The Analytic Theory of Automorphic Forms

Contents

1 Introduction

The twentieth century has shown that modular forms and their generalisations, automorphic forms, represent one of the most powerful theories in number theory. In the simplest cases, these are a priori analytic or geometric objects, but they encode vast amounts of information about the integers and interesting arithmetic quantities. Many methods were developed for investigating automorphic forms and extracting this information, inspired by analysis, algebraic geometry, representation theory, etc. This essay presents the starting point of the analytic theory of automorphic forms.

A modular form for $\mathrm{SL}_2(\mathbb{Z})$ is a holomorphic function on the complex upper half plane, which behaves well under the action of $\mathrm{SL}_2(\mathbb{Z})$ by Möbius transformations. As presented in the next section, the upper half plane is isomorphic to the quotient of $\mathrm{SL}_2(\mathbb{R})$ by its compact subgroup $SO(2)$. Under this identification, the action of $\mathrm{SL}_2(\mathbb{Z})$ is simply multiplication on the left. For more versatility in applications and possibilities of generalisation, it is useful to "lift" a modular form to a function on the full group $\mathrm{SL}_2(\mathbb{R})$ that is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$ (see [Bor08, §5.14]). We will therefore analyse functions defined on the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$, which are called automorphic functions.

A very fruitful way of understanding functions defined on a group is to consider the group's regular representation. More precisely, there is a Haar measure on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ which defines an L^2 space of functions. The group $\mathrm{SL}_2(\mathbb{R})$ acts on this Hilbert space by the right regular representation ρ, explicitly

$$\rho_g f(x) = f(xg),$$

for $g \in \mathrm{SL}_2(\mathbb{R})$ and f a square-integrable function. The main goal is to provide a "spectral" decomposition of this representation, which is almost a decomposition into irreducible representations. To understand why that might be useful, let us consider a similar case described by Fourier analysis.

We can compare our situation with the additive group $\mathbb{R}$ of real numbers. Its quotient $\mathbb{R}/\mathbb{Z}$ by the discrete subgroup $\mathbb{Z}$ is isomorphic to the circle and thus compact. The right regular representation of $\mathbb{R}$ on $L^2(\mathbb{R}/\mathbb{Z})$ decomposes into a direct sum of irreducible representations by the Fourier series expansion. More precisely, for each $n \in \mathbb{Z}$, the subspace generated by $x \mapsto e^{2\pi i n x}$ is $\mathbb{R}$-invariant, and these generate the whole Hilbert space. Discrete decompositions are typical for compact quotients. Yet, returning to our original situation, the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ is not compact, and this introduces a "continuous" part in its decomposition.

A continuous decomposition is similar to the Fourier analysis of the whole group $\mathbb{R}$. The decomposition here is given by Fourier inversion or by the Plancherel theorem for the Fourier transform. Fourier inversion identifies a Schwarz function on $\mathbb{R}$ with the integral (morally speaking, a continuous sum) of functions $t \mapsto e^{2\pi i \xi t}$ over $\xi \in \mathbb{R}$, with coefficients given by the Fourier transform $\hat{f}(\xi)$. This is not a decomposition into invariant subspaces, since the functions $e^{2\pi i \xi t}$ are not square-integrable. But the one-dimensional spaces they generate are invariant under the regular representation. Thus, by analogy, this is a continuous decomposition of the regular representation of $\mathbb{R}$.

In our case, the regular representation of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$ decomposes into a discrete part, resembling Fourier series, and a continuous part, resembling Fourier inversion. The discrete part is made up of the space of cusp forms and the constant functions, while the continuous part is the space generated by incomplete theta series. In analogy with Fourier inversion, the functions $e^{2\pi i \xi t}$ correspond in our continuous decomposition to Eisenstein series (generalisations of the classical ones in modular forms). One should remark that the generalised Eisenstein series have many applications beyond this decomposition, for example in Rankin-Selberg theory.

A treatment of the applications of the theory developed here would go far beyond the scope of this essay. Nevertheless, it is important to mention the Selberg trace formula. This has a significant theoretical, but also historical relevance, being the ultimate goal that Selberg wanted to achieve while working on the decomposition described above. This trace formula describes the trace of certain operators on $L^2(\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R}))$ in two different ways (it has a "spectral side" and a "geometric side") and can be regarded as a generalisation of the Poisson summation formula in the Fourier analysis analogy. It has applications in geometric and arithmetic problems, for example, proving an asymptotic law for the length of closed geodesics on hyperbolic Riemann surfaces (see [Iwa02, §10.9]).

The material presented in this essay is classical, and there are many study resources for this subject. In the presentation, we follow Lang [Lan75] and Godement [God66] quite closely[1], but use some ideas and notation from [Kub73], [Bor08] and the online notes of Paul Garrett, as well.

2 The basics of the group G and some representations

This section introduces the basic objects and tools of our investigation. Let G be the special linear group $\mathrm{SL}_2(\mathbb{R})$.

2.1 The action on the upper half plane and group decompositions

A good starting point for understanding the group $\mathrm{SL}_2(\mathbb{R})$ is the study of its action on the upper half plane $\mathbb{H} := \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$. This action is given by Möbius transformations, that is,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

This action is transitive, as can be seen from the orbit of the imaginary root $i \in \mathbb{H}$. Indeed, for any $x + iy \in \mathbb{H}$ we have

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{pmatrix} \cdot i = x + iy.$$

One can easily compute that the stabiliser of i is the orthogonal group $SO(2) := K$, so that we have an identification $G/K \cong \mathbb{H}$ of topological spaces. These remarks imply that G has a decomposition as $G = NAK$, where

$$\begin{aligned} N &= \left\{ n(x) = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} : x \in \mathbb{R} \right\}, \\ A &= \left\{ a(y) = \begin{pmatrix} \sqrt{y} & 0 \\ 0 & 1/\sqrt{y} \end{pmatrix} : y \in \mathbb{R}_{>0} \right\}, \\ K &= \left\{ k(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} : \theta \in \mathbb{R} \right\}. \end{aligned}$$

This is called the *Iwasawa decomposition* and it is unique, as can be checked by considering the action on $\mathbb{H}$ again. Note that G also acts on the boundary $\partial\mathbb{H} = \mathbb{R} \cup \{\infty\}$ and the stabiliser of ∞ is N.

[1] Lang also follows Godement, which is singular in the bibliography of this essay because he investigates the left rather than the right regular representation. We adopted the more common "left winger" approach that Lang advocates for in [Lan75, Notation]), letting $\mathrm{SL}_2(\mathbb{Z})$ act on $\mathrm{SL}_2(\mathbb{R})$ on the left.

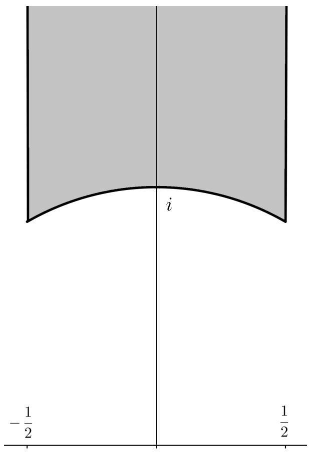

Figure 1: A fundamental domain for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$

Since $\mathrm{SL}_2(\mathbb{R})$ is a Lie group, it is naturally endowed with a (left) Haar measure, i.e. a non-trivial measure on G that is invariant under left translation. Using the Iwasawa decomposition, it is given by

$$dg = \frac{1}{y^2} dx \cdot dy \cdot d\theta,$$

and we will at times denote it by μ. In fact, the group G is unimodular, i.e., the Haar measure above is right-invariant as well, since it is a semisimple Lie group (or simply by checking).

When defining the principal series in section 2.3, we shall consider the induced Haar measure on the subgroup $P := NA$. We denote the measure simply by dp, but note that this is not right-invariant. Indeed, the modular function of P is $\Delta_P(n(x)a(y)) = y^{-1}$ (cf. [Rob83, p. 173]).

We now fix a discrete subgroup Γ of G. For simplicity, we shall always assume $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, although much of the material in this essay can be generalised to other subgroups (e.g. congruence subgroups) as well. See the concluding remarks for a brief discussion on this topic. Since the Haar measure on G is invariant (in particular under Γ), it descends to the quotient[2] $\Gamma\backslash G$. It is important to note that $\mu(\Gamma\backslash G)$ is finite. Indeed, it is a standard fact (see any book on elliptic functions) that a fundamental domain for the action of Γ on G is

$$\left\{ n(x)a(y) \mid x \in [-1/2, 1/2], y \geq \sqrt{1-x^2} \right\} \cdot K,$$

as depicted in Figure 1. Computing the measure using the formula above gives

$$\mu(\Gamma\backslash G) = \frac{\pi}{3}.$$

We denote by Γ_∞ the stabiliser of the point at infinity in the action of Γ on the upper half plane, i.e. $\Gamma_\infty = N \cap \Gamma$.

[2] This quotient is not a group, but merely the set of cosets.

Another useful type of decomposition is a kind of *Bruhat decomposition* of the group Γ (see [Kub73, p. 15]). We consider the double-cosets $\Gamma_\infty\backslash\Gamma/\Gamma_\infty$.

Lemma 2.1. *We have the disjoint union decomposition*

$$\Gamma = \pm\Gamma_\infty \cup \bigcup_{c\in\mathbb{Z}\setminus\{0\}} \bigcup_{\substack{d (mod\ c)\\(c,d)=1}} \Gamma_\infty \begin{pmatrix} * & * \\ c & d \end{pmatrix} \Gamma_\infty.$$

Proof. Note the identity

$$\begin{pmatrix} 1 & k_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & k_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a+ck_1 & * \\ c & d+ck_2 \end{pmatrix}.$$

Hence, when determining a coset, we can choose c freely and d modulo c. These choices completely determine the coset. Indeed, if c and d are fixed as above, the upper left entry must only be chosen modulo c, but the additional condition $ad \equiv 1 \pmod{c}$ determines it automatically. Finally, the upper right entry $*$ is fixed by the other entries and the determinant condition. □

2.2 Automorphic functions

An *automorphic function* is a map from the quotient $\Gamma\backslash G$ into the complex numbers. We may think of automorphic functions as functions on G that are invariant under Γ. We can define an inner product for two automorphic functions f_1 and f_2 by

$$(f_1, f_2) = \int_{\Gamma\backslash G} f_1(g)\overline{f_2(g)}dg.$$

We denote by $\mathcal{H} = L^2(\Gamma\backslash G)$ the space of square integrable automorphic functions. The group G acts on this Hilbert space by the right regular representation

$$\rho : G \longrightarrow B(\mathcal{H}), \quad \rho_g f(x) = f(xg).$$

This is a unitary representation, as is seen by a simple change of coordinates, recalling that G is unimodular.

The point ∞ is called a cusp, and it is the only cusp (up to equivalence) for the group $\mathrm{SL}_2(\mathbb{Z})$ (see [Kub73, p. 3] for definitions). Its stabiliser is Γ_∞, by definition. Recall that classical modular forms have Fourier expansions at cusps, and we can similarly define Fourier coefficients for automorphic functions (although a series expansion is in general not feasible, since we have no analytic requirements for our functions). For $g \in G$, we can induce a function on $\mathbb{H}$ by $x+iy \mapsto f(n(x)a(y)g)$, which is invariant under $x \mapsto x+1$ by Γ_∞-invariance. The usual definitions for Fourier coefficients now apply. In particular, the zero-th Fourier coefficient is defined by

$$f^0(g) = \int_0^1 f(n(x)g)dx.$$

Note that f^0 is left N-invariant (we are essentially integrating over the circle and a shift does not change the integral).

We call an automorphic function f a *cusp form* if $f^0 \equiv 0$. Note that the right regular representation ρ preserves the space of cusp forms.

