QUANTUM ERGODICITY ON HYPERBOLIC SURFACES

ALEX EASTMAN

ABSTRACT. We will survey the dynamics of hyperbolic surfaces with a main focus on quantum chaos and the quantum ergodic theorem. The reason for restricting our study of quantum chaos to hyperbolic surfaces is so that we can explicitly define many of the interesting objects. Many of the results in quantum chaos either generalize or are conjectured to generalize to negatively curved surfaces. We will begin with a review of the hyperbolic plane, in particular the upper half plane model. We will then cover geodesic flow on quotients by lattices, finishing the classical portion with a discussion of the ergodicity of the geodesic flow. Before covering the technical details of the quantum ergodic theorem we will devote a section to explaining the intuition behind the theorem. We will conclude by proving the theorem.

CONTENTS

1. Review of the Hyperbolic Plane

The hyperbolic plane can be modeled by the upper half plane with the hyperbolic metric. Orientation preserving isometries under this metric are Möbius transformations.

Definition 1.1. The upper half plane is denoted by $\mathbb{H}$.

$$\mathbb{H} = \{z = x + iy \in \mathbb{C} \mid y > 0\}$$

We will sometimes refer to the imaginary part of complex number z by y and the real part by x.

Definition 1.2. The tangent bundle is given by $T\mathbb{H} = \mathbb{H} \times \mathbb{C}$. The hyperbolic Riemannian metric for two vectors v, w tangent at a point $z \in \mathbb{H}$ is given by

$$\langle v, w \rangle_z = \frac{v\bar{w}}{\text{Im}(z)^2}$$

Date: October 3, 2020.

which is the typical inner product on $\mathbb{C}$ scaled by the imaginary component. The norm is induced in the usual way by taking the square root of the metric.

$$||v||_z^2 = \langle v, v\rangle_z$$

Definition 1.3. A path, sometimes called a curve, is a function $\phi : [0,1] \to \mathbb{H}$ which is continuous and piecewise differentiable. We can define the length of the path as

$$L(\phi) = \int_0^1 ||\phi'(t)||_{\phi(t)} dt.$$

Here $\phi'(t) \in T\mathbb{H}$ is the derivative of the path. We can define the distance between two separate points $a, b \in \mathbb{H}$ by taking the infimum of the lengths of all paths connecting them. If ϕ is a path such that $\phi(0) = a$ and $\phi(1) = b$ then

$$d(a,b) = \inf_{\phi} L(\phi).$$

A curve that obtains this minimum is called a geodesic.

Definition 1.4. $SL_2(\mathbb{R})$ is the group of real valued 2×2 matrices of determinant 1. There is an action of $SL_2(\mathbb{R})$ on $\mathbb{H}$ via the Möbius transform. For $z \in \mathbb{H}$ and $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{R})$ we define the action by

$$gz = \frac{az+b}{cz+d}.$$

The action of $-I_2$, where I_2 is the 2×2 identity matrix, is trivial. Thus, we can write the Möbius transformation as an action of $PSL_2(\mathbb{R}) = SL_2(\mathbb{R})/\{\pm I_2\}$. To prove this transformation is an action, we need to show that the imaginary part is still positive. Writing $z = x + iy$ we have

$$\begin{aligned}
\mathrm{Im}(gz) &= \mathrm{Im}\left(\frac{az+b}{cz+d}\right) \\
&= \mathrm{Im}\left(\frac{(ax+b)+(ay)i}{(cx+d)+(cy)i}\right) \\
&= \frac{\mathrm{Im}\left([(ax+b)+(ay)i][(cx+d)-(cy)i]\right)}{|cz+d|^2} \\
&= \frac{(ad-bc)y}{|cz+d|^2}.
\end{aligned}$$

Since the determinant of g is 1, we obtain the useful formula

$$\mathrm{Im}(gz) = \frac{\mathrm{Im}(z)}{|cz+d|^2}. \tag{1.5}$$

Note that $cz + d$ is never 0, otherwise z would lie along the real axis.

Lemma 1.6. *The action of $PSL_2(\mathbb{R})$ on $\mathbb{H}$ is transitive.*

Proof. We will prove this by finding a matrix $g \in PSL_2(\mathbb{R})$ that sends i to any $x + yi \in \mathbb{H}$. We will find this matrix in two steps; first we will find the matrix that scales by an amount y and then a matrix which shifts by x. The matrix scaling by y is

$$\begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} i = \frac{y^{1/2}i}{y^{-1/2}} = yi.$$

If we compose this transformation with a shift transformation, we obtain

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} yi = \frac{yi + x}{1} = x + yi.$$

Thus, the composition sends $i \to x + yi$

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} = \begin{pmatrix} y^{1/2} & xy^{-1/2} \\ 0 & y^{-1/2} \end{pmatrix}$$

□

If we find the stabilizer of a point in $\mathbb{H}$, then we can construct an isometry by taking the quotient with respect to the stabilizer. We will now find the stabilizer of i.

