Distribution results for automorphic forms, their periods and masses

Petru Constantinescu

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Mathematics

October 25, 2021

I, Petru Constantinescu, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

We explore a variety of topics in the analytic theory of automorphic forms. The main results of this thesis are about the arithmetic statistics of periods of automorphic forms and the distribution of masses of automorphic forms in the context of Quantum Chaos.

We introduce a new technique for the study of the distribution of modular symbols. Answering an average version of a conjecture due to Mazur and Rubin for $\Gamma_0(N)$ and recovering results of Petridis and Risager using a different method, we show that modular symbols are asymptotically normally distributed, We apply our technique to obtain new results for congruence subgroups of Bianchi groups. Our novel insight is to use the behaviour of the smallest eigenvalue of the Laplace operator for twisted spaces. Our approach also recovers the first and the second moment of the distribution.

In work joint with Asbjørn Nordentoft, we introduce an automorphic method for studying the residual distribution of modular symbols modulo primes. We obtain a refinement of a result of Lee and Sun, which solved an average version of another conjecture of Mazur and Rubin. In addition, we solve the full conjecture in some special cases. Furthermore, we generalise the results to quotients of general hyperbolic spaces.

Lastly, we obtain a generalisation of the Quantum Unique Ergodicity for holomorphic cusp forms, as proved by Holowinsky and Soundararajan. We show that correlations of masses coming from off-diagonal terms dissipate as the weight tends to infinity. This corresponds to classifying the possible quantum limits along any sequence of Hecke eigenforms of increasing weight.

Impact Statement

The study of modular symbols and quantum chaos lies at the crossroads of various branches of mathematics: number theory, mathematical analysis, and mathematical physics. Research on modular symbols has a long and rich history, through the work of Manin, the work of Cremona and LMFDB (L-functions and modular forms database), and the work of Goldfeld and his students. Quantum Unique Ergodicity is a fast-moving and highly active area of research within Quantum Chaos, which bridges diverse areas such as geometry, mathematical physics, dynamics, automorphic forms and arithmetic.

In this thesis we examine some deep questions in number theory, the branch of mathematics that underlies digital communication and internet security. As such it supports the UK to keep its privileged status in fundamental research in number theory. This thesis aligns with the EPSRC's strategic focus for Number Theory in its Mathematical Sciences theme. This research makes connections to neighbouring fields such as the Mathematical Analysis, via spectral theory, and Mathematical Physics.

The immediate beneficiaries of this research are other researchers in number theory, automorphic forms, dynamical systems, and mathematical physics. It is expected that the results will have high long term impact on several disciplines. The publication of the results will ensure that the impact propagates into different research communities.

Acknowledgements

First and foremost, I would like to thank my supervisor Yiannis Petridis. He has always devoted substantial time and effort to me, and I am immensely grateful for his advice, for the long discussions, both academic and otherwise, and for his encouragement and support in good and difficult moments.

I want to thank Asbjørn Nordentoft, it has been a pleasure to collaborate with him and I learned a lot from this experience. Also, I would like to thank Peter Humphries, who has been extremely generous in sharing mathematical ideas and advice.

The analytic number theory group in London has been an important part of my mathematical development and my time in London. I would like to thank my fellow PhD students, postdocs and staff for creating a friendly and stimulating group, and I am very grateful for all the time spent discussing mathematics or otherwise.

My life in London would have not been the same without all my friends from UCL and LSGNT. The many hours spent in the UCL common room having coffee or lunch were a welcome delight to my daily life as a PhD student in mathematics. In addition, I would like to thank my non-mathematical friends Nikolai and Zach, I could have not hoped for better friends to share my journey over the last few years with.

Finally, I would like to thank my family, for their continuous support that made it all possible, and Tila, for all her support for and putting up with my unconventional working hours.

This work was supported by the Engineering and Physical Sciences Research Council [EP/L015234/1], the EPSRC Centre for Doctoral Training in Geometry and Number Theory (The London School of Geometry and Number Theory), University College London.

Contents

1 Introduction **8**

1.1 Distribution results in number theory 8

1.2 Statement of the main results 12

2 Background material **19**

2.1 Geometry of the upper half-plane $\mathbb{H}$ 19

2.2 Automorphic forms 25

2.3 L-functions 32

2.4 Modular symbols 38

2.5 Quantum Unique Ergodicity 46

2.6 Geometry of hyperbolic upper half-space $\mathbb{H}^3$ 49

2.7 Higher dimensional hyperbolic spaces 51

3 Distribution of modular symbols in $\mathbb{H}^3$ **57**

3.1 Introduction 57

3.2 Modular symbols in $\mathbb{H}^3$ 61

3.3 Generating series for modular symbols 63

3.4 Perturbation theory of objects twisted by modular symbols 70

3.5 Moment generating function 77

3.6 Normal distribution of modular symbols 83

3.7 Results for imaginary quadratic fields 85

4 Residual equidistribution of modular symbols and generalisations to higher dimensions **92**

4.1 Introduction 92

4.2 Idea of proof 98

4.3 Some special cases of the conjecture of Mazur and Rubin 100
4.4 Twisted Eisenstein series for $\mathbb{H}^{n+1}$ 103
4.5 Proof of main results 112
4.6 On the size of certain cohomology groups 120

5 Dissipation of correlations of holomorphic cusp forms 128

5.1 Introduction 128
5.2 Spectral theory of weight k automorphic forms 131
5.3 Integral triple product identities 137
5.4 Bounds for Fourier coefficients 141
5.5 Shifted convolution sums 145
5.6 Proofs of Theorem 5.1.1 and Theorem 5.1.2 151

Bibliography 156

Chapter 1

Introduction

> There are five elementary arithmetical operations: addition, subtraction, multiplication, division, and...
> modular forms.
>
> Martin Eichler

> A mathematician, like a painter or poet, is a maker of patterns. If his patterns are more permanent than theirs, it is because they are made with ideas.
>
> G.H. Hardy

A central theme in number theory is the study of statistical and distributional properties of arithmetic objects, such as primes, automorphic forms or special values of L-functions. In this thesis we prove results on the distribution of periods and masses of automorphic forms. The techniques employed are diverse and include the spectral theory of automorphic forms, exponential sums, probabilistic number theory, subconvexity and perturbation theory.

1.1 Distribution results in number theory

One of the most fundamental results in Analytic Number Theory is the *Prime Number Theorem*, which gives an asymptotic for the number of primes up to any point, hence it is a key result in the distribution of primes. Let $\pi(x)$ denote the number of primes less or equal to x. The Prime Number Theorem states that

$$\pi(x) \sim \mathrm{Li}(x) \quad \text{as } x \to \infty, \tag{1.1}$$

where $\mathrm{Li}(x)$ is the logarithmic integral

$$\mathrm{Li}(x) := \int_2^x \frac{dt}{\log t}. \tag{1.2}$$

Now let q be a positive integer. We are interested how the primes distribute modulo q. We observe that if p is a prime larger than q and $p \equiv a \bmod q$, then $(a, q) = 1$, then the *Prime Number Theorem in Arithmetic Progressions* (see [24, Chapter 1] or [48, Chapter 17]) states that each residue class $a \pmod q$ with $(a, q) = 1$ contains roughly the same amount of primes. There are $\phi(q)$ such residue classes, where ϕ is the Euler totient function. More precisely, if we define

$$\pi(x; q, a) := \#\{p \text{ prime } : \ p \leq x, p \equiv a (\bmod q)\},$$

then

$$\pi(x; q, a) \sim \frac{\mathrm{Li}(x)}{\phi(q)}. \tag{1.3}$$

We say that the primes *equidistribute* in the $\phi(q)$ residue classes.

As observed by Riemann in his seminal memoir of 1859, the distribution of the primes is strongly linked to the analytic properties of the *Riemann zeta function.* For $\mathrm{Re}(s) > 1$, we define

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

Then $\zeta(s)$ admits meromorphic continuation to $s \in \mathbb{C}$. It follows that the Prime Number Theorem is equivalent to the non-vanishing of $\zeta(s)$ on the line $\mathrm{Re}(s) = 1$, see [48, Chapter 2]. The famous *Riemann Hypothesis* states that all zeros of $\zeta(s)$ in the *critical strip* $0 < \mathrm{Re}(s) < 1$ belong on the *critical line* $\mathrm{Re}(s) = 1/2$. This would imply the error term

$$|\pi(x) - \mathrm{Li}(x)| \ll x^{1/2} \log x,$$

while the best unconditional bound for the error term is of the form

$$|\pi(x) - \mathrm{Li}(x)| \ll x \exp\left(-c(\log x)^{3/5}(\log\log x)^{-1/5}\right),$$

for some explicit constant $c > 0$, see [34].

For the study of primes in residue classes modulo q, it is useful to define the *Dirichlet L-functions*. Let χ a character modulo q and for $\mathrm{Re}(s) > 1$, we define

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

These L-functions admit meromorphic continuation and functional equation, we refer to Section 2.3 for details. The Prime Number Theorem in Arithmetic Progressions is equivalent to the non-vanishing of $L(1,\chi)$, for all characters χ mod q different from the principal character. If we assume the *Grand Riemann Hypothesis* for $L(s,\chi)$, then

$$\left|\pi(x;q,a) - \frac{x}{\phi(q)}\right| \ll x^{1/2}(\log x).$$

The estimate above is non-trivial only if $q \ll x^{1/2}(\log x)^{-1}$. Unconditionally, a bound is given by the Siegel–Walfisz Theorem, see [24, p. 133]. If $q \leq (\log x)^A$, for some $A \geq 0$, then

$$\left|\pi(x;q,a) - \frac{x}{\phi(q)}\right| \ll_A \frac{x}{(\log x)^A}.$$

The examples highlighted above showcase how the distribution of the zeroes of $\zeta(s)$ and $L(s,\chi)$ have important implications in the distributions of primes. The L-functions $\zeta(s)$ and $L(s,\chi)$ are examples of GL_1 L-functions. In this thesis, we will mainly work with L-functions coming from GL_2 representations over $\mathbb{Q}$. We will use their analytic properties to obtain distribution results for automorphic forms.

1.1.1 Normal distribution in number theory

The normal distribution is one of the most fundamental concepts in mathematics, with many applications to all branches of science. Its importance is highlighted by the Central Limit Theorem, which roughly states that the average of a large number of independent variables converges to the Gaussian distribution, see [3, p. 232]. The probability density function for the standard normal distribution, i.e. with mean 0 and norm 1, is given by

$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, and its distribution function is

$$\Phi(z) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-x^2/2}dx.$$

It is a common theme in number theory to study objects which intrinsically are not 'random' or 'independent', such as the primes or the values of the Riemann zeta function, and to show that they satisfy random-like properties, such as obeying an asymptotic normal distribution.

A cornerstone in this direction is the Erdős–Kac theorem, see [105, p. 348], [29, p. 18]. It states that, with appropriate normalisations, the distribution of the number of prime divisors approaches the Gaussian distribution. We define the prime counting function $\omega(n) := \sum_{p|n} 1$. As $N \to \infty$, we have

$$\frac{1}{N}\#\left\{n \leq N : \frac{\omega(n) - \log\log N}{\sqrt{\log\log N}} \leq y\right\} \to \Phi(y).$$

This means that the values $\omega(n)$ with $n \leq N$ behave like a Gaussian distribution with mean $\log\log N$ and standard deviation $\sqrt{\log\log N}$.

Another important example is the Selberg Central Limit Theorem for the values of the Riemann zeta function on the critical line, see [84] or [107]. As $T \to \infty$, we have

$$\frac{1}{T}\text{meas}\left(\left\{T \leq t \leq 2T \;:\; \log|\zeta(1/2+it)| \leq y\sqrt{\frac{1}{2}\log\log T}\right\}\right) \to \Phi(y),$$

where we use the standard Lebesque measure on the reals. In other words, on the dyadic interval $T \leq t \leq 2T$, we observe that $\log|\zeta(1/2+it)|$ behaves like Gaussian with mean value 0 and standard deviation $\sqrt{\frac{1}{2}\log\log T}$.

One of the most powerful methods in probability to obtain convergence to the normal distribution is the *method of moments*, see [29, p. 59]. If we can show that the truncated k-th moments of a sequence $f(n)$ converge to the k-th moment of the Gaussian distribution, then we can deduce that the sequence converges to the normal distribution. We denote by

$$m_k := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^k e^{-x^2/2}dx$$

the k-th moment of the standard normal distribution. Then $m_k = 0$ for odd k and

$$m_k = \frac{k!}{2^{k/2}(k/2)!}, \quad \text{for even } k.$$

We define the truncated k-th moment as

$$\mathbb{E}_N^k f := \frac{1}{N} \sum_{n \le N} f(n)^k.$$

This means that if $|\mathbb{E}_N^k f - m_k| \to 0$ for all k as $N \to \infty$, then we obtain convergence in distribution.

An alternate way to proceed is by using the Berry–Esseen inequality, see [105, p. 349]. We define the characteristic function

$$F_N(t) := \frac{1}{N} \sum_{n \le N} e^{itf(n)},$$

which is a function on the real parameter t. We observe that the derivatives $F_N^{(k)}(0)$ produce the k-th moments defined above, hence we can view $F_N(\mathrm{t})$ as a *moment generating function.* Then we have for all $T > 0$,

$$\sup_{y \in \mathbb{R}} \left| \frac{\#\{n \ : \ f(n) \le y\}}{N} - \Phi(y) \right| \ll \frac{1}{T} + \int_{-T}^{T} \left| e^{-t^2/2} - F_N(t) \right| \frac{dt}{|t|}.$$

The advantage of using the Berry-Esseen inequality over the method of moments is that we can obtain explicit rates of convergence. However, it is often the case that the moments are easy to compute and the characteristic function or the integral above are difficult to work with.

In Chapter 3 of this thesis, we develop a new method to show that modular symbols obey asymptotically a normal distribution using the Berry–Esseen inequality. In Chapter 4, we show that modular symbols are equidistributed modulo primes. For both cases, we provide rates of convergence, which are presumably optimal.

1.2 Statement of the main results

In this section we give a brief description of the main results of this thesis. For the sake of brevity, we do not provide full definitions and hence some statements may be weaker or less complete than in the subsequent chapters. However, we pinpoint to the location where things are developed properly.

1.2.1 Normal distribution of modular symbols

Modular symbols are fundamental tools in number theory. They can be used to compute modular forms, study elliptic curves, special values of L-functions and homology and cohomology of arithmetic groups.

Let f be a holomorphic cusp form of weight 2 and level N. Then for $r \in \mathbb{Q}$, we define the *raw modular symbol* $\langle r \rangle \in \mathbb{C}$ and the plus/minus version $\langle r \rangle^{\pm} \in \mathbb{R}$

$$\langle r \rangle := \int_{\infty}^{r} f(z)dz, \quad \langle r \rangle^{+} = \frac{\langle r \rangle + \langle -r \rangle}{2}, \quad \langle r \rangle^{-} = \frac{\langle r \rangle + \langle -r \rangle}{2i}. \tag{1.4}$$

The path of integration can be taken as the vertical line connecting $r \in \mathbb{Q}$ to the cusp at ∞, and hence modular symbols can be viewed as *periods* of the cusp form f. The plus/minus modular symbols correspond to integrating the real part/imaginary part $\mathrm{Re}(f(z)dz)/\mathrm{Im}(f(z)dz)$ of the 1-form $f(z)dz$. We refer to Section 2.4.1 for a detailed exposition of the construction and properties of modular symbols.

Mazur and Rubin [66] initiated the study of the arithmetic distribution of modular symbols for congruence subgroups in order to study the excess rank of elliptic curves over cyclotomic fields. They put forward a number of conjectures that have received a lot of attention in recent years. We refer to Section 2.4.2 for an overview of the arithmetic statistics of modular symbols, where we describe these conjectures and the subsequent work they inspired, which includes parts of this thesis.

Some of these conjectures are concerned with the value distribution of the modular symbols coming from rationals with fixed denominator c. They suggest that, as $c \to \infty$, the values

$$\left\{ \langle a/c \rangle^{+} \ : \ a \in (\mathbb{Z}/c\mathbb{Z})^{*} \right\}$$

obey asymptotically a normal distribution. Additionally, they propose a law for the second moment.

Petridis and Risager [79] obtained an average version of these conjectures with an extra average over the denominators. In Chapter 3, we introduce a new technique to study the distribution of modular symbols. Our new insight is to use the Berry–Esseen inequality, an important tool in probabilistic number theory, and the perturbation theory of the smallest eigenvalue of the Laplacian. Our approach gives convergence rates for the distribution, and can be naturally extended to more general settings. We now state the result of Petridis and Risager, which can also be proved using our method. We define the sample space

$$R_d(X) = \left\{ \frac{a}{c} \mid 0 < a < c \leq X, (a,c) = 1, (c,q) = d \right\}. \tag{1.5}$$

Theorem 1.2.1. *[79] Let f be a holomorphic cusp form of weight 2 and square-free level q.*

(i) There exists a constant C_f such that, for each fixed $d|q$, the values

$$\left\{ \frac{\langle r \rangle^+}{(C_f \log X)^{1/2}} \ : \ r \in R_d(X) \right\}$$

have asymptotically a standard normal distribution as $X \to \infty$.

(ii) For each divisor d of q, there exists a constant $D_{f,d}$ such that, as $X \to \infty$,

$$\frac{1}{\#R_d(X)} \sum_{r \in R_d(X)} (\langle r \rangle^+)^2 = C_f \log X + D_{f,d} + o(1).$$

Our method can be extended to the upper half-space $\mathbb{H}^3$. We obtain the following result for Bianchi modular forms and congruence subgroups of Bianchi groups.

Theorem 1.2.2. *Let K be an imaginary quadratic number field of class number one. Let $\mathfrak{n} \lhd \mathcal{O}_K$ be a square-free ideal and F be a Bianchi modular form of weight 2 and level $\mathfrak{n}$. For $r \in K$, let $\langle r \rangle$ denote the modular symbol corresponding to F. For $\mathfrak{d}|\mathfrak{n}$, set*

$$Q_{\mathfrak{d}}(X) = \{a/c \in K \mid a \in (\mathcal{O}_K/(c))^\times \ , \ (c) + \mathfrak{n} = \mathfrak{d} \ , \ 0 < |c| < X\}.$$

(i) There exists a constant C_F such that the data

$$Q_{\mathfrak{d}}(X) \to \mathbb{R}, \quad \frac{a}{c} \mapsto \frac{\langle a/c \rangle}{\sqrt{C_F \log X}},$$

has asymptotically a standard normal distribution.

(ii) There exists a constant $D_{F,\mathfrak{d}}$ such that

$$\frac{1}{|Q_{\mathfrak{d}}(X)|} \sum_{a/c \in Q_{\mathfrak{d}}(X)} \left\langle \frac{a}{c} \right\rangle^2 = C_F \log X + D_{F,\mathfrak{d}} + o(1) \ .$$

We refer to Theorem 3.1.2 for a more general version of this theorem. We prove a version that holds for all cofinite Kleinian groups $\Gamma < \mathrm{SL}_2(\mathbb{C})$ with further restrictions on the location of the cusps.

1.2.2 Equidistribution mod p

We note that if f is a newform of weight 2 and level N, as we vary along $r \in \mathbb{Q}$, the image of $r \mapsto \langle r \rangle^{\pm}$ is a lattice in $\mathbb{R}$. Let Ω^+ and Ω^- be the real and imaginary periods of f, which are normalising factors that dilate these lattices to $\mathbb{Z}$. Consider the *normalised modular symbol*

$$[r]^{\pm} := \frac{\langle r \rangle^{\pm}}{\Omega^{\pm}} \in \mathbb{Z}.$$

Mazur and Rubin [65] also conjectured that the normalised modular symbols $[r]^+$ equidistribute modulo p as we vary along fractions with fixed denominator. Lee and Sun [57] proved an average version of the conjecture using dynamical methods. In joint work with Nordentoft, we show that modular symbols coming from a basis of newforms obey a joint equidistribution when we also average over denominators. We view $[r]^+$ and $[r]^-$ as random variables on the space $R_d(X)$ defined in (1.5) and we also allow restriction on the location of $r \in \mathbb{R}/\mathbb{Z}$. We denote by $[r]^{\pm}_f$ the normalised modular symbol coming from the cusp form f.

Theorem 1.2.3. *Let $f_1, \ldots, f_d$ a basis of newforms for $S_2(\Gamma_0(q))_{new}$. The random variables $[r]^{\pm}_{f_i}$ defined on the sample spaces $R_d(X)$ converge in distribution to the uniform distribution on $(\mathbb{Z}/p\mathbb{Z})^{2d}$ as $X \to \infty$. More precisely, for any fixed $\mathbf{a} \in (\mathbb{Z}/p\mathbb{Z})^{2d}$ and any interval $I \subset \mathbb{R}/\mathbb{Z}$, we have*

$$\frac{\#\left\{a/c \in R_d(X) \cap I \mid ([a/c]^+_{f_1}, \ldots, [a/c]^-_{f_d}) \equiv \mathbf{a} \bmod p\right\}}{\# R_d(X)} = \frac{|I|}{p^{2d}} + o(1)$$

as $X \to \infty$.

In addition, we obtain the rate of convergence to the uniform distribution by evaluating a mean-square type sum.

Theorem 1.2.4. *Fix $f \in S_2(\Gamma_0(N))_{new}$. For large enough p, there exist constants $c_p, \delta_p >$*

0 such that, as $X \to \infty$,

$$\frac{1}{p}\sum_{l\in\mathbb{Z}/p\mathbb{Z}}\left(\frac{\#\{a/c \in R_d(X) \mid [a/c]_f^{\pm} \equiv l \bmod p\}}{\#R_d(X)} - \frac{1}{p}\right)^2 \sim c_p X^{-\delta_p}.$$

Furthermore, we have for X *large enough (depending on* p*),*

$$|\{a/c \in R_d(X) \mid [a/c]_f^{\pm} \equiv l \bmod p\}| \leq |\{a/c \in R_d(X) \mid [a/c]_f^{\pm} \equiv 0 \bmod p\}|$$

with equality if and only if $l \equiv 0 \mod p$.

Remark 1.2.1. We can compute the constants c_p and δ_p explicitly, see Section 4.5.3 for more details. The second part of the theorem shows that modular symbols obey an interesting phenomenon, which we can compare to the Chebyshev biases for primes. The residue class 0 (mod p) will always contain more elements than any other residue class.

We also prove a special case of the conjecture when we consider modular symbols of fixed denominator, see Section 4.3 for more details.

Theorem 1.2.5. *Let* (N,p) *a pair of 'admissible primes'with* $p \mid N-1$. *Then there exists* $f \in S_2(\Gamma_0(N))$ *such that the values* $\{[a/c]^{\pm} \mid a \in (\mathbb{Z}/c\mathbb{Z})^*\}$ *equidistribute exactly modulo* p, *for all* $c \equiv 0 \bmod N$. *This means that for all* $l \in (\mathbb{Z}/p\mathbb{Z})$, *the equation* $[a/c]^{\pm} \equiv l(mod\ p)$ *has exactly* $\phi(c)/p$ *solutions for* $a \in (\mathbb{Z}/c\mathbb{Z})^*$.

We refer to (4.9) for the definition of an admissible pair. Note that all pairs of primes (N,p) with $N < 250$ such that $p \geq 5$ and $p|N-1$ are admissible unless $N = 31, 103, 127, 131, 181, 199, 211$.

1.2.3 Higher dimensional hyperbolic spaces

We can also extend the distribution results for modular symbols to the n-dimensional hyperbolic space $\mathbb{H}^n$. As we see in Section 2.4.1, there is a correspondence between modular symbols and elements of cuspidal cohomology $H^1_{\mathrm{cusp}}(\Gamma_0(N),\mathbb{R})$ (we assume $\Gamma_0(N)$ acts trivially on $\mathbb{R}$ in these cohomology groups). We use this point of view to obtain generalisation for $\mathbb{H}^n$. We let $\Gamma < \mathrm{SO}(n,1)$ a cofinite discrete group acting on $\mathbb{H}^n$. We refer to Section 2.7 for an introduction to the geometry of the quotient space $\Gamma\backslash\mathbb{H}^n$.

Theorem 1.2.6. *(i) Let* $\omega \in H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$. *Then the values of the map* $\Gamma_\infty\backslash\Gamma/\Gamma_\infty \to \mathbb{R}$ *given by* $\gamma \mapsto \omega(\gamma)$ *are asymptotically normally distributed with respect to a 'natural* *arithmetic ordering'of* $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$.

(ii) Let $\omega_1, \dots, \omega_d \in H^1_{\mathrm{cusp}}(\Gamma, \mathbb{Z}/p\mathbb{Z})$ *be 'linearly independent'. Then the random variables* $\gamma \mapsto (\omega_1(\gamma), \dots, \omega_d(\gamma))$ *are asymptotically uniformly distributed on* $(\mathbb{Z}/p\mathbb{Z})^d$ *with respect to the same arithmetic ordering of* $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$.

We refer Sections 2.7 and 4.1 for rigorous statements and definitions.

1.2.4 Quantum Unique Ergodicity

Mass equidistribution of eigenfunctions is a central topic in quantum chaos and number theory. Let $X = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. In [87], Rudnick and Sarnak conjectured that normalised Maaß cusp forms ϕ of eigenvalue λ obey Quantum Unique Ergodicity (QUE) as $\lambda \to \infty$. This means that, as $\lambda \to \infty$, the measures $\mu_\phi := |\phi(z)|^2 \frac{dxdy}{y^2}$ approach the uniform distribution measure $\frac{3}{\pi}\frac{dxdy}{y^2}$. Lindenstrauss [59] showed that for Hecke–Maaß forms, the only possible limiting measures are of the form $\frac{3}{\pi}c\frac{dxdy}{y^2}$, with $0 < c \leq 1$, and Soundararajan [101] completed the proof of QUE for Hecke–Maaß forms, showing that $c = 1$. In [44], Holowinsky and Soundararajan prove QUE for holomorphic Hecke eigencusp forms, which we state below.

Theorem 1.2.7. *[44] Let f be a normalised Hecke eigencusp form of weight k. For any ϕ smooth and bounded on X,*

$$\left\langle \phi y^{k/2} f, y^{k/2} f \right\rangle = \frac{1}{\mathrm{vol}(X)} \langle \phi, 1 \rangle + o(1), \quad \text{as} \quad k \to \infty.$$

In Chapter 5 we generalise these results to off-diagonal terms, where we consider two different eigencusp forms f and g of weights k_1 and k_2 respectively. We show that correlations of masses dissipate as $k_1 k_2 \to \infty$. Denote by $F_{k_1}(z) := y^{k_1/2} f(z)$ and $G_{k_2}(z) := y^{k_2/2} g(z)$, which are eigenfunctions of the Laplacians of weight k_1 and k_2 respectively. Since f and g may have different weights, we need to employ raising and lowering operators. We let $R_{k_1}^{k_2}$ be an isometry from the space of automorphic forms of weight k_1 to forms of weight k_2 given by successive applications of raising/lowering operators, see (5.22) for a precise definition. We prove the following theorem.

Theorem 1.2.8. *Let f and g be L^2-normalised holomorphic Hecke cusp forms of weights*

k_1 and k_2 respectively with $k_1 \leq k_2$. Let

$$\delta_{f=g} = \begin{cases} 1. & \textit{if } f = g; \\ 0, & \textit{otherwise.} \end{cases}$$

(i) Fix any $\phi \in C_b(\Gamma\backslash\mathbb{H})$. Along any sequences of such f and g, we have

$$\left\langle \phi\left(R_{k_1}^{k_2} F_{k_1}\right), G_{k_2} \right\rangle \to \delta_{f=g} \frac{1}{\operatorname{vol}(X)} \langle \phi, 1 \rangle \quad \textit{as} \quad k_2 \to \infty.$$

(ii) Fix l a nonnegative integer and let ϕ a square-integrable automorphic form of weight l. Then

$$\langle \phi F_k, G_{k+l} \rangle \to \delta_{f=g} \frac{1}{\operatorname{vol}(X)} \langle \phi, 1 \rangle \quad \textit{as} \quad k \to \infty.$$

Remark 1.2.2. Theorem 1.2.8(i) corresponds to a generalisation of Quantum Unique Ergodicity by classifying the possible quantum limits of Hecke cusp forms when we project back to the modular surface. That is, along any sequence of holomorphic Hecke eigenforms of increasing weight, we show there are two possible limit points. Part (ii) corresponds to going along a sequence where the difference in weight is fixed. We can moreover let the difference l grow with k, see Remark 5.1.2 for details.

Chapter 2

Background material

In this chapter we present the background material upon which we develop the later chapters. We give a brief introduction to the properties of Fuchsian groups and quotient surfaces, automorphic forms, L-functions and modular symbols. We introduce key ingredients in our toolset, such as Eisenstein series, Kloosterman sums or the spectral theory of automorphic functions. Lastly, we discuss brief generalisations to higher dimensional hyperbolic spaces.

2.1 Geometry of the upper half-plane $\mathbb{H}$

We refer to [47], [53], [98] for comprehensive accounts on the properties of the hyperbolic upper half-space and Fuchsian groups. We consider the upper half of the complex plane

$$\mathbb{H} := \{z = x + iy \ : \ x \in \mathbb{R}, \ y > 0\}.$$

The hyperbolic plane $\mathbb{H}$ becomes a Riemannian manifold when equipped with the hyperbolic metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2}.$$

The correponding volume element is given by

$$d\mu := \frac{dxdy}{y^2}.$$

The Laplace–Beltrami operator Δ acting on $\mathbb{H}$ is given by

$$\Delta = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right).$$

The group $\mathrm{SL}_2(\mathbb{R})$ acts on $\mathbb{H}$ by linear fractional transformations. If $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ and $z \in \mathbb{H}$, we have

$$\gamma(z) = \frac{az+b}{cz+d} \in \mathbb{H}.$$

We extend $\mathbb{H}$ by adding the real line and the point at infinity $\{\infty\}$, to obtain $\mathbb{H}^* := \mathbb{H} \cup \mathbb{R} \cup \{\infty\}$. The action of $\mathrm{SL}_2(\mathbb{R})$ can be extended naturally from $\mathbb{H}$ to its boundary $\mathbb{R} \cup \{\infty\}$. Since γ and $-\gamma$ describe the same motion, the group $\mathrm{PSL}_2(\mathbb{R}) := \mathrm{SL}_2(\mathbb{R})/\{\pm I\}$ describes the orientation-preserving isometries of $\mathbb{H}^*$. The Laplacian commutes with all isometries, i.e. for all $\gamma \in \mathrm{SL}_2(\mathbb{R})$ and $z \in \mathbb{H}$, we have $\Delta(f(\gamma z)) = (\Delta f)(\gamma z)$.

The non-identity linear fractional transformations are classified into three categories. Let $\gamma \in \mathrm{SL}_2(\mathbb{R}) \setminus \{\pm I\}$.

- γ is *parabolic* if $\mathrm{Tr}(\gamma) = \pm 2$. Then γ is conjugate to a matrix of the form $\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$, which corresponds to the translation motion $z \mapsto z + \lambda$. Parabolic motions have one fixed point on $\mathbb{R} \cup \{\infty\}$.

- γ is *elliptic* if $|\mathrm{Tr}(\gamma)| < 2$. Then γ is conjugate to a matrix of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, for some $0 \leq \theta < \pi$, which corresponds to rotation of angle 2θ around i. Elliptic motions have one fixed point on $\mathbb{H}$.

- γ is *hyperbolic* if $|\mathrm{Tr}(\gamma)| > 2$. Then γ is conjugate to a matrix of the form $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$, which corresponds to the dilation motion $z \mapsto \lambda^2 z$. Hyperbolic motions have two fixed points on $\mathbb{R} \cup \{\infty\}$.

Now consider Γ a *Fuchsian subgroup* of $\mathrm{SL}_2(\mathbb{R})$, that is a group Γ that acts discontinously on $\mathbb{H}$, in the sense that the orbit of any point in $\mathbb{H}$ has no limit points in $\mathbb{H}$. We assume that the group Γ is Fuchsian of the *first kind*, meaning that every point on the boundary $\mathbb{R} \cup \{\infty\}$ is a limit point of an orbit Γz, for some $z \in \mathbb{H}$.

A set $F \subset \mathbb{H}$ is a *fundamental domain* for Γ if

(i) distinct points in F° (the interior of F) are not Γ-equivalent, i.e. if $z, w \in F^\circ$ and $\gamma \in \Gamma$ such that $z = \gamma w$, then $z = w$ and $\gamma = \pm I$;

(ii) any orbit of Γ contains a point in $\overline{F}$ (the closure of F), i.e. for any $z \in \mathbb{H}$, there exists $\gamma \in \Gamma$ and $w \in \overline{F}$ such that $w = \gamma z$.

The fundamental domain is not unique, however all fundamental domains have the same volume

$$\mathrm{vol}(\Gamma) := |F| = \int_F d\mu.$$

The volume of a Fuchsian group of the first kind is always finite, we call such a group *cofinite*. A fundamental domain of Γ can be chosen as a hyperbolic polygon with vertices in $\mathbb{H}^*$ such that edges are identified in pairs by elements of Γ. The group Γ is called *co-compact* if the fundamental polygon is compact. A cofinite group Γ is co-compact if and only if it has no parabolic elements, see [47, p. 42].

Assume that Γ is not co-compact. A point $\mathfrak{a} \in \mathbb{R} \cup \{\infty\}$ is called a *cusp* of Γ if there exists a parabolic element $\gamma \in \Gamma$ such that $\gamma\mathfrak{a} = \mathfrak{a}$. The stability group of a cusp $\mathfrak{a}$ is the infinite cyclic group of parabolic motions

$$\Gamma_{\mathfrak{a}} := \{\gamma \in \Gamma \ : \ \gamma\mathfrak{a} = \mathfrak{a}\}.$$

There exists a *scaling matrix* $\sigma_{\mathfrak{a}} \in \mathrm{SL}_2(\mathbb{R})$ such that $\sigma_{\mathfrak{a}}\infty = \mathfrak{a}$. For $Y > 0$, we denote by $F_{\mathfrak{a}}(Y)$ the *cuspidal sector*

$$F_{\mathfrak{a}}(Y) := \sigma_{\mathfrak{a}}\{z = x + iy \in \mathbb{H} \ : \ 0 \leq x \leq 1, y \geq Y\}.$$

Let $\mathfrak{a}_1, \ldots \mathfrak{a}_k$ be representatives for the equivalence classes of cusps of Γ. For Y large enough, we have a fundamental domain F of the form

$$F = F(Y) \cup \bigcup_{j=1}^{k} F_{\mathfrak{a}_j}(Y),$$

where $F(Y)$ is a compact set.

Let Γ a be cofinite group. Then the quotient space $\Gamma\backslash\mathbb{H}$ can be given a complex structure and identified as a Riemann surface.

2.1.1 Double coset decomposition and Kloosterman sums

Let Γ be a cofinite group with cusps. Let $\mathfrak{a}$, $\mathfrak{b}$ be two cusps (not necessarily) distinct with corresponding scaling matrices $\sigma_{\mathfrak{a}}$ and $\sigma_{\mathfrak{b}}$. We have that

$$\sigma_{\mathfrak{a}}^{-1}\Gamma_{\mathfrak{a}}\sigma_{\mathfrak{a}} = \sigma_{\mathfrak{b}}^{-1}\Gamma_{\mathfrak{b}}\sigma_{\mathfrak{b}} = \Gamma_\infty = \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \ : \ b \in \mathbb{Z} \right\}.$$

We are interested in partitioning the set $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ into double cosets with respect to Γ_∞. We first look at the subset with fixed point at ∞, that is

$$\Omega_\infty = \left\{ \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} \right\}.$$

Now, for $c > 0$, fix some $\omega_{a/c} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$. Then if $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ is in the double coset $\Gamma_\infty\omega_{a/c}\Gamma_\infty$, then $\gamma = c$ and α and δ are determined uniquely modulo c. We have the following decomposition into a disjoint union

$$\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} = \delta_{\mathfrak{ab}}\Omega_\infty \cup \bigcup_{c>0} \bigcup_{a \bmod c} \Gamma_\infty\omega_{a/c}\Gamma_\infty, \tag{2.1}$$

where a and c run over numbers such that $\begin{pmatrix} a & * \\ c & * \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ and

$$\delta_{\mathfrak{ab}} = \begin{cases} 1, & \text{if } \mathfrak{a} \text{ and } \mathfrak{b} \text{ equivalent;} \\ 0, & \text{otherwise.} \end{cases}$$

We denote by $T_{\mathfrak{ab}}$ a system of representatives of the double cosets in $\Gamma_\infty\backslash\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}/\Gamma_\infty$ with positive lower-left entry. Also, we define

$$R_{\mathfrak{ab}} := \left\{ \frac{a}{c} \bmod 1 \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in T_{\mathfrak{ab}} \right\}.$$

Then the map

$$T_{\mathfrak{a}\mathfrak{b}} \to R_{\mathfrak{a}\mathfrak{b}}, \quad \gamma \mapsto \gamma\infty \bmod 1,$$

is well-defined and bijective. Hence we can write the decomposition (2.1) as

$$\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}} = \delta_{\mathfrak{a}\mathfrak{b}}\Omega_\infty \cup \bigcup_{r\in R_{\mathfrak{a}\mathfrak{b}}} \Gamma_\infty\omega_r\Gamma_\infty.$$

It is useful to define

$$\mathcal{C}_{\mathfrak{a}\mathfrak{b}} = \left\{ c > 0 \ : \ \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in T_{\mathfrak{a}\mathfrak{b}} \right\}. \tag{2.2}$$

Let $m, n \in \mathbb{Z}$. The *Kloosterman sum* is defined as

$$S_{\mathfrak{a}\mathfrak{b}}(m, n; c) := \sum_{\left(\begin{smallmatrix} a & * \\ c & d \end{smallmatrix}\right) \in T_{\mathfrak{a}\mathfrak{b}}} e\left(\frac{ma + nd}{c}\right). \tag{2.3}$$

Kloosterman sums are central objects in analytic number theory, for example they are used to understand the Fourier coefficients of cusp forms, as we will se later. If $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, there is only one cusp and we have the classical Kloosterman sum

$$S(m, n; c) = \sum_{ad \equiv 1 \bmod c} e\left(\frac{ma + nd}{c}\right).$$

We mention the far-reaching Weil bound

$$|S(m, n; c)| \le (m, n, c)^{1/2}\tau(c)c^{1/2}, \tag{2.4}$$

where $\tau(c)$ is the divisor function. This bound gives us square-root cancellation for Kloosterman sums and it follows from the Riemann hypothesis for curves over finite fields.

It is useful to define the Dirichlet L-series

$$L_{\mathfrak{a}\mathfrak{b}}(s, m, n) = \sum_{\left(\begin{smallmatrix} a & * \\ c & d \end{smallmatrix}\right) \in T_{\mathfrak{a}\mathfrak{b}}} e\left(\frac{ma + nd}{c}\right) c^{-2s} = \sum_{c \in \mathcal{C}_{\mathfrak{a}\mathfrak{b}}} \frac{S(m, n; c)}{c^{2s}}. \tag{2.5}$$

The analytic properties of $L_{\mathfrak{a}\mathfrak{b}}(s, m, n)$ appear in the Fourier coefficients of Eisenstein series and their analytic properties play an important role in Chapters 3 and 4.

2.1.2 Congruence groups

The most important examples of cofinite Fuchisan groups and most interesting from an arithmetic point of view are the congruence groups. The first example is the *modular group* $\mathrm{SL}_2(\mathbb{Z})$ with its fundamental domain

$$\{z = x + iy \ : \ |x| \leq 1/2, |z| \geq 1\}.$$

The quotient surface $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ is called the *modular curve* and has one cusp at ∞. We have that $\mathrm{vol}(\mathrm{SL}_2(\mathbb{Z})) = 3/\pi$.

Let N be a positive integer. The *principal congruence group of level* N is the subgroup

$$\Gamma(N) := \left\{ \gamma \in \mathrm{SL}_2(\mathbb{Z}) \ : \ \gamma \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N \right\}.$$

A subgroup Γ such $\Gamma(N) \subset \Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is called a *congruence group of level* N, where N is the smallest integer with this property. The most important example for us is $\Gamma_0(N)$, the *Hecke congruence group of level* N

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ c \equiv 0 \mod N \right\}.$$

One can see that

$$[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] = N^3 \prod_{p|N} \left(1 - p^{-2}\right) \quad \text{and} \quad [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] = N \prod_{p|N} \left(1 + p^{-1}\right),$$

see [98, p. 25] ot [47, p. 48] for details and for description of the cusps of such groups. An important particular case for us is presented by square-free integers N, when a complete set of inequivalent cusps of $\Gamma_0(N)$ is given by $1/d$ with $d|N$. The cusp $1/N$ is equivalent to the cusp at infinity. In this case, we have that

$$R_{\infty\frac{1}{d}} = \left\{ \frac{a}{c} \in \mathbb{Q}/\mathbb{Z} \ : \ (a, c) = 1, (c, q) = d \right\}.$$

Therefore we have a nice arithmetic description of double coset representatives, see the parallels with the set (1.5) defined in the introduction upon which we obtain distribution results for modular symbols.

2.2 Automorphic forms

2.2.1 Holomorphic modular forms

Let k be an integer. We definte the *weight k right action* of $\mathrm{GL}_2(\mathbb{R})$ on functions $f : \mathbb{H} \to \mathbb{C}$ by

$$(f|_k\gamma)(z) := (cz+d)^{-k} f(\gamma z)(ad-bc)^{k/2}, \quad \text{where } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{R}).$$

This is a well-defined action, in the sense that $f|_k(\gamma_1\gamma_2) = (f|_k\gamma_1)|_k\gamma_2$, for all $\gamma_1, \gamma_2 \in \mathrm{GL}_2(\mathbb{R})$.

Fix $\Gamma < \mathrm{SL}_2(\mathbb{R})$ be a cofinite group with cusps.

Definition 2.2.1. *A holomorphic modular form of weight k for Γ is a holomorphic function $f : \mathbb{H} \to \mathbb{C}$ such that*

(i) $f|_k\gamma = f$, for all $\gamma \in \Gamma$;

(ii) f is holomorphic at all cusps of Γ.

The condition (ii) means the following. For a cusp $\mathfrak{a}$ of Γ, we define $f_{\mathfrak{a}} := f|_k\sigma_{\mathfrak{a}}$. Then $f_{\mathfrak{a}}$ is invariant under $\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{a}}$, which corresponds to translations $z \to z+h$, where h is the width of the cusp $\mathfrak{a}$. Hence we can write $f_{\mathfrak{a}}(z) = F_{\mathfrak{a}}(e^{2\pi i z/h})$, where $F(q)$ is meromorphic in the domain $0 < |q| < r$, for some $r > 0$. Condition (ii) means that $F_{\mathfrak{a}}$ is holomorphic at $q = 0$. Therefore $f_{\mathfrak{a}}$ has Fourier expansion

$$f_{\mathfrak{a}}(z) = \sum_{n \geq 0} \hat{f}_{\mathfrak{a}}(n) e(nz/h).$$

If, moreover, $\hat{f}_{\mathfrak{a}}(0) = 0$, we say that f *vanishes* at the cusp $\mathfrak{a}$. If f vanishes at all cusps, we call it a *cusp form.* We denote the space of holomorphic modular forms of weight k for Γ by $M_k(\Gamma)$. In addition, we let $S_k(\Gamma)$ the subset of $M_k(\Gamma)$ denote the space of weight k cusp forms for Γ.

We note that, if f is a modular form of weight 2 for Γ, then $f(\gamma z)d(\gamma z) = f(z)dz$, for all $\gamma \in \Gamma$. Hence $f(z)dz$ is a Γ-invariant holomorphic one-form. This gives a correspondence between $S_k(\Gamma)$ and regular differential one-forms on $\Gamma\backslash\mathbb{H}$, which a key fact in the construction of modular symbols, see Section 2.4.1.

We can easily check that if $f \in S_k(\Gamma)$, then $y^{k/2}|f(z)|$ is Γ-invariant and bounded on $\mathbb{H}$ (the converse is also true, in the sense that if $y^{k/2}|f(z)|$ bounded on $\mathbb{H}$, then $f(z)$

vanishes at all cusps of Γ). More generally, we note that if $f, g \in S_k(\Gamma)$, then the function $y^k f(z)\overline{g(z)}$ is Γ-invariant. Hence we can define the *Petersson inner product* on $S_k(\Gamma)$ given by

$$\langle f, g\rangle := \int_{\Gamma\backslash\mathbb{H}} y^k f(z)\overline{g(z)} d\mu.$$

We next proceed to construct the fundamental examples of holomorphic modular forms, which are the Eisenstein and Poincaré series. To keep the notation light, we define the series associated to the cusp at ∞, but we can do similar constructions for all cusps. Let $k > 2$ and define the Eisenstein series

$$E_k(z) := \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\in\Gamma_\infty\backslash\Gamma} (cz+d)^{-k}.$$

Then $E_k(z)$ converges absolutely, and hence is holomorphic in $\mathbb{H}$, and is clearly modular by definition. It follows that $E_k \in M_k(\Gamma)$. Now fix an integer $m \geq 0$. The m-th Poincaré series of weight k is given by

$$P_m(z) := \sum_{\gamma=\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\in\Gamma_\infty\backslash\Gamma} (cz+d)^{-k} e(m\gamma z).$$

For $m = 0$, we have that $P_m(z) = E_k(z)$ and for $m > 0$, $P_m(z) \in S_k(\Gamma)$, see [48, p. 358]. Moreover, if $f \in S_k(\Gamma)$ has Fourier expansion $f(z) = \sum_{n>0} a_n e(nz)$, then a simple application of the unfolding principle, see [48, p. 359], shows that

$$\langle f, P_m\rangle = \frac{\Gamma(k-1)}{(4\pi m)^{k-1}} a_n.$$

As a consequence, we see that the Poincaré series span $S_k(\Gamma)$. It is interesting to mention that the Fourier coefficients of Poincaré series are closely related to Kloosterman sums, see [90, p. 23] for details.

We define the *Hecke operators* T_n on functions $f : \mathbb{H} \to \mathbb{C}$ by

$$T_n f(z) = \sum_{ad=n} \left(\frac{a}{d}\right)^k \sum_{0\leq b<d} f\left(\frac{az+b}{d}\right).$$

Then the following properties hold, see [46, Chapter 6], [90, p.29]:

(i) $T_{mn} = T_m T_n$ if $(m, n) = 1$;

(ii) If p prime, then

$$T_{p^n}T_p = T_{p^{n+1}} + p^{n-1}T_{p^{n-1}};$$

(iii) If $(m, N) = 1$, then T_m acts on $S_2(\Gamma_0(N))$ and it is self-adjoint, i.e

$$\langle T_m f, g\rangle = \langle f, T_m g\rangle\,, \quad \text{for all } f, g \in S_k(\Gamma_0(N));$$

(iv) If $f(z) = \sum_{n\geq 1} a(n)e(nz)$ is an eigenform of all Hecke operators T_m with eigenvalues $\lambda(m)$, then

$$a(n) = a(1)n^{\frac{k-1}{2}}\lambda(n), \quad \text{for all } n.$$

We note that if $f \in S_k(\Gamma_0(M))$, for some $M|N$, then the function $z \mapsto f(dz)$ belongs to $S_k(\Gamma_0(N))$, for any d divisor of $\frac{N}{M}$. We denote by $S_k(\Gamma_0(N))_{\text{old}}$ the space of *oldforms*, which is composed of such functions for proper divisors M of N. We denote its orthogonal complement in $S_k(\Gamma_0(N))$ by $S_k(\Gamma_0(N))_{\text{new}}$, which is the space of *newforms*. Then there exists a basis of $S_k(\Gamma_0(N))_{\text{new}}$ consisting of eigenfunctions of T_p, for all $p \nmid N$, see [11, Theorem 1.4.5].

