JOURNAL OF MODERN DYNAMICS
VOLUME 3, NO. 2, 2009, 271–309

doi: 10.3934/jmd.2009.3.271

THE VARIANCE OF ARITHMETIC MEASURES ASSOCIATED TO CLOSED GEODESICS ON THE MODULAR SURFACE

WENZHI LUO, ZEÉV RUDNICK AND PETER SARNAK

(Communicated by Jens Marklof)

ABSTRACT. We determine the variance for the fluctuations of the arithmetic measures obtained by collecting all closed geodesics on the modular surface with the same discriminant and ordering them by the latter. This arithmetic variance differs by subtle factors from the variance that one gets when considering individual closed geodesics when ordered by their length. The arithmetic variance is the same one that appears in the fluctuations of measures associated with quantum states on the modular surface.

1. INTRODUCTION

1.1. **Equidistribution theorems for closed geodesics.** Let X be a compact surface with a metric of constant negative curvature $\kappa = -1$, SX the unit tangent bundle of X, and $\Phi^t\colon SX \to SX$ the geodesic flow. We think of SX as the set of initial conditions (z,ζ), with $z \in X$ the position and ζ the direction vector.

The geodesic flow is *ergodic* with respect to Liouville measure dx, the smooth invariant measure for the flow: *generic* geodesics become equidistributed, in the sense that for Lebesgue-almost all initial conditions $x_0 \in SX$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T F(\Phi^t x_0)\,dt = \int_{SX} F(x)\,dx$$

for integrable observables on SX.

As is well known, there are infinitely many *closed* geodesics; in fact, the number $\pi(T)$ of closed geodesics of length at most T grows exponentially with T, $\pi(T) \sim e^T/T$ as $T \to \infty$ [38, 7]. For a closed geodesic C, let $\ell(C)$ be its length and μ_C the arc-length measure along C, *i.e.,* for $F \in C(X)$,

$$\int_C F\,d\mu_C := \int_0^{\ell(C)} F(\Phi^t x)\,dt, \quad x \in C\,.$$

This is a measure on SX which is invariant under the geodesic flow and of total mass $\ell(C)$. Closed geodesics become, on average, uniformly distributed with

Received October 17, 2008.
2000 *Mathematics Subject Classification:* Primary: 11F72; Secondary: 37D40.
Key words and phrases: Arithmetic variance, modular surface, closed geodesics.
WL & PS: Supported in part by NSF FRG Grant DMS-0554373.
ZR & PS: Supported in part by Grant No 2006254 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel.

respect to dx:[1] for any observable $F \in C(SX)$, we have

$$\lim_{T\to\infty} \frac{1}{\pi(T)} \sum_{\ell(C)\le T} \frac{1}{\ell(C)} \int_C F\, d\mu_C = \int_{SX} F(x)\, dx\,.$$

Lalley [19] determined the fluctuations of the numbers $\mu_C(F)/\sqrt{\ell(C)}$ for F as above and of zero mean as C varies over closed geodesics ordered by length. He showed that they are Gaussian with mean zero and variance $V(F,F)$, where V is the hermitian bilinear form on functions of zero mean given by

$$V(F_1,F_2) = \int_{-\infty}^{\infty} \left(\int_{SX} F_1(x) F_2(\Phi^t x)\, dx \right) dt\,. \tag{1.1}$$

The negative curvature guarantees that the correlations in the inner integral decay exponentially as $t \to \pm\infty$, so V is convergent [31].

The bilinear form V is positive semidefinite and is degenerate; in fact, if F_0 is smooth, then $V(F_0,F) = 0$ for all F if and only if F_0 is a derivative in the flow direction, *i.e.*, $F_0 = \frac{d}{dt}|_{t=0} f \circ \Phi^t$ for some other observable $f \in C^\infty(SX)$.

An important involution of SX is *time-reversal symmetry*,

$$\mathbf{w} \colon (z,\zeta) \mapsto (z,-\zeta),$$

which reverses the direction vector of the initial condition and satisfies $\mathbf{w}\circ\Phi^t = \Phi^{-t}\circ\mathbf{w}$. It induces an involution on the set of geodesics, taking a geodesic $C = \{\Phi^t x_0 : t \in \mathbb{R}\}$ to its time-reversal $\bar{C} = \mathbf{w}C = \{\Phi^s \mathbf{w} x_0 : s \in \mathbb{R}\}$.

Time-reversal symmetry can also be incorporated in Lalley's theorem. To do so, note that for a closed geodesic C, its time-reversed partner $\bar{C}$ is also closed and both have the same length, $\ell(C) = \ell(\bar{C})$. Grouping these together yields the measure $\mu_C^{\text{even}} := \mu_C + \mu_{\bar{C}}$ of mass $2\ell(C)$. By Lalley's theorem, the fluctuations of $\mu_C^{\text{even}}/\sqrt{2\ell(C)}$ are again Gaussian with mean zero, but with variance given by the hermitian form

$$V^{\text{even}}(F_1,F_2) = V\left(F_1^{\text{even}}, F_2^{\text{even}}\right), \tag{1.2}$$

where $F^{\text{even}} = (F + F\circ\mathbf{w})/2$ is the even part of F under $\mathbf{w}$. Note that μ_C^{even} is invariant and V^{even} is bi-invariant under the geodesic flow as well as under time-reversal symmetry $\mathbf{w}$. Both of the hermitian forms V and V^{even} on

$$L_0^2(SX) := \left\{ f \in L^2(SX) : \int_{SX} f(x)\, dx = 0 \right\}$$

can be diagonalized and computed explicitly by decomposing the regular representation of $\mathrm{PSL}_2(\mathbb{R})$ on this space; see §3.

[1] In *variable* negative curvature, one needs the Bowen–Margulis measure here. To get an equidistribution statement involving Liouville measure, one needs to weigh each geodesic by its "monodromy".

1.2. **The modular surface.** In this paper, we investigate fluctuations of measures on the *modular surface* associated with grouping together geodesics of equal *discriminant*. As is well known, any of our compact surfaces X may be uniformized as a quotient of the upper half-plane $\mathbb{H}$, equipped with the hyperbolic metric, by a Fuchsian group Γ. Furthermore, the group $G = \mathrm{PSL}_2(\mathbb{R})$ of orientation-preserving isometries of $\mathbb{H}$ acts transitively on the unit tangent bundle SX, giving an identification $SX \simeq \Gamma\backslash G$; this is reviewed in §2. The *modular surface* is obtained by taking $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$; the resulting surface is noncompact (but of finite volume) and has elliptic fixed points, but these issues will not be important for us.

Closed geodesics correspond to (hyperbolic) conjugacy classes in Γ, with the length of a closed geodesic C given in terms of the trace t of the corresponding conjugacy class by $\ell(C) = 2\log(t+\sqrt{t^2-4})/2$. In the case of the modular surface, the hyperbolic conjugacy classes correspond to (strict) equivalence classes of integer binary quadratic forms $ax^2+bxy+cy^2$ (also denoted by $[a,b,c]$) of positive discriminant $d := b^2-4ac$, with the modular group acting by linear substitutions (we need to exclude discriminants which are perfect squares). The discriminant $\mathrm{disc}(C)$ of a closed geodesic C is defined as the discriminant of the corresponding binary quadratic form.

For $d > 0$, $d \equiv 0,1 \mod 4$, and d not a perfect square, let $\bar{f}_1,\dots \bar{f}_{H(d)}$ be the classes of binary quadratic forms of discriminant d. We do not assume that $f_j = [a_j,b_j,c_j]$ is primitive, and so $H(d)$ is the Hurwitz class number [20]. Let

$$\epsilon_d = \frac{t_d+\sqrt{d}u_d}{2}, \quad t_d>0,\ u_d>0 \tag{1.3}$$

be the fundamental solution of the Pellian equation $t^2-du^2=4$. Then, as in [35, 37], associate to each $\bar{f}_j$ the Γ-conjugacy class (it is well-defined) of the matrix

$$\begin{pmatrix} \frac{t_d-b_ju_d}{2} & a_ju_d \\ -c_ju_d & \frac{t_d+b_ju_d}{2} \end{pmatrix}. \tag{1.4}$$

This gives $H(d)$ closed geodesics for each discriminant d, all of length $2\log\epsilon_d$. Let μ_d be the corresponding measure on SX:

$$\mu_d = \sum_{\mathrm{disc}(C)=d} \mu_C\,. \tag{1.5}$$

These measures are the arithmetic measures in the title of the paper. They have been studied extensively and the primary result about them is that they become equidistributed as $d\to\infty$. That is, if F is bounded and continuous on SX and has mean zero, then

$$\frac{\mu_d(F)}{H(d)2\log\epsilon_d}\to 0 \quad \text{as } d\to\infty\,.$$

Linnik [23] developed an ergodic-theoretic approach to this equidistribution problem, and recently, in [3], it was shown that this method leads to a proof of this specific result. The first proof of equidistribution is due to Iwaniec [10] and Duke [1]. Iwaniec established the requisite estimate for Fourier coefficients of holomorphic half-integral weight forms (of weight > 5/2) and Duke obtained the estimates for weight 3/2 and weight zero Maass forms. In view of our reductions in Sections §3 and §6, together these imply[2] the full equidistribution on SX.

The measures μ_d enjoy some symmetries (see [37]). First, they are invariant under time-reversal symmetry: $\mathbf{w}\mu_d = \mu_d$. Second, let $\mathbf{r}$ be the involution of $\Gamma\backslash G$ given by $g \mapsto \delta^{-1} g\delta$, where

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

(it is well defined since $\delta^{-1}\Gamma\delta = \Gamma$). In terms of the coordinates (z,ζ) on SX, $\mathbf{r}$ is the orientation-reversing symmetry

$$\mathbf{r}\colon (z,\zeta) \mapsto (-\overline{z}, -\overline{\zeta}) .$$

The measure μ_d is also invariant under $\mathbf{r}$. The involutions $\mathbf{w}$, $\mathbf{r}$ commute and their product $\mathbf{rw}$ is also an involution. Thus μ_d is invariant under the Klein four-group $H = \{I, \mathbf{r}, \mathbf{w}, \mathbf{rw}\}$. These involutions induce linear actions on $L^2(\Gamma\backslash G)$ by $f(x) \mapsto f(h(x))$, with $h \in H$ and $x \in \Gamma\backslash G$, and we denote these transformations by the same symbols. The fluctuations of the measures μ_d inherit these symmetries, and since we are particularly interested in comparing their variance with the classical variance V, we define the symmetrized classical variance V^{sym} on functions of mean zero on $\Gamma\backslash G$ by

$$V^{\mathrm{sym}}(F_1, F_2) := V(F_1^{\mathrm{sym}}, F_2^{\mathrm{sym}}), \tag{1.6}$$

where

$$F^{\mathrm{sym}} := \frac{1}{4}\sum_{h\in H} hF .$$

1.3. **Results.** We can now state our main results about the fluctuations of μ_d. We normalize these measures as

$$\tilde{\mu}_d := \frac{\mu_d}{d^{1/4}} .$$

This is essentially equivalent to normalizing by the square root of the total mass, $\sqrt{H(d)2\log\epsilon_d}$; see Remark 1.4.2. The space of natural observables for which one might compute these quantities is $L^2_0(\Gamma\backslash G)$, or at least a dense subspace thereof. This space decomposes as an orthogonal direct sum of the cuspidal subspace

$$L^2_{\mathrm{cusp}}(\Gamma\backslash G) := \left\{ f \in L^2(\Gamma\backslash G) : \int_{N\cap\Gamma\backslash N} f(nx)\, dn, \text{ for a.e. } x \in \Gamma\backslash G \right\},$$

where $N = \left\{ \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} : u \in \mathbb{R} \right\}$, and the unitary Eisenstein series [5]. The former is the major and difficult part of the space $L^2_0(\Gamma\backslash G)$, so we will concentrate exclusively on it. One can easily extend our analysis of the variance to the unitary Eisenstein series.

[2] Specifically, by (3.25), (6.1) and (6.3), equidistribution on SX is reduced to an estimation of Fourier coefficients of classical holomorphic forms of half integer weight and Maass forms of weight 1/2.

THEOREM 1.1. *Fix smooth, K-finite $F_1, F_2 \in L^2_{cusp}(\Gamma\backslash G)$. Then*

$$\lim_{Y\to\infty} \frac{1}{\#\{d : d \le Y\}} \sum_{d\le Y} \frac{\mu_d(F_1)}{d^{1/4}} = 0, \tag{1.7}$$

and there is a limiting variance

$$B(F_1,F_2) = \lim_{Y\to\infty} \frac{1}{\#\{d : d \le Y\}} \sum_{d\le Y} \frac{\mu_d(F_1)}{d^{1/4}} \overline{\frac{\mu_d(F_2)}{d^{1/4}}} \tag{1.8}$$

We call this variance B the "arithmetic variance." The structure of the bilinear form B is revealed by choosing a special basis of observables, compatible with the symmetries of the problem. Recall that the unit tangent bundle is a homogeneous space for $G = \mathrm{PSL}_2(\mathbb{R})$, and thus it is natural to decompose the space $L^2(\Gamma\backslash G)$ into the irreducible components under the G-action. In addition, there is an algebra of Hecke operators acting on this space, commuting with the G-action, hence also acting on each isotypic G-component. We take observables lying in irreducible spaces for the joint action of G and the Hecke operators – the automorphic subrepresentations of $L^2_{\mathrm{cusp}}(\Gamma\backslash G)$. Denote the decomposition of the regular representation on $L^2_{\mathrm{cusp}}(\Gamma\backslash G)$ into G- and Hecke-irreducible subspaces by

$$L^2_{\mathrm{cusp}}(\Gamma\backslash G) = \bigoplus_{j=1}^{\infty} W_{\pi_j}, \tag{1.9}$$

so π_j is a cuspidal automorphic representation.

In order to describe the arithmetic variance explicitly, we need a more detailed description of the W_{π_j}'s. To each π_j is associated an even integer k, its weight (see §3), which we indicate by π_j^k. For $k = 0$, there are infinitely many π_j^0's corresponding to Hecke–Maass cusp forms on X, while for $k > 0$ there are d_k such π_j^k (where d_k is either $\left[\frac{k}{12}\right]$ or $\left[\frac{k}{12}\right]+1$, depending on whether $k/2 = 1$ mod 6 or not), corresponding to holomorphic Hecke cusp forms of weight k. For $k < 0$, let

$$W_{\pi_j^k} = \overline{W_{\pi_j^{-k}}} = \left\{\overline{f} : f \in W_{\pi_j^{-k}}\right\}$$

for $j = 1,2,\dots,d_{-k}$, and these correspond to the antiholomorphic Hecke cusp forms. With these, we have the orthogonal decompositions

$$\begin{aligned} L^2_{\mathrm{cusp}}(\Gamma\backslash G) &= \sum_{j=1}^{\infty} W_{\pi_j} \oplus \sum_{k\ge12}\sum_{j=1}^{d_k}\left(W_{\pi_j^k} \oplus W_{\pi_j^{-k}}\right) \\ &= \sum_{j=1}^{\infty} U_{\pi_j^0} \oplus \sum_{k\ge12}\sum_{j=1}^{d_k} U_{\pi_j^k}, \end{aligned} \tag{1.10}$$

where

$$U_{\pi_j^0} = W_{\pi_j^0} \quad \text{and} \quad U_{\pi_j^k} = W_{\pi_j^k} \oplus W_{\pi_j^{-k}}. \tag{1.11}$$

One associates to each π_j as above an L-function $L(s,\pi_j)$ given by

$$L(s,\pi_j)=\sum_{n=1}^{\infty}\frac{\lambda_{\pi_j}(n)}{n^s},\quad \Re(s)>1, \tag{1.12}$$

where $\lambda_{\pi_j}(n)$ is the eigenvalue of the Hecke operator T_n acting on W_{π_j}. It is well known (Hecke–Maass) that $L(s,\pi_j)$ extends to an entire function and satisfies a functional equation relating its value at s to $1-s$. In particular, the arithmetical central value $L(\frac{1}{2},\pi_j)$ is well defined (and real).

