THETA SUMS, EISENSTEIN SERIES, AND THE SEMICLASSICAL DYNAMICS OF A PRECESSING SPIN*

JENS MARKLOF†

To appear in:

D. Hejhal, F. Chung, J. Friedman, M. Gutzwiller and A. Odlyzko (eds.)
Emerging Applications of Number Theory
(Springer, New York)

1. Introduction. The aims of this paper are twofold. First, I would like to illustrate recent results [29] on the asymptotic behaviour of theta sums of the form

$$\sum_{n=-N}^{N} f\left(\frac{n}{N}\right) \mathrm{e}^{2\pi \mathrm{i}\, n^2 x} \qquad (N \to \infty) \tag{1.1}$$

with physical examples, such as the semiclassical time evolution of a quantum spin, which has been previously discussed by Berry and Goldberg [4, 5, 6] and appears in nature, say, in the nuclear quadrupole interaction with an external axial-symmetric electric field [1, 25, 32]. This interaction is for instance responsible for the hyperfine structure of atomic spectra. Due to new experiments with quantum wave packets (see [26] for a wealth of references), there is some recent interest in theta sums in connection with so-called "fractional revivals" or "quantum echos", which shall be discussed briefly, too; compare also the contribution of Kapitanski and Rodnianski [24] in these proceedings. Analogous effects appear in optics, when light is diffracted at a periodic grating, as is described by Berry and Goldberg [4], Berry and Klein [7].

The second goal is to give a detailed account of the equidistribution of long closed horocycles on the metaplectic cover of the special linear group $\mathrm{SL}(2,\mathbb{R})$ modulo a certain arithmetic subgroup, which in turn can be used to calculate the limit distribution ($N \to \infty$) of the values of theta sums, as x varies from 0 to 1. The limit distribution is already contained in [29], here we will also calculate the next to leading order term in the asymptotic expansion with respect to the horocycle length. This asymptotic expansion may also be viewed as a semiclassical expansion of the value distribution of the auto-correlation function for the precessing spin mentioned above. The proof of the equidistribution theorem follows the lines of Sarnak [33],

*This work was supported by Gottlieb Daimler- und Karl Benz-Stiftung under Grant 2.94.36, and by DFG-Schwerpunktprogramm "Ergodentheorie, Analysis und effiziente Simulation dynamischer Systeme" under Contracts DFG–Ste 241/7–1 and 7–2. The results described in this publication constitute part of the author's dissertation [30].

†Abteilung für Theoretische Physik, Universität Ulm, D-89069 Ulm, Germany. Permanent address: School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.

who was concerned with the distribution of closed horocycles on the unit tangent bundle of a hyperbolic surface, which is $\mathrm{PSL}(2,\mathbb{R})$ modulo a discrete subgroup. His proof uses Eisenstein series of even integral weight. In our case we have to work with the metaplectic cover of $\mathrm{SL}(2,\mathbb{R})$, and thus are faced with Eisenstein series of half-integral weight, which in turn has a crucial effect on the next to leading order terms.

Acknowledgments. I thank the organizers of the IMA workshop for giving me the opportunity to participate in this interesting meeting. During the conference I had very helpful discussions concerning theta sums with J. Bolte, D.A. Hejhal, J.P. Keating, W. Luo, A.M. Odlyzko, I. Rodnianski, Z. Rudnick, P. Sarnak, F. Steiner and S. Zelditch, who stimulated further studies. I would especially like to thank P. Sarnak for his help and advice during my year in Princeton, where most of the research was carried out. I also thank Sir M.V. Berry for interesting suggestions, and K. Soundararajan for discussions on L-functions. I am grateful to A. Kaplan, C. Leichtle, W.P. Schleich, F. Steiner and P. Stifter for discussing with me about physical applications. I thank R. Aurich and A. Bäcker for their help with Figures 4 – 6.

1.1. Time evolution of a precessing spin. Let us now turn to the first topic. The time evolution of a (non-relativistic) quantum system is predicted by the Schrödinger equation

$$\mathcal{H}\,|\psi(t)\rangle = \mathrm{i}\,\hbar\,\frac{\partial}{\partial t}\,|\psi(t)\rangle, \tag{1.2}$$

where $\mathcal{H}$ denotes the Hamiltonian of the system, and $|\psi(t)\rangle$ the state vector at time t. Suppose $\mathcal{H}$ is time independent, and the initial state vector is $|\psi(0)\rangle$. Then equation (1.2) has the formal solution $|\psi(t)\rangle = \mathcal{U}(t)\,|\psi(0)\rangle$, with the time evolution operator $\mathcal{U}(t) = \exp\left(-\frac{\mathrm{i}}{\hbar}\mathcal{H}\,t\right)$. If we assume that the Hamiltonian $\mathcal{H}$ has a purely discrete spectrum, i.e.,

$$\mathcal{H}\,|n\rangle = E_n\,|n\rangle, \quad n = 0, 1, 2, 3, \dots\,, \tag{1.3}$$

then the trace of the time evolution operator reads

$$\mathrm{Tr}\,\mathcal{U}(t) = \sum_{n=0}^{\infty}\langle n|\,\mathcal{U}(t)\,|n\rangle = \sum_{n=0}^{\infty}\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_n\,t}, \tag{1.4}$$

which is not convergent in general . A regular expression can be obtained, if one considers the expectation value of $\mathcal{U}(t)$ with respect to the initial state,

$$C(t) = \langle\psi(0)|\,\mathcal{U}(t)\,|\psi(0)\rangle = \sum_{n=0}^{\infty}|c_n|^2\,\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_n\,t}, \tag{1.5}$$

with the coefficients $c_n = \langle n|\psi(0)\rangle$. The function $C(t)$ is called the *auto-correlation function* of the state $|\psi(0)\rangle$. General methods to estimate exponential sums like (1.5) have been developed by Weyl, Hardy and Littlewood, Vinogradov and others, see e.g. [43, 44]. We will restrict our attention to the case when the exponent E_n is quadratic in n, and call these sums plainly *theta sums*, sometimes also *theta series* or *theta functions*.

The simplest quantum systems, where the energy eigenvalues E_n are of this form are a point particle on a circle, or in a one-dimensional box. Even though these systems are standard textbook examples, we will see that a very complex structure of the auto-correlation function $C(t)$ emerges for initial wave packets that are extremely localized in position coordinates, and thus, due to Heisenberg's uncertainty principle (or Fourier transform, if you wish), widely spread in momentum coordinates. Let us illustrate this effect with the particle on a circle. The Schrödinger equation for the wave function $\psi(\alpha,t) = \langle\alpha|\psi(t)\rangle$ reads (in units where $\hbar = 2\times$mass of particle $=$ 1)

$$\begin{cases} -\dfrac{\partial^2}{\partial\alpha^2}\,\psi(\alpha,t) = \mathrm{i}\,\dfrac{\partial}{\partial t}\,\psi(\alpha,t) \\ \psi(\alpha+1,t) = \psi(\alpha,t). \end{cases} \tag{1.6}$$

Eigenfunctions and eigenvalues of the Hamiltonian $-\frac{\partial^2}{\partial\alpha^2}$ are $\varphi_n(\alpha) = \mathrm{e}^{2\pi\mathrm{i}\,n\alpha}$, $E_n = 4\pi^2 n^2$, $n = 0, \pm1, \pm2, \pm3, \ldots$. We can construct an initial wave packet, which is localized at some point α_0 by putting

$$\psi_N(\alpha,0) = N^{1/2} \sum_{k=-\infty}^{\infty} F\big(N(\alpha-\alpha_0-k)\big). \tag{1.7}$$

The parameter N describes the strength of localization, and the function F, which we assume to be rapidly decreasing at $\pm\infty$, determines the shape of the initial wave packet. The normalization factor is chosen to be $N^{1/2}$, since then the L^2-norm of $\psi_N(\alpha,0)$ is for large N a constant being independent of N. To be more precise, we have

$$\int_0^1 |\psi_N(\alpha,t)|^2 d\alpha = \int_{-\infty}^{\infty} |F(\alpha)|^2 d\alpha + O(N^{-R}), \tag{1.8}$$

for any R, as $N\to\infty$. A short calculation shows that the coefficients c_n are then given by $c_n = N^{-1/2}\,\mathrm{e}^{-2\pi\mathrm{i}\,n\alpha_0}\,c\left(\frac{n}{N}\right)$, where the function $c(\beta)$ is simply the Fourier transform of $F(\alpha)$, $c(\beta) = \int_{\mathbb{R}} \mathrm{e}^{-2\pi\mathrm{i}\,\beta\alpha}\,F(\alpha)d\alpha$. Hence the auto-correlation function $C(t) = C_N(t)$ of such a localized packet reads

$$C_N(t) = N^{-1} \sum_{n=-\infty}^{\infty} \left|c\left(\frac{n}{N}\right)\right|^2 \mathrm{e}^{-\mathrm{i}\,4\pi^2 n^2 t}. \tag{1.9}$$

This function is periodic in t with period $T_0 = 1/2\pi$. As we shall later see (Corollary 3.3 of Section 3), there is an asymptotic expression for $C_N(t)$ at rational values of t/T_0, i.e., $t = -\frac{p}{q}T_0$, where $p \neq 0$ and $q > 0$ are coprime integers. If we assume that the initial wave packet is smooth, i.e., $F \in \mathcal{S}(\mathbb{R})$ is of Schwartz class, and normalized in a way that

$$\int_{\mathbb{R}} |F(\alpha)|^2 d\alpha = 1,$$

then we have the relations

$$C_N(-pT_0/q) = \begin{cases} (1+\mathrm{i})\ \epsilon_p^{-1} \left(\frac{q}{p}\right)\ q^{-1/2} + O_R\big((q/N)^R\big) & \text{if } q \equiv 0 \mod 4 \\ O_R\big((q/N)^R\big) & \text{if } q \equiv 2 \mod 4 \\ \epsilon_q \left(\frac{p}{q}\right)\ q^{-1/2} + O_R\big((q/N)^R\big) & \text{if } q \text{ is odd,} \end{cases} \tag{1.10}$$

for all R, as $N \to \infty$. Here ϵ_k equals 1 if $k \equiv 1 \mod 4$, and i if $k \equiv 3 \mod 4$. $\left(\frac{a}{b}\right)$ is a generalized quadratic residue symbol, which is discussed in greater detail in Section 2.2. The last relation can be immediately deduced from the values of the Gauss sums

$$\sum_{n=0}^{q-1} \mathrm{e}^{2\pi\mathrm{i}\, n^2 p/q} = \begin{cases} (1+\mathrm{i})\ \epsilon_p^{-1} \left(\frac{q}{p}\right)\ \sqrt{q} & \text{if } q \equiv 0 \mod 4 \\ 0 & \text{if } q \equiv 2 \mod 4 \\ \epsilon_q \left(\frac{p}{q}\right)\ \sqrt{q} & \text{if } q \text{ is odd,} \end{cases} \tag{1.11}$$

which is a classical result.[1] The physical reason behind all this is that the wave packet $\psi_N(\alpha, t)$ has a big overlap with the initial one $\psi_N(\alpha, 0)$, if t/T_0 is rational with a denominator much smaller than N. This effect is called *fractional revival*, see [26] and references therein for more information on this topic, in particular on recent experiments.

Before going deeper into the subtleties of the $C_N(t)$ at irrational times t/T_0, let me briefly mention, how the above example connects to the quadrupole interaction of a nuclear or atomic spin with an axial-symmetric electrostatic field. The Hamiltonian of this interaction is given by [1, 25, 32]

$$\mathcal{H}_{\text{quad}} = A\, \frac{3\,\mathcal{I}_{\|}^2 - I(I+1)}{I(2I+1)}, \tag{1.12}$$

where I is the fixed quantum number of the total spin of the nucleus, say, and $\mathcal{I}_{\|}$ the component of the spin operator along the symmetry axis. The constant A is proportional to the electric quadrupole moment of the

[1] See for instance [27], Proposition 2.2.18. Our proof of Corollary 3.3 contained in Section 3 is more indirect. It seems that there we use only the transformation formulas of theta functions, but Gauss sums have, however, already been used in the proof of these transformation formulas.

nucleus, and proportional to the gradient of the electric field at the position of the nucleus. Since the eigenvalues of $\mathcal{I}_{\|}$ are integers $m = -I, \dots, I$, we obtain the energy values $E_m = A\,\frac{3\,m^2 - I(I+1)}{I(2I+1)}$, $m = -I, \dots, I$. The auto-correlation function is thus given by

$$C_I(t) = \mathrm{e}^{2\pi\mathrm{i}(I+1)t/3T_I} \sum_{n=-I}^{I} c_m\, \mathrm{e}^{-2\pi\mathrm{i}\, m^2 t/T_I}, \quad T_I = \frac{2\pi\, I(2I+1)}{3A}, \tag{1.13}$$

and with the analogous localized initial state, described above for the particle on a circle,

$$C_I(t) = \mathrm{e}^{2\pi\mathrm{i}(I+1)t/3T_I}\, I^{-1} \sum_{m=-\infty}^{\infty} \left| c\left(\frac{m}{I}\right)\right|^2 \mathrm{e}^{-2\pi\mathrm{i}\, m^2 t/T_I}, \tag{1.14}$$

the only restriction being now that $c(\beta)$ has to have compact support in $[-1, 1]$. We see that up to a time dependent phase factor we have recovered the familiar theta sum, and the limit that we are interested in corresponds to the limit of a nucleus with large total spin.

The sums that we have discussed so far are of the general form

$$\sum_{n=-\infty}^{\infty} f\left(\frac{n}{N}\right) \mathrm{e}^{2\pi\mathrm{i}\, n^2 x}. \tag{1.15}$$

Contrary to the behaviour at rational values of x, this function shows at irrational points extremely wild fluctuations, as N becomes large. The first ones to study the asymptotic properties of theta sums at irrational argument in great detail were Hardy and Littlewood [15, 16], using mainly the method of diophantine approximation. Their methods were later refined, see [12, 21, 22, 23] and references therein. In particular [12] contains a brief historical summary of important contributions.

1.2. The Artin billiard and theta sums. Our approach will be different, namely more geometrical. We will later see that it makes sense to replace the cut-off parameter N by $y = N^{-2}$, so that we are now interested in the limit $y \to 0^+$. Furthermore let us use the abbreviation $e(z) = e^{2\pi\mathrm{i}\, z}$. The sum in concern reads then

$$\sum_{n=-\infty}^{\infty} f(ny^{1/2})\, e(n^2 x). \tag{1.16}$$

In order to study fluctuations, we have to normalize the sum correctly. The L^2-norm of the sum is for small y (take f to be even)

$$\Big(\int_0^1 \left| \sum_{n=-\infty}^{\infty} f(ny^{1/2}) e(n^2 x)\right|^2 dx\Big)^{1/2} \sim \Big(2\int_{-\infty}^{\infty} |f(u)|^2 du\Big)^{1/2}\, y^{-1/4}, \tag{1.17}$$

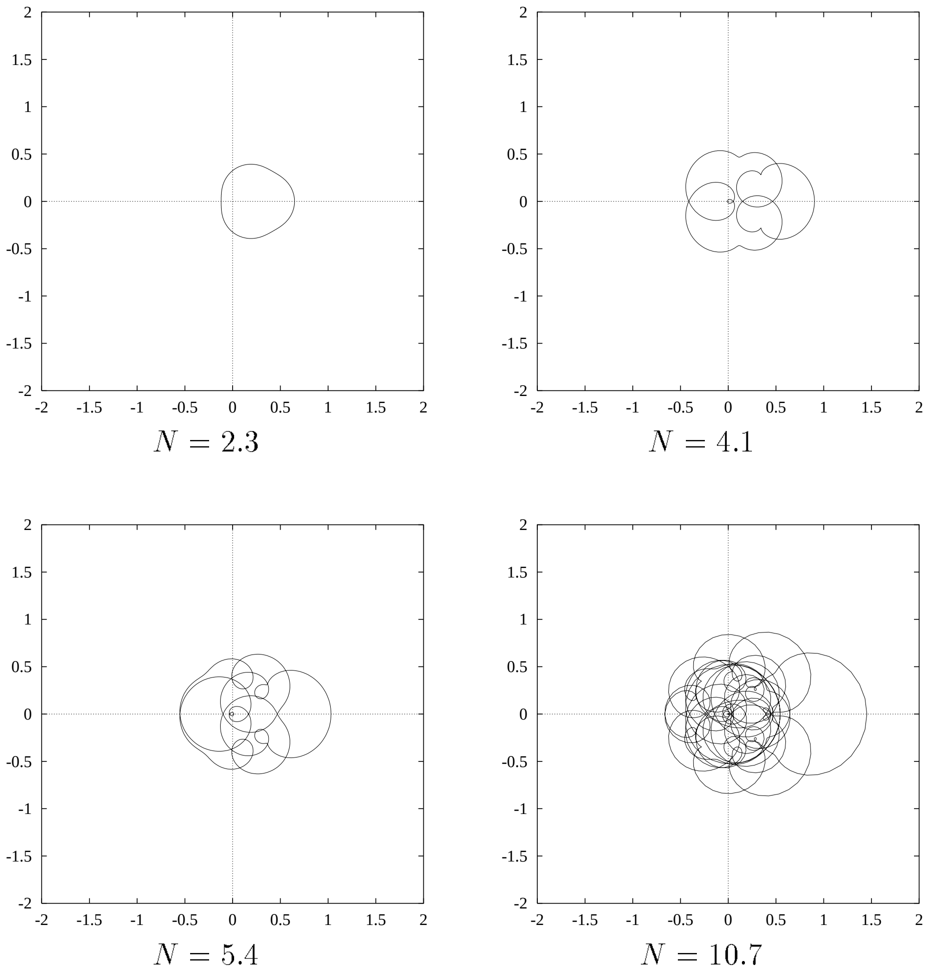


FIG. 1. *Shown are the values of the theta sum* $\{\Theta_f(x + \mathrm{i}\, N^{-2}) \,:\, x \in [0,1]\}$ *in the complex plane, for increasing* N*, where* f *is the smooth cut-off function defined in (1.20). As* $N \to \infty$*, the closed curves are becoming densely distributed in the whole complex plane, compare Proposition 3.5 of Section 3.2.*

thus we normalize by $y^{1/4}$,

$$\Theta_f(x + \mathrm{i}\, y) = y^{1/4} \sum_{n=-\infty}^{\infty} f(ny^{1/2}) e(n^2 x). \tag{1.18}$$

The advantage of this notation becomes clear, if we take the cut-off function $f(u) = \exp(-2\pi u^2)$, because then we arrive at the classical theta series

$$\theta_0(z) = y^{1/4} \sum_{n=-\infty}^{\infty} e(n^2 z), \tag{1.19}$$

where here and in the sequel $z = x + \mathrm{i}\, y$. The series is convergent if $y > 0$, i.e., if z lies in the upper half plane $\mathfrak{H} = \{z \in \mathbb{C} \,:\, \mathrm{Im}\, z > 0\}$.

