Mathematics of Quantum Chaos in 2019

Steve Zelditch

The purpose of this survey is to introduce the reader to the problems of mathematical QC (quantum chaos), moving rapidly from the origins of the subject to some of the most recent advances. Quantum chaos is now a rather mature field, and there exists a stream of expository articles over the last twenty years, including several by the present au-

Steve Zelditch is the Wayne and Elizabeth Jones Professor of Mathematics at Northwestern University. His email address is zelditch@math.northwestern.edu.

Communicated by Notices Associate Editor Chikako Mese.

For permission to reprint this article, please contact:
reprint-permission@ams.org.

DOI: <https://doi.org/10.1090/noti1958>

thor [Z17, Z18]. One of the leaders in the field, Nalini Anantharaman, has just given a plenary address on quantum chaos at the 2018 ICM, and interested readers may consult her ICM Proceedings article [A18] as well as her earlier expository article [A14]. Although this article begins at the beginning with the origins of quantum mechanics and its relation to classical mechanics, we aim to get to recent results and to avoid repeating material already nicely covered in the prior expositions.¹ The new results we discuss pertain to:

- Applications of the FUP (fractal uncertainty prin-

¹The number of references is limited to twenty; most are survey articles where the reader can find precise references.

ciple) to weak* limits of semiclassical measures (section “Uncertainty Principle and the Dyatlov-Jin Full Support Theorem”),

- L^p norms of eigenfunctions (subsection “ L^p norms”), and
- Small-scale quantum ergodicity and its applications to counting numbers of nodal domains, and proving equidistribution of L^2 mass and of nodal sets on shrinking balls (subsection “Lower bounds on numbers of nodal domains”).

Origins

The most familiar and historically the first quantum mechanical problem concerns the hydrogen atom, an electron moving around a nucleus. One might picture it as a classical 2-body problem, such as a planet going around the sun. But this picture cannot be correct, because an accelerating charge radiates energy and would spiral into the nucleus almost immediately. How can one reconcile this classical picture with the obvious fact that atoms do not immediately collapse?

Niels Bohr proposed the “old quantum theory,” which “quantizes” special periodic orbits to produce “stationary states” and “energy levels” of the hydrogen atom. These “Bohr-Sommerfeld” quantization conditions are still of considerable interest today and are the precursor to the field of semiclassical quantum mechanics connecting quantum and classical mechanics [Zw]. But the quantization conditions do not apply to more complicated systems such as the helium atom.

Schrödinger then proposed his famous equation to define “stationary states.” Instead of representing the electron as a point particle, he represented it by a vector $\psi \in \mathcal{H}$ in a Hilbert space. In the case of the hydrogen atom, $\mathcal{H} = L^2(\mathbb{R}^3, dx)$ (dx is Lebesgue measure). Schrödinger then introduced the operators of his name,

$$\hat{H}_\hbar := -\hbar^2 \Delta + V, \quad (1)$$

where $\hbar$ is a very small constant (Planck’s constant). Here, Δ is the Laplace operator and V is the potential, i.e., the multiplication operator by a function V on M . In the case of the hydrogen atom, $\Delta_g = \sum_{j=1}^3 \frac{\partial^2}{\partial x_j^2}$ is the standard Euclidean Laplacian and $V(x) = -\frac{1}{|x|}$. Schrödinger proposed that a particle with a fixed energy $E_j(\hbar)$ be represented by an L^2 -normalized eigenvector of (1), i.e.,

$$\hat{H}_\hbar \psi_{\hbar,j} = E_j(\hbar) \psi_{\hbar,j}. \quad (2)$$

He called $\psi_{\hbar,j}$ a stationary state: in some sense, it does not change with time. More precisely, the evolution of a state ψ in quantum mechanics is governed by the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} \psi(t, x) = \hat{H}_\hbar \psi(t, x), \quad \psi(0, x) = \psi_0(x).$$

When $\psi_0(x) = \psi_{\hbar,j}$ is an eigenfunction (2), then

$$\psi(t, x) = e^{\frac{-itE_j(\hbar)}{\hbar}} \psi_{\hbar,j}(x). \quad (3)$$

In general, the solution is given by applying the *propagator*

$$U_\hbar(t) := e^{-\frac{it}{\hbar} \hat{H}_\hbar}$$

to the initial state, where

$$i\hbar \frac{\partial}{\partial t} U_\hbar(t) = \hat{H}_\hbar U_\hbar(t), \quad U_\hbar(0) = I$$

is a 1-parameter group of unitary operators on $\mathcal{H}$. He also proposed that all physically relevant quantities should be matrix elements (expectation values) of a bounded operator A ,

$$\rho_{\hbar,j}(A) = \langle A \psi_{\hbar,j}, \psi_{\hbar,j} \rangle. \quad (4)$$

In the case of an eigenfunction,

$$\rho_{\hbar,j}(U_\hbar(-t) A U_\hbar(t)) = \rho_{\hbar,j}(A), \quad (5)$$

since the two factors of $e^{\frac{itE_j(\hbar)}{\hbar}}$ cancel out; the probabilities and the matrix elements are independent of t . For instance, $|\psi_{\hbar,j}(t, x)|^2 dx$ is the probability density of finding the particle of energy $E_j(\hbar)$ at the point x .

In effect, Schrödinger replaced the classical mechanics of the 2-body problem by linear algebra. By passing from classical mechanics to linear algebra, Schrödinger resolved the problem of how an electron moving around a nucleus can be moving and stationary at the same time. But the price one pays is that the intuitions of classical mechanics are lost in favor of linear algebra and functional analysis, about which most people have little intuition.

What puts classical mechanical intuition back into quantum mechanics is the behavior of the eigenfunctions and eigenvalues as the Planck constant $\hbar \rightarrow 0$. By classical mechanics is meant the time evolution (x_t, ξ_t) of points (x, ξ) in the phase space $T^*\mathbb{R}^n$ (the cotangent bundle of $\mathbb{R}^n$) under the classical Hamiltonian system

$$\begin{cases} \frac{d}{dt} x_t = \frac{\partial H}{\partial \xi}(x_t, \xi_t), \\ \frac{d}{dt} \xi_t = -\frac{\partial H}{\partial x}(x_t, \xi_t), \end{cases} \quad (6)$$

generated by the classical Hamiltonian

$$H(x, \xi) = |\xi|^2 + V(x) : T^*\mathbb{R}^n \rightarrow \mathbb{R} \quad (7)$$

associated to (1). Quantum mechanical objects tend to classical mechanical ones in the semiclassical limit $\hbar \rightarrow 0$. But which objects do they tend to? And how? In most problems, it is crucial to understand the joint behavior of the semiclassical limit $\hbar \rightarrow 0$ and the long-time evolution $t \rightarrow \infty$. This is the subject of semiclassical quantum mechanics [Zw].

The Hamiltonian flow (6) of any Hamiltonian preserves level sets (energy surfaces) $\Sigma_E = \{(x, \xi) \in T^*\mathbb{R}^n :$

$H(x, \xi) = E\}$, and classical dynamics refers to the Hamiltonian flow restricted to Σ_E . Quantum chaos refers to the special case where the classical Hamiltonian system is “chaotic” on an energy surface. The term “chaotic” is suggestive rather than precise and can be taken to mean ergodic, mixing, hyperbolic, Bernoulli, or some more specific type of chaotic or unpredictable dynamical behavior. What impact does the chaotic behavior of the classical flow have on the eigenvalues/eigenfunctions or the propagator of the quantum problem? If one holds $\hbar$ fixed, then the answer is none. One merely has the spectral decomposition of $\mathcal{H}$ into eigenspaces of (1) and the simple evolution (3) of the eigenvectors. But if one considers the behavior of the matrix elements (4) as $\hbar \rightarrow 0$ (or as the eigenvalue $E_j(\hbar) \rightarrow E$), or if one considers the joint asymptotics of $U_h(t)$ as $\hbar \rightarrow 0, t \rightarrow \infty$, then the effects of the chaotic classical dynamics are felt.