2.3 The principal series

The main goal of this essay is to understand the right-regular representation ρ of G on its quotient $\Gamma\backslash G$. This representation is partly made up of a collection of irreducible representations called the principal series. This section provides an explicit construction of these representations, following [Kub73, §7.1] in definitions and [Rob83, §18] in proofs.

Fix a complex number s. Let us consider even functions $\eta : \mathbb{C}^\times \longrightarrow \mathbb{C}$ such that $\eta(re^{i\theta}) = r^{2s-2}\eta(e^{i\theta})$, for any $r > 0$. For two such functions, η_1 and η_2, we define the inner product

$$\langle \eta_1, \eta_2 \rangle = \int_0^{2\pi} \eta_1(e^{i\theta})\overline{\eta_2(e^{i\theta})}d\theta.$$

The Hilbert space of functions η as above that have finite norm with respect to this inner product is denoted $\mathcal{P}(s)$.

We now want to view these functions as functions of $g = n(x)a(y)k(\theta) \in G$. For this, note that $N\backslash G \cong \mathbb{C}^\times$ by the map $n(x)a(y)k(\theta) \mapsto y^{-1/2}e^{i\theta}$, by considering the second row of the matrix $a(y)k(\theta)$. Then η must satisfy

$$\eta(g) = y^{1-s}\eta(k(\theta)), \quad \eta(-g) = \eta(g). \tag{2.1}$$

The first condition is equivalent to $\eta(n(x)a(y)g) = y^{1-s}\eta(g)$.

We will prove that if $\operatorname{Re} s = 1/2$, then ρ is a unitary representation of G on the space $\mathcal{P}(s)$. It suffices to show that $\left\|\rho_h(\eta)\right\| = \left\|\eta\right\|$, for all $h \in \mathrm{SL}_2(\mathbb{R})$.

Using the Iwasawa decomposition we write

$$k(\theta)h = h(\theta)k'(\theta),$$

where $h(\theta) = n(x_\theta)a(y_\theta) \in P$ and $k'(\theta) \in K$. Since η is N-invariant from the left and $|\eta(a(y)k(\theta))|^2 = y^{2-2\operatorname{Re}s}|\eta(\theta)|^2$, we have that

$$\int_0^{2\pi} |\eta(k(\theta)h)|^2 d\theta = \int_0^{2\pi} y_\theta |\eta(k'(\theta))|^2 d\theta$$

for $\operatorname{Re} s = 1/2$. We thus want to show $\int y_\theta|\eta(k'(\theta)|^2 d\theta = \int |\eta(k(\theta)|^2 d\theta$. For this, we extend the function to the whole group G and then use its unimodularity property. Let $f(pk) = \xi(p)|\eta(k)|^2$ be a function on G, where ξ is a function on $P = NA$ with compact support and normalised so that $\int_P \xi(p)dp = 1$. Thus,

$$\begin{aligned}\int_0^{2\pi} |\eta(k(\theta))|^2 d\theta &= \int_G f(g)dg = \int_G f(gh)dg \\ &= \int_P \int_0^{2\pi} f(pk(\theta)h)d\theta dp \\ &= \int_P \int_0^{2\pi} f(ph(\theta)k'(\theta)d\theta dp \\ &= \int_0^{2\pi} \int_P f(ph(\theta)k'(\theta))dpd\theta \\ &= \int_0^{2\pi} \int_P y_\theta f(pk'(\theta))dpd\theta \\ &= \int_0^{2\pi} y_\theta |\eta(k'(\theta))|^2 d\theta,\end{aligned}$$

since the modular function of P is given by $\Delta_P(n(x)a(y)) = y^{-1}$.

This collection of unitary representations parametrised by complex numbers s with $\operatorname{Re} s = 1/2$ is called the *principal series*. They are irreducible for $s \neq 1/2$, as shown in [Rob83, Theorem 18.4].

3 The continuous part

We have defined the Hilbert space $\mathcal{H}$ of automorphic functions in the last section, yet we have not introduced any explicit example of such a function. We start this section by defining incomplete theta series, thus generating a subspace of $\mathcal{H}$ that provides the so-called continuous part in the decomposition of the regular representation. This will be made explicit by mapping this subspace into the principal series and proving a Plancherel theorem for this map, effectively viewing the principal series as part of the regular representation.

3.1 Incomplete theta series

As usual, the most natural way of obtaining a function on the quotient $\Gamma\backslash G$ is by averaging over Γ. As in the case of classical Eisenstein series, to ensure convergence we pick a function that is already invariant under the infinite subgroup Γ_∞ and then average only over the cosets $\Gamma_\infty\backslash\Gamma$.

Definition 3.1. Let ψ be an even function on $N\backslash G$. The *incomplete theta series*[3] is de-

[3] These are sometimes called *pseudo-Eisenstein series* (see [Gar18]) or *theta transforms* in [Lan75]. The name incomplete theta series is used here because we follow [God66] and [Kub73], but pseudo-Eisenstein series seems to be the more modern choice.

fined as

$$\theta_\psi(g) = \frac{1}{2} \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \psi(\gamma g),$$

for $g \in G$.

Note that any function ψ on $N\backslash G$ can be decomposed into the sum of an even and an odd function. The incomplete theta series of an odd function would vanish, since the term corresponding to $\gamma \in \Gamma_\infty\backslash\Gamma$ would cancel with the term coming from $-\gamma$. Therefore, we only need to consider even functions. The 1/2 factor now simply balances off the $\pm\gamma$ terms and is sometimes left out in the literature (e.g [Lan75]).

We view $N\backslash G$ as $\mathbb{R}^2 \setminus \{0\}$ by letting G operate on row vectors as right matrix multiplication and noting that N is the stabiliser of the unit vector $(0,1)$.[4] We denote by $\mathcal{S}(N\backslash G)$ the space of functions on $N\backslash G$ that are restrictions of functions in the Schwartz space of $\mathbb{R}^2$.

When working with incomplete theta series, it suffices to consider functions on $N\backslash G$ with compact support and some analytic condition, e.g. smoothness, since these are dense in $L^2(\mathbb{R}^2)$. Nevertheless, the Fourier transform of these functions will enter the stage later, and these do not in general have compact support, but lie in the previously defined Schwartz space $\mathcal{S}(N\backslash G)$.

Lemma 3.2. *If $\psi \in C_c^\infty(N\backslash G)$, i.e. ψ is smooth and has compact support, then the incomplete theta series θ_ψ is locally finite and defines a smooth and compactly supported function on $\Gamma\backslash G$.*

Proof. Let C be a compact set in G, such that $N \cdot C$ contains the support of ψ. We now restrict the variable g to a fixed compact subset C_0. If a summand $\psi(\gamma g)$ in the series is non-zero, then $\gamma g \in NC$, that is,

$$\gamma \in \Gamma \cap N \cdot C \cdot C_0^{-1},$$

given that $g \in C_0$. Thus, as cosets,

$$\Gamma_\infty \gamma \in (\Gamma_\infty\backslash\Gamma) \cap (\Gamma_\infty\backslash NCC_0^{-1}).$$

Now $\Gamma_\infty\backslash\Gamma$ is discrete, and $\Gamma_\infty\backslash NCC_0^{-1}$ is compact by our assumptions and the fact that $\Gamma_\infty\backslash N$ is compact (isomorphic to the circle). Thus, the intersection contains only finitely many cosets, proving that the series θ_ψ is locally finite and therefore converging to a smooth function.

For proving that θ_ψ has compact support, observe again that $\psi(\gamma g)$ is non-zero only if $g \in \Gamma \cdot C$. In terms of cosets, $\Gamma g \subset \Gamma\backslash(\Gamma \cdot C)$, and the right hand side is the image of the compact set $C \subset G$ under the continuous map $G \longrightarrow \Gamma\backslash G$. □

The previous lemma immediately implies that θ_ψ is square-integrable, i.e. $\theta_\psi \in \mathcal{H}$ if $\psi \in C_c^\infty(N\backslash G)$. We denote the closed subspace spanned by all such incomplete theta

[4]This is similar to the identification in section 2.3, but differs by a rotation.

series by $\Theta \subset \mathcal{H}$. This space is invariant under the regular representation, as well as its decomposition into the constant functions and their orthogonal complement, as we prove in section 4.1. This orthogonal complement is spanned by incomplete theta series for functions ψ with their Fourier transform $\hat{\psi}$ vanishing at 0, which is what will be assumed in some of the following results.

Since θ_ψ is an automorphic function, we can define its zero-th Fourier coefficient θ_ψ^0, which will play an important role in the next sections. Therefore, it will be useful to have a more explicit formula for θ_ψ^0.

Lemma 3.3. *For ψ an even function on $N\backslash G$ and $w = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, we have*

$$\theta_\psi^0(g) = \psi(g) + \sum_{\substack{(c,d)=1 \\ c>0}} \frac{1}{c^2} \int_N \psi(wn(x)a(c^2)g)dx,$$

where the sum ranges over positive integers c and invertible residue classes modulo c, represented by d.

Proof. This is a reflection of the Bruhat decomposition. Indeed, recalling Lemma 2.1,

$$\begin{aligned}
\theta_\psi^0(g) &= \int_0^1 \psi(n(x)g)dx = \int_{\Gamma_\infty\backslash N} \frac{1}{2} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \psi(\gamma n(x)g) \\
&= \int_N \psi(n(x)g)dx + \sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma \\ \gamma\neq\pm\Gamma_\infty}} \int_{\Gamma_\infty\backslash N} \psi(\gamma n(x)g)dx \\
&= \psi(g) + \sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_\infty \\ \text{nontriv.}}} \int_N \psi(\gamma n(x)g)dx,
\end{aligned}$$

keeping in mind that ψ is even and N-invariant. Now if the lower left entry c of γ is positive, then $N\gamma N = -Nwa(c^2)N$, since

$$\begin{pmatrix} 1 & -\frac{a}{c} \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & -\frac{d}{c} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -c^{-1} \\ c & 0 \end{pmatrix}.$$

Then, interchanging matrices (see Remark 3.7),

$$\begin{aligned}
\int_N \psi(\gamma n(x)g)dx &= \int_N \psi(wa(c^2)n(x)g)dx \\
&= \int_N \psi(wn(xc^2)a(c^2)g)dx = \frac{1}{c^2}\int_N \psi(wn(x)a(c^2)g)dx.
\end{aligned}$$

Using again the parity of ψ and the Bruhat decomposition, Lemma 2.1, gives the desired formula. $\square$

3.2 The Mellin transform

As already noted at the beginning of the previous section, we shall construct a map from the space of incomplete theta series into the principal series. This is given by the Mellin transform of the zero-th Fourier coefficient.

Definition 3.4. For a function f on $N\backslash G$, define the *Mellin transform* of f by

$$L_f(g,s) = \int_0^\infty f(a(y)g)y^{-s}\frac{dy}{y},$$

where $g \in G$ and $s \in \mathbb{C}$ such that the integral converges. Similarly, if f is merely a function on $\mathbb{R}_{>0}$, we define[5]

$$L_f(s) = \int_0^\infty f(y)y^{-s}\frac{dy}{y}.$$

The (slightly modified) *Mellin transform*[6] $\widehat{\theta_\psi}$ of the constant term θ_ψ^0 is defined by

$$\widehat{\theta_\psi}(g,s) = \int_0^\infty \theta_\psi^0(a(y)g)y^{s-1}\frac{dy}{y},$$

for $g \in G$.

