ARITHMETIC QUANTUM CHAOS

JENS MARKLOF

1. Introduction

The central objective in the study of *quantum chaos* is to characterize universal properties of quantum systems that reflect the regular or chaotic features of the underlying classical dynamics. Most developments of the past 25 years have been influenced by the pioneering models on statistical properties of eigenstates (Berry 1977) and energy levels (Berry and Tabor 1977; Bohigas, Giannoni and Schmit 1984). *Arithmetic quantum chaos* (AQC) refers to the investigation of quantum system with additional arithmetic structures that allow a significantly more extensive analysis than is generally possible. On the other hand, the special number-theoretic features also render these systems non-generic, and thus some of the expected universal phenomena fail to emerge. Important examples of such systems include the modular surface and linear automorphisms of tori ('cat maps') which will be described below.

The geodesic motion of a point particle on a compact Riemannian surface $\mathcal{M}$ of constant negative curvature is the prime example of an Anosov flow, one of the strongest characterizations of dynamical chaos. The corresponding quantum eigenstates φ_j and energy levels λ_j are given by the solution of the eigenvalue problem for the Laplace-Beltrami operator Δ (or *Laplacian* for short)

$$(1) \qquad (\Delta + \lambda)\varphi = 0, \quad \|\varphi\|_{\mathrm{L}^2(\mathcal{M})} = 1,$$

where the eigenvalues

$$(2) \qquad \lambda_0 = 0 < \lambda_1 \leq \lambda_2 \leq \cdots \to \infty$$

form a discrete spectrum with an asymptotic density governed by Weyl's law

$$(3) \qquad \#\{j : \lambda_j \leq \lambda\} \sim \frac{\mathrm{Area}(\Gamma\backslash\mathbb{H})}{4\pi}\lambda, \qquad \lambda \to \infty.$$

We rescale the sequence by setting

$$(4) \qquad X_j = \frac{\mathrm{Area}(\Gamma\backslash\mathbb{H})}{4\pi}\lambda_j$$

which yields a sequence of asymptotic density one. One of the central conjectures in AQC says that, if $\mathcal{M}$ is an *arithmetic* hyperbolic surface (see Sec. 2 for examples of this very special class of surfaces of constant negative curvature), the *eigenvalues* of the Laplacian have the same
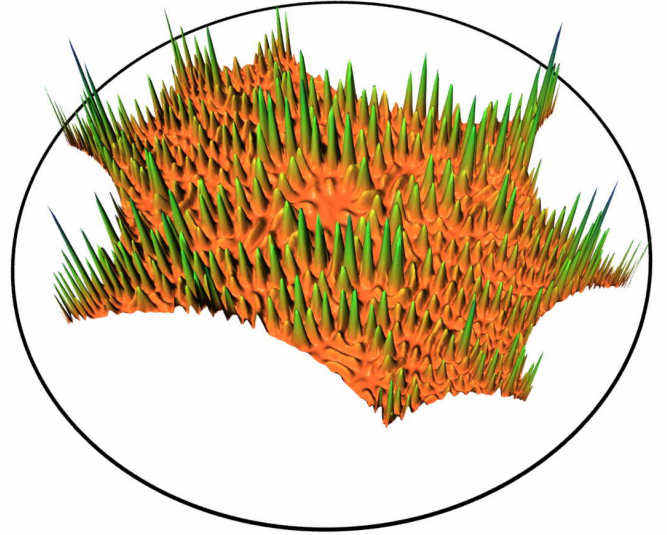
FIGURE 1. Image of the absolute value squared of an eigenfunction $\varphi_j(z)$ for a non-arithmetic surface of genus two. The surface is obtained by identifying opposite sides of the fundamental region [1].

local statistical properties as independent random variables from a Poisson process, see e.g. the surveys [12, 3]. This means that the probability of finding k eigenvalues X_j in randomly shifted interval $[X, X + L]$ of fixed length L is distributed according to the Poisson law $L^k \mathrm{e}^{-L}/k!$. The gaps between eigenvalues have an exponential distribution,

$$(5) \qquad \frac{1}{N}\#\{j \leq N : X_{j+1} - X_j \in [a, b]\} \to \int_a^b \mathrm{e}^{-s}\, ds$$

as $N \to \infty$, and thus eigenvalues are likely to appear in clusters. This is in contrast to the general expectation that the energy level statistics of *generic* chaotic systems follow the distributions of random matrix ensembles; Poisson statistics are usually associated with quantized integrable systems. Although we are at present far from a proof of (5), the deviation from random matrix theory is well understood, see Sec. 3.

Highly excited quantum eigenstates φ_j ($j \to \infty$), cf. Fig. 1, of chaotic systems are conjectured to behave locally like random wave solutions of (1), where boundary conditions are ignored. This hypothesis was put forward by Berry in 1977 and tested numerically, e.g., in the case of certain arithmetic and non-arithmetic surfaces of constant negative curvature [6, 1]. One of the implications is that eigenstates should have uniform mass on the surface $\mathcal{M}$, i.e., for any bounded continuous function $g : \mathcal{M} \to \mathbb{R}$

$$(6) \qquad \int_{\mathcal{M}} |\varphi_j|^2 g\, dA \to \int_{\mathcal{M}} g\, dA, \qquad j \to \infty,$$

where dA is the Riemannian area element on $\mathcal{M}$. This phenomenon, referred to as *quantum unique ergodicity* (QUE), is expected to hold for general surfaces of negative curvature, according to a conjecture by Rudnick and Sarnak (1994). In the case of arithmetic hyperbolic surfaces, there has been substantial progress on this conjecture in the works of Lindenstrauss, Watson and Luo-Sarnak, see Secs. 5, 6 and 7, as well as the review [13]. For general manifolds with ergodic geodesic flow, the convergence in (6) is so far established only for subsequences of eigenfunctions of density one (Schnirelman-Zelditch-Colin de Verdière Theorem, cf. [14]), and it cannot be ruled out that exceptional subsequences of eigenfunctions have singular limit, e.g., localized on closed geodesics. Such 'scarring' of eigenfunctions, at least in some weak form, has been suggested by numerical experiments in Euclidean domains, and the existence of singular quantum limits is a matter of controversy in the current physics and mathematics literature. A first rigorous proof of the existence of scarred eigenstates has recently been established in the case of quantized toral automorphisms. Remarkably, these quantum cat maps may also exhibit quantum unique ergodicity. A more detailed account of results for these maps is given in Sec. 8, see also [9, 5].