Although quantum chaos refers to quantizations of chaotic classical Hamiltonian systems, it should not be thought that chaos is the only interesting type of system. It is definitely of special interest, but the opposite extreme of “completely integrable system” is equally of interest. It is a familiar observation that the easiest dynamical systems to understand are the most chaotic ones or the most predictable (integrable) ones. One of the most interesting questions is to determine the semiclassical behavior of eigenfunctions/eigenvalues in mixed systems, which are partly integrable, partly chaotic (see [Go18]).

For which potentials V does the Hamiltonian (7) in Kinetic + Potential form generate chaotic dynamics on some energy surface $\Sigma_E \subset T^*\mathbb{R}^n$? To our knowledge, it is unknown! G. Paternain and M. Paternain have proved that such a Newtonian Hamiltonian flow cannot be highly chaotic (namely, Anosov or hyperbolic). Singular potentials generating chaotic flows do exist, and an important example is provided by a hydrogen atom in a strong magnetic field.² But there are many examples if we enlarge the setting to Riemannian manifolds, where $\mathcal{H} := L^2(M, dV_g)$ are the L^2 functions on a Riemannian manifold (M, g) with respect to its volume form dV_g . Model examples include the geodesic flow on the unit tangent bundle of a compact surface of negative curvature, and one of the most popular examples of quantum chaos is generated by the Laplacian Δ_g on a surface of constant negative curvature. This is in part due to its relevance to automorphic forms and arithmetic quantum chaos (see [S95, S11]). In this article we focus primarily on methods of PDE and semiclassical analysis rather than on L-functions and arithmetic methods.

²Thanks to Andreas Knauf for the references; he has obtained other singular examples.

Notation and Background

For the rest of the article, we specialize to the setting where $V = 0$ in (1), but Δ_g is the Laplacian on a compact Riemannian manifold (M^n, g) of dimension n , locally given by

$$\Delta_g = \frac{1}{\sqrt{|g|}} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} g^{ij} \sqrt{|g|} \frac{\partial}{\partial x_j}, \quad (8)$$

where $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$, $[g^{ij}]$ is the inverse matrix to $[g_{ij}]$, and $|g| = \det[g_{ij}]$. When M is compact, there is an orthonormal basis $\{\phi_j\}$ of eigenfunctions,

$$\Delta_g \phi_j = -\lambda_j^2 \phi_j, \quad \int_M \phi_i \phi_j dV_g = \delta_{ij} \quad (9)$$

with $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots \uparrow \infty$ repeated according to their multiplicities. When M has a nonempty boundary ∂M , one imposes boundary conditions such as Dirichlet $Bu = u|_{\partial M} = 0$ or Neumann $Bu = \partial_\nu u|_{\partial M} = 0$.

If one sets $\hbar_j = \lambda_j^{-1}$, then (9) takes the form

$$\hbar_j^2 \Delta_g \phi_j = -\phi_j, \quad (10)$$

which is (2) with $V = 0$ and with $E_j(\hbar) = -1$. The limit $\hbar_j \rightarrow 0$ is thus the same as $\lambda_j \rightarrow \infty$. It is customary in semiclassical analysis to use the form (10) and speak of the “semiclassical limit,” but also customary in PDE to use (2) and speak of the “high frequency limit.” When $V = 0$ the limits are the same.

The Weyl eigenvalue counting function is

$$N(\lambda) = \#\{j : \lambda_j \leq \lambda\} = C_n \text{Vol}(M, g) \lambda^n + O_g(\lambda^{n-1}).$$

Here and hereafter, a constant C_n (resp., C_g) is understood to depend only on the dimension (resp., metric).

The classical Hamiltonian underlying (8) is the metric norm-square of a covector

$$|\xi|_g^2 = \sum_{i,j} g^{ij}(x) \xi_i \xi_j : T^*M \rightarrow \mathbb{R}_+.$$

In the language of pseudo-differential operators, $|\xi|_g^2$ is the symbol of $-\Delta_g$. We take square roots to get the first-order operator $\sqrt{-\Delta_g}$ with symbol $|\xi|_g$. The motivation is that $\sqrt{-\Delta_g}$ generates the half-wave group

$$U(t) = \exp it\sqrt{-\Delta_g}; \quad (11)$$

i.e., we can work with the half-wave propagator (11) rather than with the Schrödinger flow. The Hamiltonian flow (6) generated by $H(x, \xi) = |\xi|_g$ is the homogenous *geodesic flow*. We usually restrict it to an energy surface $\Sigma_1 = S^*M$ (the unit cosphere bundle $\{|\xi|_g = 1\}$) to obtain the key dynamical flow

$$G^t : S^*M \rightarrow S^*M.$$

It preserves Liouville measure $d\mu_L$, the natural surface measure on S^*M .

What is “quantization”? Quantization is a procedure for converting nice functions (classical observables) $a(x, \xi)$ on phase space T^*M into operators $\text{Op}_h(a)$ on $L^2(M)$ in such a way that commutators $[\text{Op}_h(a), \text{Op}_h(b)] = \frac{h}{i}\{a, b\} + O(h^2)$, where $\{a, b\} = \sum_j (\frac{\partial a}{\partial x_j} \frac{\partial b}{\partial \xi_j} - \frac{\partial a}{\partial \xi_j} \frac{\partial b}{\partial x_j})$ is the Poisson bracket. With this (and similar requirements), quantum mechanics tends to classical mechanics as $h \rightarrow 0$. Schrödinger was the first to quantize Hamiltonians of the form $|\xi|^2 + V(x)$ as the operators (1), and H. Weyl soon after defined the *quantization* $\text{Op}_h(a)$ of a general classical observable a . For instance, the quantization of the linear function ξ_j is $P_j = \frac{h}{i} \frac{\partial}{\partial x_j}$, and the quantization of x_j is $Q_j = M_{x_j}$ (multiplication by x_j) with the commutation relations $[P_j, Q_k] = \frac{h}{i} \delta_{jk}$. The quantization also extends to the groups generated by these operators, i.e., to their exponentials $e^{itQ}, e^{itP}, e^{it\hat{H}_h}$. It is easy to see that $e^{itQ_j} f(x) = e^{itx_j} f(x)$, $e^{itP_k} f(x) = f(x + te_k)$, where e_k is the k th standard basis element of $\mathbb{R}^n$. These operators are unitary and are the simplest examples of Fourier integral operators which quantize general canonical (i.e., symplectic) transformations on cotangent bundles.

Actually, quantization is not unique, since one can conjugate any quantization by a unitary operator to obtain another. Although the Schrödinger quantization is the most common one in PDE, this article will often employ the holomorphic quantization on Bargmann–Fock space of entire holomorphic functions on $\mathbb{C}^n \simeq T^*\mathbb{R}^n$ which are in L^2 with respect to Gaussian measure. The Bargmann transform takes $L^2(\mathbb{R}^n)$ to the Bargmann–Fock space $H^2(\mathbb{C}^n, e^{-|z|^2} dz)$, where dz is the Lebesgue measure on $\mathbb{C}^n$. Bargmann–Fock space comes up naturally when one works with creation/annihilation operators $a = P + iQ$, $a^* = P - iQ$. It is somewhat simpler in that classical mechanics and quantum mechanics both take place on the phase space $T^*\mathbb{R}^n \simeq \mathbb{C}^n$; quantum mechanics is just the restriction of classical mechanics to holomorphic functions.

