QUANTUM ERGODICITY FOR RESTRICTIONS TO HYPERSURFACES

SEMYON DYATLOV AND MACIEJ ZWORSKI

Abstract. Quantum ergodicity theorem states that for quantum systems with ergodic classical flows, eigenstates are, in average, uniformly distributed on energy surfaces. We show that if N is a hypersurface in the position space satisfying a simple dynamical condition, the restrictions of eigenstates to N are also quantum ergodic.

1. Introduction

In a recent paper [6] Toth and Zelditch proved a remarkable result stating that if (M,g) is a compact manifold with an ergodic geodesic flow, then quantum ergodicity holds for restrictions of eigenfunctions to hypersurfaces satisfying a certain dynamical condition. The purpose of this note is to provide a semiclassical generalization of their result. Our approach avoids global constructions and calculations by reducing equidistribution for restrictions to the equidistribution in the ambient manifold. The geometric condition (1.1) enters to obtain a decorrelation between contributions to the restrictions coming from different parts of phase space.

For the standard quantum ergodicity result established by Shnirelman, Zelditch and Colin de Verdière, see [2],[3],[6] and references given there. We state the simplest version of the restriction result as follows.

Suppose that (M,g) is a compact Riemannian manifold with an ergodic geodesic flow $\varphi_t : S^*M \to S^*M$, and suppose that $N \subset M$ is an open smooth hypersurface. If $S^*_N M \subset S^*M$ denotes the cosphere bundle of M restricted to N and $B^*N \subset T^*N$ the coball bundle of N we let

$$\pi_1 : S^*_N M \to B^*N$$

be the restriction of an element of $S^*_N M$ to TN. It defines a unique nontrivial involution, which is the reflection across the hyperplane T^*N:

$$\gamma_1 : S^*_N M \to S^*_N M\,, \quad \pi_1 \circ \gamma_1 = \pi_1\,, \quad \gamma_1 \circ \gamma_1 = id.$$

We make the following dynamical assumption on N:

$$\begin{gathered}\text{The set of } \rho \in S^*_N M \text{ satisfying } \varphi_t(\rho) \in S^*_N M \\ \text{and } \varphi_t(\gamma_1(\rho)) = \gamma_1(\varphi_t(\rho)) \ \text{ for some } t \neq 0, \text{ has measure } 0.\end{gathered} \tag{1.1}$$

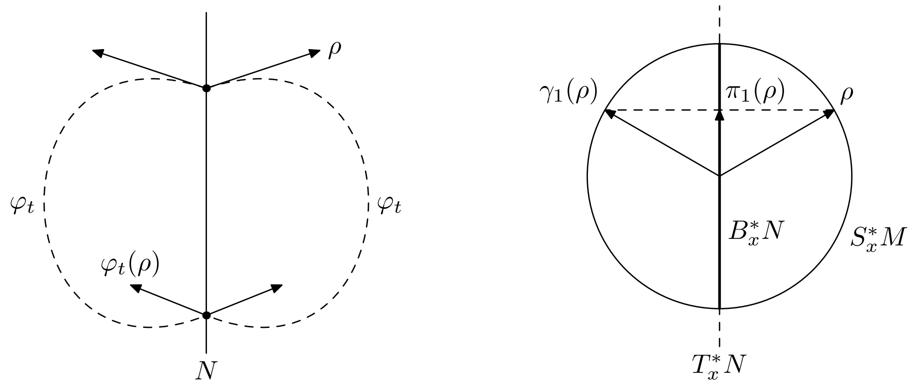

FIGURE 1. Left: the situation prohibited almost everywhere by the dynamical assumption (1.1). Right: The projection map π_1 and the reflection map γ_1 in the cotangent space over some point $x \in N$.

A natural measure on $S^*_N M$ is obtained from the Liouville measure on S^*M, μ_1. For each $x \in M$ it induces a measure on $S^*_x M$, μ_x, such that

$$\mu_1(\Omega) = \int_M \mu_x(\Omega \cap S^*_x M) d\,\mathrm{vol}_g(x), \quad \Omega \subset S^*M.$$

This defines a measure on $S^*_N M$:

$$\nu_1(\Gamma) = \frac{1}{\mu_1(S^*M)} \int_N \mu_x(\Gamma \cap S^*_x M) d\,\mathrm{vol}_{g|_N}(x), \quad \Gamma \subset S^*_N M, \tag{1.2}$$

where $g|_N$ is the metric on N induced by g. (Ω and Γ are Borel sets.)

Now let $\{u_j\}_{j=0}^\infty$ be the complete set of eigenfunctions of the Laplacian on (M, g):

$$-\Delta_g u_j = \lambda_j^2 u_j, \quad \|u_j\|_{L^2} = 1, \quad 0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots .$$

The statement of the theorem uses the standard concept of a pseudodifferential operator on a manifold – see [4, §18.2].

Theorem 1. *Let N be a smooth open hypersurface satisfying (1.1). Suppose that $A \in \Psi^0_{\mathrm{phg}}(N)$ is a classical pseudodifferential operator on N, with $\mathrm{WF}(A) \cap S^*N \Subset S^*N$, that is A compactly supported inside N. Suppose that $v_j := u_j|_N$. Then*

$$\frac{1}{\lambda^n} \sum_{\lambda_j \leq \lambda} \left| \langle A v_j, v_j \rangle_{L^2(N, d\,\mathrm{vol}_{g|_N})} - \int_{S^*_N M} \pi_1^* \sigma(A)\, d\nu_1 \right| \longrightarrow 0, \quad \lambda \to \infty, \tag{1.3}$$

where $\sigma(A)$ is the principal symbol of A (a homogeneous function of degree 0 on $T^*N \setminus \{0\}$), and the measure ν_1 is defined in (1.2).

Remark. The measure $(\pi_1)_* \nu_1$ can be explicitely calculated – see [6] and §5. Here we emphasize that it is smooth on $S^*_N M$. Its invariant meaning becomes more apparent in the semiclassical formulation below, which also easily allows more general restrictions

$au_j|_N + b\lambda^{-1}\partial_\nu u_j|_N$, for $a, b \in C^\infty(N)$. We note also that this nonsemiclassical formulation of quantum ergodicity only implies the angular equidistribution of v_j in $T_x^* N$. That is natural for the standard quantum ergodicity since u_j concentrate on S^*M but not in this case as v_j's can be microsupported anywhere in B^*N. That is remedied in the semiclassical Theorem 2.

A semiclassical version of quantum ergodicity was first provided by Helffer–Martinez–Robert [3] and Theorem 1 is a consequence of a more general semiclassical result. To make the presentation simpler we will consider a version presented in [7, Chapter 15], sufficient to deduce Theorem 1. Similarly, we will only present the result for Schrödinger operators even though (as can be seen from the proof) it holds for more general operators. The proof uses some ideas of [2, Appendix D] but we will not refer to any results from that paper. Refinements allowing energy ranges of size h should also be possible by those methods but, following [6], we present the case of fixed size energy ranges only.

Suppose

$$P(h) := -h^2 \Delta_g + V(x), \quad V \in C^\infty(M; \mathbb{R}),$$

is a semiclassical Schrödinger operator on M. We consider $P(h)$ as a self-adjoint operator acting on half-densities (see [7, Chapter 9]), $L^2(M, \Omega_M^{\frac{1}{2}})$. This is helpful when more general operators are considered.

