Small scale quantum ergodicity in negatively curved manifolds

Xiaolong Han

Department of Mathematics, The Australian National University, Canberra, ACT 2601, Australia
E-mail: Xiaolong.Han@anu.edu.au

Received 18 January 2015, revised 8 July 2015
Accepted for publication 22 July 2015
Published 13 August 2015

Abstract
In this paper, we investigate quantum ergodicity in negatively curved manifolds. We consider the symbols depending on a semiclassical parameter $h$ with support shrinking down to a point as $h \to 0$. The rate of shrinking is a power of $\log h$. This extends the asymptotic equidistribution of quantum ergodic eigenfunctions to a logarithmical scale.

Keywords: quantum ergodicity in small scales, asymptotic equidistribution of eigenfunctions, negatively curved manifolds, exponential decay of correlation
Mathematics Subject Classification: 58G25, 35P20, 37D20, 58G15

1. Introduction
Quantum ergodicity studies the quantized counterpart of a classical dynamical system that is ergodic. In this paper, we consider the geodesic flow in a manifold. Let $(\mathcal{M}, g)$ be a compact and smooth Riemannian manifold of dimension $n$ without boundary. Denote a state as $(x, \xi)$ in the cotangent bundle $T^*\mathcal{M}$. Let $H(x, \xi) = |\xi|^2_{g}$ be the Hamiltonian, where $|\cdot|$ is the induced metric in $T^*_x\mathcal{M}$. Then the geodesic flow $G_t$ is generated by $X_H$, the Hamiltonian vector field of $H$. It preserves the canonical symplectic form $\omega = d\xi \wedge dx$, hence preserves the Liouville volume form $\mu = e^{\omega/n!}$.

Let $S^*\mathcal{M} = \{(x, \xi) \in T^*\mathcal{M} : |\xi|_{g} = 1\}$, i.e. the energy layer at $H(x, \xi) = 1$. Then $G_t$ preserves $S^*\mathcal{M}$, and induces an invariant Liouville measure $\mu_t$ on $S^*\mathcal{M}$. We say that $G_t$ is ergodic on $S^*\mathcal{M}$ with respect to $\mu_t$ if any invariant subset of $S^*\mathcal{M}$ under $G_t$ has $\mu_t$-measure 0 or $\mu_t(S^*\mathcal{M})$. We normalize the Liouville measure and let $\mu_t = \mu_t/\mu_t(S^*\mathcal{M})$. 
The quantum system of \((S^*M, G_t)\) involves the eigenfunctions of the (positive) Laplace-Beltrami operator \(\Delta = \Delta_c\); write \(\{u_j\}_{j=0}^{\infty}\) as an orthonormal basis of eigenfunctions (ONBE) of \(\Delta\) with eigenvalues \(\lambda_j^2\), i.e. \(\Delta u_j = \lambda_j^2 u_j\).

The quantum ergodic theorem of Šnirel’–Zelditch–Colin de Verdière [Sn, Ze1, CdV] illustrates the correspondence between the classical system (geodesic flow on \(S^*M\)) and the quantum system (ONBE) when \(G_t\) is ergodic on \(S^*M\) with respect to the Liouville measure \(\mu_t\).

**Theorem 1.1 (Quantum ergodicity, high-energy version).** Assume that \(G_t\) is ergodic on \(S^*M\) with respect to the Liouville measure \(\mu_t\). Then

\[
S_\ell(\lambda, a) = \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} \left| \langle \text{Op}(a) u_j, u_j \rangle - \mu_t(a) \right| = o(1) \quad \text{as } \lambda \to \infty,
\]

for any orthonormal basis of eigenfunctions \(\{u_j\}_{j=0}^{\infty}\).

We call it the high-energy version since it is on the high-energy limit \(\lambda^2 \to \infty\). Here,

- \(N(\lambda) = \#\{\lambda_j \leq \lambda\}\) is the eigenfrequency counting function; Weyl’s law yields that
  \[
  N(\lambda) = \frac{\mu(B^*M)}{(2\pi)^n} \lambda^n + O(\lambda^{n-1}),
  \]
  where \(B^*M = \{ (x, \xi) \in T^*M : |\xi|_g \leq 1 \}\) is the coball bundle of \(M\). See e.g. Hörmander [Ho].
- \(\langle \cdot, \cdot \rangle\) is the inner product in \(L^2(M)\) with respect to the Riemannian volume \(\text{Vol} = \text{Vol}_g\), induced by the metric \(g\):
  \[
  \langle f, g \rangle = \int_M f \bar{g} \, d\text{Vol} \quad \text{for } f, g \in L^2(M).
  \]
- \(\text{Op}(a)\) is a classical pseudodifferential operator of order 0 with principal symbol \(a\).
- \(\mu_t(a)\) is the space-average of \(a\) with respect to the Liouville measure \(\mu_t\) on \(S^*M\):
  \[
  \mu_t(a) = \int_{S^*M} a \, d\mu_t.
  \]

If all the sectional curvatures of \(M\) are negative everywhere, then we say that \(M\) is negatively curved. In a negatively curved manifold, it is well known that the geodesic flow \(G_t\) is ergodic on \(S^*M\) with respect to the Liouville measure. (See e.g. Katok and Hasselblatt [KH].) Moreover, \(G_t\) is an Anosov flow and displays stronger properties than ergodicity, e.g. central limiting, strong-mixing, and exponentially decay of correlations, etc. (See section 3 for more discussion.) In this case, one expects a more qualitative version of theorem 1.1. Define the \(p\)-moment

\[
S_p(\lambda, a) = \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} \left| \langle \text{Op}(a) u_j, u_j \rangle - \mu_t(a) \right|^p \quad \text{for } p > 0.
\]

Zelditch [Ze2] explored Ratner [R]’s central limit theorem of \(G_t\) to improve theorem 1.1 in negatively curved manifolds to

\[
S_p(\lambda, a) = O\left( (\log \lambda)^{-\frac{p}{2}} \right).
\]

See also Schubert [Sc] in a setting with more general Hamiltonian. Sarnak [Sa1] conjectured that on compact hyperbolic surfaces,
for any small $\epsilon > 0$. In fact, his conjecture is on the individual terms:

$$\left| \langle \text{Op}(a)u_j, u_j \rangle - \mu_1(a) \right| = O\left( \lambda_j^{-\frac{1}{4} + \epsilon} \right),$$

which would imply that

$$\langle \text{Op}(a)u_j, u_j \rangle \to \mu_1(a) \quad \text{as} \quad j \to \infty.$$

Quantum unique ergodicity (QUE) conjecture states that (1.3) is valid on any compact manifold of (possibly variable) negative curvature, see Rudnick and Sarnak [RS]. The statement is on the asymptotic behavior of the whole ONBE $\{u_j\}_{j=0}^{\infty}$. However, it is very difficult to prove such result. Restricting to certain orthonormal bases of eigenfunctions, QUE has been verified in special cases when $(\mathcal{M}, g)$ is arithmetic, by Lindenstrauss [Lin], Silbermann and Venkatesh [SV], Holowinsky and Soundararajan [HS], and Brooks and Lindenstrauss [BL].

In the case when $G_t$ is only ergodic, theorem 1.1 implies that (1.3) holds for ‘almost all’ (that is, full density as defined in the following) eigenfunctions in any ONBE. We define the density $D$ of a subsequence $J = \{j_k\} \subset \mathbb{N}$ as

$$D(J) = \lim_{N \to \infty} \frac{\# \{j_k < N\}}{N} \quad \text{if it exists.}$$

When $D = 0$ ($>0$ or $= 1$), we call such subsequence a zero (positive or full) density subsequence. The full density subsequence is also called density one subsequence. It follows from a very standard argument to extract a full density subsequence of $\{u_j\}$ for which (1.3) is true, using $S_j(\lambda, a) = o(1)$ for some $p > 0$. See e.g. [Ze1, lemma 3].

**Corollary 1.2.** Assume that $G_t$ is ergodic on $S^*\mathcal{M}$ with respect to the Liouville measure $\mathbf{\mu}_t$. Given any ONBE $\{u_j\}_{j=0}^{\infty}$, there exists a full density subsequence of eigenfunctions $\{u_{j_k}\} \subset \{u_j\}$ such that

$$\langle \text{Op}(a)u_{j_k}, u_{j_k} \rangle \to \mu_1(a) \quad \text{as} \quad k \to \infty,$$

for each classical pseudodifferential operator $\text{Op}(a)$ of order 0. Then $\{u_{j_k}\}$ is called a quantum ergodic subsequence of eigenfunctions.

This corollary hence states that the semiclassical measure defined by the full density subsequence $\{u_{j_k}\}$ in $T^*\mathcal{M}$ coincides with the Liouville measure $\mathbf{\mu}_t$ on $S^*\mathcal{M}$. It, however, does not exclude the possibility that there might exist a sparse subsequence of $\{u_j\}$ for which (1.3) is not valid. In fact, Hassell [Ha] provides an example that supports such a subsequence, that is, the generic Bunimovich stadia. This also verifies that classical ergodicity of $G_t$ cannot imply QUE.

Let $\Omega \subset \mathcal{M}$ be a Borel subset with measure-zero boundary. Then by Portmanteau theorem (see [So1, theorem 6.2.5]), we have

$$\int_{\Omega} |u_{j_k}|^2 \, d\text{Vol} \to \frac{\text{Vol}(\Omega)}{\text{Vol}(\mathcal{M})} \quad \text{as} \quad k \to \infty.$$

Therefore, for this quantum ergodic subsequence of eigenfunctions, the $L^2$ mass of $\{u_{j_k}\}$ display equidistribution in $\mathcal{M}$ as $k \to \infty$. (Theorem 1.1 actually concludes asymptotic equidistribution on $S^*\mathcal{M}$.) Here, the set $\Omega$ is fixed and we say that it is of scale $O(1)$. In the
classical-quantum correspondence, an eigenfunction $u_j$ represents a stable state of a freely moving particle in $\mathbb{M}$, and the $L^2$ mass distribution of $u_j$ is interpreted as the probability density of finding a particle in $\mathbb{M}$. Therefore, (1.4) indicates that for almost all stable states, the probability density tends to the normalized Riemannian volume $d\text{Vol} / \text{Vol}(\mathbb{M})$.

The $L^2$ distribution characterization of eigenfunctions has a lot of applications beyond its connection with quantum physics. In some applications, one is interested in the mass distribution in regions of smaller scales. For example, let $B(x, r)$ be the geodesic ball centered at $x \in \mathbb{M}$ with radius $r$. Consider the eigenfunction $u_j$ in $B(x, \lambda_j^{-1})$. One can rescale $B(x, \lambda_j^{-1})$ to the ball with radius 1. Then the operator $\Delta - \lambda_j^2$ is a small perturbation of the Euclidean Laplacian in the unit ball. Hence, the classical elliptic PDE theory applies, e.g. the mean value inequality (see Li and Schoen [LS]):

$$\|u_j\|_{L^2(B(x, r/2)))} \leq C \lambda_j^{k/2} \int_{B(x, r)} |u_j|^2 d\text{Vol}.$$  

So the $L^2$ mass of $u_j$ in $B(x, \lambda_j)$ controls its supreme value in the half ball. This approach has been used in the nodal set estimates of eigenfunctions, see e.g. Donnelly and Fefferman [DF], Colding and Minicozzi [CM]. However, in such a scale $\lambda_j^{-1}$ there is no fast oscillation, therefore generally no results on mass equidistribution can be expected. Hence, we ask the following question.