There have been a number of other fruitful interactions between quantum chaos and number theory, in particular the connections of spectral statistics of integrable quantum systems with the value distribution properties of quadratic forms, and analogies in the statistical behaviour of energy levels of chaotic systems and the zeros of the Riemann zeta function. We refer the reader to [8] and [2], respectively, for information on these topics.

2. Hyperbolic surfaces

Let us begin with some basic notions of hyperbolic geometry. The hyperbolic plane $\mathbb{H}$ may be abstractly defined as the simply connected two-dimensional Riemannian manifold with Gaussian curvature -1. A convenient parametrization of $\mathbb{H}$ is provided by the complex upper half plane, $\mathfrak{H} = \{x + \mathrm{i}y : x \in \mathbb{R}, \ y > 0\}$, with Riemannian line and volume elements

$$(7) \qquad ds^2 = \frac{dx^2 + dy^2}{y^2}, \qquad dA = \frac{dx\,dy}{y^2},$$

respectively. The group of orientation-preserving isometries of $\mathbb{H}$ is given by fractional linear transformations

$$(8) \quad \mathfrak{H} \to \mathfrak{H}, \qquad z \mapsto \frac{az + b}{cz + d}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{R}),$$

where $\mathrm{SL}(2, \mathbb{R})$ is the group of 2×2 matrices with unit determinant. Since the matrices 1 and -1 represent the same transformation, the group of orientation-preserving isometries can be identified with $\mathrm{PSL}(2, \mathbb{R}) := \mathrm{SL}(2, \mathbb{R})/\{\pm 1\}$. A finite-volume hyperbolic surface may now be represented as the quotient $\Gamma \backslash \mathbb{H}$, where $\Gamma \subset \mathrm{PSL}(2, \mathbb{R})$ is a Fuchsian group of the first kind. An *arithmetic* hyperbolic surface
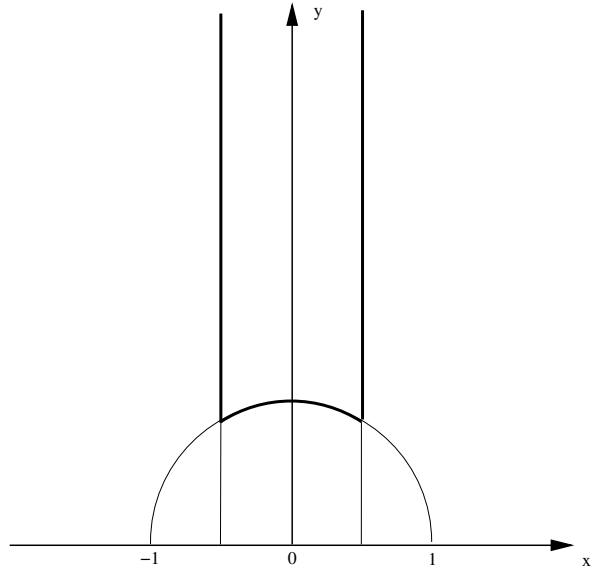

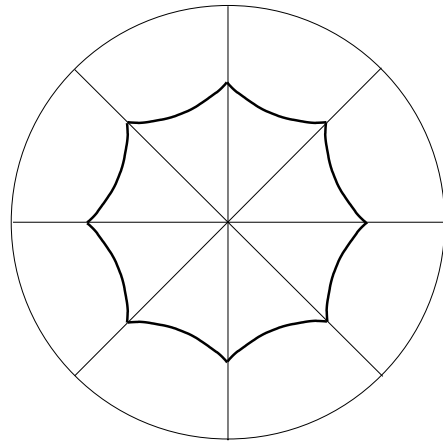FIGURE 2. Fundamental domain of the modular group $\mathrm{PSL}(2, \mathbb{Z})$ in the complex upper half plane.

FIGURE 3. Fundamental domain of the regular octagon in the Poincarè disc.

(such as the modular surface) is obtained, if Γ has, loosely speaking, some representation in $n \times n$ matrices with integer coefficients, for some suitable n. This is evident in the case of the modular surface, where the fundamental group is the *modular group* $\Gamma = \mathrm{PSL}(2, \mathbb{Z}) = \{\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{PSL}(2, \mathbb{R}) : a, b, c, d \in \mathbb{Z}\}/\{\pm 1\}$.

A fundamental domain for the action of the modular group $\mathrm{PSL}(2, \mathbb{Z})$ on $\mathfrak{H}$ is the set

$$(9) \qquad \mathcal{F}_{\mathrm{PSL}(2,\mathbb{Z})} = \left\{ z \in \mathfrak{H} : |z| > 1, -\frac{1}{2} < \mathrm{Re}\, z < \frac{1}{2} \right\},$$

see fig. 2. The modular group is generated by the translation $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} : z \mapsto z + 1$ and the inversion $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} : z \mapsto -1/z$. These generators identify sections of the boundary of $\mathcal{F}_{\mathrm{PSL}(2,\mathbb{Z})}$. By gluing the fundamental domain along identified edges we obtain a realization of the modular surface, a non-compact surface with one cusp at $z \to \infty$, and two conic singularities at $z = \mathrm{i}$ and $z = \frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$.