THEOREM 1.2. *Both V^{sym} and B are diagonalized by the decomposition (1.10), and on each subspace $U_{\pi_j^k}$ we have that*

$$B|_{U_{\pi_j^k}}=c(k)L(\frac{1}{2},\pi_j^k)V^{sym}|_{U_{\pi_j^k}},$$

where $c(0)=6/\pi$ and $c(k)=1/\pi$ if $k>0$.

1.4. **Remarks.**

1.4.1. The hermitian forms $V^{\rm sym}$ and B can be computed explicitly on each $U_{\pi_j^k}$ (see §3). Time-reversal symmetry **w** forces $V^{\rm sym}$ to vanish on $U_{\pi_j^k}$ for $k=2$ mod 4. Also, orientation-reversal symmetry **r** fixes the weight zero spaces $U_{\pi_j^0}$, and hence takes the generating vector (see §3) $\phi_j^0\in\pi_j^0$ into $\pm\phi_j^0$. Corresponding to this sign, we call $U_{\pi_j^0}$ even or odd. According to §3, $V^{\rm sym}$ is completely determined on $U_{\pi_j^k}$ by its value on the generating vector; hence it follows that $V^{\rm sym}|_{U_{\pi_j}}\equiv 0$ for the odd π_j^0's. In the above cases where $V^{\rm sym}|_{U_{\pi_j^k}}$ vanishes, the sign ϵ_{π_j} of the functional equation of $L(s,\pi_j)$ is -1, and hence the central L-value $L(\frac{1}{2},\pi_j)=0$ for reasons of symmetry. In the other cases ($k=0 \mod 4$ and π_j^0 even), $\epsilon_{\pi_j}=1$ and $V^{\rm sym}|_{U_{\pi_j}}\neq 0$. One expects that in these cases, $L(\frac{1}{2},\pi_j)\neq 0$ as well. However, if we pass from $\Gamma=\mathrm{PSL}_2(\mathbb{Z})$ to a congruence subgroup, where our analysis can be carried over with similar results, then there will be π's corresponding to holomorphic forms for which $L(\frac{1}{2},\pi)=0$ for number-theoretic reasons, specifically the conjecture of Birch and Swinnerton-Dyer [41]. In this case the restriction of the arithmetic variance to such a subspace will vanish for reasons far deeper than just symmetry.

1.4.2. The normalization $\mu_d(F)/d^{1/4}$ is natural from the arithmetic point of view. To be consistent with the previous normalization, we should use the square root of the total mass $\sqrt{H(d)2\log\epsilon_d}$ of the measure. By Dirichlet's class number formula for d fundamental, when $H(d)=h(d)$ is the ordinary class number (and similar formulae for all d),

$$h(d)\log\epsilon_d=\sqrt{d}L(1,\chi_d)\,. \tag{1.13}$$

The fluctuations of $L(1,\chi_d)$ are mild and well-understood [4], and hence the normalizations are essentially the same. In any case, one could use methods as in

[11, Chapter 26] to remove the weights $L(1,\chi_d)$ and deduce Theorem 1.1 with this other normalization.

1.4.3. In §3, we show, in a more abstract context, that the space of linear forms on an irreducible unitary representation of G which are invariant under both the geodesic flow and time-reversal symmetry is at most one-dimensional, and how to incorporate orientation-reversal symmetry. This shows that the form that the arithmetic and "classical" variance take is universal. That is, for any family of such invariant measures, the variance B', if it exists, is determined completely in each irreducible representation of G by $B'(v_0, v_0)$, where v_0 is either a spherical vector or a lowest (or highest) weight vector in the representation.

1.4.4. The geometric problem is to order the μ_d by the *length* of any of the geodesic components of the measure. We do not know how to do this. What we can do is to compute the variance of the μ_d's when ordered by the discriminant d. From the arithmetic point of view, this ordering is anyway the most natural one. For many considerations, these two orderings of μ_d yield quite different answers (see [37]). However, for the fluctuations, we believe they are similar.

The difficulty in proving the same result of the μ_d's ordered by t_d (or ϵ_d) is apparent already for $F_1 = F_2 = f$, where f is a holomorphic cusp form of weight $m \equiv 0 \mod 4$. In this case, according to the formula of Kohnen and Zagier [16], we have for d a fundamental discriminant

$$\frac{|\mu_d(f)|^2}{\sqrt{d}} = *L(\frac{1}{2}, f \otimes \chi_d) \tag{1.14}$$

(with $*$ explicit and under control). Thus we would need to understand the averages

$$\sum_{t_d \le Y} L(\frac{1}{2}, f \otimes \chi_d). \tag{1.15}$$

The first, but big, step in this direction would be to understand

$$\sum_{t \le Y} L(\frac{1}{2}, f \otimes \chi_{t^2-4}) \tag{1.16}$$

(see [32] for an execution of such an analysis on a simpler problem). This appears to be beyond the well-developed techniques for averaging special values of L-functions in families. We leave it as an interesting open problem.

1.4.5. The recent work [34] giving lower bounds for moments of special values of L-functions in families, together with (1.14), shows that the fluctuations of $\mu_d(F)/d^{1/4}$ are not Gaussian, at least not in the sense of convergence of moments.

1.4.6. The arithmetic variance B in Theorem 1.1 is the same as the quantum variance for the fluctuations of high energy eigenstates on the modular surface that were calculated in [24] and [42]. We expect that the variance for the μ_d's, when ordered by length, will be the same as B. This would give a semiclassical periodic orbit explanation for the singular finding [24] that the quantum variance is B rather than V^{even}. It points yet again, just as for the local spectral statistics (see the survey [36]), to the source of the singular behavior of the quantum fluctuations in arithmetic surfaces being the high multiplicity of the length spectrum. Similar phenomena are found for the quantized cat map [17, 18].

1.5. **Outline of the paper.** We end with an outline of the paper and the proof of Theorem 1.1. In §2, we give some background connecting the dynamics on the modular surface with the group structure on $\mathrm{SL}_2(\mathbb{R})$. In §3, we show that up to a scalar multiple, there is at most one linear form on the smooth vectors of an irreducible unitary representation of $\mathrm{SL}_2(\mathbb{R})$ that is invariant under the action of the diagonal subgroup (corresponding to the geodesic flow) and the element $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ corresponding to time-reversal symmetry. We show that such a linear form is determined by its value on a "minimal" vector—a spherical vector in the case of a principal series representation and a lowest/highest weight vector for holomorphic/antiholomorphic discrete series representations. We then bring in invariance under orientation-reversal and apply the results to show that the bilinear forms V^{sym} and B are determined by their values on Maass forms and holomorphic modular forms.

In §4, we present some background on half-integral weight forms, and in §5 we discuss Rankin–Selberg theory for these, giving a mean-square result for Fourier coefficients along positive integers by modifying work of Matthes [27] for weight zero forms.

In §6, we review the results of Maass [25], Shintani [39], Kohnen [14, 15] and Katok–Sarnak [13], relating periods along closed geodesics to Fourier coefficients of theta-lifts. This allows us to express $\mu_d(F)$ in terms of Fourier coefficients of half-integral weight forms on $\Gamma_0(4)$; the precise normalizations in terms of the inner products of the forms and their theta-lifts are crucial here. This is where the factor $L(1/2,\pi)$ appears. These results put us in a position to use the Rankin–Selberg theory of §5 to determine the variance B, which we do in §7.

2. BACKGROUND ON PERIODS

2.1. **The upper half-plane and its unit tangent bundle.** We recall the hyperbolic metric on the tangent bundle of the upper half-plane $\mathbb{H} = \{z = x + iy : y > 0\}$. We identify the tangent space at $z \in \mathbb{H}$ with the complex numbers, $T_z\mathbb{H} \simeq \mathbb{C}$. The hyperbolic metric on $T_z\mathbb{H}$ is then given by

$$\langle \xi, \eta \rangle_z := \frac{\Re(\xi\bar{\eta})}{y^2},$$

and the unit tangent bundle $S\mathbb{H}$ is then identified with

$$\{(z,\zeta)\in\mathbb{H}\times\mathbb{C} : |\zeta| = \Im(z)\}.$$

2.1.1. *Isometries.* A unimodular matrix $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ acts on the upper half-plane $\mathbb{H}$ via $z \mapsto (az+b)/(cz+d)$. Set

$$j(g,z) = cz+d.$$

The differential of the map is $g'(z) = (ad-bc)/(cz+d)^2 = (cz+d)^{-2} = 1/j(g,z)^2$. The induced map on the tangent bundle $T\mathbb{H}$ is then

$$(z,\xi) \mapsto (g(z), g'(z)\xi).$$

Note that this is an action: if $g,h\in\mathrm{SL}_2(\mathbb{R})$, then $g(h(z,\zeta)) = (gh)(z,\zeta)$. A computation shows that we get an isometry of $\mathbb{H}$:

$$\langle \xi,\eta\rangle_z = \langle g'(z)\xi, g'(z)\eta\rangle_{g(z)}.$$

2.1.2. *Group theory.* Define matrices

$$n(x) = \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}, \quad a(y) = \begin{pmatrix} y^{1/2} & \\ & y^{-1/2} \end{pmatrix}, \quad \kappa(\phi) = \begin{pmatrix} \cos(\phi/2) & \sin(\phi/2) \\ -\sin(\phi/2) & \cos(\phi/2) \end{pmatrix}.$$

The rotation $\kappa(\phi)$ preserves the base point $i=\sqrt{-1}\in\mathbb{H}$. Note that

$$\kappa(\phi+2\pi) = -\kappa(\phi),$$

and thus we get the same element in $\mathrm{PSL}_2(\mathbb{R})$.

Setting $g_{x,y,\phi} = n(x)a(y)\kappa(\phi)$, we find

$$g_{x,y,\phi}(i,i) = (x+iy, iye^{i\phi}),$$

such that using the basepoint $(i,i)\in S\mathbb{H}$ of the upward pointing unit vector at $i=\sqrt{-1}\in\mathbb{H}$, we get a bijection

$$\mathrm{PSL}_2(\mathbb{R}) \simeq S\mathbb{H}, \qquad g\mapsto g(i,i).$$

We may then identify functions on $\mathrm{PSL}_2(\mathbb{R})$ and on $S\mathbb{H}$; if $F(z,\zeta)$ is a function on $S\mathbb{H}$, we may define $\tilde{F}$ on $\mathrm{SL}_2(\mathbb{R})$ by

$$\tilde{F}(g) := F(g(i,i)),$$

so $\tilde{F}(g_{x,y,\phi}) = F(x+iy, iye^{i\phi})$.

2.1.3. *Geodesics.* The geodesic flow on $S\mathbb{H}$ is defined by $\Phi^t : (z,\zeta)\mapsto (z(t),\zeta(t))$, which is the endpoint of the (unit speed) geodesic starting at z in direction $\zeta = iye^{i\phi}$. It turns out that on $\mathrm{PSL}_2(\mathbb{R})$, the geodesic flow is multiplication on the right by $\begin{pmatrix} e^{t/2} & \\ & e^{-t/2} \end{pmatrix}$; that is,

$$\Phi^t(z,\zeta) = g_{x,y,\phi}\begin{pmatrix} e^{t/2} & \\ & e^{-t/2} \end{pmatrix}(i,i).$$

Indeed, for an initial position $(z,\zeta)\in S\mathbb{H}$, we write $(z,\zeta)=g(i,i)$, and the the geodesic $\vec{\gamma}(t)=\Phi^t(z,\zeta)$ starting at (z,ζ) will be the translate by $g_{x,y,\phi}$ of the geodesic $\vec{\gamma}_0(t)$ starting at the initial condition (i,i): $\vec{\gamma}(t)=g_{x,y,\phi}\vec{\gamma}_0(t)$. A computation shows that

$$\vec{\gamma}_0(t)=(e^t i,e^t i)=\begin{pmatrix} e^{t/2} & \\ & e^{-t/2}\end{pmatrix}(i,i),$$

and therefore

$$\vec{\gamma}(t)=g_{x,y,\phi}\begin{pmatrix} e^{t/2} & \\ & e^{-t/2}\end{pmatrix}(i,i).$$

2.1.4. *Time-reversal symmetry.* A fundamental symmetry of phase space $S\mathbb{H}$ is time-reversal $(z,\zeta)\mapsto(z,-\zeta)$. Using it, one has a symmetry of the set of geodesics, corresponding to reversing the orientation. In $\mathrm{PSL}_2(\mathbb{R})$, it is given as $g\mapsto gw$, where $w=\begin{pmatrix}0&-1\\1&0\end{pmatrix}$. Indeed, if $(z,\zeta)=g(i,i)\in S\mathbb{H}$, then

$$(z,-\zeta)=g(i,-i)=g\begin{pmatrix}0&-1\\1&0\end{pmatrix}(i,i)$$

2.1.5. *Orientation-reversal.* Orientation-reversal, $(z,\zeta)\mapsto(-\overline{z},-\overline{\zeta})$, is another fundamental symmetry. On $\mathrm{PSL}_2(\mathbb{R})$, it is given by the map

$$g\mapsto\delta g\delta,\quad \delta=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

2.1.6. *K-types.* Let k be an integer. Suppose that $F\colon S\mathbb{H}\to\mathbb{C}$ satisfies

$$F(z,e^{i\alpha}\zeta)=e^{ik\alpha}F(z,\zeta).$$

Then the corresponding function $\tilde{F}$ on $\mathrm{PSL}_2(\mathbb{R})$ satisfies

$$\tilde{F}(g\kappa(\alpha))=e^{ik\alpha}\tilde{F}(g),$$

i.e., it transforms under the right action of the maximal compact $K=SO(2)/\{\pm I\}$ via the character $\kappa(\alpha)\mapsto e^{ik\alpha}$. As an example, we start with a function f on $\mathbb{H}$ and define $F_f(z,\zeta)=\zeta^k f(z)$.

2.2. **Quotients.** Let $\Gamma\subset\mathrm{PSL}_2(\mathbb{R})$ be a Fuchsian group , $M=\Gamma\backslash\mathbb{H}$, and SM the unit tangent bundle to M. The identification $S\mathbb{H}\simeq\mathrm{PSL}_2(\mathbb{R})$ descends to an identification

$$SM\simeq\Gamma\backslash S\mathbb{H}\simeq\Gamma\backslash\mathrm{PSL}_2(\mathbb{R}).$$

2.2.1. *Automorphy conditions.* Let $k\geq 0$ be an integer and $f:\mathbb{H}\to\mathbb{C}$ a function on the upper half-plane satisfying the (weak) automorphy condition

$$f(\gamma(z))=(cz+d)^{2k}f(z),\qquad \forall\gamma=\begin{pmatrix}a&b\\c&d\end{pmatrix}\in\Gamma. \tag{2.1}$$

We define F_f on $S\mathbb{H}$ by

$$F_f(z,\zeta):=\zeta^k f(z).$$

Then

$$F_f(\gamma(z,\zeta))=F_f(z,\zeta),\qquad\forall\gamma\in\Gamma,$$

that is, F_f is Γ-invariant, so it descends to a function on $SM = \Gamma\backslash S\mathbb{H}$, and via the identification $F \mapsto \tilde{F}$ gives a Γ-invariant function $\tilde{F}_f$ on $\mathrm{PSL}_2(\mathbb{R})$:

$$\tilde{F}_f(\gamma g) = \tilde{F}_f(g), \qquad \forall \gamma \in \Gamma.$$

Moreover, F_f has K-type k, since from the definition we find

$$F_f(z, e^{i\alpha}\zeta) = (e^{i\alpha}\zeta)^k f(z) = e^{ik\alpha} F_f(z,\zeta),$$

and therefore the function $\tilde{F}_f$ on the group $\mathrm{PSL}_2(\mathbb{R})$ transforms under the right action of $K = SO(2)/\{\pm I\}$ by the character $\kappa(\alpha) \mapsto e^{ik\alpha}$.