We shall first look at the Mellin transform $L_f(g,s)$ and prove some important analytic results. We can easily see that the integral in the definition of the Mellin transform converges absolutely for any $s \in \mathbb{C}$ if the corresponding function f has compact support. In other words, if $\psi \in C_0^\infty(N\backslash G)$, then $L_\psi(g,s)$ is entire. Nevertheless, as already mentioned, the Fourier transform of ψ need not have compact support and merely lies in the Schwartz space. To treat this case we make the following reduction to Mellin transforms of functions on $\mathbb{R}$. If $\psi \in \mathcal{S}(N\backslash G)$ and $g \in G$, then we may define $f(y) = \psi(a(y)g)$ so that $L_\psi(g,s) = L_f(s)$ and $f \in \mathcal{S}(\mathbb{R})$.

Lemma 3.5. *If $f \in \mathcal{S}(\mathbb{R})$, then $L_f(s)$ is a meromorphic function on $\mathbb{C}$ with possible simple poles at most at $s \in \{0,1,2,3,\&c.\}$. If $f(0) = 0$, then $L_f(s)$ is holomorphic at* 0.

Proof. First note that the integral

$$\int_1^\infty f(y)y^{-s-1}dy$$

converges absolutely for any $s \in \mathbb{C}$, since f decays faster than any polynomial at infinity, thus defining a holomorphic function. For the rest of the integral, we integrate by parts to obtain

$$\int_0^1 f(y)y^{-s-1}dy = \frac{f(y)y^{-s}}{-s}\Big|_0^1 + \frac{1}{s}\int_0^1 f'(y)y^{-s}dy.$$

[5]Changing the variable s to $-s$ would give the standard definition of the Mellin transform.
[6]This is called the *zeta transform* in [Lan75, p. 243] and the *Laplace transform* in [God66].

Since f is bounded, the integral on the right converges absolutely for $\operatorname{Re} s < 1$, whereas the integral on the left converges absolutely only for $\operatorname{Re} s < 0$. We therefore enlarged the domain of definition, and applying partial integration as above inductively, we continue $L_f(s)$ meromorphically to all $s \in \mathbb{C}$, with possible poles at 0, 1, 2, and so on.

Assuming that $f(0) = 0$, observe that

$$\int_0^1 f(y)y^{-1}dy = f(y)\log(y)\Big|_0^1 = \lim_{y\to 0} f(y)\log(y) = \lim_{y\to 0} \frac{f(y)}{y} \cdot y\log(y) = f'(0).$$

Therefore, L_f has a removable singularity at $s = 0$. $\square$

Thus, there is a correspondence between asymptotics of f and the holomorphy of L_f. In a similar manner, smoothness translates by the Mellin transform into rapid decay on vertical strips. More precisely, we say that a function g defined on a vertical strip $\sigma_0 < \operatorname{Re} s < \sigma_1$ *decreases rapidly on every vertical line, uniformly in the strip*, if for every positive integer n, there exists a positive constant C_n such that

$$|g(\sigma + it)| \le \frac{C_n}{1 + t^{2n}},$$

for each $\sigma + it$ in the given strip. If there are finitely many poles of g in the strip and the estimate holds outside open neighbourhoods of the poles, then the same terminology applies.

Lemma 3.6. *Let $f \in \mathcal{S}(\mathbb{R})$ and assume $f(0) = 0$. Then L_f decreases rapidly on every vertical line, uniformly in each strip of finite width, outside a neighbourhood of poles.*

Proof. Assume first that $\operatorname{Re} s < 0$. We have by partial integration

$$L_f(s) = \frac{f(y)y^{-s}}{-s}\Big|_0^\infty + \frac{1}{s}\int_0^\infty f'(y)y^{-s}dy = \frac{1}{s}L_{f'}(s-1),$$

by our assumptions on the decay of f. Continuing similarly and noting that all derivatives of f lie in the Schwartz space, we obtain

$$L_f(s) = \frac{(-1)^n}{s(s-1)\dots(s-n)} L_{f^{(n+1)}}(s-(n+1)).$$

If we fix the real part of s, the first factor, i.e. the fraction, behaves like $1/|\operatorname{Im} s|^{n+1}$ in absolute value, as $\operatorname{Im} s \to \infty$. The second factor is bounded on finite strips, since

$$|L_{f^{(n+1)}}(s-n-1)| \le \int_0^\infty |f^{(n+1)}(y)y^{-s+n+1-1}|dy = \int_0^\infty |f^{(n+1)}(y)|y^{-\operatorname{Re} s+n}dy,$$

which converges (since $\operatorname{Re} s < 0$) and depends only on the real part of s.

To treat the case $\operatorname{Re} s > 0$, note that, by the meromorphic continuation, the identity above holds for all s. The bound is recovered for $\operatorname{Re} s > 0$ by choosing n large enough, so that the given strip shifted by $-(n+1)$ is contained in $\operatorname{Re} s < 0$ and the bound above can be used. $\square$

We can derive more properties of the Mellin transform by noting it is simply the Fourier transform on the multiplicative group $\mathbb{R}_{>0}$. This can be easily seen by changing the variable $y = e^x$, so that

$$L_f(\sigma + it) = \int_0^\infty f(y) y^{-(\sigma+it)} \frac{dy}{y} = \int_{-\infty}^\infty f(e^x) e^{-\sigma x} e^{-itx} dx,$$

which is the Fourier transform of the function $f(e^x)e^{-\sigma x}$. The Plancherel formula for the Fourier transform produces

$$\frac{1}{2\pi} \int_{-\infty}^\infty L_{f_1}(\sigma_1 + it) \overline{L_{f_2}(\sigma_2 + it)} dt = \int_0^\infty f_1(y) y^{-\sigma_1} \overline{f_2(y)} y^{-\sigma_2} \frac{dy}{y}. \tag{3.1}$$

Remark 3.7. When working with the Mellin transform on the group, note that $L_f(nak, s) = L_f(ak, s)$, since

$$\begin{pmatrix} \sqrt{y} & \\ & 1/\sqrt{y} \end{pmatrix} \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} = \begin{pmatrix} 1 & xy \\ & 1 \end{pmatrix} \begin{pmatrix} \sqrt{y} & \\ & 1/\sqrt{y} \end{pmatrix},$$

and f is assumed to be (left) N-invariant. Noting the N-invariance of the zero-th Fourier coefficient, this argument is valid for the Mellin transform of an incomplete theta series, as well.

We shall now make explicit the connection between the incomplete theta series and the principal series through the Mellin transform.

Lemma 3.8. *If $\psi \in C_c^\infty(N\backslash G)$, then the Mellin transform $\widehat{\theta_\psi}$ converges absolutely in the region* $\mathrm{Re}\, s > 1$. *Whenever it is defined, $\widehat{\theta_\psi}(s)$ lies in the space $\mathcal{P}(s)$.*

Proof. By Lemma 3.2, the incomplete theta series θ_ψ has compact support. Since $n(x)a(y)$ escapes any compactum as y tends to infinity, it follows by definition that $\theta_\psi^0(a(y)g)$ vanishes as $y \to \infty$, say for $y > B > 0$. Then

$$\widehat{\theta_\psi}(g, s) = \int_0^B \theta_\psi^0(a(y)g) y^{s-1} \frac{dy}{y} \le \left\| \theta_\psi^0(a(y)g) \right\|_\infty \int_0^B y^{s-2} dy,$$

which is finite for $\mathrm{Re}\, s > 1$. Remark 3.7 and a simple change of variables, i.e. $y \mapsto yy'$, shows that $\widehat{\theta_\psi}(a(y')g, s) = y'^{1-s} \widehat{\theta_\psi}(g, s)$, and the parity of $\widehat{\theta_\psi}$ follows from that of ψ. Thus, $\widehat{\theta_\psi}$ lies in $\mathcal{P}(s)$. $\square$

3.3 Inner product formulae

In the last subsection we obtained a map from the incomplete theta series to the principal series given by the Mellin transform. The goal now is to prove the Plancherel theorem for this decomposition. In other words, we will prove for incomplete theta series satisfying a certain condition the formula

$$\left\| \theta_\psi \right\|_{\Gamma\backslash G}^2 = \frac{1}{4\pi i} \int_0^\infty \left\| \widehat{\theta_\psi}(*, 1/2 + it) \right\|_{\mathcal{P}(1/2+it)}^2 dt. \tag{3.2}$$

To obtain this result, we will need formulae for different inner products and we will need to analytically continue the Mellin transform $\widehat{\theta_\psi}$, which is a priori defined for $\operatorname{Re} s > 1$, to the domain $\operatorname{Re} s \geq \frac{1}{2}$.

First, we need to introduce the method of *unfolding* (or unwinding). This is a basic and essential tool in the analytic theory of automorphic forms. Let f and g be compactly supported functions[7] on $\Gamma\backslash G$. Then, Fubini's theorem implies that

$$\int_{\Gamma\backslash G} f(h) \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} g(\gamma h)dh = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\Gamma\backslash G} f(h)g(\gamma h)dh.$$

We may think of integrating over $\Gamma\backslash G$ as integrating over a fundamental domain F in G. Using the invariance of the measure and the automorphy of f, we obtain

$$\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_F f(h)g(\gamma h)dh = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\gamma^{-1}F} f(h)g(h)dh = \int_{\Gamma_\infty\backslash G} f(h)g(h)dh,$$

since piecing together the translates $\gamma^{-1}F$ over $\gamma \in \Gamma_\infty\backslash\Gamma$ gives a fundamental domain for $\Gamma_\infty\backslash G$.

Theorem 3.9. *Let ψ and ψ' be two functions on $N\backslash G$ with θ_ψ and $\theta_{\psi'}$ their associated incomplete theta series, respectively. Then*

$$(\theta_\psi, \theta_{\psi'}) = \frac{1}{4\pi i}\int_K \int_{\operatorname{Re} s=\sigma} \widehat{\theta_\psi}(k,s)\overline{L_{\psi'}(k,\overline{s})}ds d\theta,$$

for some $\sigma > 1$.

Proof. Starting with the definition and unfolding, we have

$$\begin{aligned}
(\theta_\psi, \theta_{\psi'}) &= \int_{\Gamma\backslash G} \theta_\psi(g)\overline{\theta_{\psi'}(g)}dg = \frac{1}{2}\int_{\Gamma_\infty\backslash G} \theta_\psi(g)\overline{\psi'(g)}dg \\
&= \frac{1}{2}\int_K\int_A\int_{\Gamma_\infty\backslash N} \theta_\psi(nak)\psi'(nak)\frac{dxdyd\theta}{y^2} = \frac{1}{2}\int_K\int_A \theta_\psi^0(ak)\overline{\psi'(ak)}\frac{dyd\theta}{y^2} \\
&= \frac{1}{2}\int_K\int_A \theta_\psi^0(ak)y^{\sigma-1}\cdot\overline{\psi'(ak)}y^{-\sigma}\frac{dy}{y}d\theta.
\end{aligned}$$

Using the Mellin transform Plancherel formula (3.1), and taking into consideration the slight modification of the Mellin transform of θ_ψ, we obtain

$$(\theta_\psi, \theta_{\psi'}) = \frac{1}{4\pi i}\int_K \int_{-\infty}^{\infty} \widehat{\theta_\psi}(k,s)\overline{L_{\psi'}(k,\overline{s})}dt d\theta,$$

where $t = \operatorname{Im} s$. ∎

In the former theorem, let us now switch the integrals and notice that the K-integral is now essentially the inner product of $\widehat{\theta_\psi}$ and $L_{\psi'}$ in the sense of the principal series. In view of proving the Plancherel formula (3.2), we need to shift the contour of the s-integral to $\operatorname{Re} s = \frac{1}{2}$ and then relate the Mellin transform of ψ' to $\widehat{\theta_{\psi'}}$.

[7] We can ask for weaker conditions, as long as the application of Fubini's theorem is valid.