An interesting example of a *compact* arithmetic surface is the 'regular octagon', a hyperbolic surface of genus two. Its fundamental domain is shown in fig. 3 as a subset of the Poincarè disc $\mathfrak{D} = \{z \in \mathbb{C} : |z| < 1\}$, which yields an alternative parametrization of the hyperbolic plane $\mathbb{H}$. In these coordinates, the Riemannian line and volume element read

$$(10) \quad ds^2 = \frac{4(dx^2 + dy^2)}{(1 - x^2 - y^2)^2}, \qquad dA = \frac{4 \, dx \, dy}{(1 - x^2 - y^2)^2}.$$

The group of orientation-preserving isometries is now represented by $\mathrm{PSU}(1,1) = \mathrm{SU}(1,1)/\{\pm 1\}$, where

$$(11) \quad \mathrm{SU}(1,1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \bar{\beta} & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C}, \; |\alpha|^2 - |\beta|^2 = 1 \right\},$$

acting on $\mathfrak{D}$ as above via fractional linear transformations. The fundamental group of the regular octagon surface is the subgroup of all elements in $\mathrm{PSU}(1,1)$ with coefficients of the form

$$(12) \quad \alpha = k + l\sqrt{2}, \qquad \beta = \left(m + n\sqrt{2}\right)\sqrt{1 + \sqrt{2}},$$

where $k, l, m, n \in \mathbb{Z}[\mathrm{i}]$, i.e., Gaussian integers of the form $k_1 + \mathrm{i}k_2$, $k_1, k_2 \in \mathbb{Z}$. Note that not all choices of $k, l, m, n \in \mathbb{Z}[\mathrm{i}]$ satisfy the condition $|\alpha|^2 - |\beta|^2 = 1$. Since all elements $\gamma \neq 1$ of Γ act fix-point free on $\mathbb{H}$, the surface $\Gamma \backslash \mathbb{H}$ is smooth without conic singularities.

In the following we will restrict our attention to a representative case, the modular surface with $\Gamma = \mathrm{PSL}(2, \mathbb{Z})$.

3. EIGENVALUE STATISTICS AND SELBERG TRACE FORMULA

The statistical properties of the rescaled eigenvalues X_j (cf. (4)) of the Laplacian can be characterized by their distribution in small intervals

$$(13) \quad \mathcal{N}(x, L) := \# \{j : x \leq X_j \leq x + L\},$$

where x is uniformly distributed, say, in the interval $[X, 2X]$, X large. Numerical experiments by Bogomolny, Georgeot, Giannoni and Schmit, as well as Bolte, Steil and Steiner (see refs. in [3]) suggest that the X_j are asymptotically Poisson distributed:

Conjecture 1. *For any bounded function $g : \mathbb{Z}_{\geq 0} \to \mathbb{C}$ we have*

$$(14) \quad \frac{1}{X} \int_X^{2X} g\big(\mathcal{N}(x, L)\big) \, dx \to \sum_{k=0}^{\infty} g(k) \frac{L^k e^{-L}}{k!},$$

as $T \to \infty$.

One may also consider larger intervals, where $L \to \infty$ as $X \to \infty$. In this case the assumption on the independence of the X_j predicts a central limit theorem. Weyl's law (3) implies that the expectation value is asymptotically, for $T \to \infty$,

$$(15) \quad \frac{1}{X} \int_X^{2X} \mathcal{N}(x, L) \, dx \sim L.$$

This asymptotics holds for any sequence of L bounded away from zero (e.g. L constant, or $L \to \infty$).

Define the variance by

$$(16) \quad \Sigma^2(X, L) = \frac{1}{X} \int_X^{2X} \big(\mathcal{N}(x, L) - L\big)^2 \, dx.$$

In view of the above conjecture, one expects $\Sigma^2(X, L) \sim L$ in the limit $X \to \infty$, $L/\sqrt{X} \to 0$ (the variance exhibits a less universal behaviour in the range $L \gg \sqrt{X}$,[1] cf. [12]), and a central limit theorem for the fluctuations around the mean:

Conjecture 2. *For any bounded function $g : \mathbb{R} \to \mathbb{C}$ we have*
$$(17)$$
$$\frac{1}{X} \int_X^{2X} g\left(\frac{\mathcal{N}(x, L) - L}{\sqrt{\Sigma^2(x, L)}}\right) dx \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-\frac{1}{2}t^2} dt,$$

as $X, L \to \infty$, $L \ll X$.

The main tool in the attempts to prove the above conjectures has been the Selberg trace formula. It relates sums over eigenvalues of the Laplacians to sums over lengths of closed geodesics on the hyperbolic surface. The trace formula is in its simplest form in the case of compact hyperbolic surfaces; we have

$$(18) \quad \sum_{j=0}^{\infty} h(\rho_j) = \frac{\mathrm{Area}(\mathcal{M})}{4\pi} \int_{-\infty}^{\infty} h(\rho) \tanh(\pi\rho) \, \rho \, d\rho$$

$$+ \sum_{\gamma \in H_*} \sum_{n=1}^{\infty} \frac{\ell_\gamma \, g(n\ell_\gamma)}{2 \sinh(n\ell_\gamma/2)},$$

where H_* is the set of all primitive oriented closed geodesics γ, and ℓ_γ their lengths. The quantity ρ_j is related to the eigenvalue λ_j by the equation $\lambda_j = \rho_j^2 + \frac{1}{4}$. The trace formula (18) holds for a large class of even test functions h. E.g. it is sufficient to assume that h is infinitely differentiable, and that the Fourier transform of h,

$$(19) \quad g(t) = \frac{1}{2\pi} \int_{\mathbb{R}} h(\rho) \, e^{-\mathrm{i}\rho t} d\rho,$$

has compact support. The trace formula for non-compact surfaces has additional terms from the parabolic elements in the corresponding group, and includes also sums over the resonances of the continuous part of the spectrum. The non-compact modular surface behaves in many ways like a compact surface. In particular, Selberg showed that the number of eigenvalues embedded in the continuous spectrum satisfies the same Weyl law as in the compact case [13].