This procedure of equipping a cotangent bundle with a complex structure makes sense on any real analytic manifold.³ We can complexify M to a complex manifold $M_{\mathbb{C}}$, which is canonically equivalent to T^*M (or rather to a ball bundle $B_{\epsilon_0}^*M$) and identify T^*M with $M_{\mathbb{C}}$ using the imaginary time exponential map $E(x, \xi) = \exp_x i\xi$. Using E we endow T^*M with a complex structure, adapted to a real analytic metric g . More precisely, it lives on a ball bundle $B_{\epsilon_0}^*M$ of some radius ϵ_0 , defined by $|\xi|_g \leq \epsilon_0$, known as a Grauert tube (due to L. Lempert–R. Szöke and V. Guillemin–M. Stenzel; see [Z18]).

Each eigenfunction ϕ_j admits a holomorphic extension $\phi_j^{\mathbb{C}}$ to $M_{\mathbb{C}}$. Thus, $\phi_j^{\mathbb{C}}(E(x, \xi))$ is a holomorphic function

on the cotangent ball bundle B_{ϵ}^*M . For instance, if $M = S^1$ (the unit circle), then $M_{\mathbb{C}} \simeq S^1 \times \mathbb{R}$, $\phi_j(x) = \sin jx$, and $\phi_j^{\mathbb{C}}(z) = \sin jz$. As this example indicates, $\phi_j^{\mathbb{C}}$ is exponentially growing (or decaying) as $j \rightarrow \infty$. Therefore we L^2 -normalize it on each “sphere bundle” $\partial M_{\epsilon} \simeq S_{\epsilon}^*M$, obtaining the sequence $\left\{ \frac{\phi_j^{\mathbb{C}}}{\|\phi_j^{\mathbb{C}}\|_{L^2(\partial M_{\epsilon})}} \right\}_{j \geq 1}$ on phase space.

Quantization, observables, and expectation values. As mentioned above, the observable quantities in quantum mechanics are the “expectation values” (4)

$$\rho_j(A) := \langle A\phi_j, \phi_j \rangle \quad (12)$$

of a bounded self-adjoint operator A relative to an energy eigenstate (i.e., eigenfunction) ϕ_j . Expectation values are used to probe the behavior of eigenfunctions in the semiclassical limit: that is, we study the asymptotics of (12) for general “test operators” A as $j \rightarrow \infty$ to study the semiclassical asymptotics of eigenfunctions. As this statement suggests, the main problems involve sequences $\{\phi_{j_k}\}$ of eigenfunctions rather than individual ones and their possible limit behavior.

There will not exist good asymptotics for (12) for a general bounded operator A . To obtain useful asymptotics, one needs to choose special types of bounded operators A . The most popular and standard choice in the Schrödinger representation is that $A \in \Psi^0(M)$, i.e., A is a pseudo-differential operator of order zero on $L^2(M)$, either homogeneous or semiclassical (see [Zw] for background). Indeed, pseudo-differential operators were virtually invented for purposes such as this, to possess nice semiclassical behavior. A key property of $A \in \Psi^0(M)$ is that it possesses a principal symbol $\sigma_A \in C^\infty(T^*M)$, which is homogeneous of degree zero when A is a homogeneous pseudo-differential operator. It can then be identified with a smooth function on the unit cotangent bundle S^*M , and one thinks of A as the quantization of σ_A . In the holomorphic representation one uses Toeplitz operators $\Pi\sigma\Pi$ where Π is the orthogonal projection to the holomorphic functions and σ is a symbol.

The first thing to know about the expectation values (or matrix elements) (12) is that they induce distributions $d\Phi_j$ on T^*M . In fact, with a proper choice of quantization, they induce probability measures on S^*M by the rule

$$\langle \text{Op}(a)\phi_j, \phi_j \rangle = \int_{S^*M} a(x, \xi) d\Phi_j.$$

In other words, the linear functional $a \in C^\infty(S^*M) \rightarrow \text{Op}(a) \in \Psi^0(M) \rightarrow \rho_j(\text{Op}(a))$ is a positive linear functional of a and therefore defines a measure $d\Phi_j$. In various contexts, it has been called a “microlocal lift,” a “quantum lift,” a “Wigner distribution,” or a “microlocal defect measure.” The key property is that $d\Phi_j$ lives in phase space T^*M ; in the homogeneous setting, it lives on S^*M .

³On general C^∞ manifolds one uses “almost analytic continuation.”

In the holomorphic quantization, $\text{Op}(a) = \Pi a \Pi$, where Π is the orthogonal projection to $H^2(M_\epsilon)$, the L^2 holomorphic functions on the Grauert tube. Then

$$\rho_j(\Pi a \Pi) = \int_{\partial M_\epsilon} a(z) d\Phi_j(z),$$

with

$$d\Phi_j(z) := \left| \frac{\phi_j^c(z)}{\|\phi_j^c\|_{L^2(\partial M_\epsilon)}} \right|^2 d\mu_\epsilon, \quad (13)$$

where $d\mu_\epsilon$ is the (Liouville) volume form on ∂M_ϵ induced by the symplectic volume form on T^*M . Note that $d\Phi_j$ is manifestly a positive, smooth probability measure on $\partial M_\epsilon \simeq S^*_\epsilon M$. We opt to use the holomorphic representation because (13) is most elementary in this setting. The measures (13) are called “Husimi distributions” in the physics literature.

Weak* limit problem. We define

$$\mathcal{Q} := \text{weak*}-\text{limits of the sequence } \{d\Phi_j\}. \quad (14)$$

A probability measure $d\nu$ is a weak* limit of a subsequence $\{d\Phi_{j_k}\}_{k=1}^\infty$ if $\int_{S^*M} f d\Phi_{j_k} \rightarrow \int_{S^*M} f d\nu$ for every $f \in C(S^*M)$. In the holomorphic setting, one has the rather explicit formulae (13) for the $d\Phi_j$.

The weak* limit problem is to determine all of the weak* limits. The only simple characterization of the limits is:

Proposition 1. *If M is a compact manifold, then $\mathcal{Q} \subset \mathcal{M}_I$, where $\mathcal{M}_I$ is the compact convex set of G^t -invariant probability measures on S^*M for the geodesic flow. The limits are time-reversal invariant if the eigenfunctions are real-valued.*

Any weak* limit μ of $\{\rho_j\}$ is an invariant measure for G^t ; i.e., $\mu(E) = \mu(G^t E)$ for all Borel subsets $E \subset S^*M$. This is because, by (5), ρ_j is an invariant state for the automorphism $A \rightarrow U(t)AU(-t)$. The proposition follows by Egorov’s theorem, which says that $\rho_j(U(t)\text{Op}(a)U(-t)) = \rho_j(\text{Op}(a \circ G^t)) + o(1)$ as $j \rightarrow \infty$ (see [Zw]).

There are many invariant probability measures, and it is difficult to characterize those that arise as quantum limits. Some examples of invariant measures are:

- (1) Normalized Liouville measure $d\mu_L$.
- (2) A periodic orbit measure μ_γ defined by $\int \sigma d\mu_\gamma = \frac{1}{L_\gamma} \int_\gamma \sigma ds$, where L_γ is the length of γ . The corresponding sequence of eigenfunctions is sometimes said to “scar” along γ .
- (3) A delta-function along an invariant Lagrangian manifold $\Lambda \subset S^*M$. The associated eigenfunctions are viewed as *localizing* along Λ .
- (4) A more general measure that is singular with respect to $d\mu_L$. There are many examples in the negatively curved case.

Definition 2. A subsequence $\{\phi_{j_k}\}_{k=1}^\infty$ of an orthonormal basis is called QE (quantum ergodic) if $d\Phi_{j_k} \rightarrow d\mu_L$ in the weak* sense.

