QUANTUM ERGODICITY AND MIXING OF EIGENFUNCTIONS

STEVE ZELDITCH

ABSTRACT. This article surveys mathematically rigorous results on quantum ergodic and mixing systems, with an emphasis on general results on asymptotics of eigenfunctions of the Laplacian on compact Riemannian manifolds.

Quantum ergodicity and mixing belong to the field of Quantum Chaos, which studies quantizations of 'chaotic' classical Hamiltonian systems. The basic questions are, how does the chaos of the classical dynamics impact on the eigenvalues eigenfunctions of the quantum Hamiltonian $\hat{H}$ and on and long time dynamics generated by $\hat{H}$?

These problems lie at the foundations of the semi-classical limit, i.e. the limit as the Planck constant $\hbar \to 0$ or the energy $E \to \infty$. More generally, one could ask what impact any dynamical feature of a classical mechanical system (e.g. complete integrability, KAM, ergodicity) has on the eigenfunctions and eigenvalues of the quantization.

Over the last 30 years or so, these questions have been studied rather systematically by both mathematicians and physicists. There is an extensive literature comparing classical and quantum dynamics of model systems, such as comparing the geodesic flow and wave group on a compact (or finite volume) hyperbolic surface, or comparing classical and quantum billiards on the Sinai billiard or the Bunimovich stadium, or comparing the discrete dynamical system generated by a hyperbolic torus automorphism and its quantization by the metaplectic representation. As these models indicate, the basic problems and phenomena are richly embodied in simple, low-dimensional examples in much the same way that two-dimensional toy statistical mechanical models already illustrate complex problems on phase transitions. The principles established for simple models should apply to far more complex systems such as atoms and molecules in strong magnetic fields.

The conjectural picture which has emerged from many computer experiments and heuristic arguments on these simple model systems is roughly that there exists a length scale in which quantum chaotic systems exhibit universal behavior. At this length scale, the eigenvalues resemble eigenvalues of random matrices of large size and the eigenfunctions resemble random waves. A small sample of the original physics articles suggesting this picture is [B, BGS, FP, Gu, H, A].

This article reviews some of the rigorous mathematical results in quantum chaos, particularly the rigorous results on eigenfunctions of quantizations of classically ergodic or mixing systems. They support the conjectural picture of random waves up to two moments, i.e. on the level of means and variances. A few results also exist on higher moments in very special cases. But from the mathematical point of view, the conjectural links to random matrices or random waves remain very much open at this time. A key difficulty is that the length scale on which universal behavior should occur is very far below the resolving power of any known

Date: January 25, 2005.
Research partially supported by NSF grant #DMS 0302518.

mathematical techniques, even in the simplest model problems. The main evidence for the random matrix and random wave connections comes from numerous computer experiments of model cases in the physics literature. We will not review numerical results here, but to get a well-rounded view of the field it is important to understand the computer experiments (see [BSS, Bar, KH] for some examples).

The model quantum systems that have been most intensively studied in mathematical quantum chaos are Laplacians or Schrödinger operators on compact (or finite volume) Riemannian manifolds, with or without boundary, and quantizations of symplectic maps on compact Kähler manifolds. Similar techniques and results apply in both settings, so for the sake of coherence we concentrate on the Laplacian on a compact Riemannian manifold with 'chaotic' geodesic flow and only briefly allude to the setting of 'quantum maps'. Additionally, two main kinds of methods are in use: (i) methods of semi-classical (or microlocal) analysis, which apply to general Laplacians, and (ii) methods of number theory and automorphic forms, which apply to arithmetic models such as arithmetic hyperbolic manifolds or quantum cat maps. Arithmetic models are far more 'explicitly solvable' than general chaotic systems, and the results obtained for them are far sharper than the results of semi-classical analysis. This article is primarily devoted to the general results obtained by semi-classical analysis; for results in arithmetic quantum chaos, we refer to [M].

1. Wave group and geodesic flow

The model quantum Hamiltonians we will discuss are Laplacians Δ on compact Riemannian manifolds (M, g) (with or without boundary). The classical phase space in this setting is the cotangent bundle T^*M of M, equipped with its canonical symplectic form $\sum_i dx_i \wedge d\xi_i$. The metric defines the Hamiltonian $H(x, \xi) = |\xi|_g = \sqrt{\sum_{ij=1}^n g^{ij}(x)\xi_i\xi_j}$ on T^*M, where $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$, $[g^{ij}]$ is the inverse matrix to $[g_{ij}]$. We denote the volume density of (M, g) by $dVol$ and the corresponding inner product on $L^2(M)$ by $\langle f, g\rangle$. The unit (co-) ball bundle is denoted $B^*M = \{(x, \xi) : |\xi| \leq 1\}$.

The Hamiltonian flow Φ^t of H is the geodesic flow. By definition, $\Phi^t(x, \xi) = (x_t, \xi_t)$, where (x_t, ξ_t) is the terminal tangent vector at time t of the unit speed geodesic starting at x in the direction ξ. Here and below, we often identify T^*M with the tangent bundle TM using the metric to simplify the geometric description. The geodesic flow preserves the energy surfaces $\{H = E\}$ which are the co-sphere bundles $S^*_E M$. Due to the homogeneity of H, the flow on any energy surface $\{H = E\}$ is equivalent to that on the co-sphere bundle $S^*M = \{H = 1\}$. (This homogeneity could be broken by adding a potential $V \in C^\infty(M)$ to form a semi-classical Schrödinger operator $-\hbar^2\Delta + V$, whose underlying Hamiltonian flow is generated by $|\xi|_g^2 + V(x)$.)

The quantization of the Hamiltonian H is the square root $\sqrt{\Delta}$ of the positive Laplacian,

$$\Delta = -\frac{1}{\sqrt{g}} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} g^{ij} g \frac{\partial}{\partial x_j}$$

of (M, g). Here, $g = \det[g_{ij}]$. We choose to work with $\sqrt{\Delta}$ rather than Δ since the former generates the wave

$$U_t = e^{it\sqrt{\Delta}},$$

which is the quantization of the geodesic flow Φ^t.

By the last statement we mean that U_t is related to Φ^t in several essentially equivalent ways:

(1) singularities of waves, i.e. solutions $U_t\psi$ of the wave equation, propagate along geodesics;
(2) U_t is a Fourier integral operator (= quantum map) associated to the canonical relation defined by the graph of Φ^t in $T^*M \times T^*M$;
(3) Egorov's theorem holds.

We only define the latter since it plays an important role in studying eigenfunctions. As with any quantum theory, there is an algebra of observables on the Hilbert space $L^2(M, dvol_g)$ which quantizes T^*M. Here, $dvol_g$ is the volume form of the metric. The algebra is that $\Psi^*(M)$ of pseudodifferential operators ψDO's of all orders, though we often restrict to the subalgebra Ψ^0 of ψDO's of order zero. We denote by $\Psi^m(M)$ the subspace of pseudodifferential operators of order m. The algebra is defined by constructing a quantization Op from an algebra of symbols $a \in S^m(T^*M)$ of order m (polyhomogeneous functions on $T^*M\backslash 0$) to Ψ^m. The map Op is not unique. In the reverse direction is the symbol map $\sigma_A : \Psi^m \to S^m(T^*M)$ which takes an operator $Op(a)$ to the homogeneous term a_m of order m in a. For background we refer to [HoIII, DSj].

Egorov's theorem for the wave group concerns the conjugations

$$\alpha_t(A) := U_t A U_t^*, \quad A \in \Psi^m(M). \tag{1}$$

Such a conjugation defines the quantum evolution of observables in the Heisenberg picture, and since the early days of quantum mechanics it was known to correspond to the classical evolution

$$V_t(a) := a \circ \Phi^t \tag{2}$$

of observables $a \in C^\infty(S^*M)$. Egorov's theorem is the rigorous version of this correspondence: it says that α_t defines an order-preserving automorphism of $\Psi^*(M)$, i.e. $\alpha_t(A) \in \Psi^m(M)$ if $A \in \Psi^m(M)$, and that

$$\sigma_{U_t A U_t^*}(x, \xi) = \sigma_A(\Phi^t(x, \xi)) := V_t(\sigma_A), \quad (x, \xi) \in T^*M\backslash 0. \tag{3}$$

This formula is almost universally taken to be the definition of quantization of a flow or map in the physics literature.

The key difficulty in quantum chaos is that it involves a comparison between long-time dynamical properties of Φ^t and U_t through the symbol map and similar classical limits. The classical dynamics defines the 'principal symbol' behavior of U_t and the 'error' $U_t A U_t^* - Op(\sigma_A \circ \Phi^t)$ typically grows exponentially in time. This is just the first example of a ubiquitous 'exponential barrier' in the subject.

2. Eigenvalues and eigenfunctions of Δ

The eigenvalue problem on a compact Riemannian manifold

$$\Delta \varphi_j = \lambda_j^2 \varphi_j, \qquad \langle \varphi_j, \varphi_k \rangle = \delta_{jk}$$

is dual under the Fourier transform to the wave equation. Here, $\{\varphi_j\}$ is a choice of orthonormal basis of eigenfunctions, which is not unique if the eigenvalues have multiplicities > 1.