Lemma 1.7.

$$STAB_{PSL_2(\mathbb{R})}(i) = PSO_2(\mathbb{R})$$

Proof. Letting $g(i) = i$ and applying (1.5) yields $|ci + d| = 1$. Since c and d are real, we have $c^2 + d^2 = 1$. Therefore, there exists some θ such that $c = \sin(\theta)$ and $d = \cos(\theta)$. This observation reduces the problem to solving the following equation:

$$\begin{aligned} i &= \frac{b + ai}{\cos(\theta) + \sin(\theta)i} \\ &= \frac{b + ai}{\cos(\theta) + \sin(\theta)i} \cdot \frac{\cos(\theta) - \sin(\theta)i}{\cos(\theta) - \sin(\theta)i} \\ &= (b\cos(\theta) + a\sin(\theta)) + (a\cos(\theta) - b\sin(\theta))i \end{aligned}$$

The only solution is $a = \cos(\theta)$ and $b = -\sin(\theta)$. This solution shows that elements of the stabilizer are

$$g = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

which, as a subset of $PSL_2(\mathbb{R})$, is precisely $PSO_2(\mathbb{R})$. Equivalently, we have shown that the stabilizer under the action of $SL_2(\mathbb{R})$ is $SO_2(\mathbb{R})$. □

Thus, we obtain the isometry

$$\mathbb{H} \cong PSL_2(\mathbb{R})/PSO_2(\mathbb{R}) \cong SL_2(\mathbb{R})/SO_2(\mathbb{R}).$$

Theorem 1.8. *Each element of $PSL_2(\mathbb{R})$ acts on $\mathbb{H}$ by an isometry with respect to the distance defined in Definition 1.3.*

Proof. We need to show that $L(g \circ \phi) = L(\phi)$ for any path ϕ and $g \in PSL_2(\mathbb{R})$. This follows from showing that the integrand in Definition 1.3 does not change under the action of g. The differential of g, viewed as a map from $\mathbb{H} \to \mathbb{H}$, is given by

$$g'(z) = \frac{1}{(cz + d)^2}.$$

Using (1.5) gives

$$||g'(v)||_{g(z)} = \frac{|v|}{|cz+d|^2 \operatorname{Im}(g(z))} = ||v||_z.$$

This last equation states that the lengths of tangent vectors are unchanged when the action of g is extended to the tangent space. Therefore, the integrand is unchanged and we have an isometry. □

Definition 1.9. The unit tangent bundle is given by the vectors of length 1.

$$T^1\mathbb{H} = (z,v) \in T\mathbb{H} : |v| = 1$$

Since the action, when extended to the tangent bundle, does not change the length of the tangent vectors, $g'(z)$ gives an action on the unit tangent bundle. Explicitly, the action of g on a unit tangent vector (z,v) is

$$g \cdot (z,v) \to (g \cdot z, g'(z)) = \left(\frac{az+b}{cz+d}, \frac{v}{(cz+d)^2}\right).$$

Lemma 1.10. *The action of $PSL_2(\mathbb{R})$ on $T^1\mathbb{H}$ is simply transitive.*

Proof. Recall from Lemma 1.6 that the action on $\mathbb{H}$ is transitive. Thus, it suffices to consider unit tangent vectors (i,v). We can then look at the action of the stabilizer of i. The action of an element of $PSO_2(\mathbb{R})$ on the tangent vector v is

$$g'(i)v = \frac{v}{(i\sin(\theta)+\cos(\theta))^2} = (\cos(2\theta) - i\sin(2\theta))v$$

The condition that $g'(i)v = v$ gives us that $\theta \equiv 0 \pmod{\pi}$. As g is an element of $PSO_2(\mathbb{R})$, we must have $g = \pm I_2$. □

Corollary 1.11. *$T^1\mathbb{H}$ is isomorphic to $PSL_2(\mathbb{R})$.*

This identification has two steps. First, we choose an arbitrary point in $T^1\mathbb{H}$ to associate with the identity $\pm I_2 \in PSL_2(\mathbb{R})$. The most natural choice, given the lemmas already proven, is to choose (i,i). Then the element of $PSL_2(\mathbb{R})$ associated to an arbitrary vector (z,v) is the unique element g sending $(i,i) \to (z,v)$.

In the interest of saving space, we will state a few facts about geodesics on the hyperbolic plane without proof, only giving a summary of the main ideas. For more details see [9]

Lemma 1.12. *Between every two points, there exists a unique geodesic.*

To prove this lemma, we would first show that the unique geodesic between two points with the same real part is the vertical line connecting them. For another set of two points z_1, z_2, having different real parts, we can find a Möbius transform g such that $\operatorname{Re}(g \cdot z_1) = \operatorname{Re}(g \cdot z_2)$. Lemma 1.8 implies that the geodesic is given by the inverse of g on the vertical line.

Lemma 1.13. *Geodesics have the form given by the equation $x = c$ for some constant c, or by $(x-c)^2 + y^2 = r^2$ for some constants c and r. Equivalently, geodesics are semi-circles with centers on the real axis or vertical lines.*

This lemma is proven by showing that the image of a vertical line under a Möbius transform is either another vertical line, or a semi-circle centered on the origin. Although Möbius transformations do not contain all isometries of $\mathbb{H}$, the uniqueness of geodesics guarantees there are no other forms.

2. Geodesic Flow and Quotients by Latices.

Given a unit tangent vector $(z, v) \in T^1\mathbb{H}$, there is a unique, unit speed geodesic passing through z in the direction of v. By unit speed, we mean that our path ϕ has been reparameterized so that $||\phi'(t)||_{\phi(t)} = 1$ and $\phi : [0 : L(\phi)] \to \mathbb{H}$. When v points along the vertical line $x = c$ for some real valued c, this geodesic is the vertical line. Otherwise the geodesic is the unique semi-circle centered on the origin whose unit tangent vector at z is v.

Definition 2.1. The geodesic flow is the set of maps $g_t : T^1\mathbb{H} \to T^1\mathbb{H}$ determined by following the geodesic at z in the direction of v for time t.