Lastly, we mention the *Atkin–Lehner involutions*. Let e a divisor of N and let W_e any integral matrix of the form $\begin{pmatrix} ae & b \\ cN & de \end{pmatrix}$ which has determinant e. Then the matrix W_e normalises $\Gamma_0(N)$ and we have that $W_e^2 \in e\Gamma_0(N)$, hence the action of the matrix W_e is an involution. Moreover, the action of W_e on $S_k(\Gamma_0(N))$ commutes with the action of all Hecke operators.

2.2.2 Maaß forms

In this subsection we highlight a different kind of automorphic functions that have more straightforward transformation rules, but are not necessarily holomorphic. Maaß forms are key ingredients in the hammonic analysis of $\Gamma\backslash\mathbb{H}$. Our main reference for this subsection is [47].

Fix Γ a cofinite Fuchsian group. We denote by $\mathcal{A}(\Gamma\backslash\mathbb{H})$ the space of *automorphic forms* with respect to Γ, that is functions $f : \mathbb{H} \to \mathbb{C}$ that satisfy

$$f(\gamma z) = f(z), \quad \text{for all } \gamma \in \Gamma.$$

Hence these are functions on the Riemann surface $\Gamma\backslash\mathbb{H}$. Let $\mathcal{L}(\Gamma\backslash\mathbb{H})$ denote the space of automorphic forms that are square-integrable. On $\mathcal{L}(\Gamma\backslash\mathbb{H})$ we define the inner product

$$\langle f, g\rangle := \int_{\Gamma\backslash\mathbb{H}} f(z)\overline{g(z)}d\mu. \tag{2.6}$$

A function $f \in \mathcal{A}(\Gamma\backslash\mathbb{H})$ that is an eigenfunction of the Laplace operator Δ is called a *Maaß form*. Denote by $\lambda(s) := s(1-s)$. Then we denote by $\mathcal{A}_s(\Gamma\backslash\mathbb{H})$ the space of Maaß forms with eigenvalue $\lambda(s)$, i.e. automorphic forms which satisfy

$$\Delta f(z) = \lambda(s) f(z).$$

One of the goals of this section is to show the decomposition of $\mathcal{L}(\Gamma\backslash\mathbb{H})$ in terms of eigenfunctions of Δ. It is useful to define

$$\mathcal{B}(\Gamma\backslash\mathbb{H}) := \{f \in \mathcal{L}(\Gamma\backslash\mathbb{H}) \ : \ f \text{ and } \Delta f \text{ smooth and bounded}\}.$$

Then Δ acts on $\mathcal{B}(\Gamma\backslash\mathbb{H})$ and $\mathcal{B}(\Gamma\backslash\mathbb{H})$ is dense in $\mathcal{L}(\Gamma\backslash\mathbb{H})$.

2.2.2.1 Eisenstein Series

Our first example of a Maaß form is the (nonholomorphic) Eisenstein series. Fix $\mathfrak{a}$ a cusp of Γ. For $\mathrm{Re}(s) > 1$, we define

$$E_{\mathfrak{a}}(z, s) := \sum_{\gamma\in\Gamma_{\mathfrak{a}}\backslash\Gamma} (\mathrm{Im}\,\sigma_{\mathfrak{a}}^{-1}\gamma z)^s.$$

Then $E_{\mathfrak{a}}(z, s)$ is an eigenfunction of Δ with eigenvalue $\lambda(s)$, but it is not square-integrable. If $\mathfrak{a}$ and $\mathfrak{b}$ are cusps for Γ, we have the Fourier expansion [47, p. 66]

$$E_{\mathfrak{a}}(\sigma_{\mathfrak{b}} z, s) = \delta_{\mathfrak{a}\mathfrak{b}} y^s + \phi_{\mathfrak{a}\mathfrak{b}}(s) y^{1-s} + \sum_{n\neq 0} y^{1/2}\phi_{\mathfrak{a}\mathfrak{b}}(n, s) K_{s-1/2}(2\pi|n|y)e(nx), \tag{2.7}$$

where

$$\begin{aligned} \phi_{\mathfrak{a}\mathfrak{b}}(s) &= \pi^{1/2}\frac{\Gamma(s-1/2)}{\Gamma(s)} L_{\mathfrak{a}\mathfrak{b}}(s, 0, 0),\\ \phi_{\mathfrak{a}\mathfrak{b}}(n, s) &= \pi^s \Gamma(s)^{-1}|n|^{s-1/2} L_{\mathfrak{a}\mathfrak{b}}(s, 0, n), \end{aligned} \tag{2.8}$$

where the series $L_{\mathfrak{a}\mathfrak{b}}(s, m, n)$ is defined in (2.5) and $K_s(y)$ is the K-Bessel function as defined in [47, Appendix B.4].

The Eisenstein series $E_{\mathfrak{a}}(z,s)$ admits meromorphic continuation to $s \in \mathbb{C}$ and functional equation, see [47, Chapter 6]. There are finitely many poles in $\mathrm{Re}(s) \geq 1/2$ and they are all simple and belong to the real segment $(1/2, 1]$. If s_j is such a pole, we denote the residue

$$u_{\mathfrak{a},s_j}(z) := \mathrm{Res}_{s=s_j} E_{\mathfrak{a}}(z,s).$$

Then $u_{\mathfrak{a},s_j}(z)$ is a Maaß form and moreover it is square-integrable. The residue functions play an important role in the spectral decomposition of $\mathcal{L}(\Gamma\backslash\mathbb{H})$. In Chapter 3, we use the analytic properties of $u_{\mathfrak{a},s_j}$, where we work with an additional continuous family of twists coming from modular symbols.

The point $s = 1$ is a simple pole of $E_{\mathfrak{a}}(z,s)$ and the residue is the constant function

$$u_{\mathfrak{a},1}(z) = \frac{1}{\mathrm{vol}(\Gamma)}, \quad \text{for all cusps } \mathfrak{a}.$$

The *Selberg Eigenvalue Conjecture* predicts that for congruence subgroups $\Gamma_0(N)$, the first non-zero eigenvalue of Δ acting on $\mathcal{L}(\Gamma_0(N)\backslash\mathbb{H})$ is at least $1/4$. By the above, this would imply that $s = 1$ is the only pole of $E_{\mathfrak{a}}(z,s)$ in the region $\mathrm{Re}(s) \geq 1/2$. However, there exist non-congruence subgroups Γ such that Δ has eigenvalues arbitrarily close to 0, see [47, p. 182].

We denote by $\mathcal{R}_{s_j}(\Gamma\backslash\mathbb{H})$ the space spanned by the residues of all Eisenstein series at $s = s_j$, hence the dimension of $\mathcal{R}_{s_j}(\Gamma\backslash\mathbb{H})$ is at most the number of inequivalent cusps of Γ. Also, we denote by $\mathcal{R}(\Gamma\backslash\mathbb{H})$ the subspace spanned by all residues s_j in the interval $1/2 < s_j \leq 1$, which is called the *residual spectrum.* Therefore we have the orthogonal decomposition

$$\mathcal{R}(\Gamma\backslash\mathbb{H}) = \bigoplus_{1/2<s_j\leq 1} \mathcal{R}_{s_j}(\Gamma\backslash\mathbb{H}).$$

We now briefly mention the functional equation for the Eisenstein series. Let $\mathfrak{a}_1, \dots, \mathfrak{a}_h$ be representatives for the inequivalent cusps of Γ. Denote by $\mathcal{E}(z,s)$ the column vector of Eisenstein series $E_{\mathfrak{a}_j}(z,s)$ and the $h \times h$ scattering matrix $\Phi(s) = (\phi_{\mathfrak{a}\mathfrak{b}}(s))$, where $\phi_{\mathfrak{a}\mathfrak{b}}(s)$ apper in the 0-Fourier coefficients (2.8). Then we have

$$\mathcal{E}(z,s) = \Phi(s)\mathcal{E}(z, 1-s). \tag{2.9}$$

Moreover, the scattering matrix satisfies

$$\Phi(s)\Phi(1-s) = \mathrm{Id}.$$

We now introduce the *incomplete Eisenstein series.* Let ψ be a smooth function and compactly supported on $\mathbb{R}^+$. We then define

$$E_{\mathfrak{a}}(z|\psi) := \sum_{\gamma\in\Gamma_{\mathfrak{a}}\backslash\Gamma} \psi(\operatorname{Im}\sigma_{\mathfrak{a}}^{-1}\gamma z). \tag{2.10}$$

Then $E_{\mathfrak{a}}(z|\psi)$ is a bounded automorphic function, and hence $E_{\mathfrak{a}}(z|\psi) \in \mathcal{L}(\Gamma\backslash\mathbb{H})$. However, it is not a Maaß form since it is not an eigenfunction of Δ. We denote by $\mathcal{E}(\Gamma\backslash\mathbb{H})$ the space of incomplete Eisenstein series.

We can represent $E_{\mathfrak{a}}(z|\psi)$ as a contour integral of the standard Eisenstein series. Let $\Psi(s)$ be the Mellin transform of ψ given by

$$\Psi(s) = \int_0^\infty \psi(y)y^{s-1}dy.$$

Then, for $\sigma > 1$, we have

$$E_{\mathfrak{a}}(z|\psi) = \frac{1}{2\pi i}\int_{(\sigma)} E_{\mathfrak{a}}(z,s)\Psi(-s)ds.$$

We see that the Laplacian Δ acts on $\mathcal{E}(\Gamma\backslash\mathbb{H})$. We can decompose $\mathcal{E}(\Gamma\backslash\mathbb{H})$ in terms of residual spectrum $\mathcal{R}(\Gamma\backslash\mathbb{H})$ and the Eisenstein series at $\mathrm{Re}(s) = 1/2$. Let $\{u_j\}$ be an orthonormal basis for $\mathcal{R}(\Gamma\backslash\mathbb{H})$. We quote [47, Theorem 7.3]. Every $f \in \mathcal{E}(\Gamma\backslash\mathbb{H})$ has the expansion

$$f(z) = \sum_j \langle f, u_j\rangle\, u_j(z) + \sum_{\mathfrak{a}} \frac{1}{4\pi}\int_{-\infty}^{\infty} \langle f, E_{\mathfrak{a}}(\cdot, 1/2+ir)\rangle\, E_{\mathfrak{a}}(\cdot, 1/2+ir)dr. \tag{2.11}$$

2.2.2.2 Cusp forms

We note that if $f \in \mathcal{L}(\Gamma\backslash\mathbb{H})$, for each cusp $\mathfrak{a}$, we have a Fourier expansion

$$f(\sigma_{\mathfrak{a}}z) = f_{\mathfrak{a}}(y) + \sum_{n\neq 0} \hat{f}_{\mathfrak{a}n}(y)e(nx).$$

We let $\mathcal{C}(\Gamma\backslash\mathbb{H})$ the subspace of $\mathcal{B}(\Gamma\backslash\mathbb{H})$ of automorphic forms that vanish at all cusps, i.e.

$$\mathcal{C}(\Gamma\backslash\mathbb{H}) = \{f \in \mathcal{B}(\Gamma\backslash\mathbb{H}) \ : \ \hat{f}_{\mathfrak{a}} = 0, \quad \text{for all cusps } \mathfrak{a}\}.$$

Elements of $\mathcal{C}(\Gamma\backslash\mathbb{H})$ are called *cusp forms.* Then Δ acts on $\mathcal{C}(\Gamma\backslash\mathbb{H})$ and we have the orthogonal decomposition

$$\mathcal{B}(\Gamma\backslash\mathbb{H}) = \mathcal{E}(\Gamma\backslash\mathbb{H}) \bigoplus \mathcal{C}(\Gamma\backslash\mathbb{H}). \tag{2.12}$$

The Laplacian Δ has pure point spectrum on $\mathcal{C}(\Gamma\backslash\mathbb{H})$, see [47, Chapter 4]. Therefore this space is spanned by eigenfunctions of Δ, which are called *Maaß cusp forms.* Hence there exists an orthonormal basis $\{u_j\}$ of $\mathcal{C}(\Gamma\backslash\mathbb{H})$ composed of Maaß cusp forms such that, for all $f \in \mathcal{C}(\Gamma\backslash\mathbb{H})$, we have

$$f(z) = \sum_j \langle f, u_j\rangle \, u_j(z).$$

Using (2.12) and the results from the previous subsection, we obtain the harmonic spectral decomposition of $\mathcal{B}(\Gamma\backslash\mathbb{H})$.

If φ is a Maaß cusp form with eigenvalue $1/4 + r^2 = \lambda(1/2 + ir)$, then we have the Fourier expansion of the type

$$\varphi(z) = \sqrt{y}\sum_{n\neq 0} \rho(n) K_{ir}(2\pi|n|y)e(nx), \tag{2.13}$$

where $\rho(n)$ are the Fourier coefficients, see [11, p. 106]. The theory of Hecke operators for Maaß cusp forms is very similar to that of holomorphic forms, see [28, Section 6]. If φ is a Maaß form which is also a Hecke eigenform for all Hecke operators T_n, when we call φ a *Hecke–Maaß eigen cusp form.* Its Fourier coefficients are given in terms of Hecke eigenvalues $\lambda(n)$:

$$\rho(n) = \rho((-1)^{\operatorname{sgn}(n)})\lambda(|n|)|n|^{-1/2}.$$

As in the holomorphic case, for congruence groups $\Gamma_0(N)$, there exists an orthonormal basis of $\mathcal{C}(\Gamma_0(N)\backslash\mathbb{H})$ composed of Hecke–Maaß cusp forms.

Remark 2.2.1. In Chapter 5 we work with a slightly more general *automorphic forms of* *weight k*. These are functions $f : \mathbb{H} \to \mathbb{C}$ which transform by

$$f(\gamma z) = j_\gamma(z)^k f(z), \quad \text{for all } \gamma \in \Gamma, \tag{2.14}$$

where $j_\gamma(z) = \dfrac{cz+d}{|cz+d|}$ with $\gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix}$. We denote by $\mathcal{L}_k(\Gamma\backslash\mathbb{H})$ the space of square-integrable weight k automorphic forms. In Section 5.2 we explain in detail the spectral decomposition of $\mathcal{L}_k(\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H})$ in terms of *Maaß forms of weight k*, which are eigenfunctions of the Laplcian of weight k given by

$$\Delta_k = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - ik \frac{\partial}{\partial x}.$$

2.3 L-functions

2.3.1 Definitions and basic properties

We begin by reviewing some properties of L-functions. Our main references for this section are [48, Chapter 5], [11, Chapter 1] and [46, Chapter 7].

Definition 2.3.1. *A meromorphic function $L(s,f)$ is called an L-function if it satisfies the following properties.*

- *$L(s,f)$ admits a Dirichlet series with an Euler product of degree d*

  $$L(s,f) = \sum_{n\geq 1} \frac{a_f(n)}{n^s} = \prod_p \prod_{j=1}^{d} \left(1 - \frac{\alpha_j(p)}{p^s}\right)^{-1},$$

  which is absolutely convergent for $\mathrm{Re}(s) > 1$*.*

- *We write*

  $$L_\infty(s,f) = N^{s/2} \prod_{j=1}^{m} \Gamma_\mathbb{R}(s+\mu_j),$$

  where $\Gamma_\mathbb{R}(s) := \pi^{-s/2}\Gamma(s/2)$, $N = N(f) \in \mathbb{Z}^+$ denotes the conductor and $\mu_j \in \mathbb{C}$ are some parameters with $\mathrm{Re}(\mu_j) > 1$*. Then the completed L-function*

  $$\Lambda(s,f) := L(s,f)L_\infty(s,f)$$

  admits analytic continuation to $s \in \mathbb{C}$ with poles at most at $s = 0$ and $s = 1$.

Moreover, it satisfies the functional equation

$$\Lambda(s,f) = \kappa\Lambda(1-s,f),$$

where $\kappa = \kappa(f)$ is a complex number of absolute value 1 (the root number).

If $L(s,f)$ is a L-function, then we define the *analytic conductors* $C(f)$ and $C(f,s)$

$$C(f) = N\prod_{j=1}^{d}(1+|\mu_j|) \quad \text{and} \quad C(s,f) = N\prod_{j=1}^{d}(1+|\mu_j+s|). \tag{2.15}$$

We say that the L-function $L(s,f)$ satisfy the Ramanujan–Petersson conjecture if $|\alpha_i(p)| = 1$ for $p|N$ and $|\alpha_i(p)| \leq 1$ otherwise. As we will see later, this assumption has important consequences. However, the Ramanjuan–Petersson conjecture was only shown to hold for some particular families of L-functions, even though we expect it to hold for all L-functions coming from automorphic representations as part of Langlands philosophy.

A central theme in analytic number theory is the study of special values of L-functions and the analytic properties of $L(s,f)$. The most important conjecture in this direction is the *Grand Riemann Hypothesis* (GRH for short), a conjectural statement about the zeros of L-functions with many important implications [48, Chapter 5.9]. It states that all zeros of an L-function in the critical strip $0 < \mathrm{Re}(s) < 1$ are on the critical line $\mathrm{Re}(s) = 1/2$.

We next highlight some examples of L-functions. We begin with the *Riemann zeta function* $\zeta(s)$ given by

$$\zeta(s) = \sum_{n\geq 1}\frac{1}{n^s} = \prod_p \frac{1}{1-p^{-s}},$$

for $\mathrm{Re}(s) > 1$. It is closely related to the Riemann xi function

$$\xi(s) = \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s),$$

which is entire for all $s \in \mathbb{C}$ and has functional equation $\xi(s) = \xi(1-s)$.

Our next examples are the *Dirichlet L-functions* $L(s,\chi)$. Let χ be a primitive character modulo q and define

$$L(s,\chi) = \sum_{n\geq 1}\frac{\chi(n)}{n^s} = \prod_p \frac{1}{1-\chi(p)p^{-s}}$$

for $\mathrm{Re}(s) > 1$. Then $L(s,\chi)$ is an L-function of degree 1, conductor q and gamma factors

$$L_\infty(s,\chi) = q^{s/2}\Gamma_{\mathbb{R}}(s+\delta),$$

where $\delta = 0$ if $\chi(-1) = 1$ and $\delta = 1$ if $\chi(-1) = -1$. Its root number is $\kappa(\chi) = \tau(\chi)/\sqrt{q}$, where $\tau(\chi)$ is Gauss sum associated to the character χ. The L-functions $L(s,\chi)$ correpond to automorphic representations of GL_1 over $\mathbb{Q}$.

We next move to L-functions arising from cuspidal representations of GL_2 over $\mathbb{Q}$, which correspond to holomorphic cusp or or Maaß cusp forms. Let $f \in S_k(\Gamma_0(N))$ be a holomorphic cusp form of weight k and level N. Let $f(z) = \sum_{n\geq 1} a_f(n)e(nz)$ be the Fourier expansion at infinity. We normalise by writing

$$a_f(n) = a_f(1)\lambda_f(n)n^{(k-1)/2}.$$

Thus, if f is a Hecke cusp form, then $\lambda_f(n)$ are the Hecke eigenvalues. For $\mathrm{Re}(s) > 1$, we define

$$L(s,f) = \sum_{n\geq 1}\frac{\lambda_f(n)}{n^s} = \prod_p \left(1 - \frac{\alpha_f(p)}{p^s}\right)^{-1}\left(1 - \frac{\beta_f(p)}{p^s}\right)^{-1}.$$

Then $L(s,f)$ is an L-function of degree 2 and conductor N. By definition, we have the factorisation of Hecke polynomials

$$1 - \lambda_f(p)p^{-s} + p^{-2s} = \left(1 - \alpha_f(p)p^{-s}\right)\left(1 - \beta_f(p)p^{-s}\right).$$

By work of Deligne [25], we know that $|\alpha_f(p)| = |\beta_f(p)| = 1$, for all primes p not dividing N, hence $L(s,f)$ satisfies the Ramanujan–Petersson conjecture. The gamma factors of $L(s,f)$ are given by

$$L_\infty(f,s) = N^{s/2}\Gamma_{\mathbb{R}}(s+(k-1)/2)\Gamma_{\mathbb{R}}(s+(k+1)/2).$$

This implies that

$$C(f) = N\frac{k+1}{2}\frac{k+3}{2} \asymp Nk^2.$$

Similarly, let φ be a Maaß form of level N which is an eigenfunction of the Laplacian with eigenvalue $\lambda(1/2+ir) = 1/4+r^2$, for some $r \in \mathbb{R}$. As in (2.13), we write the Fourier expansion at ∞ as

$$\varphi(z) = \sqrt{y} \sum_{n \neq 0} \rho(n) K_{ir}(2\pi |n| y) e(nx)$$

and we associate the L-function

$$L(s, \varphi) := \sum_{n \geq 1} \frac{\rho(n)}{n^s} = \prod_p \left(1 - \rho(p) p^{-s} + p^{-2s}\right)$$

of degree 2 and conductor N. Gamma factors are given by

$$L_\infty(s, \varphi) = N^{s/2} \Gamma_{\mathbb{R}}(s + \delta + ir) \Gamma_{\mathbb{R}}(s + \delta - ir),$$

where $\delta = 0$ if ϕ even and $\delta = 1$ if ϕ odd. Therefore the analytic conductor is

$$C(\varphi) = N(1 + \delta + |r|)^2 \asymp Nr^2.$$

In contrast to holomorphic cusp forms, it is not known that $L(s, \varphi)$ satisfy the Ramanujan–Petersson conjecture. The best known bound is due to Kim and Sarnak [55], who prove

$$|\alpha_\varphi(p)|, |\beta_\varphi(p)| \leq p^{7/64}.$$

2.3.2 Rankin–Selberg convolution

For simplicity of exposition, we work with $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. The theory can be generalised to all congruence groups, however extra care needs to be taken for local factors at primes dividing the level.

Let f and g cusp forms, holomorphic or Maaß. We then have the Rankin–Selberg convolution $L(s, f \times g)$ given by the the following Euler product for $\mathrm{Re}(s) > 1$:

$$\prod_p \left(1 - \frac{\alpha_f(p)\alpha_g(p)}{p^s}\right)^{-1} \left(1 - \frac{\alpha_f(p)\beta_g(p)}{p^s}\right)^{-1} \left(1 - \frac{\beta_f(p)\alpha_g(p)}{p^s}\right)^{-1} \left(1 - \frac{\beta_f(p)\beta_g(p)}{p^s}\right)^{-1}.$$

It admits analytic continuation to all $s \in \mathbb{C}$ and it has a simple pole at $s = 1$ if and only if $f = g$. We can check that, for $\mathrm{Re}(s) > 1$, we have

$$L(s, f \times g) = \zeta(2s) \sum_{n \geq 1} \frac{\lambda_f(n)\lambda_g(n)}{n^s}.$$

Now assume f and g are holomorphic cups forms of weights k_1 and k_2 respectively, with $k_1 \leq k_2$. Then $L_\infty(s, f\times g)$ is given by

$$\Gamma_{\mathbb{R}}\left(s+\frac{k_1+k_2}{2}\right)\Gamma_{\mathbb{R}}\left(s+\frac{k_1+k_2}{2}-1\right)\Gamma_{\mathbb{R}}\left(s+\frac{k_2-k_1}{2}\right)\Gamma_{\mathbb{R}}\left(s+\frac{k_2-k_1}{2}+1\right). \tag{2.16}$$

This implies that

$$C(f\times g)\asymp (k_1+k_2)^2(1+k_2-k_1)^2. \tag{2.17}$$

We refer to [48, p. 133] for expressions of $L_\infty(s, f\times g)$ in the case that f and g are both Maaß forms or they are mixed holomorphic and Maaß.

If f, g are holomorphic cusp forms of weight k, we have the very useful integral representation

$$(4\pi)^{1-s-k}\Gamma(s+k-1)a_f(1)\overline{a_g(1)}\frac{L(s, f\times g, s)}{\zeta(2s)} = \int_{\Gamma\backslash\mathbb{H}} y^k f(z)\overline{g(z)}E(z,s)d\mu, \tag{2.18}$$

which follows from the now classical Rankin–Selberg unfolding method. We can obtain a similar statement for Maaß forms, also see Proposition 5.3.1 for a statement about mixed weights.

When $f=g$, we define the symmetric square L-function

$$L(s, \mathrm{sym}^2 f) := \prod_p \left(1-\frac{\alpha_f(p)^2}{p^s}\right)^{-1}\left(1-\frac{\beta_f(p)^2}{p^s}\right)^{-1}\left(1-\frac{1}{p^s}\right)^{-1} = \frac{1}{\zeta(s)}L(f\times f, s)\ .$$

The Gamma factors are

$$L_\infty(s, \mathrm{sym}^2 f) = \Gamma_{\mathbb{R}}(s+1)\Gamma_{\mathbb{R}}(s+k-1)\Gamma_{\mathbb{R}}(s+k),$$

if f is holomorphic of weight k and

$$L_\infty(s, \mathrm{sym}^2 f) = \Gamma_{\mathbb{R}}(s+1)\Gamma(s+2ir)\Gamma(s+2ir),$$

if f is Maaß form with eigenvalue $1/4+r^2$. As a consequence of (2.18), we see that for $f\in S_k(\Gamma)$, we have [48, (5.101)]

$$|a_f(1)|^2 = \frac{\zeta(2)(4\pi)^k\langle f, f\rangle}{\Gamma(k)L(1, \mathrm{sym}^2 f)\mathrm{vol}(\Gamma)}\ . \tag{2.19}$$

In particular, this implies that $L(1, \mathrm{sym}^2 f) > 0$. The size of $L(1, \mathrm{sym}^2 f)$ plays an important role in Chapter 5 in relation to the QUE for holomorphic cusp forms.

Similar to the definition of $L(s, f \times g)$, we can define the *triple convolution L-function* $L(s, \phi_1 \times \phi_2 \times \phi_3)$, where ϕ_i are holomorphic/Maaß cusp forms. This will be a degree 8 L-function. If $\phi_i \in S_{k_i}(\Gamma)$ are L^2-normalised Hecke cusp forms such that $k_1 + k_2 = k_3$, then Watson [110] obtains the beautiful integral representation formula

$$\int_{\Gamma\backslash\mathbb{H}} y^{k_3}\phi_1(z)\phi_2(z)\overline{\phi_3(z)}d\mu = \frac{1}{8}\frac{\Lambda(1/2, \phi_1 \times \phi_2 \times \phi_3)}{\Lambda(1, \mathrm{sym}^2\phi_1)\Lambda(1, \mathrm{sym}^2\phi_2)\Lambda(1, \mathrm{sym}^2\phi_3)}. \tag{2.20}$$

Ichino [45] has generalised this to general automorphic representation of GL_2 over number fields. In Section 5.3 we will use an equivalent form of (2.20) for mixed Maaß forms and holomorphic cusp forms of different weights.

2.3.3 Subconvexity results

It is of great interest to obtains bounds for the growth of L-functions $L(s, f)$ on the critical line $\mathrm{Re}(s) = 1/2$. We begin by mentioning the *convexity bound* [48, (5.21)]

$$L\left(\frac{1}{2} + it, f\right) \ll_\epsilon C(1/2 + it, f)^{\frac{1}{4}+\epsilon},$$

which follows from the Phragmén–Lindelöf convexity principle. Any result of the form

$$L\left(\frac{1}{2} + it, f\right) \ll C(1/2 + it, f)^{\frac{1}{4}-\delta},$$

for some $\delta > 0$, is called a *subconvexity* result. Subconvexity of L-functions has many implications for equidistribution and QUE type problems, see for example [27]. We note that, assuming that the Riemann hypothesis holds for $L(s, f)$, we have

$$L\left(\frac{1}{2} + it, f\right) \ll_\epsilon C(1/2 + it, f)^{\epsilon}. \tag{2.21}$$

The conjecture (2.21) about $L(s, f)$ is called the *Lindelöf hypothesis.*

The first subconvexity result is by Hardy–Littlewood–Weyl for the Riemann zeta function, see [48, p. 204]. They show that

$$\zeta\left(\frac{1}{2} + it\right) \ll_\epsilon (1 + |t|)^{\frac{1}{6}+\epsilon}.$$

The current best result is by Bourgain [7], who improved the exponent to 53/342. The literature on subconvexity results is vast and fast expanding, see for example [28], [69], [81] for GL_1 and GL_2 results and [58], [71] for results in more general settings.

In particular, we mention the following result of Soundararajan [102]. If we assume that the Ramanujan–Petersson conjecture holds for $L(s, f)$, then

$$L\left(\frac{1}{2}, f\right) \ll_\epsilon \frac{C(f)^{1/4}}{(\log(C(f)))^{1-\epsilon}}.$$

This is called a *weak subconvexity* result, since there is no power saving in the exponent, only a logarithmic saving. However, this is enough to apply it towards QUE for holomorphic cusp forms, as seen in [44]. In Chapter 5, this weak subconvexity result plays an important role in our work for dissipation of masses of holomorphic cusp forms.

2.4 Modular symbols

Modular symbols are certain periods of weight 2 cusp forms introduced by Birch and Manin and they are an indispensable tool for studying (twisted) L-functions of holomorphic cusp forms [62], [64] and for computing modular forms [20], study elliptic curves and homology and cohomology of arithmetic groups. For instance, see [103] for explicit computations of the space $S_k(\Gamma_0(N))$ using modular symbols.

2.4.1 Construction of modular symbols

We refer to [103], [20, Chapter 2], [22] for detailed descriptions on the constructions and properties of modular symbols. We first observe that if $z, w \in \mathbb{H}$ are Γ-equivalent, then the family of smooth paths from z to w in $\mathbb{H}$ determines a unique homology class in $H_1(\Gamma\backslash\mathbb{H}, \mathbb{Z})$. In fact, the class depends only on γ and we have the surjective map

$$\Phi : \Gamma \to H_1(\Gamma\backslash\mathbb{H}, \mathbb{Z}), \quad \gamma \mapsto \{\infty, \gamma\infty\},$$

which induces the canonical isomorphism

$$H_1(\Gamma\backslash\mathbb{H}, \mathbb{Z}) \cong \Gamma/[\Gamma, \Gamma] \ .$$

As in [89], we denote by $H^1_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C})$ the space of *cuspidal 1-forms*, that is differererentials on $\Gamma\backslash\mathbb{H}$ that vanish on cusps. We have the Eichler–Shimura isomorphism

$$S_2(\Gamma)\otimes\overline{S_2(\Gamma)}\to H^1_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C}),\quad (f,\overline{g})\mapsto f(z)dz+\overline{g(z)dz}.$$

Hence any 1-form $\omega=f(z)dz+\overline{g(z)dz}$ is composed of the 'holomorphic' part $f(z)dz$ and 'antiholomorphic' part $\overline{g(z)dz}$. However, if we are interested in harmonic differentials, then we have $\overline{f(\overline{z})}=g(z)$. In view of this relation, not much information is lost by ignoring the antiholomorphic part and we can use in practice just the holomorphic part $f(z)dz$.

We note that any cuspidal form is cohomologous to a form of compact support, i.e. if α is a cusp form, there exists $\widetilde{\alpha}\in H^1_c(\Gamma\backslash\mathbb{H},\mathbb{C})$ such that

$$\int_{\Phi(\gamma)}\alpha=\int_{\Phi(\gamma)}\widetilde{\alpha},\quad\text{for all }\gamma\in\Gamma\ ,$$

and we have the isomorphism

$$H^1_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C})\cong H^1_c(\Gamma\backslash\mathbb{H},\mathbb{C}),$$

we refer to [77, Proposition 2.1] for a detailed construction of this isomorphism. We have the exact Poincaré pairing between homology and cohomology

$$H_1(\Gamma\backslash\mathbb{H},\mathbb{C})\times H^1_c(\Gamma\backslash\mathbb{H},\mathbb{C})\to\mathbb{C},\quad (C,\alpha)\mapsto\int_C\alpha\ .$$

Moreover, every class in $H^1_c(\Gamma\backslash\mathbb{H},\mathbb{C})$ has exactly one harmonic representative, see [89, p. p. 8]. Therefore, if we denote by $\text{HAR}_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C})$ the space of harmonic cuspidals 1-forms on $\Gamma\backslash\mathbb{H}$, we have

$$H^1_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C})\cong H^1_c(\Gamma\backslash\mathbb{H},\mathbb{C})\cong\text{HAR}_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C}).$$

With this in mind, for $\gamma\in\Gamma$ and $\alpha\in H^1_{\text{cusp}}(\Gamma\backslash\mathbb{H},\mathbb{C})$, we define the *modular symbol* $\langle\gamma,\alpha\rangle$ as

$$\langle\gamma,\alpha\rangle:=\int_{\Phi(\gamma)}\alpha=\int_z^{\gamma z}\alpha,\tag{2.22}$$

for any $z \in \mathbb{H}^*$. From this definition, we can easily see that, for $\gamma_1, \gamma_2 \in \Gamma$,

$$\langle \gamma_1\gamma_2, \alpha \rangle = \int_z^{\gamma_1\gamma_2 z} \alpha = \int_P^{\gamma_2 P} \alpha + \int_{\gamma_2 P}^{\gamma_1\gamma_2 P} \alpha = \langle \gamma_1, \alpha \rangle + \langle \gamma_2, \alpha \rangle \ .$$

We note that if α is a cuspidal form, then for any parabolic $\gamma \in \Gamma$,

$$\langle \gamma, \alpha \rangle = \int_z^{\gamma z} \alpha = 0 \ .$$

In particular, $\langle \gamma, \alpha \rangle = 0$, for all $\gamma \in \Gamma_{\mathfrak{a}}$, for all cusps $\mathfrak{a}$. The additive property of modular symbols means that we can view modular symbols as elements of $H^1_{\mathrm{cusp}}(\Gamma, \mathbb{C})$.

When α is a real-valued 1-form, for real parameters ϵ, we can define the family of unitary characters

$$\chi_\epsilon : \Gamma \to S^1, \quad \gamma \mapsto \exp(2\pi i \epsilon \langle \gamma, \alpha \rangle).$$

One of the key ideas in Chapter 3 is to use these characters to twist the Laplacian or the Eisenstein series, and then use perturbation theory for small ϵ, in order to deduce results about the distribution of modular symbols.

As observed in [89], all characters of Γ that vanish on parabolics are of the form χ_ϵ. These characters can be viewed as elements of the cohomology group $H^1_{\mathrm{cusp}}(\Gamma, \mathbb{R}/\mathbb{Z})$. This point of view is useful for generalisations to higher dimensional hyperbolic spaces. In Chapter 4 we will study the distribution of characters in $H^1_{\mathrm{cusp}}(\Gamma, \mathbb{R}/\mathbb{Z})$, where Γ is a cofinite group acting on $\mathbb{H}^n$.

Returning to the 2-dimensional case, for $f \in S_2(\Gamma)$, if we take $\alpha = f(z)dz$, we simply denote

$$\langle \gamma, f \rangle := \langle \gamma, \alpha \rangle \, .$$

We recall the definition (1.4) for modular symbols $\langle r \rangle$, where $r \in \mathbb{Q}$:

$$\langle r \rangle := \int_\infty^r f(z)dz,$$

where the path of integration connects the cusp at ∞ with the cusp $r \in \mathbb{Q}$. We remark that our definitions for modular symbols are closely related. Indeed, there exists a cusp $\mathfrak{a}$

such that r and $\mathfrak{a}$ are Γ-equivalent, i.e. $r = \gamma\sigma_{\mathfrak{a}}\infty$, for some $\gamma \in \Gamma$. Then

$$\langle r\rangle = \int_{\infty}^{\gamma\sigma_{\mathfrak{a}}\infty} f(z)dz = \int_{\infty}^{\sigma_{\mathfrak{a}}\infty} f(z)dz + \int_{\sigma_{\mathfrak{a}}\infty}^{\gamma\sigma_{\mathfrak{a}}\infty} f(z)dz = \int_{\infty}^{\sigma_{\mathfrak{a}}\infty} f(z)dz + \langle \gamma, f\rangle . \quad (2.23)$$

Therefore the definitions agree up to a shift given by a period integral depending only on the cusp $\mathfrak{a}$.

We denote the plus/minus modular symbols

$$\langle r\rangle^{+} = \frac{\langle r\rangle + \langle -r\rangle}{2} \in \mathbb{R}, \quad \langle r\rangle^{-} = \frac{\langle r\rangle + \langle -r\rangle}{2i} \in \mathbb{R}.$$

This corresponds to integrating the real/imaginary part of the 1-from $f(z)dz$, i.e.

$$\langle r\rangle^{+} = \int_{\infty}^{r} \mathrm{Re}(f(z)dz), \quad \langle r\rangle^{-} = \int_{\infty}^{r} \mathrm{Im}(f(z)dz).$$

The image of the map

$$\mathbb{Q} \to \mathbb{R}, \quad r \mapsto \langle r\rangle^{\pm},$$

is a lattice in $\mathbb{R}$ given by $\Omega_f^{\pm}\mathbb{Z}$. We say that $\Omega_f^{\pm}$ is the *real/imaginary period* of f. Hence, for $r \in \mathbb{Q}$ we can define the *normalised modular symbol*

$$[r]^{\pm} := \frac{\langle r\rangle^{\pm}}{\Omega_f^{\pm}} \in \mathbb{Z}.$$

In Chapter 4, we study the distribution of $[r]^{\pm}$ modulo primes. We will use the notation $\mathfrak{m}_f^{\pm}(r)$ instead of $[r]^{\pm}$ for normalised modular symbols.

We mention some basic properties of modular symbols $\langle r\rangle^{\pm}$, see [66, p. 5]:

(i) $\langle r+1\rangle^{\pm} = \langle r\rangle^{\pm} = \pm\langle -r\rangle^{\pm}$;

(ii) For any $\gamma \in \Gamma$, $\langle r\rangle^{\pm} = \langle \gamma r\rangle^{\pm} - \langle \gamma\infty\rangle^{\pm}$;

(iii) Let c and N positive integers, such that, if $n = \gcd(c, N)$ and $N = ne$, then e and n coprime. Let w_e the eigenvalue of the Atkin–Lehner involution W_e acting on $f \in S_2(\Gamma_0(N))$. Then if $a, d \in \mathbb{Z}$ such that $ade \equiv -1 \bmod c$, then

$$\langle a/c\rangle^{\pm} = -w_e\langle d/c\rangle^{\pm}.$$

Now let $f \in S_k(\Gamma)$ be a holomorphic cusp form of weight k with Fourier expansion

$$f(z) = \sum_{n\geq 1} a_f(n) n^{(k-1)/2} e^{2\pi i n z}.$$

For $r \in \mathbb{R}$, we define the *additive twist* by r of the L-function $L(s,f)$ to be

$$L(s,f,r) := \sum_{n\geq 1} \frac{a_f(n)e(nr)}{n^s}, \quad \text{for } \operatorname{Re}(s) > 1.$$

Let $r = a/c \in \mathbb{Q}$ and $d \in (\mathbb{Z}/c\mathbb{Z})^{\times}$ such that $ad \equiv 1 \bmod c$. We define the completed L-function

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

It admits meromorphic continuation to $s \in \mathbb{C}$ and functional equation:

$$\Lambda(s,f,a/c) = \Lambda(1-s,f,d/c).$$

From the definition of the completed L-function, for $f \in S_2(\Gamma)$, we can write

$$L(1/2,f,r) = 2\pi i \langle r \rangle. \tag{2.24}$$

Therefore modular symbols can be understood in terms of central values of additively twisted L-functions!

2.4.2 Arithmetic statistics of modular symbols

The study of of distribution properties of modular symbols was pioneered by Goldfeld in the 90's, inspired by a connection to Szpiro's conjecture (which relates the conductor and the discriminant of elliptic curves). In his work [37], [38], he introduced the following Eisenstein series twisted by modular symbols, known today as the *Goldfeld Eisenstein series*. For $f, g \in S_2(\Gamma)$ and $m, n \in \mathbb{Z}$, they are defined as

$$E_{\mathfrak{a}}^{m,n}(z,s) := \sum_{\gamma \in \Gamma_{\mathfrak{a}} \backslash \Gamma} \langle \gamma, f \rangle^m \overline{\langle \gamma, g \rangle}^n \operatorname{Im}(\sigma_{\mathfrak{a}}^{-1}\gamma z)^s.$$

In his thesis, O'Sullivan [75] proved the meromorphic continuation and the functional equation for $E_{\mathfrak{a}}^{m,n}(z,s)$. Petridis and Risager [77] used the analytic properties of the

Goldfeld Eisenstein series to prove that modular symbols are asymptotically normally distributed with respect to a certain arithmetic ordering.

In 2016, Mazur, Rubin and Stein proposed the study of the arithmetic distribution of modular symbols for congruence subgroups in order to study the excess rank of elliptic curves over abelian extensions. More specifically, let E an elliptic curve over $\mathbb{Q}$ of conductor N and $F/\mathbb{Q}$ a finite abelian extension. Assuming the Birch–Swinnerton-Dyer conjecture, we have

$$\operatorname{rank}_{\mathbb{Z}} E(F) = \operatorname{rank}_{\mathbb{Z}} E(\mathbb{Q}) + \sum_{\substack{\chi:\mathrm{Gal}(F/\mathbb{Q})\to\mathbb{C}^* \\ \chi\neq 1}} \operatorname{ord}_{s=1} L(E,\chi,s),$$

see [66, p. 6]. This implies that the study of excess rank is closely related to the vanishing properties of the central values of the twisted Hasse–Weil L-function $L(E,\chi,1)$. These values are related to modular symbols by the Birch–Stevens formula, see [67, p. 10]. If χ is a primitive Dirichlet character of conductor c, then

$$\frac{\tau(\chi)L(E,\overline{\chi},1)}{\Omega^{\epsilon}} = \sum_{a\in(\mathbb{Z}/m\mathbb{Z})^*} \chi(a)\langle a/c\rangle^{\epsilon},$$

where $\epsilon := \chi(-1)$ is the sign of the character χ and $\tau(\chi)$ is the corresponding Gauss sum. This motivated Mazur and Rubin [66] to formulate a number of conjectures about the distribution of modular symbols $\langle a/c\rangle^{\pm}$ with fixed denominator c. These conjectures have received a lot of attention in recent years, see the work of Petridis–Risager [79], Bettin–Drappeau [5], Blomer et al. [6, Chapter 9], Diamantis et al. [26], Lee–Sun [57], Sun [104], Nordentoft [74].

Let $f \in S_2(\Gamma_0(q))$. We define the usual mean and variance for fixed level c for plus modular symbols:

$$\mathbb{E}(f,c) = \frac{1}{\phi(c)} \sum_{\substack{a \bmod c \\ (a,c)=1}} \langle a/c\rangle^{+}, \quad \mathrm{Var}(f,c) = \frac{1}{\phi(c)} \sum_{\substack{a \bmod c \\ (a,c)=1}} \left(\langle a/c\rangle^{+} - \mathbb{E}(f,c)\right)^2 .$$

We now state the conjectures of Mazur and Rubin about the behaviour of the asymptotic normal distribution of modular symbols and the behaviour of the variance.

Conjecture 2.4.1. *(i) There exists a constant C_f such that the limiting distribution of the data*

$$\left\{ \frac{\langle a/c\rangle^{+}}{(C_f \log c)^{1/2}} \;:\; (c,q)=d,\; a\in(\mathbb{Z}/c\mathbb{Z})^* \right\}$$

is the standard normal distribution as $c \to \infty$.

(ii) For each divisor d *of* q*, there is a constant* $D_{f,d}$ *such that*

$$\lim_{\substack{c\to\infty \\ (c,q)=d}} (\mathrm{Var}(f,c) - C_f \log c) = D_{f,d}.$$

The constant C_f is called the *variance slope*, while $D_{f,d}$ is the *variance shift.*

Petridis and Risager [79] obtain an average version of Conjecture 2.4.1 for square-free level q, where they consider the distribution of modular symbols on the larger set

$$R_d(X) = \left\{ \frac{a}{c} \mid 0 < a < c \leq X, (a,c) = 1, (c,q) = d \right\}.$$

We recall Theorem 1.2.1 from the Introduction.

Theorem 2.4.1. *[79] Let* $f \in S_2(\Gamma_0(q))$*, where* q *is a square-free integer.*

(i) The values

$$\left\{ \frac{\langle r \rangle^+}{(C_f \log X)^{1/2}} \ : \ r \in R_d(X) \right\}$$

have asymptotically a standard normal distribution as $X \to \infty$.

(ii) As $X \to \infty$,

$$\frac{1}{\# R_d(X)} \sum_{r \in R_d(X)} (\langle r \rangle^+)^2 = C_f \log X + D_{f,d} + o(1).$$

Petridis and Risager work with the spectral theory of automorphic forms and Goldfeld Eisenstein series. They give explicit formulas for C_f and $D_{f,d}$ and obtain results for any cofinite Fuchsian group Γ. Using dynamics of continued fractions, Lee–Sun [57] give an alternative proof of Theorem 1.2.1, building upon work of Baladi–Vallée [2] on dynamical methods. Their method is however restricted to $\Gamma_0(N)$ and does not give a formula for the variance.

In Chapter 3, we develop a new method to obtain distribution results for modular symbols and recover Theorem 1.2.1. While still making use of the spectral theory of Eisenstein series as in the work of Petridis–Risager, we apply the perturbation theory on character varieties to obtain significantly easier proofs. Also, instead of using the method of moments for proving convergence in distribution, we make use of the moment generator function and the Berry–Esseen inequality to obtain the limiting distribution with almost optimal error terms. Furthermore, our approach can naturally recover the first and second moments of the distribution and has the advantage that it can be naturally extended to modular symbols in $\mathbb{H}^3$ and higher dimensions.

We remark that Conjecture 2.4.1(i) seems to be very hard and out of reach of both spectral and dynamical methods. One reason is that the size of the family of modular symbols in the conjecture is small, around square-root the size of the family in Theorem 1.2.1. A version of Conjecture 2.4.1(ii) with c ranging over primes was achieved by Blomer, Fouvry, Kowalski, Michel, Milićević and Sawin in [6, Theorem 9.2], using very deep algebraic geometric methods.

These results have been generalised to higher weight modular forms. Using spectral methods, Nordentoft [74] obtains normal distribution for central values of additively twisted L-functions associated to cusp forms of general weight k and level N. In addition, Bettin–Drappeau [5] use dynamical methods to obtain results for general weight k for $\Gamma = \Gamma_0(1)$, and additional results for the distribution of quantum modular forms.

We also mention a conjecture of Mazur, Rubin and Stein about the partial first moment of modular symbols. They conjectured that if $0 < x < 1$ and $f(z) = \sum_{n=1}^{\infty} a_n q^n$, then

$$\lim_{c\to\infty} \frac{1}{c} \sum_{a=1}^{cx} [a/c]^+ = \sum_{n=1}^{\infty} \frac{a_n \sin(\pi n x)}{n^2}.$$

This result was proved by Diamantis, Hoffstein, Kıral and Lee [26] (for arbitrary level N and with explicit rate of convergence) and by Sun [104] for square-free level using dynamical methods.