The classical symbol of $P(h)$ is given by

$$p(x, \xi) = |\xi|_g^2 + V(x), \quad (x, \xi) \in T^*M,$$

and p defines the Hamiltonian flow,

$$\varphi_t := \exp(tH_p) \ : \ p^{-1}(E) \longrightarrow p^{-1}(E), \quad E \in \mathbb{R}.$$

We make the following assumption on a range on energies:

$$\text{For } E \in [a, b],\ dp|_{p^{-1}(E)} \neq 0, \text{ and the flow } \varphi_t : p^{-1}(E) \to p^{-1}(E) \text{ is ergodic,} \tag{1.4}$$

where ergodicity is with respect to the Liouville measure μ_E on $p^{-1}(E)$.

Now, let N be a smooth open hypersurface in M. We define the following analogue of $S_N^* M$:

$$\Sigma_E := p^{-1}(E) \cap \pi^{-1}(N), \tag{1.5}$$

where $\pi : T^*M \to M$ is the natural projection. We note that Σ_E is a smooth hypersurface in $p^{-1}(E)$ if

$$V(x) = E \implies dV(x) \notin N_x^* N, \tag{1.6}$$

and for simplicity we make this assumption for $E \in [a, b]$. For $E > 0$ it is satisfied when $V \equiv 0$, and that is the setting of Theorem 1.

By restricting elements of Σ_E to TN we obtain a map

$$\pi_E : \Sigma_E \to B_E := \pi_E(\Sigma_E) \subset T^*N, \tag{1.7}$$

which is a diffeomorphism almost everywhere. It defines a unique nontrivial involution

$$\gamma_E : \Sigma_E \to \Sigma_E, \quad \pi_E \circ \gamma_E = \pi_E, \quad \gamma_E \circ \gamma_E = id.$$

The assumption on N is analogous to the assumption (1.1):

$$\begin{gathered} \text{For } E \in [a,b], \text{ the set of } \rho \in \Sigma_E \text{ satisfying } \varphi_t(\rho) \in \Sigma_E \\ \text{and } \varphi_t(\gamma_E(\rho)) = \gamma_E(\varphi_t(\rho)) \ \text{ for some } t \neq 0, \text{ has measure } 0. \end{gathered} \tag{1.8}$$

We denote by $u_j(h)$ a normalized eigenfunction of $P(h)$ with an eigenvalue $E_j(h)$,

$$P(h)u_j(h) = E_j(h), \quad \|u_j(h)\|_{L^2(M,\Omega_M^{\frac12})} = 1.$$

To formulate the next theorem we need to restrict half-densities to N and that requires a choice. Suppose $f \in C^\infty(M)$, $f|_N = 0$, $df|_N \neq 0$. Informally, the restriction is now defined using, $|dx|^{\frac12} = |dy|^{\frac12}|df|^{\frac12}$, $x \in M$, $y \in N$. More precisely if, in local coordinates, $x = (x', x_n)$, $N = \{x_n = 0\}$ then, in the half-density notation of [7, §9.1],

$$\big(u(x)|dx|^{\frac12}\big)|_N := u(x',0)|dx'|^{\frac12}\left|\frac{\partial f}{\partial x_n}(x',0)\right|^{-\frac12}. \tag{1.9}$$

Theorem 2. *Let $P(h) = -h^2\Delta_g + V(x)$ be a Schrödinger operator satisfying (1.4) and N be a smooth open hypersurface satisfying (1.8). Suppose that $A \in \Psi^0(N, \Omega_N^{\frac12})$ is a compactly supported semiclassical pseudodifferential operator. For $Q \in \Psi^m(M, \Omega_M^{\frac12})$ define $v_j := Qu_j(h)|_N$, where the restriction operator on half densities is defined in (1.9). Then*

$$h^n \sum_{E_j\in[a,b]} \left| \langle Av_j, v_j\rangle_{L^2(N,\Omega_N^{\frac12})} - \int_{\Sigma_{E_j}} \pi_{E_j}^* \sigma(A)|\sigma(Q)|^2 d\nu_{E_j}\right| \longrightarrow 0, \quad h \to 0, \tag{1.10}$$

where $\sigma(A) \in S^0(T^*N)$ is the symbol of A, $\sigma(Q) \in S^m(T^*M)$ is the symbol of Q, and

$$\nu_E = \frac{1}{\mu_E(p^{-1}(E))}\frac{1}{|H_p f|}\frac{(\sigma|_{\Sigma_E})^{n-1}}{(n-1)!} \tag{1.11}$$

with μ_E the Liouville measure and f defining the restriction of half-densities in (1.9).

The now standard argument due to Colin de Verdière and Zelditch and described in [7, Theorem 15.5] shows that this result provides pointwise convergence for a density one subsequence.

The dynamical condition of Toth–Zelditch [6] is stated using Poincaré return times but the analysis in that paper shows that it is equivalent to our condition. The paper [6] provides interesting examples for which it is satisfied.

ACKNOWLEDGEMENTS. We would like to thank Nicolas Burq and Stéphane Nonnenmacher for helpful comments on the first version of this paper. The partial support by National Science Foundation is also gratefully acknowledged.

2. SEMICLASSICAL PRELIMINARIES

We will use the calculus of semiclassical pseudodifferential operators described in [7, §9.3, §14.2]. For a compact manifold, X (which could different from the compact manifold M considered in Section 1), the class $\Psi^m(X)$ denotes operators of order m, so that, for instance $-h^2\Delta_g \in \Psi^2(M)$. We have the symbol map, σ, appearing in the following exact sequence

$$0 \longrightarrow h\Psi^{m-1}(X) \longrightarrow \Psi^m(X) \xrightarrow{\sigma} S^m(T^*X)/hS^{m-1}(T^*X) \longrightarrow 0,$$

where S^m denotes the standard space of symbols. The quantization map Op_h : $S^m(T^*X) \to \Psi^m$ satisfies

$$\sigma(\mathrm{Op}_h(a)) = a \mod hS^{m-1}(T^*X).$$

We also introduce the class of *compactly microlocalized pseudodifferential operators*, $\Psi^{\mathrm{comp}}(X)$: $A \in \Psi^{-\infty}(X)$ is in $\Psi^{\mathrm{comp}}(X)$ if for some $\chi \in C_{\mathrm{c}}^\infty(T^*X)$,

$$\mathrm{Op}_h(1-\chi)A \in h^\infty\Psi^{-\infty}(X).$$

For this class the definition of $\mathrm{WF}_h(A)$ given in [7, §8.4] applies. From the same section we take the definition of microlocal equality of operators.

Following [1, §2.3], [5, §3], and [7, §11.2] we consider Fourier integral operators quantizing a canonical transformation $\kappa : U_1 \to U_2$, $U_1 \Subset T^*X$ and $U_2 \Subset T^*Y$, κ defined on a neighbourhood of U_1: we say that an operator $F : L^2(X) \to L^2(Y)$, quantizes κ if for any $A \in \Psi^{\mathrm{comp}}(Y)$ with $\mathrm{WF}_h(A) \Subset U_2$,

$$F^*AF = B, \quad B \in \Psi^{\mathrm{comp}}(X), \quad \sigma(B) = \kappa^*\sigma(A). \tag{2.1}$$

We further require that F be microlocally unitary in the sense that $F^{-1} = F^*$ microlocally near $U_1 \times U_2$. If F quantizes κ, then the operator F^* quantizes κ^{-1}.