**Question 1.3 (Small scale mass equidistribution).** Let $\rho \in (0, 1)$. Given any ONBE, does there exist a full density subsequence $\{u_{\rho k}\}$ such that

$$\int_{B(x, r_\rho)} |u_{\rho k}|^2 d\text{Vol} = \frac{\text{Vol}(B(x, r_\rho))}{\text{Vol}(\mathbb{M})} + o(r_\rho^2) \quad \text{as } k \to \infty \tag{1.5}$$

for $r_\rho = r(\lambda_k) = \lambda_k^{-\rho}$ and all $x \in \mathbb{M}$?

This statement, if true, asserts that the full density subsequence $\{u_{\rho k}\}$ is $L^2$ mass equidistributed in the scale $\lambda_k^{-\rho}$. For example, on the torus $\mathbb{T}^n$, the whole ONBE $\{e^{i n \cdot x}\}$ is equidistributed in any small scale; for small scale mass equidistribution of an arbitrary ONBE on $\mathbb{T}^n$, see Hezari and Riviére [HR2] and the reference therein.

In negatively curved manifolds, we prove small scale mass equidistribution when the scale $r(\lambda_k) = (\log \lambda_k)^{-\alpha}$ for some $\alpha > 0$: Given any ONBE, for a fixed point $x_0 \in \mathbb{M}$, we show that (1.5) is valid for a full density subsequence (depending on $x_0$) in the scale $(\log \lambda_k)^{-\alpha}$ for $\alpha < 1/(2n)$ (see corollary 1.8); for all the points on $\mathbb{M}$, we show that the two sides of (1.5) are uniformly comparable for a full density subsequence in the scale $(\log \lambda_k)^{-\alpha}$ for $\alpha < 1/(3n)$ (see corollary 1.9).

**Remark.** (1) In negatively curved manifolds at a given point $x$, (1.5) is automatically true if $\rho = 0$ by (1.4). Here, we assume that the injectivity radius of $\mathbb{M}$ is much greater than 1 (and throughout the whole paper).

(2) We focus on the negatively curved manifolds since more results on the dynamical properties of the geodesic flows are available in this case, particularly, Liverani [Liv]'s exponential decay of correlation. It is doubtful that (1.4) can be improved to smaller scales in a manifold with only ergodicity assumption on the geodesic flow.

(3) It is a natural question to ask whether the convergence in (1.5) is uniform for $x \in \mathbb{M}$. Such uniform convergence is crucial for its application, because one can use equidistribution to study the eigenfunctions locally in each small region and then combine them to a global estimate in the whole manifold by uniformity.
It is convenient to work in the semiclassical setting. In this setting, the quantum ergodicity theorem is proved by Helffer, Martinez, and Robert [HMR]. Here, we state the result from [DyGu, theorem 5].

**Theorem 1.4 (Quantum ergodicity, a semiclassical version).** Assume that $G$ is ergodic on $S^*\mathcal{M}$ with respect to the Liouville measure $\mu$. Let $\{u_j\}_{j=0}^\infty$ be an ONBE of the quantized Laplacian, $h^2\Delta$, that is, $h^2\Delta u_j = E_j \mu_j$. Then

$$V_1(h, a) = h^{n-1} \sum_{E_j \in [1.1+h]} \left| \left\{ \mathcal{O}_p(a) u_j, u_j \right\} - \mu_j(a) \right| = o(1) \quad \text{as } h \to 0,$$

where $\mathcal{O}_p(a)$ is the semiclassical pseudodifferential operator with principal symbol $a \in S(\mathcal{M})$ (see its definition in section 2).

One can choose $\lambda^2 = h^3 E_j$ to recover the high-energy version quantum ergodicity in theorem 1.1 and corollary 1.2. With this connection between the two versions of quantum ergodicity, we see that question 1.3 actually involves choosing $\lambda$. But the smoothened version of such a function $a$ would belong in nice symbol class only if $\rho \in [0, 1/2]$, that is, $S_\rho(\mathcal{M})$ (see section 2). So answering question 1.3 for $\rho \geq 1/2$ would be very challenging. (It is in fact already challenging for any polynomial scale $h^p$ for $p > 0$ as our results are all in the logarithmic scales.)

In negatively curved manifolds, we in fact are able to prove quantum ergodicity for larger classes of symbols than $S(\mathcal{M})$. To this end, we have the following definition.

**Definition (p-admissibility).** Let $\rho \in [0, 1/2]$ and $0 < h_0 \ll 1$. We say that $\delta(h) \geq 0$ is $\rho$-admissible if $\delta(h)$ satisfies that $h^\rho \leq \delta(h) \leq 1$ for all $h \in (0, h_0]$.

**Example.** $\delta(h) = [\log h]^{-\alpha}$ for $\alpha > 0$ is $\rho$-admissible for all $\rho \in (0, 1/2)$.

We then define the symbol class $S_{\delta(h)}(\mathcal{M})$ as follows.

**Definition ($S_{\delta(h)}(\mathcal{M})$ symbol classes).** Let $\delta(h)$ be $\rho$-admissible for some $\rho \in [0, 1/2]$. We say that $a \in C^\infty(T^*\mathcal{M})$ belongs to $S_{\delta(h)}(\mathcal{M})$ if and only if for each multi-indices $\alpha, \beta$, there exists a uniform constant $C_{\alpha, \beta}$ independent of $h$ such that

$$\sup_{x \in \mathcal{M}, \xi \in T^*_x \mathcal{M}} |d^\alpha \xi^\beta | \leq C_{\alpha, \beta} \delta(h)^{-|\alpha| - |\beta|} (\xi^{-\beta})^{-|\beta|},$$

where $(\xi) = (1 + |\xi|^2)^{1/2}$.

Our main theorem, small scale quantum ergodicity theorem, states that theorem 1.4 can be extended to symbols in $S_{\delta(h)}(\mathcal{M})$, given that $\delta(h) \to 0$ in a certain logarithmical rate as $h \to 0$.

**Theorem 1.5 (Small scale quantum ergodicity).** Let $(\mathcal{M}, g)$ be negatively curved. For $a > 0$ and $0 \leq \beta < 1$, or $\alpha = 0$ and $\beta = 1$,

suppose that $\delta(h) = [\log h]^{-\alpha}$. Then for any ONBE $\{u_j\}_{j=0}^\infty$ and all $a \in S_{\delta(h)}(\mathcal{M})$,

$$V_2(h, a) = h^{n-1} \sum_{E_j \in [1.1+h]} \left| \left\{ \mathcal{O}_p(a) u_j, u_j \right\} - \mu_j(a) \right|^2 = O((\log h)^{-\beta}) \quad \text{as } h \to 0.$$

(1.7)

**Remark.** Notice that when $\alpha = 0$, we have $\delta(h) = 1$ and $S_{\delta(h)}(\mathcal{M}) = S(\mathcal{M})$. Then theorem 1.5 recovers (1.2) from [Ze2].
We next consider some special symbols in $S_{\delta(h)}(\mathbb{M})$ that are used to derive small scale mass equidistribution.

**Definition (δ-microlocalized symbols).** Let $\delta(h)$ be $\rho$-admissible for some $\rho \in (0, 1/2)$. Let $b \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$. In a local chart that contains $z_0 = (x_0, \xi_0) \in S^*\mathbb{M}$, write the local coordinate of $z = (x, \xi) \in T^*\mathbb{M}$ as $(\tilde{x}, \tilde{\xi}) \in \mathbb{R}^n$. We say that a compactly supported smooth function $a^b_{\xi}(x, \xi; h)$ is a $\delta$-microlocalized symbol if it is locally defined by

$$a^b_{\xi}(x, \xi; h) = b\left(\frac{\tilde{x} - \tilde{x}_0}{\delta(h)}, \frac{\tilde{\xi} - \tilde{\xi}_0}{\delta(h)}\right)\phi(|\tilde{\xi}|_h), \quad (1.8)$$

where $\eta = \eta/|\eta|_{10}$, and $\phi \in C_0^\infty((-1/2, 1/2))$ has small compact support and equals 1 on $(-1/4, 1/4)$.

**Remark.** A $\delta$-microlocalized symbol is in the symbol class $S_{\delta(h)}(\mathbb{M})$. Since the base function $b$ is in $C_0^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$, the support of $a^b_{\xi}$ restricted on $S^*\mathbb{M}$ shrinks to $z_0 \in S^*\mathbb{M}$ at the same rate $\delta(h)$ in every direction; the cutoff function $\phi$ ensures that $a^b_{\xi}$ is extended to a well-defined symbol in a neighborhood of $S^*\mathbb{M}$ in $T^*\mathbb{M}$. We are free to choose the base function $b$ in building $a^b_{\xi}(x, \xi; h)$.

Considering $\delta$-microlocalized symbols, small scale quantum ergodicity theorem states

**Theorem 1.6.** Let $(\mathbb{M}, g)$ be negatively curved. For

$$0 < \alpha < \frac{1}{2(2n - 1)} \text{ and } 0 < \beta < 1 - 2\alpha(2n - 1), \text{ or } \alpha = 0 \text{ and } \beta = 1,$$

suppose that $\delta(h) = |\log h|^{-\alpha}$. Then for any ONBE $(u_j)^{\infty}_{j=0}$ as $h \to 0$,

$$V_2(h, a^b_{\xi}) = h^{n-1} \sum_{E \in [1,1+h]} \left| \text{Op}_h(a^b_{\xi}u_j, u_j) - \mu_i(a^b_{\xi}) \right|^2 = C \delta(h)^{2(2n-1)}|\log h|^{-\beta}, \quad (1.9)$$

uniformly for $z_0 \in S^*\mathbb{M}$, where $a^b_{\xi}$ is defined in (1.8) and $C$ depends on $b$, $\phi$, and $\mathbb{M}$.

**Remark.** We have the term $\delta(h)^{2(2n-1)}$ in the right-hand side of (1.9) because $\mu_i(a^b_{\xi}) = O(\delta(h)^{2n-1})$ in the left-hand side.

To consider the mass equidistribution of eigenfunctions in small scales of $\mathbb{M}$ (instead of on $S^*\mathbb{M}$), we need the $\delta$-localized symbols.

**Definition (δ-localized symbols).** Let $\delta(h)$ be $\rho$-admissible for some $\rho \in (0, 1/2)$. Let $b \in C_0^\infty(\mathbb{R}^n)$. In a local chart that contains $z_0 = (x_0, \xi_0) \in S^*\mathbb{M}$, write the local coordinate of $z = (x, \xi) \in T^*\mathbb{M}$ as $(\tilde{x}, \tilde{\xi}) \in \mathbb{R}^n$. We say that a compactly supported smooth function $a^b_\xi(x, \xi; h)$ is a $\delta$-localized symbol if it is locally defined by

$$a^b_\xi(x, \xi; h) = b\left(\frac{\tilde{x} - \tilde{x}_0}{\delta(h)}\right)\phi(|\tilde{\xi}|_h), \quad (1.10)$$

Similar to theorem 1.6, we have the following theorem, with the shrinking rate of the symbols slightly better than the one in theorem 1.5.
Theorem 1.7. Let $(\mathcal{M}, g)$ be negatively curved. For

$$0 < \alpha < \frac{1}{2n} \quad \text{and} \quad 0 \leq \beta < 1 - 2\alpha,$$

suppose that \(\delta(h) = |\log h|^{-\alpha}\). Then for any ONBE \(\{u_j\}_{j=0}^\infty\) as \(h \to 0\),

\[
V_2(h,a_j^b) = h^{n-1} \sum_{E_j \in [1,1+h]} \left| \left\langle \text{Op}_h(a_j^b)u_j, u_j \right\rangle - \mu_j(a_j^b) \right|^2 = C \delta(h)^{2\alpha} |\log h|^{-\beta},
\]

(1.11)

uniformly for \(x_0 \in \delta^n \mathcal{M}\), where \(a_j^b\) is defined in (1.10) and \(C\) depends on \(b\), \(\phi\), and \(\mathcal{M}\).