3.4 Eisenstein series

We now start start our investigation of the relation between the Mellin transforms L_ψ and $\widehat{\theta_\psi}$. The binding link is the Eisenstein series E_ψ, which comes up naturally when considering the K-integral in Theorem 3.9. Note first that $L_{\psi'}(a(t)k,\bar{s}) = t^{\bar{s}}L_{\psi'}(k,\bar{s}) = \overline{t^s}L_{\psi'}(k,\bar{s})$. We have

$$\begin{aligned}
\int_K \widehat{\theta_\psi}(k,s)\overline{L_{\psi'}(k,\bar{s})}d\theta &= \int_K\int_A \theta_\psi^0(ak)y^{s-1}\frac{dy}{y}\overline{L_{\psi'}(k,\bar{s})}d\theta \\
&= \int_K\int_A \theta_\psi^0(ak)\overline{L_{\psi'}(ak,\bar{s})}\frac{dyd\theta}{y^2} \\
&= \int_K\int_A\int_{\Gamma_\infty\backslash N} \theta_\psi(nak)\overline{L_{\psi'}(nak,\bar{s})}\frac{dxdyd\theta}{y^2} \\
&= \int_{\Gamma_\infty\backslash G} \theta_\psi(g)\overline{L_{\psi'}(g,\bar{s})}dg \\
&= \int_{\Gamma\backslash G} \theta_\psi(g)\overline{\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} L_{\psi'}(\gamma g,\bar{s})}dg = \left(\theta_\psi, \sum_{\Gamma_\infty\backslash\Gamma} L_{\psi'}(\gamma g,\bar{s})\right),
\end{aligned}$$

by unfolding.

Definition 3.10. For a function ψ on $N\backslash G$, we define the *Eisenstein series* $E_\psi(g,s)$ by

$$E_\psi(g,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} L_\psi(\gamma g,s).$$

The computation at the beginning of this section proves the following.

Lemma 3.11. *We have*

$$\int_K \widehat{\theta_\psi}(k,s)\overline{L_{\psi'}(k,\bar{s})}d\theta = (\theta_\psi, E_\psi(\bar{s})),$$

wherever the expressions exist.

Another motivation for defining the Eisenstein series is that we can apply Mellin inversion to ψ, so that, by interchanging sum and integral, the incomplete theta series is equal to an integral over the Eisenstein series (see [God66, Section 4]). Thus, it is clear that a good understanding of Eisenstein series would give us more information about the incomplete theta series. This is precisely our strategy in the next sections.

For proving convergence of the Eisenstein series we need to note that the cosets $\Gamma_\infty\backslash\Gamma$ are parametrised by the lower row. Indeed, fixing c and d, the integer solutions to the equation $ad-bc=1$ are $\{(a+kx, b+kd) \mid k\in\mathbb{Z}\}$, for any given solution (a,b). Accordingly,

$$\begin{pmatrix}1 & k\\ 0 & 1\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}a+kc & b+kd\\ c & d\end{pmatrix}.$$

Notation 3.12. For functions ψ defined on $N\backslash G$, we use square brackets to indicate that we view ψ as a function on $\mathbb{R}^2$. That is, for $g \in G$,

$$\psi(g) = \psi[(0,1)\cdot g].$$

Lemma 3.13. *If $\psi \in \mathcal{S}(N\backslash G)$, then the Eisenstein series $E_\psi(g,s)$ converges for* $\operatorname{Re} s > 1$.

Proof. Shifting ψ by g, we may assume that $g = e$, i.e. the identity matrix. If we let $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right)$, then

$$\begin{aligned} L_\psi(\gamma,s) &= \int_0^\infty \psi[(0,1)a(y)\gamma]y^{-s}\frac{dy}{y} \\ &= \int_0^\infty \psi[c/\sqrt{y}, d/\sqrt{y}]y^{-s}\frac{dy}{y} \\ &= \int_0^\infty \psi[c\sqrt{y}, d\sqrt{y}]y^{s}\frac{dy}{y}. \end{aligned}$$

We now need to sum over all (primitive) pairs (c,d) of integers. To control this sum, we partition these pairs into annuli of integer radius and width 1, with respect to the supremum norm on $\mathbb{R}^2$.

If $\operatorname{Re} s =: \sigma$, then[8] $\psi[\xi] \ll 1/\|\xi\|_\infty^{2(\sigma+\varepsilon)}$, for any $\varepsilon > 0$, since ψ is in the Schwartz space. Let now $m \in \mathbb{Z}^2$ be in the annulus of radius n and width 1, i.e. $n \le \|m\|_\infty < n+1$. We split our integral $L_\psi(\gamma,s)$ into the sum

$$\int_0^{1/n^2} + \int_{1/n^2}^\infty .$$

The first integral can be bounded by

$$\int_0^{1/n^2} |\psi[\sqrt{y}m]|y^\sigma \frac{dy}{y} \ll \frac{y^\sigma}{\sigma}\Big|_0^{1/n^2} \ll \frac{1}{n^{2\sigma}}.$$

The second integral has the bound

$$\int_{1/n^2}^\infty |\psi[\sqrt{y}m]|y^\sigma \frac{dy}{y} \ll \int_{1/n^2}^\infty \frac{y^{\sigma-1}}{y^{\sigma+\varepsilon}\|m\|^{2\sigma+2\varepsilon}}dy \ll \frac{1}{n^{2\sigma}}.$$

There are $\ll n$ such integer points m in the annulus. Thus, their contribution to the Eisenstein series can be bounded by $n^{1-2\sigma}$. Summing over all positive integers n proves convergence in the given range. $\square$

[8]The following notation will be used throughout this essay. For two functions f and g we have $f \ll g$ if there is a constant $C > 0$ such that $|f| \le C \cdot g$.

3.5 The analytic continuation and functional equation

The most important aspect of Eisenstein series is their analytic continuation and their functional equation. In our simplified case, i.e. for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, these follow essentially from an application of Poisson summation, similar to the Riemann zeta function. For this, we need to look at the Fourier transform of ψ.

Viewing ψ as a function on $\mathbb{R}^2$, we have the usual Fourier transform

$$\hat{\psi}(z) = \int_{\mathbb{R}^2} \psi[\xi] e^{-2\pi i z\cdot\xi} d\xi.$$

In our computations we will consider the function $\xi \mapsto \psi[\xi y g]$. By a change of variables, its Fourier transform is equal to the function $z \mapsto \hat{\psi}[zy^{-1}g^{-t}] \cdot y^{-2}$, where g^{-t} is the transpose of the inverse of g. Now the Poisson summation formula implies that

$$\sum_{m\in\mathbb{Z}^2} \psi[myg] = y^{-2} \sum_{m\in\mathbb{Z}^2} \hat{\psi}[my^{-1}g^{-t}].$$

We prove the analytic continuation of the Eisenstein series under the extra condition that $\hat{\psi}[0] = 0$. This restriction has an important meaning, namely that the incomplete theta series θ_ψ is orthogonal to the constant functions in $\mathcal{H}$, as explained in section 4.1.

Theorem 3.14. *Let $\psi \in \mathcal{S}(N\backslash G)$ be an even function, such that $\psi[0] = 0$ and $\hat{\psi}[0] = 0$. Define*

$$E^*_\psi(g,s) = \zeta(2s) E_\psi(g,s),$$

where ζ is the meromorphic Riemann zeta function. Then E^* is an entire function in s, satisfying the functional equation

$$E^*_\psi(g,s) = E^*_{\hat{\psi}}(g^{-t}, 1-s).$$

Proof. As explained in the proof of Lemma 3.13, we have

$$L_\psi(\gamma g, s) = \int_0^\infty \psi[(\sqrt{y}c, \sqrt{y}d)g] y^s \frac{dy}{y},$$

for $\gamma = \left(\begin{smallmatrix} * & * \\ c & d \end{smallmatrix}\right) \in \Gamma$, and

$$E_\psi(g,s) = \sum_{m\in\mathbb{Z}^2 \text{ prim}} \int_0^\infty \psi[m\sqrt{y}g] y^s \frac{dy}{y},$$

where the sum ranges over primitive lattice points. In the domain of absolute convergence, we multiply E_ψ by $\zeta(2s) = \sum_n n^{-2s}$ to obtain

$$\begin{aligned}
\zeta(2s)E_\psi(g,s) &= \sum_{n=1}^\infty \sum_{m \text{ prim}} \int_0^\infty \psi[m\sqrt{y}g] \left(\frac{y}{n^2}\right)^s \frac{dy}{y} \\
&= \int_0^\infty \sum_{m\in\mathbb{Z}^2} \psi[m\sqrt{y}g] y^s \frac{dy}{y},
\end{aligned}$$

since any lattice point is a multiple of a primitive point. We now split the integral and apply Poisson summation to get

$$\begin{aligned}\zeta(2s)E_\psi(g,s) &= \int_0^1 \sum_{m\in\mathbb{Z}^2} \psi[m\sqrt{y}g]y^s\frac{dy}{y} + \int_1^\infty \sum_{m\in\mathbb{Z}^2} \psi[m\sqrt{y}g]y^s\frac{dy}{y} \\ &= \int_0^1 \sum_m \psi[m\sqrt{y^{-1}}g^{-t}]y^{s-1}\frac{dy}{y} + \int_1^\infty \sum_m \psi[m\sqrt{y}g]y^s\frac{dy}{y} \\ &= \int_1^\infty \sum_m \psi[m\sqrt{y}g^{-t}]y^{1-s}\frac{dy}{y} + \int_1^\infty \sum_m \psi[m\sqrt{y}g]y^s\frac{dy}{y}.\end{aligned}$$

Taking the transpose of the inverse is an involution, and $\hat{\hat{\psi}} = \psi$, since ψ is an even function. Thus, the sum of integrals above is invariant under the change $(\psi, g, s) \mapsto (\hat{\psi}, g^{-t}, 1-s)$. By our assumptions, the series in both integrals have vanishing terms at $m = 0$. Since both functions ψ and $\hat{\psi}$ are Schwartz functions, $\psi[myg] \ll (\|m\|y)^{-k}$ for any $k \in \mathbb{N}$, so that the series $\sum_{m\neq 0}\psi[myg]$ decays rapidly as a function of y, and similarly for $\hat{\psi}$. Therefore, both integrals converge absolutely for any s and define entire functions. □

We would like a more elegant and symmetric notation. Letting

$$w = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

we have $g^{-t} = wgw^{-1}$. Define $\check{\psi}(g) = \check{\psi}[(0,1)g] = \hat{\psi}[(0,1)gw]$. Then,

$$\sum_{m\in\mathbb{Z}^2} \hat{\psi}[myg^{-t}] = \sum_{m\in\mathbb{Z}^2} \hat{\psi}[mwg^{-t}w^{-1}yw] = \sum_{m\in\mathbb{Z}^2} \check{\psi}[mgy],$$

where in the second step we make the variable change $m \mapsto mw$, which merely permutes $\mathbb{Z}^2$. Therefore, we may write the functional equation in Theorem 3.14 as

$$E^*_\psi(g,s) = E^*_{\check{\psi}}(g, 1-s).$$

On the symmetry line $\operatorname{Re} s = \frac{1}{2}$, we can rewrite this as

$$\zeta(2s)E_\psi(g,s) = \zeta(2\bar{s})E_{\check{\psi}}(g,\bar{s}). \tag{3.3}$$

We shall now make use of these properties of Eisenstein series to prove the analytic continuation of the Mellin transform $\widehat{\theta_\psi}$.