The standard example is given by $F(t) = e^{-itP(h)}$, where $P(h) = -h^2\Delta_g + V(x) \in \Psi^2(M)$ (or a more general operator) which quantizes the Hamiltonian flow $\varphi_t := \exp(tH_p)$.

We say that a tempered operator (see [7, §8.4]) $G : L^2(X) \to L^2(Y)$, is *compactly microlocalized* if for some $A \in \Psi^{\mathrm{comp}}(X)$ and $B \in \Psi^{\mathrm{comp}}(X)$,

$$AGB - G \in h^\infty\Psi^{-\infty}. \tag{2.2}$$

In that case we can define $\mathrm{WF}_h(G) \subset T^*X \times T^*Y$, by taking the twisted WF_h of its Schwartz kernel, K_G:

$$\mathrm{WF}_h(G) := \{(x, -\xi; y, \eta) \ : \ (x, \xi; y, \eta) \in \mathrm{WF}_h(K_G)\}.$$

If F quantizes some canonical transformation κ, then $\mathrm{WF}_h(F)$ lies inside the graph of κ.

We recall from [7, Theorem 14.9] that for $P(h) = -h^2\Delta + V(x)$ (and by the same methods for more general operators) and $f \in C_c^\infty(\mathbb{R})$,

$$f(P(h)) \in \Psi^{\mathrm{comp}}(M), \quad \sigma\big(f(P(h))\big) = f(p). \tag{2.3}$$

As before, let $(u_j(h))_{j\in\mathbb{N}}$ be the full orthonormal system of eigenfunctions of $P(h)$ with eigenvalues $E_j(h)$. From [7, Theorem 15.3] applied to the operator $f(P(h))A$, where $A \in \Psi^m(M)$ and $f \in C_c^\infty(\mathbb{R})$, we obtain

$$(2\pi h)^n \sum_j f(E_j)\langle Au_j, u_j\rangle = \int_{T^*M} f(p)\sigma(A)\, d\mu_\sigma + \mathcal{O}(h), \tag{2.4}$$

where μ_σ is the symplectic measure, $\mu_\sigma = \sigma^n/n!$.

We conclude this section with two lemmas. The first one, in the spirit of [2, Appendix D], gives estimates using L^2 norms of symbols:

Lemma 1. *There exists a constant C such that for each $a' < a < b < b'$ and each $A \in \Psi^m(M)$,*

$$(2\pi h)^n \sum_{E_j \in [a,b]} \|Au_j\|_{L^2}^2 \leq \|\sigma(A)\|^2_{L^2(p^{-1}([a',b']))} + \mathcal{O}(h) \tag{2.5}$$

where the L^2 norm of $\sigma(A)$ is taken with respect to the measure μ_σ.

More generally, if $N \subset M$ is a fixed smooth submanifold (of any dimension), then there exists a constant C such that for each $\widetilde{A} \in \Psi^m(N)$,

$$h^n \sum_{E_j \in [a,b]} \|\widetilde{A}(u_j|_N)\|_{L^2}^2 \leq C\|\sigma(\widetilde{A})\|^2_{L^2(\pi(p^{-1}([a',b'])\cap T_N^*M))} + \mathcal{O}(h). \tag{2.6}$$

Here T_N^*M is the cotangent bundle of M restricted to N and $\pi : T_N^*M \to T^*N$ is the projection.

Remark. We note that in the case when $\widetilde{A} = 1$ we recover the bound

$$h^n \sum_{E_j \in [a,b]} \|u_j|_N\|_{L^2}^2 \leq C\,. \tag{2.7}$$

It is essential to average as for individual eigenfuctions the bound $Ch^{\frac{n-k}{2}}$ is optimal.

Proof. To show (2.5), take $f \in C_c^\infty(a',b')$ such that $0 \le f \le 1$ everywhere and $f = 1$ on $[a,b]$. Then we write by (2.4),

$$(2\pi h)^n \sum_{E_j\in[a,b]} \|Au_j\|_{L^2}^2 \le (2\pi h)^n \sum_j f(E_j)\langle A^*Au_j,u_j\rangle$$
$$= \int_{T^*M} f(p)|\sigma(A)|^2\,d\mu_\sigma + \mathcal{O}(h) \le \int_{p^{-1}([a',b'])} |\sigma(A)^2|\,d\mu_\sigma + \mathcal{O}(h).$$

To show (2.6), denote by $R_N : C^\infty(M) \to C^\infty(N)$ the restriction operator and note that

$$h^n \sum_{E_j\in[a,b]} \|\widetilde{A}(u_j|_N)\|_{L^2}^2 \le h^n \sum_j |f(E_j)|^2 \|\widetilde{A}R_N u_j\|_{L^2}^2 = h^n \sum_j \|\widetilde{A}R_N f(P)u_j\|_{L^2}^2$$
$$= h^n \|\widetilde{A}R_N f(P)\|_{\mathrm{HS}}^2.$$

The Hilbert–Schmidt norm on the right-hand side is equal to the L^2 norm of the Schwartz kernel K of $\widetilde{A}R_N f(P)$. Note that $K = \mathcal{O}(h^\infty)$ away from the diagonal of N embedded in $N \times M$. To estimate K near the diagonal, we choose local coordinates $x = (x', x'')$, $x' \in \mathbb{R}^k$, $x'' \in \mathbb{R}^{n-k}$, where $k = \dim N$, on M near some point of N, in which N is given by $\{x'' = 0\}$. If $\tilde a$ is the full symbol of $\widetilde{A}$ in these coordinates (in the standard quantization) and $\tilde b$ is the full symbol of the pseudodifferential operator $f(P(h))$, then we can write

$$K(x,y) = (2\pi h)^{-n-k} \int e^{\frac{i}{h}(x\cdot\eta - z\cdot\eta + z\cdot\xi' - y\cdot\xi)}\tilde a(x,\eta)\tilde b(z,0,\xi)\,dzd\eta d\xi,$$

here $y,\xi \in \mathbb{R}^n$ and $x,z,\eta \in \mathbb{R}^k$. By the unitarity of the (semiclassical) Fourier transform, the $L^2_{x,y}$ norm of $K(x,y)$ is equal to the $L^2_{x,\xi}$ norm of

$$K_1(x,\xi) = (2\pi h)^{-n/2-k} \int e^{\frac{i}{h}(x\cdot\eta - z\cdot\eta + z\cdot\xi')}\tilde a(x,\eta)\tilde b(z,0,\xi)\,dzd\eta.$$

The method of stationary phase shows that

$$K_1(x,\xi) = (2\pi h)^{-n/2} e^{\frac{i}{h}x\cdot\xi'}(\tilde a(x,\xi')\tilde b(x,0,\xi) + \mathcal{O}_{C^\infty}(h)).$$

Now, $h^{n/2}$ times the L^2 norm of K_1 is bounded by a constant times the L^2 norm of $\tilde a$ on the set $\pi(\operatorname{supp}\tilde b \cap T^*_N M)$, with an $\mathcal{O}(h)$ remainder. □

