MICROLOCAL LIMITS OF PLANE WAVES AND EISENSTEIN FUNCTIONS

SEMYON DYATLOV AND COLIN GUILLARMOU

ABSTRACT. We study microlocal limits of plane waves on noncompact Riemannian manifolds (M,g) which are either Euclidean or asymptotically hyperbolic with curvature -1 near infinity. The plane waves $E(z,\xi)$ are functions on M parametrized by the square root of energy z and the direction of the wave, ξ, interpreted as a point at infinity. If the trapped set K for the geodesic flow has Liouville measure zero, we show that, as $z\to+\infty$, $E(z,\xi)$ microlocally converges to a measure μ_ξ, in average on energy intervals of fixed size, $[z,z+1]$, and in ξ. We express the rate of convergence to the limit in terms of the classical escape rate of the geodesic flow and its maximal expansion rate — when the flow is Axiom A on the trapped set, this yields a negative power of z. As an application, we obtain Weyl type asymptotic expansions for local traces of spectral projectors with a remainder controlled in terms of the classical escape rate.

For a compact Riemannian manifold (M,g) of dimension d whose geodesic flow is ergodic with respect to the Liouville measure μ_L, *quantum ergodicity* (QE) of eigenfunctions [Sh, Ze87, CdV] states that any orthonormal basis $(e_j)_{j\in\mathbb{N}}$ of eigenfunctions of the Laplacian with eigenvalues z_j^2, has a density one subsequence (e_{j_k}) that converges microlocally to μ_L in the following sense: for each symbol $a\in C^\infty(T^*M)$ of order zero,

$$\langle \mathrm{Op}_{h_{j_k}}(a)e_{j_k},e_{j_k}\rangle_{L^2(M)}\to\frac{1}{\mu_L(S^*M)}\int_{S^*M}a\,d\mu_L. \tag{1.1}$$

Here S^*M stands for the unit cotangent bundle, $\mathrm{Op}_h(a)$ denotes the pseudodifferential operator obtained by quantizing a (see Section 3.1), and we put $h_j=z_j^{-1}$. The proof uses the following integrated form of quantum ergodicity [HeMaRo]:

$$h^{d-1}\sum_{h^{-1}\leq z_j\leq h^{-1}+1}\Big|\langle \mathrm{Op}_h(a)e_j,e_j\rangle_{L^2}-\frac{1}{\mu_L(S^*M)}\int_{S^*M}a\,d\mu_L\Big|\to 0\quad\text{ as }h\to 0. \tag{1.2}$$

See Appendix D for a short self-contained proof of this result using the methods of this paper.

In the present paper, we consider a non-compact complete Riemannian manifold (M,g) and show that generalized eigenfunctions of the Laplacian on M known in scattering theory as *distorted plane waves* or *Eisenstein functions*, converge microlocally on average, similarly to (1.2), with the limiting measure μ_ξ depending on the direction of the plane wave ξ – see Theorem 1. We also give estimates on the rate of convergence in terms of classical quantities defined from the geodesic flow on M – see Theorem 2.

Our microlocal convergence of plane waves is similar in spirit to the QE results (1.1) and (1.2). However, unlike the case of QE where *ergodicity* of the geodesic flow is essential, our result is based on a different phenomenon, roughly described as *dispersion* of plane waves.

This difference manifests itself in the proofs as follows: instead of averaging an observable along the geodesic flow as in the standard proof of quantum ergodicity, we propagate it. See Section 2 for an outline of the proofs of Theorems 1 and 2.

Geometric assumptions near infinity. The manifold M has dimension $d = n + 1$. For our results to hold, we need to make several assumptions on the geometry of (M, g) near infinity and on the spectral decomposition of its Laplacian Δ. They are listed in Section 4 and we check in Sections 6 and 7 that they are satisfied in each of the following two cases:

(1) there exists a compact set $K_0 \subset M$ such that $(M \setminus g_0, K_0)$ is isometric to $\mathbb{R}^{n+1} \setminus B(0, R_0)$ with the Euclidean metric for some $R_0 > 0$; here $B(0, R_0)$ denotes the ball centered at 0 of radius R_0,

(2) (M, g) is an *asymptotically hyperbolic* manifold in the sense that it admits a smooth compactification $\overline{M}$ and there exists a smooth boundary defining function x such that in a collar neighbourhood of the boundary $\partial\overline{M}$, the metric has the form

$$g = \frac{dx^2 + h(x)}{x^2}. \tag{1.3}$$

where $h(x)$ is a smooth 1-parameter family of metrics on $\partial\overline{M}$ for $x \in [0, \varepsilon)$. We further assume that g has sectional curvature -1 in a neighbourhood of $\partial\overline{M}$.

In case (1), we call (M, g) *Euclidean near infinity*, while in case (2), we call it *hyperbolic near infinity*. Case (2) in particular includes convex co-compact hyperbolic quotients $\Gamma\backslash\mathbb{H}^{n+1}$ – see Appendix A. Other possible geometries are discussed in Section 2.1.

Distorted plane waves/Eisenstein functions. Let Δ be the (nonnegative) Laplace–Beltrami operator on M. In the study of the relation between classical dynamics and high energy behavior it is natural to use the semiclassically rescaled operator $h^2\Delta$, with $h > 0$ small parameter tending to zero.

The operator $h^2\Delta$ has continuous spectrum on a half-line $[c_0 h^2, \infty)$ (here c_0 is 0 for the Euclidean and $n^2/4$ for the hyperbolic case), parametrized by *distorted plane waves* (or *Eisenstein functions* in the hyperbolic case) $E_h(\lambda, \xi) \in C^\infty(M)$, satisfying for $\lambda \in \mathbb{R}$,

$$(h^2\Delta - \lambda^2 - c_0 h^2)E_h(\lambda, \xi) = 0. \tag{1.4}$$

Here ξ lies on the boundary $\partial\overline{M}$ of a compactification $\overline{M}$ of M. We can think of an element of $\partial\overline{M}$ as the direction of escape to infinity for a non-trapped geodesic; then ξ is the direction of the outgoing part of the plane wave $E_h(\lambda, \xi)$ at infinity.

For instance, in the case of manifolds Euclidean near infinity, $c_0 = 0$, $\partial\overline{M} = \mathbb{S}^n$ is the sphere, and for m near infinity,

$$E_h(\lambda, \xi; m) = e^{\frac{i\lambda}{h}\xi\cdot m} + E_{\mathrm{inc}},$$

where E_{inc} is incoming in the sense that it satisfies a Sommerfeld radiation condition, or equivalently, that it lies in the image of $C_0^\infty(\mathbb{R}^{n+1})$ under the free (incoming) resolvent $R_0(\lambda/h)$ of the Laplacian on the Euclidean space $\mathbb{R}^{n+1}$. These conditions provide a unique characterization of $E_h(\lambda, \xi)$. We can also write $E_h(\lambda, \xi) = E(\lambda/h, \xi)$, where $E(z, \xi)$ is the nonsemiclassical plane wave, and rewrite the results below in terms of the parameter z, as in the abstract.

We will freely use the notions of semiclassical analysis as found for example in [Zw], and reviewed in Section 3. We denote elements of the cotangent bundle T^*M by (m,ν), where $m\in M$ and $\nu\in T^*_mM$. The semiclassical principal symbol of $h^2\Delta$ is equal to $p(m,\nu)=|\nu|_g^2$, where $|\nu|_g$ is the length of $\nu\in T^*_mM$ with respect to the metric g. Therefore, the plane wave E_h should be concentrated on the unit cotangent bundle (see [Zw, Theorem 5.3])

$$S^*M:=\{(m,\nu)\in T^*M\mid |\nu|_g=1\}.$$

If $g^t:T^*M\to T^*M$ denotes the geodesic flow, then the Hamiltonian flow of p is $e^{tH_p}=g^{2t}$.

Semiclassical limits of E_h when the trapped set has measure zero. In scattering theory trajectories which never escape to infinity play a special role as they can be observed only indirectly in asymptotics of plane waves. The *incoming tail* (resp. *outgoing tail*) $\Gamma_-\subset S^*M$ (resp. $\Gamma_+\subset S^*M$) of the flow is defined as follows: a point (m,ν) lies in Γ_- (resp. Γ_+) if and only if the geodesic $g^t(m,\nu)$ stays in some compact set for $t\geq 0$ (resp. $t\leq 0$). The *trapped set* $K:=\Gamma_+\cap\Gamma_-$ is the set of points (m,ν) such that the geodesic $g^t(m,\nu)$ lies entirely in some compact subset of S^*M.

Our first result states that if $\mu_L(K)=0$, then plane waves $E_h(\lambda,\xi)$ converge on average to some measures supported on the closure of the set of trajectories converging to ξ in $\overline{M}$:

Theorem 1. *Let (M,g) be a Riemannian manifold satisfying the assumptions of Section 4 and suppose that the trapped set has Liouville measure $\mu_L(K)=0$. For Lebesgue almost every $\xi\in\partial\overline{M}$, there exists a Radon measure μ_ξ on S^*M such that for each compactly supported h-semiclassical pseudodifferential operator $A\in\Psi^0(M)$, we have as $h\to 0$,*

$$h^{-1}\Big\|\langle AE_h(\lambda,\xi),E_h(\lambda,\xi)\rangle_{L^2(M)}-\int_{S^*M}\sigma(A)\,d\mu_\xi\Big\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])}\to 0. \tag{1.5}$$

The measure μ_ξ has support

$$\operatorname{supp}(\mu_\xi)\subset\overline{\{(m,\nu)\in S^*M\mid \lim_{t\to+\infty}g^t(m,\nu)=\xi\}}, \tag{1.6}$$

and disintegrates the Liouville measure in the sense that there exists a smooth measure $d\xi$ on $\partial\overline{M}$ such that, if μ_L is the Liouville measure generated by $\sqrt{p}=|\nu|_g$ on S^*M, then

$$\int_{\partial\overline{M}}\mu_\xi\,d\xi=\mu_L. \tag{1.7}$$

The limiting measure μ_ξ is defined in Section 4.3. Implicit in (1.7) is the statement that for any bounded Borel $S\subset S^*M$, we have $\mu_\xi(S)\in L^1_\xi(\partial\overline{M})$.

In the case when $\mathrm{WF}_h(A)\cap\Gamma_-=\emptyset$ (in particular when g is non-trapping), we actually have a full expansion of $\langle AE_h,E_h\rangle$ in powers of h, with remainders bounded in $L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])$ – see (5.14).

The now standard argument of Colin de Verdière and Zelditch (see for example the proof of [Zw, Theorem 15.5]) shows that there exists a family of Borel sets $\mathcal{A}(h)\subset\partial\overline{M}\times[1,1+h]$

such that the ratio of the measure of $\mathcal{A}(h)$ to the measure of the whole $\partial\overline{M}\times[1,1+h]$ goes to 1 as $h\to 0$, and for each $A\in\Psi^0(M)$ as in Theorem 1 with $\sigma(A)$ independent of h,

$$\langle AE_h(\lambda,\xi),E_h(\lambda,\xi)\rangle_{L^2(M)}\to\int_{S^*M}\sigma(A)\,d\mu_\xi \text{ uniformly in } (\lambda,\xi)\in\mathcal{A}(h). \tag{1.8}$$

This statement can be viewed as an analogue of the quantum ergodicity fact (1.1), though as explained above, it is produced by a different phenomenon.

