Spectral Theory of $\Gamma\backslash SL(2,\mathbb{R})$
Copyright 2001 by Daniel Bump

These are lecture notes for three lectures in Jerusalem, March 2001. To key the reader, the transparency page numbers numbers are displayed in boxes.

We've tried to give a pithy account of the Spectral theory, emphasizing representation theory, with proofs reduced to their essentials. More complete proofs and details of many aspects can be found in Bump, *Automorphic Forms and Representations*, Chapter 2 and Chapter 3, Section 2.

One omission is that we only prove the analytic continuation and functional equations of the Eisenstein series for $SL(2,\mathbb{Z})$. This case is atypical and therefore somewhat misleading, as the work of Sarnak and Philips showed. Thus one should ponder also the nonarithmetic case, for which we recommend Appendix IV in Langlands [LES]. We also prove the Selberg trace formula only for compact quotients, and we omit several important related matters such as truncation and the Maass-Selberg relations.

At the end we consider the Jacquet-Langlands correspondence.

$\boxed{1}$

$G = SL(2,\mathbb{R})$ acts on $\mathfrak{H} = \{x+iy | y>0\}$:

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

The stabilizer of i is $K = SO(2)$ so $\mathfrak{H} \cong G/K$. The *noneuclidean Laplacian*

$$\Delta = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right). \tag{1}$$

is a G-invariant differential operator. Let Γ be a discrete cocompact subgroup of $G = SL(2,\mathbb{R})$. Then $X = \Gamma\backslash\mathfrak{H}$ is a compact Riemann surface. Δ is symmetric with respect to the invariant metric $y^{-2}(dx \wedge dy)$:

$$\langle \Delta f, g\rangle = \langle f, \Delta g\rangle\,, \qquad f,g \in C^\infty(\Gamma\backslash\mathfrak{H})$$

(Stokes). It extends to a self-adjoint differential operator on $L^2(\Gamma\backslash\mathfrak{H}, dx \wedge dy/y^2)$. (Proof later.)

$\boxed{2}$

Example. $1 \neq \gamma \in \Gamma$ is *hyperbolic* if its eigenvalues are real, *elliptic* if complex of absolute value 1. Γ is *hyperbolic* if each $1 \neq \gamma \in \Gamma$ is hyperbolic. Let X be a compact Riemann surface of genus ≥ 2. Its universal cover is $\mathfrak{H}$ and $\Gamma = \pi_1(X)$ acts with quotient X. These examples are precisely the hyperbolic groups.

Let $\mathfrak{g}$ be the Lie algebra of G. Thus $\mathfrak{g} = \{g \in \mathrm{Mat}_2(\mathbb{R}) | \mathrm{tr}(g) = 0\}$ with $[X,Y] = XY - YX$ (matrix mult). The universal enveloping algebra $U(\mathfrak{g})$ is the free algebra on $\mathfrak{g}$ modulo the relations

$$[X,Y] - (X\cdot Y - Y \cdot X) = 0$$

($\cdot =$ mult in $U(\mathfrak{g})$). The center of $U(\mathfrak{g})$ is $\mathbb{C}[\Delta]$,

$$-4\Delta = \hat{H}^2 + 2\hat{R}\hat{L} + 2\hat{L}\hat{R}, \tag{2}$$

$$\hat{R} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \hat{L} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \hat{H} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

This is the Laplace-Beltrami operator. It gives rise to the Laplacian, as we will explain.

$\boxed{3}$

Let $\rho : G \to C^\infty(G)$ be the right regular representation: $(\rho(g)f)(h) = f(hg)$. This has a Lie algebra version: if $X \in \mathfrak{g}$ let

$$(Xf)(h) = \frac{d}{dt} f(he^{tX})|_{t=0}. \tag{3}$$

It is a Lie algebra rep: $[X,Y]$ has the same effect as $X \circ Y - Y \circ X$ ($\circ =$ composition of operators). So it extends to a representation of $U(\mathfrak{g})$. Because Δ is in the center, it commutes with left and right translation.

Functions on $\mathfrak{H}$ are the same as functions on G right invariant by K. Since Δ commutes with right translation, it preserves right invariance by K. So Δ acts on functions on $\mathfrak{H}$, and it agrees with the Laplacian defined by (1).

$\boxed{4}$

The right invariance of Δ implies that Δ commutes with certain integral operators. These are easier to study because they are compact operators, in fact trace class.

Let $\mathcal{H}_\circ$ be the space of K-biinvariant, compactly supported smooth functions on G. It is a ring without unit under convolution.

Theorem 1 (Gelfand). *The ring $\mathcal{H}_\circ$ is commutative.*

Proof. Matrix transposition preserves K so it induces an involution ι on $\mathcal{H}_\circ$ such that $\iota(\phi * \psi) = \iota(\psi) \circ \iota(\phi)$, where $(\iota f)(g) = f(g^t)$. Every double coset in $K\backslash G/K$ has a diagonal representative. So ι is the identity map. ∎

This implies the representation of G induced from the trivial representation of K is multiplicity-free. Therefore (G, K) is called a *Gelfand pair.*

$\boxed{5}$

In this case K has a stronger property. Since K is commutative, any irreducible representation of K is a character χ, and the ring

$$\mathcal{H}_\chi = \{\phi \in C_c^\infty(G) | \phi(k_1 g k_2) = \chi(k_1)\, \phi(g)\, \chi(k_2)\}.$$

is commutative.

Thus K is a *strong multiplicity one subgroup* meaning that *every* irreducible representation of K induces a multiplicity free representation of G. This is a rarer property than Gelfand subgroup.

For example a maximal compact subgroup of a reductive algebraic group over a locally compact field F is a Gelfand subgroup. Rarely has it the strong multiplicity one property. In particular $SO(n)$ is a Gelfand subgroup of $SL(n, \mathbb{R})$ but only when $n = 2$ is it a strong multiplicity one subgroup. On the other hand there are a few important examples of strong multiplicity one subgroups: $GL(n-1)$ in $GL(n)$ and $SO(n-1)$ in $SO(n)$.

$\boxed{6}$

Any representation $\pi : G \to \mathrm{End}(V)$ (V a Banach or Frechet space) gives a representation of $\mathcal{H}_\circ$:

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

(Haar integral.) In particular the right regular representation ρ of G gives rise to a representation of $\mathcal{H}_\circ$ on $L^2(\Gamma\backslash G)$ which we will denote by T.

$$T_\phi f(x) = \int_G \phi(g)\, f(xg)\, dg.$$

This is convolution with $g \to \phi(g^{-1})$.

Let

$$K_\phi(x,y) = \sum_{\gamma\in\Gamma} \phi(x^{-1}\gamma y). \tag{4}$$

At first we regard (x,y) as an element of $G \times G$. Since ϕ is compactly supported for x and y restricted a compact set, only finitely many γ contribute.

$\boxed{7}$

A change of variables shows that $K_\phi(x,y)$ is invariant if either x or y is changed on the left by an element of γ, so we may regard the kernel as defined on either $G \times G$ or on $\Gamma\backslash G \times \Gamma\backslash G$, and it is a continuous function.

Theorem 2. *We have*

$$(T_\phi f)(x) = \int_{\Gamma\backslash G} K_\phi(x,y)\, f(y)\, dy. \tag{5}$$

Proof. The left side equals

$$\int_G \phi(y)\, f(xy)\, dy = \int_G \phi(x^{-1}y)\, f(y)\, dy = \sum_{\gamma\in\Gamma} \int_{\Gamma\backslash G} \phi(x^{-1}\gamma y)\, f(y)\, dy,$$

where we have used $f(\gamma y) = f(y)$. Interchanging sum and integral gives (5). ∎

$\boxed{8}$

Recall $X = \Gamma\backslash\mathfrak{H}$. The kernel K_ϕ is well defined as a function on $X \times X$.

If H is a Hilbert space, an operator $T : H \to H$ is *compact* if T maps bounded sets into compact sets.

Theorem 3. *T_ϕ is a compact operator.*

Proof. The kernel K_ϕ is continuous on the compact space $X \times X$, so it is certainly in $L^2(X \times X)$. The well-known theorem of Hilbert and Schmidt asserts that if X is any locally compact Borel measure space such that $L^2(X)$ is a separable Hilbert space then integral operator

$$(Tf)(x) = \int_X K(x,y)\, f(y)\, dy$$

with the kernel $K \in L^2(X \times X)$ is compact. ∎

$\boxed{9}$

If

$$\phi(g^{-1}) = \overline{\phi(g)} \tag{6}$$

then $K_\phi(x,y)$ is symmetric and T_ϕ is self adjoint.

Theorem 4 (Spectral theorem for compact operators). *Let T be a compact self-adjoint operator on a separable Hilbert space H. Then H has an orthonormal basis ϕ_i $(i = 1, 2, 3, \cdots)$ of eigenvectors of T, so that $T\phi_i = \mu_i \phi_i$. The eigenvalues $\mu_i \to 0$ as $i \to \infty$.* ∎

Thus if (6) is true then T is a self-adjoined compact operator whose nonzero eigenvalues μ_i of $T_\phi \to 0$. The Hilbert-Schmidt property implies more: $\sum |\mu_i|^2 < \infty$. Later we will see that more is true: $\sum |\mu_i| < \infty$. This means that T_ϕ is *trace class*. This fact is important because of the Selberg trace formula.

$\boxed{10}$

Theorem 5. *$L^2(X)$ has a basis consisting of eigenvectors of Δ.*

Proof. The operators T_ϕ with ϕ satisfying (6) are a commuting family of self-adjoint compact operators so they can be simultaneously diagonalized. By the spectral theorem the nonzero eigenspaces are finite-dimensional; there is no nonzero vector on which the operators T_ϕ are all zero, since ϕ can be chosen to be positive, of mass one and concentrated near the identity in which case $T_\phi f$ approximates f. Therefore the simultaneous eigenspaces of $\mathcal{H}_\circ$ are finite dimensional.

Let V be such an eigenspace. Since Δ commutes with the T_ϕ, it preserves V, and since it is symmetric it induces a self-adjoint transformation on V. Choose an orthonormal basis for each such V consisting of eigenvectors of Δ and put these together for all V. ∎

$\boxed{11}$

A point c is called a *regular singular point* for

(7) $\quad (r-c)^2\, g''(r) + (r-c)\, P(r-c)\, g'(r) + Q(r-c)\, g(r) = 0$

with $P(r-c)$ and $Q(r-c)$ analytic at $r = c$. Let α and β be the roots of the *indicial equation* $\alpha^2 + (P(0)-1)\alpha + Q(0) = 0$.

Proposition 1. *(i) If $\alpha - \beta$ is not an integer, then (7) has two solutions $(r-c)^\alpha\, g_1(r-c)$ and $(r-c)^\beta\, g_2(r-c)$, where g_1 and g_2 are analytic and nonzero at $r = c$.*
(ii) If $\alpha - \beta$ is a nonnegative integer, there exists a solution $(r-c)^\alpha\, g_1(r-c)$ and another

$$(r-c)^\beta g_2(r-c) + C\, \log(r-c)\, (r-c)^\alpha\, g_1(r-c),$$

with $g_i(r-c)$ analytic at $r = c$, and $g_2(0) = 0$ if and only if $\alpha = \beta$. ■

$\boxed{12}$

Let λ be a fixed complex number. We investigate K-bi-invariant functions on G such that $\Delta f = \lambda f$. Of course, such a function cannot be in $\mathcal{H}_\circ$ since it is not compactly supported. Since

$$G = \bigcup_{y \geq 1} K \begin{pmatrix} y^{1/2} & \\ & y^{-1/2} \end{pmatrix} K$$

f is determined by the function w on $(0,1)$ such that

$$w(r) = f\begin{pmatrix} y^{1/2} & \\ & y^{-1/2} \end{pmatrix},$$

where $r = (y-1)/(y+1)$. The eigenvalue property amounts to the differential equation

(8) $$w''(r) + \frac{1}{r} w'(r) + \frac{4\lambda}{(1-r^2)^2} w(r) = 0.$$

This has regular singular points at $(0,1)$ and there are two solutions of interest. One is nicely behaved at 0, the other at 1.