From the lemma we recover the standard fact that for each $a < b$, there exists a constant C such that

$$\#\{j \ :\ E_j \in [a,b]\} \le Ch^{-n}. \tag{2.8}$$

To formulate the next lemma we define

$$\operatorname{Diag}(T^*M) := \{(\rho,\rho) : \rho \in T^*M\} \subset T^*M \times T^*M.$$

Lemma 2. *Suppose that $G : L^2(M) \to L^2(M)$ is a compactly microlocalized tempered operator in the sense of (2.2), and that $f \in C_c^\infty(\mathbb{R})$. Then for G satisfying*

$$\mathrm{WF}_h(G) \cap \mathrm{Diag}(T^*M) = \emptyset,$$

we have

$$\sum_j f(E_j)\langle Gu_j, u_j\rangle = \mathcal{O}(h^\infty). \tag{2.9}$$

Proof. The left-hand side of (2.9) is equal to the trace of $Gf(P(h))$. We can write G as a finite sum of operators of the form X_1GX_2, where $X_1, X_2 \in \Psi^{\mathrm{comp}}$ satisfy $\mathrm{WF}_h(X_1) \cap \mathrm{WF}_h(X_2) = \emptyset$. Then by the cyclicity of the trace,

$$\mathrm{Tr}(X_1GX_2f(P)) = \mathrm{Tr}(X_2f(P)X_1G) = \mathcal{O}(h^\infty),$$

as $X_2f(P)X_1 \in h^\infty\Psi^{-\infty}$. □

3. Decorrelation for Fourier integral operators

In the proof of Theorem 2 we will encounter expressions involving $\langle Fu_j, u_j\rangle$, where $(u_j(h))_{j\in\mathbb{N}}$ is the full orthonormal system of eigenfunctions of $P(h) = -h^2\Delta_g + V(x)$ with eigenvalues $E_j(h)$, and F is a compactly microlocalized semiclassical Fourier integral operator. This section shows that the sum of such terms over j in an $\mathcal{O}(1)$ sized spectral window is negligible when the canonical relation of F satisfies a 'nonreturning' assumption; we call this phenomenon *decorrelation* for Fourier integral operators.

Assume F is a compactly microlocalized tempered operator $L^2(M) \to L^2(M)$, in the sense of (2.2), and

$$\|F\|_{L^2\to L^2} = \mathcal{O}(1),\ \mathrm{WF}_h(F) \subset \{(\rho, \kappa(\rho)) : \rho \in K_1\}, \tag{3.1}$$

where $\kappa : V_1 \to V_2$ is a canonical transformation, $V_1, V_2 \subset T^*M$ are open sets, and $K_j \subset V_j$ are compact sets such that $\kappa(K_1) = K_2$. (In our case, F will be a Fourier integral operator, but this is not required in the proof.)

For each $t \in \mathbb{R}$, define the t-exceptional set,

$$\mathcal{E}_\kappa(t) := \{\rho \in K_1 \cap \varphi_{-t}(K_1) : \varphi_t(\kappa(\rho)) = \kappa(\varphi_t(\rho))\},\quad \varphi_t := \exp(tH_p). \tag{3.2}$$

The decorrelation result is given as follows:

Lemma 3. *Suppose that $a < b$ are fixed, and that there exists $t_0 > 0$ such that*

$$\mu_\sigma\Big(p^{-1}([a,b]) \cap \bigcup_{|t|\geq t_0} \mathcal{E}_\kappa(t)\Big) = 0, \tag{3.3}$$

where $\mathcal{E}_\kappa(t)$ is given by (3.2) and μ_σ is the symplectic measure.

Then for each F satisfying (3.1),

$$h^n \sum_{E_j \in [a,b]} |\langle F u_j, u_j \rangle| \to 0 \text{ as } h \to 0. \tag{3.4}$$

Proof. Take $T > t_0$ and denote

$$\widetilde{K}_T := \bigcup_{t_0 \le |t| \le T} \mathcal{E}_\kappa(t). \tag{3.5}$$

Then $\widetilde{K}_T$ is a compact subset of U_1 and $\mu_\sigma(\widetilde{K}_T \cap p^{-1}([a,b])) = 0$. Therefore, there exists an open set $\widetilde{U}_T \subset U_1$ and constants $a' < a$ and $b' > b$ such that

$$\widetilde{K}_T \subset \widetilde{U}_T, \qquad \mu_\sigma(\widetilde{U}_T \cap p^{-1}([a',b'])) \le T^{-1}.$$

Take $X_T \in \Psi^{\mathrm{comp}}(M)$ satisfying $|\sigma(X_T)| \le 1$, $\mathrm{WF}_h(X_T) \subset \widetilde{U}_T$, and $X_T = 1$ microlocally near $\widetilde{K}_T$. Since F is bounded on $L^2(M)$, $|\langle F X_T u_j, u_j\rangle| \le C \|X_T u_j\|_{L^2}$. Hence (2.5) and (2.8) give

$$\begin{aligned} h^n \sum_{E_j \in [a,b]} |\langle F X_T u_j, u_j \rangle| &\le C\Big(h^n \sum_{E_j \in [a,b]} \|X_T u_j\|_{L^2}^2 \Big)^{1/2} \\ &\le C(\|\sigma(X_T)\|^2_{L^2(p^{-1}([a',b']))} + \mathcal{O}_T(h))^{1/2} \le C(T^{-1} + \mathcal{O}_T(h))^{1/2}, \end{aligned} \tag{3.6}$$

where C denotes a constant independent of T and h.

We now analyse the contribution of $F_1 := F(1 - X_T)$. For that define

$$\langle F_1 \rangle_T := \frac{1}{T} \int_0^T e^{itP(h)/h} F_1 e^{-itP(h)/h}\, dt.$$

For each eigenfunction u_j, we have

$$\langle F_1 u_j, u_j \rangle = \langle \langle F_1 \rangle_T u_j, u_j \rangle.$$

We now take some $f \in C_c^\infty(\mathbb{R})$ such that $0 \le f \le 1$ everywhere and $f = 1$ near $[a,b]$. Then by (2.8),

$$\begin{aligned} h^n \sum_{E_j \in [a,b]} |\langle F_1 u_j, u_j \rangle| &= h^n \sum_{E_j \in [a,b]} |\langle \langle F_1 \rangle_T u_j, u_j \rangle| \\ &\le C\Big(h^n \sum_j f(E_j) \|\langle F_1 \rangle_T u_j\|_{L^2}^2 \Big)^{1/2}. \end{aligned} \tag{3.7}$$

We now write

$$\begin{aligned} h^n \sum_j f(E_j) \|\langle F_1 \rangle_T u_j\|^2_{L^2} &= h^n \sum_j f(E_j) \langle \langle F_1 \rangle_T^* \langle F_1 \rangle_T u_j, u_j \rangle \\ = \frac{1}{T^2} \int_0^T \int_0^T h^n \sum_j f(E_j) &\langle e^{i(s-t)P(h)/h} F_1^* e^{i(t-s)P(h)/h} F_1 u_j, u_j \rangle \, dt ds. \end{aligned}$$

Since $\|e^{i(s-t)P(h)/h}F_1^*e^{i(t-s)P(h)/h}F_1\|_{L^2\to L^2}$ is bounded uniformly in t,s,h, we estimate the integral over the region $|t-s|\leq t_0$ using the upper bound on the number of eigevalues, (2.8),

$$\frac{1}{T^2}\int_{\substack{0\leq t,s\leq T\\|t-s|\leq t_0}} h^n\sum_j f(E_j)\langle e^{i(s-t)P(h)/h}F_1^*e^{i(t-s)P(h)/h}F_1u_j,u_j\rangle\,dtds\leq CT^{-1},\tag{3.8}$$

where C is again a constant independent of T and h.