Example 2.2. The geodesic flow of $(i, i) \in T^1\mathbb{H}$ is given by

$$g_t((i, i)) = (e^t i, e^t i).$$

It is enough to check that $(e^t i, e^t i)$ has unit speed. Using Definition 1.2, we get

$$\left|\left|\frac{d}{dt}e^t i\right|\right|_{e^t i} = \frac{e^t}{e^t} = 1.$$

We can see that we also get the same flow as the action of the matrix

$$\begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}.$$

If we use the identification of Corollary 1.11, we can find the geodesic flow of an arbitrary vector $(z, v) = g \cdot (i, i)$. Since g maps the unit speed geodesic defined by some reference vector to the unit speed geodesic defined by (z, v), it follows that g and g_t must commute. This observation allows us to write

$$g_t((z, v)) = g_t(g \cdot (i, i)) = g \cdot (g_t((i, i))) = g \cdot \left(\begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} \cdot (i, i) \right)$$

If we define the matrix

$$a_t = \begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2} \end{pmatrix}, \tag{2.3}$$

then we can write

$$g_t((z, v)) = (g a_t^{-1}) \cdot (i, i). \tag{2.4}$$

There is a point to make about why we have chosen to write this as the inverse of matrix. We have the association $T^1\mathbb{H} \cong PSL_2(\mathbb{R})$. The strength of this association is that it allows us to view any dynamics on $T^1\mathbb{H}$ as dynamics on $PSL_2(\mathbb{R})$. We already have Möbius transformations as a left action of $PSL_2(\mathbb{R})$ on itself and using the inverse is necessary to define a right action of the full group $PSL_2(\mathbb{R})$.

We denote the geodesic flow on $PSL_2(\mathbb{R})$ by

$$R_{a_t}(g) = g a_t^{-1}.$$

Ultimately the dynamics on $T^1\mathbb{H}$ is not particularly interesting. For any g in $PSL_2(\mathbb{R})$, the orbit exhibits no recurrence. Fortunately, we can obtain interesting dynamics when we look at a quotient of $PSL_2(\mathbb{R})$ by a discrete subgroup Γ. We will also need the quotient to have a finite volume.

Definition 2.5. The hyperbolic area and volume forms for $\mathbb{H}$ and $T^1\mathbb{H}$ are, respectively,

$$dA = \frac{1}{y^2}dxdy \tag{2.6}$$

$$dm = \frac{1}{y^2}dxdyd\theta. \tag{2.7}$$

Here θ is the angle of the unit tangent vector.

Theorem 2.8. *dA and dm are invariant under the actions of $PSL_2(\mathbb{R})$.*

Proof. The proof follows from computing the Jacobian and performing a substitution in the integral. □

Definition 2.9. A Fuchsian group is a discrete subgroup $\Gamma \leq PSL_2(\mathbb{R})$. A lattice is a Fuchsian group with the additional condition that a fundamental domain of $\Gamma\backslash PSL_2(\mathbb{R})$ has finite measure with respect to the measure m induced by the volume form dm. If the quotient space is compact, the lattice is said to be uniform. A hyperbolic surface is a quotient $\Gamma\backslash\mathbb{H}$ where Γ is a Fuchsian group.

As a reminder, a fundamental domain F is a measurable subset of $PSL_2(\mathbb{R})$ such the orbit of any $g \in PSL_2(\mathbb{R})$ under Γ intersects with F exactly once. This definition can be weakened to be larger or smaller by a null set.

The notion of the fundamental domain gives a natural way to extend the dynamics on the hyperbolic plane to dynamics on a hyperbolic surface $M = \Gamma\backslash\mathbb{H}$. If we let $X = \Gamma\backslash\mathbb{H}$, then we can still define a geodesic flow by

$$R_{a_t}(x) = xa_t^{-1}.$$

We can think of the geodesic flow on X as following the geodesic flow on $T^1\mathbb{H}$ until it intersects the boundary of F flowing outwards. We will put aside for now some technicalities about which boundary points are included or excluded from the fundamental domain. Once the flow reaches the boundary flowing outwards, there is exactly one element of Γ so that the image is on the boundary flowing inwards. To summarize, we follow the geodesic flow until reaching the boundary at which point we act by an element of Γ to return to the fundamental domain.

Example 2.10. An example will help visualize this. Consider $\Gamma = PSL_2(\mathbb{Z}) = SL_2(\mathbb{Z})/\{\pm I_2\}$, a discrete subgroup. A fundamental domain for the action of $PSL_2(\mathbb{Z})$ on $\mathbb{H}$ is given by

$$F = \{z \in \mathbb{H} \mid |z| \geq 1, |\mathrm{Re}(z)| \leq \frac{1}{2}\}.$$

This is not technically a fundamental domain as we need to include some of the boundary points but this will not matter for our purposes. A proof that this is a fundamental domain can be found in [1]. A plot of the fundamental domains is given in Figure 1.