$\boxed{13}$

Let $\lambda = -s$ where s is any sufficiently large positive real number. At $r = 1$ the roots of the indicial equation are $\frac{1}{2}(1 \pm \sqrt{1+4s})$. Only one is positive, so there is a unique (up to multiple) solution g to (8) which vanishes near the boundary. We will use it to study the resolvant of the Laplacian.

Lemma 1. *$g(r)$ has a logarithmic singularity at $r = 0$.*

Proof. The roots of the indicial equation at $r = 0$ are 0 with multiplicity 2, so one solution has a logarithmic singularity, another is analytic. If g does not have the logarithmic singularity, then $g(r)$ is real and analytic on $[-1, 1]$ hence has a maximum or minimum. At such a point $g'(r) = 0$ and since $\lambda = -s < 0$ eq. (8) implies that g and g'' have the same sign, impossible at the maximum or minimum because $g(-1) = g(1) = 0$. ■

$\boxed{14}$

Let

$$g(z, \zeta) = g\left(\left|\frac{z-\zeta}{z-\bar{\zeta}}\right|\right).$$

$z, \zeta \in \mathfrak{H}$. This is a *Green's function.*

Theorem 6.

$$\left[-y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + s\right] g(z,\zeta) = 0;$$

$$g(z,\zeta) \textit{ is singular on the diagonal } z = \zeta;$$

$$g(z,\zeta) = g(\zeta, z);$$

$$g(z,\zeta) \to 0 \textit{ as } y \to 0;$$

$$g\big(h(z), h(\zeta)\big) = g(z,\zeta), \qquad h \in SL(2,\mathbb{R}). \tag{9}$$

Proof. These properties follow immediately from the properties of $g(z)$. ■

$\boxed{15}$

Since g has a logarithmic singularity at 0 it can be normalized so $g(r) - \frac{1}{2\pi}\log(r)$ is bounded as $r \to 0$. It follows from Proposition 1 that $g'(r) - \frac{1}{2\pi r}$ is analytic near $r = 0$.

Theorem 7. *If $f \in C_c^\infty(\mathfrak{H})$ then*

$$\int_{\mathfrak{H}} g(z,\zeta)\left[-\eta^2\left(\frac{\partial^2}{\partial\xi^2}+\frac{\partial^2}{\partial\eta^2}\right)+s\right] f(\zeta)\,\frac{d\xi\wedge d\eta}{\eta^2} = f(z).$$

Proof. Let $w = (z-\zeta)/(z-\overline{\zeta}) = u+iv \in \mathfrak{D}$. Let $F : \mathfrak{D} \to \mathbb{C}$ be defined by $F(w) = f(z)$. In the w coordinates we must prove

$$\int_{\mathfrak{D}} g(|w|)\left[-\left(\frac{\partial^2}{\partial u^2}+\frac{\partial^2}{\partial v^2}\right)+\frac{4s}{(1-|w|^2)^2}\right] F(w)\,du\wedge dv = F(0).$$

$\boxed{16}$

Let B_r be the disk of radius r, and let $1 < R$ be large enough that the support of F is contained in B_R. With $\Delta_\mathrm{e} = \partial^2/\partial u^2 + \partial^2/\partial v^2$ the left side equals

$$\lim_{\epsilon\to 0}\int_{B_R - B_\epsilon} g(|w|)\left[-\Delta_\mathrm{e}+\frac{4s}{(1-|w|^2)^2}\right] F(w)\,du\wedge dv.$$

Use Green's theorem (Stokes). For a planar domain Ω,

$$\int_\Omega (g\,\Delta^\mathrm{e} f \;-\; f\Delta^\mathrm{e} g)\,dx\wedge dy \;=\; \int_{\partial\Omega}\left(g\left(\frac{\partial f}{\partial x}\,dy - \frac{\partial f}{\partial y}\,dx\right) - f\left(\frac{\partial g}{\partial x}\,dy - \frac{\partial g}{\partial y}\,dx\right)\right).$$

Boundary is traversed counterclockwise.

$\boxed{17}$

We obtain

$$\lim_{\epsilon\to 0}\int_{B_R - B_\epsilon} F(w)\left[-\left(\frac{\partial^2}{\partial u^2}+\frac{\partial^2}{\partial v^2}\right)+\frac{4s}{(1-|w|^2)^2}\right]$$

$$g(|w|)\,du\wedge dv +$$

$$\lim_{\epsilon\to 0}\int_{C_\epsilon} F(w)\left(\frac{\partial g(|w|)}{\partial u}\,dv - \frac{\partial g(|w|)}{\partial v}\,du\right) -$$

$$\lim_{\epsilon\to 0}\int_{C_\epsilon} g(|w|)\left(\frac{\partial F(w)}{\partial u}\,dv - \frac{\partial F(w)}{\partial v}\,du\right),$$

where C_ϵ is the path circling the origin counterclockwise around the circle with radius ϵ. (There should also be terms integrating around C_R, but these are zero because they lie outside the support of F.)

The first term vanishes by Theorem 6. The last term vanishes because the length of the arc shrinks faster than g blows up (logarithmically).

$\boxed{18}$

To evaluate the middle term let $w = re^{i\theta}$.

$$\frac{\partial g(|w|)}{\partial u}\,dv - \frac{\partial g(|w|)}{\partial v}\,du = r\,g'(r)\,d\theta.$$

We obtain

$$\lim_{\epsilon\to 0}\int_0^{2\pi} F(\epsilon e^{i\theta})\,d\theta\,\epsilon\,g'(\epsilon) = F(0),$$

since $g'(\epsilon) \sim 1/(2\pi\epsilon)$ as $\epsilon \to 0$. ■

Lemma 2. *There exists a constant C such that for all $0 < r < 1$ the number of $\gamma \in \Gamma$ such that*

$$\frac{|z-\zeta|}{|z-\overline{\zeta}|} < r$$

is less than $Cr^2/(1-r^2)$.

$\boxed{19}$

Proof. The volume of the ball B_r of radius r with center at the origin in the noneuclidean metric (3.27) is $4\pi r^2/(1-r^2)$. Let F be a fundamental domain for the action of Γ. The union of the images under the Cayley transform $z \mapsto (z-\zeta)/(z-\overline{\zeta})$ of $\gamma(F)$, where γ runs through the elements of Γ such that $(\gamma z-\zeta)/(\gamma(z)-\overline{\zeta})$ roughly covers B_r, so the number of such γ is approximately $\mathrm{vol}(F)^{-1}\,4\pi r^2/(1-r^2)$. ■

Proposition 2. *The series*

$$G(z,\zeta) = \sum_{\gamma\in\{\pm 1\}\backslash\Gamma} g\big(z,\gamma(\zeta)\big) = \sum_{\gamma\in\{\pm 1\}\backslash\Gamma} g\big(\gamma(z),\zeta\big).$$

is absolutely convergent.

Proof. The typical term is

$$g\big(z, \gamma(\zeta)\big) = g\left(\frac{\gamma^{-1}z - \zeta}{\gamma^{-1}z - \overline{\zeta}}\right).$$

$\boxed{20}$

By Lemma 2, the number of γ such that

$$\frac{\gamma^{-1}z - \zeta}{\gamma^{-1}z - \overline{\zeta}} \in B_R$$

is of the order $R^2/(1 - R^2)$, and the function $g(r)$ is of the order $(1-r)^\alpha$ near $r = 1$, where $\alpha = \frac{1}{2}(1 \pm \sqrt{1+4s}) > 1$; consequently, the convergence issue reduces to the convergence of the integral

$$\int_0^1 (1-r)^\alpha \left(\frac{d}{dr}\frac{r^2}{1-r^2}\right) dr,$$

which is finite. ∎

$G(z, \zeta)$ is the *automorphic Green's function.* We will see that it is an integral kernel for the resolvant of the Laplacian.

$\boxed{21}$

Theorem 8. *G is defined and analytic for all values of (z, ζ) except where $\zeta = \gamma(z)$ for some $\gamma \in \Gamma$. We have*

$$\left[-y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + s\right] G(z, \zeta) = 0;$$

$$G\big(h(z), h(\zeta)\big) = G(z, \zeta), \qquad h \in G,$$

$$G(z, \zeta) = G\big(\gamma(z), \zeta\big) = G\big(z, \gamma(\zeta)\big), \qquad \gamma \in \Gamma,$$

$$G(z, \zeta) = \overline{G(\zeta, z)},$$

and $G(z, \zeta) \sim \frac{1}{2\pi}\log|z - \zeta|$ near $z = \zeta$. For $f \in C^\infty(\Gamma\backslash\mathfrak{H})$

$$(10) \qquad \int_{\Gamma\backslash\mathfrak{H}} G(z, \zeta)\left[-\eta^2\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}\right) + s\right] f(\zeta)\, \frac{d\xi \wedge d\eta}{\eta^2} = f(z).$$

$\boxed{22}$

Proof. Most of these properties follow from the corresponding properties of g. We prove (10). By a partition of unity argument we can find a function $F \in C_c^\infty(\mathfrak{H})$ such that $f(z) = \sum_{\gamma\in\Gamma} F(\gamma z)$. Substituting this and the definition of G and using (9) gives

$$\sum_{\gamma,\delta\in\Gamma} \int_{\Gamma\backslash\mathfrak{H}} g\big(\delta(z), \gamma(\zeta)\big) \left[-\eta^2 \left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2}\right) + s\right] F(\gamma(\zeta)) \frac{d\xi\wedge d\eta}{\eta^2}.$$

One of the summations may be collapsed with the integration to give

$$\sum_{\gamma,\delta\in\Gamma} \int_{\mathfrak{H}} g(\delta(z), \zeta) \left[-\eta^2 \left(\frac{\partial^2}{\partial\xi^2} + \frac{\partial^2}{\partial\eta^2}\right) + s\right] F(\zeta) \frac{d\xi\wedge d\eta}{\eta^2} = F(z).$$

This completes the proof. ∎

$\boxed{23\text{–}24}$

Theorem 9. *(i) The eigenvalues λ_i of Δ on $L^2(\Gamma\backslash\mathfrak{H})$ tend to ∞, and satisfy $\sum \lambda_i^{-2} < \infty$.*
(ii) The Laplacian Δ has an extension to a self-adjoint operator on the Hilbert space $L^2(\Gamma\backslash\mathfrak{H})$.

Proof. By Theorem 4 let ϕ_i be a basis of $H = L^2(\Gamma\backslash\mathfrak{H})$ consisting of eigenvectors of Δ, with corresponding eigenvalues λ_i.

Let $s > 0$ be a positive constant. As is easily checked, the logarithmic singularity along the diagonal is not sufficient to cause divergence of the integral

$$\int_{\Gamma\backslash\mathfrak{H}} \int_{\Gamma\backslash\mathfrak{H}} |G(z,\zeta)|^2 \frac{dx\wedge dy}{y^2} \frac{d\xi\wedge d\eta}{\eta^2} < \infty.$$

Thus the corresponding integral operator, which we shall denote T, is Hilbert-Schmidt. If ϕ is an eigenfunction of Δ with eigenvalue λ, then it follows from (10) that it is also an eigenfunction of T with eigenvalue $(\lambda+s)^{-1}$. Since T is Hilbert-Schmidt, $\sum_i (\lambda_i + s)^{-2} < \infty$, whence $\sum_i \lambda_i^{-2} < \infty$.