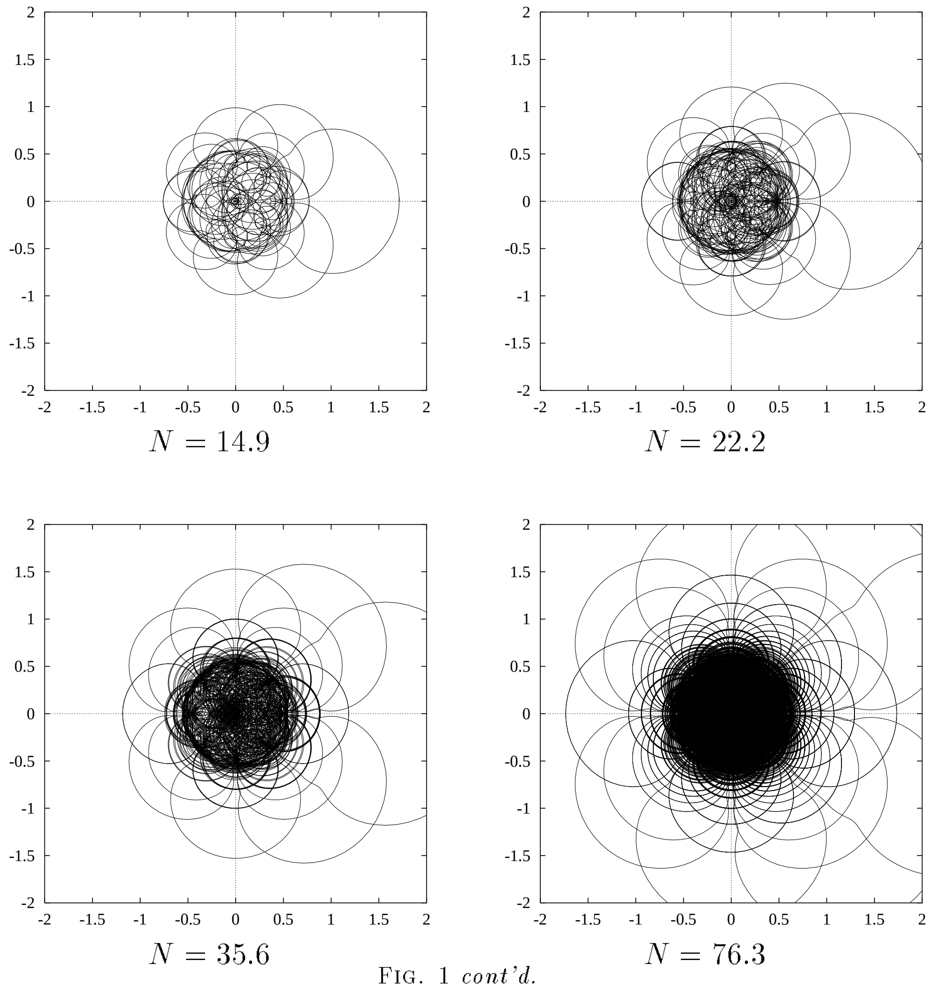


FIG. 1 *cont'd.*

The relation between the theta function $\Theta_f(z)$ and the auto-correlation functions $C(t)$ are $C_N(t) = N^{-1/2}\ \Theta_f(-t/T_0 + \mathrm{i}\,N^{-2})$ for the particle on a circle, and $C_I(t) = \mathrm{e}^{2\pi\mathrm{i}(I+1)t/3T_I}\, I^{-1/2}\ \Theta_f(-t/T_I + \mathrm{i}\,I^{-2})$ for the precessing spin. Figure 1 illustrates the increasing complexity of the theta function $\Theta_f(x+\mathrm{i}\,N^{-2})$ as N increases (y decreases). Shown are the sets $\{(\Theta_f(x+\mathrm{i}\,N^{-2}) : x \in [0,1]\}$ in the complex plane, for different (randomly chosen) values of N. We have chosen the cut-off function

$$f(u) = \begin{cases} \mathrm{e}^{-\frac{1}{1-u^2}} & \text{for } |u| \leq 1 \\ 0 & \text{otherwise,} \end{cases} \tag{1.20}$$

which is infinitely differentiable and of compact support. So the theta sum in concern reads

$$\Theta_f(x+\mathrm{i}\,N^{-2}) = N^{-1/2} \sum_{n=-[N]}^{[N]} \mathrm{e}^{-\frac{N^2}{N^2-n^2}}\, e(n^2 x), \tag{1.21}$$

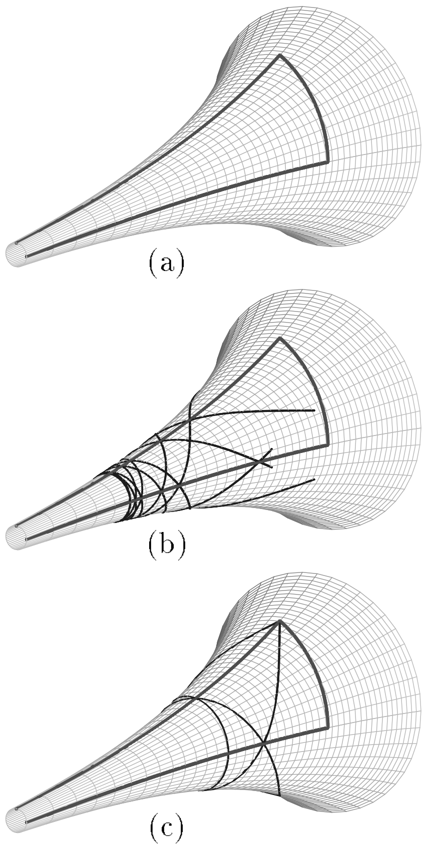


FIG. 2. *(a) The Artin billiard on the pseudosphere, embedded in three-dimensional euclidean space. (b) Example for a non-periodic geodesic. (c) Example for a periodic geodesic. The Artin billiard is non-compact, since the horn on the left hand side runs to infinity, but has finite area. The classical dynamics of this system are strongly chaotic; especially the ergodic properties of the flow are later used to prove limit theorems for the theta functions.*

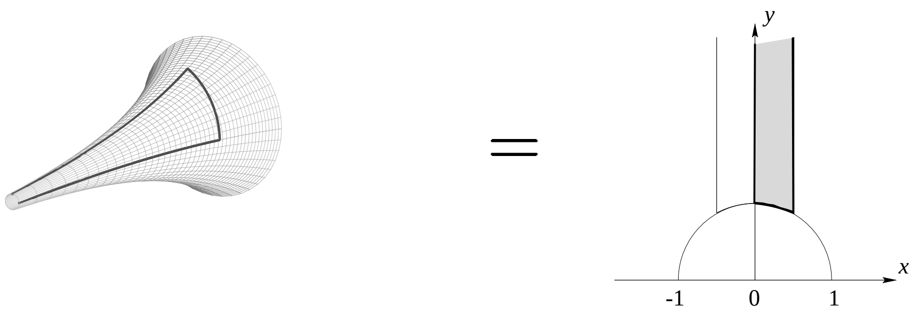


FIG. 3. *The Artin billiard can be realized both as a billiard on the pseudosphere (left) and as a billiard on the hyperbolic plane $\mathfrak{H}$ (right), where the billiard domain is half of the fundamental domain of the modular group.*

where $[N]$ is the greatest integer less or equal to N.

The main goal of this paper is to show that there is a limit distribution of the density of the closed curlicues of Figure 1. To be more precise, let us take some open subset $\mathcal{B}$ of the complex plane $\mathbb{C}$. Then we can prove – assuming that $\mathcal{B}$ is generic in a certain sense[2] – that the limit

$$\lim_{y\to 0} \left|\{x \in [0,1] \ : \ \Theta_f(x+\mathrm{i}\,y) \in \mathcal{B}\}\right| \tag{1.22}$$

exists, and will calculate its value. Here $|\{\ \cdot\ \}|$ denotes the Lebesgue measure of the set $\{\ \cdot\ \}$. This limit theorem is proved using ergodic theory on a three-dimensional space $\mathcal{M}_1$, which is closely related to the phase space[3] of a free point particle constrained to the domain of the Artin billiard [2], with elastic reflection at the walls, see Figure 2. This billiard is not a plane billiard, but built on a surface of revolution, called *pseudosphere*, since its Gaussian curvature is constant but negative, or *tractroid*, since the corresponding rotated curve is the *tractrix* $x = \operatorname{arcosh}(1/y) - \sqrt{1-y^2}$. The billiard can thus be embedded in three-dimensional euclidean space, but for calculations it is more convenient to use the half plane model $\mathfrak{H}$ of hyperbolic space. The Artin billiard can in fact be mapped isometrically onto half of the fundamental domain of the modular group $\mathrm{SL}(2,\mathbb{Z})$ (Figure 3), which shall be discussed in more detail in the next section. As mentioned before, the phase space of the Artin billiard is not exactly the space $\mathcal{M}_1$ we will have to work with; but our favorite space $\mathcal{M}_1$ has 48 discrete symmetries, and the fundamental cell which is left after desymmetrization is exactly the phase space of the Artin billiard. In other words, $\mathcal{M}_1$ is a 48th cover of the phase space of the billiard, compare Section 2.2 for details. The crucial observation will then be (Section 3) that the theta function $\Theta_f(z)$ can be extended to a function $\Theta_f(z,\phi)$ depending on an additional real variable, such that $\Theta_f(z) = \Theta_f(z,0)$, and that $\Theta_f(z,\phi)$ may be viewed as a function on $\mathcal{M}_1$. The coordinate $z = x + \mathrm{i}\,y$ describes the position of the particle, and ϕ the orientation of its momentum vector. This connection between theta sums $\Theta_f(z,\phi)$ and a chaotic dynamical system brings us into the position to study the value distribution of $\Theta_f(z,\phi)$ by means of ergodic theory. The existence of the limit (1.22) can for example be understood by the equidistribution of long closed horocycles[4] in the phase space of the Artin billiard (or rather $\mathcal{M}_1$).

Billiards like the Artin billiard, which tessellate the hyperbolic plane, are favorite objects of investigation in the theory of quantum chaos [14], in particular, because the periodic orbit theory for these systems is free from semiclassical corrections, as was discovered by Selberg [35].[5] We will later

[2] Allowed are for example discs with generic radii (i.e. except for countably many).

[3] To be correct, we mean the three-dimensional hypersurface of constant energy in phase space.

[4] A horocycle describes the wave front of a plane wave in the hyperbolic plane.

[5] The theory is worked out in detail in [17, 18]. For an excellent introduction, espe-

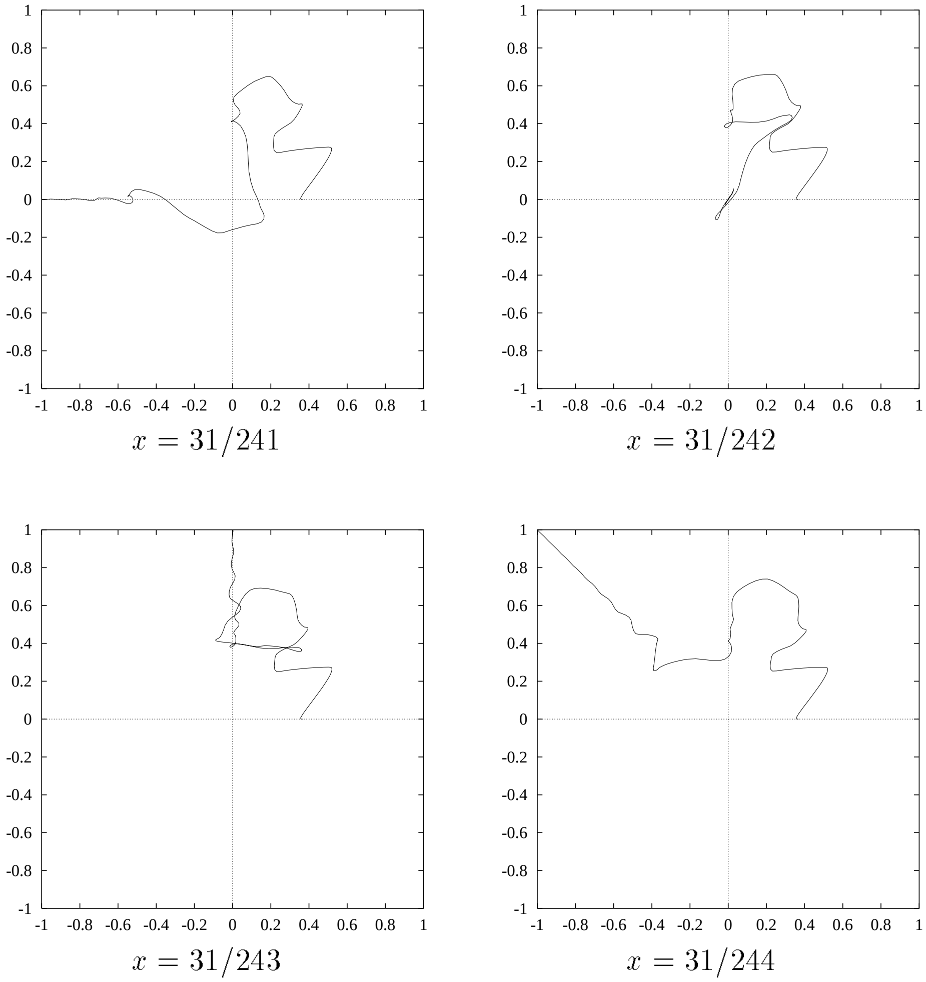


FIG. 4. *Shown are the sets* $\{\Theta_f(x + \mathrm{i}\, \mathrm{e}^{-l}) \,:\, l \geq 0\}$ *for fixed* x *in the complex plane. For rational values of* x*, the curve will eventually run to either* $\pm\infty$, 0, $\pm \mathrm{i}\,\infty$ *or* $(\pm 1 \pm \mathrm{i})\infty$, *as for* $x = \frac{31}{241}$, $x = \frac{31}{242}$. $x = \frac{31}{243}$, $x = \frac{31}{244}$, *respectively. The prefactor depends only on the residue* $\bmod 4$ *of numerator and denominator of* x *(Corollary 3.3). The corresponding trajectories in the Artin billiard are scattering orbits, which disappear into the infinitely long horn after a few reflections at the walls.*

use Eisenstein series, which can be viewed as scattering solutions of the quantum billiard, to prove the equidistribution of long closed horocycles. In the case of the Artin billiard and other non-compact hyperbolic surfaces the proof is due to Sarnak [33]; we will have to generalize it slightly, since we are here dealing with a covering of the phase space. That is, instead of Eisenstein series of even integral weight we shall turn to Eisenstein series of half-integral weight, where we mainly refer to the work of Shimura [39].

Via the connection to the Artin billiard, we can also understand – in a more geometric or dynamical way – the asymptotic behaviour of the

cially with respect to questions related to quantum chaos, see [3].

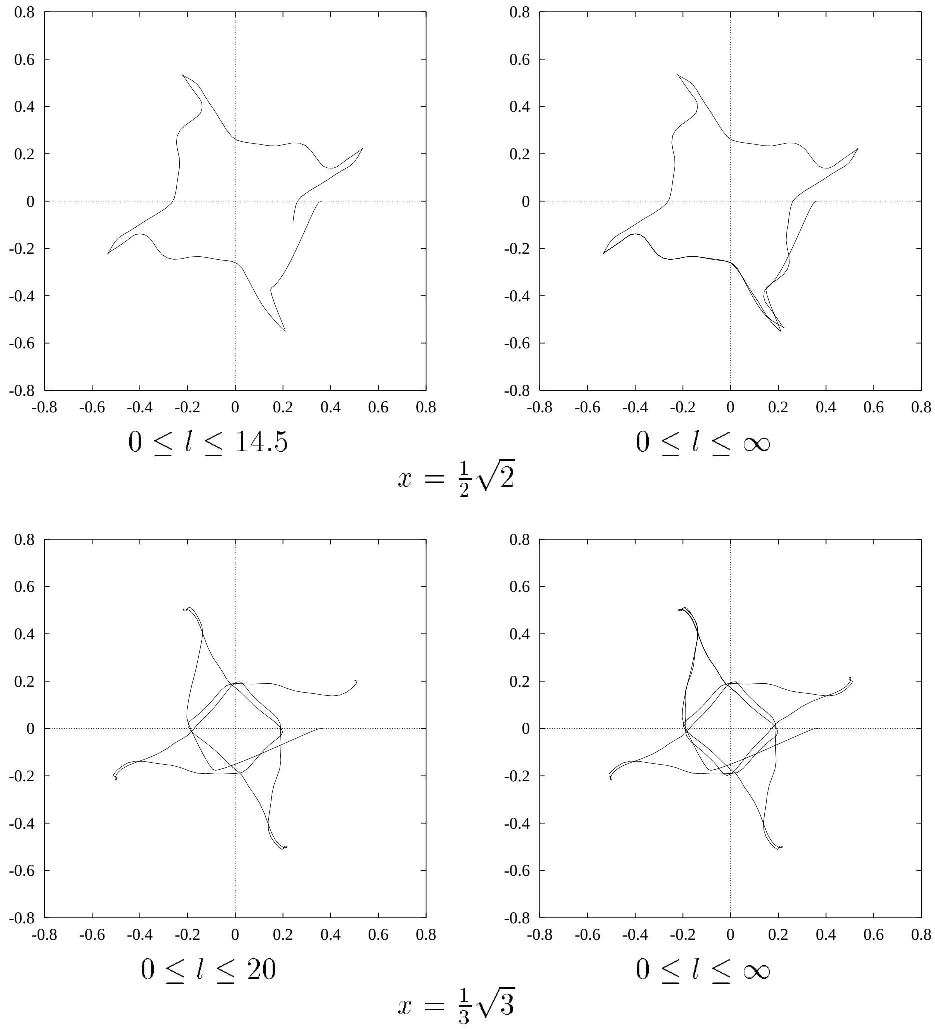
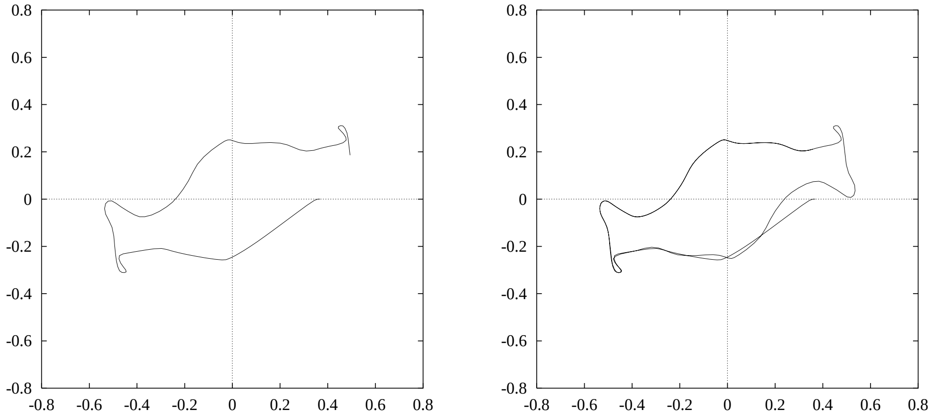


FIG. 5. *For quadratic irrational values x the corresponding trajectories in the Artin billiard are attracted by periodic orbits, hence the sets $\{\Theta_f(x + \mathrm{i}\,\mathrm{e}^{-l}) \,:\, l \ge 0\}$ show periodic structures for $x = \frac{1}{2}\sqrt{2}$, $x = \frac{1}{3}\sqrt{3}$, and the golden mean $x = \frac{1}{2}(\sqrt{5} - 1)$.*

theta sum $\Theta_f(z, \phi = 0)$ at rational points $z = p/q + \mathrm{i}\, y$, which is stated in equation (1.10) and in Corollary 3.3. It turns out that the set

$$\{(p/q + \mathrm{i}\,\mathrm{e}^{-l}, 0) \in \mathcal{M}_1 \,:\, 0 \le l \le L\} \tag{1.23}$$

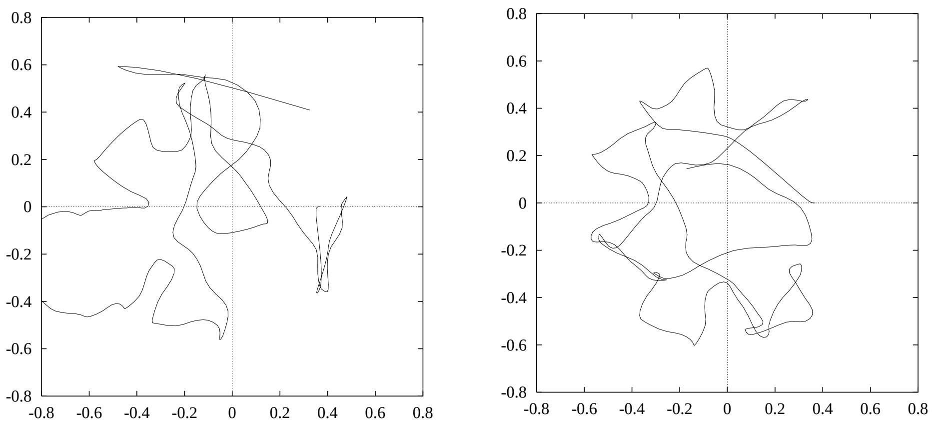
represents a trajectory of length L in the Artin billiard, which eventually (for $L \to \infty$) disappears into the infinitely long horn, after a number of reflections at the walls. The bigger the denominator q, the longer one has to wait, before the trajectory runs straight into the horn. The asymptotic relation for rational argument can then be deduced from the well understood behaviour of a theta function in the horn (Proposition 3.3). It is, however, different for every sheet of the cover $\mathcal{M}_1$ of the billiard phase space, which explains the different cases that may occur (Figure 4).