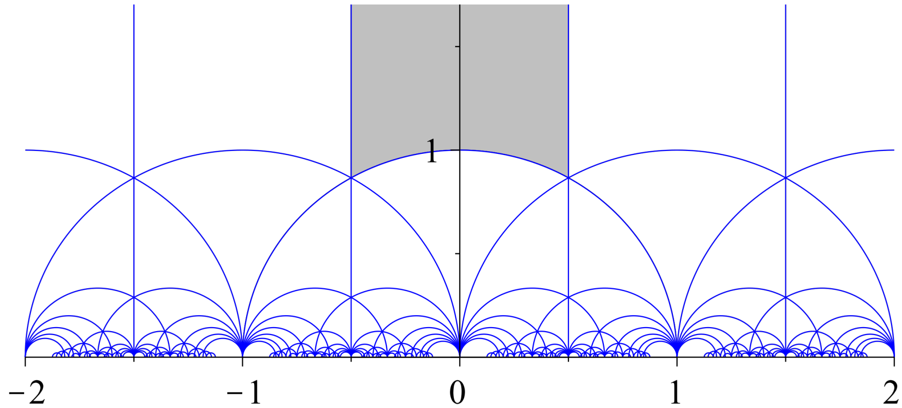

FIGURE 1. Examples of fundamental domains of $\Gamma = PSL_2(\mathbb{Z})$. The fundamental domain F is given by the shaded region. Image sourced from [2].

3. Ergodicity

This section will cover the definition of an ergodic transformation, and cite two important results we will need going into the quantum ergodicity section.

Definition 3.1. Let (M, μ) be a measure space with μ a probability measure. A measure preserving function f is said to be ergodic if for every Borel set $X \subset M$, $f(X) = X$ implies that either $\mu(X) = 0$ or $\mu(X) = 1$.

A common, intuitive way to understand an ergodic transformation is that time averages are equal to spacial averages. This idea is formalized in a key theorem on ergodicity.

Theorem 3.2 (Birkhoff's Ergodic Theorem)**.** *Let f be an ergodic transformation on (M, μ). Then for any continuous function $\phi : M \to \mathbb{R}$ and almost every $x \in M$ we have that the orbit of x is equidistributed, meaning*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x)) = \int_M \phi d\mu$$

Theorem 3.3. *Let $\Gamma \leq PSL_2(\mathbb{R})$ be a lattice. Let $X = \Gamma\backslash PSL_2(\mathbb{R})$. Then any non trivial element of the geodesic flow (a map R_{a_t} for $t \neq 0$) is ergodic with respect to m_x. Here m_x is the measure on X induced by the volume form dm.*

4. Quantum Ergodicity Overview

The precise statement of the quantum ergodicity (QE) theorem involves a large number of definitions and its proof requires many hard lemmas. To avoid getting lost in the lemmas we will dedicate a short, less technical section ahead of the final section to motivate the theorem. In essence, we will try to give an intuitive explanation of why we might expect that eigenvectors of the Laplacian should tend to be equidistributed in the high eigenvalue limit when the geodesic flow is ergodic.

In classical mechanics, we are concerned with the trajectory of a particle which is given by a flow on the phase space. Earlier in the paper, we covered classical trajectories on a hyperbolic surface, which are realized as geodesic flows. The equivalent problem in quantum mechanics is to find eigenfunctions φ_j of the Schrödinger operator. To simplify we will consider a zero potential system, which reduces the problem to studying eigenvectors of the Laplacian. We expect, from the correspondence principle, that the behavior of classical trajectories should be reflected in the high eigenvalue limit of φ_j. The question that the QE theorem answers is "what effect does the ergodicity of the geodesic flow have on the distribution of high eigenvalue eigenfunctions?" One way to answer this question would be to construct an analog of Theorem 3.2 in quantum mechanics in the limit of high eigenvalues. Since the eigenfunctions are analogous to the classical trajectories, we can expect them to be equidistributed. Informally, this is

$$\int_M f|\phi_j|^2 d\mu \to \int_M f d\mu.$$

as $j \to \infty$ for all f in $C^\infty(M)$. We will prove a more formal result for a subsequence of φ_j as a direct corollary of the QE theorem. Before we continue, we should give the explicit form of the Laplacian on $\mathbb{H}$.

Definition 4.1. The Laplacian on $\mathbb{H}$ is given by

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

We obtain this form of the Laplacian by rescaling the Laplacian on $\mathbb{C}$ using the metric from Definition 1.2.

5. The Quantum Ergodic Theorem on Hyperbolic Surfaces

We begin with a quick review of Lie algebras. The Lie algebra of $G = SL_2(\mathbb{R})$ is given by

$$\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R}) = \{X \in M_2(\mathbb{R}) : \forall t \in \mathbb{R} \quad \exp(tX) \in G\}. \tag{5.1}$$

Using the property of the matrix exponential that $\det(\exp(X)) = \exp(\mathrm{Tr}(X))$, we can rewrite $\mathfrak{g}$ as

$$\mathfrak{g} = \{X \in M_2(\mathbb{R}) : \mathrm{Tr}(X) = 0\}. \tag{5.2}$$

We can easily find a basis for the Lie algebra. $M_2(\mathbb{R})$ has dimension 4, and the traceless condition reduces the dimension of $\mathfrak{sl}_2(\mathbb{R})$ to 3. An example of a basis is

$$H = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \quad U^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad U^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{5.3}$$

Note that the matrix exponential

$$\exp(tH) = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} = a_t^{-1}$$

gives the geodesic flow on $PSL_2(\mathbb{R})$ by multiplication on the right, as previously defined in (2.4). The names U^+ and U^- refer to a type of flow called horocycle flow . We will not cover horocycle flow in this paper. For more information, see chapter 11 in [9].