Remark. In a recent preprint [HR1], Hezari and Rivière proved theorem 1.7 for \(V_p(h,a_j^b)\) of \(p \geq 2\) with the same \(\delta(h)\) when \(b\) is a cutoff function. They then used it to study \(L^p\) norm and nodal set estimates of the full density quantum ergodic subsequence of eigenfunctions. We refer to their paper for details; see also Sogge [So2].

Using theorem 1.7, we answer question 1.3 for any fixed point in the manifold.

Corollary 1.8. Let \((\mathcal{M}, g)\) be negatively curved and \(x_0 \in \mathcal{M}\). Assume that

$$0 \leq \alpha < \frac{1}{2n} \quad \text{and} \quad r(\lambda) = (\log \lambda)^{-\alpha}.$$

Given any ONBE \(\{u_j\}_{j=0}^\infty\) with \(\Delta u_j = \lambda_j^2 u_j\), there exists a full density subsequence \(\{u_{j_k}\}\) such that

\[
\int_{B(x_0,r_{j_k})} |u_{j_k}|^2 \, d\text{Vol} = \frac{\text{Vol}(B(x_0,r_{j_k}))}{\text{Vol}(\mathcal{M})} + o(r_{j_k}^n) \quad \text{as} \quad k \to \infty
\]

(1.12)

for \(r_{j_k} = r(\lambda_{j_k}) = (\log \lambda_{j_k})^{-\alpha}\).

Notice that in the above corollary, the subsequence \(\{u_{j_k}\}\) depends on the point \(x_0\). To get a uniform result, we use a covering argument to get a weaker result as uniform compatibility. In the process, we concede certain loss on the shrinking rate.

Corollary 1.9. Let \((\mathcal{M}, g)\) be negatively curved. Assume that

$$0 \leq \alpha < \frac{1}{3n} \quad \text{and} \quad r(\lambda) = (\log \lambda)^{-\alpha}.$$

Given any ONBE \(\{u_j\}_{j=0}^\infty\) with \(\Delta u_j = \lambda_j^2 u_j\), there exists a full density subsequence \(\{u_{j_k}\}\) such that

\[
c \text{Vol}(B(x,r_{j_k})) \leq \int_{B(x,r_{j_k})} |u_{j_k}|^2 \, d\text{Vol} \leq C \text{Vol}(B(x,r_{j_k})) \quad \text{as} \quad k \to \infty,
\]

(1.13)

uniformly for all \(x \in \mathcal{M}\), where the positive constants \(c\) and \(C\) depends only on \(\mathcal{M}\).

1.1. Outline and organization

We prove small scale quantum ergodicity in the same spirit as in theorems 1.1 and 1.4. The key ingredients are semiclassical trace formula and Egorov’s theorem. The former requires to
treat symbols in a more general class other than $C^\infty$ ones independent of $h$. The latter, connecting time evolution of classical observables (the symbol $a$ in $T^*M$) and quantum observables (semiclassical pseudodifferential operator $Op_h(a)$ on $L^2(M)$), also concerns symbols depending on $h$. Such correspondence is valid up to the Ehrenfest time $T_E \approx [\log h]$ (Bouzouina and Robert [BR]). In the process, we trace all the dependence on $h$.

Thus, small scale quantum ergodicity is reduced to estimating the time-average of the quantum observable, which is controlled by the time-average of its principal symbol. Liverani [Liv]’s exponential decay of correlations then can be used to give a quantitative estimate on the time-average of the symbol in terms of its Hölder norm; indeed, the time-average has polynomial decay in time. By properly choosing the $\delta(h)$ and thus the symbol class $S^\delta_h(M)$, we can prove theorem 1.5. Since Ehrenfest time $T_E$ is of order $[\log h]$, we can only select such symbols with Hölder norm of order $[\log h]^{-\alpha}$ for some $\alpha > 0$. This is essentially the reason why we have logarithmically shrinking rates.

We organize this paper as follows. In sections 2 and 3, we review semiclassical analysis and geodesic flows in negatively curved manifolds, respectively. In section 4, we prove small scale quantum ergodicity theorems, i.e. theorems 1.5–1.7. In section 5, we prove small scale mass equidistribution in corollaries 1.8 and 1.9; we also point out some further investigation on small scale quantum ergodicity, particularly on question 1.3.

1.2. A note on the previous works
Besides the above mentioned previous results on quantum ergodicity, there are other exciting development in this area recently. We refer to Zelditch [Ze3] and Sarnak [Sa2] for its current stage.

2. Semiclassical analysis
In this section, we review semiclassical analysis that will be used to prove small scale quantum ergodicity. Most of the notations and facts below are fairly standard. We refer to Zworski [Zw] for a complete treatment in this subject.

2.1. Phase space
$M$ can be a bounded open set in $\mathbb{R}^n$ or a compact Riemannian manifold. An element, called a state, in the cotangent bundle $T^*M$ is denoted as $z = (x, \xi)$ with $x \in M$ and $\xi \in T^*_xM$.

2.2. Symbol classes
Let $m \in \mathbb{R}$ and $\rho \in [0, 1/2)$. The symbol class $S^m_\rho(M)$ is defined as follows: $a \in C^\infty(T^*M)$ belongs to $S^m_\rho(M)$ if and only if for each multi-indices $\alpha, \beta$, there exists a uniform constant $C_{\alpha, \beta}$ independent of $h$ such that

$$\sup_{x \in M, \xi \in T^*_xM} |\partial^{\alpha}_x \partial^{\beta}_\xi a| \leq C_{\alpha, \beta} h^{-\rho(|\alpha|+|\beta|)}(|\xi|^{m-|\beta|)}.$$ 

where $(\xi) = (1 + |\xi|^2)^{1/2}$.

1. If $\rho = 0$, we denote $S^m_0(M)$ by $S^m(M)$.
2. If $m = 0$, we denote $S^0_\rho(M)$ by $S_\rho(M)$.
(3) If \(a\) has compact support which satisfies the above estimate, we say that \(a \in \mathcal{S}_h^{\text{comp}}(\mathbb{M})\).

Clearly, \(\mathcal{S}_h^{\text{comp}}(\mathbb{M}) \subset \mathcal{S}_h^m(\mathbb{M})\) for all \(m \in \mathbb{R}\).

(4) We denote \(\mathcal{S}_h^{\text{ad}}(\mathbb{M}) = \cap_{m \in \mathbb{R}} \mathcal{S}_h^m(\mathbb{M})\) and \(\mathcal{S}_h^{\text{op}}(\mathbb{M}) = \cup_{m \in \mathbb{R}} \mathcal{S}_h^m(\mathbb{M})\).

These classes are independent of the choice of coordinates in \(\mathbb{M}\). Moreover, the seminorms \(|\cdot|_{m, \beta}\) in \(\mathcal{S}_h(\mathbb{M})\) is defined by the best constant \(C_{m, \beta}\) that can be used in the above inequality.

**Example.** Let \(\delta(h)\) be \(\rho\)-admissible. Then \(\mathcal{S}_{\delta(h)}(\mathbb{M}) \subset \mathcal{S}_h(\mathbb{M})\); moreover, the \(\delta\)-microlocalized symbols and \(\delta\)-localized symbols are in \(\mathcal{S}_{\delta(h)}(\mathbb{M}) \cap \mathcal{S}_h^{\text{comp}}(\mathbb{M})\).

### 2.3. Quantizations in \(\mathbb{R}^n\)

Every classical observable \(a\) in the phase space \(T^*\mathbb{R}^n\) corresponds to a quantum observable \(\mathcal{O}_h(a)\) as a semiclassical pseudodifferential operator acting on \(L^2\) functions in \(\mathbb{R}^n\). In this note, we use Weyl quantization.

**Definition (Weyl quantization).** Given \(a \in \mathcal{S}_h^\rho(\mathbb{R}^n), \rho \in [0, 1/2]\), we define the Weyl quantization

\[
\mathcal{O}_h(a)(x) = \frac{1}{(2\pi\hbar)^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i(x-y) \cdot \eta} \rho \left(\frac{x+y}{2}, \eta; \hbar\right) a(y) \, dy \, dx
\]

for \(u \in \mathcal{S}(\mathbb{R}^n)\).

**Remark.** \(\mathcal{O}_h(a)\) is self-adjoint if \(a\) is a real-valued symbol.

### 2.4. Quantization in \(\mathbb{M}\)

We now define \(\Psi^m_h(\mathbb{M})\) of semiclassical pseudodifferential operators with symbols in \(\mathcal{S}_h^\rho(\mathbb{M}), \rho \in [0, 1/2]\); then we establish the correspondence of \(A \in \Psi^m_h(\mathbb{M})\) and its semiclassical principal symbol \(a\). See Zworski [Zw, section 14.2] for more details, and also Dyatlov and Guillarmou [DyGu, section 3.1] for symbols in \(\mathcal{S}_h^\rho(\mathbb{M})\). The correspondence is one-to-one modulo lower order terms. Denote

\[
a = \sigma(A) : \Psi^m_h(\mathbb{M}) \rightarrow \mathcal{S}_h^\rho(\mathbb{M})/h^{1-2\rho}\mathcal{S}_h^{m-1}(\mathbb{M}),
\]

and its right inverse, a non-canonical quantization map for \(a \in \mathcal{S}_h^m(\mathbb{M})\):

\[
A = \mathcal{O}_h(a) : \mathcal{S}_h^m(\mathbb{M}) \rightarrow \Psi^m_h(\mathbb{M}).
\]

\(\sigma(A)\) is called the principal symbol of \(A\). It is modulo \(h^{1-2\rho}\mathcal{S}_h^{m-1}(\mathbb{M})\) unique under change of quantizations and change of local coordinates. Following the same fashion in section 2.2,

(1) if \(\rho = 0\), we denote \(\Psi^m_h(\mathbb{M})\) by \(\Psi^m(\mathbb{M})\);
(2) if \(m = 0\), we denote \(\Psi_h^m(\mathbb{M})\) by \(\Psi_h(\math{M})\);
(3) if \(a\) has compact support, we say that \(\mathcal{O}_h(a) \in \Psi^\text{comp}_h(\mathbb{M}), \Psi^\text{comp}_h(\mathbb{M}) \subset \Psi^m_h(\mathbb{M})\) for all \(m \in \mathbb{R}\). Moreover, if \(A \in \Psi^\text{comp}_h(\mathbb{M})\), then \(A = \mathcal{O}_h(a)\) for some \(a \in \Psi^\text{comp}_h(\mathbb{M})\);
(4) we denote \(\Psi_h^\text{ad}(\mathbb{M}) = \cap_{m \in \mathbb{R}} \Psi_h^m(\mathbb{M})\) and \(\Psi_h^{\text{op}}(\mathbb{M}) = \cup_{m \in \mathbb{R}} \Psi_h^m(\mathbb{M})\).

The usual operations involving semiclassical pseudodifferential operators are as follows. Let \(A \in \Psi_h^m(\mathbb{M})\) and \(B \in \Psi_h^n(\mathbb{M})\).