$0 \leq l \leq 10$ $0 \leq l \leq \infty$

$x = \frac{1}{2}(\sqrt{5}-1)$

FIG. 5 *cont'd.*



$0 \leq l \leq 35$ $0 \leq l \leq 35$

$x = 2^{-1/3}$ $x = 1/\mathrm{e}$

FIG. 6. *The curves* $\{\Theta_f(x + \mathrm{i}\,\mathrm{e}^{-l}) \,:\, l \geq 0\}$ *for* $x = 2^{-1/3}$ *and* $x = 1/\mathrm{e}$. *The number* $x = 2^{-1/3}$ *might be a candidate for a generic example, i.e., it would correspond to an ergodic trajectory in the billiard.*

It is even more interesting to consider the theta sum $\Theta_f(z, \phi = 0)$ evaluated along the trajectory

$$\{(x + \mathrm{i}\,\mathrm{e}^{-l}, 0) \in \mathcal{M}_1 \,:\, 0 \leq l \leq L\}, \tag{1.24}$$

where x is now a quadratic irrational number, i.e., a root of a quadratic polynomial of the form $c_2 x^2 + c_1 x + c_0$ with rational coefficients. In this case one can show that the trajectory (1.24) is exponentially fast attracted by a periodic orbit of the billiard, and thus the function $\Theta_f(x + \mathrm{i}\,\mathrm{e}^{-L}, 0)$ is asymptotically periodic in L, where the period is the same as the one of the periodic orbit (Figure 5). For a detailed discussion of the connection between closed geodesics and binary quadratic forms, see Sarnak [34].

As in the case of horocycles, we can use the ergodic theorem for a generic trajectory in the billiard to calculate the limit

$$\lim_{L\to\infty} \frac{1}{L} \left| \{ l \in [0,L] \; : \; \Theta_f(x + \mathrm{i}\,\mathrm{e}^{-l}) \in \mathcal{B} \} \right| \tag{1.25}$$

for generic values of x (the region $\mathcal{B}$ is chosen as before), which has the same value as (1.22), and is thus independent of the particular choice of a generic x. Not a single example for a generic x is known, Figure 6 displays two curves for good candidates $x = 2^{-1/3}$ and $x = 1/\mathrm{e}$. There is a well known connection between the continued fraction expansion of x and the behaviour of the corresponding geodesic, see, e.g. [10, 36], which can be used to find a symbolic dynamics for the geodesic flow on the Artin billiard or the modular surface.[6] It illuminates in particular the relation of our geometrical approach to the methods of diophantine approximation involving Farey series, which are used by Jurkat and van Horne [21, 22, 23], as well as to the renormalization technique based on the Gauss map applied by Berry and Goldberg [4].

1.3. Notation. It is important that throughout the paper the correct branches of the logarithm are chosen. We recall the basic terminology.

For complex numbers $z \neq 0$ and α, we define $z^\alpha = |z|^\alpha \, \mathrm{e}^{\mathrm{i}\,\alpha \arg(z)}$, where the argument is chosen such that $-\pi < \arg(z) \leq \pi$. As a consequence of this definition we have the following rules: $z^{\alpha+\beta} = z^\alpha z^\beta$, $z^{m\alpha} = (z^\alpha)^m$ for $m \in \mathbb{Z}$, and $(zw)^\alpha = z^\alpha w^\alpha$, if $\arg(z)$, $\arg(w)$ and $\arg(z) + \arg(w)$ are all contained in the interval $(-\pi, \pi]$, for suitable choices of the argument. In the sequel we will understand by arg always the main branch $-\pi < \arg(z) \leq \pi$ of the argument function.

This definition means in particular that the square-root $z^{1/2}$ of a complex number z is always to be chosen such that $-\pi/2 < \arg(z^{1/2}) \leq \pi/2$.

2. The special linear group $\mathrm{SL}(2,\mathbb{R})$ and its universal cover.

2.1. $\mathrm{SL}(2,\mathbb{R})$ and some discrete subgroups. The action of

$$\mathrm{SL}(2,\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \; : \; a,b,c,d \in \mathbb{R}, \; ad - bc = 1 \right\}$$

on the upper half plane $\mathfrak{H} = \{ z \in \mathbb{C} \; : \; \operatorname{Im} z > 0 \}$ is defined via fractional linear transformations, i.e.,

$$g : z \mapsto gz = \frac{az+b}{cz+d}, \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R}). \tag{2.1}$$

[6] The related *continued fraction* or *Gauss map* can then be used to define Ruelle's dynamical zeta function for the Artin billiard or the modular surface in the framework of the thermodynamic formalism, see Chang and Mayer's contribution [8] in these proceedings.

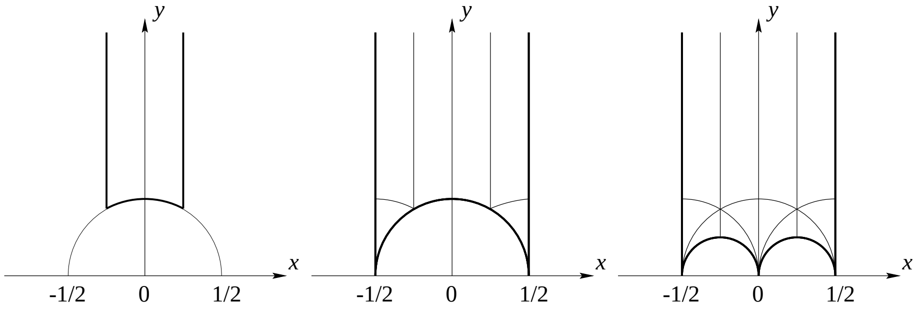


FIG. 7. *Fundamental regions of the boosted modular group* Γ_{M}*, the theta group* Γ_θ*, and the congruence group* $\Gamma_0(4)$*.*

We will now give a collection of certain discrete subgroups of $\mathrm{SL}(2,\mathbb{R})$, compare e.g. Hejhal [18], Chapter Eleven.

The *modular group*

$$\mathrm{SL}(2,\mathbb{Z})=\left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix} :\ a,b,c,d\in\mathbb{Z},\ ad-bc=1\right\}$$

is generated by the elements $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$. A fundamental region of $\mathrm{SL}(2,\mathbb{Z})$ in $\mathfrak{H}$ is $\mathcal{F}_{\mathrm{SL}(2,\mathbb{Z})}=\{z\in\mathfrak{H}\,:\,|x|<1/2,\,|z|>1\}$ (Figure 3).

We will also have to deal with the boosted modular group

$$\Gamma_{\mathrm{M}}=\left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix} :\ a,2b,c,d\in\mathbb{Z},\ ad-bc=1,\ c\equiv 0 \mod 2\right\},$$

which is conjugate to $\mathrm{SL}(2,\mathbb{Z})$,

$$\Gamma_{\mathrm{M}}=\begin{pmatrix} 2^{-1/2} & 0\\ 0 & 2^{1/2}\end{pmatrix}\mathrm{SL}(2,\mathbb{Z})\begin{pmatrix} 2^{1/2} & 0\\ 0 & 2^{-1/2}\end{pmatrix}.$$

Γ_{M} is therefore generated by the elements $\left(\begin{smallmatrix}1&1/2\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&-1/2\\2&0\end{smallmatrix}\right)$. A fundamental region of Γ_{M} in $\mathfrak{H}$ is $\mathcal{F}_{\Gamma_{\mathrm{M}}}=\{z\in\mathfrak{H}\,:\,|x|<1/4,\,|z|>1/2\}$ (Figure 7).

The *theta group*

$$\Gamma_\theta=\left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in\Gamma_{\mathrm{M}} :\ ab\in\mathbb{Z},\ cd\equiv 0 \mod 4\right\}$$
$$=\left\{\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in\Gamma_{\mathrm{M}} : \begin{pmatrix} a & 2b\\ c/2 & d\end{pmatrix}\equiv\begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix}\text{or}\begin{pmatrix} 0 & 1\\ 1 & 0\end{pmatrix}\mathrm{mod}\ 2\right\}$$

is generated by the elements $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&-1/2\\2&0\end{smallmatrix}\right)$. A fundamental region of Γ_θ in $\mathfrak{H}$ is $\mathcal{F}_{\Gamma_\theta}=\{z\in\mathfrak{H}\,:\,|x|<1/2,\,|z|>1/2\}$ (Figure 7). Γ_θ is a subgroup of index 3 in Γ_{M}, and we have the coset decomposition

$$\Gamma_{\mathrm{M}}=\Gamma_\theta\,\cup\,\Gamma_\theta\,\gamma_{1/2}\,\cup\,\Gamma_\theta\,\gamma_{1/2}^2, \tag{2.2}$$

where $\gamma_{1/2} = \left(\begin{smallmatrix} 0 & -1/2 \\ 2 & -1 \end{smallmatrix}\right)$ maps the cusp at $1/2$ onto the cusp at ∞.

The *congruence group*

$$\Gamma_0(4) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \ : \ c \equiv 0 \mod 4 \right\}$$

is generated by the elements $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ -4 & 1 \end{smallmatrix}\right)$. A fundamental region of $\Gamma_0(4)$ in $\mathfrak{H}$ is (Figure 7)

$$\mathcal{F}_{\Gamma_0(4)} = \{ z \in \mathfrak{H} \ : \ |x| < 1/2,\ |z - 1/4| > 1/4,\ |z + 1/4| > 1/4 \}. \tag{2.3}$$

$\Gamma_0(4)$ is a normal subgroup of index 6 in Γ_{M}, and normal of index 2 in Γ_θ, we have the coset decomposition

$$\Gamma_\theta = \Gamma_0(4) \ \cup \ \Gamma_0(4)\, \gamma_0, \tag{2.4}$$

where

$$\gamma_0 = \begin{pmatrix} 0 & -1/2 \\ 2 & 0 \end{pmatrix} \tag{2.5}$$

maps the cusp at 0 onto the cusp at ∞. The fact that $\Gamma_0(4)$ is normal in Γ_{M} and Γ_θ follows from

$$\gamma_{1/2}^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \gamma_{1/2} = \begin{pmatrix} -2b + d & a/2 + b - c/4 - d/2 \\ -4b & a + 2b \end{pmatrix} \in \Gamma_0(4)$$

and

$$\gamma_0^{-1} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \gamma_0 = \begin{pmatrix} d & -c/4 \\ -4b & a \end{pmatrix} \in \Gamma_0(4).$$

For the congruence subgroup

$$\Gamma_1(4) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4) \ : \ a \equiv d \equiv 1 \mod 4 \right\},$$

we find $\Gamma_0(4) = \Gamma_1(4) \ \cup \ \Gamma_1(4) \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, so $\Gamma_1(4)$ is of index 12 in Γ_{M}, and normal of index 4 in Γ_θ. $\Gamma_1(4)$ and $\Gamma_0(4)$ do not differ in their action on $\mathfrak{H}$, hence one may choose the same fundamental region as for $\Gamma_0(4)$, $\mathcal{F}_{\Gamma_1(4)} = \mathcal{F}_{\Gamma_0(4)}$ (Figure 7).

2.2. The universal cover of $\mathrm{SL}(2,\mathbb{R})$**.** The universal covering group of $\mathrm{SL}(2,\mathbb{R})$ can be defined as

$$\widetilde{\mathrm{SL}}(2,\mathbb{R}) = \Big\{ [g, \beta_g] \ : \ g \in \mathrm{SL}(2,\mathbb{R}),$$
$$\beta_g \text{ a function on } \mathfrak{H} \text{ s.t. } \mathrm{e}^{\mathrm{i}\,\beta_g(z)} = \frac{cz+d}{|cz+d|} \Big\}, \tag{2.6}$$

the product of two elements being given by

$$[g,\beta_g][g',\beta'_{g'}]=[gg',\beta''_{gg'}],\quad \beta''_{gg'}(z)=\beta_g(g'z)+\beta'_{g'}(z). \tag{2.7}$$

The inverse of $[g,\beta_g]$ is then $[g,\beta_g]^{-1}=[g^{-1},\beta'_{g^{-1}}]$ with the function $\beta'_{g^{-1}}(z)=-\beta_g(g^{-1}z)$.

$\widetilde{\mathrm{SL}}(2,\mathbb{R})$ is a Lie group with underlying manifold $\mathfrak{H}\times\mathbb{R}$. We may identify $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ with $\mathfrak{H}\times\mathbb{R}$ via $[g,\beta_g]\mapsto(z,\phi)=(g\,\mathrm{i},\beta_g(\mathrm{i}))$. The action of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ on $\mathfrak{H}\times\mathbb{R}$ is then canonically defined by

$$[g,\beta_g](z,\phi)=(gz,\phi+\beta_g(z)). \tag{2.8}$$

In order to define discrete subgroups of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ consider

$$j_g(z)=\frac{\theta_0(gz)}{\theta_0(z)},\qquad z\in\mathfrak{H},\quad g\in\mathrm{SL}(2,\mathbb{R}), \tag{2.9}$$

with the classical theta function

$$\theta_0(z)=y^{1/4}\sum_{n\in\mathbb{Z}}e(n^2z). \tag{2.10}$$

The $j_g(z)$ form a multiplier system, since

$$j_{gh}(z)=\frac{\theta_0(ghz)}{\theta_0(z)}=\frac{\theta_0(ghz)}{\theta_0(hz)}\,\frac{\theta_0(hz)}{\theta_0(z)}=j_g(hz)j_h(z). \tag{2.11}$$

It is well known [38] that, for $g\in\Gamma_0(4)$, $j_g(z)$ takes the form

$$j_g(z)=\epsilon_d^{-1}\left(\frac{c}{d}\right)\left(\frac{cz+d}{|cz+d|}\right)^{1/2}=\epsilon_d^{-1}\left(\frac{c}{d}\right)\mathrm{e}^{\mathrm{i}(\arg(cz+d))/2}. \tag{2.12}$$

Here ϵ_d is defined by

$$\epsilon_d=\begin{cases}1 & \text{if } d\equiv 1 \mod 4\\ \mathrm{i} & \text{if } d\equiv 3 \mod 4,\end{cases} \tag{2.13}$$

and $\left(\frac{c}{d}\right)$ denotes the generalized quadratic residue symbol defined as follows. For an odd prime p and an integer a prime to p, the *quadratic residue symbol* $\left(\frac{a}{p}\right)$ has the value $+1$, if a is a square modulo p, i.e., if $x^2\equiv a \mod p$ has a solution, and -1, if this is not the case. One has $\left(\frac{a}{p}\right)\equiv a^{\frac{p-1}{2}} \mod p$. For an integer a and an odd integer b the *generalized quadratic residue symbol* $\left(\frac{a}{b}\right)$ is characterized by the following properties (cf. [38]):

(i) $\left(\frac{a}{b}\right)=0$ if $(a,b)\neq 1$.
(ii) If b is an odd prime, $\left(\frac{a}{b}\right)$ coincides with the ordinary quadratic residue symbol.
(iii) If $b>0$, $\left(\frac{\cdot}{b}\right)$ defines a character modulo b.
(iv) If $a\neq 0$, $\left(\frac{a}{\cdot}\right)$ defines a character modulo a divisor of $4a$, whose conductor is the conductor of $\mathbb{Q}(\sqrt{a})$ over $\mathbb{Q}$.
(v) $\left(\frac{a}{-1}\right)=\operatorname{sgn} a$.
(vi) $\left(\frac{0}{\pm 1}\right)=1$.

Since the symbol is multiplicative, we have in particular for $b=\prod_j p_j$, with primes p_j (not necessarily distinct)

$$\left(\frac{a}{b}\right)=\begin{cases}\prod_j\left(\frac{a}{p_j}\right) & (b>0)\\ \operatorname{sgn}(a)\prod_j\left(\frac{a}{p_j}\right) & (b<0),\end{cases}$$

so that the generalized symbol can be written as a product of familiar quadratic residue symbols. If $a=u^2t$, where t is square-free, we have

$$\left(\frac{u^2t}{b}\right)=\left(\frac{t}{b}\right).$$

For more information on the properties of the generalized quadratic residue symbol, see [27].

Notice that

$$j_g(z)^2=\frac{cz+d}{|cz+d|},\qquad \text{for } g\in\Gamma_1(4). \tag{2.14}$$

Furthermore we have (recall the definition (2.5) of the element γ_0)

$$j_{\gamma_0}(z)=j_{\gamma_0^{-1}}(z)=\mathrm{e}^{-\mathrm{i}\,\pi/4}\left(\frac{z}{|z|}\right)^{1/2}=\mathrm{e}^{-\mathrm{i}\,\pi/4}\;\mathrm{e}^{\mathrm{i}\arg(z)/2}, \tag{2.15}$$

which results from the basic transformation formula of theta series, which in turn is a direct consequence of Poisson's summation formula (compare also Proof # 2 of Proposition 3.1, Section 3).