The Laplacian Δ_g is called QUE if one orthonormal basis of eigenfunctions is QE. To the author’s knowledge, the only setting where QUE is known to occur is for Hecke eigenfunctions of arithmetic hyperbolic manifolds and related settings (E. Lindenstrauss; see [S11]).

The simplest example is that of the circle $M = S^1$. The complex eigenfunctions are e^{inx} , $n \in \mathbb{Z}$. We always assume the eigenfunctions are real-valued and therefore consider $\frac{1}{\sqrt{2}} \cos nx$, $\frac{1}{\sqrt{2}} \sin nx$. It is an immediate consequence of the Riemann–Lebesgue lemma that, for any $f \in C(S^1)$,

$$\int_{S^1} f(\sin nx)^2 dx \rightarrow \int_{S^1} f dx.$$

Here, and henceforth, $\int$ denotes the average; i.e., one normalizes a measure to have unit mass.

On the other hand, this calculation fails if we use the eigenfunctions $e^{\pm inx}$. The failure is the simplest illustration of quantum ergodicity versus quantum integrability. In phase space terms, e^{inx} corresponds to the Lagrangian submanifold $\xi = n$ in $T^*S^1 = S^1 \times \mathbb{R}_\xi$. The unit cotangent bundle $|\xi| = 1$ has two connected components, so that the geodesic flow $G^t(x, \xi) = (x + t\xi, \xi)$ is not ergodic. But the real eigenfunctions correspond to the quotient of $T^*\mathbb{R}$ under $\xi \rightarrow -\xi$, gluing together the two components, and the geodesic flow is ergodic on the quotient.

There is a profound difference between studying the weak* limit problem in phase space T^*M and in configuration (physical) space M , where the testing operators are just multiplications. The latter is best thought of as testing for “flatness,” i.e., uniform distribution on the base manifold M rather than on the unit cotangent bundle. On a flat torus, for instance,

$$\text{Op}(a)e^{i\langle x, \lambda \rangle} = a(x, \lambda)e^{i\langle x, \lambda \rangle}.$$

Hence

$$\langle \text{Op}(a)e^{i\langle x, \lambda \rangle}, e^{i\langle x, \lambda \rangle} \rangle = \int_{T^n} a\left(x, \frac{\lambda}{|\lambda|}\right) dx.$$

Thus, eigenfunctions concentrate on the invariant tori $\xi = \lambda/|\lambda|$, although they have modulus 1 on the torus.

It is natural to wonder if a weak* limit measure μ might have some a priori regularity properties. Of course, invariance under G^t implies smoothness in the direction of geodesics, so the question is regularity in the transverse direction. The short answer is no. For instance, Jakobson–Zelditch proved that any invariant measure for the geodesic flow on the standard sphere S^n is a weak* limit of some sequence of eigenfunctions; there are many generalizations of this result [Z17]. There is a profound result, due

to Anantharaman and Anantharaman-Nonnenmacher (see [A14, A18]) giving lower bounds on the dynamical entropy of the limit measures; this is essentially a regularity result.

Quantum ergodicity of eigenfunctions when G^t is ergodic. In this section, we assume that the geodesic flow of (M, g) is ergodic. Let

$$\omega(A) := \int_{S^*M} \sigma_A d\mu_L. \quad (15)$$

The following is due to A. I. Schnirelman, S. Zelditch, Y. Colin de Verdière, S. Zelditch–M. Zworski, P. Gerard–E. Leichtnam (see [A14, A18, Z17, Zw, So14] for background).

Theorem 3. *Let (M, g) be a compact Riemannian manifold (possibly with boundary), and let $\{\lambda_j, \phi_j\}$ be the spectral data of $\sqrt{-\Delta_g}$. Then, if the geodesic flow G^t is ergodic on $(S^*M, d\mu_L)$, there exists a subsequence S of density one ($D^*(S) = 1$) such that*

$$\lim_{j \rightarrow \infty, j \in S} \langle A\phi_j, \phi_j \rangle \rightarrow \omega(A) \quad \text{for all } A \in \Psi^0(M).$$

Density one means that $\frac{1}{N(\lambda)} \#\{j \in S : \lambda_j \leq \lambda\} \rightarrow 1$ as $\lambda \rightarrow \infty$.

The proof splits into two parts: (i) a variance estimate and (ii) extraction of a density one subsequence by a diagonal argument. The key quantities in (i) are the quantum variances

$$V_A(\lambda) := \frac{1}{N(\lambda)} \sum_{j: \lambda_j \leq \lambda} |\langle A\phi_j, \phi_j \rangle - \omega(A)|^2.$$

The general result relating variances and classical mechanics is the following theorem, an improvement of Theorem 3 due to Zelditch and T. Sunada (see [Z17] for background).

Theorem 4. *Let (M, g) be a compact Riemannian manifold (possibly with boundary), and let $\{\lambda_j, \phi_j\}$ be the spectral data of $\sqrt{-\Delta_g}$. Then the geodesic flow G^t is ergodic on $(S^*M, d\mu_L)$ if and only if, for every $A \in \Psi^0(M)$, $\lim_{\lambda \rightarrow \infty} V_A(\lambda) = 0$, and moreover*

$$(\forall \epsilon)(\exists \delta) \lim_{\lambda \rightarrow \infty} \frac{1}{N(\lambda)} \sum_{\substack{j \neq k: \lambda_j, \lambda_k \leq \lambda \\ |\lambda_j - \lambda_k| < \delta}} |\langle A\phi_j, \phi_k \rangle|^2 < \epsilon.$$

The fact that $V_A(\lambda) \rightarrow 0$ implies that, for each A , there exists a subsequence $\{j_k\}_{k=1}^\infty \subset \mathbf{N}$ of density one for which $\rho_{j_k}(A) \rightarrow 0$. This is because all terms are positive. The fact that the density one subsequence can be chosen independently of A requires a diagonal argument using a countable dense subset of symbols. This diagonal argument fails if one tries to prove QE for observables like multiplication by the indicator function $\mathbf{1}_E$ of a general Borel set E , since L^∞ is not separable.

Variance estimates. The fact that $V_A(\lambda) \rightarrow 0$ raises the question of the “rate of quantum ergodicity.” The best that can be expected is that $V_A(\lambda) \leq C\lambda^{-1}$. More generally, one can study averages $V_A^p(\lambda)$ of p th power variances $|\langle A\phi_j, \phi_j \rangle - \omega(A)|^{2p}$. Quantitative estimates have been studied by several authors, giving the following results:

- For negatively curved manifolds, $V_A(\lambda) \leq C(\log \lambda)^{-1}$; more generally $V_A^p(\lambda) \leq C(\log \lambda)^{-p}$ (Zelditch, 1994; R. Schubert (2006, 2008)). There are generalizations to exponential functions of $|\langle A\phi_j, \phi_j \rangle - \omega(A)|$ (Anantharaman–Rivière, 2012).
- In the arithmetic setting such as the quotient of the upper-half-plane $\mathbb{H}^2/\mathrm{SL}(2, \mathbb{Z})$, $V_A(\lambda) \sim C_A \lambda^{-1}$, where C_A is an explicit constant (Luo–Sarnak (2004), P. Zhao (2010), Zhao–Sarnak (2019)). See [Z17] for references.

The proofs involve estimates on decay of correlations for the geodesic flow, i.e., the rate of decay of $\int_{S^*M} f(G^t(\zeta)) \times f(\zeta) d\mu_L(\zeta)$ when $f \in C^\infty(S^*M)$ has mean value zero. The last result is amazingly accurate. It uses arithmetic techniques that have no analogue for general hyperbolic surfaces or other manifolds. Logarithmic estimates in the negatively curved case reflect the exponential growth of the geodesic flow and are now ubiquitous in quantum chaos. It is a fundamental problem to break this exponential barrier—if, indeed, it is even possible.

