions) as is explained beautifully in the introduction of [15]. This can for instance be used to derive the theoretical distribution underlying Figure 1.2, known as the *Maxwell–Boltzman distribution*. These axioms were referred to by the Greek-French composer Iannis Xenakis as "*one of the "logical poems", which the human intelligence creates in order to trap the superficial incoherencies of physical phenomena*", [101, page 13]. The normal distribution plays a central role in his theory of

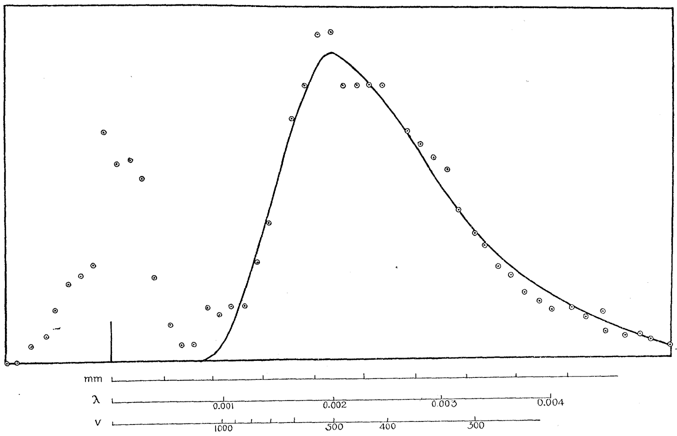

Figure 1.2: Velocity spectrum of a metallic vapor (Reprinted figure with permission from John A. Eldridge, Phys. Rev., vol. 30, p. 933, 1927. Copyright 2020 by the American Physical Society, `https://link.aps.org/doi/10.1103/PhysRev.30.931`)

music, which he called *Stochastic Music*. Xenakis was the first to introduce randomness and statistical methods into classical music. Using the randomness that mathematics could provide he found a profound and novel kind of beauty prior unknown to mankind. In the orchestral work *Pithoprakta* (meaning "action through probability" in Greek), Xenakis aims to capture the sensation of a roaring crowd and the sound of the wind in the trees. To approach this he uses an analogy between the movement of molecules in a gas and the movement of sounds through a pitch range. Because the molecules in a gas behave according to the Maxwell–Boltzman distribution, this lead Xenakis to create the *glissandi* of the strings using the normal distribution. A graphic plot by the composer of the glissandi can be found in Figure 1.3, which when listened to possess a supreme beauty.

All of the above examples should be a testament to the wide applicability and frequent occurrence of the normal distribution in all aspects of human life. The normal distribution also arises many places in pure mathematics (whether this is to be considered as part of "human life", we will leave for the reader to decide). In number theory in particular, one could mention Selberg's work on the distribution of the zeta function (Selberg never published his results, see instead [53]) or the work of Erdös and Kac [29] on the distribution of the number of prime divisors of integers. In this thesis, we encounter at least two different settings where the normal distribution arises. One is in Paper A where we study the distribution of certain numbers called *modular symbols* connected to the "doughnut"-shaped *elliptic curves*. The data that one gets in this case is shown in Figure 1.4 and we immediately recognize the shape! The other occurrence of the normal distribution in this thesis is in the Papers F and G, which deal with arithmetic aspects of *quantum chaos*. In this case it has been predicted by physicist that quantum particles (for classically ergodic systems) should behave like "random waves" when the energy is large. The randomness in this setting

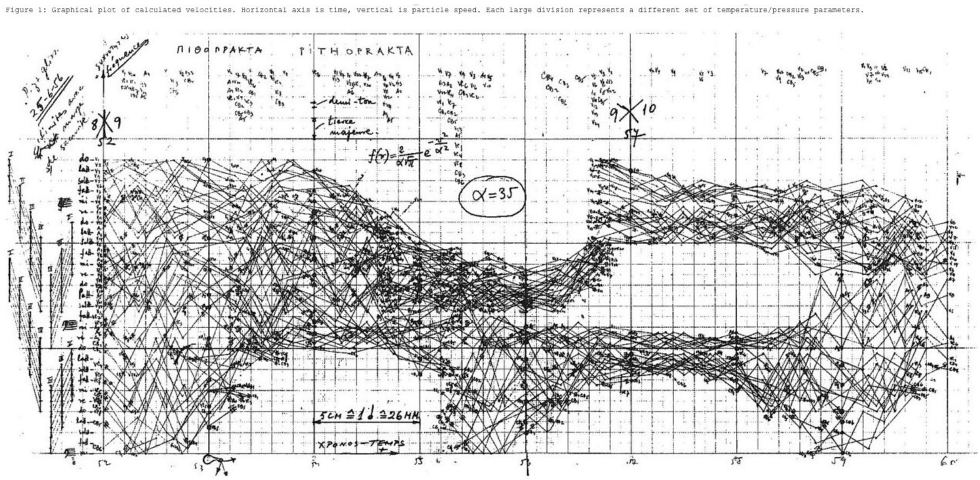

Figure 1.3: Graphic plot of the *glissandi* in *Pithoprakta* by the composer (Taken from `https://music7703lsu.wordpress.com/2017/04/02/pithoprakta-by-iannis-xenakis/`)

is exactly predicted to be that which comes from the normal distribution (see [5] for details).

Obviously, one might encounter distribution which are not "bell"-shaped nor limited to a finite number of outcomes. A very relevant example being the contagion graphs for a pandemic in a population with no immunity (nor social distancing), where the distribution is expected to follow a *logistical curve.* We will also encounter other kinds of distributions in Papers B and E.

1.1.1 Reduced fractions and Gauß points

This thesis is concerned with a number of problems in arithmetic statistics, which all take place on the so-called *modular curve.* This is a very important mathematical object, which shape reminds one of a parsnip as illustrated in Figure 1.5. We will study the distribution of a variety of objects that "live" on the surface of the modular curve and their connections to some very intricate objects called *periods of automorphic forms.*

In order to get a feeling for the flavor of the problems dealt with in this thesis, we will now go into the details of one specific example. The ideas surrounding periods of automorphic forms seem hard to explain in layman terms, but we will try to illustrate the underlying philosophy in a quite simple setting. The point that we would like to emphasise is that although the distribution problems that we care about are often very hard to get a grip on, they are connected to automorphic periods, which in many cases are easier to handle. This then has implications for the distribution questions that we wanted to understand in the first place (in some cases we are able to solve them completely!). We admit that this is very abstract and hard to grasp, but the overall philosophy can be summarized as follows: *we cannot understand a*

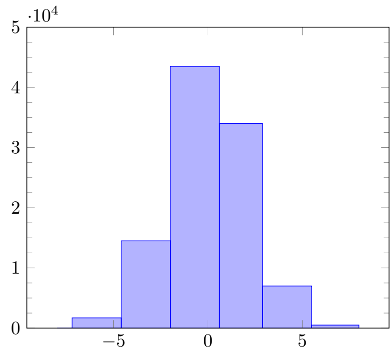

Figure 1.4: Histogram of modular symbols for the elliptic curve $E = 11A1$ and denominator 100003 (Created using unpublished data by Mazur and Rubin [72, page 34])

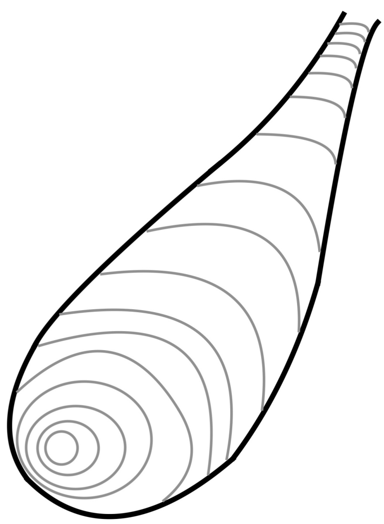

Figure 1.5: The “parsnip”-shaped modular curve

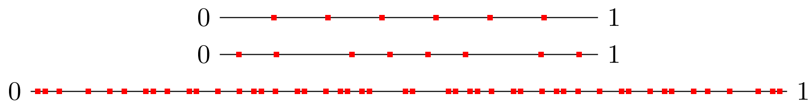

Figure 1.6: Reduced fractions with denominator resp. 7, 20 and 105

certain problem directly, so we relate to something else, which we understand better. In explaining the example, we will only make use of elementary arithmetics, simple manipulations of vectors in the (Euclidean) plane and a tiny bit of clock (or modular) arithmetics. Hopefully with a background in high school mathematics and an open mind, one can get something valuable out of what follows. For a more concise version of this "toy example", consult the introduction of Paper B.

The thing that will serve as a simple model for automorphic periods are *reduced fractions*. These are fractions of the form $\frac{n}{m}$, where n, m are natural numbers, which do not share a common factor other than 1. So for instance $\frac{2}{3}$ and $\frac{5}{7}$ are reduced fractions, whereas $\frac{3}{6}$ is not (since 3 and 6 share the factor 3). We will consider reduced fractions which lie in the interval between 0 and 1 with fixed denominator, and study how they distribute on this interval. For example when the denominator is respectively $7, 20$ and 105, we get pictures as shown in Figure 1.6. It looks like the reduced fractions distribute quite evenly on the interval! To test this in a different way, we can take some fixed subinterval and count the percentage of points that land within this interval. If we for instance consider the interval from $\frac{1}{4}$ to $\frac{1}{2}$, then for denominator 123798 the percentage is 24.995% and for denominator 524234242 one gets 25.000%. We observe that the percentages are very close to 25%, which is exactly the size of the interval from $1/4$ to $1/2$ relative to the whole interval. In fact, it can be proved (quite easily) mathematically that if you fix some interval between 0 and 1, then the percentage of reduced fractions with a fixed denominator lying in this interval, will tend to the size of the interval as the denominator grows. This phenomenon we describe by saying that reduced fractions *equidistribute* on the interval from 0 to 1.

What we will draw from the above is that reduced fractions are quite nicely behaved and relatively easy to understand. Below we will describe an intricate "recipe", which from reduced fractions will produce some new, very interesting points called *Gauß points*. In order to define these, we have to go to a slightly more complicated setup; we move from the line to the plane. The complexity of the Gauß points has its root in a complicated interaction between multiplication and addition. It will undoubtably be hard to see the bigger picture when following this "recipe", but in the end a very nice picture will arise. So hold tight!

For now we will focus on reduced fractions with denominator 7. First of all we consider powers of the number 3 and the residues that one gets when performing division by 7:

$$3^1 = 3 = 7 \cdot 0 + \mathbf{3}, \quad 3^2 = 9 = 7 \cdot 1 + \mathbf{2}, \quad 3^3 = 27 = 7 \cdot 3 + \mathbf{6},$$
$$3^4 = 81 = 7 \cdot 11 + \mathbf{4}, \quad 3^5 = 243 = 7 \cdot 24 + \mathbf{5}, \quad 3^6 = 729 = 7 \cdot 104 + \mathbf{1}.$$

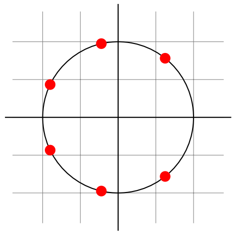

Figure 1.7: Reduced fraction with denominator 7

We see that the residues that occur among the first 6 powers are $3, 2, 6, 4, 5, 1$, which interestingly are all possible residues except 0. This property makes 3 special from the perspective of 7 and means in technical terms that 3 is a *primitive root mod* 7. We use this to assign a number to the reduced fractions with denominator 7 according to the position of the nominator in the above list. So for instance, we assign #6 to 1/7 since 1 is the residue of the *sixth* power of 3 when divided by 7. This leads to the following pairs:

$$\frac{1}{7} \leftrightarrow \#6, \quad \frac{2}{7} \leftrightarrow \#2, \quad \frac{3}{7} \leftrightarrow \#1, \quad \frac{4}{7} \leftrightarrow \#4, \quad \frac{5}{7} \leftrightarrow \#5, \quad \frac{6}{7} \leftrightarrow \#3.$$

Now given a number k between 1 and 5, we construct a new set of fractions as follows: Given a reduced fraction (with denominator 7), we take the associated number as above, multiply it by k, take the fraction with this nominator and denominator 6 and add this fraction to the reduced fraction we started with. For $k = 1$ and $k = 2$, this leads to;

$$\boxed{k=1} \quad \frac{1\cdot 6}{6} + \frac{1}{7} = \frac{8}{7}, \quad \frac{1\cdot 2}{6} + \frac{2}{7} = \frac{13}{21}, \quad \frac{1\cdot 1}{6} + \frac{3}{7} = \frac{25}{42}, \ldots$$
$$\boxed{k=2} \quad \frac{2\cdot 6}{6} + \frac{1}{7} = \frac{15}{7}, \quad \frac{2\cdot 2}{6} + \frac{2}{7} = \frac{20}{21}, \quad \frac{2\cdot 1}{6} + \frac{3}{7} = \frac{16}{21}, \ldots$$

We will now use these fractions to define points on a circle. To do this we identify the interval from 0 to 1 with the circumference of the unit circle in the plane such that 0 and 1 are both identified with the point with coordinates $(1, 0)$. This means that the reduced fractions with denominator 7 are now identified with the points shown in Figure 1.7. Now for each value of k, we plot the constructed fractions on the circle using *clock arithmetics*; if the time is 11 o'clock and 2 hours pass, then a watch does not go to 13, but to 1 instead. We do the same; for instance, we plot $\frac{8}{7}$ on the circle by starting at the point with coordinates $(1, 0)$ and then going around the circle (in an anti-clockwise direction) a total of one and one-seventh times. For each of the fractions above we plot them on the circle this way and then draw a vector (or line) from the origin (or center of the circle) to this point. The resulting vectors for $k = 1$ and $k = 2$ are show in Figure 1.8.

Now we do something which takes us away from the circle: For each $k = 1, \ldots, 5$, we *add* the 6 vectors that we have constructed; that is, we follow the direction all of them combined define. The end-points one gets this way define 5 new points (one for

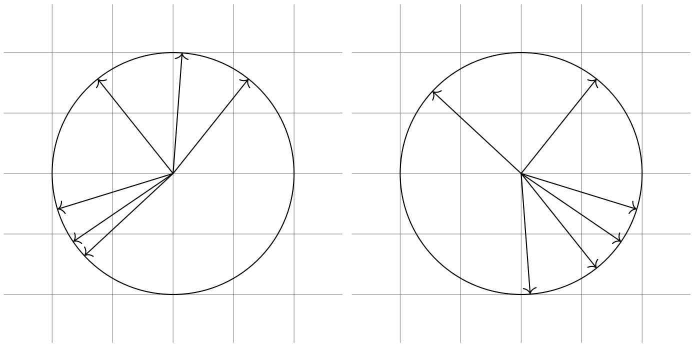

Figure 1.8: Vectors for $k = 1$ and $k = 2$

each $k = 1, \ldots, 5$), which we call the *Gauß points with denominator* 7. This last step of the "recipe" is illustrated in Figure 1.9. We admit that the steps above are quite involved and complicated to follow. This makes the following observation even more amazing; the Gauß points constructed from denominator 7 all lie on the same circle! This fact is very far from obvious from just staring at the "recipe", and was firstly proved by the great Carl Friedrich Gauß (which explains the naming of these points), see [47, Chapter 8].

The Gauß points in Figure 1.9 were constructed starting from reduced fractions with denominator 7, but we can follow a similar recipe with 7 replaced by a different prime number p (here it is actually important that the denominator is prime). This way we get a total of $p - 2$ Gauß points for each prime denominator p. Again the Gauß points will all lie on the same circle (this time with radius $\sqrt{p}$). If you have paid attention to the above, it will hardly be a surprise that we now ask the following question: *How do Gauß points distribute on the circle as the denominator becomes larger and larger?* This is a very hard question, which maybe comes as no surprise due to the complexity of the "recipe" described above. In Figure 1.10 are shown the Gauß points constructed from reduced fractions with denominator, respectively 17 and 59. We observe that although the Gauß points appear to be rather sporadic compared to reduced fractions, they again distribute quite evenly on the circle. One might speculate whether the Gauß points should also equidistribute on the circle as the denominator p grows. It turns out that this suspicion is correct, and the mathematical result is known as *Equidistribution of Gauß Sums*. As opposed to the case of reduced fractions above, this result is extremely profound and was proved by Deligne and Katz [55] in the 1980's using some very advanced mathematics.

In this setting the distribution of the Gauß points is the interesting question that we would like to understand, but this is very difficult. These Gauß points are however connected to reduced fractions, which are much easier to understand. This mirrors the situation in the theory of automorphic forms; here one studies some very complicated objects called *L-functions*, which we think of as analogues to the above defined Gauß points. L-functions are in general very hard to understand, but in some cases one

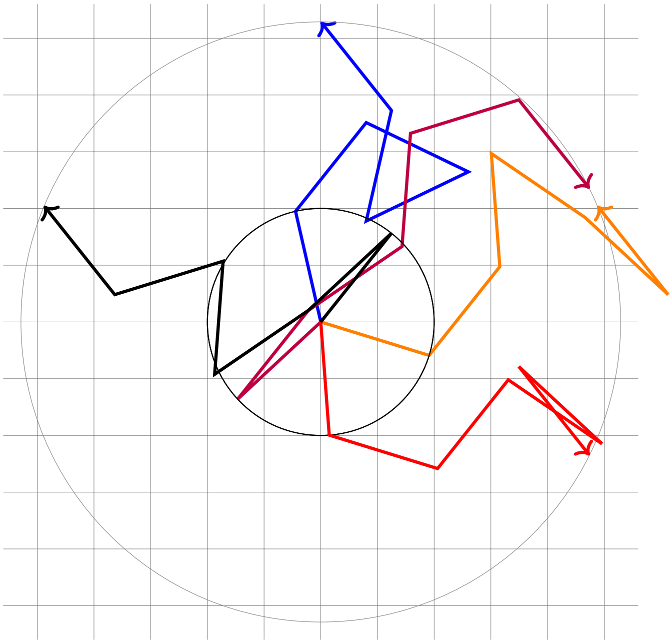

Figure 1.9: Construction of Gauß points with denominator 7

can relate these complicated objects to certain periods of automorphic forms. At least in the examples that occur in this thesis, we are able to obtain results about the automorphic periods. This then sheds light on the complicated objects that we wanted to understand in the first place.

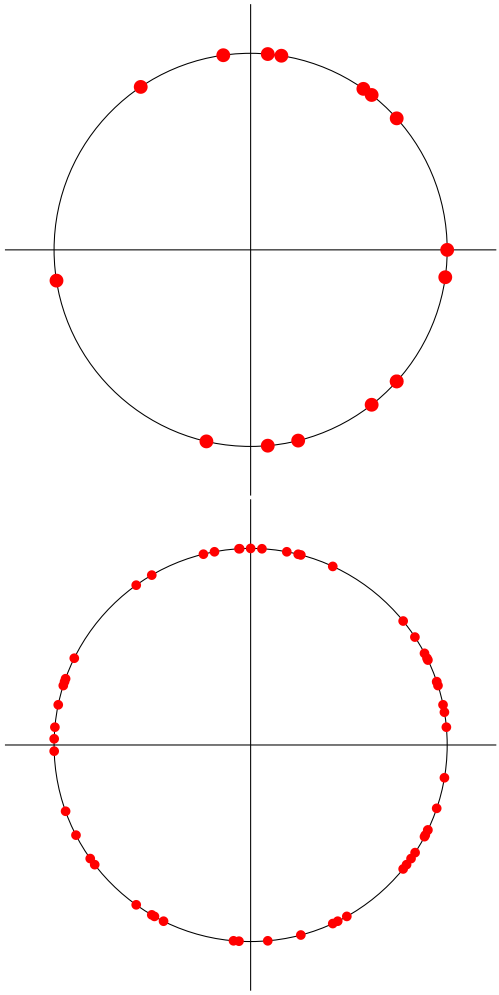

Figure 1.10: Gauß points constructed from denominators 17 and 59

CHAPTER 2

PERIODS IN THE THEORY OF AUTOMORPHIC FORMS

We will now go on to explain in more detail the kind of periods that arise in this thesis. The necessary background on classical modular forms or general automorphic forms will not be provided, but we will rather refer to the following excellent sources, respectively [49], [104] and [36], [16].

Periods play a prominent role in the theory of automorphic forms and have done so since the beginning through their connection to L-functions. As we well see below, periods of automorphic forms have turned out to have connections to a variety of other fields in number theory, and mathematics in general. In Section 2.1 we will give a short overview of the history of automorphic forms with an emphasize on the role of periods (and their relations to L-functions). Such periods seem to lack a unifying formal definition and many different notions exist throughout the literature. In Section 2.2 we will explain exactly what we mean by a *period of an automorphic form* in this thesis. For now we will just think of a period as an integral representation of the L-function of an automorphic form.

2.1 A short history of periods

It is fair to say that the (systematic) theory of automorphic forms began with the work of Hecke on modular forms building on the investigations of Jacobi, Dirichlet, Riemann, Ramanujan and many others (see [26] for an introduction to the history). Hecke introduced the important *Hecke operators* and obtained analytic continuation, functional equation and Euler product for the L-functions associated to a *Hecke cusp form* $f \in \mathcal{S}_k(\Gamma_0(N))$ of weight k and level N by the following (period) integral representation:

$$\Gamma\left(s+\frac{k-1}{2}\right)(2\pi)^{-s-\frac{k-1}{2}}L(f,s)=\int_0^\infty f(iy)y^{s+\frac{k-1}{2}}\frac{dy}{y},$$

where $L(f,s)=\sum a_f(n)n^{-s-\frac{k-1}{2}}$ for $\operatorname{Re} s>1$ with $a_f(n)$ the Fourier coefficients of f. Hecke's proof of the functional equation was inspired by Riemann's second proof of the functional equation for the zeta function, which also has an automorphic period (of a theta-series) at its core.

The next important development with regards to periods is the work of Rankin [82] and Selberg [88]. Their idea was to study the analytic properties of the L-function $L_f(s) = \sum_{n\geq 1} |a_f(n)|^2 n^{-(s+k-1)}$ for $f \in \mathcal{S}_k(\Gamma_0(N))$ via the following integral representation:

$$(4\pi)^{-(s+k-1)}\Gamma(s+k-1)L_f(s) = \int_{\Gamma_0(N)\backslash\mathbb{H}} y^k |f(z)|^2 E(z,s)d\mu(z),$$

where $E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \mathrm{Im}(\gamma z)^s$ is the non-holomorphic Eisenstein series (of level N). As an important application Rankin obtained an asymptotic formula for the sum $\sum_{n\leq X} |a_f(n)|^2$, and furthermore this method inspired Deligne in his proof of the final piece of the *Weil Conjectures*.

2.1.1 After Langlands

Langlands letter to Weil [58] from 1967 changed forever the landscape of automorphic forms and bridged the worlds of Galois representations and modular forms. He laid out the far-reaching web of conjectures known as the *Langlands Program*, which established *automorphic representations* as the proper generalization of modular forms (see [31] for an accessible introduction). From this perspective, a classical modular form corresponds to an automorphic representation of the group GL_2 over $\mathbb{Q}$, which means considering the representation theory of the adèlic group $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$. Langlands introduced the notion of an automorphic representation of any reductive algebraic group over a number field and defined an associated L-function. As part of the theory, Jacquet and Godement [33] gave an adèlic period formula for these L-functions for general number fields and general GL_n (for GL_1 this had already been achieved by Tate in his thesis [95]). Furthermore, the Rankin–Selberg method was generalized in the adèlic setting by Jacquet, Piatetski-Shapiro and Shalika [51] to general $\mathrm{GL}_m \times \mathrm{GL}_n$.

Following this, another important development in the theory of periods of automorphic forms is the work of Waldspurger [98]. He obtained a period formula for the Rankin–Selberg L-function $L(\pi\otimes\Theta_\chi, 1/2)$ where π in an automorphic representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ and Θ_χ is the theta-series associated to a Hecke character χ on $E^\times\backslash\mathbb{A}_E^\times$, where $E/\mathbb{Q}$ is a quadratic extension. The formula reads

$$\frac{|\mathcal{P}_\chi(\phi)|^2}{\langle\phi,\phi\rangle} = c_{\chi,\phi}\frac{L\left(\pi\otimes\Theta_\chi, 1/2\right)}{L(\mathrm{sym}^2\pi, 1)}, \tag{2.1.1}$$

where ϕ is *any* nonzero vector in π, $c_{\chi,\phi}$ is a finite product of local factors, $L(\mathrm{sym}^2\pi, s)$ is the symmetric square L-function of π and the automorphic period is given by

$$\mathcal{P}_\chi(\phi) := \int_{\mathbb{A}_\mathbb{Q}^\times E^\times\backslash\mathbb{A}_E^\times} \phi(x)\chi^{-1}(x)\,dx,$$

for an embedding $\mathbb{A}_\mathbb{Q}^\times E^\times\backslash\mathbb{A}_E^\times \hookrightarrow \mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$. In the special case where ϕ is an Eisenstein series such formulas had been discovered earlier by Hecke and Siegel. In the case of Eisenstein series the proofs are elementary, but in general one needs the full power of the representation theoretic language.

A profound generalization of Waldspurger's period formula has been proposed by Gan, Gross and Prasad [32] (or more precisely the global refinements due to Ichino and Ikeda [46] and N. Harris [39]), who have conjectured a period formula for Rankin–Selberg L-functions on $\mathrm{SO}(V) \times \mathrm{SO}(W)$ for any quadratic space V over a number field F and $W \subset V$ a non-degenerate hyperplane. By an accidental isomorphism the formula (2.1.1) corresponds exactly to the case where W is 2-dimensional over $\mathbb{Q}$.

Motivated by questions in quantum chaos (see Section 3.3 below for more details), Watson [99] obtained a period formula for *triple convolution L-functions*:

$$\left| \int_{\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} \phi_1(z)\phi_2(z)\phi_3(z)\frac{dxdy}{y^2} \right|^2 = \frac{\Lambda(\phi_1 \otimes \phi_2 \otimes \phi_3, 1/2)}{\Lambda(\mathrm{sym}^2\phi_1, 1/2)\Lambda(\mathrm{sym}^2\phi_2, 1/2)\Lambda(\mathrm{sym}^2\phi_3, 1/2)}, \tag{2.1.2}$$

where ϕ_i are Hecke–Maaß cusp forms of level 1 and the quotient consists of completed L-functions. Notice that if ϕ_3 is an Eisenstein series, this reduces (essentially) to the classical Rankin–Selberg formula. The formula (2.1.2) has been generalized to general automorphic representations of GL_2 over number fields by Ichino [45] (resolving a refinement of a conjecture due to Jacquet). By another accidental isomorphism, this resolves the conjecture of Gan–Gross–Prasad in the case of $\mathrm{SO}_4 \times \mathrm{SO}_3$.

In order to keep the exposition relatively short, we have left out many important results on periods (most prominently the work of Gross and Zagier [37]). The relation between automorphic periods and L-functions remain mysterious in general. There is however an emerging philosophy describing a *relative Langlands* picture as in the work of Sakellaridis and Venkatesh [86], which gives some conceptual framework for the role of periods in the theory of automorphic forms.

2.1.2 Cohomology and modular symbols

A different kind of periods of automorphic forms emerged in the work of Eichler and Shimura on cohomological models for modular forms. The *Eichler–Shimura isomorphism* for $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ discrete, cofinite subgroup is the isomorphism:

$$\mathcal{S}_k(\Gamma) \otimes \overline{\mathcal{S}_k(\Gamma)} \xrightarrow{\sim} H^1_P(\Gamma, V_{k-2}(\mathbb{C})), \tag{2.1.3}$$

where $V_{k-2}(\mathbb{C})$ is the space of degree $k-2$ homogenous polynomials in two variables equipped with a certain action of Γ and H^1_P denotes parabolic cohomology (see [90, Chapter 8] for details). More precisely, the isomorphism is given by mapping $(f, \overline{g})$ to the element of cohomology defined by $\gamma \mapsto \sigma_f(\gamma) + \sigma_{\overline{g}}(\gamma)$ where

$$\sigma_f(\gamma)(X, Y) = \int_{\gamma\infty}^{\infty} f(z)(X + zY)^{k-2}dz. \tag{2.1.4}$$

These polynomials are what Eichler calls the *period polynomials* of f and Shimura [89] refers to the coefficients of the period polynomials $\{\int_{\gamma\infty}^{\infty} f(z)z^j dz\}_{0\leq j\leq k-2}$ as the *periods of $f \in \mathcal{S}_k(\Gamma)$*.

Inspired by this Manin [68] (and independently Birch) introduced what are called *modular symbols*. For Hecke congruence subgroups $\Gamma_0(N)$, we define the associated modular symbols as the $\mathbb{Q}$-vector space generated by pairs of cusps $\{\mathfrak{a}, \mathfrak{b}\}$, $\mathfrak{a}, \mathfrak{b} \in \mathbb{Q} \cup \{\infty\}$ modulo two explicit relations (given in terms of the action of $\Gamma_0(N)$ on the cusps). These symbols provide a combinatorial model for the (rational) homology of modular curves (relative to the cusps), and have been used with great success for computing modular forms (see [20]). We will mainly be interested in what is known as the *modular symbols map*, which is obtained via the Poincaré pairing between homology and cohomology. More precisely given $f \in \mathcal{S}_2(\Gamma_0(N))$, we get a 1-form $f(z)dz$ and consider the map

$$\mathbb{Q} \ni r \mapsto \langle r, f \rangle := 2\pi i \int_r^{i\infty} f(z)dz \in \mathbb{C}. \tag{2.1.5}$$

Modular symbols (and the above pairing) have been used with great success in the study of the arithmetics of L-functions (see [3], [70] and the references therein) due to the Birch–Stevens formula: For χ a primitive Dirichlet character modulo q with $(q, N) = 1$, we have

$$\tau(\overline{\chi})L(f, \chi, 1/2) = \sum_{a \in (\mathbb{Z}/q\mathbb{Z})^\times} \chi(a)\langle a/q, f \rangle, \tag{2.1.6}$$

where $L(f, \chi, s) = \sum_{n \geq 1} \lambda_f(n)\chi(n)n^{-s}$ is the (naively) twisted L-function.

These cohomological ideas have been pushed much further. We will not go further into this but simply point to [57].

2.2 A notion of periods for automorphic forms

We will now present a definition of what we mean by a "period of an automorphic form", which encompasses all the examples alluded to in the previous section. This notion will be general enough to encompass the kind of periods encountered in this thesis. But (hopefully) narrow enough so that it still offers a useful perspective on the topics of the papers in this thesis.

Definition 2.2.1. *Let $\Gamma \subset G$ be a lattice in an S-arithmetic group and $Y \subset \Gamma\backslash G$ a subset equipped with a measure ν. Let V_Y be a vector space consisting of measurable functions on Y, and fix a basis $\{\psi_i\}$ for V_Y. Then given a (sufficiently nice) function F on $\Gamma\backslash G$, we define the numbers $\int_Y F\psi_i d\nu$ as the* **periods of F along Y** *(with respect to the basis $\{\psi_i\}$).*

This notion is a slight generalization of the one presented by Venkatesh in the first paragraph of [97]. This is (intentionally) a very general definition as it for instance depends on the choice of space V_Y and basis $\{\psi_i\}$. This thesis is a testament to the scope of such periods of automorphic forms, and we will encounter connections to topics ranging from reciprocity formulas, quantum modular forms, equidistribution, quantum chaos and the theory of automorphic L-functions. In the context of equidistribution problems, it is natural to consider the setup where ν is a probability measure and $V_Y = L^2(Y,\nu)$, which is exactly the definition given in [97]. We will encounter many different examples of periods of automorphic forms; modular symbols, additive twists, period polynomials, Heegner traces, shifted convolution sums, pairings between homology and cohomology. In the next chapter we will explain the history and problems surrounding these different kind of periods and how the papers of this thesis fit into the story. Below we have collected the data, which when plugged into Definition 2.2.1 gives rise to these periods. We will consider four different main settings, each corresponding to a specific choice of G, Γ, F, Y, ν and V_Y. In each of these cases there are a number of different choices of bases, which will all be important in different settings.

(1) $G = \mathrm{PSL}_2(\mathbb{R})$, $\Gamma = \Gamma_0(N)$, F is the lift to $\Gamma\backslash G$ of $f \in \mathcal{S}_k(\Gamma_0(N))$,

$$Y = Y_q = \left\{ \left(\begin{smallmatrix} y & a/q \\ 0 & 1 \end{smallmatrix}\right) \mid y \in \mathbb{R}_{>0}, a \in (\mathbb{Z}/q\mathbb{Z})^\times \right\}.$$

where q is a positive integer and $d\nu = \sum_{i=1}^{\varphi(q)} dy_i$, with $y_1, \ldots, y_{\varphi(q)}$ the y-variables of each of the components (one for each $a \in (\mathbb{Z}/q\mathbb{Z})^\times$). We will be interested in the periods in the setting where V_Y consists of functions of the form $\sum_{i=1}^{\varphi(q)} P_i(y_i)$ with P_i polynomials of degree at most $k-2$. In this case we have three interesting bases.

(P1.1) We can consider the basis consisting of maps of the form

$$\left(\begin{smallmatrix} y & a/q \\ 0 & 1 \end{smallmatrix}\right) \mapsto \mathbf{1}_{a=a_0} y^{j-k/2},$$

where $\mathbf{1}_{a=a_0}$ is an indicator function for $a_0 \in (\mathbb{Z}/q\mathbb{Z})^\times$ and $0 \le j \le k-2$. In this case the periods correspond to special values of the *additive twists* of the L-function of f. When $k=2$, one recovers the modular symbols (2.1.5).

(P1.2) Alternatively, we can consider the basis consisting of

$$\left(\begin{smallmatrix} y & a/q \\ 0 & 1 \end{smallmatrix}\right) \mapsto \mathbf{1}_{a=a_0} (a/q + iy)^j y^{-k/2},$$

for $a_0 \in (\mathbb{Z}/q\mathbb{Z})^\times$ and $0 \le j \le k-2$. This corresponds to the coefficients of the period polynomials of f in the sense of Eichler (2.1.4).

(P1.3) Finally, we can consider the basis consisting of

$$\left(\begin{smallmatrix} y & a/q \\ 0 & 1 \end{smallmatrix}\right) \mapsto \chi(a) y^{j-k/2},$$

for χ a Dirichlet character modulo q and $0 \le j \le k-2$. This corresponds to the special values of the automorphic L-function $L(f \otimes \chi, s)$ (with some fudge factors if $(N,q) > 1$ and/or χ is non-primitive) using the Birch–Stevens formula (2.1.6).

All of these choices of periods are interesting in their own right as we will see in Papers A, B, C and D; the first one gives rise to a normal distribution and residual equidistribution, the second one is natural from the perspective of cohomology, and the last one is evidently important in the theory of automorphic forms.

(2) $G = \mathrm{PSL}_2(\mathbb{R})$, $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, F is the lift of the non-holomorphic Eisenstein series $E(z, 1/2 + it)$ to $\Gamma\backslash G$, and

$$Y = Y_D = \{z_{\mathfrak{a}} \mid \mathfrak{a} \in \mathrm{Cl}_D\},$$

where Cl_D denotes the class group of the imaginary quadratic field of discriminant $D < 0$ and $z_{\mathfrak{a}}$ is the associated *Heegner point* (considered as points in $\Gamma\backslash G$ in an appropriate way). We equip Y with the counting measure and let V_Y be functions on Cl_D. In this case we have two dual bases.

(P2.1) One is given by maps of the form $z_{\mathfrak{a}} \mapsto \chi(\mathfrak{a})$, where χ is a class group character of Cl_D. The periods with respect to this basis are related to class group L-functions by a special case of Waldspurger's formula (2.1.1) (see also (3.3.3) below).

(P2.2) The other basis is given by the maps $z_{\mathfrak{a}} \mapsto \mathbf{1}_{\mathfrak{a}=\mathfrak{a}_0}$ where $\mathfrak{a}_0 \in \mathrm{Cl}_D$. The periods in this context are treated in Paper F and can be bounded using methods from the theory of exponential sums.

There is a similar picture for negative discriminants using *Heegner cycles*, which are certain closed geodesics associated to elements of the class group of real quadratic fields.

(3) $G = \mathrm{PSL}_2(\mathbb{R})$, $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, F is the lift of a Hecke–Maaß cusp form ϕ to $\Gamma\backslash G$, and

$$Y = Y_D = \{\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}} \subset \mathbb{H} \mid \mathfrak{a} \in \mathrm{Cl}_D^+\},$$

with $D > 0$ a positive fundamental discriminant, Cl_D^+ the (narrow) class group of discriminant D and $\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}}$ a certain hyperbolic orbifold constructed in [25] (again considered as a subset of $\Gamma\backslash G$ in an appropriate way). As above, we equip Y with the counting measure and put V_Y equal to functions on Cl_D^+.

(P3.1) In this case we will be interested in the basis given by $\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}} \mapsto \chi(\mathfrak{a})$, where χ is a class group character of Cl_D^+. The periods with respect to this basis are related to the central values of Rankin–Selberg L-functions $L(\phi \otimes \Theta_\chi, 1/2)$, which is the main result proved in Paper E.

These periods are furthermore used in Paper E to obtain sparse equidistribution results for the hyperbolic orbifolds $\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}}$.

(4) $G = \mathrm{PSL}_2(\mathbb{R}) \times \mathrm{PSL}_2(\mathbb{R})$, $\Gamma = \mathrm{PSL}_2(\mathbb{Z}) \times \mathrm{PSL}_2(\mathbb{Z})$, $F = (F_k, F_k)$ where F_k is the lift of a holomorphic cusp form of weight k and level 1 to $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PSL}_2(\mathbb{R})$, and $Y \cong \mathrm{PSL}_2(\mathbb{Z})\backslash\mathrm{PSL}_2(\mathbb{R})$ is the diagonal equipped with the (left) Haar measure of G and $V_Y = L^2_{\mathrm{cusp}}(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}, \frac{dxdy}{y^2})$, where we view functions on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ as functions on Y the standard way. In this case there are two interesting bases.

(P4.1) The first basis is given by Hecke–Maaß cusp forms. In this case it follows by the formula of Ichino and Watson (2.1.2) that the periods are connected to triple convolution L-functions.

(P4.2) The other basis is constructed from Poincaré series;

$$P_{h,m}(z) = \sum_{\gamma \in \Gamma_\infty \backslash \mathrm{SL}_2(\mathbb{Z})} e(m\gamma z) h(\operatorname{Im} \gamma z),$$

for h smooth and compactly supported and $m \in \mathbb{Z}$. In this case the periods are connected to shifted convolution sums (see Section 3.3.3 below). In particular, this connection is a key input in Paper G.

Remark 2.2.2. It is worth mentioning that all of the above periods can be written down nicely in the adèlic language and are in some cases more naturally defined in this setting. We have however sticked to the classical language, since this makes the comparison with the papers of this thesis clearer.

CHAPTER 3

SOME PROBLEMS IN ARITHMETIC STATISTICS

It is an amazing fact that periods of automorphic forms show up in many different parts of mathematics. In particular, the focus of this thesis is the connections between automorphic periods and arithmetic statistics. As will be apparent later in this thesis, automorphic periods can often can be studied using a variety of powerful tools. This can then be used to obtain information about the statistical questions, we wanted to understand in the first place (that being rational points on elliptic curves, automorphic L-function, equidistribution or quantum chaos). This observation is part of an emerging philosophy that one can use "geometric" methods to study periods of automorphic forms. This shows maybe most prominently in the work of Venkatesh [97] and Michel and Venkatesh [75], where they use (among other things) ergodic methods to obtain subconvexity bounds of certain L-functions. Interestingly, subconvexity bounds for L-functions were originally studied using techniques from analytic number theory motivated by their connections (through automorphic periods) to certain distribution problems. Venkatesh in [97], however showed that one can "turn the tables" and study the L-functions using the period representations. Recently, a variety of other methods have been used successfully to study automorphic periods; spectral theory [81], dynamical systems [59], [8] and micro-local analysis [76].

The results of the papers in this thesis touch upon a variety of questions of statistical nature, which can be separated into three main topics; **modular symbols**, **geometric invariants of quadratic fields** and **arithmetic quantum chaos**. For the two latter, there exist a vast literature including many excellent surveys. The situation is quite different for the results relating to modular symbols, where no complete overview of the current state of affairs seems to exist. Thus we will put our emphasize on the history of the distribution of modular symbols below, and refer to other sources for in-depth surveys on the study of the distribution of geometric invariants of quadratic fields and arithmetic quantum chaos.

3.1 Around the distribution of modular symbols

In the 90's inspired by a connection to Szpiro's conjecture (relating the conductor and the discriminant of elliptic curves), Goldfeld initiated the study of the distribution of modular symbols [34], [35]. Furthermore, he introduced the following Eisenstein series twisted by modular symbols, known as the *Goldfeld Eisenstein series*;

$$\sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(N)} \langle\gamma\infty, f\rangle^n \operatorname{Im}(\gamma z)^s,$$

where $f \in \mathcal{S}_2(\Gamma_0(N))$ is a cusp form of weight 2 and level N and $\Gamma_\infty = \langle\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\rangle$. In his thesis, O'Sullivan [77] proved meromorphic continuation and functional equation in the case $n = 1$. The first to obtain precise analytic information about these series for all n were Petridis and Risager, who used this information to prove that modular symbols $\langle a/q, f\rangle$ are asymptotically normally distributed when ordered by a^2+q^2 and appropriately normalized [78, Theorem A].

Theorem 3.1.1 (Petridis and Risager, 2004)**.** *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a cusp form of weight 2 and level N. Then as $Q \to \infty$, the distribution of the numbers*

$$\left\{\frac{\operatorname{Re}\langle a/q, f\rangle}{\sqrt{\log(a^2+q^2)}} \mid \sqrt{a^2+q^2} \leq Q, (a,q) = 1, N|q\right\}$$

tend to a normal distribution with mean 0 and a certain explicit variance.

Petridis and Risager used their methods to obtain a number of results [80], [83], [79], and the ideas of Goldfeld also inspired the notion of a *second order modular form* [17].

In 2016 (unaware of the results of Petridis and Risager), Mazur and Rubin (and Stein) initiated the study of the arithmetic statistics of modular symbols. This time their motivation was to obtain heuristics for the following question regarding the Diophantine stability of elliptic curves $E/\mathbb{Q}$:

(Q1) *How often is* $\operatorname{rank}_{\mathbb{Z}} E(K) > \operatorname{rank}_{\mathbb{Z}} E(\mathbb{Q})$
as $K/\mathbb{Q}$ ranges over abelian extensions?

Assuming the BSD-conjecture (in a sufficiently general form), we have

$$\operatorname{rank}_{\mathbb{Z}} E(K) = \operatorname{ord}_{s=1} L(E_{/K}, s),$$

where $L(E_{/K}, s)$ is the Hasse–Weil L-function of the base-change of E to K. For $K/\mathbb{Q}$ abelian, we have the following factorization:

$$L(E_{/K}, s) = \prod_{\chi\in\widehat{\operatorname{Gal}(K/\mathbb{Q})}} L(E,\chi,s),$$

where $L(E,\chi,s) = \sum a_E(n)\chi(n)n^{-s}$ is the χ-twisted L-function of E. This implies that the question **(Q1)** is related to the vanishing properties of the central values of the twisted L-functions $L(f_E, \chi, 1/2)$, where χ is a Dirichlet character and f_E is the weight 2 cusp form corresponding to E. These L-functions are in turn related to modular symbols through the Birch–Stevens formula (2.1.6), as we saw above.

This led Mazur and Rubin to study the statistics of modular symbols with an arithmetic ordering. Using this, Mazur and Rubin develop a heuristics, from which they obtain a (conjectural) answer to **(Q1)** in certain families. In particular, they recover predictions due to David, Fearnley and Kisilevsky [21] obtained using *Random Matrix Theory.* The ordering originally used by Petridis and Risager came naturally out of the Goldfeld Eisenstein series, whereas from an arithmetic point of view it is more convenient to order the modular symbols by the denominator of the cusps. Mazur and Rubin conjectured among other things that with this ordering the modular symbols should be asymptotically normally distributed.

Conjecture 3.1.2 (Mazur and Rubin, 2016)**.** *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a cusp form of weight 2 and level N. Then as $q \to \infty$, the distribution of the numbers*

$$\left\{ \frac{\mathrm{Re}\langle a/q, f\rangle}{\sqrt{\log q}} \mid a \in (\mathbb{Z}/q\mathbb{Z})^\times \right\}$$

tend to a normal distribution with mean 0.

This conjecture seems to be extremely hard. First of all, the size of the family in the conjecture above is the square-root of the family appearing in Petridis and Risager (relative to the largest denominator appearing in the families). Secondly, calculating the second moment for this family amounts (using Parseval and Birch–Stevens (2.1.6)) to calculating the second moment of Dirichlet twists $L(f, \chi, 1/2)$, which was only achieved very recently by the combination of the work of Blomer, Fouvry, Kowalski, Michel, Milićević and Sawin [14], [10], [56]. The methods are extremely profound using the full power of spectral and algebro-geometric methods.

It was soon after Mazur and Rubin's work realized by Petridis and Risager that their methods from [78] could be adapted to the arithmetic setting as well. They were able to obtain an average version of Conjecture 3.1.2 with the further refinement that they could restrict the cusp to a fixed interval. Finally, they also obtained a beautiful formula for the variance. The following result is [81, Theorem 1.7].

Theorem 3.1.3 (Petridis and Risager, 2018)**.** *Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a cusp form of weight 2 and square-free level N. Fix an interval $I \subset \mathbb{R}/\mathbb{Z}$. Then as $Q \to \infty$, the distribution of the numbers*

$$\left\{ \frac{\mathrm{Re}\langle a/q, f\rangle}{\sqrt{\log q}} \mid a \in (\mathbb{Z}/q\mathbb{Z})^\times, 0 < q \leq Q, a/q \in I \right\}$$

tend to a normal distribution with mean 0 *and variance*

$$\mathcal{C}_f = 6/\pi^2 \prod_{p \mid N} (1 + p^{-1})^{-1} L(\mathrm{sym}^2 f, 1).$$

Furthermore the methods of Petridis and Risager apply to cusp forms of any discrete, cofinite subgroup of $\mathrm{SL}_2(\mathbb{R})$ with cusps.

Later, Lee and Sun [59] obtained a different proof of Theorem 3.1.3 using dynamical methods. Their approach relies on *Manin's trick*, which expresses a general modular symbol in a convenient basis called *Manin symbols*. This basis is related to continued fraction expansions, which enabled Sun and Lee to employ the dynamics of such expansions as in the work of Baladi and Vallée [4]. This method is however restricted to the case of arithmetic subgroups and furthermore their method does not give a formula for the variance (see [59, page 23]).

Mazur and Rubin (together with Stein) also put forth a number of other conjectures in [73] regarding modular symbols, many of which have been (partially) resolved [22], [94], [11, Thm. 9.2]. We will not touch upon these interesting results in this thesis but instead refer to [73, Sec. 4] for a nice overview of the known results.

3.1.1 Residual distribution of modular symbols

At a talk at the number theory seminar at the California Institute of Technology, Mazur put forth a conjecture concerning the mod p distribution of modular symbols (see the unpublished note [71]). More precisely, let $f_E \in \mathcal{S}_2(\Gamma_0(N))$ be a cusp form of weight 2 and level N corresponding to an elliptic curve $E/\mathbb{Q}$. Then it is known that

$$\mathfrak{m}_E^{\pm}(a/q) := \frac{1}{\Omega^{\pm}} \left(\langle a/q, f_E\rangle \pm \langle -a/q, f_E\rangle\right) \in \mathbb{Z} \tag{3.1.1}$$

for all $a/q \in \mathbb{Q}$, where $\Omega^{\pm}$ are the Néron periods of E (see [73, Sec. 1]). Now one can ask how the values of $\mathfrak{m}_E^{\pm}(a/q)$ distribute among the residue classes modulo primes p. Mazur and Rubin conjecture that the residues should equidistribute under some assumptions on p and E.

Conjecture 3.1.4 (Mazur and Rubin, 2016)**.** *Let p be a prime such that the residual representation of E mod p is surjective and p is an ordinary and good prime of E. Then the distribution of the residues*

$$\left\{\mathfrak{m}_E^{\pm}(a/q) \bmod p \mid a \in (\mathbb{Z}/q\mathbb{Z})^{\times}\right\}$$

tend to the uniform distribution on $\mathbb{Z}/p\mathbb{Z}$ as $q \to \infty$.

Lee and Sun were able to adapt their dynamical approach to this setting as well, and as above they obtained an averages version of this conjecture, see [59, Theorem C.3].

Theorem 3.1.5 (Lee and Sun, 2019)**.** *Let p and E be as above. Then the distribution of the residues*

$$\left\{\mathfrak{m}_E^{\pm}(a/q) \bmod p \mid 0 < q \leq Q\,, a \in (\mathbb{Z}/q\mathbb{Z})^{\times}\right\}$$

tend to the uniform distribution on $\mathbb{Z}/p\mathbb{Z}$ as $Q \to \infty$,.

In Paper D (joint with Constantinescu) we give a new automorphic proof of this result and obtain a number of refinements and (most importantly) extend the results to hyperbolic spaces of arbitrary dimension. In certain special cases we can actually prove the full conjecture using a connection to *Eisenstein congruences.*

3.1.2 Generalizations of the conjectures of Mazur and Rubin

We will now discuss certain generalizations of Theorems 3.1.3 and 3.1.5. From the automorphic perspective it is natural to ask whether there exists generalizations to higher weight cusp forms, Maaß forms, GL_2 over general number fields or different algebraic groups. Whereas from the dynamical perspective other generalizations are appealing (as we will see below).

The case of higher weight

Regarding a generalization to higher weight, the first question is what the appropriate analogue of modular symbols should be, so that one will see a normal distribution appearing. Many different generalizations of modular symbols exist in the higher weight case, and at some point it was believed that the period polynomials attached to a cusp form would be the correct analogue. We recall from Section 2.1.2 that period polynomials are defined as follows:

$$\sigma_f(\gamma)(X,Y) = \int_{\gamma\infty}^{\infty} f(z)(X+zY)^{k-2}dz,$$

for $f \in \mathcal{S}_k(\Gamma_0(N))$ and $\gamma \in \Gamma_0(N)$. This notion is the natural generalization of modular symbols from a cohomological perspective in view of the Eichler–Shimura isomorphism (2.1.3). The coefficients of these polynomials correspond to the periods (**P1.2**) above. The analytic properties of period polynomials have been studied a lot recently, see [23] for a survey of results. These works have mainly studied the properties of the zeroes of $\sigma_f(S)$ as f varies (where $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$), which have been shown to satisfy an analogue of the Riemann Hypothesis. In Paper B we study the distribution of period polynomials and their zeroes when f is fixed and γ varies. In particular, we determine the limiting joint distribution of the coefficients of period polynomials, which turns out to be very far from normal.

In order to get a proper generalization of Theorem 3.1.3, the key turned out to be the Birch–Stevens formula (2.1.6). For higher weight cusp forms $f \in \mathcal{S}_k(\Gamma_0(N))$ what take the place of modular symbols are *additive twists*, defined as:

$$L(f,r,s) := \sum_{n\geq 1} a_f(n)e(nr)n^{-s},$$

for $\operatorname{Re} s > (k+1)/2$ and $r \in \mathbb{Q}$, where $a_f(n)$ are the Fourier coefficients of f and $e(x) = e^{2\pi i x}$. They satisfy analytic continuation and functional equation relating $s \leftrightarrow k-s$ (see Section 3.3 in Paper A for details). The special values $s = 1, \dots, k-1$ of the additive twists correspond exactly to the periods (**P1.1**) above. In this case we have the following generalization of the Birch–Stevens formula: For a primitive Dirichlet character χ modulo q, we have

$$\tau(\overline{\chi})L(f,\chi,1/2) = \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} \chi(a)L(f,a/q,k/2), \tag{3.1.2}$$

where $L(f,\chi,s) = \sum \lambda_f(n)\chi(n)n^{-s}$ is the (naively) twisted L-function of f and $\tau(\overline{\chi})$ is a Gauß sum. This formula is exactly encoding the relation between the two periods (**P1.1**) and (**P1.3**).

To explain why one should expect additive twists to be asymptotically normally distributed, one has to look at the problem from a "moments of L-functions" perspective: If we consider the moments of the additive twists, use an approximate functional equation and look at the contribution of the diagonal term, then one sees exactly the moments of the normal distribution appearing from the combinatorics. Assuming that the off-diagonal terms are negligible (which is only known for the 1st and 2nd complex moments by [11]) the normal distribution will follow by the method of moments. This heuristics applies equally well for all weights and also for Hecke–Maaß cusp forms. We hope to pursue these ideas further in the future.

In Paper A we prove that indeed additive twists are asymptotically normally distributed (again with the extra average) using an extension of the approach of Petridis and Risager. We also provide interesting applications to automorphic L-functions using the Birch–Stevens formula. Independently, Bettin and Drappeau [8] obtained a different proof of the normal distribution result using dynamical methods. Their method however only works for cusp forms with trivial level. The starting point for them is the fact that for $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$, the map $\mathbb{Q} \ni r \mapsto L(f, r, k/2)$ defines a *quantum modular form* in the sense of Zagier [105]. Quantum modularity in this context means concretely that we have

$$L(f, r, k/2) = L(f, -1/r, k/2) + h(r), \tag{3.1.3}$$

for all $r \in \mathbb{Q}\backslash\{0\}$, where $h : \mathbb{R} \to \mathbb{R}$ is some continuous function, which in this case can be estimated easily. Now iterating this we get (using that h is even) that

$$L(f, r, k/2) = h(r) + h(T(r)) + \ldots + h(T^m(r)),$$

for some $m \geq 1$, where $T(x) = \{\frac{1}{x}\}$ is the Gauß map. Now one can employ the dynamics of the Gauß map using an extension of the methods in [4] to get a normal distribution result. The property (3.1.3) corresponds to quantum modularity with respect to the matrix $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, which is the reason why the method does not directly generalize to non-trivial level. In return, the method of Bettin and Drappeau applies to a much larger class of quantum modular forms of level 1. In particular, they are able to show that the central values of the *Estermann function* are asymptotically normally distributed. This is a major achievement since the method of moments does not work in this case (as follows from the moment calculation of Bettin [7]).

Further symmetries, quantum modular forms and reciprocity formulas

Quantum modular forms were introduced by Zagier [105] motivated by certain symmetries satisfied by invariants appearing in quantum field theory. In words, a quantum modular form of level N and weight k is a function on $\mathbb{Q}$, which is "almost modular" with respect to the action of $\Gamma_0(N)$ given by the weight k slash operator (i.e. $f_{|k,\gamma}(r) := j(\gamma, r)^{-k} f(\gamma r)$). More precisely "almost modular" means that the discrepancy:

$$g_\gamma(r) := f(r) - j(\gamma, r)^{-k} f(\gamma r), \quad r \in \mathbb{Q},$$

for $\gamma \in \Gamma_0(N)$ fixed, can be extended to a continuous/analytic/smooth function on $\mathbb{R}\backslash\{\gamma^{-1}\infty\}$ (the choice of regularity varies from case to case). In Paper C, extending the results of [8], we prove that the central value of additive twists $L(f, \cdot, k/2)$ with $f \in \mathcal{S}_k(\Gamma_0(N))$ define a quantum modular form of weight 0 and level N (with regularity=continuous). This can be seen as a further (approximate) symmetry that is satisfied by the periods (**P1.1**).

In earlier work of Bettin [6], it is proved that the Estermann function defines a quantum modular form, which is used to give a refinement of a reciprocity relation discovered by Conrey [18]. The reciprocity relation is related to twisted second moments of Dirichlet L-functions and takes the following form:

$$\sum_{\chi \bmod q} |L(\chi, 1/2)|^2 \chi(l) \rightsquigarrow \sum_{\chi \bmod l} |L(\chi, 1/2)|^2 \chi(-q), \tag{3.1.4}$$

for primes $q \neq l$, where the sums run over Dirichlet characters modulo q, respectively l. Recently, many other reciprocity relations for moments of L-functions have been discovered, see [13], [12], [2].

Noticing that $|L(\chi, 1/2)|^2 = L(E \otimes \chi, 1/2)$, where E is an appropriate Eisenstein series (in representation theoretic language E should correspond to the isobaric sum $1 \boxplus 1$), it is natural to expect that a generalization of the reciprocity law should exist when we replace E by a GL_2 *cuspidal* automorphic representation. In Paper C we show that this is indeed the case for holomorphic cusp forms. Furthermore, building on these ideas we have work in progress, which aims to generalize the reciprocity relations above to Hecke–Maaß cusp forms. Combining this with the methods in [8], this might furthermore lead to a normal distribution result for additive twists of Hecke–Maaß cusp forms.

Residual distribution for Lorentz groups

A question that begs an answer is whether the automorphic method can be adapted to deal with residual distribution of modular symbols as in Conjecture 3.1.4. This is the topic of Paper D, where we show that this indeed can be done using analytic properties of twisted Eisenstein series. This can be seen as a discrete version of the method of Petridis and Risager introduced in [78]. Furthermore, this method allows for a generalization to quotients of higher dimensional hyperbolic spaces $\mathbb{H}^{n+1}$ or more precisely cohomology classes in $H^1(\Gamma, \mathbb{F}_p)$ where $\Gamma \subset \mathrm{SO}(n+1, 1)$ is a discrete and cofinite subgroup with cusps.

Remark 3.1.6. A different possible generalization is to study the residual distribution of the coefficients of period polynomials. These coefficients are known to have good arithmetic properties and it would thus be interesting to obtain results about their residual distribution. Nothing seems to be known at present.

As a final remark, we will emphasize that the dynamical and automorphic approach supplement each other as they allow for generalizations in different directions. The results in Papers A, B and C and D apply equally well to general discrete and cofinite subgroups of $\mathrm{SL}_2(\mathbb{R})$, and in a different direction it has been shown by Constantinescu [19] that the automorphic method generalizes very naturally to the case of Bianchi modular forms (i.e. GL_2 over imaginary quadratics). Interestingly, the proof in [19] avoids the method of moments and uses instead the *Berry–Esseen inequality* to obtain the distribution result. On the other hand, the dynamical methods can be used to obtain distribution results for very general quantum modular forms, as mentioned above.

3.2 Distribution of geometric invariants of quadratic fields

We will now describe a different circle of ideas, surrounding geometric invariants of quadratic fields. For a beautiful and informative summary of the history, we will recommend [74]. Below we will highlight some developments, which are relevant for this thesis.

In the 80's, Duke [24] discovered a connection between some classical questions about quadratic forms and the theory of automorphic forms. He observed that certain periods of automorphic forms were closely related to the distribution of integral points of quadratic forms. Duke studied the distribution on the modular surface $X_0(1) := \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ of *Heegner points* and *Heegner cycles* associated to class groups of quadratic fields. Recall that the (wide) class group of fundamental discriminant D is defined as the quotient of fractional ideals of K by principal fractional ideals, where $K = \mathbb{Q}(\sqrt{D})$ is the quadratic field of discriminant D. Given a negative fundamental discriminant D, we can associate to each element $\mathfrak{a} \in \mathrm{Cl}_D$ of the class group a point on the modular surface $z_\mathfrak{a} \in X_0(1)$ known as the Heegner point (associated to $\mathfrak{a}$). Similarly, for positive discriminants $D > 0$ and $\mathfrak{a} \in \mathrm{Cl}_D$, we can associate a closed geodesic $\gamma_\mathfrak{a} \subset \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, which we call the Heegner cycle associated to $\mathfrak{a}$. Duke proved that as $|D| \to \infty$ the Heegner points (respectively cycles) equidistribute on the modular surface with respect to the hyperbolic measure. This means more concretely that for any continuity set $A \subset X_0(1)$, we have

$$\frac{\#\{\mathfrak{a} \in \mathrm{Cl}_D \mid z_\mathfrak{a} \in A\}}{\#\mathrm{Cl}_D} = \frac{\mathrm{vol}(A)}{\mathrm{vol}(X_0(1))} + o(1),$$

as $-D \to \infty$ (with D a negative fundamental discriminant), where the volume is with respect to the hyperbolic measure $\frac{dxdy}{y^2}$. Similarly, Duke proved that for $D > 0$;

$$\frac{\sum_{\mathfrak{a}\in\mathrm{Cl}_D} \int_{\gamma_\mathfrak{a}\cap A} 1\, ds}{\sum_{\mathfrak{a}\in\mathrm{Cl}_D} \int_{\gamma_\mathfrak{a}} 1\, ds} = \frac{\mathrm{vol}(A)}{\mathrm{vol}(X_0(1))} + o(1),$$

as $D \to \infty$, where ds is the hyperbolic line element (in this case one can actually refine the result to an equidistribution result on the unit tangent bundle of $X_0(1)$).

These questions turn out to be the dual formulations of the classical problem of determining the distribution of integral points of quadratic forms on the appropriate level set as explained in [74, Section 1]. Such problems have been intensively studied by Linnik [62] in the 60's using his ergodic method with great success. The ergodic method is however only able to obtain the equidistribution results under a certain congruence condition (for positive discriminants an ergodic proof *without* the congruence condition has now been given [28]). Duke was the first to remove this condition.

The approach of Duke is to use harmonic analysis; by Weyl's criterion it is enough to show cancelation in the Weyl sums

$$\sum_{\mathfrak{a}\in \mathrm{Cl}_D} \phi(z_{\mathfrak{a}}),$$

where $D < 0$ and ϕ is a Hecke–Maaß cusp form or an Eisenstein series (and similarly for $D > 0$). These Weyl sums turn out to be connected to Fourier coefficients of half-integral Maaß forms as was proved by Maaß [67] (and revisited by Katok and Sarnak [54]). Duke managed to bound the Fourier coefficients of half-integral Maaß forms using a breakthrough of Iwaniec [48].

A slightly different approach is to employ the formula of Waldspurger (2.1.1); if we pick the automorphic representation π to be the one associated to a Hecke–Maaß cusp form of level 1 and let the Hecke character χ be the trivial one, then the automorphic period in (2.1.1) reduces exactly to the Weyl sums studied by Duke. This implies that equidistribution of Heegner points (and cycles) follows from a subconvexity bound for certain twisted L-functions. Furthermore, this approach opened up for a number of different generalizations. First of all, by choosing non-trivial Hecke characters (and using some easy Fourier analysis), one could deduce *sparse equidistribution* for Heegner points associated to subgroups of the class group assuming subconvexity estimates for certain Rankin–Selberg L-functions. Such subconvexity bounds were obtained by Harcos and Michel [38] using the *amplification method* from analytic number theory. A different case of sparse equidistribution is to consider the distribution of Heegner points with level structure and allow the level to change with the discriminant as was carried out by Liu, Masri and Young [63]. All of these applications requires a version of Waldspurger's formula completely explicit in all parameters. Such versions are available due to the work of many people (see [30] and the references therein). Variants of Duke's theorem have been studied extensively [69], [38], [27], [1]. Very recently a generalization was studied by Duke, Imamoḡlu and Tóth [25], which is the starting point for Paper E.

3.2.1 The work of Duke, Imamoḡlu and Tóth

In [25], Duke, Imamoḡlu and Tóth revisit the geometric invariants associated to positive discriminant $D > 0$. They define certain geometric invariants associated to elements of the *narrow* class group Cl_D^+. The narrow class group is defined as the quotient of fractional ideals of K with principle ideals with a generator of positive norm. If there is a unit of norm -1 then $\mathrm{Cl}_D^+ = \mathrm{Cl}_D$ but otherwise the two versions of the class group are different. Associated to $\mathfrak{a} \in \mathrm{Cl}_D^+$ there is a certain hyperbolic orbifold $\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}}$ whose boundary is the Heegner cycle $\gamma_{\mathfrak{a}}$ mentioned above. The surfaces $\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}}$ are very interesting objects and were introduced with the hope of shedding light on the class groups of real quadratic fields.

The authors study the distribution of these surfaces when projected to the modular surface $X_0(1) = \Gamma\backslash\mathbb{H}$, where $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. They manage to show that for any sequence of genera $H_D \subset \mathrm{Cl}_D^+$ (i.e. a coset of $(\mathrm{Cl}_D^+)^2$) and any continuity set $A \subset X_0(1)$, we have

$$\frac{\sum_{\mathfrak{a}\in H_D} \mathrm{vol}(\Gamma_\mathfrak{a}\backslash\mathcal{N}_\mathfrak{a}\cap \Gamma A)}{\sum_{\mathfrak{a}\in H_D} \mathrm{vol}(\Gamma_\mathfrak{a}\backslash\mathcal{N}_\mathfrak{a})} = \frac{\mathrm{vol}(A)}{\mathrm{vol}(X_0(1))} + o(1), \tag{3.2.1}$$

as $D \to \infty$, where the volume is with respect to the measure $\frac{dxdy}{y^2}$. Equidistribution for the surfaces of the whole class group turns out to be trivial since when projected down they cover $X_0(1)$ evenly (i.e. the above is true *without* error term).

By using simple Fourier analysis, the equidistribution statement (3.2.1) follows if one can show cancellation in the Weyl sum for this distribution problem twisted by class group characters. These twisted Weyl sums can be shown to be related to the product of two Fourier coefficients of half-integral Hecke–Maaß cusp forms using a generalization of the work of Katok and Sarnak [54]. These Fourier coefficients have been bounded by Duke and thus (3.2.1) follows.

A general limitation when using the Katok–Sarnak approach to these equidistribution questions is that one can only deal with Weyl sums twisted by genus characters. To deal with twists by general class group characters, one has to use the connection to central values of Rankin–Selberg L-functions as in the work of Waldspurger. In Paper E (joint work with Humphries) we give an adèlic interpretation of the twisted Weyl sums and manage to relate them to L-functions. These Weyl sums are exactly the periods (**P3.1**) above. As an application we obtain a sparse equidistribution result for the hyperbolic orbifolds.

3.3 Arithmetic quantum chaos

We will now turn to the last topic touched upon in this thesis. The point of departure is a classical problem of quantum physics going all the way back to Bohr, Sommerfeld and Einstein; to understand the behavior of quantizations of classical Hamiltonians (the *correspondence principle*). In the case of the hyperbolic arithmetic surface $X_0(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, we will be interested in the quantization of the classical Hamiltonian generating the geodesic flow, which gives rise to the hyperbolic Laplacian $\Delta = -y^2(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2})$. It is known that the geodesic flow is ergodic and thus chaotic, which puts us in the realm of *quantum chaos*. The eigenfunction of the Laplacian are exactly the Hecke–Maaß forms (forgetting the continuous spectrum for the moment), and the problem becomes to understand their behavior as the energy (i.e. eigenvalue) goes to infinity. In particular, the *Quantum Ergodicity Conjecture* (QE) states that for a density 1 subsequence of L^2-normalized eigenfunctions $\{\phi\}$ with Laplace eigenvalues $\{\lambda_\phi\}$, we should have for any sufficiently nice test function $\psi : X_0(1) \to \mathbb{C}$ that

$$\langle \psi, |\phi|^2\rangle \to \frac{3}{\pi}\langle \psi, 1\rangle,$$

as $\lambda_\phi \to \infty$, where $\langle\cdot,\cdot\rangle$ denotes the Petersson inner product. This was resolved by Zeldich [106] in a much more general setting. More generally, the *Random Wave Conjecture* of Berry [5] predicts that the eigenfunctions of the Laplacian should behave like " Gaußian random waves". This has deep implications for the possible localization of the eigenfunctions.

Following these lines of thinking, Rudnick and Sarnak [85] conjectured that for $X_0(1)$ mass equidistribution should actually hold for the *full* sequence of Laplace eigenfunction, which came to be known as the *Quantum Uniqueness Ergodicity Conjecture* (QUE). Secondly, Rudnick and Sarnak conjectured that the sup norms of Hecke–Maaß cusp forms ϕ should satisfy

$$\sup_{z\in C} \|\phi\|_\infty \ll_{\varepsilon,C} t_\phi^\varepsilon$$

for any $\varepsilon > 0$, where $C \subset X_0(1)$ is compact and $\lambda_\phi = 1/4+t_\phi^2$ is the Laplace eigenvalue of ϕ. This is known as the *Sup Norm Conjecture*. Both of these conjectures are in accordance with the Random Wave Conjecture.

The QUE conjecture for $X_0(1)$ was famously proved by Lindenstrauss [61] using ergodic methods (with an additional key input by Soundararajan [92]). The *Sup Norm Conjecture* is still wide open. The best result to date is due to Iwaniec and Sarnak [50] who obtained the bound $\ll_\varepsilon t_\phi^{5/24+\varepsilon}$ improving on the *convexity exponent* 1/4 (which can be obtained easily).

3.3.1 Mass equidistribution for holomorphic forms

The same questions can also be asked for holomorphic cusp forms (although there is no clear physical interpretation in this case). This time the QUE conjecture predicts that for Hecke eigenforms $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$, we have

$$\langle \psi, y^k |f|^2\rangle \to \frac{3}{\pi}\langle \psi, 1\rangle, \tag{3.3.1}$$

as $k \to \infty$ (where we use the normalization $\langle y^k|f|^2, 1\rangle = 1$) . As an important corollary it was shown by Rudnick [84] that QUE for holomorphic forms implies that the zeroes of Hecke eigenforms equidistribute on $X_0(1)$.

QUE for holomorphic forms was proved by a combination of the works of Holowinsky and Soundararajan [41], [93], [42], see also [91] for a nice and more in-depth survey. The starting point for both the approach of Holowinsky and of Soundararajan is Weyl's criterion; in order to conclude QUE for holomorphic forms it is enough to show (3.3.1) for a basis of $L^2(X_0(1))$. Holowinsky considered the generating set consisting of Poincare series defined by

$$P_{h,m}(z) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} e(m\gamma z)h(\mathrm{Im}\,\gamma z),$$

where $h : (0,\infty) \to (0,\infty)$ is smooth with compact support and $m \in \mathbb{Z}$ (for $m = 0$ the Poincare series are known as *incomplete Eisenstein series*). This gives rise to the periods (**P4.2**) defined above. It is a fact observed firstly by Luo and Sarnak [64] that bounding the periods $\langle P_{h,m}, y^k|f|^2\rangle$ reduces to bounding shifted convolution sums of length k of the form $\sum_{n\asymp k} \lambda_f(n)\lambda_f(n+m)$. This was achieved by Holowinsky for most f (even after taking absolute values of the Hecke eigenvalues) using sieve methods.

Soundararajan considered instead the basis consisting of Hecke–Maaß cusp forms $\{\phi\}$, which are exactly the periods (**P4.1**). In this case, it follows from the formula of Watson and Ichino (2.1.2) that the square of the periods $\langle y^k|f|^2, \phi\rangle$ are equal (up to some manageable factors) to the triple convolution L-function $L(f \otimes f \otimes \phi, 1/2)$. Soundararajan succeeded in obtaining what he calls a *weak subconvexity bound* for such L-functions, which again is strong enough to resolve mass equidistribution for most f. As it turns out, the exceptions in the works of Soundararajan and Holowinsky complemented each other, which lead to the full resolution of QUE for holomorphic cusp forms.

In his investigations of quantum chaos, Zelditch [107] introduced the *quantum variance* associated to a compact Riemannian manifold (M, g):

$$\sum_{\lambda_\phi \leq X} \left|\langle \psi, |\phi|^2\rangle - \langle \psi, 1\rangle\right|^2,$$

where $\psi : M \to \mathbb{C}$ is a test function and ϕ are eigenfunctions for the Laplace–Beltrami operator on M with eigenvalues λ_ϕ. Luo and Sarnak [65] studied the analogous quantum variance for holomorphic forms on the modular surface and were able to obtain an asymptotic formula for the quantum variance when averaging over k:

$$\sum_{k \asymp K} \sum_{f \in H_k} L(1, \mathrm{sym}^2 f)\langle \psi_1, y^k|f|^2\rangle \overline{\langle \psi_2, y^k|f|^2\rangle} = B_\omega(\psi_1, \psi_2)K + o_{\psi_1,\psi_2}(K),$$

as $K \to \infty$, where ψ_1, ψ_2 are sufficiently nice cuspidal test functions. The main term B_ω is a very interesting Hermitian form diagonalized by Hecke–Maaß cusp forms ϕ with $B_\omega(\phi, \phi)$ being equal to a constant times $L(\phi, 1/2)$. Interestingly, in their investigations Luo and Sarnak employ both choices of periods (**P4.1**) and (**P4.2**). As a surprising corollary, they deduce the deep fact that $L(\phi, 1/2) \geq 0$.

3.3.2 The case of Eisenstein series

One can ask the same questions as above for the non-holomorphic Eisenstein series;

$$E(z, s) = \sum_{\gamma \in \Gamma_\infty \backslash \mathrm{SL}_2(\mathbb{Z})} \mathrm{Im}(\gamma z)^s,$$

on the line of symmetry $s = 1/2 + it$, which corresponds to the continuous spectrum of the Laplacian. The statements require suitable modifications. The case of QUE for Eisenstein series was resolved early on by Luo and Sarnak [64] who proved that

$$\frac{1}{2\log t}\langle \psi, |E(\cdot, 1/2 + it)|^2\rangle \to \frac{3}{\pi}\langle \psi, 1\rangle,$$

as $t \to \infty$ for ψ sufficiently nice. In this case the periods that appear are much easier to handle than in the Maaß case.

The problem of sup norm bounds for Eisenstein series has also been studied. In this case the method of Iwaniec and Sarnak yields the bound

$$\sup_{z \in C} |E(z, 1/2 + it)| \ll_{C,\varepsilon} t^{5/12+\varepsilon}, \tag{3.3.2}$$

with $C \subset \mathbb{H}$ compact. Note that in this case, the convexity bound is $\ll_{C,\varepsilon} t^{1/2+\varepsilon}$. Young [103] using a slight modification of the method of Iwaniec and Sarnak, obtained the improved exponent $3/8 + \varepsilon$ in the Eisenstein case. Blomer [9] realized that the classical work of Titchmarsh [96] using exponential sum methods could be applied and obtained the Weyl type exponent $1/3 + \varepsilon$. In Paper F, we upgrade Blomer's bound to a *uniform* sup norm bound with an explicit dependence of C.

It was observed by Sarnak [87, (4.19)] that the sup norm problem has applications to subconvexity bounds for L-functions using the following formula due to Hecke:

$$L_K(s,\chi) = \frac{2^{s+1}\zeta(2s)|D|^{-s/2}}{\omega_D} \sum_{\mathfrak{a}\in \mathrm{Cl}_D} \chi(\mathfrak{a}) E(z_{\mathfrak{a}}, s), \tag{3.3.3}$$

where Cl_D is the class group of discriminant $D < 0$ and $\omega_D \in \{2, 4, 6\}$ (there is a similarly formula for real quadratic fields due to Siegel). The right-hand side of (3.3.3) is exactly the periods (**P2.1**) above. In Paper F, we use this to obtain (hybrid) subconvexity bounds for class group L-functions.

3.3.3 Mass distribution at small scales

It is natural to ask whether mass equidistribution also holds at smaller scales. This means that we want to study mass equidistribution on shrinking sets or (what amounts to the same thing) when we allow our test function ψ to vary with the spectral parameter. This problem has been studied in the physics literature and it seems to suggest that that mass equidistribution should hold all the way down to the scale of the *de Broglie wavelength* (see [40]). This means that we expect equidistribution to hold when we shrink our test function at a rate above $1/\sqrt{\lambda}$, where λ is the eigenvalue. Small scale equidistribution has been studied extensively in many aspect [102], [44]. In particular, in the holomorphic case Lester, Matomäki, and Radziwiłł [60] study QUE for shrinking sets and the distribution of zeroes high in the cusp.

In Paper G we contribute to the question of small scale mass distribution in the holomorphic case and consider the setting of shrinking "balls around infinity". Among other things we calculate the quantum variance when the test function ψ is "squeezed" up towards the cusp ∞. In this setting the connection to L-functions seizes to exist for the periods (**P4.1**). Instead we consider certain "squeezed periods" of the type (**P4.2**) corresponding to the Poincaré series basis, which again are related to shifted convolution sums. The results follows from solving this shifted convolution problem.

CHAPTER 4

SUMMARIES OF PAPERS

Paper A: *Central values of additive twists of cuspidal L-functions*

In this paper we study the distribution of central values of additive twists of holomorphic cusp forms. This is a generalization of the results of Petridis and Risager [81] on the arithmetic statistics of modular symbols to higher weight holomorphic cusp forms.

To explain our results, let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a cusp form of even weight k and level q. Then we define the *additive twist by* $x \in \mathbb{R}$ as the following Dirichlet series which converges absolutely for $\operatorname{Re} s > (k+1)/2$:

$$L(f,x,s) := \sum_{n\geq 1} \frac{a_f(n)e(nx)}{n^s},$$

where $a_f(n)$ are the Fourier coefficients of f. When $x \in \mathbb{Q}$, the additive twist $L(f,x,s)$ admits analytic continuation satisfying a functional equation relating $s \leftrightarrow k-s$.

We will study the distribution of the central values $s = k/2$ as x varies through rational numbers ordered by their denominators. More precisely, we consider $L(f,\cdot,k/2)$ as a random variable defined on the outcome space

$$T(X) := \{a/c \mid 0 < a < c \leq X, (a,c) = 1, q|c\}$$

endowed with the uniform probability measure. This defines a sequence of random variable, which we show (when appropriately normalized) converges in distribution to a standard normal.

Theorem 4.0.1 (Paper A, Theorem 1.1)**.** *Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a cusp form of even weight k and level q. Then for any fixed box $\Omega \subset \mathbb{C}$, we have*

$$\begin{aligned}\mathbb{P}_{T(X)}\left(\frac{L(f,a/c,k/2)}{(C_f \log c)^{1/2}} \in \Omega\right) :=& \frac{\#\left\{a/c \in T(X) \mid \frac{L(f,a/c,k/2)}{(C_f \log c)^{1/2}} \in \Omega\right\}}{\#T(X)} \\ =& \mathbb{P}(\mathcal{N}_\mathbb{C}(0,1) \in \Omega) + o(1),\end{aligned} \tag{4.0.1}$$

as $X \to \infty$, where $\mathcal{N}_{\mathbb{C}}(0,1)$ denotes the standard complex normal distribution and the variance is given by

$$C_f := \frac{(4\pi)^k \|f\|^2}{(k-1)!\,\mathrm{vol}(\Gamma_0(q))}, \tag{4.0.2}$$

with $\|f\|$ the Petersson-norm of f and $\mathrm{vol}(\Gamma_0(q))$ the hyperbolic volume of $\Gamma_0(q)\backslash\mathbb{H}$.
(Here $\mathbb{P}(\mathcal{N}_{\mathbb{C}}(0,1)\in\Omega)$ denotes the probability of the event $\mathcal{N}_{\mathbb{C}}(0,1)\in\Omega$.)

This result has applications to automorphic L-functions; the proof proceeds by the method of moments and as a by-product we obtain an asymptotic expression for high moments of additive twists. By combining this with the Birch–Stevens formula, we obtain a calculation of a certain "wide" moment of automorphic L-functions.

More precisely given a newform $f \in \mathcal{S}_k(\Gamma_0(q))$ of weight k and level q and a primitive Dirichlet character χ of conductor co-prime to q, we define the *multiplicatively twisted L-function of f*;

$$L(f\otimes\chi, s) := \sum_{n\geq 1} \frac{\lambda_f(n)\chi(n)}{n^s}, \quad \mathrm{Re}\, s > 1,$$

where $\lambda_f(n)$ denotes the nth Hecke eigenvalue of f. This series admits analytic continuation to the entire complex plane satisfying a functional equation $s \leftrightarrow 1-s$. Because of the co-primality condition the above Dirichlet series defines (the finite part of) the L-function of the automorphic representation $\pi_f \otimes \chi$. We then get the following asymptotic formula of a "wide" family of L-functions.

Corollary 4.0.2 (Paper A, Corollary 1.9). *Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a newform of even weight k and level q and n a non-negative integer. Then we have for all $\varepsilon > 0$*

$$\sum_{\substack{0<c\leq X,\\ (c,q)=1}} \frac{1}{\varphi(c)^{2n-1}} \sum_{\substack{\chi_1,\ldots,\chi_{2n} \bmod c,\\ \chi_1\cdots\chi_{2n}=\chi_{\mathrm{principal}}}} \varepsilon_{\chi_1,\ldots,\chi_n} \prod_{i=1}^{2n} \nu(f,\chi_i^*, c/c(\chi_i)) L(f\otimes\chi_i^*, 1/2)$$
$$= P_n(\log X)X^2 + O_\varepsilon(X^{4/3+\varepsilon}), \tag{4.0.3}$$

where $\chi^* \bmod c(\chi)$ denotes the primitive character inducing χ, $\chi_{\mathrm{principal}}$ is the principal character $\bmod c$, P_n is a certain degree n polynomial with leading coefficient

$$\frac{q(2C_f)^n\, n!}{\pi\,\mathrm{vol}(\Gamma_0(q))},$$

$\varepsilon_{\chi_1,\ldots,\chi_n} = \chi_1(-1)\cdots\chi_n(-1)$ is a sign, and ν is an arithmetic weight given by

$$\nu(f,\chi,n) := \tau(\overline{\chi}) \sum_{\substack{n_1n_2n_3=n,\\ (n_1,q)=1}} \chi(n_1)\mu(n_1)\overline{\chi}(n_2)\mu(n_2)\lambda_f(n_3)n_3^{1/2}. \tag{4.0.4}$$

Actually, the results apply more generally to any discrete and cofinite subgroup Γ of $\mathrm{SL}_2(\mathbb{R})$ with cusps. The proof has as a key input the analytic properties of resolvent operators and is inspired by the approach introduced by Petridis and Risager [78]. There are however some substantial new difficulties which arises for higher weights, due essentially to the fact that $\Gamma \ni \gamma \mapsto L(f, \gamma\infty, k/2)$ is *not* additive for higher weight. This leads to the analysis being much more involved and essential new ideas were needed.

Remark 4.0.3. One natural question to ask (which has been asked on many occasions) is whether one can generalize the above results to additive twists of Maaß forms. There are some serious obstacles when trying to generalize the approach taken in this paper to the Maaß case, since holomorphicity plays a crucial role. We do however have some ideas on how to use the dynamical methods of Bettin and Drappeau [8] to the deal with the Maaß case as well. This is work in progress (with Drappeau).

Remark 4.0.4. As mentioned above, we rely on the analytic properties of Eisenstein series to calculate the moments of the additive twists, but there is another possible approach; Bettin [7] succeeded in calculating all moments of the Esterman function (which is an analogue of the additive twists of L-functions of holomorphic cusp forms) using a classical approximate functional equation approach. It would be interesting to see whether this can be done for additive twists of cuspidal automorphic forms as well. This would give another approach to the Maaß case and maybe allow one to obtain results for function fields as well.

Paper B: *On the distribution of periods of holomorphic cusp forms and zeroes of period polynomials*

In this paper we study the distribution of *period polynomials* attached to higher weight cusp forms. Let $f \in \mathcal{S}_k(\Gamma_0(N))$ with $k \geq 4$ even. Then we define the period polynomials of f as

$$r_{f,\gamma}(X) := \int_{\gamma\infty}^{\infty} f(z)(z-X)^{k-2} dz,$$

for $\gamma \in \Gamma_0(N)$. These polynomials are the natural cohomological generalization of modular symbols to higher weight because of the Eichler–Shimura isomorphism. In this paper we study the properties of $r_{\gamma,f}$ as γ varies and f is fixed. In analogy with Paper A, we order the matrices γ by the size of their lower left entry. First of all we study the location of the zeroes and obtain a result saying that the zeroes all cluster together close to each other.

Theorem 4.0.5 (Paper B, Theorem 1.6)**.** *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a newform of even weight $k \geq 6$ and level N. Then $r_{f,\gamma}$ is a polynomial of degree $k-2$ for any $\gamma \in \Gamma_0(N)$. Furthermore, all zeroes x_0 of $r_{f,\gamma}$ satisfy*

$$x_0 = a/c + O_k((|a/c|+1)^{(k-4)/(k-2)} c^{-2/(k-2)}),$$

where a, c are the entries in the left column of γ (i.e. $\gamma\infty = a/c$).

This result complements results of Jin, Ma, Ono and Soundararajan [52], which handle the case of $\gamma = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ fixed and f varying.

Inspired by conjectures of Mazur and Rubin [73] on modular symbols, we also study the joint distribution of the coefficients of the period polynomial. The distribution is highly degenerate as it is the transformation of two independent random variables uniformly distributed on $\mathbb{R}/\mathbb{Z}$, and thus in particular, very far from normal. The transformation involved is a special value of additive twists of the L-function of f.

More precisely we put

$$\Omega_c := \{a/c \in \mathbb{Q} \mid a, c \in \mathbb{Z}_{\geq 0}, (a,c) = 1, 0 \leq a < c\}, \tag{4.0.5}$$

and

$$\begin{aligned} u_f(a/c) &= (u_{f,0}(a/c), u_{f,1}(a/c), \ldots, u_{f,k-2}(a/c)) \\ &:= \left(\int_{a/c}^{\infty} f(z)dz, \int_{a/c}^{\infty} f(z)zdz, \ldots, \int_{a/c}^{\infty} f(z)z^{k-2}dz \right)^T \in \mathbb{C}^{k-1}, \end{aligned} \tag{4.0.6}$$

which one checks are the coefficients of the period polynomial associated to γ (up to explicit constants), where $\gamma\infty = a/c$. Then we get the following result regarding the distribution.

Theorem 4.0.6 (Paper B, Theorem 1.1)**.** *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a cusp form of even weight $k \geq 4$ and level N. Then we have for any fixed box $A \subset \mathbb{C}^{k-1}$ that*

$$\frac{\#\{\frac{a}{c} \in \Omega_c \mid \frac{u_f(a/c)}{C_k c^{k-2}} \in A\}}{\#\Omega_c} = \mathbb{P}\left(F(Y,Z) \in A\right) + o(1)$$

as $c \to \infty$ with $N|c$, where Y, Z are two independent random variables both distributed uniformly on $[0,1)$, $F : [0,1) \times [0,1) \to \mathbb{C}^{k-1}$ is given by

$$F(y,z) := L(f, y, k-1)\left(1, z, \ldots, z^{k-2}\right)^T,$$

and $C_k = \frac{i\Gamma(k-1)}{(2\pi)^{k-2}}$.

(Here $\mathbb{P}\left(F(Y,Z) \in A\right)$ denotes the probability of the event $F(Y,Z) \in A$).

We also obtain results for general cofinite, discrete subgroups of Γ, but then we have to take an extra average over the lower left entries of the matrices γ or equivalently over the denominators of the twists a/c. The idea of the proof is to write the coefficients of the period polynomial as a linear combination of special values of additive twists $L(f, a/c, l)$ with $l = 1, \ldots k-1$. Now the results follows from a rather careful analysis of the analytic properties of additive twists and uniform bounds for Kloosterman sums.

Paper C: *A note on additive twists, reciprocity laws and quantum modular forms*

In this paper we study generalizations of certain reciprocity laws proved by Conrey in an unpublished paper [18] and connections to *quantum modular forms*. We study the case of holomorphic cusp forms of even weight and obtain a reciprocity relation of the following kind, for $q \neq l$ primes:

$$\sum_{\substack{\chi \bmod q,\\ \chi \text{ primitive}}} \tau(\overline{\chi})L(f,\chi,k/2)\chi(l) \rightsquigarrow \sum_{\substack{\chi \bmod l,\\ \chi \text{ primitive}}} \tau(\overline{\chi})L(f,\chi,k/2)\chi(-q), \tag{4.0.7}$$

where $L(f,\chi,s)$ denotes the analytic continuation of $\sum_{n\geq 1} a_f(n)\chi(n)n^{-s}$ with $a_f(n)$ the Fourier coefficients of $f \in \mathcal{S}_k(\Gamma_0(1))$, $\tau(\overline{\chi})$ is a Gauß sum and χ runs through primitive Dirichlet characters of conductor, respectively q and l.

We also obtain results when l, q are replaced by arbitrary integers and where we allow f to have non-trivial level. In this case the result is not as clean, but takes the following form.

Theorem 4.0.7 (Paper C, Theorem 2.1)**.** *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a primitive newform of even weight k and level N with eigenvalue ω_f under the Fricke involution. Then we have the following reciprocity relation for any pair of integers $0 < l < q$ with $(q, Nl) = 1$;*

$$\begin{aligned} &\frac{1}{\varphi(q)} \sum_{\chi \bmod q} \nu(f,\chi^*,q/c(\chi))L(f,\chi^*,k/2)\chi(l) \\ &\qquad - \frac{\omega_f}{\varphi(lN)} \sum_{\chi \bmod lN} \nu(f,\chi^*,lN/c(\chi))L(f,\chi^*,k/2)\chi(-q) \\ &\qquad\qquad = L(f,k/2) + O_f(l/q), \end{aligned} \tag{4.0.8}$$

where $\chi^* \bmod c(\chi)$ denotes the primitive character inducing χ, $L(f,\chi,s)$ is as above and the arithmetic weights ν are given by

$$\nu(f,\chi,n) := \tau(\overline{\chi}) \sum_{\substack{n_1n_2n_3=n,\\ (n_1,N)=1}} \chi(n_1)\mu(n_1)\overline{\chi}(n_2)\mu(n_2)a_f(n_3)n_3^{1-k/2}.$$

The proof is inspired by the approach of Bettin in [6], and begins by relating the twisted sums in the reciprocity relation to additive twists;

$$L(f,a/q,k/2) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \nu(f,\chi^*,q/c(\chi))L(f,\chi^*,k/2)\chi(a), \tag{4.0.9}$$

with $\chi^* \bmod c(\chi)$ and ν as above. Now the proof boils down to the fact that $L(f,\cdot,k/2)$ as a function on $\mathbb{Q}$ defines a *quantum modular form* in the sense of Zagier [105]. This is a result of independent interest.

Theorem 4.0.8 (Paper B, Theorem 4.4)**.** *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a cusp form of even weight k and level N. Then the map $\mathbb{Q} \to \mathbb{C}$ defined by*

$$\mathbb{Q} \ni r \mapsto L(f,r,k/2)$$

is a quantum modular form of weight zero for $\Gamma_0(N)$. More precisely, for $\gamma \in \Gamma_0(N)$ and $r \in \mathbb{Q}$ with $\gamma r \neq \infty$ and $\gamma\infty \neq \infty$, we have

$$\begin{aligned} &L(f,\gamma r,k/2)-L(f,r,k/2)\\ &= L(f,\gamma\infty,k/2) + \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{(c^{-1}j(\gamma,r))^j}\frac{(-2\pi i)^{-j}\Gamma(k/2+j)}{\Gamma(k/2)}L(f,r,k/2+j)\\ &\quad + \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{(cj(\gamma,r))^j}\frac{(-2\pi i)^{j}\Gamma(k/2-j)}{\Gamma(k/2)}L(f,\gamma\infty,k/2-j), \end{aligned} \tag{4.0.10}$$

where c is the lower-left entry of γ.

If we put $N = 1$ and $\gamma = S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, then the left hand encodes exactly the difference between the left- and right hand side of the reciprocity relation (4.0.7). Now quantum modularity implies that this difference is archimedean in nature and thus it can be estimated in terms of the archimedean size of q/l, which yields the wanted reciprocity relation. For general level, we use quantum modularity with respect to the Fricke involution defined as;

$$Wf(z) := N^{-k/2}z^{-k}f(-1/(Nz)). \tag{4.0.11}$$

Remark 4.0.9. A natural question is whether one can extend the results to additive twists of Maaß forms. We have work in progress that hopes to answer this question in the affirmative. Secondly, we observe that the above reciprocity formula corresponds to "GL_2 twisted by GL_1". It would be interesting to obtain reciprocity formulas in the case $\mathrm{GL}_n \times \mathrm{GL}_1$ for some $n \neq 2$. Especially the simplest case $n = 1$ begs for an answer.

Paper D: *Residual equidistribution of modular symbols and cohomology classes for quotients of hyperbolic n-space*

In this paper we study the residual distribution of modular symbols and more generally the residual distribution of cohomology classes for quotients of higher dimensional hyperbolic space. The question of residual distribution of modular symbols seems to appear for the first time in unpublished work of Mazur and Rubin [71].

For $f \in \mathcal{S}_2(\Gamma_0(N))$ a Hecke-eigenform, we define the modular symbols map as

$$\langle r, f\rangle := 2\pi i\int_r^\infty f(z)dz,$$

for $r \in \mathbb{Q}$. Now if we fix a prime p, then it is a fact that there exists numbers $\Omega^\pm$ such that

$$\mathfrak{m}^\pm_{f,p}(a/q) = \frac{\langle a/q, f\rangle \pm \langle -a/q, f\rangle}{\Omega^\pm} \in \mathbb{Z}$$

for all $a/q \in \mathbb{Q}$ with $N|q$, but not all values are divisible by p. In this setting it is conjectured by Mazur and Rubin that the values of $\mathfrak{m}^\pm_{f,p}$ modulo p on the sets $\{a/q \mid 0 < a < q, (a,q) = 1\}$ tend to the uniform distribution on $\mathbb{Z}/p\mathbb{Z}$ as $q \to \infty$.

This conjecture was settled by Lee and Sun [59] using dynamical methods after taking an extra average over q. In this paper we give a simple automorphic proof of this statement, which furthermore allows for a number of refinements and generalizations. In some very special cases we are also able to resolve the full conjecture without the extra average.

Theorem 4.0.10 (Paper D, Theorem 3.1)**.** *Let $N \geq 5$ and $p \mid N-1$ be odd primes. Then there exists a new form $f \in \mathcal{S}_2(\Gamma_0(N))$ of weight 2 and level N such that the values of $\mathfrak{m}^+_{f,p}$ on $\{\frac{a}{q} \mid (a,q)=1, 0<a<q\}$ equidistribute modulo p as $q \to \infty$ with $N|q$.*

To state the result we obtain in general, let $f_1, \dots, f_d$ be a Hecke basis for $\mathcal{S}_2(\Gamma_0(N))$ and consider the following random variable

$$\mathfrak{m}_{N,p}(a/q) := (\mathfrak{m}^+_{f_1,p}(a/q), \dots, \mathfrak{m}^-_{f_d,p}(a/q), a/q)$$

defined on the outcome space $\Omega_Q = \{a/q \mid 0 < a < q \leq Q, (a,q)=1, N|q\}$ (endowed with the uniform probability measure). Then we have the following *simultaneous* distribution result.

Theorem 4.0.11 (Paper D, Theorem 1.2)**.** *The random variables $\mathfrak{m}_{N,p}$ defined on the outcome spaces Ω_Q are asymptotically uniformly distribution on $(\mathbb{Z}/p\mathbb{Z})^{2d} \times (\mathbb{R}/\mathbb{Z})$ as $Q \to \infty$.*

Furthermore, we obtain results for higher dimensional hyperbolic spaces. So let $\Gamma \subset \mathrm{SO}(n+1,1)$ be a discrete subgroup such that $\Gamma\backslash\mathbb{H}^{n+1}$ has finite volume and a cusp at ∞ (here we use the usual action of $\mathrm{SO}(n+1,1)$ on $\mathbb{H}^{n+1}$). Then we will study the distribution of unitary characters of Γ, which is exactly computed by the cohomology group $H^1(\Gamma, \mathbb{R}/\mathbb{Z})$. We will now describe a slightly simplified case of our most general result.

Fix a prime p and let $\omega_1, \dots, \omega_d$ be an $\mathbb{F}_p$-basis for $H^1_{\Gamma_\infty}(\Gamma, \mathbb{Z}/p\mathbb{Z})$ (corresponding to unitary characters of order p trivial on the stabilizer Γ_∞ of ∞ which we assume for simplicity to be exactly the translation by $\mathbb{Z}^n$). As in the case $n=1$, we can associate an invariant $\gamma\infty \in \mathbb{R}^n \cup \{\infty\}/\mathbb{Z}^n$ using the action of Γ on the cusps of $\Gamma\backslash\mathbb{H}^{n+1}$. Now given $X > 0$, we consider

$$\omega(\gamma) := (\omega_1(\gamma), \dots, \omega_d(\gamma), \gamma\infty)$$

as a random variable with values in $(\mathbb{Z}/p\mathbb{Z})^d \times (\mathbb{R}^n/\mathbb{Z}^n)$ defined on the outcome space

$$T_\Gamma(X) = \{\gamma \in \Gamma_\infty\backslash\Gamma/\Gamma_\infty \mid 0 < c_\gamma \leq X\},$$

endowed with the uniform probability measure. Here c_γ is a certain invariant of γ, which for $n=1$ reduced to the absolute value of the lower-left entry of γ (actually this remains true using the *Vahlens model* of $\mathrm{SO}(n+1,1)$). Then we obtain the following result.

Theorem 4.0.12 (Paper D, Theorem 1.4)**.** *The random variables ω defined on the outcome spaces $T_\Gamma(X)$ are asymptotically uniformly distributed on $(\mathbb{Z}/p\mathbb{Z})^d \times (\mathbb{R}^n/\mathbb{Z}^n)$*

as $X \to \infty$. This means in concrete terms that for any residue $a \in (\mathbb{Z}/p\mathbb{Z})^d$ and $B \subset \mathbb{R}^n/\Lambda$, we have

$$\frac{\#\left\{\gamma \in T_\Gamma(X) \mid (\omega_1(\gamma), \dots, \omega_d(\gamma)) = a, \gamma\infty \in B\right\}}{\# T_\Gamma(X)} = \frac{1}{p^d} \cdot |B| + o(1)$$

as $X \to \infty$.

To phrase the results we obtain for general elements of $H^1_{\Gamma_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ requires some more notation, and we will refer to the paper itself for details. The proof relies on realizing the generating series for the Weyl sums for the distribution question above as the Fourier coefficients of twisted Eisenstein series. This idea can be seen as a discrete analogue of the method pioneered by Petridis and Risager in [78]. We also obtain results when we order the matrices of Γ by trace, in which case the proof proceeds by an application of an appropriate trace formula.