Remark 2.4.2. In [66], Mazur and Rubin also proposed the study of θ-coefficients. These are given by sums of modular symbols of fixed denominator, where the numerators are induced by elements fixed in cyclic Galois extensions of conductor m. Let ζ_m a primitive m-th root of unity and $\sigma_a \in \mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$ given by $\sigma_a(\zeta_m) = \zeta_m^a$. Let $F \subset \mathbb{Q}(\zeta_m)$ and $\delta \in \mathrm{Gal}(F/\mathbb{Q})$. Then the *$\theta$-coefficients* are given by

$$\theta^{\pm}_{F,\delta} := \sum_{\substack{a\in(\mathbb{Z}/m\mathbb{Z})^* \\ \sigma_a|_F=\delta}} [a/m]^{\pm}$$

Mazur and Rubin are interested in the distribution of the values

$$\{\theta_{F,\delta} \mid F/\mathbb{Q} \text{ cyclic of degree } d \text{ and conductor } m, \delta \in \mathrm{Gal}(F/\mathbb{Q}) \text{ generic}\}.$$

Numerical analysis show that these values do not obey a normal distribution and very little progress was achieved in this direction.

Mazur and Rubin [65] also formulated a conjecture about the distribution of the normalised modular symbols $[r]^+$. They suggested that $[r]^+$ equidistribute modulo p as we vary along fractions with fixed denominator.

Conjecture 2.4.2. *Let p be a prime. The values $\{[a/c]^+ \mid a \in (\mathbb{Z}/c\mathbb{Z})^*\}$ equidistribute modulo p as $c \to \infty$, that is, for any $l \in (\mathbb{Z}/p\mathbb{Z})$,*

$$\frac{\#\{a \in (\mathbb{Z}/c\mathbb{Z})^* \mid [a/c]^+ \equiv l \bmod p\}}{\phi(c)} = \frac{1}{p} + o(1) \quad \textit{as } c \to \infty.$$

Recently, an average version of this conjecture was settled by Lee and Sun [57, Theorem I] using dynamical methods. In Chapter 4 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 [57], we obtain an average version of the mod p conjecture of Mazur and Rubin (and its generalisations), but with further refinements. We refer to Section 1.2.2 in the Introduction or Section 4.1 in Chapter 4 for statements of our results regarding equidistribution modulo p of modular symbols.

2.5 Quantum Unique Ergodicity

Mass equidistribution of eigenfunctions is a central topic in quantum chaos and number theory. We refer to [91] for an excellent survey on physical interpretations of QUE and developements for hyperbolic arithmetic manifolds. As a starting point, let M be a compact negatively curved Riemannian manifold. Then the geodesic flow is ergodic on its contangent bundle. It is of interest the study the behaviour of masses of eigenfunctions ϕ_j of the Laplace–Beltrami operator Δ as the eigenvalue tends to infinity. The following thoerem was proved by Šnirelman [108] and Colin de Verdière [16].

Theorem 2.5.1 (Quantum Ergodicity)**.** *Let M be a compact negatively curved Riemannian manifold with the standard measure μ and the Laplace–Beltrami operator Δ. Let $\{\phi_j\}$ an orthonormal basis of $L^2(M)$ formed of eigenfunctions of Δ with corresponding* *eigenvalues* λ_j. *Then*

$$|\phi_{j_k}|^2\mu \to \mu$$

in weak- sense as* $k \to \infty$, *for some full density subsequence* j_k.

We turn our attention to the arithmetic modular surface $X = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. If we ignore for the moment the continuous spectrum, we restrict our attention the the eigenfuncions of the Laplacian are the Hecke–Maaß cusp forms. Then the Quantum Ergodicity Conjecture was proved in this case by Zelditch [113].

Motivated by this, Rudnick and Sarnak [87] conjectured that mass equidistribution should hold for the full sequence of Laplace eigenfunctions. Lindenstrauss [59] famously proved this conjecture, with a key input from Soundararajan [101].

Theorem 2.5.2 (Quantum Unique Ergodicity). *Fix any* $\psi \in C_b(X)$. *Then for any sequence of Hecke–Maaß cusp forms* ϕ_j *of eigenvalues* λ_j, *we have that*

$$\left\langle \psi, |\phi_j|^2 \right\rangle \to \frac{3}{\pi} \left\langle \psi, 1 \right\rangle \quad \text{as } \lambda_j \to \infty.$$

The literature contains many examples of analogues of this problem. In [61], Luo–Sarnak prove Quantum Unique Ergodicity for the Eisenstein series. They show that for any measurable sets A and B with $\mu(B) > 0$, we have

$$\lim_{t\to\infty} \frac{\int_A |E(z, 1/2 + it)|^2 d\mu}{\int_B |E(z, 1/2 + it)|^2 d\mu} = \frac{\mu(A)}{\mu(B)}.$$

Moreover, they actually compute the asymptotic growth explicitly

$$\int_A |E(z, 1/2 + it)|^2 d\mu \sim \frac{6}{\pi}\mu(A) \log t.$$

Holowinsky and Soundararajan [44], [43], [102] prove an analogue of the Quantum Unique Ergodicity for holomorphic cusp forms.

Theorem 2.5.3 (Holowinsky–Soundararajan). *Let* f *be a holomorphic Hecke cusp form of weight* k *that is* L^2*-normalised such that* $\int_X y^k |f(z)|^2 d\mu(z) = 1$ *and let* $F_k = y^{k/2} f(z)$. *Fix any* ϕ *smooth and bounded on* X. *Then we have*

$$\int_X y^k |f(z)|^2 \phi(z) \frac{dxdy}{y^2} \to \frac{3}{\pi} \int_X \phi(z) \frac{dxdy}{y^2} \quad \text{as} \quad k \to \infty;$$

equivalently, this can be rewritten as

$$\langle \phi F_k, F_k \rangle \to \frac{1}{\mathrm{vol}(X)} \langle \phi, 1 \rangle \quad as \quad k \to \infty.$$

Nelson generalised their results to congruence groups [72] and to compact surfaces [73]. Zelditch [114] and Jakobson [49], [50] looked at quantum ergodicity for the contangent bundle $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$.

Holowinky and Soundararajan build upon two independent papers. Using sieve theory, Holowinsky [43] obtains bounds for shifted convolution sums of multiplicative functions. His approach fails to succeed when $L(1, \mathrm{sym}^2 f)$ is unusually large. Soundararajan [102] uses a weak subconvexity result that fails when $L(1, \mathrm{sym}^2 f)$ is unusually small. Combining the two approaches gives the desired result.

In Chapter 5, we generalise these results to off-diagonal terms, where we consider two different eigencusp forms f and g of weights k_1 and k_2 respectively. We show that correlations dissipate as $\max(k_1, k_2) \to \infty$. We refer to Theorems 5.1.1 and 5.1.2 for full statements of our results. We obtain a result about joint distribution of masses in the context of QUE, a subject with interesting recent results, see for example the work of Brooks [8] on distribution of off-diagonal Eisenstein series $\langle \phi E(\cdot, r), E(\cdot, r'') \rangle$ or Brooks–Lindenstrauss [9] on joint quasimodes of the Laplacian.

Our new ingredient is to incorporate the spectral theory of weight k automorphic functions to the method of Holowinsky–Soundararajan. Denote by $\mathcal{L}_k(X)$ the space of square-integrable weight k automorphic forms. Using the spectral decomposition of $\mathcal{L}_{k_2-k_1}(X)$, it is enough to obtain bounds for $\left\langle \phi y^{\frac{k_1}{2}} f, y^{\frac{k_2}{2}} g \right\rangle$, where ϕ is a Maaß cusp form of weigh $k_2 - k_1$. We have two approaches, depending on the size of

$$S(f, g) := L(1, \mathrm{sym}^2 f) L(1, \mathrm{sym}^2 g). \tag{2.25}$$

Firstly, we can compute directly the inner products, using Rankin–Selberg unfolding for the Eisenstein series and Ichino's formula for the Maaß cusp form case, see Section 5.3. The formulas will involve central values of L-functions, to which we apply the weak subconvexity results of Soundararajan. This will win if $S(f, g)$ is large.

Alternatively, we can expand the inner products in terms of the Fourier expansions. We need bounds for the Fourier coefficients of weight k automorphic forms, which we compute in Section 5.4. This approach boils down to bounding shifted convolution sums, where we apply the results of Holowinsky, see Section 5.5. This will win if $S(f,g)$ is sufficiently small.

2.6 Geometry of hyperbolic upper half-space $\mathbb{H}^3$

We refer to [32, Chapters 1-2] for a valuable exposition of the geometry of the hyperbolic 3-space and of the groups acting on it. We define the three-dimensional hyperbolic space $\mathbb{H}^3$ as

$$\mathbb{H}^3 := \mathbb{C} \times (0,\infty) = \{(z,y) \mid z \in \mathbb{C}, y > 0\} = \{(x_1,x_2,y) \mid x_1,x_2 \in \mathbb{R},\ y > 0\}\ .$$

We denote the points in $\mathbb{H}^3$ by

$$P = (z,y) = z + yj, \quad \text{where } z = x_1 + ix_2,\ j = (0,0,1)\ .$$

We equip $\mathbb{H}^3$ with the hyperbolic metric coming from the line element:

$$ds^2 = \frac{dx_1^2 + dx_2^2 + dy^2}{y^2}\ . \tag{2.26}$$

The volume element is given by

$$dv = \frac{dx_1 dx_2 dy}{y^3}.$$

The hyperbolic Laplace–Beltrami operator is given by

$$\Delta = y^2\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial y^2}\right) - y\frac{\partial}{\partial y}\ . \tag{2.27}$$

The group $\mathrm{PSL}_2(\mathbb{C})$ acts on $\mathbb{H}^3$ via isometries. The action of $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}_2(\mathbb{C})$ is given by

$$(z,y) \mapsto \left(\frac{(az+b)\overline{(cz+d)} + a\bar{c}y^2}{|cz+d|^2 + |c|^2y^2}\ ,\ \frac{y}{|cz+d|^2 + |c|^2y^2}\right)\ . \tag{2.28}$$

Let $\Gamma \leq \mathrm{PSL}_2(\mathbb{C})$ be any cofinite Kleinian group with cusps. The theory of such objects is thoroughly developed in [32, Chapter 2]. Let $\mathfrak{a} \in \mathbb{P}^1(\mathbb{C})$ be a cusp for Γ with scaling matrix $\sigma_{\mathfrak{a}} \in \mathrm{PSL}_2(\mathbb{C})$ 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

$$\Gamma'_{\mathfrak{a}} = \Gamma_{\mathfrak{a}} \cap \sigma_{\mathfrak{a}} \left\{ \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \ : \ b \in \mathbb{C} \right\} \sigma_{\mathfrak{a}}^{-1} .$$

We note that $\Gamma'_{\mathfrak{a}}$ consists of the parabolic elements in $\Gamma_{\mathfrak{a}}$ together with I.

There exists a lattice $\Lambda_{\mathfrak{a}} \leq \mathbb{C}$ 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 $|\mathcal{P}_{\mathfrak{a}}|$. We define $\Lambda_{\mathfrak{a}}^*$ the dual lattice of $\Lambda_{\mathfrak{a}}$:

$$\Lambda_{\mathfrak{a}}^* = \{ \mu \in \mathbb{C} \ : \ \langle \mu, \lambda \rangle \in \mathbb{Z} \text{ for all } \lambda \in \Lambda_{\mathfrak{a}} \} , \tag{2.29}$$

where $\langle \cdot, \cdot \rangle$ is the usual scalar product on $\mathbb{R}^2 = \mathbb{C}$.

Since Γ is a Kleinian group, there exists a constant $c_{\mathfrak{ab}} > 0$ defined by

$$c_{\mathfrak{ab}} := \min \left\{ |c| \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1} \Gamma \sigma_{\mathfrak{b}}, \ c \neq 0 \right\} . \tag{2.30}$$

Say $\mathfrak{a}_1, \cdots, \mathfrak{a}_h \in \mathbb{P}^1(\mathbb{C})$ are representatives for the Γ-classes of cusps. For $Y > 0$, we define the cuspidal sectors

$$\mathcal{F}_{\mathfrak{a}_i}(Y) = \sigma_{\mathfrak{a}_i} \{ z + yj \ : \ z \in \mathcal{P}_{\mathfrak{a}_i} \ , \ y \geq Y \} .$$

Then for Y_0 large enough, there exists a fundamental domain $\mathcal{F}$ which we can write as a disjoint union

$$\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_{\mathfrak{a}_1}(Y_0) \cup \cdots \cup \mathcal{F}_{\mathfrak{a}_h}(Y_0) , \tag{2.31}$$

where $\mathcal{F}_0$ is a compact set.

We denote by $T_{\mathfrak{ab}}$ a system of representatives $\begin{pmatrix} * & * \\ c & * \end{pmatrix}$ of the double cosets in

$$\sigma_{\mathfrak{a}}^{-1} \Gamma'_{\mathfrak{a}} \sigma_{\mathfrak{a}} \backslash \sigma_{\mathfrak{a}}^{-1} \Gamma \sigma_{\mathfrak{b}} / \sigma_{\mathfrak{b}}^{-1} \Gamma'_{\mathfrak{b}} \sigma_{\mathfrak{b}}$$

with $c \neq 0$ and

$$T_{\mathfrak{ab}}(X) = \left\{ \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in T_{\mathfrak{ab}} \ : \ 0 < |c| \leq X \right\} .$$

Also, we define

$$R_{\mathfrak{ab}} := \left\{ \frac{a}{c} \bmod \mathcal{P}_{\mathfrak{a}} \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in T_{\mathfrak{ab}} \right\} .$$

Lemma 2.6.1. *The map*

$$\begin{aligned} T_{\mathfrak{ab}} &\to R_{\mathfrak{ab}} \\ \gamma &\mapsto \gamma\infty \quad \bmod \mathcal{P}_{\mathfrak{a}} \end{aligned}$$

is $[\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}]$*-to-one.*

Proof. We follow the lines of [79, Proposition 2.2] or [47, p. 50], where it is shown that the map is one-to-one in the two-dimensional case. Let $\gamma, \gamma' \in T_{\mathfrak{ab}}$ with

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \gamma' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$$

and $r = \gamma\infty$, $r' = \gamma'\infty$. We may assume $r, r' \in \mathcal{P}_{\mathfrak{a}}$. Then the matrix $\gamma'' = \gamma'^{-1}\gamma \in \sigma_{\mathfrak{b}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ has lower left entry $c'' = -ac' + a'c$.

If $c'' \neq 0$, then

$$|-r + r'| = \left|\frac{c''}{cc'}\right| > 0 .$$

Therefore $r \neq r'$, hence $r \not\equiv r' \bmod \mathcal{P}_{\mathfrak{a}}$.

If $c'' = 0$, then $r = r'$ and $\gamma'' \in (\sigma_{\mathfrak{b}}^{-1}\Gamma\sigma_{\mathfrak{b}})_\infty = \sigma_{\mathfrak{b}}^{-1}\Gamma_{\mathfrak{b}}\sigma_{\mathfrak{b}}$. Since we assume $\gamma, \gamma' \in T_{\mathfrak{ab}}$, there are $[\sigma_{\mathfrak{b}}^{-1}\Gamma_{\mathfrak{b}}\sigma_{\mathfrak{b}} : \sigma_{\mathfrak{b}}^{-1}\Gamma'_{\mathfrak{b}}\sigma_{\mathfrak{b}}] = [\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}]$ possible choices for γ''. □

2.7 Higher dimensional hyperbolic spaces

We introduce the upper half-space (Poincaré) model $\mathbb{H}^{n+1}$ for the $(n+1)$-dimensional hyperbolic space. 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. We make use of a specific model for the group of isometries given in terms of a certain Clifford algebra. Our main references for this section are [1], [30] and [31].

2.7.1 Clifford algebra

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 non-degenerate quadratic 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

$$e_i^2 = q(e_i), \quad e_i e_j = -e_j e_i, \quad \text{where} \quad i, j = 1, \cdots, n, \ i \neq j,$$

where $e_1, \dots, e_n$ is a q-orthonormal 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

$$e_M := e_{i_1} \cdot \dots \cdot e_{i_k}, \quad e_\emptyset := 1 \in \mathcal{C}(q).$$

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

$$\overline{e_M} := (-1)^{|M|(|M|+1)/2} e_M, \quad e_M^* := (-1)^{|M|(|M|-1)/2} e_M, \quad \text{where } M \in \mathcal{E}_n.$$

These satisfy

$$\overline{vw} = \overline{w}\ \overline{v}, \quad (vw)^* = w^* v^*, \quad \text{for all } v, w \in \mathcal{C}(q).$$

From now on we assume that $q = -I_n$, the negative definite unit form, and $e_1, \dots, e_n$ the standard basis. 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 $\{1, e_1, \dots, e_n\}$. 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 $\{1, e_1, \dots, e_n\}$.

For $x = \sum_{M \in \mathcal{E}_n} \lambda_M e_M \in \mathcal{C}_n$, we define the norm

$$|x| := \left(\sum_{M \in \mathcal{E}_n} \lambda_M^2 \right)^{1/2}. \tag{2.32}$$

We note that for $x \in V_n$, we have $|x|^2 = \langle x, x \rangle$. Now, if $\Lambda < V_n$ is a lattice, we define the dual lattice as

$$\Lambda^* := \{w \in V_n \mid \langle v, w\rangle \in \mathbb{Z} \text{ for all } v \in \Lambda\}.$$

We now define the following model of hyperbolic $(n+1)$-space:

$$\mathbb{H}^{n+1} := \{x_0 + x_1 e_1 + \cdots + x_n e_n \mid x_0, x_1, \ldots, x_{n-1} \in \mathbb{R}, x_n > 0\} .$$

We have the maps $x : \mathbb{H}^{n+1} \to V_{n-1}$ and $y : \mathbb{H}^{n+1} \to (0, \infty)$ given by

$$x(P) := x_0 + x_1 e_1 + \cdots + x_{n-1} e_{n-1}, \quad y(P) := x_n,$$

where $P = x_0 + x_1 e_1 + \cdots + x_n e_n \in \mathbb{H}^{n+1}$. We can think of $x(P)$ as an element of $\mathbb{R}^n$ via the above. Then from (2.32) 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:

$$ds^2 = \frac{dx_0^2 + dx_1^2 + \cdots + dx_n^2}{x_n^2} , \tag{2.33}$$

which makes $\mathbb{H}^{n+1}$ a Riemannian manifold with constant negative curvature -1. The volume element is given by

$$dv = \frac{dx_0 dx_1 \ldots x_n}{x_n^{n+1}}.$$

The *hyperbolic Laplace–Beltrami operator* is given by

$$\Delta = x_n^2 \left(\frac{\partial^2}{\partial x_0^2} + \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2} \right) - (n-1) x_n \frac{\partial}{\partial x_n} \tag{2.34}$$

in this model.

2.7.2 Vahlen group

We will use the above upper-half space model to describe the group of (oriented) isometries $\mathrm{Isom}^+(\mathbb{H}^{n+1})$ 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 [31, p. 648], we define the *Vahlen group* SV_n to be

$$\mathrm{SV}_n := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathcal{C}_n) \;\middle|\; \begin{array}{c} \text{(i) } a,b,c,d \in T_n \cup \{0\} \\ \text{(ii) } \overline{a}b, \overline{c}d \in V_n \\ \text{(iii) } ad^* - bc^* = 1 \end{array} \right\}. \tag{2.35}$$

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 it is a non-trivial fact that SV_n is a group under matrix multiplication with inverse

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d^* & -b^* \\ -c^* & a^* \end{pmatrix}. \tag{2.36}$$

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 2.7.1. *[31, Theorem 1.3] Let* $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SV}_{n-1}$ *and* $P \in \mathbb{H}^{n+1}$*. Then* $cP + d \in T_n$ *and we define*

$$\gamma P := (aP+b)(cP+d)^{-1} \in \mathbb{H}^{n+1}, \tag{2.37}$$

where multiplication and taking inverses is possible since we work with elements of the multiplicative group T_n*. 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* $\mathrm{SV}_{n-1}/\{I, -I\} \cong \mathrm{Isom}^+(\mathbb{H}^{n+1})$.

What is convenient about this description of $\mathrm{Isom}^+(\mathbb{H}^{n+1})$ 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 2.7.1. *[31, 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*

$$x(\gamma P) = \frac{(ax+b)(\overline{cx+d}) + a\overline{c}y^2}{|cx+d|^2 + |c|^2 y^2} \quad \text{and} \quad y(\gamma P) = \frac{y}{|cx+d|^2 + |c|^2 y^2}. \tag{2.38}$$

Remark 2.7.2. Our model for the hyperbolic $(n+1)$-space is consistent with other descriptions from the literature. For example, one can consider the Klein model $\mathbb{K}^{n+1}$ on which isometries are described by $\mathrm{SO}(n+1, 1)$. Then there exists an bijection $\Phi : \mathbb{H}^{n+1} \to \mathbb{K}^{n+1}$

and an isomorphism $\Psi : \mathrm{SV}_{n-1}/\{\pm I\} \xrightarrow{\sim} \mathrm{SO}(n+1,1)^0$ which commutes with the respective actions, i.e. $\Phi(\gamma \cdot P) = \Psi(\gamma)\Phi(P)$, for all $\gamma \in \mathrm{SV}_{n-1}$ and $P \in \mathbb{H}^{n+1}$. Here $\mathrm{SO}(n+1,1)^0$ is the connected component of the identity element in $\mathrm{SO}(n+1,1)$. We refer to [30, Section 5] for detailed descriptions of different models of the hyperbolic space.

2.7.3 Hyperbolic quotients

Let $\Gamma < \mathrm{SV}_{n-1}$ be a discrete subgroup of motions such that the surface $\Gamma\backslash\mathbb{H}^{n+1}$ has finite hyperbolic volume. We say that $\mathfrak{a} \in \mathbb{R}^n \cup \{\infty\}$ is a cusp for Γ if it is fixed by a non-identity element in Γ. Then there exists a scaling matrix $\sigma_{\mathfrak{a}} \in \mathrm{SV}_{n-1}$ such that $\sigma_{\mathfrak{a}}\infty = \mathfrak{a}$. We let $\Gamma_{\mathfrak{a}} := \{\gamma \in \Gamma \mid \gamma\mathfrak{a} = \mathfrak{a}\}$ 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

$$\sigma_{\mathfrak{a}}^{-1}\Gamma'_{\mathfrak{a}}\sigma_{\mathfrak{a}} = \left\{ \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \mid \lambda \in \Lambda_{\mathfrak{a}} \right\} \ .$$

We let $\mathcal{P}_{\mathfrak{a}}$ be a fundamental parallelogram for $\Lambda_{\mathfrak{a}}$ with Euclidean area $\mathrm{vol}(\Lambda_{\mathfrak{a}})$.

We define the *dual lattice* of $\Lambda_{\mathfrak{a}}^*$ as follows:

$$\Lambda_{\mathfrak{a}}^* := \{\mu \in \mathbb{R}^n \mid \langle \mu, \lambda \rangle \in \mathbb{Z} \text{ for all } \lambda \in \Lambda_{\mathfrak{a}}\} \ , \tag{2.39}$$

where $\langle \cdot, \cdot \rangle$ is the usual scalar product on $\mathbb{R}^n$.

For a cusp $\mathfrak{a}$ and $Y > 0$, we define the *cuspidal sector*

$$\mathcal{F}_{\mathfrak{a}}(Y) := \sigma_{\mathfrak{a}}\{(x,y) \mid x \in \mathcal{P}_{\mathfrak{a}}, y > Y\} \ .$$

Then for Y large enough, there exists a fundamental domain $\mathcal{F}$ for $\Gamma\backslash\mathbb{H}^{n+1}$ and inequivalent cusps $\mathfrak{a}_1, \cdots, \mathfrak{a}_h \in \mathbb{R}^n \cup \{\infty\}$ such that we can write $\mathcal{F}$ as the disjoint union

$$\mathcal{F} = \mathcal{F}_0 \sqcup \mathcal{F}_{\mathfrak{a}_1}(Y) \sqcup \cdots \sqcup \mathcal{F}_{\mathfrak{a}_h}(Y) \ , \tag{2.40}$$

where $\mathcal{F}_0$ is a compact set, see [100, p. 8] or [89, p. 5].

For notational convenience, from now on we will focus only on 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 (4.5) in precise terms. First we note that all elements in $\Gamma'_\infty \backslash \Gamma / \Gamma'_\infty$ share the same lower left entry. Thus it makes sense to define

$$T_\Gamma(X) := \left\{ \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in \Gamma'_\infty \backslash \Gamma / \Gamma'_\infty \mid 0 < |c| \leq X \right\},$$

where $|c|$ denotes the Clifford norm (2.32). This is the natural generalisation of the outcome space considered by Petridis–Risager in [79, p. 1002]. In (4.22) below, we provide an asymptotic formula for the size of $T_\Gamma(X)$. We put

$$C(\Gamma) := \left\{ c \in T_n \mid \exists a, b, d \in T_n : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \right\}. \tag{2.41}$$

If $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma$ then from the definition of the action (2.37), we see that $\gamma\infty = ac^{-1}$, where $\gamma\infty$ is defined as the limit of γP as P tends to the cusp at ∞. Also, from [31, Lemma 1.4], we know that $ac^{-1} \in V_{n-1}$.

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 $X \to \infty$.

Chapter 3

Distribution of modular symbols in $\mathbb{H}^3$

This chapter is mainly based on [18].

3.1 Introduction

Let K be a quadratic number field of class number one, $\mathcal{O}_K$ its ring of integers and $\mathfrak{n}$ a non-zero ideal of $\mathcal{O}_K$. In a series of papers [22], [23], [21], Cremona uses modular symbols to study the arithmetic correspondence between isogeny classes of elliptic curves defined over K of conductor $\mathfrak{n}$ and Hecke cusp forms of weight 2 for the congruence subgroup $\Gamma_0(\mathfrak{n})$. More precisely, the Hasse–Weil L-function $L(E,s)$ of an elliptic curve E and the L-function $L(F,s)$ attached to a cusp form F are conjectured to be the same as part of the 'Langlands philosophy'. Modular symbols are given by central values $L(F,\psi,1)$, where ψ is an additive twist, and they can be used to compute numerically the central value $L(F,1)$, which agrees with the value $L(E,1)$ predicted by the Birch–Swinnerton-Dyer conjecture. We prove that when $\mathfrak{n}$ is a square-free ideal of $\mathcal{O}_K$ and F a newform of weight 2 and level $\mathfrak{n}$, modular symbols coming from F obey asymptotically the standard normal distribution when ordered and normalised appropriately.

We develop a new method to obtain distribution results for modular symbols. While still making use of the spectral theory of Eisenstein series as in the work of Petridis–Risager, we apply the perturbation theory on character varieties to obtain significantly easier proofs. Also, instead of using the method of moments for proving convergence in distribution, we make use of the moment generating function and the Berry–Esseen inequality to obtain the limiting distribution with almost optimal error terms. Furthermore, our approach can naturally recover the first and second moments of the distribution and has the advantage that it can be naturally extended to modular symbols in $\mathbb{H}^3$.

We now describe the set-up for the case of cofinite groups Γ of $\mathrm{PSL}_2(\mathbb{R})$, as in the work of Petridis–Risager. Let $\mathfrak{a}$ and $\mathfrak{b}$ be two cusps (not necessarily equivalent) with scaling matrices $\sigma_{\mathfrak{a}}$ and $\sigma_{\mathfrak{b}}$. We define general modular symbols as

$$\langle r\rangle_{\mathfrak{ab}} = \int_{\mathfrak{b}}^{\sigma_{\mathfrak{a}} r} \alpha,$$

where α is a harmonic 1-form and

$$r \in T_{\mathfrak{ab}}(X) = \left\{ \frac{a}{c} \bmod 1 \ , \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_\infty \backslash \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}/\Gamma_\infty \ , \ 0 < c < X \right\}.$$

Petridis–Risager obtain the following average results of Conjecture 2.4.1.

Theorem 3.1.1 (Petridis–Risager [79]). *There exist explicit constants C_f, $D_{f,\mathfrak{ab}}$ such that*

(a) (Normal distribution) The values of

$$T_{\mathfrak{ab}}(X) \to \mathbb{R}, \quad \frac{a}{c} \mapsto \frac{\langle a/c\rangle}{\sqrt{C_f \log c}}$$

have asymptotically a standard normal distribution as $X \to \infty$.

(b) (Second moment) As $X \to \infty$,

$$\frac{\sum_{r\in T_{\mathfrak{ab}}(X)} \langle r\rangle_{\mathfrak{ab}}^2}{\# T_{\mathfrak{ab}}(X)} = C_f \log X + D_{f,\mathfrak{ab}} + o(1).$$

Here is a statement for our results. There is a natural action of $\mathrm{PSL}_2(\mathbb{C})$ on $\mathbb{H}^3$ via isometries. Let $\Gamma \leq \mathrm{PSL}_2(\mathbb{C})$ be a cofinite discrete subgroup. For each cusp $\mathfrak{a}$, we denote by $\Gamma'_{\mathfrak{a}}$ the set of parabolic elements in Γ that fix $\mathfrak{a}$. Then there exists a lattice $\Lambda_{\mathfrak{a}} \leq \mathbb{C}$ 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 note that we require this extra notation since, unlike the two dimensional case, we only know that $\Gamma'_{\mathfrak{a}}$ is a subgroup of finite index of the stabilizer subgroup $\Gamma_{\mathfrak{a}}$ and that for two cusps $\mathfrak{a}$ and $\mathfrak{b}$, the period lattices $\Lambda_{\mathfrak{a}}$ and $\Lambda_{\mathfrak{b}}$ may be different.

Now, for $\mathfrak{a}, \mathfrak{b}$ two cusps for Γ (not necessarily distinct), we define

$$R_{\mathfrak{ab}}(X) = \left\{ r = \frac{a}{c} \bmod \Lambda_{\mathfrak{a}} \ , \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma'_{\mathfrak{a}}\sigma_{\mathfrak{a}} \backslash \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}/\sigma_{\mathfrak{b}}^{-1}\Gamma'_{\mathfrak{b}}\sigma_{\mathfrak{b}} \ , \ 0 < |c| < X \right\} .$$

We prove the following theorem.

Theorem 3.1.2. *Let α be a real-valued, Γ-invariant, cuspidal one-form.*

(a) (Normal distribution) For every $a, b \in [-\infty, \infty]$ with $a \leq b$, and any $\epsilon > 0$, for X large enough,

$$\frac{\#\left\{ r \in R_{\mathfrak{ab}}(X)\ ,\ \frac{\langle r \rangle_{\mathfrak{ab}}}{\sqrt{C_\alpha \log X}} \in [a, b] \right\}}{\# R_{\mathfrak{ab}}(X)} = \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{t^2}{2}\right) dt + O_\epsilon\left((\log X)^{-1/2+\epsilon}\right),$$

where

$$C_\alpha = \frac{4\|\alpha\|_2^2}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)}\ . \tag{3.1}$$

(b) (First moment) There exists a constant $\delta > 0$ such that

$$\frac{\sum_{r \in R_{\mathfrak{ab}}(X)} \langle r \rangle_{\mathfrak{ab}}}{\# R_{\mathfrak{ab}}(X)} = \int_{\mathfrak{b}}^{\mathfrak{a}} \alpha + O\left(X^{-\delta}\right)\ .$$

(c) (Second moment) There exists an explicit constant $D_{\alpha,\mathfrak{ab}}$, called the variance shift, and a constant $\delta > 0$ such that

$$\frac{\sum_{r \in R_{\mathfrak{ab}}(X)} \langle r \rangle_{\mathfrak{ab}}^2}{\# R_{\mathfrak{ab}}(X)} = C_\alpha \log X + D_{\alpha,\mathfrak{ab}} + O\left(X^{-\delta}\right), \quad \text{as } X \to \infty\ .$$

Remark 3.1.1. The error term in Theorem 3.1.2(a) is expected to be optimal up to ϵ, see [5]. It seems to be difficult to obtain a good error term using the method of moments approach.

Remark 3.1.2. Theorem 3.1.2(b) is a generalisation of [79, Cor. 7.3] with $x = 1$, where Petridis–Risager obtain stronger results about first moment with additional restrictions on the set $R_{\mathfrak{ab}}(X)$.

Remark 3.1.3. We do not obtain an explicit value for $D_{\alpha,\mathfrak{ab}}$, but we can write it in terms of the coefficients of a certain Taylor expansion, see (3.48) for more details. For the case of $\mathbb{H}^2$, the variance shift was explicitly calculated in [79].

We obtain the following corollary for imaginary quadratic number fields. Let K be a quadratic imaginary field of class number one and $\mathfrak{n}$ a square-free ideal. Let $F \in S_2(\Gamma(\mathfrak{n}))$ be a cuspidal newform of weight 2 and level $\mathfrak{n}$, which is a vector-valued function $F : \mathbb{H}^3 \to$

$\mathbb{C}^3$. For $r \in K$, we define the modular symbol

$$\langle r \rangle = \int_{i\infty}^{r} F \cdot \beta \in \mathbb{R} ,$$

where β is a specific fixed basis for the invariant 1-forms. We rigorously introduce these objects in Section 7.

Corollary 3.1.1. *Let K be a quadratic number field of class number one. Let $\mathfrak{n} \lhd \mathcal{O}_K$ a square-free ideal with generator $\langle n \rangle = \mathfrak{n}$ and $F \in S_2(\Gamma_0(\mathfrak{n}))$. For $\mathfrak{d}|\mathfrak{n}$, set*

$$Q_{\mathfrak{d}}(X) = \{a/c \mid a \in (\mathcal{O}_K/\langle c \rangle)^{\times} \ , \ \langle c, n \rangle = \mathfrak{d} \ , \ 0 < |c| < X\}.$$

(a) There exists a constant C_F such that the data

$$K \cap Q_{\mathfrak{d}}(X) \to \mathbb{R}, \quad \frac{a}{c} \mapsto \frac{\langle a/c \rangle}{\sqrt{C_F \log X}}$$

has asymptotically a standard normal distribution.

(b) There exists a constant $D_{F,\mathfrak{d}}$ such that

$$\frac{1}{|Q_{\mathfrak{d}}(X)|} \sum_{a/c \in Q_{\mathfrak{d}}(X)} \left\langle \frac{a}{c} \right\rangle^2 = C_F \log X + D_{F,\mathfrak{d}} + o(1) \ .$$

Remark 3.1.4. We provide explicit value for C_F in terms of the Petersson norm of F and our base quadratic imaginary field K, see (3.55).

The structure of this chapter is as follows. In Section 3.2 we introduce th elementary properties of modular symbols associated to cuspidal one-forms on $\Gamma\backslash\mathbb{H}^3$.

In Section 3.3 we study the Eisenstein series and Poincaré series twisted by modular symbols. We introduce the generating series $L_{\mathfrak{ab}}(s,\epsilon)$ and obtain some of their essential analytic properties. We also provide upper bounds for modular symbols.

In Section 3.4 we study the perturbation theory of the space $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$, where χ_ϵ is a unitary character given by modular symbols. We obtain Taylor expansions for the smallest eigenvalue of the Laplacian $\lambda_0(\epsilon)$ and for $s_0(\epsilon)$, the first pole of $L_{\mathfrak{ab}}(s,\epsilon)$. Moreover, we study the behaviour of the residue of $L_{\mathfrak{ab}}(s,\epsilon)$ at $s_0(\epsilon)$.

In Section 3.5 we relate the moment generating function for the distribution of modular symbols to our generating series $L_{\mathfrak{ab}}(s,\epsilon)$. We recover the first two moments of the distribution. In addition, we show that $R_{\mathfrak{ab}}$ is equidistributed in the period lattice $\Lambda_{\mathfrak{a}}$.

In Section 3.6 we prove that modular symbols are normally distributed. We use the Berry–Esseen inequality and the perturbation theory results developed earlier.

In Section 3.7 we obtain results for congruence subgroups of $\mathrm{PSL}_2(\mathcal{O}_K)$, where K is a quadratic imaginary number field of class number one. We relate modular symbols to special values of L-functions coming from newforms of weight 2 and level $\mathfrak{n}$, where $\mathfrak{n}$ is a square-free ideal of $\mathcal{O}_K$. We develop some properties of these L-functions.

3.2 Modular symbols in $\mathbb{H}^3$

We refer to Section 2.6 for an introduction to the properties of the quotient space $\Gamma\backslash\mathbb{H}^3$, where $\Gamma < \mathrm{SL}_2(\mathbb{C})$ a cofinite subgroup with cusps. We denote by $\mathbb{H}^* := \mathbb{H}^3 \cup \mathbb{C} \cup \{\infty\}$ the extended upper half-space and consider the compactified quotient space $X_\Gamma = \Gamma\backslash\mathbb{H}^*$. If $A, B \in \mathbb{H}^*$ are Γ-equivalent, i.e. there exists some $\gamma \in \Gamma$ such that $B = \gamma(A)$, then the family of smooth paths from A to B in $\mathbb{H}^*$ determines a unique homology class in $H_1(X_\Gamma, \mathbb{Z})$. In fact, the class depends only on γ and we have the surjective map

$$\Phi : \Gamma \to H_1(X_\Gamma, \mathbb{Z}), \quad \gamma \mapsto \{\infty, \gamma\infty\}$$

which induces the canonical isomorphism

$$H_1(X_\Gamma, \mathbb{Z}) \cong \Gamma/[\Gamma, \Gamma] \ .$$

We consider the de Rham cohomology group $H^1_{\mathrm{dR}}(X_\Gamma, \mathbb{C})$ and inside of it we have $H^1_c(X_\Gamma, \mathbb{C})$ consisting of cohomology classes represented by forms of compact support. Every member of $H^1_c(X_\Gamma, \mathbb{C})$ has a unique harmonic representative. We provide a sketch argument showing that $H_1(X_\Gamma, \mathbb{C})$ and $H^1_c(X_\Gamma, \mathbb{C})$ are dual to each other.

Note that in general X_Γ may not be a manifold, since Γ may contain elements of finite order (X_Γ is called an orbifold). However, it is a result of Selberg [94, p. 482] that if $\Gamma < \mathrm{GL}_n(\mathbb{C})$ is a finitely generated subgroup, then Γ has a torsion free subgroup Γ' of finite index. Then $X_{\Gamma'}$ is a manifold and the finite quotient group $\bar{\Gamma} := \Gamma/\Gamma'$ acts on it. We have the exact Poincaré pairing between homology and cohomology for $X_{\Gamma'}$

$$H_1(X_{\Gamma'}, \mathbb{C}) \times H^1_c(X_{\Gamma'}, \mathbb{C}) \to \mathbb{C}, \quad (C, \alpha) \mapsto \int_C \alpha \ .$$

In this duality, if we restrict to forms invariant under $\bar{\Gamma}$, we recover $H^1_c(X_\Gamma, \mathbb{C})$ and can show that there is also an exact duality between $H_1(X_\Gamma, \mathbb{C})$ and $H^1_c(X_\Gamma, \mathbb{C})$. For more details, see [22, p. 43].

Definition 3.2.1. *A harmonic 1-form* $\alpha = f_1 dx_1 + f_2 dx_2 + f_3 dy$ *on* $\Gamma\backslash\mathbb{H}^3$ *is a cuspidal 1-form if*

1. *α is rapidly decreasing at all cusps;*
2. *for each cusp $\mathfrak{a}$ and $y \geq 0$,*

$$\int_{\mathcal{P}_\mathfrak{a}} f_{\mathfrak{a},i} dx_1 dx_2 = 0 \ , \quad i = 1, 2, 3 \ ,$$

where $\sigma_\mathfrak{a}^* \alpha = f_{\mathfrak{a},1} dx_1 + f_{\mathfrak{a},2} dx_2 + f_{\mathfrak{a},3} dy$ is the pullback by $\sigma_\mathfrak{a}$.

As in [89], we denote the space of cuspidal 1-forms by $H^1_{\text{cusp}}(X_\Gamma, \mathbb{C})$. We note that any cuspidal form is cohomologous to a form of compact support, i.e. if α is a cusp form, there exists $\widetilde{\alpha} \in H^1_c(X_\Gamma, \mathbb{C})$ such that

$$\int_{\Phi(\gamma)} \alpha = \int_{\Phi(\gamma)} \widetilde{\alpha}, \quad \text{for all } \gamma \in \Gamma \ ,$$

and we have the isomorphism

$$H^1_{\text{cusp}}(X_\Gamma, \mathbb{C}) \simeq H^1_c(X_\Gamma, \mathbb{C}) \ .$$

A detailed construction of the above isomorphism can be found in [77, Proposition 2.1]. With this in mind, for $\gamma \in \Gamma$ and $\alpha \in H^1_{\text{cusp}}(X_\Gamma, \mathbb{C})$, we define the modular symbol $\langle \gamma, \alpha \rangle$ as

$$\langle \gamma, \alpha \rangle := \int_{\Phi(\gamma)} \alpha = \int_{P_0}^{\gamma P_0} \alpha \tag{3.2}$$

for any $P_0 \in \mathbb{H}^*$. From this definition, we can easily see that, for $\gamma_1, \gamma_2 \in \Gamma$,

$$\langle \gamma_1 \gamma_2, \alpha \rangle = \int_P^{\gamma_1 \gamma_2 P} \alpha = \int_P^{\gamma_2 P} \alpha + \int_{\gamma_2 P}^{\gamma_1 \gamma_2 P} \alpha = \langle \gamma_1, \alpha \rangle + \langle \gamma_2, \alpha \rangle \ .$$

We note that if α is a cuspidal form, then for any parabolic $\gamma \in \Gamma$,

$$\langle \gamma, \alpha \rangle = \int_{P_0}^{\gamma P_0} \alpha = 0 \ .$$

In particular, $\langle \gamma, \alpha \rangle = 0$, for all $\gamma \in \Gamma'_{\mathfrak{a}}$, for all cusps $\mathfrak{a}$.

We remark that our definition for the modular symbol $\langle \gamma, \alpha \rangle$ agrees with the previous definition $\langle r \rangle_{\mathfrak{a}\mathfrak{b}}$. Indeed, if $\gamma \in \sigma_{\mathfrak{a}}^{-1} \Gamma \sigma_{\mathfrak{b}}$ with $r = \gamma\infty$, then

$$\langle r \rangle_{\mathfrak{a}\mathfrak{b}} = \int_{\mathfrak{b}}^{\sigma_{\mathfrak{a}}\gamma\infty} \alpha = \int_{\mathfrak{b}}^{\sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1}\mathfrak{b}} \alpha = \left\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1}, \alpha \right\rangle \ . \tag{3.3}$$

If $\alpha \in H^1_{\text{cusp}}(X_\Gamma, \mathbb{C})$ is real-valued, we have a family of unitary characters $\chi_\epsilon : \Gamma \to S^1$ defined by

$$\chi_\epsilon(\gamma) := \exp\left(2\pi i \epsilon \langle \gamma, \alpha \rangle\right) \ . \tag{3.4}$$

If $\alpha, \beta \in H^1_{\text{cusp}}(X_\Gamma, \mathbb{C})$ with $\alpha = f_1 dx_1 + f_2 dx_2 + f_3 dy$ and $\beta = g_1 dx_1 + g_2 dx_2 + g_3 dy$, we define the pointwise inner-product

$$[\alpha, \beta] := y^2 (f_1 \overline{g_1} + f_2 \overline{g_2} + f_3 \overline{g_3}) \ . \tag{3.5}$$

Since α and β are Γ-invariant 1-forms, one can see that $[\alpha, \beta]$ is a Γ-invariant function from $\mathbb{H}^3$ to $\mathbb{C}$. In particular, since α is rapid decreasing in the cusps, we conclude that $[\alpha, \alpha]$ is bounded on $\mathbb{H}^3$, which in turn implies that

$$|f_i(P)| \ll \frac{1}{y}, \quad \text{for all } P \in \mathbb{H}^3, i = 1, 2, 3. \tag{3.6}$$

Now, for $\alpha, \beta \in H^1_{\text{cusp}}(X_\Gamma, \mathbb{C})$, we define the Petersson inner product

$$\langle \alpha, \beta \rangle := \int_{\Gamma \backslash \mathbb{H}^3} [\alpha, \beta] dv \ , \tag{3.7}$$

and the L^2-norm

$$\|\alpha\|_2^2 := \langle \alpha, \alpha \rangle \ . \tag{3.8}$$

3.3 Generating series for modular symbols

In this section we define a generating series for modular symbols $L_{\mathfrak{a}\mathfrak{b}}(s, \epsilon)$. This we relate to the twisted Eisenstein series and Poincaré series by characters and derive some of their essential analytic properties.

3.3.1 Twisted Eisenstein series by modular symbols

We define the twisted Eisenstein series

$$E_{\mathfrak{a}}(P,s,\epsilon)=\sum_{\gamma\in\Gamma'_{\mathfrak{a}}\backslash\Gamma}\overline{\chi_\epsilon(\gamma)}y(\sigma_{\mathfrak{a}}^{-1}\gamma P)^s\ , \tag{3.9}$$

where χ_ϵ is defined as in (3.4).

The theory of Eisenstein series in $\mathbb{H}^3$ (without a twist) is developed in [32, Chapter 3] and [32, chapter 6.1]. We have to modify it slightly since we consider twisted Eisenstein series, so we follow the steps in Selberg's Göttingen lecture notes [94, p. 638-654]. They are absolutely convergent for $\mathrm{Re}(s)>2$. In the area of absolute convergence they satisfy

$$E_{\mathfrak{a}}(\gamma P,s,\epsilon)=\chi_\epsilon(\gamma)E_{\mathfrak{a}}(P,s,\epsilon)\ ,$$
$$-\Delta E_{\mathfrak{a}}(P,s,\epsilon)=s(2-s)E_{\mathfrak{a}}(P,s,\epsilon)\ .$$

We note that the function $P\mapsto E_{\mathfrak{a}}(\sigma_{\mathfrak{b}}P,s,\epsilon)$ is invariant under the action of the lattice $\Lambda_{\mathfrak{b}}$ corresponding to $\sigma_{\mathfrak{b}}^{-1}\Gamma'_{\mathfrak{b}}\sigma_{\mathfrak{b}}=(\sigma_{\mathfrak{b}}^{-1}\Gamma\sigma_{\mathfrak{b}})'_\infty$. We would like to write a Fourier expansion with respect to the dual lattice $\Lambda_{\mathfrak{b}}^*$. With this in mind, for $\mu_1\in\Lambda_{\mathfrak{a}}^*$, $\mu_2\in\Lambda_{\mathfrak{b}}^*$, we define the twisted generating series by

$$L_{\mathfrak{a}\mathfrak{b}}(s,\mu_1,\mu_2,\epsilon):=\sum_{\gamma\in T_{\mathfrak{a}\mathfrak{b}}}\frac{\overline{\chi_\epsilon(\sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1})}e\left(\left\langle\mu_1,\frac{a}{c}\right\rangle+\left\langle\mu_2,\frac{d}{c}\right\rangle\right)}{|c|^{2s}}\ , \tag{3.10}$$

where the sum is over $\gamma=\begin{pmatrix}a&b\\c&d\end{pmatrix}\in T_{\mathfrak{a}\mathfrak{b}}$. If $\mu_1=\mu_2=0$, we just denote $L_{\mathfrak{a}\mathfrak{b}}(s,0,0,\epsilon)=:$ $L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon)$.