Estimates for the remainder. We next provide a quantitative version of Theorem 1, namely an estimate of the left-hand side of (1.5). We define the set $\mathcal{T}(t)$ of geodesics trapped for time $t>0$ as follows: let K_0 be a compact geodesically convex subset of M containing a neighborhood of the trapped set K, then (see also Section 5.2)

$$\mathcal{T}(t):=\{(m,\nu)\in S^*M\mid m\in K_0,\ g^t(m,\nu)\in K_0\}. \tag{1.9}$$

A quantity which will appear frequently with some parameter $\Lambda>0$ is the following interpolated measure

$$r(h,\Lambda):=\sup_{0\leq\theta\leq 1}h^{1-\theta}\mu_L\big(\mathcal{T}\big(\theta\Lambda^{-1}|\log h|\big)\big), \tag{1.10}$$

where $h>0$ is small. This converges to 0 as $h\to 0$ when $\mu_L(K)=0$ and it interpolates between h (when $\theta=0$) and the Liouville measure of the set of geodesics that remain trapped for time $\Lambda^{-1}|\log h|$ (when $\theta=1$). When the measure $\mu_L(\mathcal{T}(t))$ decays exponentially in t, as in (1.14), $r(h,\Lambda)$ can be replaced by simply $\mathcal{O}(h)+\mu_L(\mathcal{T}(\Lambda^{-1}|\log h|))$. The $\mathcal{O}(h)$ term here is natural because of the influence of the subprincipal part of the operator A.

We next define the *maximal expansion rate* as follows (see also (3.17)):

$$\Lambda_{\max}:=\limsup_{|t|\to+\infty}\frac{1}{|t|}\log\sup_{(m,\nu)\in\mathcal{T}(t)}\|dg^t(m,\nu)\|. \tag{1.11}$$

We can estimate the left-hand side of (1.5) in terms of the (interpolated) measure of the set of all trajectories trapped for the Ehrenfest time. If we pair with a test function in ξ instead of taking the L^1_ξ norm, then the estimate becomes stronger, corresponding to the set of all trajectories trapped for *twice* the Ehrenfest time:

Theorem 2. *Let (M,g) be as in Theorem 1. Take $\Lambda_0>\Lambda_{\max}$. Then for each compactly supported h-semiclassical pseudodifferential operator $A\in\Psi^0(M)$ and for each $f\in C^\infty(\partial\overline{M})$,*

$$h^{-1}\Big\|\langle AE_h,E_h\rangle-\int_{S^*M}\sigma(A)\,d\mu_\xi\Big\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])}=\mathcal{O}(r(h,2\Lambda_0)), \tag{1.12}$$

$$h^{-1}\Big\|\int_{\partial\overline{M}}f(\xi)\Big(\langle AE_h,E_h\rangle-\int_{S^*M}\sigma(A)\,d\mu_\xi\Big)\,d\xi\Big\|_{L^1_\lambda([1,1+h])}=\mathcal{O}(r(h,\Lambda_0)). \tag{1.13}$$

The proof of Theorem 2 actually gives an expansion of $\langle AE_h,E_h\rangle$ in powers of h, with remainder $\mu_L(\mathcal{T}(\Lambda^{-1}|\log h|))$ instead of $r(h,\Lambda)$ – see (5.35) and the proofs of Propositions 5.11 and 5.13. This full expansion is cumbersome to write down, therefore we only do it for the trace estimates (1.16) below.

Remainder in terms of pressure. When the trapped set K has Liouville measure 0 and is uniformly partially hyperbolic in the sense of Appendix B.1, we estimate using [Yo]

$$\mu_L(\mathcal{T}(t)) = \mathcal{O}(e^{t(P(J^u)+\varepsilon)}), \tag{1.14}$$

for each $\varepsilon > 0$, where $P(J^u) \leq 0$ is the topological pressure of the unstable Jacobian – see Appendix B.1. When K is a hyperbolic basic set (Axiom A flow), then $P(J^u) < 0$ by [BoRu], and the remainders in (1.12) and (1.13) are then polynomial in h:

$$r(h,\Lambda) = \mathcal{O}(h + h^{-(P(J^u)+\varepsilon)/\Lambda}),$$

and one can get rid of ε here by slightly changing Λ_0. In the special case where g has constant sectional curvature -1 near K, the bounds in (1.12) and (1.13) become $\mathcal{O}(h + h^{(n-\delta)/2-})$ and $\mathcal{O}(h + h^{n-\delta-})$, respectively, where K has Hausdorff dimension $\dim_H(K) = 2\delta + 1$. See Appendix B.2 for details.

In all cases, if K is nonempty, then it has Minkowski dimension at least 1; since $g^{t/2}(\mathcal{T}(t))$ contains an $e^{-\Lambda_0 t/2}$ sized neighborhood of K, we have

$$\mu_L\big(\mathcal{T}((2\Lambda_0)^{-1}|\log h|)\big) \gtrsim h^{n/2},\ \mu_L\big(\mathcal{T}(\Lambda_0^{-1}|\log h|)\big) \gtrsim h^n. \tag{1.15}$$

Local Weyl asymptotics for spectral projectors. It is possible to express the spectral measure of $h^2\Delta$ in terms of the distorted plane waves (see (4.5)), and using (1.13), we obtain an expansion in powers of h for local traces of spectral projectors up to an explicit remainder. We only write it here for the case where the flow is partially uniformly hyperbolic with $P(J^u) < 0$, but a more general result with the Liouville measure of $\mathcal{T}(\Lambda_0^{-1}|\log h|)$ holds – see Theorem 4 in Section 5.3. Below, we fix a quantization procedure Op_h on M mapping compactly supported symbols to compactly supported operators.

Theorem 3. *Let (M,g) be as in Theorem 1, let $\Lambda_0 > \Lambda_{\max}$ and assume that the trapped set K is uniformly partially hyperbolic with $\mu_L(K) = 0$ and that the topological pressure $P(J^u)$ of the unstable Jacobian on K is negative. Then there exist differential operators[1] L_j of order $2j$ on T^*M, with $L_0 = 1$, such that for each compactly supported zeroth order classical symbol a, we have for each $s > 0$ and $N \in \mathbb{N}$*

$$\operatorname{Tr}(\operatorname{Op}_h(a)\,\mathbb{1}_{[0,s]}(h^2\Delta)) = (2\pi h)^{-n-1}\sum_{j=0}^N h^j \int_{|\nu|_g^2\leq s} L_j a\, d\mu_\omega + h^{-n}\mathcal{O}\big(h^{-\frac{P(J^u)}{\Lambda_0}} + h^N\big) \tag{1.16}$$

where μ_ω is the standard volume form on T^*M and $\mathbb{1}_{[0,s]}(h^2\Delta)$ denotes the spectral projector of $h^2\Delta$ onto the frequency window $[0,s]$. The remainder is uniform in s when s varies in a compact subset of $(0,\infty)$.

In particular, if g has constant sectional curvature -1 near K and the Hausdorff dimension of K is given by $2\delta + 1$, then the remainder in (1.16) becomes $\mathcal{O}(h^{-\delta-})$, for N large enough.

Applications. In a separate paper [DyGu], we show that Theorem 3 implies new asymptotics for the spectral shift function (or scattering phase) with remainders in terms of $P(J^u)$

[1] In this paper, the symbols L_j will denote different operators in different propositions.

when the trapped set has Liouville measure 0 and the manifold is Euclidean near infinity with uniformly partially hyperbolic geodesic flow near K.

Previous works. Let us briefly discuss the history of Quantum Ergodicity (QE) and explain its relation to the present paper. The original QE statement was proved by Shnirelman [Sh], Zelditch [Ze87], and Colin de Verdière [CdV] in the microlocal case, by Helffer–Martinez–Robert [HeMaRo] in the semiclassical case (with the integrated estimate using an $\mathcal{O}(h)$ spectral window like in the present paper, rather than the $\mathcal{O}(1)$ window used in the microlocal case), and by Gérard–Leichtnam [GéLe] and Zelditch–Zworski [ZeZw] for manifolds with boundary (ergodic billiards). Quantum ergodicity for boundary values and restrictions of eigenfunctions to hypersurfaces was studied by Hassell–Zelditch [HaZe], Burq [Bu05], Toth–Zelditch [ToZe10, ToZe11], and by Dyatlov–Zworski [DyZw].

The first result on noncompact manifolds, namely for embedded eigenvalues and Eisenstein functions on surfaces with cusps, was proved by Zelditch [Ze91]. For the special case of arithmetic hyperbolic surfaces, a stronger statement of Quantum Unique Ergodicity (QUE), saying that the whole sequence of eigenstates microlocally converges to the Liouville measure, was proved by Lindenstrauss [Li] and Soundararajan [So] for Hecke–Maass forms and by Luo–Sarnak [LuSa] and Jakobson [Ja] for Eisenstein functions. For further information on the topic, the reader is directed to the recent reviews [No, Sa, Ze09].

As remarked above, our result differs from the above works in that it uses dispersion of plane waves instead of the ergodicity of the geodesic flow. This dispersion phenomenon was used to study microlocal limits of plane waves on convex co-compact hyperbolic quotients satisfying $\delta < n/2$ by Guillarmou–Naud in [GuNa], and on surfaces with cusps at complex energies by Dyatlov [Dy2]. Both [GuNa] and [Dy2] guarantee microlocal convergence of the Eisenstein functions that is uniform in λ and ξ, rather than the (weaker) $L^1_{\lambda,\xi}$ estimates of the current paper; these statements are formally similar to QUE, while our statement is formally similar to QE. In [GuNa], uniform in λ and ξ estimates are possible because Lagrangian states, when propagated by the Schrödinger group $U(t)$, would disperse faster than they fail to be approximated semiclassically, a phenomenon similar to the one studied by Nonnenmacher–Zworski [NoZw]. In fact, it is plausible that the result of [GuNa] is true when the condition $\delta < n/2$ is replaced by the negative pressure condition of [NoZw]. As for [Dy2], the energy being away from the real line makes the measure corresponding to E_h exponentially increasing, rather than invariant, along the flow; then the result of propagation of E_h^0 by $U(t)$ is multiplied by $e^{-\nu t}$ for a certain $\nu > 0$, and decays in L^2 as $t \to +\infty$ (it is then more correct to say that this paper relies on damping of plane waves rather than dispersion).

We see that the uniform convergence in [GuNa] and [Dy2] is possible because one has better control on the propagated Lagrangian states. Such better control is directly related to having a polynomial bound on the scattering resolvent. In the less restricted situation of our paper, however, it is not clear if such a bound would hold; therefore, we need to average in λ and ξ to pass to trace (or, strictly speaking, Hilbert–Schmidt norm) estimates, just as in the proof of Quantum Ergodicity.

The expansions for local traces of the spectral measure as in Theorem 3 were studied by Robert–Tamura [RoTa] for nontrapping perturbations of the Euclidean space, yielding a full expansion in powers of h in that setting.

2. Outline of the proofs

In this section, we explain the ideas of the proofs of Theorems 1 and 2, in the case of manifolds Euclidean near infinity. We also describe the structure of the paper.

We start with Theorem 1. Take $t > 0$; we will use $\lim_{t\to+\infty}\lim_{h\to 0}$ limits, therefore remainders that decay in h with constants depending on t will be negligible. Since E_h is a generalized eigenfunction of the Laplacian (1.4), we have

$$E_h(\lambda,\xi) = e^{-it\lambda^2/(2h)}U(t)E_h(\lambda,\xi). \tag{2.1}$$

Here $U(t) = e^{ith\Delta/2}$ is the semiclassical Schrödinger propagator, quantizing the geodesic flow g^t. Since E_h does not lie in $L^2(M)$, we cannot apply the operator $U(t)$ to it; however, (2.1) can be made rigorous, with an $\mathcal{O}(h^\infty)$ error, by using appropriate cutoffs – see Lemma 3.10. We will not write these cutoffs here for the sake of brevity.