Paper E: *Sparse equidistribution of hyperbolic orbifolds*

In this paper we study a certain refinement of a recent distribution result on geometric invariants of real quadratic fields due to Duke, Imamoḡlu and Tóth [25]. The set-up is as follows. Let $E/\mathbb{Q}$ be a real quadratic field of discriminant $D > 0$ with narrow class group Cl_D^+. Associated to $\mathfrak{a} \in \mathrm{Cl}_D^+$ there is an oriented, closed geodesic $\gamma_\mathfrak{a}$ on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, which we will call the *Heegner cycle* of $\mathfrak{a}$. In [25] a hyperbolic orbifold $\Gamma_\mathfrak{a}\backslash\mathcal{N}_\mathfrak{a}$ is defined which boundary is exactly $\gamma_\mathfrak{a}$. It is shown that if we choose a genus $H_D \subset \mathrm{Cl}_D^+$ (i.e. a coset of $(\mathrm{Cl}_D^+)^2$) for each $D > 0$, then the orbifolds $\{\Gamma_\mathfrak{a}\backslash\mathcal{N}_\mathfrak{a} \mid \mathfrak{a} \in H_D\}$ equidistribute (with respect to the hyperbolic measure) as $D \to \infty$ when projected to the modular surface $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$.

To prove this, we use Weyl's criterion; it is enough to show cancelation in the Weyl sums

$$\sum_{\mathfrak{a} \in H_D} \int_{\Gamma_\mathfrak{a}\backslash\mathcal{N}_\mathfrak{a}} \phi(z) \frac{dxdy}{y^2}$$

where ϕ is either a Hecke–Maaß form or $E(z, 1/2 + it)$. By some simple Fourier analysis it suffices to show cancelation in Weyl sums for all of Cl_D^+, but twisted by genus characters. The proof in [25] proceeds by relating the relevant Weyl sums to Fourier coefficients of half-integral Maaß forms using an extension of the methods of Katok and Sarnak [54]. It is a general limitation to the Katok–Sarnak approach that one can only treat twists by genus characters.

The main insight of this paper is that one can relate Weyl sums twisted by general class group characters to central values of Rankin–Selberg L-functions in the spirit of Waldspurger. The proof of this proceeds by relating the twisted Weyl sums to certain automorphic periods, which can then be related to Rankin–Selberg L-functions using the refinement of Waldspurger's work due to Martin and Whitehouse [69]. This requires a careful local analysis at the archimedean place. This allows us to prove the following refinement of the results of [25].

Theorem 4.0.13 (Paper E, Theorem 1.1). *Fix $\delta \geq 0$. For each positive fundamental discriminant D choose a coset CH with $H = H_D$ a subgroup of Cl_D^+ satisfying $\#H \gg D^{-\delta} h_D^+$ and $C \in \mathrm{Cl}_D^+$. Then for each fixed continuity set $B \subset \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$,*

$$\frac{\sum_{\mathfrak{a}\in CH} \mathrm{vol}(\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}} \cap \mathrm{SL}_2(\mathbb{Z})B)}{\sum_{\mathfrak{a}\in CH} \mathrm{vol}(\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}})} = \frac{\mathrm{vol}(B)}{\pi/3} + o_{B,\delta}(1)$$

as D tends to infinity for $\delta < \frac{625}{3309568} \approx 0.0001888$ unconditionally and for $\delta < \frac{1}{4}$ assuming the generalised Lindelöf hypothesis. Here the volume is with respect to the hyperbolic measure $\frac{dxdy}{y^2}$.

Remark 4.0.14. Our approach opens up for even more refinements. If one could define a level q version of the hyperbolic orbifolds $\Gamma_{\mathfrak{a}}\backslash\mathcal{N}_{\mathfrak{a}}$ and prove the required properties, then one could prove sparse equidistribution in the level aspect as in [63]. This is work in progress (with Humphries).

Paper F: *Hybrid subconvexity for class group L-functions and uniform sup norm bounds of Eisenstein series*

This paper is concerned with what we call the *uniform sup norm problem*, which asks for bounds of the type

$$\sup_{z\in C} E(z, 1/2 + it)| \ll_{C,\varepsilon} (|t| + 1)^{\theta+\varepsilon}, \tag{4.0.12}$$

with an explicit dependence on $C \subset \mathbb{H}$. If we assume that $y \gg 1$, then it follows quite easily from work of Young [103] that we have the bound $y^{1/2}(|t| + 1)^{3/8+\varepsilon}$. The main technical contribution of the paper is the following improvement.

Theorem 4.0.15 (Paper F, Theorem 1.6). *For $z \in \mathcal{F}$, the standard fundamental domain for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$, we have*

$$E(z, 1/2 + it) \ll_{\varepsilon} y^{1/2}(|t| + 1)^{1/3+\varepsilon}, \tag{4.0.13}$$

for any $\varepsilon > 0$.

Without the explicit dependence on y, this result is due to Blomer [9] building on work of Titchmarsh [96]. Note that the factor $y^{1/2}$ is the same as what you get from Young's method, and this is actually optimal: we prove that any uniform sup norm bound for Eisenstein series of the form $y^{\delta}t^{\theta}$ has to satisfy $\delta \geq 1/2$, since for y very large the constant term in the Fourier expansion becomes dominant.

The motivation for studying the uniform sup norm problem is that it has applications to subconvexity of class group L-functions. The connection between the sup norm problem and subconvexity was noticed along time ago by Sarnak [87, (4.19)], but our results together with the recent work of Hu and Saha [43] seem to be the first times this connected has been utilized. The idea in our paper is to use a classical formula due to Hecke to transfer a sup norm estimate to a subconvexity bound. To do this we need to prove certain upper bounds for sums over Heegner points (we avoid any use of *Duke's Theorem* and actually the versions available in the literature do not suffice for our purposes), and from this we obtain the following.

Corollary 4.0.16 (Paper F, Corollary 1.7)**.** *Let $K/\mathbb{Q}$ be a quadratic extension (real or imaginary) of discriminant D and χ a (wide) class group character of K. Then*

$$L_K(1/2+it,\chi) \ll_\varepsilon |D|^{1/4+\varepsilon}\,(|t|+1)^{1/3+\varepsilon}, \tag{4.0.14}$$

and

$$\sum_{\chi\in\widehat{\mathrm{Cl}(K)}} |L_K(1/2+it,\chi)|^2 \ll_\varepsilon |D|^{1/2+\varepsilon}\,(|t|+1)^{2/3+\varepsilon}, \tag{4.0.15}$$

for any $\varepsilon>0$, where $\widehat{\mathrm{Cl}(K)}$ denotes the class group characters of K.

This hybrid subconvexity bound beats the current record due to Wu [100] in certain regimes of t and D. Combining the two yields the following "state of the art" hybrid subconvexity bound.

Corollary 4.0.17 (Paper F, Corollary 1.10)**.** *Let $K/\mathbb{Q}$ be a quadratic extension of discriminant D and χ a (wide) class group character of K. Then we have*

$$L_K(1/2+it,\chi) \ll_\varepsilon \begin{cases} |D|^{1/4+\varepsilon}\,(|t|+1)^{1/3+\varepsilon}, & \text{for } t>|D|^{3/74} \\ \left(|D|^{1/4}\,(|t|+1)^{1/2}\right)^{1-1/40}, & \text{for } t\le |D|^{3/74} \end{cases}, \tag{4.0.16}$$

for any $\varepsilon>0$.

Paper G: *Small scale equidistribution of Hecke eigenforms at infinity*

In this paper we study small scale mass distribution of holomorphic cusp forms. More precisely, we investigate the distribution of mass of holomorphic cusp forms high in the cusp.

Let $f\in\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ be a cusp form of weight k and level 1, and let ψ be a test function. Then we are interested in whether we have

$$\langle M_{(k-1)^\theta}\psi, y^k|f|^2\rangle \sim \frac{3}{\pi}\langle M_{(k-1)^\theta}\psi, 1\rangle$$

as $k\to\infty$, for some fixed $\theta>0$ where M_H is a certain *squeezing operator* defined by

$$M_H\psi(x+iy) := \psi(x+iy/H), \quad x+iy\in\mathcal{F},$$

where $\mathcal{F}$ is the standard fundamental domain for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. When $\theta=0$, this has been proved by Holowinsky and Soundararajan [42] and is known as QUE for holomorphic forms. Physicists expect that mass equidistribution should hold for all $\theta<1$, with $\theta=1$ corresponding to the *de Broglie wavelength*.

Our first result is that for $\theta\ge 1$ mass equidistribution fails, as predicted. Secondly we show that for test functions ψ supported on $B:=\{x+iy\in\mathcal{F}\mid y>1\}$ and sufficiently nice, mass equidistribution holds on average over $f\in\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ and k all the way down to the Planck scale (that is above the scale of the *de Broglie wavelength*).

Theorem 4.0.18 (Paper G, Theorem 1.2)**.** *Let $0 < \theta < 1$ and ψ. Then for sufficiently nice test function ψ supported in B, we have*

$$\begin{aligned}\sum_{k \asymp K} \sum_{f \in H_k} L(1, \mathrm{sym}^2 f) |\langle M_{(k-1)^\theta}\psi, y^k |f|^2\rangle - \frac{3}{\pi}\langle M_{(k-1)^\theta}\psi, 1\rangle|^2 \\ = o\left(\sum_{k \asymp K} \sum_{f \in H_k} L(1, \mathrm{sym}^2 f) |\langle M_{(k-1)^\theta}\psi, 1\rangle|^2\right), \qquad (4.0.17)\end{aligned}$$

where H_k is a Hecke basis for $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$.

The left hand side (4.0.17) is known as the *quantum variance* and was introduced by Zeldich [106]. If we restrict further to test function ψ, which are orthogonal to 1, then we can actually obtain an asymptotic formula for the quantum variance.

Theorem 4.0.19 (Paper G, Theorem 1.3(i))**.** *Let $0 < \theta < 1$. There exists a Hermitian form B_θ defined on sufficiently nice test functions supported on B and orthogonal to 1 such that*

$$\begin{aligned}\sum_{k \asymp K} \sum_{f \in H_k} L(1, \mathrm{sym}^2 f) |\langle M_{(k-1)^\theta}\psi, y^k |f|^2\rangle|^2 \\ \sim B_\theta(\psi, \psi)\left(\int u(y) y^{-\theta} dy\right) K^{1-\theta},\end{aligned}$$

for ψ as above.

For $\theta = 0$, this was proved by Luo and Sarnak [65]. The Hermitian forms B_θ satisfy a very interesting "phase transition" at $\theta = 1/2$. We summarize the properties of B_θ as follows.

Theorem 4.0.20 (Paper G, Theorem 1.3(ii)-(iv))**.**

1. *The Hermitian forms B_θ have three different regimes in the sense that B_θ is constant on each of the three intervals $0 < \theta < 1/2$, $\theta = 1/2$ and $1/2 < \theta < 1$.*

2. *The decomposition of test function into the cuspidal and the Eisenstein part is orthogonal with respect to B_θ for all $0 < \theta < 1$. Furthermore B_θ restricted to the Eisenstein part is independent of θ.*

3. *The Hermitian forms B_θ can be extended continuously to the larger set $1_B C_{0,0}^\infty(M)$. If ϕ_i are Hecke–Maaß cusp forms with eigenvalue $s_i(1-s_i)$, then the Hermitian form satisfies $B_\theta(1_B\phi_1, 1_B\phi_2) = 0$ unless ϕ_1, ϕ_2 are both even. If ϕ_i are both even, then*

$$B_\theta(1_B\phi_1, 1_B\phi_2) = 4\pi \sum_{m,n \geq 1} \frac{\tau_1((m,n))\lambda_{\phi_1}(m)\lambda_{\phi_2}(n)}{(mn)^{1/2}} I_\theta^{s_1,s_2}(m,n),$$

where

$$I_\theta^{s_1,s_2}(m,n) = \int_{\max(m,n)}^\infty K_{s_1-1/2}(2\pi y)\overline{K_{s_2-1/2}(2\pi y)} f_{\theta,m,n}(y)\frac{dy}{y} \qquad (4.0.18)$$

with

$$f_{\theta,m,n}(y) = \begin{cases} 1, & \textit{if } 0 < \theta < 1/2, \\ e^{-2\pi^2 y^2 (m^2+n^2)}, & \textit{if } \theta = 1/2, \\ 0, & \textit{if } \theta > 1/2. \end{cases}$$

In particular, we obtain the following quite surprising consequence of number theoretic nature.

Corollary 4.0.21 (Paper G, Corollary 1.4)**.** *If ϕ is an even Hecke–Maaß cusp form with eigenvalue $s_\phi(1-s_\phi)$ and Hecke eigenvalues $\lambda_\phi(n)$, then*

$$\sum_{m,n\geq 1} \frac{\tau_1((m,n))\lambda_\phi(m)\lambda_\phi(n)}{(mn)^{1/2}} \int_{\max(m,n)}^{\infty} |K_{s_\phi - 1/2}(2\pi y)|^2 \frac{dy}{y} \geq 0. \tag{4.0.19}$$

Remark 4.0.22*.* Whether the left-hand side of (4.0.19) has any relation to L-functions is unclear, but would obviously make Corollary 4.0.21 very interesting.

Bibliography

[1] Menny Aka, Manuel Luethi, Philippe Michel, and Andreas Wieser. Simultaneous supersingular reductions of CM elliptic curves. *arXiv e-prints*, page arXiv:2005.01537, May 2020.

[2] Nickolas Andersen and Eren Mehmet Kı ral. Level reciprocity in the twisted second moment of Rankin-Selberg L-functions. *Mathematika*, 64(3):770–784, 2018.

[3] Avner Ash and Glenn Stevens. Modular forms in characteristic l and special values of their L-functions. *Duke Math. J.*, 53(3):849–868, 1986.

[4] Viviane Baladi and Brigitte Vallée. Euclidean algorithms are Gaussian. *J. Number Theory*, 110(2):331–386, 2005.

[5] M. V. Berry. Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature. *J. Phys. A*, 35(13):3025–3038, 2002.

[6] Sandro Bettin. On the reciprocity law for the twisted second moment of Dirichlet L-functions. *Trans. Amer. Math. Soc.*, 368(10):6887–6914, 2016.

[7] Sandro Bettin. High moments of the estermann function. *Algebra and Number Theory*, 13, 01 2017.

[8] Sandro Bettin and Sary Drappeau. Limit laws for rational continued fractions and value distribution of quantum modular forms. *arXiv e-prints*, page arXiv:1903.00457, Mar 2019.

[9] Valentin Blomer. Epstein zeta-functions, subconvexity, and the purity conjecture. *Journal of the Institute of Mathematics of Jussieu*, pages 1–16, 2018.

[10] Valentin Blomer, Étienne Fouvry, Emmanuel Kowalski, Philippe Michel, and Djordje Milićević. On moments of twisted L-functions. *Amer. J. Math.*, 139(3):707–768, 2017.

[11] Valentin Blomer, Étienne Fouvry, Emmanuel Kowalski, Philippe Michel, Djordje Milićević, and Will Sawin. The second moment theory of families of L-functions. *arXiv e-prints*, page arXiv:1804.01450, Apr 2018.

[12] Valentin Blomer and Rizwanur Khan. Twisted moments of L-functions and spectral reciprocity. *Duke Math. J.*, 168(6):1109–1177, 2019.

[13] Valentin Blomer, Xiaoqing Li, and Stephen D. Miller. A spectral reciprocity formula and non-vanishing for L-functions on $\mathrm{GL}(4) \times \mathrm{GL}(2)$. *J. Number Theory*, 205:1–43, 2019.

[14] Valentin Blomer and Djordje Milićević. The second moment of twisted modular L-functions. *Geom. Funct. Anal.*, 25(2):453–516, 2015.

[15] Wł odzimierz Bryc. *The Normal Distribution: Characterizations with Applications*, volume 100 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1995. Characterizations with applications.

[16] Daniel Bump. *Automorphic forms and representations*, volume 55 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[17] G. Chinta, N. Diamantis, and C. O'Sullivan. Second order modular forms. *Acta Arith.*, 103(3):209–223, 2002.

[18] J. Brian Conrey. The mean-square of Dirichlet L-functions. *arXiv e-prints*, page arXiv:0708.2699, Aug 2007.

[19] Petru Constantinescu. Distribution of Modular Symbols in $\mathbb{H}^3$. *International Mathematics Research Notices*, 09 2020. rnaa241.

[20] J. E. Cremona. *Algorithms for modular elliptic curves*. Cambridge University Press, Cambridge, second edition, 1997.

[21] Chantal David, Jack Fearnley, and Hershy Kisilevsky. Vanishing of L-functions of elliptic curves over number fields. In *Ranks of elliptic curves and random matrix theory*, volume 341 of *London Math. Soc. Lecture Note Ser.*, pages 247–259. Cambridge Univ. Press, Cambridge, 2007.

[22] Nikolaos Diamantis, Jeffrey Hoffstein, Eren Mehmet Kıral, and Min Lee. Additive twists and a conjecture by Mazur, Rubin and Stein. *J. Number Theory*, 209:1–36, 2020.

[23] Nikolaos Diamantis and Larry Rolen. Period polynomials, derivatives of L-functions, and zeros of polynomials. *Res. Math. Sci.*, 5(1):Paper No. 9, 15, 2018.

[24] W. Duke. Hyperbolic distribution problems and half-integral weight Maass forms. *Invent. Math.*, 92(1):73–90, 1988.

[25] William Duke, Özlem Imamoḡlu, and Árpád Tóth. Geometric invariants for real quadratic fields. *Ann. of Math. (2)*, 184(3):949–990, 2016.

[26] Bas Edixhoven, Gerard van der Geer, and Ben Moonen. Modular forms. In *Modular forms on Schiermonnikoog*, pages 1–11. Cambridge Univ. Press, Cambridge, 2008.

[27] Manfred Einsiedler, Elon Lindenstrauss, Philippe Michel, and Akshay Venkatesh. Distribution of periodic torus orbits and Duke's theorem for cubic fields. *Ann. of Math. (2)*, 173(2):815–885, 2011.

[28] Manfred Einsiedler, Elon Lindenstrauss, Philippe Michel, and Akshay Venkatesh. The distribution of closed geodesics on the modular surface, and Duke's theorem. *Enseign. Math. (2)*, 58(3-4):249–313, 2012.

[29] P. Erdös and M. Kac. The Gaussian law of errors in the theory of additive number theoretic functions. *Amer. J. Math.*, 62:738–742, 1940.

[30] Brooke Feigon, Kimball Martin, and David Whitehouse. Periods and nonvanishing of central L-values for $\mathrm{GL}(2n)$. *Israel J. Math.*, 225(1):223–266, 2018.

[31] Solomon Friedberg. What is... the Langlands program? *Notices Amer. Math. Soc.*, 65(6):663–665, 2018.

[32] Wee Teck Gan, Benedict H. Gross, and Dipendra Prasad. Symplectic local root numbers, central critical L values, and restriction problems in the representation theory of classical groups. Number 346, pages 1–109. 2012. Sur les conjectures de Gross et Prasad. I.

[33] Roger Godement and Hervé Jacquet. *Zeta functions of simple algebras*. Lecture Notes in Mathematics, Vol. 260. Springer-Verlag, Berlin-New York, 1972.

[34] Dorian Goldfeld. The distribution of modular symbols. In *Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997)*, pages 849–865. de Gruyter, Berlin, 1999.

[35] Dorian Goldfeld. Zeta functions formed with modular symbols. In *Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996)*, volume 66 of *Proc. Sympos. Pure Math.*, pages 111–121. Amer. Math. Soc., Providence, RI, 1999.

[36] Dorian Goldfeld. *Automorphic forms and L-functions for the group* $\mathrm{GL}(n, \mathrm{R})$, volume 99 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2015. With an appendix by Kevin A. Broughan, Paperback edition of the 2006 original [MR2254662].

[37] Benedict H. Gross and Don B. Zagier. Heegner points and derivatives of L-series. *Invent. Math.*, 84(2):225–320, 1986.

[38] Gergely Harcos and Philippe Michel. The subconvexity problem for Rankin-Selberg L-functions and equidistribution of Heegner points. II. *Invent. Math.*, 163(3):581–655, 2006.

[39] R. Neal Harris. The refined Gross-Prasad conjecture for unitary groups. *Int. Math. Res. Not. IMRN*, (2):303–389, 2014.

[40] Dennis A. Hejhal and Barry N. Rackner. On the topography of Maass waveforms for $\mathrm{PSL}(2, \mathbf{Z})$. *Experiment. Math.*, 1(4):275–305, 1992.

[41] Roman Holowinsky. Sieving for mass equidistribution. *Ann. of Math. (2)*, 172(2):1499–1516, 2010.

[42] Roman Holowinsky and Kannan Soundararajan. Mass equidistribution for Hecke eigenforms. *Ann. of Math. (2)*, 172(2):1517–1528, 2010.

[43] Yueke Hu and Abhishek Saha. Sup-norms of eigenfunctions in the level aspect for compact arithmetic surfaces, II: newforms and subconvexity. *arXiv e-prints*, page arXiv:1905.06295, May 2019.

[44] Peter Humphries. Equidistribution in shrinking sets and L^4-norm bounds for automorphic forms. *Math. Ann.*, 371(3-4):1497–1543, 2018.

[45] Atsushi Ichino. Trilinear forms and the central values of triple product L-functions. *Duke Math. J.*, 145(2):281–307, 2008.

[46] Atsushi Ichino and Tamutsu Ikeda. On the periods of automorphic forms on special orthogonal groups and the Gross-Prasad conjecture. *Geom. Funct. Anal.*, 19(5):1378–1425, 2010.

[47] Kenneth Ireland and Michael Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[48] Henryk Iwaniec. Fourier coefficients of modular forms of half-integral weight. *Invent. Math.*, 87(2):385–401, 1987.

[49] Henryk Iwaniec. *Topics in classical automorphic forms*, volume 17 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1997.

[50] Henryk Iwaniec and Peter Sarnak. L^∞ norms of eigenfunctions of arithmetic surfaces. *Ann. of Math. (2)*, 141(2):301–320, 1995.

[51] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika. Rankin-Selberg convolutions. *Amer. J. Math.*, 105(2):367–464, 1983.

[52] Seokho Jin, Wenjun Ma, Ken Ono, and Kannan Soundararajan. Riemann hypothesis for period polynomials of modular forms. *Proc. Natl. Acad. Sci. USA*, 113(10):2603–2608, 2016.

[53] David Joyner. *Distribution theorems of L-functions*, volume 142 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow; John Wiley & Sons, Inc., New York, 1986.

[54] Svetlana Katok and Peter Sarnak. Heegner points, cycles and Maass forms. *Israel J. Math.*, 84(1-2):193–227, 1993.

[55] Nicholas M. Katz. *Sommes exponentielles*, volume 79 of *Astérisque*. Société Mathématique de France, Paris, 1980. Course taught at the University of Paris, Orsay, Fall 1979, With a preface by Luc Illusie, Notes written by Gérard Laumon, With an English summary.

[56] E. Kowalski, Ph. Michel, and W. Sawin. Bilinear forms with Kloosterman sums and applications. *arXiv e-prints*, page arXiv:1511.01636, Nov 2015.

[57] J.-P. Labesse and J. Schwermer, editors. *Cohomology of arithmetic groups and automorphic forms*, volume 1447 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1990.

[58] R. P. Langlands. Letter to Andre Weil, January 1967. Available at `www.sunsite.ubc.ca/DigitalMathArchive/Langlands/intro.html`.

[59] Jungwon Lee and Hae-Sang Sun. Dynamics of continued fractions and distribution of modular symbols. *arXiv e-prints*, page arXiv:1902.06277, Feb 2019.

[60] Stephen Lester, Kaisa Matomäki, and Maksym Radziwił ł. Small scale distribution of zeros and mass of modular forms. *J. Eur. Math. Soc. (JEMS)*, 20(7):1595–1627, 2018.

[61] Elon Lindenstrauss. Invariant measures and arithmetic quantum unique ergodicity. *Ann. of Math. (2)*, 163(1):165–219, 2006.

[62] Yu. V. Linnik. *Ergodic properties of algebraic fields*. Translated from the Russian by M. S. Keane. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 45. Springer-Verlag New York Inc., New York, 1968.

[63] Sheng-Chi Liu, Riad Masri, and Matthew P. Young. Subconvexity and equidistribution of Heegner points in the level aspect. *Compos. Math.*, 149(7):1150–1174, 2013.

[64] Wen Zhi Luo and Peter Sarnak. Quantum ergodicity of eigenfunctions on $\mathrm{PSL}_2(\mathbf{Z})\backslash\mathbf{H}^2$. *Inst. Hautes Études Sci. Publ. Math.*, (81):207–237, 1995.

[65] Wenzhi Luo and Peter Sarnak. Quantum variance for Hecke eigenforms. *Ann. Sci. École Norm. Sup. (4)*, 37(5):769–799, 2004.

[66] Aidan Lyon. Why are Normal Distributions Normal? *The British Journal for the Philosophy of Science*, 65(3):621–649, 09 2013.

[67] Hans Maass. Über die räumliche Verteilung der Punkte in Gittern mit indefiniter Metrik. *Math. Ann.*, 138:287–315, 1959.

[68] Ju I Manin. Periods of parabolic forms and p-adic Hecke Series. *Mathematics of the USSR-Sbornik*, 21(3):371–393, apr 1973.

[69] Kimball Martin and David Whitehouse. Central L-values and toric periods for GL(2). *Int. Math. Res. Not. IMRN*, (1):Art. ID rnn127, 141–191, 2009.

[70] B. Mazur. On the arithmetic of special values of L functions. *Invent. Math.*, 55(3):207–240, 1979.

[71] Barry Mazur and Karl Rubin. Relatively few rational points. Number Theory Seminar at Cal Tech May 2016. 2016.

[72] Barry Mazur and Karl Rubin. The statistical behavior of modular symbols and arithmetic conjectures, November 2016. Talk given at Fields medal symposium, Toronto 2016, `http://people.math.harvard.edu/~mazur/papers/heuristics.Toronto.12.pdf`.

[73] Barry Mazur and Karl Rubin. Arithmetic conjectures suggested by the statistical behavior of modular symbols. *arXiv e-prints*, page arXiv:1910.12798, October 2019.

[74] Philippe Michel and Akshay Venkatesh. Equidistribution, L-functions and ergodic theory: on some problems of Yu. Linnik. In *International Congress of Mathematicians. Vol. II*, pages 421–457. Eur. Math. Soc., Zürich, 2006.

[75] Philippe Michel and Akshay Venkatesh. The subconvexity problem for GL_2. *Publ. Math. Inst. Hautes Études Sci.*, (111):171–271, 2010.

[76] Paul D. Nelson and Akshay Venkatesh. The orbit method and analysis of automorphic forms. *arXiv e-prints*, page arXiv:1805.07750, May 2018.

[77] Cormac O'Sullivan. Properties of Eisenstein series formed with modular symbols. *J. Reine Angew. Math.*, 518:163–186, 2000.

[78] Yiannis N. Petridis and Morten S. Risager. Modular symbols have a normal distribution. *Geom. Funct. Anal.*, 14(5):1013–1043, 2004.

[79] Yiannis N. Petridis and Morten S. Risager. The distribution of values of the Poincaré pairing for hyperbolic Riemann surfaces. *J. Reine Angew. Math.*, 579:159–173, 2005.

[80] Yiannis N. Petridis and Morten S. Risager. Hyperbolic lattice-point counting and modular symbols. *J. Théor. Nombres Bordeaux*, 21(3):719–732, 2009.

[81] Yiannis N. Petridis and Morten S. Risager. Arithmetic statistics of modular symbols. *Invent. Math.*, 212(3):997–1053, 2018.

[82] R. A. Rankin. Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions. I. The zeros of the function $\sum_{n=1}^{\infty}\tau(n)/n^s$ on the line $\mathrm{Re}s = 13/2$. II. The order of the Fourier coefficients of integral modular forms. *Proc. Cambridge Philos. Soc.*, 35:351–372, 1939.

[83] Morten S. Risager. On the distribution of modular symbols for compact surfaces. *Int. Math. Res. Not.*, (41):2125–2146, 2004.

[84] Zeév Rudnick. On the asymptotic distribution of zeros of modular forms. *Int. Math. Res. Not.*, (34):2059–2074, 2005.

[85] Zeév Rudnick and Peter Sarnak. The behaviour of eigenstates of arithmetic hyperbolic manifolds. *Comm. Math. Phys.*, 161(1):195–213, 1994.

[86] Yiannis Sakellaridis and Akshay Venkatesh. Periods and harmonic analysis on spherical varieties. *Astérisque*, (396):viii+360, 2017.

[87] Peter Sarnak. Arithmetic quantum chaos. In *The Schur lectures (1992) (Tel Aviv)*, volume 8 of *Israel Math. Conf. Proc.*, pages 183–236. Bar-Ilan Univ., Ramat Gan, 1995.

[88] Atle Selberg. Bemerkungen über eine Dirichletsche Reihe, die mit der Theorie der Modulformen nahe verbunden ist. *Arch. Math. Naturvid.*, 43:47–50, 1940.

[89] Goro Shimura. On the periods of modular forms. *Math. Ann.*, 229(3):211–221, 1977.

[90] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*, volume 11 of *Publications of the Mathematical Society of Japan*. Princeton University Press, Princeton, NJ, 1994. Reprint of the 1971 original, Kanô Memorial Lectures, 1.

[91] K. Soundararajan. Quantum unique ergodicity and number theory. In *Proceedings of the International Congress of Mathematicians. Volume II*, pages 357–382. Hindustan Book Agency, New Delhi, 2010.

[92] Kannan Soundararajan. Quantum unique ergodicity for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. *Ann. of Math. (2)*, 172(2):1529–1538, 2010.

[93] Kannan Soundararajan. Weak subconvexity for central values of L-functions. *Ann. of Math. (2)*, 172(2):1469–1498, 2010.

[94] H-S. Sun. A proof of the conjecture of Mazur-Rubin-Stein. *To appear in Bull. Korean Math. Soc*, 2020.

[95] John Torrence Tate, Jr. *FOURIER ANALYSIS IN NUMBER FIELDS AND HECKE'S ZETA-FUNCTIONS*. ProQuest LLC, Ann Arbor, MI, 1950. Thesis (Ph.D.)–Princeton University.

[96] Edward C. Titchmarsh. On Epstein's Zeta-Function. *Proc. London Math. Soc. (2)*, 36:485–500, 1934.

[97] Akshay Venkatesh. Sparse equidistribution problems, period bounds and subconvexity. *Ann. of Math. (2)*, 172(2):989–1094, 2010.

[98] J.-L. Waldspurger. Sur les coefficients de Fourier des formes modulaires de poids demi-entier. *J. Math. Pures Appl. (9)*, 60(4):375–484, 1981.

[99] Thomas Crawford Watson. *Rankin triple products and quantum chaos*. ProQuest LLC, Ann Arbor, MI, 2002. Thesis (Ph.D.)–Princeton University.

[100] Han Wu and Nickolas Andersen. Explicit subconvexity for GL_2 and some applications. *arXiv e-prints*, page arXiv:1812.04391, December 2018.

[101] Iannis Xenakis. *Formalized Music: Thought and Mathematics in Composition*. Indiana University Press, 1971.

[102] Matthew P. Young. The quantum unique ergodicity conjecture for thin sets. *Adv. Math.*, 286:958–1016, 2016.

[103] Matthew P. Young. A note on the sup norm of Eisenstein series. *Q. J. Math.*, 69(4):1151—-1161, 2018.

[104] Don Zagier. *Modular forms in one variable*. Unpublished, 1991.

[105] Don Zagier. Quantum modular forms. In *Quanta of maths*, volume 11 of *Clay Math. Proc.*, pages 659–675. Amer. Math. Soc., Providence, RI, 2010.

[106] Steven Zelditch. Uniform distribution of eigenfunctions on compact hyperbolic surfaces. *Duke Math. J.*, 55(4):919–941, 1987.

[107] Steven Zelditch. On the rate of quantum ergodicity. I. Upper bounds. *Comm. Math. Phys.*, 160(1):81–92, 1994.

Part II

Papers

PAPER A
CENTRAL VALUES OF ADDITIVE TWISTS OF CUSPIDAL L-FUNCTIONS

CENTRAL VALUES OF ADDITIVE TWISTS OF CUSPIDAL L-FUNCTIONS

ASBJØRN CHRISTIAN NORDENTOFT

Abstract. Additive twists are important invariants associated to holomorphic cusp forms; they encode the Eichler–Shimura isomorphism and contain information about automorphic L-functions. In this paper we prove that central values of additive twists of the L-function associated to a holomorphic cusp form f of even weight k are asymptotically normally distributed. This generalizes (to $k \geq 4$) a recent breakthrough of Petridis and Risager concerning the arithmetic distribution of modular symbols. Furthermore we give as an application an asymptotic formula for the averages of certain 'wide' families of automorphic L-functions, consisting of central values of the form $L(f \otimes \chi, 1/2)$ with χ a Dirichlet character.

1. Introduction

In this paper we study the statistics of central values of additive twists of the L-functions of holomorphic cusp forms (of arbitrary even weight). Additive twists of cuspidal L-functions are important invariants; on the one hand they show up in the parametrization of the Eichler–Shimura isomorphism and on the other hand additive twists shed light on central values of Dirichlet twists of cuspidal L-functions.

We prove that when arithmetically ordered, the central values of the additive twists of a cuspidal L-function are asymptotically normally distributed. As an application we calculate the asymptotic behavior (as $X \to \infty$) of the averages of certain 'wide' families of automorphic L-functions;

$$\sum_{\pi \in \mathcal{F}_n,\, \mathrm{cond}(\pi) \leq X}^{*} L(\pi, 1/2), \tag{1.1}$$

where $\mathrm{cond}(\pi)$ denotes the conductor of the automorphic representation π, the asterisk on the sum denotes a certain weighting and the families consist of isobaric sums of twists;

$$\mathcal{F}_n = \{(\pi_f \otimes \chi_1) \boxplus \cdots \boxplus (\pi_f \otimes \chi_{2n}) \mid \chi_1 \cdots \chi_{2n} = \mathbf{1}\},$$

where $\chi_1, \ldots, \chi_{2n}$ are automorphic representations of $\mathrm{GL}_1(\mathbb{A}_\mathbb{Q})$, $\mathbf{1}$ denotes the trivial automorphic representation of $\mathrm{GL}_1(\mathbb{A}_\mathbb{Q})$ and π_f is the automorphic representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ associated to a *fixed* holomorphic newform f (suppressed in the notation). For the precise statements of our main results, see Theorem 1.1 and Corollary 1.9 below.

Date: December 3, 2020.

2010 *Mathematics Subject Classification.* 11F67(primary), and 11M41(secondary).

Our distribution result is a higher weight analogue of a recent result of Petridis and Risager [29], which settled (an averaged version of) a conjecture due to Mazur and Rubin [23] concerning the normal distribution of modular symbols (a different proof was later given by Lee and Sun [21] using dynamical methods).
The conjecture of Mazur and Rubin concerns the arithmetic distribution of the modular symbol map;

$$\{\infty, \mathfrak{a}\} \mapsto \langle \mathfrak{a}, f\rangle := 2\pi i \int_{\infty}^{\mathfrak{a}} f(z)dz,$$

where $f \in \mathcal{S}_2(\Gamma_0(q))$ is a cusp form of weight 2 and level q and $\{\infty, \mathfrak{a}\}$ is the homology class of curves between the cusps ∞ and $\mathfrak{a}$. Petridis and Risager prove that this map is asymptotically normally distributed when ordered by the denominator of the cusp $\mathfrak{a}$ and appropriately normalized [29, Theorem 1.10]. See Section 1.3 below for more background on the conjectures of Mazur and Rubin and their motivation.

1.1. **Statement of results.** We will now state a special case of our main result and refer to Theorem 5.1 below for the most general version. Let $f \in S_k(\Gamma_0(q))$ be a cusp form of (arbitrary) even weight k and level q with Fourier expansion (at ∞) given by

$$f(z) = \sum_{n\geq 1} a_f(n) q^n, \quad q = e^{2\pi i z}.$$

Then we define the *additive twist* (by $r \in \mathbb{R}$) of the L-function associated to f as

$$L(f, r, s) := \sum_{n\geq 1} \frac{a_f(n) e(nr)}{n^s},$$

where $e(x) = e^{2\pi i x}$. The additive twists converge absolutely for $\operatorname{Re} s > \frac{k+1}{2}$ and when $r \in \mathbb{Q}$, they admit analytic continuation. If furthermore r is $\Gamma_0(q)$ equivalent to ∞, we have simple functional equations relating $s \leftrightarrow k - s$ (see Section 3.3 below for details). For $k = 2$ the additive twists coincide with modular symbols (see Remark 1.5 below for details).

Now we will explain what we mean by saying that additive twists are *asymptotically normally distributed*: Given $X > 0$, we consider $L(f, \cdot, k/2)$ as a random variable defined on the following outcome space endowed with the uniform probability measure;

$$T(X) := \{a/c \in \mathbb{Q} \mid 0 < a < c \leq X,\ (a, c) = 1,\ q \mid c\}. \tag{1.2}$$

Then our main theorem is the following.

Theorem 1.1. *Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a cusp form of even weight k and level q. Then for any fixed box $\Omega \subset \mathbb{C}$, we have*

$$\begin{aligned}\mathbb{P}_{T(X)}\left(\frac{L(f, a/c, k/2)}{(C_f \log c)^{1/2}} \in \Omega\right) :=& \frac{\#\{a/c \in T(X) \mid \frac{L(f,a/c,k/2)}{(C_f \log c)^{1/2}} \in \Omega\}}{\#T(X)} \\ =& \mathbb{P}(\mathcal{N}_{\mathbb{C}}(0,1) \in \Omega) + o(1),\end{aligned} \tag{1.3}$$

as $X \to \infty$*, where* $\mathcal{N}_{\mathbb{C}}(0,1)$ *denotes the standard complex normal distribution and the variance is given by*

$$C_f := \frac{(4\pi)^k ||f||^2}{(k-1)! \operatorname{vol}(\Gamma_0(q))}, \tag{1.4}$$

with $||f||$ *the Petersson-norm of* f *and* $\operatorname{vol}(\Gamma_0(q))$ *the hyperbolic volume of* $\Gamma_0(q)\backslash\mathbb{H}$*. (Here* $\mathbb{P}(\mathcal{N}_{\mathbb{C}}(0,1) \in \Omega)$ *denotes the probability of the event* $\mathcal{N}_{\mathbb{C}}(0,1) \in \Omega$*.)*

Remark 1.2. With our methods, we can generalize the above theorem in three aspects, which can all be combined:

(1) We can consider more general outcome spaces corresponding to twists at an arbitrary cusp. Note that $T(X)$ corresponds exactly to additive twists at cusps, which are $\Gamma_0(q)$-equivalent to ∞. See (5.1) below for the definition of the outcome space for general cusps.
(2) We can consider cusp forms for a general discrete and co-finite subgroup of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at ∞.
(3) We can consider the joint distribution of an orthogonal basis of cusp forms.

See Theorem 5.1 below for the most general version of our main theorem, incorporating all of the three above aspects.

Remark 1.3. The constant C_f is a higher weight analogue of the *variance slope* defined by Mazur and Rubin (see [29, Theorem 1.9]). Note that C_f is independent of the embedding $f \hookrightarrow \mathcal{S}_k(\Gamma_0(N))$.

Remark 1.4. Independently, a different proof of Theorem 1.1 in the special case of trivial level $q = 1$ was obtained by Bettin and Drappeau [2] using dynamical methods similar to those used by Sun and Lee. It is still an open problem to extend the dynamical approach to deal with general level, but in return the dynamical approach of Bettin and Drappeau applies to much more general *quantum modular forms* in the sense of Zagier [34]. It is unclear whether the automorphic methods of this paper can be generalized to deal with quantum modular forms as well. The similarities and differences between the automorphic and dynamical approach deserve further exploration.

Remark 1.5. In more concrete terms the above theorem says that for any fixed real numbers $x_1 < x_2$ and $y_1 < y_2$, we have

$$\frac{\#\left\{\frac{a}{c} \in T(X) \mid x_1 \leq \operatorname{Re}\left(\frac{L(f,a/c,k/2)}{(C_f \log c)^{1/2}}\right) \leq x_2, y_1 \leq \operatorname{Im}\left(\frac{L(f,a/c,k/2)}{(C_f \log c)^{1/2}}\right) \leq y_2\right\}}{\#T(X)} \to \frac{1}{2\pi}\int_{x_1}^{x_2}\int_{y_1}^{y_2} e^{-(x^2+y^2)/2} dx dy,$$

as $X \to \infty$, which is exactly the formulation used in [29]. One can see that restricting to the case $k = 2$ and letting $f \in \mathcal{S}_2(\Gamma_0(q))$, Theorem 1.1 above recovers [29, Theorem 1.10]. Here one has to use the integral representation (3.10) for the additive twist of the L-function, which shows $\langle \mathfrak{a}, f\rangle = L(f, r_{\mathfrak{a}}, 1)$, where $r_{\mathfrak{a}} \in \mathbb{R}$ represents the cusp $\mathfrak{a}$ (i.e. $r_{\mathfrak{a}}$ is fixed by the parabolic subgroup $\Gamma_{\mathfrak{a}}$).

1.2. **Moment calculations.** The proof of Theorem 1.1 uses the method of moments. The calculation of the moments of additive twists is of independent interest and is used in the application to automorphic L-functions, in Corollary 1.9 below. In the course of the paper we will evaluate a number of different moments. In particular we will prove the following result at the end of Section 5.5, which is exactly what we need to conclude Corollary 1.9.

Theorem 1.6. *Let $f \in S_k(\Gamma_0(q))$ be a cusp form of even weight k and level q and n a non-negative integer. Then we have*

$$\sum_{\substack{0<a<c\leq X\\(qa,c)=1}} |L(f,a/c,k/2)|^{2n} = P_n(\log X)X^2 + O_\varepsilon(X^{4/3+\varepsilon}), \tag{1.5}$$

where P_n is a polynomial of degree n with leading coefficient

$$\frac{q(2C_f)^n\, n!}{\pi \operatorname{vol}(\Gamma_0(q))},$$

with C_f as in (1.4) above.

Remark 1.7. The above moments correspond to additive twists at cusps which are $\Gamma_0(q)$-equivalent to the cusp 0 (the set of all such twist is denoted $T_{\infty 0}$ in (5.1) below).

Remark 1.8. The determination of the moments follows from the analytic properties of a certain Eisenstein series $E^{m,n}(z,s)$ generalizing series introduced by Goldfeld in [14] and [15]. Determining the location of the dominating pole, the corresponding pole order and leading Laurent coefficient of the original Goldfeld Eisenstein series was firstly achieved by Petridis and Risager in [28] using perturbation theory and the analytic properties of the hyperbolic resolvent. This allowed them to prove normal distribution for a certain more geometrically flavored ordering of the modular symbols (ordered by c^2+d^2, where c, d are the lower entries of the matrix γ). In order to prove (an averaged version of) the conjecture of Mazur and Rubin, Petridis and Risager [29] essentially had to derive the analytic properties of the constant Fourier coefficient of $E^{m,n}(z,s)$. This is reminiscent of the Shahidi–Langlands method [13, Section 8]. The strategy of proof in this paper is inspired by the overall strategy introduced by Petridis and Risager.

1.3. **Applications to automorphic L-functions.** The motivation behind the conjectures of Mazur and Rubin was to gain information about the vanishing/non-vanishing of the central values of the twisted L-functions $L(E,\chi,1)$ where $E/\mathbb{Q}$ is an elliptic curve and χ is a Dirichlet character. By a sufficiently general version of the Birch–Swinnerton-Dyer conjecture, this is related to the following problem in Diophantine stability (see [23] for details):

How likely is it that $\operatorname{rank}_{\mathbb{Z}} E(K) > \operatorname{rank}_{\mathbb{Z}} E(\mathbb{Q})$
as K ranges over abelian extensions of $\mathbb{Q}$?

If χ is a primitive Dirichlet character modulo c, then the Birch–Stevens formula [23, Theorem 2.3] relates these central values to modular symbols;

$$\tau(\overline{\chi})L(E,\chi,1) = \sum_{a\in(\mathbb{Z}/c\mathbb{Z})^\times} \overline{\chi}(a)\langle a/c, f_E\rangle,$$

where f_E is the weight 2 newform corresponding to E via modularity and $\tau(\overline{\chi})$ is a Gauss-sum. This led Mazur and Rubin to the study of the distribution of modular symbols, and based on computational experiments, they made a number of conjectures about the distribution of modular symbols, one of which predicted a normal distribution. In this paper we will not contribute to the other conjectures put forth by Mazur and Rubin, consult instead [8] and [3, Theorem 9.2].

Following these lines of thinking, we apply our methods to the study of families constructed from certain twisted L-functions. Given a newform $f \in \mathcal{S}_k(\Gamma_0(q))$ of weight k and level q and a primitive Dirichlet character χ of conductor co-prime to q, we define the *multiplicatively twisted L-function of f*;

$$L(f \otimes \chi, s) := \sum_{n \geq 1} \frac{\lambda_f(n)\chi(n)}{n^s}, \quad \operatorname{Re} s > 1,$$

(where $\lambda_f(n)$ denotes the nth Hecke eigenvalue of f), which admits analytic continuation satisfying a functional equation. Note that because of the co-primality condition the above Dirichlet series defines (the finite part of) the L-function corresponding to the automorphic representation $\pi_f \otimes \chi$ (justifying the notation).
The study of averages of multiplicative twists of cuspidal L-functions has a long history (see for instance the work of Rohrlich [30], Duke, Friedlander and Iwaniec [9], and Chinta [5]). Recently Blomer, Fouvry, Kowalski, Michel, Milićević and Sawin [3] have given an extensive account of the second moment theory for such twists. We are able to obtain new results for these automorphic L-functions:
Combining Theorem 1.6 and the (generalized) Birch–Stevens formula (see Lemma 6.1 below), we obtain asymptotic formulas for the following (arithmetically weighted) averages of certain 'wide' families of multiplicatively twisted L-functions, making (1.1) precise.

Corollary 1.9. *Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a newform of even weight k and level q and n a non-negative integer. Then we have for all $\varepsilon > 0$*

$$\sum_{\substack{0<c\leq X,\\ (c,q)=1}} \frac{1}{\varphi(c)^{2n-1}} \sum_{\substack{\chi_1,\dots,\chi_{2n} \bmod c,\\ \chi_1\cdots\chi_{2n}=\chi_{\text{principal}}}} \varepsilon_{\chi_1,\dots,\chi_n} \prod_{i=1}^{2n} \nu(f,\chi_i^*, c/c(\chi_i)) L(f\otimes \chi_i^*, 1/2)$$

$$= P_n(\log X)X^2 + O_\varepsilon(X^{4/3+\varepsilon}), \tag{1.6}$$

where $\chi^* \bmod c(\chi)$ denotes the primitive character inducing χ, $\chi_{\text{principal}}$ is the principal character $\bmod c$, P_n is the degree n polynomial from (1.5),

$$\varepsilon_{\chi_1,\dots,\chi_n} = \chi_1(-1)\cdots\chi_n(-1)$$

is a sign, and ν is an arithmetic weight given by

$$\nu(f,\chi,n) := \tau(\overline{\chi}) \sum_{\substack{n_1n_2n_3=n,\\ (n_1,q)=1}} \chi(n_1)\mu(n_1)\overline{\chi}(n_2)\mu(n_2)\lambda_f(n_3)n_3^{1/2}. \tag{1.7}$$

In particular for $n=1$, the above corollary reduces to an average second moment (see Corollary 6.3 below), which was calculated without the extra averaging in [3, Theorem 1.17]. Interestingly our methods completely avoid the use of approximate functional equations.

Remark 1.10. Observe that the sum in the arithmetic weight ν can be expressed as the triple convolution

$$(\delta_{(\cdot,q)=1} \times \chi \times \mu) * (\overline{\chi} \times \mu) * (\lambda_f \times (\cdot)^{1/2}).$$

This was exploited by Bruggeman and Diamantis in [4] in their study of Fourier coefficients of the Goldfeld Eisenstein series $E^{1,0}(z,s)$. Furthermore for χ primitive, we have

$$\nu(f,\chi^*,c/c(\chi)) = \nu(f,\chi,1) = \tau(\overline{\chi}),$$

and in general, we have the bound $|\nu(f,\chi,n)| \ll (c(\chi)n)^{1/2}$ using Deligne's bound for the Hecke eigenvalues.

Remark 1.11. S. Bettin [1] has considered the Eisenstein case of the above theorems, which amounts to studying the Estermann function defined as

$$D(a/c,s) := \sum_{n\geq 1} \frac{d(n)e(na/c)}{n^s},$$

where $d(n)$ is the divisor function and $a/c \in \mathbb{Q}$. Bettin managed to calculate all moments averaging over $a \in (\mathbb{Z}\backslash c\mathbb{Z})^\times$ using an approximate functional equation. He similarly applied his results to studying certain iterated moments of central values of Dirichlet L-functions.

Acknowledgements. I would like to express my gratitude to my advisor Morten Risager for suggesting this problem to me and for our countless stimulating discussions. I would also like to thank Yiannis Petridis for his time and insight.

2. Method of proof

Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a cusp form of even weight k and level q. In this section we will describe the overall strategy of the proof of Theorem 1.1. Our approach is an extension of the techniques developed by Petridis and Risager in [28] and [29]. We would like to point out that many technical difficulties show up when $k \geq 4$ and some essential new ideas were needed. This includes using the lowering and raising operators in the analysis of the recursion formula (see for instance the proof of Lemma 4.8), the automorphic completion step as described in Section 5.1 and the use of what we call *N-shifted Golfeld Eisenstein series* defined in (5.8).

2.1. **The strategy.**

We will use a classical result of Fréchet and Shohat [31, p.17] known as the *method of moments*; in order to get the sought-after convergence in distribution, it is enough to show that (after a suitable normalization) the asymptotic moments (as $X \to \infty$) of the central values;

$$\sum_{a/c \in T(X)} L(f,a/c,k/2)^m \overline{L(f,a/c,k/2)}^n,$$

agree with those of the standard complex normal distribution.
By a standard complex analysis argument, this can be reduces to understanding the analytic properties of the following Dirichlet series

$$D^{m,n}(f,s) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(q)/\Gamma_\infty} \frac{L(f,\gamma\infty,k/2)^m\overline{L(f,\gamma\infty,k/2)}^n}{(c_\gamma)^{2s}}, \tag{2.1}$$

where $\gamma\infty = a_\gamma/c_\gamma$ and a_γ, c_γ denote the upper-left and lower-left entry of γ, respectively. Note that $L(f,\gamma\infty,k/2)$ and c_γ are indeed well-defined on the double coset $\Gamma_\infty\backslash\Gamma_0(q)/\Gamma_\infty$ and that $\gamma\mapsto\gamma\infty$ defines a bijection

$$\Gamma_\infty\backslash\Gamma_0(q)/\Gamma_\infty \to \{a/c\in\mathbb{Q} \mid 0<a<c,\ (a,c)=1,\ q\mid c\}\cup\{\infty\}.$$

For the convenience of the reader and since we will care about the error-terms in our applications (see Corollary 1.9 above), we have in Appendix A included a detailed exposition of the contour integration argument with explicit error-terms that is used in this step.

We will derive the analytic properties of $D^{m,n}(f,s)$ by studying the following generalized *Goldfeld Eisenstein series*;

$$E^{m,n}(z,s) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(q)} L(f,\gamma\infty,k/2)^m\overline{L(f,\gamma\infty,k/2)}^n\,\mathrm{Im}(\gamma z)^s, \tag{2.2}$$

where $\Gamma_\infty = \{\left(\begin{smallmatrix}1&n\\0&1\end{smallmatrix}\right) \mid n\in\mathbb{Z}\}$ is the stabilizer of ∞ in $\Gamma_0(q)$. The series (2.1) and (2.2) are connected since the constant term in the Fourier expansion of $E^{m,n}(z,s)$ (at ∞) is given by

$$\frac{\pi^{1/2}y^{1-s}\Gamma(s-1/2)}{\Gamma(s)}D^{m,n}(f,s).$$

For the proof see Lemma 5.6 below. This will allow us to pass analytic information from $E^{m,n}(z,s)$ to $D^{m,n}(f,s)$ as one does in the Langlands–Shahidi method.

In order to get information about the analytic properties of $E^{m,n}(z,s)$, we will use ideas from an unpublished paper by Chinta and O'Sullivan [6]; we express $E^{m,n}(z,s)$ as a linear combination of certain Poincaré series $G_{A,B,l}(z,s)$, which are weight l automorphic forms. This is known as *automorphic completion* and will allow us to employ the spectral theory of automorphic forms.
In particular we will use the analytic properties of the *higher weight resolvent operators* to recursively understand the pole order at $s=1$ and the leading Laurent coefficient of $G_{A,B,l}(z,s)$. The overall strategy can be illustrated as follows:

1. Analytic properties of higher weight resolvent operators

$\downarrow$ Induction argument

2. Analytic properties of the Poincaré series $G_{A,B,l}(z,s)$

$\downarrow$ A formula for the central value of additive twists

3. Analytic properties of $E^{m,n}(z,s)$

$\downarrow$ Fourier expansion

4. Analytic properties of $D^{m,n}(f,s)$

$\downarrow$ Contour integration

5. Asymptotic moments of $L(f,a/c,k/2)$

$\downarrow$ Fréchet–Shohat (method of moments)

6. Normal distribution of $L(f,a/c,k/2)/(C_f \log c)^{1/2}$.

The rest of the paper is structured as follows; in Section 3 we will introduce the needed background on weight k Laplacians and additive twists. In Section 4, we will study the analytic properties of the Poincaré series $G_{A,B,l}(z,s)$. In Section 5, we will prove the normal distribution of additive twists; in order to keep the exposition as simple as possible, we will restrict to the case of a single cusp form and additive twists corresponding to the cusp ∞ and then explain how to extend the methods to the general setting. In Section 6, we will present applications to certain 'wide' families of automorphic L-functions. Finally in Appendix A, we included a version of contour integration with explicit error-terms for the convenience of the reader.

3. Background: Weight k Laplacians and additive twists

In this section we will recall some standard facts about higher weight Laplacians and additive twists of modular L-functions. We will work with a general discrete and co-finite subgroup Γ of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 (see [16, Section 2] for definitions). But one does not loose much by restricting to the case $\Gamma = \Gamma_0(q) := \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \text{ s.t. } q \mid c_\gamma\}$ of *Hecke congruence subgroups*.

3.1. **Weight k Laplacians.** We will refer to [10, Chapter 4], [25, section 2.1.2] and [16, Chapter 4] for a more comprehensive account on automorphic Laplace operators. Let k be an even integer. The space of *automorphic functions of weight k* with respect to Γ are (measurable) functions $g : \mathbb{H} \to \mathbb{C}$ satisfying

$$g(\gamma z) = j_\gamma(z)^k g(z), \quad \text{for all } \gamma \in \Gamma,$$

where $j_\gamma(z) := j(\gamma, z)/|j(\gamma, z)| = (cz+d)/|cz+d|$ with c, d the bottom-row entries of γ. Note that we have the cocycle relation;

$$j_{\gamma_1\gamma_2}(z) = j_{\gamma_1}(\gamma_2 z) j_{\gamma_2}(z). \tag{3.1}$$

Given an automorphic function g of weight k, we define the *Petersson norm* by

$$||g||^2 := \int_{\Gamma\backslash\mathbb{H}} |g(z)|^2 d\mu(z),$$

where $d\mu(z) = dxdy/y^2$ is the hyperbolic measure on $\mathbb{H}$. From this we define the *Hilbert space of all square integrable weight k automorphic functions*;

$$L^2(\Gamma, k) := \{g \text{ automorphic of weight } k \text{ with respect to } \Gamma \text{ s.t. } ||g|| < \infty\}.$$

(modulo the kernel of $||\cdot||$) with inner-product given by

$$\langle g, h\rangle := \int_{\Gamma\backslash\mathbb{H}} g(z)\overline{h(z)} d\mu(z),$$

for $g, h \in L^2(\Gamma, k)$. Maaß defined *raising*- and *lowering operators*, which maps between spaces of different weights. In terms of local coordinates they are given by

$$K_k := (z - \overline{z})\frac{\partial}{\partial z} + \frac{k}{2},$$
$$L_k := (z - \overline{z})\frac{\partial}{\partial \overline{z}} + \frac{k}{2}$$

for $z \in \mathbb{H}$ (acting on, say, smooth automorphic functions).
The raising and lowering operators are adjoint to each other in the following sense;

$$\langle K_k g_k, h_{k+2}\rangle = \langle g_k, L_{k+2} h_{k+2}, \rangle \tag{3.2}$$

and satisfy the following product rule;

$$\begin{aligned} K_{k+l}(g_k g_l) &= (K_k g_k) g_l + g_k (K_l g_l), \\ L_{k+l}(h_k h_l) &= (L_k h_k) h_l + h_k (L_l h_l), \end{aligned} \tag{3.3}$$

where g_k, g_l, h_k, h_l are smooth automorphic functions of appropriate weights.

Remark 3.1. In most modern expositions the raising operator is denoted R_k (see [10, Chapter 4]), but in order to avoid confusion with the resolvent operator, we follow the notation of Fay [12]. We note that our definition of the lowering operator is equal to minus the one of Fay.

Using these two operators, the *weight k Laplacian* can be defined as follows acting on smooth automorphic functions;

$$\Delta_k := K_{k-2}L_k + \lambda(k/2) = L_{k+2}K_k + \lambda(-k/2), \tag{3.4}$$

where $\lambda(s) = s(1-s)$. The operator Δ_k defined on the space of all smooth and rapidly decaying automorphic functions, defines a non-negative and essentially self-adjoint operator on the Hilbert space $L^2(\Gamma, k)$. We denote (by abuse of notation) the unique self-adjoint extension also by Δ_k with domain $D(\Delta_k)$ dense in $L^2(\Gamma, k)$. We define a *Maaß form* as an eigenfunction $\varphi \in D(\Delta_0)$ of Δ_0.
One sees by direct computation that for $f \in \mathcal{S}_k(\Gamma)$, we have $(z \mapsto y^{k/2} f(z)) \in D(\Delta_k) \subset L^2(\Gamma, k)$ and

$$L_k y^{k/2} f(z) = 0. \tag{3.5}$$

Combining this with (3.4), we see that $y^{k/2}f(z)$ is an eigenfunction for Δ_k with eigenvalue $\lambda(k/2) = (1-k/2)k/2$.

Using the raising- and lowering operators one can show that

$$\text{spec}\Delta_k \subset [1/4, \infty) \cup \{\lambda_0 = 0, \lambda_1, \dots, \lambda_m\} \cup \{\lambda(1), \dots, \lambda(k/2)\} \tag{3.6}$$

where the three sets above correspond to respectively the continuous part constituted by the Eisenstein series, the constant eigenfunction together with the so-called *exceptional Maaß forms* (Maaß forms with eigenvalue $0 < \lambda_i < 1/4$) and finally holomorphic cusp forms $\mathcal{S}_j(\Gamma)$ with $2 \leq j \leq k$ and $j \equiv k\,(2)$. Recall that by the work of Selberg when $\Gamma = \Gamma_0(q)$ there are always (an abundance of) embedded Maaß forms (with eigenvalue $\lambda \geq 1/4$).
It is a famous conjecture of Selberg that there are no exceptional Maaß forms when $\Gamma = \Gamma_0(q)$ is a Hecke congruence subgroup. Kim and Sarnak [19] have proved that the smallest eigenvalue $\lambda_1 > 0$ for a Hecke congruence subgroup satisfies;

$$\lambda_1 \geq \frac{1}{4} - \left(\frac{7}{64}\right)^2.$$

We define the *singular set of* Γ as;

$$\mathcal{P}_\Gamma := \{s_0 = 1, s_1, \dots, s_m\}, \tag{3.7}$$

where $s_i > 1/2$ and $\lambda(s_i) = \lambda_i, i = 0, \dots, m$ are the exceptional eigenvalues (together with the trivial eigenvalue $\lambda = 0$). When Γ is clear from context, we will shorten notation and write $\mathcal{P} = \mathcal{P}_\Gamma$. The quantity s_1 (where we define $s_1 = 1/2$ if $\mathcal{P} = \{1\}$) will turn out to control the error-terms of our moment calculation in (1.5). Observe that the bound of Kim and Sarnak shows that for Hecke congruence subgroups, we have $\operatorname{Re} s_1 \leq 39/64$.

3.2. **Weight k resolvent operators.** Associated to the weight k Laplacian, we have the associated *resolvent operator*, which defines a meromorphic operator;

$$R(\cdot, k) : \{s \in \mathbb{C} \mid \operatorname{Re} s > 1/2\} \to \mathcal{B}(L^2(\Gamma, k)),$$

where $\mathcal{B}(L^2(\Gamma, k))$ denotes the space of bounded operators on the Hilbert space $L^2(\Gamma, k)$. The resolvent operator is (uniquely) characterized by the property:

$$(\Delta_k - \lambda(s))R(s,k) = \text{Id}_{L^2(\Gamma,k)}, \quad \text{for all } s \in \{s' \in \mathbb{C} \mid \operatorname{Re} s' > 1/2\} \backslash (\mathcal{P} \cup \{1, \dots, k/2\}).$$

The analytic properties of weight k resolvent operators have been studied intensively by Fay in [12]. We will however not use any of these deep results.

It follows from general properties of resolvent operators and (3.6) that $R(s,k)$ defines a meromorphic operator in the half-plane $\operatorname{Re} s > 1/2$ with poles contained in the set $\mathcal{P} \cup \{1, \dots, k/2\}$ (which is why we called $\mathcal{P}$ singular). Furthermore for any $\lambda_0 = \lambda(w_0)$ with $\operatorname{Re} w_0 > 1/2$, we have the following representation in a neighborhood of w_0;

$$R(s,k) = \frac{P_{\lambda_0,k}}{s - w_0} + R_{\text{reg},w_0}(s,k), \tag{3.8}$$

where $P_{\lambda_0,k}$ is the projection to the eigenspace of Δ_k corresponding to the eigenvalue λ_0 (which might be empty) and $R_{\mathrm{reg},w_0}(s,k)$ is regular at $s = w_0$.
Finally we also quote the following useful bound on the norm of the resolvent [16, Appendix A].

Lemma 3.2. *For $s \in \{s' \in \mathbb{C} \mid \mathrm{Re}\, s' > 1/2\} \backslash (\mathcal{P} \cup \{1, \ldots, k/2\})$, we have*

$$||R(s,k)|| \leq \frac{1}{\mathrm{dist}(\lambda(s), \mathrm{spec}(\Delta_k))},$$

where $|| \cdot ||$ is the operator norm and $\mathrm{dist}(\cdot,\cdot)$ *is the distance function.*

3.3. **Additive twists of cuspidal L-functions.** Fix a discrete and co-finite subgroup Γ of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at infinity of width 1 and let $f \in \mathcal{S}_k(\Gamma)$ be a cusp form of even weight k with Fourier expansion (at ∞) given by

$$f(z) = \sum_{n \geq 1} a_f(n) q^n.$$

Then we define the *additive twist* (by $r \in \mathbb{R}$) of the L-function of f;

$$L(f,r,s) := \sum_{n \geq 1} \frac{a_f(n) e(nr)}{n^s},$$

where $e(z) = e^{2\pi i z}$. We also define $L(f, \infty, s) \equiv 0$.
For all $r \in \mathbb{R}$, the above Dirichlet series converges absolutely for $\mathrm{Re}\, s > (k+1)/2$ by Hecke's bound [16, Theorem 3.2];

$$\sum_{n \leq X} |a_f(n)|^2 \ll_f X^k. \tag{3.9}$$

Furthermore if r corresponds to a cusp (i.e. r is fixed by a parabolic subgroup of Γ), then $L(f,r,s)$ admits analytic continuation to the entire complex plane.
Associated to additive twists by real numbers of the form $a/c = \gamma\infty$ with $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, we define the completed L-function as;

$$\Lambda(f, a/c, s) := \Gamma(s) \left(\frac{c}{2\pi}\right)^s L(f, a/c, s).$$

These completed L-functions admit analytic continuation, which satisfy the following functional equation (generalizing [18, Lemma 1.1], see also [20, Section A.3]).

Proposition 3.3. *For $\gamma \in \Gamma$, the completed L-function $\Lambda(f, a/c, s)$ admits analytic continuation to the entire complex plane, which satisfies the functional equation*

$$\Lambda(f, a/c, s) = (-1)^{k/2} \Lambda(f, -d/c, k - s),$$

where $a/c = \gamma\infty$ and $-d/c = \gamma^{-1}\infty$.

Proof. We mimic Hecke's proof of analytic continuation and functional equation of cuspidal L-functions. In the range of absolute convergence of $L(f, a/c, s)$, we have the following period integral representation;

$$\int_0^\infty f(a/c + iy/c) y^s \frac{dy}{y} = \Lambda(f, a/c, s). \tag{3.10}$$

If we let $z_\gamma := -d/c - 1/(iyc)$ then we have the following two relations;

$$\gamma z_\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} z_\gamma = a/c + iy/c, \quad j(\gamma, z_\gamma) = -(iy)^{-1}.$$

This yields

$$\begin{aligned} &\int_0^\infty f(a/c + iy/c) y^s \frac{dy}{y} \\ &= \int_0^1 f(a/c + iy/c) y^s \frac{dy}{y} + (-1)^{k/2} \int_0^1 f(-d/c + iy/c) y^{k-s} \frac{dy}{y}, \end{aligned} \tag{3.11}$$

using modularity of f and a change of variable $y \mapsto 1/y$. Now we get analytic continuation to the entire complex plane by the vanishing of f at the cusps $a/c = \gamma\infty$ and $-d/c = \gamma^{-1}\infty$. Furthermore (3.11) yields the functional equation immediately. This completes the proof. □

Remark 3.4. In the special case $\Gamma = \Gamma_0(q)$, the above proposition applies to additive twists by rational numbers a/c where $(a, c) = 1$ and $q|c$. The functional equation for twists by general rational numbers is much more involved, see [8, Theorem 3.1].

3.3.1. *The convexity bound for additive twists.* As a basic application of the functional equation, we will derive a preliminary bound for the central value ($s = k/2$) of additive twists by a/c, which are Γ-equivalent to ∞, using the Phragmén–Lindelöf principle [17, Theorem 5.53]. This is known as the *convexity bound.*

By the absolute convergence of $L(f, a/c, s)$ for $\operatorname{Re} s > (k+1)/2$, we get for any $\varepsilon > 0$

$$\Lambda(f, a/c, (k+1)/2 + \varepsilon + it) \ll_\varepsilon c^{k/2+1/2+\varepsilon},$$

where the implied constant also depends on f (here we also use Stirling's approximation, which shows that $\Gamma(k/2 + 1/2 + \varepsilon + it)$ is bounded in t). By the functional equation, we derive similarly that

$$\Lambda(f, a/c, (k-1)/2 - \varepsilon + it) \ll_\varepsilon c^{k/2+1/2+\varepsilon}.$$

Finally by the period integral representation (3.10), we get the bound

$$\Lambda(f, a/c, s) \ll_c 1,$$

for $(k-1)/2 - \varepsilon \le \operatorname{Re} s \le (k+1)/2 + \varepsilon$.
Thus the Phragmen-Lindelöf principle applies and we conclude that

$$L(f, a/c, k/2) \ll_\varepsilon c^{k/2+1/2+\varepsilon} c^{-k/2} = c^{1/2+\varepsilon}. \tag{3.12}$$

Although this is a crude bound, it shows together with

$$\#\left\{(a, c) \mid 0 \le a < c, 0 < c \le X, \begin{pmatrix} a & * \\ c & * \end{pmatrix} \in \Gamma\right\} \ll_\Gamma X^2$$

(see [16, (2.37)]) that the main generating series $D^{m,n}(f, s)$, defined in (2.1) for $\Gamma = \Gamma_0(q)$ and in (5.4) for general Γ, converges absolutely (and locally uniformly) in some half-plane $\operatorname{Re} s \gg_{m,n} 1$ to an analytic function.We will later see (Corollary 5.8) that in fact additive twists satisfy a Lindelöf type bound; $L(f, a/c, k/2) \ll_\varepsilon c^\varepsilon$ for all $\varepsilon > 0$.

3.3.2. *Additive twists and the Eichler–Shimura isomorphism.* The additively twisted L-functions show up in many papers in the analytic theory of L-functions in the disguise of the Voronoi-summation formula (which is equivalent to the functional equation for additive twists), but they also have arithmetic significance in themselves as they appear in the *Eichler–Shimura isomorphism.* We recall how this isomorphism is constructed following [33, Section 8.2].

Let f be a cusp form of weight k and level N and let $\gamma \in \Gamma_0(N)$. Then we can associate the following $(k-1)$-dimensional real vector

$$u_f(\gamma) = \left(\mathrm{Re}\int_{\gamma\infty}^{\infty} f(z)dz, \mathrm{Re}\int_{\gamma\infty}^{\infty} f(z)zdz, \ldots, \mathrm{Re}\int_{\gamma\infty}^{\infty} f(z)z^{k-2}dz\right).$$

The map $u_f : \Gamma \to \mathbb{R}^{k-1}$ defines a parabolic co-cycle in group cohomology, i.e. an element of $Z^1_P(\Gamma, X)$ in Shimura's notation where $X = \mathbb{R}^{k-1}$ is a certain Γ-module. From this we get a map

$$f \mapsto \{\text{cohomology class of } u_f\} \in H^1_P(\Gamma, X),$$

which by [33, Theorem 8.4] induces an $\mathbb{R}$-linear isomorphism from $S_k(\Gamma)$ to the parabolic cohomology group $H^1_P(\Gamma, X)$. This is what is known as the Eichler–Shimura isomorphism and it can also be described in terms of the *period polynomials*, which were introduced by M. Eichler [11];

$$\sigma_{f,\gamma}(X) := \frac{1}{(k-1)!}\int_{\gamma\infty}^{\infty} f(z)(z-X)^{k-2}dz.$$

Note that the entries of $u_f(\gamma)$ are the real parts of the coefficients of $\sigma_{f,\gamma}(X)$, up to a scaling by binomial-coefficients. The theory of period polynomials has been used to prove important rationality results for (multiplicative twists of) cuspidal L-functions [22].

Now for any $0 \leq l \leq k-2$, we have

$$\begin{aligned}\int_{\gamma\infty}^{\infty} f(z)z^l dz &= i\int_0^{\infty} f(a/c+iy)(a/c+iy)^l dy \\ &= \sum_{j=0}^{l}\binom{l}{j}(a/c)^{l-j}\frac{j!}{(-2\pi i)^{j+1}}L(f, a/c, j+1), \qquad (3.13)\end{aligned}$$

which shows that the special values of additive twists encode the Eichler–Shimura isomorphism. This formula was the starting point for the author in [27], where the distribution of the Eichler–Shimura map is determined.

4. Poincaré series defined from antiderivatives of cusp forms

In this section, we will construct a certain Poincaré series $G_{A,B,l}(z,s)$ starting from a fixed holomorphic cusp form. Then we will study the analytic properties of these Poincaré series, which will be crucial in proving our main results. The method introduced for studying these Poincaré series might have independent interest.

Let Γ be a co-finite, discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 and $f \in \mathcal{S}_k(\Gamma)$ a cusp form of even weight k. Then we define for $n \geq 1$;

$$I_n(z) = I_n(z;f) := \int_{i\infty}^{z} \int_{i\infty}^{z_{n-1}} \cdots \int_{i\infty}^{z_2} \int_{i\infty}^{z_1} f(z_0) dz_0 dz_1 \ldots d_{z_{n-1}}$$

and $I_0(z) := f(z)$. It is clear that we have $I'_{n+1} = I_n$ and thus I_n is the n-fold antiderivative of f which vanishes at ∞. By taking derivatives, we see that

$$I_n(z) = \frac{(-1)^{n-1}}{(n-1)!} \int_{i\infty}^{z} f(w)(w-z)^{n-1} dw. \tag{4.1}$$

Furthermore we let A, B denote two multi-sets (sets where elements have multiplicities) with all elements contained in $\{0, \ldots, k/2\}$. We call such a multi-set *positive* if all elements are positive or if the multi-set is empty. We let $|A|$ and $|B|$ denote the sizes of the multi-sets counted with multiplicity.
For A, B multi-sets of the above type and l an even integer, we define

$$G_{A,B,l}(z,s) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j_\gamma(z)^{-l} \left(\prod_{a \in A} \frac{I_a(\gamma z)}{(-2i)^a} \right) \left(\prod_{b \in B} \frac{\overline{I_b(\gamma z)}}{(2i)^b} \right) \mathrm{Im}(\gamma z)^{s+\alpha(A,B)}, \tag{4.2}$$

where

$$\alpha(A,B) := \left(\sum_{a \in A} k/2 - a \right) + \left(\sum_{b \in B} k/2 - b \right).$$

We will see below that these series converges absolutely when $\mathrm{Re}\, s \gg 1$. We observe that by (3.1) these Poincaré series are (formally) automorphic;

$$G_{A,B,l}(\gamma z, s) = j_\gamma(z)^l G_{A,B,l}(z,s), \quad \gamma \in \Gamma.$$

The scaling $\alpha(A,B)$ has the nice property that

$$G_{A \cup \{0\},B,l}(z,s) = y^{k/2} f(z) G_{A,B,l-k}(z,s), \tag{4.3}$$

$$G_{A,B \cup \{0\},l}(z,s) = y^{k/2} \overline{f(z)} G_{A,B,l+k}(z,s), \tag{4.4}$$

which follows from the modularity of f.
Observe that with A and B as above, we always have $\alpha(A,B) \geq 0$, which will be crucial in many argument. We also have the following symmetry

$$\overline{G_{A,B,l}(z,s)} = G_{B,A,-l}(z,\overline{s}). \tag{4.5}$$

This shows that it is enough to consider the case $l \geq 0$.

Firstly we will show that (4.2) defines an element of $L^2(\Gamma, l)$ in some half-plane following unpublished work of Chinta and O'Sullivan [6].

Lemma 4.1. *For $|A| + |B| > 0$ the series $G_{A,B,l}(z,s)$ converges absolutely (and locally uniformly in z and s) in the half-plane*

$$\mathrm{Re}\, s > 1 + |A| + |B|$$

to an element of $L^2(\Gamma, l)$.

Proof. By Hecke's bound on the coefficients of cusp forms $|a_f(n)| \ll n^{k/2}$ (coming from (3.9)), we have

$$I_n(z) \ll \sum_{m=1}^{\infty} m^{k/2-n} e^{-2\pi m y} \ll \frac{(k/2-n)!}{y^{k/2+1-n}}, \tag{4.6}$$

using that $r! = \Gamma(r+1) = \int_0^\infty e^{-x} x^r dx$. This gives

$$\begin{aligned} &\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \left| j_\gamma(z)^{-l} \left(\prod_{a\in A} \frac{I_a(\gamma z)}{(-2i)^a}\right) \left(\prod_{b\in B} \frac{\overline{I_b(\gamma z)}}{(2i)^b}\right) \mathrm{Im}(\gamma z)^{s+\alpha(A,B)} \right| \\ &\ll_{A,B} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^{\sigma-|A|-|B|} \\ &= E\left(z, \sigma-|A|-|B|\right), \end{aligned} \tag{4.7}$$

where $s = \sigma + it$. Since the non-holomorphic Eisenstein series converges absolutely for $\mathrm{Re}\, s > 1$, we get that $G_{A,B,l}(z,s)$ convergences absolutely (and locally uniformly in s and z) in the desired half-plane.
For any cusp $\mathfrak{b}$ of Γ we have by [29, Lemma 3.2]

$$\sum_{\mathrm{Id}\neq\gamma\in\Gamma_\infty\backslash\Gamma} (\mathrm{Im}\,\gamma\sigma_{\mathfrak{b}} z)^w \ll y^{1-\mathrm{Re}\, w},$$

as $y \to \infty$ and uniformly for $\mathrm{Re}\, w \in [1+\varepsilon, Y]$ with $Y > 1$ and $\varepsilon > 0$. Thus we get using (4.7) the following bound at any cusp $\mathfrak{b}$ of Γ and $\sigma = \mathrm{Re}\, s \in [1+\varepsilon, Y]$

$$\begin{aligned} &G_{A,B,l}(\sigma_{\mathfrak{b}} z, s) \\ &\ll_{A,B} \left(\prod_{a\in A} |I_a(\sigma_{\mathfrak{b}} z)|\right) \left(\prod_{b\in B} |I_b(\sigma_{\mathfrak{b}} z)|\right) \mathrm{Im}(\sigma_{\mathfrak{b}} z)^{\sigma+\alpha(A,B)} + y^{1-(\sigma-|A|-|B|)}. \end{aligned}$$

Furthermore (4.6) shows that $I_n(z) \ll e^{-\pi y}$ as $y \to \infty$. At other cusps we see that $I_n(\sigma_{\mathfrak{b}} z)$ is bounded as $y \to \infty$ and we conclude that the contribution from the identity is $\ll y^{-\sigma}$ as $y \to \infty$ in these cases. Thus we conclude for $\mathrm{Re}\, s > 1 + |A| + |B|$ that $G_{A,B,l}(z,s) \to 0$, as $z \to \mathfrak{b}$ for any cusp $\mathfrak{b}$. This implies that $G_{A,B,l}(z,s) \in L^2(\Gamma, l)$. □

4.1. **The recursion formula.** In order to understand the pole structure of $G_{A,B,l}(z,s)$, we will use certain recursion formulas involving the resolvent and the raising and lowering operators. First of all we will record how the raising and lowering operators act on the constituents of Poincaré series.

Lemma 4.2. *Let $h : \mathbb{H} \to \mathbb{C}$ be a smooth function and l an even integer. Then we have*

$$\begin{aligned} &K_l[h(\gamma z)\operatorname{Im}(\gamma z)^s j_\gamma(z)^{-l}] \\ &\qquad= \left(2i\frac{\partial h}{\partial z}(\gamma z)\operatorname{Im}(\gamma z)^{s+1} + (s+l/2)h(\gamma z)\operatorname{Im}(\gamma z)^s\right) j_\gamma(z)^{-l-2}, \\ &L_l[h(\gamma z)\operatorname{Im}(\gamma z)^s j_\gamma(z)^{-l}] \\ &\qquad= \left(2i\frac{\partial h}{\partial \overline{z}}(\gamma z)\operatorname{Im}(\gamma z)^{s+1} - (s-l/2)h(\gamma z)\operatorname{Im}(\gamma z)^s\right) j_\gamma(z)^{-l+2} \end{aligned}$$

for any $\gamma \in \mathrm{PSL}_2(\mathbb{R})$.

Proof. Using the intertwining relation;

$$K_l\left(j_\gamma(z)^{-l}F(\gamma z)\right) = j_\gamma(z)^{-l-2}(K_l F)(\gamma z),$$

valid for any smooth function $F : \mathbb{H} \to \mathbb{C}$, we reduce the problem to proving the following identity;

$$K_l h(z)y^s \stackrel{?}{=} \left(y\left(i\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) + \frac{k}{2}\right) h(z)y^s,$$

and similar for the lowering operator. This can be done by a straightforward calculation. □

This yields the following useful formula.

Corollary 4.3. *Let $h : \mathbb{H} \to \mathbb{C}$ be a smooth function and l an even integer. Then we have*

$$\begin{aligned} (\Delta_l - \lambda(s))[h(\gamma z)\operatorname{Im}(\gamma z)^s j_\gamma(z)^{-l}] =& -4\frac{\partial^2 h}{\partial z \partial \overline{z}}(\gamma z)\operatorname{Im}(\gamma z)^{s+2} j_\gamma(z)^{-l} \\ &- 2i(s-l/2)\frac{\partial h}{\partial z}(\gamma z)\operatorname{Im}(\gamma z)^{s+1} j_\gamma(z)^{-l} \\ &+ 2i(s+l/2)\frac{\partial h}{\partial \overline{z}}(\gamma z)\operatorname{Im}(\gamma z)^{s+1} j_\gamma(z)^{-l} \end{aligned}$$

for $s \in \mathbb{C}$ and $\gamma \in \mathrm{SL}_2(\mathbb{R})$.

From the above we will deduce the main recursion formula which will allow us to inductively understand the pole structure of $G_{A,B,l}(z,s)$. To write down the formula we will introduce the following convenient notation for $a \in A$;

$$A_a := (A \backslash \{a\}) \cup \{a-1\}.$$

In this notation we have for positive multisets A the following useful relation;

$$\frac{\partial}{\partial z}\prod_{a\in A} I_a(z) = \sum_{a\in A}\prod_{a'\in A_a} I_{a'}(z),$$

by the Leibniz rule. Thus by summing over $\gamma \in \Gamma_\infty \backslash \Gamma$ and using Lemma 4.1, we arrive at the following lemma.

Lemma 4.4. *Let A, B be positive multisets and $G_{A,B,l}(z,s)$ as above. Then we have*

$$K_l G_{A,B,l}(z,s) = (s + \alpha(A,B) + l/2) G_{A,B,l+2}(z,s) - \sum_{a \in A} G_{A_a,B,l+2}(z,s), \tag{4.8}$$

$$L_l G_{A,B,l}(z,s) = -(s + \alpha(A,B) - l/2) G_{A,B,l-2}(z,s) + \sum_{b \in B} G_{A,B_b,l-2}(z,s) \tag{4.9}$$

and

$$\begin{aligned} G_{A,B,l}(z,s) = R(s+\alpha(A,B), l) \Bigg(&- \sum_{a \in A, b \in B} G_{A_a,B_b,l}(z,s) \\ &+ (s + \alpha(A,B) - l/2) \sum_{a \in A} G_{A_a,B,l}(z,s) \\ &+ (s + \alpha(A,B) + l/2) \sum_{b \in B} G_{A,B_b,l}(z,s) \Bigg), \end{aligned} \tag{4.10}$$

valid apriori for $\operatorname{Re} s > 1 + |A| + |B|$.

This lemma will turn out to be extremely useful.

Remark 4.5. The recursion formula (4.10) is the reason why we have $2i$ and $-2i$ in the denominators in the definition of $G_{A,B,l}(z,s)$ and why we have the shift $\alpha(A,B)$.

Define *the total weight* of A, B (and of $G_{A,B,l}(z,s)$) as the quantity

$$\Sigma(A,B) := \sum_{a \in A} a + \sum_{b \in B} b.$$

Then we observe that all Poincaré series on the right-hand side in the recursion formula (4.10) have strictly smaller total weight than the one on the left-hand side. This will allow us to do an inductive argument on the total weight, when determining the pole structure of the Poincaré series.

As a first application of Lemma 4.4, we will show meromorphic continuation of $G_{A,B,l}(z,s)$ to $\operatorname{Re} s > 1/2$. Firstly we will handle the case $l = 0$ using (4.10). This case is easiest to handle since the poles of $R(s,0)$ all satisfy $\operatorname{Re} s \leq 1$. Then we will use (4.8) and (4.9) to get the result for general (even) weights l.

Proposition 4.6. *Let A, B be two multi-sets such that $|A| + |B| > 0$ and l an even integer. Then the Poincaré series $G_{A,B,l}(z,s)$ admits meromorphic continuation to the half-plane* $\operatorname{Re} s > 1/2$ *satisfying the following;*

(i) $G_{A,B,l}(z,s)$ defines an element of $L^2(\Gamma, l)$ at all regular points.
(ii) The poles of $G_{A,B,l}(z,s)$ in $1/2 < \operatorname{Re} s \leq 1$ are contained in $\mathcal{P}$ (defined as in (3.7)).
(iii) $G_{A,B,l}(z,s)$ is regular for $\operatorname{Re} s > 1$.

Proof. We prove the claims by an induction on the total weight $\Sigma(A,B)$. If $\Sigma(A,B) = 0$ then by (4.3) and (4.4), we can write

$$G_{A,B,l}(z,s) = (y^{k/2} f(z))^{|A|} (y^{k/2} \overline{f(z)})^{|B|} E_{l-k(|A|-|B|)}(z,s). \tag{4.11}$$

Now $E_{l-k(|A|-|B|)}(z,s)$ is meromorphic in $\operatorname{Re} s > 1/2$ with poles contained in $\mathcal{P}$ and is regular for $\operatorname{Re} s > 1$ (see [10, Chapter 4]). Furthermore since $f(z)$ decays rapidly at all cusps, the above defines an element of $L^2(\Gamma, l)$ at all regular points.

Now assume $\Sigma(A,B) > 0$. By using (4.3) and (4.4) we may assume that A, B are positive multi-sets. Firstly we consider the case $l = 0$.
By (4.10), we can write

$$\begin{aligned}&G_{A,B,0}(z,s)\\ =&R(s+\alpha(A,B),0)\left(\textit{linear combinations of } G_{A',B',l}(z,s)\textit{'s with } \Sigma(A',B') < \Sigma(A,B)\right),\end{aligned}$$

where by the induction hypothesis, all terms inside the parenthesis satisfy the properties $(i), (ii), (iii)$. Since the resolvent operator $R(s+\alpha(A,B),0)$ is regular in the half-plane $\operatorname{Re} s > 1$ and meromorphic in $\operatorname{Re} s > 1/2$ with poles contained in $\mathcal{P}$, the wanted properties follow for $G_{A,B,0}(z,s)$ as well. Observe that, if $\alpha(A,B) \geq 1$, then the resolvent is actually regular for $\operatorname{Re} s > 1/2$.

Now to get the claim for all positive weights l, we do an induction on the weight. For $l \geq 0$, the identity (4.8) gives

$$G_{A,B,l+2}(z,s) = \frac{K_l G_{A,B,l}(z,s) + \sum_{a\in A} G_{A_a,B,l+2}(z,s)}{(s+\alpha(A,B)+l/2)}.$$

We know by the induction hypothesis that all the Poincaré series on the right-hand side of the above satisfy $(i), (ii), (iii)$ of this proposition. So since $s+\alpha(A,B)+l/2$ is non-zero for $\operatorname{Re} s > 1/2$, we see that also $G_{A,B,l+2}(z,s)$ satisfies $(i), (ii), (iii)$.
A similar argument applies to negative weights using (4.9). This finishes the induction and thus the proof. □

This allows us to extend the range of validity of Lemma 4.4 by uniqueness of analytic continuation.

Corollary 4.7. *The equations (4.8), (4.9) and (4.10) are valid in the half-plane* $\operatorname{Re} s > 1/2$ *as equalities of meromorphic functions.*

4.2. **Bounds on the pole order at** $s = 1$. Next step is to determine the pole order at $s = 1$ of $G_{A,B,l}(z,s)$. In this section we will prove certain bounds on the pole order. We will proceed by induction relying on the formulas (4.8), (4.9) and (4.10). We firstly need the following key lemma.

Lemma 4.8. *Let* A, B *be positive multi-sets and* l *an even integer. Then we have for* $l \geq 0$;

$$\langle G_{A,B,lk}(z,s), (y^{k/2}f(z))^l\rangle = \frac{\sum_{a\in A}\langle G_{A_a,B,lk}(z,s), (y^{k/2}f(z))^l\rangle}{s+\alpha(A,B)+lk/2-1}, \tag{4.12}$$

and for $-l \leq 0$;

$$\langle G_{A,B,-lk}(z,s), (y^{k/2}\overline{f(z)})^l\rangle = \frac{\sum_{b\in B}\langle G_{A,B_b,-lk}(z,s), (y^{k/2}\overline{f(z)})^l\rangle}{s+\alpha(A,B)+lk/2-1}. \tag{4.13}$$

Proof. Assume $l \geq 0$ then by the identity (4.9), we have

$$\langle G_{A,B,lk}(z,s), (y^{k/2}f(z))^l \rangle$$
$$= \frac{\langle K_{kl-2}G_{A,B,lk-2}(z,s) + \sum_{a\in A} G_{A_a,B,lk}(z,s), (y^{k/2}f(z))^l \rangle}{s + \alpha(A,B) + lk/2 - 1}.$$

By the adjointness properties of the raising and lowering operators (3.2), we get

$$\langle K_{kl-2}G_{A,B,lk-2}(z,s), (y^{k/2}f(z))^l \rangle = \langle G_{A,B,lk-2}(z,s), L_{lk}(y^{k/2}f(z))^l \rangle = 0,$$

using (3.5). This yields the desired formula. The case $-l \leq 0$ is proved similarly using (4.8). □

From this we conclude the following key result.

Proposition 4.9. *The pole order of $G_{A,B,l}(z,s)$ at $s = 1$ is bounded by*

$$\min(\#\{a \in A \mid a = k/2\}, \#\{b \in B \mid b = k/2\}) + 1.$$

Proof. We will do an induction on the total weight $\Sigma(A,B)$. If $\Sigma(A,B) = 0$ then the result is clear by the properties of the non-holomorphic Eisenstein series. In general by applying modularity (as in (4.3) and (4.4)), we may assume that both A and B are positive. By the symmetry (4.5) we may also assume that $|A| \geq |B|$.

We proceed by induction on the total weight; assume that the total weight is positive; $\Sigma(A,B) > 0$ and that we have proved the claim for all smaller $\Sigma(A,B)$-values.

We begin with the case $l = 0$. The recursion formula (4.10) gives the following;

$$G_{A,B,0}(z,s) = R(s + \alpha(A,B), 0)\left(-\sum_{a\in A, b\in B} G_{A_a,B_b,l}(z,s) + \cdots\right),$$

where the terms inside the parenthesis satisfy the claim of the proposition by the induction hypothesis. If $\alpha(A,B) > 0$ then the claim also follows for $G_{A,B,0}(z,s)$, since the resolvent operator $R(s + \alpha(A,B), 0)$ is regular at $s = 1$.

If $\alpha(A,B) = 0$, then we must have

$$A = \underbrace{\{k/2, \ldots, k/2\}}_{n}, \quad B = \underbrace{\{k/2, \ldots, k/2\}}_{m}$$

for some $n \geq m \geq 0$.

Now we claim that $\langle G_{A,B,0}(z,s), 1 \rangle$ has a pole of order at most $m + 1$.

To see this we do an induction on m. If $m = 0$, then by Lemma 4.8, we see directly that

$$\langle G_{A,B,0}(z,s), 1 \rangle = 0.$$

If $m > 0$ then we get by Lemma 4.8

$$\langle G_{A,B,0}(z,s), 1 \rangle = \frac{m\langle G_{A,B_{k/2},0}(z,s), 1 \rangle}{s - 1}$$

and by the induction hypothesis, $G_{A,B_{k/2},0}(z,s)$ has a pole of order at most m, which proves the claim.

We observe that if $G_{A,B,0}(z,s)$ has a pole of order greater than $m + 1$, then by

(4.10) and the induction hypothesis there has to be an increase in the pole order coming from the pole in the singular expansion of the resolvent (3.8). This implies that the leading Laurent coefficient is constant. But we just showed that $\langle G_{A,B,0}(z,s),1\rangle$ has a pole of order at most $m+1$. This finishes the induction in the case $l=0$.

By using (4.8) and (4.9) as in the proof of Proposition 4.6, we get by induction the pole bound for all even weights l as well. This finishes the induction and hence the proof. □

4.3. **Finding the leading pole.** For $m \neq n$, Proposition 4.9 yields the desired bound needed to prove Theorem 1.1 (see (5.11) below). Next step is to determine the exact pole order and leading Laurent coefficient of $G_{A,A,0}(z,s)$ at $s=1$ when

$$A = \{\underbrace{k/2, \ldots, k/2}_{n}\}.$$

By Proposition 4.9 the pole order is bounded by $n+1$ and we will see that this bound is sharp.

Theorem 4.10. *Let*

$$A = \{\underbrace{k/2, \ldots, k/2}_{n}\}.$$

Then $G_{A,A,0}(z,s)$ has a pole of order $n+1$ at $s=1$ with leading Laurent coefficient

$$\frac{(n!)^2 ||f||^{2n}}{((k-1)!)^n \mathrm{vol}(\Gamma)^{n+1}}.$$

Proof. We do an induction on n. For $n=0$ the claim follows by the analytic properties of the non-holomorphic Eisenstein series [16, (6.33)].
Now assume $n \geq 1$. First of all we see by (4.10) that

$$\begin{aligned} &G_{A,A,0}(z,s) \\ &= R(s,0)\left(-n^2 G_{A_{k/2},A_{k/2},0}(z,s) + nsG_{A_{k/2},A,0}(z,s) + nsG_{A,A_{k/2},0}(z,s)\right) \end{aligned}$$

By the bounds on the pole order from Proposition 4.9, all the terms inside the parentheses above have a pole of order at most n. This shows that if $G_{A,A,0}(z,s)$ has a pole of order $n+1$, then the leading pole is contained in the image under the projection onto the constant subspace, since (as above) the increase in the pole order has to come from the resolvent.
We will show that indeed

$$\langle G_{A,B,0}(z,s),1\rangle / \langle 1,1\rangle,$$

has a pole of order $n+1$ at $s=1$ with the claimed leading Laurent coefficient.

Applying Lemma 4.8 twice and using the pole bound from Proposition 4.9, we get;

$$\begin{aligned}&\langle G_{A,A,0}(z,s),1\rangle\\&=\frac{\langle nG_{A,A_{k/2},0}(z,s),1\rangle}{s-1}\\&=\frac{\langle n\sum_{a\in A_{k/2}}G_{A,(A_{k/2})_a,0}(z,s),1\rangle}{(s-1)s}\\&=\frac{n\langle G_{A,(A_{k/2})_{k/2-1},0}(z,s),1\rangle+(\textit{pole of order at most } n-1 \textit{ at } s=1)}{(s-1)s},\end{aligned}$$

where

$$(A_{k/2})_{k/2-1}=\{\underbrace{k/2,\dots,k/2}_{n-1},k/2-2\}.$$

By repeated applications of Lemma 4.8 (and Proposition 4.9), we arrive at

$$\langle G_{A,A,0}(z,s),1\rangle=\frac{n\langle G_{A,A'\cup\{0\},0}(z,s),1\rangle+(\textit{pole of order at most } n-1 \textit{ at } s=1)}{(s-1)s\cdots(s+k/2-2)}$$

where

$$A'=\{\underbrace{k/2,\dots,k/2}_{n-1}\}.$$

Now by applying modularity as in (4.4), we get

$$\langle G_{A,A'\cup\{0\},0}(z,s),1\rangle=\langle G_{A,A',k}(z,s),y^{k/2}f(z)\rangle.$$

By a similar repeated application of Lemma 4.8 (now with $l=k$), we arrive at

$$\begin{aligned}&\langle G_{A,A,0}(z,s),1\rangle\\&=\frac{n^2\langle G_{A',A',k}(z,s),y^k|f(z)|^2\rangle+(\textit{pole of order at most } n-1 \textit{ at } s=1)}{(s-1)s\cdots(s+k/2-2)\cdot(s+k/2-1)\cdots(s+k-2)}.\end{aligned}$$

By the induction hypothesis, we know that $G_{A',A',0}(z,s)$ has a pole of order n at $s=1$ with leading Laurent coefficient given by

$$\frac{((n-1)!)^2||f||^{2n-2}}{((k-1)!)^{n-1}\mathrm{vol}(\Gamma)^n}.$$

Thus we see that

$$\begin{aligned}\langle G_{A,A,0}(z,s),1\rangle/\langle 1,1\rangle=&\frac{n^2\left\langle\frac{((n-1)!)^2||f||^{2n-2}}{((k-1)!)^{n-1}\mathrm{vol}(\Gamma)^n},y^k|f(z)|^2\right\rangle}{(k-1)!(s-1)^{n+1}\mathrm{vol}(\Gamma)}\\&+(\textit{pole of order at most } n \textit{ at } s=1),\end{aligned}$$

which yields the wanted. ☐

With this theorem established we can improve Proposition 4.9 in the following special case.

Corollary 4.11. *Let*

$$A = \{\underbrace{k/2, \dots, k/2}_{n}\}$$

and $l \neq 0$ a non-zero even integer. Then the order of the pole of $G_{A,A,l}(z,s)$ at $s=1$ is at most n.

Proof. By the symmetry (4.5), it is enough to prove it for $l > 0$. We prove it by induction on l. For $l = 2$ we get by (4.9)

$$G_{A,A,2}(z,s) = \frac{K_0 G_{A,A}(z,s) + nG_{A_{k/2},A,2}(z,s)}{s}.$$

From Theorem 4.10 we know that the leading Laurent coefficient of $G_{A,A,0}(z,s)$ is constant, and thus it is annihilated by K_0. Furthermore we know by Proposition 4.9 that $G_{A_{k/2},A,2}(z,s)$ has a pole of order at most n at $s = 1$. Thus we conclude that also $G_{A,A,2}(z,s)$ has a pole of order at most n at $s = 1$.
Now assume $l > 2$. We get again by (4.9) the following;

$$G_{A,A,l+2}(z,s) = \frac{K_l G_{A,A,l}(z,s) + nG_{A_{k/2},A,l+2}(z,s)}{s + l/2}.$$

Thus by the induction assumption and Proposition 4.9, we see that also $G_{A,A,l+2}(z,s)$ has a pole of order at most n at $s = 1$. This finishes the induction and hence the proof. □

4.4. **Growth on vertical lines.** In this section we will prove bounds on the L^2-norm of $G_{A,B,l}(z,s)$ with s in a horizontal strip, bounded away from the singular set $\mathcal{P}$. This we will use to get bounds on vertical lines for the main generating series $D^{m,n}(f,s)$ defined in (2.1), which is needed in order to apply Theorem A.2 in the appendix.
We will firstly consider the case of total weight zero; $\Sigma(A,B) = 0$. We will use the idea used in the proof of [28, Lemma 3.1]. Following Petridis and Risager, we will in the proof assume that Γ has only one cusp for simplicity. The same argument applies in the general case.