We quote [32, Theorem 3.4.1] to obtain Fourier expansion of $E_{\mathfrak{a}}(\sigma_{\mathfrak{b}}P,s,\epsilon)$:

$$E_{\mathfrak{a}}(\sigma_{\mathfrak{b}}P,s,\epsilon)=\delta_{\mathfrak{a}\mathfrak{b}}[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]y^s+\phi_{\mathfrak{a}\mathfrak{b}}(s,\epsilon)y^{2-s}+\sum_{0\neq\mu\in\Lambda_{\mathfrak{b}}^*}|\mu|^{s-1}\phi_{\mathfrak{a}\mathfrak{b}}(s,\mu,\epsilon)\ y\ K_{s-1}(2\pi|\mu|y)\ e(\langle\mu,z\rangle)\ , \tag{3.11}$$

where

$$\phi_{\mathfrak{a}\mathfrak{b}}(s,\epsilon):=\frac{\pi}{|\mathcal{P}_{\mathfrak{b}}|(s-1)}L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon),\quad \phi_{\mathfrak{a}\mathfrak{b}}(s,\mu,\epsilon):=\frac{2\pi^s}{|\mathcal{P}_{\mathfrak{b}}|\Gamma(s)}L_{\mathfrak{a}\mathfrak{b}}(s,0,\mu,\epsilon) \tag{3.12}$$

and K denotes the K-Bessel function.

If $\mathfrak{a}_1, \cdots, \mathfrak{a}_h \in \mathbb{P}^1(\mathbb{C})$ are the inequivalent cusps for $\Gamma\backslash\mathbb{H}^3$, we define

$$E_i(P,s,\epsilon) := \frac{1}{[\Gamma_{\mathfrak{a}_i} : \Gamma'_{\mathfrak{a}_i}]} E_{\mathfrak{a}_i}(P,s,\epsilon) \quad \text{and} \quad \phi_{ij}(s,\epsilon) = \frac{1}{[\Gamma_{\mathfrak{a}_i} : \Gamma'_{\mathfrak{a}_i}]} \phi_{\mathfrak{a}_i\mathfrak{a}_j}(s,\epsilon) \ .$$

We let

$$\mathcal{E}(P,s,\epsilon) := \begin{pmatrix} E_1(P,s,\epsilon) \\ \vdots \\ E_h(P,s,\epsilon) \end{pmatrix} \quad \text{and} \quad \Phi(s,\epsilon) := (\phi_{ij}(s,\epsilon)) \ .$$

We call Φ the scattering matrix. Then both $\mathcal{E}(P,s,\epsilon)$ and $\Phi(s,\epsilon)$ have meromorphic continuation to all of $\mathbb{C}$. The following functional equation is satisfied:

$$\mathcal{E}(P,2-s,\epsilon) = \Phi(2-s,\epsilon)\mathcal{E}(P,s,\epsilon) \ .$$

Also, poles of $\mathcal{E}(P,s,\epsilon)$ occur only where $\Phi(s,\epsilon)$ has poles and vice versa. In the region $\operatorname{Re} s > 1$, there are only finitely many simple poles, and they are on the interval $1 < s \leq 2$ of the real line. If $1 < \sigma \leq 2$ is a pole of $E_{\mathfrak{a}}(P,s,\epsilon)$, we define

$$u_{\mathfrak{a},\sigma}(P,\epsilon) = \operatorname{Res}_{s=\sigma} E_{\mathfrak{a}}(P,s,\epsilon) \ . \tag{3.13}$$

We denote by $L^2(\Gamma\backslash\mathbb{H}^3,\chi_\epsilon)$ the space of all square-integrable functions f over $\Gamma\backslash\mathbb{H}^3$ that satisfy $f(\gamma P) = \chi_\epsilon(\gamma)f(P)$, for all $\gamma \in \Gamma$. Then $u_{\mathfrak{a},\sigma}(P,\epsilon) \in L^2(\Gamma\backslash\mathbb{H}^3,\chi_\epsilon)$ and moreover

$$(\Delta + \sigma(2-\sigma))u_{\mathfrak{a},\sigma}(\cdot,\epsilon) = 0 \ . \tag{3.14}$$

We study the spectral theory of $L^2(\Gamma\backslash\mathbb{H}^3,\chi_\epsilon)$ in Section 3.4.1. The spectrum of $-\Delta$ on $L^2(\Gamma\backslash\mathbb{H}^3,\chi_\epsilon)$ contains a finite number of discrete eigenvalues in $[0,1)$, call them $0 \leq \lambda_0(\epsilon) \leq \lambda_1(\epsilon) \leq \cdots \leq \lambda_k(\epsilon) < 1$. Then $E_{\mathfrak{a}}(P,s,\epsilon)$ is meromorphic for $\operatorname{Re}(s) > 1$ and has possible poles at $s_j(\epsilon)$ corresponding to $\lambda_j(\epsilon)$, so that $s_j(\epsilon)(2-s_j(\epsilon)) = \lambda_j(\epsilon)$.

3.3.2 Twisted Poincaré series by modular symbols

We now introduce the twisted Poincaré series, extending the definition of Sarnak in [88]. We will use them to obtain an integral representation for the series $L_{\mathfrak{ab}}(s,0,\mu,\epsilon)$ and to find the residue of $L_{\mathfrak{ab}}(s,0,\mu,0)$ at $s=2$.

For $\mu \in \Lambda_{\mathfrak{a}}^*$, we define

$$E_{\mathfrak{a},\mu}(P,s,\epsilon) := \sum_{\gamma \in \Gamma_{\mathfrak{a}}' \backslash \Gamma} \overline{\chi_\epsilon(\gamma)} y(\sigma_{\mathfrak{a}}^{-1}\gamma P)^s e^{-2\pi|\mu| y(\sigma_{\mathfrak{a}}^{-1}\gamma P)} e(\langle z(\sigma_{\mathfrak{a}}^{-1}\gamma P), \mu \rangle) \ . \tag{3.15}$$

We observe that for $\mathrm{Re}(s) > 2$, the series converges absolutely, since it is certainly dominated by the Eisenstein series. Also, since the function $y(\sigma_{\mathfrak{a}}^{-1}P)^s e^{-2\pi|\mu| y(\sigma_{\mathfrak{a}}^{-1}P)} e(\langle z(\sigma_{\mathfrak{a}}^{-1}P), \mu \rangle)$ is $\Gamma_{\mathfrak{a}}'$-invariant, it follows that $E_{\mathfrak{a},\mu}(P,s,\epsilon)$ satisfies

$$E_{\mathfrak{a},\mu}(\gamma P,s,\epsilon) = \chi_\epsilon(\gamma) E_{\mathfrak{a},\mu}(P,s,\epsilon)$$

and that $E_{\mathfrak{a},\mu}(\sigma_{\mathfrak{b}} P,s,\epsilon)$ is $\Lambda_{\mathfrak{b}}$-invariant. Additionally, it is easy to check that for $\mathrm{Re}(s) > 2$ and $\mu \neq 0$,

$$E_{\mathfrak{a},\mu}(P,s,\epsilon) \in L^2(\Gamma \backslash \mathbb{H}, \chi_\epsilon). \tag{3.16}$$

An easy computation shows that

$$(\Delta + s(2-s)) E_{\mathfrak{a},\mu}(P,s,\epsilon) = 2\pi|\mu|(1-2s) E_{\mathfrak{a},\mu}(P,s+1,\epsilon), \tag{3.17}$$

which can be rewritten as

$$E_{\mathfrak{a},\mu}(P,s,\epsilon) = 2\pi|\mu|(1-2s) R(s(2-s),\epsilon)(E_{\mathfrak{a},\mu}(P,s+1,\epsilon)), \tag{3.18}$$

where $R(\lambda,\epsilon)$ is the resolvent of Δ on $L^2(\Gamma \backslash \mathbb{H}^3, \chi_\epsilon)$ at λ. We have that $R(s(2-s),\epsilon)$ is meromorphic for $\mathrm{Re}(s) > 1$ and has possible poles at $s_j(\epsilon)$. Hence, from (3.16) and (3.18), it follows that $E_{\mathfrak{a},\mu}(P,s,\epsilon)$ may be analytically continued to $\mathrm{Re}(s) > 1$, with possible poles at $s_j(\epsilon)$.

Next, we want to use the Poincaré series to obtain an integral representation for the generating series $L_{\mathfrak{ab}}(s,0,\mu,\epsilon)$.

Lemma 3.3.1. *Let $\mu \in \Lambda_{\mathfrak{b}}^* \setminus \{0\}$ and $\mathrm{Re}(s), \mathrm{Re}(w) > 2$. Then we have the integral representation*

$$L_{\mathfrak{ab}}(s,0,\mu,\epsilon) = \frac{|\mathcal{P}|(4\pi|\mu|)^{w-1}}{2\pi^{s+1/2}|\mu|^{s-1}} \frac{\Gamma(s)\Gamma(w-1/2)}{\Gamma(w+s-2)\Gamma(w-s)} \int_{\Gamma \backslash \mathbb{H}^3} E_{\mathfrak{a}}(P,s,\epsilon) \overline{E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)} dv \ .$$

Proof. We use a standard unfolding technique together with (3.11) and (3.15) to obtain

$$\begin{aligned}\int_{\Gamma\backslash\mathbb{H}^3} E_{\mathfrak{a}}(P,s,\epsilon)\overline{E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)}dv &= \int_0^\infty \int_{\mathcal{P}_{\mathfrak{b}}} E_{\mathfrak{a}}(\sigma_{\mathfrak{b}}P,s,\epsilon)y^w e^{-2\pi|\mu|y}e(-\langle z,\mu\rangle)\frac{dx_1dx_2dy}{y^3}\\ &= \int_0^\infty y^w e^{-2\pi|\mu|y}|\mathcal{P}_{\mathfrak{b}}||\mu|^{s-1}\phi_{\mathfrak{ab}}(s,\mu,\epsilon)\ y\ K_{s-1}(2\pi|\mu|y)\frac{dy}{y^3}\\ &= L_{\mathfrak{ab}}(s,0,\mu,\epsilon)\frac{2\pi^s}{\Gamma(s)}|\mu|^{s-1}\frac{\sqrt{\pi}}{(4\pi|\mu|)^{w-1}}\frac{\Gamma(w+s-2)\Gamma(w-s)}{\Gamma(w-1/2)},\end{aligned}$$

where we have used [47, p. 205] for the integral of the Bessel function. □

Remark 3.3.1. Since Poincaré series are orthogonal to constants, or using the above calculation, it follows that, for $\mu \in \Lambda_{\mathfrak{b}}^* \setminus \{0\}$,

$$\int_{\Gamma\backslash\mathbb{H}^3} E_{\mathfrak{b},\mu}(P,s,0)dv = 0\ .$$

Next we want to use Lemma 3.3.1 to find the analytic properties of $L_{\mathfrak{ab}}(s,0,\mu,0)$ at $s = 2$.

Lemma 3.3.2. *For $\mu \in \Lambda_{\mathfrak{b}}^*$, the series $L_{\mathfrak{ab}}(s,0,\mu,\epsilon)$ admits meromorphic continuation to $s \in \mathbb{C}$. At $s = 2$, $L_{\mathfrak{ab}}(s,0)$ has a pole with residue*

$$\operatorname{Res}_{s=2} L_{\mathfrak{ab}}(s,0) = \frac{|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]}{\pi\operatorname{vol}(\Gamma\backslash\mathbb{H}^3)}$$

while for $\mu \neq 0$, $L_{\mathfrak{ab}}(s,0,\mu,0)$ is holomorphic at $s = 2$.

Proof. Since the Eisenstein series $E_{\mathfrak{a}}(P,s,\epsilon)$ admits meromorphic continuation to $s \in \mathbb{C}$, its Fourier coefficients admit meromorphic continuation as well. Hence from (3.11) and (3.12), we obtain meromorphic continuation for $L_{\mathfrak{ab}}(s,0,\mu,\epsilon)$.

When $\mu = 0$, we make use of the Maaß–Selberg relations in $\mathbb{H}^3$ [32, p. 110], which tell us the behaviour of truncated Eisenstein series. We define

$$E_{\mathfrak{a}}^Y(P,s,\epsilon) := \begin{cases} E_{\mathfrak{a}}(z,s,\epsilon) - \delta_{\mathfrak{ab}}[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}](\operatorname{Im}\sigma_{\mathfrak{b}}^{-1}P)^s - \phi_{\mathfrak{ab}}(s,\epsilon)(\operatorname{Im}\sigma_{\mathfrak{b}}^{-1}P)^{2-s}, & \text{if } P \in \mathcal{F}_{\mathfrak{b}}(Y),\\ E_{\mathfrak{a}}(P,s,\epsilon), & \text{if } P \in \mathcal{F}(Y),\end{cases}$$

where we have the disjoint union of the fundamental domain $\mathcal{F} = \mathcal{F}(Y) \cup \mathcal{F}_{\mathfrak{a}_1}(Y) \cup \cdots \cup \mathcal{F}_{\mathfrak{a}_h}(Y)$ as in (2.31). Thus, by removing the cuspidal contribution, $E_{\mathfrak{a}}(P,s,\epsilon)$ is square-integrable. The Maaß–Selberg relations give us a formula for $\langle E^Y(P,s,0), E^Y(P,t,0)\rangle$, see [32, Theorem 3.3.6]:

$$\begin{aligned}\left\langle E^Y(P,s,0), E^Y(P,t,0)\right\rangle =& \frac{1}{s+\bar{t}}\delta_{\mathfrak{a}\mathfrak{b}}|\mathcal{P}_{\mathfrak{a}}|[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]Y^{s+\bar{t}} \\ &+\frac{1}{s-\bar{t}}|\mathcal{P}_{\mathfrak{a}}|\overline{\phi}_{\mathfrak{a}\mathfrak{b}}(\bar{t},0)Y^{s-\bar{t}} - \frac{1}{s-\bar{t}}\frac{|\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]}{[\Gamma_{\mathfrak{b}}:\Gamma'_{\mathfrak{b}}]}\phi_{\mathfrak{a}\mathfrak{b}}(s,0)Y^{-s+\bar{t}} \\ &-\sum_{i=1}^{h}\frac{|\mathcal{P}_{\mathfrak{a}_i}|}{[\Gamma_{\mathfrak{a}_i}:\Gamma'_{\mathfrak{a}_i}]}\frac{1}{s+\bar{t}}\phi_{\mathfrak{a}\mathfrak{a}_i}(s,0)\overline{\phi}_{\mathfrak{b}\mathfrak{a}_i}(\bar{t},0)Y^{-s-\bar{t}}.\end{aligned}$$

We know that $E_{\mathfrak{a}}^Y(P,s,0)$ has a simple pole at $s=2$, and by taking $Y\to\infty$ as in [47, Proposition 6.13], one obtains

$$\operatorname{Res}_{s=2} E_{\mathfrak{a}}(\sigma_{\mathfrak{b}}P,s,0) = \frac{|\mathcal{P}_{\mathfrak{a}}|[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]}{\operatorname{vol}(\Gamma\backslash\mathbb{H}^3)}\ . \tag{3.19}$$

The conclusion follows from relating $L_{\mathfrak{a}\mathfrak{b}}(s,0)$ to the 0-th Fourier coefficient of $E_{\mathfrak{a}}(P,s,0)$, as it can be seen from (3.11) and (3.12).

Now, when $\mu\neq 0$, then we know that $L_{\mathfrak{a}\mathfrak{b}}(s,0,\mu,0)$ has at most one simple pole at $s=2$. Using the integral representation from Lemma 3.3.1, this residue would have $\langle 1, E_{\mathfrak{b},\mu}(P,\overline{w},0)\rangle$ as a factor, and by the remark above, this vanishes. □

3.3.3 Bounds for modular symbols

In this section we prove upper bounds for modular symbols, in similar fashion to [51, Proposition 3.3] or [77, Proposition 2.6].

Theorem 3.3.1. *If* $\gamma = \begin{pmatrix} * & * \\ c & * \end{pmatrix} \in T_{\mathfrak{a}\mathfrak{b}}$, *then* $\left\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1},\alpha\right\rangle \ll |\log|c|| + 1$.

Proof. We define the antiderivative of α:

$$F_{\mathfrak{a}}(P) = \int_{\mathfrak{a}}^{P} \alpha\ . \tag{3.20}$$

Since α is cuspidal, it follows that it is rapidly decreasing at cusps, and hence F is well-defined on $\mathbb{H}^3\cup\{\text{cusps}\}$. We note that

$$F_{\mathfrak{a}}(P) = \int_{\mathfrak{a}}^{P}\alpha = \int_{j\infty}^{\sigma_{\mathfrak{a}}^{-1}P}\sigma_{\mathfrak{a}}^{*}\alpha.$$

We note that $F'_{\mathfrak{a}} := F_{\mathfrak{a}}\circ\sigma_{\mathfrak{a}}$ is invariant under the translations in $\Lambda_{\mathfrak{a}}$. Since α is rapidly decreasing at the cusp $\mathfrak{a}$, it follows that $F_{\mathfrak{a}}(P)$ is bounded for $y(\sigma_{\mathfrak{a}}^{-1}P) > Y_0$ with Y_0 chosen as in (2.31).

Writing $\sigma_{\mathfrak{a}}^{*}\alpha = f_{\mathfrak{a},1}dx_1 + f_{\mathfrak{a},2}dx_2 + f_{\mathfrak{a},3}dy$, we conclude that

$$\begin{aligned} F_{\mathfrak{a}}(P) &= \int_{j\infty}^{\sigma_{\mathfrak{a}}^{-1}P} f_{\mathfrak{a},1}dx_1 + f_{\mathfrak{a},2}dx_2 + f_{\mathfrak{a},3}dy \\ &= \int_{\infty}^{y(\sigma_{\mathfrak{a}}^{-1}P)} f_{\mathfrak{a},3}(z,y)dy \quad (\text{for some } z \in \mathcal{P}_{\mathfrak{a}}) \\ &= \int_{\infty}^{Y_0} f_{\mathfrak{a},3}(z,y)dy + \int_{Y_0}^{y(\sigma_{\mathfrak{a}}^{-1}P)} f_{\mathfrak{a},3}(z,y)dy \\ &\ll 1 + |\log y(\sigma_{\mathfrak{a}}^{-1}P)|. \end{aligned}$$

The contributions from dx_1 and dx_2 drop since we are integrating along the y-axis, and the last inequality follows from the fact that $f_{\mathfrak{a},3}(z,y) \ll 1/y$, see (3.6). We deduce that for $\delta \in \Gamma$,

$$\begin{aligned} \langle \delta, \alpha \rangle &= F_{\mathfrak{a}}(\delta P) - F_{\mathfrak{a}}(P) \\ &\ll |\log\big(y(\sigma_{\mathfrak{a}}^{-1}\delta P)\big)| + |\log\big(y(\sigma_{\mathfrak{a}}^{-1}P)\big)| + 1 \ . \end{aligned}$$

Pick $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}$ and $P = \sigma_{\mathfrak{b}}(0,0,1)$. Then the equation above implies that

$$\big\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1}, \alpha \big\rangle \ll |\log\big(|c|^2 + |d|^2\big)| + 1 \ .$$

The lower left element c is constant in a double coset in $\sigma_{\mathfrak{a}}^{-1}\Gamma'_{\mathfrak{a}}\sigma_{\mathfrak{a}}\backslash\sigma_{\mathfrak{a}}^{-1}\Gamma\sigma_{\mathfrak{b}}/\sigma_{\mathfrak{b}}^{-1}\Gamma'_{\mathfrak{b}}\sigma_{\mathfrak{b}}$ and clearly $|c| \geq c_{\mathfrak{ab}}$. Hence we can choose a representative $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in this double coset such that $|d| \ll |c|$ and we conclude that

$$\big\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1}, \alpha \big\rangle \ll |\log|c|| + 1 \ .$$

□

3.4 Perturbation theory of objects twisted by modular symbols

In this section we study the dependency on ϵ of the space $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$. Recall that $\chi_\epsilon(\gamma) = \exp(2\pi i\epsilon \langle\gamma, \alpha\rangle)$, so by studying the analytic properties and perturbation theory in ϵ of $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$, we obtain information about the value distribution of modular symbols. If we denote by $\lambda_0(\epsilon)$ the first eigenvalue of $-\Delta$ on $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$, we will see that, for ϵ small enough, $\lambda_0(\epsilon)$ is analytic in ϵ and we obtain the first few terms in the Taylor expansion around $\epsilon = 0$. We also study the behaviour of the residue of $L_{\mathfrak{ab}}(s, \epsilon)$ at $s_0(\epsilon)$, where $s_0(\epsilon)(2 - s_0(\epsilon)) = \lambda_0(\epsilon)$.

3.4.1 Spectral theory of the space $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$

Denote by $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$ the space of square integrable functions on $\Gamma\backslash\mathbb{H}^3$ with respect to the hyperbolic metric, satisfying

$$f(\gamma P) = \chi_\epsilon(\gamma) f(P)\ .$$

For $f, g \in L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$, we note that $f\overline{g}$ is Γ-invariant. Hence we define the inner product

$$\langle f, g\rangle := \int_{\Gamma\backslash\mathbb{H}^3} f\overline{g}\ dv\ .$$

We let $\mathcal{D}(\epsilon) \subset L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$ be the subspace consisting of all C^2-functions such that $\Delta f \in L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$. For $f, g \in C^1(\mathbb{H})$, as in [32, p. 136], we define

$$\mathbf{Gr}(f, g) := y^2(f_{x_1}\overline{g_{x_1}} + f_{x_2}\overline{g_{x_2}} + f_y\overline{g_y}) = [df, dg]\ , \tag{3.21}$$

where we have used the notation introduced in (3.5). Then for all $f, g \in \mathcal{D}(\epsilon)$, $\mathbf{Gr}(f, g)$ is Γ-invariant. Moreover, the following theorem holds, see [32, Theorem 4.1.7].

Theorem 3.4.1. *For all $f, g \in \mathcal{D}(\epsilon)$,*

$$\int_{\Gamma\backslash\mathbb{H}^3} (-\Delta f)\overline{g} dv = \int_{\Gamma\backslash\mathbb{H}^3} \mathbf{Gr}(f, g) dv\ .$$

In particular, $-\Delta : \mathcal{D}(\epsilon) \to L^2(\Gamma\backslash\mathbb{H}, \chi_\epsilon)$ is a symmetric and positive operator. We denote by $\widetilde{L}(\epsilon)$ the closure of Δ acting on $\mathcal{D}(\epsilon)$.

The theory developed in [32, Chapter 5] for $L^2(\Gamma\backslash\mathbb{H}^3)$ can be straightforwardly generalised to $L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$. The operator $\widetilde{L}(\epsilon)$ is nonnegative, its spectrum consists of a discrete part and a continuous part. Let

$$0 \leq \lambda_0(\epsilon) \leq \lambda_1(\epsilon) \leq \cdots \lambda_n(\epsilon) < 1$$

be the eigenvalues in the interval $[0, 1)$ counted with their multiplicities.

The first eigenvalue is zero if and only if $\epsilon = 0$, in which case it is simple and the eigenspace is generated by the constant function. We write $\lambda_n(\epsilon) = s_n(\epsilon)(2 - s_n(\epsilon))$, where we choose $1 \leq s_n(\epsilon) \leq 2$ for $0 \leq \lambda_n(\epsilon) \leq 1$.

Recall that since α is cuspidal, there exists some compactly supported 1-form $\widetilde{\alpha}$ such that

$$\langle \gamma, \widetilde{\alpha} \rangle = \langle \gamma, \alpha \rangle \qquad \text{for all } \gamma \in \Gamma \ .$$

With this in mind, we define

$$U_{\mathfrak{a}}(P, \epsilon) := \exp\left(2\pi i \epsilon \int_{\mathfrak{a}}^{P} \widetilde{\alpha}\right) \tag{3.22}$$

and consider the unitary operators

$$\begin{aligned} U_{\mathfrak{a}}(\epsilon) : L^2(\Gamma\backslash\mathbb{H}^3) &\to L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon), \\ f &\mapsto U_{\mathfrak{a}}(\cdot, \epsilon) f \ . \end{aligned}$$

We also define

$$L(\epsilon) := U_{\mathfrak{a}}(\epsilon)^{-1} \widetilde{L}(\epsilon) U_{\mathfrak{a}}(\epsilon) \ . \tag{3.23}$$

This implies that $L(\epsilon) = \Delta$ outside the support of $\widetilde{\alpha}$. This will be crucial later in the thesis, particularly in the proof of Lemma 3.4.3.

This construction ensures that the operator $L(\epsilon)$ acts on the fixed space $L^2(\Gamma\backslash\mathbb{H}^3)$ and that $L(\epsilon)$ and $\widetilde{L}(\epsilon)$ are unitary equivalent. This implies that $\mathrm{Spec}(L(\epsilon)) = \mathrm{Spec}(\widetilde{L}(\epsilon))$.

Write $\widetilde{\alpha} = f_1 dx_1 + f_2 dx_2 + f_3 dy$. Using the fact that

$$\frac{\partial U_{\mathfrak{a}}(P, \epsilon)}{\partial x_1} = 2\pi i \epsilon f_1(P) U_{\mathfrak{a}}(P, \epsilon)$$

and the other two similar corresponding derivatives with respect to x_2 and y, we observe that

$$\begin{aligned} L(\epsilon)h =& U_{\mathfrak{a}}(P,\epsilon)^{-1}\left(y^2\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}+\frac{\partial^2}{\partial y^2}\right)-y\frac{\partial}{\partial y}\right)(U_{\mathfrak{a}}(P,\epsilon)h) \\ =& \Delta h+4\pi i\epsilon y^2\left(f_1\frac{\partial h}{\partial x_1}+f_2\frac{\partial h}{\partial x_2}+f_3\frac{\partial h}{\partial y}\right)+2\pi i\epsilon y^2\left(\frac{\partial f_1}{\partial x_1}+\frac{\partial f_2}{\partial x_2}+\frac{\partial f_3}{\partial y}\right)h \\ &-4\pi^2\epsilon^2 y^2(f_1^2+f_2^2+f_3^2)h-2\pi i\epsilon y f_3 h. \end{aligned}$$

We conclude that

$$L(\epsilon)h=\Delta h+\epsilon L^{(1)}h+\epsilon^2 L^{(2)}h, \tag{3.24}$$

where

$$\begin{aligned} L^{(1)}h &= 2\pi i\left(y^2\left(\frac{\partial f_1}{\partial x_1}+\frac{\partial f_2}{\partial x_2}+\frac{\partial f_3}{\partial y}\right)-yf_3\right)+4\pi i[dh,\alpha]\ , \\ L^{(2)}h &= -4\pi^2\epsilon^2[\alpha,\alpha]h\ . \end{aligned}$$

In particular we note that $L(\epsilon)$ is independent of the choice of the cusp $\mathfrak{a}$ and as a function of ϵ is a polynomial of degree two.

From now on, we fix Y_0 in (2.31) large enough such that $\widetilde{\alpha}$ vanishes on cuspidal sectors $\mathcal{F}_{\mathfrak{a}}(Y_0)$, for all cusps $\mathfrak{a}$. Fix $Y>Y_0$. We choose $h\in C^\infty(\mathbb{R}^+)$ such that $h(y)=0$ for $y\le Y$ and $h(y)=1$ for $y\ge Y+1$. Then for $s\in\mathbb{C}$ and $P\in\mathcal{F}$ we define

$$h_{\mathfrak{a}}(P,s):=\begin{cases} h(y(\sigma_{\mathfrak{a}}^{-1}P))y(\sigma_{\mathfrak{a}}^{-1}P)^s & \text{if } P\in\mathcal{F}_{\mathfrak{a}}(Y_0), \\ 0 & \text{if } P\in\mathcal{F}\setminus\mathcal{F}_{\mathfrak{a}}(Y_0). \end{cases}$$

We extend $h_{\mathfrak{a}}(\cdot,s)$ to a Γ-invariant C^∞-function defined for $s\in\mathbb{C}$ and $P\in\mathbb{H}^3$.

We also define

$$\Omega_\epsilon=\{s\in\mathbb{C}\mid\ \mathrm{Re}(s)>1,\ s(2-s)\notin\mathrm{Spec}(-L(\epsilon))\}\ . \tag{3.25}$$

We have the following Lemma, similar to [32, Lemma 6.1.4], [80, Lemma 2.1] or [79, Lemma 3.1]:

Lemma 3.4.1. *For $s\in\Omega_\epsilon$, there exists a unique $D_a(P,s,\epsilon)$ such that*

$$(L(\epsilon)+s(2-s))D_{\mathfrak{a}}(P,s,\epsilon)=0,\quad D_{\mathfrak{a}}(P,s,\epsilon)-h_{\mathfrak{a}}(P,s)\in L^2(\Gamma\backslash\mathbb{H}^3)\ . \tag{3.26}$$

Moreover, $D_{\mathfrak{a}}(P,s,\epsilon)$ is holomorphic in $s \in \Omega_\epsilon$ and real analytic in ϵ.

Proof. If such a solution exists, we write

$$g_{\mathfrak{a}}(P,s,\epsilon) = D_{\mathfrak{a}}(P,s,\epsilon) - h_{\mathfrak{a}}(P,s) \in L^2(\Gamma\backslash\mathbb{H}^3)\ .$$

We apply $(L(\epsilon)+s(2-s))$ to deduce

$$(L(\epsilon)+s(2-s))g_{\mathfrak{a}}(P,s,\epsilon) = H_{\mathfrak{a}}(P,s,\epsilon)\ , \tag{3.27}$$

where

$$H_{\mathfrak{a}}(P,s,\epsilon) = -(L(\epsilon)+s(2-s))h_{\mathfrak{a}}(P,s)\ . \tag{3.28}$$

We note that $H_{\mathfrak{a}}$ is a Γ-invariant C^∞-function in the variable P, which is moreover of compact support when restricted to $\mathcal{F}$. It also depends holomorphically on $s \in \Omega_\epsilon$. Moreover, since $L(\epsilon)$ is equal to Δ outside the support of $\widetilde{\alpha}$, we observe that $H_{\mathfrak{a}}$ is independent from ϵ, so that we can write it as $H_{\mathfrak{a}}(P,s)$.

We can now use (3.28) as a definition for $H_{\mathfrak{a}}(P,s)$, and for $s \in \Omega_\epsilon$, we can apply the resolvent operator defined as

$$R(s,\epsilon) = (L(\epsilon)+s(2-s))^{-1}$$

to obtain a unique function

$$g_{\mathfrak{a}}(P,s,\epsilon) = R(s,\epsilon)H_{\mathfrak{a}}(P,s) \in L^2(\Gamma\backslash\mathbb{H}^3)\ .$$

Since there exist only finitely many values of $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$ for which $s(2-s)$ is an eigenvalue of $-\Delta = -L(0)$ and we know that $L(\epsilon)$ is a polynomial in ϵ given by (3.24), we can use the arguments in [52, p. 66–67] to conclude that the resolvent $R(s,\epsilon)$ is holomorphic for $s \in \Omega_\epsilon$ and depends real analytically on ϵ. □

Remark 3.4.1. For $\mathrm{Re}(s) > 2$, the equation (3.26) agrees with

$$D_{\mathfrak{a}}(P,s,\epsilon) = U_{\mathfrak{a}}(\epsilon)^{-1}E_{\mathfrak{a}}(P,s,\epsilon)\ .$$

Therefore, the conclusions of Lemma 3.4.1 hold for the Eisenstein series in the region

$s \in \Omega_\epsilon$.

3.4.2 Behavior of $\lambda_0(\epsilon)$ and the residue of $L_{\mathfrak{ab}}(s, \epsilon)$ at $s_0(\epsilon)$

We know that $\lambda_0(0) = 0$ is a simple eigenvalue for $L(0) = \Delta$. It is possible to apply Kato's perturbation theory for finite dimensional spaces [52, p. 68–70] for our operator $L(\epsilon)$ of the form (3.24), as explained in [83, Section 4]. We conclude that for ϵ in a small interval around 0, $\lambda_0(\epsilon)$ is real analytic in ϵ and also $\lambda_0(\epsilon)$ is a simple eigenvalue.

We let $u_0(P, \epsilon) \in L^2(\Gamma\backslash\mathbb{H}^3, \chi_\epsilon)$ be the normalised corresponding eigenfunction of $-\widetilde{L}(\epsilon)$, i.e.

$$-\widetilde{L}(\epsilon)u_0(P, \epsilon) = \lambda_0(\epsilon)u_0(P, \epsilon) \quad \text{and} \quad \int_{\Gamma\backslash\mathbb{H}^3} |u_0(P, \epsilon)|^2 dv = 1 \ . \tag{3.29}$$

We want to study the behaviour of $\lambda_0(\epsilon)$ around $\epsilon = 0$. We adapt the proof of [82, Lemma 2.1].

Lemma 3.4.2. *We have that $\lambda_0'(0) = 0$ and*

$$\lambda_0''(0) = \frac{8\pi^2}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} \|\alpha\|_2^2 \ .$$

Proof. We apply Theorem 3.4.1 with $f(P) = g(P) = u_0(P, \epsilon)$ to obtain

$$\lambda_0(\epsilon) = \int_{\Gamma\backslash\mathbb{H}^3} \mathbf{Gr}(u_0(\cdot, \epsilon), u_0(\cdot, \epsilon))dv = \int_{\Gamma\backslash\mathbb{H}^3} y^2 \left(\left|\frac{\partial u_0}{\partial x_1}\right|^2 + \left|\frac{\partial u_0}{\partial x_2}\right|^2 + \left|\frac{\partial u_0}{\partial y}\right|^2 \right) \frac{dx_1 dx_2 dy}{y^3}. \tag{3.30}$$

In particular, we note that $\lambda_0(\epsilon) \geq 0$ and $\lambda_0(\epsilon) = 0$ if and only if the $u_0(P, \epsilon)$ is constant iff $\epsilon = 0$.

We differentiate (3.30) with respect to ϵ, yielding

$$\lambda_0'(\epsilon) = 2 \int_{\Gamma\backslash\mathbb{H}^3} \mathbf{Gr}\left(\frac{\partial u_0}{\partial \epsilon}, u_0(\cdot, \epsilon)\right) dv \ . \tag{3.31}$$

Setting $\epsilon = 0$ we deduce that $\lambda_0'(0) = 0$ since $u_0(P, 0)$ is a constant function. Differentiating once again,

$$\lambda_0''(\epsilon) = 2 \int_{\Gamma\backslash\mathbb{H}^3} \left(\mathbf{Gr}\left(\frac{\partial^2 u_0}{\partial \epsilon^2}, u_0(\cdot, \epsilon)\right) + \mathbf{Gr}\left(\frac{\partial u_0}{\partial \epsilon}, \frac{\partial u_0}{\partial \epsilon}\right)\right) dv \ . \tag{3.32}$$

We define

$$w(P) := \left.\frac{\partial u_0}{\partial \epsilon}\right|_{\epsilon=0} .$$

Hence (3.32) and (3.7) give us

$$\lambda_0''(0) = 2\int_{\Gamma\backslash\mathbb{H}^3} \mathbf{Gr}(w,w)dv = 2\|dw\|_2^2. \tag{3.33}$$

since the mixed term vanished because $u_0(\cdot,0)$ is constant.

Since $u_0(P,\epsilon) \in L^2(\Gamma\backslash\mathbb{H}^3,\chi_\epsilon)$, we know that $u_0(\gamma P,\epsilon) = \chi_\epsilon(\gamma)u_0(P,\epsilon)$. Differentiating this equation with respect to ϵ and then setting $\epsilon = 0$, we obtain that for all $\gamma \in \Gamma$,

$$w(\gamma P) = w(P) + \frac{2\pi i \langle\gamma,\alpha\rangle}{\sqrt{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)}} , \tag{3.34}$$

where we have used the fact that $u_0(P,0) = 1/\sqrt{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)}$. Moreover, since we know that $\lambda_0(0) = \lambda_0'(0) = 0$, we know from (3.29) that

$$\Delta w = 0 . \tag{3.35}$$

If we define $\beta = dw - 2\pi i \mathrm{vol}(\Gamma\backslash\mathbb{H}^3)^{-1/2}\alpha$, then β is a harmonic, Γ-invariant 1-form such that for all $P \in \mathbb{H}^3$ and $\gamma \in \Gamma$

$$\int_P^{\gamma P} \beta = 0 . \tag{3.36}$$

In other words, this means that $\langle\gamma,\beta\rangle = 0$, for all $\gamma \in \Gamma$, and since we have a perfect pairing and β is a harmonic differential, this implies $\beta = 0$. The result then follows from (3.33). □

Remark 3.4.2. From the proof above, we can deduce that w is of the form

$$w(P) = \frac{2\pi i}{\sqrt{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)}}\int_Q^P \alpha + C_Q \quad \text{for some } Q \in \mathbb{H}^*,$$

where C_Q is a constant.

Corollary 3.4.1. *Let*

$$C_\alpha = \frac{4\|\alpha\|_2^2}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} .$$

Then

$$s_0(\epsilon) = 2 - \pi^2 C_\alpha \epsilon^2 + O(\epsilon^3) \ .$$

Proof. It follows immediately from Lemma 3.4.2 and the fact that $\lambda_0(\epsilon) = s_0(\epsilon)(2 - s_0(\epsilon))$.

□

Lemma 3.4.3. *We have that*

$$\operatorname{Res}_{s=s_0(\epsilon)} L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon) = \frac{|\mathcal{P}_\mathfrak{a}||\mathcal{P}_\mathfrak{b}|[\Gamma_\mathfrak{a} : \Gamma'_\mathfrak{a}]}{\pi \mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} + \epsilon \frac{2i|\mathcal{P}_\mathfrak{a}||\mathcal{P}_\mathfrak{b}|[\Gamma_\mathfrak{a} : \Gamma'_\mathfrak{a}]}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} \int_\mathfrak{a}^\mathfrak{b} \alpha + O(\epsilon^2) \ .$$

Proof. From the Fourier expansion (3.11) of the Eisenstein series, we deduce

$$\int_{\mathcal{P}_\mathfrak{b}} E_\mathfrak{a}(\sigma_\mathfrak{b} P, s, \epsilon) dx_1 dx_2 = \delta_{\mathfrak{a}\mathfrak{b}} |\mathcal{P}_\mathfrak{b}| [\Gamma_\mathfrak{a} : \Gamma'_\mathfrak{a}] y^s + \frac{\pi}{s-1} L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon) y^{2-s}.$$

We look at the residue at $s = s_0(\epsilon)$ on both sides of the equality to obtain

$$\begin{aligned} \frac{\pi y^{2-s_0(\epsilon)}}{s_0(\epsilon)-1} \operatorname{Res}_{s=s_0(\epsilon)} L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon) &= \operatorname{Res}_{s=s_0(\epsilon)} \int_{\mathcal{P}_\mathfrak{b}} E_\mathfrak{a}(\sigma_\mathfrak{b} P, s, \epsilon) dx_1 dx_2 \\ &= \int_{\mathcal{P}_\mathfrak{b}} u_\mathfrak{a}(\sigma_\mathfrak{b} P, \epsilon) dx_1 dx_2 \ , \end{aligned}$$

where

$$u_\mathfrak{a}(P,\epsilon) := \operatorname{Res}_{s=s_0(\epsilon)} E_\mathfrak{a}(P,s,\epsilon) \ .$$

Since $s_0(\epsilon) = 2 + O(\epsilon^2)$, it follows that

$$\left. \frac{\partial \left(\operatorname{Res}_{s=s_0(\epsilon)} L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon) \right)}{\partial \epsilon} \right|_{\epsilon=0} = \frac{1}{\pi} \int_{\mathcal{P}_\mathfrak{b}} v_\mathfrak{a}(\sigma_\mathfrak{b} P) dx_1 dx_2 \ ,$$

where

$$v_\mathfrak{a}(P) := \left. \frac{\partial u_\mathfrak{a}(P,\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} .$$

We define

$$w_\mathfrak{a}(P,\epsilon) = U_\mathfrak{a}(P,\epsilon)^{-1} u_\mathfrak{a}(P,\epsilon). \tag{3.37}$$

Then $w_\mathfrak{a}(P,\epsilon) \in L^2(\Gamma\backslash\mathbb{H}^3)$ and it is an eigenfunction of $L(\epsilon)$ with eigenvalue $\lambda_0(\epsilon)$. Differentiating (3.37) with respect to ϵ and then setting $\epsilon = 0$, we get

$$v_\mathfrak{a}(P) = 2\pi i \ w_\mathfrak{a}(P,0) \int_\mathfrak{a}^P \widetilde{\alpha} + \left. \frac{\partial w_\mathfrak{a}(P,\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} .$$

We note that when P is in the cuspidal sector $\mathcal{F}_{\mathfrak{a}}(Y_0)$, the right-hand side of the equality above is zero, since $\widetilde{\alpha}$ is compactly supported, and $L(\epsilon) = \Delta$ in this region, hence $w_{\mathfrak{a}}(\cdot, \epsilon)$ is constant in this region.

Now, as in (3.19), we know that $u_{\mathfrak{a}}(P, 0) = \frac{|\mathcal{P}_{\mathfrak{a}}|[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)}$, so by using the results in Lemma 3.4.2 and Remark 3.4.2, we deduce that

$$v_{\mathfrak{a}}(P) = \frac{2\pi i|\mathcal{P}_{\mathfrak{a}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} \int_{\mathfrak{a}}^{P} \alpha \ .$$

Since α is a cuspidal one-form, from definition we know that

$$\int_{\mathcal{P}_{\mathfrak{b}}} \left(\int_{\mathfrak{b}}^{\sigma_{\mathfrak{b}} P} \alpha \right) dx_1 dx_2 = \int_{\mathcal{P}_{\mathfrak{b}}} \left(\int_{j\infty}^{P} \sigma_{\mathfrak{b}}^{*} \alpha \right) dx_1 dx_2 = 0 \ ,$$

hence

$$\left. \frac{\partial \left(\mathrm{Res}_{s=s_0(\epsilon)} L_{\mathfrak{ab}}(s,\epsilon)\right)}{\partial \epsilon} \right|_{\epsilon=0} = \frac{2i|\mathcal{P}_{\mathfrak{a}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} \int_{\mathcal{P}_{\mathfrak{b}}} \int_{\mathfrak{a}}^{\sigma_{\mathfrak{b}} P} \alpha = \frac{2i|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} \int_{\mathfrak{a}}^{\mathfrak{b}} \alpha \ .$$

□

3.5 Moment generating function

In this section we study the exponential sum

$$\begin{aligned} \sum_{\gamma \in T_{\mathfrak{ab}}(X)} \chi_{\epsilon}(\sigma_{\mathfrak{a}} \gamma \sigma_{\mathfrak{b}}^{-1}) &= \sum_{\gamma \in T_{\mathfrak{ab}}(X)} \exp\!\left(2\pi i \epsilon \left\langle \sigma_{\mathfrak{a}} \gamma \sigma_{\mathfrak{b}}^{-1}, \alpha \right\rangle\right) \\ &= [\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}] \sum_{r \in R_{\mathfrak{ab}}(X)} \exp(2\pi i \epsilon \langle r \rangle_{\mathfrak{ab}}), \end{aligned} \tag{3.38}$$

which is the moment generating function for the distribution of modular symbols. Moments of modular symbols can be obtained by looking at the derivatives at $\epsilon = 0$ of this sum. We relate this sum to the generating series $L_{\mathfrak{ab}}(s, \epsilon)$. We write the first few terms in the Taylor expansion around $\epsilon = 0$, thus obtaining expressions for the first and second moments of the distribution of modular symbols. Additionally, we show that the values in the set $R_{\mathfrak{ab}}(X)$ become equidistributed modulo the lattice $\Lambda_{\mathfrak{a}}$ as $X \to \infty$.

Firstly, we need the following lemma about bounds on vertical lines for $L_{\mathfrak{ab}}(s, 0, \mu, \epsilon)$.

Lemma 3.5.1. *Fix some $\delta > 0$. If $1 + \delta < \mathrm{Re}(s) < 2 + \delta$ and $s(2 - s)$ bounded away from* *spectrum of $L(\epsilon)$, then, uniformly in ϵ,*

$$L_{\mathfrak{a}\mathfrak{b}}(s,0,\mu,\epsilon) \ll_\delta (1+|\mu|)^{2-\mathrm{Re}(s)+\delta}|s| \ . \tag{3.39}$$

Proof. For $\mu = 0$, we use a similar argument of that in [94, p. 655] (which follows from the Maaß–Selberg relations in $\mathbb{H}^3$ [32, p. 110]). We have that $|\phi_{\mathfrak{a}\mathfrak{b}}(s,\epsilon)| = O(1)$ in the region $\mathrm{Re}(s) > 1+\eta$ and away from the spectrum of $L(\epsilon)$. Now, the result follows from (3.12).

When $\mu \neq 0$ and $\mathrm{Re}(s) > 1$, we use Lemma 3.3.1. Choose $w = 2+2\delta+it$, where $s = \sigma + it$. Then Stirling's formula gives us that the contribution from the Gamma factors is $O(|s|)$.

Next, we want to study the contribution from the integral. We use Lemma 3.4.1 to deduce that for $\mathrm{Re}(s) > 1$ and $s(2-s)$ bounded away from the spectrum of $L(\epsilon)$,

$$\begin{aligned}\int_{\mathcal{F}} |E_{\mathfrak{a}}(P,s,\epsilon)\overline{E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)}|dv &= \int_{\mathcal{F}} |D_{\mathfrak{a}}(P,s,\epsilon)E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)|dv \\ &\leq \int_{\mathcal{F}} |h_{\mathfrak{a}}(P,s)E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)|dv + \int_{\mathcal{F}} |(D_{\mathfrak{a}}(P,s,\epsilon)-h_{\mathfrak{a}}(P,s))E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)|dv \ .\end{aligned}$$

The second integral is bounded by

$$\|g(\sigma_{\mathfrak{a}}^{-1}P,s,\epsilon)\|_{L^2}\|E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)\|_{L^2} \ll 1 \ .$$

It remains to study the first integral. It suffices to concentrate on the cuspidal sector $\mathcal{F}_{\mathfrak{a}}(Y)$ since $h_{\mathfrak{a}}(P,s)$ vanishes everywhere else. We get

$$\int_{\mathcal{F}_{\mathfrak{a}}(Y)} |h_{\mathfrak{a}}(P,s)E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)|dv = \int_Y^\infty \int_{\mathcal{P}_{\mathfrak{a}}} |y^s E_{\mathfrak{b},\mu}(\sigma_{\mathfrak{a}}P,\overline{w},\epsilon)| \ dv \ .$$

Now, with our choice of $w = 2+2\delta+it$, we see that $E_{\mathfrak{b},\mu}(\sigma_{\mathfrak{a}}P,\overline{w},\epsilon)$ decays exponentially in the cusp, so the integral above is indeed bounded. This in turn implies that

$$\int_{\mathcal{F}} |E_{\mathfrak{a}}(P,s,\epsilon)\overline{E_{\mathfrak{b},\mu}(P,\overline{w},\epsilon)}|dv \ll 1 \ ,$$

and hence we obtain the desired upper bound for $L_{\mathfrak{a}\mathfrak{b}}(0,\mu,s,\epsilon)$. □

We obtain the following expression for the moment generating function by using a similar method to [77, Section 4].