2.2.2. *Closed geodesics on M*. We consider closed geodesics on M, that is an initial condition $(z_0,\zeta_0) \in S\mathbb{H}$ such that there is some $T > 0$ and $\gamma \in \Gamma$ with

$$\Phi^T(z_0,\zeta_0) = \gamma(z_0,\zeta_0).$$

Writing $(z_0,\zeta_0) = g_0(i,i)$ for a unique $g_0 \in \mathrm{PSL}_2(\mathbb{R})$ we find that

$$\Phi^T(z_0,\zeta_0) = g_0 \begin{pmatrix} e^{T/2} & \\ & e^{-T/2} \end{pmatrix} (i,i) = \gamma g_0(i,i)$$

and hence that

$$\gamma = \pm g_0 \begin{pmatrix} e^{T/2} & \\ & e^{-T/2} \end{pmatrix} g_0^{-1} \tag{2.2}$$

(the equality is in $\mathrm{PSL}_2(\mathbb{R})$, that is the matrices agree up to a sign).

Changing the initial condition (z_0,ζ_0) to a Γ-equivalent one $(z_1,\zeta_1) = \delta(z_0,\gamma_0)$, $\delta \in \Gamma$ (so we get the same point in $SM = \Gamma\backslash S\mathbb{H}$) replaces γ by its conjugate $\delta\gamma\delta^{-1}$. Thus we get a well-defined conjugacy class γ_C corresponding to the geodesic C. The conjugacy class is *hyperbolic*, as its trace satisfies $|\mathrm{tr}\,\gamma_C| = 2\cosh(T/2) > 2$.

2.3. **A correspondence with binary quadratic forms.** An binary quadratic form $f(x,y) = ax^2 + bxy + cy^2$ (also denoted by $[a,b,c]$) is called *integral* if a,b,c are integers, and is called *primitive* if $\gcd(a,b,c) = 1$. The discriminant of f is $b^2 - 4ac$. The modular group $\mathrm{SL}_2(\mathbb{Z})$ acts on the set of integral binary quadratic forms by substitutions, and this action preserves the discriminant.

There is a bijection between $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of primitive binary quadratic forms of positive (nonsquare) discriminant and primitive hyperbolic conjugacy classes in $\mathrm{PSL}_2(\mathbb{Z})$ defined as follows. Given a primitive hyperbolic element

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

the corresponding form is

$$B(\gamma) = \frac{\mathrm{sign}(a+d)}{\gcd(b, d-a, -c)}[b, d-a, -c], \tag{2.3}$$

which is primitive by definition, and has discriminant

$$\mathrm{disc}(B(\gamma)) = \frac{(\mathrm{tr}\,\gamma)^2 - 4}{\gcd(b, d-a, -c)^2}.$$

Moreover,

$$B(-\gamma) = B(\gamma), \qquad B(\gamma^{-1}) = -B(\gamma).$$

Given a primitive integral binary quadratic form $f = [a, b, c]$ of positive nonsquare discriminant $d = b^2 - 4ac$, let (t_0, u_0) be the fundamental solution of the Pell equation $t^2 - du^2 = 4$ with $t_0 > 0$, $u_0 > 0$ (which exists since we assume $d > 0$ is not a perfect square). Define the matrix

$$\gamma(f) := \begin{pmatrix} \frac{t_0 - bu_0}{2} & au_0 \\ -cu_0 & \frac{t_0 + bu_0}{2} \end{pmatrix},$$

which is hyperbolic of trace $t_0 = \sqrt{du^2 + 4} > 2$ and is primitive. Then $B(\gamma(f)) = f$, and gives a bijection between primitive hyperbolic conjugacy classes in $\mathrm{PSL}_2(\mathbb{Z})$ and equivalence classes of primitive binary quadratic forms of nonsquare positive discriminant.

2.4. **Periods.** Consider a (primitive, oriented) closed geodesic on M; it is determined by a primitive hyperbolic conjugacy class $\gamma \in \Gamma$. Let C be the lift of the closed geodesic to to the unit tangent bundle SM. For any function F on SM, we define the period of F along C by choosing a point on the lifted geodesic (z_0, ζ_0) (that is an initial condition) and setting

$$\int_C F := \int_0^T F \circ \Phi^t(z_0, \zeta_0)\, dt,$$

where $T > 0$ is the length of the geodesic, that is, the first time that $\Phi^T(z_0, \zeta_0) = \gamma(z_0, \zeta_0)$.

2.4.1. *An alternative expression for the period.* To a hyperbolic matrix $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, associate a binary quadratic form (not necessarily primitive)

$$Q_\gamma(z) = cz^2 + (d - a)z - b = j(\gamma, z)\left(z - \gamma(z)\right).$$

Note that $Q_{-\gamma} = -Q_\gamma$.

The two zeros $w_\pm$ of Q_γ are the the fixed points of γ, which are the intersection with real axis of the semicircle in the upper half-plane which determines the closed geodesic. By (2.2), the fixed points $w_\pm$ of γ on the boundary are $g_0(0)$ and $g_0(\infty)$: Indeed, $\gamma(w) = w$ if and only if

$$\begin{pmatrix} e^{T/2} & \\ & e^{-T/2} \end{pmatrix} g_0^{-1}(w) = g_0^{-1}(w),$$

that is, if and only if $e^T g_0^{-1}(w) = g_0^{-1}(w)$, and since $T \neq 0$, this forces $g_0^{-1}(w) = 0, \infty$. Thus we find

$$Q_\gamma(z) = C(z - g_0(0))(z - g_0(\infty)).$$

Let $f : \mathbb{H} \to \mathbb{C}$ satisfy the automorphy condition (2.1) of weight $2k$ for Γ, and set $F = F_f : (z, \zeta) \mapsto \zeta^k f(z)$, which is a Γ-invariant function on $S\mathbb{H}$, that is a function on SM which transforms under $SO(2)$ with K-type k. Let

$$r_k(f, \gamma) = \int_{z_0}^{\gamma z_0} f(z) Q_\gamma(z)^{k-1}\, dz,$$

where z_0 lies on the semicircle between the fixed points of γ and the contour of integration[3] is along the geodesic arc linking z_0 and γz_0.

Let

$$D_\gamma := \operatorname{tr}(\gamma)^2 - 4 = \operatorname{disc}(Q_\gamma)$$

be the discriminant of the quadratic form Q_γ. Then $r_k(f,\gamma)$ is simply related to the period of f on the geodesic defined by γ [12, Proposition 4]:

$$r_k(f,\gamma) = \left(-\operatorname{sign}(\operatorname{tr}(\gamma))\sqrt{D_\gamma}\right)^{k-1}\int_C F. \tag{2.4}$$

Thus, in terms of the corresponding binary quadratic form $B(x,y) = B(\gamma)(x,y)$ (2.3), we get

$$\int_C F = \frac{1}{(\operatorname{disc} B)^{(k-1)/2}}\int_{z_0}^{\gamma z_0} f(z)B(1,-z)^{k-1}\,dz =: J(B). \tag{2.5}$$

Note that the right-hand side above makes sense also for nonprimitive forms, and is dilation invariant: $J(tB) = J(B)$.

3. SYMMETRY CONSIDERATIONS

3.1. **Background on the representation theory of** $\mathrm{SL}_2(\mathbb{R})$**.** Let π be an irreducible infinite-dimensional unitary representation of $\mathrm{SL}_2(\mathbb{R})$ on a Hilbert space $\mathscr{H}$ that factors through $G = \mathrm{PSL}_2(\mathbb{R})$. Let $K = SO(2)$, and let $\mathscr{H}^{(K)}$ be the space of K-finite vectors in $\mathscr{H}$, *i.e.*, vectors whose translates by K span a finite-dimensional subspace. Then $\mathscr{H}^{(K)}$ is dense in $\mathscr{H}$ and consists of smooth vectors, and the Lie algebra $\mathfrak{sl}_2$ acts on $\mathscr{H}^{(K)}$ by $d\pi$, the differential of the action of G.

According to Bargmann's classification of such π's (we follow the exposition in Lang [21]), there are orthogonal one-dimensional subspaces $\mathscr{H}_n$, with n even, which are K-invariant and together span $\mathscr{H}^{(K)}$. To be more precise, we consider two cases:

i) There is no highest or lowest K-type; this is the spherical, or Maass case:

$$\mathscr{H}^{(K)} = \bigoplus_{n \text{ even}} \mathscr{H}_n, \tag{3.1}$$

with $\mathscr{H}_n$ one-dimensional for n even, say $\mathscr{H}_n = \mathbb{C}\phi_n$, and the ϕ_n satisfy

$$\begin{aligned} d\pi(W)\phi_n &= in\phi_n \\ d\pi(E^-)\phi_n &= (s+1-n)\phi_{n-2} \\ d\pi(E^+)\phi_n &= (s+1+n)\phi_{n+2}, \end{aligned} \tag{3.2}$$

where

$$H = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \qquad V = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad W = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} \tag{3.3}$$

are the standard basis of the Lie algebra $\mathfrak{sl}_2(\mathbb{R})$,

$$E^\pm = H \pm iV \tag{3.4}$$

[3] If f is holomorphic, the integral is independent of the contour.

are in the complexified Lie algebra $\mathfrak{sl}_2(\mathbb{C})$, and $d\pi(E^{\pm})$ are the weight raising/lowering operators. Here $s \in \mathbb{C}$ is a parameter which, since we assume that π is unitary, lies on the imaginary axis $i\mathbb{R}$ or in the interval $(-1,1)$. Note that since we are assuming the representation factors through $G = \mathrm{PSL}_2(\mathbb{R})$, only even weights appear.

ii) $\mathcal{H}$ has a lowest or highest K-type.

In the first case, there is an even positive integer $m_0 > 0$ such that

$$\mathcal{H}^{(K)} = \bigoplus_{\substack{m=m_0 \\ m \text{ even}}}^{\infty} \mathcal{H}_m, \tag{3.5}$$

with $\mathcal{H}_m$ one-dimensional, say $\mathcal{H}_m = \mathbb{C}\phi_m$ and the ϕ_m satisfy (3.2) with $s = m_0 - 1$. In particular, ϕ_{m_0} is annihilated by the lowering operator:

$$d\pi(E^-)\phi_{m_0} = 0\,. \tag{3.6}$$

These π's correspond to holomorphic forms of even weight.

In the case there is a highest K-type, there is a negative even integer $m_0 < 0$ such that

$$\mathcal{H}^{(K)} = \bigoplus_{\substack{m=-\infty \\ m \text{ even}}}^{m_0} \mathcal{H}_m\,. \tag{3.7}$$

Again, $\mathcal{H}_m = \mathbb{C}\phi_m$ for $m \le m_0$ even, so ϕ_m satisfy (3.2) with $s = -m_0 - 1$ and the highest weight vector ϕ_{m_0} is annihilated by the raising operator:

$$d\pi(E^+)\phi_{m_0} = 0\,. \tag{3.8}$$

In case (i), we denote by ϕ_π the K-invariant (spherical) vector ϕ_0. We normalize it in such a way that $\langle \phi_0, \phi_0 \rangle = 1$, and then it is unique up to multiplication by a complex scalar of unit modulus. In case (ii), we denote by ϕ_π the similarly normalized lowest/highest weight vector ϕ_{m_0}. We will call these ϕ_π's "minimal vectors" of the representation.

3.2. **Linear forms.** We consider linear forms η on $\mathcal{H}^{(K)}$ that are invariant under the "geodesic flow" and "time-reversal symmetry", that is,

- η is annihilated by $d\pi(H)$, where $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathfrak{sl}_2$ is the infinitesimal generator of the group of diagonal matrices A:

$$\eta(d\pi(H)v) = 0, \qquad \forall v \in \mathcal{H}^{(K)} \tag{3.9}$$

(we say that η is A-invariant)[4].

- η is fixed by $\pi\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right)$:

$$\eta\left(\pi\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} v\right) = \eta(v), \qquad \forall v \in \mathcal{H}^{(K)} \tag{3.10}$$

[4]This choice of terminology is imprecise since $\pi(A)$ need not preserve the space of K-finite vectors on which η is a-priori defined.

(we say that η is invariant under time-reversal symmetry).

PROPOSITION 3.1. *Let π be an irreducible infinite-dimensional unitary representation of $\mathrm{SL}_2(\mathbb{R})$ on a Hilbert space $\mathscr{H}$ which factors through $G = \mathrm{PSL}_2(\mathbb{R})$. Then the space of linear forms η on $\mathscr{H}^{(K)}$ invariant under A and $\mathbf{w}$ is at most one-dimensional, and any such form is completely determined by its action on a "minimal" vector ϕ_π. In the case (ii) of discrete series, the space of A-invariant forms is one-dimensional, when $m = 2 \mod 4$, none of them is $\mathbf{w}$-invariant, and if $m = 0 \mod 4$, then any A-invariant form is automatically $\mathbf{w}$-invariant. In the spherical case, the space of linear forms invariant under A and $\mathbf{w}$ is one-dimensional.*

This is shown by giving an explicit formula for $\eta(\phi_n)$ in term of $\eta(\phi_\pi)$. Since the cases (i) and (ii) have slightly different features, we deal with them separately.

In case (i), $\phi_\pi = \phi_0$ is the spherical vector. We are assuming that η is invariant under time-reversal symmetry, that is that (3.10) holds. Since

$$\pi\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\phi_n = -\phi_n$$

if $n \equiv 2 \mod 4$, due to (3.2) and $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \exp(\frac{\pi}{2}W)$, it follows from (3.10) that

$$\eta(\phi_n) = 0, \quad \text{if } n \equiv 2 \mod 4\,. \tag{3.11}$$

Now $2H = E^+ + E^-$, and from (3.2) we have

$$2d\pi(H)\phi_n = \left(d\pi(E^+) + d\pi(E^-)\right)\phi_n = (s+1-n)\phi_{n-2} + (s+1+n)\phi_{n+2}\,.$$

Hence

$$\eta(2d\pi(H)\phi_n) = (s+1-n)\eta(\phi_{n-2}) + (s+1+n)\eta(\phi_{n+2})\,. \tag{3.12}$$

But the LHS of (3.12) is zero, since we are assuming (3.9). Hence for n even, and in particular $n \equiv 2 \mod 4$, we have

$$(n-s-1)\eta(\phi_{n-2}) = (n+s+1)\eta(\phi_{n+2})\,.$$

It follows for $m \geq 4$, $m \equiv 0 \mod 4$ that

$$\eta(\phi_m) = \eta(\phi_{-m}) = \frac{(1-s)(5-s)\cdot\dots\cdot(m-3-s)}{(3+s)(5+s)\cdot\dots\cdot(m-1+s)}\eta(\phi_0)\,. \tag{3.13}$$

This, together with (3.11), determines η on $\mathscr{H}^{(K)}$ explicitly in terms of $\eta(\phi_0)$.