Definition 5.4. For an element of the Lie algebra $X \in \mathfrak{g}$, we can define a differential operator $D_X : C^\infty(G)) \to C^\infty(G))$ by

$$D_X f(g) = \frac{d}{dt} f(g \cdot \exp(tX))_{|t=0}.$$

We can define an inner product on two smooth compactly supported functions f_1 and f_2 defined on G by

$$\langle f_1, f_2 \rangle = \int f_1(g)\overline{f_2(g)}dg. \tag{5.5}$$

This differential operator has a few key properties. The first property is linearity: for $X, Y \in \mathfrak{g}$ and real numbers α and β, we have

$$D_{\alpha X + \beta Y} = \alpha D_X + \beta D_Y. \tag{5.6}$$

Next, it preserves the Lie bracket.

$$D_X D_Y - D_Y D_X = D_{[X,Y]} \tag{5.7}$$

Here the Lie bracket is given by the commutator $[X, Y] = XY - YX$. Lastly, the adjoint is given by

$$D_X^\star = D_{-X^\dagger} \tag{5.8}$$

Here $X^\dagger$ denotes the Hermitian conjugate of X. Each property follows quickly from the definition and the properties of matrix exponentials.

By linearity, we can extend the differentiation to $\mathfrak{g} \otimes \mathbb{C}$ by

$$D_Z = D_X + iD_Y$$

for $Z = X + iY \in \mathfrak{g} \otimes \mathbb{C}$.

We can now define the Casimir operator Ω, which will help us understand the Laplacian on $\mathbb{H}$.

Definition 5.9. The Casimir operator is given by

$$\Omega = D_H D_H + \frac{1}{2} D_{U^+} D_{U^-} + \frac{1}{2} D_{U^-} D_{U^+}.$$

The Casimir operator has the important property that it commutes with all differential operators.

Lemma 5.10. *Ω commutes with all differential operators.*

Proof. Since D_H, D_{U^+}, and D_{U^-} form a basis of $\mathfrak{g}$, the result follows from showing Ω commutes with each by linearity. □

Definition 5.11. Let $K = SO_2(\mathbb{R})$. Any element in K can be written as

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

A function $f : G \to G$ is K-invariant if, for all $k \in K$ and we have $f(g \cdot k) = f(g)$

Lemma 5.12. *Recall from Lemma 1.7 that $G/K \cong \mathbb{H}$. The restriction of Ω to K-invariant functions coincides with the Laplacian on $\mathbb{H}$.*

The first step of the proof is to find the Iwasawa decomposition of $SL_2(\mathbb{R})$, which allows us to write any element g as

$$g = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} = u_x a_y k_\theta. \tag{5.13}$$

We already did most of the work to prove the decomposition by studying the action of $SL_2(\mathbb{R})$ on $\mathbb{H}$ and $T^1\mathbb{H}$ in part 1. In particular, Lemmas 1.6 and 1.7 give the way to write g as a product $u_x a_y k_\theta$ and Lemma 1.10 gives the uniqueness. This decomposition gives us coordinates (x, y, θ) where the x and y coordinates correspond to the coordinates used in Definition 4.1 for the Laplacian. The next step would be to write D_H, D_{U^+}, and D_{U^-} in terms of these coordinates, but we will exclude this step for the sake of brevity. A rigorous proof is given in [6].

We will now consider the group $G = PSL_2(\mathbb{R})$.

Definition 5.14. The space of K-eigenfunctions of weight 2n is the set

$$A_{2n} = \{f \in \mathbb{C}^\infty(\Gamma\backslash G) : f(g \cdot k_\theta) = e^{2in\theta} f(g).$$

There are a couple of remarks about this set. The first remark is that the set A_0 is just the set of K-invariant functions on $\Gamma\backslash G$. For A_{2n}, we can realize them as eigenspaces of a differential operator D_W, where

$$W = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = U^+ - U^-. \tag{5.15}$$

This operator corresponds to circular flow. To see the correspondence to circular flow, we take the matrix exponential of Wt

$$\exp(Wt) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

and see that $\exp(Wt) = k_t$. We can prove that $f \in A_{2n}$ is an eigenfunction of D_W with eigenvalue $2in$. We have

$$\begin{aligned} D_W f &= \frac{d}{dt} f(g \cdot \exp(Wt))_{|t=0} \\ &= \frac{d}{dt}(e^{2int} f)_{|t=0} \\ &= 2inf. \end{aligned}$$

This identification of weight spaces with eigenfunctions of D_W allows us to prove an important property of the weight spaces.

Lemma 5.16. *For $f \in A_{2n}$ and $h \in A_{2m}$ we have $f \cdot h \in A_{2(n+m)}$*

Proof. The result follows immediately from applying the product rule.

$$\begin{aligned} D_W(f \cdot h) &= D_W(f) \cdot h + f \cdot D_W(h) \\ &= 2inf \cdot h + 2imf \cdot h \\ &= 2i(n+m) f \cdot h \end{aligned}$$

□

Immediate corollaries of the association of the weight spaces with eigenfunctions of D_W follow a Fourier decomposition.

Lemma 5.17. *A_{2n} and A_{2m} are orthogonal when $n \neq m$.*

Lemma 5.18.

$$\overline{\bigoplus_{n\in\mathbb{Z}} A_{2n}} = C^\infty(\Gamma\backslash G)$$

Definition 5.19. A function $f \in C^\infty(\Gamma\backslash G)$ is K-finite if there exists a natural number N such that

$$f \in \bigoplus_{n=-N}^{N} A_{2n}.$$

It will be very useful to define differential operators that can turn a function in one weight space into a function in another weight space. To accomplish this goal we will define the raising and lowering operators.

Definition 5.20. The raising operator is the differential operator given by the matrix

$$E^+ = \frac{1}{2}\begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}.$$

The lowering operator is given by the matrix

$$E^- = \frac{1}{2}\begin{pmatrix} 1 & -i \\ -i & -1 \end{pmatrix}.$$

We need to verify that the raising and lowering operators actually do raise and lower the weight space of a function.