The individual eigenfunctions are difficult to study directly, and so one generally forms the spectral projections kernel,

$$E(\lambda, x, y) = \sum_{j:\lambda_j \leq \lambda} \varphi_j(x)\varphi_j(y). \tag{4}$$

Semi-classical asymptotics is the study of the $\lambda \to \infty$ limit of the spectral data $\{\varphi_j, \lambda_j\}$ or of $E(\lambda, x, y)$. The (Schwartz) kernel of the wave group can be represented in terms of the spectral data by

$$U_t(x, y) = \sum_j e^{it\lambda_j} \varphi_j(x)\varphi_j(y),$$

or equivalently as the Fourier transform $\int_{\mathbb{R}} e^{it\lambda} dE(\lambda, x, y)$ of the spectral projections. Hence spectral asymptotics is often studied through the large time behavior of the wave group.

The link between spectral theory and geometry, and the source of Egorov's theorem for the wave group, is the construction of a parametrix (or WKB formula) for the wave kernel. For small times t, the simplest is the Hadamard parametrix,

$$U_t(x, y) \sim \int_0^\infty e^{i\theta(r^2(x,y) - t^2)} \sum_{k=0}^\infty U_k(x, y)\theta^{\frac{d-3}{2} - k} d\theta \qquad (t < \text{inj}(M, g)) \tag{5}$$

where $r(x, y)$ is the distance between points, $U_0(x, y) = \Theta^{-\frac{1}{2}}(x, y)$ is the volume 1/2-density, $inj(M, g)$ is the injectivity radius, and the higher Hadamard coefficients are obtained by solving transport equations along geodesics. The parametrix is asymptotic to the wave kernel in the sense of smoothness, i.e. the difference of the two sides of (5) is smooth. The relation (5) may be iterated using $U_{tm} = U_t^m$ to obtain a parametrix for long times. This is obviously complicated and not necessarily the best long time parametrix construction, but it illustrates again the difficulty of a long time analysis.

2.1. Weyl law and local Weyl law.

A fundamental and classical result in spectral asymptotics is Weyl's law on counting eigenvalues:

$$N(\lambda) = \#\{j : \lambda_j \leq \lambda\} = \frac{|B_n|}{(2\pi)^n} Vol(M, g)\lambda^n + O(\lambda^{n-1}). \tag{6}$$

Here, $|B_n|$ is the Euclidean volume of the unit ball and $Vol(M, g)$ is the volume of M with respect to the metric g. An equivalent formula which emphasizes the correspondence between classical and quantum mechanics is:

$$TrE_\lambda = \frac{Vol(|\xi|_g \leq \lambda)}{(2\pi)^n}, \tag{7}$$

where Vol is the symplectic volume measure relative to the natural symplectic form $\sum_{j=1}^n dx_j \wedge d\xi_j$ on T^*M. Thus, the dimension of the space where $H = \sqrt{\Delta}$ is $\leq \lambda$ is asymptotically the volume where its symbol $|\xi|_g \leq \lambda$.

The remainder term in Weyl's law is sharp on the standard sphere, where all geodesics are periodic, but is not sharp on (M, g) for which the set of periodic geodesics has measure zero (Duistermaat-Guillemin [DG], Ivrii). When the set of periodic geodesics, has measure zero (as is the case for ergodic systems), one has

$$N(\lambda) = \#\{j : \lambda_j \leq \lambda\} = \frac{|B_n|}{(2\pi)^n} Vol(M, g)\lambda^n + o(\lambda^{n-1}). \tag{8}$$

For background, see [HoIV] ch. XXIX. The remainder is then of small order than the derivative of the principal term, and one then has asymptotics in shorter intervals:

$$N([\lambda, \lambda + 1]) = \#\{j : \lambda_j \in [\lambda, \lambda + 1]\} = n\frac{|B_n|}{(2\pi)^n} Vol(M, g)\lambda^{n-1} + o(\lambda^{n-1}). \tag{9}$$

Physicists tend to write $\lambda \sim h^{-1}$ and to average over intervals of this width. Then mean spacing between the eigenvalues in this interval is $\sim C_n Vol(M, g)^{-1}\lambda^{-(n-1)}$, where C_n is a constant depending on the dimension.

An important generalization is the *local Weyl law* concerning the traces $TrAE(\lambda)$ where $A \in \Psi^m(M)$. It asserts that

$$\sum_{\lambda_j \leq \lambda} \langle A\varphi_j, \varphi_j\rangle = \frac{1}{(2\pi)^n} \int_{B^*M} \sigma_A dx d\xi)\lambda^n + O(\lambda^{n-1}). \tag{10}$$

There is also a pointwise local Weyl law:

$$\sum_{\lambda_j \leq \lambda} |\varphi_j(x)|^2 = \frac{1}{(2\pi)^n} |B^n|\lambda^n + R(\lambda, x), \tag{11}$$

where $R(\lambda, x) = O(\lambda^{n-1})$ uniformly in x. Again, when the periodic geodesics form a set of measure zero in S^*M, one could average over the shorter interval $[\lambda, \lambda + 1]$. Combining the Weyl and local Weyl law, we find the surface average of σ_A is a limit of traces:

$$\begin{aligned} \omega(A) &:= \frac{1}{\mu(S^*M)} \int_{S^*M} \sigma_A d\mu \\ &= \lim_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} \langle A\varphi_j, \varphi_j\rangle \end{aligned} \tag{12}$$

Here, μ is the *Liouville measure* on S^*M, i.e. the surface measure $d\mu = \frac{dxd\xi}{dH}$ induced by the Hamiltonian $H = |\xi|_g$ and by the symplectic volume measure $dxd\xi$ on T^*M.

2.2. **Problems on asymptotics eigenfunctions.** Eigenfunctions arise in quantum mechanics as stationary states, i.e. states ψ for which the probability measure $|\psi(t, x)|^2 dvol$ is constant where $\psi(t, x) = U_t\psi(x)$ is the evolving state. This follows from the fact that

$$U_t\varphi_k = e^{it\lambda_k}\varphi_k \tag{13}$$

and that $|e^{it\lambda_k}| = 1$. They are the basic modes of the quantum system. One would like to know the behavior as $\lambda_j \to \infty$ (or $\hbar \to 0$ in the semi-classical setting) of invariants such as:

(1) Matrix elements $\langle A\varphi_j, \varphi_j\rangle$ of observables in this state;
(2) Transition elements $\langle A\varphi_i, \varphi_j\rangle$ between states;
(3) Size properties as measured by L^p norms $||\varphi_j||_{L^p}$;

(4) Value distribution as measured by the distribution function $Vol\{x \in M : |\varphi_j(x)|^2 > t\}$.
(5) Shape properties, e.g. distribution of zeros and critical points of φ_j.

Let us introduce some problems which have motivated much of the work in this area.

PROBLEM **1.** *Let $\mathcal{Q}$ denote the set of 'quantum limits', i.e. weak* limit points of the sequence $\{\Phi_k\}$ of distributions on the classical phase space S^*M, defined by*

$$\int_X a d\Phi_k := \langle Op(a)\varphi_k, \varphi_k\rangle$$

where $a \in C^\infty(S^*M)$.

The set $\mathcal{Q}$ is independent of the definition of Op. It follows almost immediately from Egorov's theorem that $\mathcal{Q} \subset \mathcal{M}_I$, where $\mathcal{M}_I$ is the convex set of invariant probability measures for the geodesic flow. Furthermore, they are time-reversal invariant, i.e. invariant under $(x, \xi) \to (x, -\xi)$ since the eigenfunctions are real-valued.

To see this, it is helpful to introduce the linear functionals on Ψ^0

$$\rho_k(A) = \langle Op(a)\varphi_k, \varphi_k\rangle. \tag{14}$$

We observe that $\rho_k(I) = 1$, that $\rho_k(A) \geq 0$ if $A =\geq 0$ and that

$$\rho_k(U_t A U_t^*) = \rho_k(A). \tag{15}$$

Indeed, if $A \geq 0$ then $A = B^*B$ for some $B \in \Psi^0$ and we can move B^* to the right side. Similarly (15) is proved by moving U_t to the right side and using (13). These properties mean that ρ_j is an *invariant state* on the algebra Ψ^0. More precisely, one should take the closure of Ψ^0 in the operator norm. An invariant state is the analogue in quantum statistical mechanics of an invariant probability measure.

