1

Introduction to (A)QUE

1.1 (Arithmetic) Quantum Unique Ergodicity

In order to formulate the conjecture, we recall some notation (see [7, Chap. 9] for a detailed treatment).

- The upper half-plane model for the hyperbolic plane is

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

- The group $\mathrm{SL}_2(\mathbb{R})$ acts transitively on $\mathbb{H}$ via Möbius transformations: the matrix $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ acts via

$$g : z \longmapsto g \bullet z = \frac{az+b}{cz+d}.$$

- Any subgroup $\Gamma \leqslant \mathrm{SL}_2(\mathbb{R})$ defines an associated quotient space $M = \Gamma\backslash\mathbb{H}$ under this action, and the quotient space inherits a measure vol_M from the volume measure $\mathrm{dvol}_M = \frac{1}{y^2}\,\mathrm{d}x\,\mathrm{d}y$ on $\mathbb{H}$, meaning that

$$\int f(x+\mathrm{i}y)\,\mathrm{dvol}_M = \int f(x+\mathrm{i}y)\frac{1}{y^2}\,\mathrm{d}x\,\mathrm{d}y.$$

  This allows us to speak of function spaces like $L^2(M) = L^2_{\mathrm{vol}_M}(M)$.
- A sequence of measures $m_1, m_2, \ldots$ is said to converge weak* to a measure m, denoted by $m_i \underset{\text{weak}^*}{\longrightarrow} m$ as $i \to \infty$, if

$$\int f\,\mathrm{d}m_i \longrightarrow \int f\,\mathrm{d}m$$

  as $i \to \infty$ for any continuous function f with compact support.

- The *Laplacian* on M is the operator

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

  and a function $\phi \in C^\infty(M) \cap L^2(M)$ is an *eigenfunction* for Δ with *eigenvalue* λ if $\Delta\phi = \lambda\phi$. An eigenvalue is *normalized* if $\|\phi\|_2 = 1$.

Conjecture 1.1 (QUE; Rudnick–Sarnak). Let Γ be a discrete subgroup of $\mathrm{SL}_2(\mathbb{R})$ such that $M = \Gamma\backslash\mathbb{H}$ is compact. If $\{\phi_i \mid i \in \mathbb{N}\}$ are normalized eigenfunctions for Δ in $C^\infty(M)$ with corresponding eigenvalues $\{\lambda_i \mid i \in \mathbb{N}\}$ such that $\lambda_i \to \infty$ as $i \to \infty$, then

$$|\phi_i|^2 \,\mathrm{dvol}_M \underset{\text{weak*}}{\longrightarrow} \mathrm{dvol}_M \tag{1.1}$$

as $i \to \infty$.

The motivation for this conjecture comes from physical considerations, but it has wide-ranging mathematical meaning. We address the motivation via a series of questions.

1.1.1 Why is $\boldsymbol{\Delta}$ the differential operator studied?

(1) If $\nabla(f)$ denotes the total derivative of a function $f \in C_c^\infty(M)$ (that is, an infinitely differentiable function with compact support), then it can be shown that $\Delta(f)$ is equal to $-\nabla^*\overline{\nabla}(f)$, where ∇^* is the adjoint operator, and $\overline{\nabla}$ is the closed operator defined by ∇. This slightly mysterious observation suggests that Δ is a natural operator, and that its eigenvalues are negative.

(2) The operator Δ is the restriction of the *Casimir operator* ω, which is a differential operator of degree two on $\mathrm{SL}_2(\mathbb{R})$ with unique invariance properties. In fact ω restricted to the space of functions on $\mathbb{H} = \mathrm{SL}_2(\mathbb{R})/\mathrm{SO}(2)$ coincides with Δ (this will be discussed in detail in the course and the notes). Here $\mathrm{SO}(2)$ denotes the special orthogonal group of matrices of the form $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = k_\theta$.

(3) In Schrödinger's quantum theory, the motion of a free (spinless, non-relativistic) quantum particle, moving in the absence of external forces on M, satisfies the equation

$$\mathrm{i}\frac{\partial\psi}{\partial t} = \Delta\psi.$$

This defines a *unitary evolution*, meaning that $\|\psi(\cdot,t)\|_{L^2(\mathrm{vol}_M)}$ is independent of t – so without loss of generality we may normalize and assume that $\|\psi(\cdot,t)\|_{L^2(\mathrm{vol}_M)} = 1$. The Born interpretation gives an empirical meaning to the "wave function" ψ by interpreting $|\psi|^2$ as the distribution of the position of the particle in the sense of probability. The eigenfunction equation $\Delta\psi = \lambda\psi$ corresponds to studying a particle with a given energy $-\lambda$. Thus the QUE conjecture concerns itself with the high-energy limit (also called the semi-classical limit). In fact the QUE conjecture implies a strengthening of the uncertainty principle: If ψ has a given large energy, then not only is the position of the particle uncertain, it is in fact almost equidistributed.