Lemma 5.21.

$$D_{E^\pm(A_{2n})} \subseteq A_{2(n\pm 1)}.$$

Proof. Since the weight spaces are the eigenspaces of D_W, the result will follow from finding the commutator $[D_W, D_{E^\pm}]$. Using (5.7), it suffices to find the commutator of the matrices.

$$\begin{aligned}
[W, E^\pm] &= WE^\pm - E^\pm W \\
&= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & \pm i \\ \pm i & -1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & \pm i \\ \pm i & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} \pm i & -1 \\ -1 & \mp i \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \mp i & 1 \\ 1 & \pm i \end{pmatrix} \\
&= \begin{pmatrix} \pm i & -1 \\ -1 & \mp i \end{pmatrix} \\
&= \pm 2iE^\pm
\end{aligned}$$

Now we need to show that if f is an eigenfunction of D_W with eigenvalue $2in$, then $D_{E^\pm} f$ is an eigenfunction with eigenvalue $2i(n \pm 1)$. We can use the commutator relation to obtain

$$\begin{aligned}
D_W D_{E^\pm} f &= D_{E^\pm} D_W f + [D_W, D_{E^\pm}]f \\
&= 2in D_{E^\pm} f \pm 2i D_{E^\pm} f \\
&= 2i(n \pm 1) D_{E^\pm} f.
\end{aligned}$$

□

An important property of the raising and lowering operators is that, together with D_W, they form a basis for $\mathfrak{g} \otimes \mathbb{C}$. Using linearity we can find more useful expressions for the Casimir operator. Once simplified, we can write Ω as

$$\Omega = D_{E^+} D_{E^-} - \frac{1}{4} D_W^2 + \frac{i}{2} D_W = D_{E^-} D_{E^+} - \frac{1}{4} D_W^2 - \frac{i}{2} D_W. \tag{5.22}$$

We can now construct the microlocal lift. Our goal is to construct a lift that is both invariant under the geodesic flow and projects to a probability measure on $M = \Gamma\backslash\mathbb{H}$.

Definition 5.23. Let φ be a normalized eigenfunction of the Laplacian with eigenvalue $\lambda = \frac{1}{4} + r^2$. We can extend this eigenfunction to a K-invariant function φ_0 on G. We inductively define

$$\varphi_{2n+2} = \frac{1}{ir + \frac{1}{2} + n} D_{E^+}(\varphi_{2n}), \text{ if } n \geq 0$$
$$\varphi_{2n-2} = \frac{1}{ir + \frac{1}{2} - n} D_{E^-}(\varphi_{2n}), \text{ if } n \leq 0$$

The coefficients are chosen to normalize the raising and lowering operators. Since Lemma 5.12 shows that the restriction of Ω to K-invariant functions coincides with the Laplacian, we have $\Omega\varphi_0 = \lambda\varphi_0$. Recall from Lemma 5.10 that Ω commutes with all differential operators, in particular D_{E^+} and D_{E^-}. This observation gives us $\Omega\varphi_{2n} = \lambda\varphi_{2n}$ for all n.

Definition 5.24. Let f be a K-finite function. The microlocal lift is a linear functional given by

$$I_\varphi(f) = \langle f \sum_{n\in\mathbb{Z}} \varphi_{2n}, \varphi_0 \rangle.$$

Definition 5.25. We can define a probability measure on $\Gamma\backslash G$ in the following way. First, let $N = N(\lambda)$ be some integer function of the eigenvalue. We will choose $N = [r]$. Then we can define

$$\psi = \frac{1}{2N+1} \sum_{n=-N}^{N} \varphi_{2n} \tag{5.26}$$

The φ_{2n} are normalized and orthogonal, hence ψ is normalized. We obtain a probability measure on $\Gamma\backslash G$ as

$$\nu = |\psi|^2(g)dg. \tag{5.27}$$

We now want to prove that the microlocal lift is close to ν when the eigenvalue, or equivalently r, is large. The following lemma is the first step in examining the asymptotics of the microlocal lift.

Lemma 5.28. *Let f be a K-finite function. Then we have*

$$\left| I_\varphi(f) - \int_{\Gamma\backslash G} f d\nu \right| = O(r^{-1/2})$$

In particular this lemma tells us that a subsequence of I_φ has a weak limit that is a probability measure as $\lambda \to \infty$.

Proof. The essential idea of the proof is as follows. First we rewrite the integral of $f d\nu$ as an inner product involving the φ_{2n}. Then, using the K-finite Property, the orthogonality of the weight spaces, and the raising and lowering operators, we can get a bound on the error.

Begin by rewriting the integral with respect to $d\nu$.

$$\int f d\nu = \langle f\psi, \psi\rangle = \frac{1}{2N+1}\sum_{n,m=-N}^{N}\langle f\varphi_{2n}, \varphi_{2m}\rangle$$

Since f is K-finite, there exists some L such that $f \in \bigoplus_{l=-L}^{L} A_{2l}$. Let f_{2l} be the component of f in A_{2l}. Using Lemma 5.16, we get that the product of f_{2l} and ϕ_{2n} is an element of $A_{2(l+n)}$. The orthogonality condition gives that the inner product is only non-zero when $l+n=m$. This condition gives the bound $|m-n| \le L$ by

$$\begin{aligned} -L \le l &\le L \\ -L+n \le l+n &\le L+n \\ -L+n \le m &\le L+n \\ -L \le m-n &\le L. \end{aligned}$$