Lemma 4.12. *Let $\varepsilon > 0$ and $s = \sigma + it$ satisfying $1/2 + \varepsilon \leq \sigma \leq 3/2$ and $\mathrm{dist}(s, \mathcal{P}) \geq \varepsilon$. Let A, B be multi-sets such that $|A| + |B| > 0$ and $\Sigma(A,B) = 0$ and let l be an even integer. Then we have the following bound;*

$$||G_{A,B,l}(z,s)|| \ll_\varepsilon 1,$$

where the implied constant might depend on $|A|, |B|, l$.

Proof. By the assumption $\Sigma(A,B) = 0$, we can write

$$G_{A,B,l}(z,s) = (y^{k/2} f(z))^{|A|} (y^{k/2} \overline{f(z)})^{|B|} E_{l'}(z,s),$$

with l' appropriately adjusted.
Let $\mathcal{F}$ be fundamental domain for $\Gamma\backslash\mathbb{H}$ with a cusp at infinity. For $\mathrm{Re}\, s > 1/2$ and $z \in \mathcal{F}$, we write (following Colin de Verdière [7]);

$$E_{l'}(z,s) = h(y) y^s + g(z,s),$$

where $g(z,s) \in L^2(\mathcal{F})$ and $h(y) \in C^\infty(0,\infty)$ is smooth with $h(y) = 1$ near the cusp at ∞.
Since $E_{l'}(z,s)$ is a formal eigenfunction for the Laplacian, we have

$$\begin{aligned}(\Delta_{l'} - \lambda(s))g(z,s) &= (\Delta_{l'} - \lambda(s))(E_{l'}(z,s) - h(y)y^s) \\ &= \lambda(s)h(y)y^s + sh'(y)y^{s+1} + h''(y)y^{s+2} - \lambda(s)h(y)y^s \\ &= sh'(y)y^{s+1} + h''(y)y^{s+2}.\end{aligned}$$

Now we extend $g(z,s)$ periodically to an element of $L^2(\Gamma, l')$. Then the above yields

$$g(z,s) = R(s,l')(sh'(y)y^{s+1} + h''(y)y^{s+2}),$$

i.e. $g(z,s)$ equals the resolvent applied to a function with compact support.
Now by the bound on the norm of the resolvent from Lemma 3.2, we get

$$||g(z,s)|| \leq \frac{||sh'(y)y^{s+1} + h''(y)y^{s+2}||}{\mathrm{dist}(\lambda(s), \mathrm{spec}\Delta_{l'})}.$$

Since $\Delta_{l'}$ is self adjoint, all eigenvalues are real. Thus using the assumption $\mathrm{dist}(s,\mathcal{P}) \geq \varepsilon$, we get

$$\mathrm{dist}(\lambda(s), \mathrm{spec}\Delta_{l'}) \gg |\operatorname{Im}(\lambda(s))| + \varepsilon = (2\sigma - 1)|t| + \varepsilon.$$

This gives

$$||g(z,s)|| \ll \frac{||sh'(y)y^{s+1} + h''(y)y^{s+2}||}{(2\sigma-1)|t| + \varepsilon} \ll \frac{|s|}{|t| + \varepsilon} \ll_\varepsilon 1.$$

Now by the above, we have

$$\begin{aligned}&||G_{A,B,l}(z,s)|| \\ &\leq ||(y^{k/2}f(z))^{|A|}(y^{k/2}\overline{f(z)})^{|B|}h(y)y^s|| + ||(y^{k/2}f(z))^{|A|}(y^{k/2}\overline{f(z)})^{|B|}g(z,s)||.\end{aligned}$$

The second term is bounded by what we showed above and by the rapid decay of f, the first term is bounded uniformly in s as well. Thus we conclude $||G_{A,B,l}(z,s)|| \ll_\varepsilon 1$ as wanted. □

With this done, we can do the general case by induction on the total weight $\Sigma(A,B)$ using the recursion formula (4.10) and the bound on the operator norm of the resolvent in Lemma 3.2.

Proposition 4.13. *Let $\varepsilon > 0$ and $s = \sigma + it$ satisfying $1/2 + \varepsilon \leq \sigma \leq 3/2$ and $\mathrm{dist}(s,\mathcal{P}) \geq \varepsilon$. Let A, B be multi-sets satisfying $|A| + |B| > 0$ and l an even integer. Then we have*

$$||G_{A,B,l}(z,s)|| \ll_\varepsilon 1,$$

where the implied constant depends on $|A|, |B|, l$.

Proof. We proceed by induction. Above we have done the base case so we may assume that $\Sigma(A,B) > 0$. By applying modularity we may assume that A and B are positive (since $y^{k/2}f(z)$ is bounded). Now by (4.10) and Lemma 3.2, we get

$$||G_{A,B,l}(z,s)|| \leq \frac{||\textit{RHS of (4.10)}||}{\mathrm{dist}(\lambda(s + \alpha(A,B)), \mathrm{spec}(\Delta_l))}.$$

By the induction assumption and the triangle inequality, we see that

$$||\textit{RHS of (4.10)}|| \ll_\varepsilon |t| + 1,$$

using

$$s + \alpha(A,B) \pm l/2 \ll |t| + 1,$$

where the implied constant depends on $|A|, |B|, l$.
Now since the spectrum of Δ_l is real, we get

$$\mathrm{dist}(\lambda(s+\alpha(A,B)), \mathrm{spec}(\Delta_l)) \gg_\varepsilon |t(2\sigma - 1)| + \varepsilon \gg_\varepsilon |t| + \varepsilon,$$

using the assumption $\mathrm{dist}(s, \mathcal{P}) \geq \varepsilon$. This gives

$$||G_{A,B,l}(z,s)|| \ll_\varepsilon \frac{|t|+1}{|t|+\varepsilon} \ll_\varepsilon 1,$$

as wanted. □

5. Central values of additive twists

In this section we will use the results from the preceding section to study the central values of additive twists. To state our main theorem in the most general version, we will need to work with more general twists than the ones described in the introduction (as was alluded to in Remark 1.2). To do this we need to introduce some notation:

Given a discrete, co-finite subgroup Γ of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 and two cusps $\mathfrak{a}$ and $\mathfrak{b}$ of Γ (not necessarily distinct), we define the following set (following [29]);

$$T_{\Gamma,\mathfrak{ab}} = T_{\mathfrak{ab}} := \left\{ r = a/c \bmod 1 \mid \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\infty \backslash \sigma_\mathfrak{a}^{-1} \Gamma \sigma_\mathfrak{b} / \Gamma_\infty, c > 0 \right\}, \tag{5.1}$$

where Γ_∞ denotes the parabolic subgroup of Γ fixing ∞ and $\sigma_\mathfrak{a}$ denotes a (fixed) scaling matrix of $\mathfrak{a}$ (see [16, (2.1)] for background). Observe that $T_{\infty\mathfrak{b}}$ contains exactly the additive twists by the cusps Γ-equivalent to $\mathfrak{b}$ (thought of as real numbers).
Any $r \in T_{\mathfrak{ab}}$ uniquely determines an element in the double quotient $\Gamma_\infty \backslash \sigma_\mathfrak{a}^{-1} \Gamma \sigma_\mathfrak{b} / \Gamma_\infty$ [29, Proposition 2.2]. Thus given $r \in T_{\mathfrak{ab}}$, we can define $c(r)$ as the left-lower entry of any such representative. Using this we define

$$T_{\Gamma,\mathfrak{ab}}(X) = T_{\mathfrak{ab}}(X) := \{ r \in T_{\mathfrak{ab}} \mid c(r) \leq X \}. \tag{5.2}$$

We observe that for $\Gamma = \Gamma_0(q)$, we get $T_{\infty\infty}(X) = T(X)$ with $T(X)$ defined as in (1.2). We will below continue to use the shorthand $T(X) = T_{\infty\infty}(X)$, when there is no danger for confusion.

Using this notation we can now state the most general statement that we can prove with our methods.

Theorem 5.1. *Let Γ be a discrete and co-finite subgroup of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at infinity of width 1, $\mathfrak{b}$ a cusp of Γ, k an even integer and $f_1, \ldots, f_d$ an orthogonal basis for the space of weight k cusps forms $S_k(\Gamma)$ with respect to the Petersson inner* *product. Then for any fixed box* $\Omega \subset \mathbb{C}^d$*, we have*

$$\begin{aligned}
&\mathbb{P}_{T_{\infty\mathfrak{b}}(X)}\left(\left(\frac{L(f_i,r,k/2)}{(C_{f_i}\log c(r))^{1/2}}\right)_{1\leq i\leq d}\in\Omega\right)\\
&:=\frac{\#\left\{r\in T_{\infty\mathfrak{b}}(X)\mid \left(\frac{L(f_i,r,k/2)}{(C_{f_i}\log c(r))^{1/2}}\right)_{1\leq i\leq d}\in\Omega\right\}}{\#T_{\infty\mathfrak{b}}(X)}\\
&=\mathbb{P}\left((Y_1,\dots,Y_d)^T\in\Omega\right)+o(1)
\end{aligned}$$

as $X\to\infty$*, where* $Y_1,\dots,Y_d$ *are mutually independent random variables all of which are distributed with respect to the standard complex normal distribution* $\mathcal{N}_{\mathbb{C}}(0,1)$ *and*

$$C_f=\frac{(4\pi)^k||f||^2}{(k-1)!\,\mathrm{vol}(\Gamma)}, \tag{5.3}$$

with $||f||$ *the Petersson-norm of* f *and* $\mathrm{vol}(\Gamma)$ *the hyperbolic volume of* $\Gamma\backslash\mathbb{H}$*. (Here* $\mathbb{P}((Y_1,\dots,Y_d)^T\in\Omega)$ *denotes the probability of the event* $(Y_1,\dots,Y_d)^T\in\Omega$*.)*

Recall from Section 2 that our strategy of proof is to use the method of moments. To obtain asymptotic formulas for the moments of additive twists, we will be studying the associated generating series. For additive twists at arbitrary cusps this generating series is defined as follows;

$$D_{\mathfrak{b}}^{m,n}(f,s):=\sum_{r\in T_{\infty\mathfrak{b}}}\frac{L(f,r,k/2)^m\overline{L(f,r,k/2)}^n}{c(r)^{2s}}. \tag{5.4}$$

We study this generating series by studying the associated Goldfeld Eisenstein series (defined in (5.15) below), which is linked to the Poincaré series $G_{A,B,l}(z,s)$ via a formula for the central values of additive twists that we will prove shortly.

To make the proof more readable, we will restrict to the case of $\mathfrak{b}=\infty$ and a single cusp form $f\in\mathcal{S}_k(\Gamma)$. In Section 5.5 and Section 5.6, we will then explain how the proof can be extended to the general case.

Remark 5.2. To make our argument work, we will need to know apriori that (5.4) converges absolutely in some half-plane $\mathrm{Re}\,s>\sigma_0$. By the argument given in Section 3.3.1, it is enough to show that

$$L(f,r,k/2)\ll c(r)^K,$$

for some $K\geq 0$. Since we do not have a nice functional equation for general additive twists (see the discussion in Remark 3.4 above), the easiest way to achieve this seems to be to combine the (generalized) Birch–Stevens formula (see (6.2) below) with the convexity bound for the twisted central values $L(f,\chi,1/2)$. We will not go into the details, but just make it clear that one can easily show an apriori polynomial bound of $L(f,r,k/2)$. We will later see that actually $L(f,r,k/2)\ll_\varepsilon c(r)^\varepsilon$ for all $\varepsilon>0$ (see Corollary 5.8 below).

5.1. **A formula for the central value.** In this section we will prove the promised formula, which expresses the generalized Goldfeld series $E^{m,n}(z,s)$ defined in (2.2) as a sum of the Poincaré series $G_{A,B,l}(z,s)$ studied in the preceding section. This generalizes to higher weight what Bruggeman and Diamantis [4] for weight 2 call *automorphic completion.*

So let Γ be a discrete and co-finite subgroup with a cusp at ∞ of width 1 and fix a cusp form $f\in\mathcal{S}_k(\Gamma)$ of even weight k. Then we will be interested in the central values $L(f,r,k/2)$ of the additive twists by $r\in T=T_{\infty\infty}$, which we will try to relate to the anti-derivatives of f (denoted by I_n above).
The starting point is the period integral representation of $L(f,\gamma\infty,s)$ with $\gamma\in\Gamma$ given in (3.11). A slight variation of this yields with $a/c=\gamma\infty$ the following;

$$i(2\pi)^{-k/2}\Gamma(k/2)L(f,a/c,k/2)=\int_{\gamma\infty}^{i\infty} f(w)\left(\frac{w-a/c}{i}\right)^{(k-2)/2}dw.$$

Observe that the integrand above is holomorphic and thus it follows by the vanishing of f at the cusps that we can shift the contour and arrive at

$$\begin{aligned}&(-2\pi i)^{-k/2}\Gamma(k/2)L(f,a/c,k/2)\\ &=\int_{\gamma\infty}^{\gamma z} f(w)(w-a/c)^{(k-2)/2}dw+\int_{\gamma z}^{i\infty} f(w)(w-a/c)^{(k-2)/2}dw\end{aligned}\tag{5.5}$$

for any $z\in\mathbb{H}$.

This expression will allow us to prove the following formula for the central value (here it is crucial that k is even).

Lemma 5.3. *Let $z\in\mathbb{H}$ and $\gamma=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\Gamma$. Then we have*

$$\begin{aligned}L(f,\gamma\infty,k/2)=&\Bigg((-1)^{k/2}\sum_{0\le j\le (k-2)/2}\frac{(\frac{k-2}{2})!}{j!}c^{-j}j(\gamma,z)^{-j}I_{k/2-j}(\gamma z)\\ &+\sum_{0\le j\le (k-2)/2}(-1)^j\frac{(\frac{k-2}{2})!}{j!}c^{-j}j(\gamma,z)^{j}I_{k/2-j}(z)\Bigg)\frac{(-2\pi i)^{k/2}}{\Gamma(k/2)},\end{aligned}\tag{5.6}$$

where I_n is the n-fold anti-derivative of f defined in (4.1).

Proof. We treat the two integrals in (5.5) separately.
Using the fact that

$$a/c=\gamma\infty=\gamma z+\frac{c^{-1}}{j(\gamma,z)},$$

we get

$$\begin{aligned}
&\int_{\gamma z}^{i\infty} f(w)(w-a/c)^{(k-2)/2}dw \\
&= \int_{\gamma z}^{i\infty} f(w)(w-\gamma z+\gamma z-a/c)^{(k-2)/2}dw \\
&= \int_{\gamma z}^{i\infty} f(w)\left(w-\gamma z-\frac{c^{-1}}{j(\gamma,z)}\right)^{(k-2)/2} dw \\
&= (-1)^{k/2} \sum_{0\le j\le (k-2)/2} \frac{((k-2)/2)!}{j!} c^{-j} j(\gamma,z)^{-j} I_{k/2-j}(\gamma z)
\end{aligned}$$

using the integral representation (4.1) of $I_j(z)$. To treat the other integral we use the identity

$$w-a/c = \gamma\gamma^{-1}w - a/c = -\frac{c^{-1}}{j(\gamma,\gamma^{-1}w)},$$

which yields

$$\begin{aligned}
\int_{\gamma\infty}^{\gamma z} f(w)(w-a/c)^{(k-2)/2}dw &= \int_{\gamma\infty}^{\gamma z} f(\gamma\gamma^{-1}w)\left(-\frac{c^{-1}}{j(\gamma,\gamma^{-1}w)}\right)^{(k-2)/2} dw \\
&= \int_{i\infty}^{z} f(\gamma w)\left(-\frac{c^{-1}}{j(\gamma,w)}\right)^{(k-2)/2} j(\gamma,w)^{-2}dw
\end{aligned}$$

after the change of variable $w\mapsto \gamma^{-1}w$. Now by using modularity of f and the following identity

$$\frac{j(\gamma,w)}{c} = w-z+\frac{j(\gamma,z)}{c},$$

the above equals

$$\begin{aligned}
&(-1)^{(k-2)/2}\int_{i\infty}^{z} f(w)j(\gamma,w)^{k-2}\left(-\frac{c^{-1}}{j(\gamma,w)}\right)^{(k-2)/2} dw \\
=&(-1)^{(k-2)/2}\int_{i\infty}^{z} f(w)\left(w-z+\frac{j(\gamma,z)}{c}\right)^{(k-2)/2} dw \\
=& \sum_{0\le j\le (k-2)/2} (-1)^j \frac{((k-2)/2)!}{j!} c^{-j} j(\gamma,z)^{j} I_{k/2-j}(z).
\end{aligned}$$

□

Remark 5.4. The above formula is very closely related to the fact that the additive twists $L(f,\cdot,k/2)$ define a *quantum modular form* in the sense of Zagier [34]. The quantum modularity of additive twists for level 1 cusp forms combined with dynamical methods, enabled Bettin and Drappeau in [2] to give a different proof of the normal distribution of additive twists (in the special case of level 1). In [26] the author, inspired by this connection, proved quantum modularity for central values of additive twists of cusp forms of general level and used this to prove a certain 'reciprocity law' for multiplicative twists $L(f,\chi,1/2)$.

5.2. **Analytic properties of Goldfeld Eisenstein series.** What we would like to do now is the following; take the formula in Lemma 5.3, sum over $\gamma \in \Gamma_\infty\backslash\Gamma$, use the identity

$$c = \frac{j(\gamma, z) - \overline{j(\gamma, z)}}{2iy} \tag{5.7}$$

and then finally use the binomial formula to express $E^{m,n}(z, s)$ as a sum of the Poincaré series $G_{A,B,l}(z, s)$. The only slight complication is that when $k \geq 4$ we have negative powers of c in our formula for the central values. In order to bypass this we multiply by a power c^N on both sides of (5.6) for some even $N \geq (k-2)/2$. With this in mind, we define the following *N-shifted Goldfeld Eisenstein series*;

$$E^{m,n}(z, s; N) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} c^N L(f, \gamma\infty, k/2)^m \overline{L(f, \gamma\infty, k/2)}^n \operatorname{Im}(\gamma z)^s. \tag{5.8}$$

By (3.12) we know that the above series converges absolutely (and locally uniformly) for $\operatorname{Re} s \gg 1$. The series $E^{m,n}(z, s; N)$ also has a very nice Fourier expansion at ∞ with constant term related to $D^{m,n}(f, s)$ as we will see below.

We will now derive the analytic properties of $E^{m,n}(z, s; N)$ from the results of the preceding sections. This is a major step towards our main result.

Proposition 5.5. *Let $N \geq (n+m)(k-2)/2$ be an even integer. Then the Eisenstein series $E^{m,n}(z, s; N)$ admits meromorphic continuation to $\operatorname{Re} s > N/2 + 1/2$ satisfying the following;*

(i) *$E^{m,n}(z, s; N)$ is regular for $\operatorname{Re} s > N/2 + 1$ and all poles in the strip*

$$N/2 + 1/2 < \operatorname{Re} s \leq N/2 + 1$$

are contained in the set $\{p + N/2 \mid p \in \mathcal{P}\}$.

(ii) *The pole order of $E^{m,n}(z, s; N)$ at $s = N/2 + 1$ is bounded by $\min(m, n) + 1$.*

(iii) *$E^{n,n}(z, s; N)$ has a pole at $s = N/2 + 1$ of order $n + 1$ with leading Laurent coefficient*

$$(4\pi)^{nk} y^{-N/2} \frac{\binom{N}{N/2}}{2^N} \frac{(n!)^2 ||f||^{2n}}{((k-1)!)^n \operatorname{vol}(\Gamma)^{n+1}}.$$

Proof. By (5.6) we can write

$$c^N L(f, \gamma\infty, k/2)^n \overline{L(f, \gamma\infty, k/2)}^m$$

as a linear combination of terms of the type

$$h(z) j(\gamma, z)^t \overline{j(\gamma, z)}^{t'} c^{N'} I_{k/2-j_1}(\gamma z) \cdots I_{k/2-j_{m'}}(\gamma z) \overline{I_{k/2-j_{m'+1}}(\gamma z)} \cdots \overline{I_{k/2-j_{n'+m'}}(\gamma z)}$$

where $h : \mathbb{H} \to \mathbb{C}$ is a smooth function (this will be a product of $I_j(z)$ for $1 \leq j \leq k/2$), t, t' are integers, $m' \leq m$, $n' \leq n$ are non-negative integers and finally N' is a *non-negative* integer. By inspecting (5.6), we see that t, t' and N' satisfy

$$t + t' + N' = N - 2 \sum_{v=1}^{m'+n'} j_v.$$

Now we use (5.7) and expand using the binomial formula (here it is essential that $N' \geq 0$) to get terms of the type

$$h(z)j(\gamma,z)^t\overline{j(\gamma,z)}^{t'} I_{k/2-j_1}(\gamma z)\cdots I_{k/2-j_{m'}}(\gamma z)\overline{I_{k/2-j_{m'+1}}(\gamma z)}\cdots\overline{I_{k/2-j_{m'+n'}}(\gamma z)}$$

where now

$$t+t' = N - 2\sum_{v=1}^{m'+n'} j_v. \tag{5.9}$$

Now we multiply by $\operatorname{Im}(\gamma z)^s$ and use the identity

$$j(\gamma,z)^t\overline{j(\gamma,z)}^{t'}\operatorname{Im}(\gamma z)^s = y^{(t+t')/2}j_\gamma(z)^{t-t'}\operatorname{Im}(\gamma z)^{s-(t+t')/2}.$$

Thus summing over $\gamma\in\Gamma_\infty\backslash\Gamma$, we can express $E^{m,n}(z,s;N)$ (for $\operatorname{Re}s$ large enough) as a linear combination of terms of the type

$$h(z)G_{A,B,l}(z,s-N/2) \tag{5.10}$$

where the h's are smooth functions (more precisely; products of powers of y's and $I_j(z)$'s), $|A|\leq m, |B|\leq n$ and l is even (which follows from (5.9)).
Notice that (5.9) fits beautifully with the the factor $\alpha(A,B)$ in the definition of $G_{A,B,l}(z,s)$, which is why we get the argument $s-N/2$ for all terms.

Now it follows directly from Proposition 4.6 that $E^{m,n}(z,s;N)$ has meromorphic continuation to $\operatorname{Re}s > N/2+1/2$ satisfying property (i) of Proposition 5.5. Furthermore by Proposition 4.9 it follows that the Poincaré series $G_{A,B,l}(z,s-N/2)$ has a pole of order at most $\min(m,n)+1$ at $s=N/2+1$. Thus the same is true for $E^{m,n}(z,s;N)$.

Now finally let us consider the diagonal case $m=n$. We see by Corollary 4.11 and Proposition 4.9, that all terms (5.10) have a pole of order at most n, except the one with

$$A=B=\{\underbrace{k/2,\ldots,k/2}_{n}\}$$

and $l=0$. Now let us calculate the coefficient of $G_{A,A,0}(z,s-N/2)$ in the expansion of $E^{n,n}(z,s;N)$; we have

$$(2\pi)^{nk}c^N|I_{k/2}(\gamma z)|^{2n} = \frac{(2\pi)^{nk}}{(2y)^N}|I_{k/2}(\gamma z)|^{2n}\sum_{v=1}^{N}(-1)^v\binom{N}{v}j(\gamma,z)^v\overline{j(\gamma,z)}^{N-v}.$$

Now we multiply by $\operatorname{Im}(\gamma z)^s$ and sum over $\gamma\in\Gamma_\infty\backslash\Gamma$. By the pole bound from Corollary 4.11, we see that only the term with $v=N/2$ above can contribute with a pole of order $n+1$ at $s=N/2+1$. Thus we can write

$$\begin{aligned} E^{n,n}(z,s;N) =&(4\pi)^{nk}y^{-N/2}\frac{\binom{N}{N/2}}{2^N}G_{A,A,0}(z,s-N/2)\\ &+ (\textit{terms with a pole of order at most } n \textit{ at } s=N/2+1), \end{aligned}$$

where

$$A=\{\underbrace{k/2,\ldots,k/2}_{n}\}.$$

(The extra factor of 2^{nk} comes from $2i$ and $-2i$ in the denominator in the definition of $G_{A,A,0}(z,s)$). Now the result follows directly from Theorem 4.10. □

5.3. **Analytic properties of $D^{m,n}(f,s)$.** Using the above we can now extract analytic information about $D^{m,n}(f,s)$ using that it is essentially the constant term in the Fourier expansion of $E^{m,n}(z,s;N)$ at ∞.

Lemma 5.6. *Let $N \geq 0$ be an even integer. Then the constant term in the Fourier expansion of $E^{m,n}(z,s;N)$ (at ∞) is equal to*

$$\frac{\pi^{1/2}y^{1-s}\Gamma(s-1/2)}{\Gamma(s)}D^{m,n}(f,s-N/2).$$

Proof. By the double coset decomposition (see [16, Theorem 2.7]), we have

$$\Gamma_\infty\backslash\Gamma/\Gamma_\infty \leftrightarrow \left\{(c,d) \mid 0 \leq d < c, \begin{pmatrix} * & * \\ c & d \end{pmatrix} \in \Gamma\right\} \cup \{(0,1)\}$$

Now since $L(f,\gamma\infty,k/2)$ is well-defined in the above double coset and $L(f,\infty,k/2) = 0$ per definition, we can write

$$\begin{aligned} &E^{m,n}(z,s;N) \\ &= \sum_{c>0}\sum_{0\leq d<c} c^N L(f,\gamma_{c,d}\infty,k/2)^m \overline{L(f,\gamma_{c,d}\infty,k/2)}^n \sum_{l\in\mathbb{Z}} \frac{y^s}{|c(z+l)+d|^{2s}}, \end{aligned}$$

where $\gamma_{c,d}$ is any representative of (c,d) in $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$. Now the result follows by computing the inner sum using Poisson summation as in [16, Section 3.4]. □

With this lemma at our disposal, we can easily derive the analytic properties of $D^{m,n}(f,s)$ from the results already established.

Theorem 5.7. *Let $m,n \geq 0$ be non-negative integers. Then the Dirichlet series $D^{m,n}(f,s)$ admits meromorphic continuation to $\operatorname{Re} s > 1/2$ satisfying the following.*

(i) *$D^{m,n}(f,s)$ is regular for $\operatorname{Re} s > 1$ and the poles in the strip*

$$1/2 < \operatorname{Re} s \leq 1$$

are contained in the singular set $\mathcal{P}$.

(ii) *The pole order of $D^{m,n}(f,s)$ at $s=1$ is bounded by $\min(m,n)+1$.*

(iii) *$D^{n,n}(f,s)$ has a pole of order $n+1$ at $s=1$ with leading Laurent coefficient*

$$\frac{(n!)^2(4\pi)^{nk}||f||^{2n}}{\pi((k-1)!)^n\operatorname{vol}(\Gamma)^{n+1}}.$$

(iv) *For $s = \sigma + it$ with $1/2+\varepsilon \leq \sigma \leq 2$ and $\operatorname{dist}(\lambda(s),\mathcal{P}) \geq \varepsilon$, we have the following bound*

$$D^{m,n}(f,s) \ll_\varepsilon (1+|t|)^{1/2},$$

where the implied constant may depend on m,n.

Proof. Fix some even integer $N \geq (n+m)(k-2)/2$. By Lemma 5.6 we have

$$D^{m,n}(f, s-N/2) = \frac{\Gamma(s)}{\pi^{1/2} y^{1-s} \Gamma(s-1/2)} \int_0^1 E^{m,n}(z, s; N) dx.$$

Thus we conclude directly from Proposition 5.5 the following properties; meromorphic continuation of $D^{m,n}(f, s)$, the claim about the position of the possible poles and the bound on the order of the pole at $s = 1$.
Now to treat the case $m = n$, we recall that the leading Laurent coefficient of $E^{n,n}(z, s; N)$ at $s = N/2 + 1$ is constant. Thus we see directly from Proposition 5.5 that $D^{n,n}(f, s)$ has a pole of order $n+1$ at $s = 1$ with leading Laurent coefficient

$$\frac{\Gamma(N/2+1)}{\pi^{1/2} y^{1-(N/2+1)} \Gamma(N/2+1/2)} (4\pi)^{nk} y^{-N/2} \frac{\binom{N}{N/2}}{2^N} \frac{n!^2 ||f||^{2n}}{((k-1)!)^n \mathrm{vol}(\Gamma)^{n+1}}.$$

Using that for even N, we have

$$\Gamma(N/2+1/2) = \frac{\pi^{1/2}(N-1)\cdots 3\cdot 1}{2^{N/2}}, \quad \Gamma(N/2+1) = (N/2)!,$$
$$\frac{\binom{N}{N/2}}{2^N} = \frac{N\cdot(N-1)\cdots 1}{2^{N/2}(N/2)!\cdot 2N\cdot 2(N-1)\cdots 2} = \frac{(N-1)\cdots 3\cdot 1}{2^{N/2}(N/2)!},$$

the claim about the leading Laurent coefficient follows.

To get the claim about growth on vertical lines, we need to somehow bring the bounds on the L^2-norms of $G_{A,B,l}(z, s)$ from Proposition 4.13 into play.
First step is to integrate $D^{m,n}(f, s-N/2)$ with respect to y over some finite segment, say $[1, 2]$, which gives

$$D^{m,n}(f, s-N/2) = \frac{\Gamma(s)}{\pi^{1/2}\Gamma(s-1/2)} \int_1^2 \int_0^1 y^{s-1} E^{m,n}(z, s; N) dx dy.$$

By the proof of Proposition 5.5 we can write the above as a linear combination of terms of the type

$$\frac{\Gamma(s)}{\Gamma(s-1/2)} \int_1^2 \int_0^1 h(z) G_{A,B,l}(z, s-N/2) dx dy,$$

with $h(z)$ some smooth function. Since $h(z)$ is bounded in the region $[0,1]\times[1,2]$, the Cauchy-Schwarz inequality implies the following;

$$\begin{aligned}
&\int_1^2 \int_0^1 y^{1-s} h(z) G_{A,B,l}(z, s-N/2) dx dy \\
&\ll \left(\left(\int_1^2 \int_0^1 y^{4-2\sigma} |h(z)|^2 \, dx dy \right) \cdot \left(\int_1^2 \int_0^1 |G_{A,B,l}(z, s-N/2)|^2 \frac{dx dy}{y^2} \right) \right)^{1/2} \\
&\ll_h ||G_{A,B,l}(z, s-N/2)||.
\end{aligned}$$

Thus for $s = \sigma + it$ with

$$1/2 + N/2 + \varepsilon \leq \sigma < 3/2 + N/2$$

and $s - N/2$ bounded at least ε away from $\mathcal{P}$, we get by Proposition 4.13 that

$$D^{m,n}(f, s - N/2) \ll_\varepsilon \frac{|\Gamma(s)|}{|\Gamma(s - 1/2)|},$$

where the implied constant may depend on m, n and f.
Now Stirling's formula implies for s in the given range;

$$\frac{\Gamma(s)}{\Gamma(s-1/2)} \ll_N (1 + |t|)^{1/2},$$

and thus

$$D^{m,n}(f, s) \ll_\varepsilon (1 + |t|)^{1/2}$$

for $s = \sigma + it$ with $1/2 + \varepsilon \leq \sigma < 1$ and s being ε-bounded away from $\mathcal{P}$. □

This result implies a Lindelöf type bound in the c-aspect for additive twists.

Corollary 5.8. *Let $r \in T = T_{\infty\infty}$ and $\varepsilon > 0$. Then we have*

$$L(f, r, k/2) \ll_\varepsilon c(r)^\varepsilon$$

where the implied constant may depend on f.

Proof. By combining Landau's Lemma [17, Lemma 5.56] and Theorem 5.7 above, we see that

$$D^{n,n}(f, s) = \sum_{r \in T_\Gamma} \frac{|L(f, r, k/2)|^{2n}}{c(r)^{2s}}$$

converges absolutely for $\operatorname{Re} s > 1$ for all $n \geq 0$. In particular this implies for all $\varepsilon > 0$ that

$$\frac{|L(f, r, k/2)|^{2n}}{c(r)^{2+\varepsilon}}$$

is bounded as $c(r) \to \infty$. This yields;

$$L(f, r, k/2) \ll_{\varepsilon,n} c(r)^{1/n+\varepsilon},$$

as wanted. □

5.4. **Normal distribution.** In this section we will show that the central values

$$L(f, r, k/2), \quad r \in T = T_{\infty\infty}$$

with a suitable normalization are normally distributed when ordered by the size of $c(r)$. This is done by determining all asymptotic moments and then appealing to a classical result of Fréchet and Shohat [31, Theorem B on p. 17] known as the *methods of moments.*
To evaluate the asymptotic moments, we firstly apply Theorem A.2 of the appendix to the Dirichlet series $D^{m,n}(f, s)$. This allows us to prove the following theorem.

Theorem 5.9. *Let $f \in S_k(\Gamma)$ be a cusp form of even weight k and let m, n be non-negative integers. Then we have*

$$\sum_{r \in T(X)} L(f, r, k/2)^m \overline{L(f, r, k/2)}^n \ll X^2 \log(X)^{\min(m,n)}. \tag{5.11}$$

Let n be a non-negative integer. Then we have

$$\sum_{r\in T(X)} |L(f,r,k/2)|^{2n} = P_n(\log X)X^2 + O_\varepsilon(X^{\max(4/3,2s_1)+\varepsilon}), \tag{5.12}$$

where $s_1 \in \mathcal{P}$ corresponds to the smallest positive eigenvalue of Δ as in (3.7) (here $s_1 = 1/2$ if $\mathcal{P} = \{1\}$) and P_n is a polynomial of degree n with leading coefficient

$$\frac{2^n n!}{\pi \operatorname{vol}(\Gamma)}(C_f)^n,$$

with C_f as in (5.3).

Proof. The bound (5.11) follows directly from Theorem 5.7 using a standard contour integration argument (even though the coefficients of $D^{m,n}(f,s)$ are not positive numbers, we do not need to be careful, since we do not care about error-term).
For $m = n$ we apply Theorem A.2 to $D^{m,n}(f,s)$ with

$$S_{\text{poles}} = 2\mathcal{P} = \{2s_0 = 2, 2s_1, \dots, 2s_m\}, \quad a = 1 \quad \text{and} \quad A = 1/2.$$

We know from Theorem 5.7 that $s \mapsto D^{n,n}(f, s/2)$ has a pole at $s = 2$ of order $n+1$ with leading Laurent coefficient

$$b_{n+1} = \frac{2^{n+1}(n!)^2(4\pi)^{nk}||f||^{2n}}{\pi((k-1)!)^n \operatorname{vol}(\Gamma)^{n+1}},$$

where the extra factor of 2^{n+1} comes from $(s/2-1)^{-n-1} = 2^{n+1}(s-2)^{-n-1}$.
Thus it follows directly from Theorem A.2 that we get the wanted asymptotic formula (1.5) with error-term

$$O_\varepsilon(X^{\max((1+2\cdot 1/2)/(1+1/2),2\operatorname{Re} s_1)+\varepsilon}) = O_\varepsilon(X^{\max(4/3,2s_1)+\varepsilon}).$$

□

Remark 5.10. If we instead considered smooth moments, we would get the slightly better error-term $O_\varepsilon(X^{2s_1+\varepsilon})$ using Theorem A.1 in the appendix.

Remark 5.11. One can check that in the weight 2 case, the main term agrees with Petridis and Risager [29, Corollary 7.7].

From the above we can deduce the asymptotic moments of $L(f, a/c, k/2)$ by partial summation.

Corollary 5.12. *Let $m \neq n$ be non-negative integers. Then we have*

$$\frac{\sum_{r\in T(X)}\left(\frac{L(f,r,k/2)}{(C_f \log c(r))^{1/2}}\right)^m \left(\frac{\overline{L(f,r,k/2)}}{(C_f \log c(r))^{1/2}}\right)^n}{\#T(X)} \to 0, \tag{5.13}$$

as $X \to \infty$.
Let $n \geq 0$ be a non-negative integer. Then we have

$$\frac{\sum_{r\in T(X)}\left|\frac{L(f,r,k/2)}{(C_f \log c(r))^{1/2}}\right|^{2n}}{\#T(X)} \to 2^n n!, \tag{5.14}$$

as $X \to \infty$.

Proof. The corollary follows immediately from partial summation using Theorem 1.6 and the asymptotic formula [29, Lemma 3.5];

$$\#T(X) \sim \frac{X^2}{\pi \operatorname{vol}(\Gamma)}.$$

□

Recall that the coordinates $\begin{pmatrix} Y \\ Z \end{pmatrix}$ of a standard complex normal distribution (or equivalently a standard 2-dimensional normal distribution with diagonal variance-matrix) has moments given by

$$E(Y^m Z^n) = \begin{cases} (m-1)!!(n-1)!! & \text{if } m \text{ and } n \text{ are even} \\ 0 & \text{otherwise} \end{cases},$$

where $(n-1)!! = (n-1)\cdot(n-3)\cdots 1$. By taking linear combinations of the moments in Corollary 5.12, it follows that the asymptotic moments of

$$\begin{pmatrix} \operatorname{Re} \frac{L(f,r,k/2)}{(C_f \log c(r))^{1/2}} \\ \operatorname{Im} \frac{L(f,r,k/2)}{(C_f \log c(r))^{1/2}} \end{pmatrix}, \quad r \in T(X)$$

as $X \to \infty$ are the same as those of the 2-dimensional standard normal. This fact and the above corollary allow us to prove Theorem 1.1.

Proof of Theorem 1.1. We would like to use the result of Fréchet and Shohat coming from probability theory [31, p. 17] mentioned before. To make it fit into the probability theoretical framework of the Fréchet–Shohat Theorem, we consider for each $X > 0$ the 2-dimensional random variable

$$\begin{pmatrix} Y_X(r) \\ Z_X(r) \end{pmatrix} = \begin{pmatrix} \operatorname{Re} \frac{L(f,r,k/2)}{(C_f \log c(r))^{1/2}} \\ \operatorname{Im} \frac{L(f,r,k/2)}{(C_f \log c(r))^{1/2}} \end{pmatrix}, \quad r \in T(X)$$

where the outcome space $T(X)$ is endowed with the discrete σ-algebra and the uniform probability measure. Note that the Fréchet–Shohat Theorem, as stated in [31, p. 17], is only directly applicable for 1-dimensional distribution functions, but we can get around this by using the Cramér–Wold Theorem, which says that it is enough to check that the moments of all linear combinations of the coordinates (marginal distributions) converge to the expected. To be precise; it follows from Corollary 5.12 that for $(a,b) \in \mathbb{R}^2 \backslash (0,0)$ the moments of the random variables $aY_X + bZ_X$ converges to the moments of a normal distribution with mean 0 and variance $a^2 + b^2$ as $X \to \infty$. Thus it follows from the Fréchet–Shohat Theorem that the random variables $aY_X + bZ_X$ converges in distribution to the normal distribution with mean 0 and variance $a^2 + b^2$ as $X \to \infty$ (the normal distribution is uniquely determined by its moments).

Now by the Cramér–Wold Theorem (see [31, p. 18]) it follows that $\begin{pmatrix} Y_X \\ Z_X \end{pmatrix}$ converges in distribution to the 2-dimensional standard normal distribution as $X \to \infty$. □

5.5. **Additive twists at a general cusp.** We will now explain how to deal with additive twists associated to general cusps $\mathfrak{b}$ and in particular how to prove Theorem 1.6.

In the case of weight 2, the modular symbol $\langle \gamma, f \rangle$ is well-defined for $\gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_{\mathfrak{b}}$ with $\mathfrak{b}$ any cusp of Γ, which implies that you get a nice Fourier expansion of $E^{m,n}(z,s)$ at every cusp. This is however *not* true for additive twists $L(f, \gamma\infty, k/2)$ of L-functions of cusp forms f of weight $k \geq 4$.

Thus in order to access the cusp $\mathfrak{b}$, we need to consider *the generalized Goldfeld series at* $\mathfrak{b}$;

$$E^{m,n,\mathfrak{b}}(z,s) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} L(f, \gamma\mathfrak{b}, 1/2)^m \overline{L(f, \gamma\mathfrak{b}, 1/2)}^n (\operatorname{Im} \gamma z)^s. \tag{5.15}$$

The constant term of the Fourier expansion of $E^{m,n,\mathfrak{b}}(z,s)$ at $\mathfrak{b}$ is exactly given by

$$\frac{\pi^{1/2} y^{1-s} \Gamma(s-1/2)}{\Gamma(s)} D_{\mathfrak{b}}^{m,n}(f,s)$$

with $D_{\mathfrak{b}}^{m,n}(f,s)$ defined as in (5.4).

Now by a slight modification of Lemma 5.3, we conclude that for $r \in T_{\infty\mathfrak{b}}$, we have

$$\begin{aligned} L(f,r,k/2) = & \Bigg(\sum_{0 \leq j \leq (k-2)/2} \frac{(\frac{k-2}{2})!}{j!} \left(\frac{z - \mathfrak{b}}{j(\gamma,\mathfrak{b}) j(\gamma,z)} \right)^j (-1)^{k/2-j} I_{k/2-j}(\gamma z) \\ & + \int_{\mathfrak{b}}^{z} f(w) \left(\frac{(w-\mathfrak{b})(j(\gamma,z)(w-\overline{z}) - \overline{j(\gamma,z)}(w-z))}{2iy j(\gamma,\mathfrak{b})} \right)^{k/2-1} dw \Bigg) \frac{(-2\pi i)^{k/2}}{\Gamma(k/2)}. \end{aligned} \tag{5.16}$$

Observe that $j(\gamma, \mathfrak{b})$ is well-defined for $\gamma \in \Gamma_\infty \backslash \Gamma / \Gamma_{\mathfrak{b}}$. Thus we can define the *Goldfeld Eisenstein series at* $\mathfrak{b}$;

$$E^{m,n,\mathfrak{b}}(z,s;N) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma,\mathfrak{b})^N L(f, \gamma\mathfrak{b}, 1/2)^m \overline{L(f, \gamma\mathfrak{b}, 1/2)}^n (\operatorname{Im} \gamma z)^s,$$

whose Fourier expansion at $\mathfrak{b}$ has constant term equal to;

$$\frac{\pi^{1/2} y^{1-s} \Gamma(s-1/2)}{\Gamma(s)} D_{\mathfrak{b}}^{m,n}(f, s - N/2).$$

Similarly to the case $\mathfrak{b} = \infty$ we can write

$$j(\gamma, \mathfrak{b}) = j(\gamma, z) \frac{\mathfrak{b} - \overline{z}}{2iy} - \overline{j(\gamma,z)} \frac{\mathfrak{b} - z}{2iy}$$

where we consider $\mathfrak{b}$ as a real number. Combining this trick with (5.16), we see that we can express $E_{\mathfrak{b}}^{m,n}(z,s;N)$ in terms of the Poincaré series $G_{A,B,l}(z,s)$ and we conclude by an argument as in the case $\mathfrak{b} = \infty$ the following.

Theorem 5.13. *Let Γ be a discrete and co-finite subgroup of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at ∞ of width 1, $\mathfrak{b}$ a cusp of Γ and $f \in \mathcal{S}_k(\Gamma)$ a cusp form of even weight k. Then we have*

$$\sum_{r \in T_{\infty\mathfrak{b}}(X)} L(f,r,k/2)^m \overline{L(f,r,k/2)}^n \ll X^2 (\log X)^{\min(m,n)}$$

and

$$\sum_{r\in T_{\infty\mathfrak{b}}(X)} |L(f,r,k/2)|^{2n} = P_n(\log X)X^2 + O_\varepsilon(X^{\max(4/3,2s_1)+\varepsilon}), \tag{5.17}$$

with P_n as in Theorem 5.9 and s_1 as in (3.7).

Using the above we deduce easily Theorem 5.1 in the case of (marginalizing to) a single cusp form $f \in \mathcal{S}_k(\Gamma)$ with twists at a general cusp $\mathfrak{b}$. The argument being exactly as in the case $\mathfrak{b} = \infty$.

Furthermore Theorem 5.13 allows us to prove Theorem 1.6, which we will need in order to obtain an asymptotic formula for the averages of certain families consisting of automorphic L-functions of the form $L(f\otimes\chi, 1/2)$. The proof of this result (Corollary 1.9) will be given in Section 6 below.

Proof of Theorem 1.6. In the case where $\Gamma = \Gamma_0(q)$ and $\mathfrak{b}$ corresponds to the real number 0, we have coming from [16, p. 47] the following scaling matrix;

$$\sigma_0 = \sigma_{\mathfrak{b}} = \begin{pmatrix} 0 & -1/\sqrt{q} \\ \sqrt{q} & 0 \end{pmatrix}.$$

This implies that

$$T_{\infty 0} = \{r = \frac{a}{c} \bmod 1 \mid (a,c)=1, (c,q)=1\}$$

and $c(r) = c\sqrt{q}$.
Thus we conclude from Theorem 5.13 with $\Gamma = \Gamma_0(q)$, $f \in \mathcal{S}_k(\Gamma_0(q))$ and $\mathfrak{b} = 0$;

$$\begin{aligned}
&\sum_{\substack{0<c\leq X,\\ (c,q)=1}} \sum_{a\in(\mathbb{Z}/c\mathbb{Z})^\times} |L(f,a/c,k/2)|^{2n} \\
&= \sum_{r\in T_{\infty 0}(\sqrt{q}X)} |L(f,r,k/2)|^{2n} \\
&= \frac{q(2C_f)^n\, n!}{\pi \operatorname{vol}(\Gamma_0(q))} (\log X)^n X^2 + \sum_{i=0}^{n-1} \beta^d_{f,i} (\log X)^i X^2 + O_\varepsilon(X^{\max(4/3,2s_1)+\varepsilon}).
\end{aligned}$$

By the approximation towards Selberg's conjecture by Kim and Sarnak [16, p. 167], we know that $\operatorname{Re} s_1 \leq 39/64 < 2/3$, which yields exactly Theorem 1.6. □

5.6. **The joint distribution of additive twists of a basis of cusp forms.** Instead of considering a single cusp form f, we can consider an orthogonal basis $f_1,\dots,f_d$ for $\mathcal{S}_k(\Gamma)$ with respect to the Petersson inner product. We will restrict to the case $\mathfrak{b} = \infty$ and then the discussion in the previous section carries directly over to this setting as well.
In this case for any two sequences $\underline{g} = (g_1,\dots,g_m), \underline{h} = (h_1,\dots,h_n)$ with $g_j, h_j \in \{f_1,\dots,f_d\}$, we define the corresponding Goldfeld Eisenstein series

$$E^{\underline{g},\underline{h}}(z,s) := \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \left(\prod_{j=1}^m L(g_j,\gamma\infty,k/2)\right)\left(\prod_{j=1}^n \overline{L(h_j,\gamma\infty,k/2)}\right)(\operatorname{Im}\gamma z)^s.$$

The constant term in the Fourier expansion of $E^{\underline{g},\underline{h}}(z,s)$ (at ∞) is given by

$$\frac{\pi^{1/2}y^{1-s}\Gamma(s-1/2)}{\Gamma(s)}D^{\underline{g},\underline{h}}(s),$$

where

$$D^{\underline{g},\underline{h}}(s):=\sum_{r\in T_{\infty\infty}}\frac{\left(\prod_{j=1}^{m}L(g_j,r,k/2)\right)\left(\prod_{j=1}^{n}\overline{L(h_j,r,k/2)}\right)}{c(r)^{2s}}.$$

We can express the Goldfeld Eisenstein series as a linear combinations of certain Poincaré series, which generalizes $G_{A,B,l}(z,s)$ above. Consider tuples of integers $\underline{u}=(u_1,\dots,u_{m'})$, $\underline{v}=(v_1,\dots,v_{n'})$ with $m'\le m, n'\le n$ and $0\le u_i,v_j\le k/2$ and define

(5.18)

$$G_{\underline{u},\underline{v},l}(z,s):=\sum_{\gamma\in\Gamma_\infty\backslash\Gamma}j_\gamma(z)^{-l}\left(\prod_{j=1}^{m'}\frac{I_{u_j}(\gamma z;g_j)}{(-2i)^{u_j}}\right)\left(\prod_{j=1}^{n'}\frac{\overline{I_{v_j}(\gamma z;h_j)}}{(2i)^{v_j}}\right)\mathrm{Im}(\gamma z)^{s+\alpha(\underline{u},\underline{v})},$$

where

$$\alpha(\underline{u},\underline{v})=\left(\sum_j k/2-u_j\right)+\left(\sum_j k/2-v_j\right).$$

Then the analogue of Proposition 4.9 holds for the above Poincaré series as well, with essentially the same proof. Using this, it can be shown by the methods from the preceding sections that $D^{\underline{g},\underline{h}}(s)$ has a pole of order at most $\min(m,n)+1$ at $s=1$. Furthermore when $m=n$, we have

$$D^{\underline{g},\underline{h}}(s)=\left(\sum_{\sigma,\sigma'\in S_n}\prod_{j=1}^{n}\langle g_{\sigma(j)},h_{\sigma'(j)}\rangle\right)\frac{(4\pi)^{nk}}{(s-1)^{n+1}\pi((k-1)!)^n\mathrm{vol}(\Gamma)^{n+1}}$$
$$+\,(\textit{pole order at most } n \textit{ at } s=1),$$

where S_n denotes the group of permutation on n letters. Observe that this generalizes our previous results since $|S_n|=n!$. In particular $D^{\underline{g},\underline{h}}(s)$ has a pole of order $n+1$ exactly if $\underline{g}$ and $\underline{h}$ are permutations of each other.

Now consider the random variable

$$\begin{pmatrix}Y_{1,X}\\ Z_{1,X}\\ \vdots\\ Y_{d,X}\\ Z_{d,X}\end{pmatrix}$$

on the outcome space $T(X)$ endowed with uniform probability measure, defined by

$$Y_{j,X}(r)=\mathrm{Re}\,L(f_j,r,k/2)/\sqrt{C_{f_j}\log c(r)},$$

$$Z_{j,X}(r)=\mathrm{Im}\,L(f_j,r,k/2)/\sqrt{C_{f_j}\log c(r)}$$

for $r\in T(X)$ and $j=1,\dots,d$. Then by the above we can evaluate all asymptotic moments and show using a combination of the results of Fréchet–Shohat and Cramér that as $X \to \infty$, this random variable converges in distribution to d independent standard complex normal distributions.

By combining the methods described in this and the preceding section, one concludes the proof of Theorem 5.1.

6. Applications to $L(f \otimes \chi, 1/2)$

We now apply our results to the averages of certain families constructed from the multiplicative twists $L(f \otimes \chi, 1/2)$ and thus giving a proof of Corollary 1.9. The connection between multiplicative and additive twists is for primitive characters given by the Birch–Stevens formula [23, Theorem 2.3], but some cleverness has to be applied in order to deal with non-primitive characters.

Our results apply to a newform $f \in \mathcal{S}_k(\Gamma_0(q))$ of even weight k and level q with Fourier expansion (at ∞) given by

$$f(z) = \sum_{n\geq 1} \lambda_f(n) n^{(k-1)/2} q^n,$$

where $\lambda_f(n)$ denotes the nth Hecke eigenvalue of f. In what follows it is essential that f is an eigenform for all Hecke operators.

Associated to such a newform $f \in \mathcal{S}_k(\Gamma_0(q))$ and a Dirichlet character χ mod c, we define the (naively) twisted L-function;

$$L(f, \chi, s) := \sum_{n\geq 1} \frac{\lambda_f(n)\chi(n)}{n^{\frac{k-1}{2}+s}},$$

which admits analytic continuation and a functional equation. Note that this is *not* necessarily equal to (the finite part) of the L-function of the automorphic representation $\pi_f \otimes \chi$ (where π_f is the automorphic representation corresponding to f). However in the special case when $(q, c) = 1$, then this is actually true and we will write $L(f \otimes \chi, s) = L(f, \chi, s)$.

6.1. **Averages of multiplicative twists.**

The first step is to establish a connection between additive twists and multiplicative ones. The formula below is a generalization of the Birch–Stevens formula [23, Theorem 2.3] to non-primitive characters.

Proposition 6.1 (Birch–Stevens formula for non-primitive characters)**.**
Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a newform of weight k and level q and χ a Dirichlet character mod c. *Then we have*

$$\nu(f, \chi^*, c/c(\chi)) L(f, \chi^*, 1/2) = \sum_{a \in (\mathbb{Z}/c\mathbb{Z})^\times} \overline{\chi}(a) L(f, a/c, k/2), \tag{6.1}$$

and

$$L(f, a/c, k/2) = \frac{1}{\varphi(c)} \sum_{\chi \bmod c} \chi(a) \nu(f, \chi^*, c/c(\chi)) L(f, \chi^*, 1/2), \tag{6.2}$$

where χ^* mod $c(\chi)$ *denotes the unique primitive character that induces* χ *and*

$$\nu(f, \chi, n) := \tau(\overline{\chi}) \sum_{\substack{n_1 n_2 n_3 = n,\\ (n_1, q)=1}} \chi(n_1)\mu(n_1)\overline{\chi}(n_2)\mu(n_2)\lambda_f(n_3) n_3^{1/2}.$$

Proof. For $\mathrm{Re}\, s > 1$, we have because of absolute convergence;

$$\sum_{a\in(\mathbb{Z}/c\mathbb{Z})^\times} \overline{\chi}(a)L(f,a/c,s+(k-1)/2) = \sum_{n\geq 1}\frac{\lambda_f(n)}{n^s}\left(\sum_{a\in(\mathbb{Z}/c\mathbb{Z})^\times}\overline{\chi}(a)e(na/c)\right). \tag{6.3}$$

The inner sum is a Gauss sum and by [32, Lemma 3], we get

$$\sum_{a\in(\mathbb{Z}/c\mathbb{Z})^\times}\overline{\chi}(a)e(na/c) = \tau(\overline{\chi^*})\sum_{d|(n,c/c(\chi))} d\,\overline{\chi^*}\left(\frac{c}{c(\chi)d}\right)\mu\left(\frac{c}{c(\chi)d}\right)\chi^*\left(\frac{n}{d}\right),$$

where χ^* mod $c(\chi)$ denotes the unique primitive character that induces χ. Plugging this into (6.3), interchanging the sums and putting $n = dl$, we arrive at

$$\begin{aligned}&\sum_{a\in(\mathbb{Z}/c\mathbb{Z})^\times}\overline{\chi}(a)L(f,a/c,s+(k-1)/2)\\ &= \tau(\overline{\chi}^*)\sum_{d|c/c(\chi)}\overline{\chi^*}\left(\frac{c}{c(\chi)d}\right)\mu\left(\frac{c}{c(\chi)d}\right)d\sum_{l>0}\frac{\lambda_f(dl)}{(dl)^s}\chi^*(l).\end{aligned}$$

Now we use that f is a newform, which implies that

$$\lambda_f(ld) = \sum_{\substack{h|(l,d),\\ (h,q)=1}}\mu(h)\lambda_f\left(\frac{l}{h}\right)\lambda_f\left(\frac{d}{h}\right).$$

With $m = l/h$ and $\delta = d/h$, we get

$$\begin{aligned}&\tau(\overline{\chi^*})\sum_{\substack{\delta h|c/c(\chi)\\ (h,q)=1}}\overline{\chi^*}\left(\frac{c}{c(\chi)\delta h}\right)\mu\left(\frac{c}{c(\chi)\delta h}\right)\frac{\lambda_f(\delta)}{\delta^{s-1}}h^{1-2s}\chi^*(h)\mu(h)\sum_{m>0}\frac{\chi^*(m)\lambda_f(m)}{m^s}\\ =&\tau(\overline{\chi^*})L(f\otimes\chi^*,s)\sum_{\substack{\delta h|c/c(\chi),\\ (h,q)=1}}\overline{\chi^*}\left(\frac{c}{c(\chi)\delta h}\right)\mu\left(\frac{c}{c(\chi)\delta h}\right)\frac{\lambda_f(\delta)}{\delta^{s-1}}h^{1-2s}\chi^*(h)\mu(h).\end{aligned}$$

Since the sum above is finite, we can extend the equality to $s = 1/2$ by analytic continuation. The second equality of this lemma follows from the first by orthogonality of characters. □

Remark 6.2. A similar formula has been considered previously by Merel in [24, Théorème 1], where the formula is applied in a more algebraic context.

Using this formula, Corollary 1.9 is an immediate consequence of Theorem 1.6.

Proof of Corollary 1.9. Since $(q,c) = 1$, it follows that $(q,c(\chi)) = 1$ for all Dirichlet characters χ appearing on the left-hand side of (1.6), where $c(\chi)$ denotes the conductor of χ. This implies by the above discussion that we have $L(f,\chi^*,1/2) = L(f\otimes\chi^*,1/2)$. The corollary now follows from Theorem 1.6 by expressing the additive twists in terms of $L(f\otimes\chi^*,1/2)$ using Lemma 6.1, interchanging the sums, using orthogonality of Dirichlet characters and the fact that

$$\overline{\nu(f,\chi^*,c/c(\chi))L(f\otimes\chi^*,1/2)} = \chi(-1)\nu(f,\overline{\chi}^*,c/c(\overline{\chi}))L(f\otimes\overline{\chi}^*,1/2).$$

□

In the special case $n = 1$, we derive the following result, which is an average version of the second moment calculation in [3, Theorem 1.17] with improved error-term.

Corollary 6.3. *Let $f \in \mathcal{S}_k(\Gamma_0(q))$ be a newform of even weight k and level q. Then we have*

$$\sum_{c \leq X, (q,c)=1} \frac{1}{\varphi(c)} \sum_{\chi \bmod c} |\nu(f, \chi^*, c/c(\chi))|^2 |L(f \otimes \chi^*, 1/2)|^2$$

$$= \frac{q(4\pi)^k ||f||^2}{\pi(k-1)! \operatorname{vol}(\Gamma_0(q))^2} (\log X) X^2 + \beta_{f,1} X^2 + O_\varepsilon(X^{4/3+\varepsilon}) \tag{6.4}$$

with χ^* mod $c(\chi)$ and ν as above and $\beta_{f,1}$ a constant.

Appendix A. Contour integration with explicit error-terms

This appendix follows closely an unpublished note of M. Risager. We are grateful to Risager for allowing us to include it here.

A.1. **Smooth cut-offs.** For a compactly supported smooth function ψ on $[0, \infty)$, we define the Mellin transform as

$$\hat{\psi}(s) = \int_0^\infty \psi(y) y^{s-1} dy$$

which converges absolutely for $\operatorname{Re} s > 0$. We have the following inversion formula

$$\psi(y) = \frac{1}{2\pi i} \int_{(c)} \hat{\psi}(s) y^{-s} ds$$

valid for $c > 0$. By the compact support of ψ and repeated partial integration, we get the bound

$$\hat{\psi}(s) \ll_N |s|^{-N} \tag{A.1}$$

for any $N \geq 1$.

Now let $(c_n)_{n \geq 1}$ be a sequence of positive real numbers such that $c_n \to \infty$ as $n \to \infty$ and $(a_n)_{n \geq 1}$ a sequence of complex numbers such that the Dirichlet series

$$D(s) = \sum_{n \geq 1} \frac{a_n}{(c_n)^s}, \tag{A.2}$$

converges absolutely in the half-plane $\operatorname{Re} s > \sigma_0 > 0$. Assume further that $D(s)$ admits meromorphic continuation to $\operatorname{Re} s > a - \varepsilon > 0$ for some $\varepsilon > 0$ and that there is a finite number of poles in the half-plane $\operatorname{Re} s > a - \varepsilon$. Denote these poles by

$$S_{\text{poles}} = \{s_0, \ldots, s_M\} \tag{A.3}$$

with $\operatorname{Re} s_0 = \sigma_0 \geq \ldots \geq \operatorname{Re} s_M$ and let the singular expansion at $s = s_m, 0 \leq m \leq M$ be given by

$$D(s) = \sum_{j=0}^{d_m} \frac{b_{m,j}}{(s - s_m)^j} + r_m(s), \tag{A.4}$$

where $r_m(s)$ is regular at $s = s_m$. Assume further that we have the following bound on the growth

$$D(s) \ll (1+|t|)^A \tag{A.5}$$

valid for $a \leq \operatorname{Re} s \leq \sigma_0 + \varepsilon$. Under these condition we have the following theorem.

Theorem A.1. *Assume that the Dirichlet series $D(s)$ in (A.2) satisfies the condition above. Then for a compactly supported smooth function ψ on $[0,\infty)$, we have for $X > 0$*

$$\sum_{n\geq 1} a_n \psi\left(\frac{c_n}{X}\right) = \sum_{m=0}^{M} P_m(\log X) X^{s_m} + O\left(X^a \int_{-\infty}^{\infty} |\hat{\psi}(a+it)|(1+|t|)^A dt\right), \tag{A.6}$$

where P_m are explicit polynomials of respective degrees $d_m - 1$ given by

$$P_m(x) = \sum_{k=1}^{d_m-1} \frac{1}{k!}\left(\sum_{l=0}^{d_m-1-k} \frac{\hat{\psi}^{(l)}(s_m)}{l!} b_{m,k+l+1}\right) x^k. \tag{A.7}$$

Proof. Follows directly from Mellin inversion by moving the contour to $\operatorname{Re} s = a$. □

In many respects the above smooth cut-off sum may be the more natural result but it is desirable to also obtain asymptotic formulas for the sharp cut-off;

$$\sum_{c_n \leq X} a_n.$$

For this we let ψ be a smooth approximation of the indicator function of $[0,1]$. Below we present an explicit such construction and work out the exact error-terms.

A.2. **Sharp cut-offs.** We will now restrict to the case where $a_n \geq 0$ for all n. First step is to construct a smooth approximation to the indicator function $1_{[0,X]}$. The first step is a smooth approximation of the Dirac measure at $t = 0$. So let $\varphi : \mathbb{R} \to \mathbb{R}_{\geq 0}$ be a smooth function supported in $[-1,1]$ with $\int_{-1}^{1} \varphi(t)dt = 1$.
Then for $\delta < 1/2$ we define

$$\varphi_\delta(t) = \delta^{-1}\varphi(t/\delta),$$

which is supported in $[-\delta,\delta]$ and satisfies $\int_{-1}^{1} \varphi_\delta(t)dt = 1$. This will serve as an approximation of the Dirac measure at $t = 0$.
From this we define the functions $\psi_{\delta,\pm} : \mathbb{R}_+ \to \mathbb{R}$ as the following convolutions

$$\psi_{\delta,\pm}(y) = \int_0^\infty 1_{[0,1\pm\delta]}(yt)\varphi_\delta(t-1)dt.$$

Observe that this defines a smooth function and that the support of $\psi_{\delta,+}$ is contained in $[0,(1+\delta)/(1-\delta)]$ and for $y < 1$ we have $\psi_{\delta,+}(y) = 1$ and similarly $\psi_{\delta,-}(y) = 0$ for $y > 1$ and $\psi_{\delta,-}(y) = 1$ for $y \in [0,(1-\delta)/(1+\delta)]$. The functions $\psi_{\delta,\pm}(y)$ will serve as respectively an upper and lower bound for the indicator function $1_{[0,X]}$. In the end we will use the parameter δ to balance the error-terms.

Now let us bound the error-term

$$\int_{-\infty}^{\infty} |\widehat{\psi_{\delta,\pm}}(a+it)|(1+|t|)^A dt. \tag{A.8}$$

For this we need estimates for $\widehat{\psi_{\delta,\pm}}(a+it)$. By using the definition of $\psi_{\delta,\pm}$ we get

$$\psi_{\delta,\pm}^{(n)}(y) \ll_n (y\delta)^{-n}.$$

For $n \geq 1$ the support of $\psi_{\delta,\pm}^{(n)}(y)$ is contained in

$$[(1-\delta)/(1+\delta),(1+\delta)/(1-\delta)].$$

Thus we get by repeated partial integration

$$\begin{aligned}\widehat{\psi_{\delta,\pm}}(a+it) &\ll \int_0^\infty |\psi_{\delta,\pm}^{(n)}(y)| \frac{y^{a+n}}{|a+it|^n} dy \\ &\ll \int_{(1-\delta)/(1+\delta)}^{(1+\delta)/(1-\delta)} \frac{1}{(\delta y)^n} \frac{y^{a+n}}{|a+it|^n} dy \\ &\ll \frac{\delta}{(\delta|a+it|)^n}\end{aligned}$$

for any $n \geq 1$. By interpolation this is true for all $r \in \mathbb{R}_{\geq 1}$. Now we need to choose a small n in order to make δ^{1-n} small but large enough so that the integral in (A.8) converges, i.e. $r = A+1+\varepsilon$. This yields

$$\int_{-\infty}^{\infty} |\widehat{\psi_{\delta,\pm}}(a+it)|(1+|t|)^A dt \ll_\varepsilon \delta^{-A-\varepsilon}.$$

By a straight forward computation, we see that for $0 \leq m \leq M$

$$\psi_{\delta,\pm}^{(n)}(s_m) = (-1)^n n! s_m^{n-1} + O(\delta^{n+1}).$$

Now since

$$\sum_n a_n \psi_{\delta,-}(c_n/X) \leq \sum_{c_n \leq X} a_n \leq \sum_n a_n \psi_{\delta,+}(c_n/X),$$

we get by Theorem A.1

$$\sum_{c_n \leq X} a_n = \sum_{m=0}^{M} P_m(\log X) X^{s_m} + O_\varepsilon(\delta X^{\sigma_0+\varepsilon} + \delta^{-A-\varepsilon} X^a)$$

where

$$P_m(x) = \sum_{k=1}^{d_m-1} \frac{1}{k!} \left(\sum_{l=0}^{d_m-1-k} (-1)^l s_m^{l-1} b_{m,k+l+1} \right) x^k.$$

To balance the error-term we put

$$\delta = X^{-(\sigma_0-a)/(1+A)}$$

which yields an error-term $\ll_\varepsilon X^{(a+\sigma_0 A)/(1+A)+\varepsilon}$. Absorbing P_m for $m \neq 0$ into the error-term, we arrive at the following theorem.

Theorem A.2. *Let $D(s)$ be the Dirichlet series (A.2) satisfying (A.3), (A.4) and (A.5). If we assume that the coefficients a_n are non-negative, then we have the following;*

$$\sum_{c_n \leq X} a_n = P(\log X) X^{s_0} + O_\varepsilon(X^{\max((a+\sigma_0 A)/(1+A), \mathrm{Re}\, s_1)+\varepsilon}), \tag{A.9}$$

where $P = P_0$ is a polynomial of degree $d_0 - 1$ with leading coefficient

$$\frac{b_{0,d_0}}{s_0(d_0-1)!}.$$

Remark A.3. If the coefficients a_n are not assumed non-negative, then one needs the extra assumption

$$\sum_{X<c_n\leq(1+\delta)X} |a_n| \ll \delta X^{\sigma_0+\varepsilon} + X^{\mathrm{Re}\, s_1+\varepsilon}, \tag{A.10}$$

but then the conclusion of Theorem A.2 still holds. This assumption is needed to control the error term coming from;

$$\sum_{n\leq X} a_n - \sum_{n\geq 1} a_n \psi_{\delta,+}(n/X).$$

References

[1] Sandro Bettin, *High moments of the estermann function*, Algebra and Number Theory **13** (2017).

[2] Sandro Bettin and Sary Drappeau, *Limit laws for rational continued fractions and value distribution of quantum modular forms*, arXiv e-prints (2019), arXiv:1903.00457.

[3] Valentin Blomer, Étienne Fouvry, Emmanuel Kowalski, Philippe Michel, Djordje Milićević, and Will Sawin, *The second moment theory of families of L-functions*, arXiv e-prints (2018), arXiv:1804.01450.

[4] Roelof Bruggeman and Nikolaos Diamantis, *Fourier coefficients of Eisenstein series formed with modular symbols and their spectral decomposition*, J. Number Theory **167** (2016), 317–335. MR 3504050

[5] Gautam Chinta, *Analytic ranks of elliptic curves over cyclotomic fields*, J. Reine Angew. Math. **544** (2002), 13–24. MR 1887886

[6] Gautam Chinta and Cormac O'Sullivan, *Non-holomorphic poincaré series constructed from derivatives and antiderivatives of cusp forms and bounds on period polynomial*, Unpublished (2002).

[7] Yves Colin de Verdière, *Pseudo-laplaciens. II*, Ann. Inst. Fourier (Grenoble) **33** (1983), no. 2, 87–113. MR 699488

[8] Nikolaos Diamantis, Jeffrey Hoffstein, Mehmet Kıral, and Min Lee, *Shifted convolutions and a conjecture by Mazur, Rubin and Stein*, arXiv e-prints (2018), arXiv:1807.02506.

[9] W. Duke, J. B. Friedlander, and H. Iwaniec, *Bounds for automorphic L-functions. III*, Invent. Math. **143** (2001), no. 2, 221–248. MR 1835388

[10] ———, *The subconvexity problem for Artin L-functions*, Invent. Math. **149** (2002), no. 3, 489–577. MR 1923476

[11] Martin Eichler, *Eine Verallgemeinerung der Abelschen Integrale*, Sammelband zu Ehren des 250. Geburtstages Leonhard Eulers, Akademie-Verlag, Berlin, 1959, pp. 112–115. MR 0122993

[12] John D. Fay, *Fourier coefficients of the resolvent for a Fuchsian group*, J. Reine Angew. Math. **293/294** (1977), 143–203. MR 0506038

[13] Stephen S. Gelbart and Stephen D. Miller, *Riemann's zeta function and beyond*, Bull. Amer. Math. Soc. (N.S.) **41** (2004), no. 1, 59–112. MR 2015450

[14] Dorian Goldfeld, *The distribution of modular symbols*, Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997), de Gruyter, Berlin, 1999, pp. 849–865. MR 1689548

[15] ———, *Zeta functions formed with modular symbols*, Automorphic forms, automorphic representations, and arithmetic (Fort Worth, TX, 1996), Proc. Sympos. Pure Math., vol. 66, Amer. Math. Soc., Providence, RI, 1999, pp. 111–121. MR 1703748

[16] Henryk Iwaniec, *Spectral methods of automorphic forms*, second ed., Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002. MR 1942691

[17] Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR 2061214

[18] M. Jutila, *Lectures on a method in the theory of exponential sums*, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, vol. 80, Published for the Tata Institute of Fundamental Research, Bombay; by Springer-Verlag, Berlin, 1987. MR 910497

[19] Henry H. Kim, *Functoriality for the exterior square of* GL_4 *and the symmetric fourth of* GL_2, J. Amer. Math. Soc. **16** (2003), no. 1, 139–183, With appendix 1 by Dinakar Ramakrishnan and appendix 2 by Kim and Peter Sarnak. MR 1937203

[20] E. Kowalski, P. Michel, and J. VanderKam, *Rankin-Selberg L-functions in the level aspect*, Duke Math. J. **114** (2002), no. 1, 123–191. MR 1915038

[21] Jungwon Lee and Hae-Sang Sun, *Dynamics of continued fractions and distribution of modular symbols*, arXiv e-prints (2019), arXiv:1902.06277.

[22] Ju I Manin, *PERIODS OF PARABOLIC FORMS AND $ p$-ADIC HECKE SERIES*, Mathematics of the USSR-Sbornik **21** (1973), no. 3, 371–393.

[23] Barry Mazur and Karl Rubin, *Arithmetic conjectures suggested by the statistical behavior of modular symbols*, arXiv e-prints (2019), arXiv:1910.12798.

[24] Loïc Merel, *Symboles de Manin et valeurs de fonctions L*, Algebra, arithmetic, and geometry: in honor of Yu. I. Manin. Vol. II, Progr. Math., vol. 270, Birkhäuser Boston, Inc., Boston, MA, 2009, pp. 283–309. MR 2641193

[25] Philippe Michel, *Analytic number theory and families of automorphic L-functions*, Automorphic forms and applications, IAS/Park City Math. Ser., vol. 12, Amer. Math. Soc., Providence, RI, 2007, pp. 181–295. MR 2331346

[26] Asbjorn Christian Nordentoft, *A note on additive twists, reciprocity laws and quantum modular forms*, arXiv e-prints (2019), arXiv:1909.09665.

[27] ———, *On the distribution of the Eichler-Shimura map and the zeroes of period polynomials*, arXiv e-prints (2019), arXiv:1909.09622.

[28] Y. N. Petridis and M. S. Risager, *Modular symbols have a normal distribution*, Geom. Funct. Anal. **14** (2004), no. 5, 1013–1043. MR 2105951

[29] Yiannis N. Petridis and Morten S. Risager, *Arithmetic statistics of modular symbols*, Invent. Math. **212** (2018), no. 3, 997–1053. MR 3802302

[30] David E. Rohrlich, *Nonvanishing of L-functions for* $\mathrm{GL}(2)$, Invent. Math. **97** (1989), no. 2, 381–403. MR 1001846

[31] Robert J. Serfling, *Approximation theorems of mathematical statistics*, John Wiley & Sons, Inc., New York, 1980, Wiley Series in Probability and Mathematical Statistics. MR 595165

[32] Goro Shimura, *On the holomorphy of certain Dirichlet series*, Proc. London Math. Soc. (3) **31** (1975), no. 1, 79–98. MR 0382176

[33] ———, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan, vol. 11, Princeton University Press, Princeton, NJ, 1994, Reprint of the 1971 original, Kanô Memorial Lectures, 1. MR 1291394

[34] Don Zagier, *Quantum modular forms*, Quanta of maths, Clay Math. Proc., vol. 11, Amer. Math. Soc., Providence, RI, 2010, pp. 659–675. MR 2757599

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, Copenhagen 2100, Denmark

E-mail address: `acnordentoft@outlook.com`

PAPER B
ON THE DISTRIBUTION OF PERIODS OF HOLOMORPHIC CUSP FORMS AND ZEROES OF PERIOD POLYNOMIALS

ON THE DISTRIBUTION OF PERIODS OF HOLOMORPHIC CUSP FORMS AND ZEROES OF PERIOD POLYNOMIALS

ASBJØRN CHRISTIAN NORDENTOFT

Abstract. In this paper we determine the limiting distribution of the image of the Eichler–Shimura map or equivalently the limiting joint distribution of the coefficients of the period polynomials associated to a fixed cusp form. The limiting distribution is shown to be the distribution of a certain transformation of two independent random variables both of which are equidistributed on the circle $\mathbb{R}/\mathbb{Z}$, where the transformation is connected to the additive twist of the cuspidal L-function. Furthermore we determine the asymptotic behavior of the zeroes of the period polynomials of a fixed cusp form. We use the method of moments and the main ingredients in the proofs are additive twists of L-functions and bounds for both individual and sums of Kloosterman sums.

1. Introduction

Understanding the special values of L-functions is a notoriously hard problem and has deep arithmetic content due to the conjectures of Birch–Swinnerton-Dyer and Bloch–Kato. As a striking example of the connection between L-functions and arithmetics, Kolyvagin [11] proved that if $E/\mathbb{Q}$ is an elliptic curve such that the central value of the Hasse–Weil zeta function $L(E,1)$ is *non-zero*, then the set of rational points $E(\mathbb{Q})$ is *finite*.
Periods of automorphic forms have been an indispensable tool in the study of L-functions since the beginning of the theory (Hecke, Rankin–Selberg, Shimura, Manin) and continue to be so to this day ([7], [19], [14]). This paper is concerned with the distribution properties of automorphic periods, which in many cases are much more well-behaved and easier to handle than the values of the L-functions themselves (see below for a toy example of this phenomena).

The *Eichler–Shimura map* defines an isomorphism between the space of weight k holomorphic cusp forms and a parabolic cohomology group introduced by Eichler, by sending a cusp form to its periods. Our main result (see Theorem 1.1) describes the asymptotic distribution of the periods of a fixed cusp form or equivalently the asymptotic joint distribution of the coefficients of period polynomials. Furthermore we use our methods to derive an asymptotic expression for the zeroes of the period polynomials of a fixed cusp form (see Theorem 1.6), supplementing recent work of Jin, Ma, Ono, and Soundararajan [9], see also [4].

Date: December 4, 2020.

2010 *Mathematics Subject Classification.* 11F67(primary), and 11L05(secondary).

For $k = 2$ the period polynomials degenerate to constants and are known as *modular symbols* introduced by Birch and Manin. Petridis and Risager [17], [18] showed that modular symbols appropriately ordered are asymptotically normally distributed. From a cohomological point of view, the period polynomials are the natural generalization of modular symbols, but in this paper we show however that for $k \geq 4$ the coefficients of the period polynomials behave very differently from modular symbols.

1.1. **A toy example.** To illustrate the relation between periods and L-functions, let us consider a toy example of such automorphic periods given by the rational values of the complex exponential; $e^{2\pi i r}, r \in \mathbb{Q}$. These periods are connected to *Gauss sums*, which will serve as analogues of L-functions in this discussion;

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

where q is a positive integer and $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$ is a primitive Dirichlet character (see (1.1) below for one possible justification for the analogy between Gauss sums and L-functions). Gauss sums are in fact intimately connected to L-functions since $i^{-\kappa}\tau(\chi)q^{-1/2}$ is the *root number* of the Dirichlet L-function $L(\chi, s)$, where $\kappa = \frac{1-\chi(-1)}{2}$. More precisely the functional equation for Dirichlet L-functions takes the form;

$$\Lambda(\chi, s) := \Gamma\left(\frac{s+\kappa}{2}\right)\left(\frac{q}{\pi}\right)^{s/2} L(\chi, s) = \frac{i^{-\kappa}\tau(\chi)}{q^{1/2}} \Lambda(\overline{\chi}, 1-s).$$

To illustrate the difference in difficulty between dealing with periods and L-functions, we will consider the problem of determining the distribution of respectively the rational values of the complex exponential and the Gauss sums. It is easy to show that the periods themselves;

$$P_q := \{e^{2\pi i a/q} \mid (a, q) = 1\}$$

equidistribute on the unit circle as $q \to \infty$ (notice that this is not completely trivial because of the co-primality condition). In this case the Weyl sums for the distribution problem are Ramanujan sums, which can be evaluated explicitly.

On the other hand Gauss showed that $\tau(\chi)$ always has absolute value equal to $q^{1/2}$. But understanding the value distribution of

$$L_q := \{\tau(\chi) \mid \chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}, \text{primitive Dirichlet character}\},$$

as $q \to \infty$ turned out to be a much more difficult problem. This problem was solved by Katz [10] who showed (for q prime) that the Gauss sums also equidistribute (now on the circle with radius $q^{1/2}$) using deep input from algebraic geometry.

This example illustrates in a very simple setting the difference in difficulty between dealing with automorphic periods and L-functions themselves.

1.2. **The periods of holomorphic cusp forms.** In this paper we study periods of holomorphic cusp forms. The most famous example of a cusp form is probably the modular Δ-function introduced by Ramanujan as the following q-series;

$$\Delta(z) := q \prod_{n \geq 1} (1-q^n)^{24} = \tau(1)q + \tau(2)q^2 + \tau(3)q^3 + \ldots, \quad q = e^{2\pi i z}.$$

In this case, given a primitive Dirichlet character $\chi : (\mathbb{Z}/q\mathbb{Z})^\times \to \mathbb{C}$, we define the twisted L-function;

$$L(\Delta, \chi, s) := \sum_{n \geq 1} \chi(n) \tau(n) n^{-s},$$

which converges absolutely for $\operatorname{Re} s > 13/2$ and admits analytic continuation with a functional equation relating $s \leftrightarrow 12 - s$. In this case the special values $s = 1, \ldots, 11$ can be written as a twisted rational linear combination of *the periods of* Δ;

$$\pi^{-m} L(\Delta, \chi, m) = \sum_{\substack{-q/2 < a < q/2, \\ 0 \leq l \leq 10}} c(a/q, l, m)\, \chi(a) \underbrace{\int_{a/q}^{i\infty} \Delta(z) z^l dz}_{\text{periods}}, \tag{1.1}$$

where $m \in \{1, \ldots, 11\}$ and $c(a/q, l, m) \in \mathbb{Q}$ (see [13] for details, where this is used to prove rationality results for $L(\Delta, \chi, m)$ and to construct p-adic L-functions). Notice the similarity between this formula for the twisted special values and the formula for Gauss sums in the toy example above.
We will study the distribution of the periods of holomorphic cusp forms appearing in (1.1) or equivalently of the image of the Eichler–Shimura map.

To be more precise let $\mathcal{S}_k(\Gamma_0(N))$ denote the space of cusp forms of even weight k and level N. To each cusp form $f \in \mathcal{S}_k(\Gamma_0(N))$ and each $\gamma \in \Gamma_0(N)$, the Eichler–Shimura map associates the following $(k-1)$-dimensional complex vector consisting of the periods of f;

$$\begin{aligned} u_f(\gamma) &= (u_{f,0}(\gamma), u_{f,1}(\gamma), \ldots, u_{f,k-2}(\gamma)) \\ &:= \left(\int_{\gamma\infty}^{\infty} f(z) dz, \int_{\gamma\infty}^{\infty} f(z) z dz, \ldots, \int_{\gamma\infty}^{\infty} f(z) z^{k-2} dz \right)^T \in \mathbb{C}^{k-1}, \end{aligned} \tag{1.2}$$

where ^{T} denotes matrix transpose and $\gamma\infty = a/c$ with a, c the left-upper and -lower entry of γ. The map $u_f : \Gamma_0(N) \to \mathbb{C}^{k-1}$ can be shown to satisfy a 1-cocycle relation with respect to a certain action of $\Gamma_0(N)$ on $\mathbb{C}^{k-1}$, which we will make precise below in Section 2.1. Thus u_f defines an element of the cohomology group $H^1(\Gamma_0(N), M)$, where M is given by $\mathbb{C}^{k-1}$ equipped with the just mentioned action of $\Gamma_0(N)$. The association $f \mapsto u_f$ is a constituent of the *Eichler–Shimura isomorphism* as we will see below.
When ordered by the denominator of the cusp $\gamma\infty$, we show that the limiting distribution of $u_f(\gamma)$ is the distribution of a certain transformation of two independent random variables both of which are uniformly distributed on the circle $\mathbb{R}/\mathbb{Z}$ (see Theorem 1.1 and Theorem 1.8 below for the precise statements).

1.3. **Results for** $\Gamma_0(N)$**.** Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a cusp form of weight k with Fourier expansion;

$$f(z) = \sum_{n \geq 1} a_f(n) q^n, \quad q = e^{2\pi i z}.$$

Then for each $x \in \mathbb{R}$, we define the following Dirichlet series called the *additive twist by* x of the L-function of f;

$$L(f,x,s) := \sum_{n\geq 1} \frac{a_f(n)e(nx)}{n^s}, \tag{1.3}$$

where $e(x) = e^{2\pi i x}$. This Dirichlet series converges absolutely for any $x \in \mathbb{R}$ when $\operatorname{Re} s > (k+1)/2$ by Hecke's bound;

$$\sum_{n\leq X} |a_f(n)|^2 \ll_f X^k, \tag{1.4}$$

which is known to hold for cusp forms for general Fuchsian groups of the first kind. When x corresponds to a cusp (i.e. $x \in \mathbb{Q}$), the additive twist by x satisfies analytic continuation to the entire complex plane and if x is $\Gamma_0(N)$ equivalent to ∞, we also have a functional equation relating s and $k-s$ (see Section 2.3 for details).
For $c > 0$ such that $N|c$, we consider the periods u_f as a $(k-1)$-dimensional complex random variable defined on the outcome space;

$$\Omega_c := \{a/c \in \mathbb{Q} \mid a, c \in \mathbb{Z}_{\geq 0}, (a,c) = 1, 0 \leq a < c\}, \tag{1.5}$$

endowed with the uniform probability measure, where $u_f(a/c) := u_f(\gamma)$ for $a/c = \gamma\infty$ (i.e. a, c are the left upper- and lower entries of $\gamma \in \Gamma_0(N)$).
Our main result is that the limiting distribution as $c \to \infty$ (when appropriately normalized) is the tranformation of two independent distributions on the circle.

Theorem 1.1. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a cusp form of even weight $k \geq 4$ and level N. Then we have for any fixed box $A \subset \mathbb{C}^{k-1}$ that*

$$\begin{aligned}\mathbb{P}_{\Omega_c}\left(\frac{u_f(a/c)}{C_k c^{k-2}} \in A\right) :=& \frac{\#\{\frac{a}{c} \in \Omega_c \mid \frac{u_f(a/c)}{C_k c^{k-2}} \in A\}}{\#\Omega_c} \\ =& \mathbb{P}\left(F(Y,Z) \in A\right) + o(1)\end{aligned} \tag{1.6}$$

as $c \to \infty$ with $N|c$, where Y, Z are two independent random variables both distributed uniformly on $[0,1)$, $F : [0,1) \times [0,1) \to \mathbb{C}^{k-1}$ is given by

$$F(y,z) := L(f,y,k-1)\left(1, z, \ldots, z^{k-2}\right)^T,$$

and $C_k = \frac{i\Gamma(k-1)}{(2\pi)^{k-2}}$.
(Here $\mathbb{P}\left(F(Y,Z) \in A\right)$ denotes the probability of the event $F(Y,Z) \in A$).

Remark 1.2. As was noted in [1, Section 1.4.1] the individual distribution of the critical values of $L(f,\gamma\infty,s)$ for $s \neq k/2$ are not that interesting since for $\operatorname{Re} s > (k+1)/2$ the critical values are rational values of a continuous function and consequently the limiting distribution is just the pullback by this continuous function of the Lebesgue measure on the circle $\mathbb{R}/\mathbb{Z}$, since reduced fractions equidistribute (and similarly for $\operatorname{Re} s < (k+1)/2$ using the functional equation). In order to handle the distribution of the Eichler–Shimura map (or equivalently the coefficients of period polynomials), we however need to control the dependence between the different critical values of $L(f,\gamma\infty,s)$ and maps of the type $\gamma \mapsto (\gamma\infty)^j$. In the end, the specific shape of the limiting distribution amounts to the non-trivial cancellation in sum of Kloosterman sums with uniformity in the frequencies and thus non-trivial input is needed.

Remark 1.3. Given an orthogonal basis $f_1, \dots, f_d$ for $\mathcal{S}_k(\Gamma_0(N))$, we can also compute the joint distribution of

$$u_{k,N} := (u_{f_1}, \dots, u_{f_d})^T \in \mathbb{C}^{d(k-1)},$$

when appropriately normalized, with a similar proof. We have however restricted the exposition to a single cusp form f for notational simplicity. For the complete orthogonal basis the result is that the random variables defined from $\frac{(2\pi/c)^{k-2}}{i\Gamma(k-1)} u_{k,N}$ converge in distribution (in the same sense as in Theorem 1.1 above) to the random variable

$$F_{k,N}(Y, Z),$$

where Y, Z are two independent and uniformly distributed random variables on $[0, 1)$ and $F_{k,N} : [0, 1) \times [0, 1) \to \mathbb{C}^{d(k-1)}$ is given by

$$F_{k,N}(y, z) := \Big(L(f_1, Y, k-1), \dots, L(f_1, Y, k-1) z^{k-2}, \dots, L(f_d, Y, k-1) z^{k-2} \Big)^T \in \mathbb{C}^{d(k-1)}.$$

In particular it is worth noticing that $u_{f_i}(\gamma)$ and $u_{f_j}(\gamma)$ for $i \neq j$ are highly dependent as opposed to the case $k = 2$ (see [15, Theorem 5.1]).

Remark 1.4. If $f \in \mathcal{S}_k(\Gamma_0(N))$ then it follows from work of Jin, Ma, Ono and Soundararajan [9, Theorem 1.2] that for $k \geq 6$ the period polynomials $r_{f,S}(\sqrt{N}X)$ (see (1.7) for a definition) converge coefficient for coefficient to $X^{k-2} - 1$ as $N \to \infty$.

Remark 1.5. The author [15] and independently Bettin and Drappeau [1] (for level 1) have considered the distribution of central values of additive twists of L-functions of cusp forms of arbitrary even weight and showed that they are normally distributed. As was also noted in [15, Section 3.3.2] the coefficients of the period polynomial can be expressed as linear combinations of critical values of additive twists (including the central value). However the left-most critical value at $s = 1$ will be the dominating term, which is why we see that the distribution degenerates (and in particular is not normal).

1.4. **Zeroes of period polynomials.** The vector u_f encodes the periods of $f \in \mathcal{S}_k(\Gamma_0(N))$, which were introduced in a slightly different setting by M. Eichler in his study of parabolic cohomology [5]. He defined the *period polynomials* associated to f as

$$r_{f,\gamma}(X) := \int_{\gamma\infty}^{\infty} f(z)(z - X)^{k-2} dz = \sum_{j=0}^{k-2} X^j (-1)^j \binom{k-2}{j} \int_{\gamma\infty}^{\infty} f(z) z^{k-2-j} dz, \tag{1.7}$$

where $\gamma \in \Gamma_0(N)$. Note that the periods of f are equal to the coefficients of this polynomial (up to a scaling by factorials). The Eichler–Shimura isomorphism can also be described intrinsically and naturally in terms of period polynomials as was

done in [16]. Our results can be interpreted as determining the joint distribution of the coefficients of the period polynomials.
Recently there has been a lot of study in the analytic properties of period polynomials, especially the location of the zeroes of $r_{f,S}$, where $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ (see [4] for a complete list of references). The results of this paper should be seen more in relation with these results rather than with those of Petridis and Risager [18].

For $f \in \mathcal{S}_k(\Gamma_0(N))$ a newform of even weight $k \geq 6$, we can use our methods to understand the zeroes of $r_{f,a/c}$ asymptotically as $c \to \infty$. The assumptions on f are made in order to ensure that $L(f, x, k-1)$ is non-zero for all $x \in \mathbb{R}$.

Theorem 1.6. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a newform of even weight $k \geq 6$ and level N. Then $r_{f,\gamma}$ is a polynomial of degree $k-2$ for any $\gamma \in \Gamma_0(N)$. Furthermore all zeroes x_0 of $r_{f,\gamma}$ satisfy*

$$x_0 = a/c + O_k((|a/c| + 1)^{(k-4)/(k-2)} c^{-2/(k-2)}),$$

where a, c are the entries in the left column of γ (i.e. $\gamma\infty = a/c$).

Remark 1.7. Analogously Jin, Ma, Ono and Soundararajan [9, Theorem 1.2] building on works of others (see [4]) determined the zeroes of $r_{f,S}$ as either the weight k or level N tend to infinity. In their case the zeroes satisfy a version of the Riemann Hypothesis, of which no analogue seems to exist in our setting.

1.5. **Results for general cofinite Fuchsian groups.** We also obtain results for a general cofinite, discrete subgroup Γ of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at infinity of width 1 (see [8, Chapter 2] for definitions), but we have to take an extra average. Given a cusp form $f \in \mathcal{S}_k(\Gamma)$, we can similarly define the additive twists $L(f, x, s)$ of the associated L-function, which satisfy the same properties as in the case of Hecke congruence groups, as we will explain in Section 2.3 below.
To state our results, we introduce the following set;

$$T_{\leq 1} = T_{\leq 1,\Gamma} := \{r = \gamma\infty \in \mathbb{R} \mid \gamma \in \Gamma/\Gamma_\infty, 0 \leq r < 1\}. \tag{1.8}$$

This is a slight modification of the set $T = T_\Gamma$ defined in [18], which parametrizes the double coset $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$. In this paper we need to choose a representative, since $u_f(\gamma)$ is not invariant under the action of Γ_∞ from the left. One would get similar results by choosing different representatives.
Using the argument in the proof of [18, Proposition 2.2], we see that to any $r \in T_{\leq 1}$ there is a unique $\gamma \in \Gamma/\Gamma_\infty$ with lower-left entry $c > 0$ such that $r = \gamma\infty$ and we define $c(r) := c$. Now for $X > 0$, we consider u_f as a random variable on the outcome space;

$$\tilde{\Omega}_X := \{r \in T_{\leq 1} \mid c(r) \leq X\}. \tag{1.9}$$

endowed with the uniform probability measure. In this setting our result is the following.

Theorem 1.8. *Let $f \in \mathcal{S}_k(\Gamma)$ be a cusp form of even weight $k \geq 4$. Then we have for any fixed box $A \subset \mathbb{C}^{k-1}$ that*

$$\begin{aligned} \mathbb{P}_{\tilde{\Omega}_X}\left(\frac{u_f(r)}{C_k c(r)^{k-2}} \in A\right) :=& \frac{\#\{r \in \tilde{\Omega}_X \mid \frac{u_f(r)}{C_k c(r)^{k-2}} \in A\}}{\#\tilde{\Omega}_X} \\ =& \mathbb{P}\left(F(Y,Z) \in A\right) + o(1) \end{aligned} \tag{1.10}$$

as $X \to \infty$, where Y, Z are two independent random variables both distributed uniformly on $[0,1)$, $F : [0,1) \times [0,1) \to \mathbb{C}^{k-1}$ and C_k as in Theorem 1.1.

Acknowledgement

I would like to express my gratitude to Dorian Goldfeld and Columbia University for their hospitality and to my advisor Morten Risager and Riccardo Pengo for valuable suggestions. Finally I would like to thank the referees for their careful readings, which improved the quality of the paper.

2. Preliminaries and Background

In this section we will introduce some background on respectively the Eichler–Shimura isomorphism, bounds on sums of Kloosterman sums and finally additive twists of modular L-functions.

2.1. Background on the Eichler–Shimura isomorphism.

The purpose of this section is to show how the periods of f appear "in nature". We will see that from a cohomological point of view, u_f defines the natural higher weight analogue of modular symbols. We will refer to [22] for a comprehensive background.
Let G be any group and let M be a left $\mathbb{Z}[G]$-module. Then one can define cohomology groups;

$$H^i(G,M) := Z^i(G,M)/B^i(G,M),$$

consisting of a quotient of certain maps

$$u : \underbrace{G \times \ldots \times G}_{i} \to M,$$

corresponding to a specific choice of injective resolution.
In particular for $i = 1$ we have the following explicit description;

$$Z^1(G,M) = \{u : G \to M \mid u(g_1 g_2) = u(g_1) + g_1 u(g_2),\ \forall g_1, g_2 \in G\},$$

$$B^1(G,M) = \{v : G \to M \mid \exists x_v \in M \text{ such that } v(g) = (g-1)x_v,\ \forall g \in G\}.$$

Now fix a subset $P \subset G$ and consider

$$Z^1_P(G,M) := \{u \in Z^1(G,M) \mid u(p) \in (p-1)M, \forall p \in P\},$$

which we note still contains the boundaries $B^1(G,M)$. From this we define the first P-cohomology group as;

$$H^1_P(G,M) := Z^1_P(G,M)/B^1(G,M).$$

In our case we consider $G = \Gamma$, a discrete, co-finite, torsion-free subgroup of $\mathrm{PSL}_2(\mathbb{R})$, and let P be the set of parabolic elements of Γ. We note that parabolic cohomology groups carry a natural Hecke action.

Now consider $M = V_{k-2}(\mathbb{C}) \cong \mathrm{Sym}^n(\mathbb{C}^2)$, the space of homogenous polynomials in two variables of degree $k-2$ with coefficients in $\mathbb{C}$, equipped with the following left-action of Γ;

$$(\gamma.P)(X,Y) := P((X,Y)\gamma) = P((aX+cY, bX+dY)),$$

for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ and $P \in V_{k-2}(\mathbb{C}) \subset \mathbb{C}[X,Y]$. From this data we form Eichler's parabolic cohomology group $H^1_P(\Gamma, V_{k-2}(\mathbb{C}))$.
Given a cusp form $f \in \mathcal{S}_k(\Gamma)$ of weight k, we can define a map $\sigma_f : \Gamma \to V_{k-2}(\mathbb{C})$ as;

$$\sigma_f(\gamma)(X,Y) := \int_{\gamma\infty}^{\infty} f(z)(Xz+Y)^{k-2}dz,$$

and it can be shown that $\sigma_f \in Z^1_P(\Gamma, V_{k-2}(\mathbb{C}))$. We similarly define

$$\sigma_{\overline{f}} \in Z^1_P(\Gamma, V_{k-2}(\mathbb{C})),$$

for $\overline{f} \in \overline{\mathcal{S}_k(\Gamma)}$ an anti-holomorphic cusp form of weight k. Note that when $k=2$, σ_f is exactly *the modular symbol map* of [18, (1.1)].
The main theorem of Eichler–Shimura [22, Proposition 6.2.3,Proposition 6.2.5] is now that the $\mathbb{C}$-linear map;

$$\begin{aligned} \mathcal{S}_k(\Gamma) \oplus \overline{\mathcal{S}_k(\Gamma)} &\to H^1_P(\Gamma, V_{k-2}(\mathbb{C})) \\ (f, \overline{g}) &\mapsto (\gamma \ \mapsto \sigma_f(\gamma) + \sigma_{\overline{g}}(\gamma)) \end{aligned}$$

is an isomorphism, which carries a natural action of the Hecke algebra as explained in [21, Section 8.3] (see also the seminal paper [2] for a purely algebraic proof of these facts).

Observe that the periods that we will study in this paper $u_f(\gamma)$ are (up to simple scaling by binomial coefficients) given by the coefficients of $\sigma_f(\gamma)(X,Y)$, and one can use the above to define an (equivalent) action of Γ on $\mathbb{C}^{k-1}$ directly (similar to the action of Γ on $\mathbb{R}^{k-1}$ described in [21, Chapter 8]), which was alluded to in the introduction. Thus we see that from a cohomological point of view the periods u_f define a natural generalization of modular symbols when $f \in \mathcal{S}_k(\Gamma)$, $k \geq 4$.
Furthermore we notice the following obvious connection with the period polynomials defined in (1.7);

$$r_{f,\gamma}(X) = \sigma_f(\gamma)(1, -X).$$

The reason why we used the definition (1.7) of the period polynomials was to make the connection to the results listed in [4] clear.

Remark 2.1. In fact there is a notion of *modular symbols* associated to $\mathcal{S}_k(\Gamma)$ for all weights k [22, Section 1.2], and one can show that the parabolic cohomology groups $H^1_P(\Gamma, V_{k-2}(\mathbb{C}))$ are isomorphic to the *cuspidal modular symbols* (see [22, Theorem 5.2.1] for details)

2.2. **Spectral bounds of sums of Kloosterman sums.** An important ingredient when proving our main results is the cancellation in Kloosterman sums. For arithmetic subgroups we have very strong bounds for individual Kloosterman sums from Weil's work on the Riemann Hypothesis over finite fields, but for general Fuchsian groups of the first kind, we only have non-trivial bounds when we average over the moduli. Below we will collect the results we will need on Kloosterman sums.