QUE in terms of time and space averages. In the ergodic case, one knows that $\mu_L \in \mathcal{Q}$ (14) and that it is the limit of “almost all” eigenfunctions of an orthonormal basis. What is known about possible exceptional subsequences of density zero?

This is known as the QUE problem. It may be put in a functional analytic context as follows: Let $\langle A \rangle := \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T U(t)AU(-t) dt$. This is the diagonal part of A with respect to the basis $\{\phi_j\}$. For $A \in \Psi^0(M)$, let $\omega(A)$ be as defined in (15). QE implies that $\langle A \rangle = \omega(A)I + K$, where all matrix elements of the diagonal operator K tend to zero along a density one subsequence.

Problem 5. Suppose the geodesic flow G^t of (M, g) is ergodic on S^*M . Is K a compact operator? Compactness of K implies that $\langle K\phi_j, \phi_j \rangle \rightarrow 0$; hence $\langle A\phi_j, \phi_j \rangle \rightarrow \omega(A)$ along the entire sequence. Equivalently, is $\mathcal{Q} = \{d\mu_L\}$?

In this case, $\sqrt{-\Delta}$ is said to be QUE (quantum uniquely ergodic). Rudnick–Sarnak conjectured that $\sqrt{-\Delta}$ of negatively curved manifolds are QUE, i.e., that for any orthonormal basis of eigenfunctions, the Liouville measure is the only quantum limit.

There do exist ergodic situations in which quantum limit measures other than Liouville occur. It was proved by A.

Hassell [Ha16] that the well-known Bunimovich stadium is an example. Numerical results of the physicist E. J. Heller had strongly suggested this result, but it is difficult to prove. **Mixed systems and converse QE.** The weak* limit problem has mainly been studied in the case of completely integrable systems (such as flat tori or surfaces of revolution) and ergodic systems (such as manifolds of negative curvature). Most Hamiltonian systems are “mixed” in that they are neither integrable nor ergodic. It is hard to analyze such systems because the “integrable component” can be Cantor-type sets whose indicator functions cannot be quantized in an obvious way. For KAM systems, this Cantor set is a kind of Cantor foliation by invariant tori, and it seems very plausible that a positive proportion of eigenfunctions concentrate on these tori and are not QE. Results of this kind have been recently established by S. Gomes [Go18] and by Gomes–Hassell.

Problem 6. Do there exist (M, g) with nonergodic geodesic flow for which $\sqrt{-\Delta_g}$ is quantum ergodic (i.e., is Theorem 3 valid)?

An example due to B. Gutkin is a race-course stadium. There is a vague conjecture that “localization of eigenfunctions” should be rare and that they should be “diffuse in phase space” in quite general settings. The results of Gomes–Hassell indicate that QE is not a generic property.

Length scales of QE. Another natural question is what is the smallest length scale $r(\lambda_j)$ on which QE takes place; i.e., whether Theorem 3 holds when the pseudo-differential operator is replaced by a sequence $A = A_j$ whose principal symbols σ_{A_j} are supported on balls $B(p, r(\lambda_j))$ of radii $r(\lambda_j) \rightarrow 0$ in phase space or in configuration space. The natural length scales in quantum dynamics are:

- (i) the wavelength scale $\hbar_j = \lambda_j^{-1}$,
- (ii) logarithmic length scales $(\log \lambda_j)^{-\gamma}$ (where $\gamma > 0$),
- (iii) the mean level spacing between eigenvalues.

The Heisenberg (and other) uncertainty principles limit the smallness of the scale on which one can relate classical and quantum mechanics. The spacing in (iii) is too small for current techniques to approach, although it is much discussed in the physics literature. On the wave length scale, eigenfunctions only oscillate a fixed number (e.g., 1,000) of times, and it is hard to conceive of a semiclassical limit on this scale, but one might imagine limit theorems on the scale $\lambda_j^{-1}(\log \lambda_j)^\gamma$. There is a sequence of recent results on small-scale QE, due to X. Han [Han18], H. Hezari–G. Rivière [HR16], P. Humphries, and others. The first three authors prove QE in a strong sense on log-scale shrinking balls. Humphries [Hum18] has rather surprising counterexamples to small scale QE on balls of the scale $\lambda_j^{-1}(\log \lambda_j)^\gamma$. Robert Chang and the author adapted the proofs of Hezari–Rivière and Han of log-scale quantum

ergodicity to the complex setting of Grauert tubes for the Husimi measures (13) (see [Z18] for references).

QER: Quantum ergodic restriction theorems. Let $H \subset M$ be a hypersurface in a manifold with ergodic geodesic flow. Then the set S_H^*M of unit-covectors with footpoint on H is a “cross section” of the geodesic flow in S^*M . The quantum analogue is the space of Cauchy data $(\phi_j|_H, \partial_\nu \phi_j|_H)$ of eigenfunctions on the hypersurface. The first return map of geodesics to the cross section is an ergodic symplectic map, suggesting that the Cauchy data should be ergodic in $L^2(H)$ (QER). Such QER phenomena are useful in nodal set problems, especially in the case where $\dim M = 2$, $\dim H = 1$, because one can relate nodal sets of ϕ_j on M and zeros of ϕ_j on H , which are much easier to study.

QER has been proved by H. Christianson, J. A. Toth, and the author following a proof in the case where $H = \partial M$ due to Hassell and the author (see [Z17]). We state the result in a special setting for later applications to nodal sets, that of negatively curved surfaces possessing an isometric involution $\sigma: M \rightarrow M$ with nonempty fixed point set $\text{Fix}(\sigma)$, a finite union of closed geodesics, which divides $M = M_+ \cup M_-$ into two components. Such a Riemann surface is called a “real Riemann surface”: it is the complexification of $\text{Fix}(\sigma)$. The isometry σ acts on $L^2(M, dA_g)$, and we define $L_{\text{even}}^2(M)$, respectively, $L_{\text{odd}}^2(M)$, to denote the subspace of even functions $f(\sigma x) = f(x)$, respectively, odd elements $f(\sigma x) = -f(x)$. The even/odd eigenfunctions with respect to σ satisfy Dirichlet/Neumann boundary conditions on $M_\pm$.

Theorem 7. Let (M, g) be a real Riemann surface with isometric involution σ as above, and let $\gamma = \text{Fix}(\sigma)$ divide $M = M_+ \cup M_-$ into two components. More generally, let M be any compact surface with boundary $\gamma = \partial M$ with ergodic geodesic flow, and let $\gamma = \partial M$. Then, for a subsequence of even (Neumann) eigenfunctions of density one,

$$\int_\gamma f \phi_j^2 ds \rightarrow \frac{4}{2\pi \text{Area}(M)} \int_\gamma f(s) ds,$$

for any $f \in C(\gamma)$. Similarly for normal derivatives of Dirichlet (odd) eigenfunctions.

In fact, there is a complementary general result: for a generic hypersurface H of any (M^n, g) , there exists a density one subsequence such that $\{\phi_j|_H\}_{j=1}^\infty$ is QE on H . The generic condition is known as “asymmetry” (see [TZ17, Z18]). However, the curves in Theorem 7 are symmetric, and all odd eigenfunctions vanish on γ .

Random wave/ONB models. A general heuristic (M. V. Berry) is that the ONB (orthonormal basis) of eigenfunctions behaves like a sequence of independent Gaussian “random waves” of fixed frequency, such as random spherical harmonics on the sphere with fixed degree N . Since

the eigenfunctions form an ONB, the author prefers a random ONB model in which one chooses a random ONB of spherical harmonics of each degree N . Random waves/ONBs have many properties that are unknown and possibly false for ONBs of eigenfunctions, but the heuristic has inspired many interesting results on random nodal domains (Nazarov–Sodin), random topology (Canzani–Sarnak–Wiegmann, Gayet–Welschinger), and QUE of random ONBs (Zelditch, Chatterjee–Galkowski, Bourgade–Yau, R. Chang (see [Z17] for some references)).