1.1.2 Why are there eigenfunctions?

(1) If M is compact, then the operator $(I - \Delta)^{-1}$ is a compact operator on $L^2(M)$. It follows that $L^2(M)$ is spanned by the eigenfunctions of Δ, and for every $\lambda \in \mathbb{R}$ the corresponding eigenspace

$$\{\phi \mid \Delta\phi = \lambda\phi\} \tag{1.2}$$

is finite-dimensional.

(2) If M is not compact, then $(I-\Delta)^{-1}$ is not a compact operator. In this case there may in general not be any eigenfunctions at all. However, if $\Gamma = \mathrm{SL}_2(\mathbb{Z})$ (or a congruence subgroup, defined below), then one can again show that Δ has infinitely many eigenfunctions in $C^\infty(M) \cap L^2(M)$, and that the eigenspaces (1.2) are once again finite-dimensional. Rudnick and Sarnak also conjectured that in this case (1.1) should hold.

1.1.3 What other reasons are there to study eigenfunctions of Δ?

Apart from the quantum-mechanical interpretation in Section 1.1.1(4), the eigenfunctions of the Laplacian arise in many parts of mathematics.

(1) On compact quotients, they give the most canonical orthonormal basis of $L^2(M)$. This is part of the theory of harmonic analysis (the appropriate generalization of Fourier analysis) on M.

(2) The eigenfunctions, which are also called *Maass cusp forms*, are intimately related to L-functions in number theory.

1.1.4 The Result

Conjecture 1.1 is, in full generality, open. However, there are some important cases for which it is known. In order to describe these, we need to make a few more definitions. We call Γ a *congruence lattice over* $\mathbb{Q}$ if either

- Γ is a *congruence subgroup* of $\mathrm{SL}_2(\mathbb{Z})$, meaning that

$$\Gamma \supseteq \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma \equiv I_2 \bmod N\}$$

  for some $N \geqslant 1$; or

- Γ is a lattice derived from a Eichler order in an $\mathbb{R}$-split quaternion division algebra over $\mathbb{Q}$.

The first type has the advantage of being quite concrete, and includes familiar examples like $\Gamma = \mathrm{SL}_2(\mathbb{Z})$; the second type has the advantage that in those cases the lattice is *uniform*, meaning that the quotient space $\Gamma\backslash\mathbb{H}$ is compact. In either case, it is possible (and we will do this later) to define a collection of additional operators $\{T_n\}$, called Hecke operators, which commute with Δ and with each other. These operators therefore act on the finite-dimensional eigenspaces (1.2), and are simultaneously diagonalizable. A *Hecke–Maass cusp form* is a joint eigenfunction $\phi \in C^\infty(M) \cap L^2(M)$ of Δ and of all the Hecke operators T_n for $n \geqslant 2$.

Lindenstrauss [10] and Soundarajan [15] together have shown the following, which we refer to as *arithmetic quantum unique ergodicity* (AQUE).

Theorem 1.2. *Let $M = \Gamma\backslash\mathbb{H}$, with Γ a congruence lattice over $\mathbb{Q}$. Then*

$$|\phi_i|^2 \,\mathrm{dvol}_M \underset{weak^*}{\longrightarrow} \mathrm{dvol}_M$$

as $i \to \infty$ for any sequence of Hecke–Maass cusp forms for which the Maass eigenvalues $\lambda_i \to -\infty$ as $i \to \infty$.

We briefly summarize some of the history leading up to this result. In 2001 Watson [16] showed this under the assumption of the Generalized Riemann Hypothesis (GRH), also obtaining under this assumption the optimal rate of convergence. In 2006 Lindenstrauss [10] obtained the result unconditionally, using ergodic methods, for lattices derived from Eichler orders and (almost) for congruence subgroups of $\mathrm{SL}_2(\mathbb{Z})$. For the latter case, Lindenstrauss showed that any weak*-limit is of the form $c\,\mathrm{dvol}_M$ for some $c \in [0,1]$ – in other words *escape of mass* to infinity was not ruled out. In 2009 Soundarajan [15] established, in a short paper of ten pages, that any weak*-limit is a probability measure – that is, escape of mass is not possible. Combined with [10], this proved Theorem 1.2.