Let Γ be a co-finite, discrete subgroup of $\mathrm{PSL}_2(\mathbb{R})$ with a cusp at infinity of width 1. Then we define the Kloosterman sum with frequencies m, n and modulus c (the lower-left entry of some matrix $\gamma \in \Gamma$) as;

$$S(m,n;c) := \sum_{\begin{pmatrix} a & * \\ c & d \end{pmatrix} \in \Gamma_\infty \backslash \Gamma / \Gamma_\infty} e\left(m\frac{d}{c} + n\frac{a}{c}\right). \tag{2.1}$$

It can be shown that

$$\#\left\{ \begin{pmatrix} a & * \\ c & d \end{pmatrix} \in \Gamma_\infty \backslash \Gamma / \Gamma_\infty \mid 0 \leq c \leq X \right\} \ll X^2,$$

which yields the following trivial bound

$$S(m,n;c) \ll c^2,$$

uniformly in m, n, see [8, Proposition 2.8]. If $\Gamma = \Gamma_0(N)$ is a Hecke congruence group, we can do much better by Weil's bound;

$$|S(m,n;c)| \leq d(c)c^{1/2}(m,n,c)^{1/2}, \tag{2.2}$$

where d is the divisor function. The point is now that if we average over the moduli c, we can also detect cancelation in Kloosterman sums for general Γ.

2.2.1. *Spectral theory of Kloosterman sums.* The most powerful tools for obtaining bounds for sums of Kloosterman sums come from the spectral theory of automorphic forms following an approach initiated by Selberg. We refer to [8] for a comprehensive background on the spectral theory of automorphic forms.

In this approach the spectrum of the automorphic Laplacian $\Delta = \Delta_\Gamma$ plays a prominent role, which in local coordinates is given by

$$\Delta_\Gamma = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right).$$

It can be shown that Δ_Γ with domain given by smooth and bounded functions on $\Gamma\backslash\mathbb{H}$, defines a non-negative, unbounded operator with a unique self-adjoint extension (which we also denote $\Delta = \Delta_\Gamma$). We observe that $\lambda = 0$ is always an eigenvalue of Δ_Γ corresponding to the constant function. Furthermore the famous *Selberg conjecture* predicts that for congruence subgroups $\Gamma_0(N)$ the first non-zero eigenvalue is $\geq 1/4$. It is known that there exist non-congruence subgroups Γ such that Δ_Γ has non-zero eigenvalues arbitrarily close to 0 as explained in [8, (11.15)].

For $n = 0$ the Kloosterman sums reduce to a generalization of the classical Ramanujan sums and the mth Fourier coefficient of the Eisenstein series;

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

is exactly

$$\Gamma(s)\zeta(2s)^{-1} \sum_{c>0} \frac{S(m,0;c)}{c^{2s}},$$

where the sum is taken over lower-left entries of matrices in Γ. Recall that by the general theory of Eisenstein series due to Selberg, $E(z,s)$ has its rightmost pole at $s=1$, which is a simple pole with residue $\mathrm{vol}(\Gamma)^{-1}$, [8, Proposition 6.13]. All the other finitely many poles in $1/2 < \mathrm{Re}\, s < 1$ are also simple and the residues are eigenfunctions for Δ.
First of all lets see how to use the analytic properties of Eisenstein series to understand the asymptotic size of the outcome space $\tilde{\Omega}_X$: This is possible since we have a bijection

$$\Gamma_\infty\backslash\Gamma/\Gamma_\infty \leftrightarrow T_{\leq 1,\Gamma} \cup \{\infty\},$$

with $T_{\leq 1,\Gamma}$ as defined in (1.8). Thus we see that the constant term in the Fourier expansion of $E(z,s)$ is exactly the generating series for $T_{\leq 1,\Gamma}$. Since the pole of $E(z,s)$ is a constant, the constant term in the Fourier expansion of $E(z,s)$ also has a simple pole (with the same residue). Now by a standard complex analysis argument we get

$$\#\tilde{\Omega}_X = \frac{X^2}{\mathrm{vol}(\Gamma)} + O(X^{2-\delta_\Gamma}), \tag{2.3}$$

for some $\delta_\Gamma > 0$ depending on the spectral gap for Γ, with $\tilde{\Omega}_X$ as in (1.9).
Furthermore since the pole at $s=1$ of the Eisenstein series has constant residue, it follows that for $m \neq 0$ the Dirichlet series

$$\sum_c \frac{S(m,0;c)}{c^{2s}},$$

where the sum is over lower-left entries of matrices in Γ, has analytic continuation to $\mathrm{Re}\, s > \mathrm{Re}\, s_1 \geq 1/2$ where $\lambda_1 = s_1(1-s_1)$ is the smallest non-zero eigenvalue. From this one easily proves

$$\sum_{c\leq X} S(m,0;c) \ll_\Gamma |m|^{1/2} X^{2-\delta_\Gamma},$$

for some $\delta_\Gamma > 0$ (see [18, (3.6)]).
For $mn \neq 0$ the corresponding Dirichlet series

$$\sum_c \frac{S(m,n;c)}{c^{2s}},$$

shows up in the Fourier coefficients of the Poincaré series

$$P_m(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} e(m\gamma z)(\mathrm{Im}\,\gamma z)^s,$$

as was brilliantly used by Goldfeld and Sarnak in [6] to obtain bounds on sums of Kloosterman sums. Using analytic properties of the resolvent of Δ_Γ, they show that $P_m(z,s)$ has meromorphic continuation with possible poles only at the spectrum of Δ_Γ and from this they obtain bounds for sums of Kloosterman sums. For our applications the dependence on m,n is essential, but this dependence is not clear from the statement of their theorem [6, Theorem 2]. However using [6, Remark 1] one can easily adapt their arguments to deduce the bound

$$\sum_{c\leq X} S(m,n;c) \ll_\Gamma mnX^{2-\delta_\Gamma}, \tag{2.4}$$

for some $\delta_\Gamma > 0$ depending on the spectral gap of Γ. We will omit the details.

2.3. **Additive twists.** The idea behind the proofs of the main theorems is to relate the periods of $f \in \mathcal{S}_k(\Gamma)$ to critical values of additive twists of the L-function of f. The additive twists are defined as

$$L(f, r, s) := \sum_{n \geq 1} \frac{a_f(n) e(nr)}{n^s},$$

where $r \in \mathbb{R}$ and $e(x) = e^{2\pi i x}$ which apriori converges for $\mathrm{Re}\, s > (k+1)/2$ by Hecke's bound (1.4). If r corresponds to a cusp of Γ then $L(f, r, s)$ admits analytic continuation by the integral representation;

$$L(f, r, s) = \frac{(2\pi)^s}{\Gamma(s)} \int_0^\infty f(r + iy) y^s \frac{dy}{y}.$$

Furthermore if $r = a/c = \gamma\infty$ with

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma,$$

the completed L-function satisfies the following functional equation;

$$\begin{aligned} \Lambda(f, a/c, s) &:= \Gamma(s) \left(\frac{c}{2\pi}\right)^s L(f, a/c, s) \\ &= (-1)^{k/2} \Lambda(f, -d/c, k - s), \end{aligned} \tag{2.5}$$

where $-d/c = \overline{r} = \gamma^{-1}\infty$ (see for instance [12, Section A.3]).
The relation between the periods of f and additive twists is given by the following.

Lemma 2.2. *Let $l \in \mathbb{Z}_{\geq 0}$ be a non-negative integer, $\gamma \in \Gamma$ and f as above. Then we have*

$$\int_{\gamma\infty}^\infty f(z) z^l dz = \sum_{j=0}^{l} \binom{l}{j} (a/c)^{l-j} (-2\pi i)^{-j-1} \Gamma(j+1) L(f, a/c, j+1), \tag{2.6}$$

where $a/c = \gamma\infty$.

Proof. By a straight forward computation we have

$$\begin{aligned} \int_{\gamma\infty}^\infty f(z) z^l dz &= i \int_0^\infty f(a/c + it)(a/c + it)^l dt \\ &= \sum_{j=0}^{l} i^{j+1} (a/c)^{l-j} \int_0^\infty f(a/c + it) t^j dt \\ &= \sum_{j=0}^{l} i^{j+1} (a/c)^{l-j} (2\pi)^{-j-1} \Gamma(j+1) L(f, a/c, j+1), \end{aligned}$$

as wanted. □

It turns out that the dominating term for all of these periods will be the left-most critical value $L(f, a/c, 1)$. This is hinted to by the following proposition.

Proposition 2.3. *For $a/c \in T_{\leq 1,\Gamma}$ (i.e. a, c are respectively, the left upper and left lower entries of some matrix in Γ such that $0 < a/c < 1$), we have the following bounds;*

(i) $L(f, a/c, \sigma) \ll 1$ *for* $\sigma \geq k/2 + 1$,
(ii) $L(f, a/c, k/2) \ll_\varepsilon c^\varepsilon$,
(iii) $L(f, a/c, \sigma) \ll c^{k-2\sigma}$ *for* $\sigma \leq k/2 - 1$,

as $c \to \infty$.

Proof. **Case** *(i)* For $\sigma \geq k/2+1$ we get by Hecke's bound (1.4) the following uniform bound;

$$L(f, a/c, \sigma) \ll \sum_{n \geq 1} \frac{|a_f(n)|}{n^\sigma} \leq \sum_{n \geq 1} \frac{|a_f(n)|}{n^{k/2+1}} < \infty,$$

which is independent of a/c and $\sigma \geq k/2 + 1$.

Case *(ii)* The bound on the central value was proved by the author [15, Corollary 5.8].

Case *(iii)* Finally for $\sigma \leq k/2 - 1$, we get by the functional equation (2.5) the following;

$$L(f, a/c, \sigma) = \frac{(-1)^{k/2}\Gamma(k-\sigma)(2\pi)^{-k+\sigma}}{\Gamma(\sigma)(2\pi)^{-\sigma}} c^{k-2\sigma} L(f, -d/c, k-\sigma),$$

and since $k - \sigma \geq k/2+1$ the result follows from *(i)*. Observe that we avoid the poles of the Γ-function in the numerator. □

3. On the zeroes of the period polynomials

In this section we will apply the bounds in Proposition 2.3 to determine the asymptotic behavior of the zeroes of the period polynomials associated to a fixed cusp form as the denominator of the cusp varies.

Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a fixed newform of even weight $k \geq 6$. Consider the period polynomials associated to f;

$$r_{f,\gamma}(X) = \int_{\gamma\infty}^{\infty} f(z)(z-X)^{k-2} dz = b_{f,k-2}(\gamma) X^{k-2} + \ldots + b_{f,0}(\gamma),$$

where $\gamma \in \Gamma$ and

$$\begin{aligned} b_{f,l}(\gamma) &= (-1)^l \binom{k-2}{l} \int_{\gamma\infty}^{\infty} f(z) z^{k-2-l} dz \\ &= \sum_{j=0}^{k-2-l} \frac{(-1)^l \binom{k-2}{l}\binom{k-2-l}{j}}{(-2\pi i)^{j+1}} (a/c)^{k-2-l-j} \Gamma(j+1) L(f, a/c, j+1) \end{aligned}$$

are the coefficients of $r_{f,\gamma}$. We have the following bound on the Fourier coefficients of f due to Deligne [3];

$$|a_f(n)| \leq d(n) n^{(k-1)/2},$$

where d is the divisor function. This implies that

$$\sum_{n\geq 2}\frac{|a_f(n)|}{n^{k-1}}\leq \sum_{n\geq 2}\frac{d(n)}{n^{(k-1)/2}}=\zeta((k-1)/2)^2-1\leq \zeta(5/2)^2-1=0.799...<1.$$

This shows that $L(f,x,k-1)$ is bounded both from above and away from zero uniformly in $x\in\mathbb{R}$. Combining this observation with the functional equation for additive twists we conclude that

$$\begin{aligned} b_{f,k-2}(\gamma) &= \frac{L(f,\gamma\infty,1)}{-2\pi i} \\ &= \frac{(-1)^{k/2}i\Gamma(k-1)}{(2\pi)^{k-1}}L(f,\gamma^{-1}\infty,k-1)c^{k-2}\neq 0. \end{aligned} \tag{3.1}$$

Thus $r_{f,\gamma}$ is actually a polynomial of degree $k-2$ and normalizing it so that it becomes a monic polynomial the coefficients become;

$$\tilde{b}_l(\gamma)=\tilde{b}_{f,l}(\gamma):=b_{f,l}(\gamma)/b_{f,k-2}(\gamma),\quad l=0,\dots,k-2.$$

We can now prove the promised asymptotic expression for the zeroes of $r_{f,\gamma}$ as $c\to\infty$.

Proof of Theorem 1.6. Let $r=\gamma\infty=a/c$. Using (3.1) we see that $b_{f,k-2}(\gamma)\gg_k c^{k-2}$. Combining this with the expression (2.6) and the bounds from Proposition 2.3, we conclude the following;

$$\begin{aligned} \tilde{b}_l(\gamma) &= (-1)^l\binom{k-2}{l}r^{\,k-2-l}+O_k\left(\sum_{j=1}^{k-2-l}|r|^{k-2-l-j}\frac{|L(f,r,j+1)|}{|b_{f,k-2}(\gamma)|}\right) \\ &= (-1)^l\binom{k-2}{l}r^{\,k-2-l}+O_k\left(\sum_{j=1}^{k-2-l}|r|^{k-2-l-j}c^{\max(0,k-2j-2)}c^{-(k-2)}\right). \end{aligned}$$

One easily checks that $|r|^{k-2-l-j}c^{\max(0,k-2j-2)}c^{-(k-2)}\ll |r|^{k-4-l}c^{-2}$ for all $j=1,\dots,k-2-l$ using that $|r|\geq c^{-1}$. Thus we conclude

$$\tilde{b}_l(\gamma)=(-1)^l\binom{k-2}{l}r^{\,k-2-l}+O_k(|r|^{k-4-l}c^{-2}), \tag{3.2}$$

and in particular $\tilde{b}_l(\gamma)\ll_k |r|^{k-2-l}$.

Now we will show that any zero x_0 of $r_{f,\gamma}$ is bounded by $O_k(|r|)$. So assume that a zero x_0 of $r_{f,\gamma}$ satisfies $|x_0|\geq |r|$. Then using (3.2), we get the bound

$$|x_0|^{k-2}=|-\tilde{b}_{k-3}(\gamma)x_0^{k-3}-\ldots-\tilde{b}_0(\gamma)|\ll_k |r||x_0|^{k-3},$$

which implies $x_0\ll_k |r|$ as wanted.

Now combining $x_0\ll_k |r|$ with (3.2), we conclude for any root x_0 of $r_{f,\gamma}$ we have that

$$0=x_0^{k-2}+\tilde{b}_{k-3}(\gamma)x_0^{k-3}+\ldots+\tilde{b}_0(\gamma)=(x_0-r)^{k-2}+O_k((|r|+1)^{k-4}c^{-2}),$$

which implies that $|x_0-r|\ll_k (|r|+1)^{(k-4)/(k-2)}c^{-2/(k-2)}$ as wanted. $\square$

If we restrict to $\gamma\in\Gamma_\infty\backslash\Gamma_0(N)$ such that $r=a/c=\gamma\infty\in\Omega_c$ (i.e. $0<r=a/c<1$) we conclude that the zeroes of $r_{f,\gamma}$ satisfy

$$x_0=a/c+O_k(c^{-2/(k-2)}).$$

4. On the distribution of the Eichler–Shimura map

In this section we will prove Theorem 1.1 and Theorem 1.8 using the method of moments. More precisely this is done by firstly computing all the moments of the random variable u_f on respectively Ω_c and $\tilde{\Omega}_X$ and then applying a result from probability theory due to Fréchet–Shohat to determine the limiting distribution.

4.1. **Computation of the moments of u_f.**

To state our results we let (as above)

$$f(z) = \sum_{n\geq 1} a_f(n) q^n,$$

be the Fourier expansion of a cusp form $f \in \mathcal{S}_k(\Gamma)$. Then we define the following Dirichlet series for $\alpha, \beta \in \mathbb{Z}_{\geq 0}$;

$$\begin{aligned} L_{f,\alpha,\beta}(s) &:= \sum_{\substack{n_1,\ldots,n_{\alpha+\beta}>0 \\ n_1+\ldots+n_\alpha = n_{\alpha+1}+\ldots+n_{\alpha+\beta}}} \frac{a_f(n_1)\cdots a_f(n_\alpha)\overline{a_f(n_{\alpha+1})}\cdots\overline{a_f(n_{\alpha+\beta})}}{(n_1\cdots n_{\alpha+\beta})^s} \\ &= \int_0^1 L(f,x,s)^{\alpha,\beta} dx, \end{aligned} \tag{4.1}$$

which converges absolutely for $\operatorname{Re} s > (k+1)/2$ by Hecke's bound (1.4), where we use the notation $z^{\alpha,\beta} = z^\alpha \overline{z}^\beta$.
For $\Gamma = \Gamma_0(N)$ a Hecke congruence group, we get the following calculation of the moments.

Theorem 4.1. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a cusp form of even weight $k \geq 4$. Then for any non-negative integers;*

$$\alpha_0, \ldots, \alpha_{k-2}, \beta_0, \ldots, \beta_{k-2},$$

not all zero and $c \equiv 0\,(N)$, we have that

$$\begin{aligned} &\frac{1}{\varphi(c)} \sum_{\substack{0\leq a<c, \\ (a,c)=1}} \prod_{j=0}^{k-2} \left(\frac{(2\pi/c)^{k-2}}{\Gamma(k-1)i} \int_{a/c}^{\infty} f(z) z^j dz \right)^{\alpha_j,\beta_j} \\ &= \frac{L_{f,\alpha,\beta}(k-1)}{1+\sum_{j=0}^{k-2} j\cdot(\alpha_j+\beta_j)} + O_{\varepsilon,\alpha,\beta,f}(c^{-1/6+\varepsilon}), \end{aligned} \tag{4.2}$$

where $\alpha = \alpha_0 + \ldots + \alpha_{k-2}$ and $\beta = \beta_0 + \ldots + \beta_{k-2}$.

For a general cofinite Fuchsian group Γ, we have to take an extra average in order to calculate the moments.

Theorem 4.2. *Let Γ be a cofinite Fuchsian group with a cusp at ∞ of width 1 and let $f \in \mathcal{S}_k(\Gamma)$ be a cusp form of even weight $k \geq 4$. Then for any non-negative integers;*

$$\alpha_0, \ldots, \alpha_{k-2}, \beta_0, \ldots, \beta_{k-2},$$

not all zero, we have that

$$\begin{aligned}\frac{1}{\#\tilde{\Omega}_X}\sum_{r\in\tilde{\Omega}_X}\prod_{j=0}^{k-2}\left(\frac{(2\pi/c(r))^{k-2}}{\Gamma(k-1)i}\int_r^\infty f(z)z^j dz\right)^{\alpha_j,\beta_j}\\ =\frac{L_{f,\alpha,\beta}(k-1)}{1+\sum_{j=0}^{k-2}j\cdot(\alpha_j+\beta_j)}+O_{\alpha,\beta}(X^{-\delta_\Gamma}),\end{aligned} \tag{4.3}$$

for some $\delta_\Gamma > 0$ depending on the spectral gap of Γ, where $\alpha = \alpha_0 + \ldots + \alpha_{k-2}$ and $\beta = \beta_0 + \ldots + \beta_{k-2}$.

Remark 4.3. Observe that the main terms above are exactly what we expect from the statements of Theorem 1.1 and Theorem 1.8, since $L_{f,\alpha,\beta}(k-1)$ is precisely the (α,β)-moment of $L(f,Y,k-1)$ with Y equidistributed on $[0,1)$ (see (4.1)) and

$$\int_0^1 z^{\alpha_1+\beta_1}z^{2(\alpha_2+\beta_2)}\ldots z^{(k-2)(\alpha_{k-2}+\beta_{k-2})}dz=\frac{1}{1+\sum_{j=0}^{k-2}j\cdot(\alpha_j+\beta_j)}.$$

Proof of Theorem 4.1 and Theorem 4.2. In the following all implied constants may depend on f, α and β. In view of (2.6) we can express the periods of f as a linear combination of critical values of the additive twists $L(f,r,s)$ and by the functional equation, we have the equality

$$L(f,r,1)=c(r)^{k-2}\frac{\Gamma(k-1)}{(2\pi)^{k-2}}L(f,\overline{r},k-1)$$

with $r=\gamma\infty$ and $\overline{r}=\gamma^{-1}\infty$. Using Proposition 2.3 this implies that

$$\begin{aligned}&\prod_{j=0}^{k-2}\left(\frac{(2\pi/c(r))^{k-2}}{\Gamma(k-1)i}\int_r^\infty f(z)z^j dz\right)^{\alpha_j,\beta_j}\\ &=L(f,\overline{r},k-1)^{\alpha,\beta}r^M+O(c(r)^{-2})\end{aligned} \tag{4.4}$$

where $z^{\alpha,\beta}=z^\alpha\overline{z}^\beta$ and

$$M=M(\alpha_1,\ldots,\alpha_{k-2},\beta_1,\ldots,\beta_{k-2}):=\sum_{j=0}^{k-2}j\cdot(\alpha_j+\beta_j).$$

In order to deal with the term r^M, we apply a standard smooth approximation. So let $\varphi:\mathbb{R}\to\mathbb{R}_{\geq 0}$ be a smooth function with compact support in $(0,1)$ such that $\int_0^1\varphi(x)dx=1$. Then we define the following approximation to the Dirac measure at $x=0$;

$$\varphi_\delta(x):=\delta^{-1}\varphi(x/\delta),$$

where $\delta>0$ is some small constant to be chosen appropriately. Notice that φ_δ is supported in $(0,\delta)$ and satisfies $\int_{\mathbb{R}}\varphi_\delta(x)dx=1$. We think of φ_δ as a function on the circle $\mathbb{R}/\mathbb{Z}$ by extending its values on $[0,1)$ periodically.
Associated to the periodic functions $h_j:\mathbb{R}/\mathbb{Z}\to\mathbb{R}$ given by $h_j(x)=x^j$ for $x\in[0,1)$ with $j\in\mathbb{Z}_{\geq 0}$, we define the following smooth approximation;

$$h_{j,\delta}:=h_j*\varphi_\delta,$$

where $*$ denotes the (additive) convolution product on $\mathbb{R}/\mathbb{Z}$. The approximation $h_{j,\delta}$ satisfies the following standard properties (which can be proved easily by partial integration using the compact support of φ_δ);

$$\widehat{h_{j,\delta}}(l) \ll_A \frac{1}{(\delta(1+|l|))^A}, \quad \widehat{h_{j,\delta}}(0) = \widehat{h_j}(0) = \frac{1}{j+1}, \tag{4.5}$$

where $A > 0$ and $\widehat{h_{j,\delta}}$ denotes the Fourier transform on $\mathbb{R}/\mathbb{Z}$. And furthermore

$$h_{j,\delta}(x) = h_j(x) + \int_{x-\delta}^{x} (h_j(t) - h_j(x))\varphi_\delta(x-t)dt = h_j(x) + O_j(\delta),$$

for $\delta \leq x < 1$ using that h_j is differentiable. This estimate fails for $0 \leq x \leq \delta$, but it is standard to show that the contribution from $r \in \tilde{\Omega}_X$ (respectively $r \in \Omega_c$) with $r < \delta$ is negligible and will not affect the error terms. More precisely it is obvious that $\{r \in \Omega_c \mid r < \delta\} \ll \delta c$ and by using the cancellation in Kloosterman sums one can easily show that $\{r \in \tilde{\Omega}_X \mid r < \delta\} \ll \delta X^2$.
The upshot is that we can replace r^M by the approximation $h_{M,\delta}$ at the cost of changing the error term in (4.4) to $O(\delta + c(r)^{-2})$ (at least when we average over Ω_c, respectively $\tilde{\Omega}_X$).

Finally we replace $h_{M,\delta}$ by its Fourier expansion to arrive at the following expression for the main term;

$$\begin{aligned} &L(f,\overline{r},k-1)^{\alpha,\beta} h_{M,\delta}(r) \\ &= \sum_{l\in\mathbb{Z}} \widehat{h_{M,\delta}}(l) e(lr) L(f,\overline{r},k-1)^{\alpha,\beta} \\ &= \sum_{l\in\mathbb{Z}} \widehat{h_{M,\delta}}(l) \sum_{n_1,\ldots,n_{\alpha+\beta}>0} \frac{a_f(n_1)\cdots a_f(n_\alpha)\overline{a_f(n_{\alpha+1})}\cdots\overline{a_f(n_{\alpha+\beta})}}{(n_1\cdots n_{\alpha+\beta})^{k-1}} \\ &\quad \times e(lr + \overline{r}(n_1+\ldots+n_\alpha - n_{\alpha+1} - \ldots - n_{\alpha+\beta})), \end{aligned} \tag{4.6}$$

using that $L(f,\overline{r},k-1)$ is absolutely convergent and so is the Fourier expansion of $h_{M,\delta}$ in view of (4.5).

Now the case where $\Gamma = \Gamma_0(N)$ is a Hecke congruence group, we average (4.4) over $r \in \Omega_c$. Since all of the r-dependence is in the exponential, we see the Kloosterman sums entering the picture. The main contribution comes from the *diagonal terms* corresponding to $l = 0$ and $n_1 + \ldots + n_\alpha = n_{\alpha+1} + \ldots + n_{\alpha+\beta}$, which contribute

$$L_{f,\alpha,\beta}(k-1)\widehat{h_{M,\delta}}(0) = L_{f,\alpha,\beta}(k-1)\frac{1}{M+1}. \tag{4.7}$$

In order to handle the off-diagonal contributions, we apply Weil's bound (2.2), which bounds the off-diagonal terms by the following;

$$\begin{aligned}
&\ll \frac{d(c)c^{1/2}}{\varphi(c)} \sum_{l\neq 0} \sum_{n_1,\dots,n_{\alpha+\beta}} |\widehat{h_{M,\delta}}(l)| \frac{|a_f(n_1)\cdots a_f(n_{\alpha+\beta})|}{(n_1\cdots n_{\alpha+\beta})^{k-1}} \left(l, c, \sum_{i=1}^{\alpha} n_i - \sum_{j=\alpha+1}^{\alpha+\beta} n_j\right)^{1/2} \\
&\ll_{\varepsilon,\alpha,\beta} \frac{c^{1/2+\varepsilon}}{\varphi(c)} \left(\sum_{l\neq 0} |\widehat{h_{M,\delta}}(l)|\right) \\
&\qquad \cdot \left(\sum_{n_1,\dots,n_{\alpha+\beta}} \frac{|a_f(n_1)\cdots a_f(n_{\alpha+\beta})|}{(n_1\cdots n_{\alpha+\beta})^{k-1}} \max(n_1,\dots,n_{\alpha+\beta})^{1/2-\varepsilon}\right) \\
&\ll_{\varepsilon,\alpha,\beta} \frac{c^{1/2+\varepsilon}}{\varphi(c)} \sum_{l\neq 0} |\widehat{h_{M,\delta}}(l)|,
\end{aligned}$$

using Hecke's bound (1.4) to show finiteness of the sum over $n_1,\dots,n_{\alpha+\beta}$. Combining the above with the fact that $\varphi(c) \gg_\varepsilon c^{1-\varepsilon}$, we arrive at the following;

$$\frac{1}{\varphi(c)} \sum_{\substack{0\leq a<c \\ (a,c)=1}} \prod_{j=0}^{k-2} \left(\frac{(2\pi/c)^{k-2}}{i\Gamma(k-1)} \int_{a/c}^{\infty} f(z) z^j dz\right)^{\alpha_j,\beta_j}$$

$$=L_{f,\alpha,\beta}(k-1)\widehat{h_{M,\delta}}(0) + O_\varepsilon\left(\delta + c^{-2} + c^{-1/2+\varepsilon} \sum_{l\neq 0} |\widehat{h_{M,\delta}}(l)|\right). \tag{4.8}$$

Next we apply (4.5) with $A = 2+\varepsilon$ to ensure convergence of the sum $\sum_{l\neq 0} |\widehat{h_{M,\delta}}(l)|$ and arrive at the following error term $O_\varepsilon(\delta + c^{-2} + c^{-1/2+\varepsilon}\delta^{-2-\varepsilon})$. Finally we choose $\delta = c^{-1/6}$ to balance the error terms.

The argument when Γ is a general cofinite Fuchsian group is similar, only now we average (4.4) over $r \in \tilde{\Omega}_X$. In this case we also see Gauss sums enter the picture since we have

$$\sum_{r\in\tilde{\Omega}_X} e(nr + m\overline{r}) = \sum_{0<c\leq X} S_\Gamma(m,n;c),$$

where the sum is taking over lower left entries c of matrices in Γ and $S_\Gamma(m,n;c)$ is a (generalized) Kloosterman sum defined by (2.1).

Again the main contribution is given by (4.7). When dealing with the off-diagonal contribution, we first of all have to trivially bound the terms in (4.6) with

$$\min(n_1,\dots,n_{\alpha+\beta}) > X^{\delta_1},$$

for some $\delta_1 > 0$ to be chosen appropriately. This is necessary since the dependence on the frequencies in (2.4) is not as strong as in Weil's bound (actually this extra step is only needed when $k=4$).

Now using the trivial bound for the exponentials, this truncation yields

$$\frac{1}{\#\tilde{\Omega}_X}\sum_{r\in\tilde{\Omega}_X} L(f,\overline{r},k-1)^{\alpha,\beta} h_{M,\delta}(r)$$

$$=\sum_{l\in\mathbb{Z}}\widehat{h_{M,\delta}}(l)\sum_{0<n_1,\dots,n_{\alpha+\beta}<X^{\delta_1}}\frac{a_f(n_1)\cdots a_f(n_\alpha)\overline{a_f(n_{\alpha+1})}\cdots\overline{a_f(n_{\alpha+\beta})}}{(n_1\cdots n_{\alpha+\beta})^{k-1}}$$

(4.9)

$$\times\frac{1}{\#\tilde{\Omega}_X}\sum_{r\in\tilde{\Omega}_X} e(lr+\overline{r}(n_1+\ldots+n_\alpha-n_{\alpha+1}-\ldots-n_{\alpha+\beta}))+O(X^{-\delta_1(k-3)/2}).$$

Now we apply the bound for sums of Kloosterman sums (2.4) which yields the following bound for the remaining off-diagonal contribution from (4.9);

$$\ll_{\alpha,\beta} X^{-\delta_\Gamma}\left(\sum_l |\widehat{h_{M,\delta}}(l)|\cdot|l|\right)$$

$$\times\sum_{0<n_1,\dots,n_{\alpha+\beta}<X^{\delta_1}}\frac{|a_f(n_1)\cdots a_f(n_{\alpha+\beta})|}{(n_1\cdots n_{\alpha+\beta})^{k-1}}\max(n_1,\dots,n_{\alpha+\beta})$$

$$\ll_{\alpha,\beta} X^{-\delta_\Gamma}\left(\sum_l |\widehat{h_{M,\delta}}(l)|\cdot|l|\right)\max(1,X^{-\delta_1(k-5)/2}),$$

using also that $\#\tilde{\Omega}_X\gg X^2$ by (2.3).
Now we apply (4.5) with $A=3+\varepsilon$ to ensure finiteness of the first sum above and then choose δ and δ_1 to balance the error terms. This yields a power savings, which we will not make explicit. This finishes this case as well. □

4.2. **Determining the limiting distribution.** In order to conclude the proofs of Theorem 1.1 and Theorem 1.8, we need to setup our problem in a probability theoretical framework.
Let $f\in\mathcal{S}_k(\Gamma)$ be as above and consider the following normalization of the periods of f;

$$\tilde{u}_{f,j}(r):=\frac{(2\pi/c(r))^{k-2}}{i\Gamma(k-1)}u_{f,j}(r)=\frac{(2\pi/c(r))^{k-2}}{i\Gamma(k-1)}\int_r^\infty f(z)z^j dz,\quad j=0,\dots,k-2,$$

where $r=\gamma\infty$ with $\gamma\in\Gamma$. According to whether Γ is a congruence subgroup or not, we consider for each $c\equiv 0\,(N)$ (respectively $X>0$) the renormalized periods;

$$\tilde{u}_f:=(\tilde{u}_{f,0},\dots,\tilde{u}_{f,k-2}),$$

as random variables defined on the outcome space Ω_c (respectively $\tilde{\Omega}_X$) endowed with the discrete σ-algebra and the uniform probability measure. Then one can easily check as in Remark 4.3 that Theorem 4.1 (respectively Theorem 4.2) implies that as $c\to\infty$ (respectively $X\to\infty$), the moments of the random variables $\tilde{u}_f$ converge to those of the random variable

$$F(Y,Z)=(F_0(Y,Z),\dots,F_{k-2}(Y,Z))^T,$$

where Y, Z are two independent random variables uniformly distributed with respect to the Lebesgue measure on $[0,1)$ and $F:[0,1)\times[0,1)\to\mathbb{C}^{k-1}$ is given (as in Theorem 1.1) by

$$F(y,z)=L(f,y,k-1)(1,z,\dots,z^{k-2})^T\in\mathbb{C}^{k-1}.$$

As an example let us consider the (complex) moment of $F(Y,Z)$ corresponding to the tuple $((1,1),(1,1),\dots,(1,1))$;

$$\mathbb{E}\left(F(Y,Z)^{((1,1),(1,1),\dots,(1,1))}\right):=\int_0^1\int_0^1|F_0(y,z)|^2\cdots|F_{k-2}(y,z)|^2dydz$$
$$=\left(\int_0^1|L(f,y,k-1)|^{2(k-1)}dy\right)\left(\int_0^1 z^{0+2+\ldots+2(k-2)}dz\right)$$
$$=\left(\sum_{\substack{n_1+\ldots+n_{k-1}\\=n_k+\ldots+n_{2(k-1)}}}\frac{a_f(n_1)\dots\overline{a_f(n_{2(k-1)})}}{(n_1\cdots n_{2(k-1)})^{k-1}}\right)\cdot\frac{1}{1+(k-2)(k-1)},$$

which we see match the corresponding moment in Theorem 4.1 and Theorem 4.2.

In order to conclude that the random variables associated with $\tilde{u}_f$ converge in distribution to $F(Y,Z)$ as $c\to\infty$ (respectively $X\to\infty$), we will combine three results from probability theory due to Fréchet–Shohat, Cramér–Wold and Carleman respectively. A similar but slightly simpler argument was carried out in [15, Section 5.4].

Proof of Theorem 1.1 and Theorem 1.8. Given a sequence of 1-dimensional real random variables $(X'_n)_{n\geq1}$ such that all moments exist and converge as $n\to\infty$ to the moments of some other random variable Y' then it follows from the Fréchet–Shohat Theorem [20, p. 17] that if Y' is uniquely determined by its moments then the random variables $(X'_n)_{n\geq1}$ converge *in distribution* to Y'.

Our random variables are however multidimensional so we have to combine the Fréchet–Shohat Theorem with a result of Cramér and Wold [20, p. 18], which says that if $(X'_n)_{n\geq1}$ is a sequence of $(d+1)$-dimensional real random variables;

$$X'_n=(X'_{n,0},\dots,X'_{n,d}),$$

and $Y'=(Y'_0,\dots,Y'_d)$ is a $(d+1)$-dimensional random variable such that

$$t_0X'_{n,0}+\ldots+t_dX'_{n,d}$$

converge in distribution as $n\to\infty$ to

$$t_0Y'_0+\ldots+t_dY'_d$$

for any $(d+1)$-tuple $(t_0,\dots,t_d)\in\mathbb{R}^{d+1}$, *then* X'_n converges in distribution to Y' as $n\to\infty$.

Thus by combining Fréchet–Shohat and Cramér–Wold with our calculation of the moments in Theorem 4.1 (respectively Theorem 4.2), it is enough to show that for any (say non-trivial) linear combination, the following random variable;

$$
\begin{aligned}
(4.10)\qquad t_0 \operatorname{Re} L(f,Y,k-1) + t_1 \operatorname{Re} L(f,Y,k-1)Z+ \\
\ldots + t_{k-2} \operatorname{Re} L(f,Y,k-1)Z^{k-2} \\
+ t_{k-1} \operatorname{Im} L(f,Y,k-1) + t_k \operatorname{Im} L(f,Y,k-1)Z+ \\
\ldots + t_{2k-3} \operatorname{Im} L(f,Y,k-1)Z^{k-2}
\end{aligned}
$$

is uniquely determined by its moments. By a condition due to Carleman (see (4.11) below), this boils down to showing that the moments are sufficiently bounded from above, which is clear in our case since Z is bounded by 1 and

$$|L(f,Y,k-1)| \leq \sum_{n\geq 1} \frac{|a_f(n)|}{n^{k-1}} < \infty,$$

both with probability one. To sum up and be precise; if we denote by α_{2m} the $2m$'th moment of (4.10), then we have

$$\sum_{m\geq 1} \alpha_{2m}^{-1/2m} \geq \sum_{m\geq 1} \left(\left(c(t_0,\ldots,t_{2k-3}) \sum_{n\geq 1} \frac{|a_f(n)|}{n^{k-1}} \right)^{2m} \right)^{-1/2m} = \infty, \tag{4.11}$$

where $c(t_0,\ldots,t_{2k-3})$ is a certain constant depending on $t_0,\ldots,t_{2k-3}$. Thus it follows from the Carleman condition [20, p. 46] that the random variable (4.10) is uniquely determined by its moments. Thus we conclude the proof of Theorem 1.1 and Theorem 1.8 using the results of Fréchet–Shohat and Cramér–Wold mentioned above. □

References

1. Sandro Bettin and Sary Drappeau, *Limit laws for rational continued fractions and value distribution of quantum modular forms*, arXiv e-prints (2019), arXiv:1903.00457.
2. Pierre Deligne, *Formes modulaires et représentations l-adiques*, Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363, Lecture Notes in Math., vol. 175, Springer, Berlin, 1971, pp. Exp. No. 355, 139–172. MR 3077124
3. ______, *La conjecture de Weil. II*, Inst. Hautes Études Sci. Publ. Math. (1980), no. 52, 137–252. MR 601520
4. Nikolaos Diamantis and Larry Rolen, *Period polynomials, derivatives of L-functions, and zeros of polynomials*, Res. Math. Sci. **5** (2018), no. 1, Paper No. 9, 15. MR 3760367
5. Martin Eichler, *Eine Verallgemeinerung der Abelschen Integrale*, Sammelband zu Ehren des 250. Geburtstages Leonhard Eulers, Akademie-Verlag, Berlin, 1959, pp. 112–115. MR 0122993
6. Dorian Goldfeld and Peter Sarnak, *Sums of Kloosterman sums*, Invent. Math. **71** (1983), no. 2, 243–250. MR 689644
7. Atsushi Ichino, *Trilinear forms and the central values of triple product L-functions*, Duke Math. J. **145** (2008), no. 2, 281–307. MR 2449948
8. Henryk Iwaniec, *Spectral methods of automorphic forms*, second ed., Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002. MR 1942691
9. Seokho Jin, Wenjun Ma, Ken Ono, and Kannan Soundararajan, *Riemann hypothesis for period polynomials of modular forms*, Proc. Natl. Acad. Sci. USA **113** (2016), no. 10, 2603–2608. MR 3482847

10. Nicholas M. Katz, *Sommes exponentielles*, Astérisque, vol. 79, Société Mathématique de France, Paris, 1980, Course taught at the University of Paris, Orsay, Fall 1979, With a preface by Luc Illusie, Notes written by Gérard Laumon, With an English summary. MR 617009
11. Victor A. Kolyvagin, *Finiteness of $E(\mathbf{Q})$ and $CH(E,\mathbf{Q})$ for a subclass of Weil curves*, Izv. Akad. Nauk SSSR Ser. Mat. **52** (1988), no. 3, 522–540, 670–671. MR 954295
12. Emmanuel Kowalski, Philippe Michel, and Jeffrey VanderKam, *Rankin-Selberg L-functions in the level aspect*, Duke Math. J. **114** (2002), no. 1, 123–191. MR 1915038
13. Ju I Manin, *Periods of parabolic forms and p-adic Hecke Series*, Mathematics of the USSR-Sbornik **21** (1973), no. 3, 371–393.
14. Paul D. Nelson and Akshay Venkatesh, *The orbit method and analysis of automorphic forms*, arXiv e-prints (2018), arXiv:1805.07750.
15. Asbjorn Nordentoft, *Central values of additive twists of modular L-functions*, arXiv e-prints (2018), arXiv:1812.08378.
16. Vicenctiu Pasol and Alexandru A. Popa, *Modular forms and period polynomials*, Proc. Lond. Math. Soc. (3) **107** (2013), no. 4, 713–743. MR 3108829
17. Yiannis N. Petridis and Morten S. Risager, *Modular symbols have a normal distribution*, Geom. Funct. Anal. **14** (2004), no. 5, 1013–1043. MR 2105951
18. ———, *Arithmetic statistics of modular symbols*, Invent. Math. **212** (2018), no. 3, 997–1053. MR 3802302
19. Yiannis Sakellaridis and Akshay Venkatesh, *Periods and harmonic analysis on spherical varieties*, Astérisque (2017), no. 396, viii+360. MR 3764130
20. Robert J. Serfling, *Approximation theorems of mathematical statistics*, John Wiley & Sons, Inc., New York, 1980, Wiley Series in Probability and Mathematical Statistics. MR 595165
21. Goro Shimura, *Introduction to the arithmetic theory of automorphic functions*, Publications of the Mathematical Society of Japan, vol. 11, Princeton University Press, Princeton, NJ, 1994, Reprint of the 1971 original, Kanô Memorial Lectures, 1. MR 1291394
22. Gabor Wiese, *Computational arithmetic of modular forms*, Notes from the International Autumn School on Computational Number Theory, Tutor. Sch. Workshops Math. Sci., Birkhäuser/Springer, Cham, 2019, pp. 63–170. MR 3932019

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, Copenhagen 2100, Denmark
E-mail address: acnordentoft@outlook.com

PAPER C
A NOTE ON ADDITIVE TWISTS, RECIPROCITY LAWS AND QUANTUM MODULAR FORMS

A NOTE ON ADDITIVE TWISTS, RECIPROCITY LAWS AND QUANTUM MODULAR FORMS

ASBJØRN CHRISTIAN NORDENTOFT

Abstract. We prove that the central values of additive twists of a cuspidal L-function define a quantum modular form in the sense of Zagier, generalizing recent results of Bettin and Drappeau. From this we deduce a reciprocity law for the twisted first moment of multiplicative twists of cuspidal L-functions, similar to reciprocity laws discovered by Conrey for the twisted second moment of Dirichlet L-functions. Furthermore we give an interpretation of quantum modularity at infinity for additive twists of L-functions of weight 2 cusp forms in terms of the corresponding functional equations.

1. Introduction

In an unpublished paper [5, Theorem 10], Conrey discovered a certain reciprocity law satisfied by the twisted second moment of Dirichlet L-functions. The reciprocity law relates the following two twisted moments;

$$\sum_{\chi \bmod q} |L(\chi,1/2)|^2\chi(l) \rightsquigarrow \sum_{\chi \bmod l} |L(\chi,1/2)|^2\chi(-q), \tag{1.1}$$

for primes $q \neq l$. Conrey's results were then generalized and refined by Young [12] and Bettin [1]. In this paper we prove a reciprocity law for twists of GL_2-cusp forms, which in the simplest case relates the following two twisted first moments;

$$\sum_{\substack{\chi \bmod q,\\ \chi \text{ primitive}}} \tau(\overline{\chi})L(f,\chi,k/2)\chi(l) \rightsquigarrow \sum_{\substack{\chi \bmod l,\\ \chi \text{ primitive}}} \tau(\overline{\chi})L(f,\chi,k/2)\chi(-q), \tag{1.2}$$

where $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ is a cusp form of weight k and level 1, $\tau(\overline{\chi})$ is a Gauss-sum and $q, l > 2$ are primes. The exact moments involved for general level N and general q, l are more involved (see Theorem 2.1 below).

Our proof uses an interpretation of the twisted moments (1.2) in terms of additive twists of L-functions. The *additive twists* associated to the L-function of a weight k cusp form $f \in \mathcal{S}_k(\Gamma_0(N))$ are defined as

$$L(f,a/c,s) := \sum_{n\geq 1} \frac{a_f(n)e(na/c)}{n^s},$$

for $\mathrm{Re}\, s > (k+1)/2$, where $a_f(n)$ denote the Fourier coefficients of f, $a/c \in \mathbb{Q}$ and $e(x) := e^{2\pi i x}$. The Dirichlet series above admit analytic continuation to the entire

Date: March, 2020.

2010 *Mathematics Subject Classification.* 11F67(primary).

complex plane, which satisfy functional equations (see Section 3.1 below for details). The reciprocity law will now follow from the fact that the central value at $s = k/2$ of the additive twists define a *quantum modular form* of weight zero in the sense of Zagier [14] (see Theorem 4.4 below for the precise statement). The quantum modularity result is of independent interest.

The inspiration for this approach comes from Bettin's work [1], which gives an interpretation of the twisted second moment (1.1) in terms of the central value of the *Estermann zeta function* defined for $\operatorname{Re} s > 1$ by

$$D(a, c; s) := \sum_{n \geq 1} \frac{d(n)e(na/c)}{n^s},$$

where $d(n)$ denotes the number of divisors of n and $a, c \in \mathbb{Z}$. One can think of this as an additive twist of $\zeta(s)^2$, which in turn is the L-function associated to the GL_2-object $\frac{\partial}{\partial s} E(z, s)_{s=1/2}$. Bettin's results [1, Theorem 1] can be interpreted as showing that $D(a/c; 1/2) := D(a, c; 1/2)$ defines a quantum modular form as a function of $a/c \in \mathbb{Q}$. Recently Bettin and Drappeau [2, Lemma 8.3] showed quantum modularity in the case of level 1 cusp forms, which they ingeniously combined with dynamical methods to determine the limiting distribution of the central values of additive twists of L-functions of level 1 cusp forms [2, Corollary 1.5]. The results of this paper extend the quantum modularity proved by Bettin and Drappeau to general discrete and cofinite subgroups of $\mathrm{SL}_2(\mathbb{R})$ with a cusp at infinity (with an appropriate definition of quantum modularity). Quantum modularity for general levels will be needed if one wants to extend the methods of Bettin and Drappeau to general level (see the remark on page 8 of [2]).

Finally we discuss quantum modularity at ∞ and show that for weight 2 cusp forms, this is in a precise sense equivalent to the functional equation at the central point for the additive twists of the associated L-function.

Remark 1.1. A different proof of the Gaussian behavior of central values of additive twists [2, Corollary 1.5] was obtained independently by the author [11] using the theory of Eisenstein series twisted by modular symbols. The methods in [11] furthermore apply to general Fuchsian groups of the first kind with a cusp at ∞, and thus in particular to general level.

Acknowledgements

I would like to thank Dorian Goldfeld and Columbia University for their hospitality and the reviewer for useful comments.

2. Statement of results

In order to state our results, let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a primitive newform of even weight k and level N. Let ω_f denote the eigenvalue under the Fricke involution W (see (4.3) below) and let

$$f(z) = \sum_{n \geq 1} a_f(n) q^n, \quad q = e^{2\pi i z},$$

be the Fourier expansion. Then given a Dirichlet character χ, we consider the following twisted L-function;

$$L(f,\chi,s) := \sum_{n\geq 1} \frac{a_f(n)\chi(n)}{n^s}, \quad \mathrm{Re}\, s > (k+1)/2, \tag{2.1}$$

which admits analytic continuation satisfying a functional equation relating $s \leftrightarrow k-s$. Our result is the following reciprocity law for a certain appropriately weighted twisted first moment of these L-functions at the central point $s = k/2$.

Theorem 2.1. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a primitive newform of even weight k and level N. Then we have the following reciprocity relation for any pair of integers $0 < l < q$ with $(q, Nl) = 1$;*

$$\begin{aligned} \frac{1}{\varphi(q)} \sum_{\chi \bmod q} & \nu(f,\chi^*,q/c(\chi))L(f,\chi^*,k/2)\chi(l) \\ & - \frac{\omega_f}{\varphi(lN)} \sum_{\chi \bmod lN} \nu(f,\chi^*,lN/c(\chi))L(f,\chi^*,k/2)\chi(-q) \\ & = L(f,k/2) + O_f(l/q), \end{aligned} \tag{2.2}$$

where χ^* mod $c(\chi)$ denotes the primitive character inducing χ, $L(f,\chi,s)$ is as in (2.1) and the arithmetic weights ν are given by

$$\nu(f,\chi,n) := \tau(\overline{\chi}) \sum_{\substack{n_1 n_2 n_3 = n, \\ (n_1,N)=1}} \chi(n_1)\mu(n_1)\overline{\chi}(n_2)\mu(n_2)a_f(n_3)n_3^{1-k/2}.$$

Remark 2.2. To see where the arithmetic weights $\nu(f,\chi,n)$ come from, consult Lemma 3.1 below. Note that for primitive characters χ, the weight is exactly given by the Gauss-sum $\tau(\overline{\chi})$.

In the simplest case where q, l are both prime and the level is 1, we get the following precise version of (1.2).

Corollary 2.3. *Let $f \in \mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ be a primitive newform of even weight k and level 1. Then we have the following reciprocity relation for any primes $q > l > 2$;*

$$\begin{aligned} \frac{1}{q-2} \sum_{\substack{\chi \bmod q \\ \chi \text{ primitive}}} \tau(\overline{\chi})L(f,\chi,k/2)\chi(l) - \frac{1}{l-2} \sum_{\substack{\chi \bmod l \\ \chi \text{ primitive}}} \tau(\overline{\chi})L(f,\chi,k/2)\chi(-q) \\ = L(f,k/2) + O_f(l/q + 1/\sqrt{l}), \end{aligned}$$

where the sums are taken over all primitive characters modulo q and modulo l, respectively.

3. Twisted first moments and additive twists

In this section we will recall some standard properties of additive twists of cuspidal L-functions and furthermore for congruence subgroups show a connection to the twisted first moments in (1.2) using (an extension of) a formula due to Birch and Stevens.

3.1. **Additive twists.** We refer to [11, Section 3.3] and [7, Section A.2] for a more detailed account on additive twists.
Let Γ be a discrete and co-finite subgroup of $\mathrm{SL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 (see [6, Chapter 2] for definitions) and consider a cusp form $f \in \mathcal{S}_k(\Gamma)$ of even weight k with Fourier expansion (at ∞) given by

$$f(z) = \sum_{n\geq 1} a_f(n)q^n.$$

Then we define the *additive twist by* $r \in \mathbb{R}$ of the L-function of f as

$$L(f,r,s) := \sum_{n\geq 1} \frac{a_f(n)e(nr)}{n^s}, \tag{3.1}$$

which converges absolutely (and locally uniformly) for $\mathrm{Re}\, s > (k+1)/2$ by Hecke's bound;

$$\sum_{n\leq X} |a_f(n)|^2 \ll X^k. \tag{3.2}$$

Thus the Dirichlet series (3.1) defines a continuous function in r when s is fixed with $\mathrm{Re}\, s > (k+1)/2$.
Furthermore if r corresponds to a cusp of Γ (i.e. r is fixed by a parabolic subgroup of Γ), then $L(f,r,s)$ admits analytic continuation to the entire complex plane as a function of s, given by the following integral representation;

$$L(f,r,s) = \frac{(2\pi)^s}{\Gamma(s)} \int_0^\infty f(r+iy)y^{s-1}dy, \tag{3.3}$$

well-defined for all $s \in \mathbb{C}$.
Finally if r is in the Γ-orbit of ∞, say $r = \gamma\infty$ with $\gamma \in \Gamma$, then we have the following functional equation;

$$\begin{aligned}\Lambda(f,\gamma\infty,s) &:= \left(\frac{c}{2\pi}\right)^s \Gamma(s)L(f,\gamma\infty,s)\\ &= (-1)^{k/2}\Lambda(f,\gamma^{-1}\infty,k-s),\end{aligned} \tag{3.4}$$

where c is the lower-left entry of γ.

3.2. **The Birch–Stevens formula.** In the special case where $\Gamma = \Gamma_0(N)$, the set of cusps corresponds to $\mathbb{Q} \cup \{\infty\}$ and the classical Birch–Stevens formula [10, Theorem 2.3] expresses the central value $L(f,\chi,k/2)$ for a primitive Dirichlet character χ mod q in terms of additive twists;

$$\tau(\overline{\chi})L(f,\chi,k/2) = \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} L(f,a/q,k/2)\overline{\chi(a)}, \tag{3.5}$$

where $\tau(\overline{\chi})$ denotes the Gauss-sum of $\overline{\chi}$.
For $k=2$ the central values $L(f,a/q,1)$ are known as *modular symbols* and the above equality has been used for computations of the central values of L-functions of base-changes of elliptic curves over $\mathbb{Q}$.
If we furthermore assume that f is a primitive newform (in particular an eigenform for all Hecke operators), then we have the following generalization to non-primitive characters χ.

Lemma 3.1. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a primitive newform of even weight k and level N and χ mod q a Dirichlet character. Then we have*

$$\nu(f,\chi^*,q/c(\chi))L(f,\chi^*,k/2) = \sum_{a\in(\mathbb{Z}/q\mathbb{Z})^\times} L(f,a/q,k/2)\overline{\chi(a)}, \tag{3.6}$$

where χ^* mod $c(\chi)$ denotes the primitive character inducing χ and the weight is given by

$$\nu(f,\chi,n) := \tau(\overline{\chi}) \sum_{\substack{n_1n_2n_3=n,\\ (n_1,N)=1}} \chi(n_1)\mu(n_1)\overline{\chi}(n_2)\mu(n_2)a_f(n_3)n_3^{1-k/2}.$$

For a proof see [11, Proposition 6.1].

Remark 3.2. This formula was also the essential ingredient for Bruggeman and Diamantis in [4, Theorem 1.1], where they give an explicit formula for the constant Fourier coefficient of Eisenstein series twisted by modular symbols.

From the above formula we conclude, using orthogonality of characters on the finite group $(\mathbb{Z}/q\mathbb{Z})^\times$, the following identity.

Corollary 3.3. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a primitive newform of even weight k and level N. Then we have*

$$L(f,a/q,k/2) = \frac{1}{\varphi(q)} \sum_{\chi \bmod q} \nu(f,\chi^*,q/c(\chi))L(f,\chi^*,k/2)\chi(a), \tag{3.7}$$

with χ^* mod $c(\chi)$ and ν as above.

Thus for $l = o(q)$, it follows from (2.2) that

$$L(f,l/q,k/2) - L(f,-q/(Nl),k/2) = L(f,k/2) + o(1),$$

as $q \to \infty$. This is in sharp contrast with the average behavior since we know from the distribution result [11, Theorem 1.1] that the numbers $L(f,l/q,k/2)$ with $(l,q) = 1$ and $q \asymp Q$ are of magnitude $\sqrt{\log Q}$ with probability one.

4. Quantum modularity of additive twists

The notion of *quantum modular forms* was introduced by Zagier in [14] with one of the first examples appearing in earlier work with Lawrence [8] on certain symmetries of quantum invariants of 3-knots (hence the name). In Zagier's original definition, quantum modular forms are maps $\mathbb{P}^1(\mathbb{Q})\backslash S \to \mathbb{C}$ with S a finite set which satisfy a variation of the modular transformation rule with respect to congruence subgroups $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})$ acting on $\mathbb{P}^1(\mathbb{Q})$. One should think of the equivalence classes of $\mathbb{P}^1(\mathbb{Q})$ under the action of $\Gamma_0(N)$ as the boundary components of the symmetric space $\Gamma_0(N)\backslash\mathbb{H}$.
In this paper we study quantum modularity for general co-finite, discrete subgroups $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 (see [6, Section 2] for definitions of these notions). Although quantum modularity with respect to general such Γ has not been explicitly defined before in the literature, the definition is implicit in the introduction of [14]. To make this precise, denote by $s(\Gamma) \subset \mathbb{P}^1(\mathbb{R}) = \mathbb{R} \cup \{\infty\}$ the set of cusps of Γ (i.e. fix-points of parabolic subgroups of Γ). In particular if $\Gamma = \Gamma_0(N)$, we have

$s(\Gamma) = \mathbb{Q} \cup \{\infty\} = \mathbb{P}^1(\mathbb{Q})$.
Then we have the following definition.

Definition 4.1. *Let Γ be a discrete, co-finite subgroup of $\mathrm{SL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 and consider a map $f : s(\Gamma)\backslash S \to \mathbb{C}$ with S a finite set. Then f is a* **quantum modular form** *of weight k for Γ if the following holds; for any fixed $\gamma \in \Gamma$ the function $g_\gamma(r) : s(\Gamma)\backslash(S \cup \gamma^{-1}S) \to \mathbb{C}$ defined by*

$$g_\gamma(r) := f(\gamma r) - j(\gamma, r)^k f(r), \tag{4.1}$$

extends to a continuous function $\mathbb{P}^1(\mathbb{R})\backslash(S \cup \gamma^{-1}S) \to \mathbb{C}$.

Remark 4.2. Here *continuous* can be replaced by different notions of regularity ($\mathcal{C}^1$, smooth, real-analytic, ...), and indeed Zagier in [14] requires the discrepancy (4.1) to extend to a real-analytic function.

Remark 4.3. Note that with our definition of quantum modularity, all restrictions of continuous maps $\mathbb{P}^1(\mathbb{R})\backslash S \to \mathbb{C}$ (with S a finite set) trivially satisfy the condition of being a quantum modular form.

4.1. **Proof of quantum modularity.** In this section we present a proof of the quantum modularity for the central values of additive twists of cuspidal L-functions. The proof uses the integral representation of the additive twist (3.3) and is similar in spirit to the treatment of Eichler integrals associated to half-integral cusp forms by Bringmann and Rolen in [3]. One can also consider the Eichler integrals of an integral weight k cusp form f, which corresponds to the special value $L(f, r, k-1)$ of the additive twists. For $k > 2$ this is however a trivial example with our definition of quantum modularity since this is the restriction of a continuous function[1], because of the absolute convergence of the Dirichlet series (3.1) at the special value $s = k - 1$ (as was also noted in [2, Section 1.4.1]).
It follows from [11, Theorem 1.4] that $L(f, r, k/2)$ considered as a function of the twisting parameter $r \in s(\Gamma)$ is unbounded on any interval and thus it is not the restriction of a continuous function $\mathbb{P}^1(\mathbb{R})\backslash S \to \mathbb{C}$ with S a finite set. Our result is that this does however define a quantum modular form.

Theorem 4.4 (Quantum modularity). *Let Γ be a discrete, co-finite subgroup of $\mathrm{SL}_2(\mathbb{R})$ with a cusp at ∞ of width 1 and let $f \in \mathcal{S}_k(\Gamma)$ be a cusp form of even weight k. Then the map $s(\Gamma)\backslash\{\infty\} \to \mathbb{C}$ defined by*

$$r \mapsto L(f, r, k/2)$$

[1]Eichler integrals of integral weight cusp forms were some of the earliest examples of quantum modular forms considered by Zagier [13, Section 11] (see also Lee [9]). In these works the discrepancy (4.1) is required to extend to a polynomial instead of just a continuous function. Now one gets a non-trivial result since the Eichler integrals are certainly not restrictions of polynomials (not even restriction of smooth functions).

is a quantum modular form of weight zero for Γ. *More precisely for* $\gamma \in \Gamma$ *and* $r \in s(\Gamma)\backslash\{\infty\}$ *with* $\gamma r \neq \infty$ *and* $\gamma\infty \neq \infty$, *we have*

$$\begin{aligned} L(f,\gamma r,k/2)-&L(f,r,k/2)\\ &= L(f,\gamma\infty,k/2) + \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{(c^{-1}j(\gamma,r))^j} \frac{(-2\pi i)^{-j}\Gamma(k/2+j)}{\Gamma(k/2)} L(f,r,k/2+j)\\ &\quad + \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{(cj(\gamma,r))^j} \frac{(-2\pi i)^{j}\Gamma(k/2-j)}{\Gamma(k/2)} L(f,\gamma\infty,k/2-j), \end{aligned} \tag{4.2}$$

where c *is the lower-left entry of* γ.

Remark 4.5. Theorem 4.4 in the special case $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ is exactly [2, Lemma 8.3].

Proof. We have to show continuity of the discrepancy (4.1) for all $\gamma \in \Gamma$. First of all if $\gamma\infty = \infty$ then it is easy to see that

$$L(f,\gamma r,k/2) = L(f,r,k/2),$$

for all $r \in s(\Gamma)\backslash\{\infty\}$ since f is 1-periodic. Thus quantum modularity for such γ is clear.

Recall that by (3.2), the additive twists $L(f,x,k/2+j)$ for $j \geq 1$ define continuous functions in $x \in \mathbb{R}$. Thus for γ fixed the two sums on the right-hand side of (4.2) both extend to continuous functions $\mathbb{R}\backslash\{\gamma^{-1}\infty\} \to \mathbb{C}$. Thus it suffices to prove the identity (4.2). To do this, we begin with the following integral representation;

$$\begin{aligned} L(f,\gamma r,k/2) &= \frac{(2\pi)^{k/2}}{\Gamma(k/2)} \int_0^\infty f(\gamma r + iy) y^{k/2-1} dy\\ &= \frac{(-2\pi i)^{k/2}}{\Gamma(k/2)} \int_{\gamma r}^{i\infty} f(z)(z-\gamma r)^{k/2-1} dz, \end{aligned}$$

where the integral is taken along the vertical line from γr to $i\infty$. Now the integrand is holomorphic and we can apply Cauchy's theorem to write

$$L(f,\gamma r),k/2) = \frac{(-2\pi i)^{k/2}}{\Gamma(k/2)} \left(\int_{\gamma r}^{\gamma\infty} f(z)(z-\gamma r)^{k/2-1} dz + \int_{\gamma\infty}^{i\infty} f(z)(z-\gamma r)^{k/2-1} dz \right).$$

We will now treat the two integrals separately. In the first integral we do the change of variable $z \mapsto \gamma z$ and use the following simple formulas;

$$\gamma z - \gamma r = \frac{z-r}{j(\gamma,z)j(\gamma,r)}, \quad j(\gamma,z) = j(\gamma,r) + c(z-r),$$

to arrive at

$$\begin{aligned}
&\int_{\gamma r}^{\gamma\infty} f(z)(z-\gamma r)^{k/2-1}dz\\
&=\int_{r}^{i\infty} f(\gamma z)(\gamma z-\gamma r)^{k/2-1}\frac{dz}{j(\gamma,z)^2}\\
&=\left(\frac{c}{j(\gamma,r)}\right)^{k/2-1}\int_r^{i\infty} f(z)\left((z-r)(z-r+\frac{j(\gamma,r)}{c})\right)^{k/2-1}dz\\
&=\sum_{j=0}^{k/2-1}\binom{k/2-1}{j}\left(\frac{c}{j(\gamma,r)}\right)^j\int_r^{i\infty} f(z)(z-r)^{k/2+j-1}dz\\
&=\sum_{j=0}^{k/2-1}\binom{k/2-1}{j}\left(\frac{c}{j(\gamma,r)}\right)^j\frac{\Gamma(k/2+j)}{(-2\pi i)^{k/2+j}}L(f,r,k/2+j).
\end{aligned}$$

A similar treatment of the other integral gives

$$\begin{aligned}
&\int_{\gamma\infty}^{i\infty} f(z)(z-\gamma\infty+(\gamma\infty-\gamma r))^{k/2-1}dz\\
&=\sum_{j=0}^{k/2-1}\binom{k/2-1}{j}\left(\frac{1}{cj(\gamma,r)}\right)^j\int_{\gamma\infty}^{i\infty} f(z)(z-\gamma\infty)^{k/2-j-1}dz\\
&=\sum_{j=0}^{k/2-1}\binom{k/2-1}{j}\left(\frac{1}{cj(\gamma,r)}\right)^j\frac{\Gamma(k/2-j)}{(-2\pi i)^{k/2-j}}L(f,\gamma\infty,k/2-j),
\end{aligned}$$

which finishes the proof. □

Remark 4.6. For $k=2$ we observe that the right-hand side of (4.2) is just a constant. From this it follows immediately that the central value of the additive twists (i.e. modular symbols) define a *strong quantum modular form* in the sense of Zagier [14].

Remark 4.7. The proof and statement of Theorem 4.4 is very similar to [11, Lemma 5.2], which was used to reduce the study of Eisenstein series twisted by additive twists to the study of certain "completions" in the sense of [4, page 6]. This was a key step for the author in [11] in order to determine the distribution of central values of additive twists and thus there seem to be some similarities with the methods in [2], which would be interesting to understand better.

If $\Gamma=\Gamma_0(N)$ is a congruence group and we assume that $f\in\mathcal{S}_k(\Gamma_0(N))$ is a primitive newform, we get a similar result for the Fricke involution defined as;

$$Wf(z):=N^{-k/2}z^{-k}f(H_Nz), \tag{4.3}$$

where

$$H_N=\begin{pmatrix}0&-1\\N&0\end{pmatrix}.$$

The proof is essentially the same and we will omit it.

Theorem 4.8. *Let $f \in \mathcal{S}_k(\Gamma_0(N))$ be a primitive newform of even weight k and level N. Then we have for all $r \in \mathbb{Q}\backslash\{0\}$ that*

$$
\begin{aligned}
&L(f,-1/(Nr),k/2) - \omega_f L(f,r,k/2) \\
&= L(f,k/2) + \omega_f \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{r^j} \frac{(-2\pi i)^{-j}\Gamma(k/2+j)}{\Gamma(k/2)} L(f,r,k/2+j) \\
&\quad + \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{(Nr)^j} \frac{(-2\pi i)^{j}\Gamma(k/2-j)}{\Gamma(k/2)} L(f,k/2-j),
\end{aligned}
\tag{4.4}
$$

where $\omega_f = \pm 1$ is the eigenvalue of f under the Fricke involution W defined in (4.3).

4.2. **Quantum modularity at ∞ and the functional equation.** A natural question to ask is whether we can extend our quantum modular forms to ∞. This is equivalent to whether we can assign a value at ∞ such that for all $\gamma \in \Gamma$, the right-hand side of (4.2) converges as respectively $r \to \infty$ and $r \to \gamma^{-1}\infty$ to the left-hand side of (4.2) with respectively $r = \infty$ and $r = \gamma^{-1}\infty$. In the special case when $k = 2$ this can be done by putting $L(f,\infty,1) = 0$. It turns out that quantum modularity at ∞ amounts to the functional equation for $L(f,\gamma\infty,s)$ at the central point.

Theorem 4.9. *Let Γ be a discrete, co-finite subgroup of $SL_2(\mathbb{R})$ with a cusp at ∞ of width 1 and let $f \in \mathcal{S}_2(\Gamma)$. Then the map $s(\Gamma) \to \mathbb{C}$ defined by*

$$r \mapsto L(f,r,1)$$

where we put $L(f,\infty,1) = 0$, is a quantum modular form of weight zero for Γ.

Proof. First of all if $\gamma\infty = \infty$, it is clear. Furthermore when $k = 2$, (4.2) reduces to

$$L(f,\gamma r,1) - L(f,r,1) = L(f,\gamma\infty,1), \tag{4.5}$$

and since the right-hand side is constant we can ignore convergence completely. Using Theorem 4.4 we only need to check that (4.5) still holds at $r = \infty$ and $r = \gamma^{-1}\infty$. The first case is immediate and the second case reduces to

$$L(f,\gamma^{-1}\infty,1) \stackrel{?}{=} -L(f,\gamma\infty,1),$$

which is exactly the functional equation (3.4) at the central point $s = k/2 = 1$. This finishes the proof. □

Remark 4.10. This seems to be a very special phenomena for $k = 2$ and we have numerical data suggesting that it is not true for the Ramanujan Δ-function. The author was however not able to disprove it for higher weights. Notice that for $k > 2$ one gets poles at $r = \gamma^{-1}\infty$ from both sums on the right-hand side of (4.4), but we are not able to rule out the possibility that these poles cancel out (corresponding to the question of whether the special value $L(f,x,k/2+1)$ is differentiable at points $x \in \Gamma\infty$). It would be interesting to investigate quantum modularity at ∞ for other quantum modular forms in the literature.

5. From quantum modularity to reciprocity laws

In this final section, we will prove the reciprocity laws Theorem 2.1 and Corollary 2.3 using the quantum modularity for the Fricke involution of additive twists proved above.

5.1. **Proof of Theorem 2.1 and Corollary 2.3.** By combining Corollary 3.3 and Theorem 4.8 with $r = -q/(lN)$, we get an explicit formula for the left-hand side of (2.2). The expression for the error-term on the right-hand side of (2.2) follows by the following estimate of the right-hand side of (4.4);

$$\begin{aligned} & L(f,k/2) + \omega_f \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{r^j} \frac{(-2\pi i)^{-j}\Gamma(k/2+j)}{\Gamma(k/2)} L(f,r,k/2+j) \\ & + \sum_{j=1}^{k/2-1} \frac{\binom{k/2-1}{j}}{(N^2 r)^j} \frac{(-2\pi i)^{j}\Gamma(k/2-j)}{\Gamma(k/2)} L(f,k/2-j) \\ & = L(f,k/2) + O_f\left(\sum_{j=1}^{k/2-1} r^{-j} \right) \\ & = L(f,k/2) + O_f(r^{-1}), \end{aligned}$$

where we again used the following uniform bound;

$$|L(f,r,k/2+j)| \leq \sum_{n\geq 1} \frac{|a_f(n)|}{n^{k/2+1}} < \infty,$$

for $j \geq 1$. This proves Theorem 2.1.
Furthermore if l and q are both primes then the only non-primitive character modulo l and q are the principal characters. Now using Deligne's bound $|a_f(n)| \leq d(n) n^{(k-1)/2}$ on the Fourier coefficients of f, we derive

$$\nu(f, \chi_0^*, q/c(\chi_0)) = \nu(f,1,q) = a_f(q) q^{1-k/2} - 1 \ll q^{1/2},$$

and similarly for l. Using that

$$\frac{1}{\varphi(q)} = \frac{1}{q-2} + O(q^{-2}),$$

for $q > 2$ prime, we conclude the proof of Corollary 2.3.

References

[1] Sandro Bettin, *On the reciprocity law for the twisted second moment of Dirichlet L-functions*, Trans. Amer. Math. Soc. **368** (2016), no. 10, 6887–6914. MR 3471080

[2] Sandro Bettin and Sary Drappeau, *Limit laws for rational continued fractions and value distribution of quantum modular forms*, arXiv e-prints (2019), arXiv:1903.00457.

[3] Kathrin Bringmann and Larry Rolen, *Half-integral weight Eichler integrals and quantum modular forms*, J. Number Theory **161** (2016), 240–254. MR 3435727

[4] Roelof Bruggeman and Nikolaos Diamantis, *Fourier coefficients of Eisenstein series formed with modular symbols and their spectral decomposition*, J. Number Theory **167** (2016), 317–335. MR 3504050

[5] J. Brian Conrey, *The mean-square of Dirichlet L-functions*, arXiv e-prints (2007), arXiv:0708.2699.

[6] Henryk Iwaniec, *Spectral methods of automorphic forms*, second ed., Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002. MR 1942691

[7] E. Kowalski, P. Michel, and J. VanderKam, *Rankin-Selberg L-functions in the level aspect*, Duke Math. J. **114** (2002), no. 1, 123–191. MR 1915038

[8] Ruth Lawrence and Don Zagier, *Modular forms and quantum invariants of* 3*-manifolds*, vol. 3, 1999, Sir Michael Atiyah: a great mathematician of the twentieth century, pp. 93–107. MR 1701924

[9] Seewoo Lee, *Quantum modular forms and Hecke operators*, Res. Number Theory **4** (2018), no. 2, Art. 18, 13. MR 3797180

[10] Barry Mazur and Karl Rubin, *Arithmetic conjectures suggested by the statistical behavior of modular symbols*, 2019.

[11] Asbjorn Nordentoft, *Central values of additive twists of modular L-functions*, arXiv e-prints (2018), arXiv:1812.08378.

[12] Matthew P. Young, *The reciprocity law for the twisted second moment of Dirichlet L-functions*, Forum Math. **23** (2011), no. 6, 1323–1337. MR 2855052

[13] D. Zagier, *From quadratic functions to modular functions*, Number theory in progress, Vol. 2 (Zakopane-Kościelisko, 1997), de Gruyter, Berlin, 1999, pp. 1147–1178. MR 1689565

[14] Don Zagier, *Quantum modular forms*, Quanta of maths, Clay Math. Proc., vol. 11, Amer. Math. Soc., Providence, RI, 2010, pp. 659–675. MR 2757599

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, Copenhagen 2100, Denmark

E-mail address: acnordentoft@outlook.com

PAPER D
RESIDUAL EQUIDISTRIBUTION OF MODULAR SYMBOLS AND COHOMOLOGY CLASSES FOR QUOTIENTS OF HYPERBOLIC N-SPACE

RESIDUAL EQUIDISTRIBUTION OF MODULAR SYMBOLS AND COHOMOLOGY CLASSES FOR QUOTIENTS OF HYPERBOLIC n-SPACE

PETRU CONSTANTINESCU AND ASBJØRN CHRISTIAN NORDENTOFT

ABSTRACT. We provide a simple automorphic method using Eisenstein series to study the equidistribution of modular symbols modulo primes, which we apply to prove an average version of a conjecture of Mazur and Rubin. More precisely, we prove that modular symbols corresponding to a Hecke basis of weight 2 cusp forms are asymptotically jointly equidistributed mod p while we allow restrictions on the location of the cusps. Additionally, we prove the full conjecture in some particular cases using a connection to Eisenstein congruences. We also obtain residual equidistribution results for modular symbols where we order by the length of the corresponding geodesic. Finally, and most importantly, our methods generalise to equidistribution results for cohomology classes of finite volume quotients of n-dimensional hyperbolic space.

1. INTRODUCTION

Modular symbols are certain periods of weight 2 cusp forms introduced by Birch and Manin and are important objects in number theory. Modular symbols are an indispensable tool for studying (twisted) L-functions of holomorphic cusp forms [29], [31] and for computing modular forms [9]. Modular symbols define elements of certain cohomology groups and the results of this paper thus fit into a bigger picture of the study of (co)homology of arithmetic groups, which has received a lot of attention recently [2], [5] due to their deep connections with number theory coming from [43].

Recently, Mazur and Rubin initiated the study of the arithmetic distribution of modular symbols and put forward a number of conjectures [32], which have received a lot of attention recently [38], [27], [50], [3], [35], [13], [8]. One of these conjectures (see [33]) predicts that (normalised) modular symbols should equidistribute among the residue classes modulo p. An average version of this conjecture was settled by Lee and Sun [27, Theorem I] recently using dynamical methods.

In this paper we introduce a new automorphic method for studying the mod p distribution of modular symbols, which also applies to more general cohomology classes. As is the case in [27], we obtain an average version of the mod p conjecture of Mazur and Rubin (and its generalisations), but with further refinements. Using different arguments, we can actually prove the full conjecture in some special cases (specific p and specific cusp forms), see Section 3.

Date: December 4, 2020.

2010 *Mathematics Subject Classification.* 11F67(primary).

Our automorphic methods enable us to deal with a much more general setup compared to the work of Lee and Sun and thus we obtain a number of new results:

(1) First of all we obtain *joint* equidistribution for the mod p values of modular symbols (appropriately normalised) associated to a Hecke basis of weight 2 cusp forms restricted to cusps which lie in a *fixed* interval of $\mathbb{R}/\mathbb{Z}$.
(2) We show that the values of (unnormalised) modular symbols restricted to cusps lying in a fixed interval of $\mathbb{R}/\mathbb{Z}$ equidistribute mod 1.
(3) We also obtain equidistribution results for modular symbols ordered by the length of the geodesic associated to the corresponding matrices (as opposed to the denominator of the cusp).
(4) Lastly (and most interestingly) we extend the equidistribution results to classes in the cohomology of general finite volume quotients of higher dimensional hyperbolic spaces.

We note that in the case of higher dimensional hyperbolic spaces there is interesting torsion in the cohomology. The breakthrough of Scholze [43] established that such torsion classes have associated Galois representations. This was actually our original motivation for studying the higher dimensional cases. Furthermore Bergeron and Venkatesh [2] have conjectured that at least in the three dimensional case there is an abundance of torsion in the cohomology. In this paper we are able to shed light on the distribution properties of these cohomology classes. In Section 8 we will survey what is know about the dimensions of the cohomology groups, which our results apply to.

Remark 1.1. Our method is automorphic in nature and relies on the theory of Eisenstein series. It can be seen as a discrete version of the method introduced by Petridis and Risager in [36] for studying the distribution of modular symbols. They consider the perturbation of the family of characters χ^ε as $\varepsilon \to 0$, whereas we consider the discrete family χ^m for $m \in \mathbb{Z}$. In particular, we find it is interesting that residual equidistribution of modular symbols is an almost direct consequence of the meromorphic continuation of twisted Eisenstein series.

1.1. **Results for modular symbols.** Let us state the result in the simplest case for the 2 dimensional hyperbolic space in an arithmetic setup. We define the *modular symbol map* associated to a weight 2 and level N cusp form $f \in \mathcal{S}_2(\Gamma_0(N))$ as the map

$$\mathbb{Q} \ni r \mapsto \langle r, f \rangle := 2\pi i \int_r^{i\infty} f(z)dz, \tag{1.1}$$

where the contour integral is taken along a vertical line. One way to think about this map is as the Poincaré pairing on $\Gamma_0(N)\backslash\mathbb{H}^2$ between the 1-form $2\pi i f(z)dz$ and the homology class of paths containing the geodesic from r to $i\infty$.

Now assume that f is a Hecke newform with associated elliptic curve $E/\mathbb{Q}$ and $N \geq 3$. If we let Ω_+ and Ω_- be the real and imaginary Néron periods of E, then it is a fact that

$$\frac{1}{\Omega_\pm}\left(\langle r, f \rangle \pm \langle -r, f \rangle\right) \in \mathbb{Q} \tag{1.2}$$

for all $r \in \mathbb{Q}$ (see [32, Sec. 1]).

Given a prime p there is a minimal p-adic evaluation $v_p^\pm(f)$ of (1.2) among all $a/q \in \mathbb{Q}$ with $N|q$ (since the image under the modular symbols map is a lattice).

We define

$$\mathfrak{m}^{\pm}_{f,p}(r) := \frac{p^{-v_p^{\pm}(f)}}{\Omega_{\pm}} \left(\langle r, f\rangle \pm \langle -r, f\rangle\right),$$

which is p-integral for $r = a/q$ with $N|q$, but not all divisible by p. Then given a basis of newforms $f_1, \dots, f_d$, we can consider the map $\mathbb{Q} \to (\mathbb{Z}/p\mathbb{Z})^{2d} \times (\mathbb{R}/\mathbb{Z})$ given by

$$\mathfrak{m}_{N,p}(r) := (\mathfrak{m}^{+}_{f_1,p}(r), \mathfrak{m}^{-}_{f_1,p}(r), \dots, \mathfrak{m}^{+}_{f_d,p}(r), \mathfrak{m}^{-}_{f_d,p}(r), r)$$

as a random variable defined on the outcome space

$$\Omega_{Q,N} := \{a/q \mid 0 < a < q \le Q, (a,q) = 1, N|q\}$$

endowed with the uniform probability measure. Then we have the following equidistribution result.

Theorem 1.2. *The random variables $\mathfrak{m}_{N,p}$ defined on the outcome spaces $\Omega_{Q,N}$ converge in distribution to the uniform distribution on $(\mathbb{Z}/p\mathbb{Z})^{2d} \times (\mathbb{R}/\mathbb{Z})$ as $Q \to \infty$. This means in concrete terms that for any fixed $\mathbf{a} \in (\mathbb{Z}/p\mathbb{Z})^{2d}$ and any interval $I \subset \mathbb{R}/\mathbb{Z}$, we have*

$$\frac{\#\left\{a/q \in \Omega_{Q,N} \cap I \mid (\mathfrak{m}^{+}_{f_1,p}(a/q), \dots, \mathfrak{m}^{-}_{f_d,p}(a/q)) \equiv \mathbf{a} \bmod p\right\}}{\#\Omega_{Q,N}} = \frac{|I|}{p^{2d}} + o(1)$$

as $Q \to \infty$.

Secondly we can consider the map $\mathbb{Q} \to (\mathbb{R}/\mathbb{Z})^{2d+1}$ given by

$$\mathfrak{m}_{N,\mathbb{R}/\mathbb{Z}}(r) = (\mathrm{Re}\langle r, f_1\rangle, \mathrm{Im}\langle r, f_1\rangle, \dots, \mathrm{Im}\langle r, f_d\rangle, r), r \in \mathbb{Q}, \tag{1.3}$$

as a random variable defined also on $\Omega_{Q,N}$ as above. It follows from a classical result of Schneider [42] in transcendental theory that $\mathrm{Re}\langle \cdot, f_n\rangle, \mathrm{Im}\langle \cdot, f_n\rangle$ for $n = 1, \dots, d$ all take some non-rational values. It is therefore tempting to think that the values should equidistribute on the circle $\mathbb{R}/\mathbb{Z}$, which is exactly what we prove.

Theorem 1.3. *The random variables $\mathfrak{m}_{N,\mathbb{R}/\mathbb{Z}}$ defined on the outcome spaces $\Omega_{Q,N}$ converge in distribution to the uniform distribution on $(\mathbb{R}/\mathbb{Z})^{2d+1}$ as $Q \to \infty$. This means in concrete terms that for any fixed product of intervals $\prod_{n=1}^{2d+1} I_n \subset (\mathbb{R}/\mathbb{Z})^{2d+1}$, we have*

$$\frac{\#\left\{a/q \in \Omega_{Q,N} \cap I_{2d+1} \mid (\mathrm{Re}\langle a/q, f_1\rangle, \dots, \mathrm{Im}\langle a/q, f_d\rangle) \in \prod_{n=1}^{2d} I_n\right\}}{\#\Omega_{Q,N}} = \prod_{n=1}^{2d+1} |I_n| + o(1)$$

as $Q \to \infty$.

We observe that the modular symbols map gives rise to a map $\Gamma_0(N) \to \mathbb{C}$ by putting $\langle \gamma, f\rangle := \langle \gamma\infty, f\rangle$, where $\gamma\infty = a/c$ with a, c the left upper and lower entries of $\gamma \in \Gamma_0(N)$. By shifting the contour and doing a change of variable we see that

$$\langle \gamma_1\gamma_2, f\rangle = \langle \gamma_1, f\rangle + 2\pi i \int_{\gamma_1\infty}^{\gamma_1\gamma_2\infty} f(z)dz = \langle \gamma_1, f\rangle + \langle \gamma_2, f\rangle,$$

which shows that modular symbols define an additive character on $\Gamma_0(N)$ and thus an element of (the cuspidal part of) the cohomology group $H^1(\Gamma_0(N), \mathbb{C})$. Furthermore, by the integrality conditions, we see that the normalised modular symbols $\mathfrak{m}_{f,p}^{\pm}$ define elements of $H^1(\Gamma_0(N), \mathbb{F}_p)$. This view point is useful for generalisations.

Remark 1.4. We note that in [27], the slightly larger outcome space $\{a/q \mid 0 < a < q \leq Q, (a,q) = 1\}$ is considered (following Mazur and Rubin), that is, without the condition that $N|q$. In fact, equidistribution on this outcome space does *not* hold in the generality above. One has to exclude some bad primes p (see Remark 3.2 below). Our methods can also deal with this larger outcome space, by considering the Fourier expansion of Eisenstein series at different cusps, as is done in [38]. The outcome space $\Omega_{Q,N}$ above is, however, very natural from the cohomological perspective and for simplicity we will restrict to this case.

1.2. **Distribution of cohomology classes.** More generally, let $\mathrm{SO}(n+1,1)$ be the special orthogonal group with signature $(n+1,1)$, which we identity with the group of isometries of the $(n+1)$-dimensional upper half space $\mathbb{H}^{n+1}$. Now, for a co-finite subgroup with cusps $\Gamma < \mathrm{SO}(n+1,1)$, we will study the distribution of unitary characters of Γ or, equivalently, cohomology classes in $H^1(\Gamma, \mathbb{R}/\mathbb{Z})$. These cohomology groups have been studied in many contexts ([41], [16, Chap. 7]) and especially the case $n = 2$ is very appealing as it corresponds to Kleinian groups due to the exceptional isomorphism $\mathrm{SO}(3,1) \cong \mathrm{SL}_2(\mathbb{C})$.

1.2.1. *Results with arithmetic ordering.* Let $\Gamma \subset \mathrm{SO}(n+1,1)$ be as above and assume that the associated symmetric space $\Gamma\backslash\mathbb{H}^{n+1}$ has a cusp at ∞. Let $\Gamma'_\infty \subset \Gamma$ be the parabolic subgroup fixing the cusp at ∞. Note that since Γ is discrete, there exists a lattice $\Lambda < \mathbb{R}^n$ such that Γ'_∞ is exactly the group of motions corresponding to translations by Λ. We will study the distribution of unitary characters trivial on Γ'_∞ or, equivalently, elements of the cohomology group $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$.

Our distribution results are with respect to a natural arithmetic ordering on $\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$ which generalises the ordering in the definition of $\Omega_{Q,N}$ above. To define this, we use the model $\mathrm{SV}_{n-1} \cong \mathrm{Iso}^+(\mathbb{H}^{n+1})$, where SV_{n-1} is the Vahlen group consisting of 2×2 matrices over a specific Clifford algebra, introduced in [1] (see Section 4.2 below for a detailed construction). This model provides a natural generalisation to $n > 2$ of the familiar models $\mathrm{SV}_0 = \mathrm{SL}_2(\mathbb{R})$ and $\mathrm{SV}_1 = \mathrm{SL}_2(\mathbb{C})$. The Vahlen model has been used before to study automorphic forms on $\mathbb{H}^{n+1}$, for example by Elstrodt, Grunewald, and Mennicke [15] to prove a generalisation of the Selberg Conjecture regarding the first non-zero eigenvalue of the Laplacian and by Södergren [48] for proving equidistribution of horospheres on $\mathbb{H}^{n+1}$.

We will order by the norm of the lower left entry in this matrix, which generalises the arithmetic ordering considered in the literature for the cases $n = 1$ and $n = 2$, see [38], [8], [35]. We define the following outcome space:

$$T_\Gamma(X) = \{\gamma \in \Gamma'_\infty\backslash\Gamma/\Gamma'_\infty \mid 0 < |c_\gamma| < X\}, \tag{1.4}$$

where $\gamma = \begin{pmatrix} a_\gamma & b_\gamma \\ c_\gamma & d_\gamma \end{pmatrix} \in \mathrm{SV}_{n-1}$ in the Vahlen group model and $|\cdot|$ denotes the norm on the Clifford algebra. The ordering defining $T_\Gamma(X)$ can also be described relatively naturally using the standard model for $\mathrm{SO}(n+1,1)$, see Remark 4.6 for details.

Now let $\omega_1, \ldots, \omega_d$ be elements of $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ in *general position*, meaning for any $(n_1, \ldots, n_d) \in \mathbb{Z}^d$, we have

$$n_1\omega_1 + \ldots + n_d\omega_d = 0 \in H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z}) \Leftrightarrow \Big(n_i\omega_i = 0 \in H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z}), \forall i = 1, \ldots, d\Big).$$

As an example one can pick $\omega_1, \ldots, \omega_d$ to be a basis for the non-torsion part of the cohomology group or a $\mathbb{F}_p$-basis for $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{F}_p)$. We notice that the image of ω_i is either dense in $\mathbb{R}/\mathbb{Z}$ or finite (recall that ω_i defines an additive character $\Gamma \to \mathbb{R}/\mathbb{Z}$). In the first case we put $J_i = \mathbb{R}/\mathbb{Z}$ and in the latter case we put $J_i = \mathbb{Z}/m_i\mathbb{Z}$, where m_i is the cardinality of the image of ω_i. We equip $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$ with the obvious choices of probability measures, Lebesque and uniform respectively. Finally associated to $\gamma \in \Gamma'_\infty \backslash \Gamma / \Gamma'_\infty$, we define the invariant $\gamma\infty \in (\mathbb{R}^n \cup \{\infty\})/\Lambda$, see Section 4.4 for more details.

Now given $X > 0$, we consider

$$\omega(\gamma) := (\omega_1(\gamma), \ldots, \omega_d(\gamma), \gamma\infty)$$

as a random variable with values in $\prod_{i=1}^d J_i \times (\mathbb{R}^n/\Lambda)$ defined on the outcome space $T_\Gamma(X)$ endowed with the uniform probability measure. Then we have the following equidistribution result.

Theorem 1.5. *The random variables ω defined on the outcome spaces $T_\Gamma(X)$ are asymptotically uniformly distributed on $\prod_{i=1}^d J_i \times (\mathbb{R}^n/\Lambda)$ as $X \to \infty$. This means in concrete terms that for any fixed (continuity) subsets $A_i \subset J_i$ and $B \subset \mathbb{R}^n/\Lambda$, we have*

$$\frac{\#\left\{\gamma \in T_\Gamma(X) \mid (\omega_1(\gamma), \ldots, \omega_d(\gamma)) \in \prod_{i=1}^d A_i, \gamma\infty \in B\right\}}{\#T_\Gamma(X)} = \prod_{i=1}^d \frac{|A_i|}{|J_i|} \cdot \frac{|B|}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} + o(1)$$

as $X \to \infty$.

Remark 1.6. The assumption on the existence of cusps is essential in Theorem 1.5 since we rely on the theory of Eisenstein series. Besides this, our methods are pretty robust and apply to non-arithmetic subgroups equally well.

Remark 1.7. Notice that the number of choices of cohomology classes in $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ in general position is infinite unless $\Gamma/\langle[\Gamma, \Gamma], \Gamma'_\infty\rangle$ is torsion. See Section 8 for results on the size of $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$.

1.2.2. *Results when ordered by length of geodesics.* We can also obtain equidistribution of the cohomology classes if we order by the length of the associated geodesics. We denote by $\mathrm{Conj}_{\mathrm{hyp}}(\Gamma)$ the set of conjugacy classes in Γ which do not correspond to the identity, parabolic or elliptic elements. Then, for each $\{\gamma\} \in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma)$, there is a unique corresponding closed geodesic on $\Gamma\backslash\mathbb{H}^{n+1}$ whose length we denote by $l(\gamma)$.

Theorem 1.8. *Let $\omega = (\omega_1, \ldots, \omega_d)$ be defined from a set of cohomology classes in general position as above. The random variables ω defined on conjugacy classes ordered by the length of the geodesics are asymptotically equidistributed on $\prod_{i=1}^d J_i$. This means in concrete terms that for any fixed (continuity) subsets $A_i \subset J_i$, we have*

$$\frac{\#\{\{\gamma\} \in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma) \mid l(\gamma) \leq X, (\omega_1(\gamma), \ldots, \omega_d(\gamma)) \in \prod_{i=1}^d A_i\}}{\#\{\{\gamma\} \in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma) \mid l(\gamma) \leq X\}} = \prod_{i=1}^d \frac{|A_i|}{|J_i|} + o(1)$$

as $X \to \infty$.

Remark 1.9. In the case of Theorem 1.8, we can remove the assumption that Γ has cusps. In fact the proof becomes more complicated in the presence of cusps.

Acknowledgements

We would like to thank Paul Gunnells, Yiannis Petridis and Morten Risager for useful discussions and feedback.

2. Proof sketch in the case of $\mathbb{H}^2$

The key ideas of the proofs of our main theorems are quite simple, having at their core the analytic continuation of twisted Eisenstein series and twisted trace formulas respectively. We will sketch the proof of Theorem 1.2 in the simplest case, which is the one dealt with in [27], where we consider only one cusp form and no restrictions on the location of $r = a/q$ in $\mathbb{R}/\mathbb{Z}$.

Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a newform of weigth 2 and level N and let $\mathfrak{m}^{\pm}_{f,p} : \Gamma_0(N) \to \mathbb{F}_p$ be the associated normalised modular symbols defined above. Recall that this defines a non-trivial additive character. We would like to show that the values of $\mathfrak{m}^{\pm}_{f,p}$ on the set $\Omega_{Q,N} = \{a/q \mid 0 < a < q \leq Q, (a,q) = 1, N|q\}$ equidistribute on $\mathbb{Z}/p\mathbb{Z}$ as $Q \to \infty$.

To do this we introduce for any $l \in (\mathbb{Z}/p\mathbb{Z})^\times$ the unitary character $\chi_l : \Gamma_0(N) \to \mathbb{C}^\times$ defined by;

$$\chi_l(\gamma) := e^{2\pi i \mathfrak{m}^{\pm}_{f,p}(\gamma) l/p}, \quad \gamma \in \Gamma_0(N).$$

By Weyl's Criterion [24, page 487] in order to conclude equidistribution, it suffices to detect cancelation in the Weyl sums; that is to prove for all $l \in (\mathbb{Z}/p\mathbb{Z})^\times$ that

$$\sum_{a/q \in \Omega_{Q,N}} \chi_l(a/q) \ll X^{2-\delta},$$

for some $\delta > 0$ where $\chi_l(a/q) := \chi_l(\gamma)$ with $\gamma \in \Gamma_0(N)$ such that $\gamma\infty = a/q$.

The key observation is now that the generating series for these Weyl sums appears very naturally as the constant term of an appropriate Eisenstein series. The cancelation in the Weyl sums is now a simple analytic consequence of the analytic properties of the corresponding Eisenstein series.

To be precise; associated to χ_l we have the following twisted Eisenstein series:

$$E(z,s,\chi_l) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(N)} \overline{\chi_l}(\gamma) \operatorname{Im}(\gamma z)^s,$$

where $\Gamma_\infty = \langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$. This Eisenstein series defines a holomorphic function for $\operatorname{Re} s > 1$ and by the work of Selberg [44, Chap. 39] admits meromorphic continuation to the entire complex plane with a pole at $s = 1$ if and only if χ_l is trivial. Note that in general the character χ_l might not come from an adelic one, but Selberg's theory applies equally well.

Now a standard calculation using Poisson summation shows that the constant term of the Fourier expansion of $E(z,s,\chi_l)$ is given by

$$y^s + \frac{\pi^{1/2} y^{1-s} \Gamma(s-1/2)}{\Gamma(s)} L_l(s),$$

with

$$L_l(s) := \sum_{c>0, N|c} \left(\sum_{0<d<c, (c,d)=1} \overline{\chi_l}\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\right) \right) c^{-2s},$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is a matrix in $\Gamma_0(N)$ with lower entries c, d. We observe that $L_l(s)$ is exactly the generating series for the Weyl sums above, as promised.

Now from the meromorphic continuation of the Eisenstein series itself, we also get meromorphic continuation of the generating series $L_l(s)$, and since χ_l is non-trivial we conclude that $L_l(s)$ is analytic for $\mathrm{Re}\, s > 1 - \delta$ for some $\delta > 0$. Thus we get the wanted cancelation in Weyl sums using the standard machinery from complex analysis if we can get bounds on vertical lines of $L_l(s)$. It turns of that such bounds follow from the general bound for matrix coefficients also due to Selberg, and thus we are done.

This shows how to deduce equidistribution of modular symbols using Eisenstein series. The proof for classes in the first cohomology of quotients of higher dimensional hyperbolic spaces uses the same idea, although some parts of the argument require some more technical work.

In order to obtain equidistribution results when restricting the cusps to a specific interval $I \subset \mathbb{R}/\mathbb{Z}$, we will have to use all the Fourier coefficients of the Eisenstein series as is done in [38]. To deal with equidistribution for modular symbols defined on conjugacy classes ordered by the length of the associated geodesic, we will use twisted Selberg trace formulas to study the corresponding Weyl sums.

3. Some special cases of the conjecture of Mazur and Rubin

In this section we explain a specific case of the conjecture of Mazur and Rubin on the residual distribution of modular symbols, which we can resolve without taking an extra average. We consider modular symbols of Hecke congruence subgroups $\Gamma_0(N)$ modulo primes $p > 2$ dividing $\varphi(N)$, which are connected to congruences between Eisenstein series and cusp forms. These types of congruences have been studied extensively before in number theory [30, Section 9], [31]. We will assume for simplicity that $N \geq 5$ is prime (in particular $\Gamma_0(N)$ is torsion-free) as is done in [30].

From the perspective of cohomology, the congruence phenomena manifests itself through the fact that Dirichlet characters modulo N define cohomology classes. The distribution for this specific cohomology class is much easier to understand, and we would thus like to connect it to a modular symbol of a cusp form. Thus we have to somehow rule out that this cohomology class is a linear combination of ones coming from modular symbols associated to different cusp forms. Using a multiplicity one result, we get the desired conclusion. The precise result we can prove is the following.

Theorem 3.1. *Let p and $N \geq 5$ be odd primes such that $p|(N-1)$. Then there exists a newform $f \in \mathcal{S}_2(\Gamma_0(N))$ of weight 2 and level N such that the values of $\mathfrak{m}_{f,p}^{+}$ on $\{\frac{a}{q} \mid (a,q) = 1, 0 < a < q\}$ equidistribute modulo p as $q \to \infty$ with $N|q$.*

Proof. Given a prime $p \mid N-1$ with $p > 2$, we know that the space of order p Dirichlet characters modulo N is one dimensional as an $\mathbb{F}_p$-vector space (since $(\mathbb{Z}/N\mathbb{Z})^\times$ is cyclic). Given a generator χ mod N of order p we get an element of $H^1(\Gamma_0(N), \mathbb{F}_p)$

defined by $\gamma \mapsto \chi(a_\gamma)$ (which we denote by σ_χ), where we identify $\mathbb{F}_p$ with the image of χ and a_γ is the upper-left entry of γ.