Setting

$$(20) \quad h(\rho) = \chi_{[X, X+L]}\big(\tfrac{\mathrm{Area}(\mathcal{M})}{4\pi}(\rho^2 + \tfrac{1}{4})\big),$$

where $\chi_{[X, X+L]}$ is the characteristic function of the interval $[X, X+L]$, we may thus view $\mathcal{N}(X, L)$ as the left-hand side

---

[1]The notation $A \ll B$ means there is a constant $c > 0$ such that $A \leq cB$

of the trace formula. The above test function h is however not admissible, and requires appropriate smoothing. Luo and Sarnak (cf. [13]) developed an argument of this type to obtain a lower bound on the average number variance,

$$(21) \qquad \frac{1}{L} \int_0^L \Sigma^2(X, L') dL' \gg \frac{\sqrt{X}}{(\log X)^2}$$

in the regime $\sqrt{X}/\log X \ll L \ll \sqrt{X}$, which is consistent with the Poisson conjecture $\Sigma^2(X, L) \sim L$. Bogomolny, Levyraz and Schmit suggested a remarkable limiting formula for the two-point correlation function for the modular surface (cf. [3, 4]), based on an analysis of the correlations between multiplicities of lengths of closed geodesics. A rigorous analysis of the fluctuations of multiplicities is given by Peter, cf. [4]. Rudnick [10] has recently established a smoothed version of Conjecture 2 in the regime

$$(22) \qquad \frac{\sqrt{X}}{L} \to \infty, \quad \frac{\sqrt{X}}{L \log X} \to 0,$$

where the characteristic function in (20) is replaced by a certain class of smooth test functions.

All of the above approaches use the Selberg trace formula, exploiting the particular properties of the distribution of lengths of closed geodesics in arithmetic hyperbolic surfaces. These will be discussed in more detail in the next section, following the work of Bogomolny, Georgeot, Giannoni and Schmit, Bolte, and Luo and Sarnak, see [3, 12] for references.

4. DISTRIBUTION OF LENGTHS OF CLOSED GEODESICS

The classical prime geodesic theorem asserts that the number $N(\ell)$ of primitive closed geodesics of length less than ℓ is asymptotically

$$(23) \qquad N(\ell) \sim \frac{e^\ell}{\ell}.$$

One of the significant geometrical characteristics of arithmetic hyperbolic surfaces is that the number of closed geodesics with the same length ℓ grows exponentially with ℓ. This phenomenon is easiest explained in the case of the modular surface, where the set of lengths ℓ appearing in the lengths spectrum is characterized by the condition

$$(24) \qquad 2\cosh(\ell/2) = |\operatorname{tr}\gamma|$$

where γ runs over all elements in $\operatorname{SL}(2, \mathbb{Z})$ with $|\operatorname{tr}\gamma| > 2$. It is not hard to see that any integer $n > 2$ appears in the set $\{|\operatorname{tr}\gamma| : \gamma \in \operatorname{SL}(2, \mathbb{Z})\}$, and hence the set of distinct lengths of closed geodesics is

$$(25) \qquad \mathcal{L} = \{2\operatorname{arcosh}(n/2) : n = 3, 4, 5, \ldots\}.$$

The number of distinct lengths less than ℓ is therefore asymptotically for large ℓ

$$(26) \qquad N'(\ell) = \#(\mathcal{L} \cap [0, \ell]) \sim e^{\ell/2}.$$

Eqs. (26) and (23) say that *on average* the number of geodesics with the same lengths is at least $\asymp e^{\ell/2}/\ell$.

The prime geodesic theorem (23) holds equally for all hyperbolic surfaces with finite area, while (26) is specific to the modular surface. For general arithmetic surfaces, we have the upper bound

$$(27) \qquad N'(\ell) \le c e^{\ell/2}$$

for some constant $c > 0$ that may depend on the surface. Although one expects $N'(\ell)$ to be asymptotic to $\frac{1}{2}N(\ell)$ for generic surfaces (since most geodesics have a time reversal partner which thus has the same length, and otherwise all lengths are distinct), there are examples of non-arithmetic Hecke triangles where numerical and heuristic arguments suggest $N'(\ell) \sim c_1 e^{c_2 \ell}/\ell$ for suitable constants $c_1 > 0$ and $0 < c_2 < 1/2$, cf. [4]. Hence exponential degeneracy in the length spectrum seems to occur in a weaker form also for non-arithmetic surfaces.

A further useful property of the length spectrum of arithmetic surfaces is the *bounded clustering property*: there is a constant C (again surface dependent) such that

$$(28) \qquad \#(\mathcal{L} \cap [\ell, \ell+1]) \le C$$

for all ℓ. This fact is evident in the case of the modular surface; the general case is proved by Luo and Sarnak, cf. [12].

5. QUANTUM UNIQUE ERGODICITY

The unit tangent bundle of a hyperbolic surface $\Gamma\backslash\mathbb{H}$ describes the physical phase space on which the classical dynamics takes place. A convenient parametrization of the unit tangent bundle is given by the quotient $\Gamma\backslash\operatorname{PSL}(2, \mathbb{R})$—this may be seen be means of the Iwasawa decomposition for an element $g \in \operatorname{PSL}(2, \mathbb{R})$,

$$(29) \quad g = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} \begin{pmatrix} \cos\theta/2 & \sin\theta/2 \\ -\sin\theta/2 & \cos\theta/2 \end{pmatrix},$$

where $x + iy \in \mathfrak{H}$ represents the position of the particle in $\Gamma\backslash\mathbb{H}$ in half-plane coordinates, and $\theta \in [0, 2\pi)$ the direction of its velocity. Multiplying the matrix (29) from the left by $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and writing the result again in the Iwasawa form (29), one obtains the action

$$(30) \qquad (z, \phi) \mapsto \left(\frac{az+b}{cz+d}, \theta - 2\arg(cz+d) \right),$$

which represents precisely the geometric action of isometries on the unit tangent bundle.