Take a compactly supported and compactly microlocalized semiclassical pseudodifferential operator A on M; then by (2.1),

$$\langle AE_h, E_h\rangle = \langle AU(t)E_h, U(t)E_h\rangle = \langle A^{-t}E_h, E_h\rangle, \tag{2.2}$$

where $A^{-t} := U(-t)AU(t)$ is a pseudodifferential operator with principal symbol $\sigma(A)\circ g^{-t}$. (It is not compactly supported, but we ignore this issue here.) We now use the following decomposition of plane waves:

$$E_h = \chi_0 E_h^0 + E_h^1,\ E_h^0 = e^{\frac{i\lambda}{h}\xi\cdot m},\ E_h^1 = -R_h(\lambda)F_h,\ F_h := (h^2\Delta - \lambda^2)\chi_0 E_h^0.$$

Here E_h^0 is the outgoing part of the plane wave, defined in a certain neighborhood of infinity and solving (1.4) there, while χ_0 is a cutoff function equal to 1 near infinity and supported inside the domain of E_h^0; then

$$F_h = [h^2\Delta, \chi_0]E_h^0$$

is compactly supported and we can apply to it the semiclassical scattering resolvent $R_h(\lambda)$. Note that here $R_h(\lambda)$ is the incoming resolvent; in particular, it is bounded $L^2 \to L^2$ for $\operatorname{Im}\lambda < 0$. (The situation is more complicated in the case of manifolds hyperbolic near infinity, in particular the domain of E_h^0 and the cutoff χ_0 will depend on ξ.) For $\lambda = 1 + \mathcal{O}(h)$, the function F_h is microlocalized inside the set

$$W_\xi := \{(m,\nu) \mid m \in \operatorname{supp}(d\chi_0),\ \nu = \xi\} \subset S^*M.$$

In general, we cannot expect the resolvent $R_h(\lambda)$ to be polynomially bounded in h, and thus cannot determine the wavefront set of E_h^1. However, we will show the following weaker propagation of singularities statement: the function

$$\widetilde{E}_h^1(\lambda,\xi) := \frac{E_h^1(\lambda,\xi)}{1 + \|E_h(\lambda,\xi)\|_{L^2(K_0)}},$$

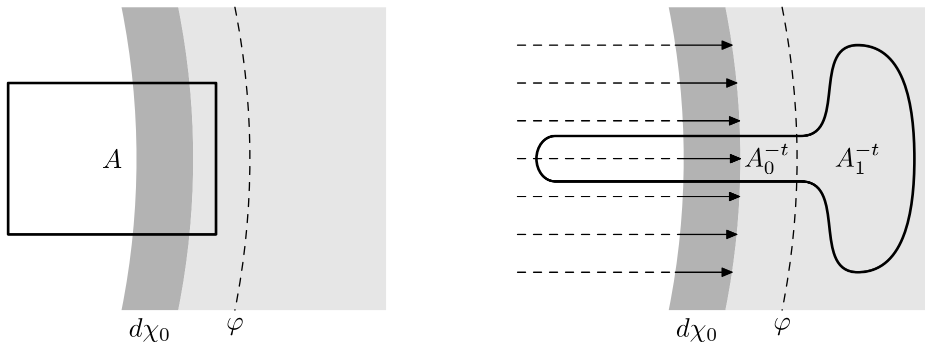

FIGURE 1. A phase space picture of the main argument. The right side of each picture represents infinity; $\chi_0 = 1$ in the lighter shaded region and $d\chi_0$ is supported in the darker shaded region, while $\varphi = 1$ to the left of the vertical dashed line. The horizontal dashed lines on the right represent the wavefront set of $\widetilde{E}_h^1$; they terminate at the solid arrows, which denote the set W_ξ.

where $K_0 \subset M$ is a sufficiently large compact set, is polynomially bounded in h and for each $(m, \nu) \in \mathrm{WF}_h(\widetilde{E}_h^1)$, the geodesic $g^t(m, \nu)$ is either trapped as $t \to +\infty$ or passes through W_ξ for some $t \geq 0$. For the case of manifolds Euclidean near infinity, this statement follows directly from the explicit formula for the scattering resolvent on the free Euclidean space; for manifolds hyperbolic near infinity, we use the microlocal properties of the resolvent established in [Va11]. See assumption (A6) in Section 4.2, Section 6.2, and Proposition 7.4 for details.

If A and $1-\chi_0$ are both supported in the ball of radius R, let $\varphi \in C_0^\infty(M)$ be independent of t and equal to 1 in the ball of radius $R+1$. Then we write

$$A^{-t} = A_0^{-t} + A_1^{-t},\ A_0^{-t} := A^{-t}\varphi,\ A_1^{-t} := A^{-t}(1-\varphi).$$

Now, each $(m, \nu) \in \mathrm{WF}_h(A_1^{-t})$ has the following properties: $|m| \geq R+1$, and for $(m', \nu') = g^{-t}(m, \nu)$, $|m'| \leq R$. (See Figure 1.) Therefore, the geodesic $g^s(m, \nu)$ escapes to infinity for $s \geq 0$ and never passes through W_ξ; it follows from the discussion of the wavefront set of $\widetilde{E}_h^1$ in the previous paragraph that

$$\|A_1^{-t} E_h^1\|_{L^2} = \mathcal{O}(h^\infty(1 + \|E_h\|_{L^2(K_0)})).$$

Therefore, we can write

$$\langle AE_h, E_h\rangle = \langle A_1^{-t}\chi_0 E_h^0, \chi_0 E_h^0\rangle + \langle A_0^{-t} E_h, E_h\rangle + \mathcal{O}(h^\infty(1 + \|E_h\|^2_{L^2(K_0)})). \tag{2.3}$$

The first term on the right-hand side is explicit, as we have a formula for E_h^0; we can calculate for Lebesgue almost every ξ and $\lambda = 1 + \mathcal{O}(h)$,

$$\lim_{t\to+\infty}\lim_{h\to 0}\langle A_1^{-t}\chi_0 E_h^0(\lambda, \xi), \chi_0 E_h^0(\lambda, \xi)\rangle = \int_{S^*M} a\, d\mu_\xi. \tag{2.4}$$

It then remains to estimate the second and third terms on average in λ and ξ. For this, we use the relation (4.5) of distorted plane waves to the spectral measure of the Laplacian to get for any bounded compactly supported pseudodifferential operator B,

$$h^{-1}\|BE_h(\lambda,\xi)\|^2_{L^2_{m,\xi,\lambda}(M\times\partial\overline{M}\times[1,1+h])}\leq Ch^n\|B\,\mathbb{1}_{[1,(1+h)^2]}(h^2\Delta)\|^2_{\mathrm{HS}}. \tag{2.5}$$

Here HS denotes the Hilbert–Schmidt norm. One can estimate the right-hand side of (2.5) uniformly in h – see Lemma 3.11 and the proof of Proposition 4.5. Then $h^{-1}\|E_h\|^2_{L^2(K_0)}$, when integrated over $\lambda\in[1,1+h]$ and ξ, is bounded uniformly in h; this removes the third term on the right-hand side of (2.3).

Finally, the average in λ,ξ of the second term on the right-hand side of (2.3) can be bounded, modulo an $\mathcal{O}_t(h)$ remainder, by the L^2 norm $\|\sigma(A_0^{-t})\|_{L^2(S^*M)}$ of the restriction of the principal symbol of A_0^{-t} to the energy surface S^*M, with respect to the Liouville measure. Now, $\sigma(A_0^{-t})=(\sigma(A)\circ g^{-t})\varphi$ converges to zero as $t\to+\infty$ at any point which is not trapped in the backwards direction. Since we assumed $\mu_L(K)=0$, by the dominated convergence theorem $\|\sigma(A_0^{-t})\|_{L^2(S^*M)}$ converges to zero as $t\to+\infty$; this finishes the proof of Theorem 1.

For the estimate (1.12) in Theorem 2, we need to take t up to the Ehrenfest time:

$$t=t_e:=\Lambda_0^{-1}\log(1/h)/2,$$

replacing the $\lim_{t\to+\infty}\lim_{h\to0}$ limit in the argument of Theorem 1 by just the $\lim_{h\to0}$ limit, but with t depending on h. The operator A^{-t} is then still pseudodifferential, though in a mildly exotic class. To avoid a quantization procedure uniform at infinity, we give an iterative argument, propagating A for a fixed time for $\sim\log(1/h)$ steps, applying t-independent cutoffs and removing the microlocally negligible terms at each step. The proof then works as before, with the term $\langle A_0^{-t}E_h,E_h\rangle$ bounded by the Liouville measure of the support of the full symbol of A_0^{-t}, which depends on h and is contained in $g^t(\mathcal{T}(t))$, where $\mathcal{T}(t)$ is defined in (1.9); this proves (1.12). The interpolated quantity $r(h,\Lambda)$ from (1.10) appears because of the subprincipal terms in (2.4).

For (1.13), we have to propagate to twice the Ehrenfest time: $t=2t_e$. The operator A^{-t} is not pseudodifferential, but we can use (2.1) to write

$$\langle A_0^{-t}E_h,E_h\rangle=\langle U(-t/2)AU(t/2)\cdot U(t/2)\varphi U(-t/2)E_h,E_h\rangle. \tag{2.6}$$

The operators $U(-t/2)AU(t/2)$ and $U(t/2)\varphi U(-t/2)$ are both pseudodifferential in a mildly exotic class; multiplying them, we get a pseudodifferential operator whose full symbol is supported inside $g^{t/2}(\mathcal{T}(t))$, and thus (2.6) can be estimated by the Liouville measure of this set, giving the remainder (1.13).

A problem arises when trying to get a rate of convergence in (2.4) for t up to twice the Ehrenfest time. We are unable to propagate the Lagrangian state $E_h^0(\lambda,\xi)$ pointwise in ξ and λ for time t, therefore we do not get an L^1_ξ estimate in (1.13). However, for $f\in C^\infty(\partial\overline{M})$ we can still approximate the integral

$$\int_{\partial\overline{M}}f(\xi)\langle A_1^{-t}\chi_0E_h^0,\chi_0E_h^0\rangle\,d\xi \tag{2.7}$$

as follows. Define the operator

$$\Pi_f^0(\lambda) := \int_{\partial\overline{M}} f(\xi)(\chi_0 E_h^0(\lambda,\xi)) \otimes (\chi_0 E_h^0(\lambda,\xi))\, d\xi.$$

Here $\otimes$ denotes the Hilbert tensor product; that is, if $u, v \in C^\infty(M)$, then $u \otimes v$ is the operator with the Schwartz kernel

$$K_{u\otimes v}(m, m') = u(m)\overline{v(m')}. \tag{2.8}$$

We can show that for $\widetilde{X} \in \Psi^{\mathrm{comp}}(M)$ satisfying certain conditions, $\widetilde{X}\Pi_f^0\widetilde{X}^*$ is a Fourier integral operator associated to the canonical relation

$$\{(m,\nu; m',\nu') \mid (m,\nu) \in S^*M,\ (m',\nu') = g^s(m,\nu) \text{ for some } s \in (-T_0, T_0)\},$$

for a fixed $T_0 > 0$ depending on $\widetilde{X}$. (For comparison, for the spectral measure of $h^2\Delta$ we would have to formally take all possible values of s, which would destroy any hope on microlocally approximating it when the geodesic flow is chaotic.) We can then write

$$\widetilde{X}\Pi_f^0(\lambda)\widetilde{X}^* = (2\pi h)^n \int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)} U(s) B_s\, ds,$$

where B_s is a smooth family of pseudodifferential operators, compactly supported in $s \in (-T_0, T_0)$ – see Lemma 5.12. We then write the integral (2.7) as

$$\mathrm{Tr}(U(-t)AU(t)(1-\varphi)\Pi_f^0(\lambda)) = \mathrm{Tr} \int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)} U(-t)AU(t)(1-\varphi)U(s)B_s\, ds$$

$$= \mathrm{Tr} \int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)} U(-t/2)AU(t/2) \cdot U(t/2)(1-\varphi)U(s)B_sU(-s-t/2) \cdot U(s)\, ds.$$

The operators $U(-t/2)AU(t/2)$ and $U(t/2)(1-\varphi)U(s)B_sU(-s-t/2)$ are pseudodifferential in a mildly exotic class; thus their product is also pseudodifferential and (bearing in mind that s varies in a bounded set) one gets a microlocal expansion for (2.7) through a local trace formula for Schrödinger propagators – see Lemma 3.12 and Proposition 5.13.