To get an explicit expression for $j_h(z)$ for all elements h of the theta group $\Gamma_\theta=\Gamma_0(4)\cup\Gamma_0(4)\gamma_0$, let us calculate $j_{g\gamma_0}$ with $g\in\Gamma_0(4)$. Since $j_{g\gamma_0}(z)=j_{-g\gamma_0}(z)$, we may assume without loss of generality that $g\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ with $d>0$. Then

$$\begin{aligned}j_{g\gamma_0}(z)&=j_g(\gamma_0 z)j_{\gamma_0}(z)\\&=\epsilon_d^{-1}\left(\frac{c}{d}\right)\mathrm{e}^{\mathrm{i}(\arg(-c/4z+d))/2}\;\mathrm{e}^{-\mathrm{i}\,\pi/4}\;\mathrm{e}^{\mathrm{i}\arg(z)/2}\\&=\mathrm{e}^{-\mathrm{i}\,\pi/4}\;\epsilon_d^{-1}\left(\frac{c}{d}\right)\mathrm{e}^{\mathrm{i}(\arg(4dz-c))/2}.\end{aligned}$$

Hence

$$j_{g\gamma_0}(z) = \mathrm{e}^{-\mathrm{i}\,\pi/4}\ \epsilon_d^{-1} \left(\frac{c}{d}\right) \left(\frac{4dz - c}{|4dz - c|}\right)^{1/2}, \qquad d > 0. \tag{2.16}$$

By virtue of (2.11) – (2.14) and (2.4) we can view

$$\Delta_1(4) = \left\{[g, \beta_g] \ : \ g \in \Gamma_1(4),\ \beta_g \text{ a function on } \mathfrak{H} \text{ s.t. } \mathrm{e}^{\mathrm{i}\,\beta_g(z)/2} = j_g(z)\right\}$$

as a discrete subgroup of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$. Another example for a discrete subgroup is

$$\Delta_0(4) = \Delta_1(4) \ \cup\ \Delta_1(4)[-1,\pi] \ \cup\ \Delta_1(4)[1,2\pi] \ \cup\ \Delta_1(4)[-1,3\pi], \tag{2.17}$$

which can be written as the set

$$\Delta_0(4) = \left\{[g, \beta_g] \ : \ g \in \Gamma_0(4),\ \beta_g \text{ a fct. on } \mathfrak{H} \text{ s.t. } \mathrm{e}^{\mathrm{i}\,\beta_g(z)} = \frac{cz+d}{|cz+d|}\right\}.$$

$\Delta_1(4)$ is clearly normal in $\Delta_0(4)$. Third, we can construct the discrete subgroup

$$\Delta_\theta = \left\{[g, \beta_g] \ : \ g \in \Gamma_\theta,\ \beta_g \text{ a function on } \mathfrak{H} \text{ s.t. } \mathrm{e}^{\mathrm{i}\,\beta_g(z)} = \frac{cz+d}{|cz+d|}\right\},$$

that is,

$$\Delta_\theta = \Delta_0(4) \ \cup\ \Delta_0(4)[\gamma_0, \arg], \tag{2.18}$$

and remark that $[\gamma_0, \arg]^2 = [-1, \pi]$, and $[\gamma_0, \arg]^{-1} = [\gamma_0^{-1}, \arg(\,\cdot\,) - \pi]$.

Both $\Delta_0(4)$ and $\Delta_1(4)$ are normal in Δ_θ, and the easiest way to see this is as follows: we have to show that, when $[g, \beta_g] \in \Delta_1(4)$, also its conjugate $[\gamma_0, \arg]^{-1}[g, \beta_g][\gamma_0, \arg]$ is contained in $\Delta_1(4)$. Since it is already known that $\gamma_0^{-1} g \gamma_0 \in \Gamma_1(4)$ (see last section), it remains to be shown

$$\mathrm{e}^{\mathrm{i}[\arg(g\gamma_0 z) - \pi]/2}\ j_g(\gamma_0 z)\ \mathrm{e}^{\mathrm{i}\,\arg(z)/2} = j_{\gamma_0^{-1} g \gamma_0}(z).$$

Notice that by (2.15) the left hand side equals

$$\mathrm{e}^{\mathrm{i}\,\pi/4}\ j_{\gamma_0^{-1}}(g\gamma_0 z)\ \mathrm{e}^{-\mathrm{i}\,\pi/2}\ j_g(\gamma_0 z)\ \mathrm{e}^{\mathrm{i}\,\pi/4}\ j_{\gamma_0}(z)$$

which in turn is equal to $j_{\gamma_0^{-1} g \gamma_0}(z)$, giving the claim.

With the help of one of those discrete subgroups, we can construct a manifold of finite volume, which looks locally like $\mathfrak{H} \times \mathbb{R}$. We do this by taking the quotient

$$\mathcal{M} = \Delta \backslash \widetilde{\mathrm{SL}}(2,\mathbb{R}) = \{\Delta \tilde{g} \ : \ \tilde{g} \in \widetilde{\mathrm{SL}}(2,\mathbb{R})\}, \tag{2.19}$$

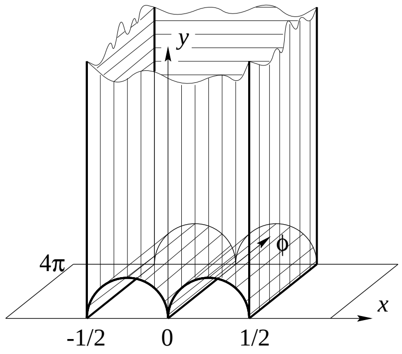


FIG. 8. *The fundamental region* $\mathcal{F}_{\Delta_1(4)}$ *of* $\Delta_1(4)$ *in* $\mathfrak{H} \times \mathbb{R}$.

where Δ is the discrete subgroup. This means, a point of the manifold $\mathcal{M}$ represents the class of all points of $\widetilde{\mathrm{SL}}(2,\mathbb{R}) \simeq \mathfrak{H} \times \mathbb{R}$, which are equivalent under Δ. $\mathcal{M}$ is indeed a manifold, because the discrete subgroups we are concerned with here, act without fixed points on $\mathfrak{H} \times \mathbb{R}$.

We will denote the manifolds corresponding to $\Delta_1(4)$, $\Delta_0(4)$, and Δ_θ, by

$$\mathcal{M}_1 = \Delta_1 \backslash \widetilde{\mathrm{SL}}(2,\mathbb{R}), \quad \mathcal{M}_0 = \Delta_0 \backslash \widetilde{\mathrm{SL}}(2,\mathbb{R}), \quad \mathcal{M}_\theta = \Delta_\theta \backslash \widetilde{\mathrm{SL}}(2,\mathbb{R}),$$

respectively.

Let us denote by Z_1 the discrete subgroup generated by the element $[-1,\pi]$. Then, by virtue of the decomposition (2.17), we have

$$\Delta_1(4) = Z_1 \backslash \Delta_0(4) = \Delta_0(4)/Z_1. \tag{2.20}$$

It does not matter, on which side we divide by Z_1, since Z_1 is contained in the center of $\Delta_0(4)$. $\Delta_1(4)$ is therefore normal of index 4 in $\Delta_0(4)$ (as noted before) so that we can view $\mathcal{M}_1$ as a 4-fold cover of $\mathcal{M}_0$, i.e., $\mathcal{M}_0 = \mathbb{Z}_4 \backslash \mathcal{M}_1$. In a similar way one finds that $\mathcal{M}_1$ and $\mathcal{M}_0$ are finitely covered by $\mathcal{M}_\theta$, for $\Delta_1(4)$ and $\Delta_0(4)$ are normal in Δ_θ. The corresponding automorphism group, which identifies the different parts of the cover, is generated by the transformation $[\gamma_0, \arg]$ and thus isomorphic to $\mathbb{Z}_8$,[7] compare the remarks after equation (2.18).

Later on, we will also view the manifolds $\mathcal{M}_0$ and $\mathcal{M}_\theta$ as quotients of $\mathrm{PSL}(2,\mathbb{R}) \simeq \mathrm{SL}(2,\mathbb{R})/\{\pm 1\}$, with $\Gamma_0(4)$ and Γ_θ, respectively. In order to see why this is possible, let us identify $\mathrm{PSL}(2,\mathbb{R})$ with $\widetilde{\mathrm{SL}}(2,\mathbb{R})/Z_1$, since then

$$\mathcal{M}_0 = \Gamma_0(4) \backslash \mathrm{PSL}(2,\mathbb{R}), \qquad \mathcal{M}_\theta = \Gamma_\theta \backslash \mathrm{PSL}(2,\mathbb{R}).$$

[7]To be more precise, the automorphism group is generated by the element $\Delta_1(4)[\gamma_0, \arg]$, and since $(\Delta_1(4)[\gamma_0, \arg])^8 = \Delta_1(4)$, the group is cyclic of order eight, hence isomorphic to $\mathbb{Z}_8$.

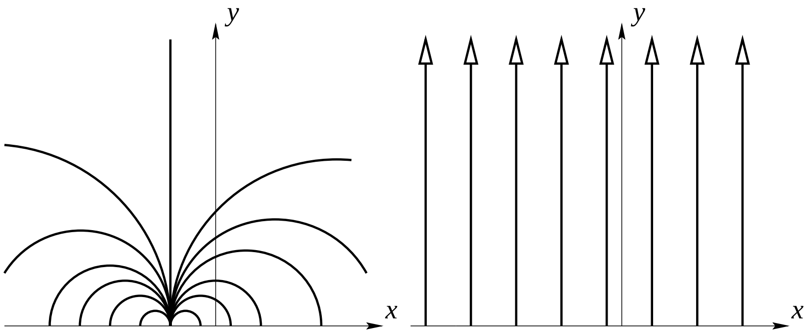


FIG. 9. *Two examples for a family of geodesics with common limit point. The limit point of the family shown left is on the real axis, the one on the right hand side is at infinity.*

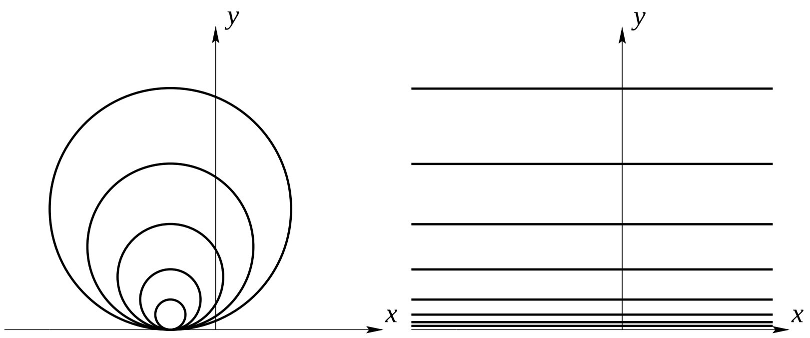


FIG. 10. *Two examples for a family of horocycles with common limit point. As in the previous figure the limit points are on the real axis and at infinity, respectively.*

It is convenient to choose the following fundamental regions in $\mathfrak{H} \times \mathbb{R}$: $\mathcal{F}_{\Delta_1(4)} = \mathcal{F}_{\Gamma_1(4)} \times [0, 4\pi)$, $\mathcal{F}_{\Delta_0(4)} = \mathcal{F}_{\Gamma_0(4)} \times [0, \pi)$, $\mathcal{F}_{\Delta_\theta} = \mathcal{F}_{\Gamma_\theta} \times [0, \pi)$, see Figure 8.

2.3. Geodesics and horocycles. The unit tangent bundle $\mathrm{T}_1\,\mathfrak{H}$ of the hyperbolic plane $\mathfrak{H}$ can be parameterized by a complex coordinate $z \in \mathfrak{H}$, and an angle θ, which describes the orientation of an arrow attached to z. The bundle $\mathrm{T}_1\,\mathfrak{H}$ is therefore naturally identified with $\mathfrak{H} \times (\mathbb{R}/2\pi\mathbb{Z})$.[8] The action of the isometry group $\mathrm{PSL}(2,\mathbb{R})$ on $\mathrm{T}_1\,\mathfrak{H}$ is canonically given by

$$g : (z, \theta) \mapsto (gz, \theta - 2\beta_g(z)), \tag{2.21}$$

see, e.g., [33]. gz means the usual action via fractional linear transformations (2.1), the transformation of θ follows from the transformation property of a tangent vector dz under $g \in \mathrm{SL}(2,\mathbb{R})$. The transformation recipe (2.21) looks more familiar, if we substitute $\theta = -2\phi$, since then the action

[8]Speaking in dynamical terms, $\mathrm{T}_1\,\mathfrak{H}$ is the hypersurface of constant energy in the phase space of a point particle sliding on the hyperbolic plane with unit speed.

of $\mathrm{PSL}(2,\mathbb{R})$ (view the latter again as $\widetilde{\mathrm{SL}}(2,\mathbb{R})/Z_1$) is,

$$g:(z,\phi)\mapsto(gz,\phi+\beta_g(z)),\tag{2.22}$$

compare (2.8). It thus makes sense to identify $\mathrm{PSL}(2,\mathbb{R})\simeq\widetilde{\mathrm{SL}}(2,\mathbb{R})/Z_1\simeq\mathfrak{H}\times(\mathbb{R}/\pi\mathbb{Z})$ with the unit tangent bundle $\mathrm{T}_1\,\mathfrak{H}$, because then the action of $\mathrm{PSL}(2,\mathbb{R})$ on $\mathrm{T}_1\,\mathfrak{H}\simeq\mathrm{PSL}(2,\mathbb{R})$ is given by group multiplication.

The previous argumentation applies also to the unit tangent bundle $\mathrm{T}_1\,M$ of a hyperbolic manifold $M=\Gamma\backslash\mathfrak{H}$.[9] Here, $\mathrm{T}_1\,M$ can be identified with $\Gamma\backslash(\mathrm{T}_1\,\mathfrak{H})$ and thus with $\Gamma\backslash\,\mathrm{PSL}(2,\mathbb{R})$. Therefore our investigation of flows on the manifolds $\mathcal{M}_0$ and $\mathcal{M}_\theta$ translates to the study of geodesic and horocycle flows on the unit tangent bundle of $\Gamma_0\backslash\mathfrak{H}$ and $\Gamma_\theta\backslash\mathfrak{H}$. In the most important case of the manifold $\mathcal{M}_1$, we are, however, faced with a 4-fold cover of such a tangent bundle, which means that we have to generalize the concept of the geodesic flow (Figure 9) and the horocycle flow (Figure 10) on tangent bundles to finite covers. To this end we recall that the geodesic flow and the horocycle flow on $\mathrm{T}_1\,M\simeq\Gamma\backslash\,\mathrm{PSL}(2,\mathbb{R})$ can be represented by the right translations

$$\mathrm{PSL}(2,\mathbb{R})\to\mathrm{PSL}(2,\mathbb{R}),\qquad g\mapsto g\begin{pmatrix}\mathrm{e}^{-t/2} & 0\\ 0 & \mathrm{e}^{t/2}\end{pmatrix},\tag{2.23}$$

and

$$\mathrm{PSL}(2,\mathbb{R})\to\mathrm{PSL}(2,\mathbb{R}),\qquad g\mapsto g\begin{pmatrix}1 & t\\ 0 & 1\end{pmatrix},\tag{2.24}$$

respectively. This motivates our following definition of flows on $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ and $\Delta\backslash\widetilde{\mathrm{SL}}(2,\mathbb{R})$, which we still denote as *geodesic* or *horocycle flows.*

Let $\{\Phi^t\}_{t\in\mathbb{R}}$ be a one-parameter subgroup in $\widetilde{\mathrm{SL}}(2,\mathbb{R})$, i.e., $\Phi^s\Phi^t=\Phi^{s+t}$. Assume $\operatorname{tr}\Phi^t=2\cosh(t/2)$. (By the trace of an element in $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ we mean the trace of its $\mathrm{SL}(2,\mathbb{R})$ component.) Then Φ^t is conjugate to

$$\Phi_0^t=\left[\begin{pmatrix}\mathrm{e}^{-t/2} & 0\\ 0 & \mathrm{e}^{t/2}\end{pmatrix},0\right],$$

i.e., we find a $\tilde g$ such that $\Phi^t\tilde g=\tilde g\,\Phi_0^t$. We define the *geodesic flow on* $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ as the right translation by Φ_0^t,

$$\Phi^t:\widetilde{\mathrm{SL}}(2,\mathbb{R})\to\widetilde{\mathrm{SL}}(2,\mathbb{R}),\qquad \tilde g\mapsto\tilde g\,\Phi_0^t.\tag{2.25}$$

In the same fashion we define the *horocycle flow*

$$\Psi^t:\widetilde{\mathrm{SL}}(2,\mathbb{R})\to\widetilde{\mathrm{SL}}(2,\mathbb{R}),\qquad \tilde g\mapsto\tilde g\,\Psi_0^t,\tag{2.26}$$

[9] $\mathrm{T}_1\,M$ is the hypersurface of constant energy in the phase space of a point particle sliding on a closed hyperbolic surface.

with

$$\Psi_0^t = \left[\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, 0 \right].$$

The associated one-parameter subgroups $\{\Psi^t\}_{t\in\mathbb{R}}$ are characterized by the condition $\operatorname{tr}\Psi^t = 2$.

We call the set $\{\tilde{g}\,\Phi_0^t\}_{t\in[0,T]}$ a *geodesic of length* T. $\{\tilde{g}\,\Psi_0^t\}_{t\in[0,T]}$ shall be called a *horocycle of length* T.

For a discrete subgroup Δ we define the geodesic flow on the quotient by

$$\Phi^t : \Delta\backslash\widetilde{\mathrm{SL}}(2,\mathbb{R}) \to \Delta\backslash\widetilde{\mathrm{SL}}(2,\mathbb{R}), \qquad \Delta\tilde{g} \mapsto \Delta\tilde{g}\,\Phi_0^t. \tag{2.27}$$

A geodesic $\{\Delta\tilde{g}\,\Phi_0^t\}_{t\in[0,T]}$ is closed, if and only if $\Phi^T = \tilde{g}\,\Phi_0^T\,\tilde{g}^{-1}$ is in Δ. T will then be the period of the closed geodesic, and the smallest $t>0$ for which $\Phi^t \in \Delta$ will be its *primitive* period. The same terminology applies to closed horocycles.

The following proposition describes the ergodicity of the geodesic flow, which implies that a generic geodesic becomes equidistributed on $\mathcal{M}_1$ (see Corollary 2.1 below). In the case when the manifold in concern is the unit tangent bundle of a hyperbolic surface (such as $\mathcal{M}_0$ or $\mathcal{M}_\theta$), the proof is due to Hopf [20]. It is not hard to see that his proof will still hold for $\mathcal{M}_1$.

PROPOSITION 2.1. *Let $F \in \mathrm{L}^1(\mathcal{M}_1)$. Then for almost all $\tilde{g}\in\mathcal{M}_1$*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T F(\tilde{g}\Phi_0^t)dt = \frac{1}{\mu(\mathcal{M}_1)}\int_{\mathcal{M}_1} F(\tilde{g})d\mu(\tilde{g})$$

with respect to Haar measure $d\mu(\tilde{g}) = dx\,dy\,d\phi/y^2$.

REMARKS. In the case when F is uniformly continuous, it is well known [20], cf. also [9], p. 114, that if equidistribution holds for some $\tilde{g}$, it will hold as well for all other points $\tilde{h}$ on the stable manifold of $\tilde{g}$.[10]

This means in particular that the relation

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T F(x+\mathrm{i}\,\mathrm{e}^{-t},0)dt = \frac{1}{\mu(\mathcal{M}_1)}\int_{\mathcal{M}_1} F(\tilde{g})d\mu(\tilde{g}) \tag{2.28}$$

holds for almost all x, with respect to Lebesgue measure on $\mathbb{R}$.

By a standard density argument, the function F can even be taken as a characteristic function.

COROLLARY 2.1. *Let $\chi_{\mathcal{D}}$ be the characteristic function of an open subset $\mathcal{D}\subset\mathcal{M}_1$ with boundary of measure $\mu(\partial\mathcal{D})=0$, then*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \chi_{\mathcal{D}}(x+\mathrm{i}\,\mathrm{e}^{-t},0)dt = \frac{\mu(\mathcal{D})}{\mu(\mathcal{M}_1)} \tag{2.29}$$

[10] i.e. for all points $\tilde{h}$ of the form $\tilde{h} = \tilde{g}\,\Phi_0^v \left[\begin{pmatrix} 1 & 0 \\ u & 1 \end{pmatrix}, \beta \right]$ with $\beta(z) = \arg(uz+1)$, for any (fixed) $u, v \in \mathbb{R}$.

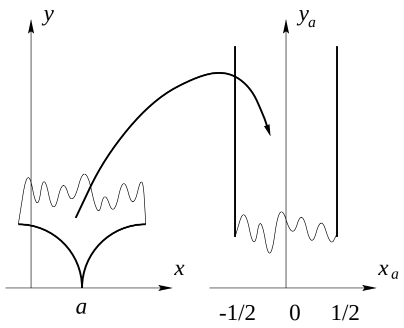


FIG. 11. *In order to study asymptotic relations in a cusp at some point a, it is convenient to introduce a new coordinate system with coordinates $z_a = x_a + \mathrm{i}\, y_a$ running transversal and parallel with respect to the cusp. These new coordinates can be chosen in a way that the cusp looks like the standard cusp at infinity with unit width.*

holds for almost all x.