Theorem 3.15. *Let $\psi \in C_c^\infty(N\backslash G)$ be an even function, such that $\hat{\psi}[0] = 0$. Then*

$$\widehat{\theta_\psi}(g,s) = L_\psi(g, 1-s) + \frac{\zeta(2-2s)}{\zeta(2s)} L_{\check{\psi}}(g, 1-s).$$

The right hand side of the identity is meromorphic on $\mathbb{C}$ and holomorphic in the region $\operatorname{Re} s \geq 1/2$, *providing an analytic continuation for $\widehat{\theta_\psi}$.*

Proof. We prove this theorem in several steps. First, using the explicit formula in 3.3, we find

$$\widehat{\theta_\psi}(g,s) = \int_0^\infty \psi(a(y)g)y^{s-1}\frac{dy}{y} + \sum_{\substack{(c,d)=1\\ c>0}} \frac{1}{c^2}\int_0^\infty \int_N \psi(wn(x)a(c^2)a(y)g)dx \cdot y^{s-1}\frac{dy}{y}$$

$$= L_\psi(g,1-s) + \sum_{\substack{(c,d)=1\\ c>0}} \frac{1}{c^{2s}}\int\int \psi(wn(x)a(y)g)y^{s-1}dx\frac{dy}{y},$$

after a change of variables. Recall that the (c,d) sum ranges over positive integers c and d modulo c, coprime to c. Therefore,

$$\sum_{\substack{(c,d)=1\\ c>0}} \frac{1}{c^{2s}} = \sum_{n=1}^\infty \frac{\varphi(n)}{n^{2s}} = \frac{\zeta(2s-1)}{\zeta(2s)},$$

by the theory of arithmetic functions (here φ is Euler's totient function) and L-series. We further compute, by switching matrices and changing variables,

$$\int_{\mathbb{R}}\int_0^\infty \psi(wn(x)a(y)g)y^{s-1}\frac{dy}{y}dx = \int_{\mathbb{R}}\int_0^\infty \psi(wa(y)n(x)g)y^{s}\frac{dy}{y}dx$$
$$= \int_{\mathbb{R}}\int_0^\infty \psi(a(y)wn(x)g)y^{-s}\frac{dy}{y}dx = \int_{\mathbb{R}} L_\psi(wn(x)g,s)dx.$$

Therefore, for $\operatorname{Re} s > 1$ (where all the manipulations above make sense),

$$\widehat{\theta_\psi}(g,s) = L_\psi(g,1-s) + \frac{\zeta(2s-1)}{\zeta(2s)}\int_{\mathbb{R}} L_\psi(wn(x)g)dx. \tag{3.4}$$

We now need to obtain a different expression for the last term in the previous identity. We consider the zero-th Fourier coefficient of the Eisenstein series

$$E^0_\psi(g,s) = \int_{\Gamma_\infty\backslash N}\sum_{\gamma\in\Gamma_\infty\backslash\Gamma} L_\psi(\gamma n(x)g,s)dx,$$

and use essentially the same computations as in the proof of Lemma 3.3 to obtain

$$\frac{1}{2}E^0_\psi(g,s) = L_\psi(g,s) + \sum_{\substack{(c,d)=1\\ c>0}} \int_N L_\psi(wa(c^2)n(x)g,s)dx$$

$$= L_\psi(g,s) + \sum_{\substack{(c,d)=1\\ c>0}} \int_N L_\psi(a(c^{-2})wn(x)g,s)dx$$

$$= L_\psi(g,s) + \sum_{\substack{(c,d)=1\\ c>0}} \frac{1}{c^{2s}}\int_N L_\psi(wn(x)g,s)dx$$

$$= L_\psi(g,s) + \frac{\zeta(2s-1)}{\zeta(2s)}\int_N L_\psi(wn(x)g,s)dx.$$

We may multiply the last identity by $\zeta(2s)$ and Theorem 3.14 implies that

$$\frac{1}{2}\zeta(2-2s)E^0_{\check{\psi}}(g,1-s) = \frac{1}{2}\zeta(2s)E^0_{\psi}(g,s) = \zeta(2s)L_{\psi}(g,s)+\zeta(2s-1)\int_{\mathbb{R}} L_{\psi}(wn(x)g,s)dx \quad (3.5)$$

is an entire function. On the other hand, the same computations as above show that

$$\frac{1}{2}\zeta(2-2s)E^0_{\check{\psi}}(g,1-s) = \zeta(2-2s)L_{\check{\psi}}(g,1-s)+\zeta(1-2s)\int_N L_{\check{\psi}}(g,wn(x)g,1-s)dx.$$

We now distinguish functions on $N\backslash G$ by type, as in [Lan75]. A function f on $N\backslash G$ has type s if the equation $f(a(y)k(\theta)) = y^s f(k(\theta))$ holds for all arguments. Lemma 3.8 implies that $\widehat{\Theta_{\psi}}$ has type $1-s$ and, by similar methods (essentially a change of variables), one can see that $L_{\psi}(g,s)$ has type s. Therefore, equation (3.4) implies that $\zeta(2s-1)\int_{\mathbb{R}} L_{\psi}(wn(x)g)dx$ has type $1-s$.

The basic observation that we will use here is that a decomposition into a sum of functions with different types is unique. More precisely, if f_1 and f_2 have type s and g_1, g_2 have type s', then $f_1+g_1 = f_2+g_2$ implies that $f_1 = f_2$ and $g_1 = g_2$. This can readily be seen since the equality of the two sums implies that $y^s(f_1-f_2) = y^{s'}(g_2-g_1)$, and this holds for any y.

Putting together equations (3.5) and (3.5) gives us such a decomposition into a sum of functions with type $1-s$ and s. Taking the last two paragraphs into consideration, we can deduce that

$$\zeta(2-2s)L_{\check{\psi}}(g,1-s) = \zeta(2s-1)\int_{\mathbb{R}} L_{\psi}(wn(x)g,s)dx,$$

which together with (3.4) proves the formula announced in the theorem.

Regarding the analytic continuation, note that Lemma 3.5 implies that the first term in the formula, $L_{\psi}(g,1-s)$, is meromorphic on $\mathbb{C}$ and holomorphic on $\operatorname{Re} s \geq 0$ (in fact, it is holomorphic everywhere, since ψ has compact support). In the second term, the Mellin transform $L_{\check{\psi}}(g,1-s)$ is also meromorphic and holomorphic on $\operatorname{Re} s \geq 0$. The (simple) pole of $\zeta(2-2s)$ at $s = \frac{1}{2}$ is killed by the pole of $\zeta(2s)$, so that the fraction $\zeta(2-2s)/\zeta(2s)$ is holomorphic for $\operatorname{Re} s \geq \frac{1}{2}$, except at possible zeros of $\zeta(2s)$. It is a classical result (i.e. the prime number theorem) that $\zeta(2s)$ has no zeros on the line $\operatorname{Re} s = \frac{1}{2}$, which implies the desired holomorphy of $\widehat{\Theta_{\psi}}$. □

Recall our desire to prove the Plancherel formula (3.2) by shifting the contour in Theorem 3.9. Having obtained the analytic continuation of $\widehat{\Theta_{\psi}}$, we only need to prove rapid decay on vertical strips. By Lemma 3.6, the Mellin transforms L_{ψ} and $L_{\check{\psi}}$ have this property, and since

$$\frac{1}{\zeta(2s)} \ll \log^7|t|,$$

for $s = \sigma + it, \sigma \geq \frac{1}{2}$, we easily obtain the following corollary.

Corollary 3.16. *If $\psi \in C_c^{\infty}(N\backslash G)$ is an even function, then $\widehat{\Theta_{\psi}}$ is rapidly decreasing on any vertical line, uniformly on finite strips, in the region* $\operatorname{Re} s \geq \frac{1}{2}$.

3.6 The decomposition

We are now ready to prove the Plancherel formula (3.2).

Theorem 3.17. *Let θ_ψ be the incomplete theta series of an even function $\psi \in C_c^\infty(N\backslash G)$ with $\hat{\psi}[0] = 0$. We have*

$$\|\theta_\psi\|^2_{\Gamma\backslash G} = \frac{1}{4\pi i}\int_{\substack{\operatorname{Re} s=\frac{1}{2}\\ \operatorname{Im} s>0}} \|\widehat{\theta_\psi}(s)\|^2_K ds.$$

Proof. We start with Theorem 3.9, i.e.

$$\|\theta_\psi\|^2 = \frac{1}{4\pi i}\int_K\int_{\operatorname{Re} s=\sigma} \widehat{\theta_\psi}(k,s)\overline{L_\psi(k,\overline{s})}dsd\theta,$$

for some $\sigma > 1$. By the meromorphic continuation of the Mellin transform and its properties, Theorem 3.15, Corollary 3.16, and Lemma 3.6, we may shift the contour of integration to $\sigma = \frac{1}{2}$. Since the integrands are rapidly decreasing, we can interchange integrals, and by Lemma 3.11 we obtain

$$\begin{aligned}
4\pi i\|\theta_\psi\|^2 &= \int_{\operatorname{Re} s=\frac{1}{2}}\int_K \widehat{\theta_\psi}(k,s)\overline{L_\psi(k,\overline{s})}d\theta ds\\
&= \int_{\operatorname{Im} s>0}(\theta_\psi, E_\psi(\overline{s})) + \int_{\operatorname{Im} s<0}(\theta_\psi, E_\psi(\overline{s}))\\
&= \int_{\operatorname{Im} s>0}(\theta_\psi, E_\psi(\overline{s})) + \int_{\operatorname{Im} s>0}\left(\theta_\psi, E_{\check{\psi}}(\overline{s})\zeta(2\overline{s})/\zeta(2s)\right)\\
&= \int_{\operatorname{Im} s>0}\left(\theta_\psi, E_\psi(\overline{s}) + E_{\check{\psi}}(\overline{s})\zeta(2\overline{s})/\zeta(2s)\right).
\end{aligned}$$

Reading the computations that lead up to Lemma 3.11, i.e. folding the integral, and using the formula in Theorem 3.15, we get

$$\begin{aligned}
4\pi i\|\theta_\psi\|^2 &= \int_{\operatorname{Im} s>0}\int_K \widehat{\theta_\psi}(k,s)\operatorname{conj}\left(L_\psi(\overline{s}) + \frac{\zeta(2\overline{s})}{\zeta(2s)}L_{\check{\psi}}(\overline{s})\right)d\theta ds\\
&= \int_{\operatorname{Im} s>0}\int_K \widehat{\theta_\psi}(k,s)\overline{\widehat{\theta_\psi}(k,s)}d\theta ds = \int_{\operatorname{Im} s>0}\|\widehat{\theta_\psi}(s)\|^2,
\end{aligned}$$

keeping in mind that integration occurs on the line $\operatorname{Re} s = \frac{1}{2}$, where $1 - s = \overline{s}$. □

Returning to the right regular representation, we see that projecting the incomplete theta series to its principal series components commutes with the representation. Indeed,

$$\widehat{\rho_h\theta_\psi}(g,s) = \int_0^\infty \rho_h\theta_\psi(a(y)g)y^{s-1}\frac{dy}{y} = \int_0^\infty \theta_\psi(a(y)gh)y^{s-1}\frac{dy}{y} = \rho_h\widehat{\theta_\psi}(g,s).$$

Denote by Θ_0 the closed subspace of Θ generated by the incomplete theta series satisfying the conditions in Theorem 3.17. We may interpret Theorem 3.17 as decomposing the right regular representation on Θ_0 as a "continuous" sum of irreducible (unitary) representations in the principal series.

Finally, we make note of an inversion formula for this decomposition. Unfortunately, it only applies for incomplete theta series, not for all functions in Θ_0. Recalling the remark before equation (3.1) that the Mellin transform is the Fourier transform for the multiplicative group $\mathbb{R}_{>0}$, we derived the Plancherel formula in this section by manipulating the Plancherel formula of the Mellin transform. There is a corresponding inversion formula, as well, namely

$$\psi(a(y)g) = \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=\sigma} L_\psi(g,s)y^s ds,$$

where L_ψ converges on the line $\mathrm{Re}\, s = \sigma$. Applying this formula to the definition of incomplete theta series and using the unfolding method we obtain

$$\theta_\psi(g) = \frac{1}{2\pi i}\sum_{\gamma\in\Gamma_\infty\backslash\Gamma}\int_{\mathrm{Re}\, s=\sigma} L_\psi(g,s)ds = \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=\sigma} E_\psi(g,s)ds, \tag{3.6}$$

where $\sigma > 1$ for the convergence of the Eisenstein series.