2.1. **Other possible geometric assumptions.** Our results should be true for asymptotically hyperbolic manifolds without the constant curvature assumption near infinity. The main difficulty here is constructing a good semiclassical parametrix for the Eisenstein function $E_h(\lambda, \xi)$ near $\xi \in \partial\overline{M}$; this can be done by WKB approximation, and the phase is a Busemann function $\phi_\xi(m)$ near ξ, however one would need a good understanding of the regularity of $\phi_\xi(m)$ as $m \to \xi$. This is in a way related to the high-frequency parametrix of [MeSBVa] in the non-trapping setting. For the asymptotically Euclidean of asymptotically conic ends, this might be more complicated as we would need a parametrix of $E_h(\lambda, \xi)$ in a large neighbourhood of $\xi \in \partial\overline{M}$, essentially in a region with closure containing a ball of radius $\pi/2$ in $\partial\overline{M}$. In particular, the Lagrangian supporting the semiclassical parametrix of $E_h(\lambda, \xi)$ would not a priori be projectable far from ξ, which would make the construction more technical. We leave these questions for future research.

The convergence result in Theorem 1 should be true in the case where M has a boundary, for instance $M = \mathbb{R}^{n+1} \setminus \Omega$ with Ω a piecewise smooth obstacle. In fact, it should be straightforward to check that the method of proof applies when combined with the idea of [ZeZw], based on the fact that the region in phase space near the boundary where the dynamics is complicated is of Liouville measure 0 (since we assume $\mu_L(K) = 0$). To get a good remainder in that setting would be more involved since one would need to care about the amount of mass of plane waves staying in the regions near the boundary where the dynamics is complicated, as we propagate up to Ehrenfest time. A reasonable case to start with is that of strictly convex obstacles.

2.2. **Structure of the paper.** In Section 3, we review certain notions of semiclassical analysis and derive several technical lemmata; in particular, in Section 3.2, we review the local theory of semiclassical Lagrangian distributions and Fourier integral operators and in Section 3.3 we study microlocal properties of Schrödinger propagators, including the Hilbert–Schmidt norm bound (Lemma 3.11). In Section 4, we formulate the general assumptions on the studied manifolds and derive some immediate corollaries; Section 4.1 contains the geometric assumptions and the definition of the trapped set and Section 4.2 contains the analytic assumptions on distorted plane waves. In Section 4.3 we construct the limiting measures μ_ξ and in Section 4.4 we prove averaged estimates on Eisenstein functions.

In Section 5, we give the proofs of our main theorems. Section 5.1 contains the proof of Theorem 1, Section 5.2 contains the proof of the estimate (1.12) in Theorem 2, while Section 5.3 contains the proof of the estimate (1.13) in Theorem 2. Section 5.3 also contains the Tauberian argument proving an expansion of the local trace of a spectral projector (Theorem 4). Sections 6 and 7 study the Euclidean and hyperbolic near infinity manifolds, respectively, and show that the general assumptions of Section 4 are satisfied in these cases.

Appendix A provides a formula for the limiting measures in the case of a convex co-compact hyperbolic quotient, which generalizes the limiting measure of [GuNa] to the case $\delta \geq n/2$. Appendix B discusses the classical escape rate, in particular explaining (1.14). Appendix C gives a self-contained proof of Egorov's theorem up to the Ehrenfest time (Proposition 3.9). Finally, Appendix D contains a short proof of (a special case of) quantum ergodicity in the semiclassical setting, which is simpler than that of [HeMaRo] because it does not rely on [DuGu, PeRo].

3. Semiclassical preliminaries

In this section, we review the methods of semiclassical analysis needed for our argument. Most of the constructions listed below are standard: pseudodifferential operators, wavefront sets, local theory of Fourier integral operators, and Egorov's theorem. However, Section 3.3 contains the propagation result for generalized eigenfunctions (Lemma 3.10) and a Hilbert–Schmidt norm estimate in an $\mathcal{O}(h)$ spectral window (Lemma 3.11), which the authors were unable to find in previous literature.

We will also need Egorov's theorem up to the Ehrenfest time (Proposition 3.9); while several versions of this fact are available, we could not find a detailed proof for the case of manifolds and when the Ehrenfest time is defined via the maximal expansion rate of the flow. For this reason, and also because we insert cutoffs in between the propagators, we give a proof of Proposition 3.9 in Appendix C.

3.1. **Notation.** In this subsection, we briefly review certain notation used in semiclassical analysis. The reader is referred to [Zw] (especially Chapter 14 on semiclassical calculus on manifolds) or [DiSj] for a detailed introduction to the subject.

The phase space. Let M be a d-dimensional manifold without boundary. We denote points in M by the letter m and elements of the cotangent bundle T^*M by (m,ν), where $\nu\in T^*_mM$. Following [Va11, Section 2], we consider the fiber-radial compactification $\overline{T}^*M$ of T^*M. The boundary of $\overline{T}^*M$, denoted by $\partial\overline{T}^*M$ and called the *fiber infinity* (unlike [Va11], we do not use the notation S^*M for fiber infinity — we reserve it for the unit cotangent bundle $\{|\nu|_g=1\}\subset T^*M$), is associated with the cosphere bundle over M and the interior of $\overline{T}^*M$ is associated with T^*M. Take some smooth inner product on the fibers of T^*M; if $|\nu|$ is the norm of a covector $(m,\nu)\in T^*M$ generated by this inner product and $\langle\nu\rangle=\sqrt{1+|\nu|^2}$, then $\langle\nu\rangle^{-1}$ is a boundary defining function on $\overline{T}^*M$.

We will mostly use compactly microlocalized operators, for which the fiber-radial compactification is not necessary. However, it will come up in the elliptic estimate (Proposition 3.2) and in the proof of the propagation of singularities result for plane waves on asymptotically hyperbolic manifolds (Proposition 7.4).

Symbol classes. For any $k\in\mathbb{R}$ and any $\rho\in[0,1/2)$, we consider the symbol class $S^k_\rho(M)$ defined as follows: a smooth function $a(m,\nu;h)$ on $T^*M\times[0,h_0)$ lies in $S^k_\rho(M)$ if and only if for each compact $K\subset M$ and each multiindices α,β, there exists a constant $C_{\alpha\beta K}$ such that for h small enough,

$$\sup_{m\in K,\ \nu\in T^*_mM}|\partial^\alpha_m\partial^\beta_\nu a(m,\nu;h)|\leq C_{\alpha\beta K}h^{-\rho(|\alpha|+|\beta|)}\langle\nu\rangle^{k-|\beta|}. \tag{3.1}$$

These classes are independent of the choice of coordinates on M. Note that we do not fix the behaviour of the symbols as $m\to\infty$. The important special case is $\rho=0$, which includes the classical symbols studied in [Va11]. The class $S^k_0(M)$, denoted simply by $S^k(M)$, would be sufficient for the convergence Theorem 1. The classes S^k_ρ with $\rho>0$ will be important for obtaining the remainder estimate of Theorem 2; these classes arise when propagating symbols in S^k_0 for short logarithmic times, as in Proposition 3.9.

Since plane waves are microlocalized on the cosphere bundle, away from the fiber infinity, we will most often work with the classes S^{comp}_ρ, consisting of compactly supported functions satisfying (3.1); we have $S^{\mathrm{comp}}_\rho\subset S^k_\rho$ for all k.

Pseudodifferential operators. Following [Zw, Section 14.2], we can define the algebra $\Psi^k_\rho(M)$ of pseudodifferential operators with symbols in $S^k_\rho(M)$. (The properties of the symbol classes S^k_ρ required for the construction of [Zw, Section 14.2] are derived as in [Zw, Section 4.4]; see also [GrSj, Chapter 3].) As before, denote $\Psi^k=\Psi^k_0$. Since our symbols can grow arbitrarily fast as $m\to\infty$, we do not make any a priori assumptions on the behavior of elements of Ψ^k_ρ near the infinity in M. However, we require that all operators $A\in\Psi^k(M)$ be *properly supported*; namely, the restriction of each of the projection maps $\pi_m,\pi_{m'}:M\times M\to M$ to the support of the Schwartz kernel $K_A(m,m')$ of A is a proper map. Then each element of $\Psi^k(M)$ acts $H^s_{h,\mathrm{loc}}(M)\to H^{s-k}_{h,\mathrm{loc}}(M)$, where $H^s_{h,\mathrm{loc}}(M)$ denotes the space of distributions locally in the semiclassical Sobolev space H_h^s (see for example [Zw, Section 7.1] for the definition of semiclassical Sobolev spaces).

We have the semiclassical principal symbol map

$$\sigma : \Psi_\rho^k(M) \to S_\rho^k(M)/h^{1-2\rho}S_\rho^{k-1}(M)$$

and its right inverse, a non-canonical quantization map

$$\mathrm{Op}_h : S_\rho^k(M) \to \Psi_\rho^k(M).$$

The standard operations of pseudodifferential calculus with symbols in S_ρ^k have an $\mathcal{O}(h^{1-2\rho})$ remainder instead of the $\mathcal{O}(h)$ remainder valid for the class S_0^k. More precisely, we have for $A \in \Psi_\rho^k(M)$ and $B \in \Psi_\rho^{k'}(M)$,

$$\begin{aligned}\sigma(A^*) &= \overline{\sigma(A)} + \mathcal{O}(h^{1-2\rho})_{S_\rho^{k-1}(M)},\\ \sigma(AB) &= \sigma(A)\sigma(B) + \mathcal{O}(h^{1-2\rho})_{S_\rho^{k+k'-1}(M)},\\ \sigma([A,B]) &= -ih\{\sigma(A),\sigma(B)\} + \mathcal{O}(h^{2(1-2\rho)})_{S_\rho^{k+k'-2}(M)}.\end{aligned}$$

Here $\{\cdot,\cdot\}$ stands for the Poisson bracket. The $\mathcal{O}(\cdot)$ notation is used in the present paper in the following way: we write $u = \mathcal{O}_z(F)_{\mathcal{X}}$ if the norm of the function, or the operator, u in the functional space $\mathcal{X}$ is bounded by the expression F times a constant depending on the parameter z.