(1) Let \(A^*\) be the adjoint operator of \(A\) in \(L^2(\mathbb{M})\). Then
\[
\sigma(A^\ast) = \overline{\sigma(A)} + O_{\mathcal{M}}(h^{1-2\rho}).
\] (2.1)

\[
\sigma(AB) = \sigma(A)\sigma(B) + O_{\mathcal{M}}(h^{1-2\rho}).
\] (2.2)

\[
\sigma([A, B]) = -i\hbar\{\sigma(A), \sigma(B)\} + O_{\mathcal{M}}(h^{2(1-2\rho)}),
\]

\[
\{a, b\} = \frac{\partial a}{\partial x} \frac{\partial b}{\partial \xi} - \frac{\partial a}{\partial \xi} \frac{\partial b}{\partial x}.
\]

We also need the following \(L^2\) boundedness of semiclassical pseudodifferential operators in \(\Psi_M(M)\), see e.g. [Zw, theorem 4.23].

**Theorem 2.1.** \([L^2\text{ boundedness}]\) Let \(a \in S_{h}(M)\) for \(\rho \in [0, 1/2]\). Then

\[
\|\mathcal{O}_{\mathcal{A}}(a)u\|_{L^2(M)} \leq C\|u\|_{L^2(M)},
\]

where \(C\) depends on finite number of seminorms of \(a\), and is independent of \(h\).

### 2.5. Semiclassical measures

For each eigenfunction \(u_j(h)\) of \(h^2\Delta\), we define the distribution associated with \(u_j(h)\) as

\[
\mathcal{W}_j(a) = \langle \mathcal{O}_{\mathcal{A}}(a)u_j(h), u_j(h) \rangle \quad \text{for } a \in S(M).
\]

We see that \(\mathcal{W}_j\) depends on local coordinates, partition of unity, etc. However, as \(h \to 0\), the accumulation points of \(\mathcal{W}_j\) are independent of such choices. We define the semiclassical measures as the limit points of \(\{\mathcal{W}_j\}\). Note that \(h^2\Delta u_j(h) = E_j u_j(h)\), that is,

\[
(h^2\Delta - E_j)u_j(h) = 0.
\]

Since the semiclassical symbol of \(h^2\Delta - E_j\) is \(|\xi|^2 - E_j\), from Zworski [Zw, theorems 5.3 and 5.4], we have

**Theorem 2.2.** Let \(0 < c_1 < c_2 < \infty\). Then any semiclassical measure for \(E_j \in [c_1, c_2]\) is supported in the energy shell

\[
\{(x, \xi) \in T^*M : |\xi|^2 \in [c_1, c_2] \},
\]

and is invariant under the geodesic flow \(G_t\).

In particular, the semiclassical measures corresponding to eigenvalues \(E_j \in [1, 1 + h]\) is supported on the energy layer \(S^*M\). From this point of view, quantum ergodicity studies the impact of geodesic flow \(G_t\) on the semiclassical measures. Theorem 1.4 asserts that if \(G_t\) is ergodic on \(S^*M\), then the difference of \(\mathcal{W}_j\) and the Liouville measure \(\mu_1\) converges to 0 in Cesàro summation for \(E_j \in [1, 1 + h]\).

### 2.6. Egorov’s theorem

The geodesic flow \(G_t : (x(0), \xi(0)) \to (x(t), \xi(t))\) in \(T^*M\) is generated by the Hamilton equation
\[ \frac{dx}{dt} = \frac{\partial H}{\partial \xi}(x, \xi), \quad \frac{d\xi}{dt} = -\frac{\partial H}{\partial x}(x, \xi), \]

in which \( H(x, \xi) = |\xi|^2 \). Then the time evolution of a classical symbol \( a \) satisfies

\[ \frac{d}{dt} a(G_t(x, \xi)) = \{H, a\}(G_t(x, \xi)), \]

where \( \{H, a\} \) is the Poisson bracket. The quantum time evolution of \( \Op_h(a) \) is associated with the unitary Fourier integral operator \( -i \partial_t \mathcal{U}_t \mathcal{E}^{-i (a \star G_t)} \), called the Schrödinger propagator of \( \hbar^2 \Delta \). Egorov’s theorem states that the quantum time evolution of \( \Op_h(a) \) can be approximated by the quantization of classical time evolution of \( a \), within finite time:

\[ \mathcal{U}(-t) \circ \Op_h(a) \circ \mathcal{U}(t) \approx \Op_h(a \star G_t). \]

Such correspondence connects the classical observable \( a \) and the quantum observable \( \Op_h(a) \) under time evolution. Precisely, if \( a \in \mathcal{S}(\mathbb{M}) \) and \( |t| \leq T < \infty \), then

\[ \| \mathcal{U}(-t) \circ \Op_h(a) \circ \mathcal{U}(t) - \Op_h(a \star G_t) \|_{L^2(\mathbb{M})} = O(h) \quad \text{as } h \to 0. \]

2.7 Semiclassical trace formula

Here we use Schubert [Sc, proposition 1] in the present context.

**Proposition 2.3.** Suppose that \( \psi \) is a smooth function on \( \mathbb{R} \) such that the Fourier transform \( \hat{\psi} \) has compact support in a small neighbourhood of 0 which contains no period of a periodic orbit of the geodesic flow \( G_t \) on \( S^*\mathbb{M} \). Then for every \( a \in \mathcal{S}(\mathbb{M}) \), we have

\[ \int_{\mathbb{R}^{n-1}} |\psi(1 - \frac{E_1}{h}) (\Op_h(a)u_1, u_1) - \frac{\hat{\psi}(0)(2\pi)^{-n/2}}{\mu_1(a)}| e^{\pm \Theta(a) v} \leq C h \|a\|_{L^{2n+8}}, \]

where the constant \( C \) depends on \( \psi \) and \( \mathbb{M} \).

3. Geodesic flows in negatively curved manifolds

In this section, we gather some facts on the geodesic flow \( G_t \) in a negatively curved manifold \( (\mathbb{M}, g) \). Recall that \( H(x, \xi) = |\xi|^2 \). \( G_t \) is Anosov on \( S^*\mathbb{M} \), that is, the tangent bundle \( T^*\mathbb{M} \) splits into \( G_t \)-invariant sub-bundles

\[ E^u(v) \oplus E^s(v) \oplus \mathbb{R} X_0(v) \quad \text{for } v \in T^*\mathbb{M}. \]

Here, \( E^u \) and \( E^s \) are the unstable and stable subspaces, respectively. They satisfy

\[ \begin{cases} \|dG_tv\| \leq C e^{-\lambda t} \|v\| & \forall v \in E^u, t \geq 0; \\ \|dG_tv\| \leq C e^{\lambda t} \|v\| & \forall v \in E^s, t \leq 0, \end{cases} \quad \text{(3.1)} \]

where \( G_t : z \to G_t(z), dG_t : T^*\mathbb{M} \to T^*G_t\mathbb{M} \) is the differential, and \( \|\cdot\| \) is the norm defined in \( T^*\mathbb{M} \) (e.g. by the Sasaki metric, see Ballmann [B]). The sub-bundles are integrable and induce stable and unstable foliations. We refer to Katok and Hasselblatt [KH] for background.
3.1. Egorov’s theorem until the Ehrenfest time

In a negatively curved manifold \((M, g)\), we use the above information on its geodesic flow to describe the long-time evolution \(a \ast G_t\) of the classical observable \(a\) and its quantum long-time evolution in the Egorov’s theorem. See Bouzouina and Robert [BR] in a more general setting.

Let \(E = \{ (x, \xi) \in T^\ast M : 1/2 \leq |\xi|, 3/2 \}\) be an energy shell in \(T^\ast M\). So \(E \supset S^\ast M\). Also, \(G_t\) preserves \(E\). From Anantharaman and Nonnenmacher [AN, section 5.2] that

\[
\sup_{z \in E} |a_t^G(z)| < C_G e^{l|t|} \quad \text{for } t \in \mathbb{R},
\]

where \(l\) can be chosen as a number that is greater than the maximal expansion rate of \(G_t\) in \(E\).

Now consider a symbol \(a \in \mathcal{S}_h^{\text{comp}}(M)\), \(\rho \in [0, 1/2)\), compactly supported in \(E\). Clearly \(a \ast G_t\) is still compactly supported in \(E\). Moreover, we have

\[
\sup_{z \in E} |a_t^G(z)| < C_{a, \rho} e^{(\rho + \varepsilon)|t| \log h} \quad \text{for } t \in \mathbb{R}.
\]

If \(\varepsilon \in [0, 1/2 - \rho)\), setting

\[
T_E = \frac{\log h}{l},
\]

then

\[
\sup_{|t| < \varepsilon T_E} |a_t^G(z)| < C_{a, \rho} e^{-\varepsilon T_E} \quad \text{for } |t| < \varepsilon T_E.
\]

This means that \(a \ast G_t \in \mathcal{S}_h^{\text{comp}}(M)\) if \(|t| < \varepsilon T_E\). Since \(\rho + \varepsilon \in [0, 1/2)\), \(a \ast G_t\) is still in a nice symbol class. We call \(T_E\) the Ehrenfest time. Then Egorov’s theorem in section 2.6, connecting the time evolution of a classical observable \(a\) and its quantum counterpart \(\mathcal{O}_h(a)\), can be extended to more general symbols until Ehrenfest time. See Anantharaman [Ana, theorem 4.2.4].

**Theorem 3.1 (Egorov’s theorem until the Ehrenfest time).** Let \(a \in \mathcal{S}_h^{\text{comp}}(M)\), \(\rho \in [0, 1/2)\), compactly supported in \(E\). Set \(\varepsilon \in [0, 1/2 - \rho)\). Then

\[
\sup_{|t| < \varepsilon T_E} \left\| U(-t) \ast \mathcal{O}_h(a) \ast U(t) - \mathcal{O}_h(a \ast G_t) \right\|_{L^2(M) \to L^2(M)} = O(h^{1 - 2\rho - 2\varepsilon})
\]

as \(h \to 0\).

3.2. Rate of ergodicity

In a negatively curved manifold \((M, g)\), the geodesic flow \(G_t\) on \(S^\ast M\) is Anosov, hence is ergodic with respect to the Liouville measure \(\mu_t\). See Anosov [Ano]. Let \(f \in L^2(S^\ast M)\). Define the time-average of \(f\) as

\[
\text{Av}_T(f) = \frac{1}{T} \int_0^T f \ast G_t \, dt.
\]

Then the von Neumann mean ergodic theorem states that

\[
\|\text{Av}_T(f) - \mu_t(f)\|_{L^2(S^\ast M)} = o_T(1) \quad \text{as } T \to \infty.
\]
To get a more quantitative version of the above convergence, we need the exponential decay of correlations from Liverani [Liv, corollary 2.5]:

**Theorem 3.2 (Exponential decay of correlations).** For each $\gamma \in (0, 1)$, there exist $c, C > 0$ depending on $\gamma$ such that for each $f, g \in C^r(S^*M)$,

$$
\left| \int_{S^*M} f \ast G_i \, d\mu_i - \mu_i(f) \mu_i(g) \right| \leq C e^{-ct} \|f\|_r \|g\|_r.
$$

(3.2)

Here, $\|f\|_r$ is the Hölder norm of a function $f \in C^r(S^*M)$, the space of Hölder continuous functions on $S^*M$. From this result we derive the following theorem.

**Theorem 3.3 (Rate of ergodicity).** Let $\gamma \in (0, 1)$ and $f \in C^r(S^*M)$. There exist $C > 0$ depending on $\gamma$ and $M$ such that,

$$
\|\text{Av}_T(f) - \mu_1(f)\|_{L^2(S^*M)} \leq \frac{C \|f\|_r}{\sqrt{T}}.
$$

(3.3)

**Remark.** The left-hand side of (3.3) is the variance in the central limiting theorem (see Zelditch [Ze2]). So the above theorem asserts that the variance tends to zero in a polynomial rate of time $T^{-1/2}$ for Hölder continuous functions.