Uncertainty Principle and the Dyatlov–Jin Full Support Theorem

A recent breakthrough on the QUE problem for hyperbolic surfaces is the following result due to Dyatlov–Jin [DJ18]. We use semiclassical notation where $\hbar_j = \lambda_j^{-1}$ and denote eigenfunctions by $u_{\hbar_j}$ or more simply $u_{\hbar}$ as in (10).

Theorem 8. *Let (M, g) be a compact hyperbolic surface. Let $a \in C_0^\infty(T^*M)$ with $a|_{S^*M}$ not identically zero. Let $u_{\hbar}$ be an eigenfunction of eigenvalue $\hbar^{-2}$ and let $\|u_{\hbar}\|_{L^2} = 1$. Then there exist constants $\hbar_0(a)$ and C_a independent of $\hbar$ so that, for $\hbar \leq \hbar_0(a)$,*

$$\|\text{Op}_{\hbar}(a)u_{\hbar}\|_{L^2} \geq C_a.$$

Here, $\text{Op}_{\hbar}(a)$ is the semiclassical pseudo-differential operator with symbol a . If $a(x, \xi) = V(x)$ is a multiplication operator, one gets that $\int_B |u_{\hbar}|^2 dV \geq C_B > 0$, i.e., a uniform lower bound (in $\hbar$) of the L^2 mass on all balls.

Corollary 9. *All quantum limits of sequences of eigenfunctions on compact hyperbolic surfaces have full support in S^*M , i.e., charge every open set.*

The corollary does not imply that every quantum limit is Liouville (see Definition 2). For instance, a quantum limit could be a convergent sum of delta-functions along a dense set of closed geodesics.

One of the main ingredients in the proof of Theorem 8 is the so-called FUP (fractal uncertainty principle). It is related to a classical problem of Landau–Slepian–Pollack related to the uncertainty principle of quantum mechanics: Can there exist a function $f \in L^2$ that is concentrated on an interval A such that its Fourier transform $\mathcal{F}f$ is concentrated on an interval B ? To make this precise, let $P_A = \mathbf{1}_A$ and $Q_B = \mathcal{F}^* \mathbf{1}_B \mathcal{F}$. If there exists f that is ϵ -concentrated on A in the sense that $\int_{X \setminus A} |f|^2 \leq \epsilon^2 \int_X |f|^2$ and such that $\mathcal{F}f$ is δ -concentrated on B , then $1 - \epsilon - \delta \leq \|P_A Q_B\|$.

The operator $Q_B P_A Q_B$ is self-adjoint and trace-class, and

$$\lambda_1 = \|P_A Q_B\|^2.$$

Thus, $\sqrt{\lambda_1}$ is the cosine of the angle between the ranges of P_A and Q_B . Moreover, $\lambda_1 < 1$. In fact, if $0 \leq \alpha, \beta \leq 1$ and $(\alpha, \beta) \neq (1, 0), (0, 1)$, then there exists f with

$\|f\| = 1$ such that $\|P_A f\| = \alpha$, $\|Q_B f\| = \beta$ if and only if $\cos^{-1} \alpha + \cos^{-1} \beta \geq \cos^{-1} \sqrt{\lambda_1}$.

The FUP in the sense of Bourgain, Dyatlov, Jin, and Zahl is a kind of generalization where one replaces intervals by regular porous fractal sets. Some definitions:

- For any set X let $X(s) = X + [-s, s]$.
- Given $\nu \in (0, 1)$ and $0 < \alpha_0 < \alpha_1$, say that $\Omega \subset \mathbb{R}$ is ν -porous on scales α_0 to α_1 if for each interval I of size $|I| \in [\alpha_0, \alpha_1]$ there exists a subinterval $J \subset I$ with $|J| = \nu|I|$ such that $J \cap \Omega = \emptyset$.
- Ahlfors–David regular: Let $\delta \in [0, 1]$, $C_R \geq 1$, and $0 < \alpha_0 < \alpha_1$. Say that $X \subset \mathbb{R}$ is δ -regular with constant C_R on scales α_0 to α_1 if there exists a Borel measure μ_X on $\mathbb{R}$ supported in X such that (i) for any interval I with $|I| \in [\alpha_0, \alpha_1]$, $\mu_X(I) \leq C_R |I|^\delta$; (ii) if I is centered at $a \in X$, then $\mu_X(I) \geq C_R^{-1} |I|^\delta$.
- Porous sets are embedded into Ahlfors–David regular sets of some dimension $\delta < 1$.

The FUP for δ -regular sets states the following.

Proposition 10. *Let $B(h)$ be a semiclassical FIO on $L^2(\mathbb{R})$ of the form*

$$B(h)f(x) = h^{-\frac{1}{2}} \int_{\mathbb{R}} e^{i\Phi(x,y)/h} b(x,y) f(y) dy$$

with $b \in C_0^\infty(U)$ and $\partial_{xy}^2 \Phi \neq 0$ on U . Suppose that $X, Y \subset \mathbb{R}$ are Ahlfors–David δ -regular. Then there exists $\beta > 0$ so that

$$\|\mathbf{1}_{X(h)} B(h) \mathbf{1}_{Y(h)}\|_{L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})} \leq Ch^\beta.$$

Dyatlov–Jin apply this result to a porous set arising in the dynamics of the geodesic flow on a compact hyperbolic surface. Such a flow is hyperbolic, and it preserves two foliations transverse to the flow direction: one foliation consists of stable leaves (horocycles) contracted by the flow, the other of unstable leaves (antihorocycles) expanded by the flow. Roughly speaking, these two foliations play the role of the position, respectively, momentum, axes in the case of the Fourier transform. If there is an open “hole” A in the support of the quantum limit measure, then the flowout $\bigcup_{t \in \mathbb{R}} G^t(A)$ of the hole is also a hole, and the quantum limit measure must be supported in its complement. This complement is a regular porous set and is very “sparse” in the Liouville-measure sense. But until Anantharaman’s entropy result (see [A18]), it could not be proved that eigenfunctions do not concentrate on a closed geodesic, much less a low measure set. The FUP is the new tool that shows that the support is full, i.e., dense, even though the limit measure could still be a convergent sum of delta-functions on a dense set of closed geodesics.

Applications of QE

In view of the amount of effort expended to prove QE and QUE in various settings and on various length scales, it is

natural to ask, what applications does QE have to the harmonic analysis of eigenfunctions? What features of eigenfunctions should be sensitive to ergodicity? Here are some intuitions.

- Ergodicity (or chaos) causes eigenfunctions to rapidly oscillate everywhere and in all directions. Therefore, their nodal (zero) sets should be uniformly distributed on M and (in some suitable sense) in phase space. Moreover, ergodic eigenfunctions should have a lot of nodal domains; i.e., the connected components of M where ϕ_j is positive or negative should be rather small and there should exist many of them.
- The oscillations are on the wavelength scale $\hbar_j = \lambda_j^{-1}$. Therefore, if one “blows up” a small ball of radius $r_j(\hbar) \gg \hbar_j$ so that the number of wavelengths is still growing across the ball, one should still see quantum ergodicity, i.e., oscillation in all directions and at all points.
- Ergodicity should prohibit exceptional concentration in any region of M ; therefore L^p norms of ergodic eigenfunctions should be rather small compared to extremal cases. Moreover, their L^2 norms on any fixed $\hbar$ -independent ball $B \subset M$ should have a uniform lower bound in $\hbar_j$ (depending on B).
- Cauchy data $(\phi_j|_H, \partial_\nu \phi_j|_H)$ of eigenfunctions on hypersurfaces $H \subset M$ should be quantum ergodic along H ; ∂_ν is the normal derivative. For generic H , the restrictions $\gamma_H \phi_j := \phi_j|_H$ should have quantum ergodic behavior similar to that on the global manifold. This is because the classical analogue of Cauchy data of eigenfunctions is the cross section S_H^*M (the set of unit covectors to M with footpoint on H) of $G^t : S^*M \rightarrow S^*M$. The corresponding first return map $\Phi : S_H^*M \rightarrow S_H^*M$ is ergodic on S_H^*M .