1.2 Introduction to (Measure Rigidity of) the Geodesic Flow

Recall that the unit tangent bundle $\mathrm{T}^1\mathbb{H}$ of the hyperbolic plane is isomorphic to $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\{\pm I_2\}$, by identifying the matrix $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with the point

$$\left(g\centerdot\mathrm{i}, \frac{1}{(cz+d)^2}\mathrm{i}\right) \in \mathrm{T}^1\mathbb{H}.$$

Under this isomorphism, the geodesic flow (which, by definition, follows the geodesic determined by the arrow $(z,v) \in \mathrm{T}^1\mathbb{H}$ with unit speed, as illustrated in Figure 1.1) corresponds on $\mathrm{PSL}_2(\mathbb{R})$ to right-multiplication by $\begin{pmatrix} e^{t/2} & \\ & e^{-t/2} \end{pmatrix}$ for $t \in \mathbb{R}$ (see [7, Chap. 9] for a detailed treatment).

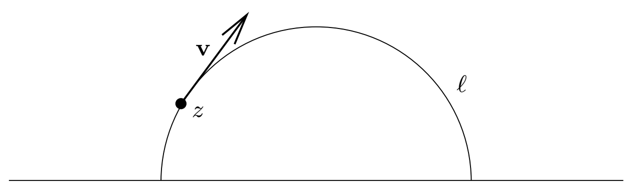

Fig. 1.1. The unique geodesic ℓ defined by a pair $(z, \mathbf{v})$.

This flow (that is, action of $\mathbb{R}$) naturally descends to the quotient

$$\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R}),$$

or to any other quotient $\Gamma\backslash \mathrm{SL}_2(\mathbb{R})$. Recall that, by definition, Γ is a lattice if $X = \Gamma\backslash \mathrm{SL}_2(\mathbb{Z})$ supports an $\mathrm{SL}_2(\mathbb{R})$-invariant probability measure m_X, which we always assume. The measure m_X is also called the Haar measure of X and, if projected to $M = X/\mathrm{SO}(2)$, gives the normalized volume measure vol_M.

It is interesting to note that there are dense orbits of the geodesic flow on X. In fact, for almost every starting point $x \in X$ the orbit is *equidistributed* with respect to m_X, meaning that

$$\frac{1}{T}\int_0^T f\left(x\begin{pmatrix} e^{t/2} & \\ & e^{-t/2} \end{pmatrix}\right) \mathrm{d}t \longrightarrow \int_X f \,\mathrm{d}m_X$$

as $T \to \infty$ for any $f \in C_c(X)$. This is a consequence of the ergodicity of m_X with respect to the geodesic flow and Birkhoff's pointwise ergodic theorem. We mention these important but basic concepts from ergodic theory only in passing, as they will not be used in these lectures (see [7, Th. 2.30] for the pointwise ergodic theorem, [7, Sect. 4.4.2] for a discussion of generic points, [7, Sect. 9.5] for an account of Hopf's proof of ergodicity for the geodesic flow, and [7, Sect. 11.3] for an explanation of the 'Mautner phenomena' and ergodicity of the geodesic flow).

It is also interesting to note that there are many periodic orbits for the geodesic flow. For example, the matrix

$$\gamma = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

is diagonalizable by some $k \in \mathrm{SO}(2)$ and has positive eigenvalues, so that

$$\mathrm{SL}_2(\mathbb{Z})k\begin{pmatrix} e^{t_0/2} & \\ & e^{-t_0/2} \end{pmatrix} = \mathrm{SL}_2(\mathbb{Z})\gamma k = \mathrm{SL}_2(\mathbb{Z})k$$

for some $t_0 > 0$, showing that $\mathrm{SL}_2(\mathbb{Z})k$ is periodic (as illustrated in Figure 1.2, where the usual fundamental domain for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ is used; the details behind this form of illustration may be found in [7, Ch. 9]).

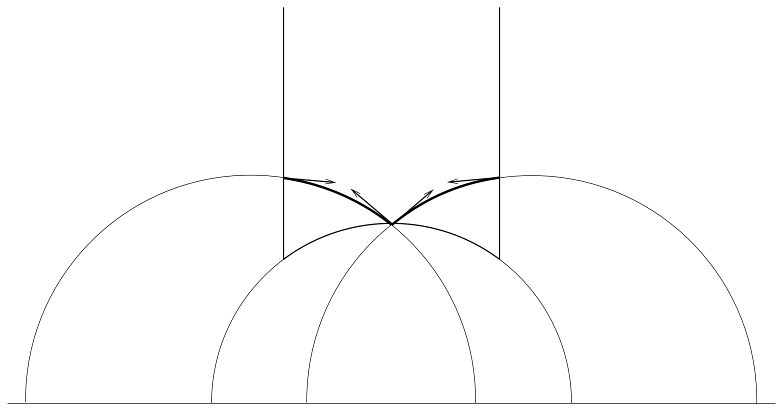

Fig. 1.2. A periodic orbit for the geodesic flow on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{R})$.