**Proof of theorem 3.3.** Without loss of generality, we may assume that $\mu_1(f) = 0$. Compute that

$$
\|\text{Av}_T(f)\|_{L^2(S^*M)}^2 = \int_{S^*M} \left( \frac{1}{T} \int_0^T f \ast G_i \, dt \right) \left( \frac{1}{T} \int_0^T \tilde{f} \ast G_i \, dt \right) \, d\mu_i
$$

$$
= \frac{1}{T^2} \int_0^T \int_0^T \int_{S^*M} f \ast G_i \ast G_i \, d\mu_i \, dt \, ds
$$

$$
= \frac{1}{T^2} \int_0^T \int_0^T \int_{S^*M} (f \ast G_i \ast G_i \ast G_{-s}) \, d\mu_i \, dt \, ds \quad \text{since } G_{-s} \text{ is symplectic}
$$

$$
= \frac{1}{T^2} \int_0^T \int_0^T \int_{S^*M} G_{-s} \ast \tilde{f} \, d\mu_i \, dt \, ds
$$

$$
\leq \frac{C \|f\|_r^2}{T^2} \int_0^T \int_0^T e^{-c|t-s|} \, dt \, ds \quad \text{by theorem 3.2}
$$

$$
\leq \frac{C \|f\|_r^2}{T}.
$$

and the theorem follows.

4. Small scale quantum ergodicity: proofs of theorems 1.5–1.7

In this section, we prove small scale quantum ergodicity.

4.1. Proof of theorem 1.5

Let $\delta(h)$ be $\rho$-admissible for some $\rho \in [0, 1/2)$. 3275
**Step 1.** Given $a \in S_{0\hbar}(M)$, we first assume that

$$a$$

is compactly supported in $E$ and $\mu_\gamma(a) = 0$.

Recall that $E = \{ 1/2 \leq |t|_1 \leq 3/2 \}$. Then $a \in S_{\rho^+}^{\text{comp}}(M)$ so the results in section 3.1 apply. That is, $a \ast G_t \in S_{\rho^+}^{\text{comp}}(M)$ if $e \in [0, 1/2 - \rho)$ and $|t| \leq \epsilon T_E$. Therefore, the time average $\text{Av}_\gamma(a) \in S_{\rho^+}^{\text{comp}}(M)$ for $|T| \leq \epsilon T_E$.

Next, we see that for any $\gamma \in (0, 1)$,

$$\|a\|_\gamma \leq C \delta(h)^{-\gamma},$$

where $C$ depends on finite number of seminorms of $a$. Then the rate of ergodicity in theorem 3.3 implies that

$$\|\text{Av}_\gamma(a)\|_{L^2(S^*M)} \leq C \|a\|_\gamma \leq C \delta(h)^{-\gamma}. $$

We set $T = \epsilon T_E$. The above inequality continues as

$$\|\text{Av}_{\epsilon T_E}(a)\|_{L^2(S^*M)} \leq C \frac{\delta(h)^{-\gamma}}{\sqrt{|\epsilon| \log h/|t|}} \leq C \frac{t}{\epsilon} \frac{\delta(h)^{-\gamma}}{\sqrt{|\log h|}}. \quad (4.1)$$

For the semiclassical pseudodifferential operator $\text{Op}_h(a)$, we define

$$\text{Av}^L_T(\text{Op}_h(a)) = \frac{1}{T} \int_0^T U(-t)\text{Op}_h(a)U(t) \, dt.$$  

To evaluate $V_2(h, a)$ in theorem 1.5, note that

$$U(t)u_j = e^{-i\hbar E_j}u_j = e^{-iE_j/\hbar}u_j,$$

then

$$\langle \text{Av}^L_T(\text{Op}_h(a))u_j, u_j \rangle = \langle \text{Op}_h(a)u_j, u_j \rangle \quad \text{for all } t \in \mathbb{R}. $$

Compute that

$$V_2(h, a) = \hbar^{n-1} \sum_{E_j \in [1, 1+h]} \left| \langle \text{Op}_h(a)u_j, u_j \rangle \right|^2$$

$$= \hbar^{n-1} \sum_{E_j \in [1, 1+h]} \left| \langle \text{Av}^L_T(\text{Op}_h(a))u_j, u_j \rangle \right|^2$$

$$\leq \hbar^{n-1} \sum_{E_j \in [1, 1+h]} \left( \text{Av}^L_T(\text{Op}_h(a))^* \text{Av}^L_T(\text{Op}_h(a))u_j, u_j \right) \quad (4.2)$$

by Cauchy–Schwarz inequality, where we drop the absolute value sign since each term in the last summation is non-negative.

Since $\text{Av}_\gamma(a) \in S_{\rho^+}^{\text{comp}}(M)$ for $e \in [0, 1/2 - \rho)$ and $|t| \leq \epsilon T_E$, using Egorov’s theorem until Ehrenfest time in theorem 3.1.
Together with the similar estimate on the adjoint, \( A_v^U(\mathcal{O}_a(t)) = O(h^{1-2\beta-2\epsilon}) \), we have

\[
\left\| A_v^U(\mathcal{O}_a(t)) \right\|_{L^2(M)} = O(h^{1-2\beta-2\epsilon}),
\]

using the \( L^2 \) boundedness in theorem 2.1 of operators in \( \Psi_p(M) \). We see that using (2.2)

\[
\mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \in \Psi^c(M)
\]

with principal symbol

\[
|\mathcal{A}_v^U(\mathcal{O}_a)|^2 \in S^c(M).
\]

Continue to estimate \( V_2(h, a) \). By Weyl’s law (see Duistermaat and Guillemin [DuGu]) that 

\[
\#(E_j \in [1, 1 + h]) = O(h^{1-n}),
\]

we have that

\[
(4.2) \leq h^{n-1} \sum_{E_j \in [1, 1 + h]} \left\{ A_v^U(\mathcal{O}_a) * A_v^U(\mathcal{O}_a) \right\} (u_j, u_j)
\]

\[
= h^{n-1} \sum_{E_j \in [1, 1 + h]} \left\{ \mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \right\} (u_j, u_j) + O(h^{1-2\beta-2\epsilon})
\]

\[
\leq h^{n-1} \sum_j \psi \left( \frac{1 - E_j}{h} \right) \left\{ \mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \right\} (u_j, u_j) + O(h^{1-2\beta-2\epsilon}).
\]

Here, we choose the positive function \( \psi \) such that \( \psi \in C_0^\infty(\mathbb{R}) \) and \( \psi \geq 1 \) in \([-1, 0]\) so

\[
\psi \left( \frac{1 - E_j}{h} \right) = 1 \text{ when } E_j \in [1, 1 + h];
\]

moreover, we can always assume that \( \text{supp} \psi \) contains no periods of \( G \), since \( \mathcal{M} \) has injectivity radius much greater than 1.

Then applying semiclassical trace formula proposition 2.3, we have

\[
(4.2) = h^{n-1} \sum_j \psi \left( \frac{1 - E_j}{h} \right) \left\{ \mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \mathcal{O}_a(\mathcal{A}_v^U(\mathcal{O}_a)) \right\} (u_j, u_j) + O(h^{1-2\beta-2\epsilon})
\]

\[
= \mu_1(\mathcal{A}_v^U(\mathcal{O}_a))^2 + C \mathcal{H} \| \mathcal{A}_v^U(\mathcal{O}_a) \|^2_{L^{2+\epsilon}} + O(h^{1-2\beta-2\epsilon})
\]

\[
= \mu_1(\mathcal{A}_v^U(\mathcal{O}_a))^2 + O(h^{1-2\beta-2\epsilon}).
\]

From (4.1) we have

\[
\mu_1(\mathcal{A}_v^U(\mathcal{O}_a))^2 = \mathcal{A}_v^U(\mathcal{O}_a)^2_{L^2(S^*M_0)} \leq C \frac{\delta(h)^{-2\epsilon}}{\epsilon \log h}.
\]

The theorem evidently follows when \( \alpha = 0 \) and \( \beta = 1 \). If
$$\alpha > 0 \quad \text{and} \quad \delta(h) = |\log h|^{-\alpha},$$
then $a \in S^{\text{comp}}_{(h)}(\mathbb{M}) \subset S^{\text{comp}}(\mathbb{M})$ for all $\rho \in (0, 1/2)$. We therefore derive that

$$V_2(h, a) = C\rho_2(|A\nu;\mathcal{L}_2(a)|^2) + \mathcal{O}(h^{1-2(2r+8\rho+c)}) = C|\log h|^{2\gamma-1} \quad \text{as} \quad h \to 0$$

by picking $\rho$ and $\epsilon$ small enough. Notice that the rate of ergodicity in theorem 3.3 is valid for all $\gamma \in (0, 1)$. Hence, for

$$0 \leq \beta < 1,$$

one can find $\gamma \in (0, 1)$ such that

$$\beta \leq 1 - 2\alpha \gamma < 1,$$

Therefore,

$$|\log h|^{2\gamma-1} = O(|\log h|^{-\beta}).$$

**Step 2.** Let $a \in S^{\delta}(\mathbb{M})$, where $\delta(h) = |\log h|^{-\alpha}$ for $\alpha \geq 0$. We have that $a \in S^{\delta}(\mathbb{M})$ for all $\rho \in (0, 1/2)$. Consider the symbol

$$\tilde{a} = (a - \mu_t(a))\phi(|x|).$$

Then $\tilde{a}$ is compactly supported in the energy shell $E$ since $\text{supp} \phi \subset (-1/2, 1/2)$; moreover

$$\mu_t(\tilde{a}) = \int_{S^*\mathbb{M}} (a - \mu_t(a))\phi(|x|) \, d\mu_t = \int_{S^*\mathbb{M}} a \, d\mu_t - \mu_t(a) = 0,$$

because $\phi(|x|) = 1$ near $S^*\mathbb{M}$. So theorem 1.5 applies to the symbol $\tilde{a}$ by the previous step:

$$V_2(h, \tilde{a}) = h^{\alpha-1} \sum_{E_j \in (1, 1+h]} \left| \left\langle \text{Op}_h(\tilde{a})u_j, u_j \right\rangle \right|^2 = O(|\log h|^{-\beta}) \quad \text{as} \quad h \to 0.$$

To prove theorem 1.5 for the symbol $a$, note the difference

$$\left| \left\langle \text{Op}_h(a)u_j, u_j \right\rangle - \left\langle \text{Op}_h(\tilde{a})u_j, u_j \right\rangle \right| = \left| \left\langle \text{Op}_h((a - \mu_t(a))\phi(|x|))u_j, u_j \right\rangle \right| \leq \| \text{Op}_h((a - \mu_t(a))\phi(|x|) - 1)u_j \|_{L^2(\mathbb{M})} := \| Qu_j \|_{L^2(\mathbb{M})},$$

in which the semiclassical pseudodifferential operator $Q$ has symbol

$$q = (a - \mu_t(a))(\phi(|x|) - 1) \in S^{\delta}_{(\mathbb{M})}.$$

We see that $q$ is supported in $T^*\mathbb{M} \setminus E$. We also know that the symbol of $h^2\Delta - E_j; \quad p = |x|^2 - E_j$, when $E_j \in [1, 1+h]$ is elliptic in $T^*\mathbb{M} \setminus \mathcal{E}$, that is, $|x|^2 - E_j \geq c > 0$. Let $r = q/p$ and $R = \text{Op}_h(r)$. Then $r \in S^{\delta}(\mathbb{M})$ and by (2.2)

$$Q = RP + E, \quad \text{where} \quad E \in h^{-2}S^{\delta}(\mathbb{M}).$$

Note that the operators in the above equation depend on the energy $E_j$. Since $P_{uj} = 0$, we have from theorem 2.1

$$\| Qu_j \|_{L^2(\mathbb{M})} \leq \| RP_{uj} \|_{L^2(\mathbb{M})} + \| E_{uj} \|_{L^2(\mathbb{M})} = O(h^{1-2\rho}).$$