Conversely, (3.11) and (3.13) with $\eta(\phi_0) = 1$ define a unique A- and $\mathbf{w}$-invariant linear form on $\mathscr{H}$, which we denote by ξ_{π,ϕ_π}. So, in this case, the space of such linear forms is one-dimensional and any such form η satisfies

$$\eta = \eta(\phi_0)\xi_{\pi,\phi_\pi}\,.$$

We turn to case (ii) and show that the space of A-invariant linear forms on $\mathscr{H}^{(K)}$ is one-dimensional. Consider the lowest weight case: take the lowest weight vector ϕ_{m_0}, $m_0 > 0$ and even. From (3.2) and (3.6), we have

$$2d\pi(H)\phi_{m_0} = \left(d\pi(E^+) + d\pi(E^-)\right)\phi_{m_0} = 2m_0\phi_{m_0+2}\,.$$

Hence, assuming η is A-invariant, we get that

$$\eta(\phi_{m_0+2}) = 0\,.$$

Furthermore, for $m > m_0$ even and by (3.2), we have

$$(m - m_0)\eta(\phi_{m-2}) = (m + m_0)\eta(\phi_{m+2})\,.$$

Hence

$$\eta(\phi_m) = 0, \qquad \text{for } m \geq m_0, \quad m \equiv m_0 + 2 \mod 4 \tag{3.14}$$

and

$$\eta(\phi_{m_0+k}) = \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (\frac{k}{2} - 1)}{(m_0 + 1) \cdot (m_0 + 3) \cdot \dots \cdot (m_0 + \frac{k}{2} - 1)} \eta(\phi_{m_0}) \tag{3.15}$$

for $k \equiv 0 \mod 4$, $k \geq 4$.

Thus the space of A-invariant linear forms on $\mathscr{H}^{(K)}$ is one-dimensional. It is spanned by $\xi_{\mathscr{H}_\pi,\phi_\pi}$ where $\xi_{\mathscr{H}_\pi,\phi_\pi}(\phi_{m_0}) = 1$ and is defined by (3.14) and (3.15). Again, any A-invariant linear form η on $\mathscr{H}^{(K)}$ satisfies

$$\eta = \eta(\phi_\pi)\xi_{\mathscr{H}_\pi,\phi_\pi}\,. \tag{3.16}$$

The case of highest weight vectors and A-invariant forms is the same.

If we now impose the further condition that η be **w**-invariant for the case (ii) representations, then invariance under time-reversal symmetry gives, as in (3.11), that

$$\eta(\phi_m) = 0, \qquad \text{for } m \equiv 2 \mod 4\,.$$

This, coupled with (3.14), means that if $m_0 \equiv 2 \mod 4$, then $\eta = 0$. That is, if $m_0 \equiv 2 \mod 4$, then there is no nonzero linear form invariant under A and **w**.

If $m_0 \equiv 0 \mod 4$, then from our discussion, every A-invariant linear form is automatically **w**-invariant and in this case such linear forms satisfy (3.16).

3.3. **Orientation-reversal symmetry.** We now examine the role of an additional possible symmetry, "orientation-reversal" **r**. It need not act on irreducible representations of $\mathrm{PSL}_2(\mathbb{R})$. What we do is, given an irreducible unitary representation π on a Hilbert space $\mathscr{H}$, we consider Hilbert spaces $\mathscr{U}$, which in the spherical case is the original representation $\mathscr{H}$ and in the case of the discrete series $\mathscr{H}^m$, where there is a lowest weight vector of weight $m > 0$, we define

$$\mathscr{U} = \mathscr{H}^{+m} \oplus \mathscr{H}^{-m}$$

to be the direct sum of the irreducible representations with lowest weight m and that with highest weight $-m$. We write $\mathscr{U}^{(K)}$ for the dense subspace of K-finite vectors in $\mathscr{U}$.

An orientation-reversing symmetry of $\mathscr{U}$ is a unitary map **r** of $\mathscr{U}$ which is an involution, that is,

$$\mathbf{r}^2 = I \tag{3.17}$$

satisfying

$$\mathbf{r}\pi(W) = -\pi(W)\mathbf{r} \tag{3.18}$$

and

$$\mathbf{r}\pi(E^+) = \pi(E^-)\mathbf{r}\,. \tag{3.19}$$

As a consequence of (3.19) and (3.17), we have

$$\mathbf{r}\pi(E^-) = \pi(E^+)\mathbf{r}\,. \tag{3.20}$$

Moreover, $\mathbf{r}$ commutes with the A-action whose infinitesimal generator is $H = \frac{1}{2}(E^+ + E^-)$ by (3.19) and (3.20), and with time-reversal symmetry, that is, with[5] $\exp(\frac{\pi}{2}\pi(W))$ by virtue of (3.18).

As our basic example, we consider the orientation-reversal involution on the function space $L^2(\Gamma\backslash G)$ given by

$$\mathbf{r}f(x) := f(\delta x \delta^{-1}), \quad \delta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The relations (3.18) and (3.19) hold, since for the Lie algebra elements H, V and W of (3.3), we have

$$\delta W = -W\delta, \quad \delta H = H\delta, \quad \delta V = -V\delta\,.$$

3.4. **Action of r on weight vectors.** We first note that due to the commutation relation (3.18), $\mathbf{r}$ must reverse weights, that is

$$\mathbf{r}\phi_n = c_n \phi_{-n}\,,$$

with $|c_n| = 1$ since $\mathbf{r}$ is unitary, and $c_n c_{-n} = 1$ since $\mathbf{r}^2 = I$. In particular, in the spherical case when there is a vector ϕ_0 of weight 0, we must have

$$\mathbf{r}\phi_0 = \epsilon\phi_0, \quad \epsilon = \pm 1\,. \tag{3.21}$$

We say the spherical representation $\mathscr{U}$ is *even* if if the sign is + and *odd* if the sign is −.

In the case of the discrete series representations $\mathscr{U} = \mathscr{H}^{+m} \oplus \mathscr{H}^{-m}$, $m > 0$, we choose a lowest weight vector $\phi_m \in \mathscr{H}^{+m}$ of unit length, which we call the minimal (or generating) vector, that $\mathbf{r}\phi_m$ is a unit vector of weight $-m$, and normalize a choice of highest weight vector of unit length by taking

$$\phi_{-m} := \mathbf{r}\phi_m\,. \tag{3.22}$$

We claim that the choice of minimal vector ϕ_0 in the spherical case and ϕ_m in the discrete series case uniquely determine $\mathbf{r}$.

Indeed, starting with ϕ_m, the lowest weight vector for $m > 0$, we get from (3.2) for $k > 0$,

$$\phi_{m+2k} = \frac{1}{c(s;m,k)}\pi(E^+)^k\phi_m, \quad c(s;m,k) = \prod_{j=0}^{k-1}(s+m+1+2j),$$

and for the highest weight vector $\phi_{-m} = \mathbf{r}\phi_m$,

$$\phi_{-m-2k} = \frac{1}{c(s;m,k)}\pi(E^-)^k\phi_{-m}$$

[5]The first π is the constant 3.1415...!

Therefore, using (3.19),

$$\begin{aligned}\mathbf{r}\phi_{m+2k} &= \frac{1}{c(s;m,k)}\mathbf{r}\pi(E^+)^k\phi_m \\ &= \frac{1}{c(s;m,k)}\pi(E^-)^k\mathbf{r}\phi_m \\ &= \frac{1}{c(s;m,k)}\pi(E^-)^k\phi_{-m} = \phi_{-m-2k},\end{aligned}$$

and likewise,

$$\mathbf{r}\phi_{-m-2k} = \phi_{m+2k}.$$

That is, for the discrete series, $\mathbf{r}$ exactly interchanges ϕ_n and ϕ_{-n}:

$$\mathbf{r}\phi_n = \phi_{-n}, \quad |n| \geq m, \quad n = m \mod 2. \tag{3.23}$$

In the case of the spherical representations, the same analysis shows that

$$\mathbf{r}\phi_n = \epsilon\phi_{-n}, \quad n \in 2\mathbb{Z}, \tag{3.24}$$

where $\epsilon = \pm 1$ is determined by (3.21).

3.5. **r-invariant functionals.** If η is a linear functional on $\mathscr{U}$, invariant under the action of A and time-reversal symmetry $\mathbf{w}$, then the functional $\eta^{\mathbf{r}}: v \mapsto \eta(\mathbf{r}v)$ is also invariant under A and $\mathbf{w}$, since $\mathbf{r}$ commutes with A and with $\mathbf{w}$. We wish to determine when $\eta^{\mathbf{r}} = \eta$.

PROPOSITION 3.2. *The space of linear functionals on $\mathscr{U}^{(K)}$ which are invariant under A, $\mathbf{w}$, and $\mathbf{r}$ is at most one-dimensional. In the spherical case, there are no such functionals for odd representations and the space is one-dimensional in the even case, every functional invariant under A and $\mathbf{w}$ being automatically $\mathbf{r}$-invariant. For the discrete series, there are no such functionals for weight $m = 2 \mod 4$, and for weight $m = 0 \mod 4$, the space of A-invariant functionals is two-dimensional, each is automatically invariant under $\mathbf{w}$, and the subspace of $\mathbf{r}$-invariant functionals is one-dimensional.*

Proof. We start with the spherical case. There is a one-dimensional space of functionals invariant under A and $\mathbf{w}$, and we take the unique one satisfying

$$\eta(\phi_0) = 1\,.$$

Hence $\eta^{\mathbf{r}}$, being itself invariant under A and $\mathbf{w}$, must be a multiple of η, and because $\mathbf{r}^2 = I$, we have

$$\eta^{\mathbf{r}} = \pm\eta\,.$$

We claim the sign is determined by the sign in (3.21), that is, if $\mathbf{r}\phi_0 = \epsilon\phi_0$, then

$$\eta^{\mathbf{r}} = \epsilon\eta\,.$$

It suffices to check this on the spherical vector ϕ_0, that is, to show $\eta^{\mathbf{r}}(\phi_0) = \epsilon$. Indeed, we have

$$\eta^{\mathbf{r}}(\phi_0) = \eta(\mathbf{r}\phi_0) = \eta(\epsilon(\phi_0) = \epsilon\eta(\phi_0) = \epsilon,$$

as required. Thus, in the odd case, $\eta^{\mathbf{r}} = -\eta$ and there are no A- and $\mathbf{w}$-invariant functionals which are $\mathbf{r}$-invariant, and in the even case, $\eta^{\mathbf{r}} = \eta$ and every A- and $\mathbf{w}$-invariant functional is automatically $\mathbf{r}$-invariant.

In the discrete series case, there are no functionals invariant under A and $\mathbf{w}$ if $m = 2 \mod 4$, hence we only consider the case $m = 0 \mod 4$. In that case, there are unique A-invariant functionals η_+ on $\mathcal{H}^{+m}$ and η_- on $\mathcal{H}^{-m}$ satisfying

$$\eta_+(\phi_m) = 1, \quad \eta_-(\phi_{-m}) = 1,$$

and these are automatically invariant under time-reversal symmetry. Hence the space of A-invariant functionals on $\mathcal{U} = \mathcal{H}^{+m} \oplus \mathcal{H}^{-m}$ is two-dimensional, consisting of linear combinations

$$\eta = c_+\eta_+ \oplus c_-\eta_-$$

and these are automatically invariant under time-reversal symmetry. They are uniquely determined by their action on the lowest and highest weight vectors ϕ_m and $\phi_{-m} = \mathbf{r}\phi_m$:

$$c_\pm = \eta(\phi_{\pm m})$$

Since $\eta^{\mathbf{r}}$ is also A-invariant, we have

$$\eta^{\mathbf{r}} = c'_+\eta_+ + c'_-\eta_-.$$

Now $\eta^{\mathbf{r}} = \eta$ if and only if $c'_+ = c_+$ and $c'_- = c_-$. We claim that this happens if and only if $c_+ = c_-$, which will show that the space of A-invariant functionals which are $\mathbf{r}$-invariant is exactly one-dimensional in this case. Indeed, we have

$$c'_+ = \eta^{\mathbf{r}}(\phi_m) = \eta(\mathbf{r}\phi_m) = \eta(\phi_{-m}) = c_-$$
$$c'_- = \eta^{\mathbf{r}}(\phi_{-m}) = \eta(\mathbf{r}\phi_{-m}) = \eta(\phi_m) = c_+,$$

and so $c'_\pm = c_\pm$ if and only if $c_+ = c_-$, as claimed. □

3.6. **Bilinear forms.** We apply the uniqueness of linear forms to bi-invariant sesquilinear forms on $\mathcal{U} \times \mathcal{U}$. Let $T(v, v')$ be such a form, that is, it is linear in v, conjugate-linear in v', and invariant under A, $\mathbf{w}$ and $\mathbf{r}$ in each variable separately. For instance, we can take

$$T(v, v') = \sum_{j=1}^{J} \eta_j(v)\overline{\eta'_j(v')},$$

where η_j, η'_k are invariant linear forms. From the prior discussion,

$$T(v, v') = 0$$

if π is of type (ii) with $m_0 \equiv 2 \mod 4$. Otherwise, T is completely determined by the value $T(\phi_\pi, \phi_\pi)$ at the minimal vector ϕ_π. In fact, T is the product of linear forms

$$T(v, v') = T(\phi_\pi, \phi_\pi)\xi_{\mathcal{U},\phi_\pi}(v)\overline{\xi_{\mathcal{U},\phi_\pi}(v')},$$

where $\xi_{\mathcal{U},\phi_\pi}$ is the unique invariant linear form taking value 1 at the minimal vector ϕ_π.

3.7. **Application to the classical and arithmetic variances.** We apply these remarks to the measures μ_d and to the classical variance V. We consider the discrete decomposition of the regular representation of $G = \mathrm{PSL}_2(\mathbb{R})$ on $L^2_{\mathrm{cusp}}(\Gamma\backslash G)$. For an irreducible subrepresentation, form the space $\mathcal{U}_\pi$ as above.

3.7.1. The arithmetic measure μ_d is a linear form on $\mathcal{U}_\pi$ invariant under A, $\mathbf{w}$, and $\mathbf{r}$. Hence $\mu_d(F) \equiv 0$ if π is a discrete series with weight $m_0 \equiv 2 \mod 4$, and otherwise

$$\mu_d = \mu_d(\phi_\pi)\xi_{\mathcal{U}_\pi,\phi_\pi}\,. \tag{3.25}$$

Hence, if F_1 and F_2 are in $\mathcal{U}_{\pi_1}$ and $\mathcal{U}_{\pi_2}$, the sesquilinear μ_d sums take the form

$$\sum_{d\le Y}\frac{\mu_d(F_1)}{\sqrt{d}}\frac{\mu_d(F_2)}{\sqrt{d}} = \xi_{\mathcal{U}_{\pi_1},\phi_{\pi_1}}(F_1)\overline{\xi_{\mathcal{U}_{\pi_2},\phi_{\pi_2}}(F_2)}\sum_{d\le Y}\frac{\mu_d(\phi_{\pi_1})\mu_d(\phi_{\pi_2})}{d}\,.$$

This gives a universal reduction for computation of the variance of μ_d to the cases $F_1 = \phi_{\pi_1}$, $F_2 = \phi_{\pi_2}$.

3.7.2. The classical variance V is, by its definition, diagonalized by the irreducibles in the decomposition of $L^2(\Gamma\backslash G)$. We define projections onto the set of $\mathbf{w}$-invariant functions

$$F^{\mathrm{even}} := \frac{1}{2}(F + F^{\mathbf{w}})$$

and onto the set of functions invariant under both $\mathbf{w}$ and $\mathbf{r}$,

$$F^{\mathrm{sym}} := \frac{1}{4}(F + F^{\mathbf{w}} + F^{\mathbf{r}} + F^{\mathbf{wr}}) = \frac{1}{2}(F^{\mathrm{even}} + (F^{\mathrm{even}})^{\mathbf{r}}).$$

Set

$$V^{ev}(F_1,F_2) = V(F_1^{\mathrm{even}},F_2^{\mathrm{even}}),\quad V^{\mathrm{sym}}(F_1,F_2) = V(F_1^{\mathrm{sym}},F_2^{\mathrm{sym}}).$$

We wish to completely determine V^{sym} and V^{even}.