The next proposition is concerned with the equidistribution of closed horocycles on $\mathcal{M}_1$. In Section 5 we will also give the next to leading order terms, for the price of a smaller class of allowed test functions. While here we can use piecewise continuous functions which do not grow too strongly in the cusp, we are later forced to restrict ourselves to infinitely differentiable functions with compact support. In order to characterize the growth condition in the cusps, we introduce a new set of coordinates (z_a, ϕ_a) for each cusp,

$$\begin{cases} (z_0, \phi_0) = (-(4z)^{-1}, \phi + \arg z) \\ (z_{1/2}, \phi_{1/2}) = (-(4z-2)^{-1}, \phi + \arg(z - 1/2)) \\ (z_\infty, \phi_\infty) = (z, \phi). \end{cases} \tag{2.30}$$

In these coordinates (z_a, ϕ_a), each cusp is represented as the standard cusp at ∞ with unit width (Figure 11).

PROPOSITION 2.2. *Let F be piecewise continuous on $\mathcal{M}_1$, and assume $F(z,\phi) = O(y_a^\sigma)$ for some $\sigma < 1$, as $y_a \to \infty$ $(a = 0, 1/2, \infty)$. Then for*

$$\tilde{g} = \left[\begin{pmatrix} T^{-1/2} & 0 \\ 0 & T^{1/2} \end{pmatrix}, 0 \right]$$

we have

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T F(\tilde{g}\Psi_0^t)dt = \frac{1}{\mu(\mathcal{M}_1)} \int_{\mathcal{M}_1} F(\tilde{g})d\mu(\tilde{g}).$$

This proposition is proved in [29], compare also Sarnak [33], Eskin and McMullen [11] (Theorem 7.1), and Hejhal [19] for the equidistribution of long closed horocycles on the unit tangent bundle of a non-compact hyperbolic surface. In coordinates of $\mathfrak{H}\times\mathbb{R}$ the last relation reads

$$\lim_{y\to 0}\int_0^1 F(z,0)dx=\frac{1}{\mu(\mathcal{M}_1)}\int_{\mathcal{M}_1} F(z,\phi)\frac{dx\,dy\,d\phi}{y^2}. \tag{2.31}$$

In Theorem 5.1 we will give some information about the rate of convergence.

3. Theta sums. The main reason for our interest in the manifold $\mathcal{M}_1$ in this paper is the following proposition.

PROPOSITION 3.1. *For every $f\in\mathcal{S}(\mathbb{R})$ there exists a function $\Theta_f\in\mathrm{C}^\infty(\mathcal{M}_1)$ such that*

$$\Theta_f(z,0)=y^{1/4}\sum_{n\in\mathbb{Z}} f(ny^{1/2})\,e(n^2x).$$

We will give two alternative proofs. The first one is more instructive, and is contained in [29]. The advantage of the second one is that it can be readily generalized to functions, which are not of Schwartz class.

Proof #1. The ν-th Hermite function h_ν reads

$$h_\nu(t)=(2^{\nu-1}\nu!)^{-1/2}H_\nu(2\pi^{1/2}t)\,\mathrm{e}^{-2\pi t^2}, \tag{3.1}$$

with the Hermite polynomial

$$H_\nu(t)=(-1)^\nu\,\mathrm{e}^{t^2}\,\frac{d^\nu}{dt^\nu}\,\mathrm{e}^{-t^2}.$$

The h_ν form a orthonormal basis of $\mathrm{L}^2(\mathbb{R})$ with respect to the standard scalar product, i.e., $(h_\mu,h_\nu)=\delta_{\mu\nu}$. Let $\hat{f}(\nu)=(f,h_\nu)$ be the ν-th Hermite coefficient. Since $f\in\mathcal{S}(\mathbb{R})$ the Hermite expansion

$$\sum_{\nu=0}^N \hat{f}(\nu)h_\nu(t)\to f(t) \tag{3.2}$$

converges uniformly in t due to the bounds

$$|\hat{f}(\nu)|\leq C_m(2\nu+1)^{-m},\quad \text{for any } m, \tag{3.3}$$

and

$$|h_\nu(t)|\leq\begin{cases} C\big((2\nu+1)^{1/3}+|4\pi t^2-(2\nu+1)|\big)^{-1/4} & \text{for } 2\pi t^2\leq 2\nu+1\\ C\,\mathrm{e}^{-\gamma t^2} & \text{for } 2\pi t^2>2\nu+1,\end{cases} \tag{3.4}$$

with positive constants C_m, C, γ. (For this and much more, see [42].)

Let us consider the theta function

$$\theta_\nu(z)=y^{1/4}\sum_{n\in\mathbb{Z}} h_\nu(ny^{1/2})\,e(n^2x), \tag{3.5}$$

with the transformation property [40][11]

$$\theta_\nu(gz) = j_g(z)^{2\nu+1}\theta_\nu(z), \quad \text{for every } g \in \Gamma_0(4), \tag{3.6}$$

$j_g(z) = \epsilon_d^{-1}\left(\frac{c}{d}\right)\left(\frac{cz+d}{|cz+d|}\right)^{1/2}$. We conclude that the function $\theta_\nu(z,\phi) = \theta_\nu(z)\,\mathrm{e}^{-\mathrm{i}(2\nu+1)\phi/2}$ is invariant under $\Delta_1(4)$, i.e., is a function on $\mathcal{M}_1$.

Put

$$\Theta_f(z,\phi) = y^{1/4}\sum_{n\in\mathbb{Z}} f_\phi(ny^{1/2})\, e(n^2x), \tag{3.7}$$

where

$$f_\phi(t) = \sum_{\nu=0}^{\infty} \hat{f}(\nu)\,\mathrm{e}^{-\mathrm{i}(2\nu+1)\phi/2}\, h_\nu(t), \tag{3.8}$$

since then $\Theta_f(z) = \Theta_f(z,0)$. The series $f_\phi(ny^{1/2})$ in (3.7) converges uniformly in n (z and ϕ are fixed), we can therefore exchange the order of summation to obtain

$$\Theta_f(z,\phi) = \sum_{\nu=0}^{\infty} \hat{f}(\nu)\theta_\nu(z,\phi), \tag{3.9}$$

which is a C^∞ function on $\mathcal{M}_1$. □

Proof #2. In Proof #1 we have used the expansion (3.8) to prove Proposition 3.1. A different representation of $f_\phi(t)$ can be obtained, for $f_\phi(t)$ is a solution of the time dependent Schrödinger equation for the harmonic oscillator,

$$\left(-\frac{1}{4\pi}\frac{\partial^2}{\partial t^2} + 4\pi t^2\right) f_\phi(t) = 2\,\mathrm{i}\,\frac{\partial}{\partial\phi} f_\phi(t) \tag{3.10}$$

(here the variable ϕ plays the role of the time). As a consequence we can write $f_\phi(t)$ as

$$f_\phi(t) = \int_{-\infty}^{\infty} G_\phi(t,t') f(t') dt', \tag{3.11}$$

with the Green function

$$G_\phi(t,t') = 2^{1/2} e(-\sigma_\phi/8)|\sin\phi|^{-1/2} e\left[\frac{(t^2+t'^2)\cos\phi - 2t\,t'}{\sin\phi}\right],$$

[11] This transformation formula for the series $\theta_\nu(z)$ can be quickly deduced from the transformation formula (2.9) for the classical theta series $\theta_0(z)$, if one uses Maass operators to map $\theta_\nu(z)$ to $\theta_{\nu+2}(z)$, which in turn are closely related to the creation and annihilation operators of the harmonic oscillator. The classical series $\theta_0(z)$ can be viewed as a ground or vacuum state of the Shale-Weil representation. For more background information, see [27].

where $\sigma_\phi = 2\nu + 1$ when $\nu\pi < \phi < (\nu+1)\pi$.

Poisson's summation formula yields

$$\Theta_f(z,\phi) = y^{-1/4} \sum_{n\in\mathbb{Z}} \int_{-\infty}^{\infty} e\left(t^2 \frac{x}{y} - \frac{n\,t}{y^{1/2}}\right) f_\phi(t)dt, \tag{3.12}$$

and by virtue of the relation

$$G_{\arg z}\left(n\,\frac{y^{1/2}}{2|z|}, t\right) = \mathrm{e}^{-\mathrm{i}\,\pi/4}\, y^{-1/4} \left(\frac{y}{4|z|^2}\right)^{-1/4} e\left(n^2 \frac{x}{4|z|^2} + t^2\frac{x}{y} - \frac{n\,t}{y^{1/2}}\right),$$

we recover the transformation formula

$$\Theta_f(-(4z)^{-1}, \phi + \arg z) = \mathrm{e}^{-\mathrm{i}\,\pi/4}\,\Theta_f(z,\phi). \tag{3.13}$$

The element of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ corresponding to this transformation is $[\gamma_0, \arg]$. Notice that repeated application yields $\Theta_f(z,\phi+\pi) = -\mathrm{i}\,\Theta_f(z,\phi)$, since $[\gamma_0,\arg]^2 = [-1,\pi]$.[12] Since the group Δ_θ is generated by $[\gamma_0,\arg]$ and $[\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right), 0]$, for which the transformation formula is trivially $\Theta_f(z+1,\phi) = \Theta_f(z,\phi)$, we know the transformation properties of Θ_f under the discrete group Δ_θ,

$$\Theta_f\big(\tilde g(z,\phi)\big) = \chi(\tilde g)\Theta_f(z,\phi), \qquad \tilde g \in \Delta_\theta, \tag{3.14}$$

where the character $\chi(\tilde g)$ is an 8th root of unity. The crucial observation is that $\chi(\tilde g)$ does not depend on the choice of the test function f, so in order to work out the value of $\chi(\tilde g)$ for some special $\tilde g$, it suffices to consider only the well known transformation properties of the classical theta function $\theta_0(z,\phi)$, see Section 2.2. We conclude that $\chi(\tilde g) = 1$ for every $\tilde g \in \Delta_1(4)$. This completes the second proof of Proposition 3.1. □

We state the additional transformation formulae, which we found in Proof #2, as a corollary. Recall the coset decompositions (2.18) and (2.17); these may be summarized as

$$\Delta_\theta = \bigcup_{\nu=1}^{8} \Delta_1(4)[\gamma_0,\arg]^\nu. \tag{3.15}$$

Corollary 3.1. *For $\tilde g \in \Delta_\theta$, we have the transformation formula*

$$\Theta_f\big(\tilde g(z,\phi)\big) = \chi(\tilde g)\Theta_f(z,\phi), \qquad \tilde g \in \Delta_\theta,$$

where the value of the character $\chi(\tilde g)$ is given by

$$\chi(\tilde g) = \mathrm{e}^{-\mathrm{i}\,\pi\nu/4},$$

[12] In the discussion of the transformation properties of the classical theta series $\theta_0(z)$, we have already stated formula (3.13) indirectly in equation (2.15). The formula $\Theta_f(z,\phi+\pi) = -\mathrm{i}\,\Theta_f(z,\phi)$ is also a direct consequence of the relation $f_{\phi+\pi}(t) = -\mathrm{i}\, f_\phi(t)$.

if $\tilde{g}$ belongs to the coset $\Delta_1(4)[\gamma_0, \arg]^\nu$.

The theta functions Θ_f span a subspace of $\mathrm{L}^2(\mathcal{M}_1)$ which is isometric to the space of even functions in $\mathrm{L}^2(\mathbb{R})$. The following statements are proved in [29].

PROPOSITION 3.2. *Let $f \in \mathrm{L}^2(\mathbb{R})$. Then there exists a function $\Theta_f \in \mathrm{L}^2(\mathcal{M}_1)$ such that*

$$\|\Theta_f - \sum_{\nu=0}^{N} \hat{f}(\nu)\, \theta_\nu\|_{\mathrm{L}^2(\mathcal{M}_1)} \to 0$$

as $N \to \infty$, where $\hat{f}(\nu)$ are the Hermite coefficients of f.

REMARK. If f is even, i.e., $f(t) = f(-t)$, then

$$\frac{1}{16\pi^2}\langle \Theta_f, \theta_\nu \rangle = \frac{1}{16\pi^2} \int_{\mathcal{M}_1} \Theta_f(z,\phi)\, \overline{\theta_\nu(z,\phi)}\, d\mu(z,\phi) = \hat{f}(\nu), \tag{3.16}$$

$$\frac{1}{4\pi}\|\Theta_f\|_{\mathrm{L}^2(\mathcal{M}_1)} = \|f\|_{\mathrm{L}^2(\mathbb{R})}. \tag{3.17}$$

3.1. Asymptotic behaviour at rational argument. The next proposition gives the asymptotic behaviour of $\Theta_f(z,\phi)$ in the cusps at $(0,\phi)$, $(1/2,\phi)$, and (∞,ϕ). Recall the definition (2.30) of the cusp coordinates (z_a,ϕ_a), $a = 0,\ 1/2,\ \infty$.

PROPOSITION 3.3. *Let $f \in \mathcal{S}(\mathbb{R})$. Then*

$$\Theta_f(z,\phi) = \begin{cases} \mathrm{e}^{\mathrm{i}\,\pi/4}\, f_{\phi_0}(0)\, y_0^{1/4} + O_N(y_0^{-N}), & y_0 \to \infty \\ O_N(y_{1/2}^{-N}), & y_{1/2} \to \infty \\ f_{\phi_\infty}(0)\, y_\infty^{1/4} + O_N(y_\infty^{-N}), & y_\infty \to \infty, \end{cases}$$

for any N.

Proof (compare [29]). The asymptotic relation in the cusp at (∞,ϕ) is evident. The behaviour at $(0,\phi)$ may be explained in a similar manner, after one has applied transformation formula (3.13). The second relation then follows from

$$\begin{aligned} &\Theta_f(z,\phi) = 2^{1/2}\Theta_f(4z-2,\phi) - \Theta_f(z-1/2,\phi) \\ &= 2^{1/2}\,\mathrm{e}^{\mathrm{i}\,\pi/4}\,\Theta_f(z_{1/2}/4,\phi_{1/2}) - \mathrm{e}^{\mathrm{i}\,\pi/4}\,\Theta_f(z_{1/2},\phi_{1/2}) = O_N(y_{1/2}^{-N}). \end{aligned} \tag{3.18}$$

□

Proposition 3.3 implies the following asymptotic relations:

$$\Theta_f(z,\phi) = \begin{cases} \langle f_\phi \rangle\, y^{-1/4} + O_N(y^N), & x = 0, \quad y \to 0 \\ O_N(y^N), & x = 1/2, \quad y \to 0 \\ f_\phi(0)\, y^{1/4} + O_N(y^{-N}), & y \to \infty, \end{cases} \tag{3.19}$$

with the average $\langle \cdot \rangle$ defined by

$$\langle f_\phi \rangle = \int_{-\infty}^{\infty} f_\phi(t)\, dt \tag{3.20}$$

These relations immediately follow from Proposition 3.3, if one substitutes

$$\begin{cases} y_0 = \frac{y}{|2z|^2} = \frac{1}{4y} & (x=0) \\ \phi_0 = \phi + \arg z = \phi + \pi/2 & (x=0) \\ y_{1/2} = \frac{y}{|2z-1|^2} = \frac{1}{4y} & (x=1/2). \end{cases} \tag{3.21}$$

Thus for $x=0$ we have

$$\Theta_f(z,\phi) = 2^{-1/2}\, \mathrm{e}^{\mathrm{i}\,\pi/4}\, f_{\phi+\pi/2}(0)\; y^{-1/4} + O_N(y^N), \tag{3.22}$$

which can be simplified using (3.11)

$$f_{\phi+\alpha}(t) = \int_{-\infty}^{\infty} G_\alpha(t,t') f_\phi(t') dt', \tag{3.23}$$

in particular, the function $f_{\phi+\pi/2}$ is the Fourier transform of f_ϕ,

$$f_{\phi+\pi/2}(t) = 2^{1/2}\, \mathrm{e}^{-\mathrm{i}\,\pi/4} \int_{-\infty}^{\infty} e(-2tt') f_\phi(t')\, dt. \tag{3.24}$$

This proves the first relation. Clearly, this relation represents the familiar estimate

$$\sum_{n\in\mathbb{Z}} f_\phi(n/R) = \langle f_\phi \rangle\; R + O_N(R^{-N}), \qquad R\to\infty,$$

which, of course, can be proved directly, e.g., by the Poisson summation formula.

It is well known that every rational point $x = p/q$ on the boundary of the upper half plane $\mathfrak{H}$ can be mapped to ∞ by some element of the modular group [37]. Hence every boundary point $(p/q, \phi)$ is $\Delta_1(4)$-equivalent to a point $(0,\phi')$, $(1/2,\phi')$, or (∞,ϕ'), by some element of $\Delta_1(4)$. In order to see which rational point corresponds to which cusp, let us assume that p and q are coprime. It is then clear that only one of the three equations

$$0 = \frac{a\,\frac{p}{q} + b}{c\,\frac{p}{q} + d} \quad \text{(i)}, \qquad 1/2 = \frac{a\,\frac{p}{q} + b}{c\,\frac{p}{q} + d} \quad \text{(ii)}, \qquad \infty = \frac{a\,\frac{p}{q} + b}{c\,\frac{p}{q} + d} \quad \text{(iii)}, \tag{3.25}$$

has a solution with

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(4),$$

namely, (i) if q is odd, (ii) if $q \equiv 2 \mod 4$, and (iii) if $q \equiv 0 \mod 4$, respectively. Hence the point $(p/q, \phi)$ is $\Delta_1(4)$-equivalent to $\big(0, \phi + \beta_g(p/q)\big)$, $\big(1/2, \phi + \beta_g(p/q)\big)$, or $\big(\infty, \phi + \beta_g(p/q)\big)$, respectively.

Since the theta group Δ_θ identifies the cusps at 0 and ∞, the boundary points $(p/q, \phi)$, again with $(p, q) = 1$, are Δ_θ-equivalent to $\big(1/2, \phi + \beta_g(p/q)\big)$, or $\big(\infty, \phi + \beta_g(p/q)\big)$, as $q \equiv 2 \mod 4$ or not. As before, let us introduce cusp coordinates by

$$(z_{p/q}, \phi_{p/q}) = \big(\gamma_{p/q} z, \phi + \beta_{\gamma_{p/q}}(z)\big) = \tilde{\gamma}_{p/q}\,(z, \phi), \tag{3.26}$$

where $\gamma_{p/q}$ maps the cusp at p/q to the standard cusp at ∞, as before. If p/q is Γ_θ-equivalent to ∞ (in the case $q \not\equiv 2 \mod 4$), we agree to choose $\gamma_{p/q}$ in Γ_θ, i.e., $\tilde{\gamma}_{p/q} \in \Delta_\theta$. Then we have the following corollary of Proposition 3.3.

Corollary 3.2. *Let $f \in \mathcal{S}(\mathbb{R})$, and p, q coprime integers. Then*

$$\Theta_f(z, \phi) = \begin{cases} O_N(y_{p/q}^{-N}) & \text{if } q \equiv 2 \mod 4 \\ \chi(\tilde{\gamma}_{p/q})^{-1}\, f_{\phi_{p/q}}(0)\, y_{p/q}^{1/4} + O_N(y_{p/q}^{-N}) & \text{if } q \not\equiv 2 \mod 4, \end{cases}$$

for any N, as $y_{p/q} \to \infty$. The implied constant does not depend on p or q.