We can then write

$$\langle f\psi, \psi\rangle = \frac{1}{2N+1}\sum_{n,m=-N}^{N} 1_{|n-m|\le L}\langle f\varphi_{2n}, \varphi_{2m}\rangle. \tag{5.29}$$

We will return to this equation later in the proof. For now, it is important to note that this equation will allow us to change from a sum from $-N$ to N, to a sum from $-L$ to L. We can rewrite $\langle f\varphi_{2n}, \varphi_{2m}\rangle$ using the raising operator.

$$\begin{aligned} \langle f\varphi_{2n}, \varphi_{2m}\rangle &= \frac{1}{ir-\frac{1}{2}+n}\langle f D_{E^+}\varphi_{2n-2}, \varphi_{2m}\rangle \\ &= \frac{1}{ir-\frac{1}{2}+n}\left(\langle D_{E^+}(f\varphi_{2n-2}), \varphi_{2m}\rangle - \langle D_{E^+}(f)\varphi_{2n-2}, \varphi_{2m}\rangle\right) \end{aligned}$$

Here the first line comes from the definition of φ_{2n}, and the second line comes from the product rule for the differential operator E^+. The Cauchy-Schwarz inequality bounds the second term to $O(r^{-1})$. We can then use (5.8) to move the raising operator to the right side of the inner product.

$$\begin{aligned} \langle f\varphi_{2n}, \varphi_{2m}\rangle &= \frac{1}{ir-\frac{1}{2}+n}\langle f\varphi_{2n-2}, D_{E^-}(\varphi_{2m})\rangle + O(r^{-1}) \\ &= \frac{-ir+\frac{1}{2}-m}{ir-\frac{1}{2}+n}\langle f\varphi_{2n-2}, \varphi_{2m}\rangle + O(r^{-1}) \\ &= \langle f\varphi_{2n-2}, \varphi_{2m-2}\rangle + O(r^{-1}) \end{aligned}$$

We can then iterate this process to get

$$\begin{aligned}\langle f\varphi_{2n}, \varphi_{2m}\rangle &= \langle f\varphi_{2n-2m}, \varphi_0\rangle + O(mr^{-1}) \\ &= \langle f\varphi_{2n-2m}, \varphi_0\rangle + O(r^{-1/2}).\end{aligned}$$

Returning to (5.29) yields

$$\begin{aligned}\langle f\psi, \psi\rangle &= \frac{1}{2N+1}\sum_{n,m=-N}^{N} 1_{|n-m|\leq L}\langle f\varphi_{2n}, \varphi_{2m}\rangle \\ &= \frac{1}{2N+1}\sum_{n,m=-N}^{N} 1_{|n-m|\leq L}\left(\langle f\varphi_{2n-2m}, \varphi_0\rangle + O(r^{-1/2})\right) \\ &= \sum_{l=-L}^{L}\frac{2N+1-|l|}{2N+1}\left(\langle f\varphi_{2l}, \varphi_0\rangle + O(r^{-1/2})\right) \\ &= \langle f\sum_{l=-L}^{L}\varphi_{2l}, \varphi_0\rangle + O(r^{-1/2}) \\ &= I_\varphi(f) + O(r^{-1/2}).\end{aligned}$$

□

We now want to show that the measure ν is asymptotically invariant under the geodesic flow.

Lemma 5.30. *For any K-finite function f,there exists a differential operator D_L, which is independent of r, such that*

$$I_\varphi((rD_H + D_L)f) = 0.$$

For a proof of this lemma see [5]. We will now explain how this lemma gives asymptotic invariance. The linearity of the microlocal lift gives

$$\begin{aligned}0 &= I_\varphi((rD_H + D_L)f) \\ &= rI_\varphi(D_H f) + I_\varphi(D_L f).\end{aligned}$$

Thus, we have

$$I_\varphi(D_H f) = O_f(r^{-1}). \tag{5.31}$$

If we combine this fact with the previous lemma, which gives us a probability measure $\tilde{\nu}$ in the $\lambda \to \infty$ limit, we have

$$\int_{\Gamma\backslash G} D_H f d\tilde{\nu} = 0. \tag{5.32}$$

If we let $\tilde{f}(g) = f(ga_t^{-1})$ for some t, then

$$\begin{aligned} D_H \tilde{f} &= \frac{d}{ds} f(ga_s^{-1} a_t^{-1})_{|s=0} \\ &= \frac{d}{ds} f(ga_{s+t}^{-1})_{|s=0} \\ &= \frac{d}{dt} f(ga_t^{-1}). \end{aligned}$$

Applying (5.32) to $\tilde{f}$ yields

$$\frac{d}{dt} \int_{\Gamma \backslash G} f(ga_t^{-1}) d\tilde{\nu} = 0.$$

Integrating this equation shows that the measure of any function is invariant under the geodesic flow.

We are now almost ready to prove the QE theorem. First, we will need one last lemma.

Lemma 5.33. *Let $\{\varphi_j\}_{j \in \mathbb{N}}$ be an orthonormal basis of eigenfunctions of the Laplacian, with eigenvalues $\lambda_j = \frac{1}{4} + r_j^2$. We will write $X = \Gamma \backslash G$. Then for any K-finite function $f \in C^\infty(X)$ and any $\epsilon > 0$, we have*

$$\frac{1}{N(L,\epsilon)} \sum_{j : |\lambda_j - L| < \epsilon} I_{\varphi_j}(f) \to \frac{1}{m_X(X)} \int_X f(g) dg$$

as $L \to \infty$. Here $N(L,\epsilon)$ is the number of eigenvalues within ϵ of L.

$$N(L,\epsilon) = \#\{j : |\lambda_j - L| < \epsilon\}$$

This lemma tells us that the sequence of microlocal lifts converges to the normalized measure induced by the volume form defined in (2.7). Proving this lemma relies on writing the microlocal lift of a function f as the expectation value of some quantized operator $Op(f)$ in the quantum state φ. For a proof see [3]. The QE theorem is a stronger version of lemma in which the convergence is absolute.