Lemma 3.5.2. *There exists an absolute constant $\nu > 0$ depending on the spectral gap of Δ such that, uniformly for ϵ small enough,*

$$\sum_{\gamma \in T_{\mathfrak{a}\mathfrak{b}}(X)} \overline{\chi}_\epsilon(\sigma_\mathfrak{a} \gamma \sigma_\mathfrak{b}^{-1}) = \frac{X^{2s_0(\epsilon)}}{s_0(\epsilon)} \operatorname{Res}_{s=s_0(\epsilon)} L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon) \left(1 + O(X^{-\nu})\right) .$$

Proof. Let $\phi_U : \mathbb{R} \to \mathbb{R}$ be a family of smooth nonincreasing functions with

$$\phi_U(t) = \begin{cases} 1 & \text{if } t \leq 1 - 1/U, \\ 0 & \text{if } t \geq 1 + 1/U, \end{cases} \tag{3.40}$$

and $\phi_U^{(j)}(t) = O(U^j)$ as $U \to \infty$. For $\operatorname{Re}(s) > 0$, we consider the Mellin transform

$$R_U(s) = \int_0^\infty \phi_U(t) t^s \frac{dt}{t} . \tag{3.41}$$

We can easily see that

$$R_U(s) = \frac{1}{s} + O\left(\frac{1}{U}\right) \quad \text{as } U \to \infty \tag{3.42}$$

and for any $c > 0$

$$R_U(s) = O\left(\frac{1}{|s|}\left(\frac{U}{1+|s|}\right)^c\right) \quad \text{as } |s| \to \infty , \tag{3.43}$$

where the last estimate follows from repeated partial integration. Now we use the Mellin inversion to obtain

$$\begin{aligned} \sum_{\gamma \in T_{\mathfrak{a}\mathfrak{b}}} \overline{\chi}_\epsilon(\sigma_\mathfrak{a} \gamma \sigma_\mathfrak{b}^{-1}) \, \phi_U\left(\frac{|c|^2}{X^2}\right) &= \sum_{\gamma \in T_{\mathfrak{a}\mathfrak{b}}} \overline{\chi}_\epsilon(\sigma_\mathfrak{a} \gamma \sigma_\mathfrak{b}^{-1}) \frac{1}{2\pi i} \int_{\operatorname{Re}(s)=3} \frac{X^{2s}}{|c|^{2s}} R_U(s) ds \\ &= \frac{1}{2\pi i} \int_{\operatorname{Re}(s)=3} L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon) X^{2s} R_U(s) ds . \end{aligned}$$

Next, we recall Lemma 3.5.1 and equation (3.43) to deduce that the last integral is absolutely convergent. We want to move the line of integration to $\operatorname{Re}(s) = h$, where

$$h = \frac{2\max(s_1(0), 1) + 2}{3}.$$

Then for ϵ small enough, $s_1(\epsilon) < h < s_0(\epsilon)$. We integrate along a box of height T and let $T \to \infty$. Indeed, the polynomial growth on vertical lines of $L_{\mathfrak{a}\mathfrak{b}}(s,\epsilon)$ guaranteed by

Lemma 3.5.1, together with equation (3.43), give us

$$\lim_{T\to\infty}\int_{\substack{\mathrm{Re}(s)=3\\ |t|\geq T}} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds = \lim_{T\to\infty}\int_{\substack{\mathrm{Re}(s)=h\\ |t|\geq T}} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds = 0 \ ,$$

and

$$\lim_{T\to\infty}\int_{\substack{h\leq\mathrm{Re}(s)<2\\ \mathrm{Im}(s)=T}} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds = \lim_{T\to\infty}\int_{\substack{h\leq\mathrm{Re}(s)<2\\ \mathrm{Im}(s)=-T}} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds = 0 \ .$$

We conclude that

$$\begin{aligned}\frac{1}{2\pi i}\int_{\mathrm{Re}(s)=3} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds =& \frac{1}{2\pi i}\int_{\mathrm{Re}(s)=h} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds \\ &+ \mathrm{Res}_{s=s_0(\epsilon)}\left(L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)\right) \ .\end{aligned}$$

Setting $c=3$ in (3.43), we observe that

$$\int_{\mathrm{Re}(s)=h} L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)ds \ll X^{2h}U^3 \ .$$

Now, (3.42) gives us

$$\mathrm{Res}_{s=s_0(\epsilon)}\left(L_{\mathfrak{ab}}(s,\epsilon)X^{2s}R_U(s)\right) = \frac{X^{2s_0(\epsilon)}}{s_0(\epsilon)}\left(\mathrm{Res}_{s=s_0(\epsilon)} L_{\mathfrak{ab}}(s,\epsilon) + O\left(\frac{1}{U}\right)\right) \ . \tag{3.44}$$

Since we want this to be the main contribution, we choose $U = X^a$, where

$$a = \frac{2-\max(s_1(0),1)}{4}.$$

With this choice, for ϵ small enough, we get

$$\sum_{\gamma\in T_{\mathfrak{ab}}} \overline{\chi}_\epsilon(\sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1})\ \phi_U\left(\frac{|c|^2}{X^2}\right) = \frac{X^{2s_0(\epsilon)}}{s_0(\epsilon)}\left(\mathrm{Res}_{s=s_0(\epsilon)} L_{\mathfrak{ab}}(s,\epsilon) + O(X^{-a})\right) \ . \tag{3.45}$$

Setting $\epsilon=0$, using Lemma 3.3.2, we obtain

$$\sum_{\gamma\in T_{\mathfrak{ab}}} \phi_U\left(\frac{|c|^2}{X^2}\right) = X^4\left(\frac{|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}}:\Gamma'_{\mathfrak{a}}]}{2\pi\mathrm{vol}(\Gamma\backslash\mathbb{H}^3)} + O(X^{-a})\right) \ .$$

We now choose ϕ_U^1 and ϕ_U^2 as in (3.40) 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

$$\sum_{\gamma \in T_{\mathfrak{ab}}} \phi_U^1 \left(\frac{|c|^2}{X^2} \right) \leq \sum_{\gamma \in T_{\mathfrak{ab}}(X)} 1 \leq \sum_{\gamma \in T_{\mathfrak{ab}}} \phi_U^2 \left(\frac{|c|^2}{X^2} \right) ,$$

so the previous two equations give us

$$\# T_{\mathfrak{ab}}(X) = X^4 \left(\frac{|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{2\pi \mathrm{vol}(\Gamma \backslash \mathbb{H}^3)} + O(X^{-a}) \right) . \tag{3.46}$$

Also, from the definition of ϕ_U,

$$\sum_{\gamma \in T_{\mathfrak{ab}}} \overline{\chi}_\epsilon(\sigma_{\mathfrak{a}} \gamma \sigma_{\mathfrak{b}}^{-1}) \, \phi_U \left(\frac{|c|^2}{X^2} \right) = \sum_{\gamma \in T_{\mathfrak{ab}}(X)} \overline{\chi}_\epsilon(\sigma_{\mathfrak{a}} \gamma \sigma_{\mathfrak{b}}^{-1}) + O \left(\# \left\{ \gamma \in T_{\mathfrak{ab}} \ : \ 1 - \frac{1}{U} \leq \frac{|c|^2}{X^2} \leq 1 + \frac{1}{U} \right\} \right) . \tag{3.47}$$

But now we use (3.46) to bound the size of the error term

$$\# \left\{ \gamma \in T_{\mathfrak{ab}} \ : \ 1 - \frac{1}{U} \leq \frac{|c|^2}{X^2} \leq 1 + \frac{1}{U} \right\} = T_{\mathfrak{ab}} \left(X \sqrt{1 + \frac{1}{U}} \right) - T_{\mathfrak{ab}} \left(X \sqrt{1 - \frac{1}{U}} \right) = O \left(X^{4-a/2} \right) .$$

The conclusion follows from (3.45) and (3.47). □

Let

$$F(\epsilon) = \mathrm{Res}_{s = s_0(\epsilon)} \, L_{\mathfrak{ab}}(s, \epsilon)$$

and we write its Taylor expansion around $\epsilon = 0$ as $F(\epsilon) = \sum_{k \geq 0} C_k \epsilon^k$. So far we have shown that

$$C_0 = \frac{|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{\pi \mathrm{vol}(\Gamma \backslash \mathbb{H}^3)} \quad \text{and} \quad C_1 = \frac{2i|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{\mathrm{vol}(\Gamma \backslash \mathbb{H}^3)} \int_{\mathfrak{a}}^{\mathfrak{b}} \alpha \, .$$

We note that the coefficients C_k were essentially computed by Petridis–Risager in [79], allowing them to obtain all moments for modular symbols.

Corollary 3.5.1. *If $\epsilon \geq X^{-\nu/4}$, for some $\nu > 0$ depending on the spectral gap, then*

$$\frac{1}{\# T_{\mathfrak{ab}}(X)} \sum_{\gamma \in T_{\mathfrak{ab}}(X)} \chi_\epsilon(\sigma_{\mathfrak{a}} \gamma \sigma_{\mathfrak{b}}^{-1}) = 1 + \epsilon \left(2\pi i \int_{\mathfrak{b}}^{\mathfrak{a}} \alpha \right) + \epsilon^2 \left(-2\pi^2 \log X C_\alpha + D_{\alpha, \mathfrak{ab}} \right) + O(X^{-\nu}) \, ,$$

where

$$D_{\alpha,\mathfrak{ab}} = 2\pi^2 C_\alpha + \frac{C_2}{C_0} \ . \tag{3.48}$$

Remark 3.5.1. From the formula above we observe that computing the variance shift $D_{\alpha,\mathfrak{ab}}$ is equivalent to finding the second term in Laurent series expansion of $\frac{\partial^2}{\partial\epsilon^2} L_{\mathfrak{ab}}(s,\epsilon)\Big|_{\epsilon=0}$, or in other words finding the first two terms in the Laurent expansion of the Goldfeld Eisenstein series $E^2_{\mathfrak{a}}(P,s)$. For the case of $\mathbb{H}^2$ this is done in [79] and their methods could be extended to work in $\mathbb{H}^3$ as well.

As a consequence of our work so far, we can show that $R_{\mathfrak{ab}}(X)$ is equidistributed in the fundamental domain $\mathcal{P}_{\mathfrak{a}}$ as $X \to \infty$.

Proposition 3.5.1. *There exists $\nu > 0$ depending on the spectral gap for Δ, such that for all $\mu \in \Lambda^*_{\mathfrak{a}}$,*

$$\sum_{r \in R_{\mathfrak{ab}}(X)} e(\langle \mu, r\rangle) = \delta_0(\mu) \frac{|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{2\pi \mathrm{vol}(\Gamma\backslash\mathbb{H}^3)[\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}]} X^4 + O\left((1+|\mu|)X^{4-\nu}\right) \ .$$

In particular, for any continuous function $h : \mathbb{C}/\Lambda_{\mathfrak{a}} \to \mathbb{C}$,

$$\frac{\sum_{r\in R_{\mathfrak{ab}}(X)} h(r)}{\# R_{\mathfrak{ab}}(X)} \to \int_{\mathbb{C}/\Lambda_{\mathfrak{a}}} h(z) dz \quad \text{as} \quad X \to \infty \ .$$

Proof. From Lemma 2.6.1, the generating series for the exponential sum is

$$\sum_{r\in R_{\mathfrak{ab}}(X)} \frac{e(\langle \mu, r\rangle)}{|c|^{2s}} = \frac{1}{[\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}]} \sum_{\gamma \in T_{\mathfrak{ab}}} \frac{e\left(\langle \mu, \gamma\infty\rangle\right)}{|c|^{2s}} = \frac{1}{[\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}]} L_{\mathfrak{ab}}(s,\mu,0,0) \ .$$

By inverting γ in the series above, we note that $L_{\mathfrak{ab}}(s,\mu,0,,0) = L_{\mathfrak{ba}}(s,0,-\mu,0)$. We use a contour integration argument similar to the one in the proof of Lemma 3.5.2. The polynomial growth of $L_{\mathfrak{ba}}(s,0,-\mu,0)$ on vertical lines is guaranteed by Lemma 3.5.1, whilst by Lemma 3.3.2 we know that $L_{\mathfrak{ba}}(s,0,\mu,0)$ has a pole at $s=2$ if and only if $\mu = 0$. Finally, from (3.46) we know that

$$\# R_{\mathfrak{ab}}(X) = X^4 \left(\frac{|\mathcal{P}_{\mathfrak{a}}||\mathcal{P}_{\mathfrak{b}}|[\Gamma_{\mathfrak{a}} : \Gamma'_{\mathfrak{a}}]}{2\pi \mathrm{vol}(\Gamma\backslash\mathbb{H}^3)[\Gamma_{\mathfrak{b}} : \Gamma'_{\mathfrak{b}}]} + O(X^{-\nu}) \right) \ .$$

The second claim follows from the generalised Weyl equidistribution criterion. ☐

3.6 Normal distribution of modular symbols

We now have all the ingredients to prove that modular symbols have asymptotically a normal distribution. We make use of the Berry–Esseen inequality and of our results about the behaviour of $s_0(\epsilon)$ and $L_{\mathfrak{ab}}(s,\epsilon)$.

We recall the Berry–Esseen inequality, see [105, Theorem II.7.16].

Theorem 3.6.1. *If X is a real valued random variable and $T>0$, then*

$$\sup_{z\in\mathbb{R}}\left|\int_{-\infty}^{z} e^{-t^2/2}dt-\mathbb{P}(X<z)\right| \ll \frac{1}{T}+\int_{-T}^{T}\left|\frac{e^{-t^2/2}-\mathbb{E}(\exp(itX))}{t}\right| dt\ . \tag{3.49}$$

For $\gamma\in T_{\mathfrak{ab}}(X)$, we define the random variable

$$A_\gamma=\sqrt{\frac{1}{C_\alpha \log X}}\left\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1},\alpha\right\rangle \tag{3.50}$$

where γ is chosen uniformly at random from $T_{\mathfrak{ab}}(X)$.

We fix $t:=2\pi\epsilon\sqrt{C_\alpha\log X}$. Then, by definition,

$$\mathbb{E}(\exp(itA_\gamma))=\frac{1}{\#T_{\mathfrak{ab}}(X)}\sum_{\gamma\in T_{\mathfrak{ab}}(X)}\chi_\epsilon(\sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1})\ .$$

Fix some $\delta>0$. We choose $T=(\log X)^{1/2-\delta}$ and apply Theorem 3.6.1 for the random variables A_γ. We split the integral on the right-hand side of (3.49) into three ranges, depending on the size of t. All the implied constants are uniform in ϵ (and hence in t).

1. *Small* $|t|$. Suppose $|t|\le X^{-\delta}$, for some small δ. Using $\exp(i\theta)=1+O(\theta)$ and the bounds for $\left\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1},\alpha\right\rangle$ provided by Theorem 3.3.1, we obtain

$$\begin{aligned}\mathbb{E}(\exp(itA_\gamma))&=1+O\left(\frac{t}{\#T_{\mathfrak{ab}}(X)\sqrt{\log X}}\sum_{\gamma\in T_{\mathfrak{ab}}(X)}\left|\left\langle \sigma_{\mathfrak{a}}\gamma\sigma_{\mathfrak{b}}^{-1},\alpha\right\rangle\right|\right)\\&=1+O\left(t\sqrt{\log X}\right)\ .\end{aligned}$$

   Also, when $|t|\le X^{-\delta}$, we see that

$$e^{-t^2/2}=1-\frac{t^2}{2}+O(t^4)=1+O\left(tX^{-\delta}\right)\ .$$

Therefore

$$\int_{|t|\le X^{-\delta}} \left| \frac{e^{-t^2/2} - \mathbb{E}(\exp(itA_\gamma))}{t} \right| dt \ll \int_{|t|\le X^{-\delta}} \sqrt{\log X}\ dt \ll X^{-\delta/2} .$$

2. *Medium* $|t|$. Suppose $X^{-\delta} \le |t| \le (\log X)^\delta$, where $\delta > 0$. Using that

$$s_0(\epsilon) = 2 - \pi^2 C_\alpha \epsilon^2 + O(\epsilon^3),$$

we see that

$$\begin{aligned}
\mathbb{E}(\exp(itA_\gamma)) &= \frac{2X^{2s_0(\epsilon)-4}}{s_0(\epsilon)} (1 + O(\epsilon)) \\
&= \exp\left(\log X(-2\pi^2 C_\alpha \epsilon^2 + O(\epsilon^3))\right)(1 + O(\epsilon)) \\
&= e^{-t^2/2}(1 + O(\epsilon^3 \log X) + O(\epsilon)) \\
&= e^{-t^2/2} + O\left(\frac{e^{-t^2/2}|t|^3}{\sqrt{\log X}} + \frac{e^{-t^2/2}}{(\log X)^{1/2-\delta}} \right) .
\end{aligned}$$

Hence the contribution from such t is

$$\begin{aligned}
&\int_{X^{-\delta}<|t|<(\log X)^\delta} \left| \frac{e^{-t^2/2} - \mathbb{E}(\exp(itA_\gamma))}{t} \right| dt \\
\ll &\int_{X^{-\delta}<|t|<(\log X)^\delta} \left(\frac{e^{-t^2/2}t^2}{(\log X)^{1/2}} + \frac{e^{-t^2/2}}{|t|(\log X)^{1/2-\delta}} \right) dt \\
\ll &(\log X)^{-1/2+\delta} .
\end{aligned}$$

3. *Large* $|t|$. Suppose $(\log X)^\delta \le |t| \le (\log X)^{1/2-\delta}$. Similarly as in the previous case,

$$\mathbb{E}(\exp(itA_\gamma)) \ll e^{-t^2/2+O(|t|^3(\log X)^{-1/2})} \ll e^{-t^2/4} \ll e^{-(\log X)^{\delta/2}} .$$

Therefore, the contribution from large $|t|$ is bounded by

$$\int_{(\log X)^\delta \le |t| \le (\log X)^{1/2-\delta}} \frac{e^{-(\log X)^{\delta/2}}}{|t|} dt \ll (\log X)^{-1/2} .$$

Putting everything together, we conclude the result in Theorem 3.1.2(a). Parts (b) and (c) of Theorem 3.1.2, where the results about the first and second moments are stated, follow easily from Corollary 3.5.1.

Remark 3.6.1. The method exposed in this chapter can be generalised to cohomology classes $H^1_{\text{cusp}}(\Gamma, \mathbb{R})$, where $\Gamma < \mathrm{SO}(n+1,1)$ cofinite with cusps. Therefore we can prove the following theorem. We refer to Section 2.7 or Chapter 4 for detailed description of the notation.

Theorem 3.6.2. *Let ω be in the free part of $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R})$. Then the random variable $\gamma \mapsto \omega(\gamma)$ defined on the sample space $T_\Gamma(X)$ is asymptotically normally distributed. More precisely, there exists a constant C_ω such that, or every $a, b \in [-\infty, \infty]$ with $a \leq b$, we have*

$$\frac{\#\left\{\gamma \in T_\Gamma(X)\ ,\ \frac{\omega(\gamma)}{\sqrt{C_\omega \log X}} \in [a,b]\right\}}{\#T_\Gamma(X)} \to \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left(-\frac{t^2}{2}\right) dt, \quad \text{as } X \to \infty.$$

3.7 Results for imaginary quadratic fields

So far, we have described our results for the general case of a Kleinian group Γ. In this section we apply our results to Bianchi groups and their congruence subgroups. Let K a quadratic number field with discriminant d_K. The arithmetic properties the groups $\mathrm{PSL}_2(\mathcal{O}_K)$ and their congruence subgroups, as well as the geometry of the corresponding quotient spaces, are thoroughly described in [32, Chapter 7], while the theory of Eisenstein series for $\Gamma = \mathrm{PSL}_2(\mathcal{O}_K)$ is developed in [32, Chapter 8].

The ring of integers $\mathcal{O}_K$ has the $\mathbb{Z}$-basis consisting of 1 and ω, where

$$\omega = \frac{d_K + \sqrt{d_K}}{2}\ .$$

We denote by $\mathcal{P}_K$ a fundamental domain for this lattice.

The zeta function $\zeta_K(s)$ of K is for $\mathrm{Re}(s) > 1$ defined by

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{N(\mathfrak{a})^s}\ ,$$

where the sum is over the non-zero ideals of $\mathcal{O}_K$ and the norm of $\mathfrak{a}$ is $N(\mathfrak{a}) = |\mathcal{O}_K/\mathfrak{a}|$.

As mentioned in the introduction, Cremona has several results about modular symbols associated to quadratic imaginary number fields. He uses them to compute spaces of modular forms and to establish an arithmetic correspondence between elliptic curves and cusp forms, see [22], [23], [21]. For consistency reasons, we will use the notation used in his work.

For technical reasons, we assume the K has class number one. This is not a vital restriction, but it allows us to obtain nice arithmetic descriptions of the cusps and easier formulae relating modular symbols to L-functions. Let $\mathfrak{n}$ be a nonzero ideal in the ring of integers $\mathcal{O}_K$. We work with the congruence subgroup

$$\Gamma_0(\mathfrak{n}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}_2(\mathcal{O}_K) \ : \ c \in \mathfrak{n} \right\} .$$

A basis for the left-invariant differential 1-forms on $\mathbb{H}^3$ is chosen to be

$$\beta = \left(-\frac{dz}{y}, \frac{dy}{y}, \frac{d\overline{z}}{y} \right) . \tag{3.51}$$

Let $F : \mathbb{H}^3 \to \mathbb{C}^3$ be a vector-valued function which we can write as $F = (F_0, F_1, F_2)$, then we define the differential 1-form

$$F \cdot \beta := \frac{1}{y} \left(-F_0 dz + F_1 dy + F_2 d\overline{z} \right) . \tag{3.52}$$

Definition 3.7.1. *Let $F : \mathbb{H}^3 \to \mathbb{C}^3$ be a vector-valued function and $\gamma \in \mathrm{GL}_2(\mathbb{C})$. Then we define a new function $(F|\gamma) : \mathbb{H}^3 \to \mathbb{C}^3$ by*

$$(F|\gamma)(P) := F(\gamma P) j(\gamma; P) ,$$

where

$$j(\gamma; P) = \frac{1}{|r|^2 + |s|^2} \begin{pmatrix} r^2 & -2rs & s^2 \\ r\overline{s} & |r|^2 - |s|^2 & -\overline{r}s \\ \overline{s}^2 & 2\overline{rs} & \overline{r}^2 \end{pmatrix}$$

with $r = \overline{cz + d}$ and $s = \overline{c}y$.

This definition ensures that the differential $F \cdot \beta$ is invariant under γ if and only if $F|\gamma = F$.

Definition 3.7.2. *A cusp form of weight 2 for $\Gamma_0(\mathfrak{n})$ is a vector-valued function $F : \mathbb{H}^3 \to \mathbb{C}^3$ such that*

1. *$F \cdot \beta$ is a harmonic 1-form;*

2. $F|\gamma = F$*, for all* $\gamma \in \Gamma_0(\mathfrak{n})$*;*

3. For all $\gamma \in \mathrm{PSL}_2(\mathcal{O}_K)$ *and* $y \geq 0$*,*

$$\int_{\mathcal{P}_K} (F|\gamma)(z, y)dz = 0 \ .$$

We denote the space of cusp forms of weight 2 for $\Gamma_0(\mathfrak{n})$ by $S(\mathfrak{n})$. We note that $F \in S(\mathfrak{n})$ if and only if $F \cdot \beta$ is a cuspidal 1-form for $X_0(\mathfrak{n}) := \Gamma_0(\mathfrak{n})\backslash\mathbb{H}^*$, where $\mathbb{H}^* = \mathbb{H}^3 \cup K \cup \{\infty\}$. In fact, the map

$$\begin{aligned} S(\mathfrak{n}) &\to H^1_{\mathrm{cusp}}(X_0(\mathfrak{n}), \mathbb{C}) \\ F &\mapsto F \cdot \beta \end{aligned}$$

is an isomorphism.

For $F \in S(\mathfrak{n})$, we have the Fourier expansion

$$F = (F_0, F_1, F_2) = \sum_{0 \neq \alpha \in \mathcal{O}_K} c(\alpha) y^2 \mathbf{K}\left(\frac{4\pi|\alpha|y}{\sqrt{|d_K|}}\right) \psi\left(\frac{\alpha z}{\sqrt{d_K}}\right) \tag{3.53}$$

where $\psi(z) = e(z + \overline{z})$ and

$$\mathbf{K}(y) = \left(-\frac{i}{2}K_1(y), K_0(y), \frac{i}{2}K_1(y)\right)$$

for $y > 0$ and K_0, K_1 the K-Bessel functions.

The theory of cusp forms and associated L-functions, Hecke operators, newforms etc. is similar to the classical Atkin–Lehner theory over $\mathbb{Q}$. We briefly recall the elements we need for our exposition.

For primes π in $\mathcal{O}_K$ which do not divide the level $\mathfrak{n}$, the Hecke operator T_π sends the cusp form with Fourier coefficients $c(\alpha)$ to one with coefficients $c'(\alpha)$, where $c'(\alpha) = N(\pi)(\alpha\pi) + c(\alpha/\pi)$, where $c(\alpha) = 0$ if $\alpha \notin \mathcal{O}_K$. As in the classical case, a newform in $S(\mathfrak{n})$ is an eigenform for all Hecke operators T_π, for π not dividing $\mathfrak{n}$, which is not induced by a form in $S(\mathfrak{m})$, for any level $\mathfrak{m}$ properly dividing $\mathfrak{n}$.

Secondly, let $\mathfrak{e}$ a divisor of $\mathfrak{n}$ and e is a generator for $\mathfrak{e}$. Then the Atkin–Lehner operator $W_\mathfrak{e}$ on $S(\mathfrak{n})$ is given by the action of any matrix of the form $\begin{pmatrix} ae & b \\ cN & de \end{pmatrix}$ which has determinant e. Then this operator is an involution and it commutes with the action of all Hecke operators.

Let ϵ be a unit in $\mathcal{O}_K^*$ and I_ϵ denote the matrix $\begin{pmatrix} \epsilon & 0 \\ 0 & 1 \end{pmatrix}$. The action of I_ϵ on $\mathbb{H}^3$ sends (z, y) to $(\epsilon z, y)$ and if $F \in S(\mathfrak{n})$ has Fourier coefficients $c(\alpha)$, then $F|I_\epsilon$ has Fourier coefficients $c(\epsilon\alpha)$. Since $\begin{pmatrix} \epsilon^2 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} \epsilon & 0 \\ 0 & \epsilon^{-1} \end{pmatrix}$ give birth to the same action, but the latter belongs to $\Gamma_0(\mathfrak{n})$, we must have that $c(\alpha) = c(\epsilon^2\alpha)$, for all units $\epsilon \in \mathcal{O}_K^*$. Hence if ϵ is a generator for the unit group $\mathcal{O}_K^*$, then I_ϵ induces an involution of $S(\mathfrak{n})$ which commutes with the Hecke operators, hence we can split $S(\mathfrak{n})$ into two eigenspaces

$$S(\mathfrak{n}) = S^+(\mathfrak{n}) \oplus S^-(\mathfrak{n}) \ .$$

Newfroms in $S^+(\mathfrak{n})$ are called plusforms, and their Fourier coefficients satisfy $c(\alpha) = c(\epsilon\alpha)$, for all $\alpha \in \mathcal{O}_K^*$. Hence they depend only on the ideal (α). So if $F \in S^+(\mathfrak{n})$, we attach to F the L-function

$$L(F, s) = \sum_{\mathfrak{a}} \frac{c(\mathfrak{a})}{N(\mathfrak{a})^s} \ .$$

Since the Fourier coefficients $c(\mathfrak{a})$ are multiplicative, we obtain the Euler product

$$L(F, s) = \prod_{\mathfrak{p}} (1 - c(\mathfrak{p})N(\mathfrak{p})^{-s} + \chi(\mathfrak{p})N(\mathfrak{p})^{1-2s})^{-1}, \quad \text{where } \chi(\mathfrak{p}) = \begin{cases} 0 & \text{if } \mathfrak{p} \mid \mathfrak{n} \ , \\ 1 & \text{if } \mathfrak{p} \nmid \mathfrak{n} \ . \end{cases}$$

Similar to classical case, one can deduce the Ramanujan bound $|c(\mathfrak{p})| \leq 2N(\mathfrak{p})^{1/2}$, from which it follows that $L(F, s)$ converges for $\mathrm{Re}(s) > 3/2$.

We now consider additive twists of this L-function. Fix $r = a/c \in K$. If F is a plusform, then we define $L(F, s, r)$ as

$$L(F, s, r) := \sum_{0 \neq \alpha \in \mathcal{O}_K} \frac{c(\alpha)}{N(\alpha)^s} \psi\left(\frac{\alpha r}{\sqrt{d_K}}\right) = \sum_{(\alpha)} \frac{c((\alpha))}{N((\alpha))^s} \widetilde{\psi}\left(\frac{\alpha r}{\sqrt{d_K}}\right)$$

where the second sum is over all ideals (α) and

$$\widetilde{\psi}(z) := \frac{1}{|\mathcal{O}_K^*|} \sum_{\epsilon \in \mathcal{O}_K^*} \psi(\epsilon z)$$

is invariant over generators of an ideal.

We form the Mellin transform of F by multiplying by y^{2s-2} and integrating along a vertical imaginary axis. For $s \in \mathbb{C}$ and $r \in K$, we define

$$\Lambda(F, s, r) := \int_r^{j\infty} y^{2s-2} F \cdot \beta = \int_0^\infty y^{2s-2} F_1(r, y) \frac{dy}{y} \ .$$

The rapid decay of $F(z, y)$ in the cusps ensures that $\Lambda(F, s, r)$ is an entire function of $s \in \mathbb{C}$.

We note that if F is a plusform, then we can write modular symbols as central values of twisted L-function:

$$\langle r \rangle = \Lambda(F, 1, r) = \int_r^\infty F \cdot \beta \ . \tag{3.54}$$

We obtain analytic continuation and functional equation for $L(F, s, r)$.

Lemma 3.7.1. *Let F be a plusform in $S(\mathfrak{n})$, where $\mathfrak{n}$ is a square-free ideal in $\mathcal{O}_K$. Then*

(a) For $\mathrm{Re}(s) > 3/2$, *we have*

$$\Lambda(F, s, r) = \frac{1}{4} \left(\frac{|c|\sqrt{|d_K|}}{2\pi} \right)^{2s} \Gamma(s)^2 \ L(F, s, r) \ .$$

(b) Write $\mathfrak{n} = \mathfrak{e}\mathfrak{f}$, where $\mathfrak{f} = \mathfrak{n} + (c)$. Let $\mathfrak{e} = (e)$. Denote by w_e the eigenvalue of the Fricke involution W_e acting on F. Then we have the following functional equation:

$$\Lambda(F, s, a/c) = -w_e N(\mathfrak{e})^{1-s} \Lambda\left(F, 2-s, -\frac{\overline{ea}}{c}\right) \ ,$$

where $\overline{ea}$ is the inverse of ea in $(\mathcal{O}_K/(c))^*$.

(c) With the same notation, we have $\langle a/c \rangle = -w_\epsilon \langle -\overline{ea}/c \rangle$.

We now quote [21, p. 415] and note that if F is a plusform in $S_2(\mathfrak{n})$, then the image of the map

$$I_F : \Gamma_0(\mathfrak{n}) \to \mathbb{C} \ , \quad I_F(\gamma) = \int_A^{\gamma A} F \cdot \beta$$

is a discrete, nontrivial subgroup of $\mathbb{R}$, hence of the form $\Omega(F)\mathbb{Z}$, for some real $\Omega(F)$. In [21], Cremona provides an algorithm for computing $\Omega(F)$. We show that for a fixed newform F, the values in the image of the map I_F are normally distributed with the required normalisation and ordering.

We have the following description of equivalent $\Gamma_0(\mathfrak{n})$-equivalent points in K, as in [22, Proposition 4.2.2] or [23, Lemma 2.2.7]:

Proposition 3.7.1. *Let $\frac{p_1}{q_1}, \frac{p_2}{q_2} \in K$ be written in their lowest terms. The following are equivalent:*

1. *There exists $\gamma \in \Gamma_0(\mathfrak{n})$ such that $\gamma\left(\frac{p_1}{q_1}\right) = \frac{p_2}{q_2}$;*
2. *There exists $u \in \mathcal{O}_K^*$ such that $s_1 q_2 \equiv u^2 s_2 q_1 (mod\ (q_1 q_2) + \mathfrak{n})$, where $p_k s_k \equiv 1 (mod\ (q_k))$, for $k = 1, 2$.*

Hence we can provide the following description for the inequivalent cusps for $\Gamma_0(\mathfrak{n})$, where $\mathfrak{n}$ is square-free. For each ideal $\mathfrak{d}|\mathfrak{n}$, we fix some $d \in \mathcal{O}_K$ such that $(d) = \mathfrak{d}$. Then a complete set of inequivalent cusps are given by $a_{\mathfrak{d}} = 1/d$ with $\mathfrak{d}|\mathfrak{n}$. If $\mathfrak{d} = \mathfrak{n}$, then $1/d$ is equivalent to the cusp at infinity. Moreover,

$$R_{\infty\mathfrak{d}} = \left\{ \frac{a}{c} \bmod \mathcal{P}_K \ : \ a \in (\mathcal{O}_K/(c))^*, (c) + \mathfrak{n} = \mathfrak{d} \right\} .$$

and

$$\langle r \rangle_{\infty\mathfrak{d}} = \int_{1/d}^{r} F \cdot \beta = \int_{1/d}^{j\infty} F \cdot \beta + \langle r \rangle.$$

Also, for all cusps $\mathfrak{d}$, we have that $[\Gamma_{\mathfrak{d}} : \Gamma'_{\mathfrak{d}}] = |\mathcal{O}_K^*|/2$. In particular, $|\mathcal{O}_{\mathbb{Q}(i)}^*| = 4$, $|\mathcal{O}_{\mathbb{Q}(\sqrt{-3})}^*| = 6$ and $|\mathcal{O}_K^*| = 2$ for all other quadratic imaginary number fields.

We note that we now have all the ingredients to derive Corollary 3.1.1 from Theorem 3.1.2. Indeed, from [32, Theorem 6.1.1] we see that the covolume of $\mathrm{PSL}_2(\mathcal{O}_K)$ is

$$\mathrm{vol}(\mathrm{PSL}_2(\mathcal{O}_K)) = \frac{|d_K|^2}{4\pi^2} \zeta_K(2)$$

and similarly as in the 2-dimensional case, we can deduce

$$[\mathrm{PSL}_2(\mathcal{O}_K) : \Gamma_0(\mathfrak{n})] = \prod_{\mathfrak{p}|\mathfrak{n}} (1 + |\mathfrak{p}|) .$$

Finally, the Petersson norm of F is given by

$$\|F\|^2 = \langle F \cdot \beta, F \cdot \beta \rangle = \int_{\Gamma\backslash\mathbb{H}^3} (2|F_1|^2 + |F_2|^2 + 2|F_3|^2) dv .$$

Putting all together, we deduce that the constant C_F in Corollary 3.1.1 is given by

$$C_F = \frac{4\pi^2 \|F\|^2}{|d_K|^2 \zeta_K(2) \prod_{\mathfrak{p}|\mathfrak{n}} (1 + |\mathfrak{p}|)} \ . \tag{3.55}$$

This completes the proofs of Theorem 3.1.2 and Corollary 3.1.1.

Chapter 4

Residual equidistribution of modular symbols and generalisations to higher dimensions

This chapter is mainly based on [19], joint work with Asbjørn Nordentoft.

4.1 Introduction

In this chapter we develop a number of new results regarding the distribution of modular symbols modulo primes and generalisations to higher dimensional hyperbolic spaces.

1. Firstly, 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 calculate the variance of the distribution and show a surprising bias for large p.

3. We show some particular cases of the full conjecture using connections with Eisenstein congruences.

4. As an application of our method, we obtain a residual equidistribution result for Dedekind sums.

5. Lastly, 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 [93] established that such torsion classes have associated Galois representations. This was actually our original motivation for studying the higher dimensional cases, where we can consider torsion classes. Furthermore, Bergeron and Venkatesh [4] have conjectured that, at least in the three dimensional case, there is an abundance of torsion in the relevant cohomology group. In this chapter we are able to shed light on the distribution properties of these cohomology classes. In Section 4.6 we will survey what is known about the dimensions of the cohomology groups, which our results apply to.

4.1.1 Results for modular symbols

Let us state the result in the simplest case for the two 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

$$\mathbb{Q} \ni r \mapsto \langle r, f \rangle := 2\pi i \int_r^{i\infty} f(z)dz, \tag{4.1}$$

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-normalised newform. Then by [66, Sec. 1], there exist periods $\Omega_{f,+}$ and $\Omega_{f,-}$ such that for all $a/q \in \mathbb{Q}$ with $q \equiv 0 \bmod N$, we have $\mathfrak{m}_f^{\pm}(a/q) \in \mathbb{Z}$ with full image, where

$$\mathfrak{m}_f^{\pm}(a/q) := \frac{1}{\Omega_{f,\pm}} \left(\langle a/q, f \rangle \pm \langle -a/q, f \rangle \right). \tag{4.2}$$

Given a basis of Hecke newforms $f_1, \ldots, f_d$ and a prime p, we can consider the map

$$r \mapsto \mathfrak{m}_{N,p}(r) := (\mathfrak{m}_{f_1}^{+}(r), \mathfrak{m}_{f_1}^{-}(r), \ldots, \mathfrak{m}_{f_d}^{-}(r), r) \in (\mathbb{Z}/p\mathbb{Z})^{2d} \times (\mathbb{R}/\mathbb{Z})$$

as a random variable defined on the outcome space

$$\Omega_{Q,N} := \{a/q \mid 0 < a < q \leq Q, (a,q) = 1, N|q\} \tag{4.3}$$

endowed with the uniform probability measure. Then we have the following equidistribution result.

Theorem 4.1.1. *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 $Q \to \infty$. More precisely,* *for any fixed* $\mathbf{a} \in (\mathbb{Z}/p\mathbb{Z})^{2d}$ *and any interval* $I \subset \mathbb{R}/\mathbb{Z}$*, we have*

$$\frac{\#\left\{a/q \in \Omega_{Q,N} \cap I \mid (\mathfrak{m}^+_{f_1}(a/q), \dots, \mathfrak{m}^-_{f_d}(a/q)) \equiv \mathbf{a} \bmod p\right\}}{\#\Omega_{Q,N}} = \frac{|I|}{p^{2d}} + o(1)$$

as $Q \to \infty$*.*

Remark 4.1.2. Similarly, we can prove equidistribution modulo p^n (see Theorem 4.1.8 below). This translates to the fact that the random variables $(\mathfrak{m}^+_{f_1}, \dots, \mathfrak{m}^-_{f_d})$ considered as maps $\Omega_{Q,N} \cap I \to \mathbb{Q}_p$ are asymptotically distributed with respect to the (multivariate) standard p-adic Gaußian (as defined in for instance [115]).

The next natural question is to ask how well the values equidistribute. We answer this by studying the "variance" of the residual distribution modulo p of the random variables $\mathfrak{m}^\pm_f$ on the sample space $\Omega_{Q,N}$. Furthermore, we show an analogue of Chebyshev's bias for large p, in the sense that the modular symbols are "biased" towards the residue class $0 \bmod p$.

Theorem 4.1.3. *For large enough* p*, there exist constants* $c_p, \delta_p > 0$ *such that*

$$\sum_{a \in \mathbb{Z}/p\mathbb{Z}} \left(\frac{\#\{b/q \in \Omega_{Q,N} \mid \mathfrak{m}^\pm_f(b/q) \equiv a \bmod p\}}{\#\Omega_{Q,N}} - \frac{1}{p} \right)^2 \sim c_p Q^{-\delta_p}$$

as $Q \to \infty$*. Moreover, as* $p \to \infty$*, we have that* $c_p = 2/p^2 + O(p^{-2})$ *and* $\delta_p \to 0$*.*

Furthermore, for p *large enough, we have for* Q *large enough (depending on* p*) that:*

$$\#\{b/q \in \Omega_{Q,N} \mid \mathfrak{m}^\pm_f(b/q) \equiv a \bmod p\} \leq \#\{b/q \in \Omega_{Q,N} \mid \mathfrak{m}^\pm_f(b/q) \equiv 0 \bmod p\}, \qquad (4.4)$$

with equality if and only if $a \equiv 0 \mod p$*.*

Remark 4.1.4. We explicitly evaluate the constants c_p and δ_p and moreover we obtain asymptotics for the deviation from the mean for different residue classes when p is large, see Section 4.5.3 for more details.

We can also show that some specific cases of the conjecture of Mazur and Rubin hold, that is without taking an extra average. We state here the result in the simplest case and refer to Section 4.3 for the more general case.

Theorem 4.1.5. *Let $f \in \mathcal{S}_2(\Gamma_0(11))$ be the unique Hecke normalised cusp form of weight 2 and level 11. Then the values of $\mathfrak{m}_f^+$ on $\{\frac{a}{q} \mid (a,q)=1, 0<a<q\}$ equidistribute exactly modulo 5 for all $q \equiv 0 \bmod 11$. That is, each residue class modulo 5 is covered exactly $\phi(q)/5$ times by the values $\mathfrak{m}_f^+(a/q)$.*

As a consequence of the method developed to study the special case as in Theorem 4.1.5, we deduce a residual equidistribution result for classical Dedekind sums, given by

$$s(a,q) := \sum_{k=1}^{q} ((k/q))((ak/q))$$

with (()) the *sawtooth function.* Dedekind sums are important objects in number theory, they appear for instance in the functional equation of the eta function $\eta(z)$. We allow for both an "algebraic" and "archimedian" restriction on (a,q). Our result supplements the vast literature on the archimedean distributional properties of Dedekind sums, see [36], [10] for surveys of results.

Corollary 4.1.6. *Let $N, p \geq 5$ be primes such that $p|N-1$ and $H \leq (\mathbb{Z}/N\mathbb{Z})^\times$ the unique subgroup of index p. Fix some class $a_0 \in (\mathbb{Z}/N\mathbb{Z})^\times$ and some interval $I \subset \mathbb{R}/\mathbb{Z}$. Then the values of*

$$s(a,Nq) - s(a,q) - \frac{(N-1)(a+\overline{a})}{12q}$$

(where $\overline{a}a \equiv 1 \bmod Nq$) on the outcome space

$$\big\{(a,q) \mid 0<q\leq Q, a \in (\mathbb{Z}/Nq\mathbb{Z})^\times, a \in a_0H, a/q \in I\big\}$$

are all p-integral and equidistribute mod p 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 $H^1(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$. This view point is useful for generalisations.

Remark 4.1.7. We note that in [57], 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 4.3.3 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 [79] or [18]. The outcome space $\Omega_{Q,N}$ above is, however, very natural from the cohomological perspective and for simplicity we will restrict to this case.

4.1.2 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 $H^1(\Gamma, S^1)$ (or $H^1(\Gamma, \mathbb{R}/\mathbb{Z})$). These cohomology groups have been studied in many contexts ([89], [32, Chap. 7]) and especially the case $n = 2$ is very appealing as it corresponds to Kleinian groups due to the exceptional isomorphism $\mathrm{SO}(3,1) \cong \mathrm{SL}_2(\mathbb{C})$.

4.1.2.1 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 $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ (the group cohomology classes trivial on Γ'_∞, see Section 4.6.2 for detailed definitions).

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 *Vahlen model* SV_{n-1} for the group of isometries of $\mathbb{H}^{n+1}$ consisting of 2×2 matrices over a specific Clifford algebra, introduced in [1] (see Section 2.7.2 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 $\mathrm{SV}_1 = \mathrm{SL}_2(\mathbb{C})$. We define the following outcome space:

$$T_\Gamma(X) = \left\{\gamma \in \Gamma'_\infty\backslash\Gamma/\Gamma'_\infty \mid 0 < |c_\gamma| < X\right\}, \tag{4.5}$$

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 relevant Clifford algebra. This generalizes the outcome space (4.3) above and the ones considered for $n = 1$ in [79], [74] and for $n = 2$ in [18].

Now let $\omega_1, \ldots, \omega_d$ be elements of $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ in *general position*, meaning that 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 $\mathbb{F}_p$-basis for $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{Z}/p\mathbb{Z})$. We notice that the image of any $\omega \in H^1(\Gamma, \mathbb{R}/\mathbb{Z})$ is either dense in $\mathbb{R}/\mathbb{Z}$ or finite (recall that ω defines an additive character $\Gamma \to \mathbb{R}/\mathbb{Z}$). In the first case we put $J_\omega = \mathbb{R}/\mathbb{Z}$ and in the latter case we put $J_\omega = \mathbb{Z}/m\mathbb{Z}$, where m is the cardinality of the image of ω. We equip $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$ with the obvious choices of probability measures, Lebesgue and uniform respectively. Finally associated to $\gamma \in \Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$, we define the invariant $\gamma\infty \in (\mathbb{R}^n \cup \{\infty\})/\Lambda$ using the action of $\mathrm{SO}(n+1, 1)$ on the boundary of $\mathbb{H}^{n+1}$, see Section 2.7.3 for more details. Then we have the following distribution result.

Theorem 4.1.8. *Let $\omega_1, \ldots, \omega_d \in H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ be in general position. The random variables $\gamma \mapsto (\omega_1(\gamma), \ldots, \omega_d(\gamma), \gamma\infty)$ defined on the sample spaces $T_\Gamma(X)$ are asymptotically uniformly distributed on $\prod_{i=1}^d J_{\omega_i} \times (\mathbb{R}^n/\Lambda)$ as $X \to \infty$. More precisely, for any fixed (continuity) subsets $A_i \subset J_{\omega_i}$ and $B \subset \mathbb{R}^n/\Lambda$, we have*

$$\frac{\#\left\{\gamma \in T_\Gamma(X) \mid (\omega_1(\gamma), \ldots, \omega_d(\gamma)) \in \prod_{i=1}^d A_i, \gamma\infty \in B\right\}}{\# T_\Gamma(X)} = \prod_{i=1}^d \frac{|A_i|}{|J_{\omega_i}|} \cdot \frac{|B|}{\mathrm{vol}(\mathbb{R}^n/\Lambda)} + o(1)$$

as $X \to \infty$.

Remark 4.1.9. The Vahlen model has been used before to study automorphic forms on $\mathbb{H}^{n+1}$, for example by Elstrodt, Grunewald, and Mennicke [31] to prove a generalisation of the Selberg Conjecture regarding the first non-zero eigenvalue of the Laplacian and by Södergren [100] for proving equidistribution of horospheres on $\mathbb{H}^{n+1}$.

Remark 4.1.10. 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 4.6 for a survey of results on the size of $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$.

4.1.2.2 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 4.1.11. *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 4.1.12. In the case of Theorem 4.1.11, we can remove the assumption that Γ has cusps. In fact the proof becomes more complicated in the presence of cusps.

4.2 Idea of proof

We will sketch the proof of Theorem 4.1.1 in the simplest case, which is the one dealt with in [57], where we consider only one cusp form for $\mathbb{H}^2$ and no restrictions on the location of $r = a/q$ in $\mathbb{R}/\mathbb{Z}$. 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 [77] 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 $m \in \mathbb{Z}$.

Let $f \in \mathcal{S}_2(\Gamma_0(N))$ be a Hecke eigenform of weight 2 and level N and let $\mathfrak{m}_f^\pm : \Gamma_0(N) \to \mathbb{Z}$ be the associated normalised modular symbols defined above. Recall that this defines a non-trivial additive character $\Gamma_0(N) \to \mathbb{Z}$. We would like to show that the values of $\mathfrak{m}_f^\pm$ on the set $\Omega_{Q,N} = \{a/q \mid 0 < a < q \leq Q, (a,q) = 1, N|q\}$ equidistribute mod p 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

$$\chi_l(\gamma) := e^{2\pi i \mathfrak{m}_f^{\pm}(\gamma) l/p}, \quad \gamma \in \Gamma_0(N).$$

By Weyl's Criterion [48, 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

$$\sum_{a/q \in \Omega_{Q,N}} \chi_l(a/q) = o(Q^2),$$

as $\mathbb{Q} \to \infty$, where $\chi_l(a/q) := \chi_l(\gamma)$ with $\gamma \in \Gamma_0(N)$ such that $\gamma\infty = a/q$.