It remains to estimate the integral over the region $t_0\leq|t-s|\leq T$. However, for $t_0\leq|r|\leq T$, $\mathrm{WF}_h(e^{irP(h)/h}F_1^*e^{-irP(h)/h}F_1)\subset$

$$\{(\rho,\rho')\mid \rho\in K_1\setminus\widetilde{K}_T,\ \varphi_r(\rho')\in K_1,\ \kappa(\varphi_r(\rho'))=\varphi_r(\kappa(\rho))\}.\tag{3.9}$$

The definition of $\widetilde{K}_T$ – see (3.2) and (3.5) – shows that the set in (3.9) does not intersect $\mathrm{Diag}(T^*M)$. This means that the operator $G=e^{irP(h)/h}F_1^*e^{-irP(h)/h}F_1$ satisfies the hypothesis of Lemma 2, and by (2.9) we find

$$\frac{1}{T^2}\iint_{\substack{0\leq t,s\leq T\\|t-s|\geq t_0}} h^n\sum_j f(E_j)\langle e^{i(s-t)P(h)/h}F_1^*e^{i(t-s)P(h)/h}F_1u_j,u_j\rangle\,dtds=\mathcal{O}_T(h^\infty).$$

Combining this with (3.8) and recalling (3.6) and (3.7), we get

$$h^n\sum_{E_j\in[a,b]}|\langle Fu_j,u_j\rangle|\leq(CT^{-1}+\mathcal{O}_T(h))^{1/2},$$

where C is a constant independent of T and h. By choosing T large and then h small, we obtain (3.4). □

4. Quantum ergodicity for restrictions

We will now prove Theorem 2 and we use the notation from the second (semiclassical) part of Section 1. To simplify the presentation we put $Q=Id$. The general case is similar.

We start with some geometric observations. The condition (1.6) shows that, in the notation of (1.7), $B_E:=\pi_E(\Sigma_E)\subset T^*N$, is a smooth manifold with a smooth boundary. Any $\rho\in B_E\setminus\partial B_E$ is a regular value of π_E; moreover, $\pi_E^{-1}(\rho)=\{\rho_+,\rho_-\}$, π_E is a local diffeomorphism near $\rho_\pm$, and the involution γ_E is given by $\gamma_E(\rho_\pm)=\rho_\mp$. The Hamilton vector field H_p is transversal to Σ_E at $\rho_\pm$.

To prove Theorem 2 we we can assume that $[a,b]$ is a small neighbourhood of a fixed energy level E. We then decompose any compactly supported $A\in\Psi^0(N)$ as follows:

$$A=\sum_{j=1}^J\widetilde{A}_{j,\epsilon}+A_\epsilon+(1-X_E)A,\tag{4.1}$$

where

- $X_E \in \Psi^{\rm comp}(N)$ is microlocally equal to Id near $B_E \subset T^*N$, and $\mathrm{WF}_h(X_E)$ is contained in small neighbourhood of B_E,
- $\widetilde{A}_{j,\epsilon} \in \Psi^{\rm comp}(N)$, and $\mathrm{WF}_h(\widetilde{A}_{j,\epsilon})$ is a small open subset of $B_E \setminus \partial B_E$,
- $A_\epsilon \in \Psi^{\rm comp}(N)$, $\mu_\sigma(\mathrm{WF}_h(A_\epsilon)) < \epsilon$, where μ_σ is the symplectic measure.

The estimate (2.6) in the second part of Lemma 1 shows that the contribution of $(1 - X_E)A$ is negligible, and that the contribution of A_ϵ will disappear in $\epsilon \to 0$ limit.

Hence we only need to prove Theorem 2 for terms of the form $\widetilde{A}_{j,\epsilon}$. We assume now that $\rho \in B_E \setminus \partial B_E$, and that $\widetilde{A} \in \Psi^{\rm comp}(N)$ is microlocalized in a small neighborhood $V \subset T^*N$ of ρ. Choose small $\delta > 0$ and define the set

$$U := \{\varphi_t(\tilde{x}, \tilde{\xi}) \ : \ |t| < \delta,\ (\tilde{x}, \tilde{\xi}) \in \Sigma_{E+\tau} \cap \pi^{-1}_{E+\tau}(V),\ |\tau| < \delta\}.$$

If V and δ are small enough, then we can write $U = U_1 \sqcup U_2$, where U_ℓ, $\ell = 1, 2$, are open subsets of T^*M (one of which is a neighborhood of ρ_+ and the other of ρ_-) and moreover, the maps $\kappa_\ell : U_\ell \to V \times \{|t|, |\tau| < \delta\}$,

$$\kappa_\ell \ : \ \varphi_t(\tilde{x}, \tilde{\xi}) \longmapsto (\pi_{E+\tau}(\tilde{x}, \tilde{\xi}), t, \tau)\,, \quad (\tilde{x}, \tilde{\xi}) \in \Sigma_{E+\tau} \cap U_\ell,\ |t|, |\tau| < \delta, \tag{4.2}$$

are diffeomorphisms. The maps κ_ℓ are symplectomorphisms if we consider $\{|t|, |\tau| < \delta\}$ as a subset of $T^*\mathbb{R}_t$, with τ the momentum corresponding to t.

Fix a local coordinate system $x = (x', x_n)$ on M such that $N = \{x_n = 0\}$. To simplify the symbol calculations, we additionally choose this coordinate system so that the volume form on N is given by $|dx'|$ and the volume form on M is given by $|dx|$. Consider the operator

$$\mathcal{R} : C^\infty(M) \to C^\infty(N \times \mathbb{R}_t), \quad \mathcal{R}u(t) := (e^{it(P(h)-E)/h}u)|_N, \tag{4.3}$$

then

$$hD_t\mathcal{R} = \mathcal{R}(P(h) - E). \tag{4.4}$$

Take $X_\ell \in \Psi^{\rm comp}(M)$ microlocalized inside U_ℓ, but such that

$$X_\ell = 1 \text{ microlocally near } \kappa_\ell^{-1}(\mathrm{WF}_h(\widetilde{A}) \times \{|\tau|, |t| \leq \delta/2\}).$$

Let $\tilde{\chi}(t) \in C^\infty_{\rm c}(-\delta, \delta)$ be equal to 1 near $[-\delta/2, \delta/2]$. Then

$$B_\ell := \tilde{\chi}(t)\mathcal{R}X_\ell : C^\infty(M) \to C^\infty(N \times \mathbb{R}_t)$$

are compactly microlocalized Fourier integral operators associated to κ_ℓ. This follows from an oscillatory representation of $e^{i(P(h)-E)/h}$ given in [7, §10.2]. Indeed, in coordinates $x = (x', x_n)$ and in the notation of [7, Theorem 10.4],

$$\mathcal{R}u(t, \tilde{x}) := \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}(\psi(t, \tilde{x}, 0, \eta) - y\cdot\eta)} b(t, \tilde{x}, 0, \eta; h) u(y)\, dy d\eta, \tag{4.5}$$

where

$$\psi(0, x, \eta) = x \cdot \eta,\ \partial_t\psi(t, x, \partial_x\psi) = p(x, \partial_x\psi) - E,$$
$$\varphi_t(x, \partial_x\psi(t, x, \eta)) = (\partial_\eta\psi(t, x, \eta), \eta).$$

The microlocalization inside U_ℓ means that $\partial_{\xi_n} p(\tilde{x},0,\xi) \neq 0$, and that implies that $\partial^2_{(t,x'),\eta}\psi$ is nondegenerate. Hence, $\psi(t,\tilde{x},0,\eta)$ is a generating function of κ_ℓ.