Wavefront sets. If $A : C^\infty(M) \to C^\infty(M)$ is a properly supported operator, we say that $A = \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ if A is smoothing and each of the $C^\infty(M \times M)$ seminorms of its Schwartz kernel is $\mathcal{O}(h^\infty)$. For each $A \in \Psi_\rho^k(M)$, we have $A = \mathrm{Op}_h(a) + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ for some $a \in S_\rho^k(M)$. Define the semiclassical wavefront set $\mathrm{WF}_h(A) \subset \overline{T}^*M$ of A as follows: a point $(m,\nu) \in \overline{T}^*M$ does not lie in $\mathrm{WF}_h(A)$, if there exists a neighborhood U of (m,ν) in $\overline{T}^*M$ such that each (m,ν)-derivative of a is $\mathcal{O}(h^\infty\langle\nu\rangle^{-\infty})$ in $U \cap T^*M$. The notion of the wavefront set does not depend on the choice of the quantization procedure. We have $\mathrm{WF}_h(A) = \emptyset$ if and only if $A = \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ and $\mathrm{WF}_h(A^*) = \mathrm{WF}_h(A)$, $\mathrm{WF}_h(AB) \subset \mathrm{WF}_h(A) \cap \mathrm{WF}_h(B)$. For $A, B \in \Psi_\rho^k(M)$, we say that $A = B$ microlocally in some open set $U \subset \overline{T}^*M$, if $\mathrm{WF}_h(A-B) \cap U = \emptyset$.

Operators with compact wavefront sets are called *compactly microlocalized*; those are exactly operators of the form $\mathrm{Op}_h(a) + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ for some $a \in S_\rho^{\mathrm{comp}}$. We denote by $\Psi_\rho^{\mathrm{comp}}(M)$ the class of all compactly microlocalized elements of $\Psi_\rho^k(M)$; as before, we put $\Psi^{\mathrm{comp}}(M) = \Psi_0^{\mathrm{comp}}(M)$. Compactly microlocalized operators should not be confused with *compactly supported* operators; that is, operators whose Schwartz kernels are compactly supported. That being said, most operators that we use will be both compactly supported and compactly microlocalized.

We will need a finer notion of microsupport on h-dependent sets, used in the proofs in Sections 5.2 and 5.3, for example in Proposition 5.9:

Definition 3.1. *An operator $A \in \Psi_\rho^{\mathrm{comp}}(M)$ is said to be* microsupported *on an h-dependent family of sets $V(h) \subset T^*M$, if we can write $A = \mathrm{Op}_h(a) + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$, where for each compact* *set $K \subset T^*M$, each differential operator ∂^α on T^*M, and each N, there exists a constant $C_{\alpha N}$ such that for h small enough,*

$$\sup_{(m,\nu)\in K\setminus V(h)} |\partial^\alpha a(m,\nu;h)| \leq C_{\alpha N} h^N.$$

Since the change of variables formula for the full symbol of a pseudodifferential operator contains an asymptotic expansion in powers of h, consisting of derivatives of the original symbol, Definition 3.1 does not depend on the choice of the quantization procedure Op_h. Moreover, if $A \in \Psi_\rho^{\mathrm{comp}}$ is microsupported inside some $V(h)$ and $B \in \Psi_\rho^k$, then AB, BA, and A^* are also microsupported inside $V(h)$. It follows from the definition of the wavefront set that $(m,\nu) \in T^*M$ does not lie in $\mathrm{WF}_h(A)$ for some $A \in \Psi_\rho^{\mathrm{comp}}$, if and only if there exists an h-independent neighborhood U of (m,ν) such that A is microsupported on the complement of U. Note however that A need not be microsupported on $\mathrm{WF}_h(A)$, though it will be microsupported on any h-independent neighborhood of $\mathrm{WF}_h(A)$. Finally, it can be seen by Taylor's formula that if $A \in \Psi_\rho^{\mathrm{comp}}(M)$ is microsupported in $V(h)$ and $\rho' > \rho$, then A is also microsupported on the set of all points in $V(h)$ which are at least $h^{\rho'}$ away from the complement of $V(h)$.

Ellipticity. For $A \in \Psi_\rho^k(M)$, define its *elliptic set* $\mathrm{ell}(A) \subset \overline{T}^*M$ as follows: $(m,\nu) \in \mathrm{ell}(A)$ if and only if there exists a neighborhood U of (m,ν) in $\overline{T}^*M$ and a constant C such that $|\sigma(A)| \geq C^{-1}\langle\nu\rangle^k$ in $U \cap T^*M$. The following statement is the standard semiclassical elliptic estimate; see [HöIII, Theorem 18.1.24'] for the closely related microlocal case and for example [Dy1, Section 2.2] for the semiclassical case.

Proposition 3.2. *Assume that $P \in \Psi_\rho^k(M)$, $A \in \Psi_\rho^{k'}(M)$, and $\mathrm{WF}_h(A) \subset \mathrm{ell}(P)$. Assume moreover that A is compactly supported. Then there exists a constant C and a function $\chi \in C_0^\infty(M)$ such that for each $s \in \mathbb{R}$, each $u \in H^s_{h,\mathrm{loc}}(M)$ and each N, we have*

$$\|Au\|_{H_h^s} \leq C\|\chi P u\|_{H_h^{s+k'-k}} + \mathcal{O}(h^\infty)\|\chi u\|_{H^{-N}}.$$

Moreover, if P is a differential operator, then we can take any χ such that the Schwartz kernel of A is supported in $\{\chi \neq 0\} \times \{\chi \neq 0\}$.

Semi-classical wave-front sets of distributions. An h-dependent family $u(h) \in \mathcal{D}'(M)$ is called *h-tempered*, if for each open U compactly contained in M, there exist constants C and N such that

$$\|u(h)\|_{H_h^{-N}(U)} \leq Ch^{-N}. \tag{3.2}$$

For a tempered distribution u, we say that $(m_0,\nu_0) \in \overline{T}^*M$ does not lie in the wavefront set $\mathrm{WF}_h(u)$, if there exists a neighborhood $V(m_0,\nu_0)$ in $\overline{T}^*M$ such that for each $A \in \Psi^0(M)$ with $\mathrm{WF}_h(A) \subset V$, we have $Au = \mathcal{O}(h^\infty)_{C^\infty}$. By Proposition 3.2, $(m_0,\nu_0) \notin \mathrm{WF}_h(u)$ if and only if there exists compactly supported $A \in \Psi^0(M)$ elliptic at (m_0,ν_0) such that $Au = \mathcal{O}(h^\infty)_{C^\infty}$. The wavefront set of u is a closed subset of $\overline{T}^*M$; it is empty if and only if $u = \mathcal{O}(h^\infty)_{C^\infty(M)}$. We can also verify that for u tempered and $A \in \Psi_\rho^k(M)$, $\mathrm{WF}_h(Au) \subset \mathrm{WF}_h(A) \cap \mathrm{WF}_h(u)$.

3.2. **Semiclassical Lagrangian distributions.** In this subsection, we review some facts from the theory of semiclassical Lagrangian distributions. See [GuSt, Chapter 6] or [VũNg, Section 2.3] for a detailed account, and [HöIV, Section 25.1] or [GrSj, Chapter 11] for the closely related microlocal case. However, note that we do not attempt to define the principal symbols as global invariant geometric objects; this makes the resulting local theory considerably simpler.

Phase functions. Let M be a manifold without boundary. We denote its dimension by d; in the convention used in the present paper, $d = n + 1$. As before, we denote elements of T^*M by (m, ν), $m \in M$, $\nu \in T^*_m M$. Let $\varphi(m, \theta)$ be a smooth real-valued function on some open subset U_φ of $M \times \mathbb{R}^L$, for some L; we call m *base variables* and θ *oscillatory variables.* We say that φ is a (nondegenerate) *phase function*, if the differentials $d(\partial_{\theta_1}\varphi), \dots, d(\partial_{\theta_L}\varphi)$ are linearly independent on the *critical set*

$$C_\varphi := \{(m, \theta) \mid \partial_\theta \varphi = 0\} \subset U_\varphi. \tag{3.3}$$

In this case

$$\Lambda_\varphi := \{(m, \partial_m \varphi(m, \theta)) \mid (m, \theta) \in C_\varphi\} \subset T^*M$$

is an (immersed, and we will shrink the domain of φ to make it embedded) Lagrangian submanifold. We say that φ *generates* Λ_φ.

Symbols. Let $\rho \in [0, 1/2)$. A smooth function $a(m, \theta; h)$ is called a compactly supported symbol of type ρ on U_φ, if it is supported in some compact h-independent subset of U_φ, and for each differential operator ∂^α on $M \times \mathbb{R}^L$, there exists a constant C_α such that

$$\sup_{U_\varphi} |\partial^\alpha a| \leq C_\alpha h^{-\rho|\alpha|}.$$

Similarly to Section 3.1, we write $a \in S^{\mathrm{comp}}_\rho(U_\varphi)$. For the convergence Theorem 1, we will only need the class $S^{\mathrm{comp}} := S^{\mathrm{comp}}_0$; the classes S^{comp}_ρ for $\rho > 0$ will be required in the proof of the remainder estimates of Theorem 2.

Lagrangian distributions. Given a phase function φ and a symbol $a \in S^{\mathrm{comp}}_\rho(U_\varphi)$, consider the h-dependent family of functions

$$u(m; h) = h^{-L/2} \int_{\mathbb{R}^L} e^{i\varphi(m,\theta)/h} a(m, \theta; h)\, d\theta. \tag{3.4}$$

We call u a *Lagrangian distribution* of type ρ generated by φ. Using the method of nonstationary phase, we can see that if $\operatorname{supp} a$ is contained in some h-independent compact set $K \subset U_\varphi$, then

$$\mathrm{WF}_h(u) \subset \{(m, \partial_m \varphi(m, \theta)) \mid (m, \theta) \in C_\varphi \cap K\} \subset \Lambda_\varphi. \tag{3.5}$$

The *principal symbol* of u is the function

$$\sigma_\varphi(u) \in S^{\mathrm{comp}}_\rho(\Lambda_\varphi)$$

defined by the formula

$$\sigma_\varphi(u)(m, \partial_m \varphi(m, \theta); h) = a(m, \theta; h), \ \ (m, \theta) \in C_\varphi. \tag{3.6}$$

That $\sigma_\varphi(u)$ does not depend on the choice of a producing u, up to an $\mathcal{O}(h^{1-2\rho})$ remainder, will follow from Proposition 3.3 and (3.9). As mentioned above, we will not attempt to define the principal symbol independently of the choice of φ.

Following [GrSj, Chapter 11], we introduce a certain (local) canonical form for Lagrangian distributions. Fix some local system of coordinates on M (shrinking M to the domain of this coordinate system and identifying it with a subset of $\mathbb{R}^d$) and consider

$$\Lambda_F = \{(m,\nu) \mid m = -\partial_\nu F(\nu),\ \nu \in U_F\} \subset T^*M, \tag{3.7}$$

where F is a smooth real-valued function on some open set $U_F \subset \mathbb{R}^d$, such that the image of $-\partial_\nu F$ is contained in M. Then Λ_F is Lagrangian; in fact, it is generated by the phase function $m\cdot\nu + F(\nu)$, with ν the oscillatory variable. One can also prove that each Lagrangian submanifold that does not intersect the zero section locally has the form (3.7) for an appropriate choice of the coordinate system on M. (We will not have to work with Lagrangians intersecting the zero section in this paper; the corresponding distributions have all the properties listed below, except that the normal forms (3.4) and (3.14) have to be written differently.)

If $b(\nu;h) \in S^{\mathrm{comp}}_\rho(U_F)$ and $\chi \in C_0^\infty(M)$ is equal to 1 near $-\partial_\nu F(\operatorname{supp} b)$, then we can define a Lagrangian distribution by the following special case of (3.4):

$$v(m;h) = \chi(m) h^{-d/2} \int_{U_F} e^{i(m\cdot\nu + F(\nu))/h} b(\nu;h)\, d\nu. \tag{3.8}$$

We need χ to make $v \in C_0^\infty(M)$; however, by (3.5) (or directly by the method of nonstationary phase), if we choose χ differently, then v will change by $\mathcal{O}(h^\infty)_{C_0^\infty}$.