Now we compute that
\[ V_2(h, a) = \sum_{E \in [1, 1 + h]} \left| \left( \text{Op}_h(a)u_j, u_j \right) - \mu_1(a) \right|^2 \]
\[ \leq \sum_{E \in [1, 1 + h]} \left( \left| \left( \text{Op}_h(a)u_j, u_j \right) - \left( \text{Op}_h(\tilde{a})u_j, u_j \right) \right| + \left| \left( \text{Op}_h(\tilde{a})u_j, u_j \right) \right| \right)^2 \]
\[ \leq \sum_{E \in [1, 1 + h]} \left( \left| \left( \text{Op}_h(\tilde{a})u_j, u_j \right) \right| + O(h^{1-2\gamma}) \right)^2 \]
\[ = O(\|h\|^{-\beta}) \quad \text{as } h \to 0. \]

4.2. Proof of theorem 1.6

Notice that for the symbol \( a^b_0 \) defined in (1.8),
\[ \|a^b_0\|_r \leq C\delta(h)^{-\gamma}, \]  
(4.4)
where \( C \) depends on \( b, \phi, \) and \( M \), and is uniform in \( z_0 \in S^*\mathbb{M} \).

Following the argument in the previous subsection,
\[ V_2(h, a^b_0) \leq C \frac{\delta(h)^{-2\gamma}}{|\log h|}. \]

The theorem evidently follows when \( \alpha = 0 \) and \( \beta = 1 \). If
\[ 0 < \alpha < \frac{1}{2n-1} \quad \text{and} \quad \delta(h) = |\log h|^{-\alpha}, \]
for
\[ 0 \leq \beta < 1 - 2\alpha(2n - 1), \]
one can find \( \gamma \in (0, 1) \) such that
\[ \beta \leq 1 - 2\alpha(2n + \gamma - 1), \]
so
\[ \delta(h)^{2(2n-1)} |\log h|^{-\beta} \geq |\log h|^{2\gamma - 1} = \frac{\delta(h)^{-2\gamma}}{|\log h|}. \]

Therefore,
\[ V_2(h, a^b_0) \leq C \frac{\delta(h)^{-2\gamma}}{|\log h|} = C \delta(h)^{2n-1/|\log h|^{-\beta}}. \]

Notice that (4.4) is uniform for \( z_0 \in S^*\mathbb{M} \). The theorem thus follows since \( C \) depends on \( M \) and finite number of derivatives of \( b \) and \( \phi \).

4.3. Proof of theorem 1.7

The theorem follows by noticing that \( a^b_0 \) defined in (1.8),
\[ \|a^b_0\|_r \leq C\delta(h)^{-\gamma}, \]  
(4.5)
where $C$ depends on $b, \phi$, and $\mathbb{M}$, and is uniform in $x_0 \in \mathbb{M}$. Therefore, when $\alpha$ and $\beta$ in the required ranges, we have that

$$V_\delta(h, a_h^b) \leq C \frac{\delta(h)^{-2\gamma}}{|\log h|} = C \delta(h)^{2\alpha} |\log h|^{-\beta}.$$ 

Notice that (4.5) is uniform for $x_0 \in \mathbb{M}$. The theorem thus follows since $C$ depends on $\mathbb{M}$ and finite number of derivatives of $b$ and $\phi$.

5. Small scale mass equidistribution: proofs of corollaries 1.8 and 1.9

In this section, we prove corollaries 1.8 and 1.9. Then we discuss some potential improvements on small scale quantum ergodicity.

5.1. Proof of corollary 1.8

We keep the notations that

$$h = \lambda^{-1}, \quad h^{-2}E_j = \lambda_j^2, \quad \text{and} \quad \delta(h) = r(\lambda) = (\log \lambda)^{-\beta}$$

such that (1.11) holds. Then

$$\Delta u_j = h^{-2}E_j u_j = \lambda_j^2 u_j.$$ 

We use the local coordinates induced by exponential map $\text{Exp}_{x_0} : T_{x_0} \mathbb{M} \to \mathbb{M}$. From the assumption that the injectivity radius of $\mathbb{M}$ is much greater than 1, $f = \text{Exp}_{x_0}(0, 0, 2)$ defines a local coordinate system around $x_0$, where $D(0, 2)$ is the ball centered at 0 with radius 2 in $\mathbb{R}^n$. We still use $\tilde{x}$ as the the local coordinate of $x$ when $x \in B(x_0, 2)$. So $\tilde{x}_0 = 0$.

Given a function $b \in C_0^\infty(D(0, 2))$, consider the symbol

$$a_h^b(x, \xi; h) = h \left( \frac{\tilde{x}}{\delta(h)} \right) \phi(h(\tilde{x})).$$

We have

$$\langle \text{Op}_h(a_h^b) u_j, u_j \rangle = \int_{D(0, 2)} b \left( \frac{\tilde{x}}{\delta(h)} \right) |\mu_j * f(\tilde{x})|^2 |Df(\tilde{x})| \, d\tilde{x} + O(h),$$

in which the $O(h)$ term comes from removing the frequency cutoff $\phi$; and

$$\mu_j(a_h^b) = \int_{S^{B(0, 2)}} a_h^b(x, \xi; h) \, d\mu = \frac{1}{\text{Vol}(\mathbb{M})} \int_{D(0, 2)} b \left( \frac{\tilde{x}}{r(\lambda)} \right) |Df(\tilde{x})| \, d\tilde{x}. $$

Putting them together with (1.11), we have

$$\left\{ \frac{1}{\# \{ \lambda_j \in [\lambda, \sqrt{\lambda^2 + \bar{c}}] \} } \right\}$$

$$\times \sum_{\lambda_j \in [\lambda, \sqrt{\lambda^2 + \bar{c}}]} \int_{D(0, 2)} b \left( \frac{\tilde{x}}{r(\lambda)} \right) |\mu_j * f(\tilde{x})|^2 |Df(\tilde{x})| \, d\tilde{x} = C r(\lambda)^{2\alpha} |\log \lambda|^{-\beta}.$$
where $C$ depends on $M$ and finite number of derivatives of $b$ and $\phi$ so the above quantity is bounded by

$$C_M,\phi \|b\|_{C^m} r(\lambda)^{2n}|\log \lambda|^{-\beta},$$

where $m > 0$ is a sufficiently large integer and $C_M,\phi$ now only depends on $M$ and $\phi$.

Thus a standard argument (see [Ze 3, lemma 3]) gives a full density subsequence $\{u_{\lambda,h}\} \subset \{u_j\}$ such that

$$\left| \int_{D(0,2)} b \left( \frac{\dd x}{r(\lambda,h)} \right) u_{\lambda,h} \ast f(\dd x) \, |Df(\dd x)| \, \dd \bar{x} - \frac{1}{\Vol(M)} \int_{D(0,2)} b \left( \frac{\dd x}{r(\lambda,h)} \right) |Df(\dd x)| \, \dd \bar{x} \right|$$

$$= C_M,\phi \|b\|_{C^m} r(\lambda,h)^n |\log \lambda,h|^{-\beta} \quad \text{as } k \to \infty$$

(5.1)

for some $0 < \bar{\beta} < \beta$.

For $i \in \mathbb{Z}$ and $|i| \geq 10$, define $b_i \subset C^0_0(D(0,2))$ such that $0 \leq b_i \leq 1$ and

$$b_i(\dd x) = \begin{cases} 1 & \text{for } \dd x \in D(0,1 + 1/i); \\ 0 & \text{for } \dd x \in D(0,2) \setminus D(0,1 + 2/i). \end{cases}$$

For each $b_i$, there is a full density subsequence $\{u_{\lambda,h}^{(i)}\}$ depending on $b_i$ such that (5.1) holds. Using a diagonal argument (see [Zw, theorem 15.5]), we can find a full density subsequence $\{u_{\lambda,h}\} \subset \{u_j\}$ such that

$$\left| \int_{D(0,2)} b \left( \frac{\dd x}{r(\lambda,h)} \right) u_{\lambda,h} \ast f(\dd x) \, |Df(\dd x)| \, \dd \bar{x} - \frac{1}{\Vol(M)} \int_{D(0,2)} b \left( \frac{\dd x}{r(\lambda,h)} \right) |Df(\dd x)| \, \dd \bar{x} \right|$$

$$= C_M,\phi \|b\|_{C^m} r(\lambda,h)^n |\log \lambda,h|^{-\beta} \quad \text{as } k \to \infty$$

for all $b_i$. Denote $r_h = r(\lambda,h)$. Using the above inequality, we have that for $i \geq 10$

$$\int_{B(x_0, r_h)} |u_{\lambda,h}|^2 \, d\Vol$$

$$= \int_{D(0,r_h)} |u_{\lambda,h} \ast f(\dd x)|^2 |Df(\dd x)| \, \dd \bar{x}$$

$$\leq \int_{D(0,2)} b_i \left( \frac{\dd x}{r_h} \right) |u_{\lambda,h} \ast f(\dd x)|^2 |Df(\dd x)| \, \dd \bar{x}$$

$$\leq \frac{1}{\Vol(M)} \int_{D(0,2)} b_i \left( \frac{\dd x}{r_h} \right) |Df(\dd x)| \, \dd \bar{x} + C_M,\phi \|b_i\|_{C^m} r_h^n |\log \lambda,h|^{-\beta}$$

$$\leq \frac{\Vol(B(x_0, r_h))}{\Vol(M)} \|f\|_{L^p} r_h^n |\log \lambda,h|^{-\beta} + C_M,\phi \|b_i\|_{C^m} r_h^n |\log \lambda,h|^{-\beta}.$$

The last inequality follows from the fact that

$$\int_{D(0,2)} b_i \left( \frac{\dd x}{r_h} \right) |Df(\dd x)| \, \dd \bar{x} \leq \int_{D(0,1 + 2/i) \setminus D(0,1)} |Df(\dd x)| \, \dd \bar{x} \leq C_M r_h^n |\log \lambda,h|,$$

where $C_M$ depends only on $M$. We know that
\[\|b_i\|_{c^m} \leq c i^m, \quad \text{where } c \text{ is an absolute constant.}\]

So
\[
\int_{B(x_0, r_i)} |u_{\lambda, i}|^2 \, d\text{Vol} \leq \frac{\text{Vol}(B(x_0, r_i))}{\text{Vol}(\mathbb{M})} + C_M r_i^\beta \log r_i^{-\beta} \leq \frac{\text{Vol}(B(x_0, r_i))}{\text{Vol}(\mathbb{M})} + C_M r_i^\beta (r_i^{-1} + i^m \log r_i^{-\beta}).
\]

Letting
\[i = \left\lfloor \frac{\beta}{m+1} \right\rfloor\]
gives that
\[
\int_{B(x_0, r_i)} |u_{\lambda, i}|^2 \, d\text{Vol} \leq \frac{\text{Vol}(B(x_0, r_i))}{\text{Vol}(\mathbb{M})} + o(r_i^n).
\]

Similarly, for \(i \leq -10\) choose appropriate \(b_i\). We can show that
\[
\int_{B(x_0, r_i)} |u_{\lambda, i}|^2 \, d\text{Vol} \geq \frac{\text{Vol}(B(x_0, r_i))}{\text{Vol}(\mathbb{M})} + o(r_i^n).
\]

Therefore, we prove (1.12) for this fixed point \(x_0 \in \mathbb{M}\).