Now, let u be an eigenfunction of $P(h)$ with eigenvalue $E' = E + \lambda$, where $\lambda \in [-\delta/2, \delta/2]$. Then $\mathrm{WF}_h(u) \subset p^{-1}([E-\delta/2, E+\delta/2])$ and thus

$$u|_N = (X_1 + X_2)u|_N = B_1 u|_{t=0} + B_2 u|_{t=0} \ \text{ microlocally near } \mathrm{WF}_h(\widetilde{A}).$$

Now, by (4.4) each $w_\ell := B_\ell u$ solves $hD_t w_\ell = \lambda w_\ell$ microlocally near $\mathrm{WF}_h(\widetilde{A})$ for $|t| \leq \delta/2$. Therefore, $w_\ell(t) = e^{it\lambda/h} w_\ell(0)$ microlocally near $\mathrm{WF}_h(\widetilde{A})$ for $|t| \leq \delta/2$.

Take $\chi(t) \in C_c^\infty(-\delta/2, \delta/2)$ that integrates to 1. Then

$$\langle \widetilde{A}(w_\ell|_{t=0}), w_k|_{t=0}\rangle_{L^2(N)} = \langle (\chi(t)\otimes \widetilde{A}) w_\ell, w_k\rangle_{L^2(N\times\mathbb{R}_t)} + \mathcal{O}(h^\infty).$$

Therefore,

$$\begin{aligned}
\langle \widetilde{A}(u|_N), (u|_N)\rangle_{L^2(N)} &= \sum_{\ell,k=1}^2 \langle \widetilde{A}(w_\ell|_{t=0}), w_k|_{t=0}\rangle_{L^2(N)} + \mathcal{O}(h^\infty) \\
&= \sum_{\ell,k=1}^2 \langle (\chi(t)\otimes \widetilde{A}) w_\ell, w_k\rangle_{L^2(N\times\mathbb{R}_t)} + \mathcal{O}(h^\infty) \\
&= \sum_{\ell,k=1}^2 \langle B_k^*(\chi(t)\otimes \widetilde{A}) B_\ell u, u\rangle_{L^2(M)} + \mathcal{O}(h^\infty).
\end{aligned} \tag{4.6}$$

We now need to analyse the operators $B_{k\ell} := B_k^*(\chi(t)\otimes\widetilde{A})B_\ell$. This is split into two cases. For $k = \ell$, $B_{k\ell}$ is a pseudodifferential operator. We can then use the usual quantum ergodicity [7, Theorem 15.4] which we slightly modify:

$$h^n \sum_{E_j \in [E-\delta/2, E+\delta/2]} \left| \langle B_{\ell\ell} u_j, u_j\rangle_{L^2(M)} - \fint_{p=E_j} \sigma(B_{\ell\ell})\, d\mu_{E_j} \right| \to 0\,, \quad h \to 0. \tag{4.7}$$

(To see this from [7, Theorem 15.4] apply the zero-mean quantum ergodicity to the function $B_{\ell\ell} - f(P)$, where $f(\lambda)$ is the mean integral of the symbol of $B_{\ell\ell}$ over $p^{-1}(E+\lambda)$.)

We need to compute the symbol of $B_{\ell\ell}$. For that, we use the integral representation (4.5) and the stationary phase method, applicable since $\partial^2_{(t,x'),\eta}\psi$ is nondegenerate. More precisely, the Schwartz kernel, $B_\ell^* B_\ell(z,y)$ is given by

$$\frac{1}{(2\pi h)^{2n}} \int\limits_{\mathbb{R}^n_{t,\tilde{x}}\times\mathbb{R}^n_\eta\times\mathbb{R}^n_\zeta} e^{\frac{i}{h}(\psi(t,\tilde{x},0,\eta)-\psi(t,\tilde{x},0,\zeta)+z\cdot\zeta - y\cdot\eta)} b(t,\tilde{x},0,\eta)\overline{b(t,\tilde{x},0,\zeta)}\, d\tilde{x}dtd\eta d\zeta.$$

We apply the method of stationary phase in the $\tilde{x}, t, \eta$ variables. The stationary point is given by $\eta = \zeta$, $\varphi_t(\tilde{x}, 0, \partial_x\psi_t(t,\tilde{x},0,\eta)) = (y,\eta)$, and the value of the phase at a stationary point is $(z-y)\cdot\zeta$. At $t = 0$, the leading term in the amplitude is given by $|\partial_{\xi_n} p|^{-1}$. That is, $\sigma(B_\ell^* B_\ell) = |\partial_{\xi_n} p|^{-1}$ on $T_N^* M$ near $\kappa_\ell^{-1}(\mathrm{WF}_h(\widetilde{A}) \times \{|\tau| \le \delta/2\})$. Now, we also see from (4.4) that

$$[P, B_\ell^* B_\ell] = B_\ell^* \tilde{\chi}(t) \mathcal{R}[P, X_\ell] + [P, X_\ell^*] \mathcal{R}^* \tilde{\chi}(t) B_\ell + X_\ell^* \mathcal{R}^* [hD_t, |\tilde{\chi}(t)|^2] \mathcal{R} X_\ell$$

vanishes microlocally near $\kappa_\ell^{-1}(\mathrm{WF}_h(\widetilde{A}) \times \{|\tau|, |t| \le \delta/2\})$. It follows that

$$\sigma(B_\ell^* B_\ell) \circ \kappa_\ell^{-1} = |\partial_{\xi_n} p \circ \kappa_\ell^{-1} \circ \pi_0|^{-1} \text{ near } \mathrm{WF}_h(\widetilde{A}) \times \{|\tau|, |t| \le \delta/2\},$$

where $\pi_0 : T^*(N \times \mathbb{R}_t) \to T^*(N \times \mathbb{R}_t)$ maps $(\tilde{x}, \tilde{\xi}, t, \tau)$ to $(\tilde{x}, \tilde{\xi}, 0, \tau)$. From here and by Egorov's Theorem applied to $\chi(t) \otimes \widetilde{A}$, we get $\sigma(B_{\ell\ell}) \circ \kappa_\ell^{-1} = |\partial_{\xi_n} p \circ \kappa_\ell^{-1} \circ \pi_0|^{-1} \sigma(\widetilde{A}) \chi(t)$ near $\{|\tau| \le \delta/2\}$. Then

$$\begin{aligned} ⨍_{\{p=E_j\}} \sigma(B_{\ell\ell})\, d\mu_{E_j} &= \frac{1}{\mu_{E_j}(p^{-1}(E_j))} \int_{\{\tau = E_j - E\}} \sigma(B_{\ell\ell}) \circ \kappa_\ell^{-1}\, d\tilde{x} d\tilde{\xi} dt \\ &= \frac{1}{\mu_{E_j}(p^{-1}(E_j))} \int_{\Sigma_{E_j} \cap U_\ell} |\partial_{\xi_n} p|^{-1} \pi_{E_j}^* \sigma(\widetilde{A})\, d\tilde{x} d\tilde{\xi}, \end{aligned} \tag{4.8}$$

where we parametrized Σ_{E_j} by $(\tilde{x}, \tilde{\xi}) \in B_{E_j}$.