As in [God66, §8], one can show that E_ψ is rapidly decreasing on vertical strips for $\mathrm{Re}\, s \geq 1/2$ (using the representation of $\zeta(2s)E_\psi$ as a Mellin transform in the proof of Theorem 3.14). In this range E_ψ has no poles by Theorem 3.14 (note that the pole of $\zeta(2s)$ at $s = 1/2$ is cancelled, as in formula (3.3)). Thus, we can shift the integration domain to $\mathrm{Re}\, s = 1/2$ and obtain an integral representation that can be interpreted as an inversion formula. We thus have

$$\begin{aligned}\theta_\psi(g) &= \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=\frac{1}{2}} E_\psi(g,s)ds = \frac{1}{4\pi i}\int_{\mathrm{Re}\, s=\frac{1}{2}} E_\psi(g,s) + E_\psi(g,1-s)ds \\ &= \frac{1}{4\pi i}\int_{\mathrm{Re}\, s=\frac{1}{2}} E_\psi(g,s) + E_\psi(g,1-s) = \frac{1}{4\pi i}\int_{\mathrm{Re}\, s=\frac{1}{2}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \widehat{\theta_\psi}(\gamma g,s)ds.\end{aligned}$$

The last step was obtained by summing the formula in Theorem 3.15 over the cosets $\Gamma_\infty\backslash\Gamma$ and using the functional equation of the Eisenstein series.

4 The discrete part

We have dealt with the continuous part of the regular representation in the previous section. The rest of the decomposition consists of a discrete direct sum. We have to understand the orthogonal complement of the space of incomplete theta series Θ, but also what the complement of the subspace Θ_0 is within Θ. An elegant way of investigating this is to view some of our theory through the prism of operators, as in [Lan75].

4.1 Some adjoint operators

We can view the incomplete theta series θ_ψ as the image of the function ψ under the operator θ. Similarly, denote $T^0 f = f^0$ for an automorphic function f, so that T^0 is an operator on $\mathcal{H}$. These two operators are adjoint in the following sense.

Lemma 4.1. *Let ψ be a smooth function on $N\backslash G$ with compact support and $f \in \mathcal{H}$ an automorphic function. Then*

$$(\theta_\psi, f)_{\Gamma\backslash G} = \frac{1}{2}(\psi, f^0)_{N\backslash G}.$$

In particular, if f is orthogonal to the space Θ, then f is a cusp form. Moreover, $\hat{\psi}[0] = 0$ if and only if θ_ψ is orthogonal to the constant 1 function.

Proof. This is an application of the unfolding method, which is valid under our analytic conditions. We have

$$\begin{aligned}
(\theta_\psi, f)_{\Gamma\backslash G} &= \int_{\Gamma\backslash G} \overline{f(g)} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \psi(\gamma g) dg \\
&= \int_{\Gamma_\infty\backslash G} \psi(g)\overline{f(g)} dg \\
&= \int_{N\backslash G}\int_{\Gamma_\infty\backslash G} \psi(ng)\overline{f(ng)} dn dg \\
&= \int_{N\backslash G} \psi(g)\overline{f^0(g)} dg = (\psi, f^0)_{N\backslash G}.
\end{aligned}$$

If $(\theta_\psi, f) = 0$ for all $\psi \in C_c^\infty(N\backslash G)$, then $f^0 = 0$, since we have $(\psi, f^0) = 0$ and $C_c^\infty(N\backslash G)$ is a dense subspace.

Finally, note that the invariant measure on $N\backslash G$ is the Lebesgue measure on $\mathbb{R}^2$, i.e. $dxdy$ in Iwasawa coordinates. Therefore,

$$\hat{\psi}[0] = \int_{N\backslash G} \psi(g) dg = (\psi, 1)_{N\backslash G} = (\psi, T^0(1)) = (\theta_\psi, 1).$$

The desired orthogonality relation follows. $\square$

Denote the closed subspace of cusp forms by $\mathcal{H}^0$. We have just proved the orthogonal decomposition

$$\mathcal{H} = \mathcal{H}^0 \oplus \langle 1\rangle \oplus \Theta_0.$$

Each subspace is invariant under the right regular representation. Indeed, it is readily seen that the constant functions and $\mathcal{H}^0$ are invariant, respectively, directly from the definitions. The space Θ_0 is invariant, since the right regular representation is unitary and the decomposition above is orthogonal.

The space that still needs investigation is the subspace of cusp forms $\mathcal{H}^0$. Unfortunately, we do not have a very explicit description of this space, but we can understand something important about the regular representation on it. Namely, the regular representation is completely reducible, i.e., it is the direct (orthogonal) sum of irreducible representations. Put another way, the space $\mathcal{H}^0$ is the direct sum of irreducible G-invariant subspaces[9]. To prove this, we introduce a related representation of the space $L^1(G)$ on $\mathcal{H}^0$, for which we prove that it acts by compact operators. Using Dirac sequences, we recover from this the desired decomposition for the regular representation.

4.2 Integrating representations and Dirac sequences

Given a unitary representation π of G on a Hilbert space H, we can produce a representation of the space $L^1(G)$ by an integration process. In many applications, gaining information about one of these representations leads to more insight into the other one.

For any $\phi \in C_c(G)$ and $v \in H$ define

$$\pi^1(\phi)v = \int_G \phi(g)\pi(g)v dg.$$

This is a H-valued integral, which is defined just as in real or complex integration theory (we use here the completeness of Hilbert spaces; see a discussion in [Rob83, pp. 54]). The integral exists given the continuity and compact support of ϕ. If we equip $C_c(G)$ with the convolution product, i.e.,

$$\phi_1 * \phi_2(g) = \int_G \phi_1(gh^{-1})\phi_2(h)dh,$$

then $\tilde{\rho}$ is an algebra homomorphism $C_c(G) \longrightarrow End(\mathcal{H})$. One can compute that $\left\|\pi^1(\phi)\right\| \leq \left\|\phi\right\|_1$, so that the linear map $\pi^1(\phi)$ is a bounded operator. The homomorphism π^1 extends to a (continuous) representation $L^1(G) \longrightarrow B(\mathcal{H})$, where $B(\mathcal{H})$ denotes the bounded endomorphisms of $\mathcal{H}$. Since π is unitary, one can compute that

$$\pi^1(\phi)^* = \pi^1(\phi^*),$$

where $\phi^*(g) = \overline{\phi(g^{-1})}$.

Recall that our goal is to decompose $\mathcal{H}^0$ into G-invariant irreducible subspaces. For making use of the integrated regular representation ρ^1 to this end, note that a G-invariant subspace is also $L^1(G)$-invariant. Indeed, in general, the image of a vector v under the integrated representation π^1 is the limit of sums of images $\pi(g)v$ with certain coefficients. This limit lies in any G-invariant subspace containing v, since these subspaces are (by our convention) closed.

We postpone the investigation of the integrated regular representation ρ^1 for the next section. For proving the complete reducibility of $\mathcal{H}^0$ from properties of ρ^1, we need Dirac sequences.

[9]In this essay, a subspace of a Hilbert space is always meant to be *closed*.

Definition 4.2. A *Dirac sequence* is a sequence of real-valued functions $(\eta_n)_{n\in\mathbb{N}} \subset C_c(G)$, such that all η_n are non-negative,

$$\int_G \eta_n(g)dg = 1,$$

for all n, and for any neighbourhood U of the identity in G, the support of η_n is contained in U for all n large enough.

The existence of Dirac sequences on G may be seen for instance using the local Euclidean structure. Note as well that we can produce self-adjoint Dirac sequences, by taking the sequence $\frac{1}{2}(\eta_n + \eta_n^*)$ for any given Dirac sequence η_n. Using Dirac sequences, we can approximate the identity operator.

Lemma 4.3. *If (η_n) is a Dirac sequence, then for each function $f \in \mathcal{H}$, the sequence $\rho^1(\eta_n)f$ converges to f.*

Proof. We have

$$\rho^1(\eta_n)f - f = \int_G \eta_n(g)\rho(g)f - \eta_n(g)f\,dg$$

for any n. Given $\varepsilon > 0$, we can find a neighbourhood U of the identity in G, such that $\left\|\rho(g)f - f\right\| < \varepsilon$ for all $g \in U$. By definition, for all n large enough, the support of η_n is contained in U. For these n we have

$$\left\|\rho^1(\eta_n)f - f\right\| \leq \int_U \eta_n(g)\left\|\rho(g)f - f\right\|dg \leq \varepsilon \int_U \eta_n(g) = \varepsilon.$$

This shows the desired convergence. □

For the next theorem we need to recall some properties of compact operators. An operator $T : H \longrightarrow H$ on a Hilbert space H is called *compact* if the image of the unit ball has compact closure. The operator T is called *hermitian* if it equals its adjoint T^*. The *spectral theorem* for compact operators (see [Lan75, I, §2]) states that the space H decomposes as the orthogonal sum of the eigenspaces of a compact self-adjoint operator T over all its eigenvalues. Moreover the eigenvalues of T are discrete, except for possible accumulation at 0, and the eigenspaces are finite dimensional.

Lemma 4.4. *Let π be a unitary representation of G on a Hilbert space H. Assume that there is a Dirac sequence (η_n) such that $\pi^1(\eta_n)$ is a self-adjoint compact operator for all n. Then H is the direct sum of irreducible G-invariant subspaces.*

Proof. We use the proof in [Bor08, Lemma 16.1]. We prove first that any G-invariant subspace $W \neq 0$ contains an irreducible invariant subspace. For this we use our Dirac sequence. Since $\pi^1(\eta_n)$ approximates the identity, there exists j such that $\pi^1(\eta_j)$ is nonzero on W. By the spectral theorem, we can find an eigenspace $M \neq 0$ of $\pi^1(\eta_j)$ in W. Since M is finite dimensional, we can find a minimal subspace N among the nonzero intersections of M with G-invariant subspaces of W. Let $v \in N$ be nonzero and let P be the smallest G-invariant subspace containing v. By the minimality properties, it follows that $N = P \cap M$.

We now show that P is irreducible under the G-action. Let $Q \subset P$ be a G-invariant subspace and R the orthogonal complement of Q inside P. Then R is G-invariant as well, since $\pi^1(\eta_j)$ is self-adjoint. Thus, P, Q, R are $\pi^1(\eta_j)$ invariant (as in the discussion at the beginning of the section). We have $N = (Q \cap M) \oplus (R \cap M)$ and by minimality we must have $R = 0$ or $Q = 0$. This implies the irreducibility of P.

We must now show that the entire space H decomposes into irreducible subspaces. The set of subspaces that are decomposable into direct sums of irreducible subspaces is non-empty by the preceding paragraphs and it is partially ordered by inclusion. By Zorn's lemma, we find a maximal such subspace V. If $V \neq H$, let $V^\perp \neq 0$ be the orthogonal complement of V in H, which is also G-invariant, since $\pi(g)^* = \pi(g^{-1})$ by unitarity. The argument above allows us to find an irreducible non-zero subspace $P \subset V^\perp$, which could be added to V, producing a larger direct sum of irreducible subspaces — contradiction.