Theorem 5.34. *for any K-finite function $f \in C^\infty(X)$ and any $\epsilon > 0$, we have*

$$\frac{1}{N(L,\epsilon)} \sum_{j : |\lambda_j - L| < \epsilon} \left| I_{\varphi_j}(f) - \frac{1}{m_X(X)} \int_X f(g) dg \right|^2 \to 0$$

as $L \to \infty$.

Proof. The main idea behind this proof is to consider a time average along the geodesic flow and use the earlier lemmas to prove it is asymptotically close to the microlocal lift. First, note that it is enough to consider only functions satisfying

$$\int_X f(g) dg = 0.$$

This follows from the fact that the theorem is easily shown to hold for constant functions by the linearity of the microlocal lift and the integral. We define the time average over the geodesic flow $M_T(f)$ as

$$M_T(f)(g) = \frac{1}{T} \int_0^T f(ga_t^{-1}) dt.$$

We can apply the asymptotic invariance of I_{φ_j} from Lemma 5.30 to the time average.

$$I_{\varphi_j}(M_T(f)) = I_{\varphi_j} + O(1/\lambda_j)$$

In the sum, we restrict to $|\lambda_j - L| < \epsilon$. Therefore, we can alternatively write the previous line as

$$I_{\varphi_j}(M_T(f)) = I_{\varphi_j} + O(1/L).$$

Lemma 5.28 gives a probability measure ν_j where

$$\left|I_{\varphi_j}(f) - \nu_j(f)\right| = O(1/\sqrt{L}). \tag{5.35}$$

Putting the previous two facts together gives us

$$\begin{aligned}\frac{1}{N(L,\epsilon)} \sum_{j:|\lambda_j - L|<\epsilon} \left|I_{\varphi_j}(f)\right|^2 &= \sum_{j:|\lambda_j - L|<\epsilon} \left|\int M_T(f) d\nu_j\right|^2 + O(1/\sqrt{L}) \\ &\leq \sum_{j:|\lambda_j - L|<\epsilon} \int |M_T(f)|^2 d\nu_j + O(1/\sqrt{L}).\end{aligned}$$

Applying (5.35) again gives

$$\frac{1}{N(L,\epsilon)} \sum_{j:|\lambda_j - L|<\epsilon} \left|I_{\varphi_j}(f)\right|^2 \leq \sum_{j:|\lambda_j - L|<\epsilon} \int I_{\varphi_j}\left(|M_T(f)|^2\right) d\nu_j + O(1/\sqrt{L}).$$

We can take the lim sup and apply Lemma 5.33 to the right hand side to conclude

$$\limsup \frac{1}{N(L,\epsilon)} \sum_{j:|\lambda_j - L|<\epsilon} \left|I_{\varphi_j}(f)\right|^2 = \frac{1}{m_X(X)} \int_X |M_T(f)|^2 dg.$$

Now we can finally use our ergodicity condition. Taking the limit as $T \to \infty$ and using Theorem 3.2 gives

$$\lim_{T\to\infty} \frac{1}{m_X(X)} \int_X |M_T(f)|^2 dg = \frac{1}{m_X(X)} \int_X f(g) dg$$

which by our earlier hypothesis is equal to 0. □

Finally we have the immediate corollary stating that a dense subsequence of eigenfunctions tends towards the measure m_X.

Corollary 5.36. *There exists a subsequence φ_{j_k} such that for any K-finite f*

$$I_{\varphi_{j_k}}(f) \to \frac{1}{m_X(X)} \int_X f(g) dg$$

as $k \to \infty$, and the subsequence j_k is dense in j.

Acknowledgments

Thank you to my mentor, DeVon Ingram, for suggesting the topic and for providing the time to help me understand the material.

References

[1] A. G. Kurosh, The theory of groups (Chelsea Publishing Co., New York, 1960). Translated from the Russian and edited by K. A. Hirsch. 2nd English ed. 2 volumes. (fundamental domain part)
[2] https://commons.wikimedia.org/wiki/File:ModularGroup-FundamentalDomain.svg
[3] Steven Zelditch, Uniform distribution of eigenfunctions on compact hyperbolic surfaces, Duke Math. J. 55 (1987), no. 4, 919–941. MR 916129
[4] Alexander Gorodnik, Microlocal lifts and quantum ergodicity https://people.maths.bris.ac.uk/ mazag/hyperbolic/lecture7.pdf
[5] Etienne Le Masson, Quantum chaos on hyperbolic surfaces, https://perso.u-cergy.fr/ elemasson/qchyp.pdf.
[6] M.E. and T.W., Arithmetic Quantum Unique Ergodicity, http://swc.math.arizona.edu/aws/2010/2010EinsiedlerNotes.pdf
[7] Felix J. Wong, Quantum Ergodicity and the Analysis of Semiclassical Pseudodifferential Operators, https://www.math.harvard.edu/media/wong.pdf
[8] Amie Wilkinson, Mechanisms for Chaos, http://www.math.uchicago.edu/ wilkinso/papers/VIASMnotes5Aug2019.pdf
[9] Manfred Einsiedler and Thomas Ward, Ergodic Theory with a view towards Number Theory