Now, we consider the case $k \ne \ell$. Then $B_{k\ell}$ is a Fourier integral operator with the canonical transformation $\kappa_{k\ell} := \kappa_k^{-1} \circ \kappa_\ell$. We want to apply the decorrelation result given in Lemma 3.

Using the definition (4.2) of κ_ℓ, we see that the canonical transformation $\kappa = \kappa_{k\ell}$ can be described as follows:

$$\kappa(\varphi_s(\tilde{x}, \tilde{\xi})) = \varphi_s(\gamma_{E'}(\tilde{x}, \tilde{\xi})),\ |s| < \delta,\ (\tilde{x}, \tilde{\xi}) \in \Sigma_{E'} \cap U_\ell.$$

To apply Lemma 3 we need to verify the following: there exists $t_0 > 0$ such that the set

$$\mathcal{E} := \{\rho \in U_\ell \cap \varphi_{-t}(U_\ell)\ :\ \exists\, t,\ |t| \ge t_0\,,\ \varphi_t(\kappa(\rho)) = \kappa(\varphi_t(\rho))\} \subset T^*M\,, \tag{4.9}$$

has μ_σ-measure zero. To see this, suppose that $\rho \in \mathcal{E}$, t is the corresponding time, and $s, s' \in (-\delta, \delta)$ are such that $\rho = \varphi_s(\tilde{x}, \tilde{\xi})$, $\varphi_t(\rho) = \varphi_{s'}(\tilde{x}', \tilde{\xi}')$, $(\tilde{x}, \tilde{\xi}), (\tilde{x}', \tilde{\xi}') \in \Sigma_{E'} \cap U_\ell$. Then $(\tilde{x}', \tilde{\xi}') = \varphi_{t+s-s'}(\tilde{x}, \tilde{\xi})$ and the condition $\varphi_t(\kappa(\rho)) = \kappa(\varphi_t(\rho))$ can be rewritten as

$$(\tilde{x}, \tilde{\xi}) \in \Sigma_{E'},\ \varphi_{t+s-s'}(\tilde{x}, \tilde{\xi}) \in \Sigma_{E'},\ \varphi_{t+s-s'}(\gamma_{E'}(\tilde{x}, \tilde{\xi})) = \gamma_{E'}(\varphi_{t+s-s'}(\tilde{x}, \tilde{\xi})).$$

Put $t_0 > 2\delta$, then $t + s - s' \ne 0$. It now follows from (1.8) that the set $\mathcal{E}$ from (4.9) has measure zero; by Lemma 3, the contributions of $B_{k\ell}$, $k \ne \ell$ to the sum (1.10) go to 0 as $h \to 0$.

Going back to (4.6), (4.7) and (4.8) this means for $\widetilde{A}$ satisfying our localization assumptions

$$h^n \sum_{E_j \in [a,b]} \left| \langle \widetilde{A}(u_j|_N), (u_j|_N) \rangle_{L^2(N)} - \frac{1}{V_j} \sum_{\ell=1,2} \int_{\Sigma_{E_j} \cap U_\ell} |\partial_{\xi_n} p|^{-1} \pi_{E_j}^* \sigma(\widetilde{A})\, dx' d\xi' \right| = o(1),$$

where $V_j := \mu_{E_j}(p^{-1}(E_j))$. This is (1.10) with $f = x_n$. If we chooose a different f to obtain the restriction of half-densities in (1.9), then we obtain a factor of $|\partial_{x_n} f|$:

$$\langle \widetilde{A}_f v, v\rangle_{L^2(N)} = \langle \widetilde{A}|\partial_{x_n} f|^{-\frac{1}{2}} v, |\partial_{x_n} f|^{-\frac{1}{2}} v\rangle_{L^2(N)}, \quad \sigma(\widetilde{A}_f) = |\partial_{x_n} f|^{-1}\sigma(\widetilde{A}).$$

and $H_p f = \partial_{x_n} f \partial_{\xi_n} p$ on $T^*_N M$. This completes the proof of (1.10).

5. From semiclassical estimates to high energy estimates

We conclude the paper by explaining how Theorem 1 follows from Theorem 2. We put $V \equiv 0$ and identify $L^2(M, \Omega^{\frac{1}{2}}_M)$ with $L^2(M, d\,\mathrm{vol}_g)$ by writing half-densities as $u(x)|d\,\mathrm{vol}_g|^{\frac{1}{2}}$, where $u \in L^2(M, d\,\mathrm{vol}_g)$.

Let $x = (x', x_n)$ be normal geodesic coordinates near $(0,0) \in N$, in which $N = \{x_n = 0\}$, $p(x, \xi', \xi_n) = \xi_n^2 + r(x, \xi')$, and $r(x', 0, \xi')$ is the dual of the restriction metric $g|_N$. Suppose that f satisfies $f|_N = 0$, $|df(x)|_g = 1$. In the chosen coordinates the last condition means that $\partial_{x_n} f = 1$, $f = 0$, on N. Hence the restriction of half-densities (1.9) obtained using this choice of f shows that, we obtain an identification with the restriction of functions $u|_N \in L^2(N, d\,\mathrm{vol}_{g|_N})$.

We now identify the measures appearing in Theorems 1 and 2; this can be done locally. In our coordinates, S^*M can be parametrized by (x', x_n, ξ'), $r(x', \xi') \le 1$, $\xi_n = \pm(1 - r(x, \xi'))^{\frac{1}{2}}$ (the parametrization degenerates at $r(x, \xi') = 1$). The Liouville measure is obtained by requiring $d\mu_1 \wedge dp = dx d\xi$, and

$$d\mu_1 = \frac{1}{2|\xi_n|} dx d\xi' = \frac{1}{2(1 - r(x, \xi'))^{\frac{1}{2}}} dx d\xi'.$$

In the notation of (1.2) this gives $d\mu_{(0,x')} = \det g(0, x')^{-\frac{1}{2}} d\xi' / (2(1 - r(x', 0, \xi'))^{\frac{1}{2}})$, so that

$$d\nu_1 = \frac{1}{\mu_1(S^*M)} \frac{1}{2\sqrt{1 - r(x', 0, \xi')^2}} dx' d\xi', \qquad B_1 = \{(x', \xi') : r(x', 0, \xi') \le 1\},$$

where we parametrized $S^*_N M = \{(x', 0, \xi) : \xi_n^2 + r(x', 0, \xi') = 1\}$ by $(x', \xi') \in B_1$.