If v is given by (3.8), then we can recover the symbol b by the Fourier inversion formula:

$$e^{iF(\nu)/h} b(\nu;h) = (2\pi)^{-d} h^{-d/2} \int_M e^{-im\cdot\nu/h} v(m;h)\, dm + \mathcal{O}(h^\infty)_{\mathscr{S}(\mathbb{R}^d)}, \tag{3.9}$$

here $\mathscr{S}$ denotes the space of Schwartz functions. Note that $v = \mathcal{O}(h^\infty)_{C_0^\infty}$ implies $b(\nu;h) = \mathcal{O}(h^\infty)_{C_0^\infty}$. Moreover, if $v \in C_0^\infty(M)$ satisfies (3.9) for some $b \in S^{\mathrm{comp}}_\rho(U_F)$, then v is given by (3.8) modulo $\mathcal{O}(h^\infty)_{C_0^\infty}$.

Any Lagrangian distribution can be brought locally into the form (3.8):

Proposition 3.3. *Assume that φ is a phase function, and the corresponding Lagrangian $\Lambda = \Lambda_\varphi$ can be written in the form (3.7). For $a(m,\theta;h) \in S^{\mathrm{comp}}_\rho(U_\varphi)$ and $b(\nu;h) \in S^{\mathrm{comp}}_\rho(U_F)$, denote by u_a and v_b the functions given by (3.4) and (3.8), respectively. Then:*

1. For each $a \in S^{\mathrm{comp}}_\rho(U_\varphi)$, there exists $b \in S^{\mathrm{comp}}_\rho(U_F)$ such that $u_a = v_b + \mathcal{O}(h^\infty)_{C_0^\infty}$. Moreover, we have the following asymptotic decomposition for b:

$$b(\nu;h) = \sum_{0\leq j<N} h^j L_j a(m,\theta;h) + \mathcal{O}(h^{N(1-2\rho)})_{S^{\mathrm{comp}}_\rho(U_F)}, \tag{3.10}$$

where each L_j is a differential operator of order $2j$ on U_φ, and $(m,\theta) \in C_\varphi$ is the solution to the equation $(m, \partial_m\varphi(m,\theta)) = (-\partial_\nu F(\nu), \nu)$. In particular, if $\sigma_\varphi(u)$ is given by (3.6), then

$$\sigma_\varphi(u)(-\partial_\nu F(\nu), \nu; h) = f_{\varphi F} b(\nu;h) + \mathcal{O}(h^{1-2\rho})_{S^{\mathrm{comp}}_\rho(U_F)}, \tag{3.11}$$

where $f_{\varphi F}$ is some nonvanishing function depending on φ and the choice of the coordinate system on M. Adding a certain constant to the function F, we can make $f_{\varphi F}$ independent of h.

2. For each $b \in S^{\mathrm{comp}}_\rho(U_F)$, there exists $a \in S^{\mathrm{comp}}_\rho(U_\varphi)$ such that $v_b = u_a + \mathcal{O}(h^\infty)_{C_0^\infty}$.

Proof. We follow [GrSj, Chapter 11]. For part 1, we apply the method of stationary phase to get (3.9) and then use the Fourier inversion formula. For part 2, we take some $a_0 \in S^{\mathrm{comp}}_\rho(U_\varphi)$ satisfying (3.11) and define u_{a_0} by the formula (3.4) using the symbol a_0. Then $v_b = u_{a_0} + v_{b_1}$, where v_{b_1} has the form (3.8) with the symbol $b_1 = \mathcal{O}(h^{1-2\rho})$. Repeating this process with v_{b_1} in place of v_b, we can write for each N,

$$v = \sum_{0\leq j<N} u_{a_j} + v_{b_N},$$

where each u_{a_j} has the form (3.4) with the symbol $a_j = \mathcal{O}(h^{j(1-2\rho)})$, and v_{b_N} has the form (3.8) with the symbol $b_N = \mathcal{O}(h^{N(1-2\rho)})$. If $a \sim \sum_j a_j$ is an asymptotic sum, then $v_b - u_a = \mathcal{O}(h^\infty)_{C_0^\infty}$. □

We can now give

Definition 3.4. *Let $\Lambda \subset T^*M$ be an embedded Lagrangian submanifold. We say that an h-dependent family of functions $u(m;h) \in C_0^\infty(M)$ is a (compactly supported and compactly microlocalized) Lagrangian distribution of type ρ associated to Λ, if it can be written as a sum of finitely many functions of the form (3.4), for different phase functions φ parametrizing open subsets of Λ, plus an $\mathcal{O}(h^\infty)_{C_0^\infty}$ remainder. Denote by $I^{\mathrm{comp}}_\rho(\Lambda)$ the space of all such distributions, and put $I^{\mathrm{comp}}(\Lambda) := I^{\mathrm{comp}}_0(\Lambda)$.*

We can write any $u \in I^{\mathrm{comp}}_\rho(\Lambda)$ as the sum of Lagrangian distributions associated to a given finite open covering of $\mathrm{WF}_h(u)$ in Λ; by Proposition 3.3, u is a sum of functions of the form (3.8). Moreover, if φ is a phase function and $u \in I^{\mathrm{comp}}_\rho(\Lambda_\varphi)$, then u can be written in the form (3.4) for some symbol a, plus an $\mathcal{O}(h^\infty)_{C_0^\infty}$ remainder. The symbol $\sigma_\varphi(u)$, given by (3.6), is well-defined modulo $\mathcal{O}(h^{1-2\rho})$.

The action of a pseudodifferential operator on a Lagrangian distribution is given by the following proposition, following from Proposition 3.3 and the method of stationary phase:

Proposition 3.5. *Assume that $u \in I^{\mathrm{comp}}_\rho(\Lambda)$ and $P \in \Psi^k_\rho(M)$. Then $Pu \in I^{\mathrm{comp}}_\rho(\Lambda)$. Moreover,*

1. Assume that $\Lambda = \Lambda_\varphi$ for some phase function φ. Then

$$\sigma_\varphi(Pu) = \sigma(P)|_{\Lambda_\varphi} \cdot \sigma_\varphi(u) + \mathcal{O}(h^{1-2\rho})_{S^{\mathrm{comp}}_\rho(\Lambda)}.$$

2. Assume that $\Lambda = \Lambda_F$ is given by (3.7) in some coordinate system on M. Let $b(\nu;h)$ and $b^P(\nu;h)$ be the symbols corresponding to u and Pu, respectively, via (3.8). Let also $P = \mathrm{Op}_h(p)$ for some quantization procedure Op_h. Then we have the following asymptotic decomposition for b^P:

$$b^P(\nu;h) = \sum_{0\leq j<N} h^j L_j(p(m,\nu';h)b(\nu;h))|_{\nu'=\nu,\, m=-\partial_\nu F(\nu)} + \mathcal{O}(h^{N(1-2\rho)})_{S^{\mathrm{comp}}_\rho(U_F)},$$

where each L_j is a differential operator of order $2j$ on $M \times U_F \times U_F$.

Finally, we give the following estimate of the L^2 norm of a Lagrangian distribution, following from the boundedness of the Fourier transform on L^2:

Proposition 3.6. *Assume that $u \in I_\rho^{\mathrm{comp}}(\Lambda_F)$, where Λ_F is given by (3.7). Assume that u is given by (3.8), with $b(\nu; h)$ the corresponding symbol. Then*

$$\|u(m; h)\|_{L^2} \leq C\|b(\nu; h)\|_{L^2(U_F)}. \tag{3.12}$$

Here C is a constant independent of h.

Fourier integral operators. A special case of Lagrangian distributions are Fourier integral operators associated to canonical transformations. Let M, M' be two manifolds of the same dimension d, and let κ be a symplectomorphism from an open subset of T^*M to an open subset of T^*M'. Consider the Lagrangian

$$\Lambda_\kappa = \{(m, \nu; m', -\nu') \mid \kappa(m, \nu) = (m', \nu')\} \subset T^*M \times T^*M' = T^*(M \times M').$$

A compactly supported operator $U : \mathcal{D}'(M') \to C_0^\infty(M)$ is called a (semiclassical) *Fourier integral operator* of type ρ associated to κ, if its Schwartz kernel $K_U(m, m')$ lies in $h^{-d/2} I_\rho^{\mathrm{comp}}(\Lambda_\kappa)$. We write $U \in I_\rho^{\mathrm{comp}}(\kappa)$. Note that we quantize a canonical transformation $T^*M \to T^*M'$ as an operator $\mathcal{D}'(M') \to C_0^\infty(M)$, in contrast with the standard convention, which would quantize it as an operator $\mathcal{D}'(M) \to C_0^\infty(M')$. The $h^{-d/2}$ factor is explained as follows: the normalization for Lagrangian distributions is chosen so that $\|u\|_{L^2} \sim 1$, while the normalization for Fourier integral operators is chosen so that $\|U\|_{L^2(M') \to L^2(M)} \sim 1$.

After sufficiently shrinking the domain of κ and choosing an appropriate coordinate system on M' (replacing M' with the domain of this coordinate system and identifying it with a subset of $\mathbb{R}^d$), we can find a generating function $S(m, \nu')$ for κ; that is,

$$\kappa(m, \nu) = (m', \nu') \iff \partial_m S(m, \nu') = \nu,\ \partial_{\nu'} S(m, \nu') = m'. \tag{3.13}$$

Here (m, ν') vary in some open set $U_S \subset M \times \mathbb{R}^d$. The phase function $S(m, \nu') - m' \cdot \nu'$, with ν' the oscillatory variable, parametrizes Λ_κ and for $U \in I_\rho^{\mathrm{comp}}(\kappa)$, we can write similarly to (3.8),

$$K_U(m, m') = h^{-d} \chi(m') \int_{\mathbb{R}^d} e^{\frac{i}{h}(S(m,\nu') - m' \cdot \nu')} b(m, \nu'; h)\, d\nu' + \mathcal{O}(h^\infty)_{C_0^\infty}, \tag{3.14}$$

for some symbol $b \in S_\rho^{\mathrm{comp}}(U_S)$ and any $\chi \in C_0^\infty(M')$ such that $\chi = 1$ near the set $\partial_{\nu'} S(\operatorname{supp} b)$. The function b is determined uniquely by U modulo $\mathcal{O}(h^\infty)_{S_\rho^{\mathrm{comp}}(U_S)}$, similarly to (3.9). Note that if κ is the identity map, then $S(m, \nu') = m \cdot \nu'$ and we arrive to the quantization formula for a semiclassical pseudodifferential operator.

Similarly to Proposition 3.5, we have

Proposition 3.7. *Assume that $U \in I_\rho^{\mathrm{comp}}(\kappa)$ and $P \in \Psi_\rho^k(M')$. Then $UP \in I_\rho^{\mathrm{comp}}(\kappa)$. If moreover κ is given by (3.13), $b(m, \nu'; h)$ and $b^P(m, \nu'; h)$ are the symbols corresponding to*

U and UP, respectively, via (3.14), and $P = \operatorname{Op}_h(p)$ for some quantization procedure Op_h, then we have the following asymptotic decomposition for b^P:

$$b^P(m,\nu';h) = \sum_{0\leq j<N} h^j L_j(p(m',\tilde\nu)b(m,\nu'))|_{\tilde\nu=\nu',\, m'=\partial_{\nu'}S(m,\nu')} + \mathcal{O}(h^{N(1-2\rho)})_{S^{\mathrm{comp}}_\rho(U_S)}.$$

Here each L_j is a differential operator of order $2j$ on $M'\times\mathbb{R}^d\times U_S$. In particular,

$$b^P(m,\nu';h) = p(\partial_{\nu'}S(m,\nu'),\nu';h)b(m,\nu';h) + \mathcal{O}(h^{1-2\rho})_{S^{\mathrm{comp}}_\rho(U_S)}.$$

A similar statement is true for an operator of the form PU, where $P\in\Psi^k_\rho(M)$; the terms of the asymptotic decomposition have the form $h^jL_j(p(\widetilde m,\nu)b(m,\nu'))$, where we take $\widetilde m = m$ and $\nu = \partial_m S(m,\nu')$.