Proof. Suppose first $q \equiv 2 \mod 4$. Then the cusp is equivalent to the one at $1/2$, and the result is clear.[13] Now suppose, $q \not\equiv 2 \mod 4$, i.e., the cusp is Γ_θ-equivalent to the one at ∞. Thus we choose $\tilde{\gamma}_{p/q}$ to be an element of the theta group Δ_θ. Then $\Theta_f(z, \phi) = \chi(\tilde{\gamma}_{p/q})^{-1}\,\Theta_f\big(\tilde{\gamma}_{p/q}(z, \phi)\big)$, and this equals $\chi(\tilde{\gamma}_{p/q})^{-1}\,\Theta_f(z_{p/q}, \phi_{p/q})$. ☐

Corollary 3.3. *Let $f \in \mathcal{S}(\mathbb{R})$, and $p \neq 0$, $q > 0$ coprime integers. Then, for $y \to 0$,*

$$\Theta_f(p/q + \mathrm{i}\,y, \phi)$$
$$= \begin{cases} (1+\mathrm{i})\,\epsilon_p^{-1} \left(\frac{q}{p}\right)\, \langle f_\phi \rangle\, (q^2 y)^{-1/4} + O_N\big((q^2 y)^N\big) & \text{if } q \equiv 0 \mod 4 \\ O_N\big((q^2 y)^N\big) & \text{if } q \equiv 2 \mod 4 \\ \epsilon_q \left(\frac{p}{q}\right)\, \langle f_\phi \rangle\, (q^2 y)^{-1/4} + O_N\big((q^2 y)^N\big) & \text{if } q \text{ is odd,} \end{cases}$$

for any N. The implied constant does not depend on p or q.

Proof. Corollary 3.3 is a consequence of Corollary 3.2, for the following reasons.

Case 1. Let us start with the case $q \equiv 0 \mod 4$. Cusp coordinates and standard coordinates are related via $(z_{p/q}, \phi_{p/q}) = \tilde{\gamma}(z, \phi) = [\gamma, \beta_\gamma](z, \phi)$, where $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Hence $z_{p/q} = \frac{az+b}{cz+d}$ and $\phi_{p/q} = \phi + \beta_\gamma(z)$. The cusp at p/q is $\Gamma_0(4)$-equivalent to the one at ∞, which we see by choosing $p = -d$, $q = c$. (Recall that $(c, d) = 1$ because of $ad - bc = 1$.)

[13] The implied constant does not depend on p or q, because we have defined $z_{p/q}$ to be the canonical coordinates of the standard cusp of width 1.

Case 1a. $-p \equiv 1 \mod 4$. In this case we can choose $\tilde{\gamma} = [\gamma, \beta_\gamma] \in \Delta_1(4)$, where

$$\mathrm{e}^{\mathrm{i}\,\beta_\gamma(z)/2} = j_\gamma(z) = \left(\frac{q}{-p}\right) \mathrm{e}^{\mathrm{i}(\arg(qz-p))/2}, \tag{3.28}$$

see relation (2.12). We have $\left(\frac{q}{-p}\right) = \left(\frac{q}{-1}\right)\left(\frac{q}{p}\right) = \left(\frac{q}{p}\right)$, since $q > 0$. We thus obtain for $z = p/q + \mathrm{i}\,y$

$$\mathrm{e}^{\mathrm{i}\,\beta_\gamma(p/q+\mathrm{i}\,y)/2} = \left(\frac{q}{p}\right) \mathrm{e}^{\mathrm{i}\,\pi/4}, \tag{3.29}$$

and, as $\chi(\tilde{\gamma}) = 1$ (Corollary 3.1), $y_{p/q} = (q^2 y)^{-1}$, and $\phi_{p/q} = \phi + \beta_\gamma(p/q + \mathrm{i}\,y)$,

$$\chi(\tilde{\gamma})^{-1} f_{\phi_{p/q}}(0) = \left(\frac{q}{p}\right) f_{\phi+\pi/2}(0) = 2^{1/2}\,\mathrm{e}^{-\mathrm{i}\,\pi/4} \left(\frac{q}{p}\right) \langle f_\phi \rangle. \tag{3.30}$$

Case 1b. $-p \equiv 3 \mod 4$. Now we take $\tilde{\gamma} = [\gamma, \beta_\gamma] \in \Delta_1(4)[-1, \pi]$, where

$$\mathrm{e}^{\mathrm{i}\,\beta_\gamma(z)/2} = j_\gamma(z)\,\mathrm{e}^{\mathrm{i}\,\pi/2} = \left(\frac{q}{-p}\right) \mathrm{e}^{\mathrm{i}(\arg(qz-p))/2}, \tag{3.31}$$

see relation (2.12). Hence for $z = p/q + \mathrm{i}\,y$

$$\mathrm{e}^{\mathrm{i}\,\beta_\gamma(p/q+\mathrm{i}\,y)/2} = \left(\frac{q}{p}\right) \mathrm{e}^{\mathrm{i}\,\pi/4}, \tag{3.32}$$

and, as $\chi(\tilde{\gamma}) = \mathrm{e}^{-\mathrm{i}\,\pi/2}$ (Corollary 3.1),

$$\chi(\tilde{\gamma})^{-1} f_{\phi_{p/q}}(0) = \left(\frac{q}{p}\right) f_{\phi+\pi/2}(0) = 2^{1/2}\,\mathrm{e}^{\mathrm{i}\,\pi/4} \left(\frac{q}{p}\right) \langle f_\phi \rangle. \tag{3.33}$$

We could have also derived the relation for $-p \equiv 3 \mod 4$ from the equation

$$\Theta_f(p/q + \mathrm{i}\,y, \phi) = \overline{\Theta_{\bar{f}}(-p/q + \mathrm{i}\,y, -\phi)}, \tag{3.34}$$

since $p \equiv 1 \mod 4$, which leads back to case 1a.

Case 2 ($q \equiv 2 \mod 4$) is clear, because the cusp is equivalent to the one at $1/2$.

Case 3. In the case when q is odd it is necessary to take $\tilde{\gamma} = [\gamma, \beta_\gamma] = [g\gamma_0, \beta_{g\gamma_0}]$, where $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(4)$. This means that

$$z_{p/q} = \frac{4bz - a}{4dz - c}, \tag{3.35}$$

and thus $q = d$, $4p = c$. (Still $(c,d) = 1$ implies $(p,q) = 1$.) In particular we now have

$$y_{p/q} = \frac{y}{|2dz - c/2|^2} = \frac{y}{4|qz-p|^2}. \tag{3.36}$$

Case 3a. For $q \equiv 1 \mod 4$ we know that $g \in \Gamma_1(4)$. The phase $\beta_{g\gamma_0}$ is then given by

$$\mathrm{e}^{\mathrm{i}\,\beta_{g\gamma_0}(z)/2} = j_g(\gamma_0 z)\,\mathrm{e}^{\mathrm{i}\arg(z)/2} = j_{g\gamma_0}(z)\,\mathrm{e}^{\mathrm{i}\,\pi/4} = \left(\frac{p}{q}\right)\mathrm{e}^{\mathrm{i}(\arg(4pz-q))/2},$$

see (2.16). Since $\chi(\tilde\gamma) = \mathrm{e}^{-\mathrm{i}\,\pi/4}$, we get

$$\chi(\tilde\gamma)^{-1} f_{\phi_{p/q}}(0) = 2^{1/2}\left(\frac{p}{q}\right)\langle f_\phi\rangle. \tag{3.37}$$

Case 3b. In the case $q \equiv 3 \mod 4$ we have $g \in \Gamma_1(4)(-1)$. The phase $\beta_{g\gamma_0}$ is then given by

$$\mathrm{e}^{\mathrm{i}\,\beta_{g\gamma_0}(z)/2} = j_g(\gamma_0 z)\,\mathrm{e}^{\mathrm{i}\,\pi/2}\,\mathrm{e}^{\mathrm{i}\arg(z)/2} = j_{g\gamma_0}(z)\,\mathrm{e}^{\mathrm{i}\,3\pi/4} = \left(\frac{p}{q}\right)\mathrm{e}^{\mathrm{i}(\arg(4pz-q))/2}, \tag{3.38}$$

see (2.16). Since $\chi(\tilde\gamma) = \mathrm{e}^{-\mathrm{i}\,3\pi/4}$, we eventually obtain

$$\chi(\tilde\gamma)^{-1} f_{\phi_{p/q}}(0) = 2^{1/2}\,\mathrm{i}\left(\frac{p}{q}\right)\langle f_\phi\rangle. \tag{3.39}$$

The proof of Corollary 3.3 is now complete. □

3.2. Asymptotic behaviour at generic argument. Limit theorems. We are now going to study the distribution of values of the theta function $\Theta_f(z,\phi)$ in the complex plane where x is fixed and generic, as $y \to 0$. To this end let us take an open set $\mathcal{B} \subset \mathbb{C}$ and define the distribution function

$$D_{f,\mathcal{B}}(z,\phi) = \begin{cases} 1 & \text{if } \Theta_f(z,\phi) \in \mathcal{B} \\ 0 & \text{if } \Theta_f(z,\phi) \notin \mathcal{B}. \end{cases} \tag{3.40}$$

$D_{f,\mathcal{B}}(z,\phi)$ is the characteristic function of the pre-image $\Theta_f^{-1}(\mathcal{B})$ of $\mathcal{B}$ in $\mathcal{M}_1$. We say $\mathcal{B}$ has a "nice" boundary $\partial\mathcal{B}$, if the boundary of $\Theta_f^{-1}(\mathcal{B})$ has measure zero in $\mathcal{M}_1$. A sufficient condition for $\partial\mathcal{B}$ to be nice is that $\partial\mathcal{B}$ is of Lebesgue measure zero and contains no critical values of Θ_f. By Sard's theorem [31], the set of critical values of a smooth map $\mathcal{M}_1 \to \mathbb{C}$ is itself of measure zero in $\mathbb{C}$.

Although the definition of "nice" depends on Θ_f, we indicate by the following example that most sets indeed have a nice boundary. Suppose $\mathcal{B}$

is an open disc of radius R centered at a point $w \in \mathbb{C}$. Then, except for countably many radii R, the boundary of the disc $\mathcal{B}$ will be nice, because there are only countably many R, for which the set of points $(z,\phi) \in \mathcal{M}_1$ satisfying $|\Theta_f(z,\phi) - w| = R$ has positive measure. This follows from the fact that the function $(z,\phi) \mapsto |\Theta_f(z,\phi) - w|$ is measurable.

The following limit theorem holds for generic values of x, and is a direct consequence of the ergodicity of the geodesic flow discussed in Section 2.3. It is proved in [29].

THEOREM 3.1. *Let $\mathcal{B} \subset \mathbb{C}$ be open with nice boundary, $f \in \mathcal{S}(\mathbb{R})$. Then for almost all $x \in \mathbb{R}$ (with respect to Lebesgue measure)*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T D_{f,\mathcal{B}}(x + \mathrm{i}\,\mathrm{e}^{-t}, 0)\, dt = \frac{\mu(\Theta_f^{-1}(\mathcal{B}))}{8\pi^2}.$$

It follows from previous remarks that the statement in Theorem 3.1 does not hold, when x is rational or real quadratic. In the case $x = p/q \in \mathbb{Q}$ certainly a limit only exists if $q \equiv 2 \mod 4$, and the corresponding limit distribution is a delta mass at zero (Corollary 3.3). Since for quadratic irrational x the theta sum $\Theta_f(x + \mathrm{i}\,\mathrm{e}^{-L}, 0)$ is asymptotically periodic, a limit will of course exist, but is different from the one for generic x.

It is clear that we can use the equidistribution of closed horocycles to prove also the following theorem [29].

THEOREM 3.2. *Let $\mathcal{B} \subset \mathbb{C}$ be open with nice boundary, $f \in \mathcal{S}(\mathbb{R})$. Then*

$$\lim_{y\to 0} \int_0^1 D_{f,\mathcal{B}}(x + \mathrm{i}\,y, 0)\, dx = \frac{\mu(\Theta_f^{-1}(\mathcal{B}))}{8\pi^2}.$$

Let us illustrate these results by the distribution of the absolute values of $\Theta_f(z,\phi)$ on the real line, that is we choose $\mathcal{B} = \mathcal{B}_R := \{z : |z| > R\}$. In this case it is easy to obtain an asymptotic expression for $\mu(\Theta_f^{-1}(\mathcal{B}_\mathcal{R}))$, R large. Notice that due to Proposition 3.3 the measure of the set

$$\begin{aligned} \mathcal{D}_R = \{(z,\phi) \in \mathcal{F}_{\Delta_1(4)} \,:\, y_0 > |f_{\phi_0}(0)|^{-4} R^4\} \\ \cup\ \{(z,\phi) \in \mathcal{F}_{\Delta_1(4)} \,:\, y_\infty > |f_{\phi_\infty}(0)|^{-4} R^4\} \end{aligned} \tag{3.41}$$

($\mathcal{F}_{\Delta_1(4)}$ is the fundamental region defined in Section 2.2) converges for $R \to \infty$ rapidly to what we are looking for:

$$\mu(\Theta_f^{-1}(\mathcal{B}_\mathcal{R})) = \mu(\mathcal{D}_R) + O_M(R^{-M}). \tag{3.42}$$

Elementary integration yields (provided R is so large that the sets in (3.41) are disjoint)

$$\mu(\mathcal{D}_R) = 2R^{-4} \int_0^{4\pi} |f_\phi(0)|^4 d\phi, \tag{3.43}$$

where

$$|f_\phi(0)|^4 = 4(\sin\phi)^{-2}\left|\int_{-\infty}^{\infty} e(t^2\cot\phi) f(t)dt\right|^4,$$

compare equation (3.11). After some change of variables we eventually obtain

$$\mu(\mathcal{D}_R) = 32R^{-4}\int_{-\infty}^{\infty}\left|\int_{-\infty}^{\infty} e(t^2 u) f(t)dt\right|^4 du. \tag{3.44}$$

In [29] it is shown that – under the assumption that $f_\phi(0) \neq 0$ for all $\phi \in [0, 4\pi)$ – a curlicue corresponding to a generic value of x becomes densely distributed in the *whole* complex plane. This is *not* a direct consequence of Theorem 3.1, since in general there can be regions in the complex plane which are omitted, although the limit exist. The above condition on f holds for instance, when $f(t)$ is real-valued and monotonically increasing for $t < 0$, and monotonically decreasing for $t > 0$, which is satisfied by our favorite cut-off function (1.20).

PROPOSITION 3.4. *Let $f \in \mathcal{S}(\mathbb{R})$, and assume that $f_\phi(0) \neq 0$ for all $\phi \in [0, 4\pi)$. Then for almost all $x \in \mathbb{R}$ (with respect to Lebesgue measure) the set*

$$\{\Theta_f(x + \mathrm{i}\,\mathrm{e}^{-t}, 0) \,:\, t \in [0, T]\}$$

becomes densely distributed in $\mathbb{C}$ as T tends to infinity.

The same effect can be observed for the curves which are images of closed horocycles.

PROPOSITION 3.5. *Let $f \in \mathcal{S}(\mathbb{R})$, and assume that $f_\phi(0) \neq 0$ for all $\phi \in [0, 4\pi)$. Then the set*

$$\{\Theta_f(x + \mathrm{i}\,y, 0) \,:\, x \in [0, 1]\}$$

becomes densely distributed in $\mathbb{C}$ as y tends to zero.

4. Eisenstein series of half-integral weight. Consider the Eisenstein series

$$E_\infty(z; s, k) = \sum_{\gamma\in\Gamma_\infty\backslash\Gamma_1(4)} [\mathrm{Im}(\gamma z)]^s j_\gamma(z)^{-k}, \qquad \mathrm{Re}\, s > 1, \tag{4.1}$$

of weight $k/2$, where $k \in \mathbb{Z}$. Γ_∞ denotes the subgroup

$$\Gamma_\infty = \left\{\begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \,:\, m \in \mathbb{Z}\right\}, \tag{4.2}$$

which stabilizes the cusp at infinity. Because of the decomposition $\Gamma_0(4) = \Gamma_1(4) \cup \Gamma_1(4)\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ we can also express $E_\infty(z; s, k)$ in the form

$$E_\infty(z; s, k) = \sum_{\gamma\in\pm\Gamma_\infty\backslash\Gamma_0(4)} [\mathrm{Im}(\gamma z)]^s j_\gamma(z)^{-k} \tag{4.3}$$

where

$$\pm\Gamma_\infty = \left\{ \begin{pmatrix} \pm 1 & m \\ 0 & \pm 1 \end{pmatrix} : m \in \mathbb{Z} \right\}. \tag{4.4}$$

It can easily be checked that we have the transformation formula

$$E_\infty(gz; s, k) = j_g(z)^k E_\infty(z; s, k), \qquad g \in \Gamma_0(4). \tag{4.5}$$

In particular, $E_\infty(z; s, k)$ is periodic in x, i.e., $E_\infty(z+1; s, k) = E_\infty(z; s, k)$, and Fourier expansion can be employed to find a meromorphic continuation in s to the whole complex plane, see below. Formula (4.5) means that the function

$$E_\infty(z, \phi; s, k) = E_\infty(z; s, k)\, \mathrm{e}^{-\mathrm{i}\, k\phi/2} \tag{4.6}$$

is invariant under $\Delta_1(4)$, i.e., can be viewed as a function on $\mathcal{M}_1$.