Clearly the periodic orbit is itself isomorphic to $\mathbb{R}/t_0\mathbb{Z}$, and the flow on the orbit corresponds under this isomorphism to translation on $\mathbb{R}/t_0\mathbb{Z}$. This gives rise to another type of invariant ergodic probability measure on $\Gamma\backslash\mathrm{SL}_2(\mathbb{R})$, namely the one-dimensional Lebesgue measure supported on a periodic orbit.

Taking convex combinations of m_X and one-dimensional Lebesgue measures on periodic orbits gives rise to many other invariant measures. However, these are not ergodic if they are proper convex combinations. One (of several) definitions of ergodicity for an invariant measure is extremality in the convex set of invariant probability measures. This implies, by Choquet's theorem, that any invariant probability measure on X is a convex combination* of invariant ergodic probability measures. Hence we would like to know if m_X as above and the periodic one-dimensional Lebesgue measures on periodic orbits are the only invariant ergodic probability measures for the geodesic flow. This turns out to be very far from the truth; indeed for every $d \in [1,3]$ there are many invariant ergodic probability measures for which the support† of the measure has Hausdorff dimension d.

We speak of *rigidity of invariant measures* for some group action if it is possible to give a complete classification of the invariant probability measures, and if the ergodic measures show a rich algebraic structure. This is, by the discussion, manifestly not the case for the geodesic flow. However, as the following theorem due to Lindenstrauss [10] shows, it is possible to give some

* This convex combination is really an integral over an entire probability space of ergodic measures; see [7, Ch. 6] for a detailed treatment.

† The support of a measure μ is the smallest closed set A with $\mu(A) = 1$.

(mild, and often checkable) additional conditions that characterize the Haar measure m_X. This theorem built on earlier work of Rudolph [13], Host [8] and others on the unpublished conjecture of Furstenberg concerning measures on $\mathbb{R}/\mathbb{Z}$ invariant under $x \mapsto 2x \pmod 1$ and $x \mapsto 3x \pmod 1$, and of Katok and Spatzier [9] and of Einsiedler and Katok [3] on invariant measures for higher-rank diagonalizable flows in the direction of conjectures of Furstenberg, Katok and Spatzier, and Margulis. More surprising is that in [10] ideas from Ratner's work [11] on unipotent flows were also used, an unexpected connection because on the face of it the measures have very little structure with respect to the unipotent horocycle flow.

Theorem 1.3 (Lindenstrauss). *Let Γ be a congruence lattice over $\mathbb{Q}$, let $X = \Gamma \backslash \mathrm{SL}_2(\mathbb{R})$ and let μ be an probability measure satisfying the following properties:*

[I] μ *is* invariant *under the geodesic flow,*
$[\mathsf{R}]_p$ μ *is* Hecke-p-recurrent *for a prime p, and*
[E] *the* entropy *of every ergodic component of μ is positive for the geodesic flow.*

Then $\mu = m_X$ is the Haar measure on X.

The conditions are labeled [I] for invariance, $[\mathsf{R}]_p$ for recurrence and [E] for entropy. The method behind the theorem is more general, and has lead to a number of further applications: Einsiedler, Katok and Lindenstrauss [4] applied this to obtain a partial result towards Littlewood's conjecture on simultaneous Diophantine approximation for pairs of real numbers, and Einsiedler, Lindenstrauss, Michel and Venkatesh [5] an application to the distribution of periodic orbits for the full diagonal flow on $\mathrm{SL}_3(\mathbb{Z}) \backslash \mathrm{SL}_3(\mathbb{R})$.

Finally, we note that conjecturally invariance [I] and recurrence $[\mathsf{R}]_p$ (for all primes p) should be sufficient to obtain the conclusion of Theorem 1.3. However, this is out of reach with current techniques in ergodic theory.

1.3 Outline of course

For us, Theorem 1.3 will be used as a black box; we refer to the lecture notes of the Pisa Summer School in the Clay Mathematical Proceedings by Einsiedler and Lindenstrauss [2] for an introduction to the ideas and results needed in the proof. Instead we will focus on explaining the three assumptions in Theorem 1.3, and how they may be proved in order to deduce Theorem 1.2.