3.3. **Schrödinger propagators.** In this subsection, we assume that (M,g) is a complete Riemannian manifold and $\Delta = \Delta_g$ is the corresponding (nonnegative) Laplace–Beltrami operator. Let p be the semiclassical principal symbol of $h^2\Delta \in \Psi^2(M)$; we have $p(m,\nu) = |\nu|_g^2$, where $|\nu|_g$ is the norm of the covector $\nu\in T^*_mM$ induced by g. We use the notation

$$S^*M = p^{-1}(1)\subset T^*M$$

for the unit cotangent bundle. The geodesic flow g^t on T^*M is related to the Hamiltonian flow e^{tH_p} of p by the formula $g^t = e^{tH_p/2}$. Note that g^t is a canonical transformation. The operator Δ is essentially self-adjoint on $L^2(M)$ by [Ch] and its domain is given by the Friedrichs extension. Let

$$U(t) = e^{ith\Delta/2} = e^{\frac{it}{h}(h^2\Delta/2)}$$

be the semiclassical Schrödinger propagator; it is a unitary operator on $L^2(M)$. The basic microlocal properties of $U(t)$ are given by the following

Proposition 3.8. *For each $t\in\mathbb{R}$,*

1. (Egorov's Theorem) For each compactly supported $A\in\Psi^{\mathrm{comp}}_\rho(M)$, there exists compactly supported $A^t\in\Psi^{\mathrm{comp}}_\rho(M)$ such that

$$U(t)AU(-t) = A^t + \mathcal{O}(h^\infty)_{L^2\to L^2}. \tag{3.15}$$

Moreover, $\operatorname{WF}_h(A^t)\subset g^{-t}(\operatorname{WF}_h(A))$ and $\sigma(A^t) = \sigma(A)\circ g^t + \mathcal{O}(h^{1-2\rho})$.

2. (Microlocalization) The operator $U(t)$ is microlocalized on the graph of g^{-t} in the following sense: if $A,B\in\Psi^k_\rho(M)$ are compactly supported, at least one of them is compactly microlocalized, and

$$g^t(\operatorname{WF}_h(A))\cap\operatorname{WF}_h(B) = \emptyset, \tag{3.16}$$

then $AU(t)B = \mathcal{O}(h^\infty)_{L^2\to L^2}$.

3. (Parametrix) If $A\in\Psi^{\mathrm{comp}}(M)$ is compactly supported, then $U(t)A$ is the sum of a compactly microlocalized Fourier integral operator (of type 0) associated to g^t, as defined in Section 3.2, and an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder.

The proofs are standard; part 1 can be found in [Zw, Theorem 11.1] (with the mildly exotic classes $\Psi^{\mathrm{comp}}_\rho$ handled as in Appendix C), part 2 follows directly from part 1, and part 3 is proved similarly to [Zw, Theorem 10.3]. The operator $U(t)A$ quantizes g^t, not g^{-t}, because of the convention adopted in Section 3.2 that a canonical transformation $T^*M \to T^*M'$ is quantized as an operator $\mathcal{D}'(M') \to C_0^\infty(M)$.

Egorov's theorem until the Ehrenfest time. Proposition 3.8 is valid for bounded times t; as $t \to \infty$, the constants in the estimates for the corresponding symbols will blow up. However, it is still possible to prove Egorov's Theorem for t bounded by a certain multiple of $\log(1/h)$, called the Ehrenfest time. To define this time, we fix an open bounded geodesically convex set $U \subset M$ and define the *maximal expansion rate*

$$\Lambda_{\max} := \limsup_{|t|\to\infty} \frac{1}{|t|} \log \sup_{\substack{m\in U,\, |\nu|_g=1,\\ g^t(m,\nu)\in U}} \|dg^t(m,\nu)\|. \tag{3.17}$$

Here $\|dg^t(m,\nu)\|$ is the operator norm of the differential

$$dg^t(m,\nu) : T_{(m,\nu)}T^*M \to T_{g^t(m,\nu)}T^*M$$

with respect to any given smooth norm on the fibers of $T(T^*M)$ (e.g. the norm induced by the metric g).

Since we will work on a noncompact manifold, we introduce cutoffs into the corresponding propagators:

Proposition 3.9. *Assume that $X_1, X_2 \in \Psi^0(M)$ satisfy $\|X_j\|_{L^2\to L^2} \leq 1 + \mathcal{O}(h)$ and are compactly supported inside U. Let $\varepsilon_e > 0$ and take $\Lambda_0, \Lambda_0' > 0$ such that*

$$\Lambda_0 > \Lambda_0' > (1 + 2\varepsilon_e)\Lambda_{\max}.$$

Fix $t_0 \in \mathbb{R}$. Then for each integer

$$l \in [0, \log(1/h)/(2|t_0|\Lambda_0)], \tag{3.18}$$

and each compactly supported $A \in \Psi^{\mathrm{comp}}(M)$ with

$$\mathrm{WF}_h(A) \subset \mathcal{E}_{\varepsilon_e} := \{1 - \varepsilon_e \leq |\nu|_g \leq 1 + \varepsilon_e\},$$

the compactly supported operator

$$A^{(l)} := (X_2 U(t_0))^l A (U(-t_0)X_1)^l$$

lies in $\Psi^{\mathrm{comp}}_{\rho_l}(M)$, modulo an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder, with

$$\rho_l = l|t_0|\Lambda_0'/\log(1/h) < 1/2. \tag{3.19}$$

Moreover, the $S^{\mathrm{comp}}_{\rho_l}$ seminorms of the full symbol of $A^{(l)}$ are bounded uniformly in l, in the following sense: the order k derivatives of this symbol are bounded by $Ch^{-k\rho_l}$, where C is a constant independent of h and l. The principal symbol of $A^{(l)}$ is

$$\sigma(A^{(l)}) = (\sigma(A) \circ g^{lt_0}) \prod_{j=0}^{l-1} (\sigma(X_1)\sigma(X_2)) \circ g^{jt_0} + \mathcal{O}(h^{1-2\rho_l}).$$

The wavefront set of $A^{(l)}$, for $l > 0$, is contained in $\mathrm{WF}_h(X_1) \cap \mathrm{WF}_h(X_2) \cap \mathcal{E}_{\varepsilon_e}$. Finally, if U_A and U_X are open sets such that $\mathrm{WF}_h(A) \subset U_A$ and $\mathrm{WF}_h(X_1) \cap \mathrm{WF}_h(X_2) \subset U_X$, then

$A^{(l)}$ is microsupported, in the sense of Definition 3.1, inside the set

$$V^{(l)} := g^{-lt_0}(U_A) \cap \bigcap_{j=0}^{l-1} g^{-jt_0}(U_X).$$

The set $V^{(l)}$ does not depend on h directly, however it depends on l, which is allowed to depend on h, and our microlocal vanishing statement is uniform in l.

Proposition 3.9 is the main technical tool of obtaining the polynomial remainder bound of Theorem 2; it is also the reason why the classes $\Psi_\rho^{\mathrm{comp}}$ appear. Its proof, following the methods of [AnNo, Section 5.2] and [Zw, Theorem 11.12], is given in Appendix C. We do not impose any restrictions on the set U at this point, however in our actual argument it will have to contain a neighborhood of the trapped set – see the beginning of Section 5.2.

Propagating generalized eigenfunctions. The following fact, similar to [Dy2, Proposition 3.3], will be used to propagate the Eisenstein functions by the group $U(t)$:

Lemma 3.10. *Assume that $u \in C^\infty(M)$ solves the equation*

$$(h^2\Delta - z)u = 0,\ \ |1-z| \leq Ch.$$

Let $\chi \in C_0^\infty(M)$; take $t \in \mathbb{R}$ and assume that $\chi_t \in C_0^\infty(M)$ is supported in the interior of a compact set $K_t \subset M$ and satisfies

$$d_g(\operatorname{supp}\chi, \operatorname{supp}(1-\chi_t)) > |t|. \tag{3.20}$$

Here d_g denotes Riemannian distance on M. Then

$$\chi u = \chi e^{-itz/(2h)} U(t)\chi_t u + \mathcal{O}(h^\infty \|u\|_{L^2(K_t)})_{L^2(M)}.$$

Proof. Without loss of generality, we assume that $t \geq 0$. For $0 \leq s \leq t$, define

$$u_s = \chi(u - e^{-isz/(2h)}U(s)\chi_t u).$$

We need to prove that

$$\|u_t\|_{L^2} = \mathcal{O}(h^\infty)\|u\|_{L^2(K_t)}. \tag{3.21}$$

Since $\chi = \chi\chi_t$, we have $u_0 = 0$; next,

$$\begin{aligned}2hD_s u_s &= -\chi e^{-isz/(2h)}U(s)(h^2\Delta - z)\chi_t u\\ &= -e^{-isz/(2h)}\chi U(s)[h^2\Delta, \chi_t]u.\end{aligned}$$

Let $B \in \Psi^{\mathrm{comp}}$ be compactly supported inside $K_t \times K_t$, equal to the identity microlocally near $\operatorname{supp}\chi_t \cap S^*M$, but microlocalized in a small enough neighborhood of S^*M so that by (3.20),

$$g^s(\operatorname{supp}\chi) \cap \mathrm{WF}_h(B) \cap \operatorname{supp}(1-\chi_t) = \emptyset.$$

Note that $\mathrm{WF}_h([h^2\Delta, \chi_t]) \subset \operatorname{supp}(1-\chi_t)$. Then by part 2 of Proposition 3.8,

$$\|\chi U(s)[h^2\Delta, \chi_t]Bu\|_{L^2} = \mathcal{O}(h^\infty)\|u\|_{L^2(K_t)},\ \ 0 \leq s \leq t. \tag{3.22}$$

Moreover, by Proposition 3.2

$$\|\chi U(s)[h^2\Delta, \chi_t](1-B)u\|_{L^2} = \mathcal{O}(h^\infty)\|u\|_{L^2(K_t)}. \tag{3.23}$$

Combining (3.22) and (3.23), we get $\|\partial_s u_s\|_{L^2} = \mathcal{O}(h^\infty)\|u\|_{L^2(K_t)}$; it remains to integrate in s to get (3.21). ☐

Hilbert–Schmidt norm estimates. We now prove Hilbert–Schmidt norm estimates for the product of a pseudodifferential operator with a spectral projector. (See [HöIII, Section 19.1] for the properties of Hilbert–Schmidt and trace class operators.) To simplify notation, we consider a spectral interval of size h centered at $\lambda = 1$; similar statement is true for the interval $[\lambda + c_1 h, \lambda + c_2 h]$ with $\lambda > 0$, replacing S^*M by λS^*M.