Let us introduce new coordinates

$$(z_0, \phi_0) = (\gamma_0 z, \phi + \arg z) \tag{4.7}$$

relative to the cusp at zero. The old coordinates relative to ∞ can then be expressed by the new ones through

$$(z, \phi) = (\gamma_0^{-1} z, \phi_0 + \arg(z_0) - \pi). \tag{4.8}$$

In these new coordinates the Eisenstein series reads

$$E_\infty(z, \phi; s, k) = E_\infty(\gamma_0^{-1} z_0, \phi_0 + \arg(z_0) - \pi; s, k), \tag{4.9}$$

so

$$E_\infty(z; s, k) = E_\infty(\gamma_0^{-1} z_0; s, k)\, \mathrm{e}^{-\mathrm{i}\, k \arg(z_0)/2}\ \mathrm{e}^{\mathrm{i}\, k\pi/2}. \tag{4.10}$$

We will now show that $E_\infty(z; s, k)$ is invariant under the substitution $z_0 \mapsto z_0 + 1$, and in fact has the same transformation properties in the z_0-coordinates as in the z-coordinates. To this end it is convenient to define

$$\begin{aligned} E_0(z; s, k) &= E_\infty(\gamma_0^{-1} z; s, k)\, \mathrm{e}^{-\mathrm{i}\, k \arg(z)/2} \\ &= E_\infty(\gamma_0^{-1} z; s, k)\, j_{\gamma_0^{-1}}(z)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4}. \end{aligned} \tag{4.11}$$

Then we have for $g \in \Gamma_0(4)$

$$\begin{aligned} E_0(gz; s, k) &= E_\infty(\gamma_0^{-1} g z; s, k)\, j_{\gamma_0^{-1}}(gz)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4} \\ &= E_\infty(\gamma_0^{-1} g \gamma_0\, \gamma_0^{-1} z; s, k)\, j_{\gamma_0^{-1}}(gz)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4} \\ &= E_\infty(\gamma_0^{-1} z; s, k)\, j_{\gamma_0^{-1} g \gamma_0}(\gamma_0^{-1} z)^k\, j_{\gamma_0^{-1}}(gz)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4} \\ &= E_\infty(\gamma_0^{-1} z; s, k)\, j_g(z)^k\, j_{\gamma_0^{-1}}(z)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4}, \end{aligned}$$

in summary:

$$E_0(gz;s,k)=j_g(z)^k E_0(z;s,k),\qquad g\in\Gamma_0(4). \tag{4.12}$$

Let us turn to the cusp at 1/2, and introduce the coordinates

$$(z_{1/2},\phi_{1/2})=(\gamma_{1/2}\,z,\phi+\arg(z-1/2)). \tag{4.13}$$

Inversion of this transformation yields

$$(z,\phi)=(\gamma_{1/2}^{-1}\,z,\phi_{1/2}+\arg(z_{1/2})-\pi). \tag{4.14}$$

In these new coordinates the Eisenstein series reads

$$E_\infty(z,\phi;s,k)=E_\infty(\gamma_{1/2}^{-1}\,z_{1/2},\phi_{1/2}+\arg(z_{1/2})-\pi;s,k), \tag{4.15}$$

hence

$$E_\infty(z;s,k)=E_\infty(\gamma_{1/2}^{-1}\,z_{1/2};s,k)\,\mathrm{e}^{-\mathrm{i}\,k\arg(z_{1/2})/2}\ \mathrm{e}^{\mathrm{i}\,k\pi/2}. \tag{4.16}$$

As before, it is convenient to put

$$\begin{aligned}E_{1/2}(z;s,k)&=E_\infty(\gamma_{1/2}^{-1}z;s,k)\,\mathrm{e}^{-\mathrm{i}\,k\arg(z)/2}\\&=E_\infty(\gamma_{1/2}^{-1}z;s,k)\;j_{\gamma_{1/2}^{-1}}(z)^{-k}\;\frac{j_{\gamma_{1/2}^{-1}}(z)^k}{j_{\gamma_0^{-1}}(z)^k}\ \mathrm{e}^{-\mathrm{i}\,k\pi/4},\end{aligned} \tag{4.17}$$

and to study the transformation behaviour of this function under $\Gamma_0(4)$. One can do the same calculation as before in (4.12) to see that

$$E_\infty(\gamma_{1/2}^{-1}\,gz;s,k)\;j_{\gamma_{1/2}^{-1}}(gz)^{-k}=E_\infty(\gamma_{1/2}^{-1}z;s,k)\;j_{\gamma_{1/2}^{-1}}(z)^{-k}\;j_g(z)^k,$$

and for the quotient in (4.17) we get

$$\frac{j_{\gamma_{1/2}^{-1}}(gz)}{j_{\gamma_0^{-1}}(gz)}=\frac{j_{\gamma_{1/2}^{-1}g\gamma_{1/2}}(\gamma_{1/2}^{-1}z)}{j_{\gamma_0^{-1}g\gamma_0}(\gamma_0^{-1}z)}\;\frac{j_{\gamma_{1/2}^{-1}}(z)}{j_{\gamma_0^{-1}}(z)}=\epsilon_{2b+1}^{-1}\left(\frac{-b}{2b+a}\right)\left(\frac{-b}{a}\right)\;\frac{j_{\gamma_{1/2}^{-1}}(z)}{j_{\gamma_0^{-1}}(z)}$$

where $g=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in\Gamma_0(4)$. All in all

$$E_{1/2}(gz;s,k)=\left[\epsilon_{2b+1}^{-1}\left(\frac{-b}{2ab+a^2}\right)j_g(z)\right]^k E_{1/2}(z;s,k), \tag{4.18}$$

which implies

$$E_{1/2}(z+1;s,k)=\mathrm{i}^k\,E_{1/2}(z;s,k). \tag{4.19}$$

That means the function is periodic in x with period 1, 2 or 4, if the weight is even, odd, or half-integral, respectively.

For the reader's convenience, we summarize some of the formulas discussed so far:

$$\gamma_0 = \begin{pmatrix} 0 & -1/2 \\ 2 & 0 \end{pmatrix}, \qquad \gamma_{1/2} = \begin{pmatrix} 0 & -1/2 \\ 2 & -1 \end{pmatrix}, \tag{4.20}$$

$$\begin{cases} j_g(z) = \epsilon_d^{-1} \left(\frac{c}{d}\right) \mathrm{e}^{\mathrm{i}(\arg(cz+d))/2}, \\ j_{g\gamma_0}(z) = \mathrm{e}^{-\mathrm{i}\,\pi/4}\ \epsilon_d^{-1} \left(\frac{c}{d}\right) \mathrm{e}^{\mathrm{i}(\arg(4dz-c))/2} \qquad (d>0), \\ j_{\gamma_0}(z) = j_{\gamma_0^{-1}}(z) = \mathrm{e}^{-\mathrm{i}\,\pi/4}\ \mathrm{e}^{\mathrm{i}\arg(z)/2}, \end{cases} \tag{4.21}$$

$$\begin{cases} E_0(z;s,k) = E_\infty(\gamma_0^{-1} z;s,k)\,\mathrm{e}^{-\mathrm{i}\,k\arg(z)/2}, \\ E_{1/2}(z;s,k) = E_\infty(\gamma_{1/2}^{-1} z;s,k)\,\mathrm{e}^{-\mathrm{i}\,k\arg(z)/2}, \end{cases} \tag{4.22}$$

$$\begin{cases} E_\infty(z;s,k) = E_0(\gamma_0 z;s,k)\,\mathrm{e}^{-\mathrm{i}\,k(\arg(z)-\pi)/2}, \\ E_\infty(z;s,k) = E_{1/2}(\gamma_{1/2} z;s,k)\,\mathrm{e}^{-\mathrm{i}\,k(\arg(z-1/2)-\pi)/2}, \end{cases} \tag{4.23}$$

$$\begin{cases} E_\infty(gz;s,k) = j_g(z)^k E_\infty(z;s,k), \qquad g\in\Gamma_0(4), \\ E_0(gz;s,k) = j_g(z)^k E_0(z;s,k), \qquad g\in\Gamma_0(4), \\ E_{1/2}(z+1;s,k) = \mathrm{i}^k\, E_{1/2}(z;s,k). \end{cases} \tag{4.24}$$

4.1. Whittaker functions. The Whittaker functions $W_{\kappa,\mu}(y)$ appear in the Fourier expansion of the Eisenstein series, which we shall discuss in the next section. It is defined for $y>0$, $\kappa,\mu\in\mathbb{C}$, $\mathrm{Re}(\mu-\kappa)>-1/2$ by

$$W_{\kappa,\mu}(y) = \frac{\mathrm{e}^{-y/2}\,y^{1/2+\mu}}{\Gamma(1/2+\mu-\kappa)} \int_0^\infty \mathrm{e}^{-yt}\,t^{\mu-\kappa-1/2}(1+t)^{\mu+\kappa-1/2}\,dt, \tag{4.25}$$

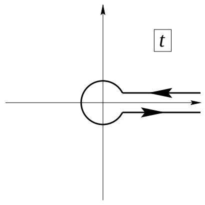
see [28]. For $\mathrm{Re}(\mu-\kappa)\le -1/2$ we define $W_{\kappa,\mu}(y)$ by analytic continuation. The following integral representation of $W_{\kappa,\mu}(y)$, valid for all $(\kappa,\mu)\in\mathbb{C}^2$, may be found in [39],[14]

$$W_{\kappa,\mu}(y) = \Gamma(1/2-\mu+\kappa)\,\mathrm{e}^{-y/2}\,y^{\kappa}\times$$
$$\times\frac{\mathrm{e}^{-\mathrm{i}\,\pi(1/2+\mu-\kappa)}}{2\pi\mathrm{i}} \int_\infty^{(0+)} \mathrm{e}^{-t}\,t^{\mu-\kappa-1/2}(1+y^{-1}t)^{\mu+\kappa-1/2}\,dt, \tag{4.26}$$

where the path of integration is chosen as follows:

[14] The relation between $W_{\kappa,\mu}(y)$ and the function $\sigma(y,\alpha,\beta)$ used there is

$$W_{\kappa,\mu}(y) = \Gamma(1/2+\mu-\kappa)^{-1}\,\mathrm{e}^{-y/2}\,y^{1/2+\mu}\sigma(y,1/2+\mu+\kappa,1/2+\mu-\kappa).$$



The integral is convergent for all $(\kappa, \mu) \in \mathbb{C}^2$, and the Gamma function has no poles in the region $\mathrm{Re}(\mu - \kappa) \leq -1/2$. Since we already know from our initial definition (4.25) that $W_{\kappa,\mu}(y)$ is also regular in the remaining domain $\mathrm{Re}(\mu-\kappa) > -1/2$, the Whittaker function $W_{\kappa,\mu}(y)$ has to be entire in κ and μ.

Whittaker's differential equation reads

$$\left(y^2 \frac{d^2}{dy^2} + \kappa y - y^2/4\right) W_{\kappa,\mu}(y) = (\mu^2 - 1/4) W_{\kappa,\mu}(y). \tag{4.27}$$

For large argument y we have the asymptotic relation

$$W_{\kappa,\mu}(y) \sim y^\kappa \,\mathrm{e}^{-y/2}, \qquad y \to \infty, \tag{4.28}$$

which holds uniformly in $(\kappa, \mu) \in K$, with K any compact subset of $\mathbb{C}^2$.

Lemma 4.1. *Let $\kappa, \mu \in \mathbb{C}$ with $\mathrm{Re}\,\mu > \max\{0, |\,\mathrm{Re}\,\kappa| - 1/2\}$. Then, for $z = x + \mathrm{i}\, y \in \mathfrak{H}$, we have*

$$\sum_{m=-\infty}^{\infty} \frac{1}{(z+m)^{1/2+\mu+\kappa}(\overline{z}+m)^{1/2+\mu-\kappa}} = \sum_{n=-\infty}^{\infty} T_n(y, \mu, \kappa)\, e(nx),$$

with

$$T_n(y,\mu,\kappa) = \begin{cases} \mathrm{i}^{-2\kappa}\ \pi^{1/2+\mu}\ |n|^{\mu-1/2}\ y^{-\mu-1/2}\ \frac{W_{-\kappa,\mu}(4\pi|n|y)}{\Gamma(1/2+\mu-\kappa)} & (n<0) \\ \mathrm{i}^{-2\kappa}\ 2^{-2\mu+1}\ \pi\ y^{-2\mu}\ \frac{\Gamma(2\mu)}{\Gamma(1/2+\mu-\kappa)\Gamma(1/2+\mu+\kappa)} & (n=0) \\ \mathrm{i}^{-2\kappa}\ \pi^{1/2+\mu}\ |n|^{\mu-1/2}\ y^{-\mu-1/2}\ \frac{W_{\kappa,\mu}(4\pi|n|y)}{\Gamma(1/2+\mu+\kappa)} & (n>0). \end{cases}$$

This lemma is proved in [39].

4.2. Fourier expansion of the Eisenstein series. The Eisenstein series $E_j(z; s, k)$, $j = 0, 1/2, \infty$, have a meromorphic continuation to the whole complex s-plane. This may be proved by considering the Fourier expansion of one of them. The Fourier expansion of $E_0(z; s, k)$ turns out to be the easiest; we will calculate it following the lines of Shimura [39]. The Fourier expansions of $E_\infty(z; k, s)$ and $E_{1/2}(z; k, s)$ are closely related and can be found in [41, 13].

By definitions (4.1) and (4.22), we have for $\mathrm{Re}\, s > 1$

$$\begin{aligned} E_0(z; s, k) &= \sum_{\gamma \in \pm\Gamma_\infty \backslash \Gamma_0(4)} [\mathrm{Im}(\gamma\gamma_0^{-1} z)]^s j_\gamma(\gamma_0^{-1} z)^{-k}\ j_{\gamma_0^{-1}}(z)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4} \\ &= \sum_{\gamma \in \pm\Gamma_\infty \backslash \Gamma_0(4)} [\mathrm{Im}(\gamma\gamma_0^{-1} z)]^s j_{\gamma\gamma_0^{-1}}(z)^{-k}\ \mathrm{e}^{-\mathrm{i}\, k\pi/4}. \end{aligned} \tag{4.29}$$

It is well known that if W is a set of representatives of $\pm\Gamma_\infty\backslash\Gamma_0(4)$, then the map

$$W \to W_0, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \{c, d\}$$

to the set W_0 of ordered pairs of integers satisfying $c \equiv 0 \mod 4$, $d > 0$, and $(c,d) = 1$, is one-to-one, which means that we can write (4.29) as (put $b = -c/4$)

$$E_0(z; s, k) = y^s 2^{-2s} \sum_{\substack{d=1 \\ d \text{ odd}}}^{\infty} \sum_{b=-\infty}^{\infty} \epsilon_d^k \left(\frac{-b}{d}\right)^k \frac{1}{|dz+b|^{2s-k/2}(dz+b)^{k/2}}. \tag{4.30}$$

Set furthermore $b = ld + m$ with $l \in \mathbb{Z}$ and $1 \le m \le d$, so

$$E_0(z; s, k) = y^s 2^{-2s} \sum_{\substack{d=1 \\ d \text{ odd}}}^{\infty} d^{-2s} \epsilon_d^k \left(\frac{-1}{d}\right)^k \sum_{m=1}^{d} \left(\frac{m}{d}\right)^k \times$$
$$\times \sum_{l=-\infty}^{\infty} \frac{1}{|z + \frac{m}{d} + l|^{2s-k/2}(z + \frac{m}{d} + l)^{k/2}}. \tag{4.31}$$

Since $|z + \frac{m}{d} + l|^{2s-k/2}(z + \frac{m}{d} + l)^{k/2} = (z + \frac{m}{d} + l)^{s+k/4}(\overline{z} + \frac{m}{d} + l)^{s-k/4}$ we can apply Lemma 4.1, yielding

$$E_0(z; s, k) = \sum_{n=-\infty}^{\infty} \alpha_n(y; s, k)\, e(nx) \tag{4.32}$$

with the Fourier coefficients

$$\alpha_n(y; s, k) = 2^{-2s} \mathrm{e}^{-\mathrm{i}\, k\pi/4} \sum_{\substack{d=1 \\ d \text{ odd}}}^{\infty} d^{-2s} \epsilon_d^{-k} \sum_{m=1}^{d} \left(\frac{m}{d}\right)^k e(nm/d) \times$$
$$\times \begin{cases} \pi^s\, |n|^{s-1} \frac{W_{-k/4,\, s-1/2}(4\pi|n|y)}{\Gamma(s-k/4)} & (n < 0) \\ \pi^{1/2} \frac{\Gamma(s)\Gamma(s-1/2)}{\Gamma(s-k/4)\Gamma(s+k/4)}\, y^{1-s} & (n = 0) \\ \pi^s\, |n|^{s-1} \frac{W_{k/4,\, s-1/2}(4\pi|n|y)}{\Gamma(s+k/4)} & (n > 0). \end{cases} \tag{4.33}$$

Here we have used the relation $\epsilon_d \left(\frac{-1}{d}\right) = \epsilon_d^{-1}$.

In the case $n = 0$ the remaining sum in (4.33) can be written as a quotient of zeta or L-functions,

$$\sum_{\substack{d=1 \\ d \text{ odd}}}^{\infty} d^{-2s} \epsilon_d^{-k} \sum_{m=1}^{d} \left(\frac{m}{d}\right)^k = \begin{cases} \frac{\zeta(2s-1)}{\zeta(2s)} \frac{1-2^{-(2s-1)}}{1-2^{-2s}} & \text{if } k/4 \in \mathbb{Z} \\ \frac{L(2s-1,\chi_4)}{L(2s,\chi_4)} & \text{if } (k-2)/4 \in \mathbb{Z} \\ \frac{\zeta(4s-2)}{\zeta(4s-1)} \frac{1-2^{-(4s-2)}}{1-2^{-(4s-1)}} & \text{if } (k-1)/2 \in \mathbb{Z} \end{cases}$$

compare [18, 39, 13]. The three cases correspond to even integral, odd integral, and half-integral weight. $L(s,\chi_4)$ is the Dirichlet L-function with real primitive character mod 4, i.e., for $\operatorname{Re} s > 0$,

$$L(s,\chi_4) = \sum_{n=1}^{\infty} \frac{\chi_4(n)}{n^s}, \tag{4.34}$$

$$\chi_4(n) = \begin{cases} 1 & \text{if } n \equiv 1 \mod 4 \\ -1 & \text{if } n \equiv 3 \mod 4 \\ 0 & \text{if } n \equiv 0 \mod 2. \end{cases} \tag{4.35}$$

In all other cases $n \neq 0$ one obtains

$$\sum_{\substack{d=1\\ d \text{ odd}}}^{\infty} d^{-2s} \epsilon_d^{-k} \sum_{m=1}^{d} \left(\frac{m}{d}\right)^k e(nm/d) =$$

$$= \begin{cases} \dfrac{1}{\zeta(2s)(1-2^{-2s})} \displaystyle\sum_{\substack{d>0 \text{ odd}\\ d|n}} d^{1-2s} & \text{if } \frac{k}{4} \in \mathbb{Z} \\ \dfrac{1}{L(2s,\chi_4)} \displaystyle\sum_{\substack{d>0 \text{ odd}\\ d|n}} \chi_4(d) d^{1-2s} & \text{if } \frac{k-2}{4} \in \mathbb{Z} \\ \dfrac{L(2s-1/2,\omega)}{\zeta(4s-1)(1-2^{-(4s-1)})} \displaystyle\sum_{\substack{a,b>0\\ \text{odd}\\ ab|u}} \mu(a)\omega(a) a^{1/2-2s} b^{2-4s} & \text{if } \frac{k-1}{2} \in \mathbb{Z}. \end{cases} \tag{4.36}$$

The positive integer u is chosen in a way such that $n = tu^2$, where $t \in \mathbb{Z}^*$ is square-free. μ is the Möbius function, and ω the character mod $4t$ defined by $\omega(a) = 0$ for a even, and

$$\omega(a) = \begin{cases} \left(\frac{t}{a}\right) & \text{if } (k-1)/2 \text{ is even} \\ \left(\frac{-t}{a}\right) & \text{if } (k-1)/2 \text{ is odd} \end{cases} \tag{4.37}$$

for a odd.

4.3. Analytic continuation of the Eisenstein series. It is clear from its Fourier expansion that $E_0(z;k,s)$ has a meromorphic continuation to the whole s-plane. By virtue of (4.22), the other two Eisenstein series $E_{1/2}(z;k,s)$ and $E_\infty(z;k,s)$ have the same analytic properties, in particular the location of their poles coincides.

Consequently, we extend the definition (4.1), (4.22) of $E_j(z;k,s)$, valid initially only for $\operatorname{Re} s > 1$, to the whole complex s-plane by analytic continuation.

We shall be especially interested in the behaviour in the half plane $\operatorname{Re} s > 1/2$. If the (generalized) Riemann hypothesis for $\zeta(s)$ and $L(s,\chi_4)$ is true, the following results can be extended to any half plane $\operatorname{Re} s > 1/4+\epsilon$, with $\epsilon > 0$.