The geodesic flow Φ^t on $\Gamma\backslash\operatorname{PSL}(2, \mathbb{R})$ is represented by the right translation

$$(31) \qquad \Phi^t : \Gamma g \mapsto \Gamma g \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}.$$

The Haar measure μ on $\operatorname{PSL}(2, \mathbb{R})$ is thus trivially invariant under the geodesic flow. It is well known that μ is not the only invariant measure, i.e., Φ^t is not uniquely ergodic, and that there is in fact an abundance of invariant measures. The simplest examples are those with uniform mass on one, or a countable collection of, closed geodesics.

To test the distribution of an eigenfunction φ_j in phase space, one associates with a function $a \in \mathrm{C}^\infty(\Gamma\backslash\mathrm{PSL}(2,\mathbb{R}))$ the quantum observable $\mathrm{Op}(a)$, a zeroth order pseudo-differential operator with principal symbol a. Using semi-classical techniques based on Friedrichs symmetrization, on can show that the matrix element

$$(32) \qquad \nu_j(a) = \langle \mathrm{Op}(a)\varphi_j, \varphi_j \rangle$$

is asymptotic (as $j \to \infty$) to a positive functional that defines a probability measure on $\Gamma\backslash\mathrm{PSL}(2,\mathbb{R})$. Therefore, if $\mathcal{M}$ is compact, any weak limit of ν_j represents a probability measure on $\Gamma\backslash\mathrm{PSL}(2,\mathbb{R})$. Egorov's theorem (cf. [14]) in turn implies that any such limit must be invariant under the geodesic flow, and the main challenge in proving quantum unique ergodicity is to rule out all invariant measures apart from Haar.

Conjecture 3 (Rudnick and Sarnak (1994) [12, 13]). *For every compact hyperbolic surface $\Gamma\backslash\mathbb{H}$, the sequence ν_j converges weakly to μ.*

Lindenstrauss has proved this conjecture for compact arithmetic hyperbolic surfaces of congruence type (such as the second example in Sec. 2) for special bases of eigenfunctions, using ergodic-theoretic methods. These will be discussed in more detail in Sec. 6 below. His results extend to the non-compact case, i.e., to the modular surface where $\Gamma = \mathrm{PSL}(2,\mathbb{Z})$. Here he shows that any weak limit of subsequences of ν_j is of the form $c\mu$, where c is a constant with values in $[0,1]$. One believes that $c = 1$, but with present techniques it cannot be ruled out that a proportion of the mass of the eigenfunction escapes into the non-compact cusp of the surface. For the modular surface, $c = 1$ can be proved under the assumption of the generalized Riemann hypothesis, see Sec. 7 and [13]. Quantum unique ergodicity also holds for the continuous part of the spectrum, which is furnished by the Eisenstein series $E(z,s)$, where $s = 1/2 + ir$ is the spectral parameter. Note that the measures associated with the matrix elements

$$(33) \qquad \nu_r(a) = \langle \mathrm{Op}(a)E(\,\cdot\,, 1/2 + ir), E(\,\cdot\,, 1/2 + ir) \rangle$$

are not probability but only Radon measures, since $E(z,s)$ is not square-integrable. Luo and Sarnak, and Jakobson have shown that

$$(34) \qquad \lim_{r\to\infty} \frac{\nu_r(a)}{\nu_r(b)} = \frac{\mu(a)}{\mu(b)}$$

for suitable test functions $a, b \in \mathrm{C}^\infty(\Gamma\backslash\mathrm{PSL}(2,\mathbb{R}))$, cf. [13].

6. Hecke operators, entropy and measure rigidity

For compact surfaces, the sequence of probability measures approaching the matrix elements ν_j is relatively compact. That is, every infinite sequence contains a convergent subsequence. Lindenstrauss' central idea in the proof of quantum unique ergodicity is to exploit the presence of Hecke operators to understand the invariance properties

of possible quantum limits. We will sketch his argument in the case of the modular surface (ignoring issues related to the non-compactness of the surface), where it is most transparent.

For every positive integer n, the Hecke operator T_n acting on continuous functions on $\Gamma\backslash\mathbb{H}$ with $\Gamma = \mathrm{SL}(2,\mathbb{Z})$ is defined by

$$(35) \qquad T_n f(z) = \frac{1}{\sqrt{n}} \sum_{\substack{a,d=1 \\ ad=n}}^{n} \sum_{b=0}^{d-1} f\left(\frac{az+b}{d}\right).$$

The set M_n of matrices with integer coefficients and determinant n can be expressed as the disjoint union

$$(36) \qquad M_n = \bigcup_{\substack{a,d=1 \\ ad=n}}^{n} \bigcup_{b=0}^{d-1} \Gamma \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$$

and hence the sum in (35) can be viewed as a sum over the cosets in this decomposition. We note the product formula

$$(37) \qquad T_m T_n = \sum_{d\,|\,\gcd(m,n)} T_{mn/d^2}.$$

The Hecke operators are normal, form a commuting family, and in addition commute with the Laplacian Δ. In the following we consider an orthonormal basis of eigenfunctions φ_j of Δ that are simultaneously eigenfunctions of all Hecke operators. We will refer to such eigenfunctions as *Hecke eigenfunctions*. The above assumption is automatically satisfied, if the spectrum of Δ is simple (i.e. no eigenvalues coincide), a property conjectured by Cartier and supported by numerical computations. Lindenstrauss' work is based on the following two observations. Firstly, all quantum limits of Hecke eigenfunctions are geodesic-flow invariant measures of positive entropy, and secondly, the only such measure of positive entropy that is recurrent under Hecke correspondences is Lebesgue measure.