We prove (ii). Let $\mathfrak{D}_\Delta$ be the linear subspace of $\mathfrak{H}$ consisting of elements of the form $\sum a_i\phi_i$ such that $\sum \lambda_i^2 |a_i|^2 < \infty$; on this space, define

$$\Delta\Big(\sum a_i\phi_i\Big) = \sum \lambda_i\, a_i\, \phi_i.$$

Since the λ_i tend to infinity, and in particular are bounded away from zero, it is not hard to check that this operator is closed and in fact self-adjoint. This completes the proof of Theorem 3.11. ■

We have shown incidentally that if s is a sufficiently large real number, then $sI + \Delta$ has an inverse, which is a compact operator. This is the *resolvant* of the Laplacian.

$\boxed{25}$

We have already shown that the integral operators T_ϕ are Hilbert-Schmidt, hence compact. More is true: they are *trace class*. A compact operator is *trace class* if it can be factored as the composite of two Hilbert-Schmidt operators. If it is self-adjoint, and has eigenvalues λ_i, it is easy to see that this is equivalent to $\sum |\lambda_i| < \infty$. Lang's book $SL(2,\mathbb{R})$ contains useful material about trace class operators.

If $\phi \in \mathfrak{H}$ satisfies (6), let f_i be a basis of $L^2(\Gamma\backslash\mathfrak{H})$ consisting of eigenfunctions of T_ϕ which are also eigenfunctions of Δ. Let μ_i be the eigenvalues of T_ϕ with respect to this basis. Making a Fourier expansion we have

$$K_\phi(x,y) = \sum \mu_i f_i(x)\,\overline{f_i(y)}. \tag{11}$$

$\boxed{26}$

Theorem 10. *T_ϕ is trace class.*

Proof. A linear combination of trace class operators is trace class. Hence it is sufficient to prove this with ϕ replaced by $\frac{1}{2}\big(\phi(g)+\overline{\phi(g^{-1})}\big)$ and by $\frac{1}{2i}\big(\phi(g)-\overline{\phi(g^{-1})}\big)$. We may thus assume that ϕ satisfies (6) and so T_ϕ is self-adjoint. Let μ_i be its (nonzero) eigenvalues. Let λ_i be the corresponding eigenvalues of Δ. Thus $\sum \lambda_i^{-2} < \infty$.

Applying Δ to $K_\phi(x,y)$ in the first variable gives a new kernel

$$(\Delta_x K_\phi)(x,y) = \sum \mu_i \lambda_i f_i(x)\,\overline{f_i(y)}.$$

by (11). Since this function is continuous, it is Hilbert-Schmidt, and so we obtain the bound

$$\sum |\mu_i \lambda_i|^2 < \infty. \tag{12}$$

Now $\sum |\mu_i| < \infty$ follows from $\sum |\lambda_i|^{-2} < \infty$ and (12) by Cauchy-Schwarz. ■

$\boxed{27}$

The trace $\operatorname{tr} T$ of a self-adjoint trace class operator T is by definition the sum of its eigenvalues.

Theorem 11. *If ϕ satisfies (6), and if μ_i are the eigenvalues of T_ϕ, the trace*

$$\operatorname{tr} T_\phi = \int_{\Gamma\backslash\mathfrak{H}} K_\phi(z,z)\,\frac{dx\wedge dy}{y^2}. \tag{13}$$

Proof. This follows from orthonormality on integrating (11). ∎

The Selberg trace formula is a more explicit formula for its trace. Let $\{\gamma\}$ denote a set of representatives for the conjugacy classes of Γ. Let $Z_\Gamma(\gamma)$ denote the centralizer in Γ of γ.

$\boxed{28}$

Theorem 12. *We have*

$$\operatorname{tr} T_\phi = \sum_{\{\gamma\}} \int_{Z_\Gamma(\gamma)\backslash G} \phi(g^{-1}\gamma g)\,dg. \tag{14}$$

Proof. We rewrite the right side of (13) as

$$\sum_{\gamma\in\Gamma} \int_G \phi(g^{-1}\gamma g)\,dg = \sum_{\{\gamma\}} \sum_{\delta\in Z_\Gamma\backslash\Gamma} \int_G \phi\big(g^{-1}\,\delta^{-1}\,\gamma\,\delta\,g\big)\,dg.$$

Combining the integral and the summation gives (14). ∎

This is a primitive form of the trace formula. To make it more useful, we introduce the *spherical functions* ω_λ.

Theorem 13. *(i) Let $\lambda \in \mathbb{C}$. Then there is a unique smooth K-bi-invariant function ω_λ on $SL(2, \mathbb{R})$ such that $\Delta\omega_\lambda = \lambda\omega_\lambda$ and $\omega_\lambda(1) = 1$.*
(ii) If $f : G \to \mathbb{C}$ is any smooth function such that $\Delta f = \lambda f$ then

$$\int_{K\times K} f(kgk')\,dk\,dk' = f(1)\,\omega_\lambda(g). \tag{15}$$

(iii) If f is right K-invariant, then

$$\int_K f(hkg)\,dk = f(h)\,\omega_\lambda(g). \tag{16}$$

Proof. To satisfy $\Delta\omega = \lambda\omega$ we need

$$w(r) = \omega\begin{pmatrix} y^{1/2} & \\ & y^{-1/2}\end{pmatrix}, \qquad r = \frac{y-1}{y+1}$$

to satisfy (8). As we have seen, this differential equation has a regular singular point at the origin, and one solution is bounded there, whereas the other has a logarithmic singularity. Hence ω_λ, if it exists, is unique.

To show that such a function exists, let f be any continuous function on G which is an eigenfunction of Δ. Then (15) is a K-bi-invariant function which is an eigenfunction of Δ; if f is right K-invariant this is equivalent to (16). If $f(1) \neq 1$, this will satisfy (i), proving existence. For example we can take $f = f_s$ where $\lambda = s(1-s)$ and

$$f_s(g) = y^s. \tag{17}$$

Formulas (15) and (16) follow since the left sides in both equations are functions satifying the hypotheses of (i) except possibly the normalization, so these are constant multiples of ω_λ. The normalization constant can be determined in either case by taking $g = 1$. ∎

$\boxed{31}$

Theorem 14. *Suppose that f is a smooth function on G which is right invariant by K and such that $\Delta f = \lambda f$. Then for $\phi \in \mathcal{H}_\circ$ we have $T_\phi f = \chi(\phi) f$ where*

$$\chi(\phi) = \int_G \phi(g)\, \omega_\lambda(g)\, dg. \tag{18}$$

Proof. By Theorem 13, we have

$$\int_K f(hkg)\, dk = f(h)\, \omega_\lambda(g). \tag{19}$$

We note that $T_\phi f$ is an average of right translates of f, and right translation commutes with left translation. Hence we may apply T_ϕ to both sides of (19) to obtain

$\boxed{32}$

$$\int_K (T_\phi f)(hkg)\, dk = f(h)\, (T_\phi \omega_\lambda)(g).$$

We take $g = 1$ in this identity. Since $T_\phi f$ is right K-invariant, the integrand on the left side becomes constant when $g = 1$ and so the left side becomes just $(T_\phi f)(h)$. On the other hand $(T_\phi \omega_\lambda)(1)$ equals the integral (18), so $T_\phi f(h) = \chi(\phi)\, f(h)$. ■

The function $\chi : \mathcal{H}_\circ \to \mathbb{C}$ is a *character* of $\mathcal{H}_\circ$.

Theorem 15. *If ϕ_1 and $\phi_2 \in \mathcal{H}_\circ$, then*

$$\chi(\phi_1 * \phi_2) = \chi(\phi_1)\, \chi(\phi_2).$$

Proof. This follows from Theorem 14 applying T_ϕ to any f, for example f_s as in (17). ■

$\boxed{33}$

To make this more explicit, we introduce

$$g(u) = e^{u/2} \int_{-\infty}^{\infty} \phi\left(\begin{pmatrix} e^{u/2} & \\ & e^{-u/2} \end{pmatrix} \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \right) dx, \tag{20}$$

and let

$$h(t) = \int g(u)\, e^{iut}\, dt \tag{21}$$

be its Fourier transform.

Theorem 16. *The functions g and h are even, and g is compactly supported. If $\lambda = \frac{1}{4} + t^2$, then*

$$\chi(\phi) = h(t). \tag{22}$$

Proof. Let f_s be as in (17) with $s = \frac{1}{2} + it$ so that $\lambda = \frac{1}{4} + t^2 = s(1-s)$. By Theorem 13,

$$\int_K f_s(kg) = \omega_\lambda(g),$$

$\boxed{34}$

and

$$\chi(\phi) = \int_G \int_K \phi(g)\, f_s(gk)\, dk\, dg.$$

Interchanging the order of integration and making the variable change $g \to gk^{-1}$, since ϕ is right k invariant, we obtain

$$\chi(\phi) = \int_G \phi(g)\, f_s(g)\, dg.$$

Now we use the coordinates

$$g = \begin{pmatrix} e^{u/2} & \\ & e^{-u/2} \end{pmatrix} \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \kappa_\theta, \tag{23}$$

$$\kappa_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}, \qquad dg = \frac{1}{2\pi}\, du\, dx\, d\theta.$$

Noting that in these coordinates $f_s(g) = e^{u/2}\, e^{iut}$, we obtain

$$\chi(\phi) = \int_\infty^\infty g(u)\, e^{itu}\, du,$$

$\boxed{35\text{–}36}$

proving (22). We note that ω_λ, and therefore the character $\chi(\phi)$ is unchanged if $s \to 1-s$, that is, if $t \to -t$. Hence (22) implies that h is an even function. By Fourier inversion, so is g. ■

Theorem 17. *We have*

$$\phi(1) = \frac{1}{4\pi} \int_{-\infty}^{\infty} t\, h(t)\ \tanh(\pi t)\, dt.$$

This is the "Plancherel formula", essentially the Fourier inversion formula on the noncommmutative group $SL(2, \mathbb{R})$.

Proof omitted. For proof see Gelfand, Graev and Piatetski-Shapiro [GGP], Chapter 2 Section 6 and Varadarajan [V] Theorem 39 on p. 205. ■

We assume now that Γ is a hyperbolic group. If $1 \neq \gamma \in \Gamma$ define $N = N(\gamma)$ by asking that γ be conjugate to

$$\begin{pmatrix} N^{1/2} & \\ & N^{-1/2} \end{pmatrix}. \tag{24}$$

for some N. Let $N_0 = N_0(\gamma)$ be such that

$$\begin{pmatrix} N_0^{1/2} & \\ & N_0^{-1/2} \end{pmatrix}$$

is conjugate to a generator of $Z_\Gamma(\gamma)$. We may obviously assume that N and N_0 are > 1. We note that $Z_G(\gamma)$ is conjugate to the diagonal subgroup. Its image in X is a closed geodesic. So the numbers $N(\gamma)$ are thus the lengths of closed geodesics, and the numbers $N_0(\gamma)$ are the lengths of prime geodesics.