5.2. Proof of corollary 1.9

Step 1. First we need a covering lemma that is similar to Colding and Minicozzi [CM, lemma 2]. For readers’ convenience, we provide a proof here.

Lemma 5.1. Let \(r > 0\) be sufficiently small. Then there exists a family of geodesic balls that covers \(\mathbb{M}\):
\[
\{ B(x_i, r) \}_{i=1}^N \supseteq \mathbb{M} \quad \text{with } N \leq c e^{-n};
\]

furthermore, the following statements are valid.

(1) Each ball \(B(x, r) \subseteq \mathbb{M}\) is covered by at most \(c_2\) balls among \(\{ B(x_i, r) \}_{i=1}^N \):
\[
B(x, r) \subset \bigcup_{m=1}^{c_2} B(x_{i_m}, r) \quad \text{for some } \{ i_1, \ldots, i_{c_2} \} \subset \{ 1, \ldots, N \}.
\]

(2) Each ball \(B(x, r) \subseteq \mathbb{M}\) contains at least one ball among \(\{ B(x_i, r/3) \}_{i=1}^N \):
\[
B(x, r) \supset B(x_i, r/3) \quad \text{for some } i \in \{ 1, \ldots, N \}.
\]

Here, the positive constants \(c_1\) and \(c_2\) depend only on \(\mathbb{M}\).

Proof of Lemma 5.1. Select a maximal disjoint family of balls \(\{ B(x_i, r/3) \}_{i=1}^N \) in \(\mathbb{M}\). Then
\[ \sum_{i=1}^{N} \Vol(B(x_i, r/3)) \leq \Vol(M). \]

So \( N \leq c_1 r^{-n} \), where \( c_1 > 0 \) depends only on \( \mathcal{M} \). It follows immediately from maximality that the doubled balls \( B(x_i, 2r/3), i = 1, ..., N \), cover \( \mathcal{M} \).

Therefore, the family \( \{ B(x_i, r) \}_{i=1}^{N} \) is also a cover of \( \mathcal{M} \).

Given any \( x \in \mathcal{M} \), suppose that

\[ B(x, r) \subseteq \bigcup_{i=1}^{k} B(x_i, r) \quad \text{and} \quad B(x, r) \cap B(x_i, r) \neq \emptyset \quad \text{for all } i = 1, ..., k. \]

Since the disjoint balls \( B(x_i, r/3), i = 1, ..., k, \) are all contained in \( B(x, 3r) \),

\[ \sum_{i=1}^{k} \Vol(B(x_i, r/3)) \leq \Vol(B(x, 3r)); \]

but for each \( i = 1, ..., k \), \( B(x, 3r) \subseteq B(x_i, 5r) \) so

\[ \Vol(B(x, 3r)) \leq \Vol(B(x_i, 5r)) \leq C_3 \Vol(B(x_i, r/3)). \]

Combing the above two inequalities, we have that \( k \leq c_2 \), where \( c_2 \) depends only on \( \mathcal{M} \). This shows (1) in the lemma.

Given any \( x \in \mathcal{M} \), suppose that

\[ x \in B(x_i, 2r/3) \quad \text{for some } i = 1, ..., N. \]

Then

\[ B(x_i, r/3) \subseteq B(x, r), \]

which proves (2) in the lemma.

\[ \square \]

**Step 2.** For \( 0 \leq \alpha < 1/(2\alpha) \), take \( r = \delta(h) = |\log h|^{-\alpha} \) in lemma 5.1; let \( \{ x_i \}_{i=1}^{N(h)} \) be the centers of the balls selected in Step 1. So \( \delta(h) \leq c_1 r^{-n} = c_1 |\log h|^{-\alpha} \). From theorem 1.7,

\[ V_2(h, a_\alpha^h) = h^{n-1} \sum_{E_j \in [1.1+H]} \left| \left\{ \text{Op}_b(a_\phi^h) u_j, u_j \right\} - \mu_t(a_\phi^h) \right|^2 = C_\delta(h) \delta(h)^{2n} O(|\log h|^{-\bar{\beta}}) \]

as \( h \to 0 \). Here, \( C \) depends on the base function \( b \), \( \phi \), and \( \mathcal{M} \), and is uniform for \( x_i \).

For \( \bar{\beta} > 0 \), write

\[ \Lambda^h_{\bar{\beta}} = \{ j : E_j \in [1, 1+h] \text{ and } \left| \langle \text{Op}_b(a_\phi^h) u_j, u_j \rangle - \mu_t(a_\phi^h) \right| \geq \delta(h)^\alpha |\log h|^{-\bar{\beta}} \}. \]

Hence,

\[ h^{n-1} \# \Lambda^h_{\bar{\beta}} \leq C |\log h|^{2\bar{\beta} - \bar{\beta}}. \]

Let
\[ \Lambda^h(h) = \bigcup_{i=1}^{N(h)} \Lambda^h_i(h). \]

Then
\[ h^{-1} \# \Lambda^h(h) \leq N(h) C \log h^{2\bar{\beta} - \beta} \leq C \log h^{[\alpha n + 2\bar{\beta} - \beta]}. \]

Write
\[ \Gamma^h(h) = [j : E_j \in [1, 1 + h]] \setminus \Lambda^h(h). \]

Thus, if \( an + 2\bar{\beta} - \beta < 0 \), then
\[ \frac{\# \Gamma^h(h)}{\# [j : E_j \in [1, 1 + h]]} = 1 - \frac{\# \Lambda^h(h)}{\# [j : E_j \in [1, 1 + h]]} \leq 1 - C \log h^{[\alpha n + 2\bar{\beta} - \beta]} \to 1 \]

as \( h \to 0 \) by Weyl’s law. Note that if \( \beta - an > 0 \), then there exists \( \tilde{\beta} > 0 \) such that
\[ an + 2\tilde{\beta} - \beta < 0. \]

But this just means that \( \alpha < 1/(3n) \) since \( 0 \leq \beta < 1 - 2an \). Note the existence of \( \tilde{\beta} \) here requires \( \alpha \) to be in a small range \([0, 1/(3n)]\) than corollary 1.8.

Now \( j \in \Gamma^h(h) \) implies that
\[ \left| \langle \text{Op}_\varphi(a^h_i)u_j, u_j \rangle - \mu_i(a^h_i) \right| < \tilde{\delta}(h)^\beta \log h^{-\tilde{\beta}} \]
for some \( 0 < \tilde{\beta} < (\beta - an)/2 \), and is uniform for all \( x_i, i = 1, \ldots, N(h) \).

**Step 3.** In a small patch of \( M \) that contains \( x_i \), similar to corollary 1.8, we use the local coordinate system \( f_i = \text{Exp}_{x_i}(D(0, 2)) \). We still use \( \bar{x} \) as the the local coordinate of \( x \) when \( x \in B(x_i, 2) \). So \( \bar{x}_i = 0 \).

Consider the symbol
\[ a^h_i(x, \bar{x}; h) = b \left( \frac{\bar{x}}{\delta(h)} \right) \phi(\|\bar{x}\|), \]
where \( b \in C^\infty_0(D(0, 2)) \). We have
\[ \langle \text{Op}_\varphi(a^h_i)u_i, u_i \rangle = \int_{D(0, 2)} b \left( \frac{\bar{x}}{\delta(h)} \right) \|u_j * f_i(\bar{x})\|^2 |Df_i(\bar{x})| \, d\bar{x} + O(h), \]
and
\[ \mu_i(a^h_i) = \int_{S^{2|D(0, 2)}} \frac{1}{\text{Vol}(\mathcal{M})} \int_{D(0, 2)} b \left( \frac{\bar{x}}{\delta(h)} \right) |Df_i(\bar{x})| \, d\bar{x}. \]

Hence, from step 2, \( j \in \Gamma^h(h) \subset [j : E_j \in [1, 1 + h]] \) implies that
\[ \left| \int_{D(0, 2)} b \left( \frac{\bar{x}}{\delta(h)} \right) |u_j * f_i(\bar{x})|^2 |Df_i(\bar{x})| \, d\bar{x} - \frac{1}{\text{Vol}(\mathcal{M})} \int_{D(0, 2)} b \left( \frac{\bar{x}}{\delta(h)} \right) |Df_i(\bar{x})| \, d\bar{x} \right| < \tilde{\delta}(h)^\beta \log h^{-\tilde{\beta}} \]
for some \( \tilde{\beta} > 0 \), and is uniform for \( x_i, i = 1, \ldots, N(h) \).
Choose $b_1 \in C^0_0(D(0,2))$, $0 \leq b_1 \leq 1$, and $b_1 = 1$ in $D(0,1)$. Then (5.2) implies that there is a positive constant $h_1 \ll 1$ such that if $j \in \Gamma^{b_1}(h)$ and $0 < h \leq h_1$, then
\[
\int_{B(x_0,\delta(h))} |u_j(x)|^2 \, dVol \\
\leq \int_{D(0,2)} b_1 \left( \frac{\bar{x}}{\delta(h)} \right) |u_j \ast f_j(\bar{x})|^2 |Df_j(\bar{x})| \, d\bar{x} \\
\leq \frac{1}{Vol(M)} \int_{D(0,2)} b_1 \left( \frac{\bar{x}}{\delta(h)} \right) |Df_j(\bar{x})| \, d\bar{x} + \delta(h)^n |\log h|^{-\beta} \\
\leq \frac{Vol(B(x_0,2\delta(h)))}{Vol(M)} + \delta(h)^n |\log h|^{-\beta} \\
\leq CVol(B(x_0,\delta(h))),
\]
which is uniform for $x_i$, $i = 1, \ldots, N(h)$. Here, $C$ depends on $b_1$, $\phi$, and $M$. From Step 1, each $B(x,\delta(h))$ is covered by at most $c_2$ balls among $\{B(x_i,\delta(h))\}_{i=1}^{N(h)}$, where $c_2$ only depends on $M$. Hence, for each $B(x,\delta(h)) \subset M$,
\[
\int_{B(x,\delta(h))} |u_j(x)|^2 \, dVol \leq CVol(B(x,\delta(h)))
\]
if $j \in \Gamma^{b_1}(h)$ and $0 < h \leq h_1$, and is uniform for $x \in M$.

Choose $b_2 \in C^0_0(D(0,1/3))$, $0 \leq b_2 \leq 1$, and $b_2 = 1$ in $D(0,1/6)$. Then (5.2) implies that there is a positive constant $h_2 \ll 1$ such that if $j \in \Gamma^{b_2}(h)$ and $0 < h \leq h_2$, then
\[
\int_{B(x,\delta(h)/3)} |u_j(x)|^2 \, dVol \\
\geq \int_{D(0,1/3)} b_2 \left( \frac{\bar{x}}{\delta(h)} \right) |u_j \ast f_j(\bar{x})|^2 |Df_j(\bar{x})| \, d\bar{x} \\
\geq \frac{1}{Vol(M)} \int_{D(0,1/3)} b_2 \left( \frac{\bar{x}}{\delta(h)} \right) |Df_j(\bar{x})| \, d\bar{x} + \delta(h)^n |\log h|^{-\beta} \\
\geq \frac{Vol(B(x,\delta(h)/6))}{Vol(M)} + \delta(h)^n |\log h|^{-\beta} \\
\geq cVol(B(x,\delta(h))),
\]
which is uniform for $x_i$, $i = 1, \ldots, N(h)$. Here, $c$ depends on $b_2$, $\phi$, and $M$. From step 1, each $B(x,\delta(h))$ contains at least one ball among $\{B(x_i,\delta(h)/3\}_{i=1}^{N(h)}$. Hence, for each $B(x,\delta(h)) \subset M$,
\[
\int_{B(x,\delta(h))} |u_j(x)|^2 \, dVol \geq cVol(B(x,\delta(h))).
\]
if $j \in \Gamma^{b_2}(h)$ and $0 < h \leq h_2$, and is uniform for $x \in M$.