These heuristics are intentionally vague and debatable. In the next section we describe some rigorous results in this direction.

L^p norms. As mentioned above, quantum ergodic sequences in the sense of Definition 2 should have relatively small sup-norms. The following theorem is a special case of [Gal19, Theorem 1]:

Theorem 11. *Let $\{\phi_{j_k}\}_{k=1}^\infty$ be a QE sequence of eigenfunctions. Then $\|\phi_{j_k}\|_{L^\infty} = o(\lambda_{j_k}^{\frac{n-1}{2}})$.*

Galkowski’s result applies more generally to sequences whose quantum limit measure is diffuse at x ; we refer to [Gal19] for the definition. Y. Canzani–J. Galkowski have proved further results in this direction. In the case of negatively curved compact manifolds, logarithmic improve-

ments on the universal bound were proved by A. Selberg (in the constant curvature case) and by P. Berard (in the variable curvature case). The improvement in Theorem 11 is that it makes no curvature assumptions, only ergodicity of the geodesic flow. A precursor to Theorem 11 was a proof of the same statement in the special case of real analytic surfaces by Sogge–Toth–Zelditch, following general results of Sogge–Zelditch (see [Z17]).

Aside from L^∞ norms, the most attention has been devoted to L^4 norms, where it is conjectured that the L^4 norm is uniformly bounded. The best evidence for this pertains to (non- L^2) Eisenstein series on arithmetic quotients [Hum18] and to dihedral L^2 Maass forms (Humphries–Khan).

Lower bounds on numbers of nodal domains. The nodal set of ϕ_j is the hypersurface

$$\mathcal{N}_{\phi_j} = \{x \in M : \phi_j(x) = 0\}.$$

The nodal domains of ϕ_j are the connected components of $M \setminus \mathcal{N}_{\phi_j}$. Let $N(\phi_j)$ be the number of its nodal domains. The Courant upper bound states that the number $N(\phi_j)$ of nodal domains of an eigenfunction ϕ_j is bounded above by j . The question arises whether a (possibly generic) (M, g) possesses any sequence of eigenfunctions for which $N(\phi_{j_k}) \rightarrow \infty$. The first result on counting nodal domains by counting intersections with a curve was proved by Ghosh–Reznikov–Sarnak for $M = \mathbb{H}^2/\mathrm{SL}(2, \mathbb{Z})$ [GRS13]. Junehyuk Jung and the author have proved a general result showing that the number of nodal domains grows with the eigenvalue in certain ergodic cases, and QE plays an essential role in the proof. The setting is the same as that in Theorem 7. The following was proved in [JZ16] and then improved to the quantitative log estimate in a later article (see [Z18]). (See the discussion preceding Theorem 7 for notation.)

Theorem 12. *Let (M, J, σ) be a compact real Riemann surface with $\mathrm{Fix}(\sigma) \neq \emptyset$ and dividing. Let g be any σ -invariant Riemannian metric. Then for any orthonormal eigenbasis $\{\phi_j\}$ of $L^2_{\mathrm{even}}(M)$, respectively, $\{\psi_j\}$ of $L^2_{\mathrm{odd}}(M)$, one can find a density 1 subset A of $\mathbb{N}$ such that*

$$\lim_{\substack{j \rightarrow \infty \\ j \in A}} N(\phi_j) = \infty, \quad \lim_{\substack{j \rightarrow \infty \\ j \in A}} N(\psi_j) = \infty.$$

In fact,

$$N(\psi_j) \geq C_g (\log \lambda_j)^K \quad \text{for all } K < \frac{1}{6}.$$

The last statement uses the log-scale QE results of Han and Hezari–Rivière [HR16]. For generic negatively curved invariant metrics, the eigenvalues are simple (multiplicity one), and therefore all eigenfunctions are either even or odd.

The same result holds for any nonpositively curved surface with concave boundary (J. Jung–Zelditch, 2016), al-

though the quantitative lower bound is still open in that case. Hezari generalized the result to any domain with ergodic billiards [H18]. Jung and Seung uk Jang proved the result in QUE cases for the entire orthonormal basis of eigenfunctions. Jakobson–Naud proved a related result on infinite area hyperbolic surfaces ([Z17, Z18]).

All of these results pertain to dimension 2. The picture in dimensions ≥ 3 seems to be quite different. There are many topological types of 3-manifolds, so we restrict attention to unit tangent bundles SM of surfaces M . There exist natural “Kaluza–Klein” or bundle-metrics g_{KK} on SM , corresponding to a connection on $SM \rightarrow M$ and a metric g_M on M . Namely, the connection defines horizontal spaces. The horizontal part of g_{KK} is isometric to g_M under the natural projection, the vertical spaces (tangent to the fibers) are orthogonal to the horizontal spaces, and the fibers are geodesics of a fixed length. Jung–Zelditch recently proved that for generic Kaluza–Klein metrics g_{KK} (i.e., generic choices of g_M or connections on SM), the nodal set is regular (no self-intersections) and the number of nodal domains for every eigenfunction is 2. Thus, for a large class of topological 3-manifolds and metrics, the number of nodal domains is opposite to that of Theorem 12.

Sketch of the proof of Theorem 12. The proof may be instructive, because it combines almost all of the known results and tools of QE. Let M be a negatively curved surface with involution σ and let γ be the fixed point set of σ . We assume γ divides $M = M_+ \cup M_-$ into two pieces interchanged by σ , so that each piece is a surface with boundary. Even/odd eigenfunctions correspond to Neumann/Dirichlet on the pieces.

- (A) The first step is to show that the number $N(\phi_j)$ of nodal domains is $\geq \frac{1}{2}N(\phi_j|_\gamma)$, the number of zeros of ϕ_j on γ . This is purely topological and essentially uses that γ is a boundary.
- (B) The key is then to prove that Neumann eigenfunctions have a lot of zeros on $\gamma = \partial M_+$; respectively, Dirichlet eigenfunctions have many zeros of $\partial_\nu \phi_j = 0$ on γ . This is where QER (Theorem 7) is used. The proof is to show that

$$\int_\beta \phi_j ds \ll \int_\beta |\phi_j| ds \quad (16)$$

on any arc $\beta \subset \gamma$. To get a log lower bound, one proves (16) for $|\beta| \leq (\log \lambda_j)^{-1}$.

The proof of (16) uses the following estimates:

- (1) For any arc $\beta \subset \gamma$, $|\int_\beta \phi_j ds| \leq C\lambda_j^{-1/2}(\log \lambda_j)^{1/4}$.
- (2) For any $f \in C(\gamma)$, $\int_\gamma f \phi_j^2 ds \geq 1$ (Theorem 7).
- (3) $\|\phi_j\|_{L^\infty} \leq C \frac{\lambda_j^{1/2}}{\sqrt{\log \lambda_j}}$.

Item (1) is sometimes called a Kuznecov sum formula estimate. It holds on any curve on any surface for a subsequence of density 1 of the eigenfunctions. Item (2) is the QER theorem. The proof shows that one may let $f = \mathbf{1}_\beta$, the indicator function of any arc. Item (3) is the Bérard–Selberg sup norm estimate.