The first property is proved by Bourgain and Lindenstrauss (2003) and refines arguments of Rudnick and Sarnak (1994) and Wolpert (2001) on the distribution of Hecke points (see [13] for references to these papers). For a given point $z \in \mathbb{H}$ the set of Hecke points is defined as

$$(38) \qquad T_n(z) := M_n z.$$

For most primes, the set $T_{p^k}(z)$ comprises $(p+1)p^{k-1}$ distinct points on $\Gamma\backslash\mathbb{H}$. For each z, the Hecke operator T_n may now be interpreted as the adjacency matrix for a finite graph embedded in $\Gamma\backslash\mathbb{H}$, whose vertices are the Hecke points $T_n(z)$. Hecke eigenfunctions φ_j with

$$(39) \qquad T_n \varphi_j = \lambda_j(n)\varphi_j$$

give rise to eigenfunctions of the adjacency matrix. Exploiting this fact, Bourgain and Lindenstrauss show that for a large set of integers n

$$(40) \qquad |\varphi_j(z)|^2 \ll \sum_{w\in T_n(z)} |\varphi_j(w)|^2,$$

i.e. pointwise values of $|\varphi_j|^2$ cannot be substantially larger than its sum over Hecke points. This and the observation that Hecke points for a large set of integers n are sufficiently uniformly distributed on $\Gamma\backslash\mathbb{H}$ as $n \to \infty$, yields the estimate of positive entropy with a quantitative lower bound.

Lindenstrauss' proof of the second property, which shows that Lebesgue measure is the only quantum limit of Hecke eigenfunctions, is a result of a currently very active branch of ergodic theory: measure rigidity. Invariance under the geodesic flow alone is not sufficient to rule out other possible limit measures. In fact there are uncountably many measures with this property. As limits of Hecke eigenfunctions, all quantum limits possess an additional property, namely recurrence under Hecke correspondences. Since the explanation of these is rather involved, let us recall an analogous result in a simpler set-up. The map $\times 2 : x \mapsto 2x \bmod 1$ defines a hyperbolic dynamical system on the unit circle with a wealth of invariant measures, similar to the case of the geodesic flow on a surface of negative curvature. Furstenberg conjectured that, up to trivial invariant measures that are localized on finitely many rational points, Lebesgue measure is the only $\times 2$-invariant measure that is also invariant under action of $\times 3 : x \mapsto 3x \bmod 1$. This fundamental problem is still unsolved and one of the central conjectures in measure rigidity. Rudolph however showed that Furstenberg's conjecture is true if one restricts the statement to $\times 2$-invariant measures of positive entropy, cf. [11]. In Lindenstrauss' work, $\times 2$ plays the role of the geodesic flow, and $\times 3$ the role of the Hecke correspondences. Although it might also here be interesting to ask whether and analogue of Furstenberg's conjecture holds, it is inessential for the proof of QUE due to the positive entropy of quantum limits discussed in the previous paragraph.

7. EIGENFUNCTIONS AND L-FUNCTIONS

An even eigenfunction $\varphi_j(z)$ for $\Gamma = \mathrm{SL}(2, \mathbb{Z})$ has the Fourier expansion

$$(41) \qquad \varphi_j(z) = \sum_{n=1}^{\infty} a_j(n) y^{1/2} K_{i\rho_j}(2\pi n y) \cos(2\pi n x).$$

We associate with $\varphi_j(z)$ the Dirichlet series

$$(42) \qquad L(s, \varphi_j) = \sum_{n=1}^{\infty} a_j(n) n^{-s}$$

which converges for $\mathrm{Re}\, s$ large enough. These series have an analytic continuation to the entire complex plane $\mathbb{C}$ and satisfy a functional equation,

$$(43) \qquad \Lambda(s, \varphi_j) = \Lambda(1 - s, \varphi_j)$$

where

$$(44) \qquad \Lambda(s, \varphi_j) = \pi^{-s} \Gamma\left(\tfrac{s + i\rho_j}{2}\right) \Gamma\left(\tfrac{s - i\rho_j}{2}\right) L(s, \varphi_j).$$

If $\varphi_j(z)$ is in addition an eigenfunction of all Hecke operators, then the Fourier coefficients in fact coincide (up to a normalization constant) with the eigenvalues of the Hecke operators

$$(45) \qquad a_j(m) = \lambda_j(m) a_j(1).$$

If we normalize $a_j(1) = 1$, the Hecke relations (37) result in an Euler product formula for the L-function,

$$(46) \qquad L(s, \varphi_j) = \prod_{p \text{ prime}} \left(1 - a_j(p) p^{-s} + p^{-1-2s}\right)^{-1}.$$

These L-functions behave in many other ways like the Riemann zeta or classical Dirichlet L-functions. In particular, they are expected to satisfy a Riemann hypothesis, i.e. all non-trivial zeros are constrained to the critical line $\mathrm{Im}\, s = 1/2$.

Questions on the distribution of Hecke eigenfunctions, such as quantum unique ergodicity or value distribution properties, can now be translated to analytic properties of L-functions. We will discuss two examples.