Let $h_0 = \min\{h_1, h_2\}$ and $G(h) = \Gamma^{b_1}(h) \cap \Gamma^{b_2}(h)$. We have that if $0 < h \leq h_0$ and $j \in G(h)$, then
\[
cVol(B(x,\delta(h))) \leq \int_{B(x,\delta(h))} |u_j(x)|^2 \, dVol \leq CVol(B(x,\delta(h)))
\]
(5.3)
uniformly for $x \in \mathbb{M}$, where the constants $c$ and $C$ depend only on $M$. Moreover,

$$\frac{\# \Gamma(h)}{\# \{ j : E_j \in [1, 1 + h] \}} \geq 1 - \frac{\# \Lambda^b(h) + \# \Lambda^b(\bar{h})}{\# \{ j : E_j \in [1, 1 + h] \}} \to 1 \text{ as } h \to 0. \quad (5.4)$$

**Step 4.** We now pass from semiclassical version in Step 3 to its high-energy version. Recall the notations that $h = \lambda^{-1}$, $h^{-2}E_j = \lambda_j^2$, and $\delta(h) = r(\lambda) = (\log \lambda)^{-n}$. Then

$$\Delta u_j = h^{-2}E_j u_j = \lambda_j^2 u_j$$

and

$$\{ j : E_j \in [1, 1 + h] \} = \{ j : \lambda_j^2 \in [h^{-2}, h^{-2} + h^{-1}] \}.$$

Divide $\mathbb{N}$ into families

$$I_0 = \{ j : \lambda_j^2 \in [0, 1) \},$$

$$I_m = \{ j : \lambda_j^2 \in [h_{m-1}^{-2}, h_m^{-2} + h_{m-1}^{-1}] \}, \quad \text{if } m \geq 1,$$

where $h_1 = 1$ and $h_{m+1}^{-2} = h_m^{-2} + h_{m-1}^{-1}$. Then the sequence

$$J = \bigcup_{m=0}^{\infty} \Gamma(h_m)$$

has full density in $\mathbb{N} = \bigcup_{m=0}^{\infty} I_m$. This is because

$$\frac{\# \Gamma(h)}{\# \{ j : \lambda_j^2 \in [h^{-2}, h^{-2} + h^{-1}] \}} = \frac{\# \Gamma(h)}{\# \{ j : E_j \in [1, 1 + h] \}} \to 1 \text{ as } h \to 0$$

from (5.4). Therefore, letting $J = \{ j \}_{k=1}^{\infty} \subset \mathbb{N}$, in the view of (5.3), we have (1.13):

$$c \text{Vol}(B(x, r(\lambda_j))) \leq \int_{B(x, r(\lambda_j))} |u_j(x)|^2 \, d\text{Vol} \leq C \text{Vol}(B(x, r(\lambda_j))) \text{ as } k \to \infty,$$

uniformly for $x \in \mathbb{M}$.

### 5.3. Further investigations

1. Periodic and non-periodic points. The results in theorems 1.6 and 1.7 are independent of whether the point $z_0 \in S^*\mathbb{M}$ or $x_0 \in \mathbb{M}$ is on a periodic trajectory or not. But the geodesic flow displays very different properties around periodic points and non-periodic points. This is reflected in Paul and Uribe [PU], where pointwise behavior of semiclassical measures is different around periodic and non-periodic points. However, one has to use more delicate information on the dynamical system $(S^*\mathbb{M}, G)$ to get qualitative results in its quantum counterpart.

2. Different shrinking rates in $x$ and in $\xi$. Comparing the symbols that we used in theorems 1.5 and 1.7, the former one has the same shrinking rates in $x \in \mathbb{M}$ and in $\xi \in S^*\mathbb{M}$, while the latter only shrinks in $x \in \mathbb{M}$, and has a slightly better shrinking rate. So generally, one can localize better in $x$ than in $\xi$ without violating the uncertainty principle. Moreover, the pointwise version of the Weyl’s law
\[ \sum_{\lambda \leq \lambda} |u_{\lambda}(x)|^2 = O(h^\rho) \]

holds for all \( x \in \mathcal{M} \). See also Christianson, Hassell and Toth [CHT], in which they used up to \( h^2 \) localization in one direction to estimate the Neumann restriction bound of eigenfunctions. These evidence suggests that better localization in \( x \in \mathcal{M} \) (or in \( \xi \)) is feasible, e.g. to consider symbols as

\[ a^{b,c}(x, \xi; h) = b \left( \frac{x - x_0}{h^\rho} \right) \phi(|\xi|_L), \quad \rho \in (0, 1). \]

(3) Manifolds with constant curvature or arithmetic hyperbolic manifolds. The main tool we used from dynamical system is the exponential rate of correlation (3.2) by Liverani [Liv]. His result, however, is on general contact Anosov flows. In theorem 3.3, the rate of ergodicity is proved to be controlled by the Hölder norm. Its improvement may lead to better localization in theorems 1.5 and 1.7. One expects that this can be done in manifolds with constant curvature or arithmetic hyperbolic manifolds, in which cases there are group symmetry or number theoretic results to exploit.

Acknowledgments

I would like to thank G. Rivièrè for informing me Liverani’s paper [Liv] and pointing out that theorem 3.3 follows from it. I also want to thank H. Hezari and S. Nonnenmacher for their comments that helped to improve the paper. It is a pleasure to thank A. Hassell for encouraging me to work on small scale quantum ergodicity and all our discussions throughout the preparation of this paper. Finally, I would like to thank the referees for very careful reading and detailed comments that helped to improve the presentation.

References

[Ana] Anantharaman N 2008 Entropy and the localization of eigenfunctions Ann. Math. 168 435–75
[Ano] Anosov D 1967 Geodesic flows on closed Riemannian manifolds of negative curvature Tr. Mat. Inst. Steklov. 90
[AN] Anantharaman N and Nonnenmacher S 2007 Half-delocalization of eigenfunctions for the Laplacian on an Anosov manifold Festival Yves Colin de Verdière Ann. Inst. Fourier 57 2465–523
[B] Ballmann W 1995 Lectures on Spaces of Nonpositive Curvature (BaselP: Birkhäuser)
[BL] Brooks S and Lindenstrauss E 2014 Joint quasimodes, positive entropy and quantum unique ergodicity Invent. Math. 198 219–59
[BR] Bouzouina A and Robert D 2002 Uniform semiclassical estimates for the propagation of quantum observables Duke Math. J. 111 223–252
[CdV] Colin de Verdière Y 1985 Ergodicité et fonctions propres du laplacien Commun. Math. Phys. 102 497–502
[CHT] Christianson H, Hassell A and Toth J Exterior mass estimates and \( L^2 \) restriction boundary for Neumann data along hypersurfaces Int. Math. Res. Not. in press (arXiv:1303.4319)
[CM] Colding T and Minicozzi W 2011 Lower bounds for nodal sets of eigenfunctions Commun. Math. Phys. 306 777–84
[DF] Donnelly H and Fefferman C 1988 Nodal sets of eigenfunctions on Riemannian manifolds Invent. Math. 93 161–83
[DuGu] Duistermaat J J and Guillemin V W 1975 The spectrum of positive elliptic operators and periodic bicharacteristics Invent. Math. 29 39–79
[DyGu] Dyatlov S and Guillarmou C 2014 Microlocal limits of plane waves and Eisenstein functions Ann. Sci. Éc. Norm. Supér. 47 371–448

[Ha] Hassell A 2010 Ergodic billiards that are not quantum unique ergodic Ann. Math. 171 605–19

[Ho] Hörmander L 1968 The spectral function of an elliptic operator Acta Math. 121 193–218

[HMR] Helffer B, Martinez A and Robert D 1987 Ergodicité et limite semi-classique Commun. Math. Phys. 109 313–26

[HR1] Hezari H and Rivière G 2015 $L^p$ norms, nodal sets, and quantum ergodicity (arXiv:1411.4078)

[HR2] Hezari H and Rivière G 2015 Quantitative equidistribution properties of toral eigenfunctions (arXiv:1503.02794)

[HS] Holowinsky R and Soundararajan K 2010 Mass equidistribution for Hecke eigenforms Ann. Math. 172 1517–28

[KH] Katok A and Hasselblatt B 1995 Introduction to the Modern Theory of Dynamical Systems (Cambridge: Cambridge University)

[Lin] Lindenstrauss E 2006 Invariant measures and arithmetic quantum unique ergodicity Ann. Math. 163 165–219

[Liv] Liverani C 2004 On contact Anosov flows Ann. Math. 159 1275–312

[LS] Li P and Schoen R 1984 $L^p$ and mean value properties of subharmonic functions on Riemannian manifolds Acta Math. 153 279–301

[PU] Paul T and Uribe T 1996 On the pointwise behavior of semi-classical measures Commun. Math. Phys. 175 229–58

[R] Ratner M 1973 The central limit theorem for geodesic flows on $n$-dimensional manifolds of negative curvature Isr. J. Math. 16 181–97

[RS] Rudnick Z and Sarnak P 1994 The behaviour of eigenstates of arithmetic hyperbolic manifolds Commun. Math. Phys. 161 195–213

[Sa1] Sarnak P 1992 Arithmetic Quantum Chaos. The Schur Lectures (Israel Mathematical Conf. Proc. vol 8) (Ramat-Gan, Israel: Bar-Ilan University) pp 183–236

[Sa2] Sarnak P 2011 Recent progress on the quantum unique ergodicity conjecture Bull. Am. Math. Soc. 48 211–28

[Sc] Schubert R 2006 Upper bounds on the rate of quantum ergodicity Ann. Henri Poincaré. 7 1085–98

[So1] Sogge C 2014 Hangzhou Lectures on Eigenfunctions of the Laplacian (Princeton, NJ: Princeton University)

[So2] Sogge C 2015 Localized $L^p$-estimates of eigenfunctions: a note on an article of Hezari and Rivière (arXiv:1503.07238)

[Sn] Šnirel’man A 1975 The asymptotic multiplicity of the spectrum of the Laplace operator Usp. Mat. Nauk 30 265–6

[SV] Silberman L and Venkatesh A 2007 On quantum unique ergodicity for locally symmetric spaces Geom. Funct. Anal. 17 960–98

[Zel1] Zelditch S 1987 Uniform distribution of eigenfunctions on compact hyperbolic surfaces Duke Math. J. 55 919–41

[Zel2] Zelditch S 1994 On the rate of quantum ergodicity: I. Upper bounds Commun. Math. Phys. 160 81–92

[Zel3] Zelditch S 2010 Recent developments in mathematical quantum chaos Current Developments in Mathematics (Somerville, MA: International) pp 115–204

[Zw] Zworski M 2012 Semiclassical Analysis (Providence, RI: American Mathematical Society)