For this we will be discussing diverse topics. For the discussion of the microlocal lift which will prove invariance, we will introduce and explain the notion of the universal enveloping algebra of the Lie algebra of $\mathrm{SL}_2(\mathbb{R})$. For establishing recurrence and positive entropy, we have to introduce the Hecke operators as the analogue of the Laplace operator on regular trees.

1.4 Problems

- Show that the Casimir operator ω generates the center of the universal enveloping algebra.
- Show that on a compact quotient $M = \Gamma\backslash\mathbb{H}$ there is a sequence of smooth Maass forms that form an orthonormal eigenbasis of $L^2(M)$.
- Show that on $M = \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ there are infinitely many Maass cusp forms.
- Show that equidistribution of large closed orbits of the subgroup $\{(g, g) : g \in \mathrm{SL}_2(\mathbb{R})\} \subseteq \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ in $\big(\mathrm{SL}_2(\mathbb{Z})\backslash \mathrm{SL}_2(\mathbb{R})\big)^2$ implies that (with the normalization discussed in the lectures) the eigenvalues of the Hecke operators T_n go to zero as $n \to \infty$.

References

1. Y. Colin de Verdière, 'Ergodicité et fonctions propres du laplacien', *Comm. Math. Phys.* **102** (1985), no. 3, 497–502. `http://projecteuclid.org/euclid.cmp/1104114465`.
2. M. Einsiedler and E. Lindenstrauss, *Diagonal actions on locally homogeneous spaces*, in *Homogeneous flows, moduli spaces and arithmetic*, in *Clay Math. Proc.* (Amer. Math. Soc., Providence, RI, 2010).
3. M. Einsiedler and A. Katok, 'Rigidity of measures—the high entropy case and non-commuting foliations', *Israel J. Math.* **148** (2005), 169–238. `http://dx.doi.org/10.1007/BF02775436`.
4. M. Einsiedler, A. Katok, and E. Lindenstrauss, 'Invariant measures and the set of exceptions to Littlewood's conjecture', *Ann. of Math. (2)* **164** (2006), no. 2, 513–560. `http://dx.doi.org/10.4007/annals.2006.164.513`.
5. M. Einsiedler, E. Lindenstrauss, P. Michel, and A. Venkatesh, *Distribution of periodic orbits on homogeneous spaces*. `http://arxiv.org/abs/math.DS/0607815`.
6. M. Einsiedler, E. Lindenstrauss, and T. Ward, *Entropy in ergodic theory and homogeneous dynamics*. `www.mth.uea.ac.uk/entropy`.
7. M. Einsiedler and T. Ward, *Ergodic theory with a view toward number theory*, in *Graduate Texts in Mathematics* (Springer-Verlag, New York, 2010). `www.mth.uea.ac.uk/ergodic`.
8. B. Host, 'Nombres normaux, entropie, translations', *Israel J. Math.* **91** (1995), no. 1-3, 419–428.
9. A. Katok and R. J. Spatzier, 'Invariant measures for higher-rank hyperbolic abelian actions', *Ergodic Theory Dynam. Systems* **16** (1996), no. 4, 751–778.
10. E. Lindenstrauss, 'Invariant measures and arithmetic quantum unique ergodicity', *Ann. of Math. (2)* **163** (2006), no. 1, 165–219. `http://dx.doi.org/10.4007/annals.2006.163.165`.
11. M. Ratner, 'Horocycle flows, joinings and rigidity of products', *Ann. of Math. (2)* **118** (1983), no. 2, 277–313.
12. Z. Rudnick and P. Sarnak, 'The behaviour of eigenstates of arithmetic hyperbolic manifolds', *Comm. Math. Phys.* **161** (1994), no. 1, 195–213. `http://projecteuclid.org/euclid.cmp/1104269797`.
13. D. J. Rudolph, '$\times 2$ and $\times 3$ invariant measures and entropy', *Ergodic Theory Dynam. Systems* **10** (1990), no. 2, 395–406.

14. A. I. Šnirel′man, 'Ergodic properties of eigenfunctions', *Uspehi Mat. Nauk* **29** (1974), no. 6(180), 181–182. http://mi.mathnet.ru/eng/umn4463.
15. K. Soundarajan, 'Quantum unique ergodicity for $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$', *Ann. of Math. (2)*. To appear.
16. T. Watson, 'Rankin triple products and quantum chaos', *Ann. of Math. (2)*. http://arxiv.org/abs/0810.0425. to appear.
17. S. Zelditch, 'Uniform distribution of eigenfunctions on compact hyperbolic surfaces', *Duke Math. J.* **55** (1987), no. 4, 919–941. http://dx.doi.org/10.1215/S0012-7094-87-05546-3.