Now, the key observation is 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 $\Gamma_\infty = \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right\rangle$. This Eisenstein series defines a holomorphic function for $\operatorname{Re} s > 1$ and by the work of Selberg [94, 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

$$L_l(s) := \sum_{c>0, N|c} \left(\sum_{0<d<c, (c,d)=1} \overline{\chi_l} \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right) \right) c^{-2s},$$

where $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is any 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 $\operatorname{Re} s > 1-\delta$ 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 scattering matrices 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 [79].

4.3 Some special cases of the conjecture of Mazur and Rubin

In this section we will consider certain special cases of the conjectures of Mazur and Rubin (and the generalization to $\mathbb{H}^3$), which we can resolve *without* taking an extra average. These special cases correspond to the fact that Hecke characters define unitary characters of congruence subgroups, which in turn are connected to Eisenstein congruences as studied intensively by Mazur in [63, Section 9] and [64].

First of all we will define the relevant cohomology classes and introduce the Hecke operators in this context. Recall that for a discrete, cofinite subgroup $\Gamma \subset \mathrm{SL}_2(k)$ with $k = \mathbb{R}$ or $\mathbb{C}$ and an element $\alpha \in \tilde{\Gamma}$ of the commensurator of Γ, we have a decomposition

$$\Gamma\alpha\Gamma = \bigsqcup_{i=1}^{d} \Gamma\alpha_i$$

for some $\alpha_1, \ldots, \alpha_d \in \tilde{\Gamma}$. Using this we define the *Hecke operator* T_α acting on the cohomology group $H^1(\Gamma, X)$ with X a trivial Γ-module as:

$$(T_\alpha \omega)(\gamma) := \sum_{i=1}^{d} \omega(\gamma_i), \tag{4.6}$$

where $\alpha_i\gamma = \gamma_i\alpha_{\sigma(i)}$ with $\gamma_i \in \Gamma$ and σ some permutation of $\{1, \ldots d\}$.

We will consider the case of congruence subgroups

$$\Gamma_0(\mathfrak{f}) = \{\gamma \in \mathrm{SL}_2(\mathcal{O}_K) \mid \gamma \equiv \left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right) \bmod \mathfrak{f}\},$$

where K is equal to $\mathbb{Q}$ or an imaginary quadratic extension thereof and $\mathfrak{f}$ is a non-trivial ideal of $\mathcal{O}_K$. In this case we have $\widetilde{\Gamma_0}(\mathfrak{f}) = \mathrm{GL}_2(K)$ and the parabolic subgroup fixing ∞ is $\Gamma'_\infty = \left(\begin{smallmatrix} \pm 1 & \mathcal{O}_K \\ 0 & \pm 1 \end{smallmatrix}\right)$. We say that a Hecke operator is *good* if it is of the form T_α, where $\alpha = \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right)$ with $\gcd(\mathfrak{f}, (a)) = 1$.

Proposition 4.3.1. *Let m be an odd integer diving $|(\mathcal{O}_K/\mathfrak{f})^\times|$. Then there exists a class $\omega \in H^1_{\Gamma'_\infty}(\Gamma_0(\mathfrak{f}), \mathbb{Z}/m\mathbb{Z})$, which is an eigenvector for all good Hecke operators and such that for all $a \in \mathbb{Z}/m\mathbb{Z}$ and $c_0 \in \mathfrak{f}$, it satisfies*

$$\frac{\#\{\gamma \in \Gamma_\infty \backslash \Gamma_0(\mathfrak{f})/\Gamma_\infty \mid c_\gamma = c_0, \omega(\gamma) = a\}}{\#\{\gamma \in \Gamma_\infty \backslash \Gamma_0(\mathfrak{f})/\Gamma_\infty \mid c_\gamma = c_0\}} = \frac{1}{m}, \tag{4.7}$$

where c_γ denotes the lower left entry of γ.

Proof. Let $\chi : (\mathcal{O}_K/\mathfrak{f})^\times \to \mathbb{C}^\times$ be a unitary Hecke character of order m. Then we define a character of $\Gamma_0(\mathfrak{f})$ by

$$\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \mapsto \chi(d). \tag{4.8}$$

This character is clearly trivial on Γ'_∞ since the order m of χ is odd, and thus (4.8) defines an element $\omega_\chi \in H^1_{\Gamma'_\infty}(\Gamma_0(\mathfrak{f}), \mathbb{Z}/m\mathbb{Z})$. For T_α a good Hecke operator, it is easy to check that α_i (with notation as in (4.6)) can all be chosen of the form $\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)$ with determinant equal to the determinant of α (thus the diagonal entries are coprime to $\mathfrak{f}$). Combining this with $\gamma_i = \alpha_i \gamma \alpha_{\sigma(i)}^{-1}$, one easily sees that

$$(T_\alpha(\omega_\chi))(\gamma) = d\omega_\chi(\gamma),$$

where $d = |\Gamma_0(\mathfrak{f}) \backslash \Gamma_0(\mathfrak{f}) \alpha \Gamma_0(\mathfrak{f})|$. This shows that ω_χ is an Hecke eigenclass with eigenvalue d, as wanted.

Finally, recall the basic fact that a set of representatives of $\Gamma_\infty \backslash \Gamma_0(\mathfrak{f})/\Gamma_\infty$ is given by

$$\left\{ \begin{pmatrix} * & * \\ c & d \end{pmatrix} \mid c \in \mathfrak{f}, d \in (\mathcal{O}_K/(c))^\times \right\}.$$

From this, the equidistribution statement (4.7) follows directly. □

It is a natural question to ask how the cohomology classes constructed above are related to the modular symbols defined in (4.2). To tackle this we need to understand so-called *Eisenstein congruences*, which have been studied intensively by Mazur [63]. We will now introduce some required terminology and refer to [63] for a detailed account: We say that a pair of primes (N,p) with $N,p \geq 5$ and $p|N-1$ is *admissible* if the local ring $\mathbb{T}_{\mathfrak{P}}$ has rank 1 over $\mathbb{Z}_p$ where $\mathbb{T}$ is the Hecke algebra of level N and $\mathfrak{P} \subset \mathbb{T}$ is the Eisenstein prime corresponding to p. In classical terms (N,p) being admissible means that there is a unique cuspidal Hecke eigenform of level N which is congruent to the Eisenstein series of weight 2 (i.e. $f \in \mathcal{S}_2(\Gamma_0(N))$ s.t. the Hecke eigenvalues satisfy $\lambda_f(l) \equiv l+1 \bmod p$ for primes $l \neq N$ and $Uf = -f$ where U is the Hecke operators at N). By a computation of Merel [68], (N,p) is admissible exactly if

$$\prod_{k=1}^{p-1}((k(N-1)/p)!)^k \tag{4.9}$$

is a p-power in $(\mathbb{Z}/N\mathbb{Z})^\times$. Note that all pairs of primes (N,p) with $N < 250$ are admissible unless $N = 31, 103, 127, 131, 181, 199, 211$ (see the remark on [63, p. 141]). In the admissible case we have the following strengthening of Proposition 4.3.1 (see [63, Chapter II, Proposition 18.8] for a very related result).

Theorem 4.3.2. *For an admissible pair of primes (N,p) with $N,p \geq 5$ and $p|N-1$, there exists a Hecke eigenform $f \in \mathcal{S}_2(\Gamma_0(N))$ of weight 2 and level N such that the values of $\mathfrak{m}_f^+$ (defined as in (4.2)) on $\{\frac{a}{q} \mid (a,q)=1, 0<a<q\}$ equidistribute exactly modulo p for $q \equiv 0 \bmod p$.*

Proof. Let χ be a Dirichlet character mod N of order $p|N-1$. Then by Proposition 4.3.1 we have an associated cohomology class $\omega_\chi \in H^1_{\Gamma_\infty}(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$ which equidistributes as above and such that $T_l\omega_\chi = (l+1)\omega_\chi$ for all primes $l \neq N$, where T_l is the Hecke operator corresponding to the matrix $\left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Furthermore, ω_χ satisfies $U\omega_\chi = -\omega_\chi$, where U is the Hecke operator at the bad prime N given by conjugation by $\left(\begin{smallmatrix} 0 & 1 \\ N & 0 \end{smallmatrix}\right)$. Also ω_χ is trivial on the stabilizer $\langle \left(\begin{smallmatrix} \pm 1 & 0 \\ 1 & \pm 1 \end{smallmatrix}\right) \rangle$ of the cusp 0 (using that the order of χ is odd) and thus ω_χ defines a parabolic cohomology class.

As there is no p-torsion in $\Gamma_0(N)$ (since $p > 3$), $H^1_P(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$ is a $2g$-dimensional vector space over $\mathbb{F}_p$ with $g = \dim_{\mathbb{C}}(\mathcal{S}_2(\Gamma_0(N)))$, which carries an action of the Hecke algebra. An element which is annihilated by $T_\ell - \ell - 1$ for all primes $\ell \neq N$ and by $U+1$ corresponds exactly to a cusp form congruent to the weight 2 Eisensten series. By the assumption that (N,p) is admissible we know that there exists a *unique* such Hecke eigenform $f \in \mathcal{S}_2(\Gamma_0(N))$. We conclude that ω_χ is a linear combination of $\mathfrak{m}_f^{\pm}$.

Finally, we recall that $H^1_P(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$ can be diagonalized by the involution ι given by conjugation with $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (here we need $p > 2$), which follows from e.g. [65, Sec. 1]. We see directly that the eigenvalue of ω_χ under the action of ι is $+1$. Thus we conclude that $\mathfrak{m}_f^+ = m \cdot \omega_\chi$ for some $m \in (\mathbb{Z}/p\mathbb{Z})^\times$. This gives the wanted result. □

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 4.3.3. Strictly speaking the conjecture of Mazur and Rubin [65] is 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, but the above seems like the natural generalization of the conjecture to this case.

Remark 4.3.4. The assumption that N is prime is essential for the results of [63] to apply. For composite level (and for imaginary quadratic fields) the situation becomes much more complicated as *multiplicity one* might fail (see e.g. [109]).

4.4 Twisted Eisenstein series for $\mathbb{H}^{n+1}$

Let $\Gamma < \mathrm{SV}_{n-1}$, Γ'_∞ and Λ be as described in Section 2.7. We now fix χ a unitary character of Γ which is trivial on Γ'_∞. From this we define the twisted Eisenstein series

$$E(P,s,\chi) = \sum_{\Gamma'_\infty \backslash \Gamma} \overline{\chi(\gamma)} y(\gamma P)^s. \tag{4.10}$$

It is absolutely convergent for $\mathrm{Re}(s) > n$ and 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 [14, Ch. 6].

For $\mu,\nu\in\Lambda^*$ and $c\in C(\Gamma)$, we define the generalised Kloosterman sum as in [31, Section 4] using the Vahlen model:

$$S(\mu,\nu,c,\chi) := \sum_{\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty} \overline{\chi}\left(\begin{pmatrix} a & b\\ c & d\end{pmatrix}\right) e\left(\langle ac^{-1},\mu\rangle + \langle dc^{-1},\nu\rangle\right) \tag{4.11}$$

$$= \sum_{\substack{\gamma\in\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty\\ c_\gamma=c}} \overline{\chi(\gamma)} e(\langle\gamma\infty,\mu\rangle + \langle(\gamma^{-1}\infty)^*,\nu\rangle), \tag{4.12}$$

where c_γ is the lower-left entry of γ in the Vahlen model. We now calculate the Fourier expansion of the Eisenstein series using the techniques developed in [32, p. 111–113] and [31, p. 676–678]. We obtain

$$\begin{aligned} E(P,s,\chi) =& [\Gamma_\infty:\Gamma'_\infty]y^s + y^{n-s}\frac{\pi^{n/2}\Gamma\left(s-\frac{n}{2}\right)}{\mathrm{vol}(\Lambda)\Gamma(s)}L(s,\chi)\\ &+\frac{2\pi^s y^{n/2}}{\mathrm{vol}(\Lambda)\Gamma(s)}\sum_{\mu\in\Lambda^*\backslash\{0\}} L(s,\mu,\chi)|\mu|^{s-n/2}K_{s-n/2}(2\pi|\mu|y)e(\langle x,\mu\rangle), \end{aligned} \tag{4.13}$$

where

$$L(s,\chi) := \sum_{\gamma\in T_\Gamma}\frac{\overline{\chi}(\gamma)}{|c_\gamma|^{2s}} = \sum_{c\in C(\Gamma)}\frac{S(0,0,c,\chi)}{|c|^{2s}}\ , \tag{4.14}$$

and for $\mu\neq 0$,

$$L(s,\chi,\mu) := \sum_{\gamma\in T_\Gamma}\overline{\chi}(\gamma)\frac{e(\langle d_\gamma c_\gamma^{-1},\mu\rangle)}{|c_\gamma|^{2s}} = \sum_{c\in C(\Gamma)}\frac{S(0,\mu,c,\chi)}{|c|^{2s}}. \tag{4.15}$$

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 (4.13) is closely related to [31, 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 [14, Ch. 6, Prop. 1.42]:

$$E(P^{\mathfrak{a}},s,\chi) = \phi_{\mathfrak{a}}(s)(y^{\mathfrak{a}})^{n-s} + \sum_{\mu\in\Lambda_{\mathfrak{a}}^*\backslash\{0\}}\phi_{\mathfrak{a}}(s,\mu)(y^{\mathfrak{a}})^{n-s}K_{s-n/2}(2\pi n|\mu|y^{\mathfrak{a}})e(\langle x^{\mathfrak{a}},\mu\rangle),$$

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 (2.40)).

Remark 4.4.1. By inverting γ in the definition of $L(s,\chi,\mu)$, we observe that

$$L(s,\chi,\mu) = \sum_{\gamma\in T_\Gamma} \overline{\chi}(\gamma)\frac{e(\langle(\gamma^{-1}\infty)^*,\mu\rangle)}{|c_\gamma|^{2s}} = \sum_{\gamma\in T_\Gamma} \chi(\gamma)\frac{e(\langle\gamma\infty,\mu\rangle)}{|c_\gamma|^{2s}}. \tag{4.16}$$

4.4.1 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 [89] and [14, 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 $\mathrm{Re}\, s > n/2$ form a subset of $\Omega(\chi)$ (exactly the non-cuspidal part of the discrete spectrum). Moreover, we can see from [14, Ch 6, p. 37] that for χ trivial, we have

$$\mathrm{Res}_{s=n} E(P, s) = \frac{[\Gamma_\infty : \Gamma'_\infty]\mathrm{vol}(\Lambda)}{\mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})} \ . \tag{4.17}$$

4.4.2 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 [32, Theorem 4.1.7] they spend several pages making it rigorous in the three dimensional case. Instead we will give an argument using the Fourier expansion and the mean value theorem for harmonic functions.

Lemma 4.4.1. *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 [14, Ch. 6, p.10] that the Fourier expansion of u takes the form

$$c_{1,\mathfrak{a}} + c_{2,\mathfrak{a}}(y^{\mathfrak{a}})^n + \sum_{\mu \in \Lambda_{\mathfrak{a}}^* \setminus \{0\}} a_{u,\mathfrak{a}}(\mu)(y^{\mathfrak{a}})^{n/2} K_{n/2}(2\pi n|\mu|y)e(\langle x, \mu \rangle).$$

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 4.4.2. *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 (4.13), we know that for $\mu \in \Lambda^* \setminus \{0\}$

$$L(s, \mu, \chi) = \frac{\Gamma(s)}{2\pi^s y^{n/2}|\mu|^{s-n/2} K_{s-n/2}(2\pi|\mu|y)} \int_{\mathcal{P}} E((x,y), s, \chi)e(-\langle x, \mu \rangle)dx,$$

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 defines 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 4.4.1, 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

$$\operatorname{Res}_{s=n} L(s,0) = \frac{\Gamma(n)}{\pi^{n/2}\Gamma\left(\frac{n}{2}\right)}\int_{\mathcal{P}} \frac{[\Gamma_\infty:\Gamma'_\infty]\operatorname{vol}(\Lambda)}{\operatorname{vol}(\Gamma\backslash\mathbb{H}^{n+1})}dx = \frac{[\Gamma_\infty:\Gamma'_\infty]\Gamma(n)\operatorname{vol}(\Lambda)^2}{\pi^{n/2}\Gamma\left(\frac{n}{2}\right)\operatorname{vol}(\Gamma\backslash\mathbb{H}^{n+1})},$$

as wanted. □

In order to obtain bounds on vertical lines for our generating series, we will use ideas due to Colin de Verdière [15], 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 [79] and [18]. 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 (2.40). 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 [47, Lemma A.4], we conclude that

$$\|H(\cdot, s, \chi)\|_{L^2} \ll_\varepsilon 1,$$

when s is bounded at least ε away from $\Omega(\chi)$. We can now write

$$E(P, s, \chi) = H(P, s, \chi) + h(y)y^s, P \in \mathcal{F} \tag{4.18}$$

where we have good control on the L^2-norm of $H(P, s, \chi)$. We will use this to obtain bounds on vertical lines for the Fourier coefficients of $E(P, s, \chi)$, mimicking [74, Section 4.4].

Proposition 4.4.3. *Let $\mu \in \Lambda^*$. 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 $\Omega(\chi)$.

Proof. We have

$$L(s, \mu, \chi) = \int_{\mathcal{P}} f_s(y, \mu)E((x, y), s, \chi)e(-\langle x, \mu\rangle)dx - \mathbf{1}_{\mu=0}[\Gamma_\infty : \Gamma'_\infty]y^s f_s(y, \mu), \tag{4.19}$$

where $\mathbf{1}_{\mu=0}$ is 1 if $\mu = 0$ and 0 otherwise and

$$f_s(y, \mu) = \begin{cases} \Gamma(s)\left(2\pi^s y^{n/2}|\mu|^{s-n/2}K_{s-n/2}(2\pi n|\mu|y)\right)^{-1}, & \mu \neq 0, \\ \Gamma(s)\left(y^{n-s}\pi^{n/2}\Gamma(s - n/2)\right)^{-1}, & \mu = 0. \end{cases}$$

The idea is now to bound the right hand side of (4.19) using (4.18). 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

$$\begin{aligned}
&\int_Y^{Y+1}\int_{\mathcal{P}} f_s(y,\mu)H((x,y),s,\chi)e(-\langle\mu,x\rangle)dxdy \\
&\leq \left(\int_Y^{Y+1}\int_{\mathcal{P}}|H((x,y),s,\chi)|^2dxdy\right)^{1/2}\left(\int_Y^{Y+1}\int_{\mathcal{P}}|f_s(y,\mu)|^2dxdy\right)^{1/2} \\
&\ll \|H(\cdot,s,\chi)\|_{L^2}\left(\int_Y^{Y+1}|f_s(y,\mu)|^2dy\right)^{1/2},
\end{aligned}$$

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 [47, (B.32)] and [47, (B.34)]) to obtain a good approximation. By applying Stirling's approximation, this gives for $s=\sigma+it$ with $t\gg 1$

$$\begin{aligned}
K_{s-n/2}(2\pi|\mu|y) &= \frac{\pi^{1/2}t^{\sigma-n/2-1/2}e^{\pi t/2}\left(\frac{t}{e}\right)^{it}}{2\sqrt{2}\sin(\pi(s-n/2))}\left(\pi|\mu|y\right)^{-s+n/2}\left(1+O_{\mu,y}(t^{-1})\right) \\
&\gg_{\mu,y} e^{-\pi t/2}t^{\sigma-n/2-1/2},
\end{aligned}$$

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 $\Omega(\chi)$, as wanted. □

Using this we deduce the following asymptotic expression using a standard complex analysis argument. See [18, p. 20–21] or [74, Appendix A] for fully detailed proofs in similar settings.

Proposition 4.4.4. *Let χ be a unitary character of Γ trivial on Γ'_∞ and $\mu \in \Lambda^*$. Then there exists a constant $\nu(\chi) > 0$ such that*

$$\sum_{\gamma \in T_\Gamma(X)} \chi(\gamma) e\left(\langle \gamma\infty, \mu\rangle\right) = \frac{X^{2s_0(\chi)}}{s_0(\chi)} \left(\mathrm{Res}_{s=s_0(\chi)} L(s,\chi,\mu) + O_{\chi,\mu}(X^{-\nu(\chi)})\right).$$

Proof. Let $\phi_U : \mathbb{R} \to \mathbb{R}$ be a family of smooth non-increasing functions with

$$\phi_U(t) = \begin{cases} 1 & \text{if } t \leq 1 - 1/U, \\ 0 & \text{if } t \geq 1 + 1/U \end{cases} \tag{4.20}$$

and $\phi_U^{(j)}(t) = O(U^j)$ as $U \to \infty$. For $\mathrm{Re}(s) > 0$, we consider the Mellin transform

$$R_U(s) = \int_0^\infty \phi_U(t) t^s \frac{dt}{t} . \tag{4.21}$$

Now we use Mellin inversion and (4.16) to obtain

$$\sum_{\gamma \in T_\Gamma} \chi(\gamma) e\left(\langle \gamma\infty, \mu\rangle\right)\ \phi_U\left(\frac{|c|^2}{X}\right) = \frac{1}{2\pi i} \int_{\mathrm{Re}(s)=n+1} L(s,\chi,\mu) X^s R_U(s) ds.$$

We move the line of integration to $\mathrm{Re}(s) = h(\chi)$ for some $h(\chi) > n/2$ 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 4.4.3 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

$$\sum_{\gamma \in T_\Gamma} \chi(\gamma) e\left(\langle \gamma\infty, \mu\rangle\right) \phi_U\left(\frac{|c|^2}{X}\right) = \frac{X^{s_0(\chi)}}{s_0(\chi)} \left(\mathrm{Res}_{s=s_0(\chi)} L(s,\chi,\mu) + O_{\chi,\mu,U}(X^{-\nu(\chi)})\right) ,$$

for some $\nu(\chi) > 0$. Also, with the appropriate choice of U, one can show that

$$\sum_{\gamma \in T_\Gamma} \chi(\gamma) e\left(\langle \gamma\infty, \mu\rangle\right) \phi_U\left(\frac{|c|^2}{X}\right) = \sum_{\gamma \in T_\Gamma(\sqrt{X})} \chi(\gamma) e\left(\langle \gamma\infty, \mu\rangle\right) + O_{\chi,\mu}(X^{n-a(\chi)}),$$

for some $a(\chi) > 0$. The conclusion follows. □

Remark 4.4.5. As a consequence of Proposition 4.4.4 and Proposition 4.4.2, 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. In particular, we conclude

$$\#T_\Gamma(X) \sim \frac{\mathrm{vol}(\Lambda)^2 \Gamma(n)}{n\pi^{n/2} \mathrm{vol}(\Gamma\backslash\mathbb{H}^{n+1})\Gamma(n/2)} X^{2n}, \tag{4.22}$$

as $X \to \infty$.

4.5 Proof of main results

In this section we will use the analytic properties of twisted Eisenstein series proved in the previous section to prove our main results.

We recall the setup from the introduction. Consider the cohomology group $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ (see Appendix 4.6 for details), which can be identified with the set of unitary characters of Γ trivial on Γ'_∞.

Definition 4.5.1. *We say that $\omega_1, \ldots, \omega_d \in H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z})$ are in* **general position** *if for any $(l_1, \ldots, l_d) \in \mathbb{Z}^d$, we have*

$$n_1\omega_1 + \ldots + n_d\omega_d = 0 \in H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z}) \Leftrightarrow \left(n_i\omega_i = 0 \in H^1_{\Gamma'_\infty}(\Gamma, \mathbb{R}/\mathbb{Z}), \forall i = 1, \ldots, d\right).$$

As an example one can pick $\omega_1, \ldots, \omega_d$ to be a $\mathbb{F}_p$-basis for $H^1_{\Gamma'_\infty}(\Gamma, \mathbb{Z}/p\mathbb{Z})$, where we consider $\mathbb{Z}/p\mathbb{Z} \subset \mathbb{R}/\mathbb{Z}$ via $\mathbb{Z}/p\mathbb{Z} \ni a \mapsto a/p$.

The image of any $\omega \in H^1(\Gamma, \mathbb{R}/\mathbb{Z})$ 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 $J_\omega = \mathbb{R}/\mathbb{Z}$ and in the latter case we put $J_\omega = \mathbb{Z}/m\mathbb{Z}$ where m is the cardinality of the image of ω. That is, J_ω is the closure of the image of ω. We equip $\mathbb{R}/\mathbb{Z}$ and $\mathbb{Z}/m\mathbb{Z}$ with respectively the Lebesgue measure and the uniform probability measure.

Proof of Theorem 4.1.8. 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 $\chi_{\underline{l}} : \Gamma \to \mathbb{C}^\times$ given by

$$\chi_{\underline{l}}(\gamma) := e\left(\omega_{\underline{l}}(\gamma)\right),$$

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* [48, p. 487] in order to conclude equidistribution of the values of

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

inside $\prod_{i=1}^d J_{\omega_i} \times (\mathbb{R}^n/\Lambda)$, we have to show cancelation in the corresponding Weyl sums:

$$\sum_{\gamma \in T_\Gamma(X)} \chi_{\underline{l}}(\gamma) e(\langle \gamma\infty, \mu \rangle),$$

where $\underline{l} \in \mathbb{Z}^d$ and $\mu \in \Lambda^*$. We see that it follows from combining Proposition 4.4.4 and Remark 4.4.1 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 4.1.8 using Weyl's Criterion. □

4.5.1 Distribution of modular symbol mod p and mod 1

Now let us see how Theorem 4.1.1 follows from Theorem 4.1.8.

Proof of Theorem 4.1.1. We restrict to $n = 1$ and $\Gamma = \Gamma_0(N)$. We see that $\mathfrak{m}_f^\pm$ with $f \in \mathcal{S}_2(\Gamma_0(N))_{\text{new}}$ give a basis for $H_P^1(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$. Thus it follows that they are in general position and thus we conclude Theorem 4.1.1 after noting that $T_{\Gamma_0(N)}(Q) = \Omega_{Q,N}$. □

A different application is to consider the distribution of un-normalized modular symbols mod 1. So let $f_1, \ldots, f_d \in \mathcal{S}_2(\Gamma_0(N))$ be a basis of Hecke-normalized new forms and consider the map $\mathbb{Q} \to (\mathbb{R}/\mathbb{Z})^{2d+1}$ given by

$$\mathbb{Q} \ni r \mapsto \mathfrak{m}_{N,\mathbb{R}/\mathbb{Z}}(r) = (\text{Re}\langle r, f_1 \rangle, \text{Im}\langle r, f_1 \rangle, \ldots, \text{Im}\langle r, f_d \rangle, r), \tag{4.23}$$

as a random variable defined on $\Omega_{Q,N}$ defined as in (4.3).

Corollary 4.5.2. *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 $Q \to \infty$. More preciely, for any fixed product of intervals $\prod_{n=1}^{2d+1} I_n \subset (\mathbb{R}/\mathbb{Z})^{2d+1}$, we have*

$$\frac{\#\left\{a/q \in \Omega_{Q,N} \cap I_{2d+1} \mid (\mathrm{Re}\langle a/q, f_1\rangle, \ldots, \mathrm{Im}\langle a/q, f_d\rangle) \in \prod_{n=1}^{2d} I_n\right\}}{\#\Omega_{Q,N}} = \prod_{n=1}^{2d+1} |I_n| + o(1)$$

as $Q \to \infty$.

Proof. From a classical result of Schneider [92] we know that the periods (or elliptic integrals) $\Omega_{f,\pm}$ appearing in (4.2) are transcendental. By the rationality of (4.2), 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 by the Eichler–Shimura isomorphism, we conclude that given a basis $f_1, \ldots, f_d$ of Hecke-normalized newforms, the associated cohomology classes $\mathrm{Re}\, f_i(z)dz$ and $\mathrm{Im}\, f_i(z)dz$ are in general position and the images of the associated characters are dense in $\mathbb{R}/\mathbb{Z}$.

Now Corollary 4.5.2 follows directly from Theorem 4.1.8. □

4.5.2 Proof of Corollary 4.1.6.

Now we see how our results can be applied to the residual distribution of Dedekind sums $s(a,q) = \sum_{k=1}^{q} ((k/q))((ak/q))$ where

$$((x)) = \begin{cases} x - \lfloor x \rfloor - 1/2, & x \notin \mathbb{Z} \\ 0, & x \in \mathbb{Z} \end{cases}$$

is the "sawtooth" function.

Proof of Corollary 4.1.6. The results of [64, Section 5] shows (after some simple manipulations) that for N, p as in Corollary 4.1.6,

$$\Gamma_0(N) \ni \left(\begin{smallmatrix} a & b \\ Nq & d \end{smallmatrix}\right) \mapsto s(a, Nq) - s(a,q) - \frac{(N-1)(a+d)}{12q}$$

defines a non-trivial element $\omega_{N,p} \in H^1_{\Gamma_\infty}(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$ with eigenvalue -1 under the involution given by conjugation by $\left(\begin{smallmatrix} 0 & 1 \\ N & 0 \end{smallmatrix}\right)$. Now let $\omega_\chi \in H^1_{\Gamma_\infty}(\Gamma_0(N), \mathbb{Z}/p\mathbb{Z})$ be the cohomology class associated to a Dirichlet character χ mod N of order p as in the proof of Theorem 4.3.1, which we recall has eigenvalue $+1$ under the conjugation action by $\left(\begin{smallmatrix} 0 & 1 \\ N & 0 \end{smallmatrix}\right)$. We observe that $\omega_\chi(\gamma) = a_0' \in \mathbb{Z}/p\mathbb{Z}$ corresponds exactly to γ having upper left entry in some fixed coset $a_0 H$ of the unique index p subgroup H of $(\mathbb{Z}/N\mathbb{Z})^\times$. Now Corollary 4.1.6 follows directly by applying Theorem 4.1.1 to $\omega_{N,p}$ and ω_χ. □

4.5.3 On the variance of the residual distribution

A natural question to ask next is how well the values equidistribute in Theorem 4.1.8. For simplicity, we will restrict to $\mathbb{H}^2$. So let $\Gamma = \Gamma_0(N)$, $f \in \mathcal{S}_2(\Gamma_0(N))$ be Hecke newform and consider the normalized modular symbols $\mathfrak{m}_f^\pm$ as above. In what follows we will suppress $\mathfrak{m}_f^\pm$ from the notation.

We consider for each $X > 0$ the random variable $Y_{p,X}$ defined on the outcome space $\mathbb{Z}/p\mathbb{Z}$ (with uniform probability measure) by

$$\mathbb{Z}/p\mathbb{Z} \ni a \mapsto \frac{\#\{\gamma \in T_\Gamma(X) \mid \mathfrak{m}_f^\pm(\gamma) \equiv a \bmod p\}}{\# T_\Gamma(X)}.$$

Clearly, we have $\mathbb{E}(Y_{p,X}) = \frac{1}{p}$ and Theorem 4.1.1 says that as $X \to \infty$, the random variable $Y_{p,X}$ converge in distribution to the Dirac measure at $\frac{1}{p}$. We will now calculate the variance, which is a natural measure for the regularity of our distribution problem:

$$\mathrm{Var}(Y_{p,X}) = \mathbb{E}((Y_{p,X} - \mathbb{E}Y_{p,X})^2) = \frac{1}{p} \sum_{a \in \mathbb{Z}/p\mathbb{Z}} \left(Y_{p,X}(a) - \frac{1}{p} \right)^2.$$

First of all we observe that for the modular symbols and primes appearing in Theorem 4.3.1, we have $\mathrm{Var}(Y_{p,X}) = 0$ for all X. On the other hand we can prove using the perturbation theory of the hyperbolic Laplacian, that as p grows, the picture is very different.

Theorem 4.5.3. *For p large enough, we have*

$$\mathrm{Var}(Y_{p,X}) = c_p X^{4s_p - 4} + O_p(X^{4s_p - 4 - \delta_p}), \tag{4.24}$$

for some $s_p, c_p, \delta_p > 0$, as $X \to \infty$. As $p \to \infty$, we have $c_p = 2/p^2 + O(p^{-3})$ and $s_p = 1 - c_f p^{-2} + O(p^{-3})$, where c_f is given by (4.29).

Furthermore, we can calculate the deviation from the mean for each individual residue class. For p large enough and $a \in \mathbb{Z}/p\mathbb{Z}$, we have:

$$\frac{\#\{\gamma \in T_\Gamma(X) \mid \mathfrak{m}_f^{\pm}(\gamma) \equiv a \bmod p\}}{\# T_\Gamma(X)} - \frac{1}{p} \sim d_{a,p} X^{2s_p - 2}, \tag{4.25}$$

as $X \to \infty$, where $d_{a,p} = \frac{2\cos\left(\frac{2\pi a}{p}\right)}{p} + O(p^{-2})$ as $p \to \infty$.

Proof. For $\varepsilon > 0$ we define the character $\chi_\varepsilon : \Gamma_0(N) \to \mathbb{C}$ defined by

$$\gamma \mapsto e^{2\pi i \mathfrak{m}_f^{\pm}(\gamma)\varepsilon}.$$

Let $\lambda_0(\varepsilon) = s_0(\varepsilon)(1 - s_0(\varepsilon))$ with $s_0(\varepsilon) > 1/2$ be the smallest *non-cuspidal* eigenvalue of the hyperbolic Laplacian acting on χ_ε-automorphic functions (i.e. $s_0(\varepsilon)$ is the right-most pole of the twisted Eisenstein series $E(z, s, \chi_\varepsilon)$). Here we put $s_0(\varepsilon) = 1/2$ if there are no residual eigenvalues. From this we define

$$s_p := \max_{a \in (\mathbb{Z}/p\mathbb{Z})^\times} s_0(a/p),$$

which will turn out to control the variance. Note that $s_p < 1$ for all p by Lemma 4.4.1.

By simple Fourier analysis on $\mathbb{Z}/p\mathbb{Z}$ we have

$$\frac{\#\{\gamma \in T_\Gamma(X) \mid \mathfrak{m}_f^{\pm}(\gamma) \equiv a \bmod p\}}{\# T_\Gamma(X)} = \frac{1}{p} \sum_{b \in \mathbb{Z}/p\mathbb{Z}} \frac{1}{\# T_\Gamma(X)} \sum_{\gamma \in T_\Gamma(X)} \chi_{b/p}(\gamma) e^{-2\pi i ab/p}. \tag{4.26}$$

By Parseval this implies

$$\sum_{a \in \mathbb{Z}/p\mathbb{Z}} \left| \frac{\#\{\gamma \in T_\Gamma(X) \mid \mathfrak{m}_f^{\pm}(\gamma) \equiv a \bmod p\}}{\# T_\Gamma(X)} \right|^2 = \frac{1}{p} \sum_{a \in \mathbb{Z}/p\mathbb{Z}} \left| \frac{\sum_{\gamma \in T_\Gamma(X)} \chi_{a/p}(\gamma)}{\# T_\Gamma(X)} \right|^2.$$

Hence we have

$$\begin{aligned} \mathrm{Var}(Y_{p,X}) &= \frac{1}{p} \sum_{a \in \mathbb{Z}/p\mathbb{Z}} \left| \frac{\#\{\gamma \in T_\Gamma(X) \mid \mathfrak{m}_f^{\pm}(\gamma) \equiv a \bmod p\}}{\# T_\Gamma(X)} \right|^2 - \frac{1}{p^2} \\ &= \frac{1}{p^2} \sum_{a \in (\mathbb{Z}/p\mathbb{Z})^\times} \left| \frac{1}{\# T_\Gamma(X)} \sum_{\gamma \in T_\Gamma(X)} \chi_{a/p}(\gamma) \right|^2, \end{aligned} \tag{4.27}$$

since the contribution in the last sum for $a \equiv 0 \bmod p$ is exactly $1/p^2$. Now by a contour integration argument as in Proposition 4.4.4, we conclude that if $s_0(a/p) = 1/2$ (i.e. there are no non-cuspidal eigevalues in $[0, 1/4)$ for the Laplacian acting on $\chi_{a/p}$-automorphic functions) then

$$\sum_{\gamma \in T_\Gamma(X)} \chi_{a/p}(\gamma) = O_\varepsilon(X^{1+\varepsilon}).$$

On the other hand if $s_0(a/p) > 1/2$, then we conclude that

$$\sum_{\gamma \in T_\Gamma(X)} \chi_{a/p}(\gamma) = c_{a,p} X^{2s_0(a/p)}(1 + O(X^{-\delta_{a,p}})),$$

for some $\delta_{a,p} > 0$ depending on the spectral gap between $\lambda_0(a/p)$ and $\lambda_1(a/p)$ and some $c_{a,p} \neq 0$ depending on the constant term of the non-cuspidal eigenfunction corresponding to $\lambda_0(a/p)$. Combining this with (4.27), we deduce the formula (4.24).

We now want to understand the large p behavior. For this we employ perturbation theory of the twisted Laplacian, as developed in [83, Section 4] and [33]. We have that the smallest eigenvalue $\lambda_0(\epsilon) = s_0(\epsilon)(1 - s_0(\epsilon))$ of the twisted Laplacian by the character χ_ϵ is real analytic in ϵ, for ϵ small enough. Moreover, we know that

$$s_0(\varepsilon) = 1 - c_f \varepsilon^2 + O(\varepsilon^3), \tag{4.28}$$

as $\varepsilon \to 0$, where

$$c_f = \frac{8\pi^2 \|f\|^2}{\mathrm{vol}(\Gamma)\Omega_{f,\pm}^2}, \tag{4.29}$$

see [18, Section 4] or [79] for more details.

Now fix $\epsilon > 0$ small enough such that (4.28) holds. We want to show that if $\theta \in [\epsilon, 1 - \epsilon]$, then $\lambda_0(\theta)$ is bounded away from 0 (and hence $s_0(\theta)$ is bounded away from 1). This follows almost directly from [33, Proposition 2.1]. Suppose the contradiction, i.e. there exists a sequence $\{\theta_j\} \subset [\epsilon, 1 - \epsilon]$ such that $\lambda_0(\theta_j) \to 0$. By a compactness argument, by passing to a subsequence, we can assume that there exists $\theta^* \in [\epsilon, 1 - \epsilon]$ such that $\theta_j \to \theta^*$. Denote by $f_j \in L^2(\Gamma\backslash\mathbb{H}, \chi_{\theta_j})$ the corresponding eigenfunctions with eigenvalues $\lambda_0(\theta_j)$. By the continuity statement in [33, Proposition 2.1], we conclude that there exists $f^* \in L^2(\Gamma\backslash\mathbb{H}, \chi_{\theta^*})$ such that a subsequence of (f_j) is L^2-convergent to f^* and $\Delta f^* = 0$. But this means that f^* is constant, and hence θ^*=0, which is a contradiction.

By conjugation, we have $s_0(\varepsilon) = s_0(-\varepsilon)$. Using the above and (4.28), we conclude that for p large enough, we have that $s_p = s_0(1/p) = s_0(-1/p) = s_0((p-1)/p)$, which combined with (4.28) gives the wanted.

Now, from (4.27), we note that the main term in the variance is given by the contributions of $a = 1$ and $a = p-1$ in the sum. By (4.22) we have $\#T_\Gamma(X) = (\pi \mathrm{vol}(\Gamma))^{-1} X^2 (1 + O(X^{-\nu}))$, for some $\nu > 0$. Furthermore, we know that the eigenfunction (and in particular its constant Fourier coefficient) corresponding to $s_0(\varepsilon)$ varies analytically with ε (for ε small enough) and we can deduce that

$$\mathrm{Res}_{s=s_0(\epsilon)} L(s, \chi_\epsilon) = \frac{1}{\pi \mathrm{vol}(\Gamma)} + O(\epsilon^2),$$

see [18] for more details. Hence, from (4.27) and Proposition 4.4.4, we deduce that

$$c_p = \frac{2}{p^2} + O(p^{-3}).$$

Finally for p large enough, we see that the main term in (4.26) comes from $b = 0$, and the second main term is given by

$$\frac{1}{p}(c_{1,p} e^{-2\pi i/p} + c_{p-1,p} e^{2\pi i/p}) X^{2 s_0(1/p) - 2},$$

which by the above gives (4.25). □

We note that the inequality (4.4) does indeed follow from (4.25).

Remark 4.5.4. We note that it should be straightforward to generalise Theorem 4.5.3 to $\mathbb{H}^n$, as the perturbation theory of the first eigenvalue of the Laplacian has been developed by Epstein [33] for $\mathbb{H}^n$.

4.5.4 Cohomology classes ordered by lengths of geodesics

We now give a proof of Theorem 4.1.11 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 [78], where Petridis–Risager show that for co-compact 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 [82].

We firstly consider the case $n = 1$. If $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ is hyperbolic, then γ is conjugate in $\mathrm{SL}_2(\mathbb{R})$ to a unique element $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix}$ 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 $\mathrm{Li}(x) = \int_2^x (\log t)^{-1} dt$ is the logarithmic integral (see [41, 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 4.4.1, 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's criterion.

We now discuss the general case n. As mentioned earlier, the first proof of the Prime Geodesic Theorem in the general case was given by Gangolli and Warner [35]. 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 [14, Ch. 7]. As a consequence, they obtain the following stronger version of Prime Geodesic Theorem for $\mathbb{H}^{n+1}$ [14, Thm. 7.37]:

$$\pi_\Gamma(X) = \sum_{n/2 < s_j \leq n} \mathrm{Li}(e^{s_j X}) + O\left(e^{(n - \frac{n}{n+2})X}\right)$$

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 $L^2(\Gamma \backslash \mathbb{H})$. 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 will leave out the details. The analysis should be similar to the case $n = 1$ and is furthermore implied to hold by Sarnak in [89, p. 6]. Similarly, Phillips and Sarnak [82] 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 [82].

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 4.1.11 follows by Weyl's Criterion as above.

4.6 On the size of certain cohomology groups

In Chapter 4 of this thesis 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 (or cohomology classes) our results actually apply to. This amounts to finding the dimensions of the relevant spaces of unitary characters or equivalently of certain cohomology groups. This last perspective is most useful when comparing it to the existing literature. We will mostly restrict to arithmetic subgroups, which we will define shortly. Then we will define the cohomology groups that are relevant and finally survey what is known about their size.

4.6.1 Congruence subgroups

We will now define what we mean by a *congruence subgroup*, which most of the results mentioned below applies to. In this case one can obtain quite explicit descriptions of the double coset $\Gamma'_\infty\backslash\Gamma/\Gamma'_\infty$ occuring in Theorem 4.1.8.

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 *principal congruence subgroup*

$$\mathrm{SV}_n(J;N) := \left\{ \left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{SV}_n(J) \mid a-1, b, c, d-1 \in NJ \right\}. \tag{4.30}$$

A subgroup $\Gamma < \mathrm{SV}_n(J)$ is called a *congruence subgroup* if $\mathrm{SV}_n(J;N) < \Gamma$, for some $N \in \mathbb{N}$. We quote [31, 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

$$D(c) := \left\{ \quad (a,d) \quad \left| \begin{array}{c} a \in J/(N \cdot V(J) \cdot c),\ d \in J/(N \cdot c \cdot V(J)), \\ a-1, d-1 \in N \cdot J,\ a\bar{c}, \bar{c}d \in N \cdot V(J) \end{array} \right. \right\}.$$

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 4.6.1. 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 of* Γ is any subgroup containing $\{\gamma \in \Gamma \mid \gamma \equiv I_{n+2} \mod N\}$ for some positive integer N, see [89, 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}. 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 [31, Section 3] for more details.

4.6.2 The first cohomology group

We refer to [98, 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 $H^1_P(\Gamma, \mathbb{R}/\mathbb{Z})$.

4.6.3 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 [94, p. 484] that Γ is finitely presented 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;

$$H^1(\Gamma,\mathbb{R}/\mathbb{Z}) \cong H^1_{\mathrm{free}}(\Gamma,\mathbb{R}/\mathbb{Z}) \oplus H^1_{\mathrm{tor}}(\Gamma,\mathbb{R}/\mathbb{Z}),$$

where the $\mathbb{R}/\mathbb{Z}$ rank of $H^1_{\mathrm{free}}(\Gamma,\mathbb{R}/\mathbb{Z})$ is the same as the dimension of $H^1(\Gamma,\mathbb{R})$ and the size of $H^1_{\mathrm{tor}}(\Gamma,\mathbb{R}/\mathbb{Z})$ 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 [40];

$$H^1(\Gamma,\mathbb{R}) \cong H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R}) \oplus H^1_{\mathrm{Eis}}(\Gamma,\mathbb{R}), \tag{4.31}$$

where $H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$ is the cuspidal part corresponding to certain automorphic forms for Γ (as we will see shortly) and $H^1_{\mathrm{Eis}}(\Gamma,\mathbb{R})$ is the (remaining) Eisenstein part, which can be canonically defined. The cuspidal part $H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$ 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 $H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$, $H^1_{\mathrm{Eis}}(\Gamma,\mathbb{R})$ and $H^1_{\mathrm{tor}}(\Gamma,\mathbb{R}/\mathbb{Z})$, 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 $H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$ is non-trivial.

4.6.4 The dimension of cohomology groups

It is a result of Kazhdan [54] 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 [70] and Lubotzky [60], 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.

4.6.4.1 Cohomology classes associated to automorphic forms

Recall the splitting (4.31) due to Harder of the cohomology into a cuspidal and an Eisenstein part. We give a brief overview of the description of $H^1_{\mathrm{cusp}}(\Gamma,\mathbb{R})$ in terms of automorphic forms, as in [89]. We recall the canonical isomorphism between $H^1(\Gamma,\mathbb{R})$ and the de Rham cohomology group $H^1_{\mathrm{dR}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R})$ consisting of 1-forms. Inside $H^1_{\mathrm{dR}}(\Gamma\backslash\mathbb{H}^{n+1},\mathbb{R})$ we define the subset of cuspidal harmonic 1-forms.

Definition 4.6.2. *A harmonic 1-form $\alpha = f_0dx_0 + f_1dx_1 + \cdots + f_ndx_n$ 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 [89, (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 [99] of GL_2 holomorphic forms of weight $(n+1)/2+1$ (for details see [89, page 21]). This gives us non-trivial examples for which Theorem 4.1.8 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 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 (4.1):

$$\langle \gamma, f(z)dz \rangle = 2\pi i \int_\infty^{a_\gamma/c_\gamma} f(z)dz = -\langle a_\gamma/c_\gamma, f \rangle .$$

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 $H^1_{\mathrm{free}}(\Gamma, \mathbb{R}/\mathbb{Z}) \cong \mathrm{Har}^1_{\mathrm{cusp}}(\Gamma\backslash\mathbb{H}^{n+1}, \mathbb{R})/L$.

4.6.4.2 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 [111, 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.

4.6.4.3 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 [96] 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 [32, Proposition 7.5.6]. The same conclusion holds for co-finite subgroups $\Gamma \leq \mathrm{SL}_2(\mathcal{O}_D)$, 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 $\Gamma \leq \mathrm{SL}_2(\mathcal{O}_{\mathcal{D}})$ with $D = -4, -3$ the picture is much more mysterious, but a lot of numerics are available in [95] and [32, Ch. 7.5].