For an irreducible π, V^{sym} vanishes on $\mathcal{U}_\pi$ if π is a discrete series representation of weight $m_0 \equiv 2 \mod 4$ and otherwise is given by

$$V^{\mathrm{sym}}(v,v') = V^{\mathrm{sym}}(\phi_\pi,\phi_\pi)\xi_{\mathcal{U}_\pi,\phi_\pi}(v)\overline{\xi_{\mathcal{U}_\pi,\phi_\pi}(v')}\,.$$

It remains to determine $V^{\mathrm{sym}}(\phi_\pi,\phi_\pi)$.

LEMMA 3.3.

i) For π spherical with parameter $s = ir$,

$$V^{sym}(\phi_0,\phi_0) = \frac{|\Gamma(\frac{1}{4}+ir)|^4}{2\pi|\Gamma(\frac{1}{2}+2ir)|^2}\langle\phi_0,\phi_0\rangle. \tag{3.26}$$

ii) For π a discrete series representation of weight $m = 0 \mod 4$ ($m > 0$),

$$V^{sym}(\phi_m,\phi_m) = \frac{1}{2}2^m\,\mathrm{Beta}\left(\frac{m}{2},\frac{m}{2}\right) \tag{3.27}$$

where Beta *is Euler's beta function.*

Proof. In the spherical case, we need to compute $V^{\mathrm{sym}}(\phi_0,\phi_0)$. For spherical representations, we saw that A-invariance and $\mathbf{w}$-invariance automatically imply $\mathbf{r}$-invariance; hence on such spherical $\mathscr{U}$,

$$V^{\mathrm{sym}}|_{\mathscr{U}\times\mathscr{U}} = V^{\mathrm{even}}|_{\mathscr{U}\times\mathscr{U}}\,.$$

Moreover, since ϕ_0 is spherical, $\mathbf{w}\phi_0=\phi_0$, and so $\phi_0^{\mathrm{even}}=\phi_0$. Thus $V^{\mathrm{sym}}(\phi_0,\phi_0)=V(\phi_0,\phi_0)$, which was computed in [24] to give

$$V^{\mathrm{sym}}(\phi_0,\phi_0)=V(\phi_0,\phi_0)=\frac{|\Gamma(\frac14+ir)|^4}{2\pi|\Gamma(\frac12+2ir)|^2}\langle\phi_0,\phi_0\rangle\,.$$

For π a discrete series representation of weight $m_0\equiv 0 \mod 4$ ($m>0$), A-invariance implies $\mathbf{w}$-invariance, hence

$$V^{\mathrm{even}}|_{\mathscr{U}\times\mathscr{U}} = V|_{\mathscr{U}\times\mathscr{U}},$$

and so

$$V^{\mathrm{sym}}(\phi_m,\phi_m)=\frac14\big(V(\phi_m,\phi_m)+V(\phi_m,\mathbf{r}\phi_m)+V(\mathbf{r}\phi_m,\phi_m)+V(\mathbf{r}\phi_m,\mathbf{r}\phi_m)\big)\,.$$

By its definition, V respects the orthogonal decomposition into irreducibles; as $\phi_m\in\pi_m$ and $\mathbf{r}\phi_m=\phi_{-m}\in\pi_{-m}$ lie in distinct irreducibles, we get $V(\phi_m,\mathbf{r}\phi_m)=0=V(\mathbf{r}\phi_m,\phi_m)$. Moreover, we have

$$V(\mathbf{r}F_1,\mathbf{r}F_2)=V(F_1,F_2)$$

for any F_1, F_2. To see this, note first that $\mathbf{r}$ is induced by the measure-preserving map $x\mapsto\delta x\delta^{-1}$ of $SX=\Gamma\backslash G$, and hence

$$\langle\mathbf{r}F_1,\mathbf{r}F_2\rangle=\int_{\Gamma\backslash G}\mathbf{r}F_1(x)\overline{\mathbf{r}F_2(x)}\,dx=\langle F_1,F_2\rangle.$$

Moreover, $\mathbf{r}$ commutes with the geodesic flow, and so

$$\begin{aligned}V(\mathbf{r}F_1,\mathbf{r}F_2)&=\int_{-\infty}^{\infty}\langle\pi\begin{pmatrix}e^{t/2}&0\\0&e^{-t/2}\end{pmatrix}\mathbf{r}F_1,\mathbf{r}F_2\rangle\,dt\\&=\int_{-\infty}^{\infty}\langle\mathbf{r}\pi\begin{pmatrix}e^{t/2}&0\\0&e^{-t/2}\end{pmatrix}F_1,\mathbf{r}F_2\rangle\,dt\\&=\int_{-\infty}^{\infty}\langle\pi\begin{pmatrix}e^{t/2}&0\\0&e^{-t/2}\end{pmatrix}F_1,F_2\rangle\,dt\\&=V(F_1,F_2)\,.\end{aligned}$$

Thus we find

$$V^{\mathrm{sym}}(\phi_m,\phi_m)=\frac12V(\phi_{-m},\phi_{-m})\,.$$

Let

$$f(x)=\langle\pi(x)\phi_{-m},\phi_{-m}\rangle,\qquad x\in G\,.$$

Then applying the raising operator E^+ via the regular representation gives an operator $\mathscr{L}_{E^+}$ satisfying

$$.\mathscr{L}_{E^+}f(x)=\langle\pi(x)(d\pi(E^+)\phi_{-m}),\phi_{-m}\rangle=0 \tag{3.28}$$

Also, by the unitarity of π,

$$f(k(\theta_1)xk(\theta_2)) = \langle \pi(k(\theta_1)xk(\theta_2))\phi_{-m}, \phi_{-m}\rangle \\ = e^{-im(\theta_1+\theta_2)}\langle \pi(x)\phi_{-m}, \phi_{-m}\rangle = e^{-im(\theta_1+\theta_2)} f(x). \tag{3.29}$$

Using the coordinates $k(\theta_1)\begin{pmatrix} e^{r/2} & 0 \\ 0 & e^{-r/2}\end{pmatrix} k(\theta_2)$ on G and the formula for $\mathscr{L}^+$ in these coordinates, we deduce from (3.28) and (3.29) that

$$f(k(\theta_1)\begin{pmatrix} e^{r/2} & 0 \\ 0 & e^{-r/2}\end{pmatrix} k(\theta_2)) = e^{-im(\theta_1+\theta_2)} g(r),$$

where g satisfies the ODE

$$2\frac{dg}{dr} = -\frac{\cosh r}{\sinh r} mg + \frac{m}{\sinh r} g, \tag{3.30}$$

and since we have normalized $\langle \phi_{-m}, \phi_{-m}\rangle = 1$, we have

$$g(0) = 1\,.$$

We integrate (3.30) and find that

$$g(r) = (\cosh\frac{r}{2})^{-m}.$$

Hence

$$V(\phi_{-m}, \phi_{-m}) = \int_{-\infty}^{\infty} (\cosh\frac{r}{2})^{-m}\, dr = 2^m \operatorname{Beta}(\frac{m}{2}, \frac{m}{2}).$$

Thus we find

$$V^{\mathrm{sym}}(\phi_m, \phi_m) = \frac{1}{2} 2^m \operatorname{Beta}(\frac{m}{2}, \frac{m}{2}).$$

□

4. HALF-INTEGRAL WEIGHT FORMS

4.1. **Basic properties.** Let Γ be a discrete subgroup of $SL_2(\mathbb{R})$ of finite covolume. Given a character $\chi\colon \Gamma \to S^1$, an automorphic function of weight k and character χ for Γ is a function $f\colon \mathbb{H} \to \mathbb{C}$ satisfying

$$f(\gamma z) = \chi(\gamma)\left(\frac{cz+d}{|cz+d|}\right)^k f(z), \quad \forall \gamma \in \Gamma$$

with suitable growth conditions at the cusps of Γ. It is cuspidal if it vanishes at the cusps.

The Laplacian of weight k is defined as

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

The Laplacian Δ_k maps forms of weight k to themselves and maps cusp forms to themselves. A Maass cusp form of weight k is a cuspidal automorphic function of weight k (for some character χ) which is an eigenfunction of Δ_k.

Let $W_{\kappa,\mu}$ be the standard Whittaker function, normalized in such a way that at infinity,

$$W_{\kappa,\mu}(y) \sim y^{\kappa} e^{-y/2}, \quad y \to \infty. \tag{4.1}$$

The asymptotic behavior of $W_{\kappa,\mu}(y)$ near $y = 0$ is

$$W_{\kappa,\mu}(y) \sim \frac{\Gamma(-2\mu)}{\Gamma(\frac{1}{2}-\mu-\kappa)} y^{\frac{1}{2}+\mu} + \frac{\Gamma(2\mu)}{\Gamma(\frac{1}{2}+\mu-\kappa)} y^{\frac{1}{2}-\mu}, \quad y \to 0 \tag{4.2}$$

for $\mu \neq 0$, while

$$W_{\kappa,0} \ll y^{1/2} \log y, \quad y \to 0. \tag{4.3}$$

The functions

$$f_k^{\pm}(z,s) := W_{\pm\frac{k}{2}, s-\frac{1}{2}}(4\pi y) e(\pm x)$$

are eigenfunctions of Δ_k with eigenvalue $\lambda = s(1-s)$.

A Maass cusp form F of weight k and eigenvalue $\lambda = s(1-s)$ has Fourier expansion

$$F(z) = \sum_{n\neq 0} \rho(n) f_k^{\operatorname{sign}(n)}(|n|z, s) = \sum_{n \neq 0} \rho(n) W_{\operatorname{sign}(n)k/2, s-\frac{1}{2}}(4\pi|n|y) e(nx).$$

The Petersson inner product is defined for a pair of (cuspidal) functions of the same weight k and character χ as

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

4.2. **Maass operators.** For any real k, define the raising operator

$$\mathsf{K}_k = \frac{k}{2} + y\left(i\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) = \frac{k}{2} + (z-\bar{z})\frac{\partial}{\partial z}$$

and the lowering operator

$$\Lambda_k = \frac{k}{2} + y\left(i\frac{\partial}{\partial x} - \frac{\partial}{\partial y}\right) = \frac{k}{2} + (z-\bar{z})\frac{\partial}{\partial \bar{z}}.$$

The raising operator K_k takes Maass forms of weight k to forms of weight $k+2$ and the lowering operator Λ_k takes Maass forms of weight k to forms of weight $k-2$.

Then

$$\mathsf{K}_k\Delta_k = \Delta_{k+2}\mathsf{K}_k, \qquad \Lambda_k\Delta_k = \Delta_{k-2}\Lambda_k.$$

The effect of the Maass operators on Petersson inner products is given as follows. If f, g have weight k and character χ, then

$$\langle \mathsf{K}_k f, \mathsf{K}_k g\rangle = \left(\lambda(s) - \lambda(-\frac{k}{2})\right)\langle f, g\rangle$$

and

$$\langle \Lambda_k f, \Lambda_k g\rangle = \left(\lambda(s) - \lambda(\frac{k}{2})\right)\langle f, g\rangle.$$

The action of the Maass operators on the eigenfunctions $f_k^{\pm}(z,s)$ is

$$\mathsf{K}_k f_k^+(z,s) = -f_{k+2}^+(z,s), \quad \mathsf{K}_k f_k^-(z,s) = (s+\frac{k}{2})(1-s+\frac{k}{2}) f_{k+2}^-(z,s) \tag{4.4}$$

and

$$\Lambda_k f_k^+(z,s) = -(s-\frac{k}{2})(1-s-\frac{k}{2}) f_{k-2}^+(z,s), \quad \Lambda_k f^-(z,s) = f_{k-2}^-(z,s). \tag{4.5}$$

4.3. **Maass operators and Fourier expansions.** We want to see the Fourier expansion of a "raised" Maass form in terms of its original.

So start with a Maass form F of weight $1/2$ and eigenvalue $\lambda = 1/4 + r^2$ with Fourier expansion

$$F(z) = \sum_{n\neq 0} \rho(n) W_{\mathrm{sign}(n)/4, ir}(4\pi|n|y) e(nx),$$

Applying the Maass raising operator $\mathsf{K}_{1/2}$, we get a form $\mathsf{K}_{1/2}F$ of weight $5/2$ whose Fourier expansion is obtained by (4.4) as

$$\mathsf{K}_{1/2}F(z) = \sum_{n=1}^{\infty} -\rho(n) W_{5/4,ir}(4\pi ny) e(nx) + \sum_{n=1}^{\infty} \left((\frac{3}{4})^2 + r^2 \right) \rho(-n) W_{-5/4,ir}(4\pi ny) e(-nx). \tag{4.6}$$

Applying the lowering operator $\Lambda_{1/2}$, we get a form $\Lambda_{1/2}F$ of weight $-3/2$ whose Fourier expansion is obtained by (4.5) as

$$\Lambda_{1/2}F(z) = \sum_{n=1}^{\infty} -\left((\frac{1}{4})^2 + r^2 \right) \rho(n) W_{-3/4,ir}(4\pi ny) e(nx) + \sum_{n=1}^{\infty} \rho(-n) W_{3/4,ir}(4\pi ny) e(-nx). \tag{4.7}$$

5. RANKIN–SELBERG THEORY

5.1. **Classical Rankin–Selberg theory.** We recall classical Rankin–Selberg theory as applied to a holomorphic form F of weight $k+1/2$ with Fourier expansion

$$F(z) = \sum_{d\geq 1} c_F(d) e(dz)\,.$$

Let $E(z,s)$ be the standard Eisenstein series for $\Gamma_0(4)$:

$$E(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} \Im(\gamma z)^s,$$

where

$$\Gamma_\infty = \{\pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} : n \in \mathbb{Z}\}.$$

The series is absolutely convergent for $\Re(s) > 1$, with an analytic continuation to $\Re(s) > 1/2$ except for a simple pole at $s = 1$, where the residue is

$$\mathrm{Res}_{s=1} E(z,s) = \frac{1}{\mathrm{vol}(\Gamma_0(4)\backslash\mathbb{H})} = \frac{1}{2\pi}.$$

One starts with the integral

$$I(s)=\int_{\Gamma_0(4)\backslash\mathbb{H}}|F(z)|^2E(z,s)y^{k+\frac{1}{2}}\,\frac{dx\,dy}{y^2},$$

which is analytic in $\Re(s)>1/2$ except for a simple pole at $s=1$ with residue

$$R(F)=\frac{\langle F,F\rangle}{2\pi}.$$

By the "unfolding trick", we have

$$I(s)=(4\pi)^{-(s+k-\frac{1}{2})}\Gamma(s+k-\frac{1}{2})\sum_{n=1}^{\infty}|\frac{c_F(n)}{n^{\frac{k-1/2}{2}}}|^2n^{-s},$$

and hence the Dirichlet series

$$D(s)=\sum_{n=1}^{\infty}|\frac{c_F(n)}{n^{\frac{k-1/2}{2}}}|^2n^{-s}$$

has a simple pole at $s=1$ with residue

$$\frac{(4\pi)^{k+\frac{1}{2}}}{\Gamma(k+\frac{1}{2})}\frac{\langle F,F\rangle}{2\pi}.$$

Consequently, we find

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}|\frac{c_F(n)}{n^{\frac{k-1/2}{2}}}|^2=\frac{(4\pi)^{k+\frac{1}{2}}}{\Gamma(k+\frac{1}{2})}\frac{\langle F,F\rangle}{2\pi}.\tag{5.1}$$

Similar considerations show that if we take forms F of weight $k+1/2$ and G of weight $\ell+1/2$ (k and ℓ possibly different) which are orthogonal, then we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{c_F(n)}{n^{\frac{k-1/2}{2}}}\frac{\overline{c_G(n)}}{n^{\frac{\ell-1/2}{2}}}=0,\tag{5.2}$$

and that if F is a Maass form of weight $1/2$ for $\Gamma_0(4)$ with Fourier expansion

$$F(x+iy)=\sum_{n\neq 0}\rho(n)W_{\mathrm{sign}(n)/4,ir}(4\pi|n|y)e(nx)$$

and G is a holomorphic form of weight $k+1/2$, then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\sqrt{n}\rho(n)\frac{\overline{c_G(n)}}{n^{\frac{k-1/2}{2}}}=0.\tag{5.3}$$

These arguments will also give the asymptotics of the sum of squares

$$\sum_{-N\le n\le N}|4\pi n\rho(n)|^2$$

of Fourier coefficients with both positive and negative indices. However, for our application we need to be able to separately sum only coefficients indexed by positive integers; that is, we require the asymptotics of the series

$$\sum_{n=1}^{N}|4\pi n\rho(n)|^2.$$

To do so, we make use of the arguments in the paper by Matthes [27], which we adapt for our case; see also [26, 28].