Finally, we prove that the multiplicities are finite. Let (σ, F) be an irreducible unitary representation occurring in H with multiplicity n_σ. We can find j such that $\sigma^1(\eta_j)$ is nonzero on F. By compactness, let λ be a nonzero eigenvalue of $\sigma^1(\eta_j)$ and m its multiplicity in F. For any representation $\tilde{\sigma}$ isomorphic to σ, λ is also an eigenvalue of $\tilde{\sigma}^1(\eta_j)$ with multiplicity m. By construction, $\pi^1(\eta_j)$ restricts to $\sigma^1(\eta_j)$ and to the other isomorphic representations. Compactness implies that the dimension of the λ-eigenspace of $\pi^1(\eta_j)$ is finite, but also at least $n_\sigma \cdot m$, by the above. Therefore, the multiplicity n_σ is finite. $\square$

To apply this lemma to our situation, we need to prove that the integrated regular representation on the space of cusp forms $\mathcal{H}^0$ acts by compact operators.

4.3 A criterion for compactness

We prove that the operators $\rho^1(\phi)$ are compact by showing they are integral operators with square-integrable kernels. An *integral operator* T is given by an expression of the form $Tf(g) = \int \mathcal{K}(g,h) f(h) dh$, where $\mathcal{K}$ is a measurable function in two variables and is called the *kernel*. If $\mathcal{K}$ is square-integrable, then the operator T is compact (see [Lan75, I, §3]). The criterion that we shall use is given in the following lemma.[10]

Lemma 4.5. *Let X be a locally compact space with positive finite measure μ, such that the σ-algebra of measurable sets is generated by a countable subalgebra. Let H be a subspace of $L^2(X)$, and T a linear map from H into the space of bounded continuous functions on X. Assume that there exists $C > 0$ such that*

$$\|Tf\|_\infty \leq C \|f\|_2,$$

[10] This is proven here in a more general setting. The use of variables x, y should *not* remind the reader of the Iwasawa decomposition.

for all $f \in H$. Then $T : H \longrightarrow L^2(X)$ is an integral operator with square-integrable kernel and thus compact.

Proof. Since T is continuous for the supremum norm on the bounded functions on X, it follows that the evaluation map $f \mapsto Tf(x)$ is continuous for all $x \in X$. By the Riesz representation theorem, there exists a function $k_x \in H$ such that for all $f \in H$ we have

$$Tf(x) = (f, k_x) = \int f(y)\overline{k_x(y)}dy.$$

Since Tf is continuous, it follows that $x \mapsto k_x$ is weakly continuous as a map $X \longrightarrow H \subset L^2(X)$, and thus weakly measurable. Each k_x is square-integrable by construction.

We now prove that there exists $\mathcal{K} \in L^2(X \times X)$ such that for almost all $x \in X$ we have

$$k_x(y) = \mathcal{K}(x, y),$$

for all y outside a set of measure zero, depending on x. This will then prove the lemma.

We first show that the map $x \mapsto (k_x, k_x)$ is measurable. To see this, note that, under our conditions on the measure, the space $L^2(X)$ is separable. Given a basis (u_j) of $L^2(X)$, we have the Fourier expansion $k_x = \sum a_j(x)u_j$. Since k_x is weakly measurable, it follows that $a_j(x)$ is measurable for all j. Therefore,

$$(k_x, k_x) = \sum |a_j(x)|^2$$

is a limit of measurable functions, whence measurable.

We have noted above that $x \mapsto (k_x, k_x)$ is also bounded and thus lies in $L^1(X)$. A similar proof to that of the Cauchy-Schwarz inequality shows that

$$\left| \int \int g(x,y)\overline{k_x(y)}dydx \right|^2 \leq \|g\|_2^2 \int \int |k_x(y)|^2 dydx,$$

for any step function g on $X \times X$ with respect to products of measurable sets. Therefore, the map

$$g \mapsto \int \int g(x,y)\overline{k_x(y)}dydx$$

is L^2-continuous. The Riesz representation theorem implies that there exists a map $\mathcal{K} \in L^2(X \times X)$ such that for all characteristic functions ψ and ψ of measurable sets on X we have

$$\int \phi(x) \int \psi(y)\overline{k_x(y)}dydx = \int \phi(x) \int \psi(y)\overline{\mathcal{K}(x,y)}dydx,$$

by taking above $g = \phi \otimes \psi$.

Therefore, for each ψ there exists a measure-zero set Z_ψ such that if $x \notin Z_\psi$, then

$$\int \psi(y)\overline{k_x(y)}dy = \int \psi(y)\overline{\mathcal{K}(x,y)}dy.$$

This relation holds for countably many ψ, for all x outside a set Z of measure zero. This implies that, for $x \notin Z$,

$$k_x(y) = \mathcal{K}(x, y)$$

for all y outside a set of measure zero, depending on x. This finally proves the lemma. □

4.4 Proof of compactness and of complete reducibility

We now prove that $\rho^1(\phi)$, which we denote by $\rho(\phi)$ in this section for simplicity, is a compact operator on $\mathcal{H}^0$ using the criterion in the previous section. Lemma 4.4 then implies the complete reducibility of the space of cusp forms.

Theorem 4.6. *If $\phi \in C_c^\infty(G)$, then there exists $C_\phi > 0$ such that for all cusp forms $f \in \mathcal{H}^0$ we have*

$$\left\|\rho(\phi)f\right\|_\infty \leq C_\phi \|f\|_2.$$

Proof. By a change of variables and the unfolding method, we have

$$\begin{aligned}
\rho(\phi)f(g) &= \int_G \phi(g^{-1}h)f(h)dh \\
&= \int_{\Gamma_\infty\backslash G} \sum_{\gamma\in\Gamma_\infty} \phi(g^{-1}\gamma h)f(\gamma h)dh \\
&= \int_{\Gamma_\infty\backslash G} \mathcal{K}_\phi(g,h)f(h)dh,
\end{aligned}$$

where

$$\mathcal{K}_\phi(g,h) = \sum_{m\in\mathbb{Z}} \phi\left(g^{-1}\left(\begin{smallmatrix} 1 & m \\ & 1\end{smallmatrix}\right)h\right).$$

The function $\phi_{g,h}(t) = \phi\left(g^{-1}\left(\begin{smallmatrix} 1 & t \\ & 1\end{smallmatrix}\right)h\right)$ is a smooth and compactly supported function on $\mathbb{R}$. We apply the Poisson summation formula to find that

$$\mathcal{K}_\phi(g,h) = \sum_{m\in\mathbb{Z}} \widehat{\phi_{g,h}}(m),$$

where $\widehat{\phi_{g,h}}$ is the Fourier transform of $\phi_{g,h}$. This is the expression that we shall bound to prove the theorem.

To distinguish Iwasawa decompositions, we write $g = n_g a_g k_g$ and $h = n_h a_h k_h$. If $a_g = a(y)$, then we can compute (see Remark 3.7) that $a_g^{-1} n(t) a_g = n(ty^{-1})$. Making a variable change, this implies that

$$\begin{aligned}
\widehat{\phi_{g,h}}(t) &= \int_{\mathbb{R}} \phi(g^{-1}n(t)h)e^{-2\pi i\lambda t}dt \\
&= \int_{\mathbb{R}} \phi(g^{-1}a_g a_g^{-1} n(t) a_g a_g^{-1} h)e^{-2\pi i\lambda t}dt \\
&= \int_{\mathbb{R}} \phi(g^{-1}a_g n(u) a_g^{-1} h)e^{-2\pi i\lambda y u} y du.
\end{aligned}$$

If we set $\omega_g = g^{-1}a_g$, $\omega_{g,h} = a_g^{-1}h$ and $\phi_{\omega,g,h}(t) = \phi\left(\omega_g n(t)\omega_{g,h}\right)$, then we have found that

$$\widehat{\phi_{g,h}}(\lambda) = y\widehat{\phi_{\omega,g,h}}(\lambda y).$$

We shall prove that the function $\widehat{\phi_{g,h}}$ is smooth and compactly supported for the parameters that are relevant for the bounding problem. First we reduce the domain of the variable g to one that is more convenient technically. Let $\mathfrak{S}$ be a so-called *Siegel set* consisting of all $g = n(x)a(y)k(\theta) \in G$ such that $n(x)$ lies in a compact set $\Omega_N \subset N$, e.g. $x \in [-1/2, 1/2]$, and $y \geq c$ for some $c > 0$. Choosing c small enough, e.g. $c = 1/2$ in our case, and considering that the image of $\mathfrak{S}$ in the quotient $G/K \cong \mathbb{H}$ covers a fundamental domain for Γ acting on $\mathbb{H}$ (recall Figure 1), we see that

$$\Gamma \cdot \mathfrak{S} = G.$$

For estimating $\rho(\phi)f(g)$ we may assume that $g \in \mathfrak{S}$, since f and thus $\rho(\phi)f$ is left Γ-invariant.

We can assume without loss of generality that $\omega_g = g^{-1}a_g$ lies in a compact set. As Garrett puts it in [Gar], a point g in a Siegel set is well approximated by its A-coordinate $a(y) = a_g$. More precisely, if $g \in \mathfrak{S}$, then $g \in a_g\Omega_G$ for some compact set Ω_G of G, depending only on $\mathfrak{S}$. Indeed, the commutation relation in Remark 3.7 states that $n(x)a(y)K = a(y)n(xy^{-1})K$. If $n(x)a(y)k(\theta) \in \mathfrak{S}$, then x lies in a compact set and $y \geq c$, so that xy^{-1} also lies in a compact set that depends only on c, i.e., $n(xy^{-1})$ lies in some compactum Ω'_N. Letting $\Omega_G = \Omega'_N K$ proves the claim.

Similarly, without loss of generality, we can assume that $\omega_{g,h} = a_g^{-1}h$ lies in a compact set. Indeed, for bounding the kernel $\mathcal{K}$, we can assume that $\phi(g^{-1}\gamma h) \neq 0$ for some $\gamma \in \Gamma_\infty$. Shifting γ if necessary, we can assume that there is a compact domain $\Omega_N \subset N$ such that $h \in \Omega_N AK$. Now if C is the compact support of ϕ, then

$$g^{-1}\gamma h = g^{-1}a_g a_g^{-1}\gamma n_h a_g a_g^{-1} a_h k_h \in C.$$

As above we may assume that $g^{-1}a_g$ lies in a compactum. Therefore,

$$\underbrace{a_g^{-1}\gamma n_h a_g}_{\in N} \cdot \underbrace{a_g^{-1} a_h}_{\in A} \cdot k_h$$

lies in some compact set, only depending on ϕ and $\mathfrak{S}$. Using the Iwasawa decomposition, we find that there is some compact Ω_A such that $a_h \in a_g \in \Omega_A$. Therefore

$$h \in \Omega_N a_g \Omega_A K = a_g \Omega'_N \Omega_A K = a_g \Omega'_G,$$

where Ω'_G is compact, again using that $g \in \mathfrak{S}$, and only depending on $\mathfrak{S}$ and ϕ.

Returning to our bounding problem, since ω_g and $\omega_{g,h}$ lie in compact sets and ϕ has a compact support, it follows that $\phi_{\omega,g,h}$ also has a compact support, and we easily see that it is a smooth function. Therefore, its Fourier transform $\widehat{\phi_{\omega,g,h}}$ is in the Schwartz class, and for any positive integer k there is a constant $C(\phi,\mathfrak{S},k) > 0$ such that

$$|\widehat{\phi_{\omega,g,h}}(\lambda)| \le C(\phi,\mathfrak{S},k)|\lambda|^{-k}.$$

This inequality can be deduced explicitly, too, by using partial integration k times. We now find the desired estimate

$$|\widehat{\phi_{g,h}}(\lambda)| = |y\widehat{\phi_{\omega,g,h}}(\lambda y)| \le C(\phi,\mathfrak{S},k)|\lambda|^{-k}y^{1-k}. \tag{4.1}$$

To obtain a bound for

$$\rho(\phi)f(g) = \int_{\Gamma_\infty\backslash G}\sum_{m\in\mathbb{Z}}\widehat{\phi_{g,h}}(m)f(h)dh,$$

we now finally use the cusp condition

$$f^0(h) = \int_{\Gamma_\infty\backslash N} f(nh)dn = 0,$$

for all h, i.e. $f \in \mathcal{H}^0$. This condition implies that the $m = 0$ term in $\rho(\phi)f(g)$ vanishes. Indeed, unfolding the first integral, the $m = 0$ term is equal to

$$\int_{\Gamma_\infty\backslash G}\int_N \phi(g^{-1}nh)f(h)dndh = \int_{N\backslash G}\int_{\Gamma_\infty\backslash N}\int_N \phi(g^{-1}nn'h)f(n'h)dndn'dh.$$

Making the variable change $nn' \mapsto n$, under which dn is invariant, this expression is equal to

$$\int_{N\backslash G}\left(\int_{\Gamma_\infty\backslash N} f(n'h)dn'\right)\cdot\left(\int_N \phi(g^{-1}nh)dn\right)dh = 0,$$

by the cusp condition.