For the cuspidal part there are some useful results giving lower bounds on the rank. First of all Rohlfs [85] showed that

$$\dim H^1_{\mathrm{cusp}}(\mathrm{SL}_2(\mathcal{O}_D), \mathbb{R}) \geq \frac{\varphi(D)}{6} - \frac{1}{2} - h(D),$$

where $h(D)$ denotes the class number of $\mathbb{Q}(\sqrt{D})$. Furthermore Şengün and Turkelli [97] 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

$$\dim H^1_{\mathrm{cusp}}(\Gamma_0(p^n), \mathbb{R}) \geq p^{6n},$$

as $n \to \infty$ (an upper bound of p^{10n} has been proved by Calegari and Emerton [12]). In the case of cocompact groups stronger results were obtained by Kionke and Schwermer [56].

4.6.5 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 [95] 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 [4]. 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 [4, 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 [93]. There seems however to be no experimental or theoretical work available in this case.

This concludes our discussion on the sizes of cohomology groups to which our results apply.

Chapter 5

Dissipation of correlations of holomorphic cusp forms

This chapter is mainly based on [17]. We would like to thank Peter Humphries for suggesting the problem to us and for his help with triple product integrals, in particular equations (5.26) and (5.27).

5.1 Introduction

Fix $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ and $X = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. Let k be an integer. We denote by $\mathcal{A}_k(\Gamma)$ the space of automorphic functions of weight k, that is functions $f : \mathbb{H} \to \mathbb{C}$ which transform as

$$f(\gamma z) = j_\gamma(z)^k f(z), \quad \text{for all } \gamma \in \Gamma, \tag{5.1}$$

where $j_\gamma(z) = \dfrac{cz+d}{|cz+d|}$ with $\gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix}$. We denote by $\mathcal{L}_k(X)$ the space of automorphic functions of weight k which are square-integrable. We see that if $f \in S_k(\Gamma)$, then $y^{k/2} f(z) \in \mathcal{L}_k(X)$.

We have the Maaß raising and lowering operators

$$K_k : \mathcal{L}_k(X) \to \mathcal{L}_{k+2}(X) \quad \text{and} \quad \Lambda_k : \mathcal{L}_k(X) \to \mathcal{L}_{k-2}(X),$$

which allow us to move between automorphic functions of different weights, see 5.5 for definitions. Hence, for even integers $k_1 \le k_2$ we define the operators

$$R_{k_1}^{k_2} : \mathcal{L}_{k_1}(X) \to \mathcal{L}_{k_2}(X), \qquad \phi \mapsto \frac{K_{k_2-2} \dots K_{k_1+2} K_{k_1} \phi}{\| K_{k_2-2} \dots K_{k_1+2} K_{k_1} \phi \|},$$

where $\|R_{k_1}^{k_2}\phi\| = 1$. We prove the following theorem.

Theorem 5.1.1. *Fix any $\phi \in C_b(\Gamma\backslash\mathbb{H})$ (a bounded function on on $\Gamma\backslash\mathbb{H}$). Let f and g be L^2-normalised holomorphic Hecke cusp forms of weights k_1 and k_2 respectively with $k_1 \leq k_2$. Let*

$$\delta_{f=g} = \begin{cases} 1, & \text{if } f = g; \\ 0, & \text{otherwise.} \end{cases}$$

Along any sequences of such f and g, we have

$$\int_{\Gamma\backslash\mathbb{H}} \phi(z) R_{k_1}^{k_2} \left(y^{k_1/2} f(z) \right) y^{k_2/2} \overline{g(z)} d\mu(z) \to \delta_{f=g} \frac{3}{\pi} \int_{\Gamma\backslash\mathbb{H}} \phi(z) d\mu(z) \quad \text{as} \quad k_2 \to \infty.$$

In other words, if $F_{k_1}(z) = y^{k_1/2} f(z)$ and $G_{k_2}(z) = y^{k_2/2} g(z)$, then

$$\left\langle \phi \left(R_{k_1}^{k_2} F_{k_1} \right), G_{k_2} \right\rangle \to \delta_{f=g} \frac{1}{\text{vol}(X)} \langle \phi, 1 \rangle \quad \text{as} \quad k_2 \to \infty.$$

Remark 5.1.1. This corresponds to a generalisation of Quantum Unique Ergodicity by classifying the possible quantum limits of Hecke cusp forms when we project back to the modular surface. That is, along any sequence of holomorphic Hecke eigenforms of increasing weight, we show there are two possible limit points.

We also consider the case where we do not raise F_{k_1} to weight k_2, but rather project into $\mathcal{L}_{k_2-k_1}(X)$. These statements are not the same, since there are extra normalising factors that play an important role.

Theorem 5.1.2. *Fix $\phi \in C_b(X)$. Let l be a nonnegative even integer. Let f and g vary along a sequence of Hecke cusp forms of weights k and $k + l$ respectively. Then*

$$\int_{\Gamma\backslash\mathbb{H}} \left(R_0^l \phi(z) \right) y^{k+l/2} f(z) \overline{g(z)} d\mu(z) \to \delta_{f=g} \frac{3}{\pi} \int_{\Gamma\backslash\mathbb{H}} \phi(z) d\mu(z) \quad \text{as} \quad k \to \infty.$$

In other words, we have

$$\left\langle \left(R_0^l \phi \right) F_k, G_{k+l} \right\rangle \to \delta_{f=g} \frac{1}{\text{vol}(X)} \langle \phi, 1 \rangle \quad \text{as} \quad k \to \infty.$$

Remark 5.1.2. In Theorem 5.1.2, we can also allow l to grow with k. Our method works if $l \leq c \log\log k$, where $c < \frac{1}{12\log 2}$.

Remark 5.1.3. It is crucial for us in Theorem 5.1.2 that ϕ is obtained from repeated iterations of raising operators. We expect the statement to hold for all $\phi \in \mathcal{L}_l(X)$. However, to achieve this we would also need to compute inner products of the type $\left\langle (R_m^l F_m) G_k, H_{k+l} \right\rangle$, which in representation theory corresponds to a triple integral of three discrete series representations. The local factors of such integrals are difficult to estimate. The local factors of triple product integrals where at least one factor comes from principal series representation (Maaß forms) were computed by Cheng [13].

We use the spectral theory of weight k automorphic functions, which we summarise thoroughly in Section 5.2. We can write a decomposition of $\mathcal{L}_k(X)$ in terms of eigenfunctions of the weight k Laplacian Δ_k. The spectral expansion will involve:

- Hecke Maaß cusp forms $R_0^k u_j$ raised to weight k;
- raised holomorphic Hecke cusp forms $R_l^k(F_l)$, for $0 < l \leq k$;
- weight k Eisenstein series $E_k\left(z, \frac{1}{2} + it\right)$.

Therefore, it is enough to compute inner products of type $\langle \phi F_{k_1}, G_{k_2} \rangle$ or $\left\langle \phi R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle$, where ϕ appears in the spectral decomposition. We proceed similarly as in the work of Holowinsky [43] and Soundararajan [102]. Our new ingredient is to incorporate the spectral theory of weight k automorphic functions to their method, which we review in Section 5.2. We have two approaches, depending on the size of

$$S(f, g) := L(1, \mathrm{sym}^2 f) L(1, \mathrm{sym}^2 g). \tag{5.2}$$

Firstly, we can compute directly the inner products, using Rankin–Selberg unfolding for the Eisenstein series and Ichino's formula for the Maaß cusp form case, see Section 5.3. The formulas will involve central values of L-functions, to which we apply the weak subconvexity results of Soundararajan. This will win if $S(f, g)$ is large.

Alternatively, we can expand the inner products in terms of the Fourier expansions. We need bounds for the Fourier coefficients of weight k automorphic forms, which we compute in Section 5.4. This approach boils down to bounding shifted convolution sums, where we apply the results of Holowinsky, see Section 5.5. This will win if $S(f, g)$ is sufficiently small. We put everything together and complete the proofs of Theorems 5.1.1 and 5.1.2 in Section 5.6.

5.2 Spectral theory of weight k automorphic forms

We quote [28, Chapter 4], [11, Chapter 2] for detailed expositions on the analytical theory of weight k automorphic forms. Let k be an integer. We denote by $\mathcal{A}_k(\Gamma)$ the space of automorphic functions of weight k, that is functions $f : \mathbb{H} \to \mathbb{C}$ that transform by

$$f(\gamma z) = j_\gamma(z)^k f(z), \quad \text{for all } \gamma \in \Gamma, \tag{5.3}$$

where $j_\gamma(z) = \dfrac{cz+d}{|cz+d|}$ with $\gamma = \begin{pmatrix} * & * \\ c & d \end{pmatrix}$. Note that we have the cocycle relation

$$j_{\gamma_1\gamma_2}(z) = j_{\gamma_1}(\gamma_2 z) j_{\gamma_2}(z), \quad \text{for all } \gamma_1, \gamma_2 \in \Gamma.$$

Let $\mathcal{L}_k(\Gamma)$ the automorphic functions of weight k that are square-integrable. On $\mathcal{L}_k(\Gamma)$ we define the inner product

$$\langle f, g \rangle = \int_{\Gamma\backslash\mathbb{H}} f(z)\overline{g(z)} d\mu. \tag{5.4}$$

We consider the Maaß raising and lowering operators acting on $C^\infty(\mathbb{H})$ (smooth functions on $\mathbb{H}$)

$$\begin{aligned} K_k &= \frac{k}{2} + y\left(i\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) = \frac{k}{2} + (z - \overline{z})\frac{\partial}{\partial z}, \\ \Lambda_k &= \frac{k}{2} + y\left(i\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) = \frac{k}{2} + (z - \overline{z})\frac{\partial}{\partial \overline{z}}. \end{aligned} \tag{5.5}$$

These operators are used to map between spaces of different weights:

$$K_k : C^\infty(\Gamma) \cap \mathcal{L}_k(\Gamma) \to C^\infty(\Gamma) \cap \mathcal{L}_{k+2}(\Gamma),$$
$$\Lambda_k : C^\infty(\Gamma) \cap \mathcal{L}_k(\Gamma) \to C^\infty(\Gamma) \cap \mathcal{L}_{k-2}(\Gamma),$$

and satisfy the following property:

$$\langle K_k f, g \rangle = - \langle f, \Lambda_{k+2} g \rangle, \tag{5.6}$$

for $f \in C^\infty(\Gamma) \cap \mathcal{L}_k(\Gamma)$ and $g \in C^\infty(\Gamma) \cap \mathcal{L}_{k+2}(\Gamma)$. Moreover, the following product rule holds:

$$\begin{aligned} K_{k+l}(g_k g_l) &= (K_k g_k) g_l + g_k (K_l g_l), \\ \Lambda_{k+l}(g_k g_l) &= (\Lambda_k g_k) g_l + g_k (\Lambda_l g_l), \end{aligned} \tag{5.7}$$

where g_k and g_l are smooth automorphic functions of weights k and l respectively.

The Laplace operator of weight k is defined by

$$\Delta_k = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - ik \frac{\partial}{\partial x}.$$

This can be written in terms of the raising and lowering operators as

$$\Delta_k = -K_{k-2}\Lambda_k - \lambda(k/2) = -\Lambda_{k+2}K_k - \lambda(-k/2), \tag{5.8}$$

where

$$\lambda(s) := s(1-s). \tag{5.9}$$

The operator Δ_k acts on $\mathcal{A}_k(\Gamma) \cap C^\infty(\Gamma)$. We define a *Maaß form* to be a smooth automorphic function of weight k which is an eigenfunction of Δ_k. Let $\mathcal{A}_k(\Gamma, s)$ denote the space of Maaß forms with eigenvalue $\lambda(s)$. We also note that, if $f(z) \in \mathcal{A}_k(\Gamma, s)$ has at most polynomial growth in cusp, it has a Fourier expansion of the form

$$f(z) = a_0(y) + \sum_{n \neq 0} a_f(n) W_{\frac{kn}{2|n|}, s-\frac{1}{2}}(4\pi|n|y) e(nx),$$

where $W_{\alpha,\beta}(z)$ is the Whittaker function, see [28] for more details.

We denote by $\mathcal{B}_k(\Gamma)$ the space of smooth automorphic functions of weight k such that $f, \Delta_k f \in \mathcal{L}_k(\Gamma)$. Then $-\Delta_k$ defines a symmetric, non-negative operator on $\mathcal{B}_k(\Gamma)$. The space $\mathcal{B}_k(\Gamma)$ is dense in $\mathcal{L}_k(\Gamma)$ and the operator $-\Delta_k$ admits a self-adjoint extension to $\mathcal{L}_k(\Gamma)$ and we can study the spectral decomposition of this space.

5.2.1 Eisenstein Series

The Eisenstein series of weight k is defined by

$$E_k(z, s) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\operatorname{Im} \gamma z)^s j_\gamma(z)^{-k}, \tag{5.10}$$

The series (5.10) converges absolutely for $\mathrm{Re}(s) > 1$ and has analytic continuation to the whole complex plane. Unless $k = 0$, $E_k(z, s)$ has no poles for $\mathrm{Re}(s) \geq 1/2$. If $k = 0$, then $E(z, s)$ has a pole at $s = 1$ with residue

$$\mathrm{Res}_{s=1} E(z, s) = \frac{3}{\pi}. \tag{5.11}$$

If s is not a pole of $E_k(z, s)$, then $E_k(z, s)$ is a weight k Maaß form with eigenvalue $\lambda(s)$, but it is not in $\mathcal{L}_k(\Gamma)$. We note that

$$K_k E_k(z, s) = \left(\frac{k}{2} + s\right) E_{k+2}(z, s), \quad \Lambda_k E_k(z, s) = \left(\frac{k}{2} - s\right) E_{k-2}(z, s).$$

Hence, if k is an even positive integer,

$$K_{k-2} \dots K_2 K_0 E(z, s) = s(s+1)\dots(s + k/2 - 1) E_k(z, s) = \frac{\Gamma(s + k/2)}{\Gamma(s)} E_k(z, s). \tag{5.12}$$

As in [49], [28], [27] or [76], the Fourier expansion of $E_k(z, s)$ is given by

$$\begin{aligned} E_k(z, s) = & y^s + \frac{(-1)^{k/2}\Gamma(s)^2}{\Gamma\left(s - \frac{k}{2}\right)\Gamma\left(s + \frac{k}{2}\right)} \phi(s) y^{1-s} \\ & + \frac{(-1)^{k/2}\Gamma(s)}{2\Gamma\left(s + \frac{|k|}{2}\right)\xi(2s)} \sum_{n>0} |n|^{s-1} \sigma_{1-2s}(|n|) W_{|k|/2, s-1/2}(4\pi|n|y) e(nx) \\ & + \frac{(-1)^{k/2}\Gamma(s)}{2\Gamma\left(s - \frac{|k|}{2}\right)\xi(2s)} \sum_{n<0} |n|^{s-1} \sigma_{1-2s}(|n|) W_{-|k|/2, s-1/2}(4\pi|n|y) e(nx), \end{aligned} \tag{5.13}$$

where

$$\begin{aligned} \phi(s) &= \frac{\xi(2s-1)}{\xi(2s)}, \\ \xi(s) &= \pi^{-s/2}\Gamma(s/2)\zeta(s) = \Gamma_{\mathbb{R}}(s)\zeta(s), \\ \sigma_\nu(n) &= \sum_{d|n} d^\nu. \end{aligned}$$

Let $\psi(y)$ be a smooth compactly supported function on $\mathbb{R}^+$. Then we define the incomplete Eisenstein series

$$E_k(z|\psi) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \psi(\mathrm{Im}\, \gamma z) j_\gamma(z)^{-k},$$

that is in $\mathcal{L}_k(\Gamma)$, but it is not a Maaß form. We denote by $\mathcal{E}_k(\Gamma)$ the space of all incomplete Eisenstein series. Then Δ_k acts on $\mathcal{E}_k(\Gamma)$ with purely continuous spectrum which covers the interval $[1/4, \infty)$ with multiplicity one. Moreover, for any $f \in \mathcal{E}_k(\Gamma)$, we have the expansion

$$f(z) = \frac{1}{4\pi}\int_{-\infty}^{\infty}\left\langle f, E_k\left(\cdot, \frac{1}{2}+it\right)\right\rangle E_k\left(\cdot, \frac{1}{2}+it\right) dt.$$

We let

$$\Psi(s) := \int_0^{\infty} \psi(y) y^s \frac{dy}{y}$$

be the Mellin transform of ψ. Hence, $\Psi(s)$ is entire and satisfies

$$\Psi(s) \ll (1+|s|)^{-A} \tag{5.14}$$

for any $A > 0$, uniformly in vertical strips. By the Mellin inversion theorem, we have

$$\psi(y) = \frac{1}{2\pi i}\int_{(\sigma)} y^{-s}\Psi(s) ds$$

for $\sigma > 1$. Using this, we observe that

$$E_k(z|\psi) = \frac{1}{2\pi i}\int_{(2)} \Psi(-s) E_k(z, s) ds. \tag{5.15}$$

5.2.2 Cusp forms

The orthogonal complement of $\mathcal{E}_k(\Gamma)$ in $\mathcal{L}_k(\Gamma)$ consists of functions whose zero Fourier coefficient vanishes, which we denote by $\mathcal{C}_k(\Gamma)$. Then Δ_k acts on $\mathcal{C}_k(\Gamma)$ with purely discrete spectrum. We now provide a description of this space.

Let $\mathcal{C}_k(\Gamma, s)$ be the space of Maaß cusp forms of weight k and eigenvalue $\lambda(s)$. Then $K_k : \mathcal{C}_k(\Gamma, s) \to \mathcal{C}_{k+2}(\Gamma, s)$ and $\Lambda_k : \mathcal{C}_k(\Gamma, s) \to \mathcal{C}_{k-2}(\Gamma, s)$. Also,

$$K_k F = 0 \iff \lambda(s) = \lambda(-k/2) \iff y^{k/2}\overline{f(z)} \text{ is holomorphic in } z,$$
$$\Lambda_k F = 0 \iff \lambda(s) = \lambda(k/2) \iff y^{-k/2} f(z) \text{ is holomorphic in } z.$$

If $\lambda(s) \neq \lambda(-k/2)$, then the map

$$\left(\lambda(s) - \lambda\left(-\frac{k}{2}\right)\right)^{-1/2} K_k : \mathcal{C}_k(\Gamma, s) \to \mathcal{C}_{k+2}(\Gamma, s)$$

is a bijective isometry. A similar statement holds for Λ_k. Now for even integers $k_1 < k_2$ and $\lambda(s) \notin \{\lambda(-k_1/2), \dots \lambda(-k_2/2+1)\}$, we define the bijective isometry $R_{k_1}^{k_2} : \mathcal{C}_{k_1}(\Gamma, s) \to \mathcal{C}_{k_2}(\Gamma, s)$ given by

$$R_{k_1}^{k_2}(s) := \prod_{\substack{k_1 \leq l < k_2 \\ l \equiv 2 \bmod 2}} \left(\lambda(s) - \lambda\left(-\frac{k}{2}\right)\right)^{-1/2} K_{k_2-2} \dots K_{k_1+2} K_{k_1}. \tag{5.16}$$

When $k \geq 0$, the eigenspace of Δ_k with eigenvalue $\lambda(k/2)$ is given by

$$\mathcal{C}_k\left(\Gamma, \frac{k}{2}\right) = \left\{ y^{k/2} f(z) \mid f \in S_k(\Gamma) \right\} \tag{5.17}$$

and

$$\mathcal{C}_{-k}\left(\Gamma, \frac{k}{2}\right) = \left\{ y^{k/2} \overline{f(z)} \mid f \in S_k(\Gamma) \right\}. \tag{5.18}$$

The eigenspaces of Δ_k in $\mathcal{C}_k(\Gamma, m/2)$ for even m in the range $0 < m \leq k$ are determined by classical cusp forms in $S_m(\Gamma)$ with repeated applications of the Maaß raising operators.

Putting everything together, we have the following theorem, see [28, Corollary 4.4].

Theorem 5.2.1. *Let k be an even positive integer. Let $\{u_j(z)\}$ be an orthonormal basis of Maaß cusp forms of $\mathcal{C}_0(\Gamma)$ with corresponding eigenvalues $\lambda(s_j)$. Also, choose $\{f_{j,m}\}$ an orthonormal basis for $S_m(\Gamma)$. Then an orthonormal basis of $\mathcal{C}_k(\Gamma)$ is given by*

$$u_{j,k}(z) := \prod_{0 \leq l < k/2} (\lambda(s_j) - \lambda(-l))^{-1/2} K_{2l}(u_j(z)),$$
$$u_{j,m,k}(z) := \prod_{m \leq l < k/2} (\lambda(m) - \lambda(-l))^{-1/2} K_{2l}\left(y^m f_{j,2m}(z)\right).$$

Remark 5.2.1. Since Selberg's eigenvalue conjecture holds for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, all the points s_j are on the line $\mathrm{Re}(s_j) = 1/2$.

Remark 5.2.2. We choose an orthonormal basis of $\mathcal{C}_k(\Gamma)$ consisting of Hecke–Maaß cusp forms, i.e. common eigenfunctions of the Laplacian Δ_k and all Hecke operators T_n. This is possible because the operators T_n commute with Δ_k, K_k and Λ_k. We denote such a basis by

$$B_k := \{u_{j,k}\} \cup \left(\bigcup_{0 < m \leq k/2} \{u_{j,m,k}\} \right). \tag{5.19}$$

We can compute the normalisation factors, as in [28, p. 508]. They are given by

$$\alpha^2(s,k) := \prod_{0\le l<k/2} (\lambda(s_j)-\lambda(-l))^{-1} = (-1)^{k/2}\frac{\Gamma(s-k/2)}{\Gamma(s+k/2)}, \tag{5.20}$$

$$\beta^2(m,k) := \prod_{m/2\le l<k/2} (\lambda(m)-\lambda(-l))^{-1} = \frac{\Gamma(m)}{\Gamma\left(\frac{k+m}{2}\right)\Gamma\left(\frac{k-m}{2}+1\right)}. \tag{5.21}$$

If $f \in S_{k_1}(\Gamma)$ and $F_{k_1} = y^{k_1/2}f(z) \in \mathcal{C}_{k_1}(\Gamma, k_1/2)$, we just denote the isometry $R_{k_1}^{k_2}(k_1/2)$ from (5.16) by $R_{k_1}^{k_2} : \mathcal{C}_{k_1}(\Gamma,k_1/2) \to \mathcal{C}_{k_2}(\Gamma,k_1/2)$ given by

$$R_{k_1}^{k_2}F_{k_1} = \beta(k_1,k_2)K_{k_2-2}.\dots K_{k_1}F_{k_1}. \tag{5.22}$$

If u_j is a cuspidal Maaß form with eigenvalue $\lambda(1/2+it_j)$, then its Fourier expansion is given by

$$u_j(z) = \sum_{n\neq 0}\frac{c_j(|n|)}{\sqrt{|n|}}W_{0,it_j}(4\pi|n|y)e(nx).$$

If u_j is a Hecke eigenform, then the Hecke eigenvalues are given by $c_j(n)/c_j(1)$, for positive n. We can relate it to the Fourier expansion of $u_{j,k}$, as in [49]:

$$\begin{aligned} u_{j,k}(z) =& \frac{(-1)^{k/2}\Gamma(1/2+it_j)}{\Gamma\left(\frac{1}{2}+\frac{k}{2}+it_j\right)}\sum_{n>0}\frac{c_j(|n|)}{\sqrt{|n|}}W_{k/2,it_j}(4\pi|n|y)e(nx) \\ &+ \frac{(-1)^{k/2}\Gamma(1/2+it_j)}{\Gamma\left(\frac{1}{2}-\frac{k}{2}+it_j\right)}\sum_{n<0}\frac{c_j(|n|)}{\sqrt{|n|}}W_{-k/2,it_j}(4\pi|n|y)e(nx). \end{aligned} \tag{5.23}$$

Now, if $f(z) \in S_{k_1}(\Gamma)$ has Fourier expansion

$$f(z) = a_f(1)\sum_{n=1}^{\infty}\lambda_f(n)n^{\frac{k_1-1}{2}}e(nz),$$

then we have the expansion

$$R_{k_1}^{k_2}(F_{k_1}(z)) = (-1)^{\frac{k_2-k_1}{2}}\beta(k_1,k_2)a_f(1)\sum_{n=1}^{\infty}\frac{\lambda_f(n)}{\sqrt{n}}W_{\frac{k_2}{2},\frac{k_1-1}{2}}(4\pi ny)e(nx), \tag{5.24}$$

where $F_{k_1}(z) = y^{k_1/2}f(z)$ as above.

5.3 Integral triple product identities

Fix f and g holomorphic cusp forms of weights k_1 and k_2 respectively with $k_1 \leq k_2$. Denote $F_{k_1} = y^{k_1/2} f(z)$ and $G_{k_2}(z) = y^{k_2/2} g(z)$. In this section we evaluate the inner products $\langle \phi F_{k_1}, G_{k_2} \rangle$, where ϕ is an automorphic form of weight $k_2 - k_1$. If ϕ is an Eisenstein series, we use the classical Rankin–Selberg integral method. If ϕ is a cusp form, we evaluate the triple product integral using Ichino's formula [45]. In both cases, it boils down to estimating central values $L(f \times g, 1/2)$ or $L(\phi \times f \times g, 1/2)$, to which we apply the subconvexity bounds of Soundararajan from [102].

We begin with the following proposition, which uses the Rankin–Selberg unfolding, see [46, Proposition 13.1].

Proposition 5.3.1. *We have*

$$(4\pi)^{1-s-\frac{k_1+k_2}{2}} \Gamma\left(s + \frac{k_1+k_2}{2} - 1\right) a_f(1)\overline{a_g(1)} \frac{L(f \times g, s)}{\zeta(2s)} = \int_X y^{(k_1+k_2)/2} f(z)\overline{g(z)} E_{k_2-k_1}(z,s) d\mu.$$

Proof. Using an unfolding argument, for $\mathrm{Re}(s) > 1$ we write the integral as

$$\begin{aligned}
\int_{\Gamma\backslash\mathbb{H}} F_{k_1}(z)\overline{G_{k_2}(z)} E_{k_2-k_1}(z) d\mu &= \int_{\Gamma\backslash\mathbb{H}} F_{k_1}(z)\overline{G_{k_2}(z)} \sum_{\gamma \in \Gamma_\infty\backslash\Gamma} (\mathrm{Im}\,\gamma z)^s j_\gamma(z)^{-(k_2-k_1)} d\mu \\
&= \sum_{\gamma \in \Gamma_\infty\backslash\Gamma} \int_{\Gamma\backslash\mathbb{H}} (\mathrm{Im}\,\gamma z)^s F_{k_1}(z)\overline{G_{k_2}(z)} j_\gamma(z)^{k_1-k_2} d\mu \\
&= \int_0^1 \int_0^\infty y^s y^{\frac{k_1+k_2}{2}} f(z)\overline{g(z)} dy dx \\
&= \int_0^1 \int_0^\infty y^{s+\frac{k_1+k_2}{2}} \sum_{n,m \geq 1} a_f(n)\overline{a_g(n)} e^{2\pi i(n-m)x} e^{-2\pi(n+m)y} dy dx \\
&= a_f(1)a_g(1) \sum_{n \geq 1} \lambda_f(n)\lambda_g(n) n^{\frac{k_1+k_2}{2}-1} \int_0^\infty y^{s+\frac{k_1+k_2}{2}} e^{-4\pi n y} dy \\
&= (4\pi)^{1-s-\frac{k_1+k_2}{2}} \Gamma\left(s + \frac{k_1+k_2}{2} - 1\right) a_f(1)\overline{a_g(1)} \sum_{n \geq 1} \lambda_f(n)\lambda_g(n) n^{-s}.
\end{aligned}$$

□

We now write the inner products involving the Eisenstein series.

Lemma 5.3.1. *Let $s = \frac{1}{2} + it$ and $\alpha = \frac{k_2-k_1}{2}$. Then*

$$|\langle E_{k_2-k_1}(\cdot, s) F_{k_1}, G_{k_2} \rangle| \ll_\epsilon \frac{(1+|t|)^{3/2} (\log k_2)^{-1+\epsilon} (1+\alpha)^{1/2}}{(L(1, \mathrm{sym}^2 f) L(1, \mathrm{sym}^2 g))^{1/2}}.$$

and, for $k_1 < k_2$,

$$\left|\left\langle E(\cdot,s)\, R_{k_1}^{k_2} F_{k_1}, G_{k_2}\right\rangle\right| \ll_\epsilon \frac{\Gamma(k_2-\alpha)^{1/2}}{\Gamma(k_2)^{1/2}\Gamma(\alpha)^{1/2}} k_2^\epsilon (1+|t|)^{3/2}\,|s(s+1)\ldots(s+\alpha-1)|\,.$$

Proof. We use Proposition 5.3.1 and (2.19) to obtain

$$\left|\left\langle E_{k_2-k_1}\left(\frac{1}{2}+it,\cdot\right) F_{k_1}, G_{k_2}\right\rangle\right| = \frac{\pi^{3/2}\Gamma\left(\frac{k_2+k_1}{2}-\frac{1}{2}+it\right) L\left(\frac{1}{2}+it, f\times g\right)}{\zeta(1+2it)\Gamma(k_1)^{1/2}\Gamma(k_2)^{1/2} L(1,\mathrm{sym}^2 f)^{1/2} L(1,\mathrm{sym}^2 g)^{1/2}}.$$

We use the weak subconvexity bound of Soundararajan [102]:

$$\left|L\left(\frac{1}{2}+it, f\times g\right)\right| \ll \frac{(k_1+k_2)^{1/2}(1+k_2-k_1)^{1/2}}{(\log k_2)^{1-\epsilon}}(1+|t|).$$

We now use that for $\sigma>0$, $|\Gamma(\sigma+it)|\le\Gamma(\sigma)$ and employ Stirling formula to deduce that $\Gamma(x+1/2)\sim\Gamma(x)\sqrt{x}$ as $x\to\infty$. Since $|\zeta(1+it)|\gg 1/\log(1+|t|)$, we obtain

$$\left|\left\langle E_{k_2-k_1}\left(\frac{1}{2}+it,\cdot\right) F_{k_1}, G_{k_2}\right\rangle\right| \ll_\epsilon \frac{\Gamma\left(\frac{k_2+k_1}{2}\right)(1+k_2-k_1)^{\frac{1}{2}}(1+|t|)^{1+\epsilon}}{(\log k_2)^{1-\epsilon}\Gamma(k_1)^{1/2}\Gamma(k_2)^{1/2} L(1,\mathrm{sym}^2 f)^{1/2} L(1,\mathrm{sym}^2 g)^{1/2}} \tag{5.25}$$

Finally we see that

$$\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)}{\Gamma(k_1)^{1/2}\Gamma(k_2)^{1/2}} = \frac{\binom{k_1+k_2-2}{k_1-1}^{1/2}}{\binom{k_1+k_2-2}{(k_1+k_2)/2-1}^{1/2}} \le 1.$$

and the first part follows.

Now, for the second part, we use the adjointness property (5.6), the product rule (5.7), together with the fact that $\Lambda_{k_2} G_{k_2}(z)=0$, to see that

$$\begin{aligned}
&\left\langle E\left(\frac{1}{2}+it,z\right) R_{k_1}^{k_2} F_{k_1}(z), G_{k_2}(z)\right\rangle \\
=&\beta(k_1,k_2)\left\langle E\left(\frac{1}{2}+it,z\right)(K_{k_2-2}\ldots K_{k_1}F_{k_1}(z)), G_{k_2}(z)\right\rangle \\
=&(-1)^{\frac{k_2-k_1}{2}}\beta(k_1,k_2)\left\langle\left(K_{k_2-k_1-2}\ldots K_0 E\left(\frac{1}{2}+it,z\right)\right) F_{k_1}(z), G_{k_2}(z)\right\rangle \\
=&(-1)^{\frac{k_2-k_1}{2}}\frac{\beta(k_1,k_2)\Gamma\left(\frac{k_2-k_1}{2}+\frac{1}{2}+it\right)}{\Gamma\left(\frac{1}{2}+it\right)}\left\langle E_{k_2-k_1}\left(\frac{1}{2}+it,z\right) F_{k_1}(z), G_{k_2}(z)\right\rangle.
\end{aligned}$$

If $k_1=k_2$, then the conclusion follows. Now assume $k_1<k_2$. Substituting $\beta(k_1,k_2)$

from (5.21) and using (5.25), we have that $\left|\left\langle E\left(\frac{1}{2}+it,\cdot\right)R_{k_1}^{k_2}F_{k_1},G_{k_2}\right\rangle\right|$ is bounded by

$$\frac{\Gamma\left(\frac{k_1+k_2}{2}\right)^{1/2}}{\Gamma(k_2)^{1/2}\Gamma\left(\frac{k_2-k_1}{2}\right)^{1/2}}\left|\prod_{0\leq j<\frac{k_2-k_1}{2}}\left(\frac{1}{2}+j+it\right)\right|\frac{(1+|t|)^{1+\epsilon}}{(\log k_2)^{1-\epsilon}S(f,g)^{\frac{1}{2}}}.$$

We use the bound $L(1,\mathrm{sym}^2 f)\gg(\log k_1)^{-1}$, see [42], and similarly for g. Hence the contribution from the last fraction is bounded by k_2^{ϵ} and the conclusion follows.

□

Next, we evaluate the inner products involving Hecke–Maaß cusp forms.

Lemma 5.3.2. *Let $\epsilon>0$. We have*

$$|\langle u_{j,k_2-k_1}F_{k_1},G_{k_2}\rangle|\ll_{\epsilon}\frac{(1+k_2-k_1)^{1/2}}{(\log k_2)^{1/2-\epsilon}L(1,\mathrm{sym}^2 f)^{1/2}L(1,\mathrm{sym}^2 g)^{1/2}}.$$

For N_ϵ large depending on ϵ, we have

$$\left|\left\langle u_j R_{k_1}^{k_2}F_{k_1},G_{k_2}\right\rangle\right|\ll_{\epsilon,t_j}\begin{cases}\dfrac{1}{(\log k_2)^{1/2-\epsilon}L(1,\mathrm{sym}^2 f)^{1/2}L(1,\mathrm{sym}^2 g)^{1/2}} & \text{if } k_2-k_1\leq N_\epsilon;\\ k_2^{-1+\epsilon} & \text{if } k_2-k_1\geq N_\epsilon.\end{cases}$$

Proof. From Ichino's formula [45], we know that

$$\left|\int_{\Gamma\backslash\mathbb{H}}u_{j,k_2-k_1}(z)F_{k_1}(z)\overline{G_{k_2}(z)}d\mu(z)\right|^2=\frac{1}{8}\frac{\Lambda(1/2,u_j\times f\times g)}{\Lambda(1,\mathrm{sym}^2 u_j)\Lambda(1,\mathrm{sym}^2 f)\Lambda(1,\mathrm{sym}^2 g)}I_\infty^*, \tag{5.26}$$

where I_∞^* is a certain local integral. When $k_1=k_2$, Watson [110] shows that $I_\infty^*=1$. For the general case, Woodbury [112] and Cheng [13] calculated for the real local place and show that $I_\infty^*=2^{-k_2+k_1}$.

We have that

$$\begin{aligned}\Lambda(s,f\times g\times u_j)=\prod_{\pm}&\Gamma_{\mathbb{R}}\left(s+\frac{k_1+k_2}{2}\pm it_j\right)\Gamma_{\mathbb{R}}\left(s+\frac{k_1+k_2}{2}-1\pm it_j\right)\\&\times\Gamma_{\mathbb{R}}\left(s+\frac{k_2-k_1}{2}\pm it_j\right)\Gamma_{\mathbb{R}}\left(s+\frac{k_2-k_1}{2}+1\pm ir\right)L(s,f\times g\times u_j).\end{aligned}$$

Then it follows that

$$|\langle u_{j,k_2-k_1}F_{k_1},G_{k_2}\rangle|^2 \ll_{t_j} \frac{\Gamma\left(\frac{k_1+k_2-1}{2}+it_j\right)\Gamma\left(\frac{k_1+k_2-1}{2}-it_j\right)L(1/2,f\times g\times u_j)}{\Gamma(k_1)\Gamma(k_2)L(1,\mathrm{sym}^2 f)L(1,\mathrm{sym}^2 g)}. \quad (5.27)$$

We use the weak subconvexity bound [102]

$$L\left(\frac{1}{2},f\times g\times u_j\right) \ll_{t_j,\epsilon} \frac{(k_1+k_2)(1+k_2-k_1)}{(\log k_2)^{1-\epsilon}}.$$

Similarly to the previous proof, we use that for $\sigma\geq 1/2$, we have that $\Gamma(\sigma+1/2)\asymp\sqrt{\sigma}\Gamma(\sigma)$ and $|\Gamma(\sigma+it_j)|\leq\Gamma(\sigma)$. Also, as before, we know that $\Gamma\left(\frac{k_1+k_2}{2}\right)^2\leq\Gamma(k_1)\Gamma(k_2)$ and then we conclude the first part of the lemma.

For the second part, we first note that

$$\begin{aligned}\left|\left\langle u_j R_{k_1}^{k_2}F_{k_1},G_{k_2}\right\rangle\right|^2 &= \beta(k_1,k_2)^2\,|\langle(K_{k_2-k_1-2}\dots K_0u_j)F_{k_1},G_{k_2}\rangle|^2\\ &= \frac{\beta(k_1,k_2)^2}{\alpha(s_j,k_2-k_1)^2}|\langle u_{j,k_2-k_1}F_{k_1},G_{k_2}\rangle|^2.\end{aligned} \quad (5.28)$$

Now fix N_ϵ large enough such that $N_\epsilon>1/\epsilon$ and $\log n<n^\epsilon$, for $n\geq N_\epsilon$. We treat two separate cases, depending on whether k_2-k_1 is smaller or larger than N_ϵ.

1. If $0\leq k_2-k_1\leq N_\epsilon$. Then from definitions of (5.20) (5.21), we see that

$$\frac{\beta(k_1,k_2)^2}{\alpha(s_j,k_2-k_1)^2}\ll_{\epsilon,t_j}1$$

   and the conclusion follows.

2. If $k_2-k_1\geq N^\epsilon$. For notation simplicity, denote $\alpha=(k_2-k_1)/2$. We also use the bounds $L(1,\mathrm{sym}^2 f)\gg(\log k_1)^{-1}$ and $L(1,\mathrm{sym}^2 g)\gg(\log k_2)^{-1}$. Now, from (5.20) and (5.21), we see that

$$\begin{aligned}\left|\frac{\beta(k_1,k_2)^2}{\alpha(1/2+it_j,k_2-k_1)^2}\right| &\ll_{t_j}\frac{\Gamma(k_1)\Gamma\left(\alpha+\frac{1}{2}\right)^2}{\Gamma\left(\frac{k_1+k_2}{2}\right)\Gamma(\alpha+1)}=\frac{\Gamma(k_1)\Gamma(\alpha+1)}{\Gamma\left(\frac{k_1+k_2}{2}\right)}\frac{\Gamma\left(\alpha+\frac{1}{2}\right)^2}{\Gamma(\alpha+1)^2}\\ &\ll\left(\frac{\frac{k_1+k_2}{2}-1}{\alpha}\right)^{-1}\frac{1}{\alpha}\ll k_2^{-1}\alpha^{-1}\end{aligned}$$

   Now the conclusion follows from (5.28).

□

5.4 Bounds for Fourier coefficients

In order to evaluate Fourier coefficients of automorphic forms of weight k, it is useful to define

$$F(k,t,y) := \frac{W_{k,it}(u)}{\Gamma\left(\frac{1}{2}+k+it\right)} + \frac{W_{-k,it}(u)}{\Gamma\left(\frac{1}{2}-k+it\right)}. \tag{5.29}$$

In [50], Jakobson evaluated this expression as

$$F(k,t,y) = 2(-1)^k \sum_{l=0}^{k} \frac{(-k)_l (k)_l y^l}{(1/2)_l 4^l l!} \frac{W_{0,l+it}(y)}{\Gamma\left(\frac{1}{2}+l+it\right)},$$

where the Pochhammer symbol $(x)_l$ is defined by

$$(x)_l := x(x+1)\dots(x+l-1); \quad (x)_0 = 1.$$

We use the fact that $W_{0,\nu}(y) = \sqrt{(y/\pi)}K_\nu(y/2)$. We apply the integral representation of the K-Bessel function [47, p. 205]

$$K_\nu(y) = \pi^{-1/2}\Gamma\left(\nu+\frac{1}{2}\right)\left(\frac{y}{2}\right)^{-\nu}\int_0^\infty (u^2+1)^{-\nu-1/2}\cos(uy)du,$$

which holds for $y > 0$ and $\mathrm{Re}(\nu) > -1/2$. From this we obtain

$$\begin{aligned}\frac{y^l W_{0,l+it}(y)}{\Gamma\left(\frac{1}{2}+l+it\right)} &\ll y^{1/2}\left|\int_0^\infty (u^2+1)^{-l-1/2-it}\cos(uy)du\right| \\ &\ll y^{1/2}\left(\frac{1+l+|t|}{y}\right)^A\left(1+\frac{1+|t|}{y}\right)^\epsilon,\end{aligned}$$

for any $\epsilon > 0$ and any integer $A \geq 0$.

Next we note that

$$\left|\frac{(-k)_l(k)_l y^l}{(1/2)_l 4^l l!}\right| = \frac{k}{k+l}\binom{k+l}{l},$$

hence using the identity

$$\sum_{l=0}^{m}\binom{k+l}{l} = \binom{k+m+1}{m},$$

we see that

$$F(k,t,y) \ll 4^k k^A \sqrt{y}\left(\frac{1+|t|}{y}\right)^A\left(1+\frac{1+|t|}{y}\right)^\epsilon. \tag{5.30}$$

Also, from [46, B. 36], we have the asymptotic for large y

$$K_\nu(y) = \left(\frac{\pi}{2y}\right)^{1/2} e^{-y}\left(1 + O\left(\frac{1+|\nu|^2}{y}\right)\right).$$

Now we are ready to give bounds for the Fourier coefficients of incomplete Eisenstein series.

Lemma 5.4.1. *Let $E_k(z|\psi)$ an incomplete Eisenstein series with Fourier expansion*

$$E_k(z|\psi) = \sum_{n\in\mathbb{Z}} a_n(y)e(nx).$$

Then

$$a_0(y) = \delta_{k=0}\frac{3}{\pi}\Psi(-1) + O\left(\sqrt{y}\right),$$

and for $n \neq 0$, we have

$$a_n(y) + a_{-n}(y) \ll 2^k k^A \sqrt{y}\tau(|n|)\left(\frac{1}{|n|y}\right)^A \left(1 + \frac{1}{|n|y}\right)^\epsilon,$$

for any $\epsilon > 0$ and any integer $A \geq 0$.

Proof. Using (5.15) and (5.13), we note that

$$a_0(y) = \frac{1}{2\pi i}\int_{(\sigma)} \Psi(-s)\left(y^s + \frac{(-1)^{k/2}\Gamma(s)^2}{\Gamma\left(s-\frac{k}{2}\right)\Gamma\left(s+\frac{k}{2}\right)}\phi(s)y^{1-s}\right) ds,$$

for some $\sigma > 1$. We want to move the line of integration to $\mathrm{Re}(s) = 1/2$ and we notice we that encounter a pole at $s = 1$ if and only if $k = 0$. Using the duplication formula $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$ and that $|\Gamma(1/2+it)|^2 = \frac{\pi}{\cosh \pi t}$, we observe that

$$\left|\frac{\Gamma\left(\frac{1}{2}+it\right)^2}{\Gamma\left(\frac{1}{2}+it+\frac{k}{2}\right)\Gamma\left(\frac{1}{2}+it-\frac{k}{2}\right)}\right| = \left|\frac{\frac{\pi}{\cosh \pi t}}{\frac{\pi}{\sin \pi\left(\frac{1}{2}+it+\frac{k}{2}\right)}}\right| \sim 1 \quad \text{as } |t| \to \infty.$$

Hence, by (5.14), we have that

$$a_0(y) = \delta_{k=0}\frac{3}{\pi}\Psi(-1) + O\left(\sqrt{y}\right). \tag{5.31}$$

Note that, by unfolding, we see that

$$\langle E_0(z|\psi), 1\rangle = \int_{-1/2}^{1/2}\int_0^\infty \psi(y)\frac{dxdy}{y^2} = \Psi(-1).$$

Similarly, for $n \neq 0$, we have that

$$a_n(y) = \frac{1}{2\pi i}\int_{-\infty}^{\infty}\Psi\left(-\frac{1}{2}-it\right)\frac{(-1)^{k/2}\Gamma\left(\frac{1}{2}+it\right)}{2\Gamma\left(\frac{1}{2}+\frac{k}{2}+it\right)\xi(1+2it)}|l|^{-\frac{1}{2}}\left(\sum_{ab=|n|}\left(\frac{a}{b}\right)^{it}\right)W_{\frac{k}{2},it}(4\pi|n|y)dt.$$

We easily see that

$$a_n(y) + a_{-n}(y) \ll \tau(|n|)|n|^{-1/2}\int_{-\infty}^{\infty}\Psi\left(-\frac{1}{2}-it\right)\frac{\Gamma\left(\frac{1}{2}+it\right)}{\xi(1+2it)}F\left(\frac{k}{2}, t, 4\pi|n|y\right).$$

The conclusion follows from (5.30). □

Next we turn our attention to the Fourier coefficients of Maaß cusp forms.

Lemma 5.4.2. *Let $u_{j,k}$ be a Maaß cusp form as defined in the previous section with eigenvalue $1/4 + t_j^2$. If its Fourier expansion is given by*

$$u_{j,k}(z) = \sum_{n\in\mathbb{Z}} a_n(y)e(nx),$$

then $a_0(y) = 0$ and for $n \neq 0$, we have that

$$a_n(y) + a_{-n}(y) \ll 2^k k^A \sqrt{y}|c_j(|n|)|\left(\frac{1+|t_j|}{|n|y}\right)^A\left(1+\frac{1+|t_j|}{|n|y}\right)^\epsilon.$$

Proof. From (5.23), we see that for $n \neq 0$, we have that

$$a_n(y) + a_n(-y) = \Gamma(1/2 + it_j)c_j(|n|)|n|^{-1/2}F(k/2, t_j, 4\pi|n|y).$$

Now the conclusion simply follows from (5.30). □

Next we develop a formula for Whittaker functions of the form $W_{\alpha+k,\alpha-\frac{1}{2}}(y)$, which is useful for expressing the Fourier coefficients of $R_{k_1}^{k_2}F_{k_1}$.