First of all we observe that σ_χ is trivial on the parabolic subgroups of $\Gamma_0(N)$; since N is prime we only have to check it for $\langle \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \rangle$ and $\langle \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right) \rangle$ corresponding to the two cusps ∞ and 0. This implies that σ_χ defines a class in the parabolic cohomology group $H^1_P(\Gamma_0(N), \mathbb{F}_p)$ (see Section 7.1 below or [46, Chapter 8] for a definition). It follows by a mod p version of Eichler–Shimura isomorphism (see [27, (3.5)]) that the associations $f \mapsto \mathfrak{m}^{\pm}_{f,p}$ defined for newforms f induce an isomorphism

$$H^1_P(\Gamma_0(N), \mathbb{F}_p) \cong \mathcal{S}_2(\Gamma_0(N))_{\mathbb{F}_p} \oplus \mathcal{S}_2(\Gamma_0(N))_{\mathbb{F}_p} \tag{3.1}$$

where $\mathcal{S}_2(\Gamma_0(N))_{\mathbb{F}_p}$ denotes the space of cusp forms of weight 2 and level N with coefficients in $\mathbb{F}_p$ (which we will just think of as the formal $\mathbb{F}_p$-vector space generated by Hecke eigenforms of weight 2 and level N).

To relate σ_χ to cusp forms, we need to consider the Hecke action. Recall that we have an action by the Hecke operators on the cohomology group defined as follows (see [45, Chapter 8.3]). Let

$$\alpha_{r,l} = \left(\begin{smallmatrix} 1 & r \\ 0 & l \end{smallmatrix}\right), r = 0, \dots, l-1 \quad \text{and} \quad \alpha_{l,l} = \left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right),$$

for $l \neq N$ prime. Then given $\gamma \in \Gamma_0(N)$, we define $\gamma_{r,l} \in \Gamma_0(N)$ by $\alpha_{r,l}\gamma = \gamma_{r,l}\alpha_{\sigma(r),l}$ for $r = 0, \dots, l$ and some $\sigma(r) \in \{0, \dots, l\}$. Then we define the operator T_l as follows:

$$(T_l\omega)(\gamma) := \sum_{r=0}^{l} \omega(\gamma_{r,l}),$$

for $\omega \in H^1(\Gamma_0(N), \mathbb{F}_p)$ and $\gamma \in \Gamma_0(N)$. At the prime $l = N$, we have the Atkin–Lehner operator U defined by

$$(U\omega)(\gamma) := \omega\left(\left(\begin{smallmatrix} 0 & 1 \\ N & 0 \end{smallmatrix}\right) \gamma \left(\begin{smallmatrix} 0 & 1 \\ N & 0 \end{smallmatrix}\right)^{-1} \right).$$

One sees directly that $U\sigma_\chi = -\sigma_\chi$. Furthermore it is a small computation that if $l|c$ (where c is the lower left entry of γ), then

$$\gamma_{r,l} = \left(\begin{smallmatrix} a+cr & * \\ cp & * \end{smallmatrix}\right), \text{ for } r = 0, \dots, l-1, \quad \text{and} \quad \gamma_{l,l} = \left(\begin{smallmatrix} a & * \\ c/p & * \end{smallmatrix}\right).$$

And if $l \nmid c$, then $\gamma_{l,l} = \left(\begin{smallmatrix} ap & * \\ c & * \end{smallmatrix}\right)$ and for $r = 0, \dots, l-1$ we have

$$\gamma_{r,l} = \begin{cases} \left(\begin{smallmatrix} (a+cr)/p & * \\ c & * \end{smallmatrix}\right), & r \equiv -a\bar{c} \mod l, \\ \left(\begin{smallmatrix} a+cr & * \\ cp & * \end{smallmatrix}\right), & \text{else.} \end{cases}$$

It follows that $T_l\sigma_\chi = (l+1)\sigma_\chi$. Now we employ a general multiplicity one result of Mazur [30, Proposition 9.2], which implies that a cusp form of weight 2 and level N is uniquely determined by its Hecke eigenvalues (i.e. the eigenvalues under the action of U and the operators T_l for $l \neq N$) modulo p, for $p \neq N$ prime. Thus from the above calculations we conclude that there exists a unique Hecke eigenform $f \in \mathcal{S}_2(\Gamma_0(N))$ such that σ_χ is a linear combination of $\mathfrak{m}^+_{f,p}$ and $\mathfrak{m}^-_{f,p}$ (considered as elements of $H^1_P(\Gamma_0(N), \mathbb{F}_p)$ the obvious way).

Finally, we recall that $H^1_P(\Gamma_0(N), \mathbb{F}_p)$ can be diagonalised by the involution ι given by conjugation with $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (here we need $p > 2$) as follows from [33, Sec. 1]. We observe that the eigenvalue of σ_χ under the action of ι is $+1$ (since the order of χ is odd).

Thus we conclude that $\mathfrak{m}_{f,p}^{+} = m \cdot \sigma_\chi$ for some $m \in \mathbb{F}_p \backslash \{0\}$ (see also [30, Proposition 9.6], which draws a closely related conclusion).

Now it is easy to see that the values of $\mathfrak{m}_{f,p}^{+}$ (with f as above) on $\{a/q \mid (a,q) = 1, 0 < a < q\}$ equidistribute modulo p as $q \to \infty$ with $N|q$. This follows directly from the fact that

$$\mathfrak{m}_{f,p}^{+}(a/q) = m \cdot \chi(a) \mod p,$$

and that the values of χ clearly equidistribute. Notice that actually the values equidistribute exactly. □

This settles the conjecture of Mazur and Rubin in these very special cases, whereas in general the conjecture seems out of reach without the extra average both with the automorphic and the dynamical approach.

Remark 3.2. Strictly speaking the conjectures of Mazur and Rubin [33] are only formulated for primes p and cusp forms corresponding to elliptic curves E where the residual representation of E mod p is surjective and p is an ordinary and good prime of E. This is not the case in the example considered above. We, however, expect the statement of Theorem 3.1 to be true for general $\mathfrak{m}_{f,p}^{\pm}$ with $f \in \mathcal{S}_k(\Gamma_0(N))$ and p prime.

4. Geometry of $\mathbb{H}^{n+1}$

We introduce two models for the $(n+1)$-dimensional hyperbolic space and the connections between them. We look at the upper half-space (Poincaré) model $\mathbb{H}^{n+1}$ and the hyperboloid (Klein) model $\mathbb{K}^{n+1}$. We briefly describe some geometric and arithmetic properties of the space $\Gamma\backslash\mathbb{H}^{n+1}$, where Γ is a cofinite discrete subgroup of isometries. Our main references for this section are [1], [14] and [15].

We denote by $\mathrm{Iso}^+(\mathbb{H}^{n+1})$ the space of orientation preserving isometries of the hyperbolic $(n+1)$-space. We say that $\gamma \in \mathrm{Iso}^+(\mathbb{H}^{n+1})$ is *elliptic* if it has exactly one fixed point in $\mathbb{H}^{n+1}$. A non-elliptic isometry is called *parabolic* if it has exactly one fixed point on the boundary $\mathbb{R}^n \cup \{\infty\}$ and *hyperbolic* if it has 2 fixed points on $\mathbb{H}^{n+1} \cup \mathbb{R}^n \cup \{\infty\}$ (hence both of them on the boundary or both of them inside $\mathbb{H}^{n+1}$). We note that our definition of a hyperbolic motion includes what is known in the literature as loxodromic motions, so any isometry is either the identity or one of the three types described. We quote [19] for a thorough discussion of the three classes.

4.1. **Definitions.** We will now describe the *upper-half space model* $\mathbb{H}^{n+1}$ for hyperbolic $(n+1)$-space. Let $q : \mathbb{R}^n \to \mathbb{R}$ a quadratic non-degenerate form and $\mathcal{C}(q)$ the associated *Clifford algebra*, i.e. the free $\mathbb{R}$-algebra on $\{e_1, \dots, e_n\}$ modulo the relations

$$e_i^2 = q(e_i), \quad e_ie_j = -e_je_i, \quad \text{where} \quad i,j = 1, \cdots, n, \ i \neq j,$$

where $e_1, \dots, e_n$ is the standard basis for $\mathbb{R}^n$. We denote by $\mathcal{E}_n$ the set of all subsets of $\{1, \dots, n\}$. Then for $M = \{i_1, \dots, i_k\} \in \mathcal{E}_n$ with $i_1 < \cdots < i_k$, we define

$$e_M := e_{i_1} \cdot \cdots \cdot e_{i_k}, \quad e_\emptyset := 1 \in \mathcal{C}(q).$$

Then one can check that $\{e_M \mid M \in \mathcal{E}_n\}$ is a $\mathbb{R}$-basis for $\mathcal{C}(q)$.

We have two linear involutions on $\mathcal{C}(q)$ given by

$$\overline{e_M} := (-1)^{|M|(|M|+1)/2} e_M, \quad e_M^* := (-1)^{|M|(|M|-1)/2} e_M, \quad \text{where } M \in \mathcal{E}_n.$$

They satisfy

$$\overline{v}\ \overline{w} = \overline{w}\ \overline{v}, \quad v^*w^* = w^*v^*, \quad \text{for all } v, w \in \mathcal{C}(q).$$

From now on we assume that $q = -I_n$, the negative definite unit form. In this case we write $\mathcal{C}_n$ for $\mathcal{C}(q)$. We denote by $V_n \subset \mathcal{C}_n$ the vector space spanned by $\{1, e_1, \ldots, e_n\}$. It is easy to see that $V_0 \cong \mathbb{R}$ and $V_1 \cong \mathbb{C}$ as $\mathbb{R}$-algebras.

V_n is equipped with the inner product

$$\langle v, w\rangle = \frac{1}{2}(v\overline{w} + \overline{v}w).$$

We note that this coincides with the standard Euclidean inner product if we identify V_n with $\mathbb{R}^{n+1}$ using the basis $\{1, e_1, \ldots, e_n\}$.

For $x = \sum_{M\in\mathcal{E}_n} \lambda_M e_M \in \mathcal{C}_n$, we define the norm

$$|x| := \left(\sum_{M\in\mathcal{E}_n} \lambda_M^2\right)^{1/2}. \tag{4.1}$$

We note that for $x \in V_n$,

$$|x|^2 = \langle x, x\rangle .$$

Now, if $\Lambda < V_n$ is a lattice, we define the dual lattice as

$$\Lambda^\circ := \{w \in V_n \mid \langle v, w\rangle \in \mathbb{Z} \text{ for all } v \in \Lambda\}.$$

We now define the following model of hyperbolic $(n+1)$-space:

$$\mathbb{H}^{n+1} := \{x_0 + x_1e_1 + \cdots + x_ne_n \mid x_0, x_1, \ldots, x_{n-1} \in \mathbb{R}, x_n > 0\} .$$

We have the maps $x : \mathbb{H}^{n+1} \to V_{n-1}$ and $y : \mathbb{H}^{n+1} \to (0, \infty)$ given by

$$x(P) := x_0 + x_1e_1 + \cdots + x_{n-1}e_{n-1}, \quad y(P) := x_n,$$

where $P = x_0 + x_1e_1 + \cdots + x_ne_n \in \mathbb{H}^{n+1}$. We can think of $x(P)$ as an element of $\mathbb{R}^n$ via the above. Then from (4.1) we see that

$$|P|^2 = |x(P)|^2 + |y(P)|^2.$$

We equip $\mathbb{H}^{n+1}$ with the hyperbolic metric coming from the line element:

$$ds^2 = \frac{dx_0^2 + dx_1^2 + \cdots + dx_n^2}{x_n^2}, \tag{4.2}$$

which makes $\mathbb{H}^{n+1}$ a Riemannian manifold with constant negative curvature -1. The volume element is given by

$$dv = \frac{dx_0dx_1 \ldots x_n}{x_n^{n+1}}.$$

The *hyperbolic Laplace–Beltrami operator* is given by

$$\Delta = x_n^2\left(\frac{\partial^2}{\partial x_0^2} + \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\right) - (n-1)x_n\frac{\partial}{\partial x_n}, \tag{4.3}$$

in this model.

4.2. **Vahlen group.** We will use the above upper-half space model to describe the group of (oriented) isometries in a way that is convenient for our purposes. We let $T_n \subset \mathcal{C}_n$ be the multiplicative subgroup generated by $V_n \setminus \{0\}$. As in [1, p. 219] or [15, p. 648], we define the *Vahlen group* SV_n to be

$$\mathrm{SV}_n := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathcal{C}_n) \;\middle|\; \begin{array}{c} \text{(i) } a,b,c,d \in T_n \cup \{0\} \\ \text{(ii) } \overline{a}b, \overline{c}d \in V_n \\ \text{(iii) } ad^* - bc^* = 1 \end{array} \right\}. \tag{4.4}$$

We can easily check that $\mathrm{SV}_0 = \mathrm{SL}_2(\mathbb{R})$ and $\mathrm{SV}_1 = \mathrm{SL}_2(\mathbb{C})$ as $\mathbb{R}$-algebras. Then SV_n is a group under matrix multiplication with inverse

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d^* & -b^* \\ -c^* & a^* \end{pmatrix}. \tag{4.5}$$

We can now define the action of SV_{n-1} on $\mathbb{H}^{n+1}$, which resembles the actions of $\mathrm{SL}_2(\mathbb{R})$ and $\mathrm{SL}_2(\mathbb{C})$ on $\mathbb{H}^2$ and $\mathbb{H}^3$, respectively, as can be seen from the following result.

Theorem 4.1 ([15], Theorem 1.3)**.** *Let* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SV}_{n-1}$ *and* $P \in \mathbb{H}^{n+1}$. *Then* $cP + d \in T_n$ *and we define*

$$\gamma P := (aP+b)(cP+d)^{-1} \in \mathbb{H}^{n+1}. \tag{4.6}$$

The map $P \mapsto \gamma P$ *is an orientation preserving isometry of* $\mathbb{H}^{n+1}$. *Moreover, all orientation preserving isometries are obtained in this way and we have the induced isomorphism* $\mathrm{SV}_{n-1}/\{I, -I\} \cong \mathrm{Iso}^+(\mathbb{H}^{n+1})$.

What is convenient about this description of $\mathrm{Iso}^+(\mathbb{H}^{n+1})$ is that one gets very familiar expressions for the coordinate-projections of the image under the action of $\gamma \in \mathrm{SV}_{n-1}$ on $P = (x, y) \in \mathbb{H}^{n+1}$.

Lemma 4.2 ([15], page 648)**.** *Let* $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SV}_{n-1}$ *and* $P = x + ye_n \in \mathbb{H}^{n+1}$. *Then*

$$x(\gamma P) = \frac{(ax+P)(\overline{cx+d}) + a\overline{c}y^2}{|cx+d|^2 + |c|^2y^2} \tag{4.7}$$

and

$$y(\gamma P) = \frac{y}{|cx+d|^2 + |c|^2y^2}. \tag{4.8}$$

4.3. **Hyperboloid model.** We recall the hyperboloid (or Klein) model for the hyperbolic $(n+1)$-space given by

$$\mathbb{K}^{n+1} := \{z \in \mathbb{R}^{n+2} \mid z_0^2 - z_1^2 - \ldots - z_{n+1}^2 = 1, z_0 > 0\},$$

where $z = (z_0, \ldots, z_{n+1})$. The line element

$$ds^2 = -dz_0^2 + dz_1^2 + \cdots + dz_{n+1}^2 \tag{4.9}$$

defines the hyperbolic metric on $\mathbb{K}^{n+1}$.

The group

$$\mathrm{SO}(n+1,1) := \left\{\gamma \in \mathrm{SL}_{n+2}(\mathbb{R}) \mid \gamma^T I_{1,n+1} \gamma = I_{1,n+1}\right\} \tag{4.10}$$

acts on $\mathbb{K}^{n+1}$ by left multiplication, where $I_{1,n+1} = \begin{pmatrix} -1 & 0 \\ 0 & I_n \end{pmatrix}$. We can identify the set of orientation preserving isometries by $\mathrm{SO}^0(n+1,1)$, where $\mathrm{SO}^0(n+1,1)$ is the component of the identity element in $\mathrm{SO}(n+1,1)$. Hence we have the identifications

$$\begin{aligned} \mathrm{Iso}^+(\mathbb{K}^{n+1}) &\cong \mathrm{SO}^0(n+1,1), \\ \mathbb{K}^{n+1} &\cong \mathrm{SO}^0(n+1,1)/\mathrm{O}(n+1). \end{aligned}$$

We have the following important result which connects the two models.

Theorem 4.3 ([14], Section 5)**.** *We can go between the two models $\mathbb{H}^{n+1}$ and $\mathbb{K}^{n+1}$ as follows.*

(i) There exists a bijection $\Phi : \mathbb{H}^{n+1} \to \mathbb{K}^{n+1}$ which is also an isometry, i.e. the pullback of the line element (4.9) *via Φ is the line element* (4.2)*.*

(ii) There exists an isomorphism $\Psi : \mathrm{SV}_{n-1}/\{\pm I\} \xrightarrow{\sim} \mathrm{SO}^0(n+1,1)$ such that Φ is Ψ-equivariant, i.e.

$$\Phi(\gamma \cdot P) = \Psi(\gamma)\Phi(P),$$

for all $\gamma \in \mathrm{SV}_{n-1}$ and $P \in \mathbb{H}^{n+1}$.

Remark 4.4. The maps Φ and Ψ are explicitly constructed in [14, Section 5]. This result allows us to move freely between the two identifications of hyperbolic $(n+1)$-space.

4.4. **Hyperbolic quotients.** Let Γ be a discrete group of hyperbolic motions such that the surface $\Gamma\backslash\mathbb{H}^{n+1}$ has finite hyperbolic volume. From Theorem 4.3 we note that we can choose freely between the two models $\Gamma < \mathrm{SV}_{n-1} \cong \mathrm{Iso}^+(\mathbb{H}^{n+1})$ or $\Gamma < \mathrm{SO}^0(n+1,1) \cong \mathrm{Iso}^+(\mathbb{K}^{n+1})$. We will mainly work with the Vahlen model since it provides nicer arithmetic descriptions.

We say that $\mathfrak{a} \in \mathbb{R}^n \cup \{\infty\}$ is a cusp for Γ if it is fixed by a parabolic element in Γ. There exists a scaling matrix $\sigma_{\mathfrak{a}} \in \mathrm{SV}_{n-1}$ such that $\sigma_{\mathfrak{a}}\infty = \mathfrak{a}$. We let $\Gamma_{\mathfrak{a}} := \{\gamma \in \Gamma : \gamma\mathfrak{a} = \mathfrak{a}\}$ be the stabilizer of $\mathfrak{a}$ in Γ. We define

$$\Gamma'_{\mathfrak{a}} := \Gamma_{\mathfrak{a}} \cap \sigma_{\mathfrak{a}} \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in \mathrm{SV}_{n-1} \right\} \sigma_{\mathfrak{a}}^{-1} .$$

We note that $\Gamma'_{\mathfrak{a}}$ consists of the parabolic elements in $\Gamma_{\mathfrak{a}}$ together with the identity.

There exists a lattice $\Lambda_{\mathfrak{a}} \leq \mathbb{R}^n$ such that

$$\sigma_{\mathfrak{a}}^{-1}\Gamma'_{\mathfrak{a}}\sigma_{\mathfrak{a}} = \left\{ \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \;:\; \lambda \in \Lambda_{\mathfrak{a}} \right\} .$$

We let $\mathcal{P}_{\mathfrak{a}}$ be a period parallelogram for $\Lambda_{\mathfrak{a}}$ with Euclidean area $\mathrm{vol}(\Lambda_{\mathfrak{a}})$.

We define the *dual lattice* of $\Lambda_{\mathfrak{a}}$ as follows:

$$\Lambda_{\mathfrak{a}}^{\circ} := \{\mu \in \mathbb{R}^n \;:\; \langle \mu, \lambda \rangle \in \mathbb{Z} \text{ for all } \lambda \in \Lambda_{\mathfrak{a}}\} , \tag{4.11}$$

where $\langle \cdot, \cdot \rangle$ is the usual scalar product on $\mathbb{R}^n$.

Say $\mathfrak{a}_1, \cdots, \mathfrak{a}_h \in \mathbb{R}^n \cup \{\infty\}$ are representatives for the equivalent-classes of cusps under the action of Γ. For $Y > 0$, we define the *cuspidal sectors* as follows:

$$\mathcal{F}_{\mathfrak{a}_i}(Y) := \sigma_{\mathfrak{a}_i}\{(x, y) : x \in \mathcal{P}_{\mathfrak{a}_i}, y \geq Y\} .$$

Then for Y large enough, there exists a fundamental domain $\mathcal{F}$ for $\Gamma\backslash\mathbb{H}^{n+1}$ which we can write as a disjoint union

$$\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_{\mathfrak{a}_1}(Y) \cup \cdots \cup \mathcal{F}_{\mathfrak{a}_h}(Y)\ , \tag{4.12}$$

where $\mathcal{F}_0$ is a compact set, see [48, p. 8] or [41, p. 5].

For notational convenience, from now on we will focus only at the cusp at ∞. We drop the subscript by denoting $\Lambda := \Lambda_\infty$, $\mathcal{P} := \mathcal{P}_\infty$ etc. Our theory can be generalised to take all cusps into account.

We will now define our outcome space (1.4) in precise terms, and describe it explicitly in some arithmetic examples. First we note that all elements in such a coset share the lower left entry. Thus it makes sense to define

$$T_\Gamma(X) := \left\{ \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \Gamma'_\infty\backslash\Gamma/\Gamma'_\infty \mid 0 < |c| \leq X \right\}$$

which is the natural generalisation of the outcome space considered by Petridis–Risager in [38, p. 1002]. In (6.7) we provide an asymptotic formula for the size of $T_\Gamma(X)$.

If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$, then by the definition of SV_{n-1} and of the action (4.6) we see that $\gamma\infty = ac^{-1} \in V_{n-1}$, where $\gamma\infty$ is defined as the limit of γP as P tends to the cusp at ∞. We observe that $\gamma\infty$ is well-defined on double cosets in $\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$ up to translations by the lattice Λ. Therefore we see that the map

$$\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty \to \mathbb{R}^n/\Lambda \cup \{\infty\}$$
$$\gamma \mapsto \gamma\infty$$

is well-defined using the identification of V_{n-1} with $\mathbb{R}^n$ as above. A simple consequence of our main theorems is that $\gamma\infty$ become equidistributed on $\mathbb{R}^n/\Lambda$ as we vary along $\gamma \in T_\Gamma(X)$ as $X \to \infty$.

Let

$$C(\Gamma) := \left\{ c \in T_n \mid \exists a, b, d \in T_n : \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma \right\}. \tag{4.13}$$

We will now provide explicit descriptions for both $C(\Gamma)$ and $\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$ in the case of *congruence subgroups*. To define these, let $J \subset \mathcal{C}_n$ be an order stable under the involutions $-$ and $*$. We put $\mathrm{SV}_n(J) := \mathrm{SV}_n \cap M_2(J)$. We also define $V(J) := J \cap V_n$ and $T(J) = J \cap T_n$. For $N \in \mathbb{N}$, we define the *principle congruence subgroup*

$$\mathrm{SV}_n(J; N) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SV}_n(J) \mid a - 1, b, c, d - 1 \in NJ \right\}. \tag{4.14}$$

A subgroup $\Gamma < \mathrm{SV}_n(J)$ is called a *congruence group* if $\mathrm{SV}_n(J; N) < \Gamma$, for some $N \in \mathbb{N}$. We quote [15, Section 4] to provide an explicit description for representatives of $\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$ in the case $\Gamma = \mathrm{SV}_n(J; N)$. In this case, $C(\Gamma) = N \cdot T(J)$ and a set of representatives for $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$ with $c \neq 0$ is given by

$$\left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SV}_n(J) \mid c \in N \cdot T(J),\ (a, d) \in D(c) \right\}$$

where

$$D(c) := \left\{ (a, d) \;\middle|\; \begin{array}{c} a \in J/(N \cdot V(J) \cdot c),\ d \in J/(N \cdot c \cdot V(J)), \\ a - 1, d - 1 \in N \cdot J,\ a\bar{c}, \bar{c}d \in N \cdot V(J) \end{array} \right\}.$$

In the more familiar cases $n = 1$ and $n = 2$, the above reduces to the following.

- $n = 1$. Then $\mathrm{SV}_0 = \mathrm{SL}_2(\mathbb{R})$, $J = \mathbb{Z}$ and $\mathrm{SV}_1(J; N) = \Gamma_1(N)$. Representatives in $\Gamma_1(N)'_\infty \backslash \Gamma_1(N)/\Gamma_1(N)'_\infty$ with $c \neq 0$ are uniquely determined by
$$\{(a, c) \mid c > 0,\ N \mid c,\ a \in (\mathbb{Z}/cN\mathbb{Z})^*,\ a \equiv 1 \bmod N\}\,.$$
If we consider $\Gamma = \Gamma_0(N)$, then representatives are uniquely determined by
$$\{(a, c) \mid c > 0,\ N \mid c,\ a \in (\mathbb{Z}/c\mathbb{Z})^*\}\,.$$
- $n = 2$. Then $\mathrm{SV}_1 = \mathrm{SL}_2(\mathbb{C})$. We take $J = \mathcal{O}_K$, where $\mathcal{O}_K$ is the ring of integers of a quadratic imaginary field K. Let $\mathfrak{n} < \mathcal{O}_K$ be an ideal. We consider congruence subgroups of the form
$$\Gamma_1(\mathfrak{n}) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathcal{O}_K) \mid a - 1, b, c, d - 1 \in \mathfrak{n} \right\},$$
$$\Gamma_0(\mathfrak{n}) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathcal{O}_K) \mid c \in \mathfrak{n} \right\}.$$
In the case $\Gamma_1(\mathfrak{n})$, representatives are uniquely provided by
$$\{(a, c) \mid c \in \mathfrak{n} \setminus \{0\},\ a \in (\mathcal{O}_K/(c \cdot \mathfrak{n}))^*,\ a - 1 \in \mathfrak{n}\}\,,$$
while for $\Gamma_0(\mathfrak{n})$ we have
$$\{(a, c) \mid c \in \mathfrak{n} \setminus \{0\},\ a \in (\mathcal{O}_K/(c))^*\}\,.$$

Remark 4.5. There is also a notion of congruence groups for $\mathrm{SO}(n+1, 1)$. To define them, let Γ be the integral automorphisms of an isotropic quadratic form of signature $(n+1, 1)$ defined over $\mathbb{Q}$. Then a *congruence subgroup* is any subgroup of Γ containing $\{\gamma \in \Gamma \mid \gamma \equiv I_{n+2} \mod N\}$ for some positive integer N, see [41, p. 7]. If $\Gamma < \mathrm{SO}^0(n+1, 1)$ is a congruence subgroup, then $\Psi^{-1}(\Gamma)$ is a congruence subgroup in SV_{n-1}. This fact will be useful when comparing our results with the results mentioned in Section 8. However the converse is not true, there exists a congruence subgroup $\Gamma < \mathrm{SV}_{n-1}$ such that $\Psi(\Gamma)$ is not a congruence subgroup in $\mathrm{SO}^0(n+1, 1)$, see [15, Section 3] for more details.

Remark 4.6. We can also describe the ordering defining $T_\Gamma(X)$ explicitly using the model $\mathrm{SO}(n+1, 1)$ for the isometry group of $\mathbb{H}^{n+1}$. In this case we have
$$|c_\gamma| = \frac{1}{2}(a_{00} + a_{0(n+1)} - a_{(n+1)0} - a_{(n+1)(n+1)}),$$
for $\gamma = (a_{ij}) \in \mathrm{SO}^0(n+1, 1)$ as in (4.10), where c_γ is the lower left entry of $\Psi^{-1}(\gamma)$, i.e in the Vahlen model.

4.5. **Conjugacy classes.** We now look at certain invariants associated to the conjugacy classes $\{\gamma\}$ of Γ alluded to in Theorem 1.8. We refer to [18, Section 5] for more details. We denote by $\mathrm{Conj}_{\mathrm{hyp}}(\Gamma)$ the set of conjugacy classes of hyperbolic elements in Γ. For each $\{\gamma\} \in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma)$, there exists a unique closed geodesic on $\Gamma\backslash\mathbb{H}^{n+1}$ whose length we denote by $l(\gamma)$. The geodesic can be defined as follows: Each conjugacy class corresponds to a free homotopy class on $\Gamma\backslash\mathbb{H}^{n+1}$ via the map $\gamma \mapsto \{P, \gamma P\} \subset \mathbb{H}^{n+1}$, for some point P, and the corresponding geodesic is the path of minimal length among all paths in that class. See [18, Sections 1 and 5] for explicit descriptions of the lengths $l(\gamma)$. Every $\gamma \in \Gamma$ can be written uniquely as $\gamma = \gamma_0^{j(\gamma)}$, where γ_0 is primitive and $j(\gamma) \in \mathbb{N}$. We put

$$\pi_\Gamma(X) := \{\{\gamma_0\} \in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma) \mid \gamma_0 \text{ primitive}, l(\gamma_0) \leq X\}, \tag{4.15}$$

which is exactly the outcome space considered in Theorem 1.8. *The Prime Geodeseic Theorem for* $\mathbb{H}^{n+1}$ gives asymptotics for $\pi_\Gamma(X)$ and was firstly proved by Gangolli [17] in the compact case and by Gangolli and Warner [18, Prop. 5.4] in the non-compact case.

5. Twisted Eisenstein series for $\mathbb{H}^{n+1}$

Let $\Gamma < \mathrm{SV}_{n-1}$, Γ'_∞ and Λ be as in the previous section. We now fix χ a unitary character of Γ which is trivial on Γ'_∞. From this we define the twisted Eisenstein series

$$E(P,s,\chi) = \sum_{\Gamma'_\infty\backslash\Gamma} \overline{\chi(\gamma)} y(\gamma P)^s. \tag{5.1}$$

It is absolutely convergent for $\mathrm{Re}(s) > n$. It satisfies

$$E(\gamma P,s,\chi) = \chi(\gamma)E(P,s,\chi),$$
$$\Delta E(P,s,\chi) = s(n-s)E(P,s,\chi).$$

We see that $E(P,s,\chi)$ is invariant under the action by the lattice Λ and hence it has a Fourier expansion. It is well-known that the constant term in the Fourier expansion has the form $y^s + \phi(s,\chi)y^{n-s}$, where $\phi(s,\chi)$ is called the *scattering matrix*. Its basic properties are well-known, see [6, Ch. 6].

For $\mu,\nu \in \Lambda^\circ$ and $c \in C(\Gamma)$, we define the generalised Kloosterman sum as in [15, Section 4] using the Vahlen model:

$$S(\mu,\nu,c,\chi) := \sum_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty} \overline{\chi}\left(\begin{pmatrix} a & b \\ c & d\end{pmatrix}\right) e\left(\langle ac^{-1},\mu\rangle + \langle dc^{-1},\nu\rangle\right). \tag{5.2}$$

We can rewrite this as

$$S(\mu,\nu,c,\gamma) = \sum_{\substack{\gamma\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty \\ c_\gamma = c}} \overline{\chi(\gamma)} e(\langle\gamma\infty,\mu\rangle + \langle(\gamma^{-1}\infty)^*,\nu\rangle), \tag{5.3}$$

where c_γ is the lower-left entry of γ in the Vahlen model. We now calculate the Fourier expansion of the Eisenstein series explicitly using (higher dimensional) Poisson summation:

$$\begin{aligned}
&E(P,s,\chi)\\
&= [\Gamma_\infty : \Gamma'_\infty]y^s + \sum_{\substack{\gamma\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty \\ c_\gamma\neq 0}} \overline{\chi(\gamma)} \sum_{l\in\Lambda} y(\gamma(x+l,y))^s \\
&= [\Gamma_\infty : \Gamma'_\infty]y^s + \frac{1}{\mathrm{vol}(\Lambda)} \sum_{\gamma\in T_\Gamma} \overline{\chi(\gamma)} \sum_{\mu\in\Lambda^\circ} \left(\int_{\mathbb{R}^n} y(\gamma(t,y))^s e(-\langle t,\mu\rangle)dt\right) e(\langle x,\mu\rangle).
\end{aligned}$$

Now by applying (4.8), we get for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SV}_{n-1}$:

$$\begin{aligned}
&\int_{\mathbb{R}^n} y(\gamma(t,y))^s e(-\langle t,\mu\rangle)dt \\
&= \int_{\mathbb{R}^n} \left(\frac{y}{|ct+d|^2+|c|^2y^2}\right)^s e(-\langle t,\mu\rangle)dt \\
&= \frac{y^s}{|c|^{2s}} \int_{\mathbb{R}^n} \left(\frac{1}{|t+c^{-1}d|^2+y^2}\right)^s e(-\langle t,\mu\rangle)dt \\
&= \frac{y^s}{|c|^{2s}} e(\langle dc^{-1},\mu\rangle) \int_{\mathbb{R}^n} \left(\frac{1}{|t|^2+y^2}\right)^s e(-\langle t,\mu\rangle)dt \\
&= \frac{y^s}{|c|^{2s}} e(\langle dc^{-1},\mu\rangle) \int_{\mathbb{R}^n} \left(\frac{1}{|t|^2+y^2}\right)^s e(-|\mu|t_0)\, dt
\end{aligned}$$

where the last equality follows by applying the orthogonal linear transformation which sends μ to $(|\mu|, 0, \dots, 0)$.

When $\mu = 0$, we obtain

$$\int_{\mathbb{R}^n} \left(\frac{1}{|t|^2+y^2}\right)^s dt = \frac{y^{n-2s}\pi^{n/2}\Gamma\left(s-n/2\right)}{\Gamma(s)},$$

while for $\mu \neq 0$

$$\int_{\mathbb{R}^n} \left(\frac{1}{|t|^2+y^2}\right)^s e(-|\mu|t_0)\, dt = \frac{2\pi^s y^{n/2-s}|\mu|^{s-n/2}}{\Gamma(s)} K_{s-n/2}(2\pi|\mu|y).$$

This follows from [15, p. 678]. Alternatively it follows from combining [49, p. 213] and [20, 6.565.4]. Hence we get

$$\begin{aligned}
E(P,s,\chi) = [\Gamma_\infty : \Gamma'_\infty]y^s + y^{n-s}\frac{\pi^{n/2}\Gamma\left(s-\frac{n}{2}\right)}{\mathrm{vol}(\Lambda)\Gamma(s)}L(s,\chi) \\
+ \frac{2\pi^s y^{n/2}}{\mathrm{vol}(\Lambda)\Gamma(s)} \sum_{\mu\in\Lambda^\circ\setminus\{0\}} L(s,\mu,\chi)|\mu|^{s-n/2}K_{s-n/2}(2\pi|\mu|y),
\end{aligned} \tag{5.4}$$

where

$$L(s,\chi) := \sum_{\gamma\in T_\Gamma} \frac{\overline{\chi}(\gamma)}{|c_\gamma|^{2s}} = \sum_{c\in C(\Gamma)} \frac{S(0,0,c,\chi)}{|c|^{2s}}\ , \tag{5.5}$$

and for $\mu \neq 0$,

$$L(s,\chi,\mu) := \sum_{\gamma\in T_\Gamma} \overline{\chi}(\gamma)\frac{e(\langle d_\gamma c_\gamma^{-1},\mu\rangle)}{|c_\gamma|^{2s}} = \sum_{c\in C(\Gamma)} \frac{S(0,\mu,c,\chi)}{|c|^{2s}}. \tag{5.6}$$

For $\chi = 1$ the trivial character, we just denote $L(s,\mu) := L(s,\mu,1)$. We note that the explicit Fourier expansion we obtain in (5.4) is closely related to [15, Thm. 9.1].

At other cusps $\mathfrak{a} \neq \infty$ of Γ, we will also need some information about the Fourier expansion. For this let $P^{\mathfrak{a}} = (x^{\mathfrak{a}}, y^{\mathfrak{a}}) = \sigma_{\mathfrak{a}}^{-1}P$ denote the coordinates at $\mathfrak{a}$. Then the Fourier expansion at $\mathfrak{a}$ is given by [6, Ch. 6, Prop. 1.42]:

$$E(P^{\mathfrak{a}},s,\chi) = \phi_{\mathfrak{a}}(s)(y^{\mathfrak{a}})^{n-s} + \sum_{\mu\in\Lambda_{\mathfrak{a}}^\circ\setminus\{0\}} \phi_{\mathfrak{a}}(s,\mu)(y^{\mathfrak{a}})^{n-s}K_{s-n/2}(2\pi n|\mu|y^{\mathfrak{a}})e(\langle x^{\mathfrak{a}},\mu\rangle),$$

where $\phi_{\mathfrak{a}}(s,\mu)$ are the Fourier coefficients, which decay rapidly in $|\mu|$ (for s fixed). In particular we observe that $E(P,s,\chi)$ is square integrable when restricted to $\mathcal{F}_{\mathfrak{a}}(Y)$ for $\mathfrak{a}\neq\infty$ (for Y sufficiently large as in (4.12)).

Remark 5.1. By inverting γ in the definition of $L(s,\chi,\mu)$, we observe that

$$\begin{aligned} L(s,\chi,\mu) &= \sum_{\gamma\in T_\Gamma}\overline{\chi}(\gamma)\frac{e(\langle(\gamma^{-1}\infty)^*,\mu\rangle)}{|c_\gamma|^{2s}} \\ &= \sum_{\gamma^{-1}\in T_\Gamma}\chi(\gamma)\frac{e(\langle\gamma\infty,\mu\rangle)}{|c_\gamma|^{2s}} \\ &= \sum_{\gamma\in T_\Gamma}\chi(\gamma)\frac{e(\langle\gamma\infty,\mu\rangle)}{|c_\gamma|^{2s}}. \end{aligned} \tag{5.7}$$

5.1. **Short discussion on spectral properties.** We say that a (measurable) function $f:\mathbb{H}^{n+1}\to\mathbb{C}$ is χ-*automorphic* if it satisfies

$$f(\gamma P)=\chi(\gamma)f(P)\ ,$$

for $P\in\mathbb{H}^{n+1}$ and $\gamma\in\Gamma$.

Denote by $L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)$ the space of square integrable χ-automorphic functions with respect to the hyperbolic metric. For $f,g\in L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)$, we note that $f\overline{g}$ is Γ-invariant. Hence we can define the inner product

$$\langle f,g\rangle := \int_{\mathcal{F}} f\overline{g}\ dv\ .$$

We let $\mathcal{D}(\chi)\subset L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)$ be the subspace consisting of all C^2-functions such that $\Delta f\in L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)$. Then one can see that $-\Delta:\mathcal{D}(\chi)\to L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)$ is a symmetric and nonnegative operator, its spectrum consists of discrete and continuous parts with finitely many discrete points in the interval $[0,n^2/4)$. Let

$$0\le\lambda_0(\chi)\le\lambda_1(\chi)\le\cdots\le\lambda_k(\chi)<n^2/4$$

be the eigenvalues in the interval $[0,n^2/4)$ (see [41] and [6, Ch. 6]). The Eisenstein series $E(z,s,\chi)$ admits meromorphic continuation to $s\in\mathbb{C}$ and satisfies the functional equation

$$E(P,n-s,\chi)=\phi(n-s,\chi)E(P,s,\chi)\ ,$$

where $\phi(s,\chi)$ is the scattering matrix. Moreover, $E(P,s,\chi)$ has poles where $\phi(s,\chi)$ has poles and viceversa. There are finitely many poles in the region $\mathrm{Re}(s)>n/2$, all of them simple and on the real line. If $n/2<\sigma_0\le n$ is a pole of $E(P,s,\chi)$, denote by u_{σ_0} its residue at σ_0. Then

$$u_{\sigma_0}\in L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)\quad\text{and}\quad \Delta u_{\sigma_0}+\sigma_0(n-\sigma_0)u_{\sigma_0}=0\ .$$

For $0\le j\le k$, let $s_j(\chi)\in(n/2,n]$ be such that $s_j(\chi)(n-s_j(\chi))=\lambda_j(\chi)$. We denote by

$$\Omega(\chi):=\{s_0(\chi),\ldots,s_k(\chi)\}.$$

Then the poles of $E(P,s,\chi)$ in $\operatorname{Re} s > n/2$ form a subset of $\Omega(\chi)$. Moreover, we can see from [6, Ch 6, p. 37] that for χ trivial, we have

$$\operatorname{Res}_{s=n} E(P,s) = \frac{[\Gamma_\infty : \Gamma'_\infty]\mathrm{vol}(\Lambda)}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})}\ . \tag{5.8}$$

5.2. **Key lemmas.** In this section we will prove certain key analytic lemmas that we will need in the proofs of our theorems. First of all we will show that we can only have $\lambda_0(\chi)=0$ when χ is trivial. Secondly we obtain meromorphic continuation of the Fourier coefficients of the twisted Eisenstein series, which will serve as generating series for our distribution problems. Finally we will prove a bound on vertical lines for these generating series.

The most conceptual way to see the first claim above is probably to use Green's identity

$$\int_{\mathcal{F}} (-\Delta u)u dv = \int_{\mathcal{F}} \nabla u.\nabla u\ dv + \int_{\partial\mathcal{F}} u(\nabla u.\mathbf{n})\mathbf{dS}.$$

If we have $\Delta u = 0$, then the first integral is 0. The third integral should vanish since contributions from "opposing sides" in the boundary of the fundamental domain should cancel each other. This would force the second integral to be 0, which means u is constant. This argument works in principle, but for example in [16, Theorem 4.1.7] they spend several pages making it rigorous. Instead we will give an argument using the Fourier expansion and the mean value theorem for harmonic functions.

Lemma 5.2. *We have that $\lambda_0(\chi)=0$ if and only if χ is trivial.*

Proof. Suppose $\lambda_0(\chi) = 0$ and let u be a corresponding eigenvector, i.e. $u \in L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi)$ and $\Delta u = 0$. Then we can consider the Fourier expansion of u at a cusp $\mathfrak{a}$ of Γ. We know from [6, Ch. 6, p.10] that the Fourier expansion of u takes the form

$$c_{1,\mathfrak{a}} + c_{2,\mathfrak{a}}(y^{\mathfrak{a}})^n + \sum_{\mu\in\Lambda_{\mathfrak{a}}^\circ\backslash\{0\}} a_{u,\mathfrak{a}}(\mu)(y^{\mathfrak{a}})^{n/2}K_{n/2}(2\pi n|\mu|y)e(\langle x,\mu\rangle).$$

From the rapid decay of the K-Bessel function we see that if $c_{2,\mathfrak{a}} \neq 0$, then u behaves like $(y^{\mathfrak{a}})^n$ close enough to $\mathfrak{a}$ and thus $\int_{F_{\mathfrak{a}}(Y)} |u(x,y)|^2 dxdy$ is divergent contradicting the fact that u is square integrable. Thus $c_{2,\mathfrak{a}} = 0$ and we conclude again using the rapid decay of the K-Bessel functions that u is bounded on $F_{\mathfrak{a}}(Y)$. Since $\mathfrak{a}$ was an arbitrary cusp we conclude that u is bounded on all of $\mathcal{F}$. Thus since χ is unitary, we conclude that u is bounded on all of $\mathbb{H}^{n+1}$. Now it follows from the *Mean Value Theorem for Harmonic Functions on* $\mathbb{H}^{n+1}$ that u is constant. By definition, $u(\gamma P) = \chi(\gamma)u(P)$, for all $\gamma\in\Gamma$ and $P\in\mathbb{H}^{n+1}$. Thus we conclude that χ is the trivial character.

Therefore, if χ is trivial the unique eigenfunction of eigenvalue 0 is the constant one, and for χ non-trivial there are no eigenfunctions of eigenvalue 0. This finishes the proof. □

We now obtain meromorphic continuation of the Fourier coefficients of the Eisenstein series and crucial information about the location of the poles.

Proposition 5.3. *The Dirichlet series $L(s,\mu,\chi)$ admits meromorphic continuation to the entire complex plane. The possible poles in the half-plane $\operatorname{Re} s > n/2$ are contained in $\Omega(\chi)$. Furthermore, there is a pole at $s = n$ exactly if χ is trivial and $\mu = 0$. In this case the residue is equal to*

$$\frac{[\Gamma_\infty : \Gamma'_\infty]\Gamma(n)\mathrm{vol}(\Lambda)^2}{\pi^{n/2}\Gamma\left(\frac{n}{2}\right)\mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})}.$$

Proof. From (5.4), we know that for $\mu \in \Lambda^\circ\backslash\{0\}$

$$L(s,\mu,\chi) = \frac{\Gamma(s)}{2\pi^s y^{n/2}|\mu|^{s-n/2}K_{s-n/2}(2\pi|\mu|y)}\int_{\mathcal{P}} E((x,y),s,\chi)e(-\langle x,\mu\rangle)dx,$$

and

$$L(s,\chi) = \frac{y^{s-n}\Gamma(s)}{\pi^{n/2}\Gamma\left(s-\frac{n}{2}\right)}\left(\int_{\mathcal{P}} E((x,y),s,\chi)dx - [\Gamma_\infty : \Gamma'_\infty]y^s\right),$$

where $\mathcal{P}$ is a fundamental parallelogram for Λ. Now for $y > 0$ fixed , the Bessel function $K_s(y)$ defines an analytic function in s, which is non-zero for some y large enough. Similarly the Gamma function define a meromorphic function. Thus we get the meromorphic continuation of $L(s,\mu,\chi)$ from that of the Eisenstein series. We also note that in the half-plane $\operatorname{Re} s > n/2$, $L(s,\mu,\chi)$ has possible poles only where $E(P,s,\chi)$ has poles, i.e. the poles are contained in $\Omega(\chi)$. By Lemma 5.2, we see that $L(s,\mu,\chi)$ is regular at $s = n$ unless χ is trivial.

If χ is trivial, we see that $L(s,\mu)$ with $\mu \neq 0$ is regular at $s = n$, since the pole of the Eisenstein series is constant. For $\mu = 0$ the residue is given by

$$\operatorname{Res}_{s=n} L(s,0) = \frac{\Gamma(n)}{\pi^{n/2}\Gamma\left(\frac{n}{2}\right)}\int_{\mathcal{P}}\frac{[\Gamma_\infty : \Gamma'_\infty]\mathrm{vol}(\Lambda)}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})}dx = \frac{[\Gamma_\infty : \Gamma'_\infty]\Gamma(n)\mathrm{vol}(\Lambda)^2}{\pi^{n/2}\Gamma\left(\frac{n}{2}\right)\mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})},$$

as wanted. □

In order to obtain bounds on vertical lines for our generating series, we will employ an idea due to Colin de Verdière [7], which employs the analytic properties of resolvent operators. Alternatively one could use Poincaré series for $\mu \neq 0$ and Maaß–Selberg for $\mu = 0$ as is done in [38] and [8]. In the end the two methods are essentially equivalent.

Let $h : \mathbb{R}^+ \to \mathbb{R}^+$ be a smooth function which is equal to $[\Gamma_\infty : \Gamma'_\infty]$ for $y > Y + 1$ and 0 for $y < Y$, where Y is as in (4.12). Then for $\operatorname{Re}(s) > n/2$ we define a χ-automorphic function on $\mathbb{H}^{n+1}$ by $P \mapsto h(y)y^s$ for $P \in \mathcal{F}$ and extended periodically (twisted accordingly by χ). Then from the above mentioned results on the Fourier expansions of the Eisenstein series at the different cusps, we see that

$$g(P,s,\chi) := E(P,s,\chi) - h(y)y^s \in L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi),$$

which satisfies for $z \in \mathcal{F}$

$$(\Delta - s(n-s))g(P,s,\chi) = -(\Delta - s(n-s))h(y)y^s = h''(y)y^{s+2} + (2s-n+1)h'(y)y^{s+1}.$$

We observe that the right hand side above is compactly supported with L^2-norm bounded by $O(|s|+1)$ for $n/2 + \varepsilon < \operatorname{Re} s < n+2$. Now we put

$$H(P,s,\chi) := R(s,\chi)(h''(y)y^{s+2} + (2s-n+1)h'(y)y^{s+1}) \in L^2(\Gamma\backslash\mathbb{H}^{n+1},\chi),$$

where $R(s,\chi) = (\Delta - s(n-s))^{-1}$ denotes the resolvent operator associated to Δ. By a general bound for the operator norm of resolvent operators [23, Lemma A.4], we conclude that

$$||H(\cdot,s,\chi)||_{L^2} \ll_\varepsilon 1,$$

when s is bounded at least ε away from the spectrum of Δ. We can now write

$$E(P,s,\chi) = H(P,s,\chi) + h(y)y^s, P \in \mathcal{F} \tag{5.9}$$

where we have good control on the L^2-norm of $H(P,s,\chi)$. We will now use this to obtain bounds on vertical lines for the Fourier coefficients of $E(P,s,\chi)$. We mimic [35, Section 4.4].

Proposition 5.4. *Let $\mu \in \Lambda^\circ$. Then we have*

$$L(s,\mu,\chi) \ll_{\varepsilon,\mu} (|s|+1)^{n/2},$$

for $n/2+\varepsilon < \operatorname{Re} s < n+2$ and s bounded at least ε away from the spectrum of Δ.

Proof. We have

(5.10)

$$L(s,\mu,\chi) = \int_{\mathcal{P}} f_s(y,\mu)E((x,y),s,\chi)e(-\langle x,\mu\rangle)dx - \mathbf{1}_{\mu=0}[\Gamma_\infty : \Gamma'_\infty]y^s f_s(y,\mu),$$

where

$$f_s(y,\mu) = \begin{cases} \frac{\Gamma(s)}{2\pi^s y^{n/2}|\mu|^{s-n/2}K_{s-n/2}(2\pi n|\mu|y)}, & \mu \neq 0, \\ \frac{\Gamma(s)}{y^{n-s}\pi^{n/2}\Gamma(s-n/2)}, & \mu = 0. \end{cases}$$

The idea is now to bound the right hand side of (5.10) using (5.9). In order to bring the information we have about $H(P,s,\chi)$ into play, we need to make an extra integration over y. So let Y be a fixed quantity such that $\{(x,y) \mid x \in \mathcal{P}, y > Y\} \subset \mathcal{F}$, then we see that

$$\begin{aligned} &\int_Y^{Y+1}\int_{\mathcal{P}} f_s(y,\mu)E((x,y),s,\chi)e(-\langle \mu,x\rangle)dxdy \\ &= \int_Y^{Y+1}\int_{\mathcal{P}} f_s(y,\mu)H((x,y),s,\chi)e(-\langle \mu,x\rangle)dxdy \\ &+ \int_Y^{Y+1}\int_{\mathcal{P}} f_s(y,\mu)h(y)y^s e(-\langle \mu,x\rangle)dxdy \end{aligned}$$

Now we observe that by Cauchy–Schwarz we have

$$\begin{aligned} &\int_Y^{Y+1}\int_{\mathcal{P}} f_s(y,\mu)H((x,y),s,\chi)e(-\langle \mu,x\rangle)dxdy \\ &\leq \left(\int_Y^{Y+1}\int_{\mathcal{P}} |H((x,y),s,\chi)|^2 dxdy\right)^{1/2} \left(\int_Y^{Y+1}\int_{\mathcal{P}} |f_s(y,\mu)|^2 dxdy\right)^{1/2} \\ &\ll ||H(\cdot,s,\chi)||_{L^2} \left(\int_Y^{Y+1} |f_s(y,\mu)|^2 dy\right)^{1/2}, \end{aligned}$$

where we use that $\{(x,y) \mid x \in \mathcal{P}, y > Y\} \subset \mathcal{F}$. To finish the proof we need an upper bound for $f_s(y,\mu)$.

For $\mu = 0$ we get by Stirling's approximation the upper bound

$$f_s(y,0) \ll_\varepsilon y^{n-\sigma}(|s|+1)^{n/2},$$

for $s = \sigma + it$ with $n/2 + \varepsilon < \sigma < n+2$.

For $\mu \neq 0$, we use the Fourier expansion for the K-Bessel function (coming from combining [23, (B.32)] and [23, (B.34)]) to obtain a good approximation. By applying Stirling's approximation, this gives for $s = \sigma + it$ with $t \gg 1$

$$\begin{aligned} K_{s-n/2}(2\pi|\mu|y) &= \frac{\pi^{1/2}t^{\sigma-n/2-1/2}e^{\pi t/2}\left(\frac{t}{e}\right)^{it}}{2\sqrt{2}\sin(\pi(s-n/2))}(\pi|\mu|y)^{-s+n/2}\left(1+O_{\mu,y}(t^{-1})\right) \\ &\gg_{\mu,y} e^{-\pi t/2}t^{\sigma-n/2-1/2}, \end{aligned}$$

where the implied constants depend continuously on y. From this we conclude that when $y \in (Y, Y+1)$, we have

$$f_s(y,\mu) \ll_\mu (1+|s|)^{n/2}.$$

Inserting this and using the bound $||H(\cdot,s,\chi)||_{L^2} \ll_\varepsilon 1$, we conclude that

$$L(s,\mu,\chi) \ll_{\varepsilon,\mu} (|s|+1)^{n/2},$$

for s bounded ε away from the spectrum of Δ, as wanted. □

6. Proof of Theorem 1.5

In this section we will use the analytic properties of twisted Eisenstein series proved in the previous section to proof our main results. First of all we deduce the following result using a standard complex analysis argument.

Proposition 6.1. *Let χ be a unitary character of Γ trivial on Γ'_∞ and $\mu \in \Lambda^\circ$. Then there exists a constant $\nu(\chi) > 0$ such that*

$$\sum_{\gamma \in T_\Gamma(X)} \chi(\gamma) e\left(\langle \gamma\infty, \mu\rangle\right) = \frac{X^{s_0(\chi)}}{s_0(\chi)}\left(\mathrm{Res}_{s=s_0(\chi)} L(s,\chi,\mu) + O_{\chi,\mu}(X^{-\nu(\chi)})\right).$$

Proof. Let $\phi_U : \mathbb{R} \to \mathbb{R}$ be a family of smooth non-increasing functions with

$$\phi_U(t) = \begin{cases} 1 & \text{if } t \leq 1 - 1/U, \\ 0 & \text{if } t \geq 1 + 1/U \end{cases} \tag{6.1}$$

and $\phi_U^{(j)}(t) = O(U^j)$ as $U \to \infty$. For $\mathrm{Re}(s) > 0$, we consider the Mellin transform

$$R_U(s) = \int_0^\infty \phi_U(t) t^s \frac{dt}{t}\ . \tag{6.2}$$

We can easily see that

$$R_U(s) = \frac{1}{s} + O\left(\frac{1}{U}\right) \quad \text{as } U \to \infty \tag{6.3}$$

and for any $N > 0$,

$$R_U(s) = O\left(\frac{1}{|s|}\left(\frac{U}{1+|s|}\right)^N\right) \quad \text{as } |s| \to \infty\ , \tag{6.4}$$

where the last estimate follows from repeated partial integration. Now we use Mellin inversion and (5.7) to obtain

$$\begin{aligned}&\sum_{\gamma\in T_\Gamma}\chi(\gamma)e\left(\langle\gamma\infty,\mu\rangle\right)\ \phi_U\left(\frac{|c|^2}{X}\right)\\&=\sum_{\gamma\in T_\Gamma}\chi(\gamma)e\left(\langle\gamma\infty,\mu\rangle\right)\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=n+1}\frac{X^s}{|c|^{2s}}R_U(s)ds\\&=\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=n+1}L(s,\chi,\mu)X^sR_U(s)ds.\end{aligned}$$

Next, we recall Proposition 5.4 and equation (6.4) to deduce that the last integral is absolutely convergent. We want to move the line of integration to $\mathrm{Re}(s)=h=h(\chi)$ for some $h>n/2$ such that $s_1(\chi)<h(\chi)<s_0(\chi)$. We use the fact that we have polynomial growth on vertical lines for $L(s,\chi,\mu)$ guaranteed by Lemma 5.4 and that $L(s,\chi,\mu)$ has only a possible pole at $s_0(\chi)$ in the region $\mathrm{Re}(s)>h(\chi)$. We conclude that

$$\begin{aligned}\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=n+1}L(s,\chi,\mu)X^sR_U(s)ds=&\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=h}L(s,\chi,\mu)X^sR_U(s)ds\\&+\mathrm{Res}_{s=s_0(\chi)}\left(L(s,\chi,\mu)X^sR_U(s)\right)\ .\end{aligned}$$

Setting $N=(n+1)/2$ in (6.4), we observe from Proposition 5.4 that

$$\int_{\mathrm{Re}(s)=h}L(s,\chi,\mu)X^sR_U(s)ds\ll X^hU^{(n+1)/2}\ . \tag{6.5}$$

Now, (6.3) gives us

$$\mathrm{Res}_{s=s_0(\chi)}\left(L(s,\chi,\mu)X^sR_U(s)\right)=\frac{X^{s_0(\chi)}}{s_0(\chi)}\mathrm{Res}_{s=s_0(\chi)}L(s,\chi,\mu)\left(1+O\left(\frac{1}{U}\right)\right)\ . \tag{6.6}$$

Since we want this to be the main contribution, we choose $U=X^{a(\chi)}$, where $a(\chi):=(s_0(\chi)-h(\chi))/(n+1)$.

Now if χ is the trivial character and $\mu=0$, we obtain

$$\sum_{\gamma\in T_\Gamma}\phi_U\left(\frac{|c|^2}{X}\right)=\frac{X^n}{n}\left(\mathrm{Res}_{s=n}L(s)+O(X^{-\delta})\right),$$

for some fixed $\delta>0$. We now choose ϕ_U^1 and ϕ_U^2 as in (6.1) with the further requirements that $\phi_U^1(t)=0$ for $t\geq 1$ and $\phi_U^2(t)=1$ for $0\leq t\leq 1$. Then

$$\sum_{\gamma\in T_\Gamma}\phi_U^1\left(\frac{|c|^2}{X}\right)\leq\sum_{\gamma\in T_\Gamma(X)}1\leq\sum_{\gamma\in T_\Gamma}\phi_U^2\left(\frac{|c|^2}{X}\right)\ ,$$

so the previous two equations and Proposition 5.3 give us

$$\#T_\Gamma(X)=\frac{X^{2n}}{n}\left(\frac{[\Gamma_\infty:\Gamma'_\infty]\mathrm{vol}(\Lambda)^2\Gamma(n)}{\pi^{n/2}\mathrm{vol}(\Gamma)\Gamma(n/2)}+O(X^{-\delta})\right)\ . \tag{6.7}$$

Now we return to (6.6). Indeed,

(6.8)

$$\sum_{\gamma\in T_\Gamma}\chi(\gamma)e\left(\langle\gamma\infty,\mu\rangle\right)\phi_U\left(\frac{|c|^2}{X}\right)=\frac{X^{s_0(\chi)}}{s_0(\chi)}\left(\mathrm{Res}_{s=s_0(\chi)}L(s,\chi,\mu)+O(X^{-a(\chi)})\right).$$

Also, from the definition of ϕ_U,

$$\sum_{\gamma\in T_\Gamma}\chi(\gamma)e\left(\langle\gamma\infty,\mu\rangle\right)\phi_U\left(\frac{|c|^2}{X}\right)$$
$$=\sum_{\gamma\in T_\Gamma(\sqrt{X})}\chi(\gamma)e\left(\langle\gamma\infty,\mu\rangle\right)+O\left(\#\left\{\gamma\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty\ :\ 1-\frac{1}{U}\le\frac{|c|^2}{X}\le1+\frac{1}{U}\right\}\right).$$

But now we use (6.7) to bound the size of the error term

$$\begin{aligned}&\#\left\{\gamma\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty\ :\ 1-\frac{1}{U}\le\frac{|c|^2}{X}\le1+\frac{1}{U}\right\}\\&=T_\Gamma\left(\sqrt{X\left(1+\frac{1}{U}\right)}\right)-T_\Gamma\left(\sqrt{X\left(1-\frac{1}{U}\right)}\right)\\&=O\left(X^{n-\nu}\right),\end{aligned}$$

for some $\nu(\chi)>0$. The conclusion follows. □

Remark 6.2. As a consequence of Proposition 6.1, we conclude that for all unitary characters χ as above there exist $\nu(\chi)>0$ such that

$$\sum_{\gamma\in T_\Gamma(X)}\chi(\gamma)e\left(\langle\gamma\infty,\mu\rangle\right)=\mathbf{1}_{\chi,\mu}\frac{\mathrm{vol}(\Lambda)^2\Gamma(n)}{n\pi^{n/2}\mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})\Gamma(n/2)}X^{2n}+O_\chi(X^{2n-\nu(\chi)}),$$

where $\mathbf{1}_{\chi,\mu}$ is 1 if $\mu=0$ and χ is trivial and 0 otherwise.

6.1. **Applications to equidistribution.** Using the the above proposition we are now ready to finish the proof of Theorem 1.5 and from this deduce Theorems 1.2 and 1.3.

We recall the setup from the definition. Consider the cohomology group $H^1_{\Gamma'_\infty}(\Gamma,\mathbb{R}/\mathbb{Z})$ (see Section 8 for details), which can be identified with the set of unitary characters of Γ trivial on Γ'_∞.

Definition 6.3. *We say that* $\omega_1,\dots,\omega_d\in H^1_{\Gamma'_\infty}(\Gamma,\mathbb{R}/\mathbb{Z})$ *are in* **general position** *if for any* $(l_1,\dots,l_d)\in\mathbb{Z}^d$*, we have*

$$n_1\omega_1+\ldots+n_d\omega_d=0\in H^1_{\Gamma'_\infty}(\Gamma,\mathbb{R}/\mathbb{Z})\Leftrightarrow\left(n_i\omega_i=0\in H^1_{\Gamma'_\infty}(\Gamma,\mathbb{R}/\mathbb{Z}),\forall i=1,\dots,d\right).$$

As an example one can pick $\omega_1,\dots,\omega_d$ to be a basis for the non-torsion part of $H^1_{\Gamma'_\infty}(\Gamma,\mathbb{R}/\mathbb{Z})$. Also we could pick a $\mathbb{F}_p$-basis for $H^1_{\Gamma'_\infty}(\Gamma,\mathbb{F}_p)$, where we consider $\mathbb{F}_p\subset\mathbb{R}/\mathbb{Z}$ via $\mathbb{F}_p\ni a\mapsto a/p$.

Observe that the image of ω_i is an additive subgroup of $\mathbb{R}/\mathbb{Z}$ and thus is either dense in $\mathbb{R}/\mathbb{Z}$ or finite. In the first case we put $J_i=\mathbb{R}/\mathbb{Z}$ and in the latter case we put $J_i=\mathbb{Z}/m_i\mathbb{Z}$ where m_i is the cardinality of the image of ω_i. That is, J_i is the closure of the image of ω_i. We equip $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$ with respectively the Lebesque measure and the uniform probability measure.

Let $\omega_1, \ldots, \omega_d \in H^1_{\Gamma'_\infty}(\Gamma_0(N), \mathbb{R}/\mathbb{Z})$ be in general position. Then for any tuple $\underline{l} = (l_1, \ldots, l_d) \in \mathbb{Z}^d$ such that $l_i\omega_i \neq 0 \in H^1_{\Gamma'_\infty}(\Gamma_0(N), \mathbb{R}/\mathbb{Z})$ for all $i = 1, \ldots, d$, we get a non-trivial element of $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ defined by

$$\omega_{\underline{l}} := l_1\omega_1 + \ldots + l_d\omega_d.$$

Now we consider the associated non-trivial unitary character $\chi_{\underline{l}} : \Gamma \to \mathbb{C}^\times$;

$$\chi_{\underline{l}}(\gamma) := e\left(\omega_{\underline{l}}(\gamma)\right),$$

where $e(x) = e^{2\pi i x}$. Observe that this is indeed well-defined and that we get an induced map $\chi_{\underline{l}} : \Gamma'_\infty \backslash \Gamma / \Gamma'_\infty \to \mathbb{C}^\times$ since $\omega_{\underline{l}}$ is trivial on Γ'_∞.

By *Weyl's Criterion* [24, p. 487] in order to conclude equidistribution of the values of

$$\omega(\gamma) := (\omega_1(\gamma), \ldots, \omega_d(\gamma), \gamma\infty)$$

inside $\prod_{i=1}^d J_i \times (\mathbb{R}^n/\Lambda)$, we have to show cancelation in the corresponding Weyl sums. These are exactly given by:

$$\sum_{\gamma \in T_\Gamma(X)} \chi_{\underline{l}}(\gamma) e(\langle \gamma\infty, \mu \rangle),$$

where $\underline{l} \in \mathbb{Z}^d$ and $\mu \in \Lambda^\circ$. We see that it follows from combining Proposition 6.1 and Remark 5.1 that we have

$$\sum_{\gamma \in T_\Gamma(X)} \chi_{\underline{l}}(\gamma) e(\langle \gamma\infty, \mu \rangle) = o\left(\sum_{\gamma \in T_\Gamma(X)} 1\right),$$

as $X \to \infty$ *unless* $\mu = 0$ and $\chi_{\underline{l}}$ is trivial. This finishes the proof of Theorem 1.5 using Weyl's Criterion.

Now let us see how Theorem 1.2 and 1.3 follow from Theorem 1.5. We restrict to $n = 1$ and $\Gamma = \Gamma_0(N)$. By the mod p-version of the Eichler–Shimura isomorphism (3.1), we see that $\mathfrak{m}_f^\pm$ with $f \in \mathcal{S}_2(\Gamma_0(N))$ gives a basis for $H^1_P(\Gamma_0(N), \mathbb{F}_p)$, and thus it follows directly that they are in general position.

Similarly, by Eichler–Shimura, we know that the cohomology classes associated to $\mathrm{Re}\, f(z)dz$ and $\mathrm{Im}\, f(z)dz$ are in general position, where $f \in \mathcal{S}_2(\Gamma_0(N))$ ranges over Hecke newforms. From a classical result of Schneider [42] we know that the Néron periods $\Omega_\pm$ are transcendental. Using the rationality (1.2), this implies that the cohomology class associated to a newform f given by

$$\Gamma_0(N) \ni \gamma \mapsto \int_{\gamma\infty}^{\infty} \mathrm{Re}(f(z)dz)$$

takes some irrational value (and similarly for $\mathrm{Im}(f(z)dz)$).

Thus we see that in these two cases Theorem 1.5 reduces to Theorem 1.2 and 1.3.

7. Proof of Theorem 1.8

We now give a proof of Theorem 1.8 showing equidistriburion of the values of cohomology classes when ordered by the lengths of the geodesics corresponding to conjugacy classes of Γ. This will be an almost direct consequence of a twisted trace formula for $\mathrm{SO}(n+1,1)$. Our method is in the spirit of [37], where Petridis–Risager show that for cocompact subgroups of $\mathrm{SL}_2(\mathbb{R})$ the values of modular symbols are asymptotically normally distributed when ordered by the length of the corresponding geodesics. This was in turn inspired by ideas of Phillips and Sarnak [39].

We firstly consider the case $n = 1$. If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{R})$ is hyperbolic, then γ is conjugate in $\mathrm{SL}_2(\mathbb{R})$ to a unique element $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{smallmatrix}\right)$ with $\lambda > 1$. Let Γ be a discrete, cofinite subgroup of $\mathrm{SL}_2(\mathbb{R})$. We know that for each hyperbolic conjugacy class $\{\gamma\} \in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma)$ there is a corresponding geodesic of length $l(\gamma) = \log \lambda^2$. It is a consequence of the twisted trace formula for Γ that for any unitary character χ of Γ, we have

$$\sum_{\substack{\{\gamma_0\}\text{ primitive} \\ l(\gamma_0)\leq X}} \chi(\gamma_0) = \sum_{s\in\Omega(\chi)} \mathrm{li}(e^{sX}) + O_\chi(e^{\frac{3}{4}X}),$$

where $\mathrm{li}(x) = \int_2^x t^{-1}dt$ is the logarithmic integral (see [22, p. 475]). Hence we obtain

$$\sum_{\substack{\{\gamma\}\in\mathrm{Conj}_{\mathrm{hyp}}(\Gamma) \\ l(\gamma)\leq X}} \chi(\gamma) \sim \sum_{\substack{\{\gamma_0\}\text{ primitive} \\ l(\gamma_0)\leq X}} \chi(\gamma_0) \sim \mathrm{li}(e^{s_0(\chi)X})$$

where the first sum is over all hyperbolic classes. Therefore, using Lemma 5.2, we obtain that for some $\nu(\chi) > 0$,

$$\frac{1}{|\{\{\gamma\}\in\mathrm{Conj}_{\mathrm{hyp}}(\Gamma) : 0 < l(\gamma) \leq X\}|} \sum_{\substack{\{\gamma\}\in\mathrm{Conj}_{\mathrm{hyp}}(\Gamma) \\ l(\gamma)\leq X}} \chi(\gamma) = \mathbf{1}_\chi + O(e^{-\nu(\chi)X}),$$

where $\mathbf{1}_\chi$ is 1 if χ is trivial and 0 otherwise. Now the proof follows using the Weyl criterion as in Section 6.1.

We now discuss the general case n. As mentioned eariler, the first proof of the Prime Geodesic Theorem in the general case was given by Gangolli and Warner [18]. The trace formula for cofinite subgroups of $\mathrm{SO}(n+1,1)$ acting on $\mathbb{H}^{n+1}$ was developed by Cohen and Sarnak in [6, Ch. 7]. As a consequence, they obtain the following stronger version of Prime Geodesic Theorem for $\mathbb{H}^{n+1}$ [6, Thm. 7.37]:

$$\pi_\Gamma(X) = \sum_{n/2<s_j\leq n} \mathrm{li}(e^{s_jX}) + O\left(e^{(n-\frac{n}{n+2})X}\right)$$

where the sum is taken over all $n/2 \leq s_j \leq n$ such that $s_j(n-s_j)$ is an eigenvalue of $-\Delta$ acting on $L^2(\Gamma\backslash\mathbb{H}^{n+1})$. Now we would like to apply a trace formula where we allow twists by characters. We did not find a place in literature where it is written down explicitly, and to keep the exposition simple we wil leave out the details. The analysis should be similar to the case $n = 1$ and is furthermore implied to hold by Sarnak in [41, p. 6]. Similarly Phillips and Sarnak [39] prove a theorem about distribution of geodesics in homology classes for quotients of $\mathbb{H}^{n+1}$, but only treat the case $n = 1$

in detail. The twisted trace formula for $\mathbb{H}^{n+1}$ that we need is exactly the same one which is implicit [39].

As in the 2 dimensional case, we would get

$$\sum_{\substack{\{\gamma\}\in \mathrm{Conj}_{\mathrm{hyp}}(\Gamma)\\ l(\gamma)\leq X}} \chi(\gamma) \sim \mathrm{li}(e^{s_0(\chi)X})$$

from which Theorem 1.8 follows by Weyl's Criterion as above.

8. On the size of certain cohomology groups

In this paper we study the distribution of certain cohomology classes which can be identified with the unitary characters of cofinite subgroups $\Gamma < \mathrm{SO}(n+1,1)$ (or equivalently $\Gamma < \mathrm{SV}_{n-1}$) with cusps. It is now a natural question to ask how many unitary characters our results actually apply to. This amounts to finding the dimensions of the relevant space of unitary characters or equivalently of certain cohomology groups. This last perspective is most useful when comparing it to the existing literature. First of all we will define the cohomology groups that are relevant and then survey what is known about their size.

8.1. **The first cohomology group.** We refer to [45, Chapter 8] for a comprehensive account. The *first cohomology group* of Γ with coefficients in a $\mathbb{Z}[\Gamma]$-module A is defined as the quotient between the corresponding *coboundaries* and *cocycles*;

$$H^1(\Gamma, A) := Z^1(\Gamma, A)/B^1(\Gamma, A),$$

where

$$Z^1(\Gamma, A) := \{\omega : \Gamma \to A \mid \omega(\gamma_1\gamma_2) = \omega(\gamma_1) + \gamma_1.\omega(\gamma_2), \forall \gamma_1, \gamma_2 \in \Gamma\}$$

and

$$B^1(\Gamma, A) := \{\omega : \Gamma \to A \mid \exists a \in A : \omega(\gamma) = \gamma.a - a, \forall \gamma \in \Gamma\}.$$

Furthermore given a subset $P \subset \Gamma$, we will be studying the first P-cohomology group of Γ with coefficients in A defined by;

$$H^1_P(\Gamma, A) := \{\omega \in H^1(\Gamma, A) \mid \omega(p) \in (p-1)A, \forall p \in P\}.$$

We will in particular study the distribution of P-cohomology group in the case where $P = \Gamma'_\infty$ is the set of parabolic elements of Γ fixing ∞ and A is given by the circle $\mathbb{R}/\mathbb{Z}$ equipped with the trivial Γ-action. In this case $H^1_P(\Gamma, \mathbb{R}/\mathbb{Z})$ computes exactly the unitary characters of Γ trivial on Γ'_∞. Now we will make some general comments on the structure and size of $H^1_P(\Gamma, \mathbb{R}/\mathbb{Z})$.

8.2. **On the structure of the cohomology groups.** We recall that for A a trivial Γ module we have

$$H^1(\Gamma, A) \cong \mathrm{Hom}_{\mathbb{Z}}(\Gamma/[\Gamma, \Gamma], A),$$

which is a special case of the *Universal Coefficients Theorem* since $H_1(\Gamma, \mathbb{Z}) \cong \Gamma/[\Gamma, \Gamma]$. From this we see that $H^1(\Gamma, \mathbb{R}/\mathbb{Z})$ can be identified with the unitary characters of Γ. It is known [44, p. 484] that Γ is finitely represented and thus $\Gamma/[\Gamma, \Gamma]$ is a finitely generated abelian group. From this we see that we have a splitting of the cohomology group $H^1(\Gamma, \mathbb{R}/\mathbb{Z})$ in a free part and a torsion part;

$$H^1(\Gamma, \mathbb{R}/\mathbb{Z}) \cong H^1_{\text{free}}(\Gamma, \mathbb{R}/\mathbb{Z}) \oplus H^1_{\text{tor}}(\Gamma, \mathbb{R}/\mathbb{Z}),$$

where the $\mathbb{R}/\mathbb{Z}$ rank of $H^1_{\text{free}}(\Gamma, \mathbb{R}/\mathbb{Z})$ is the same as the dimension of $H^1(\Gamma, \mathbb{R})$ and the size of $H^1_{\text{tor}}(\Gamma, \mathbb{R}/\mathbb{Z})$ is equal to the size of the torsion in $H_1(\Gamma, \mathbb{Z}) \cong \Gamma/[\Gamma, \Gamma]$.

We have a further Eichler–Shimura splitting of the free part due to Harder [21];

$$H^1(\Gamma, \mathbb{R}) \cong H^1_{\text{cusp}}(\Gamma, \mathbb{R}) \oplus H^1_{\text{Eis}}(\Gamma, \mathbb{R}), \tag{8.1}$$

where $H^1_{\text{cusp}}(\Gamma, \mathbb{R})$ is the cuspidal part corresponding to certain automorphic forms for Γ (as we will see shortly) and $H^1_{\text{Eis}}(\Gamma, \mathbb{R})$ is the (remaining) Eisenstein part, which can be canonically defined. The cuspidal part $H^1_{\text{cusp}}(\Gamma, \mathbb{R})$ can be identified with $H^1_P(\Gamma, \mathbb{R})$ where P is the set of all parabolic elements of Γ and furthermore all of the above splittings are compatible with the Hecke action, when Γ is arithmetic.

There has been a lot of work recently on the study of the size of respectively $H^1_{\text{cusp}}(\Gamma, \mathbb{R})$, $H^1_{\text{Eis}}(\Gamma, \mathbb{R})$ and $H^1_{\text{tor}}(\Gamma, \mathbb{R}/\mathbb{Z})$, and we will now collect the relevant results for our problem. We observe that the image of Γ'_∞ in $\Gamma/[\Gamma, \Gamma]$ is either trivial, finite or isomorphic to $\mathbb{Z}$. Thus we conclude that $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ is non-trivial as soon as, say $H^1(\Gamma, \mathbb{R}/\mathbb{Z})$ is not generated by a single element or $H^1(\Gamma, \mathbb{R})$ is non-trivial.

8.3. **The dimension of cohomology groups.** It is a result of Kazhdan [25] that for discrete, cofinite subgroups of real Lie groups of rank larger than 1, the abelianization is always torsion. In our case, since $\mathrm{SO}(n+1, 1)$ is of rank one, we can however hope to see some free part. In the case of cofinite subgroups $\Gamma \subset \mathrm{SO}(n+1, 1)$, the dimension of $H^1(\Gamma, \mathbb{R})$ (or equivalently the free part of $\Gamma/[\Gamma, \Gamma]$) is not very well understood for arbitrary n. The best lower bounds of the rank available in the literature seem to be what follows from the work of Millson [34] and Lubotzky [28], which gives that any arithmetic subgroup Γ (with a few restrictions when $n = 3, 7$) contains a subgroup such that the dimension of $H^1(\Gamma, \mathbb{R})$ is at least one. In certain arithmetic situations, we will be able to say more using a connection to automorphic forms.

8.3.1. *Cohomology classes associated to automorphic forms.* Recall the splitting (8.1) due to Harder of the cohomology into a cuspidal and an Eisenstein part. We give a brief overview of the description of $H^1_{\text{cusp}}(\Gamma, \mathbb{R})$ in terms of automorphic forms, as in [41]. We recall the canonical isomorphism between $H^1(\Gamma, \mathbb{R})$ and the de Rham cohomology group $H^1_{\text{dR}}(\Gamma\backslash\mathbb{H}^{n+1}, \mathbb{R})$ consisting of 1-forms. Inside $H^1_{\text{dR}}(\Gamma\backslash\mathbb{H}^{n+1}, \mathbb{R})$ we define the subset of cuspidal harmonic 1-forms.

Definition 8.1. *A harmonic 1-form $\alpha = f_0 dx_0 + f_1 dx_1 + \cdots + f_n dx_n$ on $\Gamma\backslash\mathbb{H}^{n+1}$ is a* **cuspidal 1-form** *if*

(1) α is rapidly decreasing at all cusps of Γ,
(2) for each cusp $\mathfrak{a}$ and $y \geq 0$, we have

$$\int_{\mathcal{P}_\mathfrak{a}} f_{\mathfrak{a},i}(x, y) dx = 0 \ , \quad i = 0, \ldots, n \ ,$$

where $\sigma_\mathfrak{a}^* \alpha = f_{\mathfrak{a},0} dx_0 + f_{\mathfrak{a},1} dx_1 \cdots + f_{\mathfrak{a},n} dx_n$.

We denote by $\mathrm{Har}^1_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R})$ the space of harmonic cuspidal 1-forms on $\Gamma\backslash\mathbb{H}^{n+1}$. Then we have the following identification

$$\mathrm{Har}^1_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R}) \cong H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R}),$$

coming from [41, (2.14)]. This reduces the task of lower bounding the dimension of $H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$ to constructing cuspidal automorphic forms. For congruence subgroups $\Gamma < \mathrm{SV}_{n-1}$, this can be achieved using certain *theta lifts* developed by Shintani [47] of GL_2 holomorphic forms of weight $(n+1)/2+1$ (for details see [41, page 21]). This gives us non-trivial examples for which Theorem 1.5 applies for any n. In the low-dimensional cases $n=1,2$ a lot more can be said, as we will see below.

Finally let us see explicitly how to construct a unitary characters from cuspidal automorphic forms. We let

$$\Phi : \Gamma \to H_1(\Gamma,\mathbb{Z}), \quad \gamma \mapsto \{\infty, \gamma\infty\}$$

which induces the canonical isomorphism $H_1(\Gamma,\mathbb{Z}) \cong \Gamma/[\Gamma,\Gamma]$. For $\gamma\in\Gamma$ and $\omega \in \mathrm{Har}^1_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R})$, we define the *Poincaré pairing*

$$\langle \gamma,\omega\rangle := 2\pi i \int_{\Phi(\Gamma)} \omega = 2\pi i \int_P^{\gamma P} \omega \quad \text{for any } P\in\mathbb{H}^{n+1}.$$

We note that that when $n=1$ and f is a classical Hecke cusp form of weight 2 for Γ, then $f(z)dz$ is indeed a harmonic cuspidal 1-form on $\Gamma\backslash\mathbb{H}^2$ and the Poincaré symbol is equal to (minus) the standard modular symbol (1.1):

$$\langle \gamma, f(z)dz\rangle = 2\pi i \int_\infty^{a_\gamma/c_\gamma} f(z)dz = -\langle a_\gamma/c_\gamma, f\rangle .$$

We observe that if $\gamma\in\Gamma$ is parabolic, then $\langle\gamma,\alpha\rangle = 0$. Hence if we define $\chi_\alpha(\gamma) := e(\langle\gamma,\alpha\rangle)$ then χ_α defines a unitary character trivial on Γ'_∞. The kernel of the map $\alpha\mapsto\chi_\alpha$ is a full rank lattice L inside $\mathrm{Har}^1_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R})$. If we assume that Γ is torsion-free, we indeed obtain the identification $H^1_{\mathrm{free}}(\Gamma,\mathbb{R}/\mathbb{Z}) \cong \mathrm{Har}^1_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R})/L$.

8.3.2. *The case of* $\mathbb{H}^2$. When $n=1$, we have explicit formulas for the dimensions of both the cuspidal and the Eisenstein part. More precisely we have coming from [51, Prop. 6.2.3] that

$$H^1_{\mathrm{cusp}}(\Gamma,\mathbb{Z}) \cong \mathbb{R}^{2g}, H^1_{\mathrm{Eis}}(\Gamma,\mathbb{R}) \cong \mathbb{R}^{2(h-1)},$$

where g is the genus and h is the number of inequivalent cusps of the Riemann surface $\Gamma\backslash\mathbb{H}^2$. In particular if $\Gamma=\Gamma_0(N)$ is a standard Hecke congruence subgroup, we know that $g \sim \frac{N\cdot\prod_{p|N}(1+p^{-1})}{12}$ and $h=\sum_{d|N}\varphi(d,N/d)$ and we conclude that we can find towers of Hecke congruence subgroups such that both the cuspidal and Eisenstein part goes to infinity.

8.3.3. *The case of* $\mathbb{H}^3$. When $n=2$ there has been a lot of activity recently and we refer to the survey of Şengün [11] for an excellent and more thorough overview. In this case no formulas are known in general for the ranks of the cuspidal and Eisenstein part and the best one can hope for are lower bounds.

Regarding the Eisenstein part, we can describe it explicitly when Γ is torsion-free. In this case, we have that $H^1_{\mathrm{Eis}}(\Gamma, \mathbb{R}) \cong \mathbb{R}^h$, where h is the number of cusps of $\Gamma\backslash\mathbb{H}^3$, see [16, Proposition 7.5.6]. The same conclusion holds for co-finite subgroups $\Gamma \le \mathrm{SL}_2(\mathcal{O}_D)$, where $\mathcal{O}_D$ is the ring of integers of the imaginary quadratic field $\mathbb{Q}(\sqrt{D})$ with $D < 0$ a fundamental discriminant not equal to $-4, -3$ (in which case there might be torsion in Γ). In the case of co-finite subgroups $\Gamma \le \mathrm{SL}_2(\mathcal{O}_{\mathcal{D}})$ with $D = -4, -3$ the picture is much more mysterious, but a lot of numerics are available in [10] and [16, Ch. 7.5].

For the cuspidal part there are some useful results giving lower bounds on the rank. First of all Rohlfs [40] showed that

$$\dim H^1_{\mathrm{cusp}}(\mathrm{SL}_2(\mathcal{O}_D), \mathbb{R}) \ge \frac{\varphi(D)}{6} - \frac{1}{2} - h(D),$$

where $h(D)$ denotes the class number of $\mathbb{Q}(\sqrt{D})$. Furthermore Şengün and Turkelli [12] proved that if D is a fundamental discriminant such that $h(D) = 1$, p is a rational prime which is inert in $\mathbb{Q}(\sqrt{D})$ and $\Gamma_0(p^n) \subset \mathrm{SL}_2(\mathcal{O}_D)$ is a congruence subgroup, then we have

$$\dim H^1_{\mathrm{cusp}}(\Gamma_0(p^n), \mathbb{R}) \ge p^{6n},$$

as $n \to \infty$ (an upper bound of p^{10n} has been proved by Calegari and Emerton [4]). In the case of cocompact groups stronger results were obtained by Kionke and Schwermer [26].

8.4. **Torsion in the (co)homology of arithmetic groups.** Now we will discuss what is known about the torsion part of $H_1(\Gamma, \mathbb{Z})$ when $\Gamma \subset \mathrm{SO}(n+1, 1)$ is a cofinite, arithmetic subgroup. In the simplest case $n = 1$, we know that all the torsion in the abeliazation comes from the torsion in the subgroup itself and thus in particular $\Gamma/[\Gamma, \Gamma]$ is torsion-free when Γ is so.

It was noticed a long time ago in unpublished work by Grunewald and Mennicke that in the case $n = 2$ there is a lot of torsion in the abeliazation of congruence subgroups. See Şengün's work [10] for some recent extensive computations.

The study of torsion in the abelianization of Γ fits into a more general framework of understanding the torsion in the homology of arithmetic groups as in the work of Bergeron and Venkatesh [2]. Bergeron and Venkatesh have conjectured that when Γ is a congruence subgroup of $\mathrm{SL}_2(\mathcal{O}_D)$ with $D < 0$ a negative fundamental discriminant, then the torsion in $\Gamma/[\Gamma, \Gamma]$ grows exponentially with the index $[\mathrm{SL}_2(\mathcal{O}_D) : \Gamma]$.

More generally the conjectures predicts that the torsion in the cohomology of symmetric spaces associated to a semisimple Lie group G will grow exponentially in towers of congruence subgroups exactly if we consider the middle dimensional cohomology and if the *fundamental rank* (or "deficiency") $\delta(G) := \mathrm{rank}(G) - \mathrm{rank}(K)$ is 1 (here K is a maximal compact). It follows from [2, 1.2] that the fundamental rank of $\mathrm{SO}(n+1, 1)$ is equal to 1 exactly if n is even. And thus we see that we will have exponential growth of the torsion of $\Gamma/[\Gamma, \Gamma]$ when Γ runs through a tower of congruence groups exactly when $n = 2$ (corresponding to Kleinian groups).

For $n > 2$ the torsion should conjecturally *not* grow exponentially, but there might still be torsion, which is equally arithmetically interesting in view of [43]. There seems however to be no experimental or theoretical work available in this case.

References

[1] Lars V. Ahlfors. Möbius transformations in $\mathbf{R}^n$ expressed through 2×2 matrices of Clifford numbers. *Complex Variables Theory Appl.*, 5(2-4):215–224, 1986.

[2] Nicolas Bergeron and Akshay Venkatesh. The asymptotic growth of torsion homology for arithmetic groups. *J. Inst. Math. Jussieu*, 12(2):391–447, 2013.

[3] Sandro Bettin and Sary Drappeau. Limit laws for rational continued fractions and value distribution of quantum modular forms. *arXiv e-prints*, page arXiv:1903.00457, Mar 2019.

[4] Frank Calegari and Matthew Emerton. Bounds for multiplicities of unitary representations of cohomological type in spaces of cusp forms. *Ann. of Math. (2)*, 170(3):1437–1446, 2009.

[5] Frank Calegari and Akshay Venkatesh. A torsion Jacquet-Langlands correspondence. *Astérisque*, (409):x+226, 2019.

[6] P. Cohen and P. Sarnak. Selberg trace formula, chapters 6 and 7, 1980.

[7] Yves Colin de Verdière. Pseudo-laplaciens. II. *Ann. Inst. Fourier (Grenoble)*, 33(2):87–113, 1983.

[8] Petru Constantinescu. Distribution of modular symbols in $\mathbb{H}^3$. *arXiv e-prints*, page arXiv:2005.07629, May 2020.

[9] J. E. Cremona. *Algorithms for modular elliptic curves*. Cambridge University Press, Cambridge, second edition, 1997.

[10] Mehmet Haluk Şengün. On the integral cohomology of Bianchi groups. *Exp. Math.*, 20(4):487–505, 2011.

[11] Mehmet Haluk Şengün. Arithmetic aspects of Bianchi groups. In *Computations with modular forms*, volume 6 of *Contrib. Math. Comput. Sci.*, pages 279–315. Springer, Cham, 2014.

[12] Mehmet Haluk Şengün and Seyfi Türkelli. Lower bounds on the dimension of the cohomology of Bianchi groups via Sczech cocycles. *J. Théor. Nombres Bordeaux*, 28(1):237–260, 2016.

[13] Nikolaos Diamantis, Jeffrey Hoffstein, Eren Mehmet Kı ral, and Min Lee. Additive twists and a conjecture by Mazur, Rubin and Stein. *J. Number Theory*, 209:1–36, 2020.

[14] J. Elstrodt, F. Grunewald, and J. Mennicke. Vahlen's group of Clifford matrices and spin-groups. *Math. Z.*, 196(3):369–390, 1987.

[15] J. Elstrodt, F. Grunewald, and J. Mennicke. Kloosterman sums for Clifford algebras and a lower bound for the positive eigenvalues of the Laplacian for congruence subgroups acting on hyperbolic spaces. *Invent. Math.*, 101(3):641–685, 1990.

[16] J. Elstrodt, F. Grunewald, and J. Mennicke. *Groups acting on hyperbolic space*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998. Harmonic analysis and number theory.

[17] Ramesh Gangolli. The length spectra of some compact manifolds of negative curvature. *J. Differential Geometry*, 12(3):403–424, 1977.

[18] Ramesh Gangolli and Garth Warner. Zeta functions of Selberg's type for some noncompact quotients of symmetric spaces of rank one. *Nagoya Math. J.*, 78:1–44, 1980.

[19] Krishnendu Gongopadhyay and Ravi S. Kulkarni. z-classes of isometries of the hyperbolic space. *Conform. Geom. Dyn.*, 13:91–109, 2009.

[20] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam, eighth edition, 2015. Translated from the Russian, Translation edited and with a preface by Daniel Zwillinger and Victor Moll, Revised from the seventh edition [MR2360010].

[21] G. Harder. On the cohomology of discrete arithmetically defined groups. In *Discrete subgroups of Lie groups and applications to moduli (Internat. Colloq., Bombay, 1973)*, pages 129–160. 1975.

[22] Dennis A. Hejhal. *The Selberg trace formula for* PSL(2, $\mathbf{R}$). *Vol. 2*, volume 1001 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1983.

[23] Henryk Iwaniec. *Spectral methods of automorphic forms*, volume 53 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, second edition, 2002.

[24] Henryk Iwaniec and Emmanuel Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[25] D. A. Každan. On the connection of the dual space of a group with the structure of its closed subgroups. *Funkcional. Anal. i Priložen.*, 1:71–74, 1967.

[26] Steffen Kionke and Joachim Schwermer. On the growth of the first Betti number of arithmetic hyperbolic 3-manifolds. *Groups Geom. Dyn.*, 9(2):531–565, 2015.

[27] Jungwon Lee and Hae-Sang Sun. Dynamics of continued fractions and distribution of modular symbols. *arXiv e-prints*, page arXiv:1902.06277, Feb 2019.

[28] Alexander Lubotzky. Free quotients and the first Betti number of some hyperbolic manifolds. *Transform. Groups*, 1(1-2):71–82, 1996.

[29] Ju. I. Manin. Parabolic points and zeta functions of modular curves. *Izv. Akad. Nauk SSSR Ser. Mat.*, 36:19–66, 1972.

[30] B. Mazur. Modular curves and the Eisenstein ideal. *Inst. Hautes Études Sci. Publ. Math.*, (47):33–186 (1978), 1977. With an appendix by Mazur and M. Rapoport.

[31] B. Mazur. On the arithmetic of special values of L functions. *Invent. Math.*, 55(3):207–240, 1979.

[32] B. Mazur and K. Rubin. Arithmetic conjectures suggested by the statistical behavior of modular symbols. *arXiv: Number Theory*, 2019.

[33] Barry Mazur and Karl Rubin. Relatively few rational points. Number Theory Seminar at Cal Tech. 2016.

[34] John J. Millson. On the first Betti number of a constant negatively curved manifold. *Ann. of Math. (2)*, 104(2):235–247, 1976.

[35] Asbjorn Christian Nordentoft. Central values of additive twists of cuspidal L-functions. *arXiv e-prints*, page arXiv:1812.08378, December 2018.

[36] Y. N. Petridis and M. S. Risager. Modular symbols have a normal distribution. *Geom. Funct. Anal.*, 14(5):1013–1043, 2004.

[37] Yiannis N. Petridis and Morten S. Risager. The distribution of values of the Poincaré pairing for hyperbolic Riemann surfaces. *J. Reine Angew. Math.*, 579:159–173, 2005.

[38] Yiannis N. Petridis and Morten S. Risager. Arithmetic statistics of modular symbols. *Invent. Math.*, 212(3):997–1053, 2018.

[39] Ralph Phillips and Peter Sarnak. Geodesics in homology classes. *Duke Math. J.*, 55(2):287–297, 1987.

[40] J. Rohlfs. On the cuspidal cohomology of the Bianchi modular groups. *Math. Z.*, 188(2):253–269, 1985.

[41] P. Sarnak. On cusp forms. II. In *Festschrift in honor of I. I. Piatetski-Shapiro on the occasion of his sixtieth birthday, Part II (Ramat Aviv, 1989)*, volume 3 of *Israel Math. Conf. Proc.*, pages 237–250. Weizmann, Jerusalem, 1990.

[42] Theodor Schneider. Arithmetische Untersuchungen elliptischer Integrale. *Math. Ann.*, 113(1):1–13, 1937.

[43] Peter Scholze. On torsion in the cohomology of locally symmetric varieties. *Ann. of Math. (2)*, 182(3):945–1066, 2015.

[44] Atle Selberg. *Collected papers. I.* Springer Collected Works in Mathematics. Springer, Heidelberg, 2014. With a foreword by K. Chandrasekharan, Reprint of the 1989 edition [MR1117906].

[45] Goro Shimura. On the holomorphy of certain Dirichlet series. *Proc. London Math. Soc. (3)*, 31(1):79–98, 1975.

[46] Goro Shimura. *Introduction to the arithmetic theory of automorphic functions*, volume 11 of *Publications of the Mathematical Society of Japan.* Princeton University Press, Princeton, NJ, 1994. Reprint of the 1971 original, Kanô Memorial Lectures, 1.

[47] Takuro Shintani. On construction of holomorphic cusp forms of half integral weight. *Nagoya Math. J.*, 58:83–126, 1975.

[48] Anders Södergren. On the uniform equidistribution of closed horospheres in hyperbolic manifolds. *Proc. Lond. Math. Soc. (3)*, 105(2):225–280, 2012.

[49] Elias M. Stein and Rami Shakarchi. *Fourier analysis*, volume 1 of *Princeton Lectures in Analysis.* Princeton University Press, Princeton, NJ, 2003. An introduction.

[50] H-S. Sun. A proof of the conjecture of Mazur-Rubin-Stein. *To appear in Bull. Korean Math. Soc*, 2020.

[51] Gabor Wiese. Computational arithmetic of modular forms. In *Notes from the International Autumn School on Computational Number Theory*, Tutor. Sch. Workshops Math. Sci., pages 63–170. Birkhäuser/Springer, Cham, 2019.

Department of Mathematics, University College London, 25 Gordon Street, London, UK, WC1H 0AY

E-mail address: petru.constantinescu.17@ucl.ac.uk

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, Copenhagen 2100, Denmark

E-mail address: acnordentoft@outlook.com

PAPER E
SPARSE EQUIDISTRIBUTION OF HYPERBOLIC ORBIFOLDS

SPARSE EQUIDISTRIBUTION OF HYPERBOLIC ORBIFOLDS

PETER HUMPHRIES AND ASBJØRN CHRISTIAN NORDENTOFT

Abstract. Duke, Imamoḡlu, and Tóth have recently constructed a new geometric invariant, a hyperbolic orbifold, associated to each narrow ideal class of a real quadratic field. Furthermore, they have shown that the projection of these hyperbolic orbifolds onto the modular surface $\Gamma\backslash\mathbb{H}$ equidistributes on average over a genus of the narrow class group as the fundamental discriminant D of the real quadratic field tends to infinity.

We study refinements of this result by exploring sparse equidistribution in the subgroup aspect; we prove equidistribution on average over small subgroups of the narrow class group as D tends to infinity.

Behind this refinement is a new interpretation of the Weyl sums arising in these equidistribution problems; we show that these Weyl sums can be expressed in terms of adèlic period integrals instead of in terms of Fourier coefficients of half-integral weight Maaß forms.

1. Introduction

1.1. **Equidistribution of hyperbolic orbifolds.** Let $E := \mathbb{Q}(\sqrt{D})$ be a real quadratic number field, where $D > 1$ is a positive fundamental discriminant. In [DIT16], Duke, Imamoḡlu, and Tóth introduced a new geometric invariant, a hyperbolic orbifold $\Gamma_A\backslash\mathcal{N}_A$, associated to each narrow ideal class A of the narrow class group Cl_D^+ of E. The group $\Gamma_A \subset \mathrm{PSL}_2(\mathbb{Z})$ is a Fuchsian group of the second kind whose construction is in terms of certain invariants of A, while $\mathcal{N}_A \subset \mathbb{H}$ is the Nielsen region of Γ_A, namely the smallest nonempty Γ_A-invariant open convex subset of $\mathbb{H}$. Furthermore, they showed that these hyperbolic orbifolds equidistribute as D tends to infinity when projected onto the modular surface $\Gamma\backslash\mathbb{H}$, where $\Gamma := \mathrm{SL}_2(\mathbb{Z})$.