On the other hand, the measure ν_E given by (1.11) is

$$d\nu_E = \frac{1}{\mu_1(S^*M)} \frac{1}{2\sqrt{1 - r(x', 0, \xi')^2}} dx' d\xi',$$

$$\Sigma_E = \{(x', 0, \sqrt{E}\xi', \sqrt{E}\xi_n) : \xi_n = \pm(1 - r(x', 0, \xi'))^{\frac{1}{2}},\ r(x', 0, \xi') \le 1\}$$

since for f with $\partial_{x_n} f = 1$, $\partial_{x'} f = 0$ on $x_n = 0$, $H_p f = 2\xi_n$.

To pass from the semiclassical result to the special case of the high energy result we put $E_j = h^2 \lambda_j^2$, $h = 1/\lambda$. The difficulty lies in controlling low frequency contributions and estimates (2.6) and (2.7) are crucial for that.

Let $\widehat{A}$ be a classical pseudodifferential operator of order 0 on N, with a compactly supported Schwartz kernel in N. (Henceforth operators with hats denote polyhomogeneous operators, while operators without hats denote semiclassical operators.) Its principal symbol $\sigma(\widehat{A})$ is a homogeneous function of degree 0 on T^*N. We define $A_\epsilon \in \Psi^0(N)$ by putting

$$A_\epsilon := \mathrm{Op}_h\big(\sigma(\widehat{A})(1-\chi(|\xi'|_{g|_N}/\epsilon))\big), \quad \chi \in C_{\mathrm{c}}^\infty(\mathbb{R}), \ \chi(t)=1, \quad |t|\le 1.$$

Theorem 2 shows that for $0<a<b$, and $v_j = u_j|_N$,

$$h^n \sum_{h\lambda_j\in[a,b]} \left| \langle A_\epsilon v_j, v_j\rangle_{L^2(N, d\,\mathrm{vol}_{g|_N})} - \int_{\Sigma_{E_j}} \pi_{E_j}^* \sigma(A_\epsilon) d\nu_{E_j} \right| \longrightarrow 0, \quad h = 1/\lambda \to 0. \tag{5.1}$$

We also have

$$\int_{\Sigma_{E_j}} \pi_{E_j}^*\sigma(A_\epsilon) d\nu_{E_j} = \int_{S_N^*M} \pi_1^*\sigma(\widehat{A}) d\nu_1 + \mathcal{O}(\epsilon),$$

and hence the result will follow (by taking $[a,b]=[1,2]$ and using a dyadic sum in λ) once we show that

$$\lambda^{-n} \sum_{h\lambda_j\in[a,b]} \left| \langle (\widehat{A}-A_\epsilon) v_j, v_j\rangle_{L^2(N, d\,\mathrm{vol}_{g|_N})} \right| = \mathcal{O}(\epsilon) + \mathcal{O}_\epsilon(h). \tag{5.2}$$

Using (2.7) this will follow from

$$\lim_{\epsilon\to 0} \limsup_{h\to 0} h^n \sum_{h\lambda_j\in[a,b]} \|(\widehat{A}-A_\epsilon) v_j(h)\|^2 = 0. \tag{5.3}$$

For this, we first claim that for any vector field X on N,

$$\|(\widehat{A}-A_\epsilon) hX\|_{L^2\to L^2} \le C\epsilon + \mathcal{O}_\epsilon(h). \tag{5.4}$$

Indeed, the left-hand side of (5.4) is $\mathcal{O}(h)$ if we put an operator in the class Ψ^{-1}_{phg} or $h\Psi^{-1}$ in place of $\widehat{A}-A_\epsilon$, which means that we can reduce to local coordinates, in which we can assume $X = \partial_{y_1}$ and the full symbol of $(\widehat{A}-A_\epsilon)hX$ in the non-semiclassical left quantization becomes, up to $\Psi^{-1}_{\mathrm{phg}} + h\Psi^{-1}$ terms,

$$r(y,\eta;h) := h\eta_1 a^0(y,\eta/|\eta|)\big((1-\chi(|\eta|)) - (1-\chi(h|\eta|_{g|_N}/\epsilon))\big).$$

However, $\partial_y^\alpha\partial_\eta^\beta r(y,\eta;h) = \mathcal{O}(\epsilon+h)\langle\eta\rangle^{-|\beta|}$ (here the first cutoff gives the $\mathcal{O}(h)$ term, while the second cutoff gives the $\mathcal{O}(\epsilon)$ term); therefore, by the L^2 boundedness of classical pseudodifferential operators, we get (5.4).

Now, let $B_0 \in \Psi^{\mathrm{comp}}(M)$ be a semiclassical pseudodifferential operator equal to the identity microlocally near the zero section of T^*M, but supported inside an $\epsilon^{1/2}$ sized neighborhood of the zero section. Then we can write

$$1 - B_0 = \sum_k (hX_k) B_0^k + \mathcal{O}_\epsilon(h)_{L^2\to L^2}$$

for some vector fields X_k (independent of ϵ) and some $B_0^k \in \Psi^{\mathrm{comp}}(M)$ (with $L^2 \to L^2$ norm $\mathcal{O}(\epsilon^{-1/2})$); by (5.4), we have

$$\|(\widehat{A} - A_\epsilon)(1 - B_0)\|_{L^2\to L^2} \leq C\epsilon^{1/2} + \mathcal{O}_\epsilon(h)$$

and thus by (2.7), the estimate (5.3) holds for $(\widehat{A} - A_\epsilon)(1 - B_0)$. Same estimate holds for $(\widehat{A} - A_\epsilon)B_0$, by recalling that $\|\widehat{A} - A_\epsilon\|_{L^2\to L^2} = \mathcal{O}(1)$ and using (2.6) together with the bound $\|\sigma(B_0)\|_{L^2} \leq C\epsilon^{n/4}$. This finishes the proof of (5.3) and thus of Theorem 1.

References

[1] S. Dyatlov, *Asymptotic distribution of quasi-normal modes for Kerr–de Sitter black holes,* to appear in Ann. Henri Poincaré, arXiv:1101.1260.

[2] S. Dyatlov and C. Guillarmou, *Microlocal limits of plane waves and Eisenstein functions,* in preparation.

[3] B. Helffer, A. Martinez, and D. Robert, *Ergodicité en limite semi-classique,* Comm. Math. Phys. **109**(1987), no. 2, 313–326.

[4] L. Hörmander, *The Analysis of Linear Partial Differential Operators III. Pseudo-differential Operators,* Springer, 1985.

[5] J. Sjöstrand and M. Zworski, *Quantum monodromy and semiclassical trace formulae* , J. Math. Pure Appl. **81**(2002), no. 1, 1–33.

[6] J.A. Toth and S. Zelditch, *Quantum ergodic restriction theorems, II: manifolds without boundary,* preprint, arXiv:1104.4531.

[7] M. Zworski, *Semiclassical analysis,* to appear in Graduate Studies in Mathematics, AMS, 2012, `http://math.berkeley.edu/~zworski/semiclassical.pdf`.

E-mail address: `dyatlov@math.berkeley.edu`

E-mail address: `zworski@math.berkeley.edu`

Department of Mathematics, Evans Hall, University of California, Berkeley, CA 94720, USA