5.2. **One-sided Rankin–Selberg theory.** Let $F(z)$ and $F'(z)$ be Maass cusp forms of weight 1/2 for $\Gamma_0(4)$ and Laplace eigenvalues $\lambda = \frac{1}{4} + r^2$, $\lambda' = \frac{1}{4} + (r')^2$ with Fourier expansion

$$F(x+iy) = \sum_{n\neq 0} \rho(n) W_{\mathrm{sign}(n)/4, ir}(4\pi|n|y)e(nx)$$
$$F'(x+iy) = \sum_{n\neq 0} \rho'(n) W_{\mathrm{sign}(n)/4, ir'}(4\pi|n|y)e(nx).$$

We define two Dirichlet series

$$L_+(s, F\times F') = \sum_{n=1}^{\infty} \frac{4\pi n \rho(n)\overline{\rho'(n)}}{n^s}$$
$$L_-(s, F\times F') = \sum_{n=1}^{\infty} \frac{4\pi n \rho(-n)\overline{\rho'(-n)}}{n^s}.$$

PROPOSITION 5.1. *Both $L_\pm(s)$ have analytic continuation to $\Re(s) > \frac{1}{2}$, except for a simple pole at $s = 1$ if F and F' are not orthogonal.*

We next compute the residue at $s = 1$ when $F' = F$:

PROPOSITION 5.2. *The residue at $s = 1$ of $L^+(s, F\times F)$ is*

$$R^+ := Res_{s=1} L_+(s, F\times F) = \frac{|\Gamma(\frac{1}{4}+ir)|^2}{|\Gamma(\frac{1}{2}+2ir)|^2}\frac{\langle F,F\rangle}{\pi} \tag{5.4}$$

and the residue of $L_-(s, F\times F)$ is

$$R^- = \frac{|\Gamma(\frac{3}{4}+ir)|^2}{|\Gamma(\frac{1}{2}+2ir)|^2}\frac{\langle F,F\rangle}{\pi}.$$

The arguments and ideas needed to establish this have been essentially provided in [27].

As a consequence, we deduce by a standard Tauberian argument the following.

COROLLARY 5.3. *Let F be as above. Then*

$$\sum_{1\le n\le N} 4\pi n|\rho(n)|^2 \sim R^+ N, \quad N\to\infty, \tag{5.5}$$

while if F and F' are orthogonal, then

$$\sum_{1\le n\le N} 4\pi n \rho(n)\overline{\rho'(n)} = o(N). \tag{5.6}$$

Proof. Applying the Wiener–Ikehara Tauberian Theorem (see [22], for example) to the Dirichlet series

$$\sum_{n=1}^{\infty}\frac{4\pi n|\rho(n)|^2}{n^s}, \quad \sum_{n=1}^{\infty}\frac{4\pi n|\rho'(n)|^2}{n^s}, \quad \text{and} \quad \sum_{n=1}^{\infty}\frac{4\pi n|\rho(n)+\rho'(n)|^2}{n^s},$$

respectively, we infer by Proposition 5.1 that, as $N \to \infty$,

$$\sum_{1\le n\le N} 4\pi n|\rho(n)|^2 \sim R_F^+ N,$$

$$\sum_{1\le n\le N} 4\pi n|\rho'(n)|^2 \sim R_{F'}^+ N,$$

and

$$\sum_{1\le n\le N} 4\pi n|\rho(n)+\rho'(n)|^2 \sim (R_F^+ + R_{F'}^+)N.$$

Thus,

$$\sum_{1\le n\le N} 4\pi n\, \Re(\rho(n)\overline{\rho'(n)}) = o(N).$$

Similarly applying the Wiener–Ikehara Theorem to the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{4\pi n|\rho(n)+i\rho'(n)|^2}{n^s},$$

we obtain

$$\sum_{1\le n\le N} 4\pi n|\rho(n)+i\rho'(n)|^2 \sim (R_F^+ + R_{F'}^+)N,$$

and consequently,

$$\sum_{1\le \le N} 4\pi n\, \Im(\rho(n)\overline{\rho'(n)}) = o(N).$$

This proves

$$\sum_{1\le n\le N} 4\pi n \rho(n)\overline{\rho'(n)} = o(N).$$

□

5.3. **A Mellin transform.** Fix $\mu = ir$ and $\nu = ir'$ and define

$$\mathcal{M}_k(s) := \int_0^\infty W_{k,\mu}(y)W_{k,\nu}(y)y^{s-2}\,dy.$$

In view of the asymptotics (4.1), (4.2), and (4.3), the integral is absolutely convergent for $\Re(s) > 0$ and hence $\mathcal{M}_k(s)$ is analytic in that region. The asymptotic behavior of $\mathcal{M}_k(s)$ is given by the following.

LEMMA 5.4. *Assume that $|k| < 1/2$. Then as $|s| \to \infty$,*

$$\mathcal{M}_k(s) = \frac{\Gamma(s-\frac{1}{2}+k)^2}{\Gamma(s)}\left(1+O_{\mu,\nu}(|s|^{-1/2})\right).$$

Proof. This is a direct generalization of Lemma 4.1 in [27], which deals with the case $r' = r$. As in [27], we use the integral representation

$$W_{k,ir}(y) = \frac{1}{2\pi i}\int_L \frac{\Gamma(\frac{1}{2}-v-ir)\Gamma(\frac{1}{2}-v+ir)\Gamma(v-k)}{\Gamma(\frac{1}{2}+ir-k)\Gamma(\frac{1}{2}-ir-k)}y^v\,dv,$$

where the path of integration runs from $-i\infty$ to $i\infty$ and is chosen in such a way that all poles of $\Gamma(v-k)$ are to the left and all poles of $\Gamma(\frac{1}{2}-v\pm ir)$ are to the right of L; this is possible since we assume that $|k| < 1/2$. Inserting this into the formula for $\mathcal{M}_k(s)$ gives

$$\mathcal{M}_k(s) = \int_0^\infty W_{k,ir'}(y)\frac{1}{2\pi i}\int_L \frac{\Gamma(\frac{1}{2}-v-ir)\Gamma(\frac{1}{2}-v+ir)\Gamma(v-k)}{\Gamma(\frac{1}{2}+ir-k)\Gamma(\frac{1}{2}-ir-k)} y^{v+s-2}\,dv\,dy.$$

Then, one uses the formula

$$\int_0^\infty W_{k,ir'}(y)e^{-y/2}y^{u-1}\,dy = \frac{\Gamma(\frac{1}{2}+u+ir')\Gamma(\frac{1}{2}+u-ir')}{\Gamma(u-k+1)}, \quad \Re(u) > -1$$

to find

$$\mathcal{M}_k(s) = \frac{1}{2\pi i}\int_L \Gamma(-\frac{1}{2}+v+s+ir')\Gamma(-\frac{1}{2}+v+s-ir') \times \frac{\Gamma(\frac{1}{2}-v-ir)\Gamma(\frac{1}{2}-v+ir)\Gamma(v-k)}{\Gamma(v+s-k)\Gamma(\frac{1}{2}+ir-k)\Gamma(\frac{1}{2}-ir-k)}\,dv.$$

One then shifts the contour of integration to the line $\Re(v) = k-1/2$, picking up a single residue at $v = k$, and estimates the remaining integral as in [27], giving

$$\mathcal{M}_k(s) = \frac{\Gamma(s-\frac{1}{2}+k-ir')\Gamma(s-\frac{1}{2}+k+ir')}{\Gamma(s)} + O_{\mu,\nu}\left(\Gamma(s-\frac{3}{2}+2k)\right).$$

The conclusion of the lemma now follows from Stirling's formula. □

Let

$$M(s) = \begin{bmatrix} \mathcal{M}_{-k}(s+1) & \mathcal{M}_{-k}(s) \\ \mathcal{M}_k(s+1) & -\mathcal{M}_k(s) \end{bmatrix}. \tag{5.7}$$

LEMMA 5.5. *$M(s)$ is analytic and nonsingular for $\Re(s) > 0$.*

Proof. Holomorphy in $\Re(s) > 0$ follows from that of $\mathcal{M}_k(s)$. As in [27], one shows that there is a recurrence relation

$$(s+1)\mathcal{M}_k(s+2) - 2k(2s+1)\mathcal{M}_k(s+1) = \frac{1}{s}\left(s^2-(\mu+\nu)^2\right)\left(s^2-(\mu-\nu)\right)\mathcal{M}_k(s). \tag{5.8}$$

By the recurrence relation (5.8), we infer that

$$(s+1)\det M(s+1) = (s\det M(s))\frac{(s^2-(\mu+\nu)^2)\,(s^2-(\mu-\nu)^2)}{s^2},$$

and

$$\det M(s) = \frac{\Gamma(s+\mu+\nu)\Gamma(s+\mu-\nu)\Gamma(s-\mu+\nu)\Gamma(s-\mu-\nu)}{\Gamma(s+n+\mu+\nu)\Gamma(s+n+\mu-\nu)\Gamma(s+n-\mu+\nu)\Gamma(s+n-\mu-\nu)} \times \frac{(s+n)\Gamma^2(s+n)}{s\,\Gamma^2(s)}\det M(s+n).$$

By using Lemma 5.4 with Stirling's formula, we deduce that[6]

$$\lim_{n\to\infty}\frac{(s+n)\Gamma^2(s+n)\det M(s+n)}{\Gamma(s+n+\mu+\nu)\Gamma(s+n+\mu-\nu)\Gamma(s+n-\mu+\nu)\Gamma(s+n-\mu-\nu)} = -2.$$

[6]Note the limit is −2 instead of 1 as in Matthes' paper [27].

Therefore we conclude that

$$\det M(s) = -2\frac{\Gamma(s+\mu+\nu)\Gamma(s+\mu-\nu)\Gamma(s-\mu+\nu)\Gamma(s-\mu-\nu)}{s\,\Gamma^2(s)},$$

and thus $\det M(s) \neq 0$ for $\Re(s) > 0$. □

5.4. **Proof of Proposition 5.1.** We define the Eisenstein series of weight −2 and 2 by

$$E_{-2}(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} \left(\frac{j(\gamma,z)}{|j(\gamma,z)|}\right)^2 \Im(\gamma z)^s$$

$$E_{-2}(z,s) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_0(4)} \left(\frac{j(\gamma,z)}{|j(\gamma,z)|}\right)^{-2} \Im(\gamma z)^s.$$

These series are absolutely convergent for $\Re(s) > 1$, with an analytic continuation (no poles) to $\Re(s) > 1/2$.

Using the Maass raising operator $\mathsf{K}_{1/2}$, we get a form $\mathsf{K}_{1/2}F$ of weight 5/2 with Fourier expansion

$$\mathsf{K}_{1/2}F(z) = \sum_{n=1}^{\infty} -\rho(n)W_{5/4,ir}(4\pi ny)e(nx) + \sum_{n=1}^{\infty}\left((\frac{3}{4})^2+r^2\right)\rho(-n)W_{-5/4,ir}(4\pi ny)e(-nx),$$

and using the lowering operator $\Lambda_{1/2}$ we get a form $\Lambda_{1/2}F$ of weight −3/2 with Fourier expansion

$$\Lambda_{1/2}F(z) = \sum_{n=1}^{\infty} -\left((\frac{1}{4})^2+r^2\right)\rho(n)W_{-3/4,ir}(4\pi ny)e(nx) + \sum_{n=1}^{\infty}\rho(-n)W_{3/4,ir}(4\pi ny)e(-nx).$$

Consider the Rankin–Selberg integrals

$$\mathscr{A}(s) = \int_{\Gamma_0(4)\backslash\mathbb{H}} F(z)\overline{F'(z)}E(z,s)\,\frac{dx\,dy}{y^2}$$

and

$$2\mathscr{B}(s) = \int_{\Gamma_0(4)\backslash\mathbb{H}} \mathsf{K}_{1/2}F(z)\overline{F'(z)}E_{-2}(z,s)\,\frac{dx\,dy}{y^2} + \int_{\Gamma_0(4)\backslash\mathbb{H}} \Lambda_{1/2}F(z)\overline{F'(z)}E_{2}(z,s)\,\frac{dx\,dy}{y^2}.$$

Since $E_{\pm 2}(z,s)$ are analytic in $\Re(s) > 1/2$, $\mathscr{B}(s)$ is analytic in $\Re(s) > 1/2$; and since $E(z,s)$ is analytic in $\Re(s) > 1/2$ except for simple pole at $s = 1$, $\mathscr{A}(s)$ is analytic for $\Re(s) > 1/2$ except for possibly a simple pole at $s = 1$, where the residue is

$$\operatorname{Res}_{s=1}\mathscr{A}(s) = \frac{\langle F,F'\rangle}{\operatorname{vol}(\Gamma_0(4)\backslash\mathbb{H})} = \frac{\langle F,F'\rangle}{2\pi},$$

so $\mathscr{A}(s)$ is analytic also at $s=1$ if F and F' are orthogonal.

By the standard unfolding trick, we find that for $\Re(s)>1$,

$$\mathscr{A}(s)=\sum_{n=1}^{\infty}\frac{4\pi n\rho(n)\overline{\rho'(n)}}{(4\pi n)^s}\int_0^{\infty}W_{1/4,ir}(y)W_{1/4,ir'}(y)y^{s-2}\,dy \\ +\sum_{n=1}^{\infty}\frac{4\pi n\rho(-n)\overline{\rho'(-n)}}{(4\pi n)^s}\int_0^{\infty}W_{-1/4,ir}(y)W_{-1/4,ir'}(y)y^{s-2}\,dy$$

and

$$2\mathscr{B}(s)=-\sum_{n=1}^{\infty}\frac{4\pi n\rho(n)\overline{\rho'(n)}}{(4\pi n)^s}\int_0^{\infty}W_{5/4,ir}(y)W_{1/4,ir'}(y)y^{s-2}\,dy \\ +\left((\frac{3}{4})^2+r^2\right)\sum_{n=1}^{\infty}\frac{4\pi n\rho(-n)\overline{\rho'(-n)}}{(4\pi n)^s}\int_0^{\infty}W_{-5/4,ir}(y)W_{-1/4,ir'}(y)y^{s-2}\,dy.$$

Setting $k=1/4$, we then have

$$\mathscr{A}(s)=(4\pi)^{-s}\left(L^+(s)\mathscr{M}_k(s)+L^-(s),\mathscr{M}_{-k}(s)\right),$$

and moreover the following holds.