$\boxed{37}$

Theorem 18. *With g the function in (20),*

$$\int_{Z_\Gamma(\gamma)\backslash G} \phi(g^{-1}\gamma g)\, dg = \frac{\log\ N_0}{N^{1/2} - N^{-1/2}}\, g(\log N). \tag{25}$$

Proof. We may assume that γ equals (24). Using Iwasawa coordinates (23) the integral is

$$\int_0^{\log N_0} du \int_{-\infty}^{\infty} \phi\left(\begin{pmatrix} 1 & -x \\ & 1 \end{pmatrix} \begin{pmatrix} N^{1/2} & \\ & N^{-1/2} \end{pmatrix} \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} \right) dx =$$
$$|\log(N_0)| \int_{-\infty}^{\infty} \phi\left(\begin{pmatrix} N^{1/2} & \\ & N^{-1/2} \end{pmatrix} \begin{pmatrix} 1 & (1-N^{-1})x \\ & 1 \end{pmatrix} \right) dx,$$

and a change of variables proves (25). ■

$\boxed{38}$

Theorem 19 (Selberg trace formula). *Let g be a smooth, even, compactly supported function and let h be its Fourier transform, defined by (21). If $\frac{1}{2}+it_i$ are the eigenvalues of Δ on $\mathfrak{H}$, and if $\log(N)$ runs through the lengths of closed geodesics of Γ, where for each N we let N_0 be the length of the corresponding prime geodesic, we have*

$$\sum h(t_i) = \frac{1}{4\pi}\int_{-\infty}^{\infty} t\,h(t)\,\tanh(\pi t)\,dt\,\mathrm{vol}(\Gamma\backslash\mathfrak{H}) \\ + \sum_N \frac{\log\,N_0}{N^{1/2}-N^{-1/2}}\,g(\log\,N).$$

Proof. Theorems 12, 17 and 18. ■

$\boxed{39}$

With similar hypotheses on g and h, let $\frac{1}{2}+t_i$ be the zeros of $\zeta(s)$. Weil proved (following similar formulae of Riemann, von Mangoldt and Ingham):

$$\sum h(t_i) = h\left(\frac{i}{2}\right) + h\left(-\frac{i}{2}\right) + \frac{1}{2\pi}\int_{-\infty}^{\infty} h(r)\frac{\Gamma'\left(\frac{1+2ir}{4}\right)}{\Gamma\left(\frac{1+2ir}{4}\right)}\,dr \\ -2\sum\frac{\log(p)}{\sqrt{p^n}}\,g\big(\log(p^n)\big).$$

Motivated by the mysterious resemblence of the Selberg trace formula to this explicit formula, Selberg introduced

$$Z(s) = \prod_{\{\gamma\}\ \mathrm{primitive}}\ \prod_{k=0}^{\infty}(1-N_0(\gamma)^{-s-k}).$$

Its zeros are s with $s(1-s)$ an eigenvalue of Δ, so $Z(s)=0$ implies $s=\frac{1}{2}+it$, t real or $s\in[0,1]$.

$\boxed{40}$

Let

$$R = \frac{1}{2}\begin{pmatrix} 1 & i \\ i & -1\end{pmatrix}, \qquad L = \frac{1}{2}\begin{pmatrix} 1 & -i \\ -i & -1\end{pmatrix},$$

(26)
$$H = \frac{1}{i} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

They are elements of the complexified Lie algebra $\mathfrak{g}_{\mathbb{C}}$. Note that $iH \in \mathrm{Lie}(K)$. These elements satisfy the same relations as $\hat{H}$, $\hat{R}$ and $\hat{L}$ defined in (2):

$$[H, R] = 2R, \qquad [H, L] = -2L, \qquad [R, L] = H.$$

This is no coincidence, for they are conjugate to $\hat{H}$, $\hat{R}$ and $\hat{L}$ in the complexified group $SL(2, \mathbb{C})$—in fact, the conjugating element (interpreted as a linear fractional transformation of the Riemann sphere) is just the Cayley transform which we've already applied to map the upper half plane into the disk.

$\boxed{41}$

The action (3) of $\mathfrak{g}$ extends by linearity to $\mathfrak{g}_{\mathbb{C}}$, and as differential operators

$$R = e^{2i\theta} \left(iy \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} + \frac{1}{2i} \frac{\partial}{\partial \theta} \right),$$

$$L = e^{-2i\theta} \left(-iy \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} - \frac{1}{2i} \frac{\partial}{\partial \theta} \right),$$

and
$$H = \frac{1}{i} \frac{\partial}{\partial \theta},$$

in the Iwasawa coordinates

$$g = \begin{pmatrix} u & \\ & u \end{pmatrix} \begin{pmatrix} y^{1/2} & xy^{-1/2} \\ & y^{-1/2} \end{pmatrix} \kappa_\theta,$$

$$\kappa_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

$\boxed{42}$

Let k be an integer, and let $C^\infty(G, k)$ be the subset of elements of $C^\infty(G)$ satisfying
$$F(g\kappa_\theta) = e^{ik\theta} F(g).$$

Since $H = \frac{1}{i} \frac{\partial}{\partial \theta}$, the element H acts by the scalar k on $C^\infty(G, k)$.

Let Γ be a cocompact discrete subgroup of $SL(2,\mathbb{R})$. We may assume that $-I \in \Gamma$. Let $\chi : \Gamma \to \mathbb{C}^\times$ be a unitary character. Let $C^\infty(\Gamma\backslash G,\chi)$ be the space of $f \in C^\infty(G)$ satisfying

$$f(\gamma g) = \chi(\gamma)\, f(g)$$

for $\gamma \in \Gamma$. Let k be such that $\chi(-I) = (-1)^k$. Let

$$C^\infty(\Gamma\backslash G,\chi,k) = C^\infty(\Gamma\backslash G,\chi) \cap C^\infty(G,k).$$

Since this is characterized as the k-eigenspace of H on $C^\infty(\Gamma\backslash G,\chi)$, $[H,R] = 2R$ and $[H,L] = -2L$ imply that R and L shift $C^\infty(\Gamma\backslash G,\chi,k)$ into $C^\infty(\Gamma\backslash G,\chi,k+2)$ and $C^\infty(\Gamma\backslash G,\chi,k-2)$.

$\boxed{43}$

On the other hand let $C^\infty(\Gamma\backslash\mathfrak{H},\chi,k)$ be the space of smooth functions on $\mathfrak{H}$ such that

$$\chi(\gamma)\, f(z) = \left(\frac{c\overline{z}+d}{|cz+d|}\right)^k f\left(\frac{az+b}{cz+d}\right),$$

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma.$$

If $f, g \in C^\infty(\Gamma\backslash\mathfrak{H},\chi,k)$, then $f\,\overline{g}$ is invariant under Γ, and so we may define

$$\langle f, g\rangle = \int_{\Gamma\backslash\mathfrak{H}} f(z)\,\overline{g(z)}\,\frac{dx\,dy}{y^2}.$$

Let $L^2(\Gamma\backslash\mathfrak{H},\chi,k)$ denote the Hilbert space completion of $C^\infty(\Gamma\backslash\mathfrak{H},\chi,k)$ with respect to this inner product. If $f \in L^2(\Gamma\backslash\mathfrak{H},\chi,k)$, then k is called its *weight*. Let

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

$\boxed{44}$

One may check Δ_k preserves $L^2(\Gamma\backslash\mathcal{H},\chi,k)$; we will soon see the reason for this. An eigenform of Δ_k on this space is called a *Maass form of weight* k.

For example, the trace formula shows that Maass forms of weight zero exist. Another, very special class of Maass forms is obtained from holomorphic modular forms. Let f be a holomorphic function on $\mathfrak{H}$ such that

$$\chi(\gamma)\, f(z) = (cz+d)^{-k}\, f\left(\frac{az+b}{cz+d}\right)$$

for $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \Gamma$. Then $y^{k/2}\, f(z)$ is a Maass form of weight k. The eigenvalue of Δ_k is $\lambda = \frac{k}{2}\left(1 - \frac{k}{2}\right)$. The Riemann Roch theorem guarantees that holomorphic modular forms exist if k is sufficiently large and of the correct parity, i.e. $\chi(-I) = (-1)^k$.

$\boxed{45}$

Theorem 20. *The spaces $C^\infty(\Gamma\backslash G, \chi, k)$ and $C^\infty(\Gamma\backslash\mathfrak{H}, \chi, k)$ are isomorphic. Indeed, an isomorphism is given by*

$$\sigma_k : C^\infty(\Gamma\backslash\mathfrak{H}, \chi, k) \to C^\infty(\Gamma\backslash G, \chi, k)$$

given by

$$(\sigma_k f)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \left(\frac{-ci+d}{|ci+d|}\right)^k f\left(\frac{ai+b}{ci+d}\right)$$

for $f \in C^\infty(\Gamma\backslash\mathfrak{H}, \chi, k)$. ■

The importance of this theorem is that we may exchange the classical spaces of Maass forms of weight k, which are right invariant by K but have a cocycle built into their multiplier systems, for functions with simpler multipliers (just χ) on which K acts by a character.

$\boxed{46}$

By Theorem 20 we should be able to transfer R, L and Δ to the spaces of Maass forms. Define the following *Maass differential operators* on $C^\infty(\Gamma\backslash\mathfrak{H}, k)$:

$$R_k = iy\,\frac{\partial}{\partial x} + y\,\frac{\partial}{\partial y} + \frac{k}{2},$$

$$L_k = -iy\,\frac{\partial}{\partial x} + y\,\frac{\partial}{\partial y} - \frac{k}{2},$$

and the *(weight k) noneuclidean Laplacian*

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

It is easily checked that

$$\Delta_k = -L_{k+2}\,R_k - \frac{k}{2}\left(1+\frac{k}{2}\right) = -R_{k-2}\,L_k + \frac{k}{2}\left(1-\frac{k}{2}\right).$$

$\boxed{47}$

Of course these operators are simply the Lie operators L, R and Δ in $U(\mathfrak{g}_{\mathbb{C}})$, whose actions on $C^\infty(\Gamma\backslash G, \chi)$ are transferred to the spaces of Maass forms by

Theorem 20. Thus R shifts the weight up by 2, and L shifts it down by 2. In general the weight can by shifted as much as you like in either direction, with one exception: it is a consequence of the Cauchy-Riemann equation that L_k kills $y^{k/2} f(z) \in C^\infty(\Gamma\backslash\mathfrak{H}, \chi, k)$ where f is a holomorphic modular form of weight k. Similarly, R kills certain vectors in $C^\infty(\Gamma\backslash\mathfrak{H}, \chi, -k)$ if $k > 0$; these come from the complex conjugates of holomorphic modular forms.

We will leave the Maass forms with weight and holomorphic modular forms aside and study the spectral theory of $\Gamma\backslash G$. Most aspects of the spectral theory which we have discussed transfer without much change to weight k. But considering functions on G instead of $\mathfrak{H}$ lets us introduce a new perspective of representation theory.

$\boxed{48\text{–}49}$

Let $\mathcal{H} = C_c^\infty(G)$. It is a ring under convolution. Unlike $\mathcal{H}$ it is not commutative.

Lemma 3. *If $0 \neq f \in L^2(\Gamma\backslash G, \chi)$ there exists $\phi \in \mathcal{H}$ such that T_ϕ is self-adjoint and $T_\phi f \neq 0$.*

Proof. By taking $\phi(g)\,dg$ to be a probability distribution concentrated near the identity we may make $T_\phi f$ as near f as we like. We may make ϕ symmetric with respect to $g \to g^{-1}$ so T_ϕ is self-adjoint. ■

Theorem 21. *Let H be a nonzero G-invariant subspace of $L^2(\Gamma\backslash G, \chi)$. Then H contains an irreducible subspace.*

Proof. (Langlands.) By the Lemma we can find ϕ such that T_ϕ is nonzero on H. Let $L \subset H$ be the eigenspace of a nonzero eigenvalue. It is finite-dimensional by Theorem 4. Let L_0 be a nonzero subspace minimal with respect to the being the intersection of L with a nonzero closed invariant subspace. Let V be the smallest closed invariant subspace such that $L \cap W = L_0$. We show V is irreducible. If not, $V = V_1 \oplus V_2$. Let $0 \neq f \in L_0$. Write $f = f_1 + f_2$ with $f_i \in V_i$. Since $0 = T_\phi f - \lambda f = (T_\phi f_1 - \lambda f_1) + (T_\phi f_2 - \lambda f_2)$ and $T_\phi f_i - \lambda f_i \in V_i$ we have $T_\phi f_i - \lambda f_i = 0$. Thus $f_i \in L \cap V_i$. By the minimality of L_0, $L_0 = L \cap V_i$ for some i, say $L_0 = L \cap V_1$. Now the minimality of V is contradicted. ■

Theorem 22. *$L^2(\Gamma\backslash G, \chi)$ decomposes as a direct sum of closed, irreducible subspaces.*

Proof. By Zorn, let S be a maximal set of orthogonal closed irreducible subspaces. Let $H = \bigoplus_{V \in S} V$. If H is proper, applying Theorem 21 to its orthogonal complement contradicts the maximality of S. ■

$\boxed{50\text{–}51}$

If (π, V) is an irreducible representation of $SL(2,\mathbb{R})$, decompose it as

$$V = \bigoplus V(k)$$

where K acts as the character $\kappa_\theta \to e^{ik\theta}$. Thus with H as in (26), $V(k)$ is the k-eigenspace of H. If $\pi(-I)$ acts by 1, only even k appear; if $\pi(-I)$ acts by -1, only odd k appear. Let $\epsilon = 0$ or 1 be the parity of $\pi(-I)$.