The asymptotics in (6) can be established by proving (6) for the choices $g = \varphi_k$, $k = 1, 2, \ldots$, i.e.,

$$(47) \qquad \int_{\mathcal{M}} |\varphi_j|^2 \varphi_k \, dA \to 0.$$

Watson discovered the remarkable relation [13]

$$(48) \qquad \left| \int_{\mathcal{M}} \varphi_{j_1} \varphi_{j_2} \varphi_{j_3} \, dA \right|^2$$
$$= \frac{\pi^4 \Lambda(\tfrac{1}{2}, \varphi_{j_1} \times \varphi_{j_2} \times \varphi_{j_3})}{\Lambda(1, \mathrm{sym}^2 \varphi_{j_1}) \Lambda(1, \mathrm{sym}^2 \varphi_{j_2}) \Lambda(1, \mathrm{sym}^2 \varphi_{j_3})}.$$

The L-functions $\Lambda(s, g)$ in Watson's formula are more advanced cousins of those introduced earlier, see [13] for details. The Riemann hypothesis for such L-functions implies then via (48) a precise rate of convergence to QUE for the modular surface,

$$(49) \qquad \int_{\mathcal{M}} |\varphi_j|^2 g \, dA = \int_{\mathcal{M}} g \, dA + O(\lambda_j^{-1/4+\epsilon}),$$

for any $\epsilon > 0$, where the implied constant depends on ϵ and g.

A second example on the connection between statistical properties of the matrix elements $\nu_j(a) = \langle \mathrm{Op}(a)\varphi_j, \varphi_j \rangle$ (for fixed a and random j) and values L-functions has appeared in the work of Luo and Sarnak, cf. [13]. Define the variance

$$(50) \qquad V_\lambda(a) = \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} |\nu_j(a) - \mu(a)|^2,$$

with $N(\lambda) = \#\{j : \lambda_j \leq \lambda\}$; cf. (3). Following a conjecture by Feingold-Peres and Eckhardt et al. (see [13] for references) for 'generic' quantum chaotic systems, one expects a central limit theorem for the statistical fluctuations of the $\nu_j(a)$, where the normalized variance $N(\lambda)^{1/2} V_\lambda(a)$ is asymptotic to the classical autocorrelation function $C(a)$, see eq. (54).

Conjecture 4. *For any bounded function $g : \mathbb{R} \to \mathbb{C}$ we have*

(51)
$$\frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} g\left(\frac{\nu_j(a) - \mu(a)}{\sqrt{V_\lambda(a)}}\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-\frac{1}{2}t^2} dt,$$

as $\lambda \to \infty$.

Luo and Sarnak prove that in the case of the modular surface the variance has the asymptotics

(52)
$$\lim_{\lambda \to \infty} N(\lambda)^{1/2} V_\lambda(a) = \langle Ba, a \rangle,$$

where B is a non-negative self-adjoint operator which commutes with the Laplacian Δ and all Hecke operators T_n. In particular we have

(53)
$$B\varphi_j = \frac{1}{2} L(\tfrac{1}{2}, \varphi_j) C(\varphi_j) \, \varphi_j,$$

where

(54)
$$C(a) := \int_{\mathbb{R}} \int_{\Gamma \backslash \text{PSL}(2,\mathbb{R})} a\big(\Phi^t(g)\big)\overline{a(g)} d\mu(g) \, dt$$

is the classical autocorrelation function for the geodesic flow with respect to the observable a [13]. Up to the arithmetic factor $\frac{1}{2}L(\tfrac{1}{2}, \varphi_j)$, eq. (53) is consistent with the Feingold-Peres prediction for the variance of generic chaotic systems. Furthermore, recent estimates of moments by Rudnick and Soundararajan (2005) indicate that Conjecture 4 is not valid in the case of the modular surface.

8. Quantum eigenstates of cat maps

Cat maps are probably the simplest area-preserving maps on a compact surface that are highly chaotic. They are defined as linear automorphisms on the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$,

(55)
$$\Phi_A : \mathbb{T}^2 \to \mathbb{T}^2,$$

where a point $\xi \in \mathbb{R}^2$ (mod $\mathbb{Z}^2$) is mapped to $A\xi$ (mod $\mathbb{Z}^2$); A is a fixed matrix in $\text{GL}(2, \mathbb{Z})$ with eigenvalues off the unit circle (this guarantees hyperbolicity). We view the torus $\mathbb{T}^2$ as a symplectic manifold, the 'phase space' of the dynamical system. Since $\mathbb{T}^2$ is compact, the Hilbert space of quantum states is an N dimensional vector space $\mathcal{H}_N$, N integer. The semiclassical limit, or limit of small wavelengths, corresponds here to $N \to \infty$.

It is convenient to identify $\mathcal{H}_N$ with $\text{L}^2(\mathbb{Z}/N\mathbb{Z})$, with inner product

(56)
$$\langle \psi_1, \psi_2 \rangle = \frac{1}{N} \sum_{Q \bmod N} \psi_1(Q)\overline{\psi_2(Q)}.$$

For any smooth function $f \in \text{C}^\infty(\mathbb{T}^2)$, define a *quantum observable*

$$\text{Op}_N(f) = \sum_{n \in \mathbb{Z}^2} \widehat{f}(n) T_N(n)$$

where $\widehat{f}(n)$ are the Fourier coefficients of f, and $T_N(n)$ are translation operators

(57)
$$T_N(n) = e^{\pi i n_1 n_2/N} t_2^{n_2} t_1^{n_1},$$

(58)
$$[t_1\psi](Q) = \psi(Q+1), \quad [t_2\psi](Q) = e^{2\pi i Q/N} \psi(Q).$$

The operators $\text{Op}_N(a)$ are the analogues of the pseudo-differential operators discussed in Sec. 5.