The next important fact about the states ρ_k is that any weak limit of the sequence $\{\rho_k\}$ on Ψ^0 is a probability measure on $C(S^*M)$, i.e. a positive linear functional on $C(S^*M)$ rather than just a state on Ψ^0. This follows from the fact that $\langle K\varphi_j, \varphi_j\rangle \to 0$ for any compact operator K, and so any limit of $\langle A\varphi_k, \varphi_k\rangle$ is equally a limit of $\langle (A+K)\varphi_k, \varphi_k\rangle$. Hence any limit is bounded by $\inf_K ||A+K||$ (the infimum taken over compact operators), and for any $A \in \Psi^0$, $||\sigma_A||_{L^\infty} = \inf_K ||A+K||$. Hence any weak limit is bounded by a constant times $||\sigma_A||_{L^\infty}$ and is therefore continuous on $C(S^*M)$. It is a positive functional since each ρ_j is and hence any limit is a probability measure. By Egorov's theorem and the invariance of the ρ_k, any limit of $\rho_k(A)$ is a limit of $\rho_k(Op(\sigma_A \circ \Phi^t))$ and hence the limit measure is invariant.

Problem I is thus to identify which invariant measures in $\mathcal{M}_I$ show up as weak limits of the functionals ρ_k or equivalently the distributions $d\Phi_k$. The weak limits reflect the concentration and oscillation properties of eigenfunctions. Here are some possibilities:

(1) Normalized Liouville measure. In fact, the functional ω of (12) is also a state on Ψ^0 for the reason explained above. A subsequence $\{\varphi_{j_k}\}$ of eigenfunctions is considered diffuse if $\rho_{j_k} \to \omega$.
(2) A periodic orbit measure μ_γ defined by $\mu_\gamma(A) = \frac{1}{L_\gamma}\int_\gamma \sigma_A ds$ where L_γ is the length of γ. A sequence of eigenfunctions for which $\rho_{k_j} \to \mu_\gamma$ obviously concentrates (or strongly 'scars') on the closed geodesic.
(3) A finite sum of periodic orbit measures.

(4) A delta-function along an invariant Lagrangian manifold $\Lambda \subset S^*M$. The associated eigenfunctions are viewed as *localizing* along Λ.
(5) A more general measure which is singular with respect to $d\mu$.

All of these possibilities can and do happen in different examples. If $d\Phi_{k_j} \to \omega$ then in particular, we have

$$\frac{1}{Vol(M)} \int_E |\varphi_{k_j}(x)|^2 dVol \to \frac{Vol(E)}{Vol(M)}$$

for any measurable set E whose boundary has measure zero. In the interpretation of $|\varphi_{k_j}(x)|^2 dVol$ as the probability density of finding a particle of energy λ_k^2 at x, this says that the sequence of probabilities tends to uniform measure.

However, $d\Phi_{k_j} \to \omega$ is much stronger since it says that the eigenfunctions become diffuse on the energy surface S^*M and not just on the configuration space M. As an example, consider the flat torus $\mathbb{R}^n/\mathbb{Z}^n$. An orthonormal basis of eigenfunctions is furnished by the standard exponentials $e^{2\pi i \langle k,x\rangle}$ with $k \in \mathbb{Z}^n$. Obviously, $|e^{2\pi i\langle k,x\rangle}|^2 = 1$, so the eigenfunctions are already diffuse in configuration space. On the other hand, they are far from diffuse in phase space, and localize on invariant Lagrange tori in S^*M. Indeed, by definition of pseudodifferential operator, $Ae^{2\pi i\langle k,x\rangle} = a(x,k)e^{2\pi i\langle k,x\rangle}$ where $a(x,k)$ is the complete symbol. Thus,

$$\langle Ae^{2\pi i\langle k,x\rangle}, e^{2\pi i\langle k,x\rangle}\rangle = \int_{\mathbb{R}^n/\mathbb{Z}^n} a(x,k)dx \sim \int_{\mathbb{R}^n/\mathbb{Z}^n} \sigma_A(x, \frac{k}{|k|})dx.$$

A subsequence $e^{2\pi i\langle k_j,x\rangle}$ of eigenfunctions has a weak limit if and only if $\frac{k_j}{|k_j|}$ tends to a limit vector ξ_0 in the unit sphere in $\mathbb{R}^n$. In this case, the associated weak* limit is $\int_{\mathbb{R}^n/\mathbb{Z}^n} \sigma_A(x,\xi_0)dx$, i.e. the delta-function on the invariant torus $T_{\xi_0} \subset S^*M$ defined by the constant momentum condition $\xi = \xi_0$. The eigenfunctions are said to localize under this invariant torus for Φ^t.

The flat torus is a model of a completely integrable system, on both the classical and quantum levels. Another example is that of the standard round sphere S^n. In this case, the author and D. Jakobson showed that absolutely any invariant measure $\nu \in \mathcal{M}_I$ can arise as a weak limit of a sequence of eigenfunctions. This reflects the huge degeneracy (multiplicities) of the eigenvalues.

On the other hand, if the geodesic flow is ergodic one would expect the eigenfunctions to be diffuse in phase space. In the next section, we will discuss the rigorous results on this problem.

Off-diagonal matrix elements

$$\rho_{jk}(A) = \langle A\varphi_i, \varphi_j\rangle \tag{16}$$

are also important as transition amplitudes between states. They no longer define states since $\rho_{jk}(I) = 0$, are no longer positive, and are no longer invariant. Indeed, $\rho_{jk}(U_tAU_t^*) = e^{it(\lambda_j-\lambda_k)}\rho_{jk}(A)$, so they are eigenvectors of the automorphism α_t of (1). A sequence of such matrix elements cannot have a weak limit unless the spectral gap $\lambda_j - \lambda_k$ tends to a limit $\tau \in \mathbb{R}$. In this case, by the same discussion as above, any weak limit of the functionals ρ_{jk} will be an eigenmeasure of the geodesic flow which transforms by $e^{i\tau t}$ under the action of Φ^t. Examples of such eigenmeasures are orbital Fourier coefficients $\frac{1}{L_\gamma}\int_0^{L_\gamma} e^{-i\tau t}\sigma_A(\Phi^t(x,\xi))dt$

along a periodic orbit. Here $\tau \in \frac{2\pi}{L_\gamma}\mathbb{Z}$. We denote by $\mathcal{Q}_\tau$ such eigenmeasures of the geodesic flow. Problem 1 has the following extension to off-diagonal elements:

PROBLEM **2.** *Determine the set $\mathcal{Q}_\tau$ of 'quantum limits', i.e. weak* limit points of the sequence $\{\Phi_{kj}\}$ of distributions on the classical phase space S^*M, defined by*

$$\int_X a d\Phi_{kj} := \langle Op(a)\varphi_k, \varphi_j \rangle$$

where $\lambda_j - \lambda_k = \tau + o(1)$ and where $a \in C^\infty(S^*M)$, or equivalently of the functionals ρ_{jk}.

As will be discussed in §4, the asymptotics of off-diagonal elements depends on the weak mixing properties of the geodesic flow and not just its ergodicity.

Matrix elements of eigenfunctions are quadratic forms. More 'nonlinear' problems involve the L^p norms or the distribution functions of eigenfunctions. Estimates of the L^∞ norms can be obtained from the local Weyl law (10). Since the jump in the the left hand side at λ is $\sum_{j:\lambda_j=\lambda} |\varphi_j(x)|^2$ and the jump in the right hand side is the jump of $R(\lambda, x)$, this implies

$$\sum_{j:\lambda_j=\lambda} |\varphi_j(x)|^2 = O(\lambda^{n-1}) \implies ||\varphi_j||_{L^\infty} = O(\lambda^{\frac{n-1}{2}}). \tag{17}$$

For general L^p-norms, the following bounds hold on any compact Riemannian manifold [Sog]:

$$\frac{||\varphi_j||_p}{||\varphi||_2} = O(\lambda^{\delta(p)}), \quad 2 \le p \le \infty. \tag{18}$$

where

$$\delta(p) = \begin{cases} n(\frac{1}{2} - \frac{1}{p}) - \frac{1}{2}, & \frac{2(n+1)}{n-1} \le p \le \infty \\ \frac{n-1}{2}(\frac{1}{2} - \frac{1}{p}), & 2 \le p \le \frac{2(n+1)}{n-1}. \end{cases} \tag{19}$$

These estimates are sharp on the unit sphere $S^n \subset \mathbb{R}^{n+1}$. The extremal eigenfunctions are the zonal spherical harmonics, which are the L^2-normalized spectral projection kernels $\frac{\Pi_N(x,x_0)}{||\Pi_N(\cdot,x_0)||}$ centered at any x_0. However, they are not sharp for generic (M, g), and it is natural to ask how 'chaotic dynamics' might influence L^p norms.

PROBLEM **3.** *Improve the estimates $\frac{||\varphi_j||_p}{||\varphi||_2} = O(\lambda^{\delta(p)})$ for (M, g) which ergodic or mixing geodesic flow.*

In [SogZ] it is proved that if a sequence of eigenfunctions attains the bounds in (17), then there must exist a point x_0 so that a positive measure of geodesics starting at x_0 in $S^*_{x_0}M$ return to x_0 at a fixed time T. In the real analytic case, all return so x_0 is a perfect recurrent point. In dimension 2, such a perfect recurrent point cannot occur if the geodesic flow is ergodic; hence $||\varphi_j||_{L^\infty} = o(\lambda^{\frac{n-1}{2}})$ on any real analytic surface with ergodic geodesic flow. This shows that none of the L^p estimates above the critical index are sharp for real analytic surfaces with ergodic geodesic flow, and the problem is the extent to which they can be improved.