Theorem 23. *Δ acts by a scalar λ on V. R maps $V(k)$ into $V(k+2)$ and similarly L maps $V(k)$ into $V(k-2)$. If $x \in V(k)$ then*

$$LRx = \left(-\lambda - \tfrac{k}{2}(1 + \tfrac{k}{2})\right) x, \tag{27}$$

$$RLx = \left(-\lambda + \tfrac{k}{2}(1 - \tfrac{k}{2})\right) x.$$

Proof. Since $[H, R] = 2R$, if $x \in V(k)$ then $HRx = 2Rx + RHx = (k+2)Rx$. Since Δ is in the center of $U(\mathfrak{g})$ it acts by a scalar λ. From

$$-4\Delta = H^2 + 2H + 4LR = H^2 - 2H + 4RL$$

we obtain (27). ■

Given a representation (π, V), the *algebraic* direct sum of the $V(k)$ is not invariant by G but it has compatible actions of $\mathfrak{g}$ and K and is called a $(\mathfrak{g}, K)$*-module*. Two representations with isomorphic $(\mathfrak{g}, K)$-modules they are *infinitessimal equivalence* and essentially the same. Let

$$\{k \in \mathbb{Z} | V(k) \neq 0\}$$

be called the set of *weights* or *K-types* of π.

Theorem 24. *If λ is not of the form $\frac{l}{2}\left(1 \pm \frac{l}{2}\right)$, where l is an integer of the same parity as ϵ, there is a unique infinitessimal equivalence class of π with Δ and $\pi(-I)$ acting by λ and $(-1)^\epsilon$. The set of K-types of π consists of all integers of parity ϵ.*

$\boxed{52}$

Proof. (27) shows that R and L do not annihilate $V(k)$, so all K-types appear. There's enough information in Theorem 23 to recover the Lie algebra action, so π is unique. For existence, find s so $s(1-s)=\lambda$. Let

$$f_{s,k}\left(\begin{pmatrix} y^{1/2} & xy^{-1/2} \\ & y^{-1/2} \end{pmatrix} \kappa_\theta\right) = y^s\, e^{ik\theta}, \tag{28}$$

$y>0$. Then

$$P(s,\epsilon)_{\mathrm{fin}} = \bigoplus_{k\equiv\epsilon (\mathrm{mod}\, 2)} \mathbb{C} f_{s,k}$$

is an irreducible $(\mathfrak{g},K)$-module realizing λ, ϵ. ■

The *principal series representation* $P(s,\epsilon)$ is constructed by induction from the Borel subgroup. It consists of all functions $f: G \to \mathbb{C}$ such that

$$f\left(\begin{pmatrix} y^{1/2} & xy^{-1/2} \\ & y^{-1/2} \end{pmatrix} g_\theta\right) = (-1)^{\epsilon\,\mathrm{sgn}(y)}\, y^s f(g),$$

and such that the restriction of f to K is square integrable. Its $(\mathfrak{g},K)$-module is $P(s,\epsilon)_{\mathrm{fin}}$.

$\boxed{53}$

Since $P(s,\epsilon)$ and $P(1-s,\epsilon)$ correspond to the same value $\lambda = s(1-s)$ of Δ, if $2s$ is not an integer congruent to ϵ modulo 2, then Theorem 24 implies that $P(s,\epsilon)$ and $P(1-s,\epsilon)$ are infinitessimally equivalent. Suppose first $\mathrm{re}(s) > \frac{1}{2}$. An intertwining map $M(s): P(s,\epsilon)_{\mathrm{fin}} \to P(1-s,\epsilon)$ is given by

$$\big(M(s)f\big)(g) = \int_{-\infty}^{\infty} f\left(\begin{pmatrix} & -1 \\ 1 & \end{pmatrix}\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} g\right) dg. \tag{29}$$

It may be checked that the integral is absolutely convergent, and that

$$M(s)\, f_{s,k} = (-i)^k \sqrt{\pi}\, \frac{\Gamma(s)\,\Gamma\big(s-\frac{1}{2}\big)}{\Gamma\big(s+\frac{k}{2}\big)\,\Gamma\big(s-\frac{k}{2}\big)}\, f_{1-s,k}.$$

This formula extends $M(s)$ to an intertwining map for all s such that $2s$ is not an integer congruent to ϵ mod 2.

$\boxed{54}$

If we embed $P(s,\epsilon)$ into $L^2(\Gamma\backslash G.\chi)$, where $\chi(-I) = (-1)^\epsilon$, then the image of $V(k)$ function corresponding to a Maass form of weight k by Theorem 20. The representation $P(s,\epsilon)$ obviously must be unitary, and this implies that either $s = \frac{1}{2} + it$, where t is real, or that $\epsilon = 0$ and s is real, $s \in (0,1)$. The first class of unitary representations are called the *unitary principal series*, the second the *complementary series*.

If $k > 0$, $k \equiv \epsilon \pmod 2$ and $\lambda = \frac{k}{2}\left(1 - \frac{k}{2}\right)$, then (27) is consistent with L vanishing on $V(k)$. In this case $P(k/2,\epsilon)$ has a subrepresentation $D^+(k)$ whose set of K-types is $\{k, k+2, k+4, \cdots\}$; it has another whose $D^-(k)$ set of K-types is $\{-k, -k-2, -k-4, \cdots\}$. If $k \geq 2$ the representations $D^\pm(k)$ are unitary, in fact square integrable. These are the *discrete series*. If we embed $D^+(k)$ into $L^2(\Gamma\backslash G, \chi)$, where $k \geq 2$, then the image of a vector in the minimal K-type $V(k)$ corresponds to a modular form of weight k.

$\boxed{55}$

If the group Γ is of cofinite volume but has a cusp, the spectral theory is complicated by a continuous spectrum, coming from the Eisenstein series. However the cuspidal spectrum is discrete. For simplicity assume that ∞ is the only cusp of Γ, and the stabilizer of ∞ in Γ is

$$\Gamma_\infty = \left\{ \begin{pmatrix} 1 & n \\ & 1 \end{pmatrix} \middle| n \in \mathbb{Z} \right\}.$$

We may choose a fundamental domain $\mathcal{F}$ of Γ contained in a "Siegel set"

$$\mathcal{S} = \left\{ \begin{pmatrix} y^{1/2} & xy^{-1/2} \\ & x^{-1/2} \end{pmatrix} \kappa_\theta \middle| |x| \leq \tfrac{1}{2}, |y| > C \right\}$$

For some C. For $\Gamma = SL(2,\mathbb{Z})$, we may take $C = \sqrt{3}/2$.

Let $L^2_0(\Gamma\backslash G)$ be the subspace of "cusp forms" $f \in L^2(\Gamma\backslash G)$ satisfying

$$\int_0^1 f\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} g\right) = 0. \tag{30}$$

$\boxed{56}$

Following Godement, let $f \in L^2_0(\Gamma\backslash G)$.

$$T_\phi f(g) \;=\; \int_G \phi(h)\, f(gh)\, dh \;=\; \int_G \phi(g^{-1}h)\, f(h)\, dh \;=\; \int_{\Gamma_\infty\backslash G} K(g,h)\, f(h)\, dh,$$

$$K(g,h) = \sum_{n\in\mathbb{Z}} \Phi_{g,h}(n) = \sum_{n\in\mathbb{Z}} \hat{\Phi}_{g,h}(n),$$

by Poisson summation, where

$$\Phi_{g,h}(t) = \phi\left(g^{-1}\begin{pmatrix}1 & t\\ & 1\end{pmatrix} h\right).$$

Since $\hat{\Phi}_{g,h}(0)$ is unchanged when $g \mapsto \left(\begin{smallmatrix}1 & t\\ & 1\end{smallmatrix}\right) g$, cuspidality implies

$$\int_{\Gamma_\infty\backslash G} \hat{\Phi}_{g,h}(0)\, f(g) = 0.$$

$\boxed{57}$

Hence we may omit this term from $K(g,h)$ without harm and

$$T_\phi f(g) = \int_{\Gamma_\infty\backslash G} K_0(g,h)\, f(h)\, dh,$$

(31) $$K_0(g,h) = \sum_{n\neq 0} \hat{\Phi}_{g,h}(n).$$

Let

$$g = \begin{pmatrix}1 & x\\ & 1\end{pmatrix}\begin{pmatrix}y^{1/2} & \\ & y^{-1/2}\end{pmatrix}\kappa_\theta,$$

$$h = \begin{pmatrix}1 & u\\ & 1\end{pmatrix}\begin{pmatrix}v^{1/2} & \\ & v^{-1/2}\end{pmatrix}\kappa_\sigma.$$

Then

(32) $$|\hat{\Phi}_{g,h}(t)| = |y\,\hat{\psi}_{\theta,\sigma,y^{-1}v}(yt)|,$$

$$\psi_{\theta,\sigma,w}(t) = \phi\left(\kappa_\theta^{-1}\begin{pmatrix}1 & t\\ & 1\end{pmatrix}\begin{pmatrix}w^{1/2} & \\ & w^{-1/2}\end{pmatrix}\kappa_\sigma\right).$$

$\boxed{58}$

We say that a function f on $\mathcal{S}$ is of *rapid decay* if f and is $\ll |y|^{-N}$ on $\mathcal{S}$ for every N, and if Df has the same property for every $D \in U(\mathfrak{g})$. We say that f is of *moderate growth* if it is $\ll |y|^{N}$ for some N, and if Df has the same property for every $D \in U(\mathfrak{g})$. We say that f is *K-finite* if it is a finite linear combination of functions f_k where $f_k(g\kappa_\theta) = e^{ik\theta} f_k(g)$. We say f is *Δ-finite* if it is a finite linear combination of eigenvalues of Δ.

If $X \in \mathfrak{g}$ then $X(T_\phi f) = T_{\phi_X} f$ where

$$\phi_X(g) = \frac{d}{dt}\phi(e^{-tX}g)\,\Big|_{t=0}, \tag{33}$$

so all conclusions we draw about $T_\phi f$ will apply equally to its derivatives $D(T_\phi f)$ with $D \in U(\mathfrak{g})$.

We consider the behavior of $K_0(g,h)$ when g is restricted to the $\mathcal{S}$. This means that $y > C$. Since ϕ is compactly supported, (32) shows that $K_0(g,h)$ vanishes unless $y^{-1}v$ is restricted to a compact set.

$\boxed{59}$

Recall that the Fourier transform of a smooth compactly supported function on $\mathbb{R}$ is of Schwartz class. We have proved that

$$|K_0(g,h)| \leq \sum_{n \neq 0} |y\,\hat{\psi}_{\theta,\sigma,y^{-1}v}(yn)|,$$

where $(\theta,\sigma,w,t) \mapsto \psi_{\theta,\sigma,w}(t)$ is compactly supported and smooth on $[0,2\pi] \times [0,2\pi] \times \mathbb{R}^\times \times \mathbb{R}$. Consequently as the series defining $K_0(g,h)$ converges absolutely and uniformly to a function which as a function of g is of rapid decay on $\mathcal{S}$. Moreover if y is restricted to $\mathcal{S}$ the set of possible v for which $K_0(g,h) \neq 0$ is bounded below.