Lemma 5.4.3. *Let $\alpha > 0$ and $k \geq 0$ an integer. Then*

$$W_{\alpha+k,\alpha-\frac{1}{2}}(y) = e^{-\frac{y}{2}} y^{\alpha} \sum_{l=0}^{k} y^{k-l} (-1)^l \binom{k}{l} \frac{\Gamma(2\alpha+k)}{\Gamma(2\alpha+k-l)}.$$

In particular, this implies that for $y > 0$ and $\alpha \geq 1$, we have

$$W_{\alpha+k,\alpha-\frac{1}{2}}(y) \ll 2^k e^{-\frac{y}{2}} y^{\alpha}((2\alpha+k)^k + y^k).$$

Proof. We proceed by induction on k. From [28, (4.21)], we see that

$$W_{\alpha,\alpha-\frac{1}{2}}(y) = y^{\alpha} e^{-y/2}.$$

We use the recursion formula [39, (9.234)]

$$W_{\lambda+1,\mu}(y) = \left(\frac{1}{2}y - \lambda\right) W_{\lambda,\mu}(y) - yW'_{\lambda,\mu}(y).$$

We we see that $W_{\alpha+k,\alpha-\frac{1}{2}}(y)$ is of the form

$$W_{\alpha+k,\alpha-\frac{1}{2}}(y) = e^{-\frac{y}{2}} y^{\alpha} \sum_{l=0}^{k} y^{k-l} P_{k,l}(\alpha)$$

where $P_{k,l}(X)$ polynomial of degree l. The recursion formula gives us that, for $1 \leq l \leq k$, we have

$$P_{k+1,l}(\alpha) = P_{k,l}(\alpha) - (2\alpha + 2k - l + 1)P_{k,l-1}(\alpha).$$

Moreover, $P_{k,0}(\alpha) = 1$ and $P_{k,k}(\alpha) = (-1)^k (2\alpha)_k$, for all k. If we write $Q_{k,l}(X) = P_{k,l}\left(\frac{X}{2}\right)$, one can check by inducion on k that

$$Q_{k,l}(X) = (-1)^l \binom{k}{l} (X+k-1)(X+k-2)\cdots(X+k-l).$$

The conclusion follows. □

5.5 Shifted convolution sums

Let $\phi \in \mathcal{L}_{k_2-k_1}(X)$ with Fourier expansion

$$\phi(z) = a_0(y) + \sum_{l \neq 0} a_l(y) e(lx).$$

We want to evaluate $\langle \phi F_{k_1}, G_{k_2} \rangle$ by applying Holowinsky's approach [43] by relating the inner product to shifted convolution sums. In this section we prove the following theorem.

Theorem 5.5.1. *Define*

$$M_{k_1,k_2}(f,g) := \frac{1}{(\log k_2)^2 L(1, \mathrm{sym}^2 f)^{1/2} L(1, \mathrm{sym}^2 g)^{1/2}} \prod_{p \leq k_2} \left(1 + \frac{|\lambda_f(p)|}{p}\right)\left(1 + \frac{|\lambda_g(p)|}{p}\right). \tag{5.32}$$

Fix $\epsilon > 0$. Then there exists a constant N_ϵ such that the following hold.

i Let u_{j,k_2-k_1} be a Hecke–Maaß form as above with eigenvalue $1/4 + t_j^2$. Then

$$\langle u_{j,k_2-k_1} F_{k_1}, G_{k_2} \rangle \ll_{t_j,\epsilon} 2^{k_2-k_1}(1 + k_2 - k_1)^{N_\epsilon} M_{k_1,k_2}(f)^{1/2} (\log k_2)^\epsilon.$$

ii For an incomplete Eisenstein series $E_{k_2-k_1}(z|\psi)$, we have that $\langle E_{k_2-k_1}(\cdot|\psi) F_{k_1}, G_{k_2} \rangle - \delta_{f=g} \frac{3}{\pi} \langle E_0(\cdot|\psi), 1 \rangle$ is bounded by

$$O_{\psi,\epsilon}\left(2^{k_2-k_1}(1 + k_2 - k_1)^{N_\epsilon} M_{k_1,k_2}(f)^{1/2} (\log k_2)^\epsilon (1 + R_{k_1,k_2}(f,g))\right),$$

where

$$R_{k_1,k_2}(f,g) = \frac{1}{k_2^{1/2} L(1, \mathrm{sym}^2 f)^{1/2} L(1, \mathrm{sym}^2 g)^{1/2}} \int_{-\infty}^{+\infty} \frac{\left|L\left(f \times g, \frac{1}{2} + it\right)\right|}{(|t| + 1)^5} dt.$$

Fix ψ smooth and compactly supported on $\mathbb{R}^+$ and $\Psi(s)$ its Mellin transform. For $Y \geq 1$, we define

$$I_\phi(Y) := \frac{1}{2\pi i} \int_{(\sigma)} \Psi(-s) Y^s \int_X E(z,s) \phi(z) F_{k_1}(z) \overline{G_{k_2}(z)} d\mu ds \tag{5.33}$$

for $\sigma > 1$.

Lemma 5.5.1. *For ϕ a fixed a Hecke–Maaß cusp form or incomplete Eisenstein series,* *we have*

$$\langle \phi F_{k_1}, G_{k_2} \rangle = c_Y^{-1} I_\phi(Y) + O_\psi\left(Y^{-1/2}\right) \ , \tag{5.34}$$

where

$$c_Y := \frac{3}{\pi}\Psi(-1)Y \ . \tag{5.35}$$

Proof. We move the contour of integration in (5.33) to the line $\mathrm{Re}(s) = 1/2$. There is a pole at $s = 1$ coming from the Eisenstein series, with residue

$$\Psi(-1)Y\left(\mathrm{Res}_{s=1} E(z,s)\right)\langle \phi F_{k_1}, G_{k_2}\rangle = c_Y \langle \phi F_{k_1}, G_{k_2}\rangle \ .$$

Therefore we obtain

$$I_\phi(Y) = c_Y \langle \phi F_{k_1}, G_{k_2}\rangle + \int_X p(z)\phi(z)F_{k_1}(z)\overline{G_{k_2}(z)}d\mu \ , \tag{5.36}$$

where

$$p(z) := \int_{(1/2)} \Psi(-s)Y^s E(z,s) ds \ .$$

On the line $\mathrm{Re}(s) = 1/2$, from [43, Lemma 2.1], we have

$$E(z,s) \ll_\epsilon \sqrt{y} + |s|^2 y^{-3/2}(1+|s|/y)^\epsilon.$$

Using the fast decay of $\Psi(s)$, we obtain $p(z) \ll \sqrt{yY}$ if $y \geq 1/2$. Going back to (5.36), if we assume $\sqrt{y}|\phi(z)|$ is bounded on X, we conclude that

$$\int_X p(z)\phi(z)F_{k_1}(z)\overline{G_{k_2}(z)}d\mu \ll_{\phi,\psi} \sqrt{Y} \ .$$

The assumption that $\sqrt{y}|\phi(z)|$ is bounded on X is true for cusp forms and incomplete Eisenstein series. □

We observe that

$$I_\phi(Y) = \frac{1}{2\pi i}\int_0^\infty \psi(Yy)y^{-2}\left(\int_{-1/2}^{1/2}\phi(z)F_{k_1}(z)\overline{G_{k_2}(z)}dx\right)dy \ . \tag{5.37}$$

This follows from using a standard unfolding argument and then applying the inverse Mellin transform.

Proposition 5.5.1. *Let $Y > 1$. For any $\epsilon > 0$, there exists a constant N_ϵ such that, for ϕ a Hecke–Maaß cusp form or incomplete Eisenstein series, we have*

$$\begin{aligned}\langle \phi F_{k_1}, G_{k_2}\rangle =& c_Y^{-1}\int_0^\infty \psi(Yy)y^{-2}\left(\int_{-1/2}^{1/2}\phi^*(z)F_{k_1}(z)\overline{G_{k_2}(z)}dx\right)dy\\ &+O\left(2^{k_2-k_1}(k_2-k_1+1)^{N_\epsilon}Y^{-1/2}\right),\end{aligned}$$

where

$$\phi^*(z) := \sum_{|l|<Y^{1+\epsilon}} a_l(y)e(lx)\ .$$

Proof. We evaluate the contribution to $I_\phi(Y)$ coming from large Fourier coefficients $a_l(y)$ of ϕ. Assume ϕ is an incomplete Eisenstein series of weight k_2-k_1. We make use of Lemma 5.4.1. The contribution coming from Fourier coefficients larger than $Y^{1+\epsilon}$ is bounded by

$$\begin{aligned}&\sum_{|l|\geq Y^{1+\epsilon}}\int_0^\infty\int_{-1/2}^{1/2}\psi(Yy)y^{-2}a_l(y)|F_{k_1}(z)||G_{k_2}(z)|dxdy\\ \ll& 2^{k_2-k_1}(k_2-k_1+1)^A\left(\int_0^\infty\int_{-1/2}^{1/2}\psi(Yy)y^{-2}|F_{k_1}(z)||G_{k_2}(z)|dxdy\right)Y^{A-1/2+\epsilon}\sum_{l>Y^{1+\epsilon}}\frac{\tau(l)}{l^A}\\ \ll& 2^{k_2-k_1}(k_2-k_1+1)^AY^{A+1/2+\epsilon}Y^{(1+\epsilon)(1-A)} \ll 2^{k_2-k_1}(k_2-k_1+1)^AY^{-1/2},\end{aligned}$$

if we choose A large enough with respect to ϵ. We note that the double integral is bounded by $O(Y)$, since

$$\int_{\Gamma\backslash\mathbb{H}}|F_{k_1}(z)||G_{k_2}(z)|d\mu(z)\leq \|F_{k_1}\|+\|G_{k_2}\|=2,$$

and we know that $y \asymp 1/Y$, $-1/2\leq x\leq 1/2$, and by [47, Lemma 2.10] we know there are $O(Y)$ copies of the fundamental domain in this region. The proof for Maaß forms follows similarly. □

For an integer l, we define

$$S_l(Y) := \int_0^\infty \psi(Yy)y^{-2}\left(\int_{-1/2}^{1/2}a_l(y)e(lx)F_{k_1}(z)\overline{G_{k_2}(z)}dx\right)dy\ . \tag{5.38}$$

Hence

$$c_Y\langle\phi F_{k_1},G_{k_2}\rangle = S_0(Y)+\sum_{0<|l|<Y^{1+\epsilon}}S_l(Y)+O\left(2^{k_2-k_1}(k_2-k_1+1)^{N_\epsilon}Y^{1/2}\right)\ . \tag{5.39}$$

We note that $S_0(Y) \equiv 0$ when ϕ is a cusp form.

Lemma 5.5.2. *Let* $Y \geq 1$ *and* $\phi = E_{k_2-k_1}(z|h)$ *an incomplete Eisenstein series of weight* $k_2 - k_1$*. Then*

$$\begin{aligned} c_Y^{-1} S_0(Y) =& \delta_{k_1=k_2} \frac{3}{\pi} \langle \phi, 1 \rangle + O(Y^{-1/2}) \\ &+ O\left(\left(Yk_2 \right)^{-1/2} (|L(\mathrm{sym}^2 f, 1) L(\mathrm{sym}^2 g, 1)|\Big)^{-1/2} \int_{-\infty}^{\infty} \frac{\left| L\left(f \times g, \frac{1}{2} + it\right) \right|}{(|t|+1)^5} dt \right). \end{aligned}$$

Proof. From the definition of $S_0(Y)$ and (5.31), we obtain

$$S_0(Y) = \left(\delta_{k_1=k_2} \frac{3}{\pi} \langle \phi, 1 \rangle + O\left(Y^{-1/2}\right) \right) \int_0^\infty \psi(Yy) y^{\frac{k_1+k_2}{2}-2} \left(\int_{-1/2}^{1/2} f(z) \overline{g(z)} dx \right) dy \ .$$

Expanding the product $f(z)\overline{g(z)}$ as a Fourier sum and computing the inner integral above, we obtain

$$S_0(Y) = \left(\delta_{k_1=k_2} \frac{3}{\pi} \langle \phi, 1 \rangle + O\left(Y^{-1/2}\right) \right) \sum_{n \geq 1} a_f(n) \overline{a_g(n)} \int_0^\infty \psi(Yy) y^{\frac{k_1+k_2}{2}-2} e^{-4\pi n y} dy \ .$$

We evaluate the integral in y using the inverse Mellin transform.

$$\begin{aligned} \int_0^\infty \psi(Yy) y^{\frac{k_1+k_2}{2}-2} e^{-4\pi n y} dy &= \int_0^\infty \left(\frac{1}{2\pi i} \int_{(\sigma)} (Yy)^s \Psi(-s) ds \right) y^{\frac{k_1+k_2}{2}-2} e^{-4\pi n y} dy \\ &= \frac{1}{2\pi i} \int_{(\sigma)} Y^s \Psi(-s) (4\pi n)^{-s-\frac{k_1+k_2}{2}+1} \Gamma\left(s + \frac{k_1+k_2}{2} - 1\right) ds, \end{aligned}$$

Hence

$$\begin{aligned} &\sum_{n \geq 1} a_f(n) \overline{a_g(n)} \int_0^\infty \psi(Yy) y^{\frac{k_1+k_2}{2}-2} e^{-4\pi n y} dy = \\ =& \frac{1}{2\pi i} a_f(1) \overline{a_g(1)} (4\pi)^{1-\frac{k_1+k_2}{2}} \int_{(\sigma)} \left(\frac{Y}{4\pi} \right)^s \Psi(-s) \frac{L(f \times g, s)}{\zeta(2s)} \Gamma\left(s + \frac{k_1+k_2}{2} - 1\right) ds \ . \end{aligned}$$

We move the contour of integration to the line $\mathrm{Re}(s) = 1/2$. We note that we pick up a pole at $s = 1$ if and only if $f = g$. In this case, we use (2.19) to compute the residue. Therefore, we obtain

$$\sum_{n \geq 1} a_f(n) \overline{a_g(n)} \int_0^\infty \psi(Yy) y^{\frac{k_1+k_2}{2}-2} e^{-4\pi n y} dy = \delta_{f=g} \frac{3}{\pi} \Psi(-1) Y + E(Y),$$

where

$$E(Y) = \frac{1}{2\pi i} a_f(1)\overline{a_g(1)}(4\pi)^{1-\frac{k_1+k_2}{2}} \times \int_{-\infty}^{\infty} \left(\frac{Y}{4\pi}\right)^{1/2+it} \Psi\left(-\frac{1}{2}-it\right) \frac{L\left(f\times g, \frac{1}{2}+it\right)}{\zeta(1+2it)} \Gamma\left(\frac{k_1+k_2}{2}-\frac{1}{2}+it\right) dt.$$

From [106, p.51], we know that $\zeta(1+it) \gg (\log t)^{-7}$. Hence, using the rapid decay of $\Psi(-s)$ guaranteed by (5.14) and expanding $a_f(1)\overline{a_g(1)}$ as in (2.19), we obtain

$$\begin{aligned} E(Y) &\ll Y^{1/2} \frac{\binom{k_1+k_2-2}{k_1-1}^{1/2}}{(k_1+k_2)^{1/2}\binom{k_1+k_2-2}{\frac{k_1+k_2}{2}-1}^{1/2}} |L(\mathrm{sym}^2 f,1)L(\mathrm{sym}^2 g,1)|^{-1/2} \int_{-\infty}^{\infty} \frac{\left|L\left(f\times g,\frac{1}{2}+it\right)\right|}{(|t|+1)^{10}} dt \\ &\ll Y^{1/2} k_2^{-1/2} |L(\mathrm{sym}^2 f,1)L(\mathrm{sym}^2 g,1)|^{-1/2} \int_{-\infty}^{\infty} \frac{\left|L\left(f\times g,\frac{1}{2}+it\right)\right|}{(|t|+1)^{10}} dt. \end{aligned}$$

□

Lemma 5.5.3. *Let ϕ be a fixed automorphic form. Then for $l \neq 0$, we have*

$$c_Y^{-1} S_l(Y) \ll \left| \frac{a_l(Y^{-1})}{L(\mathrm{sym}^2 f,1)^{1/2} L(\mathrm{sym}^2 g,1)^{1/2}} \right| \left(\frac{1}{Yk_2} \sum_{n \asymp Yk_2} |\lambda_f(n)\lambda_g(n+l)| + Y^{\epsilon}(k_1+k_2)^{-1+\epsilon} \right).$$

Proof. Expanding the Fourier sum in the definition (5.38), we obtain

$$S_l(Y) = \sum_{n\geq 1} a_f(n)\overline{a_g(n+l)} \int_0^{\infty} \psi(Yy) a_l(y) y^{\frac{k_1+k_2}{2}-2} e^{-2\pi(2n+l)dy} .$$

We note that the inner integral is only supported for $y \asymp 1/Y$. Hence

$$S_l(Y) \ll |a_l(Y^{-1})| \sum_{n\geq 1} a_f(n)\overline{a_g(n+l)} \int_0^{\infty} \psi(Yy) y^{\frac{k_1+k_2}{2}-2} e^{-2\pi(2n+l)} dy .$$

Similarly as in the proof of Lemma 5.5.2, using the inverse Mellin transform and evaluating the inner integral, we obtain

$$S_l(Y) \ll |a_l(Y^{-1})| \sum_{n\geq 1} a_f(n)\overline{a_g(n+l)} \frac{1}{2\pi i} \int_{(\sigma)} Y^s \Psi(-s)(2\pi(2n+l))^{1-s-\frac{k_1+k_2}{2}} \Gamma\left(s+\frac{k_1+k_2}{2}-1\right) ds.$$

From (2.19), we see that

$$S_l(Y) \ll \left| \frac{a_l(Y^{-1})}{L(\text{sym}^2 f, 1)^{1/2} L(\text{sym}^2 g, 1)^{1/2}} \right| \sum_{n \geq 1} |\lambda_f(n)\lambda_g(n+l)| A_{n,l}(Y) \ ,$$

where

$$A_{n,l}(Y) := \left(\frac{n^{\frac{k_1-1}{2}} (n+l)^{\frac{k_2-1}{2}}}{\left(n+\frac{l}{2}\right)^{\frac{k_1+k_2}{2}-1}} \right) \frac{1}{2\pi i} \int_{(\sigma)} \Psi(-s) \left(\frac{Y}{2\pi(2n+l)} \right)^s \frac{\Gamma(s+\frac{k_1+k_2}{2}-1)}{\Gamma(k_1)^{1/2}\Gamma(k_2)^{1/2}} ds.$$

If we interchange f and g, which we can without losing the generality, then the first term will be bounded above by 1. From Stirling's relations, any vertical strip $0 < a \leq \text{Re}(s) \leq b$ and $k > 1$, we have

$$\frac{\Gamma(s+\alpha)}{\Gamma(\alpha)} = \alpha^s \left(1 + O_{a,b}\left((|s|+1)^2 \alpha^{-1}\right)\right), \tag{5.40}$$

see [43, (19)]. Choosing the line of integration $\text{Re}(s) = \sigma = 1+\epsilon$, we obtain

$$\begin{aligned} A_{n,l}(Y) &\ll \frac{1}{2\pi i} \int_{(\sigma)} \Psi(-s) \left(\frac{Y}{2\pi(2n+l)} \right)^s \frac{\Gamma(s+\frac{k_1+k_2}{2}-1)}{\Gamma\left(\frac{k_1+k_2}{2}-1\right)} \frac{\Gamma\left(\frac{k_1+k_2}{2}-1\right)}{\Gamma(k_1)^{1/2}\Gamma(k_2)^{1/2}} ds \\ &\ll \frac{\binom{k_1+k_2-2}{k_1-1}^{1/2}}{(k_1+k_2)\binom{k_1+k_2-2}{\frac{k_1+k_2}{2}-1}^{1/2}} \left(\psi\left(\frac{Y\left(\frac{k_1+k_2}{2}-1\right)}{2\pi(2n+l)} \right) + (k_1+k_2)^{\epsilon} \left(\frac{Y}{2n+l} \right)^{1+\epsilon} \right). \end{aligned}$$

Therefore we get

$$S_l(Y) \ll \left| \frac{a_l(Y^{-1})}{L(\text{sym}^2 f, 1)^{1/2} L(\text{sym}^2 g, 1)^{1/2}} \right| \left(\frac{1}{k_2} \sum_{n \asymp Y k_2} |\lambda_f(n)\lambda_g(n+l)| + Y^{1+\epsilon}(k_1+k_2)^{-1+\epsilon} \right).$$

□

We recall [43, Theorem 1.2].

Theorem 5.5.2. *Let $\lambda_1(n)$ and $\lambda_2(n)$ be multiplicative functions such that $|\lambda_i(n)| \leq \tau(n)$. Then for any $0 < \delta < 1$ and any fixed integer $0 < |l| \leq x$, we have*

$$\sum_{n \leq x} |\lambda_1(n)\lambda_2(n+l)| \ll x(\log x)^{-2+\delta} \tau(|l|) \prod_{p \leq z} \left(1 + \frac{|\lambda_1(p)|}{p}\right) \left(1 + \frac{|\lambda_2(p)|}{p}\right)$$

where $z = \exp\left(\frac{\log x}{\delta \log\log x}\right)$.

We apply Theorem 5.5.2 with $\lambda_1 = \lambda_f$ and $\lambda_2 = \lambda_g$. The Ramanujan conjecture for holomorphic cusp forms ensures that the conditions in the statement of the theorem are satisfied. There exists a constant C_ψ such that, for all $\epsilon > 0$

$$\sum_n |\lambda_f(n)\lambda_g(n+l)|\psi\left(\frac{Y\left(\frac{k_1+k_2}{2}-1\right)}{2\pi(2n+l)}\right) \ll \sum_{n\leq C_\psi Y(k_1+k_2)} \lambda_f(n)\lambda_g(n+l)$$
$$\ll \tau(|l|)Y^{1+\epsilon}(k_1+k_2)(\log(k_1+k_2))^{-2+\epsilon}\prod_{p\leq(k_1+k_2)^\epsilon}\left(1+\frac{|\lambda_1(p)|}{p}\right)\left(1+\frac{|\lambda_2(p)|}{p}\right).$$

Case 1: ϕ is an incomplete Eisenstein series. Using Lemma 5.4.1, we have that

$$S_l(Y) + S_{-l}(Y) \ll 2^{k_2-k_1}S(f,g)^{-1/2}Y^{1/2+\epsilon}\tau(l)^2(\log k_2)^{-2+\epsilon}\prod_{p\leq k_2}\left(1+\frac{|\lambda_1(p)|}{p}\right)\left(1+\frac{|\lambda_2(p)|}{p}\right).$$

We use the trivial bound

$$\sum_{1\leq l<Y^{1+\epsilon}} \tau(l)^2 \ll Y^{1+\epsilon}$$

to see that

$$C_Y^{-1}\sum_{0<|l|<Y^{1+\epsilon}} S_l(Y) \ll 2^{k_2-k_1}Y^{1/2+\epsilon}M_{k_1,k_2}(f,g).$$

Case 2: $\phi = u_{j,k_2-k_1}$ is a Hecke–Maaß cusp form. It is very similar to the above case, where we employ Lemma 5.4.2 instead. While we sum $S_l(Y)$, we need to bound

$$\sum_{0<l<Y^{1+\epsilon}} \tau(l)c_j(l) \ll \left(\sum_{0<l<Y^{1+\epsilon}} \tau(l)^2\right)^{1/2}\left(\sum_{0<l<Y^{1+\epsilon}} c_j(l)^2\right)^{1/2} \ll Y^{1+\epsilon},$$

where the bound for the second sum over the Hecke eigenvalues follows from [47, p. 55].

To finish the proof of Theorem 5.5.1, we simply choose $Y = M_{k_1,k_2}(f,g)^{-1}$. If $M_{k_1,k_2}(f,g) > 1$, we take $Y = 1$.

5.6 Proofs of Theorem 5.1.1 and Theorem 5.1.2

Lemma 5.6.1. *If $k_1 \leq k_2$ and $\log k_1 \geq C\log k_2$, for some absolute constant C, then*

$$M_{k_1,k_2}(f,g) \ll_\epsilon (\log k_2)^{1/6+\epsilon}L(1,\mathrm{sym}^2 f)^{\frac{1}{4}}L(1,\mathrm{sym}^2 g)^{\frac{1}{4}}.$$

Proof. The key input is [44, Lemma 2] which states that

$$L(1,\mathrm{sym}^2 f) \gg (\log\log k_1)^{-3} \exp\left(\sum_{p\le k_1} \frac{\lambda_f(p^2)}{p}\right), \tag{5.41}$$

and a similar statement holds for $L(1,\mathrm{sym}^2 g)$. As in [44, Lemma 3], we use the inequality $|x| \le \frac{1}{3} + \frac{3}{4}x^2$ and the Hecke relations $\lambda_f(p^2) = \lambda_f(p)^2 - 1$ to see that

$$\begin{aligned}\sum_{p\le k_1} \frac{|\lambda_f(p)|}{p} &\le \frac{1}{3}\sum_{p\le k_1}\frac{1}{p} + \frac{3}{4}\sum_{p\le k_1}\frac{\lambda_f(p)^2}{p}\\ &= \frac{13}{12}\sum_{p\le k_1}\frac{1}{p} + \frac{3}{4}\sum_{p\le k_1}\frac{\lambda_f(p^2)}{p}\\ &\le \frac{13}{12}\log\log k_1 + \frac{3}{4}\sum_{p\le k_1}\frac{\lambda_f(p^2)}{p} + O(1).\end{aligned}$$

Now the conclusion follows from (5.41) and the fact that $\log k_1 \asymp \log k_2$. □

5.6.1 Proof of Theorem 5.1.2

From the analysis in Section 5.2, it suffices to bound $\langle u_{j,k_2-k_1}F_{k_1}, G_{k_2}\rangle$ and $\langle E_{k_2-k_1}(z|\psi)F_{k_1}, G_{k_2}\rangle$. We have who cases, depending on the size of $L(1,\mathrm{sym}^2 f)L(1,\mathrm{sym}^2 g)$.

Case (i): Suppose $L(1,\mathrm{sym}^2 f)L(1,\mathrm{sym}^2 g) \ge (\log k_2)^{-5/6}$. Then by Lemma 5.3.2, we have that

$$|\langle u_{j,k_2-k_1}F_{k_1}, G_{k_2}\rangle| \ll_\epsilon \left(1 + \frac{k_2-k_1}{2}\right)^{1/2} (\log k_2)^{-1/12+\epsilon}.$$

For the Eisenstein case, from (5.15) we know that

$$E_{k_2-k_1}(z|\psi) = \delta_{k_1=k_2}\frac{3}{\pi}\Psi(-1) + \frac{1}{2\pi i}\int_{(1/2)} \Psi(-s)E_{k_2-k_1}(z,s)ds.$$

Hence

$$\langle E_{k_2-k_1}(z|\psi)F_{k_1}, G_{k_2}\rangle = \delta_{f=g}\frac{3}{\pi}\Psi(-1) + \int_{-\infty}^{\infty} \Psi\left(-\frac{1}{2}-it\right)\left\langle E_{k_2-k_1}\left(\cdot, \frac{1}{2}+it\right)F_{k_1}, G_{k_2}\right\rangle dt.$$

Now, using Lemma 5.3.1 and the fast decay of $\Psi(s)$ given by (5.14), we see that

$$\left|\langle E_{k_2-k_1}(z|\psi)F_{k_1}, G_{k_2}\rangle - \delta_{f=g}\frac{3}{\pi}\Psi(-1)\right| \ll_\epsilon (\log k_2)^{-\frac{1}{12}+\epsilon}(1+k_2-k_1)^{1/2}.$$

Hence the conclusion follows if

$$k_2 - k_1 \leq \log k_2^{1/6-\epsilon}. \tag{5.42}$$

Case (ii): Suppose $L(1, \mathrm{sym}^2 f)L(1, \mathrm{sym}^2 g) \leq (\log k_2)^{-5/6}$. Then we deduce the previous Lemma that $M_{k_1,k_2}(f,g) \ll_\epsilon (\log k_2)^{-\frac{1}{24}+\epsilon}$. The conclusion follows from Theorem 5.5.1 as long as $k_2 - k_1 \leq c \log\log k_2$, for some constant c. If we optimise out choices, we can let any $c < \frac{1}{12\log 2} \asymp 0.12$.

5.6.2 Proof of Theorem 5.1.1

It suffices to to bound $\left\langle u_j R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle$ and $\left\langle E(z|\psi) R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle$.

We begin with the cusp form case. From Lemma 5.4.2, $\left\langle u_j R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle$ is small when $k_2 - k_1 \geq N_\epsilon$, for some N_ϵ large enough depending only ϵ. When $k_2 - k_1 \leq N_\epsilon$, we just combine Lemma 5.4.2 and Lemma 5.5.1 depending on the size of $L(1, \mathrm{sym}^2 f)L(1, \mathrm{sym}^2 g)$, as in the previous proof.

For the Eisenstein case, we use that

$$\left\langle E(z|\psi) R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle - \delta_{f=g} \frac{3}{\pi} \Psi(-1) = \int_{-\infty}^{\infty} \Psi\left(-\frac{1}{2} - it\right) \left\langle E\left(\cdot, \frac{1}{2} + it\right) R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle dt.$$

If $k_2 - k_1 \leq N_\epsilon$, the conclusion follows again easily from Lemma 5.3.1 and the bound for $\Psi(s)$ on vertical lines given by (5.14) and from Lemma 5.5.1.

If $k_2 - k_1$ goes to infinity, we need to obtain a bound for $\Psi(s)s(s+1)\ldots(s+n-1)$ in terms of n. By repeated partial integration, this boils down to estimating $\|\psi^{(n)}\|_\infty$. One problem is that these derivatives can grow arbitrarily fast in terms of n. We show that we can work with an approximation ψ_ϵ of ψ such that $\left\langle E(z|\psi_\epsilon) R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle$ is very close to $\left\langle E(z|\psi) R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle$ and such that we can control $\|\psi_\epsilon^{(n)}\|_\infty$.

We need to construct a nontrivial function of compact support ϕ for which we control the sizes of derivatives $\|\phi^{(n)}\|_\infty$, for all n. From Denjoy–Carleman Theorem [86, p. 380], we deduce that, for any $\delta > 0$, there exists $\phi \in C^\infty(\mathbb{R})$ supported on $[-1, 1]$ such that $\int_{\mathbb{R}} \phi(x)dx = 1$ and $\|\phi^{(n)}\| \ll_\delta n^{(1+\delta)n}$, for all n. From now on we consider δ fixed (we will choose it later).

For all $\epsilon > 0$, we define $\phi_\epsilon(x) = \frac{1}{\epsilon}\phi\left(\frac{x}{\epsilon}\right)$. Then clearly ϕ_ϵ is supported on $[-\epsilon, \epsilon]$ and

$\int_{\mathbb{R}} \phi_\epsilon(x)dx = 1$. Now let any $\psi \in C_b(0,\infty)$. We consider the convolution

$$\psi_\epsilon(x) := (\psi * \phi_\epsilon)(x) = \int_{\mathbb{R}} \psi(y)\phi_\epsilon(x-y)dy,$$

which is clearly compactly supported in $(0,\infty)$, for ϵ small enough. It is not hard to see that

$$\|\psi - \psi_\epsilon\|_\infty \leq \epsilon \|\psi'\|_\infty.$$

Hence, for any $u, v \in \mathcal{L}_k(X)$ such that $\|u\|_2^2 = \|v\|_2^2 = 1$, we have

$$\begin{aligned}
|\langle E(z|\psi)u, v\rangle - \langle E(z|\psi_\epsilon)u, v\rangle| &= \left| \int_X (E(z|\psi) - E(z|\psi_\epsilon))u\overline{v}d\mu(z) \right| \\
&= \left| \int_0^\infty \int_0^1 (\psi(y) - \psi_\epsilon(y))u(x)\overline{v(z)}\frac{dxdy}{y^2} \right| \\
&\ll_\psi \epsilon,
\end{aligned}$$

since there are $O_\psi(1)$ copies of the fundamental domain for which $\psi(y) - \psi_\epsilon(y) \neq 0$ and

$$\int_{\Gamma\backslash\mathbb{H}} |u\overline{v}|d\mu(z) \leq \int_{\Gamma\backslash\mathbb{H}} \frac{|u|^2 + |v|^2}{2} d\mu(z) = 1.$$

This shows it is enough to consider $\left\langle E(z|\psi_\epsilon)R_{k_1}^{k_2}F_{k_1}, G_{k_2}\right\rangle$. Clearly,

$$\left\langle E(z|\psi_\epsilon)R_{k_1}^{k_2}F_{k_1}, G_{k_2}\right\rangle = \delta_{f=g}\frac{3}{\pi}\Psi_\epsilon(-1) + \int_{-\infty}^{\infty} \Psi_\epsilon\left(-\frac{1}{2} - it\right)\left\langle E\left(\cdot, \frac{1}{2} + it\right) R_{k_1}^{k_2}F_{k_1}, G_{k_2}\right\rangle dt,$$

and we have that

$$\Psi_\epsilon(-1) = \Psi(-1) + O_\psi(\epsilon).$$

From definition of ψ_ϵ, we have $\|\psi_\epsilon^{(k)}\|_\infty \ll_\psi \dfrac{k^{(1+\delta)k}}{\epsilon^k}$. Denote by Ψ_ϵ the Mellin transform of ψ_ϵ. From repeated partial integration, we see that for $|\sigma| \leq 2$, where $s = \sigma + it$, we have

$$\Psi_\epsilon(s)s(s+1)\dots(s+k-1) \ll_\psi \left(\frac{kC_\psi}{\epsilon}\right)^k k^{\delta k},$$

for some constant C_ψ depending on the support of ψ.

For simplicity of notation, let $\alpha = \frac{k_2 - k_1}{2}$. Choose $\epsilon = \alpha^{-\delta/2}$. We apply Lemma 5.3.1

and choose $k = \alpha + 3$. We have

$$A(f, g, \psi_\epsilon) := \left| \int_{-\infty}^{\infty} \Psi_\epsilon \left(-\frac{1}{2} - it \right) \left\langle E\left(\cdot, \frac{1}{2} + it \right) R_{k_1}^{k_2} F_{k_1}, G_{k_2} \right\rangle dt \right|$$
$$\ll \frac{\Gamma\left(\frac{k_1+k_2}{2}\right)^{1/2}}{\Gamma(k_2)^{1/2}\Gamma(\alpha)^{1/2}} k_2^{1/2} \left(C_\psi \alpha\right)^{\left(1+\frac{3\delta}{2}\right)\alpha}.$$

Hence

$$\log A(f, g, \psi_\epsilon) \leq \left(\frac{1}{2} + \frac{3\delta}{2} \right) \alpha \log \alpha - \frac{\alpha}{2} \log k_2 + O(\log k_2 + \alpha).$$

The conclusion follows if we pick $\delta = 1/12$.

Bibliography

[1] L. 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] V. Baladi and B. Vallée. Euclidean algorithms are Gaussian. *J. Number Theory*, 110(2):331–386, 2005.

[3] H. Bauer. *Probability theory*, volume 23 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1996. Translated from the fourth (1991) German edition by Robert B. Burckel and revised by the author.

[4] N. Bergeron and A. Venkatesh. The asymptotic growth of torsion homology for arithmetic groups. *J. Inst. Math. Jussieu*, 12(2):391–447, 2013.

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

[6] V. Blomer, É. Fouvry, E. Kowalski, P. Michel, D. Milićević, and W. Sawin. The second moment theory of families of L-functions. *arXiv e-prints*, page arXiv:1804.01450, April 2018.

[7] J. Bourgain. Decoupling, exponential sums and the Riemann zeta function. *J. Amer. Math. Soc.*, 30(1):205–224, 2017.

[8] S. Brooks. Eisenstein quasimodes and QUE. *Ann. Henri Poincaré*, 17(3):615–643, 2016.

[9] S. Brooks and E. Lindenstrauss. Joint quasimodes, positive entropy, and quantum unique ergodicity. *Invent. Math.*, 198(1):219–259, 2014.

[10] R. W. Bruggeman. On the distribution of Dedekind sums. In *The Rademacher legacy to mathematics (University Park, PA, 1992)*, volume 166 of *Contemp. Math.*, pages 197–210. Amer. Math. Soc., Providence, RI, 1994.

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

[12] F. Calegari and M. Emerton. Bounds for multiplicities of unitary representations of cohomological type in spaces of cusp forms. *Ann. of Math. (2)*, 170(3):1437–1446, 2009.

[13] Y. Cheng. Special value formula for the twisted triple product L-function and an application to the restricted L^2-norm problem. *Forum Math.*, 33(1):59–108, 2021.

[14] P. Cohen and P. Sarnak. Selberg trace formula, chapters 6 and 7. `http://publications.ias.edu/sarnak/section/513`, 1980.

[15] Y. Colin de Verdière. Pseudo-laplaciens. II. *Ann. Inst. Fourier (Grenoble)*, 33(2):87–113, 1983.

[16] Y. Colin de Verdière. Ergodicité et fonctions propres du laplacien. *Comm. Math. Phys.*, 102(3):497–502, 1985.

[17] P. Constantinescu. Dissipation of correlatations of automorphic forms. *in preparation.*

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

[19] P. Constantinescu and A. C. Nordentoft. Residual equidistribution of modular symbols and cohomology classes for quotients of hyperbolic n-space. *arXiv e-prints*, page arXiv:2010.12403, October 2020.

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

[21] J. E. Cremona and E. Whitley. Periods of cusp forms and elliptic curves over imaginary quadratic fields. *Math. Comp.*, 62(205):407–429, 1994.

[22] J.E. Cremona. *Modular Symbols*. PhD thesis, Oxford University, 1981. unpublished thesis.

[23] J.E. Cremona. Hyperbolic tessellations, modular symbols, and elliptic curves over complex quadratic fields. *Compositio Math.*, 51(3):275–324, 1984.

[24] H. Davenport. *Multiplicative number theory*, volume 74 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, second edition, 1980. Revised by Hugh L. Montgomery.

[25] P. Deligne. La conjecture de Weil. I. *Inst. Hautes Études Sci. Publ. Math.*, (43):273–307, 1974.

[26] N. Diamantis, J. Hoffstein, E. M. Kıral, and M. Lee. Additive twists and a conjecture by Mazur, Rubin and Stein. *J. Number Theory*, 209:1–36, 2020.

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

[28] W. Duke, J. B. Friedlander, and H. Iwaniec. The subconvexity problem for Artin L-functions. *Invent. Math.*, 149(3):489–577, 2002.

[29] P. D. T. A. Elliott. *Probabilistic Number Theory II: Central Limit Theorems*. Grundlehren der mathematischen Wissenschaften 240. Springer-Verlag New York, 1 edition, 1980.

[30] J. Elstrodt, F. Grunewald, and J. Mennicke. Vahlen's group of Clifford matrices and spin-groups. *Math. Z.*, 196(3):369–390, 1987.

[31] 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.

[32] 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.

[33] C. L. Epstein. Asymptotics for closed geodesics in a homology class, the finite volume case. *Duke Math. J.*, 55(4):717–757, 1987.

[34] K. Ford. Vinogradov's integral and bounds for the Riemann zeta function. *Proc. London Math. Soc. (3)*, 85(3):565–633, 2002.

[35] R. Gangolli and G. Warner. Zeta functions of Selberg's type for some noncompact quotients of symmetric spaces of rank one. *Nagoya Math. J.*, 78:1–44, 1980.

[36] K. Girstmair. On the distribution of Dedekind sums. *Surv. Math. Appl.*, 13:251–263, 2018.

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

[38] D. 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.

[39] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products.* Elsevier/Academic Press, Amsterdam, seventh edition, 2007. Translated from the Russian, Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger, With one CD-ROM (Windows, Macintosh and UNIX).

[40] 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.

[41] D. A. Hejhal. *The Selberg trace formula for* $\mathrm{PSL}(2,\mathbf{R})$*. Vol. 2*, volume 1001 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1983.

[42] J. Hoffstein and P. Lockhart. Coefficients of Maass forms and the Siegel zero. *Ann. of Math. (2)*, 140(1):161–181, 1994. With an appendix by Dorian Goldfeld, Hoffstein and Daniel Lieman.

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

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

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

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

[47] H. 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.

[48] H. Iwaniec and E. Kowalski. *Analytic number theory*, volume 53 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[49] D. Jakobson. Quantum unique ergodicity for Eisenstein series on $\mathrm{PSL}_2(\mathbf{Z})\backslash\mathrm{PSL}_2(\mathbf{R})$. *Ann. Inst. Fourier (Grenoble)*, 44(5):1477–1504, 1994.

[50] D. Jakobson. Equidistribution of cusp forms on $\mathrm{PSL}_2(\mathbf{Z})\backslash\mathrm{PSL}_2(\mathbf{R})$. *Ann. Inst. Fourier (Grenoble)*, 47(3):967–984, 1997.

[51] J. Jorgenson and C. O'Sullivan. Unipotent vector bundles and higher-order non-holomorphic Eisenstein series. *J. Théor. Nombres Bordeaux*, 20(1):131–163, 2008.

[52] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin-New York, second edition, 1976. Grundlehren der Mathematischen Wissenschaften, Band 132.

[53] S. Katok. *Fuchsian groups*. Chicago Lectures in Mathematics. University of Chicago Press, Chicago, IL, 1992.

[54] 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.

[55] H. Kim and P. Sarnak. Refined estimates towards the Ramanujan and Selberg conjectures. *J. Amer. Math. Soc.*, 16(1):175–181, 2003.

[56] S. Kionke and J. Schwermer. On the growth of the first Betti number of arithmetic hyperbolic 3-manifolds. *Groups Geom. Dyn.*, 9(2):531–565, 2015.

[57] J. Lee and H-S. Sun. Dynamics of continued fractions and distribution of modular symbols. *arXiv e-prints*, page arXiv:1902.06277, Feb 2019.

[58] X. Li. Bounds for $\mathrm{GL}(3)\times\mathrm{GL}(2)$ L-functions and $\mathrm{GL}(3)$ L-functions. *Ann. of Math. (2)*, 173(1):301–336, 2011.

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

[60] A. Lubotzky. Free quotients and the first Betti number of some hyperbolic manifolds. *Transform. Groups*, 1(1-2):71–82, 1996.

[61] W. Z. Luo and P. 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.

[62] J. I. Manin. Parabolic points and zeta functions of modular curves. *Izv. Akad. Nauk SSSR Ser. Mat.*, 36:19–66, 1972.

[63] 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.

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

[65] B. Mazur and K. Rubin. Relatively few rational points. Number Theory Seminar at Cal Tech. 2016.

[66] B. Mazur and K. Rubin. Arithmetic conjectures suggested by the statistical behavior of modular symbols. *arXiv e-prints, To appear in Experimental Mathematics*, page arXiv:1910.12798, October 2019.

[67] B. Mazur, J. Tate, and J. Teitelbaum. On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer. *Inventiones mathematicae*, 84:1–48, 1986.

[68] L. Merel. L'accouplement de Weil entre le sous-groupe de Shimura et le sous-groupe cuspidal de $J_0(p)$. *J. Reine Angew. Math.*, 477:71–115, 1996.

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

[70] J. J. Millson. On the first Betti number of a constant negatively curved manifold. *Ann. of Math. (2)*, 104(2):235–247, 1976.

[71] R. Munshi. The circle method and bounds for L-functions—III: t-aspect subconvexity for $GL(3)$ L-functions. *J. Amer. Math. Soc.*, 28(4):913–938, 2015.

[72] P. D. Nelson. Equidistribution of cusp forms in the level aspect. *Duke Math. J.*, 160(3):467–501, 2011.

[73] P. D. Nelson. Quadratic Hecke Sums and Mass Equidistribution. *International Mathematics Research Notices*, 05 2021. rnab093.

[74] A. C. Nordentoft. Central values of additive twists of cuspidal L-functions. *J. Reine Angew. Math.*, 776:255–293, 2021.

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

[76] C. O'Sullivan. Formulas for non-holomorphic Eisenstein series and for the Riemann zeta function at odd integers. *Res. Number Theory*, 4(3):Paper No. 36, 38, 2018.

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

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

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

[80] Y.N. Petridis. Spectral deformations and Eisenstein series associated with modular symbols. *Int. Math. Res. Not.*, (19):991–1006, 2002.

[81] I. Petrow and M. P. Young. The Weyl bound for Dirichlet L-functions of cube-free conductor. *Ann. of Math. (2)*, 192(2):437–486, 2020.

[82] R. Phillips and P. Sarnak. Geodesics in homology classes. *Duke Math. J.*, 55(2):287–297, 1987.

[83] R. Phillips and P. Sarnak. The spectrum of Fermat curves. *Geom. Funct. Anal.*, 1(1):80–146, 1991.

[84] M. Radziwiłł and K. Soundararajan. Selberg's central limit theorem for $\log|\zeta(1/2+it)|$. *Enseign. Math.*, 63(1-2):1–19, 2017.

[85] J. Rohlfs. On the cuspidal cohomology of the Bianchi modular groups. *Math. Z.*, 188(2):253–269, 1985.

[86] W. Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

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

[88] P. Sarnak. The arithmetic and geometry of some hyperbolic three-manifolds. *Acta Math.*, 151(3-4):253–295, 1983.

[89] 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.

[90] P. Sarnak. *Some applications of modular forms*, volume 99 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1990.

[91] P. Sarnak. Recent progress on the quantum unique ergodicity conjecture. *Bull. Amer. Math. Soc. (N.S.)*, 48(2):211–228, 2011.

[92] T. Schneider. Arithmetische Untersuchungen elliptischer Integrale. *Math. Ann.*, 113(1):1–13, 1937.

[93] P. Scholze. On torsion in the cohomology of locally symmetric varieties. *Ann. of Math. (2)*, 182(3):945–1066, 2015.

[94] A. Selberg. *Collected papers. I.* Springer Collected Works in Mathematics. Springer, Heidelberg, 2014. With a foreword by K. Chandrasekharan, Reprint of the 1989 edition [MR1117906].

[95] M. H. Şengün. On the integral cohomology of Bianchi groups. *Exp. Math.*, 20(4):487–505, 2011.

[96] M. H. Şengün. Arithmetic aspects of Bianchi groups. In *Computations with modular forms*, volume 6 of *Contrib. Math. Comput. Sci.*, pages 279–315. Springer, Cham, 2014.

[97] M. H. Şengün and S. Türkelli. Lower bounds on the dimension of the cohomology of Bianchi groups via Sczech cocycles. *J. Théor. Nombres Bordeaux*, 28(1):237–260, 2016.

[98] G. 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.

[99] T. Shintani. On construction of holomorphic cusp forms of half integral weight. *Nagoya Math. J.*, 58:83–126, 1975.

[100] A. Södergren. On the uniform equidistribution of closed horospheres in hyperbolic manifolds. *Proc. Lond. Math. Soc. (3)*, 105(2):225–280, 2012.

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

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

[103] W. Stein. *Modular forms, a computational approach*, volume 79 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2007. With an appendix by Paul E. Gunnells.

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

[105] G. Tenenbaum. *Introduction to analytic and probabilistic number theory*, volume 163 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, third edition, 2015. Translated from the 2008 French edition by Patrick D. F. Ion.

[106] E. C. Titchmarsh. *The theory of the Riemann zeta-function.* The Clarendon Press, Oxford University Press, New York, second edition, 1986. Edited and with a preface by D. R. Heath-Brown.

[107] K.-M. Tsang. *The distribution of the values of the Riemann zeta-function.* ProQuest LLC, Ann Arbor, MI, 1984. Thesis (Ph.D.)–Princeton University.

[108] A. I. Šnirelman. Ergodic properties of eigenfunctions. *Uspehi Mat. Nauk*, 29(6(180)):181–182, 1974.

[109] P. Wake and C. Wang-Erickson. The rank of Mazur's Eisenstein ideal. *Duke Math. J.*, 169(1):31–115, 2020.

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

[111] G. 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.

[112] M. Woodbury. *On the Triple Product Formula: Real Local Calculations*, pages 275–298. 01 2017.

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

[114] S. Zelditch. Mean Lindelöf hypothesis and equidistribution of cusp forms and Eisenstein series. *J. Funct. Anal.*, 97(1):1–49, 1991.

[115] E. I. Zelenov. p-adic Gaussian random variables. *Tr. Mat. Inst. Steklova*, 306(Matematicheskaya Fisika i Prilozheniya):131–138, 2019.