The random wave model (see §6) predicts that eigenfunctions of Riemannian manifolds with chaotic geodesic flow should have the bounds $||\varphi_\lambda||_{L^p} = O(1)$ for $p < \infty$ and that $||\varphi_\lambda||_{L^\infty} < \sqrt{\log \lambda}$. But there are no rigorous estimates at this time close to such predictions. The best general estimate to date on negatively curved compact manifolds (which are models of chaotic geodesic flow) is just the logarithmic improvement $||\varphi_j||_{L^\infty} = O(\frac{\lambda^{n-1}}{\log \lambda})$ on the standard remainder term in the local Weyl law. This was known for compact hyperbolic manifolds from the Selberg trace formula, and similar estimates hold manifolds without conjugate points [Ber]. The exponential growth of the geodesic flow again causes a barrier in improving the estimate beyond the logarithm. In the analogous setting of quantum 'cat maps', which are models of chaotic classical dynamics, there exist arbitrarily large eigenvalues with multiplicities of the order $O(\frac{\lambda^{n-1}}{\log \lambda})$; the L^∞-norm of the L^2-normalized projection kernel onto an eigenspace of this multiplicity is of order the square root of the multiplicity ([FND]). This raises doubt that the logarithmic estimate can be improved by general dynamical arguments. Further discussion of L^∞-norms, as well as zeros, will be given at the end of §3 for ergodic systems.

3. Quantum ergodicity

In this section, we discuss results on the problems stated above when the geodesic flow of (M, g) is assumed to be ergodic. Let us recall that this means that Liouville measure is an ergodic measure for Φ^t. This is a spectral property of the operator V_t of (2) on $L^2(S^*M, d\mu)$, namely that V_t has 1 as an eigenvalue of multiplicity one. That is, the only invariant L^2 functions (with respect to Liouville measure) are the constant functions. This implies that the only invariant sets have Liouville measure 0 or 1 and (Birkhoff's ergodic theorem) that time averages of functions are constant almost everywhere (equal to the space average).

In this case, there is a general result which originated in the work of A. I. Schnirelman [Sh.1, Sh.2]:

Theorem 1. [Sh.1, Sh.2, Z0, Z.1, Z.3, CV, Su, GL, ZZw] *Let (M, g) be a compact Riemannian manifold (possibly with boundary), and let $\{\lambda_j, \varphi_j\}$ be the spectral data of its Laplacian Δ. Then the geodesic flow G^t is ergodic on $(S^*M, d\mu)$ if and only if, for every $A \in \Psi^o(M)$, we have:*

(1) $\lim_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} |(A\varphi_j, \varphi_j) - \omega(A)|^2 = 0.$

(2) $(\forall \epsilon)(\exists \delta) \limsup_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\substack{j\neq k: \lambda_j, \lambda_k \leq \lambda \\ |\lambda_j - \lambda_k| < \delta}} |(A\varphi_j, \varphi_k)|^2 < \epsilon$

This implies that there exists a subsequence $\{\varphi_{j_k}\}$ of eigenfunctions whose indices j_k have counting density one for which $\langle A\varphi_{j_k}, \varphi_{j_k}\rangle \to \omega(A)$. We will call the eigenfunctions in such a sequence 'ergodic eigenfunctions'. One can sharpen the results results by averaging over eigenvalues in the shorter interval $[\lambda, \lambda + 1]$ rather than in $[0, \lambda]$.

There is also an ergodicity result for boundary values of eigenfunctions on domains with boundary and with Dirichlet, Neumann or Robin boundary conditions [GL, HZ, Bu]. This corresponds to the fact that the billiard map on $B^*\partial M$ is ergodic.

The first statement (1) is essentially a convexity result. It remains true if one replaces the square by any convex function φ on the spectrum of A,

$$\frac{1}{N(E)} \sum_{\lambda_j \leq E} \varphi(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) \to 0. \tag{20}$$

Before sketching a proof, we point out a somewhat heuristic 'picture proof' of the theorem. Namely, ergodicity of the geodesic flow is equivalent to the statement that Liouville measure is an extreme point of the compact convex set $\mathcal{M}_I$. In fact, it further implies that ω is an extreme point of the compact convex set $\mathcal{E}_{\mathbb{R}}$ of invariant states for α_t of (1); see [Ru] for §6.3 for background. But the local Weyl law says that ω is also the limit of the convex combination $\frac{1}{N(E)}\sum_{\lambda_j\le E}\rho_j$. An extreme point cannot be written as a convex combination of other states unless all the states in the combination are equal to it. In our case, ω is only a limit of convex combinations so it need not (and does not) equal each term. However, almost all terms in the sequence must tend to ω, and that is equivalent to (1).

Sketch of Proof of (1) As mentioned above, this is a convexity result and with no additional effort we can consider more general sums of the form We then have

$$\sum_{\lambda_j\le E}\varphi(\langle A\varphi_k,\varphi_k\rangle-\omega(A))=\sum_{\lambda_j\le E}\varphi(\langle\langle A\rangle_T-\omega(A)\varphi_k,\varphi_k\rangle). \tag{21}$$

We then apply the Peierls–Bogoliubov inequality

$$\sum_{j=1}^{n}\varphi((B\varphi_j,\varphi_j))\le \operatorname{Tr}\varphi(B)$$

with $B=\Pi_E[\langle A\rangle_T-\omega(A)]\Pi_E$ to get:

$$\sum_{\lambda_j\le E}\varphi(\langle\langle A\rangle_T-\omega(A)\varphi_k,\varphi_k\rangle)\le \operatorname{Tr}\varphi(\Pi_E[\langle A\rangle_T-\omega(A)]\Pi_E). \tag{22}$$

Here, Π_E is the spectral projection for $\hat{H}$ corresponding to the interval $[0,E]$. From the Berezin inequality [Si, (8.18)] we then have (if $\varphi(0)=0$):

$$\begin{aligned}\frac{1}{N(E)}\operatorname{Tr}\varphi(\Pi_E[\langle A\rangle_T-\omega(A)]\Pi_E)&\le\frac{1}{N(E)}\operatorname{Tr}\Pi_E\varphi([\langle A\rangle_T-\omega(A)])\Pi_E\\&=\omega_E(\varphi(\langle A\rangle_T-\omega(A))).\end{aligned}\tag{1.6.7}$$

As long as φ is smooth, $\varphi(\langle A\rangle_T-\omega(A))$ is a pseudodifferential operator of order zero with principal symbol $\varphi(\langle\sigma_A\rangle_T-\omega(A))$. By the assumption that $\omega_E\to\omega$ we get

$$\lim_{E\to\infty}\frac{1}{N(E)}\sum_{\lambda_j\le E}\varphi(\langle A\varphi_k,\varphi_k\rangle-\omega(A))\le\int_{\{H=1\}}\varphi(\langle\sigma_A\rangle_T-\omega(A))d\mu. \tag{1.6.8}$$

As $T\to\infty$ the right side approaches $\varphi(0)$ by the dominated convergence theorem and by Birkhoff's ergodic theorem. Since the left hand side is independent of T, this implies that

$$\lim_{E\to\infty}\frac{1}{N(E)}\sum_{\lambda_j\le E}\varphi(\langle A\varphi_k,\varphi_k\rangle-\omega(A))=0$$

for any smooth convex φ on $\operatorname{Spec}(A)$ with $\varphi(0)=0$. □

As mentioned above, the statement (1) is equivalent to saying that there is a subsequence $\{\varphi_{j_k}\}$ of counting density one for which $\rho_{j_k}\to\omega$. The above proof does not and cannot settle the question whether there exist exceptional sparse subsequences of eigenfunctions of density zero tending to other invariant measures. To see this, we observe that the proof is so general that it applies to seemingly very different situations. In place of the distributions $\{\Phi_j\}$ we may consider the set μ_γ of periodic orbit measures for a hyperbolic flow on a compact manifold X. That is, $\mu_\gamma(f)=\frac{1}{T_\gamma}\int_\gamma f$ for $f\in C(X)$, where γ is a closed orbit and T_γ is its period. According to the Bowen–Margulis equidistribution theorem for closed orbits of hyperbolic flows, we have

$$\frac{1}{\Pi(T)}\sum_{\gamma:T_\gamma\leq T}\frac{1}{|\det(I-P_\gamma)|}\mu_\gamma\to\mu$$

where as above μ is the Liouville measure, where P_γ is the linear Poincaré map and where $\Pi(T)$ is the normalizing factor which makes the left side a probability measure, i.e. defined by the integral of 1 against the sum. An exact repetition of the previous argument shows that up to a sparse subsequence of γ's, $\mu_\gamma\to\mu$ individually. Yet clearly, the whole sequence does not tend to $d\mu$: for instance one could choose the sequence of iterates γ^k of a fixed closed orbit.

3.1. **Quantum ergodicity in terms of operator time and space averages.** The first part of the result above may be reformulated as a relation between operator time and space averages.

Definition *Let $A\in\Psi^0$ be an observable and define its time average to be:*

$$\langle A\rangle:=\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}U_t^*AU_tdt$$

and its space average to be scalar operator

$$\omega(A)\cdot I$$

Here, the limit is taken in the weak operator topology (i.e. one matrix element at a time). To see what is involved we consider matrix elements with respect to the eigenfunctions. We have

$$(\frac{1}{2T}\int_{-T}^{T}U_t^*AU_tdt\varphi_i,\varphi_j)=\frac{\sin T(\lambda_i-\lambda_j)}{T(\lambda_i-\lambda_j)}(A\varphi_i,\varphi_j) \tag{1.3.2}$$

from which it is clear that the matrix element tends to zero as $T\to\infty$ unless $\lambda_i=\lambda_j$. However, there is no uniformity in the rate at which it goes to zero since the spacing $\lambda_i-\lambda_j$ could be uncontrollably small.

In these terms, Theorem 1 (1) says that:

$$\langle A\rangle=\omega(A)I+K,\qquad\text{where}\qquad\lim_{\lambda\to\infty}\omega_\lambda(K^*K)\to 0, \tag{23}$$

where $\omega_\lambda(A)=TrE(\lambda)A$. Thus, the time average equals the space average plus a term K which is semi-classically small in the sense that its Hilbert-Schmidt norm square $||E_\lambda K||_{HS}^2$ in the span of the eigenfunctions of eigenvalue $\leq\lambda$ is $o(N(\lambda))$.

This is not exactly equivalent to Theorem 1 (1) since it is independent of the choice of orthonormal basis, while the previous result depends on the choice of basis. However, when all eigenvalues have multiplicity one, then the two are equivalent. To see the equivalence, note that $\langle A\rangle$ commutes with $\sqrt{\Delta}$ and hence is diagonal in the basis $\{\varphi_j\}$ of joint eigenfunctions of $\langle A\rangle$ and of U_t. Hence K is the diagonal matrix with entries $\langle A\varphi_k,\varphi_k\rangle-\omega(A)$. The condition is therefore equivalent to

$$\lim_{E\to\infty}\frac{1}{N(E)}\sum_{\lambda_j\leq E}|\langle A\varphi_k,\varphi_k\rangle-\omega(A)|^2=0.$$

Since all the terms are positive, no cancellation is possible and this condition is equivalent to the existence of a subset $\mathcal{S} \subset \mathbb{N}$ of density one such that $\mathcal{Q}_{\mathcal{S}} := \{d\Phi_k : k \in \mathcal{S}\}$ has only ω as a weak* limit point. As above, one says that the sequence of eigenfunctions is ergodic.

One could take this re-statement of Theorem 1 (1) as a semi-classical definition of quantum ergodicity. Two natural questions arise. First:

PROBLEM **4.** *Suppose the geodesic flow Φ^t of (M, g) is ergodic on S^*M. Is the operator K in*

$$\langle A \rangle = \omega(A) + K$$

a compact operator? In this case, $\sqrt{\Delta}$ is said to be QUE (quantum uniquely ergodic) If ergodicity is not sufficient for the QUE property, what extra conditions need to be added?

Compact would imply that $\langle K\varphi_k, \varphi_k \rangle \to 0$, hence $\langle A\varphi_k, \varphi_k \rangle \to \omega(A)$ along the entire sequence. Quite a lot of attention has been focussed on this problem in the last decade. It is probable that ergodicity is not by itself sufficient for the QUE property of general quantum ergodic systems. For instance, it is believed that there exist modes of asymptotic bouncing ball type which concentrate on the invariant Lagrangian cylinder (with boundary) formed by bouncing ball orbits of the Bunimovich stadium (see e.g. [KH] for more on such 'scarring'). Further, Faure-Nonnenmacher-de Bièvre have shown that QUE does not hold for the hyperbolic system defined by a quantum cat map on the torus [FND]. Since the methods applicable to eigenfunctions of quantum maps and of Laplacians have much in common, this negative result shows that there cannot exist a universal structural proof of QUE.

The principal positive result at this time is the proof by E. Lindenstrauss [Lin] of the QUE property for the orthonormal basis of Laplace-Hecke eigenfunctions eigenfunctions on arithmetic hyperbolic surfaces. It is generally believed that the spectrum of the Laplace eigenvalues is of multiplicity one for such surfaces, so this should imply QUE completely for these surfaces. Earlier partial results on Hecke eigenfunctions are due to Rudnick-Sarnak [R.S], Wolpert [W] and others. For more on Hecke eigenfunctions, see [M].

So far we have not mentioned Theorem 1 (2). In the next section we will describe a similar but more general result for mixing systems and the relevance of (2) will become clear. An interesting open problem is the extent to which (2) is actually necessary for the equivalence to classical ergodicity.

PROBLEM **5.** *Converse QE: What can be said of the classical limit of a quantum ergodic system, i.e. a system for which $\langle A \rangle = \omega(A) + K$ where K is semiclassically in the sense above, or compact?. Is it necessarily ergodic?*

Very little is known on this converse problem at present. It is known that if there exists an open set in S^*M filled by periodic orbits, then the Laplacian cannot be quantum ergodic (see [MO] for recent results and references). But no proof exists at this time that KAM systems, which have Cantor-like positive measure invariant sets, are not quantum ergodic. It is known that there exist a positive proportion of approximate eigenfunctions (quasi-modes) which localize on the invariant tori, but it has not been proved that a positive proportion of actual eigenfunctions have this localization property.

3.2. **Further problems and results on ergodic eigenfunctions.** Ergodicity is also known to have an impact on the distribution of zeros. The complex zeros in Kähler phase spaces of ergodic eigenfunctions of quantum ergodic maps become uniformly distributed with respect to the Kähler volume form [NV, SZ]. An interesting problem is whether the real analogue is true:

PROBLEM **6.** *Ergodicity and equidistribution of nodal sets. Let $\mathcal{N}_{\varphi_j} \subset M$ denote the nodal set (zero set) of φ_j, and equip it with its hypersurface volume form $d\mathcal{H}^{n-1}$ induced by g. Let (M, g) have ergodic geodesic flow, and suppose that $\{\varphi_j\}$ is an ergodic sequence of eigenfunctions. Are the following asymptotics valid?*

$$\int_{\mathcal{N}_{\varphi_j}} f d\mathcal{H}^{n-1} \sim \lambda_j \frac{1}{Vol(M,g)} \int_M f dVol.$$

This is predicted by the random wave model of §6. An equidistribution law for the complex zeros is known which gives some evidence for the validity of this limit formula. Let (M, g) be a compact real analytic Riemannian manifold and let $\varphi_j^{\mathbb{C}}$ be the holomorphic extension of the real analytic eigenfunction φ_j to the complexification $M_{\mathbb{C}}$ of M (its Grauert tube). Then if the geodesic flow is ergodic and if φ_j is an ergodic sequence of eigenfunctions, the normalized current of integration $\frac{1}{\lambda_j} Z_{\varphi_j^{\mathbb{C}}}$ over the complex zero set of $\varphi_j^{\mathbb{C}}$ tends weakly to $\overline{\partial}\partial|\xi_g|$. This current is invariant under the geodesic flow and is singular along the zero section.

Finally, we mention some results on L^∞ norms of eigenfunctions on arithmetic hyperbolic manifolds of dimensions 2 and 3. It is proved in [IS] that the joint eigenfunctions of Δ and the Hecke operators on arithmetic hyperbolic surfaces have the upper bound $\|\varphi_j\|_\infty = O_\epsilon(\lambda_j^{5/48+\epsilon})$ for all j and $\epsilon > 0$, and the lower bound $\|\varphi_j\|_\infty \geq c\sqrt{\log\log\lambda_j}$ for some constant $c > 0$ and infinitely many j. In [R.S] it is proved that there exists an arithmetic hyperbolic manifold and a subsequence φ_{j_k} of eigenfunctions with $\|\varphi_{j_k}\|_{L^\infty} \gg \lambda_{j_k}^{1/4}$, contradicting the random wave model predictions.

4. Quantum weak mixing

There are parallel results on quantizations of weak-mixing geodesic flows which are the subject of this section. First we recall the classical definition: the geodesic flow of (M, g) is weak mixing if the operator V_t has purely continuous spectrum on the orthogonal complement of the constant functions in $L^2(S^*M, d\mu)$. Hence like ergodicity it is a spectral property of the geodesic flow.

We have:

THEOREM **2.** *([Z.3,4]) The geodesic flow Φ^t of (M, g) is weak mixing if and only if the conditions (1)-(2) of Theorem 1 hold and additionally, for any $A \in \Psi^o(M)$,*

$$(\forall\epsilon)(\exists\delta) \limsup_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\substack{j\neq k: \lambda_j, \lambda_k \leq \lambda \\ |\lambda_j - \lambda_k - \tau| < \delta}} |(A\varphi_j, \varphi_k)|^2 < \epsilon \qquad (\forall\tau \in \mathbb{R})$$

The restriction $j \neq k$ is of course redundant unless $\tau = 0$, in which case the statement coincides with quantum ergodicity. This result follows from the general asymptotic formula, valid for any compact Riemannian manifold (M, g), that

$$\begin{array}{l}\frac{1}{N(\lambda)}\sum_{i\neq j,\lambda_i,\lambda_j\leq\lambda}|\langle A\varphi_i,\varphi_j\rangle|^2\left|\frac{\sin T(\lambda_i-\lambda_j-\tau)}{T(\lambda_i-\lambda_j-\tau)}\right|^2 \\ \sim ||\frac{1}{2T}\int_{-T}^{T}e^{it\tau}V_t(\sigma_A)||_2^2-|\frac{\sin T\tau}{T\tau}|^2\omega(A)^2.\end{array} \tag{24}$$

In the case of weak-mixing geodesic flows, the right hand side $\to 0$ as $T\to\infty$. As with diagonal sums, the sharper result is true where one averages over the short intervals $[\lambda, \lambda+1]$.

4.1. **Spectral measures and matrix elements.** Theorem 2 is based on expressing the spectral measures of the geodesic flow in terms of matrix elements. The main limit formula is:

$$\int_{\tau-\varepsilon}^{\tau+\varepsilon} d\mu_{\sigma_A} := \lim_{\lambda\to\infty}\frac{1}{N(\lambda)}\sum_{i,j:\ \lambda_j\leq\lambda,\ |\lambda_i-\lambda_j-\tau|<\varepsilon}|\langle A\varphi_i,\varphi_j\rangle|^2\ , \tag{25}$$

where $d\mu_{\sigma_A}$ is the spectral measure for the geodesic flow corresponding to the principal symbol of A, $\sigma_A\in C^\infty(S^*M, d\mu)$. Recall that the spectral measure of V_t corresponding to $f\in L^2$ is the measure $d\mu_f$ defined by

$$\langle V_t f, f\rangle_{L^2(S^*M)} = \int_{\mathbb{R}} e^{\mathrm{it}\,\tau} d\mu_f(\tau)\ .$$

The limit formula (25) is equivalent to the dual formula (under the Fourier transform)

$$\lim_{\lambda\to\infty}\frac{1}{N(\lambda)}\sum_{i,j:\lambda_j\leq\lambda} e^{it(\lambda_i-\lambda_j)}|\langle A\varphi_i,\varphi_j\rangle|^2 = \langle V_t\sigma_A,\sigma_A\rangle_{L^2(S^*M)}. \tag{26}$$

The proof of (26) is to consider, for $A\in\Psi^\circ$, the operator $A_t^*A\in\Psi^\circ$ with $A_t = U_t^*AU_t$. By the local Weyl law,

$$\lim_{\lambda\to\infty}\frac{1}{N(\lambda)}\operatorname{Tr} E(\lambda)A_t^*A = \langle V_t\sigma_A,\sigma_A\rangle_{L^2(S^*M)}\ .$$

The right side of (25) defines a measure dm_A on $\mathbb{R}$ and (26) says

$$\int_{\mathbb{R}} e^{it\tau} dm_A(\tau) = \langle V_t\sigma_A,\sigma_A\rangle_{L^2(S^*M)} \quad = \int_{\mathbb{R}} e^{it\tau} d\mu_{\sigma_A}(\tau).$$

Since weak mixing systems are ergodic, it is not necessary to average in both indices along an ergodic subsequence:

$$\lim_{\lambda_j\to\infty}\langle A_t^*A\varphi_j,\varphi_j\rangle = \sum_j e^{it(\lambda_i-\lambda_j)}|\langle A\varphi_i,\varphi_j\rangle|^2 = \langle V_t\sigma_A,\sigma_A\rangle_{L^2(S^*M)}. \tag{27}$$

Dually, one has

$$\lim_{\lambda_j\to\infty}\sum_{i\,:\,|\lambda_i-\lambda_j-\tau|<\varepsilon}|\langle A\varphi_i,\varphi_j\rangle|^2 \quad = \int_{\tau-\varepsilon}^{\tau+\varepsilon} d\mu_{\sigma_A}. \tag{28}$$

For QUE systems, these limit formulae are valid for the full sequence of eigenfunctions.

5. Rate of quantum ergodicity and mixing

A quantitative refinement of quantum ergodicity is to ask at what rate the sums in Theorem 1(1) tend to zero, i.e. to establish a rate of quantum ergodicity. More generally, we consider 'variances' of matrix elements. For diagonal matrix elements, we define:

$$V_A(\lambda) := \frac{1}{N(\lambda)} \sum_{j:\lambda_j \leq \lambda} |\langle A\varphi_j, \varphi_j\rangle - \omega(A)|^2. \tag{29}$$

In the off-diagonal case one may view $|\langle A\varphi_i, \varphi_j\rangle|^2$ as analogous to $|\langle A\varphi_j, \varphi_j\rangle - \omega(A)|^2$. However, the sums in (25) are double sums while those of (29) are single. One may also average over the shorter intervals $[\lambda, \lambda + 1]$.

5.1. **Quantum chaos conjectures.** First, consider off-diagonal matrix elements. One conjecture is that it is not necessary to sum in j in (28): each individual term has the asymptotics consistent with (28). This is implicitly conjectured by Feingold-Peres in [FP] (11) in the form

$$|\langle A\varphi_i, \varphi_j\rangle|^2 \simeq \frac{C_A(\frac{E_i - E_j}{\hbar})}{2\pi\rho(E)}, \tag{30}$$

where $C_A(\tau) = \int_{-\infty}^{\infty} e^{-i\tau t}\langle V_t\sigma_A, \sigma_A\rangle dt$. In our notation, $\lambda_j = \hbar^{-1}E_j$ and $\rho(E)dE \sim dN(\lambda)$. There are $\sim C\lambda^{n-1}$ eigenvalues λ_i in the interval $[\lambda_j - \tau - \epsilon, \lambda_j - \tau + \epsilon]$, so (30) says that individual terms have the asymptotics of (28).

On the basis of the analogy between $|\langle A\varphi_i, \varphi_j\rangle|^2$ and $|\langle A\varphi_j, \varphi_j\rangle - \omega(A)|^2$, it is conjectured in [FP] that

$$V_A(\lambda) \sim \frac{C_{A-\omega(A)I}(0)}{\lambda^{n-1} vol(\Omega)}.$$

The idea is that $\varphi_\pm = \frac{1}{\sqrt{2}}(\varphi_i \pm \varphi_j)$ have the same matrix element asymptotics as eigenfunctions when $\lambda_i - \lambda_j$ is sufficiently small. But then $2\langle A\varphi_+, \varphi_-\rangle = \langle A\varphi_i, \varphi_i\rangle - \langle A\varphi_j, \varphi_j\rangle$ when $A^* = A$. Since we are taking a difference, we may replace each matrix element by $\langle A\varphi_i, \varphi_i\rangle$ by $\langle A\varphi_i, \varphi_i\rangle - \omega(A)$ (and also for φ_j). The conjecture then assumes that $\langle A\varphi_i, \varphi_i\rangle - \omega(A)$ has the same order of magnitude as $\langle A\varphi_i, \varphi_i\rangle - \langle A\varphi_j, \varphi_j\rangle$. Dynamical grounds for this conjecture are given in [EFKAMM]. The order of magnitude is predicted by some natural random wave models, as discussed below in §6.

5.2. **Rigorous results.** At this time, the strongest variance result is an asymptotic formula for the diagonal variance proved by Luo-Sarnak for special Hecke eigenfunctions on the quotient $\mathbf{H}^2/SL(2,\mathbb{Z})$ of the upper half plane by the modular group [LS, Sa]. Their result pertains to holomorphic Hecke eigenforms, but the analogous statement for smooth Maass-Hecke eigenfunctions is expected to hold by similar methods, so we state the result as a Theorem/Conjecture. Note that $\mathbf{H}^2/SL(2,\mathbb{Z})$ is a non-compact finite area surface whose Laplacian Δ has both a discrete and a continuous spectrum. The discrete Hecke eigenfunctions are joint eigenfunctions of Δ and the Hecke operators T_p (see [Sa] for background).

Theorem/Conjecture **1.** [LS] *Let $\{\varphi_k\}$ denote the orthonormal basis of Hecke eigenfunctions for $\mathbf{H}^2/SL(2,\mathbb{Z})$. Then there exists a quadratic form $B(f)$ on $C_0^\infty(\mathbf{H}^2/SL(2,\mathbb{Z}))$*

such that

$$\frac{1}{N(\lambda)}\sum_{\lambda_j\leq\lambda}|\int_X f|\varphi_j|^2dvol-\frac{1}{Vol(X)}\int_X fdVol|^2=\frac{B(f,f)}{\lambda}+o(\frac{1}{\lambda}).$$

When the multiplier $f=\varphi_\lambda$ is itself an eigenfunction, Luo-Sarnak have shown that

$$B(\varphi_\lambda,\varphi_\lambda)=C_{\varphi_\lambda}(0)L(\frac{1}{2},\varphi_\lambda)$$

where $L(\frac{1}{2},\varphi_\lambda)$ is a certain L-function. Thus, the conjectured classical variance is multiplied by an arithmetic factor depending on the multiplier. A crucial fact in the proof is that the quadratic form B is diagonalized by the φ_λ.

The only rigorous result to date which is valid on general Riemannian manifolds with hyperbolic geodesic flow is the logarithmic decay [Z6]

Theorem **3.** *For any (M,g) with hyperbolic geodesic flow,*

$$\frac{1}{N(\lambda)}\sum_{\lambda_j\leq\lambda}|(A\varphi_j,\varphi_j)-\omega(A)|^{2p}=\frac{1}{(\log\lambda)^p}.$$

The logarithm as usual reflects the exponential blow up in time of remainder estimates for traces involving the wave group. It is rather doubtful that such a result is sharp. However, in the case of two-dimensional quantum cat maps, with eigenspaces of multiplicty $\lambda/\log\lambda$, there may exist orthonormal bases with rather large rates of ergodicity.

6. Random waves and orthonormal bases

We have mentioned that the random wave model provides a kind of guideline for what to conjecture about eigenfunctions of quantum chaotic system. In this final section, we briefly discuss random wave models and what they predict.

By a random wave model one means a probability measure on a space of functions. To deal with orthonormal bases rather than individual functions, one puts a probability measure on a space of orthonormal bases, i.e. on a unitary group. We denote expected values relative to a given probability measure by $\mathbf{E}$. We now consider some specific Gaussian models and what they predict about variances.

As a model for quantum chaotic eigenfunctions in plane domains, M. V. Berry suggested using the *Euclidean random wave model at fixed energy* [B]. Let $\mathcal{E}_\lambda$ denote the space of (tempered) eigenfunctions of eigenvalue λ^2 of the Euclidean Laplacian Δ on $\mathbb{R}^n$. It is spanned by exponentials $e^{i\langle k,x\rangle}$ with $k\in\mathbb{R}^n,|k|=\lambda$. The infinite dimensional space $\mathcal{E}_\lambda$ is a unitary representation of the Euclidean motion group and carries an invariant inner product. The inner product defines an associated Gaussian measure whose covariance kernel $C_\lambda(x,y)=\mathbf{E}f(x)\bar{f}(y)$ is the derivative at λ of the spectral function

$$E(\lambda,x,y)=(2\pi)^{-n}\int_{|\xi|\leq\lambda}e^{i\langle x-y,\xi\rangle}d\xi,\quad \xi\in\mathbb{R}^n. \tag{31}$$

Thus,

$$C_\lambda(x,y)=\frac{d}{d\lambda}E(\lambda,x,y)=(2\pi)^{-n}\int_{|\xi|=\lambda}e^{i\langle x-y,\xi\rangle}dS=(2\pi)^{-n}\lambda^{n-1}\int_{|\xi|=1}e^{i\lambda\langle x-y,\xi\rangle}dS, \tag{32}$$

where dS is the usual surface measure. With this definition, $C_\lambda(x,x) \sim \lambda^{n-1}$. In order to make $\mathbf{E}(f(x)^2) = 1$ consistent with normalized eigenfunctions, we divide by λ^{n-1} to define

$$\hat{C}_\lambda(x,y) = (2\pi)^{-n} \int_{|\xi|=1} e^{i\lambda\langle x-y,\xi\rangle} dS.$$

One could express the integral as a Bessel function to rewrite this as $\Gamma(\frac{n-1}{2})|\lambda|x-y||^{-\frac{n-2}{2}} J_{\frac{n-2}{2}}(\lambda|x-y|)$.

Wick's formula in this ensemble gives:

$$\mathbf{E}\varphi(x)^2\varphi(y)^2 = \frac{1}{Vol(\Omega)^2}[1 + 2C_\lambda(x,y)^2].$$

Thus, in dimension n we have:

$$\begin{aligned} \mathbf{E}[\textstyle\int\int V(x)V(y)\varphi(x)^2\varphi(y)^2 dxdy - \bar{V}^2] &= \frac{2}{Vol(\Omega)^2}\int_\Omega\int_\Omega \hat{C}_\lambda(x,y)^2 V(x)V(y)dxdy \\ &\sim \frac{1}{\lambda^{n-1}Vol(\Omega)^2}\int_\Omega\int_\Omega \frac{V(x)V(y)}{|x-y|^{n-1}}\cos(|x-y|\lambda)^2 dxdy. \end{aligned}$$

In the last line, we used the stationary phase asymptotics

$$(2\pi)^{-n}\lambda^{n-1}\int_{|\xi|=1} e^{i\lambda\langle x-y,\xi\rangle} dS \sim C_n(\lambda|x-y|)^{-\frac{n-1}{2}}\cos(|x-y|\lambda). \tag{33}$$

Thus, the variances have order $\lambda^{-(n-1)}$ in dimension n, consistent with the conjectures in [FP, EFKAMM].

This model is often used to obtain predictions on eigenfunctions of chaotic systems. By construction it is tied to Euclidean geometry and only pertains directly to individual eigenfunctions of a fixed eigenvalue. It is based on the infinite dimensional multiplicity of eigenfunctions of fixed eigenvalue of the Euclidean Laplacian on $\mathbb{R}^n$. There also exist random wave models on a curved Riemannian manifold (M,g), which model individual eigenfunctions and also random orthonormal bases [Z.2, Z.5]. Thus, one can compare the behavior of sums over eigenvalues of the orthonormal basis of eigenfunctions of Δ with that of a random orthonormal basis. Instead of taking Gaussian random combinations of Euclidean plane waves of a fixed eigenvalue, one takes Gaussian random combinations $\sum_{j:\lambda_j\in[\lambda,\lambda+1]} c_j\varphi_j$ of the eigenfunctions of (M,g) with eigenvalues in a short interval in the sense above. Equivalently, one takes random combinations with $\sum_j |c_j|^2 = 1$. These random waves are globally adapted to (M,g). The statistical results depend on the measure of the set of periodic geodesics of (M,g); thus, as discussed in [KHZ], different random wave models make different predictions about off-diagonal variances.

Fix a compact Riemannian manifold (M,g) and partition the spectrum of $\sqrt{\Delta}$ into the intervals $I_k = [k,k+1]$. Let $\Pi_k = E(k+1) - E(k)$ be the kernel of the spectral projections for $\sqrt{\Delta}$ corresponding to the interval I_k. Its kernel $\Pi_k(x,y)$ is the covariance kernel of Gaussian random combinations $\sum_{j:\lambda_j\in I_k} c_j\varphi_j$ and is analogous to $C_\lambda(x,y)$ in the Euclidean case; it is of course not the derivative $dE(\lambda,x,y)$ but the difference of the spectral projector over I_k. We denote by $N(k)$ the number of eigenvalues in I_k and put $\mathcal{H}_k = \mathrm{ran}\Pi_k$ (the range of Π_k). We define a *random* orthonormal basis of $\mathcal{H}_k$ by changing the basis of eigenfunctions $\{\varphi_j\}$ of Δ in $\mathcal{H}_k$ by a random element of the unitary group $U(\mathcal{H}_k)$ of the finite dimensional Hilbert space $\mathcal{H}_k$. We then define a random orthonormal basis of $L^2(M)$ by taking the product over all the spectral intervals in our partition. More precisely, we define the infinite dimensional unitary group

$$U(\infty) = \Pi_{k=1}^{\infty} U(\mathcal{H}_k)$$

of sequences $(U_1, U_2, \dots)$, with $U_k \in U(\mathcal{H}_k)$. We equip $U(\infty)$ with the product

$$d\nu_\infty = \Pi_{k=1}^{\infty} d\nu_k$$

of the unit mass Haar measures $d\nu_k$ on $U(\mathcal{H}_k)$: We then define a *random* orthonormal basis of $L^2(M)$ to be obtained by applying a random element $U \in U(\infty)$ to the orthonormal basis $\Phi = \{\varphi_j\}$ of eigenfunctions of $\sqrt{\Delta}$.

Assuming the set of periodic geodesics of (M, g) has measure zero, the Weyl remainder results (8) and strong Szegö limit asymptotics of [GO, LRS] give two term asymptotics for the traces $\Pi_k A \Pi_k, (\Pi_k A \Pi_k)^2$ for any pseudodifferential operator A. Combining the strong Szegö asymptotics with the arguments of [Z.5], random orthonormal bases can be proved to satisfy the following variance asymptotics:

$$(i) \quad \mathbf{E}(\textstyle\sum_{j:\lambda_j \in I_k} |(AU\varphi_j, U\varphi_j) - \omega(A)|^2 \sim (\omega(A^*A) - \omega(A)^2);$$

$$(ii) \quad \mathbf{E}(\textstyle\sum_{i \neq j: \lambda_j, \lambda_i \in I_k} \left| \frac{sinT(\lambda_i - \lambda_j - \tau)}{T(\lambda_i - \lambda_j - \tau)} \right|^2 |(AU\varphi_j, U\varphi_i)|^2$$

$$\sim \{2 \left| \frac{sin\tau T}{\tau T} \right|^2 + \frac{1}{N(k)} \textstyle\sum_{i \neq j} \left| \frac{sinT(\lambda_i - \lambda_j - \tau)}{T(\lambda_i - \lambda_j - \tau)} \right|^2 \}(\omega(A^*A) - \omega(A)^2)$$

References

[A] O. Agam, B. Altshuler and A. V. Andreev, Spectral statistics: from disordered to chaotic systems. Phys. Rev. Lett. 75 (1995), no. 24, 4389–4392

[BSS] A. Bäcker, R. Schubert and P. Stifter, Rate of quantum ergodicity in Euclidean billiards. Phys. Rev. E (3) 57 (1998), no. 5, part A, 5425–5447. (Erratum: "Rate of quantum ergodicity in Euclidean billiards". Phys. Rev. E (3) 58 (1998), no. 4, 5192).

[Bar] A. Barnett, Asymptotic rate of quantum ergodicity in chaotic Euclidean billiards (2005).

[Ber] P. Bérard, On the wave equation on a compact Riemannian manifold without conjugate points. Math. Z. 155 (1977), no. 3, 249–276.

[B] M. V. Berry, Regular and irregular semiclassical wavefunctions. J. Phys. A 10 (1977), no. 12, 2083–2091.

[BBL] E. Bogomolny, O. Bohigas and P. Leboeuf, Quantum chaotic dynamics and random polynomials. J. Statist. Phys. 85 (1996), no. 5-6, 639–679.

[BGS] O. Bohigas, M. J. Giannoni and C. Schmit, Characterization of chaotic quantum spectra and universality of level fluctuation laws. Phys. Rev. Lett. 52 (1984), no. 1, 1–4.

[Bu] N. Burq, Quantum ergodicity of boundary values of eigenfunctions: A control theory approach, to appear in Canadian Math. Bull. (math.AP/0301349).

[CV] Y.Colin de Verdière, Ergodicité et fonctions propres du Laplacien, Comm.Math.Phys. 102 (1985), 497-502.

[DSj] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit.* London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.

[DG] J. Duistermaat and V. Guillemin, The spectrum of positive elliptic operators and periodic bicharacteristics, Invent. Math. 29 (1975), 39–79.

[EFKAMM] B. Eckhardt, S. Fishman, J. Keating, O. Agam, J. Main and K. Miller, Approach to Ergodocity in Quantum Wave Functions, Phys. Rev. E 52, 5893-5903 (1995).

[FND] F. Faure, S. Nonnenmacher and S. De Bièvre, Scarred eigenstates for quantum cat maps of minimal periods. Comm. Math. Phys. 239 (2003), no. 3, 449–492

[FP] M. Feingold and A. Peres, Distribution of matrix elements of chaotic systems. Phys. Rev. A (3) 34 (1986), no. 1, 591–595.

[GL] P.Gérard and E.Leichtnam, Ergodic properties of eigenfunctions for the Dirichlet problem, Duke Math J. 71 (1993), 559-607.

[GO] V. Guillemin and K. Okikiolu, Subprincipal terms in Szegö estimates, Math. Res. Lett. 4 (1997), 173–179.

[Gu] M. C. Gutzwiller, *Chaos in classical and quantum mechanics.* Interdisciplinary Applied Mathematics, 1. Springer-Verlag, New York, 1990

[H] E. J. Heller, Bound-state eigenfunctions of classically chaotic Hamiltonian systems: scars of periodic orbits. Phys. Rev. Lett. 53 (1984), no. 16, 1515–1518.

[IS] H. Iwaniec and P. Sarnak, L^∞ norms of eigenfunctions of arithmetic surfaces. Ann. of Math. (2) 141 (1995), no. 2, 301–320.

[KHZ] L. Kaplan and E. J. Heller, Weak quantum ergodicity. Phys. D 121 (1998), no. 1-2, 1–18. Appendix by S. Zelditch.

[KH] L. Kaplan and E.J. Heller, Measuring scars of periodic orbits. Phys. Rev. E (3) 59 (1999), no. 6, 6609–6628.

[HZ] A. Hassell and S. Zelditch, Quantum ergodicity and boundary values of eigenfunctions, Comm.Math.Phys. Volume 248, Number 1 (2004) 119 - 168.

[HoIII] L. Hörmander, *The Analysis of Linear Partial Differential Operators III*, Grundlehren **274** , Springer-Verlag (1985).

[HoIV] L. Hörmander, *The Analysis of Linear Partial Differential Operators IV*, Grundlehren **275**, Springer-Verlag (1985).

[LRS] A. Laptev, D. Robert and Yu. Safarov, Remarks on the paper of V. Guillemin and K. Okikiolu: "Subprincipal terms in Szeg estimates" [Math. Res. Lett. 4 (1997), no. 1, 173–179. Math. Res. Lett. 5 (1998), no. 1-2, 57–61.

[Lin] E. Lindenstrauss, Invariant measures and arithmetic quantum ergodicity, Annals of Math. (to appear).

[LS] W.Luo and P.Sarnak, Quantum ergodicity of eigenfunctions on $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathbf{H}^2$, IHES Publ. 81 (1995), 207-237.

[L.S.2] W.Luo and P.Sarnak, Quantum variance for Hecke eigenforms, Annales Scient. de l'École Norm. Sup. 37 (2004), p. 769-799.

[MO] J. Marklof and S. O'Keefe, Weyl's law and quantum ergodicity for maps with divided phase space. With an appendix "Converse quantum ergodicity" by Steve Zelditch. Nonlinearity 18 (2005), no. 1, 277–304.

[M] J. Marklof, Arithmetic quantum chaos, this Encyclopedia.

[NV] S. Nonnenmacher and A. Voros, Chaotic eigenfunctions in phase space. J. Statist. Phys. 92 (1998), no. 3-4, 431–518.

[R.S] Z.Rudnick and P.Sarnak, The behaviour of eigenstates of arithmetic hyperbolic manifolds, Comm.Math.Phys. 161 (1994), 195-213.

[Ru] D. Ruelle, *Statistical mechanics: Rigorous results.* W. A. Benjamin, Inc., New York-Amsterdam 1969

[Sa.1] P.Sarnak, Arithmetic quantum chaos, in *The Schur Lectures (Tel Aviv, 1992), Israel Mathematical Conference Proc. Vol. 8* (1995), 183-236.

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

[SZ] B. Shiffman and S. Zelditch, Distribution of zeros of random and quantum chaotic sections of positive line bundles. Comm. Math. Phys. 200 (1999), no. 3, 661–683.

[Sh.1] A.I.Shnirelman, Ergodic properties of eigenfunctions, Usp.Math.Nauk. 29/6 (1974), 181-182.

[Sh.2] A.I.Shnirelman, On the asymptotic properties of eigenfunctions in the region of chaotic motion, addendum to V.F.Lazutkin, *KAM theory and semiclassical approximations to eigenfunctions*, Springer (1993).

[Sog] C. D. Sogge, Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds, J. Funct. Anal. 77 (1988), 123–138.

[SogZ] C. Sogge and S. Zelditch, Riemannian manifolds with maximal eigenfunction growth. Duke Math. J. 114 (2002), no. 3, 387–437

[Su] T.Sunada, Quantum ergodicity. *Progress in inverse spectral geometry*, 175–196, Trends Math., Birkhuser, Basel, 1997.

[W] S. A. Wolpert, The modulus of continuity for $\Gamma_0(m)\backslash\mathbb{H}$ semi-classical limits. Comm. Math. Phys. 216 (2001), no. 2, 313–323.

[Z0] S. Zelditch, Uniform distribution of eigenfunctions on compact hyperbolic surfaces. Duke Math. J. 55 (1987), no. 4, 919–941

[Z.1] S.Zelditch, Quantum ergodicity of C^* dynamical systems, Comm.Math.Phys. 177 (1996), 507-528.

[Z.2] ————, Quantum ergodicity on the sphere, Comm. Math. Phys., **146** (1992), 61-71.

[Z.3] ————, Quantum transition amplitudes for classically ergodic or completely integrable systems, J. Fun. Anal. **94** (1990), 415-436.

[Z.4] ———, Quantum Mixing, J.Fun.Anal.140 (1996), 68-86.

[Z.5] ———, A random matrix model for quantum mixing, Int.Math. Res.Not. 3 (1996), 115-137.

[Z6] ———, On the rate of quantum ergodicity. I. Upper bounds. Comm. Math. Phys. 160 (1994), no. 1, 81–92

[ZZw] S.Zelditch and M.Zworski, Ergodicity of eigenfunctions for ergodic billiards, Comm.Math. Phys. 175 (1996), 673-682.

Department of Mathematics, Johns Hopkins University, Baltimore, MD 21218, USA
E-mail address: szelditch@jhu.edu