If X is a compact Hausdorff space, the ring $C(X)$ of continuous functions on X is a closed subspace of $L^\infty(X)$. If Σ is a subset of $C(X)$, then Σ is called *equicontinuous* if for any $\epsilon > 0$ and for any $x \in X$ there exists a neighborhood N of x such that $|f(y) - f(x)| < \epsilon$ for all $y \in N$, $f \in \Sigma$.

$\boxed{60\text{–}62}$

The Ascoli-Arzéla Lemma. *Let Y be a compact Hausdorff space, and let $\Sigma \subset C(Y)$ be an equicontinuous set which is bounded in the L^∞ norm. Then the closure of Σ in $C(Y)$ is compact in $L^\infty(X)$.* ■

Theorem 25 (Gelfand, Graev and Piatetski-Shapiro). *If $\phi \in \mathcal{H}$, the restriction of T_ϕ to $L^2_0(\Gamma\backslash G)$ is a compact operator.*

Proof. (Godement.) We compactify $\Gamma\backslash G$ by adjoining a single point at infinity corresponding to the cusp. The decay and support properties of $K_0(g,h)$ imply an estimate of the form

$$\sup_{g\in\Gamma\backslash G} |T_\phi\, f(g)| \le C\, ||f||_1;$$

since $\mathcal{S}$ contains a fundamental domain for Γ it is sufficient to prove this for $g \in \mathcal{S}$. This in turn is dominated by $||f||_2$ since $\mathcal{S}$ is of finite volume. We extend $T_\phi f$ to the compactification of $\Gamma\backslash G$ by making it zero at the cusp. The image Σ of the unit ball in $L^2(\Gamma\backslash G)$ is therefore bounded in $L^2(\Gamma\backslash G)$. In fact, Σ is equicontinuous, because we can bound the derivatives of its elements uniformly using (33). It is thus compact in $L^\infty(\Gamma\backslash G)$ by the Ascoli-Arzéla Lemma and hence compact in $L^2(\Gamma\backslash G)$. ∎

Theorem 26. *$L_0^2(\Gamma\backslash G)$ decomposes as a direct sum of irreducible invariant subspaces. Any K-finite element of one of these spaces is of rapid decay.*

Proof. The proof of Theorems 21 and 22 are easily adapted to give compactness in the present context. For the second assertion, it is sufficient to show that if $V \subset L_0^2(\Gamma\backslash G)$ is an irreducible subspace and $f \in V(k)$ then f is of rapid decay. We may find $\phi \in \mathcal{H}$ satisfying $\phi(\kappa_\theta g \kappa_\sigma) = e^{ik(\theta+\sigma)}\phi(g)$, and such that

$$y \to \phi\begin{pmatrix} y^{1/2} & \\ & y^{-1/2} \end{pmatrix}$$

is a positive function of mass 1 concentrated near $y = 1$. Then $T_\phi f$ is near f, therefore nonzero, and it is in $V(k)$, which is one-dimensional, so it is proportional to f. It follows from the rapid decay of $K_0(g,f)$ that $T_\phi f$ is of rapid decay, thus so is f. ∎

As in the case of compact quotient, the Laplacian acts by scalars on each on irreducible one-dimensional subspace. The cuspidal spectrum behaves much as the entire spectrum in the noncompact case.

$\boxed{63}$

On the other hand the orthogonal complement of $L_0^2(\Gamma\backslash G)$ contains a continuous spectrum. To gain intuition, we consider a simpler example.

Example of a continuous spectrum. The group $\mathbb{R}$ acts on itself by conjugation, and the Laplacian $-d^2/dx^2$ is an invariant differential operator. It has eigenfunctions $f_a(x) = e^{2\pi i a x}$ with eigenvalues a^2. Any L^2 function has a Fourier expansion

$$\phi(x) = \int_{-\infty}^{\infty} \hat{\phi}(a)\, f_a(x)\, da,$$

but f_a is itself not L^2. If $T \subset \mathbb{R}$ is measurable, the Fourier transforms of L^2 functions supported on T form an invariant subspace. There are no minimal invariant subspaces, so $L^2(\mathbb{R})$ doesn't decompose as a direct sum of irreducible representations.

$\boxed{64}$

We consider the Eisenstein series of weight zero, which occur in the spectrum of $L^2(\Gamma\backslash\mathfrak{H})$. Define

$$f_s(x+iy) = y^s.$$

Or transferring this function to the group,

$$f_s\left(\begin{pmatrix} y^{1/2} & xy^{-1/2} \\ & y^{-1/2} \end{pmatrix} \kappa_\theta\right) = y^s.$$

It is the K-fixed vector in the principal series representation $P(s,0)$. The *Eisenstein series*

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

is convergent if $\mathrm{re}(s) > 1$. Selberg proved that it has meromorphic continuation to all $\mathbb{C}$. The continuous spectrum is spanned by $E(z, \frac{1}{2} + it)$. If we considered a group with multiple cusps, there would be an Eisenstein series from each cusp.

$\boxed{65}$

There are substantial differences between the arithmetic case where $\Gamma = SL(2,\mathbb{Z})$ or a congruence subgroup and the general case. In the arithmetic case, it is possible to *normalize* the Eisenstein series to get rid of most of its poles.

In the Selberg-Langlands theory, the analytic continuation of the Eisenstein series is proved simultaneously with that of the constant term

$$\int_0^1 E\left(\begin{pmatrix} 1 & t \\ & 1 \end{pmatrix} z, s\right) dt.$$

Since we are assuming there is only one cusp, this has the form $y^s + c(s)\, y^{1-s}$ where c is a meromorphic function. If $\Gamma = SL(2,\mathbb{Z})$ one computes

$$c(s) = \frac{\zeta^*(2s-1)}{\zeta^*(2s)}, \qquad \zeta^*(s) = \pi^{-s}\,\Gamma(s/2)\,\zeta(s).$$

Hence multiplying by $\zeta^*(2s)$ normalizes the Eisenstein series. But if Γ is nonarithmetic normalization is impossible.

If $\Gamma = SL(2,\mathbb{Z})$, the *normalized Eisenstein series* is

$$E^*(z,s) = \pi^s\,\Gamma(s)\,\zeta(2s)\,E(z,s) \tag{34}$$

Theorem 27. *$E^*(z,s)$ has analytic continuation to all s except $s=0,\,1$, where it has simple poles; the residue at $s=1$ is the constant function $\frac{1}{2}$. It satisfies the functional equation $E^*(z,s) = E^*(z,1-s)$.*

Proof. If $z = x+iy \in \mathcal{H}$ and $t>0$, let

$$\Theta(t) = \sum_{(m,n)\in\mathbb{Z}^2} e^{-\pi|mz+n|^2t/y}.$$

It follows from Euler's integral for Γ that

$$E^*(z,s) = \tfrac{1}{2}\int_0^\infty \big(\Theta(t)-1\big)\,t^s\,\frac{dt}{t}.$$

Define the Fourier transform:

$$\hat{f}(x) = \int_{\mathbb{R}^n} f(y)\,e^{2\pi i(x_1y_1+\ldots+x_ny_n)}\,dy.$$

The Poisson summation formula for $\mathbb{R}^n$ asserts

$$\sum_{\xi\in\mathbb{Z}^n} f(\xi) = \sum_{\xi\in\mathbb{Z}^n}\hat{f}(\xi)$$

to a Gaussian on $\mathbb{R}^2$ gives

$$\Theta(t) = t^{-1}\,\Theta(t^{-1}).$$

Using this and

$$\int_0^1 t^{s-1}\,dt = \frac{1}{s}, \qquad \int_1^\infty t^{-s}\,dt = \frac{1}{s-1},$$

$$E^*(z,s) = \tfrac{1}{2}\int_0^\infty \big(\Theta(t)-1\big)\, t^s\, \frac{dt}{t} =$$

$$\tfrac{1}{2}\int_1^\infty \big(\Theta(t)-1\big)\, t^s\, \frac{dt}{t} + \tfrac{1}{2}\int_0^1 \Theta(t)\, t^s\, \frac{dt}{t} - \frac{1}{s} =$$

$$\tfrac{1}{2}\int_0^\infty \Theta(t)\,(t^s + t^{1-s})\,\frac{dt}{t} - \frac{1}{2s} - \frac{1}{2-2s}.$$

From this the analytic continuation and functional equation follow. ■

In the general case where Γ is not arithmetic the functional equation is much more difficult to prove. Selberg and (much later) Bernstein found important general principles to get the analytic continuation. Langlands established the general theory in great generality.

$\boxed{69}$

In order to use representation theoretic methods we need to consider more general functions $f_s \in P(s,\epsilon)$. Fix $\chi : \Gamma \to \infty$ such that $\chi(-I) \equiv (-1)^\epsilon \bmod 2$ and such that $\chi|\Gamma_\infty$ is trivial. We wish to regard f_s as varying continuously with s. This may be accomplished by asking that the restriction of f_s to K be independent of s. It is best to assume this function is K-finite. Let

$$E(g,s,f_s,\chi) = \sum_{\Gamma_\infty\backslash\Gamma} \overline{\chi(\gamma)}\, f_s(\gamma g,\chi).$$

Let $\mathcal{A}(\Gamma\backslash G,\chi)$ be the space of functions on $\Gamma\backslash G$ which are K-finite, Δ-finite and of moderate growth. Such functions are called *automorphic forms*. Let $\mathcal{A}_0(\Gamma\backslash G,\chi)$ be the space of *cusp forms* satisfying (30). It follows from Theorem 26 that elements of $\mathcal{A}_0(\Gamma\backslash G,\chi)$ are of rapid decay and that $\mathcal{A}_0(\Gamma\backslash G,\chi)$ is contained in and spans $L^2(\Gamma\backslash G,\chi)$.

$\boxed{70\text{–}71}$

Then $f_s \mapsto E(g,s,f_s,\chi)$ is an embedding of the $(\mathfrak{g},K)$-module of the principal series representation $P(s,\epsilon)$ into the space $\mathcal{A}(\Gamma\backslash G,\chi)$. We study the constant term

$$E_0(g,s,f_s,\chi) = \int_0^1 E\left(\begin{pmatrix} 1 & t \\ & 1\end{pmatrix} g, s, f_s, \chi\right) dt. \tag{35}$$

We recall that an intertwining operator $M(s) : P(s,\epsilon) \to P(1-s,\epsilon)$ was defined by (29).

Theorem 28. *Assume* $\mathrm{re}(s) > 1$. *There exists an analytic function* $c(s)$ *such that for all* $f_s \in P(s,\epsilon)$

$$E_0(g,s,f_s,\chi) = f_s(g) + c(s)\,\tilde{f}_{1-s},$$

where $\hat{f}_{1-s} = M(s)\,f_s \in P(1-s,\epsilon)$.

Proof. Substitute the definition of $E(g,s,f_s,\chi)$ into (35). The coset Γ_∞ in $\Gamma_\infty\backslash\Gamma$ contributes f_s. We show that the contribution of the remaining terms is proportional to $M(s)\,f_s$. This equals

$$\int_0^1 \sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_\infty\\ \gamma\notin\Gamma_\infty}} \sum_{\delta\in\Gamma_\infty} \overline{\chi(\gamma\delta)}\, f_s\left(\gamma\delta\begin{pmatrix}1 & x\\ & 1\end{pmatrix}g\right)dx$$

$$= \sum_{\substack{\gamma\in\Gamma_\infty\backslash\Gamma/\Gamma_\infty\\ \gamma\notin\Gamma_\infty}} \overline{\chi(\gamma)} \int_{-\infty}^{\infty} f_s\left(\gamma\begin{pmatrix}1 & x\\ & 1\end{pmatrix}g\right)dx.$$

If $\gamma = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$ with $c \neq 0$ then

$$\gamma = \begin{pmatrix}c^{-1} & a\\ & c\end{pmatrix}\begin{pmatrix} & -1\\ 1 & \end{pmatrix}\begin{pmatrix}1 & d/c\\ & 1\end{pmatrix},$$

so the variable change $x \to x - d/c$ shows that

$$\int_{-\infty}^{\infty} f_s\left(\gamma\begin{pmatrix}1 & x\\ & 1\end{pmatrix}g\right)dx = |c|^{-2s}\,M(s)\,f_s,$$

and summing gives the required relation. ■

$\boxed{72}$

Selberg proved:

Theorem 29. $c(s)$ *and* $E(g,s,f_s,\chi)$ *have meromorphic continuation and the same poles. They are analytic for* $\mathrm{re}(s) > \frac{1}{2}$ *except possibly for a finite number of poles on the real axis in* $(\frac{1}{2},1]$. *On the line* $\mathrm{re}(s) = \frac{1}{2}$ *they are holomorphic.*

To the left of $\frac{1}{2}$ there may be infinitely many poles.