More precisely, for each positive fundamental discriminant D, we choose a genus G_D in the group of genera $\mathrm{Gen}_D := \mathrm{Cl}_D^+/(\mathrm{Cl}_D^+)^2$, so that G_D is a coset $C(\mathrm{Cl}_D^+)^2$ of narrow ideal classes for some $C \in \mathrm{Cl}_D^+$. Then Duke, Imamoḡlu, and Tóth prove that for every continuity set $B \subset \Gamma\backslash\mathbb{H}$,

$$\frac{\sum_{A\in G_D} \mathrm{vol}(\Gamma_A\backslash\mathcal{N}_A \cap \Gamma B)}{\sum_{A\in G_D} \mathrm{vol}(\Gamma_A\backslash\mathcal{N}_A)} = \frac{\mathrm{vol}(B)}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} + o_B(1)$$

as D tends to infinity through fundamental discriminants [DIT16, Theorem 2]. Here the volume measure on the upper half-plane $\mathbb{H} \ni z = x + iy$ is $d\mu(z) = y^{-2}\,dx\,dy$, so that $\mathrm{vol}(\Gamma\backslash\mathbb{H}) = \pi/3$.

The first author investigated a refinement of this equidistribution result in [Hum18], namely small scale equidistribution, in which the volume of B shrinks as D grows. In this paper, we study a further refinement in a different direction.

The first author is supported by the European Research Council grant agreement 670239.

1.2. **Sparse equidistribution in the subgroup aspect.** Our refinement consists of studying the equidistribution of hyperbolic orbifolds averaged over sparse subsets of Cl_D^+. Previously, we averaged over a genus G_D, which has cardinality $2^{1-\omega(D)}h_D^+$, where $h_D^+ := \#\mathrm{Cl}_D^+$ denotes the narrow class number and $\omega(D)$ denotes the number of distinct prime divisors of D, so that $\#G_D \gg_\varepsilon D^{-\varepsilon}h_D^+$ for every $\varepsilon > 0$. We instead consider an arbitrary subgroup $H = H_D$ of the narrow class group Cl_D^+ in place of the subgroup $(\mathrm{Cl}_D^+)^2$ and with an arbitrary coset CH in place of a genus G_D, where now $\#H$ may be significantly smaller than h_D^+.

Theorem 1.1. *Fix $\delta \geq 0$. For each positive fundamental discriminant D choose a coset CH with $H = H_D$ a subgroup of Cl_D^+ satisfying $\#H \gg D^{-\delta}h_D^+$ and $C \in \mathrm{Cl}_D^+$. Then for each fixed continuity set $B \subset \Gamma\backslash\mathbb{H}$,*

$$\frac{\sum_{A\in CH}\mathrm{vol}(\Gamma_A\backslash\mathcal{N}_A \cap \Gamma B)}{\sum_{A\in CH}\mathrm{vol}(\Gamma_A\backslash\mathcal{N}_A)} = \frac{\mathrm{vol}(B)}{\pi/3} + o_{B,\delta}(1)$$

as D tends to infinity for $\delta < \frac{625}{3309568} \approx 0.0001888$ unconditionally and for $\delta < \frac{1}{4}$ assuming the generalised Lindelöf hypothesis.

This result is motivated by a conjecture of Michel and Venkatesh [MV06, Conjecture 1], where the analogous statement for Heegner points is conjectured to hold for any fixed $\delta < \frac{1}{2}$, and it is noted that the generalised Lindelöf hypothesis implies such a conjecture in the range $\delta < \frac{1}{4}$. Harcos and Michel have proven this conjecture in the range $\delta < \frac{1}{23042}$ [HM06, Theorem 1.2]. More recently, a toy model of this problem was resolved by the first author [Hum20, Theorem 1.10], namely the sparse equidistribution as q tends to infinity of the points

$$\left\{\left(\frac{d}{q}, \frac{d'}{q}\right) \in \mathbb{T}^2 : d \in CH,\ dd' \equiv 1 \pmod{q}\right\}$$

in the torus $\mathbb{T}^2 = (\mathbb{R}/\mathbb{Z})^2$ indexed by a coset CH of the group $(\mathbb{Z}/q\mathbb{Z})^\times$ with q a prime and $\#H \gg q^\delta$ for some fixed $\delta > 0$; it is shown that this is as a simple consequence of a deep result of Bourgain on cancellation in certain exponential sums [Bou05].

The condition that CH be a coset of Cl_D^+ in Theorem 1.1 may be thought of as imposing the requirement that we restrict to a subset of Cl_D^+ with an *algebraic* structure. By comparing to related results on the sparse equidistribution of closed geodesics, we expect that such an algebraic condition is necessary if this subset is smaller than $D^{-\delta}h_D^+$ for some $\delta > 0$ (cf. [AE16, Theorem 4.1] and [BoKo17, Theorem 1.8]); on the other hand, we expect that Theorem 1.1 holds for subsets *without* any algebraic structure provided that the cardinality of this subset divided by $h_D^+/\log D$ tends to infinity with D (cf. [AE16, Theorem 1.2]). We also observe that by taking H to be the trivial subgroup, Theorem 1.1 implies the equidistribution of *individual* hyperbolic orbifolds as D tends to infinity along fundamental discriminants for which $h_D^+ \ll D^\delta$ for some fixed $\delta < \frac{625}{3309568}$ (cf. [Pop06, Theorem 6.5.1]).

In [DIT16, Section 4], it is observed that Theorem 1.1 is *trivial* when $H = \mathrm{Cl}_D^+$ is the whole narrow class group, for then for every fixed continuity set $B \subset \Gamma\backslash\mathbb{H}$, we have the *equality*

$$\frac{\sum_{A\in CH}\mathrm{vol}(\Gamma_A\backslash\mathcal{N}_A \cap \Gamma B)}{\sum_{A\in CH}\mathrm{vol}(\Gamma_A\backslash\mathcal{N}_A)} = \frac{\mathrm{vol}(B)}{\mathrm{vol}(\Gamma\backslash\mathbb{H})} \tag{1.2}$$

with *no* error term. Moreover, if $H = (\mathrm{Cl}_D^+)^2$, so that $CH \in \mathrm{Gen}_D$ is a genus, then we also have the equality (1.2) if J lies in the principal genus $(\mathrm{Cl}_D^+)^2 \in \mathrm{Gen}_D$, where $J \in \mathrm{Cl}_D^+$ is the narrow ideal class containing the different $\mathfrak{d} := (\sqrt{D}) = \sqrt{D}\mathcal{O}_E$, so that J^2 is the principal narrow ideal class I; this occurs if and only if D is not divisible by a prime congruent to 3 modulo 4. We show that this generalises to arbitrary subgroups.

Proposition 1.3. *Let D be a positive fundamental discriminant. Let CH be a coset of Cl_D^+ with H a subgroup of Cl_D^+ and $C \in \mathrm{Cl}_D^+$. Then we have the equality (1.2) for every fixed continuity set $B \subset \Gamma\backslash\mathbb{H}$ if $C^2 J \in H$.*

2. Hyperbolic orbifolds

2.1. **Ideals, quadratic forms, embeddings, and closed geodesics.** We begin by recording several details relating oriented ideals and narrow ideal classes of real quadratic fields, integral binary quadratic forms, embeddings of real quadratic fields into spaces of matrices, and closed geodesics on the level q modular surface $\Gamma_0(q)\backslash\mathbb{H}$. We will in this section work with general level as much as possible with future applications in mind (see Remark 4.9). Useful references for this material include [GKZ87, Section 1], [Dar94, Section 1], and [Pop06, Section 6]. We work throughout with a positive fundamental discriminant $D > 1$ and a squarefree integer q for which every prime dividing q splits in $E := \mathbb{Q}(\sqrt{D})$. We also fix once and for all a residue class r modulo $2q$ for which $r^2 \equiv D \pmod{4q}$.

2.1.1. *Oriented ideals.* Let $\mathfrak{a}$ be a nonzero fractional ideal of E. If $\mathfrak{a}$ is generated over $\mathbb{Z}$ by $\alpha_1, \alpha_2 \in E$, so that $\mathfrak{a} = \mathbb{Z}\alpha_1 + \mathbb{Z}\alpha_2$, then the (absolute) norm of $\mathfrak{a}$ is

$$\mathrm{N}(\mathfrak{a}) = \frac{|\alpha_1\sigma(\alpha_2) - \alpha_2\sigma(\alpha_1)|}{\sqrt{D}},$$

where σ denotes the nontrivial Galois automorphism of E. The ideal $\mathfrak{a}$ is said to be oriented with respect to the generators α_1, α_2 if $\alpha_1\sigma(\alpha_2) - \alpha_2\sigma(\alpha_1) > 0$ and to be of level q if

$$\frac{\mathrm{N}(\alpha_1)}{\mathrm{N}(\mathfrak{a})} \equiv 0 \pmod{q} \quad \text{and} \quad \frac{\mathrm{Tr}(\alpha_1\sigma(\alpha_2))}{\mathrm{N}(\mathfrak{a})} \equiv r \pmod{2q}.$$

We denote by $[\mathfrak{a}; \alpha_1, \alpha_2]$ the oriented ideal $\mathfrak{a}$ with respect to the generators α_1, α_2. The congruence subgroup $\Gamma_0(q) \ni \gamma$ acts on the set of such triples $[\mathfrak{a}; \alpha_1, \alpha_2]$ by acting trivially on the ideal $\mathfrak{a}$ and mapping the pair of generators $(\alpha_1, \alpha_2) \in E^2$, viewed as a row vector, to $(\alpha_1, \alpha_2)\gamma$. This action preserves oriented ideals of level q. The set

$$I := \{(\alpha) = \alpha\mathcal{O}_E \subset E : \alpha \in E,\ \alpha, \sigma(\alpha) > 0\}$$

of totally positive principal ideals — equivalently, the identity in the narrow class group — acts on an oriented ideal $[\mathfrak{a}; \alpha_1, \alpha_2]$ of level q via the map $(\alpha) \cdot [\mathfrak{a}; \alpha_1, \alpha_2] := [(\alpha)\mathfrak{a}; \alpha\alpha_1, \alpha\alpha_2]$, and this action commutes with the action of $\Gamma_0(q)$. In this way, each narrow ideal class of Cl_D^+ may be bijectively identified with equivalence classes of oriented ideals of level q modulo the action of $\Gamma_0(q)$ and I.

2.1.2. *Binary quadratic forms.* For each trio of integers $(a, b, c) \in \mathbb{Z}^3$ having greatest common divisor equal to 1 and satisfying $b^2 - 4ac = D$, $a \equiv 0 \pmod{q}$, and $b \equiv r \pmod{2q}$, we define the integral binary quadratic form

$$Q(x, y) = ax^2 + bxy + cy^2, \tag{2.1}$$

which is a primitive form of level q and discriminant $b^2 - 4ac = D$. We write $Q = [a, b, c]$ to denote this form and call such a form a Heegner form, following [Dar94]; we denote the set of such forms by $\mathcal{Q}_D(q)$. The congruence subgroup $\Gamma_0(q) \ni \gamma$ acts on integral binary quadratic forms via

$$(\gamma \cdot Q)(x, y) := Q(\gamma(x, y)),$$

where we view (x, y) as a column vector; moreover, this action preserves $\mathcal{Q}_D(q)$.

To each equivalence class $I \cdot [\mathfrak{a}; \alpha_1, \alpha_2]$ of oriented ideals of level q, we associate the Heegner form $Q = Q_{I\cdot[\mathfrak{a};\alpha_1,\alpha_2]} \in \mathcal{Q}_D(q)$ given by

$$Q(x, y) := \frac{\mathrm{N}(\alpha_1 x + \alpha_2 y)}{\mathrm{N}(\mathfrak{a})}.$$

Conversely, associated to each Heegner form $Q = [a, b, c] \in \mathcal{Q}_D(q)$ as in (2.1) is the equivalence class of oriented ideals of level q given by

$$\begin{cases} I \cdot \left[\mathfrak{a}; a, \dfrac{b - \sqrt{D}}{2}\right] & \text{if } a > 0, \\ I \cdot \left[\mathfrak{a}; -a\sqrt{D}, \dfrac{D - b\sqrt{D}}{2},\right] & \text{if } a < 0. \end{cases} \tag{2.2}$$

This map is a bijection between $\mathcal{Q}_D(q)$ and equivalence classes of oriented ideals of level q. This association descends to a bijection between narrow ideal classes of the narrow class group Cl_D^+ and equivalence classes of primitive integral binary quadratic forms of level q and discriminant D modulo the action of $\Gamma_0(q)$.

2.1.3. *Oriented embeddings.* Again let $(a, b, c) \in \mathbb{Z}^3$ have greatest common divisor equal to 1 and satisfy $b^2 - 4ac = D$, $a \equiv 0 \pmod{q}$, and $b \equiv r \pmod{2q}$. We define an embedding $\Psi : E \hookrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{Q})$ by

$$\Psi(x + \sqrt{D}y) := \begin{pmatrix} x + by & 2cy \\ -2ay & x - by \end{pmatrix} \tag{2.3}$$

for $x, y \in \mathbb{Q}$. This satisfies

$$\Psi(E) \cap \{g \in \mathrm{Mat}_{2\times 2}(\mathbb{Z}) : g_{2,1} \equiv 0 \pmod{q}\} = \Psi(\mathcal{O}_E),$$

where $\mathcal{O}_E$ denotes the ring of integers of E; that is, Ψ is an oriented optimal embedding of level q. Conversely, every oriented optimal embedding of level q arises from such a trio of integers $(a, b, c) \in \mathbb{Z}^3$. The congruence subgroup $\Gamma_0(q)$ acts on the set of optimal oriented embeddings of level q by conjugation, namely

$$(\gamma \cdot \Psi)(x + \sqrt{D}y) := \gamma^{-1}\Psi(x + \sqrt{D}y)\gamma$$

for $\gamma \in \Gamma_0(q)$, and this action preserves optimal oriented embeddings of level q.

There is a natural bijection between oriented optimal embeddings Ψ of level q as in (2.3) and Heegner forms $Q = [a, b, c]$ as in (2.1), since these are both completed determined by $(a,b,c) \in \mathbb{Z}^3$; in turn, there is a bijection with equivalence classes of oriented ideals of level q as in (2.2). Again, this descends to a bijection between narrow ideal classes A of the narrow class group Cl_D^+ and equivalence classes of oriented optimal embeddings of level q modulo the action of $\Gamma_0(q)$.

2.1.4. *Closed geodesics.* Associated to a Heegner form $Q = [a,b,c] \in \mathcal{Q}_D(q)$ as in (2.1) is an oriented geodesic in the upper half-plane connecting the two points $\frac{-b-\sqrt{D}}{2a}$ and $\frac{-b+\sqrt{D}}{2a}$, namely the Euclidean semicircle

$$\big\{z \in \mathbb{H} : a|z|^2 + b\Re(z) + c = 0\big\} \tag{2.4}$$

oriented anticlockwise if $a > 0$ and clockwise if $a < 0$.

Let $\epsilon_D > 1$ be the least unit with positive norm in $\mathcal{O}_E$, so that $\epsilon_D = u + \sqrt{D}v$ with u, v positive half-integers that satisfy Pell's equation $u^2 - Dv^2 = 1$ and minimise v among all such positive half-integral solutions. For the oriented optimal embedding Ψ of level q as in (2.3) associated to Q, define

$$\gamma_Q \coloneqq \Psi(\epsilon_D) = \begin{pmatrix} u + bv & 2cv \\ -2av & u - bv \end{pmatrix} \in \Gamma_0(q).$$

Together with $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, this generates the group of automorphs of Q,

$$\Gamma_0(q)_Q \coloneqq \{\gamma \in \Gamma_0(q) : \gamma \cdot Q = Q\}.$$

Let A be the narrow ideal class associated to the equivalence class of Heegner forms modulo $\Gamma_0(q)$ that contains Q. We let $\mathcal{C}_A(q)$ denote the reduction modulo $\Gamma_0(q)$ of the oriented geodesic segment from z_Q to $\gamma_Q z_Q$, where

$$z_Q \coloneqq \begin{cases} \dfrac{-b + i\sqrt{D}}{2a} & \text{if } a > 0, \\ \dfrac{b + i\sqrt{D}}{-2a} & \text{if } a < 0, \end{cases} \qquad \gamma_Q z_Q = \begin{cases} \dfrac{-b(u^2 + Dv^2) - 2Duv + i\sqrt{D}}{2a(u^2 + Dv^2)} & \text{if } a > 0, \\ \dfrac{b(u^2 + Dv^2) + 2Duv + i\sqrt{D}}{-2a(u^2 + Dv^2)} & \text{if } a < 0. \end{cases} \tag{2.5}$$

This is an oriented closed geodesic in $\Gamma_0(q)\backslash\mathbb{H}$.

2.2. **Hyperbolic orbifolds.** We will now briefly recall the definition of the Fuchsian groups Γ_A and the hyperbolic orbifolds $\Gamma_A\backslash\mathcal{N}_A$ as in [DIT16, Section 3]. In particular we will work with trivial level $q = 1$.

Let A be a narrow ideal class of E. Then we know that A contains an ideal of the form $\omega\mathbb{Z} + \mathbb{Z}$ with $\omega \in E$ and $\omega > \omega^\sigma$, where σ is the non-trivial Galois automorphism of E. Then we consider the *minus continuous fraction* expansion of ω:

$$w = a_0 - \cfrac{1}{a_1 - \cfrac{1}{a_2 - \cfrac{1}{a_3 - \cfrac{1}{\ddots}}}},$$

where $a_j \in \mathbb{Z}$ with $a_j \geq 2$ for $j \geq 1$. The sequence $a_0, a_1, \ldots$ is eventually period and we denote by $((n_1, \ldots, n_l))$ its unique primitive cycle (only defined up to cyclic permutation).

Now we let $S := \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T := \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and consider the following elements of Γ:

$$S_k = T^{(n_1+\ldots+n_k)} S T^{-(n_1+\ldots+n_k)},$$

for $k = 1, \ldots, l$. Then we define

$$\Gamma_A := \langle S_1, \ldots, S_l, T^{(n_1+\ldots+n_l)} \rangle,$$

which one can show is a Fuchsian group of the second kind. Let $\mathcal{N}_A$ be the associated Nielsen region, which is the smallest nonempty Γ_A-invariant open convex subset of $\mathbb{H}$. Then $\Gamma_A \backslash \mathcal{N}_A$ is a hyperbolic orbifold with boundary component equal to $\mathcal{C}_A(1)$, when projected down to $\Gamma\backslash\mathbb{H}$.

Unlike for geodesics the area of the surfaces $\Gamma_A\backslash\mathcal{N}_A$ might be different for different A. However using the explicit construction described above it follows from [DIT16, Proof of Proposition 1] that we have the following lower bound

$$\sum_{A \in CH} \operatorname{vol}(\Gamma_A \backslash \mathcal{N}_A) > \frac{\#H \log \epsilon_D}{\log(\sqrt{D}+1)}, \tag{2.6}$$

where H is a subgroup of Cl_D^+ and $C \in \mathrm{Cl}_D^+$.

3. Adèlisation of Maass cusp forms

We review some standard notions about Maaß cusp forms of weight k, level q, and principal nebentypus, with an emphasis on forms of weight 0 and the action of raising and lowering operators on such forms. We then describe the relation between such classical automorphic forms and adèlic automorphic forms, highlighting the correspondence between Whittaker functions of representations of $\mathrm{GL}_2(\mathbb{R})$ and $\mathrm{GL}_2(\mathbb{Q}_p)$, the Whittaker expansion of an adèlic automorphic form, and the Fourier expansion of a classical automorphic form. This explicit correspondence is invaluable in Section 4, where we prove an identity between integrals of Maaß forms over hyperbolic orbifolds and period integrals of adèlic automorphic forms. Useful references for this material include [DFI02, Section 4], [GH11, Chapters 3 and 4], [Sch02], and [Pop08].

3.1. **Maaß cusp forms.**

Let k be an integer, q be a positive integer, and denote by $\mathscr{C}_k(\Gamma_0(q))$ the vector subspace of $L^2(\Gamma_0(q)\backslash\mathbb{H})$ spanned by Maaß cusp forms of weight k, level q, and principal nebentypus, in the sense of [DFI02, Section 4]. Such a Maaß cusp form is a real-analytic function $f : \mathbb{H} \to \mathbb{C}$ for which

- f is an eigenfunction of the weight k Laplacian
$$\Delta_k = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + iky\frac{\partial}{\partial x},$$
so that $\Delta_k f(z) = \lambda_f f(z)$ for some $\lambda_f \in \mathbb{C}$ (and necessarily $\lambda_f \in [\frac{1}{4} - \left(\frac{7}{64}\right)^2, \infty)$),
- f is automorphic, so that $j_\gamma(z)^{-k} f(\gamma z) = f(z)$ for all $z \in \mathbb{H}$ and $\gamma \in \Gamma_0(q)$, where for $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2^+(\mathbb{R})$, the space of 2×2 matrices with real entries and positive determinant,

$$gz := \frac{az+b}{cz+d}, \qquad j_g(z) := \frac{cz+d}{|cz+d|}, \tag{3.1}$$

- f is of moderate growth, and
- f is cuspidal, so that for each cusp $\mathfrak{b}$ of $\Gamma_0(q)\backslash\mathbb{H}$,

$$\int_0^1 j_{\sigma_{\mathfrak{b}}}(z)^{-k} f(\sigma_{\mathfrak{b}} z)\, dx = 0$$

for all $y > 0$, where $\sigma_{\mathfrak{b}} \in \mathrm{SL}_2(\mathbb{R})$ is a scaling matrix for $\mathfrak{b}$.

3.1.1. *The Fourier expansion and Hecke eigenvalues of a Maaß cusp form.* The Fourier expansion at the cusp at infinity of a weight 0 Maaß cusp form $f \in \mathscr{C}_0(\Gamma_0(q))$ is

$$f(z) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \rho_f(n) W_{0,it_f}(4\pi|n|y) e(nx), \tag{3.2}$$

where $W_{\alpha,\beta}$ is a Whittaker function and $t_f \in \mathbb{R} \cup i[-\frac{7}{64}, \frac{7}{64}]$ is the spectral parameter of f, so that $\lambda_f = 1/4 + t_f^2$. If f is additionally a Hecke–Maaß *newform*, namely an eigenfunction of the n-th Hecke operator T_n for all $n \in \mathbb{N}$ as well as the reflection operator $X : \mathscr{C}_0(\Gamma_0(q)) \to \mathscr{C}_0(\Gamma_0(q))$ given by $(Xf)(z) := f(-\overline{z})$, the Fourier coefficients $\rho_f(n)$ and Hecke eigenvalues $\lambda_f(n)$ of f satisfy

- $\rho_f(1)\lambda_f(n) = \sqrt{n}\rho_f(n)$ for $n \in \mathbb{N}$,
- $\rho_f(n) = \epsilon_f \rho_f(-n)$ for $n \in \mathbb{Z}$, where $\epsilon_f \in \{1, -1\}$ is the parity of f, so that $Xf = \epsilon_f f$,
- for all $m, n \in \mathbb{N}$, the Hecke eigenvalues satisfy the multiplicativity relations

$$\lambda_f(m)\lambda_f(n) = \sum_{\substack{d\mid(m,n) \\ (d,q)=1}} \lambda_f\left(\frac{mn}{d^2}\right), \qquad \lambda_f(mn) = \sum_{\substack{d\mid(m,n) \\ (d,q)=1}} \mu(d)\lambda_f\left(\frac{m}{d}\right)\lambda_f\left(\frac{n}{d}\right),$$

- for each $p \nmid q$, there exists $\alpha_f(p) \in \mathbb{C}$ satisfying $p^{-\frac{7}{64}} \leq |\alpha_f(p)| \leq p^{\frac{7}{64}}$ such that for all $r \geq 0$,

$$\lambda_f(p^r) = \sum_{m=0}^{r} \alpha_f(p)^m \alpha_f^{-1}(p)^{r-m}. \tag{3.3}$$

- for each $p \parallel q$, there exists $\alpha_f(p) \in \{1, -1\}$ such that for all $r \geq 0$,

$$\lambda_f(p^r) = \frac{\alpha_f(p)^r}{p^{r/2}}, \tag{3.4}$$

- for each prime p for which $p^2 \mid q$, we have that $\lambda_f(p^r) = 0$ for all $r \geq 0$.

3.1.2. *Raising and lowering operators.* The weight k raising operator

$$R_k := \frac{k}{2} + (z - \overline{z})\frac{\partial}{\partial z} = \frac{k}{2} + iy\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)$$

acts on $\mathscr{C}_k(\Gamma_0(q))$ and raises the weight by 2; that is, its image lies in $\mathscr{C}_{k+2}(\Gamma_0(q))$. Similarly, the weight k lowering operator

$$L_k := -\frac{k}{2} - (z - \overline{z})\frac{\partial}{\partial \overline{z}} = -\frac{k}{2} - iy\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)$$

maps $\mathscr{C}_k(\Gamma_0(q))$ to $\mathscr{C}_{k-2}(\Gamma_0(q))$. From [GH11, Proposition 3.9.13], if $f \in \mathscr{C}_0(\Gamma_0(q))$ is a Hecke–Maaß newform of weight 0 and level q with Fourier expansion (3.2), then the Fourier expansion of the weight k Maaß cusp form $F_k \in \mathscr{C}_k(\Gamma_0(q))$ defined by

$$F_k := \begin{cases} R_{k-2} \cdots R_0 f & \text{if } k \in 2\mathbb{N} \cup \{0\}, \\ L_{k+2} \cdots L_0 f & \text{if } k \in -2\mathbb{N}, \end{cases} \tag{3.5}$$

is given by

$$F_k(z) = \frac{\Gamma\left(\frac{k+1}{2} + it\right)\Gamma\left(\frac{k+1}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right)\Gamma\left(\frac{1}{2} + it\right)} \epsilon_f \rho_f(1) \sum_{n=-\infty}^{-1} \frac{\lambda_f(|n|)}{\sqrt{|n|}} W_{-\frac{k}{2}, it_f}(4\pi|n|y)e(nx) + (-1)^{\frac{k}{2}} \rho_f(1) \sum_{n=1}^{\infty} \frac{\lambda_f(n)}{\sqrt{n}} W_{\frac{k}{2}, it_f}(4\pi n y)e(nx) \tag{3.6}$$

if $k \in 2\mathbb{N} \cup \{0\}$, while if $k \in -2\mathbb{N}$, the Fourier expansion is

$$F_k(z) = (-1)^{\frac{k}{2}} \epsilon_f \rho_f(1) \sum_{n=-\infty}^{-1} \frac{\lambda_f(|n|)}{\sqrt{|n|}} W_{-\frac{k}{2}, it_f}(4\pi|n|y)e(nx) + \frac{\Gamma\left(\frac{1-k}{2} + it\right)\Gamma\left(\frac{1-k}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right)\Gamma\left(\frac{1}{2} + it\right)} \rho_f(1) \sum_{n=1}^{\infty} \frac{\lambda_f(n)}{\sqrt{n}} W_{\frac{k}{2}, it_f}(4\pi n y)e(nx). \tag{3.7}$$

3.2. **Adèlic automorphic forms.**

3.2.1. *The adèlic lift of a Maaß cusp form.* Following [GH11, Sections 4.11 and 4.12], we describe the adèlic lift of a Maaß cusp form $f \in \mathscr{C}_k(\Gamma_0(q))$. We first lift $f \in \mathscr{C}_k(\Gamma_0(q))$ to a function $\widetilde{f} : \mathrm{GL}_2^+(\mathbb{R}) \to \mathbb{C}$ defined via

$$\widetilde{f}(g) := j_g(i)^{-k} f(gi). \tag{3.8}$$

For all $g \in \mathrm{GL}_2^+(\mathbb{R})$ and $\theta \in [0, 2\pi)$, this satisfies

$$\widetilde{f}\left(g \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\right) = e^{ik\theta} \widetilde{f}(g).$$

Next, we lift $\widetilde{f}$ to an adèlic automorphic form $\phi = \phi_f$ on $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$, where $\mathbb{A}_\mathbb{Q}$ denotes the ring of adèles of $\mathbb{Q}$. To describe this lift, we first let $K_0(q) \ni k = (k_\infty, k_2, k_3, k_5, \ldots)$ denote the congruence subgroup of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ of the form

$$K_0(q) := \left\{ k \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}) : k_\infty = 1,\ k_p = \begin{pmatrix} a_p & b_p \\ c_p & d_p \end{pmatrix} \in \mathrm{GL}_2(\mathbb{Z}_p), c_p \in p^r\mathbb{Z}_p \text{ if } p^r \parallel q \right\}. \tag{3.9}$$

Here 1 denotes the 2×2 identity matrix. We view $\mathrm{GL}_2^+(\mathbb{R}) \ni g_\infty$ as a subgroup of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ via the embedding $g_\infty \mapsto (g_\infty, 1, 1, \ldots)$. Finally, we view $\mathrm{GL}_2(\mathbb{Q}) \ni \gamma$ as a subgroup of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ via the diagonal embedding $\gamma \mapsto (\gamma, \gamma, \gamma, \ldots)$. Then via the strong approximation theorem,

$$\mathrm{GL}_2(\mathbb{A}_\mathbb{Q}) = \mathrm{GL}_2(\mathbb{Q})\, \mathrm{GL}_2^+(\mathbb{R}) K_0(q), \tag{3.10}$$

so that every $g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ can be written (non-uniquely) as $g = \gamma g_\infty k$ for some $\gamma \in \mathrm{GL}_2(\mathbb{Q})$, $g_\infty \in \mathrm{GL}_2^+(\mathbb{R})$, and $k \in K_0(q)$. The adèlic lift $\phi = \phi_f$ of a Maaß cusp form $f \in \mathscr{C}_k(\Gamma_0(q))$ is then given by

$$\phi(g) = \phi(\gamma g_\infty k) := \widetilde{f}(g_\infty). \tag{3.11}$$

This is well-defined even though the decomposition $g = \gamma g_\infty k$ is not unique. In particular,

$$\phi(g) = f(x + iy) \tag{3.12}$$

for $g = g_\infty = \left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{GL}_2^+(\mathbb{R}) \subset \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ with $y > 0$ and $x \in \mathbb{R}$.

3.2.2. *The Whittaker expansion of an adèlic automorphic form.* Let $\psi : \mathbb{Q}\backslash\mathbb{A}_\mathbb{Q} \to \mathbb{C}$ be the standard adèlic additive character defined as in [GH11, Definition 1.7.1], so that $\psi(u) = \psi_\infty(u_\infty) \prod_p \psi_p(u_p)$ for $u = (u_\infty, u_2, u_3, \ldots) \in \mathbb{A}_\mathbb{Q}$ with $\psi_\infty : \mathbb{R} \to \mathbb{C}$ the additive character $\psi_\infty(u_\infty) := e(u_\infty) := e^{2\pi i u_\infty}$ and $\psi_p : \mathbb{Q}_p \to \mathbb{C}$ the standard unramified additive character defined in [GH11, Definition 1.6.3]. The Whittaker function $W_\phi : \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}) \to \mathbb{C}$ of a cuspidal adèlic automorphic form ϕ is

$$W_\phi(g) := \int_{\mathbb{Q}\backslash\mathbb{A}_\mathbb{Q}} \phi\left(\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} g\right) \overline{\psi}(u)\, du,$$

which satisfies

$$W_\phi\left(\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} g\right) = \psi(u) W_\phi(g) \tag{3.13}$$

for all $u \in \mathbb{A}_\mathbb{Q}$ and $g \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$. The automorphic form ϕ has the Whittaker expansion

$$\phi(g) = \sum_{\alpha \in \mathbb{Q}^\times} W_\phi\left(\begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} g\right). \tag{3.14}$$

3.2.3. *The Whittaker expansion of an adèlic lift.* Let $\phi = \phi_{F_k}$ be the adèlic lift of a Maaß cusp form $F_k \in \mathscr{C}_k(\Gamma_0(q))$ of weight k associated to a Hecke–Maaß newform $f \in \mathscr{C}_0(\Gamma_0(q))$ of weight 0 as in (3.5). Then ϕ is a pure tensor lying in the vector space of a cuspidal automorphic representation $\pi = \pi_f = \pi_\infty \otimes \bigotimes_p \pi_p$ of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$, where each π_p is a generic irreducible admissible unitary representation of $\mathrm{GL}_2(\mathbb{Q}_p)$ and π_∞ is a generic irreducible unitary Casselman–Wallach representation of $\mathrm{GL}_2(\mathbb{R})$. In particular, for $g = (g_\infty, g_2, g_3, \ldots) \in \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$, we have the factorisation

$$W_\phi(g) = c_\phi W_\infty(g_\infty) \prod_p W_p(g_p), \tag{3.15}$$

where c_ϕ is a constant independent of g, each W_p is a Whittaker function in the Whittaker model $\mathcal{W}(\pi_p, \psi_p)$ of π_p, and similarly $W_\infty \in \mathcal{W}(\pi_\infty, \psi_\infty)$.

We show in Section 3.2.4 that the Whittaker functions W_p are such that for $\alpha \in \mathbb{Q}^\times$,

$$\prod_p W_p \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} = \begin{cases} \dfrac{\lambda_f(|n|)}{\sqrt{|n|}} & \text{if } \alpha = n \in \mathbb{Z} \setminus \{0\}, \\ 0 & \text{otherwise.} \end{cases}$$

In Section 3.2.5, we will show that for $n \in \mathbb{Z} \setminus \{0\}$ and $y > 0$,

$$W_\infty \begin{pmatrix} ny & 0 \\ 0 & 1 \end{pmatrix} = \begin{cases} (-1)^{\frac{k}{2}} W_{\frac{k}{2},it}(4\pi ny), & n \in \mathbb{N}, k \in 2\mathbb{N} \cup \{0\}, \\ \epsilon_f \dfrac{\Gamma\left(\frac{k+1}{2} + it\right)\Gamma\left(\frac{k+1}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right)\Gamma\left(\frac{1}{2} - it\right)} W_{-\frac{k}{2},it}(4\pi|n|y), & n \in -\mathbb{N}, k \in 2\mathbb{N} \cup \{0\}, \\ \dfrac{\Gamma\left(\frac{1-k}{2} + it\right)\Gamma\left(\frac{1-k}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right)\Gamma\left(\frac{1}{2} - it\right)} W_{\frac{k}{2},it}(4\pi ny), & n \in \mathbb{N}, k \in -2\mathbb{N}, \\ \epsilon_f (-1)^{\frac{k}{2}} W_{-\frac{k}{2},it}(4\pi|n|y), & n \in -\mathbb{N}, k \in -2\mathbb{N}. \end{cases}$$

From this, we see that the constant c_ϕ in (3.15) is equal to $\rho_f(1)$ by taking $g = g_\infty = \left(\begin{smallmatrix} y & x \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{GL}_2^+(\mathbb{R}) \subset \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ with $y > 0$ and $x \in \mathbb{R}$ in (3.14), so that by (3.12) and (3.13),

$$F_k(x+iy) = \phi(g) = c_\phi \sum_{\alpha \in \mathbb{Q}^\times} W_\infty \begin{pmatrix} \alpha y & 0 \\ 0 & 1 \end{pmatrix} \prod_p W_p \begin{pmatrix} \alpha & 0 \\ 0 & 1 \end{pmatrix} e(\alpha x),$$

and comparing this adèlic Whittaker expansion to the classical Fourier expansion at the cusp at infinity (3.6) and (3.7).

3.2.4. *Nonarchimedean Whittaker functions.* Let $\phi = \phi_{F_k}$ be the adèlic lift of a Maaß cusp form $F_k \in \mathscr{C}_k(\Gamma_0(q))$ associated to a Hecke–Maaß newform $f \in \mathscr{C}_0(\Gamma_0(q))$ as in (3.5) with q squarefree. For each prime p, the local Whittaker functions $W_p \in \mathcal{W}(\pi_p, \psi_p)$ are of a distinguished form. One can explicitly describe the values of the Whittaker function $W_p(g_p)$ for $g_p = \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with $a \in \mathbb{Q}_p^\times$; see [Sch02, Section 2.4].

- For $p \nmid q$, the representation π_p is a spherical principal series representation $\omega_p \boxplus \omega_p^{-1}$, where ω_p is an unramified character of $\mathbb{Q}_p^\times$ satisfying $p^{-\frac{7}{64}} \leq |\omega_p(p)|_p \leq p^{\frac{7}{64}}$ and $|\cdot|_p$ denotes the p-adic absolute value normalised such that $|p|_p = p^{-1}$. This character is such that $\omega_p(p)$ is equal to $\alpha_f(p)$ as in (3.3). For $a \in \mathbb{Q}_p^\times$, let $v(a) \in \mathbb{Z}$ be such that $|a|_p = p^{-v(a)}$. There is a distinguished Whittaker function, the spherical Whittaker function, that is right $\mathrm{GL}_2(\mathbb{Z}_p)$-invariant and satisfies

$$W_p \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \begin{cases} \displaystyle\sum_{m=0}^{v(a)} \omega_p(p)^m \omega_p^{-1}(p)^{v(a)-m} |a|_p^{1/2} & \text{if } 0 < |a|_p \leq 1, \text{ so that } v(a) \geq 0, \\ 0 & \text{if } |a|_p \geq p, \text{ so that } v(a) \leq -1. \end{cases}$$

- For $p \mid q$, the representation π_p is a special representation $\omega_p \mathrm{St}_p$, where ω_p is an unramified unitary character of $\mathbb{Q}_p^\times$, so that $\omega_p(p) \in \{1, -1\}$. This character is such that $\omega_p(p)$ is equal to $\alpha_f(p)$ as in (3.4). There is a distinguished Whittaker function, the Whittaker newform, that is right-invariant under the congruence subgroup of $\mathrm{GL}_2(\mathbb{Z}_p)$ consisting of elements $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ for which $c \in p\mathbb{Z}_p$ and satisfies

$$W_p \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \begin{cases} \omega_p(a)|a|_p & \text{if } 0 < |a|_p \leq 1, \text{ so that } v(a) \geq 0, \\ 0 & \text{if } |a|_p \geq p, \text{ so that } v(a) \leq -1. \end{cases}$$

3.2.5. *Archimedean Whittaker functions.* Let $\phi = \phi_{F_k}$ be the adèlic lift of a Maaß cusp form $F_k \in \mathscr{C}_k(\Gamma_0(q))$ associated to a Hecke–Maaß newform $f \in \mathscr{C}_0(\Gamma_0(q))$ as in (3.5). The local Whittaker function $W_\infty = W_\infty^k \in \mathcal{W}(\pi_\infty, \psi_\infty)$ is again of a distinguished form. Since f has weight 0, the representation π_∞ is a principal series representation of the form $\operatorname{sgn}^\kappa |\cdot|_\infty^{it} \boxplus \operatorname{sgn}^\kappa |\cdot|_\infty^{-it}$ with $\kappa = \kappa_f \in \{0,1\}$ and $t = t_f \in \mathbb{R} \cup i[-\frac{7}{64}, \frac{7}{64}]$ such that $(-1)^{\kappa_f} = \epsilon_f$ is the parity of f and t_f is the spectral parameter of f, so that $1/4 + t_f^2 = \lambda_f$ is the Laplacian eigenvalue of f. Here $|\cdot|_\infty = |\cdot|$ is the usual archimedean absolute value on $\mathbb{R}$.

The following claims are essentially implicit (albeit with some typographical errors) in the seminal work of Jacquet–Langlands [JL70, Section 2.5], as further detailed by Godement [God18, Sections 2.3–2.6]; see also [Pop08]. For the sake of completeness, we give explicit proofs.

- For each $k \in 2\mathbb{Z}$, there exists a distinguished Whittaker function $W_\infty^k \in \mathcal{W}(\pi_\infty, \psi_\infty)$, where $\pi_\infty = \operatorname{sgn}^\kappa |\cdot|_\infty^{it} \boxplus \operatorname{sgn}^\kappa |\cdot|_\infty^{-it}$, that is of weight k, so that for all $g_\infty \in \mathrm{GL}_2(\mathbb{R})$ and $\theta \in [0, 2\pi)$, it satisfies

$$W_\infty^k \left(g_\infty \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \right) = e^{ik\theta} W_\infty^k(g_\infty). \tag{3.16}$$

- For all $a \in \mathbb{R}^\times$, this distinguished Whittaker function satisfies

$$W_\infty^k \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \begin{cases} (-1)^{\frac{k}{2}} W_{\frac{k}{2}, it}(4\pi a), & a > 0, k \in 2\mathbb{N} \cup \{0\}, \\ (-1)^\kappa \dfrac{\Gamma\left(\frac{k+1}{2} + it\right)\Gamma\left(\frac{k+1}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right)\Gamma\left(\frac{1}{2} - it\right)} W_{-\frac{k}{2}, it}(4\pi|a|), & a < 0, k \in 2\mathbb{N} \cup \{0\}, \\ \dfrac{\Gamma\left(\frac{1-k}{2} + it\right)\Gamma\left(\frac{1-k}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right)\Gamma\left(\frac{1}{2} - it\right)} W_{\frac{k}{2}, it}(4\pi a), & a > 0, k \in -2\mathbb{N}, \\ (-1)^{\kappa + \frac{k}{2}} W_{-\frac{k}{2}, it}(4\pi|a|), & a < 0, k \in -2\mathbb{N}. \end{cases} \tag{3.17}$$

- For $\kappa' \in \{0,1\}$ and $\Re(s) \geq 1/2$, we have that

$$\int_{\mathbb{R}^\times} W_\infty^2 \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \operatorname{sgn}^{\kappa'}(a) |a|^{s-\frac{1}{2}} \, d^\times a = \begin{cases} -2\pi^{-s}\Gamma\left(\dfrac{s+1+it}{2}\right)\Gamma\left(\dfrac{s+1-it}{2}\right) & \text{if } \kappa \equiv \kappa' + 1 \pmod 2, \\ \left(\dfrac{1}{2} - s\right)\pi^{-s}\Gamma\left(\dfrac{s+it}{2}\right)\Gamma\left(\dfrac{s-it}{2}\right) & \text{if } \kappa \equiv \kappa' \pmod 2, \end{cases} \tag{3.18}$$

where $d^\times a := |a|^{-1} \, da$ denotes the multiplicative Haar measure on $\mathbb{R}^\times$.

- We have that

$$\int_{\mathbb{R}^\times} \left| W_\infty^2 \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} \right|^2 d^\times a = \left(\frac{1}{4} + t^2\right) \Gamma\left(\frac{1}{2} + it\right) \Gamma\left(\frac{1}{2} - it\right). \tag{3.19}$$

Proof of (3.16). Let $\pi_\infty = \mathrm{sgn}^\kappa |\cdot|_\infty^{t_1} \boxplus \mathrm{sgn}^\kappa |\cdot|_\infty^{t_2}$ be a principal series representation with $\kappa \in \{0,1\}$ and $t_1, t_2 \in \mathbb{C}$. We initially assume that $\Re(t_1) > \Re(t_2)$. For each $k \in 2\mathbb{Z}$ and $(x_1, x_2) \in \mathbb{R}^2$, let

$$\begin{aligned} \Phi^k(x_1, x_2) &\coloneqq (x_2 - \mathrm{sgn}(k) i x_1)^{|k|} e^{-\pi(x_1^2 + x_2^2)} \\ &= \begin{cases} (x_2 - ix_1)^k e^{-\pi(x_1^2+x_2^2)}, & k \in 2\mathbb{N} \cup \{0\}, \\ (x_2 + ix_1)^{-k} e^{-\pi(x_1^2+x_2^2)}, & k \in -2\mathbb{N}. \end{cases} \end{aligned} \tag{3.20}$$

Define the Godement section

$$\varphi_\infty^k(g_\infty) \coloneqq \pi^{\frac{|k|}{2}} \mathrm{sgn}^\kappa(\det g_\infty) \, |\det g_\infty|^{t_1 + \frac{1}{2}} \int_{\mathbb{R}^\times} |y|^{t_2 - t_1 - 1} \Phi^k((0, y^{-1}) g_\infty) \, d^\times y. \tag{3.21}$$

This Godement section converges absolutely and defines an element of the induced model of π_∞ of weight k; that is,

$$\varphi_\infty^k\left(\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} g_\infty\right) = \mathrm{sgn}^\kappa(a)|a|^{t_1+\frac{1}{2}} \mathrm{sgn}^\kappa(d)|d|^{t_2 - \frac{1}{2}} \varphi_\infty^k(g_\infty),$$

$$\varphi_\infty^k\left(g_\infty \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}\right) = e^{ik\theta} \varphi_\infty^k(g_\infty)$$

for all $a, d \in \mathbb{R}^\times$, $b \in \mathbb{R}$, $g_\infty \in \mathrm{GL}_2(\mathbb{R})$, and $\theta \in [0, 2\pi)$. Moreover, the Jacquet integral

$$\begin{aligned} W_\infty^k(g_\infty) &\coloneqq \int_{\mathbb{R}} \varphi_\infty^k\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} g_\infty\right) e^{-2\pi i u} \, du \\ &= \pi^{\frac{|k|}{2}} \mathrm{sgn}^\kappa(\det g_\infty) \, |\det g_\infty|^{t_1 + \frac{1}{2}} \int_{\mathbb{R}^\times} |y|^{t_2 - t_1} \int_{\mathbb{R}} \Phi^k((y^{-1}, u) g_\infty) e^{-2\pi i y u} \, du \, d^\times y \end{aligned} \tag{3.22}$$

converges absolutely and defines an element of the Whittaker model $\mathcal{W}(\pi_\infty, \psi_\infty)$ of π_∞ of weight k. The Whittaker function W_∞^k extends holomorphically as a function of the complex variables $t_1, t_2 \in \mathbb{C}$ to $(t_1, t_2) = (it, -it)$ with $t \in \mathbb{R} \cup i[-\frac{7}{64}, \frac{7}{64}]$. This holomorphic extension defines a weight k element of the Whittaker model $\mathcal{W}(\pi_\infty, \psi_\infty)$ of $\pi_\infty = \mathrm{sgn}^\kappa |\cdot|_\infty^{it} \boxplus \mathrm{sgn}^\kappa |\cdot|_\infty^{-it}$. □

Proof of (3.17). For $a \in \mathbb{R}^\times$, we have that

$$W_\infty^k \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = \pi^{\frac{|k|}{2}} \mathrm{sgn}^\kappa(a) |a|^{it + \frac{1}{2}} \int_{\mathbb{R}^\times} |y|^{-2it} \int_{\mathbb{R}} \Phi^k(y^{-1} a, u) e^{-2\pi i y u} \, du \, d^\times y \tag{3.23}$$

from (3.22). Since $\Phi^k(x_1, x_2) = \Phi^{-k}(-x_1, x_2)$, we see that

$$W_\infty^k \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = (-1)^\kappa W_\infty^{-k} \begin{pmatrix} -a & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.24}$$

Thus it suffices to prove (3.17) for $a > 0$.

For $a > 0$ and $k \in 2\mathbb{N} \cup \{0\}$, we have that

$$W_\infty^k \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = 4\pi^{\frac{k}{2}} \sqrt{a} e^{2\pi a} \int_0^\infty \cos(2ty) \int_{\mathbb{R}} u^k e^{-\pi u^2} e^{-4\pi i u \sqrt{a} \cosh y} \, du \, dy$$

upon inserting the identity (3.20) for Φ^k into (3.23), making the change of variables $u \mapsto u + iy^{-1}a$, shifting the contour of integration back to the line $\Im(u) = 0$ via Cauchy's integral theorem, and making the change of variables $y \mapsto \sqrt{a} e^y$. By [GR15, 9.241.1], the inner integral is

$$(-2\pi)^{-\frac{k}{2}} e^{-2\pi a \cosh^2 y} D_k(2\sqrt{2\pi a} \cosh y),$$

where D_k is a parabolic cylinder function. Then by [GR15, 7.731.1], we deduce that

$$W_\infty^k \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = (-1)^{\frac{k}{2}} W_{\frac{k}{2}, it}(4\pi a).$$

Similarly, for $a > 0$ and $k \in -2\mathbb{N}$, we have that

$$W_\infty^k \begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix} = 4\pi^{-\frac{k}{2}} \sqrt{a} e^{-2\pi a} \int_0^\infty \cos(2ty) \int_{\mathbb{R}} u^{-k} e^{-\pi u^2} e^{-4\pi i u \sqrt{a} \sinh y} \, du \, dy.$$

Via Parseval's identity and [GR15, 8.432.4], this is

$$\frac{4\pi^{1-\frac{k}{2}} \sqrt{a} e^{-2\pi a}}{\Gamma\left(\frac{1}{2} + it\right) \Gamma\left(\frac{1}{2} - it\right)} \int_{\mathbb{R}} u^{-k} e^{-\pi u^2} K_{2it}(4\pi \sqrt{a} u) \, du,$$

where K_{2it} is a modified Bessel function of the second kind, which, by [GR15, 6.631.3], is equal to

$$\frac{\Gamma\left(\frac{1-k}{2} + it\right) \Gamma\left(\frac{1-k}{2} - it\right)}{\Gamma\left(\frac{1}{2} + it\right) \Gamma\left(\frac{1}{2} - it\right)} W_{\frac{k}{2}, it}(4\pi a). \qquad \square$$

Proof of (3.18). We insert the identity (3.23) for W_∞^2 into the left-hand side of (3.18), make the change of variables $a \mapsto ya$ and $u \mapsto u + ia$, then shift the contour of integration back to the line $\Im(u) = 0$, yielding

$$\pi \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(a) |a|^{s+it} \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(y) |y|^{s-it} e^{2\pi y a} \int_{\mathbb{R}^\times} u^2 e^{-\pi u^2} e^{-2\pi i (y+a) u} \, du \, d^\times y \, d^\times a.$$

The innermost integral may be evaluated via integration by parts, leading to

$$\begin{aligned} &\frac{1}{2} \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(a) |a|^{s+it} e^{-\pi a^2} \, d^\times a \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(y) |y|^{s-it} e^{-\pi y^2} \, d^\times y \\ &\quad - \pi \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(a) |a|^{s+it} e^{-\pi a^2} \, d^\times a \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(y) |y|^{s+2-it} e^{-\pi y^2} \, d^\times y \\ &\quad - \pi \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(a) |a|^{s+2+it} e^{-\pi a^2} \, d^\times a \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'}(y) |y|^{s-it} e^{-\pi y^2} \, d^\times y \\ &- 2\pi \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'+1}(a) |a|^{s+1+it} e^{-\pi a^2} \, d^\times a \int_{\mathbb{R}^\times} \operatorname{sgn}^{\kappa+\kappa'+1}(y) |y|^{s+1-it} e^{-\pi y^2} \, d^\times y. \end{aligned}$$

The first three expressions vanish if $\kappa \equiv \kappa' + 1 \pmod 2$, while the last vanishes if $\kappa \equiv \kappa' \pmod 2$. The result then follows via the recurrence relation $\Gamma(s+1) = s\Gamma(s)$. $\square$

Proof of (3.19). The left-hand side of (3.19) is

$$\int_0^\infty \left(W_{1,it}(4\pi a)^2 + \left(\frac{1}{4} + t^2\right)^2 W_{-1,it}(4\pi a)^2 \right) \frac{da}{a}$$

from (3.17). The desired identity then follows from the change of variables $a \mapsto a/4\pi$ together with [GR15, 7.611.4, 8.365, and 8.368]. □

4. Weyl sums

4.1. **Weyl sums for newforms and L-Functions.** Let $f \in \mathscr{C}_0(\Gamma)$ be a Hecke–Maaß newform of level 1, and let χ be a narrow class character of E. Our goal is to relate the twisted Weyl sum

$$W_{\chi,f} := \sum_{A \in \mathrm{Cl}_D^+} \chi(A) \int_{\Gamma_A \backslash \mathcal{N}_A} f(z)\, d\mu(z) \tag{4.1}$$

to a special value of the Rankin–Selberg L-function $L(s, f \otimes \Theta_\chi)$, where Θ_χ denotes the theta series associated to χ, as in [HK20, Appendix A.1], which is a newform of weight 0, level D, nebentypus χ_D, Laplacian eigenvalue $\lambda_{\Theta_\chi} = 1/4$, and parity $\epsilon_{\Theta_\chi} = \chi(J) \in \{1, -1\}$, where J is the narrow ideal class containing the different. The automorphic form Θ_χ is a cusp form if and only if χ is complex; otherwise Θ_χ is an Eisenstein series and χ is a genus character.

Proposition 4.2. *Let $f \in \mathscr{C}_0(\Gamma)$ be a Hecke–Maaß newform of level 1 normalised such that*

$$\int_{\Gamma \backslash \mathbb{H}} |f(z)|^2\, d\mu(z) = 1,$$

and let χ be a narrow class character of E. We have that

$$|W_{\chi,f}|^2 = \begin{cases} \dfrac{2\sqrt{D}}{\left(\frac{1}{4} + t_f^2\right)^2} \dfrac{\Gamma\left(\frac{3}{4} + \frac{it_f}{2}\right)^2 \Gamma\left(\frac{3}{4} - \frac{it_f}{2}\right)^2}{\Gamma\left(\frac{1}{2} + it_f\right)\Gamma\left(\frac{1}{2} - it_f\right)} \dfrac{L\left(\frac{1}{2}, f \otimes \Theta_\chi\right)}{L(1, \mathrm{sym}^2 f)} & \text{if } \epsilon_f = -\chi(J), \\ 0 & \text{if } \epsilon_f = \chi(J), \end{cases}$$

where $t_f \in \mathbb{R} \cup i[-\frac{7}{64}, \frac{7}{64}]$ denotes the spectral parameter and $\epsilon_f \in \{1, -1\}$ denotes the parity of f.

Remark 4.3. In terms of the completed L-functions

$$\Lambda(s, f \otimes \Theta_\chi) := \pi^{-2(s+|\kappa_f - \kappa_\chi|)} \Gamma\left(\frac{s + |\kappa_f - \kappa_\chi| + it_f}{2}\right)^2 \cdot \Gamma\left(\frac{s + |\kappa_f - \kappa_\chi| - it_f}{2}\right)^2 L(s, f \otimes \Theta_\chi),$$

$$\Lambda(s, \mathrm{sym}^2 f) := \pi^{-\frac{3s}{2}} \Gamma\left(\frac{s}{2} + it_f\right) \Gamma\left(\frac{s}{2}\right) \Gamma\left(\frac{s}{2} - it_f\right) L(s, \mathrm{sym}^2 f),$$

this is

$$\frac{2\pi^2\sqrt{D}}{\left(\frac{1}{4} + t_f^2\right)^2} \frac{\Lambda\left(\frac{1}{2}, f \otimes \Theta_\chi\right)}{\Lambda(1, \mathrm{sym}^2 f)}.$$

Remark 4.4. It is instructive to consider the case where χ is a genus character associated to the pair of primitive quadratic Dirichlet characters χ_{D_1} and χ_{D_2} modulo $|D_1|$ and $|D_2|$ respectively, where D_1 and D_2 are fundamental discriminants for which $D_1 D_2 = D$. Then Θ_χ is the Eisenstein newform associated to χ_{D_1} and χ_{D_2}, as described in [You19], and so $L(s, f \otimes \Theta_\chi) = L(s, f \otimes \chi_{D_1}) L(s, f \otimes \chi_{D_2})$. Since $D > 0$, either $D_1, D_2 > 0$ or $D_1, D_2 < 0$; in the former case, we have that $\chi(J) = 1$, while $\chi(J) = -1$ in the latter case. Proposition 4.2 then gives the identity

(4.5)
$$|W_{\chi,f}|^2 = \begin{cases} \dfrac{2\sqrt{D}}{\left(\frac{1}{4} + t_f^2\right)^2} \dfrac{\Gamma\left(\frac{3}{4} + \frac{it_f}{2}\right)^2 \Gamma\left(\frac{3}{4} - \frac{it_f}{2}\right)^2}{\Gamma\left(\frac{1}{2} + it_f\right)\Gamma\left(\frac{1}{2} - it_f\right)} \dfrac{L\left(\frac{1}{2}, f \otimes \chi_{D_1}\right) L\left(\frac{1}{2}, f \otimes \chi_{D_2}\right)}{L(1, \mathrm{sym}^2 f)} & \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if } \epsilon_f = -\operatorname{sgn}(D_1) = -\operatorname{sgn}(D_2), & \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \text{if } \epsilon_f = \operatorname{sgn}(D_1) = \operatorname{sgn}(D_2). & \end{cases}$$

We see that the Weyl sum $W_{\chi,f}$ vanishes if f is even and $D_1, D_2 > 0$ or if f is odd and $D_1, D_2 < 0$; additionally, $W_{\chi,f}$ vanishes if f is odd and $D_1, D_2 > 0$, for then the root numbers of $f \otimes \chi_{D_1}$ and $f \otimes \chi_{D_2}$ are both equal to -1 [HK20, Lemma A.2], and hence $L(s, f \otimes \chi_{D_1})$ and $L(s, f \otimes \chi_{D_2})$ both vanish at $s = 1/2$. These vanishing results and the identity (4.5) when f is even and $D_1, D_2 < 0$ are in exact accordance with the work of Duke, Imamoḡlu, and Tóth [DIT16, Theorem 4 and (5.17)].

Similarly, letting $E(z, s)$ denote the Eisenstein series on $\Gamma\backslash\mathbb{H}$, which has parity 1, we wish to relate the Weyl sum

$$W_{\chi,t} := \sum_{A \in \mathrm{Cl}_D^+} \chi(A) \int_{\Gamma_A \backslash \mathcal{N}_A} E\left(z, \frac{1}{2} + it\right) d\mu(z) \tag{4.6}$$

to a special value of $L(s, \Theta_\chi) = L(s, \chi)$.

Proposition 4.7. *Let χ be a narrow class character of E. For $t \in \mathbb{R}$, we have that*

$$W_{\chi,t} = \begin{cases} \dfrac{2D^{\frac{1}{4} + \frac{it}{2}}}{\frac{1}{4} + t^2} \dfrac{\Gamma\left(\frac{3}{4} + \frac{it}{2}\right)^2}{\Gamma\left(\frac{1}{2} + it\right)} \dfrac{L\left(\frac{1}{2} + it, \Theta_\chi\right)}{\zeta(1 + 2it)} & \text{if } \chi(J) = -1, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 4.8. This is

$$\frac{2\pi D^{\frac{1}{4} + \frac{it}{2}}}{\frac{1}{4} + t^2} \frac{\Lambda\left(\frac{1}{2} + it, \Theta_\chi\right)}{\Lambda(1 + 2it)}$$

in terms of the completed L-functions

$$\Lambda(s, \Theta_\chi) := \pi^{-s-1} \Gamma\left(\frac{s+1}{2}\right)^2 L(s, \Theta_\chi), \qquad \Lambda(s) := \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s).$$

The proofs of Propositions 4.2 and 4.7 are given in Section 4.5. Our method is to first prove identities relating certain adèlic period integrals to L-functions, then show that these adèlic period integrals are equal to integrals over closed geodesics, and finally relate these integrals over closed geodesics to integrals over hyperbolic orbifolds.

Remark 4.9. The automorphic setup allows very naturally to work with general level. Thus it is tempting to try to obtain sparse equidistribution results in the level aspect in the spirit of [LMY13]. All of our arguments goes through, except that we at the moment do not have a satisfactory definition of the hyperbolic orbifolds in the higher level case. This is work in progress.

4.2. **Adèlic period integrals and choice of test vectors.** The first step towards proving Proposition 4.2 is to apply a formula due to Martin and Whitehouse [MW09], which relates certain adèlic period integrals to a ratio of special values of L-functions. In this section, we describe in some detail how the results of Martin and Whitehouse apply in our specific case. This part of the argument applies equally well for general levels, and so we will work in this generality.

4.2.1. *Waldspurger-type formulæ.* Following the pioneering work of Waldspurger [Wal85], there has been considerable work in obtaining explicit formulæ relating adèlic period integrals and central values of Rankin–Selberg L-functions, as in work of Gross [Gro88], Zhang [Zha01], Martin and Whitehouse [MW09], and File, Martin, and Pitale [FMP17], among others. The setting is as follows: let π be an automorphic representation of $\mathrm{GL}_2(\mathbb{A}_F)$ for some number field F, let E be a quadratic extension of F embedded in a quaternion algebra D defined over F, let $\Omega : E^\times\backslash\mathbb{A}_E^\times \to \mathbb{C}^\times$ be a unitary Hecke character for which $\Omega|_{\mathbb{A}_F^\times}$ is equal to the central character of π, and let ϕ be a test vector in the automorphic representation π^{D} of $\mathrm{D}^\times(\mathbb{A}_F)$ corresponding to π via the Jacquet–Langlands correspondence. Then we define the adèlic period integral

$$\mathscr{P}_\Omega^{\mathrm{D}}(\phi) := \int\limits_{\mathbb{A}_F^\times E^\times\backslash\mathbb{A}_E^\times} \phi(x)\Omega^{-1}(x)\,dx.$$

Note that implicitly this depends on a choice of embedding $\mathbb{A}_E^\times \hookrightarrow \mathrm{D}^\times(\mathbb{A}_F)$, which is suppressed in the notation, as well as a choice of normalisation of the measure dx.

An amazing result of Waldspurger [Wal85] is the following formula;

$$\frac{|\mathscr{P}_\Omega^{\mathrm{D}}(\phi)|^2}{\langle\phi,\phi\rangle} = c_{\Omega,\phi}\frac{\Lambda\left(\frac{1}{2},\pi_E\otimes\Omega\right)}{\Lambda(1,\mathrm{sym}^2\pi)}, \tag{4.10}$$

where ϕ is *any* nonzero test vector in π^{D} and $c_{\Omega,\phi}$ is a finite product of local factors. Here π_E denotes the base change of π to an automorphic representation of $\mathrm{GL}_2(\mathbb{A}_E)$; alternatively, we may write $\Lambda(s,\pi_E\otimes\Omega) = \Lambda(s,\pi\otimes\pi_\Omega)$, where π_Ω denote the automorphic induction of Ω to an automorphic representation of $\mathrm{GL}_2(\mathbb{A}_F)$.

4.2.2. *An explicit formula.* For applications in analytic number theory, it is essential that we have at our disposal a completely explicit form of Waldspurger's formula (4.10). Martin and Whitehouse [MW09, Theorem 4.1] provide such a formula for a *specific* choice of test vector $\phi \in \pi^{\mathrm{D}}$ (assuming some local compatibility between Ω and π, which were slightly relaxed by File, Martin, and Pitale [FMP17]). For this specific choice of test vector, the local factors (whose product we denoted by $c_{\Omega,\phi}$ in (4.10)) are described in [MW09, Section 4.2].

For our application, we can restrict to the case where $F = \mathbb{Q}$, $E = \mathbb{Q}(\sqrt{D})$ with $D > 1$ a positive fundamental discriminant, $\pi = \pi_f = \pi_\infty \otimes \bigotimes_p \pi_p$ is a cuspidal automorphic representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ associated to a Hecke–Maaß newform f of weight 0, principal nebentypus, and squarefree level q for which every prime diving q splits in E, and Ω is the idèlic lift of a narrow class character χ. With this choice of data, the quaternion algebra D is simply the matrix algebra $\mathrm{Mat}_{2\times 2}$, so that $\pi^{\mathrm{D}} = \pi$. We will shorten notation and write $\mathscr{P}_\Omega(\phi) := \mathscr{P}^{\mathrm{D}}_\Omega(\phi)$ in this case.

The choice of test vector $\phi \in \pi$ used in [MW09] is characterised by some local compatibilities with the Hecke character Ω and thus implicitly depends on the choice of embedding $\Psi_{\mathbb{A}_\mathbb{Q}} = (\Psi_\infty, \Psi_2, \Psi_3, \ldots) : \mathbb{A}_E \hookrightarrow \mathrm{D}(\mathbb{A}_\mathbb{Q})$. The properties that characterise the local test vectors ϕ_p and ϕ_∞ are described in [MW09, p. 172] and are as follows.

- At a nonarchimedean place p, Martin and Whitehouse pick $\phi_p \in \pi_p$ to be the unique nonzero and invariant under the units $R^\times$ of a certain order R in the local quaternion algebra (which determines ϕ_p up to scaling). In our setting, R is simply the Eichler order in $\mathrm{GL}_2(\mathbb{Q}_p)$ of reduced discriminant $p^{c(\pi_p)}$ such that
$$R \cap \Psi_p(E_p) = \Psi_p(\mathcal{O}_{E_p}),$$
where $c(\pi_p)$ denotes the conductor exponent of π_p and $E_p := E \otimes \mathbb{Q}_p$.
- At the archimedean place, we let $K_\infty \cong \mathrm{O}(2)$ be a maximal compact subgroup of $\mathrm{GL}_2(\mathbb{R})$ such that $K_\infty \cap \Psi_\infty(E_\infty^\times) \cong (\mathbb{Z}/2\mathbb{Z})^2$ is a maximal compact subgroup of $\Psi_\infty(E_\infty^\times) \cong (\mathbb{R}^\times)^2$, where $E_\infty := E \otimes \mathbb{R} \cong \mathbb{R}^2$. Martin and Whitehouse pick ϕ_∞ such that $K_\infty \cap \Psi_\infty(E_\infty^\times)$ acts (via π) on ϕ_∞ in the same way as $\Omega_\infty : E_\infty^\times \to \mathbb{C}^\times$ and ϕ_∞ lies in the *minimal* such K_∞-type in the sense of Popa [Pop08, Theorem 1] (which also uniquely determines ϕ_∞ up to scaling).

In our application, we slightly modify the choice of test vector ϕ_∞.

In order to get an explicit formula, we now need to specify an embedding $\Psi_{\mathbb{A}_\mathbb{Q}} : \mathbb{A}_E \hookrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{A}_\mathbb{Q})$ and then determine which choice of local test vectors ϕ_p the above described conditions imply.

4.2.3. *A specific test vector.* We construct an embedding $\Psi_{\mathbb{A}_\mathbb{Q}}$ using an oriented optimal embedding $\Psi : E \hookrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{Q})$ of level q as described in (2.3) associated to a Heegner form $Q = [a, b, c] \in \mathcal{Q}_D(q)$ as in (2.1). By tensoring with $\mathbb{A}_\mathbb{Q}$, we get an embedding
$$\Psi_{\mathbb{A}_\mathbb{Q}} = (\Psi_\infty, \Psi_2, \Psi_3, \Psi_5, \ldots) : \mathbb{A}_E \hookrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{A}_\mathbb{Q}).$$

Since Ψ is an *optimal* embedding, the Eichler order R is exactly the standard order of level $p^{c(\pi_p)}$ for each prime p, so that
$$\prod_p R^\times = K_0(q),$$
where $K_0(q) \subset \mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ is the congruence subgroup of level q as in (3.9). This means that we can choose the local component of ϕ at each finite prime to be the same as those of the adèlisation ϕ_{F_2} of the Maaß cusp form $F_2 = R_0 f$ of weight 2.

At the archimedean place, there is the slight complication that $\Psi_\infty : E_\infty \hookrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{R})$ is *not* the diagonal embedding. If, however, we conjugate Ψ_∞ by the matrix $\gamma_\infty \in \mathrm{GL}_2(\mathbb{R})$ given by

$$\gamma_\infty := \begin{pmatrix} b+\sqrt{D} & b-\sqrt{D} \\ -2a & -2a \end{pmatrix},$$

where $(a,b,c) \in \mathbb{Z}^3$ are associated to Ψ as in Section 2.1.3, then we obtain the diagonal embedding, namely

$$(4.11) \qquad \begin{aligned}(\gamma_\infty \cdot \Psi_\infty)(x+\sqrt{D}y, x-\sqrt{D}y) &:= \gamma_\infty^{-1}\Psi_\infty(x+\sqrt{D}y, x-\sqrt{D}y)\gamma_\infty \\ &= \begin{pmatrix} x+\sqrt{D}y & 0 \\ 0 & x-\sqrt{D}y \end{pmatrix}.\end{aligned}$$

We shall only consider oriented optimal embeddings Ψ of level q for which $\gamma_\infty \in \mathrm{GL}_2^+(\mathbb{R})$, which is to say that $a < 0$. As we shall show in Lemma 4.20, this is without loss of generality, for then $\mathscr{P}_\Omega(\phi)$ turns out to be independent of the choice of oriented optimal embedding Ψ of level q within an the equivalence class of embeddings modulo the action of $\Gamma_0(q)$, and every equivalence class of optimal embeddings of level q contains an embedding for which $a < 0$.

With this in mind, Martin and Whitehouse choose the local component of ϕ at the archimedean place to be

$$\begin{cases} \pi_\infty(\gamma_\infty)\phi_{F_0,\infty} & \text{if } \epsilon_f = \chi(J), \\ \pi_\infty(\gamma_\infty)\phi_{F_2,\infty} - \pi_\infty(\gamma_\infty)\phi_{F_{-2},\infty} & \text{if } \epsilon_f = -\chi(J), \end{cases}$$

where $\phi_{F_0,\infty}$, $\phi_{F_2,\infty}$, and $\phi_{F_{-2},\infty}$ are the local components of the adèlisations ϕ_{F_0}, ϕ_{F_2}, and $\phi_{F_{-2}}$ of $F_0 = f$, $F_2 = R_0 f$, and $F_{-2} = L_0 f$ respectively. We instead merely take the local component of ϕ at the archimedean place to be $\pi_\infty(\gamma_\infty)\phi_{F_2,\infty}$.

Altogether, the above implies that when using the embedding

$$\Psi_{\mathbb{A}_\mathbb{Q}} : \mathbb{A}_E \hookrightarrow \mathrm{Mat}_{2\times 2}(\mathbb{A}_\mathbb{Q}),$$

our test vector is

$$\phi = \pi(\gamma_\infty)\phi_{F_2},$$

where we view $\gamma_\infty \in \mathrm{GL}_2^+(\mathbb{R})$ as an element of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$.

4.3. **A formula for certain adèlic periods.** Let $\phi_{F_2} : \mathrm{GL}_2(\mathbb{A}_\mathbb{Q}) \to \mathbb{C}$ denote the adèlic lift of $F_2 := R_0 f \in \mathscr{C}_2(\Gamma_0(q))$, which is an element of the cuspidal automorphic representation $\pi = \pi_f$ of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ associated to f. Let $\Omega \in \widehat{E^\times \backslash \mathbb{A}_E^1}$ be the idèlic lift of χ, so that Ω is a unitary Hecke character that is unramified at every nonarchimedean place of E and has local components at the two archimedean places of E of the form $(\mathrm{sgn}^{\kappa_\Omega}, \mathrm{sgn}^{\kappa_\Omega})$ with $\kappa_\Omega \in \{0,1\}$ such that $(-1)^{\kappa_\Omega} = \chi(J)$. We study the adèlic period integral

$$(4.12) \qquad \mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2}) := \int\limits_{\mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times} \phi_{F_2}(\Psi_{\mathbb{A}_\mathbb{Q}}(x)\gamma_\infty)\Omega^{-1}(x)\, dx.$$

The measure dx is normalised such that $\mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times$ has volume

$$2\Lambda L(1,\chi_D) = 2L(1,\chi_D),$$

where

$$\Lambda(s,\chi_D) := \pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)L(s,\chi_D).$$

Lemma 4.13. *We have that*

$$|\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2})|^2 = \begin{cases} \dfrac{2}{q\sqrt{D}}\dfrac{\Gamma\left(\frac{3}{4}+\frac{it_f}{2}\right)^2\Gamma\left(\frac{3}{4}-\frac{it_f}{2}\right)^2}{\Gamma\left(\frac{1}{2}+it_f\right)\Gamma\left(\frac{1}{2}-it_f\right)}\dfrac{L\left(\frac{1}{2},f\otimes\Theta_\chi\right)}{L(1,\mathrm{sym}^2 f)} & \text{if } \epsilon_f = -\chi(J), \\ 0 & \text{if } \epsilon_f = \chi(J). \end{cases}$$

Proof. We apply [MW09, Theorem 4.1] with $F = \mathbb{Q}$, $E = \mathbb{Q}(\sqrt{D})$, $\varphi = \pi(\gamma_\infty)\phi_{F_2}$ (so that $\pi = \pi_f$), and Ω as above. With this choice of data, we have that $S'(\pi) = S(\Omega) = \emptyset$, $\mathrm{Ram}(\pi) = \{p : p \mid q\}$, $\Delta_F = 1$, $\Delta_E = D$, $c(\Omega) = 1$, and $\Sigma_F^\infty = \{\infty\}$ in the notation of [MW09, Theorem 4.1].

There is a slight caveat, however; [MW09, Theorem 4.1] does not quite apply, since although the automorphic form ϕ_{F_2} has the same local Whittaker functions at every nonarchimedean place to that appearing in [MW09, Theorem 4.1] (compare Section 3.2.4 to [MW09, Section 2]), the local Whittaker function at the archimedean place, as in Section 3.2.5, has a slightly different form than that appearing in [MW09, Theorem 4.1]. This issue is readily circumvented: we replace the term $C_\infty(E,\pi,\Omega)$ appearing in [MW09, Theorem 4.1] with its definition in [MW09, Section 4.2.2] in terms of local archimedean L-functions and $\widetilde{J}_{\pi_\infty}(f_\infty)$, where now

$$\widetilde{J}_{\pi_\infty}(f_\infty) = \frac{\left|\displaystyle\int_{\mathbb{R}^\times} W_\infty^2\begin{pmatrix} a & 0 \\ 0 & 1\end{pmatrix}\mathrm{sgn}^{\kappa_\Omega}(a)\,d^\times a\right|^2}{\displaystyle\int_{\mathbb{R}^\times}\left|W_\infty^2\begin{pmatrix} a & 0 \\ 0 & 1\end{pmatrix}\right|^2 d^\times a}.$$

This local distribution is just as in [MW09, Section 3.3] except that we have projected onto the local Whittaker function $W_\infty^2 \in \mathcal{W}(\pi_\infty,\psi_\infty)$ associated to ϕ_{F_2} instead of the local Whittaker function $W_\infty \in \mathcal{W}(\pi_\infty,\psi_\infty)$ for which the numerator is equal to the local archimedean L-function and additionally satisfying $W_\infty\left(\begin{smallmatrix} a & 0 \\ 0 & 1\end{smallmatrix}\right) = (-1)^{\kappa_\Omega}W_\infty\left(\begin{smallmatrix} -a & 0 \\ 0 & 1\end{smallmatrix}\right)$ for all $a \in \mathbb{R}^\times$.

With this minor modification at the archimedean place, we deduce from [MW09, Theorem 4.1] that

$$|\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2})|^2 = \frac{\pi}{2\sqrt{D}}\prod_{p\mid q}\frac{1}{1-p^{-1}}\widetilde{J}_{\pi_\infty}(f_\infty)\frac{L\left(\frac{1}{2},\pi_f\otimes\pi_\Omega\right)}{L(1,\mathrm{sym}^2\pi_f)}\int_{\mathrm{Z}(\mathbb{A}_\mathbb{Q})\,\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})}|\phi_{F_2}(g)|^2\,dg, \tag{4.14}$$

where $\mathrm{Z}(\mathbb{A}_\mathbb{Q})$ denotes the centre of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ and the measure dg is normalised such that

$$\mathrm{vol}(\mathrm{Z}(\mathbb{A}_\mathbb{Q})\,\mathrm{GL}_2(\mathbb{Q})\backslash\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})) = 2\Lambda^q(2) = \frac{\pi}{3}\prod_{p\mid q}\left(1-\frac{1}{p^2}\right) \tag{4.15}$$

since the Tamagawa number of GL_2 is 2, where

$$\Lambda^q(s) := \pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s)\prod_{p\mid q}\left(1-\frac{1}{p^s}\right).$$

From (3.18) and (3.19) with $\kappa = \kappa_f$, $\kappa' = \kappa_\Omega$, $t = t_f$, and $s = 1/2$,

(4.16)

$$\widetilde{J}_{\pi_\infty}(f_\infty) = \begin{cases} \dfrac{4}{\pi\left(\frac{1}{4}+t_f^2\right)}\dfrac{\Gamma\left(\frac{3}{4}+\frac{it_f}{2}\right)^2\Gamma\left(\frac{3}{4}-\frac{it_f}{2}\right)^2}{\Gamma\left(\frac{1}{2}+it_f\right)\Gamma\left(\frac{1}{2}-it_f\right)} & \text{if } \kappa_f \equiv \kappa_\Omega + 1 \pmod 2, \\ 0 & \text{if } \kappa_f \equiv \kappa_\Omega \pmod 2, \end{cases}$$

while

(4.17)

$$\begin{aligned}\int\limits_{\mathrm{Z}(\mathbb{A}_\mathbb{Q})\,\mathrm{GL}_2(\mathbb{Q})\backslash\,\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})} |\phi_{F_2}(g)|^2\,dg &= \frac{1}{q}\prod_{p\mid q}\left(1-\frac{1}{p}\right)\int_{\Gamma_0(q)\backslash\mathbb{H}}|(R_0f)(z)|^2\,d\mu(z) \\ &= \frac{1}{q}\prod_{p\mid q}\left(1-\frac{1}{p}\right)\left(\frac{1}{4}+t_f^2\right)\int_{\Gamma_0(q)\backslash\mathbb{H}}|f(z)|^2\,d\mu(z),\end{aligned}$$

where the first equality holds via the strong approximation theorem, (3.10), while the second equality follows from [DFI02, (4.38)]. We can check that the normalisation of measures in the first equality is correct by replacing ϕ_{F_2} with the constant function 1 and recalling (4.15) and the fact $\mathrm{vol}(\Gamma_0(q)\backslash\mathbb{H}) = \frac{\pi}{3}\nu(q)$. Combining (4.14), (4.16), and (4.17) and noting that $L(s,\pi_f\otimes\pi_\Omega) = L(s, f\otimes\Theta_\chi)$ and $L(s,\mathrm{sym}^2\,\pi_f) = L(s,\mathrm{sym}^2\,f)$, we obtain the result. □

4.4. **From adèlic period integrals to cycle integrals.** We relate the adèlic period integral (4.12) to a certain sum of cycle integrals over oriented geodesics in $\Gamma_0(q)\backslash\mathbb{H}$ indexed by narrow ideal classes. We first define these cycle integrals and show that they are well-defined. Given a Heegner form $Q \in \mathcal{Q}_D(q)$, we consider the cycle integral

$$\int_{z_Q}^{\gamma_Q z_Q}(R_0f)(z)\,\frac{dz}{\Im(z)}, \tag{4.18}$$

where z_Q and $\gamma_Q z_Q$ are as in (2.5) and the contour of integration is the geodesic segment between these two points.

Lemma 4.19. *For all $\gamma \in \Gamma_0(q)$, the cycle integral (4.18) is invariant under replacing Q with $\gamma\cdot Q$.*

For this reason, we may write (4.18) as

$$\int_{\mathcal{C}_A(q)}(R_0f)(z)\,\frac{dz}{\Im(z)}$$

without ambiguity, where $\mathcal{C}_A(q)$ denotes the oriented geodesic in $\Gamma_0(q)\backslash\mathbb{H}$ associated to a narrow ideal class A corresponding to Q as in Section 2.1.4, since this cycle integral is independent of the choice of Heegner form Q associated to A.

Proof. Suppose that $Q' = \gamma \cdot Q$ for some $\gamma \in \Gamma_0(q)$. We make the change of variables $z \mapsto \gamma^{-1} z$ in (4.18). The integrand remains unchanged since

$$(R_0 f)(\gamma z) = j_\gamma(z)^2 (R_0 f)(z), \qquad \frac{d}{dz}(\gamma z) = \frac{\Im(\gamma z)}{\Im(z)} j_\gamma(z)^{-2}.$$

It is easily checked that $\gamma^{-1}\gamma_Q\gamma = \gamma_{Q'}$, and so the new contour of integration is the geodesic segment from γz_Q to $\gamma_{Q'}\gamma z_Q$ on the semicircle (2.4) associated to Q'. Further changes of variables by powers of $\gamma_{Q'}$ rotate the contour of integration along this semicircle while leaving the integrand intact, and so there is an appropriate power of $\gamma_{Q'}$ for which the resulting geodesic segment intersects nontrivially with the geodesic segment from $z_{Q'}$ to $\gamma_{Q'} z_{Q'}$. We then break up the integral into two parts, and for the part that does not intersect this geodesic segment, we make one last change of variables by either $\gamma_{Q'}$ or $\gamma_{Q'}^{-1}$ as appropriate; recombining, we obtain (4.18) with Q' in place of Q. □

Lemma 4.20. *We have that*

$$\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2}) = -\frac{i\overline{\chi}(A_\Psi)}{\sqrt{D}} \sum_{A \in \mathrm{Cl}_D^+} \overline{\chi}(A) \int_{\mathcal{C}_A(q)} (R_0 f)(z) \, \frac{dz}{\Im(z)},$$

where $A_\Psi \in \mathrm{Cl}_D^+$ is the element of the narrow class group associated to the oriented optimal embedding Ψ. In particular, $\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2})$ is independent of the choice of oriented optimal embedding Ψ of level q within an equivalence class of embeddings modulo the action of $\Gamma_0(q)$.

The proof of Lemma 4.20 closely follows that of a similar result of Popa [Pop06, Section 6].

Proof. Since Ω is the idèlic lift of a narrow class group character, it is trivial on both $\widehat{\mathcal{O}}_E^\times$ and $\mathbb{A}_\mathbb{Q}^\times$. Furthermore, we have the inclusion $\Psi_{\mathbb{A}_\mathbb{Q}}(\widehat{\mathcal{O}}_E^\times) \subset K_0(q)$ since Ψ is an optimal embedding. As the newform f has level q, it follows that both $\pi(\gamma_\infty)\phi_{F_2}$ and Ω are well-defined on the double quotient $\mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times / \widehat{\mathcal{O}}_E^\times$. Since $\widehat{\mathcal{O}}_E^\times$ has measure 1, we deduce that

$$\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2}) = \int\limits_{\mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times / \widehat{\mathcal{O}}_E^\times} \phi_{F_2}(\Psi_{\mathbb{A}_\mathbb{Q}}(x)\gamma_\infty)\Omega^{-1}(x) \, dx.$$

Via the strong approximation theorem, we have the decomposition

$$\mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times / \widehat{\mathcal{O}}_E^\times \cong \bigsqcup_{A \in \mathrm{Cl}_D^+} A \cdot \epsilon_D^\mathbb{Z} \backslash E_\infty^1,$$

where $A = (A_v) \in \mathbb{A}_E^\times$ runs through a set of finite idèle representatives of the narrow class group Cl_D^+ (so that $A_v = 1$ if v is an archimedean place of E), which we freely identify with elements of the narrow class group, while $E_\infty^1 := \{(t, t^{-1}) \in E_\infty : t > 0\} \cong \mathbb{R}_+^\times$, which we view as a subgroup of $\mathbb{A}_E^\times$ via the embedding $(t, t^{-1}) \mapsto (t, t^{-1}, 1, 1, \ldots)$, and $\epsilon_D^\mathbb{Z} := \{(\epsilon_D^m, \epsilon_D^{-m}) \in E_\infty : m \in \mathbb{Z}\}$. Thus every $x \in \mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times / \widehat{\mathcal{O}}_E^\times$

can be written as $x = A(t, t^{-1}, 1, 1, \ldots)$ with $t \in [1, \epsilon_D)$. From this, we may write

$$\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2}) = \frac{2}{\sqrt{D}} \sum_{A \in \mathrm{Cl}_D^+} \overline{\chi}(A) \int_1^{\epsilon_D} \phi_{F_2}(\Psi_{\mathbb{A}_\mathbb{Q}}(A)\Psi_\infty(t, t^{-1})\gamma_\infty) \, d^\times t.$$

We can check that the normalisation of measures here is correct by replacing $\pi(\gamma_\infty)\phi_{F_2}$ with the constant function 1 and taking χ to be the trivial character, noting that $\mathrm{vol}(\mathbb{A}_\mathbb{Q}^\times E^\times \backslash \mathbb{A}_E^\times)$ is $2L(1, \chi_D)$, whereas $h_D^+ \log \epsilon_D = \sqrt{D} L(1, \chi_D)$ by the narrow class number formula.

Let g_A denote the inverse of the $\mathrm{GL}_2^+(\mathbb{R})$-component in the representation (3.10). By the definition (3.11) and (3.8) of the adèlic lift together with our assumption that $\det \gamma_\infty > 0$, we have that

$$\begin{aligned} &\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2}) \\ &= \frac{2}{\sqrt{D}} \sum_{A \in \mathrm{Cl}_D^+} \overline{\chi}(A) \int_1^{\epsilon_D} j_{g_A^{-1}\Psi_\infty(t,t^{-1})\gamma_\infty}(i)^{-2} (R_0 f)(g_A^{-1}\Psi_\infty(t, t^{-1})\gamma_\infty i) \, d^\times t. \end{aligned} \tag{4.21}$$

We make the change of variables $z = g_A^{-1}\Psi_\infty(t, t^{-1})\gamma_\infty i = g_A^{-1}\gamma_\infty(it^2)$; the contour of integration in (4.21) then becomes the oriented geodesic segment from $g_A^{-1}\gamma_\infty i = g_A^{-1} z_Q$ to $g_A^{-1}\gamma_Q z_Q$ parametrised by $g_A^{-1}\Psi_\infty(t, t^{-1}) z_Q$, where z_Q and $\gamma_Q z_Q$ are as in (2.5) with $Q = Q_\Psi = [a, b, c]$ the Heegner form associated to the optimal embedding Ψ. We have that

$$d^\times t = -\frac{i}{2} j_{\gamma_\infty^{-1} g_A}(z)^{-2} \frac{dz}{\Im(z)},$$

since $t^2 = -i\gamma_\infty^{-1} g_A z$ and

$$\frac{d}{dz}(gz) = \frac{\Im(gz)}{\Im(z)} j_g(z)^{-2}$$

for any $g \in \mathrm{GL}_2(\mathbb{R})$, while the cocycle relation $j_{g_1 g_2}(z) = j_{g_2}(z) j_{g_1}(g_2 z)$ implies that

$$j_{g_A^{-1}\Psi_\infty(t,t^{-1})\gamma_\infty}(i) j_{\gamma_\infty^{-1} g_A}(z) = j_{\gamma_\infty^{-1}\Psi_\infty(t,t^{-1})\gamma_\infty}(i) = 1,$$

where the last equality follows upon recalling (3.1) and (4.11). We deduce that (4.21) is equal to

$$-\frac{i}{\sqrt{D}} \sum_{A \in \mathrm{Cl}_D^+} \overline{\chi}(A) \int_{g_A^{-1} z_Q}^{g_A^{-1}\gamma_Q z_Q} (R_0 f)(z) \, \frac{dz}{\Im(z)}.$$

It is shown in [Pop06, Theorem 6.2.2 (i)] that as A runs through the narrow class group Cl_D^+, $g_A^{-1}\Psi g_A$ runs through a set of representatives of equivalence classes of oriented optimal embeddings of level q modulo the action of $\Gamma_0(q)$. Furthermore, letting Q' denote the Heegner form of level q associated to the oriented optimal embedding $g_A^{-1}\Psi g_A$, we have that the contour of integration from $g_A^{-1} z_Q$ and $g_A^{-1}\gamma_Q z_Q = \gamma_{Q'} g_A^{-1} z_Q$ is a geodesic segment of length $2 \log \epsilon_D$ of the semicircle (2.4) associated to Q', and hence

$$\int_{g_A^{-1} z_Q}^{g_A^{-1}\gamma_Q z_Q} (R_0 f)(z) \, \frac{dz}{\Im(z)} = \int_{\mathcal{C}_{A'}(q)} (R_0 f)(z) \, \frac{dz}{\Im(z)}$$

by the proof of Lemma 4.19, where $A' \in \mathrm{Cl}_D^+$ is the element of the narrow class group associated to Q'. It remains to note that by [Pop06, Theorem 6.2.2 (ii)], the oriented optimal embedding $g_A^{-1}\Psi g_A$ is associated to $AA_\Psi \in \mathrm{Cl}_D^+$, where A_Ψ is the element of Cl_D^+ corresponding to Ψ. □

4.5. **Proofs of Propositions 4.2 and 4.7.** We are finally in a position to prove Proposition 4.2. For this we restrict to $q = 1$ in the sections above.

Proof of Proposition 4.2. Lemmata 4.20 with $q = 1$ and [DIT16, Lemma 1] imply that the Weyl sum $W_{\chi,f}$ defined in (4.1) satisfies the identity

$$W_{\chi,f} = -\frac{i\overline{\chi}(A_\Psi)\sqrt{D}}{\frac{1}{4}+t_f^2}\overline{\mathscr{P}_\Omega(\pi(\gamma_\infty)\phi_{F_2})},$$

at which point the result follows from Lemma 4.13. □

The proof of Proposition 4.7 is a little simpler, since we can circumvent the adèlic formulation of this Weyl sum.

Proof of Proposition 4.7. From [DIT16, (7.3)], we have that

$$\sum_{A\in \mathrm{Cl}_D^+} \chi(A)\int_{\mathcal{C}_A(1)} (R_0E)(z,s)\,\frac{dz}{\Im(z)} = (1-\chi(J))D^{\frac{s}{2}}\frac{\Gamma\left(\frac{s+1}{2}\right)^2}{\Gamma(s)}\frac{L(s,\chi)}{\zeta(2s)}$$

for $\Re(s) > 1$. Via analytic continuation, this identity extends to $s = 1/2 + it$. The result then follows via [DIT16, Lemma 1]. □

5. Proof of Theorem 1.1

Proof of Theorem 1.1. Using the Weyl equidistribution criterion together with the lower-bound (2.6), it suffices to show that for each $f \in \mathcal{B}_0^*(\Gamma)$, and for each $t \in \mathbb{R}$,

$$\sum_{A\in CH}\int_{\Gamma_A\backslash\mathcal{N}_A} f(z)\,d\mu(z) = o_{t_f}\left(\frac{\#H}{h_D^+}\frac{\sqrt{D}L(1,\chi_D)}{h_D^+\log D}\right),$$

$$\sum_{A\in CH}\int_{\Gamma_A\backslash\mathcal{N}_A} E\left(z,\frac{1}{2}+it\right)\,d\mu(z) = o_t\left(\frac{\#H}{h_D^+}\frac{\sqrt{D}L(1,\chi_D)}{\log D}\right),$$

where we have used the narrow class number formula $h_D^+\log\epsilon_D = \sqrt{D}L(1,\chi_D)$. Via character orthogonality these expressions are respectively equal to

$$\frac{\#H}{h_D^+}\sum_{\chi\in H^\perp}\overline{\chi}(C)W_{\chi,f}, \qquad \frac{\#H}{h_D^+}\sum_{\chi\in H^\perp}\overline{\chi}(C)W_{\chi,t},$$

where the Weyl sums $W_{\chi,f}$ and $W_{\chi,t}$ are as in (4.1) and (4.6). From Propositions 4.2 and 4.7, the bounds $L(1,\mathrm{sym}^2 f) \gg 1/\log(|t_f|+3)$ and $\zeta(1+2it) \gg 1/\log(|t|+3)$, and Stirling's formula, we have that

$$|W_{\chi,f}|^2 \ll \frac{\sqrt{D}\log((|t_f|+3))}{(|t_f|+1)^3}L\left(\frac{1}{2},f\otimes\Theta_\chi\right),$$

$$|W_{\chi,t}|^2 \ll \frac{\sqrt{D}\log(|t|+3)^2}{(|t|+1)^3} L\left(\frac{1}{2}+it,\Theta_\chi\right) L\left(\frac{1}{2}-it,\Theta_\chi\right).$$

We obtain Theorem 1.1 for $\delta < \frac{1}{4}$ under the assumption of the generalised Lindelöf hypothesis, the fact that $\#H^\perp = h_D^+/\#H$, and the (ineffective) Siegel bound $L(1,\chi_D) \gg_\varepsilon D^{-\varepsilon}$. Furthermore, Theorem 1.1 for $\delta < \frac{625}{3309568}$ holds unconditionally due to the estimates

$$|W_{\chi,f}|^2 \ll_{q_1,t_f,\varepsilon} D^{1-\frac{625}{1654784}+\varepsilon},$$
$$|W_{\chi,t}|^2 \ll_t D^{1-\frac{2}{1889}}.$$

These bounds for the Weyl sums $W_{\chi,f}$ and $W_{\chi,t}$ follow from the subconvex bounds

$$L\left(\frac{1}{2}, f\otimes\Theta_\chi\right) \ll_{q_1,t_f,\varepsilon} D^{\frac{1}{2}-\frac{625}{1654784}+\varepsilon}, \tag{5.1}$$

$$L\left(\frac{1}{2}+it,\Theta_\chi\right) \ll_t D^{\frac{1}{4}-\frac{1}{1889}}. \tag{5.2}$$

The first bound is due to Harcos and Michel [HM06, Theorem 1], while this second is due to Blomer, Harcos, and Michel [BHM07, Theorem 2]. □

Remark 5.3. Improvements of the bound (5.2) exist in the literature (see, for example, [BlKh19, Theorem 1] for D prime); the obstacle in unconditionally enlarging the range of δ in Theorem 1.1 is an improvement of the bound (5.1).

Proof of Proposition 1.3. It follows from [DIT16, (2.5)] that the oriented closed geodesics $\mathcal{C}_A(1)$ and $\mathcal{C}_{JA^{-1}}(1)$ are the same curve with opposite orientations, which means that $\Gamma_A\backslash\mathcal{N}_A$ and $\Gamma_{JA^{-1}}\backslash\mathcal{N}_{JA^{-1}}$ cover $\Gamma\backslash\mathbb{H}$ evenly. Thus (1.2) holds if $JA^{-1} \in CH$ for every $A \in CH$. This condition is met precisely when $C^2J \in H$. □

References

[AE16] Manny Aka and Manfred Einsiedler, "Duke's Theorem for Subcollections", *Ergodic Theory and Dynamical Systems* **36**:2 (2016), 335–342.

[BHM07] Valentin Blomer, Gergely Harcos, and Philippe Michel, "Bounds for Modular L-Functions in the Level Aspect", *Annales Scientifiques de l'École Normale Supérieure* **40**:5 (2007), 697–740.

[BlKh19] Valentin Blomer and Rizwanur Khan, "Uniform Subconvexity and Symmetry Breaking Reciprocity", *Journal of Functional Analysis* **276**:7 (2019), 2315–2358.

[Bou05] J. Bourgain, "Mordell's Exponential Sum Estimate Revisited", *Journal of the American Mathematical Society* **18**:2 (2005), 477–499.

[BoKo17] Jean Bourgain and Alex Kontorovich, "Beyond Expansion II: Low-Lying Fundamental Geodesics", *Journal of the European Mathematical Society* **19**:5 (2017) 1331–1359.

[Dar94] Henri Darmon, "Heegner Points, Heegner Cycles, and Congruences", in *Elliptic Curves and Related Topics*, editors Hershy Kisilevsky and M. Ram Murty, CRM Proceedings & Lecture Notes **4**, American Mathematical Society, Providence, 1994, 45–59.

[DFI02] W. Duke, J. B. Friedlander, and H. Iwaniec, "The Subconvexity Problem for Artin L-Functions", *Inventiones Mathematicae* **149**:3 (2002), 489–577.

[DIT16] W. Duke, Ö. Imamoḡlu, and Á. Tóth, "Geometric Invariants for Real Quadratic Fields", *Annals of Mathematics* **184**:3 (2016), 949–990.

[FMP17] Daniel File, Kimball Martin, and Ameya Pitale, "Test Vectors and Central L-Values for GL(2)", *Algebra and Number Theory* **11**:2 (2017), 253–318.

[God18] Roger Godement, *Notes on Jacquet–Langlands' Theory. With Commentaries by Robert Langlands and Hervé Jacquet*, Classical Topics in Mathematics **8**, Higher Education Press, Beijing, 2018.

[GH11] Dorian Goldfeld and Joseph Hundley, *Automorphic Representations and L-Functions for the General Linear Group, Volume I*, Cambridge Studies in Advanced Mathematics **129**, Cambridge University Press, Cambridge, 2011.

[GR15] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products, Eighth Edition*, editors Daniel Zwillinger and Victor Moll, Academic Press, Burlington, 2015.

[Gro88] Benedict H. Gross, "Local Orders, Root Numbers, and Modular Curves", *American Journal of Mathematics* **110**:6 (1988), 1153–1182.

[GKZ87] B. Gross, W. Kohnen, and D. Zagier, "Heegner Points and Derivatives of L-Series. II", *Mathematische Annalen* **278**:1–4 (1987), 497–562.

[HM06] Gergely Harcos and Philippe Michel, "The Subconvexity Problem for Rankin–Selberg L-Functions and Equidistribution of Heegner Points. II", *Inventiones Mathematicae* **163**:3 (2006), 581–655.

[HT14] Roman Holowinsky and Nicolas Templier, "First Moment of Rankin–Selberg Central L-Values and Subconvexity in the Level Aspect", *Ramanujan Journal* **33**:1 (2014), 131–151.

[Hum18] Peter Humphries, "Equidistribution in Shrinking Sets and L^4-Norm Bounds for Automorphic Forms", *Mathematische Annalen* **371**:3–4 (2018), 1497–1543.

[Hum20] Peter Humphries, "Distributing Points on the Torus via Modular Inverses", preprint (2020), 11 pages.

[HK20] Peter Humphries and Rizwanur Khan, "On the Random Wave Conjecture for Dihedral Maaß Forms", *Geometric and Functional Analysis* **30**:1 (2020), 34–125.

[HR19] Peter Humphries and Maksym Radziwiłł, "Optimal Small Scale Equidistribution of Lattice Points on the Sphere, Heegner Points, and Closed Geodesics", preprint (2019), 43 pages.

[JL70] H. Jacquet and R. P. Langlands, *Automorphic Forms on* $\mathrm{GL}(2)$, Lecture Notes in Mathematics **114**, Springer–Verlag, Berlin, 1970.

[LMY13] Sheng-Chi Liu, Riad Masri, and Matthew P. Young, "Subconvexity and Equidistribution of Heegner Points in the Level Aspect", *Compositio Mathematica* **149**:7 (2013), 1150–1174.

[MW09] Kimball Martin and David Whitehouse, "Central L-Values and Toric Periods for GL(2)", *International Mathematics Research Notices* **2009**:1 (2009), 141–191.

[MV06] Philippe Michel and Akshay Venkatesh, "Equidistribution, L-Functions and Ergodic Theory: On Some Problems of Yu. Linnik", in *Proceedings of the International Congress of Mathematicians, Madrid 2006* **II**, editors Marta Sanz-Solé, Javier Soria, Juan Luis Varona, and Joan Verdera, European Mathematical Society, Zürich, 2006, 421–457.