LEMMA 5.6.
(5.9)

$$\mathscr{B}(s)=(4\pi)^{-s}(L_+(s)(\frac{r}{2}\mathscr{M}_k(s)-\frac{1}{2}\mathscr{M}_k(s+1))+L_-(s)(\frac{r}{2}\mathscr{M}_{-k}(s)\ +\ \frac{1}{2}\mathscr{M}_{-k}(s+1))).$$

Proof. Let $k>0$ be any half integer. In our case, $k=1/2$. We have, by unfolding the integral, that

$$\int_{\Gamma_0\backslash\mathbb{H}}K_kF(z)\overline{F'(z)}E_{-2}(z,\ s)\frac{dx\,dy}{y^2}=\int_0^{\infty}\int_0^1K_kF(z)\overline{F'(z)}y^{s-2}\,dx\,dy,$$

and

$$\int_{\Gamma_0\backslash\mathbb{H}}\Lambda_kF(z)\overline{F'(z)}E_{2}(z,\ s)\frac{dx\,dy}{y^2}=\int_0^{\infty}\int_0^1\Lambda_kF(z)\overline{F'(z)}y^{s-2}\,dx\,dy.$$

Now

$$K_kW_{sgn(n)k/2,\,ir}(4\pi|n|y)e(nx)=\Big[(k/2-2\text{sign}(n)\pi|n|y)W_{sgn(n)k/2,\,ir}(4\pi|n|y) \\ +4\pi|n|yW'_{sgn(n)k/2,\,ir}(4\pi|n|y)\Big]e(nx),$$

and

$$\Lambda_kW_{sgn(n)k/2,\,ir}(4\pi|n|y)e(nx)=\Big[(k/2-2\text{sign}(n)\pi|n|y)W_{sgn(n)k/2,\,ir}(4\pi|n|y) \\ -4\pi|n|yW'_{sgn(n)k/2,\,ir}(4\pi|n|y)\Big]e(nx).$$

Hence Lemma 5.5 follows. □

Consequently,

$$\mathscr{A}(s) = (4\pi)^{-s}\left(L_+(s)\mathscr{M}_k(s) + L_-(s)\mathscr{M}_{-k}(s)\right)$$
$$r\mathscr{A}(s) - 2\mathscr{B}(s) = (4\pi)^{-s}\left(L_+(s)\mathscr{M}_k(s+1) - L_-(s)\mathscr{M}_{-k}(s+1)\right).$$

Solving for $L_+(s)$, $L_-(s)$, we obtain

$$\begin{bmatrix} L_+(s) \\ L_-(s) \end{bmatrix} = -\frac{(4\pi)^s}{\det M(s)} M(s) \begin{bmatrix} \mathscr{A}(s) \\ r\mathscr{A}(s) - 2\mathscr{B}(s) \end{bmatrix},$$

where $M(s)$ is given by (5.7). Therefore Proposition 5.1 follows from Lemma 5.5.

5.5. **Proof of Proposition 5.2.** Consider the Rankin–Selberg integrals

$$\mathscr{I}(s) = \int_{\Gamma_0(4)\backslash\mathbb{H}} F(z)\overline{F(z)}E(z,s)\,\frac{dx\,dy}{y^2}$$

and

$$\mathscr{J}(s) = \int_{\Gamma_0(4)\backslash\mathbb{H}} \mathsf{K}_{1/2}F(z)\overline{F(z)}E_{-2}(z,s)\,\frac{dx\,dy}{y^2}.$$

Since $E_{-2}(z,s)$ is analytic in $\Re(s) > 1/2$, $\mathscr{J}(s)$ is analytic in $\Re(s) > 1/2$; and since $E(z,s)$ except for simple pole at $s = 1$, $\mathscr{I}(s)$ is analytic for $\Re(s) > 1/2$ except for a simple pole at $s = 1$, where the residue is

$$\mathrm{Res}_{s=1}\mathscr{I}(s) = \frac{\langle F,F\rangle}{\mathrm{vol}(\Gamma_0(4)\backslash\mathbb{H})} = \frac{\langle F,F\rangle}{2\pi}.$$

By the standard unfolding trick, we find that for $\Re(s) > 1$,

(5.10)
$$\mathscr{I}(s) = \sum_{n=1}^{\infty}\frac{|\rho(n)|^2}{(4\pi n)^{s-1}}\int_0^\infty W_{1/4,ir}(y)^2 y^{s-2}\,dy + \sum_{n=1}^{\infty}\frac{|\rho(-n)|^2}{(4\pi n)^{s-1}}\int_0^\infty W_{-1/4,ir}(y)^2 y^{s-2}\,dy$$

and

$$\begin{aligned}\mathscr{J}(s) = &-\sum_{n=1}^{\infty}\frac{|\rho(n)|^2}{(4\pi n)^{s-1}}\int_0^\infty W_{5/4,ir}(y)W_{1/4,ir}(y)y^{s-2}\,dy \\ &+\left((\frac{3}{4})^2 + r^2\right)\sum_{n=1}^{\infty}\frac{|\rho(-n)|^2}{(4\pi n)^{s-1}}\int_0^\infty W_{-5/4,ir}(y)W_{-1/4,ir}(y)y^{s-2}\,dy.\end{aligned} \tag{5.11}$$

Denoting by $R^{\pm}$ the residue at $s = 1$ of $L_{\pm}(s)$, we find from (5.11) that, since $\mathscr{J}(s)$ is analytic at $s = 1$,

$$\left((\frac{3}{4})^2 + r^2\right)\int_0^\infty W_{-5/4,ir}(y)W_{-1/4,ir}(y)\,\frac{dy}{y}R^- = \int_0^\infty W_{5/4,ir}(y)W_{1/4,ir}(y)\,\frac{dy}{y}R^+.$$

From the formula [6, p. 858], valid for $|\Re(\mu)| < 1/2$,

$$\begin{aligned}&\int_0^\infty W_{\kappa,\mu}(x)W_{\lambda,\mu}(x)\,\frac{dx}{x} \\ &= \frac{1}{(\kappa-\lambda)\sin 2\pi\mu}\left[\frac{1}{\Gamma(1/2-\kappa+\mu)\,\Gamma(1/2-\lambda-\mu)} - \frac{1}{\Gamma(1/2-\kappa-\mu)\,\Gamma(1/2-\lambda+\mu)}\right],\end{aligned}$$

it follows that

$$\int_0^\infty W_{\frac{5}{4},\,ir}(x)W_{\frac{1}{4},\,ir}(x)\,\frac{dx}{x}=\frac{1}{\sinh(2r\pi)}\frac{2r}{|\Gamma(1/4+ir)|^2}\,,$$

and

$$\int_0^\infty W_{-\frac{5}{4},\,ir}(x)W_{-\frac{1}{4},\,ir}(x)\,\frac{dx}{x}=\frac{1}{\sinh(2r\pi)}\frac{2r}{|\Gamma(7/4+ir)|^2}\,.$$

Hence

$$R^+=\frac{|\Gamma(1/4+ir)|^2}{|\Gamma(3/4+ir)|^2}R^-\,. \tag{5.12}$$

Computing the residue at $s=1$ of $\mathscr{I}(s)$ using (5.10) gives

$$\frac{4\pi\langle F,F\rangle}{\mathrm{vol}(\Gamma_0(4)\backslash\mathbb{H})}=R^+\int_0^\infty W_{\frac{1}{4},ir}(y)^2\,\frac{dy}{y}+R^-\int_0^\infty W_{-\frac{1}{4},ir}(y)^2\,\frac{dy}{y}. \tag{5.13}$$

We have for $|\Re(\mu)|<1/2$ [6, p. 858]:

$$\int_0^\infty W^2_{\kappa,\,\mu}(z)\,\frac{dz}{z}=\frac{\pi}{\sin 2\pi\mu}\,\frac{\psi(1/2+\mu-\kappa)-\psi(1/2-\mu-\kappa)}{\Gamma(1/2+\mu-\kappa)\,\Gamma(1/2-\mu-\kappa)},$$

where $\psi(x)=\frac{d}{dx}\log\Gamma(x)$. Therefore

$$\int_0^\infty W_{\frac{1}{4},ir}(y)^2\,\frac{dy}{y}=\frac{\pi}{i\sinh 2\pi r}\,\frac{\psi(1/4+ir)-\psi(1/4-ir)}{|\Gamma(1/4+ir)|^2} \tag{5.14}$$

and

$$\int_0^\infty W_{-\frac{1}{4},ir}(y)^2\,\frac{dy}{y}=\frac{\pi}{i\sinh 2\pi r}\,\frac{\psi(3/4+ir)-\psi(3/4-ir)}{|\Gamma(3/4+ir)|^2}. \tag{5.15}$$

Inserting (5.14), (5.15), and (5.12) into (5.13) gives

$$2\langle F,F\rangle=R^+\frac{\pi}{i\sinh(2\pi r)|\Gamma(\frac{1}{4}+ir)|^2}\left(\psi(\frac{1}{4}+ir)-\psi(\frac{1}{4}-ir)+\psi(\frac{3}{4}+ir)-\psi(\frac{3}{4}-ir)\right)$$

Taking the logarithmic derivative of the reflection formula

$$\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin(\pi x)}$$

gives

$$\psi(x)-\psi(1-x)=-\pi\cot(\pi x).$$

Hence we find

$$\begin{aligned}
\psi(\tfrac{1}{4}+ir)-\psi(\tfrac{1}{4}-ir)+\psi(\tfrac{3}{4}+ir)-\psi(\tfrac{3}{4}-ir)&\\
&=-\frac{\pi\cos\pi(\frac{1}{4}+ir)}{\sin\pi(\frac{1}{4}+ir)}+\frac{\pi\cos\pi(\frac{1}{4}-ir)}{\sin\pi(\frac{1}{4}-ir)}\\
&=\frac{\pi i\sinh(2\pi r)}{|\sin\pi(\frac{1}{4}+ir)|^2}=\frac{2\pi i\sinh(2\pi r)}{\cosh(2\pi r)}\\
&=2i\sinh(2\pi r)|\Gamma(\frac{1}{2}+2ir)|^2\,.
\end{aligned}$$

Therefore, we get

$$\langle F, F\rangle = \pi \frac{|\Gamma(\frac{1}{2}+2ir)|^2}{|\Gamma(\frac{1}{4}+ir)|^2}\cdot R^+,$$

which proves (5.4).

6. THETA LIFTS AND PERIODS

We summarize the results on theta lifts of Shintani [39] and Kohnen [14, 15] in the holomorphic case, and of Katok and Sarnak [13] in the Maass case.

6.1. **Holomorphic forms.** For a holomorphic form f of weight $2k$ (we will later take only even k) for the full modular group $\mathrm{SL}_2(\mathbb{Z})$, Shintani [39] showed that it can be lifted to a cuspidal Hecke eigenform $\theta(f,z)\in S_{k+\frac{1}{2}}(\Gamma_0(4),\chi_k)$ of weight $k+\frac{1}{2}$ with character $\chi_k\left(\begin{pmatrix} a & b\\ c & d\end{pmatrix}\right) = \left(\frac{-1}{d}\right)^k$ for $\Gamma_0(4)$, that is, transforming as

$$F(\gamma z) = \chi_k(\gamma) J_1(\gamma,z)^{2k+1} F(z), \quad \gamma\in\Gamma_0(4),$$

where $J_1(\gamma,z) = \theta_1(\gamma z)/\theta_1(z)$, $\theta_1(z) = \sum_{n=-\infty}^{\infty} e(n^2 z)$. Thus

$$J_1(\begin{pmatrix} a & b\\ c & d\end{pmatrix}, z) = \epsilon_d^{-1}(\frac{c}{d})\sqrt{cz+d},$$

with $\epsilon_d = 1$ if $d=1 \mod 4$ and $\epsilon_d = i$ if $d = 3 \mod 4$. In particular, $J_1(\gamma,z)^2 = (\frac{-1}{d})(cz+d)$.

We write the Fourier expansion of $\theta(f,z)$ as

$$\theta(f,z) = \sum_{d\geq 1} c_f(d) e(dz).$$

Then

$$c_f(d) = \sum_{\mathrm{disc}(q)=d} \int_{C(q)} f(w)(a-bw+cw^2)^{k-1}\,dw,$$

the sum over all $\mathrm{SL}_2(\mathbb{Z})$-equivalence classes of binary quadratic forms $q=[a,b,c]$ of positive discriminant $d = b^2-4ac$.

We note that from (2.5),

$$\int_{C(q)} f(w)(a-bw+cw^2)^{k-1}\,dw = d^{\frac{k-1}{2}}\int_{C(q)} f\,d\mu_{C(q)}$$

is a simple multiple of the period of f over the closed (primitive, oriented) geodesic $C(q)$ corresponding to q. Thus

$$\frac{c_f(d)}{d^{(k-\frac{1}{2})/2}} = \frac{1}{d^{1/4}}\sum_{\mathrm{disc}\,q=d}\int_{C(q)} f\,d\mu_{C(q)}$$

(the sum over all forms of discriminant d, not necessarily primitive), that is

$$\frac{c_f(d)}{d^{(k-\frac{1}{2})/2}} = \frac{\mu_d(f)}{d^{1/4}}. \tag{6.1}$$

6.1.1. *An inner product formula.* Denote by $S^+_{k+1/2}(\Gamma_0(4))$ the space of cusp forms of weight $k+1/2$ on $\Gamma_0(4)$ whose Fourier expansion $\sum_{n\geq 1} c(n)e(nz)$ satisfies the condition $c(n)=0$ unless $(-1)^k n \equiv 0,1 \mod 4$. It was proved by Kohnen [14] that the two spaces $S_{2k}(\Gamma_0(1))$ and $S^+_{k+1/2}(\Gamma_0(4))$ are isomorphic as modules over the Hecke algebra under the Shimura correspondence. Assuming k is even, let

$$f(z)=\sum_{n=1}^{\infty} a_f(n)e(nz)\in S_{2k}(\Gamma_0(1)),$$

and normalize the L-function of f as

$$L(s,f)=\sum_{n=1}^{\infty}\frac{a_f(n)}{n^{s+(2k-1)/2}}$$

so the functional equation is $s\mapsto 1-s$. Let

$$h_f(z)=\sum_{n\geq 1} c_{h_f}(n)e(nz)\in S^+_{k+1/2}(\Gamma_0(4))$$

correspond to f as above and such that $\langle h_f,h_f\rangle=1$. By the work of Kohnen [15, Theorem 3 and Corollary 1], we have

$$c_{h_f}(m)\overline{c_{h_f}(1)}=\frac{(-1)^{k/2}2^k}{\langle f,f\rangle}a_f(m)$$

and

$$|c_{h_f}(1)|^2=\frac{(k-1)!}{\pi^k}\frac{L(\frac{1}{2},f)}{\langle f,f\rangle}.$$

Hence we see that $\theta(f,z)\neq 0$ if and only if $L(\frac{1}{2},f)\neq 0$, and in this case h_f is proportional to $\theta(f,z)$:

$$\overline{c_{h_f}(1)}h_f(z)=\frac{(-1)^{k/2}2^k}{\langle f,f\rangle}\theta(f,z).$$

Thus we get a formula for the inner product of the lift: if g is another cuspidal Hecke eigenform in $S_{2k}(\Gamma_0(1))$, then

$$\langle\theta(f,\cdot),\theta(g,\cdot)\rangle=\frac{(k-1)!}{2^{2k}\pi^k}L(\frac{1}{2},f)\langle f,g\rangle. \tag{6.2}$$

6.2. **Maass forms.** Given an even Maass form ϕ for $\mathrm{SL}_2(\mathbb{Z})$, with eigenvalue $\lambda=\frac{1}{4}+(2r)^2$, the theta lift $F(z)=\theta(\phi,z)$ is a Maass form for $\Gamma_0(4)$ of weight $1/2$, that is, it transforms under $\Gamma_0(4)$ as

$$F(\gamma z)=J(\gamma,z)F(z),\quad \gamma\in\Gamma_0(4),$$

where $J(\gamma,z)=\theta(\gamma z)/\theta(z)$, $\theta(z)=y^{1/4}\theta_1(z)=y^{1/4}\sum_{n=-\infty}^{\infty}e(n^2z)$. Moreover, F is an eigenfunction of $\Delta_{1/2}$ with eigenvalue $\frac{1}{4}+r^2$.