Proof when $\Gamma = SL(2,\mathbb{Z})$ and $\chi = 1$. In this case $\epsilon = 0$, and it is sufficient to prove this when $f_s = f_{s,k}$ as in (28). In this case, we will have $c(s) = \zeta(2s-1)/\zeta(2s)$. If $k = 0$, the analytic continuation is proved in Theorem 27; the general case where k is an even integer may be obtained by applying the raising and lowering operators R and L. ■

We recommend Appendix IV of Langlands [LES] as an introduction to the general case.

$\boxed{73}$

Theorem 30. *If s is not a pole of $E(g,s,f_s,\chi)$ then $E - E_0$ is of rapid decay.*

Proof. Let $\phi \in \mathcal{H}$. The function $\tilde{E}(g) = E(g,s) - E_0(g,s)$ is not automorphic with respect to Γ but it is left invariant by Γ_∞ so as before we have

$$T_\phi f(g) = \int_{\Gamma_\infty\backslash G} K_0(g,h)\, f(h)\, dh,$$

where (31). Since f_s is assumed K-finite we may choose ϕ so that $T_\phi f_s = f_s$, in which case $T_\phi \tilde{E} = \tilde{E}$. The rapid decay now follows from the corresponding property of the kernel $K_0(g,h)$. ■

$\boxed{74}$

The asymptotic behavior of $E(g,s,f_s)$ near the cusp is therefore determined by its constant term. The constant term, we see, consists of two parts, one of the order y^s and the other of order y^{1-s}. The smallest growth evidently occurs on the line $\mathrm{re}(s) = \frac{1}{2}$, where it is of order $\sqrt{y}$. Even on this line, the Eisenstein series is not quite L^2.

Let $G_\infty = \left\{ \left(\begin{smallmatrix} 1 & x \\ & 1 \end{smallmatrix}\right) \,\middle|\, x \in \mathbb{R} \right\}$ and let ϕ be a K-finite element of $C_c^\infty(G_\infty\backslash G)$. The *incomplete theta series*

$$\Theta_\phi(g) = \sum_{\Gamma_\infty\backslash\Gamma} \phi(\gamma g)$$

is something like an Eisenstein series but it is not a Δ-Eigenfunction. If $f \in C^\infty(\Gamma\backslash G)$ let

$$f_0(g) = \int_{\Gamma_\infty\backslash G_\infty} f(ug)\, du$$

be its constant term, which is in $C^\infty(G_\infty\backslash G)$.

$\boxed{75}$

Theorem 31. *The incomplete theta series and constant term maps are adjoints; that is*

$$\int_{\Gamma\backslash G} \theta_\phi(g)\, \overline{f(g)}\, dg = \int_{G_\infty\backslash G} \phi(g)\, \overline{f_0(g)}\, dg.$$

Proof. The left side is

$$\int_{\Gamma\backslash G} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \phi(\gamma g)\, \overline{f(g)}\, dg = \int_{\Gamma_\infty\backslash G} \phi(g)\, \overline{f(g)}\, dg =$$

$$\int_{\Gamma_\infty\backslash G_\infty} \int_{G_\infty\backslash G} \phi(ug)\, \overline{f(ug)}\, du\, dg$$

which equals the right side. ∎

$\boxed{76}$

The incomplete theta series are easily seen to be square-integrable.

Theorem 32. *$L^2_0(\Gamma\backslash G)$ is the orthogonal complement of the closed subspace spanned by the incomplete theta series.*

Proof. Immediate from Theorem 31, since the cuspidal spectrum is characterized by vanishing of its constant terms. ∎

We will describe the spectral expansions for $SL(2,\mathbb{R})$. To avoid the slight complexity of having to make a Fourier expansion over K, we restrict ourselves to $L^2(\Gamma\backslash\mathfrak{H})$. Let ξ_i $(i = 1,2,3,\cdots)$ be an orthonormal basis of $L^2_0(\Gamma\backslash\mathfrak{H})$, and let ξ_0 be the constant function $1/\sqrt{\mathrm{vol}(\Gamma\backslash\mathfrak{H})}$.

For the remainder, all functions are right K-invariant, and we do not distinguish functions on G and $\mathfrak{H}$; f_s is as in (17), $E(g,s)$ is the corresponding Eisenstein series and $E^*(g,s)$ is (34).

$\boxed{77}$

Theorem 33. *If ϕ is a cusp form then*

$$\int_{\Gamma\backslash\mathfrak{H}} \phi(g)\, E(g,s)\, dg = 0 \qquad (36)$$

Even though $E(z,s)$ is not L^2 (36) converges absolutely due to the rapid decay of ϕ.

Proof. Assume first $\mathrm{re}(s) > 1$. Then (36) is

$$\int_{\Gamma\backslash\mathfrak{H}} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \phi(g)\, f_s(\gamma g)\, dg =$$

$$\int_{\Gamma_\infty\backslash\mathfrak{H}} \phi(g)\, f_s(g)\, dg =$$

$$\int_{G_\infty\backslash\mathfrak{H}} \int_{\Gamma_\infty\backslash G_\infty} \phi(ug)\, f_s(g)\, du\, dg = 0$$

since ϕ is a cusp form. The general case follows by analytic continuation. ∎

$\boxed{78}$

Theorem 34. *The function ξ_0 is square integrable and orthogonal to the cusp forms.*

Proof. The square integrability is trivial. Taking the residue at $s = 1$ in (36) with a cusp form ϕ

$$\int_{\Gamma\backslash\mathfrak{H}} \phi(g)\, dg = 0$$

so ϕ is orthogonal to the constant function. ∎

Thus the constant function ξ_0 occurs in the discrete spectrum. If besides the constant function ξ_0 the Eisenstein series has other poles in $(\frac{1}{2}, 1]$, these would also be square integrable. The residues of Eisenstein series comprise the "residual spectrum." The discrete spectrum consists of the cusp forms and residual spectrum.

$\boxed{79}$

From now on we assume $\Gamma = SL(2,\mathbb{Z})$.

Theorem 35. *Let $\psi \in L^2(\Gamma\backslash\mathfrak{H})$. Assume also that $\psi(g)\, f_{1/2}(g)$ is integrable over the Siegel set $\mathcal{S}$. Then*

$$\psi(g) = \sum_{j=0}^{\infty} \langle \psi, \xi_j \rangle\, \xi_j(g) + \frac{1}{4\pi} \int_0^\infty \left\langle \psi, E(\,\text{-}\,, \tfrac{1}{2} + it) \right\rangle E(g, \tfrac{1}{2} + it)\, dt.$$

The integrability assumption on that $\psi(g)\, f_{1/2}(g)$ guarantees that the inner products occurring in the expansion are convergent. Correctly interpreted, the expansion is valid for all $f \in L^2(\Gamma\backslash\mathfrak{H})$. This is like the Fourier inversion formula for $\mathbb{R}$, which is valid for all $L^2(\mathbb{R})$ though the usual definition of the Fourier transform as an integral is only strictly correct on $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$.

$\boxed{80}$

Proof. By Theorems 33 and 34 this is true for cusp forms. By Theorem 32 it is sufficient to prove it for incomplete theta series; that is,

$$\theta_\phi(g) = \langle\theta_\phi, \xi_0\rangle\,\xi_0(g) + \frac{1}{4\pi}\int_0^\infty \big\langle\theta_\phi, E(\,\text{-}\,, \tfrac{1}{2}+it)\big\rangle\, E(g, \tfrac{1}{2}+it)\,dt \tag{37}$$

when $\phi \in C_c^\infty(G_\infty\backslash\mathfrak{H})$. Define

$$\hat\phi(s) = \int_{G_\infty\backslash G} \phi(g)\, f_{1-s}(g)\,dg = \int_0^\infty \phi\begin{pmatrix} y^{1/2} & \\ & y^{-1/2}\end{pmatrix} y^{-s}\,\frac{dy}{y}.$$

By the Mellin inversion formula

$$\phi\begin{pmatrix} y^{1/2} & \\ & y^{-1/2}\end{pmatrix} = \int_{\sigma-i\infty}^{\sigma+i\infty} \hat\phi(s)\, y^s\,dy. \tag{38}$$

$\boxed{81}$

This is valid for all values of σ. Now

$$\phi(g) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \hat\phi(s)\, f_s(g)\,dy.$$

Indeed, both sides are left invariant by G_∞, right invariant by K, and agree on the diagonal by (38), so this follows by the Iwasawa decomposition. We take $\sigma > 1$ and sum over γg with $\gamma \in \Gamma_\infty\backslash\Gamma$ to obtain

$$\theta_\phi(g) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \hat\phi(s)\, E(g,s)\,ds.$$

Now we move the path of integration to the left. The pole at $s = 1$ contributes a residue so

$$\theta_\phi(g) = \tfrac{6}{\pi}\,\hat\phi(1) + \frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} \hat\phi(s)\, E(g,s)\,ds. \tag{39}$$

$\boxed{82}$

The constant term

$$E_0(g,s) = \int_0^1 E\left(\begin{pmatrix} 1 & x \\ & 1 \end{pmatrix} g, s\right) dx = f_s(g) + \frac{\zeta^*(2s-1)}{\zeta^*(2s)} f_{1-s}(g),$$

so

$$\text{(40)} \quad \int_{\Gamma\backslash G} \theta_\phi(g)\, E(g,s)\, dg = \sum_{\gamma\in\Gamma\backslash\Gamma_\infty} \int_{\Gamma_\infty\backslash G} \theta_\phi(\gamma g)\, E(g,s)\, dg =$$
$$\int_{\Gamma_\infty\backslash G} \theta_\phi(\gamma g)\, E_0(g,s)\, dg = \hat{\phi}(1-s) + \frac{\zeta^*(2s-1)}{\zeta^*(2s)} \hat{\phi}(s).$$

$\boxed{83}$

In (39) we use the functional equation (Theorem 27):

$$E(g, \tfrac{1}{2}+it) = \tfrac{1}{2}\left(E(g, \tfrac{1}{2}+it) + \frac{\zeta^*(1-2it)}{\zeta^*(1+2it)} E(g, \tfrac{1}{2}-it)\right).$$

Also, $\mathrm{vol}(\Gamma\backslash G) = \pi/3$, as may be shown by applying the spectral expansion as derived so far to the function 1 or by any other method, so

$$\hat{\phi}(1) = \int_{\Gamma_\infty\backslash G} \phi(g)\, dg = \int_{\Gamma\backslash G} \theta_\phi(g)\, dg = \mathrm{vol}(\Gamma\backslash G)\, \langle \theta_\phi, \xi_0\rangle\, \xi_0$$

Substituting these last two identities into (39) and applying (40) we obtain (37). ■

The Jacquet-Langlands Correspondence

If D_1 and D_2 are central division algebras over a field F, then $D_1 \otimes D_2 \cong \mathrm{Mat}_k(D_3)$ for some D_3 and k, and $D_1, D_2 \to D_3$ is an associative multiplication on the set $B(F)$ of isomorphism classes of central division algebras. Thus $B(F)$ is a group, called the *Brauer group*.