[PY19] Ian Petrow and Matthew P. Young, "A Generalized Cubic Moment and the Petersson Formula for Newforms", *Mathematische Annalen* **373**:1–2 (2019), 287–353.

[PY20] I. Petrow and M. P. Young, "The Weyl Bound for Dirichlet L-Functions of Cube-Free Conductor", to appear in *Annals of Mathematics* (2020), 40 pages.

[Pop06] Alexandru A. Popa, "Central Values of Rankin L-Series over Real Quadratic Fields", *Compositio Mathematica* **142**:4 (2006), 811–866.

[Pop08] Alexandru A. Popa, "Whittaker Newforms for Archimedean Representations of GL(2)", *Journal of Number Theory* **128**:6 (2008), 1637–1645.

[Sch02] Ralf Schmidt, "Some Remarks on Local Newforms for GL(2)", *Journal of the Ramanujan Mathematical Society* **17**:2 (2002), 115–147.

[Ven10] Akshay Venkatesh, "Sparse Equidistribution Problems, Period Bounds and Subconvexity", *Annals of Mathematics* **172**:2 (2010), 989–1094.

[Wal85] J.-L. Waldspurger, "Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie", *Compositio Mathematica* **54**:2 (1985), 173–242.

[You17] Matthew P. Young, "Weyl-Type Hybrid Subconvexity Bounds for Twisted L-Functions and Heegner Points on Shrinking Sets", *Journal of the European Mathematical Society* **19**:5 (2017), 1545–1576.

[You19] Matthew P. Young, "Explicit Calculations with Eisenstein Series", *Journal of Number Theory* **199** (2019), 1–48.

[Zha01] Shou-Wu Zhang, "Gross–Zagier Formula for GL_2", *Asian Journal of Mathematics* **5**:2 (2001), 183–290.

Department of Mathematics, University of Virginia, Charlottesville, VA 22904, USA
E-mail address: pclhumphries@gmail.com

Department of Mathematical Sciences, Copenhagen University, Universitetsparken 5, Copenhagen 2100, Denmark
E-mail address: acnordentoft@outlook.com

PAPER F
HYBRID SUBCONVEXITY FOR CLASS GROUP L-FUNCTIONS AND UNIFORM SUP NORM BOUNDS OF EISENSTEIN SERIES

HYBRID SUBCONVEXITY FOR CLASS GROUP L-FUNCTIONS AND UNIFORM SUP NORM BOUNDS OF EISENSTEIN SERIES

ASBJØRN CHRISTIAN NORDENTOFT

ABSTRACT. In this paper we study hybrid subconvexity bounds for class group L-functions associated to quadratic extensions $K/\mathbb{Q}$ (real or imaginary). Our proof relies on relating the class group L-functions to Eisenstein series evaluated at Heegner points using formulas due to Hecke. The main technical contribution is the following uniform sup norm bound for Eisenstein series;

$$E(z,1/2+it)\ll_\varepsilon y^{1/2}(|t|+1)^{1/3+\varepsilon},\quad y\gg 1,$$

extending work of Blomer and Titchmarsh. Finally we propose a uniform version of the sup norm conjecture for Eisenstein series.

1. INTRODUCTION

This paper is concerned with the family of L-functions $L_K(s,\chi)$ associated to a character χ of the (wide) class group Cl (K) of a quadratic field extension $K/\mathbb{Q}$ (real or imaginary) of discriminant D. One of our results is a hybrid subconvexity bound in terms of the discriminant D and the archimedean parameter t where $s=1/2+it$ (both for individual class group L-functions and for the second moment of the entire family). We will do this by relating the subconvexity bound for class group L-functions to sup norm bounds of Eisenstein series via formulas due to Hecke. Our second main result is what we will call a *uniform sup norm bound* of Eisenstein series.

1.1. **Class group L-functions.** The study of analytic properties of the family of class group L-functions was initiated by Duke, Friedlander and Iwaniec in [6] where they computed the second moment of class group L-functions in the limit $D\to-\infty$. Other notable works on the family of class group L-functions include [2], [7], [4] [20]. Our approach in the imaginary quadratic case is to use a classical formula of Hecke, which relates class group L-functions to Eisenstein series evaluated at Heegner points;

$$L_K(s,\chi)=\frac{2^{s+1}\zeta(2s)|D|^{-s/2}}{\omega_K}\sum_{\mathfrak{a}}\chi(\mathfrak{a})E(z_{\mathfrak{a}},s),\tag{1.1}$$

where the sum runs over a complete set of representatives for the class group of the imaginary quadratic field K of discriminant D, $z_{\mathfrak{a}}\in\mathbb{H}$ is the associated Heegner point and $\omega_K\in\{2,4,6\}$. There is a real quadratic analogue also due to Hecke (see (2.4) below). These formulas give a connection between subconvexity bounds and the so-called *sup norm problem* for Eisenstein series, which we will introduce shortly.

Date: December 4, 2020.
2010 *Mathematics Subject Classification.* 11F03(primary), and 11L07(secondary).

Remark 1.1. The connection between the sup norm problem and subconvexity estimates can be traced back to Sarnak [18, (4.19)]. However this paper together with the recent work of Hu and Saha [10] seem to be the first time sup norm results have been used to obtain new subconvexity results. Hu and Saha apply sup norm bounds of automorphic forms on quaternion algebras (in the depth aspect) to obtain subconvexity estimates in the depth aspect for $L(1/2, f \otimes \theta_\chi)$, where f is a quaternionic automorphic form and θ_χ is an essentially fixed theta series.

Remark 1.2. The formula (1.1) was also the starting point for Templier in [20], where it was combined with equidistribution of Heegner points to give an alternative computation (compared with [6]) of the second moment of the family of class group L-functions as $D \to -\infty$. Similarly Michel and Venkatesh [15] used an analogue of (1.1) in the case of cusp forms due to Zhang [27], [28] to deduce non-vanishing results for the central values of the corresponding Rankin-Selberg L-functions. The approach of Michel and Venkatesh was then applied by Dittmer, Proulx and Seybert in [5] to deduce non-vanishing for class group L-functions as well (their method only shows non-vanishing for one class group character for each K, whereas Blomer in [2] achieved a much stronger result using mollification).

1.2. **The sup norm problem.** Now let $\Gamma_0(1) = \mathrm{PSL}_2(\mathbb{Z})$ and denote by $X_0(1) := \Gamma_0(1)\backslash\mathbb{H}$ the modular curve. The sup norm problem for $X_0(1)$ is concerned with bounds of the following form for some fixed $\theta > 0$;

$$\sup_{z\in C} |u_j(z)| \ll_C t_j^\theta,$$

where u_j is a Maass form of level 1, t_j is the spectral parameter and $C \subset \mathbb{H}$ is compact. The case $\theta = 1/4 + \varepsilon$ is known as the *convexity bound* and is elementary to prove, but it is conjectured [18, Conjecture 3.10] that any $\theta > 0$ is admissible. Iwaniec and Sarnak in their seminal paper [14] were the first to go beyond the convexity bound by proving the bound $\ll_\varepsilon t_j^{5/24+\varepsilon}$.
In this paper we will focus on the analogue for the continuous spectrum which is constituted by Eisenstein series. This means that we are concerned with bounds of the type

$$\sup_{z\in C} |E(z, 1/2 + it)| \ll_C (|t| + 1)^\theta, \tag{1.2}$$

where $\theta > 0$ is fixed and C is compact. In this case the convexity bound is $\theta = 1/2 + \varepsilon$, and again the sup norm conjecture predicts that any $\theta > 0$ is admissible. Iwaniec and Sarnak's method also applies in this case and yields similarly the bound $\ll_\varepsilon (|t| + 1)^{5/12+\varepsilon}$. In [26] Young used a slight modification of the Iwaniec–Sarnak method to prove the bound $\ll_\varepsilon (|t| + 1)^{3/8+\varepsilon}$. In [3] Blomer improved this using exponential sum methods, building on earlier work of Titchmarsh [22], and proved the Weyl type bound $\ll_\varepsilon (|t| + 1)^{1/3+\varepsilon}$. Finally the sup norm problem for Eisenstein series over general number fields has been dealt with in the work of Assing [1].

Plugging Blomer's result into (1.1) yields immediately a subconvexity bound for $L_K(s, \chi)$ in the t-aspect, which recovers a result of Söhne [19] (the conductor of $L_K(s, \chi)$ is $|D|(|t| + 1)^2$, which means that the convexity bound is $\ll_\varepsilon |D|^{1/4+\varepsilon}(|t| +$

$1)^{1/2+\varepsilon}$). If one however wants a hybrid subconvexity estimate, one needs to control the D-dependence in (1.1). This leads to what we will call the *uniform sup norm problem*, which are sup norm bounds with an explicit dependence on z. In a similar vein Huang and Xu [11] studied sup norm bounds of Eisenstein series with level and obtained bounds uniform in both the spectral parameter and the level.

1.3. **Statement of results.** Our first result is the following translation between uniform sup norm bounds of the Eisenstein series $E(z,s)$ and hybrid subconvexity bounds for $L_K(s,\chi)$. Let

$$\mathcal{F} := \{z \in \mathbb{H} \mid -1/2 \leq \operatorname{Re} z \leq 0, |z| \geq 1 \text{ or } 0 < \operatorname{Re} z < 1/2, |z| > 1\}, \tag{1.3}$$

denote the standard fundamental domain for $\Gamma_0(1)$.

Theorem 1.3. *Assume the following uniform bound uniformly for all $z = x+iy \in \mathcal{F}$;*

$$E(z, 1/2+it) \ll y^{\delta}(|t|+1)^{\theta}, \tag{1.4}$$

with $1/2 \leq \delta \leq 1$ and $\theta > 0$. Then it follows that

$$L_K(1/2+it,\chi) \ll_\varepsilon |D|^{1/4+\varepsilon}(|t|+1)^{\theta+\varepsilon}, \tag{1.5}$$

for any $\varepsilon > 0$ and $\chi \in \widehat{\mathrm{Cl}\,(K)}$, a (wide) class group character of a quadratic extension $K/\mathbb{Q}$ (real or imaginary) of discriminant D.
Furthermore it also follows from (1.4) that

$$\sum_{\chi \in \widehat{\mathrm{Cl}\,(K)}} |L_K(1/2+it,\chi)|^2 \ll_\varepsilon |D|^{\delta+\varepsilon}(|t|+1)^{2\theta+\varepsilon}, \tag{1.6}$$

for any $\varepsilon > 0$.

The second part of this paper is concerned with proving a result of the type (1.4). As we will see in Section 3.1 below, the results of Young [26] imply the following.

Theorem 1.4 (M. Young). *For $z \in \mathcal{F}$, the standard fundamental domain (1.3) for $\Gamma_0(1)$, we have*

$$E(z, 1/2+it) \ll_\varepsilon y^{1/2}(|t|+1)^{3/8+\varepsilon}, \tag{1.7}$$

for any $\varepsilon > 0$.

Remark 1.5. Huang and Xu [11, Theorem 1.1] obtained the slightly stronger bound $E(z,s) \ll_\varepsilon y^{1/2} + |t|^{3/8+\varepsilon}$.

It turns out however to be a much more delicate task to upgrade Blomer's Weyl type estimate to a uniform one, which is the main technical contribution of this paper. Our result is the following.

Theorem 1.6. *For $z \in \mathcal{F}$, the standard fundamental domain (1.3) for $\Gamma_0(1)$, we have*

$$E(z, 1/2+it) \ll_\varepsilon y^{1/2}(|t|+1)^{1/3+\varepsilon}, \tag{1.8}$$

for any $\varepsilon > 0$.

Combining this bound with Theorem 1.3, we arrive at the following.

Corollary 1.7. *Let $K/\mathbb{Q}$ be a quadratic extension (real or imaginary) of discriminant D and χ a (wide) class group character of K. Then*

$$L_K(1/2+it,\chi) \ll_\varepsilon |D|^{1/4+\varepsilon}\,(|t|+1)^{1/3+\varepsilon}, \tag{1.9}$$

and

$$\sum_{\chi\in\widehat{\mathrm{Cl}\,(K)}} |L_K(1/2+it,\chi)|^2 \ll_\varepsilon |D|^{1/2+\varepsilon}\,(|t|+1)^{2/3+\varepsilon}, \tag{1.10}$$

for any $\varepsilon > 0$.

Remark 1.8. Observe that for imaginary quadratic fields, (1.10) corresponds to Lindelöf on average in the D-aspect, since $h(K) \gg |D|^{1/2-\varepsilon}$. On the other hand if $K/\mathbb{Q}$ is a real quadratic fields with class number 1, (1.10) just recovers (1.9).

Remark 1.9. As mentioned above it has been conjectured [18, Conjecture 3.10] that the following should hold for all $\varepsilon > 0$;

$$\sup_{z\in C} |E(z,1/2+it)| \ll_{\varepsilon,C} (|t|+1)^{\varepsilon}, \tag{1.11}$$

where $C \subset \mathbb{H}$ is a compact set. This implies the Lindelöf hypothesis in the t-aspect for the class group L-function. In the last section we will speculate what the uniform analogue of (1.11) should be.

1.3.1. *Hybrid subconvexity bounds for class group L-functions.* The first to obtain subconvexity for class group L-functions seems to be Söhne [19] in the t-aspect and Duke, Friedlaner and Iwaniec [7] in the D-aspect (which was then improved numerically by Blomer, Harcos and Michel [4]). The first to achieve subconvexity in both aspects simultaneously (with an unspecified exponent) was Michel and Venkatesh [16] as a consequence of their solution of the subconvexity problem for GL_2 automorphic L-functions (for general number fields). The results of Michel and Venkatesh were then later made explicit by Wu [24]. More precisely [24, Corollary 1.4] states that if π is an automorphic representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$ with (unitary) central character ω, then we have

$$L(\pi,1/2) \ll \mathbf{C}(\pi)^{1/4}\left(\frac{\mathbf{C}(\pi)}{\mathbf{C}(\omega)}\right)^{-\frac{1-2\theta}{40}} \mathbf{C}(\omega)^{-1/160}, \tag{1.12}$$

where $\mathbf{C}(\pi), \mathbf{C}(\omega)$ denote the analytic conductors of respectively π,ω and θ is any approximation towards the Ramanujan–Petersson conjecture. Let us briefly explain how to extract a subconvexity bound for class group L-functions from (1.12).

Let χ be a (wide) class group character of the quadratic extension $K/\mathbb{Q}$ of conductor D, $\theta_\chi \in \mathcal{M}_1(\Gamma_0(|D|),\chi_D)$ the theta series associated to χ (see [13, Section 14.3]) and π_χ the corresponding automorphic representation of $\mathrm{GL}_2(\mathbb{A}_\mathbb{Q})$. The analytic conductor of the automorphic representation $\pi_\chi \otimes |\cdot|^{it}_{\mathbb{A}_\mathbb{Q}}$ is given by $D(|t|+1)^2$ and the same is true for the analytic conductor of its central character. By plugging this into (1.12) above, we thus get

$$L(\pi_\chi\otimes|\cdot|^{it}_{\mathbb{A}_\mathbb{Q}},1/2) = L_K(1/2+it,\chi) \ll \left(|D|^{1/4}\,(|t|+1)^{1/2}\right)^{1-1/40},$$

which is the state of the art for hybrid subconvexity. We observe that the bound (1.9) improves on this in certain regimes of t and D. Combining the result of Wu with ours, we arrive at the following improvement.

Corollary 1.10. *Let $K/\mathbb{Q}$ be a quadratic extension of discriminant D and χ a (wide) class group character of K. Then we have*

$$L_K(1/2+it,\chi) \ll_\varepsilon \begin{cases} |D|^{1/4+\varepsilon}\,(|t|+1)^{1/3+\varepsilon}, & \text{for } t > |D|^{3/74} \\ \left(|D|^{1/4}\,(|t|+1)^{1/2}\right)^{1-1/40}, & \text{for } t \leq |D|^{3/74} \end{cases}, \tag{1.13}$$

for any $\varepsilon > 0$.

Remark 1.11. The state of the art hybrid subconvexity bound for GL_1 automorphic L-functions [23, Corollary 1.2] is very similar to the above; the best hybrid subconvexity bound is obtained by combining the results of Wu [23] and those of Söhne [19]. Notice that the bounds obtained in these two papers depend on the number field and are thus not relevant in our hybrid setting.

Remark 1.12. In the special case where χ is a genus character, we have the following factorization in terms of quadratic Dirichlet L-functions;

$$L_K(s,\chi) = L(s,\left(\tfrac{d_1}{\cdot}\right))L(s,\left(\tfrac{d_2}{\cdot}\right)),$$

where χ corresponds to the factorization $d_1 d_2 = D$. In this case it follows from [25, (1.8)] that we have the following improvement on the above;

$$L_K(1/2+it,\chi) \ll_\varepsilon |D|^{1/6+\varepsilon}(|t|+1)^{1/3+\varepsilon}.$$

Acknowledgement

I would like to express my gratitude to my advisor Morten Risager for suggesting this problem to me and for pointing me to [3] and [24], and to the referee for many useful comments, which enabled me to prove a stronger result.

2. From sup norm bounds to subconvexity

In this section we will prove Theorem 1.3. First of all we will introduce some background on quadratic fields and the formulas due to Hecke mentioned above.

2.1. Quadratic fields.

We will now recall a few standard facts about quadratic fields and refer to [13, Chapter 22], [17, Section 1] and [9, Section 2] for more background. Let $K/\mathbb{Q}$ be a quadratic extension of number fields, then we can write $K = \mathbb{Q}[\sqrt{D}]$ where D is the discriminant of K. We denote by $\mathrm{Cl}\,(K)$ the class group of K consisting of classes of fractional ideals modulo principal ideals. According to Gauss each fractional ideal class $\mathfrak{a}$ corresponds to an equivalence class of integral binary quadratic forms of discriminant D modulo integral linear transformations. When $D < 0$ we can to each $\mathfrak{a} \in \mathrm{Cl}\,(K)$ associate a Heegner point on the modular curve given by;

$$z_{\mathfrak{a}} := \frac{-b + i\sqrt{|D|}}{2a} \in X_0(1),$$

where $Q = aX^2 + bXY + cY^2$ is any representative of $\mathfrak{a}$. We denote by $h(K)$ the size of the class group and we have the following (ineffective) bound due to Siegel;

$$|D|^{1/2-\varepsilon} \ll_\varepsilon h(K) \ll_\varepsilon |D|^{1/2+\varepsilon}. \tag{2.1}$$

When $D > 0$, we can analogously to any ideal class $\mathfrak{a}$ in the (wide) class group of K associate a certain primitive, closed geodesic $C_\mathfrak{a}$ on $X_0(1)$. If $\mathfrak{a}$ corresponds to some integral binary quadratic form $Q = aX^2 + bXY + cY^2$, then $C_\mathfrak{a}$ is defined as the projection onto $X_0(1)$ of a certain arc on the semi-circle $S_Q \subset \mathbb{H}$ defined by the end-points $\frac{-b\pm\sqrt{D}}{2a}$ (see the references above for the precise definition). The hyperbolic line element on $X_0(1)$ is given by $|ds| = |dz|/y$ and $C_\mathfrak{a}$ has hyperbolic length $2\log \epsilon_K$, where ϵ_K is the fundamental unit of K. Similar to the imaginary quadratic case we have the (ineffective) bound;

$$|D|^{1/2-\varepsilon} \ll_\varepsilon h(K) \log \epsilon_K \ll_\varepsilon |D|^{1/2+\varepsilon}, \tag{2.2}$$

also due to Siegel.

2.2. **Hecke's formula for class group L-functions.** For a real or imaginary quadratic extension $K/\mathbb{Q}$ and a character χ of Cl (K), we associate the class group L-function absolutely convergent for $\operatorname{Re} s > 1$;

$$L_K(s,\chi) := \sum_{\mathfrak{a}} \chi(\mathfrak{a}) N_K(\mathfrak{a})^{-s} = \prod_{\mathfrak{p}} (1 - \chi(\mathfrak{p}) N_K(\mathfrak{p})^{-s}), \tag{2.3}$$

where N_K is the norm and the sum runs over all integral ideals of K and the product is taken over integral prime ideals of K. The class group L-functions admit analytic continuation and functional equations, which we will see shortly follows from the same properties for the non-holomorphic Eisenstein series.

The connection between class group L-functions and Eisenstein series is given by a beautiful formula due to Hecke. In the introduction we already mentioned that for imaginary quadratic extensions $K/\mathbb{Q}$, the formula reads [13, (22.58)];

$$L_K(s,\chi) = \frac{2^{s+1}\zeta(2s)|D|^{-s/2}}{\omega_K} \sum_{\mathfrak{a}} \chi(\mathfrak{a}) E(z_\mathfrak{a}, s),$$

where the sum runs over a complete set of representatives for the class group of K, $z_\mathfrak{a}$ is the associated Heegner point and $\omega_K \in \{2,4,6\}$ denotes the number of roots of unity in K.
For real quadratic fields, we have similarly the following formula [9, (7.7)];

$$L_K(s,\chi) = \frac{\zeta(2s) D^{-s/2}\Gamma(s)}{\Gamma(s/2)^2} \sum_{\mathfrak{a}} \chi(\mathfrak{a}) \int_{C_\mathfrak{a}} E(z,s) y^{-1} |dz|. \tag{2.4}$$

We observe that analytic continuation and functional equation for $L_K(s,\chi)$ now follows from the corresponding properties of the Eisenstein series [12, Theorem 6.5].

2.3. **Proof of Theorem 1.3.** In this section we will prove Theorem 1.3. To do this we will need a lemma that bounds averages over Heegner points (resp. cycles) of the function $y : X_0(1) \to \mathbb{R}_+$ defined by $y(z) := \mathrm{Im}(z_{\mathcal{F}})$, where $z_{\mathcal{F}} \in \mathbb{H}$ is the representative of $z \in X_0(1)$ which lies in $\mathcal{F}$, the standard fundamental domain (1.3) for $\Gamma_0(1)$. Observe that this function is continuous.

Lemma 2.1. *Let $K/\mathbb{Q}$ be a quadratic field of discriminant D. Then we have for any $\delta > 0$ and $\varepsilon > 0$;*

$$\sum_{\mathfrak{a}\in \mathrm{Cl}\,(K)} \begin{cases} y(z_{\mathfrak{a}})^\delta & \text{if } D<0, \\ \int_{C_{\mathfrak{a}}} y(z)^\delta\,|ds| & \text{if } D>0, \end{cases} \ll_\varepsilon |D|^{\max(\delta,1)/2+\varepsilon}.$$

Proof. Assume $D < 0$. The representative of $z_{\mathfrak{a}} \in X_0(1)$ which lies in $\mathcal{F}$, is exactly given by

$$(z_{\mathfrak{a}})_{\mathcal{F}} = \frac{-b + i\sqrt{|D|}}{2a},$$

where the integral binary quadratic form $aX^2 + bXY + cY^2$ of discriminant D corresponds to $\mathfrak{a}$ and (a, b, c) is reduced [13, (22.12)], meaning that;

$$-a < b \le a \le c \quad \text{or} \quad -a \le b \le a = c.$$

Since $\mathcal{F} \subset \{z \in \mathbb{H} \mid \mathrm{Im}\, z \ge \sqrt{3}/2\}$, we conclude that $a \ll \sqrt{|D|}$ and thus we get;

$$\begin{aligned}\sum_{\mathfrak{a}\in \mathrm{Cl}\,(K)} y(z_{\mathfrak{a}})^\delta =&|D|^{\delta/2} \sum_{a>0} \frac{\#\{a,b,c \mid b^2-4ac = D, (a,b,c) \text{ reduced}\}}{(2a)^\delta} \\ \ll&|D|^{\delta/2} \sum_{0<a\ll |D|^{1/2}} \frac{\rho_D(a)}{a^\delta},\end{aligned}$$

where $\rho_D(a) = \#\{0 < b \le 2a \mid b^2 \equiv D \bmod 4a\}$. It is well-known [13, p. 521] that ρ_D is multiplicative with $\rho_D(p^\alpha) = 1 + \chi_D(p)$ if $p \nmid D$, $\rho_D(p) = 1$ if $p|D$ and $\rho_D(p^\alpha) = 0$ if $p|D$, $\alpha > 1$, which implies the bound $\rho_D(a) \ll \sum_{d|a} 1 \ll_\varepsilon a^\varepsilon$. Thus we conclude that

$$\sum_{\mathfrak{a}\in \mathrm{Cl}\,(K)} y(z_{\mathfrak{a}})^\delta \ll_\varepsilon |D|^{\delta/2} \sum_{0<a\ll\sqrt{|D|}} \frac{a^\varepsilon}{a^\delta} \ll |D|^{1/2\max(\delta,1)+\varepsilon},$$

as wanted.

Now we turn to the case $D > 0$. We denote by Ω_D all integral binary quadratic forms of discriminant D and for $Q = aX^2 + bXY + cY^2 \in \Omega_D$, we denote by S_Q the semicircle in $\mathbb{H}$ with end-points $\frac{-b\pm\sqrt{D}}{2a}$. Then it follows from an easy lemma [8, Lemma 6] (observe that they use a different looking but equivalent measure) that;

$$\sum_{\mathfrak{a}\in \mathrm{Cl}\,(K)} \int_{C_{\mathfrak{a}}} y(z)^\delta\,|ds| = \sum_{Q\in\Omega_D} \int_{S_Q\cap\mathcal{F}} y(z)^\delta\,|ds|, \tag{2.5}$$

where $\mathcal{F}$ is the standard fundamental domain (1.3) for $\Gamma_0(1)$.
Now we take the quotient from the left by $\Gamma_\infty = \langle T\rangle$ where $T = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, which rewrites

(2.5) as the following;

$$\sum_{[Q]\in\Gamma_\infty\backslash\Omega_D}\int_{S_Q\cap\mathcal{F}^{(\infty)}} y(z)^\delta\,|ds|, \tag{2.6}$$

where $\mathcal{F}^{(\infty)} := \cup_{n\in\mathbb{Z}}T^{(n)}\mathcal{F}$ is the union of all horizontal translates of $\mathcal{F}$ (notice that the integral above does not depend on the choice of Q). Since $\mathcal{F}^{(\infty)}\subset\{z\in\mathbb{H}\mid \operatorname{Im} z\geq \sqrt{3}/2\}$, we only get contributions in (2.6) from quadratic forms $Q = aX^2+bXY+cY^2$ with $a\ll\sqrt{D}$ and furthermore we can pick representatives of $\Gamma_\infty\backslash\Omega_D$ satisfying $|b|\leq 2a$. Now we recall that $|ds| = y^{-1}|dz|$ and use the trivial fact that the Euclidean circumference of S_Q is $\ll \frac{D^{1/2}}{a}$, which implies;

$$\begin{aligned}\sum_{[Q]\in\Gamma_\infty\backslash\Omega_D}\int_{S_Q\cap\mathcal{F}^{(\infty)}} y(z)^\delta\,|ds| &= \sum_{0<a\ll D^{1/2}}\ \sum_{\substack{[Q]\in\Gamma_\infty\backslash\Omega_D,\\ Q(1,0)=a}}\int_{S_Q\cap\mathcal{F}^{(\infty)}} y(z)^{\delta-1}|dz|\\ &\ll \sum_{0<a\ll D^{1/2}}\ \sum_{\substack{[Q]\in\Gamma_\infty\backslash\Omega_D,\\ Q(1,0)=a}}\frac{D^{1/2}}{a}\left(\max_{z\in S_Q\cap\mathcal{F}^{(\infty)}} y(z)^{\delta-1}\right)\\ &\ll D^{1/2+\max(\delta-1,0)/2}\sum_{0<a\ll D^{1/2}}\frac{\rho_D(a)}{a}.\end{aligned}$$

Now the conclusion follows exactly as in the case of negative D using the bound $\rho_D(a)\ll_\varepsilon a^\varepsilon$ (which also holds for $D>0$ by the above). □

Now we are ready to prove Theorem 1.3.

Proof of Theorem 1.3. Consider the case $D<0$. By feeding (1.4) into (1.1), we see that

$$L_K(1/2+it,\chi)\ll_\varepsilon \frac{(|t|+1)^\varepsilon}{|D|^{1/4}}\sum_{\mathfrak{a}} y(z_\mathfrak{a})^\delta(|t|+1)^\theta, \tag{2.7}$$

where we used some standard estimates for ζ on $\operatorname{Re} s = 1$.
Now since we assumed $\delta\leq 1$, it follows from Lemma 2.1 that

$$L_K(1/2+it,\chi)\ll_\varepsilon |D|^{1/4+\varepsilon}(|t|+1)^{\theta+\varepsilon}.$$

as wanted.

To prove (1.6), we observe that by orthogonality, the formula (1.1) implies that

$$\sum_\chi |L_K(1/2+it,\chi)|^2 = \frac{8h(K)|\zeta(1+2it)|^2}{\omega_K^2|D|^{1/2}}\sum_{\mathfrak{a}}|E(z_\mathfrak{a},1/2+it)|^2.$$

Thus by the assumption (1.4), Siegel's bound (2.1) and standard estimates for the zeta function, we get

$$\sum_\chi |L_K(1/2+it,\chi)|^2\ll_\varepsilon (|t|+1)^{2\theta+\varepsilon}|D|^\varepsilon\sum_{\mathfrak{a}} y(z_\mathfrak{a})^{2\delta},$$

and the result follows directly from Lemma 2.1.

The proof of (1.5) for D positive is exactly the same using Lemma 2.1 and Hecke's formula (2.4) in the case $D > 0$.
In order to prove (1.6), we use orthogonality as above to get

$$\sum_{\chi} |L_K(1/2+it,\chi)|^2 \ll_\varepsilon (|t|+1)^{2\theta+\varepsilon} \frac{h(K)}{D^{1/2}} \sum_{\mathfrak{a}} \left| \int_{C_{\mathfrak{a}}} y(z)^\delta |ds| \right|^2 .$$

Now we apply Cauchy-Schwarz to bound the above by

$$(|t|+1)^{2\theta+\varepsilon} \frac{h(K)\log \epsilon_K}{D^{1/2}} \sum_{\mathfrak{a}} \int_{C_{\mathfrak{a}}} y(z)^{2\delta} |ds|,$$

and the results follows from Lemma 2.1 and Siegel's bound (2.2). □

Remark 2.2. If one believes the sup norm conjecture (1.11), Theorem 1.3 tells you in particular that the cancellations in individual Eisenstein series are strong enough to give the Lindelöf hypothesis for class group L-functions in the t-aspect. It is however conjectured that (1.2) holds for eigenfunctions on any hyperbolic surface [18, Conjecture 3.10]. So in some sense the t-aspect is not essentially arithmetic. This method is however not able to give subconvexity estimates in the D-aspect for individual L-functions. This is due to the fact that the sup norm bounds do not "see" the arithmetics of the Heegner points (it is uniform for z in a fixed compact set) and the cancellation between Eisenstein series evaluated at the different Heegner points is exactly what gives rise to subconvexity behavior in the D-aspect. In the last section (see (5.2)), we will state a uniform analogue of the conjecture (1.2), which using (1.6) does give Lindelöf on average in the D-aspect for imaginary quadratic fields.

3. Uniform sup norm bounds of Eisenstein series

In this section we will prove the hybrid bound (1.7) and (1.8) for the classical Eisenstein series. The proof of (1.7) follows directly from [26]. The proof of (1.8) requires much more work and is an adaptation (and elaboration) of the argument in [3] building on [21], which in turn is an extension of the van der Corput method [13, Section 8.3].

3.1. **Uniform bounds for Eisenstein series following Young.**

In [26] Young extends the method used by Iwaniec and Sarnak in [14] to give the first non-trivial result towards the sup norm conjecture for the modular curve. The main insight of Young was that one can choose a more efficient mollifier, which improves the bound for the continuous spectrum. The method of Iwaniec and Sarnak embeds respectively the cusp form and Eisenstein series into the entire spectrum of the modular curve. Then an application of the Selberg trace formula (with a carefully chosen test function) reduces the sup norm bound to a bound of the geometric side, which can be done with elementary means. The action of the Hecke operators plays a crucial role in the argument.
In [26] the sup norm bound is stated as a bound in the t-aspect with z in a fixed compact set, but as Young also mentions the method yields something slightly stronger

(this was also observed by Huang and Xu [11, p. 2]).
The main inequality in Young's paper is [26, (6.3)], which gives

$$|E(z,1/2+it)|^2 \ll_\varepsilon (N|t|)^\varepsilon \left(\frac{|t|}{N} + |t|^{1/2}(N+N^{1/2}y)\right), \tag{3.1}$$

where N is some parameter to be chosen appropriately. By inspecting [26, Lemma 4.1, Lemma 5.1] one sees that the restrictions on the variables are $\log N \gg (\log t)^{2/3+\delta}$ for some fixed $\delta > 0$ and $y \ll |t|^{100}$. In particular in the range $y \ll |t|^{1/4}$, we can put $N = |t|^{1/4}$ and get

$$|E(z,1/2+it)|^2 \ll_\varepsilon |t|^{3/4+\varepsilon} + |t|^{3/4+\varepsilon} + |t|^{5/8+\varepsilon} y.$$

From this we conclude

$$|E(z,1/2+it)| \ll_\varepsilon y^{1/2}|t|^{3/8+\varepsilon}, \quad 1 \ll y \ll |t|^{1/4}.$$

In the range $y \gg |t|^{1/4}$, we have the trivial bound [26, (3.2)], which yields

$$|E(z,1/2+it)| \ll_\varepsilon y^{1/2} + |t|^{3/8+\varepsilon}.$$

Combining the two, concludes the proof of Theorem 1.4.

3.2. **Titchmarsh's method for bounding Epstein zeta functions.** Now we turn to the proof of Theorem 1.6. The following serves first of all as an extension of Blomer and Titchmarsh's work but secondly as an elaboration of some of the details, which are left out in [3]. The approach expresses the non-holomorphic Eisenstein series in terms of an *Epstein zeta function*, which is then bounded using the van der Corput method from the theory of exponential sums. Originally Titchmarsh considered only Epstein zeta functions associated to diagonal matrices and there are some technical difficulties to deal with general Epstein zeta functions. Furthermore in order to get a bound uniform in the entries of the matrix defining the Epstein zeta function, one has to modify parts of the argument.

Given any positive definite matrix $Z \in \mathrm{GL}_2(\mathbb{R})$, we can consider the quadratic form $Q(\boldsymbol{x}) = \boldsymbol{x}\, Z\, \boldsymbol{x}^T$, $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$ and the associated Epstein zeta function

$$E_{\mathrm{Epstein}}(Z,s) := \sum_{\boldsymbol{x} \in \mathbb{Z}^2 \setminus (0,0)} Q(\boldsymbol{x})^{-s},$$

which satisfies the functional equation

$$\Gamma_{\mathbb{R}}(2s) E_{\mathrm{Epstein}}(Z,s) = (\det Z)^{-1/2} \Gamma_{\mathbb{R}}(2(1-s)) E_{\mathrm{Epstein}}(Z^{-1}, 1-s),$$

where $\Gamma_{\mathbb{R}}(s) := \pi^{-s/2}\Gamma(s/2)$.
Recall that this is related to the non-holomorphic Eisenstein series as follows

$$\zeta(2s)E(z,s) = y^s E_{\mathrm{Epstein}}(Z,s), \qquad Z = \begin{pmatrix} x^2+y^2 & x \\ x & 1 \end{pmatrix}, \tag{3.2}$$

which reduces the sup norm problem for Eisenstein series to bounding the Epstein zeta function. We may restrict to the case where $z \in \mathcal{F}$, the standard fundamental domain (1.3) for $X_0(1)$, which corresponds to considering only matrices of the form

$$Z = \begin{pmatrix} a & b \\ b & 1 \end{pmatrix},$$

where $a \geq 1$ and $|b| \leq 1/2$.
The trivial estimate [26, (3.2)];

$$E(z, 1/2 + it) \ll y^{1/2} + (t/y)^{1/2}$$

yields (1.8) in the range $|t|^{1/6} \ll y$ and thus in the sequel we may assume $a \ll |t|^{1/3}$ and thus also $|t| \gg 1$.

3.3. **Reduction to an exponential sum.** As in [3] we start by applying an approximate functional equation [13, Theorem 5.3] with $G(u) = e^{u^2}$, but deviate slightly by using a balanced version (corresponding to putting $X = a^{1/2}$ in [13, Theorem 5.3]). By estimating the contribution coming from the pole of $E_{\text{Epstein}}(Z, s)$ at $s = 1$ trivially, the approximate functional equation yields

$$\begin{aligned} E_{\text{Epstein}}(Z, 1/2 + it) &= \sum_{\mathbf{x} \neq 0} \frac{W_t^+(Q_+(\mathbf{x})a^{-1/2})}{Q_+(\mathbf{x})^{1/2+it}} \\ &+ \frac{\Gamma_{\mathbb{R}}(1 - 2it)}{\Gamma_{\mathbb{R}}(1 + 2it)(\det Z)^{1/2}} \sum_{\mathbf{x} \neq 0} \frac{W_t^-(Q_-(\mathbf{x})a^{1/2})}{Q_-(\mathbf{x})^{1/2-it}} + O(1) \end{aligned} \tag{3.3}$$

where $Q_\pm(\mathbf{x}) = \mathbf{x}\, Z^{\pm 1}\, \mathbf{x}^T$ and

$$W_t^\pm(y) = \frac{1}{2\pi i} \int_{(1)} e^{u^2} \frac{\Gamma_{\mathbb{R}}(2(u + 1/2 \pm it))}{\Gamma_{\mathbb{R}}(2(1/2 \pm it))} y^{-u} \frac{du}{u}.$$

The weight $W_t^\pm$ can be nicely bounded as follows; we move the contour to the line (A) with $A > 0$ and bound the integrand using Stirling's approximation as follows;

$$e^{u^2/2} \frac{\Gamma_{\mathbb{R}}(2(u + 1/2 \pm it))}{\Gamma_{\mathbb{R}}(2(1/2 \pm it))} u^{-1} \ll \frac{e^{A^2/2} e^{-b^2/2} \pi^{-A/2} e^{-A}(|t|^A + (b + A)^A)}{A + |b|} \ll_A |t|^A,$$

with $u = A + ib$ using that $e^{-b^2/2}(b + A)^A \to 0$ as $b \to \infty$. Thus we get the bound

$$W_t^\pm(y) \ll_A |t|^A / y^A \int_{-\infty}^{\infty} e^{-x^2/2} dx \ll |t|^A / y^A,$$

and more generally one deduces $\frac{\partial^n}{\partial y^n} W_t^\pm(y) \ll_A |t|^A / y^{A+n}$ as in [13, Proposition 5.4]. From this we see that the contributions in (3.3) from $\boldsymbol{x}$ such that $Q_\pm(\boldsymbol{x}) \gg a^{\pm 1/2} |t|^{1+\varepsilon}$ are negligible.

To deal with the remaining sums in (3.3), we divide the range of summation into dyadic rectangles of the form $(X_1, 2X_1) \times (X_2, 2X_2)$. Observe that we get $O(\log^2 t)$ such rectangles, which implies that it suffices to bound each of these dyadic sums individually.

For each such rectangle we get by two-dimensional partial summation;

$$\sum_{\substack{X_1\le x_1\le 2X_1\\ X_2\le x_2\le 2X_2}} \frac{W_t^+(Q_+(\mathbf{x})a^{-1/2})}{Q_+(\mathbf{x})^{1/2+it}} = F_+(2\boldsymbol{X}) \sum_{\substack{X_1\le x_1\le 2X_1\\ X_2\le x_2\le 2X_2}} e^{it\log Q_+(\boldsymbol{x})} - \int_{X_1}^{2X_1} \left(\sum_{\substack{X_1\le x_1\le x\\ X_2\le x_2\le 2X_2}} e^{it\log Q_+(\boldsymbol{x})} \right) F_+^{(1,0)}(x,2X_2)dx - \int_{X_2}^{2X_2} \left(\sum_{\substack{X_1\le x_1\le 2X_1\\ X_2\le x_2\le y}} e^{it\log Q_+(\boldsymbol{x})} \right) F_+^{(0,1)}(2X_1,y)dy + \int_{X_1}^{2X_1}\int_{X_2}^{2X_2} \left(\sum_{\substack{X_1\le x_1\le x\\ X_2\le x_2\le y}} e^{it\log Q_+(\boldsymbol{x})} \right) F_+^{(1,1)}(x_1,x_2)\,dxdy, \tag{3.4}$$

where $\boldsymbol{X}=(X_1,X_2)$, $F_+(\mathbf{x}) = W_t^+(Q_+(\mathbf{x})a^{-1/2})/Q_+(\mathbf{x})^{1/2}$ and $F_+^{(i,j)} := \frac{\partial^{i+j}F_+}{\partial x_1^i \partial x_2^j}$.
Similarly we get

$$\sum_{\substack{X_1\le x_1\le 2X_1\\ X_2\le x_2\le 2X_2}} \frac{W_t^-(Q_-(\mathbf{x})a^{1/2})}{(\det Z)^{1/2}Q_-(\mathbf{x})^{1/2-it}} = F_-(2\boldsymbol{X}) \sum_{\substack{X_1\le x_1\le 2X_1\\ X_2\le x_2\le 2X_2}} e^{it\log Q_-(\boldsymbol{x})} + \dots, \tag{3.5}$$

where $F_-(\mathbf{x}) = W_t^-(Q_-(\mathbf{x})a^{1/2})/\left((\det Z)Q_-(\mathbf{x})\right)^{1/2}$.
Now we have reduced the desired bound on the Epstein zeta function to proving a certain estimate on exponential sums. The result we need is the following.

Proposition 3.1. *For $\boldsymbol{X}=(X_1,X_2)$ satisfying $Q_+(\boldsymbol{X}) \ll a^{1/2}|t|^{1+\varepsilon}$, we have the following bound;*

$$\frac{1}{Q_+(\boldsymbol{X})^{1/2}} \sum_{\substack{X_1\le x_1\le X_1'\\ X_2\le x_2\le X_2'}} e^{it\log Q_+(\boldsymbol{x})} \ll_\varepsilon |t|^{1/3+\varepsilon}, \tag{3.6}$$

uniformly in $a\ge 1$, where $X_i \le X_i' \le 2X_i$. Similarly for $\boldsymbol{X}=(X_1,X_2)$ satisfying $Q_-(\boldsymbol{X}) \ll a^{-1/2}|t|^{1+\varepsilon}$, we have

$$\frac{1}{((\det Z)Q_-(\boldsymbol{X}))^{1/2}} \sum_{\substack{X_1\le x_1\le X_1'\\ X_2\le x_2\le X_2'}} e^{it\log Q_-(\boldsymbol{x})} \ll_\varepsilon |t|^{1/3+\varepsilon}, \tag{3.7}$$

where $X_i \le X_i' \le 2X_i$.

Remark 3.2. Observe that when proving (3.6), we may assume

$$X_1 \gg |t|^{1/3} \quad \text{and} \quad X_2 \gg |t|^{1/3}a^{1/2}, \tag{3.8}$$

and similar when proving (3.7), we may assume

$$X_1 \gg |t|^{1/3}a^{1/2} \quad \text{and} \quad X_2 \gg |t|^{1/3}, \tag{3.9}$$

since otherwise the bounds follows from the trivial estimate on the exponentials.

Now let us see how Theorem 1.6 follows from the above proposition.

Proof of Theorem 1.6 assuming Proposition 3.1. We will begin by deducing from Proposition 3.1 that $E_{\mathrm{Epstein}}(Z,s) \ll_\varepsilon (|t|+1)^{1/3+\varepsilon}$ for all Z as above; by the above reductions, it suffices to prove the same bound for each of the dyadic sums (3.4) and (3.5) with X_1, X_2 satisfying respectively $Q_\pm(\boldsymbol{X}) \ll a^{\pm 1/2}|t|^{1+\varepsilon}$. We do this by bounding each of the four terms, we get after applying partial summation separately (observe that we may assume $|t| \gg 1$).
The above estimates for W_t^+ imply $W_t^+(Q_+(\boldsymbol{x})a^{-1/2}) \ll |t|^\varepsilon$, which together with (3.6) implies that we can bound the first sum on the right-hand side of (3.4) by the following;

$$F_+(2\boldsymbol{X}) \sum_{\substack{X_1 \le x_1 \le 2X_1 \\ X_2 \le x_2 \le 2X_2}} e^{it \log Q_+(\boldsymbol{x})} \ll |t|^{1/3+\varepsilon}.$$

Similarly using $\frac{\partial^n}{\partial y^n} W_t^+(y) \ll |t|^A / y^{A+n}$ and the chain rule, we get

$$F_+^{(1,0)}(\boldsymbol{x}) \ll \frac{|t|^\varepsilon a^{1/2}}{Q_+(\boldsymbol{X})}, \quad F_+^{(0,1)}(\boldsymbol{x}) \ll \frac{|t|^\varepsilon}{Q_+(\boldsymbol{X})}, \quad F_+^{(1,1)}(\boldsymbol{x}) \ll \frac{|t|^\varepsilon a^{1/2}}{Q_+(\boldsymbol{X})^{3/2}},$$

which together with (3.6) implies

$$\int_{X_1}^{2X_1} \left(\sum_{\substack{X_1 \le x_1 \le x \\ X_2 \le x_2 \le 2X_2}} e^{it \log Q_+(\boldsymbol{x})} \right) F_+^{(1,0)}(x, 2X_2)\,dx$$
$$\ll X_1 |t|^{1/3+\varepsilon} Q_+(\boldsymbol{X})^{1/2} \frac{a^{1/2}}{Q_+(\boldsymbol{X})} \ll |t|^{1/3+\varepsilon},$$

using $X_1 a^{1/2} \ll Q_+(\boldsymbol{X})^{1/2}$, and similarly for the other one-dimensional integral. Finally a similar calculation gives

$$\int_{X_1}^{2X_1} \int_{X_2}^{2X_2} \left(\sum_{\substack{X_1 \le x_1 \le x \\ X_2 \le x_2 \le y}} e^{it \log Q_+(\boldsymbol{x})} \right) F^{(1,1)}(x,y)\,dxdy$$
$$\ll \frac{X_1 X_2 a^{1/2} Q_+(\boldsymbol{X})^{1/2} |t|^{1/3+\varepsilon}}{Q_+(\boldsymbol{X})^{3/2}},$$

which yields the desired bound for the Q_+-sum.
The sum involving Q_- can be bounded similarly using

$$F_-^{(1,0)}(\boldsymbol{x}) \ll \frac{|t|^\varepsilon}{(\det Z) Q_-(\boldsymbol{X})}, \quad F_-^{(0,1)}(\boldsymbol{x}) \ll \frac{|t|^\varepsilon a^{1/2}}{(\det Z) Q_-(\boldsymbol{X})},$$
$$F_-^{(1,1)}(\boldsymbol{x}) \ll \frac{|t|^\varepsilon a^{1/2}}{((\det Z) Q_-(\boldsymbol{X}))^{3/2}},$$

which yields the desired bound for the Epstein zeta function.
Thus we conclude that

$$E(z,1/2+it) = \frac{y^{1/2+it}}{\zeta(1+2it)} E_{\text{Epstein}}(Z,1/2+it) \ll_\varepsilon y^{1/2}(|t|+1)^{1/3+\varepsilon},$$

using $\zeta(1+2it) \gg_\varepsilon (|t|+1)^{-\varepsilon}$. This finishes the proof. □

4. A uniform bound for an exponential sum in two variables

In this section we will prove Proposition 3.1 using an extension of the ideas of Titchmarsh and Blomer building on the work of van der Corput.
Firstly we will make a simplification; if we multiply with the phase $(\det Z)^{it}$ in (3.7), the summands become;

$$e^{it\log(\det Z)} e^{it\log Q_-(\boldsymbol{x})} = e^{it\log((\det Z)Q_-(\boldsymbol{x}))},$$

where $(\det Z)Q_-(\boldsymbol{x}) = x_1^2 - 2bx_1x_2 + ax_2^2$. Since $\det Z \asymp a$, the ranges $Q_+(\boldsymbol{X}) \ll a^{1/2}|t|^{1+\varepsilon}$ and $(\det Z)Q_-(\boldsymbol{X}) \ll (\det Z)a^{-1/2}|t|^{1+\varepsilon}$ are the same just with X_1 and X_2 interchanged. Thus by symmetry the two bounds (3.6) and (3.7) are equivalent, which is exactly why we used a balanced approximate functional equation in the first place.
Thus we see that it suffices to prove (3.6) under the assumption $Q_+(\boldsymbol{x}) \ll a^{1/2}|t|^{1+\varepsilon}$. To lighten notation, we put $Q := Q_+$.

4.1. Some lemmas of Titchmarsh.

Titchmarsh [21] extended the van der Corput method for bounding exponential sums [13, Section 8.3] to two-dimensional sums. In this section we will quote some lemmas due to Titchmarsh, which we will employ later.
Through-out this section we assume that

$$f : (X_1, X_1') \times (X_2, X_2') \to \mathbb{R}$$

has algebraic partial derivatives of order one to three. We will as above use the notation $f^{(i,j)} := \frac{\partial^{i+j} f}{\partial x_1^i \partial x_2^j}$.

The first lemma is a version of Weyl differencing in the two-dimensional setting.

Lemma 4.1 (Lemma β, [21]). *Let $\rho \le \min(X_1' - X_1, X_2' - X_2)$ be a positive integer. Then we have*

$$\sum_{\substack{X_1 \le x_1 \le X_1' \\ X_2 \le x_2 \le X_2'}} e^{if(\boldsymbol{x})} \ll \frac{(X_1' - X_1)(X_2' - X_2)}{\rho}$$

$$+ \frac{(X_1' - X_1)^{1/2}(X_2' - X_2)^{1/2}}{\rho} \left(\sum_{\substack{1 \le \mu_1 \le \rho - 1 \\ 0 \le \mu_2 \le \rho - 1}} |S_1(\boldsymbol{\mu})| \right)^{1/2}$$

$$+ \frac{(X_1' - X_1)^{1/2}(X_2' - X_2)^{1/2}}{\rho} \left(\sum_{\substack{0 \le \mu_1 \le \rho - 1 \\ 1 \le \mu_2 \le \rho - 1}} |S_2(\boldsymbol{\mu})| \right)^{1/2}, \tag{4.1}$$

where $\boldsymbol{x} = (x_1, x_2)$, $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and

$$S_1(\boldsymbol{\mu}) = \sum_{\substack{X_1 \le x_1 \le X_1' - \mu_1 \\ X_2 \le x_2 \le X_2' - \mu_2}} e^{i[f(\boldsymbol{x}+\boldsymbol{\mu}) - f(\boldsymbol{x})]}, \quad S_2(\boldsymbol{\mu}) = \sum_{\substack{X_1 \le x_1 \le X_1' - \mu_1 \\ X_2 + \mu_2 \le x_2 \le X_2'}} e^{i[f(\boldsymbol{x}+(\mu_1, -\mu_2)) - f(\boldsymbol{x})]}.$$

The above lemma reduces the task to bounding the sums $S_1(\boldsymbol{\mu})$ and $S_2(\boldsymbol{\mu})$ with μ_1, μ_2 in the appropriate ranges. The idea of the van der Corput method is to reduce the bound of the sums $S_1(\boldsymbol{\mu})$ and $S_2(\boldsymbol{\mu})$ to bounding a certain integral. We will use the following extension of van der Corput's result due to Titchmarsh.

Lemma 4.2 (Lemma γ, [21]). *Let $l = \max(X_1' - X_1, X_2' - X_2)$ and assume that f satisfies*

$$|f^{(1,0)}(\boldsymbol{x})| \le \frac{3\pi}{2}, \qquad |f^{(0,1)}(\boldsymbol{x})| \le \frac{3\pi}{2}.$$

Then

$$\sum_{\substack{X_1 \le x_1 \le X_1' \\ X_2 \le x_2 \le X_2'}} e^{if(\boldsymbol{x})} = \int_{(X_1, X_1') \times (X_2, X_2')} e^{if(\boldsymbol{x})} d\boldsymbol{x} + O(l). \tag{4.2}$$

Finally we gonna bound this integral by a second derivative test.

Lemma 4.3 (Lemma ϵ, [21]). *Let $\Omega \subset \mathbb{R}^2$ be a rectangle and l its maximal side length. If $f : \Omega \to \mathbb{R}$ is a function satisfying the conditions mentioned in the beginning of the section and*

$$r \ll |f^{(2,0)}(\boldsymbol{x})| \ll r, \quad r \ll |f^{(0,2)}(\boldsymbol{x})| \ll r, \quad |f^{(1,1)}(\boldsymbol{x})| \ll r \tag{4.3}$$

$$|f^{(2,0)}(\boldsymbol{x}) f^{(0,2)}(\boldsymbol{x}) - (f^{(1,1)}(\boldsymbol{x}))^2| \gg r^2, \qquad \boldsymbol{x} \in \Omega. \tag{4.4}$$

Then

$$\int_\Omega e^{if(\boldsymbol{x})} d\boldsymbol{x} \ll \frac{1 + \log l + \log r}{r},$$

where the implied constant depends only on the angle of the rectangle to the coordinate axes.

Remark 4.4. Note that as stated, [21, Lemma ϵ] (or more precisely Lemma δ) assumes that

$$|f^{(2,0)}(\boldsymbol{x})|, |f^{(0,2)}(\boldsymbol{x})| \geq r, \quad |f^{(2,0)}(\boldsymbol{x})f^{(0,2)}(\boldsymbol{x}) - (f^{(1,1)}(\boldsymbol{x}))^2| \geq r^2,$$

that is; without an implicit constant in the lower bounds. By inspecting the proof of [21, Lemma ϵ], one however sees that Lemma 4.3 as stated above follows with the exact same proof (this observation is also implicit in [3]).

4.2. **Applying the lemmas.** With these results of Titchmarsh at our disposal, we are now ready to make some reductions in the direction of proving (3.6).
By applying Lemma 4.1 with $f(\boldsymbol{x}) = t\log Q(\boldsymbol{x})$ and $Q(\boldsymbol{x}) = ax_1^2 + 2bx_1x_2 + x_2^2$ to the left hand side of (3.6), we reduce the task to bounding sums of the following kind;

$$S'(\boldsymbol{\mu}) = \sum_{\substack{X_1 \leq x_1 \leq X_1' \\ X_2 \leq x_2 \leq X_2'}} e^{ig_{\boldsymbol{\mu}}(\boldsymbol{x})}, \tag{4.5}$$

where

$$g_{\boldsymbol{\mu}}(\boldsymbol{x}) := t(\log Q(\boldsymbol{x}+\boldsymbol{\mu}) - \log Q(\boldsymbol{x})), \tag{4.6}$$

$X_i' \leq 2X_i$ and $\boldsymbol{\mu} = (\mu_1, \mu_2) \in [0,\rho] \times [0,\rho]$ with $\rho = o(\min(X_1, X_2))$ to be chosen appropriately later.

The first step is to divide the rectangle of summation in $S'(\boldsymbol{\mu})$ into rectangles $\Delta_{p,q}$ (where p, q runs through an appropriate indexing set) each with side lengths $l_1 \times l_2$, where

$$l_1 \asymp \frac{Q(\boldsymbol{X})^{3/2}}{a|t|^{1+2\varepsilon}Q(\boldsymbol{\mu})^{1/2}}, \quad l_2 \asymp \frac{Q(\boldsymbol{X})^{3/2}}{a^{1/2}|t|^{1+2\varepsilon}Q(\boldsymbol{\mu})^{1/2}}. \tag{4.7}$$

We denote the sub-sum associated to $\Delta_{p,q}$ by $S_{p,q}(\boldsymbol{\mu})$ and observe that the number of such sub-sums is bounded by;

$$\frac{X_1X_2}{l_1l_2} \ll \frac{X_1X_2}{a^{-3/2}Q(\boldsymbol{x})^3|t|^{-2-2\varepsilon}Q(\boldsymbol{\mu})^{-1}}.$$

We will bound the sub-sums $S_{p,q}(\boldsymbol{\mu})$ individually.

Remark 4.5. There is some balancing in choosing the values l_1, l_2; one the hand l_1, l_2 have to be small enough so that $g_{\boldsymbol{\mu}}$ and its derivatives are close to being constant in $\Delta_{p,q}$ (i.e. the variation is small), and on the other hand the number of rectangles $\Delta_{p,q}$ grows reciprocally with l_1, l_2. The reason for choosing these specific values will become clear later.

4.3. **Bounds on derivatives of $g_{\boldsymbol{\mu}}$.** In this subsection we will prove upper bounds on partial derivates of $g_{\boldsymbol{\mu}}$ and a lower bound on the determinant of the Hesse-matrix of $g_{\boldsymbol{\mu}}$. Titchmarsh [21] only considers diagonal matrices and the fact that $b \neq 0$ creates some minor technical difficulties, which were also addressed by Blomer in [3]. We need to be a bit more careful since we need to consider the a-dependence as well, so our methods of computation differ a bit from those in [3]; to handle the upper bounds on the derivates we apply a Taylor expansion around μ and to lower bound the Hesse determinant we use an explicit calculation.
First of all we will need the following lemma.

Lemma 4.6. *Let* $f(\boldsymbol{x}) = t\log(Q(\boldsymbol{x}))$ *with* $Q(x_1,x_2) = ax_1^2 + 2bx_1x_2 + x_2^2$ *where* $|b| \le 1/2$ *and* $a \ge 1$. *Then we have*

$$f^{(i,j)}(\boldsymbol{x}) \ll_{i,j} \frac{a^{i/2}|t|}{Q(\boldsymbol{x})^{(i+j)/2}}, \tag{4.8}$$

where the implied constant depends on i,j *but is independent of* a,b.

Proof. Observe that $f(\boldsymbol{x})$ is the composition of the function $h(\boldsymbol{x}) := t\log(x_1^2+x_2^2)$ with the linear map

$$\boldsymbol{x} \mapsto \begin{pmatrix} (a-b^2)^{1/2} & 0 \\ b & 1 \end{pmatrix} \boldsymbol{x}^T,$$

where $a-b^2>0$ by the assumptions. Now one sees by a direct computation that

$$h^{(i,j)}(\boldsymbol{x}) = t \sum_{\substack{0\le k\le i,k\equiv i\,(2)\\ 0\le l\le j,l\equiv j\,(2)}} c_{k,l}\frac{x_1^k x_2^l}{(x_1^2+x_2^2)^{(i+j+k+l)/2}},$$

for some constants $c_{k,l}$. Thus we get the bound

$$h^{(i,j)}(\boldsymbol{x}) \ll_{i,j} \frac{|t|}{(x_1^2+x_2^2)^{(i+j)/2}}, \tag{4.9}$$

using the elementary inequality $xy \ll_\alpha x^{1/\alpha} + y^{1/(1-\alpha)}$ for $0<\alpha<1$.
By the chain rule we have

$$f^{(i,j)}(\boldsymbol{x}) = \sum_{l=0}^{i}\binom{i}{l}(a-b^2)^{(i-l)/2}b^l h^{(i-l,j+l)}((a-b^2)^{1/2}x_1, bx_1+x_2),$$

and thus the results follows from (4.9) since b is bounded. □

From this we can now prove the following bounds.

Lemma 4.7. *Let* $\boldsymbol{\mu}$, $\boldsymbol{x}$ *and* $\boldsymbol{X}$ *satisfy the constraints coming from Lemma 4.1.Then we have*

$$\left|g_{\boldsymbol{\mu}}^{(i,j)}(\boldsymbol{x})\right| \ll a^{i/2}\frac{|t|Q(\boldsymbol{\mu})^{1/2}}{Q(\boldsymbol{X})^{(i+j+1)/2}}, \tag{4.10}$$

$$\det(\mathrm{Hess}(g_{\boldsymbol{\mu}}(\boldsymbol{x})) \gg \left(a^{1/2}\frac{|t|\,Q(\boldsymbol{\mu})^{1/2}}{Q(\boldsymbol{X})^{3/2}}\right)^2. \tag{4.11}$$

Proof. It follows from a two-dimensional Taylor expansion that

$$g_{\boldsymbol{\mu}}(\boldsymbol{x}) = \sum_{\alpha\in\{(1,0),(0,1)\}} \frac{1}{\alpha!}\boldsymbol{\mu}^\alpha \int_0^1 f^\alpha(\boldsymbol{x}+t\boldsymbol{\mu})dt, \tag{4.12}$$

where we use the multi-exponential notation $(x_1,x_2)^{(i,j)} := x_1^i x_2^j$.
Using Lemma 4.6, we see that for $\alpha=(\alpha_1,\alpha_2)\in\{(1,0),(0,1)\}$, we have

$$\begin{aligned}\boldsymbol{\mu}^\alpha f^{\alpha+(i,j)}(\boldsymbol{x}) &\ll_{i,j} \boldsymbol{\mu}^\alpha\frac{|t|a^{(\alpha_1+i)/2}}{Q(\boldsymbol{X})^{(i+j+1)/2}}\\ &\ll a^{i/2}\frac{|t|Q(\mu)^{1/2}}{Q(\boldsymbol{X})^{(i+j+1)/2}},\end{aligned}$$

using that $\mu_1 a^{1/2} \ll Q(\boldsymbol{\mu})^{1/2}$, respectively $\mu_2 \ll Q(\boldsymbol{\mu})^{1/2}$.
Thus by applying $\frac{\partial^{i+j}}{\partial x_1^i \partial x_2^j}$ term by term in (4.12) and the bound above, we conclude (4.10).

To prove the last inequality, we apply the following direct computation;

$$\det(\mathrm{Hess}(g_{\boldsymbol{\mu}}(\boldsymbol{x})) = \frac{t^2(\det Q)Q(2\boldsymbol{x}+\boldsymbol{\mu})Q(\boldsymbol{\mu})}{Q(\boldsymbol{x})^2 Q(\boldsymbol{x}+\boldsymbol{\mu})^2} \gg a\frac{|t|^2 Q(\boldsymbol{\mu})}{Q(\boldsymbol{X})^3},$$

where we used $||\boldsymbol{\mu}|| = o(\min(X_1, X_2))$.

□

4.4. **Proof of Proposition 3.1.** Now we would like to apply Lemma 4.2, but obviously we need to alter $g_{\boldsymbol{\mu}}$ a bit in order for the conditions on the derivatives to be satisfied. We observe that the maximum variation in $\Delta_{p,q}$ of $g_{\boldsymbol{\mu}}^{(1,0)}$ is bounded by

$$l_1 \cdot \max_{\boldsymbol{x}\in\Delta_{p,q}} \left| g_{\boldsymbol{\mu}}^{(2,0)}(\boldsymbol{x}) \right| + l_2 \cdot \max_{\boldsymbol{x}\in\Delta_{p,q}} \left| g_{\boldsymbol{\mu}}^{(1,1)}(\boldsymbol{x}) \right| \ll |t|^{-\varepsilon},$$

where we used (4.10), and similarly for $g_{\boldsymbol{\mu}}^{(0,1)}$, in which case the variation is even smaller.
Thus for sufficiently large t the variation in each sub-sum $S_{p,q}$ is less than π, which was exactly why we chose l_1, l_2 as in (4.7). Thus (following Titchmarsh) we can, associated to each $\Delta_{p,q}$, find integers M, N such that

$$G_{\boldsymbol{\mu}}(\boldsymbol{x}) := g_{\boldsymbol{\mu}}(\boldsymbol{x}) - 2\pi M x_1 - 2\pi N x_2,$$

satisfies

$$\left| G_{\boldsymbol{\mu}}^{(1,0)}(\boldsymbol{x}) \right| \leq 3\pi/2 \quad \text{and} \quad \left| G_{\boldsymbol{\mu}}^{(0,1)}(\boldsymbol{x}) \right| \leq 3\pi/2,$$

for all $\boldsymbol{x} \in \Delta_{p,q}$. Thus we get by Lemma 4.2

$$\sum_{\boldsymbol{x}\in\Delta_{p,q}} e^{ig_{\boldsymbol{\mu}}(\boldsymbol{x})} = \sum_{\boldsymbol{x}\in\Delta_{p,q}} e^{iG_{\boldsymbol{\mu}}(\boldsymbol{x})} = \int_{\Delta_{p,q}} e^{iG_{\boldsymbol{\mu}}(\boldsymbol{x})} d\boldsymbol{x} + O(l). \tag{4.13}$$

Observe that all partial derivates of order at least two of $G_{\boldsymbol{\mu}}$ and $g_{\boldsymbol{\mu}}$ coincide.

We would like to apply Lemma 4.3, but we cannot do this directly since the required lower bounds on the order two derivatives do not hold in general. By considering different cases and doing an appropriate change of variable, we can however put us in a situation where we can apply Lemma 4.3. Titchmarsh makes similar considerations in the proof of [21, Lemma ζ] and on [21, p. 497], but his argument gets simplified by the fact that $G_{\boldsymbol{\mu}}^{(2,0)} = -aG_{\boldsymbol{\mu}}^{(0,2)}$ when $b = 0$ (which is *not* true for $b \neq 0$).
The idea to deal with the non-diagonal case is quite simply to consider two cases; if the partial derivative $G_{\boldsymbol{\mu}}^{(1,1)}$ is small then the lower bound on the Hesse-determinant forces the two other partial derivatives to be large. If on the other hand $G_{\boldsymbol{\mu}}^{(1,1)}$ is large then after a change of variable, we can force the new partial derivatives $(2,0)$ and $(0,2)$ to be large. This will allow us to prove the following key lemma.

Lemma 4.8. *With notation as above we have*

$$\int_{\Delta_{p,q}} e^{iG_{\boldsymbol{\mu}}(\boldsymbol{x})} d\boldsymbol{x} \ll |t|^{-1+\varepsilon} \frac{Q(\boldsymbol{X})^{3/2}}{a^{1/2}Q(\boldsymbol{\mu})^{1/2}}.$$

Proof. Firstly we make a change of variables to the new variables $\boldsymbol{y} = (y_1, y_2) = (a^{1/4}x_1, a^{-1/4}x_2)$, under which the integral becomes

$$\int_{\tilde{\Delta}_{p,q}} e^{i\tilde{G}_{\boldsymbol{\mu}}(\boldsymbol{y})} d\boldsymbol{y}, \tag{4.14}$$

where $\tilde{G}_{\boldsymbol{\mu}}(\boldsymbol{y}) = G_{\boldsymbol{\mu}}(a^{-1/4}y_1, a^{1/4}y_2)$ and the new rectangle $\tilde{\Delta}_{p,q}$ has side lengths $(a^{1/4}l_1) \times (a^{-1/4}l_2)$.
The reason for doing this change of variable is that by the bounds in Lemma 4.7 and the chain rule, it now follows that all order two partial derivates of $\tilde{G}_{\boldsymbol{\mu}}$ are bounded by

$$\ll |t| a^{1/2} Q(\boldsymbol{\mu})^{1/2} Q(\boldsymbol{X})^{-3/2} =: r. \tag{4.15}$$

Let $\lambda_1, \lambda_2 > 0$ be constants independent of a, b and t (large enough) such that

$$|\tilde{G}^{\alpha}_{\boldsymbol{\mu}}(\boldsymbol{y})| \le \lambda_1 r, \tag{4.16}$$

$$|\tilde{G}^{(2,0)}_{\boldsymbol{\mu}}(\boldsymbol{y})\tilde{G}^{(0,2)}_{\boldsymbol{\mu}}(\boldsymbol{y}) - (\tilde{G}^{(1,1)}_{\boldsymbol{\mu}}(\boldsymbol{y}))^2| \ge \lambda_2 r^2, \tag{4.17}$$

for $\alpha \in \{(2,0),(1,1),(0,2)\}$ and $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$. We now split into different cases depending on the sizes of the order two partial derivatives.

Case 1: Assume that $(\tilde{G}^{(1,1)}_{\boldsymbol{\mu}}(\boldsymbol{y}))^2 < \lambda_2 r^2/2$ for all $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$.
Then it follows from (4.17) that

$$|\tilde{G}^{(2,0)}_{\boldsymbol{\mu}}(\boldsymbol{y})\tilde{G}^{(0,2)}_{\boldsymbol{\mu}}(\boldsymbol{y})| > \lambda_2 r^2/2.$$

Thus we conclude using the bound (4.16) above

$$\lambda_2 r^2/2 < |\tilde{G}^{(2,0)}_{\boldsymbol{\mu}}(\boldsymbol{y})\tilde{G}^{(0,2)}_{\boldsymbol{\mu}}(\boldsymbol{y})| < \lambda_1 r |\tilde{G}^{(2,0)}_{\boldsymbol{\mu}}(\boldsymbol{y})|,$$

and thus $|\tilde{G}^{(2,0)}_{\boldsymbol{\mu}}(\boldsymbol{y})| \gg r$ and similarly for $\tilde{G}^{(0,2)}_{\boldsymbol{\mu}}(\boldsymbol{y})$. The result now follows from Lemma 4.3.

Case 2: Assume that $|\tilde{G}^{(1,1)}_{\boldsymbol{\mu}}(\boldsymbol{y})|^2 \ge \lambda_2 r^2/2$ for some $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$.
This we will show implies that for any $\delta > 0$, we have

$$|\tilde{G}^{(1,1)}_{\boldsymbol{\mu}}(\boldsymbol{y})| \ge (2^{-1/2} - \delta)\lambda_2^{1/2} r$$

for *all* $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$ when t is sufficiently large. To see this we bound the variation of $\tilde{G}^{(1,1)}_{\boldsymbol{\mu}}$ in $\tilde{\Delta}_{p,q}$; we observe that by the chain rule

$$\tilde{G}^{(i,j)}_{\boldsymbol{\mu}}(\boldsymbol{y}) = a^{(j-i)/4} G^{(i,j)}_{\boldsymbol{\mu}}(a^{-1/4}y_1, a^{1/4}y_2),$$

and thus by applying (4.10), we can bound the variation of $\tilde{G}_{\boldsymbol{\mu}}^{(1,1)}$ in $\tilde{\Delta}_{p,q}$ by;

$$\begin{aligned}
&a^{1/4}l_1 \cdot \max_{\boldsymbol{y}\in\tilde{\Delta}_{p,q}} |\tilde{G}_{\boldsymbol{\mu}}^{(2,1)}(\boldsymbol{y})| + a^{-1/4}l_2 \cdot \max_{\boldsymbol{y}\in\tilde{\Delta}_{p,q}} |\tilde{G}_{\boldsymbol{\mu}}^{(1,2)}(\boldsymbol{y})| \\
&= l_1 \cdot \max_{\boldsymbol{x}\in\Delta_{p,q}} |G_{\boldsymbol{\mu}}^{(2,1)}(\boldsymbol{x})| + l_2 \cdot \max_{\boldsymbol{x}\in\Delta_{p,q}} |G_{\boldsymbol{\mu}}^{(1,2)}(\boldsymbol{x})| \\
&\ll \frac{Q(\boldsymbol{X})^{3/2}}{aQ(\boldsymbol{\mu})^{1/2}|t|^{1+2\varepsilon}} \cdot \frac{aQ(\boldsymbol{\mu})^{1/2}|t|}{Q(\boldsymbol{X})^2} + \frac{Q(\boldsymbol{X})^{3/2}}{a^{1/2}Q(\boldsymbol{\mu})^{1/2}|t|^{1+2\varepsilon}} \cdot \frac{a^{1/2}Q(\boldsymbol{\mu})^{1/2}|t|}{Q(\boldsymbol{X})^2} \\
&\ll |t|^{-2\varepsilon}Q(\boldsymbol{X})^{-1/2},
\end{aligned}$$

which is $o(r)$ as $t \to \infty$ since $Q(\boldsymbol{X}) \ll a^{1/2}|t|^{1+\varepsilon}$ (recall the definition (4.15) of r). Now we have two further sub-cases.

Case 2.1: If

$$|\tilde{G}_{\boldsymbol{\mu}}^{(2,0)}(\boldsymbol{y})|, |\tilde{G}_{\boldsymbol{\mu}}^{(0,2)}(\boldsymbol{y})| > 2^{-2}\lambda_1^{-1}\lambda_2 r,$$

for *all* $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$, then we can apply Lemma 4.3 directly.

Case 2.2: So we may assume that, say, $|\tilde{G}_{\boldsymbol{\mu}}^{(2,0)}(\boldsymbol{y})| \leq 2^{-2}\lambda_1^{-1}\lambda_2 r$ for some $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$. As above, we see using (4.10) that the variation of $\tilde{G}_{\boldsymbol{\mu}}^{(2,0)}$ in $\tilde{\Delta}_{p,q}$ is bounded by

$$\begin{aligned}
&a^{1/4}l_1 \cdot \max_{\boldsymbol{y}\in\tilde{\Delta}_{p,q}} |\tilde{G}_{\boldsymbol{\mu}}^{(3,0)}(\boldsymbol{y})| + a^{-1/4}l_2 \cdot \max_{\boldsymbol{y}\in\tilde{\Delta}_{p,q}} |\tilde{G}_{\boldsymbol{\mu}}^{(2,1)}(\boldsymbol{y})| \\
&= a^{-1/2}l_1 \cdot \max_{\boldsymbol{x}\in\Delta_{p,q}} |G_{\boldsymbol{\mu}}^{(3,0)}(\boldsymbol{x})| + a^{-1/2}l_2 \cdot \max_{\boldsymbol{x}\in\Delta_{p,q}} |G_{\boldsymbol{\mu}}^{(2,1)}(\boldsymbol{x})| \\
&\ll \frac{Q(\boldsymbol{X})^{3/2}}{a^{3/2}Q(\boldsymbol{\mu})^{1/2}|t|^{1+2\varepsilon}} \cdot \frac{a^{3/2}Q(\boldsymbol{\mu})^{1/2}|t|}{Q(\boldsymbol{X})^2} + \frac{Q(\boldsymbol{X})^{3/2}}{aQ(\boldsymbol{\mu})^{1/2}|t|^{1+2\varepsilon}} \cdot \frac{aQ(\boldsymbol{\mu})^{1/2}|t|}{Q(\boldsymbol{X})^2} \\
&\ll |t|^{-2\varepsilon}Q(\boldsymbol{X})^{-1/2},
\end{aligned}$$

which as above is $o(r)$ as $t \to \infty$. Thus we conclude that for any $\delta' > 0$;

$$|\tilde{G}_{\boldsymbol{\mu}}^{(2,0)}(\boldsymbol{y})| \leq (2^{-2} + \delta')\lambda_1^{-1}\lambda_2 r$$

holds for *all* $\boldsymbol{y} \in \tilde{\Delta}_{p,q}$ when t is sufficiently large.

If we write

$$\boldsymbol{z} = (z_1, z_2) = (dy_1 - cy_2, dy_1 + cy_2), \tag{4.18}$$

with $cd = 1/2$, then after a change of variable the integral (4.14) becomes

$$\int_{\Omega_{p,q}} e^{ih(\boldsymbol{z})} d\boldsymbol{z},$$

where $h(\boldsymbol{z}) = \tilde{G}_{\boldsymbol{\mu}}(cz_1 + cz_2, -dz_1 + dz_2)$ and $\Omega_{p,q}$ is a new rectangle with angle $\pi/4$ to the coordinate axis and maximum side length $\ll a^{1/4}l_1 \max(c, d)$. We observe that

$$\begin{aligned}
h^{(2,0)} &= c^2\tilde{G}_{\boldsymbol{\mu}}^{(2,0)} + d^2\tilde{G}_{\boldsymbol{\mu}}^{(0,2)} - \tilde{G}_{\boldsymbol{\mu}}^{(1,1)}, \\
h^{(0,2)} &= c^2\tilde{G}_{\boldsymbol{\mu}}^{(2,0)} + d^2\tilde{G}_{\boldsymbol{\mu}}^{(0,2)} + \tilde{G}_{\boldsymbol{\mu}}^{(1,1)}, \\
h^{(1,1)} &= c^2\tilde{G}_{\boldsymbol{\mu}}^{(2,0)} - d^2\tilde{G}_{\boldsymbol{\mu}}^{(0,2)}.
\end{aligned}$$

Thus by choosing $c = \lambda_1^{1/2}\lambda_2^{-1/4}, d = \lambda_1^{-1/2}\lambda_2^{1/4}/2$ and δ, δ' sufficiently small, we get for *all* $\boldsymbol{z} \in \Omega_{p,q}$ the following bounds;

$$r \ll (2^{-1/2} - 1/2 - \delta - \delta')\lambda_2^{1/2} r \leq |h^{(2,0)}(\boldsymbol{z})|, |h^{(0,2)}(\boldsymbol{z})| \ll r, \quad |h^{(1,1)}(\boldsymbol{z})| \ll r.$$

Since the determinant of the Hesse-matrix is unchanged under the change of variable corresponding to (4.18), the result follows from Lemma 4.3. Observe that the implied constant we get from Lemma 4.3 is indeed uniform in a, b and t since the angles of the rectangles $\Omega_{p,q}$ to the coordinate axes are fixed. □

We are now ready to finish the proof of our main theorem.

Proof of Proposition 3.1 and Theorem 1.6. Combining (4.13) and Lemma 4.8, we get the following bound for all $\boldsymbol{\mu}$ as above;

$$\begin{aligned} S'(\boldsymbol{\mu}) &= \sum_{p,q} S_{p,q}(\boldsymbol{\mu}) \\ &\ll \sum_{p,q} \left(|t|^{-1+\varepsilon} \frac{Q(\boldsymbol{X})^{3/2}}{a^{1/2}Q(\boldsymbol{\mu})^{1/2}} + l_2 \right) \\ &\ll \frac{X_1 X_2}{a^{-3/2}Q(\boldsymbol{X})^3 |t|^{-2-4\varepsilon} Q(\boldsymbol{\mu})^{-1}} \cdot \left(|t|^{-1+\varepsilon} \frac{Q(\boldsymbol{X})^{3/2}}{a^{1/2}Q(\boldsymbol{\mu})^{1/2}} + \frac{Q(\boldsymbol{X})^{3/2}}{a^{1/2}|t|^{1+2\varepsilon}Q(\boldsymbol{\mu})^{1/2}} \right) \\ &\ll a^{1/2} \frac{|t|^{1+5\varepsilon}Q(\boldsymbol{\mu})^{1/2}}{Q(\boldsymbol{X})^{1/2}}, \end{aligned}$$

where we used $a^{1/2}X_1X_2 \ll Q(\boldsymbol{X})$. Plugging this into Lemma 4.1 yields;

$$\begin{aligned} &\frac{1}{Q(\boldsymbol{X})^{1/2}} \sum_{\substack{X_1 \leq x_1 \leq X_1' \\ X_2 \leq x_2 \leq X_2'}} e^{if(x_1,x_2)} \\ &\qquad \ll \frac{X_1X_2}{Q(\boldsymbol{X})^{1/2}\rho} + \frac{(X_1X_2)^{1/2}}{Q(\boldsymbol{X})^{1/2}\rho} \left(\sum_{0 \leq \mu_1, \mu_2 \leq \rho} \frac{a^{1/2}|t|^{1+5\varepsilon}Q(\boldsymbol{\mu})^{1/2}}{Q(\boldsymbol{X})^{1/2}} \right)^{1/2} \\ &\qquad \ll \frac{Q(\boldsymbol{X})^{1/2}}{a^{1/2}\rho} + \frac{|t|^{1/2+3\varepsilon}}{Q(\boldsymbol{X})^{1/4}\rho} \left(\sum_{0 \leq \mu_1, \mu_2 \leq \rho} Q(\boldsymbol{\mu})^{1/2} \right)^{1/2} \\ &\qquad \ll \frac{Q(\boldsymbol{X})^{1/2}}{a^{1/2}\rho} + \frac{|t|^{1/2+3\varepsilon}a^{1/4}}{Q(\boldsymbol{X})^{1/4}\rho} \left(\sum_{||\boldsymbol{\mu}|| \leq \rho} ||\boldsymbol{\mu}|| \right)^{1/2} \\ &\qquad \ll \frac{Q(\boldsymbol{X})^{1/2}}{a^{1/2}\rho} + \frac{|t|^{1/2+3\varepsilon}a^{1/4}\rho^{1/2}}{Q(\boldsymbol{X})^{1/4}}. \end{aligned}$$

Finally we choose an integer $\rho \asymp Q(\boldsymbol{X})^{1/2}|t|^{-1/3}a^{-1/2}$ to balance the terms, which yields the desired bound $\ll_\varepsilon |t|^{1/3+3\varepsilon}$. This choice of ρ is admissible with respect to the conditions in Lemma 4.1 since first of all

$$\rho \ll a^{1/4}|t|^{1/2+\varepsilon}|t|^{-1/3}a^{-1/2} = |t|^{1/6+\varepsilon}a^{-1/4},$$

which is less than X_1 and X_2 by (3.8) and secondly we have $\rho \gg 1$, which again follows from (3.8).
This proves Proposition 3.1 and consequently we conclude the proof of Theorem 1.6. □

5. Lower bounds for the sup norm and a conjecture

As a concluding remark we will make some consideration on the best possible bound of the type (1.4). First of all the appearance of y^δ in (1.4) is necessary in the sense that for a fixed t, the Eisenstein series is unbounded because of the constant Fourier coefficient. We will now show that the lower bound $\delta \geq 1/2$ holds for any bound of the form (1.4) and state a uniform version of the sup norm conjecture for Eisenstein series.

We have for t fixed the following crude bound for the K-Bessel function [12, p. 60];

$$K_{it}(y) \ll_t y^{-1/2}e^{-y},$$

as $y \to \infty$. Thus from the Fourier expansion of the Eisenstein series [12, Theorem 3.4];

$$E(z,s) = y^s + \varphi(s)y^{1-s} + 4\sqrt{y}\sum_{n\geq 1} \frac{K_{s-1/2}(2\pi yn)\tau_{s-1/2}(n)}{\Gamma(s)\zeta(2s)\pi^{-s}} \cos(2\pi xn),$$

we see that

$$E(z,1/2+it) = y^{1/2+it} + \varphi(1/2+it)y^{1/2-it} + O_t(e^{-\pi y}). \tag{5.1}$$

Now observe that for fixed $t \geq 1$, we can choose arbitrarily large y such that

$$1 + \varphi(1/2+it)y^{-2it} = 2,$$

using that $|\varphi(1/2+it)| = 1$.
For such y, we thus have

$$E(z,1/2+it) = y^{1/2}(2y^{it} + o_t(1)) \gg y^{1/2},$$

when t is sufficiently large. Since we can let $y \to \infty$, we conclude that any bound of the form (1.4) has to satisfy $\delta \geq 1/2$.

One might speculate that the following holds for any $\varepsilon > 0$;

(5.2) **Conjecture:** $E(z,1/2+it) \ll_\varepsilon y^{1/2}(|t|+1)^\varepsilon$,

uniformly for $z \in \mathcal{F}$, the standard fundamental domain (1.3) for $\Gamma_0(1)$. Note that this conjecture together with (1.6) implies simultaneous Lindelöf in the t-aspect and on average in the D-aspect for the family of class group L-functions of imaginary quadratic fields.

References

1. Edgar Assing, *On sup-norm bounds part II:* GL(2) *Eisenstein series*, Forum Math. **31** (2019), no. 4, 971–1006. MR 3975671
2. Valentin Blomer, *Non-vanishing of class group L-functions at the central point*, Ann. Inst. Fourier (Grenoble) **54** (2004), no. 4, 831–847. MR 2111013
3. ———, *Epstein zeta-functions, subconvexity, and the purity conjecture*, Journal of the Institute of Mathematics of Jussieu (2018), 1–16.
4. Valentin Blomer, Gergely Harcos, and Philippe Michel, *Bounds for modular L-functions in the level aspect*, Ann. Sci. École Norm. Sup. (4) **40** (2007), no. 5, 697–740. MR 2382859
5. Samuel Dittmer, Michael Proulx, and Stephanie Seybert, *Some arithmetic problems related to class group L-functions*, Ramanujan J. **37** (2015), no. 2, 257–268. MR 3341689
6. William Duke, John B. Friedlander, and Henryk Iwaniec, *Class group L-functions*, Duke Math. J. **79** (1995), no. 1, 1–56. MR 1340293
7. ———, *The subconvexity problem for Artin L-functions*, Invent. Math. **149** (2002), no. 3, 489–577. MR 1923476
8. William Duke, Özlem Imamoḡlu, and Árpád Tóth, *Cycle integrals of the j-function and mock modular forms*, Ann. of Math. (2) **173** (2011), no. 2, 947–981. MR 2776366
9. ———, *Geometric invariants for real quadratic fields*, Ann. of Math. (2) **184** (2016), no. 3, 949–990. MR 3549627
10. Yueke Hu and Abhishek Saha, *Sup-norms of eigenfunctions in the level aspect for compact arithmetic surfaces, II: newforms and subconvexity*, arXiv e-prints (2019), arXiv:1905.06295.
11. Bingrong Huang and Zhao Xu, *Sup-norm bounds for Eisenstein series*, Forum Math. **29** (2017), no. 6, 1355–1369. MR 3719306
12. Henryk Iwaniec, *Spectral methods of automorphic forms*, second ed., Graduate Studies in Mathematics, vol. 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002. MR 1942691
13. Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR 2061214
14. Henryk Iwaniec and Peter Sarnak, *L^∞ norms of eigenfunctions of arithmetic surfaces*, Ann. of Math. (2) **141** (1995), no. 2, 301–320. MR 1324136
15. Philippe Michel and Akshay Venkatesh, *Heegner points and non-vanishing of Rankin/Selberg L-functions*, Analytic number theory, Clay Math. Proc., vol. 7, Amer. Math. Soc., Providence, RI, 2007, pp. 169–183. MR 2362200
16. ———, *The subconvexity problem for* GL_2, Publ. Math. Inst. Hautes Études Sci. (2010), no. 111, 171–271. MR 2653249
17. Peter Sarnak, *Class numbers of indefinite binary quadratic forms. II*, J. Number Theory **21** (1985), no. 3, 333–346. MR 814010
18. ———, *Arithmetic quantum chaos*, The Schur lectures (1992) (Tel Aviv), Israel Math. Conf. Proc., vol. 8, Bar-Ilan Univ., Ramat Gan, 1995, pp. 183–236. MR 1321639
19. Peter Söhne, *An upper bound for Hecke zeta-functions with Groessencharacters*, J. Number Theory **66** (1997), no. 2, 225–250. MR 1473880
20. Nicolas Templier, *Heegner points and Eisenstein series*, Forum Math. **23** (2011), no. 6, 1135–1158. MR 2855045
21. Edward C. Titchmarsh, *On Epstein's Zeta-Function*, Proc. London Math. Soc. (2) **36** (1934), 485–500. MR 1575971
22. ———, *The theory of the Riemann zeta-function*, second ed., The Clarendon Press, Oxford University Press, New York, 1986, Edited and with a preface by D. R. Heath-Brown. MR 882550
23. Han Wu, *Burgess-like subconvexity for* GL_1, Compos. Math. **155** (2019), no. 8, 1457–1499. MR 3977317
24. Han Wu and Nickolas Andersen, *Explicit subconvexity for* GL_2 *and some applications*, arXiv e-prints (2018), arXiv:1812.04391.
25. Matthew P. Young, *Weyl-type hybrid subconvexity bounds for twisted L-functions and Heegner points on shrinking sets*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 5, 1545–1576. MR 3635360

26. ———, *A note on the sup norm of Eisenstein series*, Q. J. Math. **69** (2018), no. 4, 1151—-1161. MR 3908697
27. Shou-Wu Zhang, *Gross-Zagier formula for* GL_2, Asian J. Math. **5** (2001), no. 2, 183–290. MR 1868935
28. ———, *Gross-Zagier formula for* $\mathrm{GL}(2)$. *II*, Heegner points and Rankin L-series, Math. Sci. Res. Inst. Publ., vol. 49, Cambridge Univ. Press, Cambridge, 2004, pp. 191–214. MR 2083213

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark
E-mail address: acnordentoft@outlook.com

PAPER G
SMALL SCALE EQUIDISTRIBUTION OF HECKE EIGENFORMS AT INFINITY

SMALL SCALE EQUIDISTRIBUTION OF HECKE EIGENFORMS AT INFINITY

ASBJØRN C. NORDENTOFT, YIANNIS N. PETRIDIS, AND MORTEN S. RISAGER

Abstract. We investigate the equidistribution of Hecke eigenforms on sets that are shrinking towards infinity. We show that at scales finer than the Planck scale they do *not* equidistribute while at scales more coarse than the Planck scale they equidistribute on a full density subsequence of eigenforms. On a suitable set of test functions we compute the variance showing interesting transition behavior at half the Planck scale.

1. Introduction

It is a fundamental consequence of Berry's random wave conjecture [1] that we expect the eigenfunctions of the Laplace operator on a hyperbolic manifold $M = \Gamma\backslash\mathbb{H}$ to 'spread out' in the large eigenvalue limit. For a measure $d\nu'$ on $\Gamma\backslash\mathbb{H}$ and a sufficiently nice function ψ on $\Gamma\backslash\mathbb{H}$ we write

$$\langle \psi, d\nu' \rangle = \int_{\Gamma\backslash\mathbb{H}} \psi(z) d\nu'(z).$$

Let φ_λ be L^2-normalized eigenfunctions of the Laplacian with eigenvalue λ, and consider the measures

$$d\mu_\lambda = |\varphi_\lambda|^2 \, d\mu, \quad d\nu = \frac{d\mu}{\mathrm{vol}\,(\Gamma\backslash\mathbb{H})},$$

where $d\mu(z) = y^{-2}dxdy$ is the uniform measure on the surface.

The question about whether the eigenfunctions are indeed spread out is quantified by the question of whether

$$\langle \psi, d\mu_\lambda \rangle \to \langle \psi, d\nu \rangle\,, \text{ as } \lambda \to \infty \tag{1}$$

for a suitable set of test functions ψ.

For the full modular group $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ with φ_λ being Hecke–Maass forms this (and much more) was famously proved by Lindenstrauss [14] and Soundararajan [26]. Zelditch [29] had previously studied the variance sum

$$\sum_{\lambda \le \Lambda} |\langle \psi, d\mu_\lambda \rangle - \langle \psi, d\mu \rangle|^2$$

Date: December 4, 2020.

2000 *Mathematics Subject Classification.* Primary 58J50; Secondary 11F72 .

Key words and phrases. Equidistribution in shrinking sets, Quantum Variance, Shifted convolution sums.

providing weak but non-trivial upper bounds on this to conclude (1) for a full density subsequence of λ (See also [25, 28]). For the full modular group Sarnak and Zhao [24] were able to prove asymptotics for the variance sum on a suitable set of test functions, and Nelson [19–21] has recently found a way to determine the asymptotics also for arithmetic compact hyperbolic surfaces arising from maximal orders in quaternion algebras.

It is natural to ask if the equidistribution (1) still holds if we allow the support of the test function ψ to shrink as a function of λ. An interesting special case is when ψ is the indicator function of a hyperbolic ball of radius R with R going to zero as a function of λ. This is the question of equidistribution in 'shrinking sets', which has been analyzed e.g. by Young [27, Prop 1.5]. The physics literature seems to suggest (see also [5]) that we may expect equidistribution to hold all the way down to the scale of the de Broglie wavelength, which is of the order of $1/\sqrt{\lambda}$. Humphries [8] has shown that if we go below this threshold, also called the Planck scale, then there are cases where equidistribution does *not* hold.

Humphries and Khan [9] proved that individual equidistribution holds all the way to the Planck scale, if we restrict to dihedral forms, which form a very thin set of Maass forms.

It should be noted that ergodic theory methods provide equidistribution in shrinking balls for general negatively curved manifolds but typically only for a slow logarithmic rate, see e.g. [4, 6].

On the other hand for the eigenfunctions on the Euclidean torus Granville and Wigman [3] showed individual equidistribution close to the Planck scale and failure of equidistribution at scales finer than the Planck scale. This was previously proved by Lester and Rudnick [13] along a full density subsequence.

1.1. **Mass equidistribution for holomorphic Hecke cusp forms.** We may ask questions analogous to the above if we replace the eigenfunction φ_λ by $y^{k/2}f(z)$, where $f(z)$ is a holomorphic cusp form of weight k. In fact $y^{k/2}f(z)$ *is* an eigenfunction of the weight k Laplacian Δ_k for the full modular group with eigenvalue $-k/2(1-k/2)$, which is the bottom of the spectrum for Δ_k . Holowinsky and Soundararajan [7] proved that in analogy with (1) we have

$$\mu_f(\psi) := \langle \psi, d\mu_f \rangle \to \langle \psi, d\nu \rangle \,, \text{ as } k \to \infty,$$

where

$$d\mu_f = y^k \left| f(z) \right|^2 d\mu.$$

Luo and Sarnak [18] computed the quantum variance of these measures on the modular surface. More precisely they proved that for a fixed compactly supported function u on the positive reals we have

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f \in H_k} L(1, \operatorname{sym}^2 f) \left|\mu_f(\psi)\right|^2 = B_\omega(\psi, \psi) K + O_{\varepsilon,\psi}(K^{1/2+\varepsilon}).$$

Here H_k is an orthonormal basis of Hecke eigenforms, $L(s, \operatorname{sym}^2 f)$ is the symmetric square L-function of f, and ψ is a rapidly decaying smooth function of mean zero whose zero-th Fourier coefficient vanishes sufficiently high in the cusp, and $B_\omega(\psi_1, \psi_2)$

is a Hermitian form diagonalized by cusp forms. The eigenvalues of B_ω are arithmetically significant: they are $\pi/2$ times the central value of the corresponding Hecke L-function.

1.2. **Equidistribution on shrinking sets.** The question of equidistribution on shrinking sets in the holomorphic setting was considered by Lester, Matomäki, and Radziwiłł [12]. They proved an effective version of the result of Holowinsky and Soundararajan, allowing to shrink the test function at the rate of a small negative power of $\log k$.

We consider the following variant of the 'shrinking sets' problem: Let H be a large number and define the set

$$B_H = \{z \in \Gamma\backslash\mathbb{H} : \Im(z) > H\},$$

considered to be a shrinking ball around infinity. We study the distribution of compactly supported functions on B_1 squeezed into B_H using the operator M_H defined by

$$M_H\psi(z) = \psi(x + iy/H).$$

This may be formulated in a coordinate-independent way. See Section 3.1. Similar shrinking has been considered before by Ghosh and Sarnak [2] as well as by Lester, Matomäki, and Radziwiłł [12].

The length scale of B_H is of the order of H^{-1} so we might expect equidistribution to hold all the way down to $H^{-1} \gg k^{-1}$, as this is the order of the de Broglie wavelength of $y^{k/2}f(z)$.

Let $B := B_1$. We will consider the following class of functions:

$$C_0^\infty(M, B) = \{\psi \in C_0^\infty(M) \mid \operatorname{supp}\psi \subset B\},$$

where $C_0^\infty(M)$ consists of all smooth functions on $M = \Gamma\backslash\mathbb{H}$ decaying rapidly at the cusp, and such that the zero-th Fourier coefficient vanishes sufficiently high in the cusp. Given $\psi \in C_0^\infty(M, B)$, we investigate upper bounds and asymptotics for

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f \in H_k} L(1, \operatorname{sym}^2 f) \left|\mu_f(M_{H(k)}\psi) - \nu(M_{H(k)}\psi)\right|^2,$$

where $H(k) = (k-1)^\theta$ for some $0 \leq \theta < 1$, and $u : \mathbb{R}_+ \to \mathbb{R}_{\geq 0}$ is smooth with compact support. It turns out that the asymptotics depends crucially on θ.

1.3. **Mass equidistribution below and above the Planck scale.** We first prove that mass equidistribution fails on shrinking sets around infinity as above for scales finer than the Planck scale. This is consistent with the above prediction.

Proposition 1.1. *Let $\theta \geq 1$, i.e. shrinking below the Planck scale. Then there exists $\psi \in C_0^\infty(M, B)$ such that $\mu_f(M_{(k-1)^\theta}\psi) = o(\nu(M_{(k-1)^\theta}\psi))$.*

Secondly we obtain a power-saving bound for the quantum variance sum for general observables all the way down to the Planck scale. This implies that mass equidistribution holds for a density one subsequence of holomorphic cusp forms.

Theorem 1.2. *Let* $0 < \theta < 1$ *and* $\psi \in C_0^\infty(M, B)$. *Then*

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f \in H_k} L(1, \operatorname{sym}^2 f) \left|\mu_f(M_{(k-1)^\theta}\psi) - \nu(M_{(k-1)^\theta}\psi)\right|^2 = O_{\psi,u}(K^{2-2\theta-\min(1/5, 1-\theta)+\varepsilon}).$$

Since $\nu(M_{(k-1)^\theta}\psi)$ is of size about $k^{-\theta}$ this supplements [12, Theorem 1.3] as it shows that equidistribution holds on average at a much finer scale than proved individually in [12]. The precise polynomial saving of $1/5$ can probably be improved; its proof has as its input the convexity bound in the k-aspect of $L(s, \operatorname{sym}^2 f)$.

1.4. **Asymptotics of the quantum variance.** We let $C_{0,0}^\infty(M, B)$ denote functions in $C_0^\infty(M, B)$ that are orthogonal to the constant function. We note that for $\psi \in C_{0,0}^\infty(M, B)$ we have $\nu(M_{(k-1)^\theta}\psi) = 0$. If we restrict to test functions in this space we can improve on Theorem 1.2 and obtain an asymptotic result.

Theorem 1.3. *Let* $0 < \theta < 1$ *and fix* $u : \mathbb{R}_+ \to \mathbb{R}_{\geq 0}$ *smooth with compact support.*

(i) There exists a Hermitian form B_θ *on* $C_{0,0}^\infty(M, B)$ *and* $\delta_\theta > 0$ *such that*

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f \in H_k} L(1, \operatorname{sym}^2 f) \left|\mu_f(M_{(k-1)^\theta}\psi)\right|^2 = B_\theta(\psi, \psi) \left(\int u(y) y^{-\theta} dy\right) K^{1-\theta} + O_{\psi,\theta}(K^{1-\theta-\delta_\theta}),$$

for $\psi \in C_{0,0}^\infty(M, B)$.