PROPOSITION 4.1. *The Eisenstein series $E_j(z;k,s)$, $j = 0, 1/2, \infty$, are holomorphic for $\operatorname{Re} s > 1/2$, up to a simple pole at $s = 1$ in the case $k = 0$, and a simple pole at $s = 3/4$ in the case when $(|k|-1)/2$ is even.*

The residues of the poles are given by

$$\operatorname{Res} E_j(z; s=1, 0) = \frac{1}{2\pi},$$

and for $\nu = (|k|-1)/2$ even,

$$\operatorname{Res} E_0(z; s=\frac{3}{4}, k) = \begin{cases} \frac{e^{-i\pi/4}}{4\pi}\,(-1)^{\nu/2}\, h_\nu(0)\, \theta_\nu(z) & (k>0) \\ \frac{e^{i\pi/4}}{4\pi}\,(-1)^{\nu/2}\, h_\nu(0)\, \overline{\theta_\nu(z)} & (k<0) \end{cases}$$

$$\operatorname{Res} E_{1/2}(z; s=\frac{3}{4}, k) = \begin{cases} \frac{1}{4\pi}\, h_\nu(0)\, \theta_\nu(-1/4z-1/2)\, e^{-i\,k\arg(z)/2} & (k>0) \\ \frac{1}{4\pi}\, h_\nu(0)\, \overline{\theta_\nu(-1/4z-1/2)}\, e^{-i\,k\arg(z)/2} & (k<0) \end{cases}$$

$$\operatorname{Res} E_\infty(z; s=\frac{3}{4}, k) = \begin{cases} \frac{1}{4\pi}\, h_\nu(0)\, \theta_\nu(z) & (k>0) \\ \frac{1}{4\pi}\, h_\nu(0)\, \overline{\theta_\nu(z)} & (k<0) \end{cases}$$

where $\theta_\nu(z)$ denotes the theta series defined in Section 3.

Proof. First notice that $\zeta(2s)$, $L(2s,\chi_4)$, and $\zeta(4s-1)$ are zero-free for $\operatorname{Re} s > 1/2$.

In the case $k/4 \in \mathbb{Z}$, $\zeta(2s-1)$ has a simple pole at $s=1$, which is canceled by the pole of $\Gamma(s-k/4)\Gamma(s+k/4)$ if $k \neq 0$. The residue of the pole in the case $k=0$ can be readily calculated, it is the inverse of the volume of $\Gamma_0(4)\backslash\mathfrak{H}$, thus a constant.

For the case $(k-2)/4 \in \mathbb{Z}$, it suffices to note that $L(2s-1,\chi_4)$ is regular.

Let us suppose next that $(k-1)/2$ is odd integral, and $k>0$. Then the pole of $\zeta(4s-2)$ and the possible simple pole of $L(2s-1/2,\omega)$ at $s=3/4$ are canceled by the factor $\Gamma(s-k/4)$.

If $(k-1)/2$ is even integral, and $k>0$, we have a simple pole at $s=3/4$ coming from $\zeta(4s-2)$ and $L(2s-1/2,\omega)$, whenever ω is a trivial character. This happens when n is a square; let us put $n=u^2$. The residue of $E_0(z;s,k)$ at $s=3/4$ can be worked out as follows. Consider the u^2-th Fourier coefficient,

$$\alpha_{u^2}(y;s,k) = 2^{-2s}\, e^{-i\,k\pi/4}\, \frac{\zeta(2s-1/2)}{\zeta(4s-1)}\, \frac{1-2^{-(2s-1/2)}}{1-2^{-(4s-1)}} \times \\ \times \Bigg(\sum_{\substack{a,b>0 \text{ odd} \\ ab|u}} \mu(a) a^{1/2-2s} b^{2-4s} \Bigg) \pi^s\, u^{2s-2}\, \frac{W_{k/4,s-1/2}(4\pi|n|y)}{\Gamma(s+k/4)}. \tag{4.38}$$

For $s=3/4$ the sum in the brackets equals

$$\sum_{\substack{a,\,b\,>\,0\ \mathrm{odd}\\ ab|u}} \frac{\mu(a)}{a}\,\frac{1}{b} = 1, \tag{4.39}$$

and the residue of $\alpha_{u^2}(y;s,k)$ at $s=3/4$ is therefore

$$\operatorname{Res}\alpha_{u^2}(y;s=3/4,k) = \frac{\mathrm{e}^{-\mathrm{i}\,k\pi/4}}{\pi}\,(4\pi u^2 y)^{-1/4}\,\frac{W_{k/4,1/4}(4\pi u^2 y)}{\Gamma(\frac{3+k}{4})}. \tag{4.40}$$

It is convenient to substitute $\nu=(k-1)/2$.

The Whittaker function $W_{\nu/2+1/4,1/4}(y)$ is related to the Hermite polynomials via [28], pp. 324-5,

$$H_\nu(z) = 2^\nu z^{-1/2}\,\mathrm{e}^{z^2/2}\,W_{\nu/2+1/4,1/4}(z^2), \tag{4.41}$$

which leads us to

$$\begin{aligned}\operatorname{Res}\alpha_{u^2}(y;s=3/4,2\nu+1) &= \frac{\mathrm{e}^{-\mathrm{i}\,\pi/4}}{\pi}\,\frac{y^{1/4}}{2^\nu\,\frac{\nu}{2}!}\,H_\nu(2\pi^{1/2}\,u\,y^{1/2})\,\mathrm{e}^{-2\pi u^2 y}\\ &= \frac{\mathrm{e}^{-\mathrm{i}\,\pi/4}}{\pi}\,y^{1/4}\,\frac{1}{2}(-1)^{\nu/2}h_\nu(0)\,h_\nu(u\,y^{1/2})\end{aligned} \tag{4.42}$$

since

$$h_\nu(t) = (2^{\nu-1}\nu!)^{-1/2}\,H_\nu(2\pi^{1/2}\,t)\,\mathrm{e}^{-2\pi t^2}$$

and

$$h_\nu(0) = (2^{\nu-1}\nu!)^{-1/2}\,(-1)^{\nu/2}\,\frac{\nu!}{\frac{\nu}{2}!} \qquad (\nu \text{ even}).$$

The residue of the zeroth coefficient may be expressed as

$$\operatorname{Res}\alpha_0(y;s=3/4,2\nu+1) = \frac{\mathrm{e}^{-\mathrm{i}\,\pi/4}}{2\pi}\,y^{1/4}\,\frac{1}{2}(-1)^{\nu/2}h_\nu(0)^2. \tag{4.43}$$

The residues of the other two Eisenstein series are obtained via the transformation formulas (4.22), (4.23) and (3.13); the results for negative k follow from

$$E_j(z;s,-k) = \overline{E_j(z;\overline{s},k)}. \tag{4.44}$$

□

LEMMA 4.2. *The Eisenstein series $E_j(z;s,k)$ are eigenfunctions of the Laplacian*

$$\Delta_{k/2} = y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right) - \mathrm{i}\,y\,\frac{k}{2}\,\frac{\partial}{\partial x},$$

with eigenvalue $s(s-1)$, *i.e.,*

$$\Delta_{k/2} E_j(z;s,k) = s(s-1) E_j(z;s,k).$$

Proof. Apply Whittaker's differential equation (4.27). $\square$

REMARK. Since the theta functions $\theta_\nu(z)$ are residues of the Eisenstein series, they are also eigenfunctions of the Laplacian $\Delta_{k/2}$ with $s = 3/4$, i.e.,

$$\Delta_{(2\nu+1)/2}\, \theta_\nu(z) = -\frac{3}{16}\, \theta_\nu(z). \tag{4.45}$$

Therefore the theta functions $\Theta_f(z,\phi)$ are eigenfunctions of the Casimir operator of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$,

$$\mathrm{C} = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + y\, \frac{\partial^2}{\partial x\, \partial \phi}, \tag{4.46}$$

that is, $\mathrm{C}\ \Theta_f = -\frac{3}{16}\, \Theta_f$.

5. An equidistribution theorem. Let $\nu(y)$ be the probability measure that puts uniform mass on the closed horocycle at height y, i.e.,

$$\langle F, \nu(y) \rangle = \int_0^1 F(z,0) dx. \tag{5.1}$$

In the following we assume for simplicity that $F \in \mathrm{C}_{00}^\infty(\mathcal{M}_1)$, i.e., F is infinitely differentiable and has compact support.

THEOREM 5.1. *Let* $F \in \mathrm{C}_{00}^\infty(\mathcal{M}_1)$. *Then, for* $y \to 0$, *one has*

$$\langle F, \nu(y) \rangle = \langle F, \mu \rangle + \langle F, \mu_\theta \rangle\, y^{1/4} + o(y^{1/2}),$$

with

$$\langle F, \mu \rangle = \frac{1}{8\pi^2} \int_{\mathcal{M}_1} F(z,\phi)\, d\mu(z,\phi),$$

$$\langle F, \mu_\theta \rangle = \frac{1}{8\pi^2} \sum_{\nu=0}^{\infty} h_{2\nu}(0) \int_{\mathcal{M}_1} F(z,\phi)\, \mathrm{Re}\{\theta_{2\nu}(z,\phi)\} d\mu(z,\phi).$$

Proof. The equidistribution of horocycles on $\Gamma \backslash \mathrm{PSL}(2,\mathbb{R})$ is proved by Sarnak [33], and with our knowledge of the analytic properties of Eisenstein series of half-integral weight (Section 4.3) it is not hard to generalize his proof;[15] we will just give the crucial steps.

[15]We avoid difficulties, since here we are only dealing with such arithmetic groups, whose Eisenstein series have analytic properties that are related to the well known zeta or L-function. In particular, this is the reason why only y^0 and $y^{1/4}$ terms appear in the asymptotic expansion of $\langle F, \nu(y) \rangle$. The proof of Sarnak uses Selberg's theory for forms of even integral weight, and is therefore not restricted to arithmetic congruence groups.

For $\operatorname{Re} s > 1$ we define new measures[16]

$$E(s) = \int_0^\infty \nu(y)\, y^{s-2} dy, \tag{5.2}$$

we then have for $\operatorname{Re} s > 1$

$$\langle F, E(s)\rangle = \int_0^\infty \int_0^1 F(z,0)\, \overline{y^s}\, \frac{dx\, dy}{y^2}. \tag{5.3}$$

Consider the Fourier expansion for F with respect to ϕ,

$$F(z,\phi) = \sum_{k=-\infty}^{\infty} \tilde{F}(z,k)\, \mathrm{e}^{-\mathrm{i}\, k\phi/2}, \quad \tilde{F}(z,k) = \frac{1}{4\pi} \int_0^{4\pi} F(z,\phi)\, \mathrm{e}^{\mathrm{i}\, k\phi/2}\, d\phi, \tag{5.4}$$

the latter being a form of weight $k/2$, i.e., for $\gamma \in \Gamma_0(4)$, $\tilde{F}(\gamma z, k) = j_\gamma(z)^k \tilde{F}(z,k)$. After substitution of variables it becomes clear that we can write

$$\langle F, E(s)\rangle = \sum_{k=-\infty}^{\infty} \int_{\mathcal{F}} \tilde{F}(z,k)\, \overline{E_\infty(z;s,k)}\, \frac{dx\, dy}{y^2}, \tag{5.5}$$

where $\mathcal{F}$ is the fundamental region (2.3) of $\Gamma_0(4)$ in $\mathfrak{H}$, and $E_\infty(z;s,k)$ the Eisenstein series of weight $k/2$.

By Lemma 4.2 we know that

$$\begin{aligned} &\int_{\mathcal{F}} \tilde{F}(z,k)\, \overline{E_\infty(z;s,k)}\, \frac{dx\, dy}{y^2} \\ &\qquad = (\overline{s}(\overline{s}-1))^{-N} \int_{\mathcal{F}} [\Delta_{k/2}^N \tilde{F}(z,k)]\, \overline{E_\infty(z;s,k)}\, \frac{dx\, dy}{y^2} \end{aligned} \tag{5.6}$$

for any N. Using well known estimates [43] for $\zeta(s)$, $L(s,\chi)$ and their inverses $1/\zeta(s)$, $1/L(s,\chi)$, we see that $\langle F, E(s)\rangle$ is rapidly decreasing for $t = \operatorname{Im} s \to \pm\infty$, $\sigma = \operatorname{Re} s$ fixed.

Applying the Mellin inversion formula to (5.3) yields therefore

$$\begin{aligned} \langle F, \nu(y)\rangle &= \frac{1}{2\pi\mathrm{i}} \int_{2-\mathrm{i}\,\infty}^{2+\mathrm{i}\,\infty} \langle F, E(s+1)\rangle \overline{y^{-s}} d\overline{s} \\ &= \operatorname{Res}\langle F, E(s=1)\rangle + \operatorname{Res}\langle F, E(s=3/4)\rangle\, y^{1/4} \\ &\qquad + \int_{-1/2-\mathrm{i}\,\infty}^{-1/2+\mathrm{i}\,\infty} \langle F, E(s+1)\rangle \overline{y^{-s}} d\overline{s}. \end{aligned} \tag{5.7}$$

The last term is of order $o(y^{1/2})$, and the residues at the poles at 1 and 3/4 are given by Proposition 4.1. □

[16] For the proof that these objects are in fact measures, compare [33], Proposition 2.3.

Let us see what happens, if we put $F=\Theta_f$, ignoring the fact that this function does not have compact support. In this case, we a priori know the answer, $\langle \Theta_f, \nu(y)\rangle = y^{1/4} f(0)$, and it is interesting to check whether it coincides with the prediction of our theorem. Indeed, we have $\langle \Theta_f,\mu\rangle = 0$ and

$$\begin{aligned}\langle \Theta_f, \mu_\theta\rangle &= \frac{1}{8\pi^2}\sum_{\nu=0}^\infty h_{2\nu}(0)\int_{\mathcal{M}_1} \Theta_f(z,\phi)\,\mathrm{Re}\{\theta_{2\nu}(z,\phi)\} d\mu(z,\phi)\\ &= \frac{1}{16\pi^2}\sum_{\nu=0}^\infty \langle \Theta_f,\theta_{2\nu}\rangle\, h_{2\nu}(0) = \sum_{\nu=0}^\infty \hat{f}(\nu)\, h_\nu(0) = f(0), \qquad (5.8)\end{aligned}$$

compare (3.16).

REFERENCES

[1] A. Abragam, *The Principles of Nuclear Magnetism* (Oxford Univ. Press, 1961).

[2] E. Artin, Ein mechanisches System mit quasi-ergodischen Bahnen, *Abh. Math. Sem. d. Hamburgischen Univ.* **3** (1924) 170-175.

[3] N. Balazs and A. Voros, Chaos on the pseudosphere, *Phys. Rep.* **143** (1986) 109-240.

[4] M.V. Berry and J. Goldberg, Renormalisation of curlicues, *Nonlinearity* **1** (1988) 1-26.

[5] M.V. Berry, Random renormalisation in the semiclassical long-time limit of a precessing spin, *Physica* D **33** (1988) 26-33.

[6] M.V. Berry, Some quantum-to-classical asymptotics, in: *Chaos and Quantum Physics*, Les Houches 1989, M.-J. Giannoni, A. Voros, and J. Zinn-Justin, eds. (North-Holland, 1991) 251-303.

[7] M.V. Berry and S. Klein, Integer, fractional and fractal Talbot effects, *J. Modern Opt.* **43** (1996) 2139-2164.

[8] C.-H. Chang and D.H. Mayer, The transfer operator approach to Selberg's zeta function and modular and Maass wave forms for $\mathrm{PSL}(2,\mathbb{Z})$, these proceedings.

[9] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic Theory* (Springer, 1982).

[10] I. Efrat, Dynamics of the continued fraction map and the spectral theory of $\mathrm{SL}(2,\mathbb{Z})$, *Invent. Math.* **114** (1993) 207-218.

[11] A. Eskin and C. McMullen, Mixing, counting, and equidistribution in Lie groups, *Duke Math. J.* **71** (1993) 181-209.

[12] H. Fiedler, W. Jurkat and O. Körner, Asymptotic expansions of finite theta series, *Acta Arithm.* **32** (1977) 129-146.

[13] D. Goldfeld and J. Hoffstein, Eisenstein series of $\frac{1}{2}$-integral weight and the mean value of Dirichlet L-series, *Invent. Math.* **80** (1985) 185-208.

[14] M.C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer, 1990).

[15] G.H. Hardy and J.E. Littlewood, Some problems in diophantine approximation. II, *Acta Math.* **37** (1914) 193-239.

[16] G.H. Hardy and J.E. Littlewood, A further note on the trigonometrical series associated with the elliptic ϑ-functions, *Proc. Camb. Phil. Soc.* **21** (1923) 1-5.

[17] D. Hejhal, *The Selberg Trace Formula for* $\mathrm{PSL}(2,\mathbb{R})$, Vol. 1, Lect. Notes in Math. 548 (Springer, 1976).

[18] D. Hejhal, *The Selberg Trace Formula for* $\mathrm{PSL}(2,\mathbb{R})$, Vol. 2, Lect. Notes in Math. 1001 (Springer, 1983).

[19] D. Hejhal, On value distribution properties of automorphic functions along closed horocycles, *XVIth Rolf Nevanlinna Colloquium (Joensuu, 1995)*, 39-52 (de Gruyter, 1996).

[20] E. Hopf, Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung, *Ber. Verh. Sächs. Akad. Wiss. Leipzig* **91** (1939) 261-304.

[21] W.B. Jurkat and J.W. van Horne, The proof of the central limit theorem for theta sums, *Duke Math. J.* **48** (1981) 873-885.

[22] W.B. Jurkat and J.W. van Horne, On the central limit theorem for theta series, *Michigan Math. J.* **29** (1982) 65-77.

[23] W.B. Jurkat and J.W. van Horne, The uniform central limit theorem for theta sums, *Duke Math. J.* **50** (1983) 649-666.

[24] L. Kapitanski and I. Rodnianski, Does a quantum particle know the time?, these proceedings.

[25] L.D. Landau and E.M. Lifschitz, *Lehrbuch der Theoretischen Physik*, Vol. III: *Quantenmechanik*, 5. Aufl. (Akademie-Verlag, 1974).

[26] C. Leichtle, I.Sh. Averbukh, and W.P. Schleich, Multilevel quantum beats: An analytical approach, *Phys. Rev.* A **54** (1996) 5299-5312.

[27] G. Lion and M. Vergne, *The Weil Representation, Maslov Index and Theta Series*, Progr. in Math. 6 (Birkhäuser, 1980).

[28] W. Magnus, F. Oberhettinger and R.P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, 1966).

[29] J. Marklof, Limit theorems for theta sums, *Duke Math. J.*, to appear.

[30] J. Marklof, *Limit Theorems for Theta Sums with Applications in Quantum Mechanics*, Dissertation, Universität Ulm, 1997 (Shaker Verlag, Aachen, 1997).

[31] J. Milnor, *Topology from the Differentiable Viewpoint* (The Univ. Press of Virginia, 1972)

[32] R.R. Roy and B.P. Nigam, *Nuclear Physics* (John Wiley & Sons, 1967).

[33] P. Sarnak, Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series, *Comm. Pure Appl. Math.* **34** (1981) 719-739.

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

[35] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric spaces with applications to Dirichlet series, *J. Indian. Math. Soc.* **20** (1956) 49-87.

[36] C. Series, The modular surface and continued fractions, *J. London. Math. Soc.* **31** (1985) 69-80.

[37] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions* (Princeton Univ. Press, 1971).

[38] G. Shimura, On modular forms of half integral weight, *Ann. of Math.* **97** (1973) 440-481.

[39] G. Shimura, On the holomorphy of certain Dirichlet series, *Proc. London Math. Soc.* **31** (1975) 79-98.

[40] G. Shimura, On the transformation formulas of theta series, *Am. J. Math.* **115** (1993) 1011-1052.

[41] J. Sturm, Special values of zeta functions, and Eisenstein series of half integral weight, *Amer. J. Math.* **102** (1980) 219-240.

[42] S. Thangavelu, *Lectures on Hermite and Laguerre Expansions*, Math. Notes 42 (Princeton Univ. Press, 1993).

[43] E.C. Titchmarsh, *The Theory of the Riemann Zeta-Function*, 2nd ed., revised by D.R. Heath-Brown (Oxford Science Publications, 1986).

[44] I.M. Vinogradov, *Trigonometrical Sums in Number Theory* (Statist. Publ. Soc., 1975)