The Fourier expansion of F is given by [13] as

$$F(u+iv)=\sum_{d\neq 0}\rho(d)W_{\frac{\mathrm{sign}(d)}{4},ir}(4\pi|d|v)e(dv),$$

where for $d>0$,

$$\rho(d)=\frac{1}{\sqrt{8}\pi^{1/4}d^{3/4}}\sum_{\mathrm{disc}(q)=d}\int_{C(q)}\phi\,ds,$$

the sum over all $SL_2(\mathbb{Z})$-equivalence classes of binary quadratic forms $q = [a, b, c]$ of positive discriminant $d = b^2 - 4ac$ and $C(q)$ is the closed (primitive, oriented) geodesic corresponding to q. (For $d < 0$ there is an analogous formula involving Heegner points.) Thus

$$\rho(d) = \frac{1}{\sqrt{8}\pi^{1/4} d^{3/4}} \mu_d. \tag{6.3}$$

Moreover, if ψ is another Hecke–Maass eigenform, we have in view of [13, formula (5.6), p. 224] the inner product formula

$$\langle \theta(\phi,\cdot), \theta(\psi,\cdot)\rangle = \frac{3}{2}\Lambda(1/2, \phi)\langle \phi, \psi\rangle, \tag{6.4}$$

where

$$\Lambda(s, \phi) = \pi^{-s}\Gamma\left(\frac{s+2ir}{2}\right)\Gamma\left(\frac{s-2ir}{2}\right)L(s, \phi).$$

Note the above formula is still valid even if ϕ, ψ are not even.

7. PROOF OF THE MAIN THEOREM

7.1. **Expected value of μ_d.** In §6, we identified the measures $\mu_d(f)$ with Fourier coefficients of theta-lifts up to a normalization. Hence the vanishing of the mean value of $\mu_d(f)/d^{1/4}$ follows from the corresponding fact for Fourier coefficients of forms of half-integer weight.

The first statement of Theorem 1.1 follows immediately from Hecke's bound that, for any $\alpha \in \mathbb{R}$ and $\epsilon > 0$, we have

$$\sum_{n \le N} a(n)e(\alpha n) = O(N^{1/2+\epsilon}),$$

where $a(n)$ is the normalized Fourier coefficient of any holomorphic or Maass cusp form and the implicit constant depends on the form and ϵ alone. For the proof, see for example [8, page 111, Theorem 8.1] and [9, page 71, Theorem 5.3]. One makes use of the following formula ([6, page 857, 7.611]:

$$\int_0^\infty x^{-1} W_{k,\mu}(x)\, dx = \frac{\pi^{3/2} 2^k \sec(\mu\pi)}{\Gamma(\frac{3}{4} - \frac{k}{2} + \frac{\mu}{2})\Gamma(\frac{3}{4} - \frac{k}{2} - \frac{\mu}{2})}.$$

7.2. **Proof of Theorem 1.2.** We wish to compute the bilinear form

$$B(f,g) = \lim_{N\to\infty} \frac{1}{\#\{d : d \le N\}} \sum_{d \le N} \frac{\mu_d(f)\overline{\mu_d(g)}}{d^{1/2}},$$

where $f, g \in L^2_{\text{cusp}}(\Gamma\backslash G)$ are smooth and K-finite, and the sum is over discriminants. We take f and g to lie in the irreducible subspaces U_π defined in (1.11), that is, the subspaces of $L^2_{\text{cusp}}(\Gamma\backslash G)$ irreducible under the joint actions of G, the orientation-reversal symmetry $\mathbf{r}$, and under the Hecke algebra. We wish to show that for such f, g,

$$B(f,g) = 0 \tag{7.1}$$

if f, g lie in distinct (hence orthogonal) subspaces U_f, U_g and to compute $B(f,f)$. By the results of §3, it suffices to consider "minimal", or generating vectors, that is, to consider holomorphic forms or Maass forms. In particular, we need to show that for such f,

$$\lim_{N\to\infty}\frac{1}{\#\{d:d\le N\}}\sum_{d\le N}\frac{|\mu_d(f)|^2}{d^{1/2}}=c(f)L(\frac{1}{2},f)V^{\mathrm{sym}}(f,f), \tag{7.2}$$

where

$$c(f)=\begin{cases}6/\pi, & f \text{ Maass form}\\ 1/\pi, & f \text{ holomorphic}.\end{cases}$$

Since both B and V^{sym} vanish when f is holomorphic of weight $\equiv 2 \mod 4$ or an odd Maass form, it suffices to treat the cases of holomorphic forms of weight divisible by 4 and of even Maass forms. To do so, we recall that in these cases we have identified $\mu_d(f)$ with simple multiples of the Fourier coefficients of theta-lifts $\theta(f,z)$. Thus we may use Rankin–Selberg theory (Corollary 5.3) to recover (7.1) and (7.2) once we have made the correct identifications. We treat separately the case of holomorphic forms and Maass forms.

7.2.1. *Holomorphic forms.* For a cuspidal Hecke eigenform of weight $2k$, k even, the theta lift has weight $k+1/2$ with Fourier expansion $\theta(f,z)=\sum_{d\ge1}c_f(d)e(dz)$ satisfying (6.1), *i.e.,*

$$\frac{c_f(d)}{d^{(k-\frac{1}{2})/2}}=\frac{\mu_d(f)}{d^{1/4}}.$$

By standard Rankin–Selberg theory (see Section 5.1), the series

$$(4\pi)^{-(s+k-\frac{1}{2})}\Gamma(s+k-\frac{1}{2})\sum_{n=1}^{\infty}|\frac{c(n)}{n^{\frac{k-1/2}{2}}}|^2n^{-s}$$

has a simple pole at $s=1$ with residue

$$R^+(f)=\frac{\langle\theta(f),\theta(f)\rangle}{2\pi}=\frac{1}{2\pi}(4\pi)^{-k}\Gamma(k)L(\frac{1}{2},f)\langle f,f\rangle$$

by (6.2) and hence the Dirichlet series

$$D(s)=\sum_{n=1}^{\infty}|\frac{\mu_n(f)}{n^{1/4}}|^2n^{-s}$$

has a simple pole at $s=1$ with residue

$$\frac{(4\pi)^{k+\frac{1}{2}}}{\Gamma(k+\frac{1}{2})}R^+(f)=\frac{\Gamma(k)}{\sqrt{\pi}\Gamma(k+\frac{1}{2})}L(\frac{1}{2},f)\langle f,f\rangle\,.$$

Note that by (3.27)

$$V^{\mathrm{sym}}(f,f)=\frac{1}{2}2^{2k}\frac{\Gamma(k)^2}{\Gamma(2k)}\langle f,f\rangle=\pi\frac{\Gamma(k)}{\sqrt{\pi}\Gamma(k+\frac{1}{2})}\langle f,f\rangle,$$

and therefore the residue at $s=1$ of $D(s)$ is

$$\frac{1}{\pi}V^{\mathrm{sym}}(f,f)L(\frac{1}{2},f)\,.$$

Consequently, we find

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n\le N} |\frac{\mu_n(f)}{n^{1/4}}|^2 = \frac{1}{\pi} V^{\mathrm{sym}}(f,f) L(\frac{1}{2},f).$$

7.2.2. *Maass forms.* Let ϕ be an even Maass Hecke eigenform with Laplace eigenvalue $1/4+(2r)^2$ and $F=\theta(\phi,\cdot)$ its theta-lift with Fourier expansion

$$F(u+iv) = \sum_{d\ne 0} \rho(d) W_{\frac{\mathrm{sign}(d)}{4},ir}(4\pi|d|v) e(dv)$$

Recall (6.3), that for $d>0$ we have

$$\rho(d) = \frac{\sqrt{2}}{4\pi^{1/4} d^{3/4}} \mu_d.$$

Thus by Corollary 5.3, we have

$$\frac{1}{N} \sum_{1\le n\le N} \left(\frac{\mu_n(\phi)}{n^{1/4}}\right)^2 \sim \frac{2}{\sqrt{\pi}} R^+(\phi),$$

where $R^+(\phi)$ is given by (5.4).

Inserting (6.4) into (5.4) gives

$$R^+ = \frac{3}{2\pi^{3/2}} \frac{|\Gamma(\frac{1}{4}+ir)|^4}{|\Gamma(\frac{1}{2}+2ir)|^2} L(\frac{1}{2},\phi)\langle\phi,\phi\rangle. \tag{7.3}$$

We note that by (3.26),

$$\frac{|\Gamma(\frac{1}{4}+ir)|^4}{2\pi|\Gamma(\frac{1}{2}+2ir)|^2} \langle\phi,\phi\rangle = V^{\mathrm{sym}}(\phi,\phi),$$

and hence

$$R^+ = \frac{3}{\sqrt{\pi}} V^{\mathrm{sym}}(\phi,\phi) L(\frac{1}{2},\phi). \tag{7.4}$$

Therefore, we get

$$\frac{1}{N} \sum_{1\le n\le N} \left(\frac{\mu_n(\phi)}{n^{1/4}}\right)^2 \sim \frac{2}{\sqrt{\pi}} R^+ = \frac{6}{\pi} V^{\mathrm{sym}}(\phi,\phi).L(\frac{1}{2},\phi) \tag{7.5}$$

7.2.3. *Orthogonality.* Finally, the fact that $B(f,g)=0$ if the subspaces U_f, U_g are distinct follows from standard Rankin–Selberg theory when at least one of f or g is holomorphic (see (5.2) and (5.3)), while the case of both f, g being Maass forms follows from Corollary 5.3.

Acknowledgments. We would like thank Akshay Venkatesh for insightful discussions of the material related to this paper.

REFERENCES

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

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

[3] M. Einsiedler, E. Lindenstrauss, P. Michel and A. Venkatesh, "The Distribution of Periodic Torus Orbits on Homogeneous Spaces II," manuscript.

[4] P. D. T. A. Elliott, "Probabilistic Number Theory. II. Central Limit Theorems," Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 240. Springer-Verlag, Berlin-New York, 1980.

[5] R. Godement, *The decomposition of* $L^2(G/\Gamma)$ *for* $\Gamma = \mathrm{SL}(2, Z)$, 1966 Algebraic Groups and Discontinuous Subgroups (Proc. Sympos. Pure Math., Boulder, Colo. 211–224, Amer. Math. Soc., Providence, R.I.

[6] I. S. Gradshtein and I. M. Ryzhik, "Tables of Integrals, Series and Products," Fourth edition prepared by Ju. V. Geronimus and M. Ju. Ceitlin. Translated from the Russian by Scripta Technica, Inc. Translation edited by Alan Jeffrey, Academic Press, New York-London, 1965.

[7] H. Huber, *Zur analytischen Theorie der hyperbolischen Raumformen und Bewegungsgruppen*, (German), Math. Ann., **138** (1959), 1–26.

[8] H. Iwaniec, "Spectral Methods of Automorphic Forms," Second edition, Graduate Studies in Mathematics, 53, American Mathematical Society, Providence, RI; Revista Matemática Iberoamericana, Madrid, 2002.

[9] H. Iwaniec, "Topics in Classical Automorphic Forms," Graduate Studies in Mathematics, 17, American Mathematical Society, Providence, RI, 1997.

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

[11] H. Iwaniec and E. Kowalski, "Analytic Number Theory," American Mathematical Society Colloquium Publications, 53, American Mathematical Society, Providence, RI, 2004.

[12] S. Katok, *Closed geodesics, periods and arithmetic of modular forms*, Invent. Math., **80** (1985), 469–480.

[13] S. Katok and P. Sarnak, *Heegner points, cycles and Maass forms*, Israel J. Math., **84** (1993), 193–227.

[14] W. Kohnen, *Modular forms of half-integral weight on* $\Gamma_0(4)$, Math. Ann., **248** (1980), 249–266.

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

[16] W. Kohnen and D. Zagier, *Values of L-series of modular forms at the center of the critical strip*, Invent. Math., **64** (1981), 175–198.

[17] P. Kurlberg and Z. Rudnick, *On the distribution of matrix elements for the quantum cat map*, Ann. of Math. (2), **161** (2005), 489–507.

[18] P. Kurlberg, L. Rosenzweig and Z. Rudnick, *Matrix elements for the quantum cat map: Fluctuations in short windows*, Nonlinearity, **20** (2007), 2289–2304.

[19] S. P. Lalley, *Distribution of periodic orbits of symbolic and Axiom A flows*, Adv. in Appl. Math., **8** (1987), 154–193.

[20] E. Landau, "Elementary Number Theory," Translated by J. E. Goodman, Chelsea Publishing Co., New York, N.Y., 1958.

[21] S. Lang, "$\mathrm{SL}_2(R)$," Reprint of the 1975 edition. Graduate Texts in Mathematics, 105, Springer-Verlag, New York, 1985.

[22] S. Lang, "Algebraic Number Theory," second edition, Graduate Texts in Mathematics, 110, Springer-Verlag, New York, 1994.

[23] Y. V. Linnik, "Ergodic Properties of Algebraic Fields," Translated from the Russian by M. S. Keane, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 45, Springer-Verlag New York Inc., New York, 1968.

[24] W. Luo and P. Sarnak, *Quantum variance for Hecke eigenforms*, Ann. Sci. École Norm. Sup. (4), **37** (2004), 769–799.

[25] H. Maass, *Über die räumliche Verteilung der Punkte in Gittern mit indefiniter Metrik*, Math. Ann., **138** (1959), 287–315.

[26] H. Maass, "Lectures on Modular Functions of One Complex Variable," Notes by Sunder Lal. Tata Institute of Fundamental Research Lectures on Mathematics, No. 29, Tata Institute of Fundamental Research, Bombay 1964.

[27] R. Matthes, *Rankin–Selberg method for real analytic cusp forms of arbitrary real weight*, Math. Z., **211** (1992), 155–172.

[28] W. Muller, *The mean square of Dirichlet series associated with automorphic forms*, Monatsh. Math., **113** (1992), 121–159.

[29] S. Niwa, *Modular forms of half integral weight and the integral of certain theta-functions*, Nagoya Math. J., **56** (1975), 147–161.

[30] S. Rallis, *Injectivity properties of liftings associated to Weil representations*, Compos. Math., **52** (1984), 139–169.

[31] M. Ratner, *The rate of mixing for geodesic and horocycle flows*, Ergodic Theory Dynam. Systems, **7** (1987), 267–288.

[32] N. Raulf, *Asymptotics of class numbers for progressions and of fundamental discriminants*, preprint 2007.

[33] W. Roelcke, *Das Eigenwertproblem der automorphen Formen in der hyperbolischen Ebene, I, II*, (German), Math. Ann., **167** (1966), 292–337; ibid, **168** (1966), 261–324.

[34] Z. Rudnick and K. Soundararajan, *Lower bounds for moments of L-functions*, Proc. Natl. Acad. Sci. USA, **102** (2005), 6837–6838 (electronic).

[35] P. Sarnak, *Class numbers of indefinite binary quadratic forms*, J. Number Theory, **15** (1982), 229–247.

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

[37] P. Sarnak, *Reciprocal geodesics*, Analytic number theory, 217–237, Clay Math. Proc., **7**, Amer. Math. Soc., Providence, RI, 2007.

[38] 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.

[39] T. Shintani, *On construction of holomorphic cusp forms of half integral weight*, Nagoya Math. J., **58** (1975), 83–126.

[40] A. Weil, *Sur certains groupes d'opérateurs unitaires*, (French), Acta Math., **111** (1964), 143–211.

[41] A. Wiles, *The Birch and Swinnerton-Dyer conjecture*, The millennium prize problems, 31–41, Clay Math. Inst., Cambridge, MA, 2006.

[42] P. Zhao, Ph.D. thesis, the Ohio State University, in preparation.

WENZHI LUO <wluo@math.ohio-state.edu>: Department of Mathematics, The Ohio State University, 100 Math Tower, 231 West 18th Avenue, Columbus, OH 43210-1174, USA

ZEÉV RUDNICK <rudnick@post.tau.ac.il>: Raymond and Beverly Sackler School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel and School of Mathematics, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA

PETER SARNAK <sarnak@math.princeton.edu>: Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544 and School of Mathematics, Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA