PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 131, Number 9, Pages 2945–2951
S 0002-9939(02)06810-7
Article electronically published on December 30, 2002

QUANTUM UNIQUE ERGODICITY

HAROLD DONNELLY

(Communicated by Jozef Dodziuk)

ABSTRACT. Consider a compact Riemannian manifold with ergodic geodesic flow. Quantum ergodicity is generalized from orthonormal bases of eigenfunctions of the Laplacian to packets of eigenfunctions. It is shown that this more general result is sharp. Namely, there may exist exceptional packets of eigenfunctions which concentrate on a submanifold.

0. INTRODUCTION

Let M be a compact Riemannian manifold. The Laplacian Δ of M is a second order self-adjoint differential operator. Moreover, there exist orthonormal bases for L^2M which consist of eigenfunctions ϕ_k of Δ with eigenvalues λ_k, $\Delta\phi_k = -\lambda_k\phi_k$.

In the case when the geodesic flow on the unit sphere bundle of M is ergodic, Colin de Verdière [2] proved a remarkable theorem. Suppose that A is a zero'th order pseudodifferential operator with symbol a. Then there exists a subsequence ϕ_{k_i}, of density one in ϕ_k, so that, for any A,

$$\lim_{i\to\infty} \langle A\phi_{k_i}, \phi_{k_i}\rangle = \int_{S^1M} a(\omega)d\omega. \tag{0.1}$$

Here $d\omega$ is the normalized Liouville measure on the unit sphere bundle S^1M. An important open problem is whether the equality (0.1) holds for the full orthonormal basis.

We contribute to this discussion by broadening the perspective. Definition 1.1 introduces the concept of packets of eigenfunctions, which are certain finite linear combinations of eigenfunctions. Our Theorem 1.4 shows that the result of Colin de Verdière generalizes to such packets of eigenfunctions. The rest of the paper is devoted to establishing the sharpness of this more general result. Namely, there may exist exceptional subsequences, of density zero, where (0.1) fails. Theorem 2.2 constructs manifolds with ergodic geodesic flow which have nonpositive curvature and contain flat cylinders. The proof of Theorem 2.2 is an application of the method of Burns and Gerber [1]. In Theorem 3.5, we produce exceptional sequences of eigenfunction packets supported on the manifolds provided by Theorem 2.2. These packets may be localized to arbitrarily small intervals on the spectrum while maintaining a fixed normalization for the average level spacings.

Received by the editors March 15, 2002 and, in revised form, April 1, 2002.
2000 *Mathematics Subject Classification.* Primary 58G25.
Research supported by the Ellentuck Fund and the National Science Foundation.

The problem of quantum unique ergodicity has been studied extensively in both the mathematics and physics literature. A survey is provided in the Schur lectures by Sarnak [4]. The study of the stadium billiard by Heller [3] is closely related to the ideas presented here. Prior to the definitive paper by Colin de Verdière [2], there were important works by Shnirelman [5] and Zelditch [6].

The author thanks Professors J. Bourgain and T. Spencer for helpful conversations during the development of this work.

1. QUANTUM ERGODICITY FOR PACKETS OF EIGENFUNCTIONS

Suppose that M is a compact Riemannian manifold. The Laplacian Δ of M has pure point spectrum. This means that L^2M admits an orthonormal basis consisting of eigenfunctions of Δ. We will consider more general orthonormal bases, where each basis element is a finite linear combination of eigenfunctions. The following definition makes this concept precise:

Definition 1.1. An orthonormal basis ϕ_k of L^2M consists of packets of eigenfunctions if there exists a sequence of spectral projectors E_k for Δ, supported on intervals of length at most $\delta > 0$, so that $\phi_k = E_k\phi_k$. Here $\delta > 0$ is independent of k.

Let S^1M denote the unit sphere bundle in the cotangent bundle of M. The Friedrichs quantification associates to each function $a \in C^\infty(S^1M)$ a pseudodifferential operator $A = Op^F(a)$, of order zero. Moreover, if $a \geq 0$, then $A \geq 0$ as an operator on L^2M. If ϕ_k is a packet of eigenfunctions, then we define a measure μ_k by $\mu_k(a) = \langle Op^F(a)\phi_k, \phi_k\rangle$.

We follow the outline of [2], noting that each step generalizes from individual eigenfunctions to packets of eigenfunctions. In the subsequent discussion, ϕ_k will denote a fixed orthonormal basis which consists of packets of eigenfunctions. Let G_t denote the geodesic flow on the unit cosphere bundle S^1M. One has

Lemma 1.2. *For each $t > 0$, $\mu_k - G_t^*\mu_k$ converges weakly to zero, as k approaches infinity.*

Proof. Choose $-\lambda_k$ belonging to the interval which supports the spectral projector E_k. Since $\phi_k = E_k\phi_k$, it follows that

$$\left\| \left(e^{it\sqrt{-\Delta}} - e^{it\sqrt{\lambda_k}} \right) \phi_k \right\| \to 0$$

as $k \to \infty$. One uses the elementary inequality

$$\sqrt{\lambda_k + \delta} - \sqrt{\lambda_k} = \sqrt{\lambda_k}\,\left(\sqrt{1 + \delta/\lambda_k} - 1\right) < c_\delta \lambda_k^{-1/2}.$$

Now suppose that a is a smooth function on the unit cosphere bundle and $A = Op_F(a)$. If$A_t = \exp\left(-it\sqrt{\Delta}\,\right) A \exp\left(it\sqrt{\Delta}\,\right)$, it is easy to deduce that

$$\lim_{k\to\infty} \langle A_t\phi_k, \phi_k\rangle = \lim_{k\to\infty} \langle A\phi_k, \phi_k\rangle \,.$$

More precisely, we mean that the limit of the difference is zero.

The remainder of the proof is the same as for individual eigenfunctions. By Egorov's theorem $A_t - Op_F(a \circ G_t)$ is a pseudodifferential operator of the order -1.

Consequently,

$$\begin{aligned}\lim_{k\to\infty}\left\langle Op_F(a\circ G_t)\phi_k,\phi_k\right\rangle &= \lim_{k\to\infty}\left\langle A_t\phi_k,\phi_k\right\rangle\\ &= \lim_{k\to\infty}\left\langle A\phi_k,\phi_k\right\rangle = \lim_{k\to\infty}\left\langle Op_F(a)\phi_k,\phi_k\right\rangle.\end{aligned}$$

Since smooth functions are dense in the space of continuous functions, this completes the proof of Lemma 1.2. □

It follows from Lemma 1.2 that any weak limit of the measures μ_k is invariant under the geodesic flow. The basic question is whether such weak limits coincide with the normalized Liouville measure $d\omega$ on S^1M. Our next lemma will show that the Liouville measure is an asymptotic average of the μ_k. Fix any sequence of numbers λ_k so that λ_k belongs to the interval which supports the spectral projector E_k. Let N_λ denote the number of λ_k which are less than λ. One may now state

Lemma 1.3. *If $d\omega$ denotes the normalized Liouville measure on S^1M, then*

$$\lim_{\lambda\to\infty}\frac{1}{N_\lambda}\sum_{\lambda_k\le\lambda}\int a\,d\mu_k = \int a\,d\omega.$$

Proof. Since $\phi_k = E_k\phi_k$, we have

$$\begin{aligned}\left|\left\langle A\phi_k,\left(e^{t\Delta}-e^{-t\lambda_k}\right)\phi_k\right\rangle\right| &\le cte^{-t\lambda_k}\|A\phi_k\|\,\|\phi_k\|\\ &= cte^{-t\lambda_k}\left\langle A^*A\phi_k,\phi_k\right\rangle^{1/2} \le cte^{-t\lambda_k}(1+\left\langle A^*A\phi_k,\phi_k\right\rangle).\end{aligned}$$

Thus $\left|\left\langle A\phi_k,\left(e^{t\Delta}-e^{-t\lambda_k}\right)\phi_k\right\rangle\right| \le cte^{-t\lambda_k}$. It follows that

$$\mathrm{Tr}(Ae^{t\Delta}) = \sum_k\left\langle Ae^{t\Delta}\phi_k,\phi_k\right\rangle = (1+0(t))\sum_k e^{-t\lambda_k}\left\langle A\phi_k,\phi_k\right\rangle$$

as $t\downarrow 0$.

Again, we may proceed to apply the method used for orthonormal bases of eigenfunctions. One has the small time asymptotics

$$\sum_k e^{-t\lambda_k}\left\langle A\phi_k,\phi_k\right\rangle \sim \mathrm{Tr}\left(Ae^{t\Delta}\right) \sim \mathrm{Tr}\left(e^{t\Delta}\right)\int a\,d\omega.$$

Lemma 1.3 now follows from the Tauberian theorem of Karamata. □

Observe that the preceding lemmas hold for any compact Riemannian manifold. No special hypotheses are needed concerning the geodesic flow. However, the main result of this section applies only to manifolds with ergodic geodesic flow:

Theorem 1.4. *Let M be a compact Riemannian manifold. Assume that the geodesic flow on S^1M is ergodic. Suppose that ϕ_k is an orthonormal basis, for L^2M, consisting of packets of eigenfunctions. Then, for a subsequence ϕ_{k_i}, having density one,*

$$\lim_{i\to\infty}\left\langle A\phi_{k_i},\phi_{k_i}\right\rangle = \int_{S^1M} a\,d\omega. \tag{1.5}$$

In short, almost every spectral measure converges to Liouville measure.

Proof. Given $a\in C^\infty(S^1M)$, let $a_T(z) = \frac{1}{T}\int_0^T a(G_tz)dt$ and $\bar a = \int_{S^1M} a\,d\omega$. The ergodic theorem implies that a_T converges almost everywhere to $\bar a$, as $T\to\infty$. Set $\hat a_T = a_T - \bar a$. Given $\epsilon > 0$, it now follows, when T is sufficiently large, that $\int_{S^1M}|\hat a_T|\,d\omega \le \epsilon$.

By our Lemma 1.2, we have $\mu_k(a_T - a) \to 0$, as $k \to \infty$. Thus $\mu_k(\hat{a}_T)$ has the same limiting behavior as $\mu_k(a - \bar{a})$. Theorem 1.4 now follows by applying Lemma 1.3 to $|\hat{a}_T|$. The reader is referred to [2] for the additional elementary details. ☐

2. Manifolds with ergodic geodesic flow

Let M be a compact Riemannian manifold of dimension two. If M has strict negative curvature, then it is well-known that the geodesic flow on the unit cosphere bundle S^1M is ergodic. We will apply the method of [1] to construct surfaces of nonpositive curvature with ergodic geodesic flow. These surfaces will contain cylinders of zero curvature. This leads us to formulate

Definition 2.1. A flat cylinder $C(L, \alpha)$ is a product of the form $[-L, L] \times S^1$, where the metric on S^1 has length $2\pi\alpha$.

The desired surfaces are constructed by starting with two isometric surfaces of finite volume and constant negative curvature. These surfaces will be noncompact with a finite number of ends, where the topology is $R^+ \times S^1$, with a warped product metric $dr^2 + e^{-2r}d\theta^2$, $r > c$. We truncate the cusps and then extend the metric to obtain manifolds with boundary, with $K < 0$ in the interior, and the boundary consisting of a finite union of closed geodesics, where $K = 0$. Because of the warped product structure on the ends, the extension exists by elementary real analysis in one variable r. If X_1 and X_2 are the two manifolds with boundary thus obtained, we scale their metrics so that both their volumes become as small as needed. The boundary components are then joined in pairs by flat cylinders $C(L_i, \alpha_i)$. We obtain a surface M with non-negative Gaussian curvature K, with $K < 0$ outside the flat cylinders.

These surfaces satisfy the requirements of the following

Theorem 2.2. *There exist compact surfaces M with ergodic geodesic flow which contain finitely many flat cylinders $C(L_i, \alpha_i)$. In fact, the geodesic flow is Bernoulli. Moreover, we may choose the L_i to be arbitrarily large while maintaining unit volume for M.*

Proof. We apply the method of invariant cones as developed in [1]. The tangent space of S^1M decomposes into the horizontal and vertical parts. If $\xi \in TS^1M$, then we write $\xi = \xi_H + \xi_V$, where each of ξ_H, ξ_V may be identified with a vector in TM. A distribution P, which is invariant under the geodesic flow G_t and transverse to the flow, is defined for every $w \in S^1M$ by

$$P(w) = \left\{ \xi \in T_w(S^1M) \;\middle|\; \langle \xi_H, w \rangle = 0 = \langle \xi_V, w \rangle \right\}.$$

The distribution $P(w)$ contains the cones

$$K(w) = \left\{ \xi \in P(w) \;\middle|\; \langle \xi_H, \xi_V \rangle \geq 0 \right\}.$$

It follows from the nonpositive curvature of M that the cones are invariant, meaning that $dG^t(K(w)) \subseteq K(G^t w)$. The containment is strict if the basepoint w lies outside the flat cylinders $C(L_i, \alpha_i)$. We let $\mathcal{U}$ be the set of unit vectors whose basepoints lie outside the cylinders. Using the simple and explicit behavior of the geodesic flow on the cylinders, one checks that $\bigcup\limits_{t \in R} \phi^t \mathcal{U}$ has full measure in M. The ergodicity of the geodesic flow now follows from Theorem 1.4 in [1]. The main point is that the curvature is nonpositive and strictly negative outside the cylinders. ☐

3. EXCEPTIONAL SEQUENCES OF EIGENFUNCTION PACKETS

Let M be a compact Riemannian manifold of dimension two. We suppose that M contains a flat cylinder $C(L,\alpha)$ in the sense of Definition 2.1. Our first step is to show that if L is sufficiently large, then there exist ϵ-accurate quasimodes supported on $C(L,\alpha)$. It may be helpful to recall the definition of quasimodes.

Definition 3.1. A function w is called an ϵ-accurate quasimode if there exists a real number γ so that $\|(\Delta+\gamma)w\|_2 \leq \epsilon\|w\|_2$.

The quasimodes are closely related but different in detail from the eigenfunction packets considered in section one. In particular, a quasimode need not be representable as a finite sum of eigenfunctions. An eigenfunction packet will be an ϵ-accurate quasimode only if the corresponding eigenvalues are all contained in an interval of length ϵ. Perhaps the main difference is that quasimodes need not be localized in phase space, but eigenfunction packets are so localized.

Recall that $C(L,\alpha) = [-L,L] \times S^1$ where the metric on S^1 has length $2\pi\alpha$. Let $r \in [-L,L]$ and $\theta \in S^1$ be the standard coordinates. We assume $L > 1$, and choose a cut-off function $\eta \in C_0^\infty[-L,L]$, so that $\eta(r) = 1$ for $r \in [-L+1, L-1]$. Our quasimodes are defined by $w_k(r,\theta) = \eta(r)\exp(ik\theta/\alpha)$, with corresponding approximate eigenvalues $\gamma_k = k^2/\alpha^2$. One calculates

$$\|(\Delta+\gamma_k)w_k\|_2 = \|(\Delta\eta)e^{ik\theta}\|_2 \leq \epsilon\|w_k\|_2.$$

If L is sufficiently large, then an arbitrarily small ϵ can be achieved.

We normalize our quasimodes by setting $u_k = w_k/\|w_k\|_2$. For future reference, it is convenient to record the following lemmas concerning the normalized quasimodes.

Lemma 3.2. *Let E_k denote the spectral projection onto the interval $[\gamma_k - x, \gamma_k + x]$. Then*

$$\|(I-E_k)u_k\|_2 \leq \epsilon x^{-1}\|u_k\|_2.$$

Proof. One observes that

$$\|(I-E_k)u_k\|_2 \leq x^{-1}\|(\Delta+\gamma_k)(I-E_k)u_k\|_2 \leq x^{-1}\|(\Delta+\gamma_k)u_k\|_2 \leq \epsilon x^{-1}\|u_k\|_2,$$

for any $x > 0$. ☐

Lemma 3.3. *For a constant $C(\epsilon)$, depending upon ϵ,*

$$\left\|(-\Delta+1)^{-1}u_k\right\|_2 \leq C(\epsilon)\gamma_k^{-1}\|u_k\|_2.$$

Proof. Clearly $\|(-\Delta+1)^{-1}u_k\|_2 \leq \|(-\Delta+1)^{-1}E_k u_k\|_2 + \|(-\Delta+1)^{-1}(I-E_k)u_k\|_2$ $\leq (\gamma_k - x)^{-1}\|E_k u_k\|_2 + \|(I-E_k)u_k\|_2$.

Using Lemma 3.2, we deduce that $\|(-\Delta+1)^{-1}u_k\|_2 \leq [(\gamma_k - x)^{-1} + \epsilon x^{-1}]\|u_k\|_2$. Now choose the variable $x = \gamma_k\sqrt{\epsilon}/(1+\sqrt{\epsilon}\,)$. ☐

Our next result is that the normalized quasimodes violate quantum unique ergodicity. Define a zero'th order pseudodifferential operator by $A = (-\Delta+1)^{-1}D$, where $D = \eta(r)\frac{\partial^2}{\partial r^2}$. Note that the radial derivatives of our normalized quasimodes are bounded independent of k. One has

Proposition 3.4. *Although the average value of the symbol of A is non-zero, nevertheless $\lim\limits_{k\to\infty} \langle Au_k, u_k\rangle = 0$.*

Proof. One calculates

$$\begin{aligned}|\langle Au_k, u_k\rangle| = |\left\langle Du_k, (-\Delta+1)^{-1}u_k\right\rangle| &\leq \|Du_k\|_2 \|(-\Delta+1)^{-1}u_k\|_2 \\ &\leq c\|(-\Delta+1)^{-1}u_k\|_2\|u_k\|_2 \leq C(\epsilon)\gamma_k^{-1}\|u_k\|_2^2.\end{aligned}$$

Proposition 3.4 follows since $\gamma_k \to \infty$ as $k \to \infty$. □

Remark. The semiclassical limit of the u_k's is a distribution supported on the one parameter family of transverse geodesics on the cylinder.

An exceptional sequence of eigenfunction packets is an orthonormal sequence for which formula (1.5) fails. If M has ergodic geodesic flow, then such a sequence must have density zero in a complete orthonormal basis for L^2M. Our main result is that such exceptional sequences do indeed exist.

Theorem 3.5. *There exists a surface M with ergodic geodesic flow, and volume one, which supports an exceptional sequence of eigenfunction packets ψ_k. Moreover, given $\delta > 0$, we may suppose that the eigenvalues associated to each packet lie on an interval of length at most δ. The surface M will depend upon the given δ.*

Proof. The surface M with ergodic geodesic flow is provided by Theorem 2.2. Suppose that u_k is the sequence of quasimodes which we constructed earlier in this section. Let E_k be the spectral projection onto the interval $[\gamma_k - \delta/2, \gamma_k + \delta/2]$. The sequence $\psi_k = E_k u_k/\|E_k u_k\|_2$. By Lemma 3.2, $\|(I - E_k)u_k\|_2 \leq 2\epsilon\delta^{-1}\|u_k\|_2$. If $\epsilon\delta^{-1}$ is sufficiently small, then it follows from Proposition 3.4 that ψ_k is an exceptional sequence. Specifically, for the zero'th order pseudodifferential operator A of Proposition 3.4,

$$\begin{aligned}|\langle A\psi_k, \psi_k\rangle| &= \frac{1}{\|E_k u_k\|_2^2}\,|\langle AE_k u_k, E_k u_k\rangle| \\ &\leq 2|\langle AE_k u_k, E_k u_k\rangle| = 2|\langle Au_k, u_k\rangle| + O(\epsilon).\end{aligned}$$

By Proposition 3.4, $\langle Au_k, u_k\rangle \to 0$ as $k \to \infty$. But the average value of the symbol of A is a fixed non-zero constant, independent of ϵ. □

Remark. By Weyl's formula, the number of eigenvalues less than λ satisfies $N(\lambda) \sim c\,\mathrm{vol}(M)\lambda$, as $\lambda \to \infty$. If $\mathrm{vol}(M) = 1$, then the average level of spacing of the eigenvalues is normalized to be c^{-1}. By contrast, our eigenfunction packets are δ-localized, where δ is arbitrary. This explains the reason for normalizing the volume in Theorem 3.5.

REFERENCES

1. Burns, K. and Gerber, M., *Real analytic Bernoulli geodesic flows on S^2*, Ergodic theory and dynamical systems, 9 (1989), pp. 27–45. MR **90e:**58126
2. Colin de Verdière, Y., *Ergodicité et fonctions propres du laplacian*, Communications in Mathematical Physics, 102 (1985), pp. 497–502.
3. Heller, E.J., *Wavepacket dynamics and quantum chaology*, Chaos and Quantum Physics, Les Houches, 1989, North Holland Publishing Company, Amsterdam, 1991, pp. 547–664. MR **94i:**81031
4. Sarnak, P.J., *Arithmetic Quantum Chaos*, Israel Mathematics Conference Proceedings, 8, Bar–Ilan University, 1995, pp. 183–236. MR **96d:**11059

5. Schnirelman, A.I., *Ergodic Properties of Eigenfunctions*, Uspehi Mat. Nauk, 29 (1974), pp. 181–182.
6. Zelditch, S., *Uniform distribution of eigenfunctions on a compact hyperbolic surface*, Duke Mathematics Journal, 55 (1987), pp. 919–941. MR **89d:**58129

Department of Mathematics, Purdue University, West Lafayette, Indiana 47906