If D is a central division algebra over F then the dimension of D is a square d^2, and if E/F is any field extension of degree d then $E \otimes D \cong \mathrm{Mat}_d(E)$. Thus a division ring is a Galois form of a matrix ring. The composite map

$$D \to E \otimes D \cong \mathrm{Mat}_d(E) \to E,$$

the last map being either the trace or determinant, takes values in F, and gives us the *reduced trace* or *reduced norm*.

$\boxed{85}$

The Brauer group of a local or global field F admits a simple and beautiful description related to the reciprocity laws of class field theory. For example, suppose that D is a *quaternion algebra* over a local field F, that is a central simple algebra of degree 4, whose reduced norm is equivalent to the quadratic form $x^2 - ay^2 - bz^2 + abw^2$, define $(a, b) = 1$ if D is a matrix ring, -1 if D is a division ring. Then (a, b) is the *Hilbert symbol*.

If F is a global field and v is a place, let $(a, b)_v$ be the corresponding local Hilbert symbol. The *Hilbert reciprocity law* states that

$$\prod_v (a, b)_v = 1.$$

This is equivalent to the quadratic reciprocity law. It implies that there are an even number of places v such that $D_v = F_v \otimes D$ is not a matrix ring. These facts may be extended to a full description of $B(F)$.

$\boxed{86}$

Let F be a global field, A its adele ring, and D a central division algebra of degree p^2 over F, where p is a prime. Let Z be the center of $D^\times$. Let S be the finite set of places where D_v is a division ring. If $v \notin S$ we identify $D_v = \mathrm{Mat}_p(F_v)$.

Let $\mathcal{H}$ be the set of functions on $D_A^\times = \prod_v D_v^\times$ which are finite linear combinations of functions of the form $\prod_v \phi_v$ where for each v, $\phi_v : D_v^\times \to \mathbb{C}$ is smooth and compactly supported modulo Z_v, satisfies $\phi_v(z_v g_v)$ when $z_v \in Z_v$, and agrees with the characteristic function of $Z_v \mathrm{Mat}_p(\mathfrak{o}_v)$ for almost all places v of F.

$Z_A D_F^\times \backslash D_A^\times$ is compact. As with $SL(2,\mathbb{R})$, $L^2(Z_A D_F^\times \backslash D_A^\times)$ admits integral operators T_ϕ for $\phi \in \mathcal{H}$:

$$T_\phi(g) = \int_{Z_A \backslash D_A^\times} \phi(h)\, \phi(gh)\, dh.$$

$\boxed{87\text{–}88}$

Let $\{\gamma\}$ be a set of representatives for the conjugacy classes of $D_F^\times$. We denote by C_γ the centralizer of γ in $D_F^\times$. It is an algebraic group, so $C_\gamma(A) \subset D_A^\times$ will denote its points in A.

Theorem 36. *(Selberg trace formula).*

$$\text{(41)} \quad \operatorname{tr} T_\phi = \sum_{\{\gamma\}} \operatorname{vol}(C_\gamma \backslash C_\gamma(A)) \int_{C_\gamma(A) Z_A \backslash D_A^\times} \phi(g^{-1}\gamma g)\, dg.$$

Proof. The proof of Theorem 2 goes through without change, so

$$(T_\phi f)(g) = \int_{Z_A D_F^\times \backslash D_A^\times} K_\phi(g,h)\, f(h)\, dh,$$

$$K_\phi(g,h) = \sum_{\gamma \in \Gamma} \phi(g^{-1}\gamma h).$$

As with $SL(2,\mathbb{R})$, the operator T_ϕ is thus Hilbert-Schmidt, and with more work may be shown to be trace class. As in Theorem 11,

$$\operatorname{tr} T_\phi = \int_{Z_A D_F^\times \backslash D_A^\times} K_\phi(g,g)\, dg$$

Now (41) follows as in Theorem 19. ∎

The conjugacy classes of $D_F^\times$ are easily described.

Theorem 37. *If $\alpha \in D_F^\times - Z_F$, then $F(\alpha)$ is a field extension of F of degree p. Elements α and β are conjugate in $D_F^\times$ if and only if there is a field isomorphism $F(\alpha) \to F(\beta)$ such that $\alpha \mapsto \beta$. If $F(\alpha)$ is a field extension of degree p, then $F(\alpha)$ may be embedded in D_F if and only if $[F_v(\alpha) : F_v] = p$ for all $v \in S$.*

Proof. The conjugacy of α and β follows from the *Skolem-Noether Theorem* (Herstein [He], p. 99). The last statement follows from (i) $\leftrightarrow$ (ii) in Weil [W], Proposition VIII.5 on p. 253. ■

$\boxed{89}$

The trace formula can be used to prove functorial liftings in many cases. We describe the simplest example, from Gelbart and Jacquet [GJ].

Let E be another division algebra of degree p^2, and assume that the set of places where E_v is a division ring agrees with the set S of places where D_v is. If $p = 2$ this implies that D and E are isomorphic, but not in general. We will show that two spaces of automorphic forms on D and on E are isomorphic.

Suppose that $\pi = \otimes_v \pi_v$ is an irreducible constituent of $L^2(Z_A D_F^\times \backslash D_A^\times)$. Since $Z_v \backslash D_v^\times$ is compact for $v \in S$, π_v is finite-dimensional at these places. We assume that π_v is trivial when $v \in S$.

If $v \notin S$, then $D_v \cong E_v \cong \mathrm{Mat}_p(F_v)$. We may therefore identify π_v with an irreducible representation π'_v of E_v when $v \notin S$, and if $v \in S$ we take $\pi'_v = 1$. Let $\pi' = \otimes \pi'_v$. It is an irreducible representation of $E_A^\times$.

$\boxed{90}$

Theorem 38. *π' occurs in $L^2(Z_A E_F^\times \backslash E_A^\times)$.*

Proof. If $v \in S$, then $Z_v \backslash D_v^\times$ is compact, so the constant function $\phi_v^\circ(g_v) = 1$ is in $C_c^\infty(F_v)$. Let $\mathcal{H}_S$ be the subalgebra of $\mathcal{H}$ spanned by functions $\prod \phi_v$ such that $\phi_v = \phi_v^\circ$ for $v \in S$. It is isomorphic to the corresponding Hecke ring on E. Let $\phi \to \phi'$ denote this isomorphism.

By Theorem 37, noncentral conjugacy classes in $D_F^\times$ and $D_F^\times$ are both in bijection with the set of Galois equivalence classes of elements α of field extensions $[F(\alpha) : F] = p$ such that $[F_v(\alpha) : F_v] = p$ for all $p \in S$. This intrinsic characterization shows that we may identify the conjugacy classes of D_F and E_F, and compare trace formulae to get

$$\mathrm{tr}\ T_\phi = \mathrm{tr}\ T'_\phi. \tag{42}$$

This is almost but not quite as easy as we've made it sound.

$\boxed{91}$

It follows from (42) that the representations of $\mathcal{H}_S$ on the spaces $L^2(Z_A D_F^\times \prod_{v\in S} D_v^\times \backslash D_A^\times)$ and $L^2(Z_A E_F^\times \prod_{v\in S} E_v^\times \backslash E_A^\times)$ are isomorphic, and the theorem follows. ■

Underlying the final step is the fact that two representations of rings are characterized by their traces. For example if R is an algebra over a field of characteristic zero and if M_1, M_2 are finite dimensional semisimple R-modules, and if for every $\alpha \in R$ the induced endomorphisms of M_1 and M_2 have the same trace, then the modules are isomorphic. This statement is not directly applicable here but it gives the flavor.

For the remainder we take $p = 2$, and review the *Jacquet-Langlands correspondence*. Let D be as before. The Jacquet-Langlands correspondence is a lifting of automorphic representations from $D^\times$ to $GL(2, F)$.

$\boxed{92}$

There is a local correspondence for $v \in S$. $D_v^\times$ is compact modulo its center, so its irreducible representations are finite dimensional. These lift to irreducible representations of $GL(2, F_v)$ having the same central character. The lifting was constructed by Jacquet and Langlands by use of the theta correspondence. Indeed, $Z_v \backslash D_v^\times$ is a quotient of the orthogonal group $GO(4)$ while $GL(2)$ is the same as $GSp(2)$, and theta correspondence $GO(4) \leftrightarrow GSp(2)$ gives the Jacquet-Langlands correspondence. Its image is the square integrable representations (supercuspidal+Steinberg).

Jacquet and Langlands constructed a global correspondence from automorphic forms on $D^\times$ to automorphic forms on $GL(2)$ first using the converse theorem in Section 14 of their book. To prove functional equations of L-functions on $D^\times$, they use the Godement-Jacquet construction, because the Hecke integral is not available in this context.

$\boxed{93}$

Finally, they reconsidered the lifting from the point of view of the trace formula. This allowed them to characterize the image of the lift. They sketched a proof (and later Gelbart and Jacquet completed) of:

Theorem 39. *An automorphic representation π of $GL(2, A)$ is the lift of an automorphic representation of $D_A^\times$ if and only if π_v is square integrable for every $v \in S$.* ∎

The remarkable fact is that their paper of this fact uses the full range of techniques which have proved important in the subsequent 30 years: the Hecke and Godement-Jacquet constructions, the Weil representation and the trace formula.

The trace formula on $GL(2)$ is hard because of the presence of the continuous spectrum. We've avoided these problems by proving Theorem 38 instead of Theorem 39.

References

[A] J. Arthur, The trace formula and Hecke operators, in *Number theory, trace formulas and discrete groups,* A Symposium in honor of Atle Selberg held at the University of Oslo, 11–27, Aubert, Bombieri and Goldfeld ed., Academic Press (1989).

[Bo] A. Borel, *Automorphic Forms on* $SL(2,\mathbb{R})$, Cambridge (1997).

[B] D. Bump, Automorphic Forms and Representations, Cambridge (1997).

[GJ] S. Gelbart and H. Jacquet, Forms of $GL(2)$ from an analytic point of view, in A. Borel and W. Casselman, ed., *Automorphic Forms, Representations, and L-functions*, AMS Proceedings of Symposia in Pure Mathematics **33** (1979), Part I, 213–254.

[GGP] I. Gelfand, M. Graev and I. Piatetskii-Shapiro, *Representation theory and automorphic functions*, K. Hirsch trans., Saunders (1969). Reprinted by Academic Press (1990).

[H] D. Hejhal, *The Selberg trace formula for* $PSL(2,R)$*, Vol. I*, Lecture Notes in Mathematics **548**, Springer Verlag (1976).

[He] I. Herstein, *Noncommutative Rings*, Carus Monograph **15**, MAA (1968).

[I] H. Iwaniec, *Introduction to the spectral theory of automorphic forms*, Biblioteca de la Revista Matematica Iberoamericana. Madrid (1995).

[JL] H. Jacquet and R. Langlands, *Automorphic Forms on* $GL(2)$, Springer Lecture Notes in Mathematics **114** (1970).

[L] S. Lang, $SL(2,\mathbb{R})$, Addison Wesley (1975).

[LES] R. Langlands, *On the Functional Equations satisfied by Eisenstein Series*, Springer Lecture Notes in Mathematics **544** (1976).

[MW] C. Moeglin and J.-L. Waldspurger, *Décomposition Spectrale et Séries d'Eisenstein*, Birkhäuser (1993). English version, *Spectral Decomposition and Eisenstein Series*, Cambridge (1995).

[S] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series. *J. Indian Math. Soc.* (N.S.) **20** (1956), 47–87.

[Ta] T. Tamagawa, On Selberg's trace formula, *J. Fac. Sci. Univ. Tokyo Sect. I* **8** (1960) 363–386.

[Te] A. Terras, *Harmonic analysis on symmetric spaces and applications*, I., Springer-Verlag (1985).

[Va] V. Varadarajan, *An introduction to harmonic analysis on semisimple Lie groups*, Cambridge, (1999).

[V] A. Venkov, *Spectral theory of automorphic functions and its applications*, Kluwer (1990).

[W] A. Weil, *Basic Number Theory*, Springer Verlag (1967).