For the terms where $m \neq 0$, we use the bound in (4.1). Recall that for $a_g = a(y)$ and g in a fixed Siegel set $\mathfrak{S}$, if $\phi(g^{-1}\Gamma_\infty h)$ is nonzero, then $h \in a_g\Omega_G$ for some compact Ω_G depending on ϕ and $\mathfrak{S}$. We can rewrite this as $h \in \Omega_N a_g \Omega_A K$ by the same techniques above, where $\Omega_N \subset N$ and $\Omega_A \subset A$ are compact. For any positive k we have

$$\begin{aligned}\rho(\phi)f(g) &\ll \sum_{m\neq 0}\int_{\Gamma_\infty\backslash\Omega_N a_g\Omega_A K} y^{1-k}\sum_{m\neq 0}|m|^{-k}|f(h)|dh \\ &\ll y^{1-k}\int_{\Gamma_\infty\backslash\Omega_N a_g\Omega_A K}|f(h)|dh \\ &\ll y^{1-k}\mu(\Gamma\backslash G)^{1/2}\|f\|_2,\end{aligned}$$

where the last inequality employed Cauchy-Schwarz. For the change of integration domain, note that $\mu(\Omega_N a_g \Omega_A K)$ is finite, since y is bounded from below (in the Siegel set). Thus, this domain is covered by a finite number of shifted copies of a fundamental domain for $\Gamma\backslash G$. Therefore, we can bound the original integral by a constant times the integral over $\Gamma\backslash G$.

The last computation proves the desired bound, since $\mu(\Gamma\backslash G)$ is finite and y is bounded from below in the Siegel set. □

5 Concluding remarks

5.1 Relation to the hyperbolic Laplacian

The representation theoretic approach that was taken in this essay is very much related to the spectral theory of the hyperbolic Laplacian. We can restrict the space $L^2(\Gamma\backslash G)$ to functions invariant under K from the right. These functions then descend to the the space $L^2(\Gamma\backslash G/K)$. Since $G/K \cong \mathbb{H}$, these are functions on the quotient $\Gamma\backslash\mathbb{H}$. On the upper half plane, the hyperbolic Laplacian is the differential operator

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

It is invariant under the Möbius transformations of $\mathrm{SL}_2(\mathbb{R})$ and descends to the quotient $\Gamma\backslash\mathbb{H}$. The Laplacian is a priori defined on a dense subspace of $L^2(\Gamma\backslash\mathbb{H})$, but it has a Friedrichs extension to the whole Hilbert space. It can be shown that Δ commutes with the right regular representation of G, and thus, by Schur's lemma, the Laplacian acts by scalars on irreducible G-invariant subspaces.

We can define a similar operator for the space $L^2(\Gamma\backslash G)$, called the Casimir operator. In Iwasawa coordinates it takes the form

$$\mathcal{C} = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - y\frac{\partial^2}{\partial x \partial \theta}.$$

It also extends to the whole of $L^2(\Gamma\backslash G)$, acting as scalars on irreducible G-invariant subspaces. In fact, there is an isomorphism between eigenfunctions of Δ and eigenfunctions of $\mathcal{C}$, as explained in [Gel75, Example 2.3].

The spectral decomposition of the Laplacian or of the Casimir operator is analogous to the decomposition of the regular representation. This is essentially because K is compact and, thus, the regular representation restricted to K is completely reducible by Peter-Weyl (see a more in-depth discussion in [Gel75, §2]). In the discrete part of the decomposition for the Laplacian, there are the constant functions and the space of cusp forms, which has an orthonormal basis $\{e_i \mid i \in \mathbb{N}\}$ of eigenfunctions. The continuous part is described by Eisenstein series

$$E_s(z) = \sum_{\gamma \in \Gamma_\infty\backslash\Gamma} \mathrm{Im}(\gamma z)^s,$$

which are eigenfunctions of Δ, but not square-integrable. A function $f \in L^2(\Gamma\backslash\mathbb{H})$ can be decomposed as

$$f = \sum_{i \in \mathbb{N}} (f, e_i)e_i + (f, \sqrt{3/\pi})\sqrt{3/\pi} + \frac{1}{4\pi i}\int_{\mathrm{Re}\, s = \frac{1}{2}} (f, E_s)E_s ds,$$

as in [Gol15, Thm. 3.16.1]. Note here that $\sqrt{3/\pi}$ is there to normalise the constant function, since the volume of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ is $\pi/3$.

It is worth stating an important conjecture of Selberg regarding the eigenvalues of the Laplacian on $\Gamma\backslash\mathbb{H}$, now for any congruence subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$. Selberg proposed that all eigenvalues (which are real because Δ is self-adjoint) are greater or equal to 1/4. This has not yet been proven in general, but for $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ we have stronger lower bounds, such as $3\pi^2/2$ (see [Gol15, Thm. 3.7.2]). For applications and more details in the general case, see [LRS95].

5.2 Other subgroups

The spectral decomposition presented in this essay can be done analogously for other similar subgroups (such as congruence subgroups), as in [Kub73] for instance. We shall only describe here the essential differences.

As in [Kub73], we assume $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$ is a discrete group, such that the stabiliser of ∞ in the action on $\mathbb{H}$ is $\Gamma_\infty = \mathrm{SL}_2(\mathbb{Z}) \cap N$, as defined above (such a group is called by Kubota reduced at infinity). The group Γ may have several inequivalent cusps, that is, Γ-orbits of points in $\mathbb{R} \cup \{\infty\}$ with infinite cyclic stabiliser in Γ.

Let $\kappa_1, \dots \kappa_n$ be a complete set of inequivalent cusps for Γ. Note that we assume finiteness of this set, i.e. $n \in \mathbb{N}$, which is true for congruence subgroups. Since we assume infinity is a cusp, let κ_1 correspond to ∞. For each cusp, we associate a stabiliser of the cusp $\Gamma_i = \gamma_i \Gamma_\infty \gamma_i^{-1}$. For each cusp we can then define Eisenstein series and incomplete theta series by swapping Γ_∞ in the original definition for Γ_i. For example, at the cusp κ_i, we have

$$E_{i,\psi}(g,s) = \sum_{\gamma \in \Gamma_i} L_\psi(\gamma_i^{-1}\gamma g),$$

and similarly for $\theta_{i,\psi}$. At each cusp, we have a Fourier expansion (as in the classical theory), and the zero-th coefficient of an incomplete theta series at κ_j is given by

$$\theta^0_{ij,\psi}(g) = \int_0^1 \theta_{i,\psi}(\gamma_j n(x) g)\,dx.$$

Accordingly we define the Mellin transforms $\widehat{\theta_{ij,\psi}}$ of the zero-th Fourier coefficients.

The functional equation of the Eisenstein series changes to one relating the entire vector $(E_{1,\psi}, \dots, E_{n,\psi})$ to the vector $(E_{1,\hat{\psi}}, \dots, E_{n,\hat{\psi}})$ by a matrix depending on s, as in [Kub73, p. 74]. The Plancherel formula, giving the continuous spectrum, changes as well to

$$\left\|\theta_{i,\psi}\right\|^2_{\Gamma\backslash G} = \frac{1}{4\pi i} \sum_{j=1}^{n} \int_{\substack{\mathrm{Re}\, s = \frac{1}{2} \\ \mathrm{Im}\, s > 0}} \left\|\widehat{\theta_{ij,\psi}}(s)\right\|^2_K ds.$$

More details are provided in [Kub73, pp. 84].

As in the case of $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, the Plancherel formula only holds for certain functions ψ. The space of incomplete theta series decomposes into the subspace Θ_0 where the formula holds and an orthogonal complement, $\hat{\Theta}$, which is equal to the constant functions when $\Gamma = \mathrm{SL}_2(\mathbb{Z})$. More generally, this complement is generated by the residues of Eisenstein series. Without expanding on technical details, we can at least make an argument for the plausibility of the last statement using the formula (3.6), that is

$$\theta_\psi(g) = \frac{1}{2\pi i}\int_{\mathrm{Re}\, s=\sigma} E_\psi(g,s)ds.$$

Intuitively speaking, to obtain the Plancherel or inversion formula for the continuous part, we need to shift the contour of integration to $\mathrm{Re}\, s = 1/2$. Without any conditions on ψ, the Eisenstein series E_ψ may have poles in $1/2 \le \mathrm{Re}\, s \le 1$, which we pick up as residues when shifting the contour. For examples, if $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, then the Eisenstein series may have a pole at $s = 1$, with residue a constant function. The technical details are sketched out in [Kub73, pp. 92].

The main theorem regarding this decomposition is [Kub73, Theorem 7.5.7], which states that the space $\hat{\Theta}$ is a finite sum of irreducible subspaces for the regular representation. As with the decomposition for $\mathrm{SL}_2(\mathbb{Z})$, in the more general case, the residues of the Eisenstein series play an analogous role in the spectral decomposition of the hyperbolic Laplacian (see [Gar18, Corollary 1.14.1]).

References

[Bor08] Armand Borel. *Automorphic forms on SL2 (R)*. Cambridge University Press, 2008.

[Gar] Paul Garrett. *Decomposition and estimates for cuspforms*. URL: http://www-users.math.umn.edu/~garrett/m/v/cfms_decompn.pdf.

[Gar18] Paul Garrett. *Modern Analysis of Automorphic Forms By Example*. Vol. 1. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2018. DOI: 10.1017/9781316650332.

[Gel75] Stephen Gelbart. *Automorphic Forms on Adele Groups. (AM-83)*. Princeton University Press, 1975. ISBN: 9780691081564. URL: http://www.jstor.org/stable/j.ctt1b7x82z.

[God66] R. Godement. "The decomposition of $L^2(G/\Gamma)$ for $\Gamma = SL(2,Z)$". In: *Proceedings of Symposia in Pure Mathematics* 9 (1966), 211—224. DOI: 10.1090/pspum/009/0210827.

[Gol15] Dorian M. Goldfeld. *Automorphic forms and L-functions for the group GL(n, R)*. Cambridge University Press, 2015.

[Iwa02] Henryk Iwaniec. *Spectral methods of automorphic forms*. 2nd ed. American Mathematical Society, 2002.

[Kub73] Tomio Kubota. *Elementary theory of Eisenstein series*. Kodansha, 1973.

[Lan75] Serge Lang. *SL2(R)*. Addison-Wesley, 1975.

[LRS95] W. Luo, Z. Rudnik, and P. Sarnak. "On Selberg's eigenvalue conjecture". In: *Geometric and Functional Analysis* 5 (1995), pp. 387–401. DOI: 10.1007/BF01895672.

[Rob83] Alain Robert. *Introduction to the Representation Theory of Compact and Locally Compact Groups*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1983. DOI: 10.1017/CBO9780511661891.