A quantization of Φ_A is a unitary operator $U_N(A)$ on $\text{L}^2(\mathbb{Z}/N\mathbb{Z})$ satisfying the equation

(59)
$$U_N(A)^{-1} \text{Op}_N(f) U_N(A) = \text{Op}_N(f \circ \Phi_A)$$

for all $f \in \text{C}^\infty(\mathbb{T}^2)$. There are explicit formulas for $U_N(A)$ when A is in the group

(60)
$$\Gamma = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}(2, \mathbb{Z}) : ab \equiv cd \equiv 0 \bmod 2 \right\};$$

these may be viewed as analogues of the Shale-Weil or metaplectic representation for SL(2). E.g., the quantization of

(61)
$$A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$$

yields

(62)
$$U_N(A)\psi(Q)$$
$$= N^{-\frac{1}{2}} \sum_{Q' \bmod N} \exp\left[\frac{2\pi i}{N}(Q^2 - QQ' + Q'^2)\right]\psi(Q').$$

In analogy with (1) we are interested in the statistical features of the eigenvalues and eigenfunctions of $U_N(A)$, i.e., the solutions to

(63)
$$U_N(A)\varphi = \lambda\varphi, \quad \|\varphi\|_{\text{L}^2(\mathbb{Z}/N\mathbb{Z})} = 1.$$

Unlike typical quantum-chaotic maps, the statistics of the N eigenvalues

(64)
$$\lambda_{N1}, \lambda_{N2}, \ldots, \lambda_{NN} \in \text{S}^1$$

do not follow the distributions of unitary random matrices in the limit $N \to \infty$, but are rather singular [7]. In analogy with the Selberg trace formula for hyperbolic surfaces (18), there is an exact trace formula relating sums over eigenvalues of $U_N(A)$ with sums over fixed points of the classical map [7].

As in the case of arithmetic surfaces, the *eigenfunctions* of cat maps appear to behave more generically. The analogue of the Schnirelman-Zelditch-Colin de Verdière Theorem states that, for any orthonormal basis of eigenfunctions $\{\varphi_{Nj}\}_{j=1}^N$ we have, for all $f \in \text{C}^\infty(\mathbb{T}^2)$,

(65)
$$\langle \text{Op}(f)\varphi_{Nj}, \varphi_{Nj} \rangle \to \int_{\mathbb{T}^2} f(\xi)d\xi$$

as $N \to \infty$, for all j in an index set J_N of full density, i.e., $\#J_N \sim N$. Kurlberg and Rudnick [9] have characterized special bases of eigenfunctions $\{\varphi_{Nj}\}_{j=1}^N$ (termed *Hecke eigenbases* in analogy with arithmetic surfaces) for which QUE holds, generalizing earlier work of Degli Esposti, Graffi and Isola (1995). That is, (65) holds for *all* $j = 1, \ldots, N$. Rudnick and Kurlberg, and more recently Gurevich and Hadani, have established results on the rate

of convergence analogous to (49). These results are un-conditional. Gurevich and Hadani use methods from algebraic geometry based on those developed by Deligne in his proof of the Weil conjectures (an analogue of the Riemann hypothesis for finite fields).

In the case of quantum cat maps, there are values of N for which the number of coinciding eigenvalues can be large, a major difference to what is expected for the modular surface. Linear combinations of eigenstates with the same eigenvalue are as well eigenstates, and may lead to different quantum limits. Indeed Faure, Nonnenmacher and De Bièvre [5] have shown that there are subsequences of values of N, so that, for all $f \in C^\infty(\mathbb{T}^2)$,

$$(66) \qquad \langle \mathrm{Op}(f)\varphi_{Nj}, \varphi_{Nj} \rangle \to \frac{1}{2} \int_{\mathbb{T}^2} f(\xi)d\xi + \frac{1}{2}f(0),$$

i.e., half of the mass of the quantum limit localizes on the hyperbolic fixed point of the map. This is the first, and to date only, rigorous result concerning the existence of scarred eigenfunctions in systems with chaotic classical limit.

j.marklof@bristol.ac.uk

Further reading

[1] R. Aurich, F. Steiner, Statistical properties of highly excited quantum eigenstates of a strongly chaotic system. Phys. D 64 (1993), no. 1-3, 185–214.

[2] M.V. Berry, J.P. Keating, The Riemann zeros and eigenvalue asymptotics. SIAM Rev. 41 (1999), no. 2, 236–266

[3] E.B. Bogomolny, B. Georgeot, M.-J. Giannoni, C. Schmit, Arithmetical chaos. Phys. Rep. 291 (1997), no. 5-6, 219–324.

[4] E.B. Bogomolny, Quantum and arithmetical chaos, Frontiers in Number Theory, Physics and Geometry, Les Houches 2003, to appear in Springer Lecture Notes.

[5] S. De Bièvre, Recent results on quantum map eigenstates, Proceedings of QMATH9, Giens 2004, to appear.

[6] D.A. Hejhal, B.N. Rackner, On the topography of Maass waveforms for PSL(2, Z). Experiment. Math. 1 (1992), no. 4, 275–305.

[7] J.P. Keating, The cat maps: quantum mechanics and classical motion. Nonlinearity 4 (1991), no. 2, 309–341.

[8] J. Marklof, Energy level statistics, lattice point problems and almost modular functions, Frontiers in Number Theory, Physics and Geometry, Les Houches 2003, to appear in Springer Lecture Notes.

[9] Z. Rudnick, On quantum unique ergodicity for linear maps of the torus. European Congress of Mathematics, Vol. II (Barcelona, 2000), 429–437, Progr. Math., 202, Birkhuser, Basel, 2001.

[10] Z. Rudnick, A central limit theorem for the spectrum of the modular group, Park city lectures, to appear in Annales Henri Poincare.

[11] E. Lindenstrauss, Rigidity of multi-parameter actions, Israel J. Math. (Furstenberg Special Volume), to appear.

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

[13] P. Sarnak, Spectra of hyperbolic surfaces. Bull. Amer. Math. Soc. (N.S.) 40 (2003), no. 4, 441–478

[14] S. Zelditch, Quantum Ergodicity and Mixing, math-ph/0503026 (this encycledia)

School of Mathematics, University of Bristol, Bristol BS8 1TW, U.K.