(ii) The Hermitian forms B_θ *have three different regimes in the sense that* B_θ *is constant on each of the three intervals* $0 < \theta < 1/2$, $\theta = 1/2$ *and* $1/2 < \theta < 1$. *The decomposition*

$$C_{0,0}^\infty(M, B) = C_{\text{cusp}}^\infty(M, B) \oplus C_{\text{Eis}}^\infty(M, B),$$

into the cuspidal and the Eisenstein part is orthogonal with respect to B_θ *for all* $0 < \theta < 1$. *Furthermore* B_θ *restricted to* $C_{\text{Eis}}^\infty(M, B)$ *is independent of* θ.

(iii) The Hermitian forms B_θ *can be extended to the larger set* $1_B C_{0,0}^\infty(M)$ *such that the following holds: On the subset* $1_B C_{\text{cusp}}^\infty(M)$ *of functions with the zero-th Fourier coefficient vanishing, the form* B_θ *is continuous with respect to a certain Sobolev norm* $\|\cdot\|_{2,1}$. *The set* $C_{\text{cusp}}^\infty(M, B)$ *is dense in* $1_B C_{\text{cusp}}^\infty(M)$ *with respect to the same norm* $\|\cdot\|_{2,1}$.

(iv) If ϕ_i *are Hecke–Maass forms with eigenvalue* $s_i(1-s_i)$, *then the Hermitian form satisfies* $B_\theta(1_B\phi_1, 1_B\phi_2) = 0$ *unless* ϕ_1, ϕ_2 *are both even. If* ϕ_i *are both even, then*

$$B_\theta(1_B\phi_1, 1_B\phi_2) = 4\pi \sum_{m,n \geq 1} \frac{\tau_1((m,n))\lambda_{\phi_1}(m)\lambda_{\phi_2}(n)}{(mn)^{1/2}} I_\theta^{s_1,s_2}(m,n),$$

where

$$I_\theta^{s_1,s_2}(m,n) = \int_{\max(m,n)}^\infty K_{s_1-1/2}(2\pi y)\overline{K_{s_2-1/2}(2\pi y)} f_{\theta,m,n}(y) \frac{dy}{y}$$

with

$$f_{\theta,m,n}(y) = \begin{cases} 1, & \text{if } 0 < \theta < 1/2, \\ e^{-2\pi^2 y^2 (m^2+n^2)}, & \text{if } \theta = 1/2, \\ 0, & \text{if } \theta > 1/2. \end{cases}$$

For the precise form of B_θ and $\|\cdot\|_{2,1}$ we refer to (21) and (32).

Luo and Sarnak [18, p.773] proved that $L(\phi, 1/2)$ is non-negative for ϕ a Hecke–Maass cusp form by realizing it as eigenvalue of the Hermitian form B_0. One may speculate whether $B_\theta(1_B\phi, 1_B\phi)$ for $0 < \theta \leq 1/2$ is also related to central values of L-functions. Irrespectively, we may use Theorem 1.3 to prove that $B_\theta(1_B\phi, 1_B\phi) \geq 0$. Seeing this directly from the series representation in Theorem 1.3 (iv) seems difficult, and is as such surprising.

In fact this was our original motivation for extending B_θ in Theorem 1.3 (iii) to a set containing $1_B\phi$. Notice that $1_B\phi$ together with incomplete Eisenstein series provide a basis for $1_B C^\infty_{0,0}(M)$.

Corollary 1.4. *If ϕ is an even Hecke–Maass cusp form with eigenvalue $s_\phi(1-s_\phi)$ and Hecke eigenvalues $\lambda_\phi(n)$, then*

$$\sum_{m,n\geq 1} \frac{\tau_1((m,n))\lambda_\phi(m)\lambda_\phi(n)}{(mn)^{1/2}} \int_{\max(m,n)}^\infty \left|K_{s_\phi-1/2}(2\pi y)\right|^2 \frac{dy}{y} \geq 0.$$

Remark 1. Let $w : \mathbb{R} \to \mathbb{R}$ be any smooth and bounded weight function with support contained in $[1, \infty)$. Then one can similarly show by using the explicit expression for B_θ in (21) combined with Theorem 1.3 (i) for $\psi(z) = w(y)\phi(z)$ that

$$\sum_{m,n\geq 1} \frac{\tau_1((m,n))\lambda_\phi(m)\lambda_\phi(n)}{(mn)^{1/2}} \int_0^\infty \left|K_{s_\phi-1/2}(2\pi y)\right|^2 w\left(\frac{y}{m}\right) w\left(\frac{y}{n}\right) \frac{dy}{y} \geq 0.$$

Remark 2. We expect that the techniques and results in this paper will work with some modifications also for Maass cusp forms in the same way that the results in [16] are extended to the Maass case by Sarnak and Zhao [24]. For simplicity and clarity we restrict ourselves to the holomorphic case.

1.5. **The behavior of holomorphic cusp forms high in the cusp.** Ghosh and Sarnak [2] considered the distribution of the zeroes of holomorphic modular forms high in the cusp as the weight grows. By the work of Rudnick [23] mass equidistribution for holomorphic forms implies equidistribution of their zeroes in the fundamental domain. Ghosh and Sarnak observed that, although the proportion of zeroes in a shrinking ball around infinity (more precisely $H \gg \sqrt{K \log k}$) was proportional to the area of the domain, the statistical behavior of the zeroes was very different. They observed experimentally that the zeroes tend to localize on the two "real" lines $\Re z = -1/2$ and $\Re z = 0$, conjectured that 100% of the zeroes in these shrinking balls around infinity should lie on these two lines, and obtained some results in this direction. These results were then strengthened by Lester, Matomäki, and Radziwiłł [12].

The reason for the qualitative change in the behavior of holomorphic cusp forms high in the cusp has its roots in the fact that for all integers $1 \ll l \ll \sqrt{k/\log k}$, we have

$$\left(\frac{e}{l}\right)^{k-1} f(x+iy_l) = \lambda_f(l)e(xl) + O(k^{-\delta}), \tag{2}$$

where $y_l = (k-1)/4\pi l$, and $\delta > 0$ is some constant. This means that counting zeroes on the real lines reduces to detecting sign-changes of the Hecke eigenvalues $\lambda_f(l)$, which is exactly what was done in [12].

We observe that our bilinear form B_θ exhibits a *phase transition* at $\theta = 1/2$, which coincides exactly with the threshold in [2] and [12]. Combined, these results point towards the phenomenon that, although the mass of holomorphic cusp forms equidistribute all the way down to the Planck scale, the qualitative behavior changes high in the cusp at *half* the Planck scale.

The asymptotic (2) implies that $y^k|f(x+iy)|^2$ is essentially constant as x varies, at least when $y = y_l$ for some l as above. This provides intuition for the phenomena observed in this paper: $y^k|f(x+iy)|^2$ exhibits very strong cancellation with cuspidal test functions when we go to scales finer than halfway to the Planck scale. On the other hand for incomplete Eisenstein series the behavior is the same all the way down to the Planck scale, according to Theorem 1.3 (ii).

2. The variance of shifted convolution sums over a Hecke basis

An essential tool in understanding questions of equidistribution of Hecke eigenforms is understanding shifted convolution sums.

Let f be a weight k, level 1 holomorphic cuspidal Hecke eigenform, normalized such that its Fourier expansion

$$f(z) = \sum_{n=1}^{\infty} \lambda_f(n) n^{\frac{k-1}{2}} e(nz), \qquad \text{where } e(z) = e^{2\pi i z}$$

satisfies $\lambda_f(1) = 1$. The normalized Hecke eigenvalues satisfy the Hecke relations

$$\lambda_f(n)\lambda_f(m) = \sum_{d|(m,n)} \lambda_f\left(\frac{mn}{d^2}\right), \tag{3}$$

see [10, (6.38)]. Consider the shifted convolution sum

$$\begin{aligned} A_f^W(X,h) &:= \sum_{n\in\mathbb{N}} \lambda_f(n)\lambda_f(n+h) W((n+h/2)/X) \\ &= \sum_{d|h} \sum_{r\in\mathbb{N}} \lambda_f(r(r+d)) W\left(\frac{\frac{h}{d}(r+d/2)}{X}\right), \end{aligned} \tag{4}$$

where $W : \mathbb{R}_+ \to \mathbb{R}$ is smooth and supported in a compact interval, and where in the second line we have used the Hecke relations (3).

Let $\tau_1(n) = \sum_{d|n} d$, and let $L(s, \mathrm{sym}^2 f)$ be the the symmetric square L-function associated to f, i.e.

$$L(s, \mathrm{sym}^2 f) = \zeta(2s) \sum_{n=1}^{\infty} \frac{\lambda_f(n^2)}{n^s}, \qquad \text{when } \Re(s) > 1,$$

and defined on $\mathbb{C}$ by analytic continuation.

We investigate the variance of the smooth shifted convolution sums $A_f^W(X,h)$ over an orthonormal basis of Hecke eigenforms H_k and over k of size K. Let $u : \mathbb{R}_+ \to \mathbb{R}_{\geq 0}$ be a compactly supported function. We want to understand

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \frac{2\pi^2}{k-1} \sum_{f\in H_k} \frac{A_f^{W_1}(h_1, X(k))\overline{A_f^{W_2}(h_2, X(k))}}{L(1, \mathrm{sym}^2, f)}, \tag{5}$$

where

$$X(k) = (k-1)^{1-\theta}$$

for some $0 < \theta < 1$.

In order to better describe the dependence on W, h we use Sobolev norms

$$\|W\|_{l,p}^p = \sum_{0\leq i\leq l} \left\| \frac{d^i}{dy^i} W \right\|_p^p$$

$$\|W\|_{l,\infty} = \sum_{0\leq i\leq l} \left\| \frac{d^i}{dy^i} W \right\|_\infty .$$

For all compactly supported functions W we choose $a_W > 0, A_W > 1$ such that $\operatorname{supp} W \subseteq [a_W, A_W]$. For $h_1,\ h_2 \geq 1$ we denote $\|h\|_\infty = \max(h_1, h_2)$.

The main tool in understanding (5) is the Petersson formula, which states that

$$\begin{aligned} &\frac{2\pi^2}{k-1} \sum_{f\in H_k} \frac{\lambda_f(n_1)\lambda_f(n_2)}{L(1, \mathrm{sym}^2(f))} \\ &= \delta_{n_1,n_2} + 2\pi(-1)^{k/2} \sum_{c\geq 1} \frac{S(n_1, n_2; c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{n_1 n_2}}{c}\right), \end{aligned} \tag{6}$$

see e.g. [18, p. 776]. We will use the following estimate for the J-Bessel function:

$$J_{k-1}(x) \ll \left(\frac{ex}{2k}\right)^{k-1}, \tag{7}$$

see e.g. [15, p. 233].

To state our theorem we define, for functions $W_1, W_2 : \mathbb{R}_+ \to \mathbb{R}$ and $h_1, h_2 \in \mathbb{N}$,

$$B_{h_1,h_2}(W_1, W_2) = \tau_1((h_1, h_2)) \int_0^\infty W_1(h_1 y)\overline{W_2(h_2 y)} dy.$$

We now prove the following result.

Theorem 2.1. *Let $u : \mathbb{R}_+ \to \mathbb{R}_{\geq 0}$ be a smooth compactly supported weight function, and let $W_1, W_2 : \mathbb{R}_+ \to \mathbb{R}$ be smooth functions compactly supported below $A_{W_i} \geq 1$. Then*

$$\begin{aligned} &\sum_{2|k} u\left(\frac{k-1}{K}\right) \frac{2\pi^2}{k-1} \sum_{f\in H_k} \frac{A_f^{W_1}(h_1, X(k))\overline{A_f^{W_2}(h_2, X(k))}}{L(1, \mathrm{sym}^2, f)} \\ =& B_{h_1,h_2}(W_1, W_2) \frac{K^{2-\theta}}{2} \int_0^\infty u(y) y^{1-\theta} dy + O_{W_i,h_i,\theta}(K). \end{aligned}$$

The implied constant in the error term may be bounded by a constant depending only on θ times

$$(1+\|h\|_\infty)^{1+\varepsilon}(A_{W_1}A_{W_2})^C\,\|W_1\|_{C,\infty}\,\|W_2\|_{C,\infty}$$

for C sufficiently large depending on θ.

Proof. Using (4) and the Petersson formula (6) we find, for all $X\geq 1$,

$$\begin{aligned}&\frac{2\pi^2}{k-1}\sum_{f\in H_k}\frac{A_f^{W_1}(h_1,X)\overline{A_f^{W_2}(h_2,X)}}{L(1,\mathrm{sym}^2 f)}\\ &=\sum_{\substack{d_1|h_1\\ d_2|h_2}}\sum_{r_1,r_2\in\mathbb{N}}\delta_{r_1(r_1+d_1)=r_2(r_2+d_2)}W_1\left(\frac{h_1(r_1+d_1/2)}{d_1X}\right)\overline{W_2\left(\frac{h_2(r_2+d_2/2)}{d_2X}\right)}\\ &+2\pi(-1)^{k/2}\sum_{\substack{d_1|h_1\\ d_2|h_2}}\sum_{r_1,r_2\in\mathbb{N}}W_1\left(\frac{h_1(r_1+d_1/2)}{d_1X}\right)\overline{W_2\left(\frac{h_2(r_2+d_2/2)}{d_2X}\right)}\\ &\times\sum_{c\geq 1}\frac{S(r_1(r_1+d_1),r_2(r_2+d_2);c)}{c}J_{k-1}\left(\frac{4\pi\sqrt{r_1(r_1+d_1)r_2(r_2+d_2)}}{c}\right).\end{aligned}$$

We refer to the line with the Kronecker delta as the *diagonal term*, and the rest as the *off-diagonal term.*

To handle the diagonal term, we observe that for fixed positive $d_1\neq d_2$ the equation

$$r_1(r_1+d_1)=r_2(r_2+d_2) \tag{8}$$

has only finitely many positive solution. To see this we rewrite (8) as

$$(2r_1+d_1)^2-(2r_2+d_2)^2=d_1^2-d_2^2$$

Factoring the left-hand-side as $(2r_1+d_1+2r_2+d_2)(2r_1+d_1-2r_2-d_2)$ we see that any solution gives a factorization of $d_1^2-d_2^2$, and that any factorization of $d_1^2-d_2^2$ comes from at most 1 solution. This shows that there are at most $d(d_1^2-d_2^2)$ (where $d(n)$ denotes the number of divisors of n) solutions to with $d_1\neq d_2$; indeed we see that the total contribution from these terms is $O(\|h\|_\infty^\varepsilon\,\|W_1\|_\infty\,\|W_2\|_\infty)$.

For the remaining terms, i.e. $d_1=d_2=d$, $r_1=r_2=r$ we apply first Poisson summation in the r-variable and use that the Fourier transform of the function $y\mapsto W_1\left(\frac{h_1(y+d_1/2)}{d_1X}\right)\overline{W_2\left(\frac{h_2(y+d_2/2)}{d_2X}\right)}$ at r is bounded by an absolute constant times

$$|r|^{-n}\,(dX)^{-n+1}\,\|W_1(h_1\cdot)W_2(h_2\cdot)\|_{n,1}\,, \tag{9}$$

which follows from repeated integration by parts. We now see that

$$\sum_{\substack{d_1|h_1\\ d_2|h_2\\ d_1=d_2}}\sum_{r\in\mathbb{N}}W_1\left(\frac{h_1(r+d_1/2)}{d_1X}\right)\overline{W_2\left(\frac{h_2(r+d_2/2)}{d_2X}\right)}$$

equals the same expression with the sum over $r \in \mathbb{N}$ replaced by the same sum over $r \in \mathbb{Z}$ up to an error term of $O(\|h\|_\infty^{1+\varepsilon} \|W_1\|_\infty \|W_2\|_\infty)$. We then observe that

$$\sum_{d|(h_1,h_2)} \sum_{r\in\mathbb{Z}} W_1\left(\frac{h_1(r+d/2)}{dX}\right)\overline{W_2\left(\frac{h_2(r+d/2)}{dX}\right)}$$

$$= \sum_{d|(h_1,h_2)} \sum_{r\in\mathbb{Z}} \int_{-\infty}^{\infty} W_1\left(\frac{h_1(y+d/2)}{dX}\right)\overline{W_2\left(\frac{h_2(y+d/2))}{dX}\right)}e(-ry)dy$$

$$= \tau_1((h_1,h_2)) \int_{-\infty}^{\infty} W_1(h_1 y)\overline{W_2(h_2 y)}dyX + O(\|h\|_\infty^{\varepsilon} \|W_1(h_1\cdot)W_2(h_2\cdot)\|_{2,1},$$

where we have extended trivially the r-sum to all of r, then used Poisson summation and the bound (9) with $n = 2$. Now we average over k and apply Poisson summation in the k-variable. Using integration by parts on the dual side we find that for any $A > 0$, we have

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) X(k) = \frac{K^{2-\theta}}{2} \int_0^\infty u(y)y^{1-\theta}dy + O_{u,A}(K^{-A}),$$

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) = \frac{K}{2} \int_0^\infty u(y)dy + O_{u,A}(K^{-A}),$$

which yields the desired main term up to the stated error term.

For the off-diagonal terms we need to bound

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) 2\pi(-1)^{k/2} \sum_{\substack{d_1|h_1 \\ d_2|h_2 \\ r_1,r_2\in\mathbb{N}}} W_1\left(\frac{h_1(r_1+d_1/2)}{d_1(k-1)^{1-\theta}}\right)\overline{W_2\left(\frac{h_2(r_2+d_2/2)}{d_2(k-1)^{1-\theta}}\right)}$$
$$\times \sum_{c\geq 1} \frac{S(r_1(r_1+d_1), r_2(r_2+d_2); c)}{c} J_{k-1}\left(\frac{\Delta}{c}\right),$$

where $\Delta = 4\pi\sqrt{r_1(r_1+d_1)r_2(r_2+d_2)}$. We mimic the arguments of Luo and Sarnak [17, p 880–881]. We start by noticing that

(i) the summation over k is supported in $K \ll_u k \ll_u K$,

(ii) the summations over r_i are supported in

$$a_{W_i}\frac{d_i}{h_i}(k-1)^{1-\theta} \leq (r_i + d_i/2) \leq A_{W_i}\frac{d_i}{h_i}(k-1)^{1-\theta} \left(\ll_u A_{W_i}\frac{d_i}{h_i}K^{1-\theta}\right).$$

Using again $r_i(r_i+d_i) = (r_i+d_i/2)^2 - d_i^2/4$ we see that

(iii) in the support of the above sums, we have

$$\Delta \ll_u \|h\|_\infty^{\varepsilon} A_{W_1}A_{W_2}\frac{d_1 d_2}{h_1 h_2}K^{2(1-\theta)}.$$

We want to truncate the sum over c and notice that for r_1, r_2 in the support of the sums we may use the bound (7) on the Bessel function and the trivial bound on the

Kloosterman sum to see that

$$\sum_{c\geq M}\frac{S(r_1(r_1+d_1),r_2(r_2+d_2);c)}{c}J_{k-1}\left(\frac{\Delta}{c}\right)$$
$$\ll \sum_{c\geq M}\left(\frac{C_uA_{W_1}A_{W_2}d_1d_2K^{1-2\theta}}{h_1h_2c}\right)^{k-1}\ll_u \left(\frac{C_uA_{W_1}A_{W_2}d_1d_2K^{1-2\theta}}{h_1h_2M}\right)^{k-1}\frac{M}{K}.$$

We conclude that, if $M=C_uA_{W_1}A_{W_2}\frac{d_1d_2}{h_1h_2}K^{1-2\theta+\varepsilon}$, this term decays exponentially in K. Therefore

(iv) the sum in c above may be truncated at

$$c\ll_u A_{W_1}A_{W_1}\frac{d_1d_2}{h_1h_2}K^{1-2\theta+\varepsilon}$$

up to an additional error of $\ll_u \|W_1\|_\infty\|W_2\|_\infty A_{W_1}A_{W_2}K^{-A}$. We now quote lemmata 4.1 and 4.2 in [17] stating that for g a smooth function compactly supported function on $\mathbb{R}_+$ we have

$$\sum_{2|k}2\pi(-1)^{k/2}J_{k-1}(x)g(k-1)=-2\pi\int_{-\infty}^{\infty}\hat{g}(t)\sin(x\cos(2\pi t))dt, \tag{10}$$
$$\int_{-\infty}^{\infty}\hat{g}(t)\sin(x(1-2\pi^2t^2)))dt=\int_0^{\infty}\frac{g(\sqrt{2yx})}{(\pi y)^{1/2}}\sin(y+x-\pi/4)dy,$$
$$\int_{-\infty}^{\infty}\hat{g}(t)\cos(x(1-2\pi^2t^2)))dt=\int_0^{\infty}\frac{g(\sqrt{2yx})}{(\pi y)^{1/2}}\cos(y+x-\pi/4)dy. \tag{11}$$

In our case we apply (10) to the function

$$g(y)=u(y/K)W_1\left(\frac{h_1(r_1+d_1/2)}{d_1y^{1-\theta}}\right)\overline{W_2}\left(\frac{h_2(r_2+d_2/2)}{d_2y^{1-\theta}}\right).$$

This shows that the remaining part of the non-diagonal contribution can be bounded by an absolute constant times

$$\sum_{d_i|h_i}\sum_{r_i\geq 1}\sum_{c\geq 1}\left|\int_{-\infty}^{\infty}\hat{g}(t)\sin\left(\frac{\Delta}{c}\cos(2\pi t)\right)dt\right|, \tag{12}$$

with restrictions on the sums as (ii)-(iv) above. Here we have used the trivial estimate on the Kloosterman sums.

As in [17, Eq (4.4)] we now use a trigonometric identity and Taylor expansions to get

$$\begin{aligned}\sin(x\cos(2\pi t)) &= \sin\left(x(1-2\pi^2t^2) + x\sum_{n\geq 2}(-1)^n\frac{(2\pi t)^n}{(2n)!}\right)\\ &= \sin(x(1-2\pi^2t^2))\left(\sum_{0\leq n,m\leq N-1} c_{m,n}(xt^4)^{2n}t^{2m}\right)\\ &\quad + \cos(x(1-2\pi^2t^2))\left(\sum_{\substack{1\leq n\leq N\\0\leq m\leq N-1}} d_{m,n}(xt^4)^{2n-1}t^{2m}\right)\\ &\quad + O((xt^4)^{2N}+(xt^4)^{4N}+t^{2N}+t^{4N}),\end{aligned} \tag{13}$$

for any N, where $c_{m,n}$ and $d_{m,n}$ are real constants. In order to bound the term coming from the error-term above, we observe that all derivatives $g^{(m)}$ are supported in $K\cdot\operatorname{supp}(u)$ and we claim that, when r_1, r_2 satisfy (i)–(iv), we have the bound

$$g^{(m)}(y) \ll_{u,m} C_{W_1,W_2,m}K^{-m}, \tag{14}$$

where $C_{W_1,W_2,m} = \prod_{i=1,2}\|W_i\|_{m,\infty}A_{W_i}^m$.

To see why the claim is true we observe from Leibniz' rule that $g^{(m)}(y)$ is bounded by an absolute constant (depending on m) times

$$\max_{p_1+p_2+p_3=m}\left|\frac{d^{p_1}}{dy^{p_1}}u(y/K)\frac{d^{p_2}}{dy^{p_2}}W_1\left(\frac{h_1(r_1+d_1/2)}{d_1y^{1-\theta}}\right)\frac{d^{p_3}}{dy^{p_3}}\overline{W_2}\left(\frac{h_2(r_2+d_2/2)}{d_2y^{1-\theta}}\right)\right|.$$

Now we observe that by the chain rule

$$\frac{d^p}{dy^p}u(y/K) = u^{(p)}(y/K)K^{-p} \ll_{u,p} K^{-p}.$$

Using Faà di Bruno's formula for the higher derivative we see that

$$\begin{aligned}\frac{d^p}{dy^p}W_i\left(\frac{h_i(r_i+d_i/2)}{d_iy^{1-\theta}}\right) &\ll_p \|W_i\|_{p,\infty}\sum\prod_{j=1}^{p}\left(\frac{h_i(r_i+d_i/2)}{d_iy^{-\theta+1+j}}\right)^{m_j}\\ &= \|W_i\|_{p,\infty}\sum\left(\frac{h_i(r_i+d_i/2)}{d_iy^{-\theta+1}}\right)^{\sum_i m_i}y^{-p}.\end{aligned}$$

The sum is over p-tuples of integers satisfying $m_1+2m_2+\cdots pm_p = p$. Using that W_i is supported in $[a_{W_i}, A_{W_i}]$ we see that we may bound the big parenthesis in the last equation by A_{W_i}. For y in the support of g we have $y\in K\cdot\operatorname{supp}u$ so for such y we get

$$\frac{d^p}{dy^p}W_i\left(\frac{h_i(r_i+d_i/2)}{d_iy^{1-\theta}}\right) \ll_{u,p} \|W_i\|_{p,\infty}A_{W_i}^pK^{-p}.$$

Combining these bounds proves the claim (14).

From (14) it follows that $\widehat{g^{(m)}}(y) \ll_{u,m} C_{W_1,W_2,m}K^{-(m-1)}$. Additionally partial integration gives $\widehat{g^{(m)}}(t) \ll_{u,m} C_{W_1,W_2,m+l}|t|^{-l}K^{-(m+l-1)}$ so by using $\widehat{g^{(m)}}(y) =$

$(-2\pi iy)^m\hat{g}(y)$ we may conclude, by using the first bound for $|t|\leq K^{-1}$ and the second bound with $l=2$ when $|t|>K^{-1}$ that

$$\int_{-\infty}^{\infty}|\hat{g}(t)t^m|\,dt \ll_{u,m} C_{W_1,W_2,m+2}K^{-m}.$$

Using this bound we see that, when we use the Taylor expansion (13), the contribution from

(1) $((\Delta/c)t^4)^{2N}$ is $\ll_u \|h\|_\infty^\varepsilon C_{W_1,W_2,8N+2}\prod_{i=1,2}A_{W_i}^{2N+1}K^{(1-\theta)2(2N+1)-8N}$,
(2) $((\Delta/c)t^4)^{4N}$ is $\ll_u \|h\|_\infty^\varepsilon C_{W_1,W_2,16N+2}\prod_{i=1,2}A_{W_i}^{4N+1}K^{(1-\theta)2(4N+1)-16N}$,
(3) t^{2N} is $\ll_u \|h\|_\infty^\varepsilon C_{W_1,W_2,2N+2}\prod_{i=1,2}A_{W_i}^{2}K^{(3-4\theta)-2N}$,
(4) t^{4N} is $\ll_u \|h\|_\infty^\varepsilon C_{W_1,W_2,4N+2}\prod_{i=1,2}A_{W_i}^{2}K^{(3-4\theta)-4N}$.

We note that for $N=1$ all terms are $\ll_u \|h\|_\infty^\varepsilon C_{W_1,W_2,18}\prod_{i=1,2}A_{W_i}^5K$.

To bound the remaining terms involving

$$e(x,t):=\sin(x(1-2\pi^2t^2))c_{00}+\cos(x(1-2\pi^2t^2))d_{01}xt^4$$

coming from the Taylor expansion, we combine $\widehat{g^{(m)}}(y)=(-2\pi iy)^m\hat{g}(y)$ with (11) which gives the bound
(15)
$$\int_{-\infty}^{\infty}\hat{g}(t)e(\frac{\Delta}{c},t)dt \ll \max_{\substack{(m,m')=(4,1),(0,0)\\ \pm}}\left|\left(\frac{\Delta}{c}\right)^{m'}\int_0^\infty g^{(m)}\left(\sqrt{\frac{2\Delta}{c}y}\right)y^{-1/2}e^{\pm iy}dy\right|,$$

where we used Euler's formulas for sine and cosine. Now we apply partial integration to the integral with e^{iy} as one of the functions.

For r_1,r_2,Δ,c as in (ii)–(iv) (i.e. where the terms in the sum (12) might be non-vanishing) we claim that for any $n,m\in\mathbb{Z}_{\geq 0}$

$$\frac{d^n}{dy^n}\left(g^{(m)}\left(\sqrt{\frac{2\Delta}{c}y}\right)y^{-1/2}\right)\ll_{u,n,m}\frac{C_{W_1,W_2,m+n}}{K^my^{1/2+n}}.$$

To see this we note that the left-hand side is non-zero only if $\Delta y/c\asymp_u K^2$. By using the Leibniz rule and Faà di Bruno's formula we see that

$$\begin{aligned}\frac{d^n}{dy^n}\left(g^{(m)}\left(\sqrt{\frac{2\Delta y}{c}}\right)y^{-1/2}\right)&\ll_n\sum_{i=0}^n\left|\frac{d^i}{dy^i}\left(g^{(m)}\left(\sqrt{\frac{2\Delta y}{c}}\right)\right)y^{-1/2-(r-i)}\right|\\ &\ll_{u,n}\sum_{i=0}^n\left|\sum_{m_1,\dots,m_i}\left(g^{(m+m_1+\dots m_i)}\left(\sqrt{\frac{2\Delta y}{c}}\right)\right)\prod_{j=1}^i\left(\sqrt{\frac{2\Delta y}{c}}y^{-j}\right)^{m_j}\right|y^{-1/2-(n-i)}\\ &\ll_{u,n}C_{W_1,W_2,m+n}K^{-m}y^{-1/2-n}.\end{aligned}$$

Here the inner sum is over $m_1,\dots,m_i$ satisfying $m_1+2m_2+\dots+im_i=i$ and in the last line we have used (14) and that $\Delta y/c\asymp_u K^2$.

For $(m,m')=(4,1)$ in (15) we use the claim with $n=0$ and for $(m,m')=(0,0)$ we take a general n which will eventually depend on θ, and we find, by using integration by parts as described above,

$$\int_{-\infty}^{\infty} \hat{g}(t)e(\frac{\Delta}{c},t)dt \ll_u C_{W_1,W_2,4}\frac{K^{-3}\Delta^{1/2}}{c^{1/2}} + C_{W_1,W_2,n}\frac{K^{1-2n}\Delta^{n-1/2}}{c^{n-1/2}}.$$

Plugging this bound back in the sum (12) and using the restriction (ii)-(iv) gives the result by choosing n sufficiently large depending on θ. □

Remark 3. Note the resemblance between Theorem 2.1 and [22, Thm 1.3]. Whereas [22, Thm 1.3] is restricted to a range where the contribution of the individual off-diagonals are essentially trivial due to the decay of the J-Bessel function (corresponding to $\theta > 1/2$), we note that for $\theta \le 1/2$ we need to exploit additional cancellation between the J_{k-1}-Bessel functions for different k.

3. Computing the quantum variance

We now explain how the above results may be used to understand quantum variance for shrinking sets around infinity.

3.1. Squeezing sets towards cusps.

Let $M = \Gamma\backslash\mathbb{H}$ be a finite volume hyperbolic surface. Then M admits a decomposition

$$M = M_0 \cup Z_1 \cup \ldots \cup Z_l$$

where M_0 is compact and Z_i is isometric to

$$Z_i \simeq S^1\times]a_i,\infty[,$$

for some $a_i > 0$ with the metric on $S^1\times]a_i,\infty[$ equal to

$$ds^2 = \frac{dx^2+dy^2}{y^2}$$

for $(x,y) \in S^1\times]a_i,\infty[$. In the literature the regions Z_i are called horoball cusp neighboorhoods, horocusps, cuspidal zones, Siegel sets, horocyclic regions, or simply (by an abuse of notation) cusps. These subregions Z_i are unbounded regions whose boundary is a horocycle ($S^1 \times \{a_i\}$) and a point (the cusp).

Let $Z = Z_1$ which we may assume corresponds to a cusp at infinity. We now consider a measurable set $B \subseteq Z$ of volume $\mathrm{vol}(B) > 0$ and define, for every $H \geq \mathrm{vol}(B)^{-1}$ the injective map

$$\begin{aligned} S_H^B\colon B &\longrightarrow Z \\ x+iy &\longmapsto x + i\mathrm{vol}(B)\,Hy, \end{aligned}$$

pushing the region B up towards the cusp at infinity. We note that this may be formulated as a scaling along a geodesic going to the cusp thereby defining S_H^B in a coordinate free way. We let $B_H = S_H^B(B)$ and notice that by a simple change of variables

$$\begin{aligned}\mathrm{vol}\,(B_H) &= \int_{a_1}^{\infty}\int_0^1 1_B((S_H^B)^{-1}z)\frac{dxdy}{y^2}\\ &= \frac{1}{\mathrm{vol}\,(B)\,H}\int_0^{\infty}\int_{0,1} 1_B(z)\frac{dxdy}{y^2} = \frac{1}{H}.\end{aligned}$$

For $A \subseteq M$ we let

$$L^2(M,A) = \{f \in L^2(M) : \mathrm{supp}\, f \subseteq A\}$$

and define *the squeezing operator*

$$\begin{aligned} M_H^B \colon L^2(M,B) &\longrightarrow L^2(M,B_H)\\ f &\longmapsto f \circ (S_H^B)^{-1}\end{aligned}$$

i.e. $M_H^B f(z) = f(x + iy/(\mathrm{vol}\,(B)\,H))$. We note that M_H^B loosely speaking squeezes the function f into the region B_H which moves towards the cusp at infinity.

A simple change of variable computation – similar to the volume computation of $\mathrm{vol}\,(B)$ above – shows that for $\varphi \in L^2(M,B)$

$$\left\|M_H^B\varphi\right\|^2 = \frac{1}{\mathrm{vol}\,(B)\,H}\left\|\varphi\right\|^2, \quad \left\langle M_H^B\varphi, 1\right\rangle = \frac{1}{\mathrm{vol}\,(B)\,H}\left\langle \varphi, 1\right\rangle.$$

We now specialize to $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$ and and $Z = S^1\times]1,\infty[$. For $T > 1$ we let

$$B_T(\infty) = \{z \in Z : \Im(z) > T\},$$

which we consider to be a ball around the cusp at infinity. A trivial computation shows that $\mathrm{vol}\,(B_T(\infty)) = 1/T$. Fix now $T_0 > 1$ and let $B = B_{T_0}(\infty) \subseteq Z$. We observe that we have $B_H = S_H^B(B) = B_H(\infty)$. In particular – with this choice of $B = B_{T_0}(\infty)$ – the squeezed set B_H does not depend on T_0. Note, however, that the squeezing operator M_H^B still depends on the choice of T_0.

3.2. **Mass equidistribution in squeezed sets.** We now want to consider the notion of mass equidistribution in the context of the squeezed sets as above: Fix H, $\varphi \in L^2(M,B)$. It follows from the mass equidistribution theorem of Soundararajan and Holowinsky [7] that

$$\int_{B_H}(M_H^B\varphi)(z)y^k\left|f(z)\right|^2 d\mu(z) = \frac{1}{\mathrm{vol}\,(M)}\int_{B_H}(M_H^B\varphi)(z)d\mu(z) + o_{M_H^B\varphi}(1)$$

as $k \to \infty$.

We investigate what condition on H as a function of $k \to \infty$ implies that

$$\int_{B_H}(M_H^B\varphi)(z)y^k\left|f(z)\right|^2 d\mu(z) = \frac{1}{\mathrm{vol}\,(M)}\int_{B_H}(M_H^B\varphi)(z)d\mu(z) + o\left(\int_{B_H}\left|M_H^B\varphi\right|d\mu\right)$$

as $k, H \to \infty$.

Choosing $\varphi = 1_B$ this simplifies to the question of when

$$\int_{B_H} y^k\left|f(z)\right|^2 d\mu(z) = \frac{1}{H\mathrm{vol}\,(\Gamma\backslash\mathbb{H})} + o(H^{-1}),$$

as $H, k \to \infty$. However, we want to investigate also more general test functions φ.

For the rest of the paper we fix $B = B_1(\infty)$, and consider the above when $H = (k-1)^\theta$ for some $\theta > 0$, i.e. we consider

$$M_{(k-1)^\theta} := M^B_{(k-1)^\theta},$$

i.e. $M_{(k-1)^\theta} f(z) = f(x+iy/(k-1)^\theta)$. We want to investigate the mass equidistribution when the test function is squeezed via this operator. We will do this by considering the squeezing of the non-holomorphic Poincaré series

$$P_{V,h}(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} V(y(\gamma z)) e(hx(\gamma z)),$$

where $V : \mathbb{R}_+ \to \mathbb{C}$ is a smooth compactly supported function with support contained in $(1, \infty)$. In other words, we want to understand, as $k \to \infty$, the asymptotic properties of

$$\mu_f(M_{(k-1)^\theta} P_{V,h}) := \langle M_{(k-1)^\theta} P_{V,h}(z), y^k |f(z)|^2 \rangle.$$

We note that with our assumption on V the function $P_{V,h}$ is supported on B, and that these actually series span $L^2(M, B)$. In fact

$$P_{V,h}(z) = V(y) e(hx).$$

For $h_1 h_2 \neq 0$ we define

$$B_\theta(P_{V_1,h_1}, P_{V_2,h_2}) = \frac{\pi}{4} \tau_1((|h_1|, |h_2|)) \int_0^\infty V_1(\frac{y}{|h_1|}) \overline{V_2(\frac{y}{|h_2|})} f_{\theta,h_1,h_2}(y) \frac{dy}{y^2},$$

where

$$f_{\theta,h_1,h_2}(y) = \begin{cases} 1 & \text{if } 0 < \theta < 1/2, \\ e^{-2\pi^2 y^2 (h_1^2 + h_2^2)} & \text{if } \theta = 1/2, \\ 0 & \text{if } \theta > 1/2. \end{cases}$$

When $\theta = 0$ we define B_0 to be the form B_ω defined by Luo and Sarnak in [18, Eq (15)].

Theorem 3.1. *Let $u : \mathbb{R}_+ \to \mathbb{R}_{\geq 0}$ be a smooth compactly supported weight function, and let V_1, V_2 be as above. For $h_1 h_2 \neq 0$ and $0 \leq \theta < 1$, we have*

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f \in H_k} L(1, \operatorname{sym}^2 f) \mu_f(M_{(k-1)^\theta} P_{V_1,h_1}) \overline{\mu_f(M_{(k-1)^\theta} P_{V_2,h_2})}$$
$$= B_\theta(P_{V_1,h_1}, P_{V_2,h_2}) \int_0^\infty u(y) y^{-\theta} dy\, K^{1-\theta} + O_{\theta,\varepsilon,V_i,h_i}(K^{1-\theta-\delta_\theta+\varepsilon}),$$

where

$$\delta_\theta = \begin{cases} (1+\theta)/2 & \theta \in (0, 1/5), \\ 1 - 2\theta & \theta \in [1/5, 1/2), \\ 1/2 & \theta = 1/2, \\ 1 + 2\theta & \theta \in (1/2, 1). \end{cases}$$

The implied constant in the error term may be bounded by a constant depending only on θ, ϵ times

$$(1 + \|h\|_\infty)^C (A_{W_1} A_{W_2})^C \|W_1\|_{C,\infty} \|W_2\|_{C,\infty},$$

for C sufficiently large depending on θ.

Proof. We start by observing that $\mu_f(M_{(k-1)^\theta}P_{V,h})$ is a real number (as is seen by unfolding and using that $|f|^2$ is even). Therefore we have $\mu_f(M_{(k-1)^\theta}P_{V,h}) = \mu_f(M_{(k-1)^\theta}P_{V,-h})$. So we may assume $h_i > 0$ below. We also notice that if V is supported in $(1, A_V]$ then, up to an absolute constant times a power of h times a power of A_V, the functions $W_i(y) = V((4\pi y)^{-1})y^i$, $W^*(y) = V(4\pi y)^{-1})\exp(-h^2y^{-2}/8)$ all have Sobolev norms less than or equal to the corresponding Sobolev norm of V. This will be used below without mention.

The case $\theta = 0$ is [18, Thm 2]. To handle the other cases we proceed as in the proof of [17, Prop 2.1]. Doing this, noticing in the proof that the Mellin transform satisfies

$$\int_0^\infty V\left(\frac{y^{-1}}{(k-1)^\theta}\right) y^s \frac{dy}{y} = (k-1)^{-s\theta} \int_0^\infty V\left(y^{-1}\right) y^s \frac{dy}{y},$$

we find that

$$\begin{aligned}\mu_f(M_{(k-1)^\theta}P_{V,h}) =& \frac{2\pi^2}{(k-1)L(1,\mathrm{sym}^2 f)} \sum_{n\in\mathbb{N}} \lambda_f(n)\lambda_f(n+h) \\ & \times V\left(\frac{(k-1)^{1-\theta}}{4\pi(n+h/2)}\right)\left(\frac{\sqrt{n(n+h)}}{n+h/2}\right)^{k-1} \\ & + O_{V,h}(k^{-1-\theta+\varepsilon}),\end{aligned} \tag{16}$$

where the implied constant is less than a constant depending on θ, ε times

$$(1+h^B)A_V^B \left\|V\right\|_{B,\infty} \tag{17}$$

for B sufficiently large. This holds also for $h = 0$.

We now assume that $0 < \theta < 1/2$, and observe that in the above sum we may restrict to n such that $(n+h/2) \asymp k^{1-\theta}$, which implies that $(k-1)/(n+h/2)^2 = o(1)$ as $k \to \infty$. Therefore, we can employ the following Taylor expansion

$$\begin{aligned}\left(\frac{\sqrt{n(n+h)}}{n+h/2}\right)^{k-1} &= \exp\left(\frac{k-1}{2}\log\left(1-\frac{h^2}{(2n+h)^2}\right)\right) \\ &= \exp\left(-\frac{k-1}{2}\frac{h^2}{(2n+h)^2} + O\left(\frac{kh^4}{(2n+h)^4}\right)\right) \\ &= \exp\left(-\frac{k-1}{2}\frac{h^2}{(2n+h)^2}\right) + O\left(\frac{kh^4}{(2n+h)^4}\right) \\ &= \sum_{i=0}^{N-1} \frac{\left(-\frac{h^2}{2}\frac{k-1}{(2n+h)^2}\right)^i}{i!} + O_N\left(\frac{(kh^2)^N}{(2n+h)^{2N}} + \frac{kh^4}{(2n+h)^4}\right) \\ &= \sum_{i=0}^{N-1} \frac{\left(-\frac{h^2}{8}\frac{k-1}{(n+h/2)^2}\right)^i}{i!} + O_N(h^{2N}k^{2N(\theta-1/2)} + h^4k^{4\theta-3}).\end{aligned}$$

This gives us that

$$\begin{aligned}\mu_f(M_{(k-1)^\theta}P_{V,h}) =&\frac{2\pi^2}{(k-1)L(1,\mathrm{sym}^2 f)}\sum_{i=0}^{N-1}\frac{\left(-\frac{h^2}{8}(k-1)^{2\theta-1}\right)^i}{i!}A_f^{W_i}((k-1)^{1-\theta},h)\\ &+O_{V,h,N}((h^{2N}k^{2N(\theta-1/2)-\theta+\varepsilon}+h^4k^{3\theta-3+\varepsilon})+k^{-1-\theta+\varepsilon}),\end{aligned} \tag{18}$$

where $W_i(y)=V\left(\frac{1}{4\pi y}\right)y^{-2i}$ for $y\in\mathbb{R}_+$, and the implied constant is of the form (17). Since $\theta<1/2$, we can choose N large enough such that the dominating error-term in k is $k^{-1-\theta+\varepsilon}$.
We now plug (18) into the expression we want to evaluate. The terms involving the products of error terms is easily seen to be $O_{V,h,N}(K^{-2\theta+\varepsilon})$.

To bound the mixed terms we note that $(k-1)^{(2\theta-1)i}$ is largest when $i=0$, so it suffices to observe that

$$\begin{aligned}&K^{-1-\theta+\varepsilon}\sum_{2|k}u\left(\frac{k-1}{K}\right)\frac{1}{k-1}\sum_{f\in H_k}\left|A_f^{W_i}((k-1)^{1-\theta}),h\right|\\ &\ll_{V,h,N}K^{-1-\theta+\varepsilon}K^{1/2+\varepsilon}\left(\sum_{k\geq1,2|k}u\left(\frac{k-1}{K}\right)\sum_{f\in H_k}\frac{\left|A_f^{W_i}((k-1)^{1-\theta},h)\right|^2}{L(1,\mathrm{sym}^2 f)}\right)^{1/2}\\ &\ll_{V,h,N}K^{1/2-3\theta/2+\varepsilon},\end{aligned}$$

where we have used Cauchy–Schwarz, the positivity of $L(\mathrm{sym}^2 f,1)$, and Theorem 2.1. The implied constant is of the claimed form. This implies that

$$\begin{aligned}&\sum_{2|k}u\left(\frac{k-1}{K}\right)\sum_{f\in H_k}L(1,\mathrm{sym}^2 f)\mu_f(M_{(k-1)^\theta}P_{V_1,h_1})\overline{\mu_f(M_{(k-1)^\theta}P_{V_2,h_2})}\\ &\quad=\sum_{2|k}u\left(\frac{k-1}{K}\right)\frac{(2\pi^2)^2}{(k-1)^2}\sum_{0\leq i,j\leq N-1}\frac{h_1^{2i}h_2^{2j}}{i!j!}\left(-\frac{1}{8}(k-1)^{2\theta-1}\right)^{i+j}\\ &\qquad\times\sum_{f\in H_k}\frac{A_f^{W_{1,i}}(h_1,(k-1)^{1-\theta})A_f^{W_{2,j}}(h_2,(k-1)^{1-\theta})}{L(1,\mathrm{sym}^2 f)}\\ &\qquad+O_{V,h,N}(K^{1/2-3\theta/2+\varepsilon}),\end{aligned}$$

with an allowed implied constant. Now for each pair $i,j\in\{0,\dots,N-1\}$, we apply Theorem 2.1 with smooth weights $W_{1,i}$ $W_{2,j}$ and weight function

$$u_{ij}(y)=u(y)y^{(2\theta-1)(i+j)-1}.$$

This gives

$$\sum_{k\geq 1,2|k} u\left(\frac{k-1}{K}\right)\frac{(2\pi^2)^2}{(k-1)^2}\frac{h_1^{2i}h_2^{2j}}{i!j!}\left(-\frac{1}{8}(k-1)^{2\theta-1}\right)^{i+j}$$
$$\times\sum_{f\in H_k}\frac{A_f^{W_{1,i}}(h_1,(k-1)^{1-\theta})A_f^{W_{2,j}}(h_2,(k-1)^{1-\theta})}{L(1,\mathrm{sym}^2 f)}$$
$$=2\pi^2h_1^{2i}h_2^{2j}\left(\frac{-1}{8}\right)^{i+j}K^{2(i+j)(\theta-1/2)-1}\sum_{k\geq 1,2|k}u_{ij}\left(\frac{k-1}{K}\right)\frac{2\pi^2}{(k-1)}$$
$$\times\sum_{f\in H_k}\frac{A_f^{W_{1,i}}(h_1,(k-1)^{1-\theta})A_f^{W_{2,j}}(h_2,(k-1)^{1-\theta})}{L(1,\mathrm{sym}^2 f)}$$
$$=2\pi^2h_1^{2i}h_2^{2j}\left(\frac{-1}{8}\right)^{i+j}\left(\int_0^\infty u_{ij}(y)y^{1-\theta}dy\right)\tau_1((h_1,h_2))$$
$$\cdot B_{h_1,h_2}(W_{1,i},W_{2,j})K^{1-\theta-(i+j)(1-2\theta)}+O_{h_i,V_i,\theta,N,\varepsilon}\left(K^{-(i+j)(1-2\theta)+\varepsilon}\right),$$

with an implied constant of the desired form. For $(i,j)\neq(0,0)$ we see that the contribution is bounded by $O(K^\theta)$, and for $i=j=0$, we get the wanted main-term. So in this case we have an error of $O(K^{\max(\theta,1/2-3\theta/2)})$ which translates to the claimed δ_θ.

Now assume that $\theta=1/2$, which implies that $\frac{k-1}{(n+h/2)^2}\asymp 1$ for non-zero terms in the sum (16). Again by a Taylor expansion, we see that

$$\left(\frac{\sqrt{n(n+h)}}{n+h/2}\right)^{k-1}=\exp\left(-\frac{h^2}{8}\frac{k-1}{(n+h/2)^2}\right)+O\left(\frac{h^4}{k}\right),$$

which is the source of the different main term in this case. We proceed as above to write

$$\sum_{k\geq 1,2|k}u\left(\frac{k-1}{K}\right)\sum_{f\in H_k}L(1,\mathrm{sym}^2 f)\mu_f(M_{(k-1)^\theta}P_{V_1,h_1})\overline{\mu_f(M_{(k-1)^\theta}P_{V_2,h_2})}$$
$$=\sum_{k\geq 1,2|k}u\left(\frac{k-1}{K}\right)\frac{(2\pi^2)^2}{(k-1)^2}\sum_{f\in H_k}\frac{A_f^{W_1^*}(h_1,(k-1)^{1-\theta})A_f^{W_2^*}(h_2,(k-1)^{1-\theta})}{L(1,\mathrm{sym}^2 f)}$$
$$+O_{V_i,h_i}(K^{-1/4+\varepsilon}),$$

where $W_i^*(y)=V_i\left(\frac{1}{4\pi y}\right)\exp\left(-\frac{m^2}{8}/y^2\right)$ for $i=1,2$, and where the implied constant is of the desired form. Again by an application of Theorem 2.1, we get the desired main term with error-term $O_{V_i,h_i}(K^\varepsilon)$ and an implied constant of the desired form.

Finally when $\theta>1/2$, we see that for non-zero terms in the sum (16) we have $\frac{k-1}{(n+h/2)^2}\gg k^{2\theta-1}$, which implies

$$\left(\frac{\sqrt{n(n+h)}}{n+h/2}\right)^{k-1}\ll\exp\left(-ck^{2\theta-1}\right),$$

for some $c>0$ depending only on V, and hence we get exponential decay of the sum in (16). Therefore we can even get the desired bound without any averaging. By summing up we arrive at the error-term $O(\|V_1\|_\infty \|V_2\|_\infty K^{-2\theta+\varepsilon})$. □

Remark 4. The above theorem also holds, with the same proof, when we allow V_i to have support in $\mathbb{R}_+$, if we interpret $M_{(k-1)^\theta} P_{V,m}$ as the Poincaré series $P_{V_{k,\theta}}$ related to $V_{k,\theta} = V(y/(k-1)^\theta)$.

Remark 5. We now give a quick sketch of what happens in the case when $h_1 = 0$ and $\int_0^\infty V_1(y)y^{-2}dy = 0$ (i.e. in the case where $P_{V_1,0}$ is an incomplete Eisenstein series orthogonal to 1) and $h_2 \neq 0$. The translation to a shifted convolution sum as in (16) is still valid.

To analyze the resulting shifted convolution sum we imitate the proof of Theorem 2.1. In this case we use the Hecke relations (3) to write

$$A_f^{W_1}(0,X) = \sum_{d\in\mathbb{N}} \sum_{r\in\mathbb{N}} \lambda_f(r^2) W_1\left(\frac{dr}{X}\right).$$

Here $W_1(y) = V_1(1/(4\pi y)$ and $X = (k-1)^{1-\theta}$. We deal with the off-diagonal terms as above and the diagonal term from the Petersson formula becomes

$$\sum_{\substack{d_1\in\mathbb{N}, d_2|h_2\\ r_1,r_2\in\mathbb{N}}} \delta_{r_1^2 = r_2(r_2+d_2)} W_1\left(\frac{X}{d_1 r_1}\right) W_2\left(\frac{X}{\frac{h_2}{d_2}(r_2+d_2/2)}\right).$$

Now we observe that for fixed d_2 the equation $r_1^2 = r_2(r_2+d_2)$ has only finitely many solutions (r_1,r_2) and for any such solution, we have by Poisson summation

$$\begin{aligned}\sum_{d_1\in\mathbb{N}} W_1\left(\frac{X}{d_1 r_1}\right) &= \int_0^\infty W_1\left(\frac{X}{r_1 y}\right) dy + O_{h_2,A}(X^{-A}) \\ &= \frac{X}{4\pi r_1}\int_0^\infty V_1(y)\frac{dy}{y^2} + O_{h_2,A}(X^{-A}) = O_{h_2,A}(X^{-A}).\end{aligned}$$

Therefore Theorem 3.1 holds in this case with $B_\theta(P_{V_1,0}, P_{V_2,h_2}) = 0$.

Now if $h_i = 0$ and $\int_0^\infty V_i(y)y^{-2}dy = 0$, using a similar analysis (in this case we observe that since the factor $\left(\frac{\sqrt{n(n+h_i)}}{n+h_i/2}\right)^{k-1}$ equals 1 we do not have to distinguish between various regimes of θ) we find that

$$\begin{aligned}&\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f\in H_k} L(1,\mathrm{sym}^2 f)\mu_f(M_{(k-1)^\theta} P_{V_1,0})\overline{\mu_f(M_{(k-1)^\theta} P_{V_2,0})} \\ &= 2\pi^2 \sum_{2|k} \frac{u\left(\frac{k-1}{K}\right)}{k-1} \sum_{r,d_1,d_2\in\mathbb{N}} V_1\left(\frac{(k-1)^{1-\theta}}{4\pi r d_1}\right) \overline{V_2\left(\frac{(k-1)^{1-\theta}}{4\pi r d_2}\right)} \\ &\quad + O(\max(K^{1/2-3\theta/2+\varepsilon}, 1)).\end{aligned} \tag{19}$$

Analogous to [18, p. 781] we find that by using successive Euler–Maclaurin summation on the d_i sums ([11, Eq. (4.20)]), followed by Poisson summation on the r

sum and on the k sum we have that also in this case Theorem 3.1 holds with

$$B_\theta(P_{V_1,0}, P_{V_2,0}) = \frac{\pi}{4}\int_0^\infty \int_0^\infty \int_0^\infty b_2(y_1)b_2(y_2)\tilde{V}_1\left(\frac{t}{y_1}\right)\tilde{V}_2\left(\frac{t}{y_2}\right)\frac{dy_1}{y_1^2}\frac{dy_2}{y_2^2}\frac{dt}{t^2},$$

with error term $O(\max K^{1/2-3\theta/2+\varepsilon}, 1)$, and the same type of implied constant. Here $\tilde{V}_i(y) = (V_i'(y)y^2)'$ and $2b_2(y) = B_2(y - \lfloor y \rfloor)$, where $B_2(y) = y^2 - y + 1/6$ is the second order Bernoulli polynomial. Note that the y_i integrals vanishes for t sufficiently small so the t-integral converges (although not absolutely).

4. Extension of B_θ and quantum variance for more general observables

Let

$$C_0^\infty(M,B) := \left\{ \psi : M \to \mathbb{C} \text{ smooth} \;\middle|\; \begin{smallmatrix} \text{supp}\,\psi \subset B \\ \psi \text{ decays rapidly at } \infty \\ \int_0^1 \psi(z)dx = 0 \text{ for } y \text{ large enough} \end{smallmatrix} \right\},$$

where $B = \{x + iy \in M \mid y > 1\} \subset X$ is the standard horocyclic region. In this section we will extend the above variance results to the space

$$C_{0,0}^\infty(M,B) = \{\psi \in C_0^\infty(M,B) : \langle \psi, 1\rangle = 0\}.$$

For $\psi \in C_{0,0}^\infty(M,B)$ we let V_m^ψ be its mth Fourier coefficient. Note that, since ψ is supported in B, the coefficient $V_m^\psi(y)$ are supported in $y > 1$ and we have

$$\psi(z) = \sum_{m\in\mathbb{Z}} V_m^\psi(y)e(mx) = \sum_{m\in\mathbb{Z}} P_{V_m^\psi,m}(z), \tag{20}$$

where V_0^ψ has compact support, and satisfies $\int_0^\infty V_0^\psi(y)y^{-2}\,dy = 0$. Inspired by Theorem 3.1 and Remark 5 we define, for $\psi_1, \psi_2 \in C_{0,0}^\infty(M,B)$, the Hermitian form

$$\begin{aligned} B_\theta(\psi_1,\psi_2) = {} & \frac{\pi}{4}\sum_{m,n\neq 0} \tau_1((|m|, |n|)) \int_0^\infty V_m^{\psi_1}\left(\frac{y}{|m|}\right)\overline{V_n^{\psi_2}\left(\frac{y}{|n|}\right)} f_{\theta,m,n}(y)\frac{dy}{y^2} \\ & + \frac{\pi}{4}\int_0^\infty\int_0^\infty\int_0^\infty b_2(y_1)b_2(y_2)\widetilde{V_0^{\psi_1}}\left(\frac{t}{y_1}\right)\overline{\widetilde{V_0^{\psi_2}}\left(\frac{t}{y_2}\right)}\frac{dy_1}{y_1^2}\frac{dy_2}{y_2^2}\frac{dt}{t^2}. \end{aligned} \tag{21}$$

Note that if ψ_1, ψ_2 consist of a single Fourier coefficient, and if this coefficient is not just of rapid decay but of compact support then (21) agrees with the result of Theorem 3.1 and Remark 5. To see that $B_\theta(\psi_1, \psi_2)$ is well defined we argue as follows. By smoothness and rapid decay of ψ we see, using integration by parts, that

$$V_m^\psi(y) \ll_{A,B,\psi} y^{-A}m^{-B},$$

for any $A, B \geq 0$. It follows that

$$\int_0^\infty V_m^{\psi_1}\left(\frac{y}{|m|}\right)\overline{V_n^{\psi_2}\left(\frac{y}{|n|}\right)}\frac{dy}{y^2} \ll_A (|mn|)^{-A},$$

and so the first sum in (21) converges absolutely. The second term in (21) is well-defined by the discussion in Remark 5.

We observe that, when restricted to incomplete Eisenstein series, the form B_θ is independent of $0 < \theta < 1$, while for cuspidal test functions B_θ exhibits a phase transition at $\theta = 1/2$ as claimed in Theorem 1.3 (ii).

We can now show that the variance result of Theorem 3.1 can be extended to the space $C^\infty_{0,0}(M,B)$.

Theorem 4.1. *Let $u:\mathbb{R}_+\to\mathbb{R}_{\geq 0}$ be a smooth compactly supported weight function, and let $\psi\in C^\infty_{0,0}(M,B)$ and $0<\theta<1$. Then we have*

$$\begin{aligned}\sum_{2|k} u\left(\frac{k-1}{K}\right)\sum_{f\in H_k} L(1,\operatorname{sym}^2 f)\left|\mu_f(M_{(k-1)^\theta}\psi)-\nu(M_{(k-1)^\theta}\psi)\right|^2 \\ = B_\theta(\psi,\psi)\left(\int_0^\infty u(y)y^{-\theta}dy\right)K^{1-\theta}+O_{\psi,\theta}(K^{1-\theta-\delta_\theta}),\end{aligned} \tag{22}$$

for $\delta_\theta>0$ as in Theorem 3.1.

Proof. Consider a partition of unity

$$\sum_{l\geq 0} u_l(y)=1_{\geq 1}(y)=\begin{cases}1, & y>1\\ 0, & y<1\end{cases},$$

where $u_l:\mathbb{R}_+\to[0,1]$ with $\operatorname{supp} u_l\subset(3^l,2\cdot 3^{l+1})$, u_l smooth for $l>0$ and u_0 smooth on $(1,\infty)$ and $u_l^{(n)}(y)\ll_n y^{a_n}$ for some $a_n>0$ independently of l. Multiplying this partition of unity on ψ as in (20) we find

$$\psi(z)=V_0^\psi(y)+\sum_{l\geq 0,m\neq 0} V^\psi_{l,m}(y)e(mx),$$

where $V^\psi_{l,m}(y)=u_l(y)V^\psi_m(y)$ and $V^\psi_0(y)$ are smooth with compact support. We have

$$V^{\psi\;(n)}_{l,m}(y)\ll_C y^{-C}\left|m\right|^{-C}, \tag{23}$$

for any $C>0$ and independent of l. To see this we note that by the definition and partial integration

$$V^{\psi\;(n)}_{l,m}(y)=\frac{1}{(2\pi i m)^C}\sum_{j=0}^n\binom{n}{j}u_l^{(n-j)}(y)\int_0^1\left(\frac{\partial^{C+j}}{\partial x^C\partial y^j}\psi\right)(z)e(-mx)dx.$$

Now by using the rapid decay of ψ and the bound of the derivatives of u_l, we arrive at (23). This implies in particular that for every $C\geq 0$ we have $\left\|V^\psi_{l,m}\right\|_{C,\infty}\ll_{C,\psi} 3^{-Cl}\cdot\left|m\right|^{-C}$.

This implies, using Theorem 3.1, that for m_1,m_2 and $l_1,l_2\geq 0$

$$\begin{aligned}\sum_{2|k} u\left(\frac{k-1}{K}\right)\sum_{f\in H_k} L(1,\operatorname{sym}^2 f)\mu_f(M_{(k-1)^\theta}P_{V_{l_1,m_1},m_1})\mu_f(M_{(k-1)^\theta}P_{V_{l_2,m_2},m_2}) \\ = B_\theta(P_{V_{l_1,m_1},m_1},P_{V_{l_2,m_2},m_2})\left(\int_0^\infty u(y)y^{-\theta}dy\right)K^{1-\theta} \\ +O_{\psi,\theta}\left(\frac{K^{1-\theta-\delta_\theta}}{3^{l_1+l_2}((1+|m_1|)(1+|m_2|))^2}\right).\end{aligned}$$

Therefore by summing up all the contributions we get

$$\begin{aligned}
&\sum_{k,2|k} u\left(\frac{k-1}{K}\right) \sum_{f\in H_k} L(1,\mathrm{sym}^2 f)|\mu_f(M_{(k-1)^\theta}\psi)|^2 \\
&= \left(\sum_{m_1,m_2,l_1,l_2} B_\theta(P_{h_{l_1,m_1},m_1}, P_{h_{l_2,m_2},m_2})\right)\left(\int_0^\infty u(y)y^{-\theta}dy\right)K^{1-\theta} \\
&\quad + O_{\psi,\theta}\left(K^{1-\theta-\delta_\theta}\left(\sum_{\substack{l_1,l_2>0\\ m_1,m_2}} \frac{3^{-l_1-l_2}}{((1+|m_1|)(1+|m_2|))^2}\right)\right) \\
&= B_\theta(\psi,\psi)\left(\int_0^\infty u(y)y^{-\theta}dy\right)K^{1-\theta} + O_{\psi,\theta}(K^{1-\theta-\delta_\theta}),
\end{aligned}$$

which finishes the proof. □

5. Small scale quantum ergodicity around infinity

In this section we show that if we average over $f \in H_k$ and over the weight k quantum ergodicity holds for appropriately chosen sets shrinking towards infinity all the way down to the Planck scale.

Theorem 5.1. *Let $0 < \theta < 1$ and $\psi \in C_0^\infty(M,B)$. Then*

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f\in H_k} L(1,\mathrm{sym}^2 f)\left|\mu_f(M_{(k-1)^\theta}\psi) - \nu(M_{(k-1)^\theta}\psi)\right|^2 = O_{\psi,u}(K^{\max(2-2\theta-1/5,1-\theta)}).$$

Proof. Note that $\psi \in C_0^\infty(M,B)$ can be written as $\psi = \psi_0 + \psi_1$ where $\psi_1 \in C_{0,0}^\infty(M,B)$ and $\psi_0 = P_{V,0}$ is an incomplete Eisenstein series with V supported in $(1,\infty)$. Since trivially

$$\left|\mu_f(M_{(k-1)^\theta}\psi) - \nu(M_{(k-1)^\theta}\psi)\right|^2 \leq 2\sum_{i=1,2}\left|\mu_f(M_{(k-1)^\theta}\psi_i) - \nu(M_{(k-1)^\theta}\psi_i)\right|^2,$$

we may use Theorem 4.1 to see that we only need to prove Theorem 5.1 in the case where $\psi = P_{V,0}$, which we assume for the rest of the proof. In order to do so we open up the square and compute asymptotics with error terms for each of the averages over each of the terms $\left|\mu_f(M_{(k-1)^\theta}\psi)\right|^2$, $\left|\nu(M_{(k-1)^\theta}\psi)\right|^2$, $\mu_f(M_{(k-1)^\theta}\psi)\overline{\nu(M_{(k-1)^\theta}\psi)}$ and its conjugate. Since $\nu(M_{(k-1)^\theta}\psi) = (k-1)^{-\theta}\nu(\psi)$ we see that this essentially corresponds to computing the second, zero-th, and first moment of $\mu_f(M_{(k-1)^\theta}\psi)$.

We start by showing that

$$\sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{f\in H_k} L(1,\mathrm{sym}^2 f)\left|\mu_f(M_{(k-1)^\theta}\psi)\right|^2 \tag{24}$$

$$= |\nu(\psi)|^2 \frac{\zeta(2)^2}{12}\int_0^\infty u(y)y^{1-2\theta}dy\frac{K^{2-2\theta}}{2} + O_{\psi,u}(K^{1-\theta}). \tag{25}$$

To prove this we start as in Remark 5 and arrive at (19). We then evaluate the sum over $d = d_i$ using the second order Euler–Maclaurin formula and find that we have for any $X > 0$

$$\sum_d V\left(\frac{X}{rd}\right) = \int_0^\infty V\left(\frac{X}{ry}\right) dy - \int_0^\infty b_2(y)\tilde{V}\left(\frac{X}{ry}\right)\frac{dy}{y^2} \tag{26}$$
$$= \frac{X}{r}\int_0^\infty V(y)\frac{dy}{y^2} - \int_0^\infty b_2(y)\tilde{V}\left(\frac{X}{ry}\right)\frac{dy}{y^2},$$

where $2b_2(y) = B_2(y-\lfloor y \rfloor)$ and $B_2(y) = y^2-y+1/6$ is the second Bernoulli polynomial and $\tilde{V}(y) = (V'(y)y^2)'$. Here we have used that $\frac{\partial^2}{\partial y^2}V\left(\frac{X}{ry}\right) = \tilde{V}\left(\frac{X}{ry}\right)y^{-2}$. We know by the assumptions on V that the above defines a smooth function in r and that $\sum_d V\left(\frac{X}{dr}\right)$ vanishes for $r > AX$.

We can now evaluate

$$\sum_{r,d_1,d_2\in\mathbb{N}} V\left(\frac{X}{rd_1}\right)\overline{V\left(\frac{X}{rd_2}\right)}$$

by inserting (26) and evaluating the four terms coming from opening the square. The contribution coming from the absolute square of the first term on the right of (26) equals

$$X^2\left|\int_0^\infty V(y)\frac{dy}{y^2}\right|^2\sum_{1\le r\le AX}\frac{1}{r^2} = X^2\left|\int_0^\infty V(y)\frac{dy}{y^2}\right|^2\zeta(2) + O(X).$$

A change of variables combined with the fact that $b_2(v)$ is uniformly bounded shows that $\int_0^\infty b_2(y)\tilde{V}\left(\frac{X}{ry}\right)\frac{dy}{y^2} \ll_V r/X$. This implies that the remaining contributions are $O(X)$. Plugging these estimates back in (19) with $X = (k-1)^{1-\theta}/4\pi$ and using Poisson summation in the k variable proves (25).

We next show that

$$\sum_{2|k} u\left(\frac{k-1}{K}\right)\sum_{f\in H_k} L(1,\mathrm{sym}^2 f) = \frac{\zeta(2)^2}{12}\frac{K^2}{2}\int_0^\infty u(y)y\,dy + O(K^{2-\frac{1}{5}+\varepsilon}). \tag{27}$$

To approximate $L(1,\mathrm{sym}^2 f)$ we use $e^{-x} = \frac{1}{2\pi i}\int_{(\sigma)}\Gamma(s)x^{-s}ds$ to see that

$$\sum_{n=1}^\infty \frac{\lambda_f(n^2)}{n}e^{-n/T} = \frac{1}{2\pi i}\int_{(2)}\Gamma(s)\frac{L(s+1,\mathrm{sym}^2)}{\zeta(2(s+1))}T^s ds.$$

Here $T \ge 1$ is a parameter which will be chosen later. For now we assume that $T = K^a$ with $1 < a < 2$. Moving the line of integration to $\sigma = -1/2$ we pick up a pole of the Gamma function at $s = 0$ and we find that

$$\sum_{n=1}^\infty \frac{\lambda_f(n^2)}{n}e^{-n/T} = \frac{L(1,\mathrm{sym}^2)}{\zeta(2)} + I_f(T), \tag{28}$$

where $I_f(T) = \frac{1}{2\pi i}\int_{(-1/2)}\Gamma(s)\frac{L(s+1,\mathrm{sym}^2)}{\zeta(2(s+1))}T^s ds$. Using any bound of the form

$$L(s,\mathrm{sym}^2 f) \ll_A (1+|s|)^A(k^2)^{1/4-\rho}$$

for $\Re(s)=1/2$ we see, that $I_f(T) \ll_A T^{-1/2}k^{1/2-2\rho+\varepsilon}$. In fact the convexity estimate $\rho=0$ will suffice for what we need. We have

$$\sum_{n=1}^{\infty} \frac{\lambda_f(n^2)}{n} e^{-n/T} = \sum_{n \leq T^{1+\epsilon}} \frac{\lambda_f(n^2)}{n} e^{-n/T} + O_A(K^{-A}),$$

for any $A>0$. We observe also that since $\lambda_f(n^2) \ll n^{\varepsilon}$ we have $\sum_{n\in\mathbb{N}} \frac{\lambda_f(n^2)}{n} e^{-n/T} \ll T^{\varepsilon}$. Using these observations we see that

$$\begin{aligned}\sum_{2|k} u\left(\frac{k-1}{K}\right) &\sum_{f\in H_k} L(1,\operatorname{sym}^2 f) = \\ &= \zeta(2)^2 \sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{n_1,n_2 \leq T^{1+\varepsilon}} \frac{e^{-(n_1+n_2)/T}}{n_1 n_2} \sum_{f\in H_k} \frac{\lambda_f(n_1^2)\lambda_f(n_2^2)}{L(1,\operatorname{sym}^2 f)} \\ &\quad + O\left(\sum_k u\left(\frac{k-1}{K}\right) \sum_{f\in H_k} K^{\epsilon}(|I_f(T)| + |I_f(T)|^2) + K^{-A}\right).\end{aligned}$$

By using convexity ($\rho=0$) to bound $I_f(T)$ the error is $O\left(K^{2+\varepsilon}\left(\left(\frac{K}{T}\right)^{1/2} + \frac{K}{T}\right)\right)$. Up to this error term the sum we want to estimate therefore equals

$$\frac{\zeta(2)^2 K}{2\pi^2} \sum_{2|k} \tilde{u}\left(\frac{k-1}{K}\right) \sum_{n_1,n_2 \leq T^{1+\varepsilon}} \frac{e^{-(n_1+n_2)/T}}{n_1 n_2} \frac{2\pi^2}{(k-1)} \sum_{f\in H_k} \frac{\lambda_f(n_1^2)\lambda_f(n_2^2)}{L(1,\operatorname{sym}^2 f)},$$

where $\tilde{u}(y) = u(y)y$. We now use the Petersson formula (6) on the last sum. The diagonal term gives the claimed main term

$$\frac{\zeta(2)^2 K}{2\pi^2} \sum_{2|k} \tilde{u}\left(\frac{k-1}{K}\right) \sum_{n_1 \leq T^{1+\varepsilon}} \frac{e^{-2n_1/T}}{n_1^2} = \frac{\zeta(2)^3}{2\pi^2} \frac{K^2}{2} \int_0^{\infty} u(y) y dy + O(K^2/T).$$

We also need to bound the non-diagonal contribution which is done as in the proof of Theorem 2.1. This consists of a k sum with k supported around K, sums over $n_1, n_2 \leq T^{1+\varepsilon}$, and a c-sum. The c-sum can be truncated at $c \leq M$ at the expense of an error which is big O of

$$K \sum_{k\asymp K} \sum_{n_1,n_2 \leq T^{1+\varepsilon}} \frac{1}{n_1 n_2} \sum_{c>M} \left(\frac{e\Delta}{2kc}\right)^{k-1} \ll K \sum_{k\asymp K} T^{\epsilon} \left(\frac{e 4\pi T^{2+2\varepsilon}}{2kM}\right)^{k-1} \frac{M}{K},$$

where $\Delta = 4\pi n_1 n_2 \leq 4\pi T^{2+2\varepsilon}$ and we have used (7) on the Bessel function. If we choose $M = CT^{2+2\varepsilon}K^{-1+\epsilon}$ for a suitably big constant C the parentesis is $\ll K^{-\varepsilon(k-1)}$ which decays exponentially so this contribution is $O_A(K^{-A})$ for every positive A.

By using (10), as in the proof of Theorem 2.1, we see that it suffices to bound

$$K \sum_{n_1,n_2 \leq T^{1+\varepsilon}} \frac{1}{n_1 n_2} \sum_{c \leq M} \left| \int_{-\infty}^{\infty} \hat{g}(t) \sin\left(\frac{\Delta}{c}\cos(2\pi t)\right) dt \right|$$

with $g(y) = \tilde{u}(y/K)$. Here it is clear that g is supported in $y \asymp K$ and $g^m(y) \ll K^{-m}$ and we conclude as before that

$$\int_{-\infty}^{\infty} |\hat{g}(t)t^m|\, dt \ll K^{-m}.$$

We use (13) with $N = 1$ and we estimate the contribution form the error terms by

$$K \sum_{n_1,n_2 \le T^{1+\varepsilon}} \frac{1}{n_1 n_2} \sum_{c \le M} \int_{-\infty}^{\infty} |\hat{g}(t)| \left(\frac{\Delta}{c}\right)^{\alpha} |t|^{\beta}\, dt \ll K^{1-\beta} \left(\sum_{n_1 \le T^{1+\varepsilon}} n^{\alpha-1} \right)^2 \sum_{c \le M} c^{-\alpha}.$$

For the four contributions $(\alpha, \beta) = (2,8), (4,16), (0,2), (0,4)$ this gives an error term of $(K((T/K^2)^4 + (T/K^2)^8) + T^2/K^2 + T^2/K^4)K^{\varepsilon}$ which are all less that $(K + T^2/K^2)K^{\varepsilon}$ To bound the last term we see as before

$$\frac{d^n}{dy^n}\left(g^{(m)}\left(\sqrt{\frac{2\Delta}{c} y} \right) y^{-1/2} \right) \ll_{u,n,m} K^{-m} y^{-1/2-n}, \tag{29}$$

so again we find

$$\int_{-\infty}^{\infty} \hat{g}(t) e(\frac{\Delta}{c}, t) dt \ll_u \frac{K^{-3}\Delta^{1/2}}{c^{1/2}} + \frac{K^{1-2n}\Delta^{n-1/2}}{c^{n-1/2}}. \tag{30}$$

It turns out to be convenient to interpolate and use $n = 3/2$ such that the last contribution is

$$K \sum_{n_1,n_2 \le T^{1+\varepsilon}} \frac{1}{n_1 n_2} \sum_{c \le M} \int_{-\infty}^{\infty} \hat{g}(t) e(\frac{\Delta}{c}, t) dt \ll T^2 K^{-5/2+\varepsilon} + K^{-1+\varepsilon} T^2.$$

The total error therefore become $\ll K^{2+\varepsilon} \left(\frac{K}{T}\right)^{1/2} + K^{1+\varepsilon} + T^2 K^{-1+\varepsilon}$, as all other contributions are smaller. Choosing $T = K^{7/5}$ proves (27).

Lastly we use a similar strategy to prove that

$$\begin{aligned} (31)\ &\sum_{2|k} u(\frac{k-1}{K}) \sum_{f \in H_k} L(1, \mathrm{sym}^2 f) \mu_f(M_{(k-1)^{\theta}} P_{V,0}) \\ &\quad = \nu(P_{V,0}) \int_0^{\infty} u(y) y^{1-\theta} dy \frac{\zeta(2)^2}{12} \frac{K^{2-\theta}}{2} + O(K^{2-\theta-(1/4+3\theta/8)+\varepsilon} + K^{1+\varepsilon}). \end{aligned}$$

We use (28) to approximate $L(1, \mathrm{sym}^2 f)$ by $\zeta(2) \sum_{n \le T^{1+\varepsilon}} \frac{\lambda_f(n^2)}{n} e^{-n/T}$ at the cost of an error satisfying $\ll K^{2-\theta+\varepsilon} \frac{K^{1/2-2\rho}}{T^{1/2}}$. We then use (16) and the Hecke relations (3) to arrive at

$$\zeta(2) \sum_{2|k} u\left(\frac{k-1}{K}\right) \sum_{\substack{n_1 \le T^{1+\varepsilon} \\ n_2 \in \mathbb{N} \\ d | n_2}} \frac{e^{-n_1/T}}{n_1} V\left(\frac{(k-1)^{1-\theta}}{4\pi n_2} \right) \frac{2\pi^2}{(k-1)} \sum_{f \in H_k} \frac{\lambda_f(n_1^2)\lambda_f(d^2)}{L(1, \mathrm{sym}^2(f))},$$

at the expense of an additional error which is $\ll K^{1-\theta+\varepsilon}$. We then use the Petersson formula (6). The diagonal gives

$$\zeta(2)\sum_{2|k} u\left(\frac{k-1}{K}\right)\sum_{n_1\leq T^{1+\varepsilon}}\frac{e^{-n_1/T}}{n_1}\sum_{r=1}^{\infty}V\left(\frac{(k-1)^{1-\theta}}{4\pi r n_1}\right),$$

which, after using Poisson in the r variable, a change of variables, and then Poisson again in the k variable, gives the claimed main term up to an error which is $\ll K^2/T + K^{1+\varepsilon}$. The off diagonal is handled as before: We truncate the c-sum at $M = CK^{-\theta+\varepsilon}T^{1+\varepsilon}$ with C sufficiently large at the expense $\ll K^{-A}$. We then use (10) with $g(y) = u(y/K)v\left(\frac{y^{1-\theta}}{4\pi n_2}\right)$ and find that in the support of the sums $g^{(m)}(y) \ll K^{-m}$ which allows us to bound the error coming from the approximation (with $N = 1$) of $\sin(\frac{\Delta}{c}\cos(2\pi t))$ with $\Delta = 4\pi n_1 n_2$ by big O of

$$T^2K^{(1-\theta)3-8+\varepsilon} + T^4K^{(1-\theta)5-16+\varepsilon} + TK^{-1-2\theta+\varepsilon} + TK^{-3-2\theta+\varepsilon}.$$

We also find, using Faà di Bruno's formula as before that (29) and (30) holds. Using (30) with $n = 2$ we get the final error contribution to be bounded by $TK^{-3/2-2\theta+\varepsilon} + T^{3/2}K^{-1/2-5\theta/2}$. Balancing $T^{3/2}K^{-1/2-5\theta/2} = K^{2-\theta}\frac{K^{1/2}}{T^{1/2}}$ gives $T = K^{3/2-3\theta/4}$. This proves (31) as with this choice of T all error contributions are less than the claimed one.

We can now finish the proof: We open up the square of the expression on the right-hand side of the theorem and use the expressions in (24), (27), and (31). The main terms cancel and we are left with the claimed error term.

□

Remark 6. It is obvious from the above proof that a subconvexity result in the k-aspect for $L(s, \mathrm{sym}^2 f)$ would give an improvement of the exponent. In fact a non-trivial bound on the second moment of $L(s, \mathrm{sym}^2 f)$ in the weight aspect would suffice.

Theorem 5.1 shows that if $0 < \theta < 1$ then mostly (i.e. in a full-density set of $f \in H_k$) we have

$$\mu_f(M_{(k-1)^\theta}\psi) = \nu(M_{(k-1)^\theta}\psi) + o(k^{-\theta}).$$

If we go below the Planck scale, i.e. if we let $\theta \geq 1$, then this is *not* the case i.e. mass equidistribution fails.

Proposition 5.2. *Let $\theta \geq 1$ and let $V : \mathbb{R}_+ \to \mathbb{R}$ be a smooth function with compact support in $(1, \infty)$, which satisfies $\int_0^\infty V(y)dy/y^2 \neq 0$ and let ψ_V be the associated incomplete Eisenstein series. Then*

$$\mu_f(M_{(k-1)^\theta}\psi_V) = o(\nu(M_{(k-1)^\theta}\psi_V)),$$

as $k \to \infty$. This means in particular that mass equidistribution fails for shrinking annuli around infinity below the Planck scale i.e. when $\theta \geq 1$.

Proof. We use (16) and observe that the sum is identically zero since $(k-1)^{1-\theta}/(4\pi n)$ is less than one which is outside the support of V. Therefore

$$\mu_f(M_{(k-1)^\theta}\psi_V) = O_\varepsilon(k^{-1-\theta+\varepsilon}),$$

and since $\nu(M_{k^\theta}\psi_V) \asymp k^{-\theta}$ the proposition follows. $\square$

6. Further extensions of B_θ and computations at truncated eigenfunctions.

Before we extend B_θ we notice that on the set $C^\infty_{0,0}(M,B)$, B_θ is symmetric with respect to the Laplacian.

Lemma 6.1. *The map $B_\infty : C^\infty_{0,0}(M,B) \times C^\infty_{0,0}(M,B) \to \mathbb{C}$ satisfies $B_\infty(\Delta\psi,\varphi) = B_\infty(\psi,\Delta\varphi)$.*

Proof. Writing ψ as in (20) we note that $\Delta P_{V^\psi_m,m}(z) = P_{L_m V^\psi_m,m}(z)$ where $L_m = y^2\frac{d^2}{dy^2} - 4\pi^2 m^2 y^2$ and that the support of $L_m V^\psi_m$ is contained in $(1,\infty]$ if this is the case for V^ψ_m. The argument is now a straightforward modification of [18, p. 782]. $\square$

We now extend $B_\theta(\psi_1,\psi_2)$ defined in (21) on $C^\infty_{0,0}(M,B)$ to the slightly larger space $1_B C^\infty_{0,0}(M)$. This space includes for instance $1_B\cdot\phi$ where ϕ is a Hecke–Maass form, which together with the incomplete Eisenstein series of mean 0 actually span this entire space. We may define $B_\theta(\psi_1,\psi_2)$ on this slightly larger space by the same formula (21). The same arguments as after (21) shows that the infinite sum converges.

Unfortunately we do not know how to extend Lemma 6.1 to this larger space. When trying to do the obvious generalization we are faced with certain boundary terms which we cannot dismiss. This means also that, contrary to the situation when $\theta = 0$ studied by Luo and Sarnak [18], we do not know if truncated Hecke–Maass forms diagonalize B_θ for $\theta > 0$.

On the subspaces $C^\infty_{\mathrm{cusp}}(M,B) \subset C^\infty_{0,0}(M,B)$ and $1_B C^\infty_{\mathrm{cusp}}(M,B) \subset 1_B C^\infty_{0,0}(M)$ consisting of functions where the zero-th Fourier coefficient vanishes completely we can make the following analysis. It is straightforward to check that the Sobolev norm on $1_B C^\infty_{\mathrm{cusp}}(M)$ defined by

$$\|1_B\psi\|^2_{2,N} = \sum_{j\le N}\left\|1_B\frac{d^j\psi}{dx^j}\right\|^2_{L^2(M)} \tag{32}$$

is indeed a norm. Note that for $1_B\psi \in C^\infty_{\mathrm{cusp}}(M)$ we may write

$$\psi = \sum_{n\neq 0} V^{(\psi)}_m(y)e(nx),$$

and we have

$$\left\|1_B\frac{d^j\psi}{dx^j}\right\|^2_{L^2(M)} = \sum_{n\neq 0}|2\pi n|^{2j}\int_0^\infty \left|1_{y\ge 1}V^{(\psi)}_m\right|^2\frac{dy}{y^2}.$$

Proposition 6.2. *(1) The set $C^\infty_{\mathrm{cusp}}(M,B)$ is dense in $1_B C^\infty_{\mathrm{cusp}}(M)$ with respect to $\|\cdot\|_{2,N}$.*

(2) There exist an absolute constant $c > 0$ such that for $1_B\psi_i \in 1_B C^\infty_{\mathrm{cusp}}(M)$

$$|B_\theta(1_B\psi_1, 1_B\psi_2)| \le c\,\|1_B\psi_1\|_{2,1}\,\|1_B\psi_2\|_{2,1}\,.$$

(3) The form $B_\theta(\cdot,\cdot)$ is continuous on $1_B C^\infty_{\mathrm{cusp}}(M)\times 1_B C^\infty_{\mathrm{cusp}}(M)$ with respect to $\|\cdot\|_{2,1}$.

Proof. To see that $C^\infty_{\text{cusp}}(M,B)$ is dense in $1_B C^\infty_{\text{cusp}}(M)$ we approximate 1_B by a smooth cut-off as follows: Fix $w:\mathbb{R}\to\mathbb{R}_{\geq 0}$ smooth and supported in $[1/2,1]$ with $\int_0^1 w(t)dt=1$. For $0<\delta<1/2$ we define $w_\delta(t):=\delta^{-1}w(t/\delta)$. This is supported in $[\delta/2,\delta]$ and satisfies $\int_0^1 w_\delta(t)dt=1$. We then define the function $1_B^\delta:\mathbb{R}_+\to\mathbb{R}$ as the convolution of $1_{y>1}$ and w_δ i.e.

$$1_B^\delta(y):=\int_0^\infty 1_{t>1}(t)w_\delta(y-t)dt.$$

We observe that 1_B^δ is smooth and supported in $[1+\delta/2,\infty]$. It satisfies $0\leq 1_B^\delta(y)\leq 1$ and $1_B^\delta(y)=1$ for $y\geq 1+\delta$.

Let now $1_B\psi\in 1_B C^\infty_{\text{cusp}}(M)$, and observe that $1_B^\delta\psi\in C^\infty_{\text{cusp}}(M,B)$, where we use the same notation for $y\mapsto 1_B^\delta(y)$ and $x+iy\mapsto 1_B^\delta(y)$. Furthermore

$$\begin{aligned}\left\|1_B\psi-1_B^\delta\psi\right\|_{2,N}&=\left\|1_B(\psi-1_B^\delta\psi)\right\|_{2,N}\\ &\leq \max_{\substack{1\leq\Im(z)\leq 2\\ j\leq N}}\left|\frac{d^j\psi}{dx^j}(z)\right|\sqrt{N+1}\int_1^2\left|1-1_B^\delta(y)\right|\frac{dy}{y^2},\end{aligned}\tag{33}$$

which goes to zero as $\delta\to 0$. This proves that $C^\infty_{\text{cusp}}(M,B)$ is dense in $1_B C^\infty_{\text{cusp}}(M)$ with respect to $\|\cdot\|_{2,N}$.

To prove the inequality for B_θ we see from (21), the bound $\tau_1((|m|,|n|))\ll_\varepsilon |mn|^{1+\epsilon}$, and Cauchy–Schwarz on the involved integral that for $1_B\psi_i\in 1_B C^\infty_{\text{cusp}}(M)$ we have that $|B_\theta(1_B\psi_1,1_B\psi_2)|$ is bounded by a constant times

$$\sum_{m,n\neq 0}|mn|^{1+\varepsilon}\sqrt{\int_0^\infty\left|1_{y/|m|\geq 1}V_m^{(\psi_1)}\left(\frac{y}{|m|}\right)\right|^2\frac{dy}{y^2}\int_0^\infty\left|1_{y/|n|\geq 1}V_n^{(\psi_2)}\left(\frac{y}{|n|}\right)\right|^2\frac{dy}{y^2}}.$$

Doing a change of variables this splits as a product of

$$\sum_{m\neq 0}|m|^\varepsilon\sqrt{\int_0^\infty\left|1_{y\geq 1}V_m^{(\psi_1)}(y)\right|^2\frac{dy}{y^2}}$$

times the same expression for ψ_2. Dividing and multiplying the terms by $|m|^{1/2+\varepsilon}$ we can use Cauchy–Schwarz to see that this is bounded by

$$\sqrt{\sum_{m\neq 0}\frac{1}{|m|^{1+2\varepsilon}}}\sqrt{\sum_{m\neq 0}|m|^{1+4\varepsilon}\int_0^\infty\left|1_{y\geq 1}V_m^{(\psi_1)}(y)\right|^2\frac{dy}{y^2}}.$$

Comparing with (33) and (32) we see that this is bounded by a constant times $\|1_B\phi_1\|_{2,1}$, which proves the inequality for B_θ.

To see that $B_\theta(\cdot,\cdot)$ is continuous on $1_B C^\infty_{\text{cusp}}(M)$ we observe that if $1_B\psi_i^{(n_i)}\to 1_B\psi_i$ with respect to $\|\cdot\|_{2,1}$ as $n_i\to\infty$ then we can use that

$$\begin{aligned}&B_\theta(1_B\psi_1^{(n_1)},1_B\psi_2^{(n_2)})-B_\theta(1_B\psi_1,1_B\psi_2)\\ &\qquad=B_\theta(1_B\psi_1^{(n_1)}-1_B\psi_1,1_B\psi_2^{(n_2)})+B_\theta(1_B\psi_1,1_B\psi_2^{(n_2)}-1_B\psi_2),\end{aligned}$$

and the claim now follows easily from the inequality satisfied by B_θ.

□

If ϕ is a Hecke–Maass form then $1_B\phi \in 1_B C^\infty_{\text{cusp}}(M)$ and we consider we consider the expansion

$$1_B\phi(z) = \sum_{m\neq 0} P_{1_{y>1}a_m^{(\phi)},m}(z).$$

with $a_m^{(\phi)}(y) = \epsilon_{\phi,m} 2\lambda_\phi(|m|) y^{1/2} K_{s-1/2}(2\pi\,|m|\,y)$ and where $\epsilon_{\phi,m} = 1$ if ϕ is an even and $\epsilon_{\phi,m} = \operatorname{sgn}(m)$ if ϕ is odd. It follows from this and (21) that $B_\theta(1_B\phi_1, 1_B\phi_2) = 0$ if either ϕ_1 or ϕ_2 is odd. This is also the case when $\theta = 0$ as proved in [18] as follows from $L(\phi, 1/2) = 0$ for ϕ odd. If ϕ_1, ϕ_2 are both even Hecke–Maass forms with Laplace eigenvalues $s_1(1-s_1)$ and $s_2(1-s_2)$, respectively, we see that $B_\theta(1_B\phi_1, 1_B\phi_2)$ equals

$$4\pi \sum_{m,n\geq 1} \frac{\tau_1((m,n))\lambda_{\phi_1}(m)\lambda_{\phi_2}(n)}{(mn)^{1/2}} \int_{\max(m,n)}^\infty K_{s_1-1/2}(2\pi y)\overline{K_{s_2-1/2}(2\pi y)}\frac{dy}{y},$$

for $0 < \theta < 1/2$ and for $\theta = 1/2$, the number $B_{1/2}(1_B\phi_1, 1_B\phi_2)$ equals

$$4\pi \sum_{m,n\geq 1} \frac{\tau_1((m,n))\lambda_{\phi_1}(m)\lambda_{\phi_2}(n)}{(mn)^{1/2}} \int_{\max(m,n)}^\infty K_{s_1-1/2}(2\pi y)\overline{K_{s_2-1/2}(2\pi y)}e^{-2\pi^2 y^2(m^2+n^2)}\frac{dy}{y},$$

as claimed in Theorem 1.3 (iv).

It is a deep number-theoretic fact that the central value of $L(\phi_j, s)$ is non-negative. Luo and Sarnak [18] observed that this follows from noticing that these numbers are essentially the eigenvalues of the non-negative Hermitian form B_0. We are now ready to draw a similar conclusion for $B_\theta(1_B\phi, 1_B\phi)$ as computed above from the fact that B_θ is non-negative on $C^\infty_{0,0}(M,B)$ for any $0 \leq \theta < 1$. Since we only know beforehand that B_θ is non-negative on the smaller space $C^\infty_{0,0}(M,B)$, we use the continuity properties of $B_\theta(\cdot,\cdot)$.

Proof of Corollary 1.4. We have seen above that the expression on the right of Corollary 1.4 equals, up to a positive constant, the value $B_\theta(1_B\phi, 1_B\phi)$. It follows from Proposition 6.2 there exist $\{\psi_n\} \subset C_{\text{cusp}}(M,B)$ such that $\psi_n \to 1_B\phi$ with respect to $\|\cdot\|_{2,1}$. By Theorem 4.1 we may conclude, since the left-hand side of (22) is non-negative, that $B_\theta(\psi_n, \psi_n) \geq 0$. By the continuity properties of B_θ in Proposition 6.2 we conclude that $B_\theta(1_B\phi, 1_B\phi) \geq 0$ which proves the result. □

Of course one may make a conclusion analogous to that of Corollary 1.4 for the case $\theta = 1/2$ where the integrand gets multiplied by $e^{-2\pi^2 y^2(m^2+n^2)}$.

References

1. Michael V. Berry, *Regular and irregular semiclassical wavefunctions*, Journal of Physics A: Mathematical and General **10** (1977), no. 12, 2083–2091.
2. Amit Ghosh and Peter Sarnak, *Real zeros of holomorphic Hecke cusp forms*, Journal of the European Mathematical Society (2012), 465–487.
3. Andrew Granville and Igor Wigman, *Planck-scale mass equidistribution of toral Laplace eigenfunctions*, Communications in Mathematical Physics **355** (2017), no. 2, 767–802. MR 3681390
4. Xiaolong Han, *Small scale quantum ergodicity in negatively curved manifolds*, Nonlinearity **28** (2015), no. 9, 3263–3288. MR 3403398

5. Dennis A. Hejhal and Barry N. Rackner, *On the topography of Maass waveforms for PSL(2, Z)*, Experiment. Math. **1** (1992), no. 4, 275–305. MR 1257286
6. Hamid Hezari and Gabriel Rivière, *$Lp̂$ norms, nodal sets, and quantum ergodicity*, Advances in Mathematics **290** (2016), 938–966. MR 3451943
7. Roman Holowinsky and Kannan Soundararajan, *Mass equidistribution for Hecke eigenforms*, Ann. of Math. (2) **172** (2010), no. 2, 1517–1528. MR 2680499
8. Peter Humphries, *Equidistribution in shrinking sets and $L4̂$-norm bounds for automorphic forms*, Mathematische Annalen **371** (2018), no. 3-4, 1497–1543. MR 3831279
9. Peter Humphries and Rizwanur Khan, *On the random wave conjecture for dihedral Maaß forms*, Geometric and Functional Analysis **30** (2020), no. 1, 34–125. MR 4081056
10. Henryk Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics, vol. 17, American Mathematical Society, Providence, RI, 1997. MR 1474964
11. Henryk Iwaniec and Emmanuel Kowalski, *Analytic number theory*, American Mathematical Society Colloquium Publications, vol. 53, American Mathematical Society, Providence, RI, 2004. MR 2061214
12. Stephen Lester, Kaisa Matomäki, and Maksym Radziwiłł, *Small scale distribution of zeros and mass of modular forms*, J. Eur. Math. Soc. (JEMS) **20** (2018), no. 7, 1595–1627. MR 3807308
13. Stephen Lester and Zeév Rudnick, *Small scale equidistribution of eigenfunctions on the torus*, Communications in Mathematical Physics **350** (2017), no. 1, 279–300. MR 3606476
14. Elon Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity*, Ann. of Math. (2) **163** (2006), no. 1, 165–219. MR 2195133
15. Sheng-Chi Liu and Riad Masri, *Nonvanishing of Rankin-Selberg L-functions for Hilbert modular forms*, Ramanujan J. **34** (2014), no. 2, 227–236. MR 3206197
16. Wen Zhi Luo and Peter Sarnak, *Quantum ergodicity of eigenfunctions on $PSL_2(Z)\backslash H^2$*, Inst. Hautes Études Sci. Publ. Math. (1995), no. 81, 207–237. MR 1361757
17. Wenzhi Luo and Peter Sarnak, *Mass equidistribution for Hecke eigenforms*, Comm. Pure Appl. Math. **56** (2003), no. 7, 874–891. MR 1990480
18. ———, *Quantum variance for Hecke eigenforms*, Ann. Sci. École Norm. Sup. (4) **37** (2004), no. 5, 769–799. MR 2103474
19. Paul D. Nelson, *Quantum variance on quaternion algebras, I*, ArXiv:1601.02526 (2016), 1–36.
20. ———, *Quantum variance on quaternion algebras, II*, arXiv:1702.02669 (2017), 1–60.
21. ———, *Quantum variance on quaternion algebras, III*, arXiv:1903.08686 (2019), 1–32.
22. Asbjørn C. Nordentoft, Yiannis N. Petridis, and Morten S. Risager, *Averages of shifted convolution sums for Hecke eigenforms*, In preparation, 2020.
23. Zeév Rudnick, *On the asymptotic distribution of zeros of modular forms*, Int. Math. Res. Not. (2005), no. 34, 2059–2074. MR 2181743
24. Peter Sarnak and Peng Zhao, *The Quantum variance of the Modular Surface*, Annales scientifiques de l'École normale supérieure **52** (2019), no. 5, 1155–1200.
25. Roman Schubert, *Upper Bounds on the Rate of Quantum Ergodicity*, Annales Henri Poincaré **7** (2006), no. 6, 1085–1098.
26. Kannan Soundararajan, *Quantum unique ergodicity for $SL_2(Z)\backslash H$*, Ann. of Math. (2) **172** (2010), no. 2, 1529–1538. MR 2680500
27. Matthew P. Young, *The quantum unique ergodicity conjecture for thin sets*, Adv. Math. **286** (2016), 958–1016. MR 3415701
28. Steven Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces*, Duke Math. J. **55** (1987), no. 4, 919–941. MR 916129
29. ———, *On the rate of quantum ergodicity. I. Upper bounds*, Comm. Math. Phys. **160** (1994), no. 1, 81–92. MR 1262192

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark
E-mail address: acnordentoft@outlook.com

Department of Mathematics, University College London, Gower Street, London WC1E 6BT, United Kingdom
E-mail address: i.petridis@ucl.ac.uk

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark
E-mail address: risager@math.ku.dk