To prove that the restricted eigenfunctions $\phi_j|_\gamma$ have an unbounded number of sign changes as $j \rightarrow \infty$, one combines (2)–(3) to give

$$\int_\beta |\phi_j| ds \geq \frac{\|\phi_j\|_{L^2(\beta)}^2}{\|\phi_j\|_{L^\infty}} \geq C\lambda_j^{-1/2}(\log \lambda_j)^{1/2},$$

and this contradicts (1) if $\phi_j \geq 0$ on β .

Intuitively, (1) shows that ϕ_j either is small on γ or oscillates a lot on γ so that its integral is small; (2) shows that ϕ_j^2 is not small anywhere on γ . There is quite a difference between ϕ_j^2 being large everywhere and ϕ_j being large everywhere; (3) is used to soften this difference. Hence, (1) must be due to oscillations and in particular to zeros.

It is doubtful that the log estimates are sharp. It is a reasonable conjecture that in the negatively curved case, the number of zeros of $\phi_j|_\gamma$ is often roughly comparable with λ_j . So far, no techniques in nonarithmetic cases approach this number.

Log-scale equidistribution of complex zeros. The results in this section pertain to zeros of the Husimi measures (13). It turns out that phase space zeros in B^*M are simpler to study than physical space nodal sets in M but have basically the same interpretation: the least likely places for the particle to be. R. Chang and the author proved log-scale equidistribution results for phase space nodal sets both in the setting of Grauert tubes and in the setting of line bundles over Kähler manifolds (holomorphic automorphic forms). The results were partly motivated by a prior result of Lester–Matomaki–Radziwiłł for the special case of Hecke holomorphic forms for $SL(2, \mathbb{Z})$. The equidistribution law is quite different in the Grauert tube setting and in the Kähler setting, although the two are analogous (see [Z18] for references to the parallel results).

The complex zero set of $\phi_j^\mathbb{C}$ is the complex hypersurface

$$Z_j := \{\zeta \in B_{\tau_0}^*M : \phi_j^\mathbb{C}(\zeta) = 0\},$$

where $\phi_j^\mathbb{C}$ is the analytic continuation of ϕ_j to the co-ball bundle $B_{\tau_0}^*M$. The zero sets define currents $[Z_j]$ of integration in the sense that for every smooth $(n-1, n-1)$ test form $\eta \in \mathcal{D}^{n-1, n-1}(B_{\tau_0}^*M)$, we have that the pairing

$$\langle [Z_j], \eta \rangle := \int_{\mathbb{Z}_j} \eta = \int_{B_{\tau_0}^*M} \frac{i}{2\pi} \partial \bar{\partial} \log |\phi_j^\mathbb{C}|^2 \wedge \eta$$

is a well-defined closed current. In the special case $\eta =$

$f\omega^{n-1}$, the zero set defines a positive measure $|Z_j|$ by

$$\langle |Z_j|, f \rangle := \int_{Z_j} f \omega^{n-1}, \quad f \in C(B_{\tau_0}^* M).$$

In a prior article, the author proved that the “currents of integration” of QE eigenfunctions have a weak* limit law:

$$\frac{1}{\lambda_{j_k}} [Z_{j_k}] \rightharpoonup \frac{i}{\pi} \partial \bar{\partial} |\xi|_g \quad \text{weakly as currents on } B_{\tau_0}^* M$$

along a density one subsequence of eigenvalues λ_{j_k} . It is further proved that a similar convergence theorem holds on balls in $M_{\tau_0} \setminus M$ with logarithmically shrinking radii of size $(\log \lambda)^{-\gamma}$ for some explicit dimensional constant (independent of the frequency λ_j) $\gamma > 0$.

In the real domain, for (M, g) with ergodic geodesic flow, it is conjectured that nodal sets become uniformly distributed on M . Han has proved that the limit distribution is at least absolutely continuous with respect to dV_g if small scale QE holds [Han18].

References

- [A14] Anantharaman N. Le théorème d’ergodicité quantique. (French) [The quantum ergodicity theorem] *Chaos en mécanique quantique*, 101–146, Ed. Èc. Polytech., Palaiseau, 2014. MR3362293
- [A18] Anantharaman N. Delocalization of Schrödinger eigenfunctions, Proceedings ICM 2018, Vol. 1, 341–376, M. Viana (ed.), World Scientific, 2019.
- [D19] Dyatlov S. Notes on the fractal uncertainty principle, arXiv:1903.02599, 2019.
- [DJ18] Dyatlov S and Jin L. Semiclassical measures on hyperbolic surfaces have full support. *Acta Math.* 220 (2018), no. 2, 297–339. MR3849286
- [Gal19] Galkowski J. Defect measures of eigenfunctions with maximal L^∞ growth, *Ann. Institut Fourier* (to appear), arXiv:1704.01452, 2017.
- [GRS13] Ghosh A, Reznikov A, Sarnak P. Nodal domains of Maass forms I. *Geom. Funct. Anal.* 23(5):1515–1568, 2013. MR3102912
- [Go18] Gomes S. KAM Hamiltonians are not quantum ergodic, arXiv:1811.07718, 2018.
- [Han18] Han X. Distribution of the nodal sets of eigenfunctions on analytic manifolds. *J. Spectr. Theory* 8 (2018), no. 4, 1281–1293. MR3870068
- [Ha16] Hassell A. Ergodic billiards that are not quantum unique ergodic. With an appendix by the author and Luc Hillairet. *Ann. of Math. (2)* 171 (2010), no. 1, 605–619. MR2630052
- [H18] Hezari H. Applications of small scale quantum ergodicity in nodal sets. *Anal. PDE* 11 (2018), no. 4, 855–871. MR3749369
- [HR16] Hezari H, Rivière G. L^p norms, nodal sets, and quantum ergodicity. *Adv. Math.* 290 (2016), 938–966. MR3451943
- [Hum18] Humphries P. Equidistribution in shrinking sets and L^4 -norm bounds for automorphic forms. *Math. Ann.* 371, no. 3–4, (2018), 1497–1543. MR3831279
- [JZ16] Jung J, Zelditch S. Number of nodal domains and singular points of eigenfunctions of negatively curved surfaces with an isometric involution. *J. Differential Geom.* 102 (2016), no. 1, 37–66. MR3447086
- [S95] Sarnak P. Arithmetic quantum chaos. The Schur lectures (1992) (Tel Aviv), 183–236, Israel Math. Conf. Proc., 8, Bar-Ilan Univ., Ramat Gan, 1995. MR1321639
- [S11] Sarnak P. Recent progress on the quantum unique ergodicity conjecture. *Bull. Amer. Math. Soc. (N.S.)* 48 (2011), no. 2, 211–228. MR2774090
- [So14] Sogge CD. *Hangzhou lectures on eigenfunctions of the Laplacian*. Ann. of Math. Stud., 188. Princeton University Press, Princeton, NJ, 2014. MR3186367
- [TZ17] Toth JA, Zelditch S. Nodal intersections and geometric control, *J. Diff. Geom.* (to appear) arXiv:1708.05754, 2017.
- [Z17] Zelditch S. *Eigenfunctions of the Laplacian on a Riemannian manifold*. CBMS Regional Conference Series in Mathematics, 125. Amer. Math. Soc., Providence, RI, 2017. MR3729409
- [Z18] Zelditch S. *Local and global analysis of nodal sets*. Surveys in differential geometry 2017. Celebrating the 50th anniversary of the *Journal of Differential Geometry*, 365–406, Surv. Differ. Geom., 22, Int. Press, Somerville, MA, 2018. MR3838125
- [Zw] Zworski M. *Semi-classical analysis*, Graduate Studies in Mathematics, 138. Amer. Math. Soc., Providence, RI, 2012. MR2952218

Steve Zelditch

Credits

Opening figure is courtesy of Alexander Strohmaier.
Photo of author is courtesy of Northwestern University.