Lemma 3.11. *Fix $c_1, c_2 \in \mathbb{R}$ and let $\mathbb{1}_{[1+c_1h,1+c_2h]}(h^2\Delta)$ be defined by means of spectral theory. Assume that $A \in \Psi^{\mathrm{comp}}_\rho(M)$ is compactly supported. Then*

$$h^{(d-1)/2}\| \mathbb{1}_{[1+c_1h,1+c_2h]}(h^2\Delta)A\|_{\mathrm{HS}} \le C\|\sigma(A)\|_{L^2(S^*M)} + \mathcal{O}(h^{1-2\rho}). \tag{3.24}$$

Here C is a constant independent of A (if $\mathrm{WF}_h(A)$ is contained in a fixed compact set), however the constant in $\mathcal{O}(h^{1-2\rho})$ depends on A. We take the L^2 norm of $\sigma(A)$ on the energy surface S^*M with respect to the Liouville measure μ_L.

Moreover, if $\mathrm{WF}_h(A)$ is microsupported, in the sense of Definition 3.1, in some h-dependent family of sets $V(h) \subset T^*M$, then

$$h^{(d-1)/2}\| \mathbb{1}_{[1+c_1h,1+c_2h]}(h^2\Delta)A\|_{\mathrm{HS}} \le C\mu_L(V(h)\cap S^*M)^{1/2} + \mathcal{O}(h^\infty). \tag{3.25}$$

Here $\mu_L(V(h)\cap S^*M)$ denotes the volume of $V(h)\cap S^*M$ with respect to the Liouville measure on S^*M and the constant C depends on a certain S^{comp}_ρ-seminorm of the full symbol of A.

Proof. Take a function $\chi \in \mathscr{S}(\mathbb{R})$ such that $\hat\chi$ is compactly supported in some interval $(-T, T)$ and χ does not vanish on $[c_1, c_2]$ (for example, take nonzero $\psi \in C_0^\infty(\mathbb{R})$ with $\psi \ge 0$, then $|\hat\psi| > 0$ in an interval $[c_1\varepsilon, c_2\varepsilon]$; set $\chi(x) := \hat\psi(\varepsilon x)$). Then

$$\mathbb{1}_{[1+c_1h,1+c_2h]}(h^2\Delta) = Z\chi((h^2\Delta - 1)/h),$$

where Z is a certain function of $h^2\Delta$ and it is bounded on $L^2(M)$ uniformly in h. It then suffices to estimate the Hilbert–Schmidt norm of

$$B = h^{(d-1)/2}\chi((h^2\Delta - 1)/h)A = (2\pi)^{-1}h^{(d-1)/2}\int_{-T}^{T}\hat\chi(t)e^{-it/h}U(2t)A\,dt.$$

Let $A_0 \in \Psi^{\mathrm{comp}}_0(M)$ be compactly supported and equal to the identity microlocally near $\mathrm{WF}_h(A)$. By part 3 of Proposition 3.8, for each t we have

$$U(2t)A_0 = U_{2t} + R_{2t},$$

where $U_{2t} \in I^{\mathrm{comp}}(g^{2t})$ is a compactly supported Fourier integral operator and $R_{2t} = \mathcal{O}(h^\infty)_{L^2\to L^2}$. Then

$$(U(2t) - U_{2t})A = \mathcal{O}(h^\infty)_{\mathrm{HS}}. \tag{3.26}$$

Indeed, we can write the left-hand side of (3.26) as the sum of $R_{2t}A$ and $U(2t)(1 - A_0)A$; it remains to note that $R_{2t} = \mathcal{O}(h^\infty)_{L^2\to L^2}$, $\|A\|_{\mathrm{HS}}$ is polynomially bounded in h, and $\|(1 - A_0)A\|_{\mathrm{HS}} = \mathcal{O}(h^\infty)$.

By (3.26), we can replace $U(2t)$ by U_{2t} in the definition of B. Now, the Hilbert–Schmidt norm of B is equal to the $L^2(M\times M)$ norm of its Schwartz kernel K_B. Using the local normal form (3.14) for Fourier integral operators, we can write K_B, up to an $\mathcal{O}(h^\infty)_{C_0^\infty}$ remainder and an appropriate cutoff in the m' variable, as a finite sum of expressions of the form (in a fixed coordinate system on M)

$$h^{-(d+1)/2}\int_{-T}^{T}\int_{\mathbb{R}^d} e^{i(S(m,\nu';2t)-m'\cdot\nu'-t)/h} b(m,\nu',t;h)\,d\nu' dt. \tag{3.27}$$

Here $S(m,\nu';2t)$ is a generating function for g^{2t} and b is a certain symbol in S_ρ^{comp}. Moreover, b admits an asymptotic expansion in terms of the full symbol of A, by Proposition 3.7. The fact that S and b can be choosen to depend smoothly on t follows from the proof of part 3 of Proposition 3.8.

We can deduce from (3.13) that

$$g^{2t}(m,\nu)=(m',\nu') \Longrightarrow \partial_t(S(m,\nu';2t))=p(m,\nu). \tag{3.28}$$

The equation (3.28) is true for the particular generating function constructed in the proof of part 3 of Proposition 3.8. One can add any function of t to the function S and still obtain a generating function of g^{2t}; however, the amplitude b with respect to the new generating function will no longer be a symbol, as its derivatives in t will not be bounded uniformly in h. It follows from (3.28) that the function

$$\Phi(m,m',\nu',t)=S(m,\nu';2t)-m'\cdot\nu'-t$$

is a nondegenerate phase function (with m,m' as base variables and ν',t as the oscillatory variables) and generates the (immersed) Lagrangian

$$\Lambda=\{(m,\nu;m',-\nu')\mid p(m,\nu)=1,\ \exists t\in(-T,T): g^{2t}(m,\nu)=(m',\nu')\}.$$

Then (3.27) lies in $I_\rho^{\mathrm{comp}}(\Lambda)$. By the local normal form (3.8) of a Lagrangian distribution, we can write (3.27), up to an $\mathcal{O}(h^\infty)_{C_0^\infty}$ remainder and an appropriate cutoff in the (m,m') variables, as the sum of finitely many expressions of the form

$$h^{-d}\int_{\mathbb{R}^{2d}} e^{i(m\cdot\nu+m'\cdot\nu'+F(\nu,\nu'))/h}\tilde b(\nu,\nu';h)\,d\nu d\nu', \tag{3.29}$$

where F parametrizes some open subset of Λ by (3.7) and $\tilde b$ is a symbol in S_ρ^{comp}. By Proposition 3.7 and Proposition 3.3, we see that the symbol $\tilde b$ has the following asymptotic expansion in terms of the full symbol a of A:

$$\tilde b(\nu,\nu';h)=\sum_{0\le j<N} h^j L_j a(m',\nu';h)+\mathcal{O}(h^{N(1-2\rho)})_{S_\rho^{\mathrm{comp}}}, \tag{3.30}$$

where each L_j is a differential operator of order $2j$ and m,m' are given by the relation $(m,\nu,m',-\nu')\in\Lambda$; in particular, $(m',\nu')\in S^*M$.

We now use Proposition 3.6 to estimate the L^2 norm of (3.29); as B is, modulo $\mathcal{O}(h^\infty)_{\mathrm{HS}}$, a sum of operators with Schwartz kernels of the form (3.29), this would give an estimate on the Hilbert–Schmidt norm of B. For (3.24), we can write $\tilde b(\nu,\nu';h)$ as a multiple of $a(m',\nu')$ plus an $\mathcal{O}(h^{1-2\rho})$ remainder and note that (m',ν') always lies in S^*M. For (3.25), we use that $\tilde b=\mathcal{O}(h^\infty)$ outside of the preimage of $V(h)$ under the map $(\nu,\nu')\mapsto(m',\nu')$, and that $\sup|\tilde b|$ can be estimated by a certain S_ρ^{comp}-seminorm of a. □

Local traces of integrated Schrödinger propagators. We give the following version of the Schrödinger propagator trace formula in the case where there are no contributions from closed geodesics:

Lemma 3.12. *Assume that M is a d-dimensional complete Riemannian manifold and B_s is a family of compactly supported pseudodifferential operators in $\Psi_\rho^{\mathrm{comp}}(M)$, smooth and compactly supported in $s \in (-T_0, T_0)$, where $T_0 > 0$ is fixed. Assume also that all B_s are microsupported, in the sense of Definition 3.1, in some h-dependent family of bounded sets $V(h) \subset T^*M$, and the following nonreturning condition holds:*

$$(m,\nu) \in V(h),\ |s| < T_0 \Longrightarrow d((m,\nu), g^s(m,\nu)) \geq C^{-1}|s|h^\rho. \tag{3.31}$$

Here C is some constant and d denotes some smooth distance function on T^*M. Let $B_s = \mathrm{Op}_h(b(s)) + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ for some family of symbols $b(s,m,\nu)$ and some quantization procedure Op_h. Then for each N and each $\lambda > 0$, we have the trace expansion

$$\begin{aligned}&(2\pi h)^{d-1}\int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)}\,\mathrm{Tr}(U(s)B_s)\,ds\\ &= \sum_{0\leq j<N} h^j \int_{S^*M} L_j b(0,m,\lambda\nu)\,d\mu_L(m,\nu) + \mathcal{O}(h^{N(1-2\rho)})_{C^\infty_\lambda},\end{aligned} \tag{3.32}$$

where μ_L is the Liouville measure and each L_j is a differential operator of order $2j$ on $T^*M_{(m,\nu)} \times (-T_0,T_0)_s$, independent of B_s and depending smoothly on λ. In particular, $L_0 = \lambda^{d-2}$.

Proof. As in the proof of Lemma 3.11, we can reduce to computing the trace of the operator with the Schwartz kernel (in some fixed local coordinates)

$$K(m,m') = (2\pi h)^{-1}\int_{-T_0}^{T_0}\int_{\mathbb{R}^d} e^{\frac{i}{h}(S(m,\nu';s)-m'\cdot\nu'-\lambda^2 s/2)}\tilde b(m,\nu',s;h)\,d\nu' ds,$$

where $S(m,\nu';s)$ is a local generating function for g^s in the sense of (3.13) and $\tilde b(m,\nu',s;h)$ is a certain symbol in S_ρ having an asymptotic expansion in terms of the jet of b_s at the point $(\partial_{\nu'}S(m,\nu';s),\nu')$. The trace of the corresponding operator is

$$\int_M K(m,m)\,dm = (2\pi h)^{-1}\int_{-T_0}^{T_0}\int_{M\times\mathbb{R}^d} e^{\frac{i}{h}(S(m,\nu';s)-m\cdot\nu'-\lambda^2 s/2)}\tilde b(m,\nu',s;h)\,dm d\nu' ds.$$

We now use the method of stationary phase. The stationary points of the phase are solutions to the equations $g^s(m,\nu') = (m,\nu')$ and $|\nu'|_g = \lambda$; they occur at $s=0$ and may also occur for $\lambda|s| \geq r_i$, where $r_i > 0$ is the injectivity radius of M. For $\lambda|s| \geq r_i/2$, we see by (3.31) that the expression under the integral can be split into two pieces, on one of which the symbol is $\mathcal{O}(h^\infty)$ and on the other, the differential of the phase function has length at least $C^{-1}h^\rho$; by repeated integration by parts, the latter integral is $\mathcal{O}(h^\infty)$.

It remains to evaluate the contribution of the stationary set $\{s=0\}\cap\lambda S^*M$. The phase function is degenerate on these points; however, one can pass to polar coordinates $\nu' = r\omega$, with $|\omega|_g = 1$ and $r > 0$, and apply the method of stationary phase in the (r,s) variables, resulting in the expansion (3.32). See for example the proofs of [Ro, Théorème V-7 and Proposition V-8] or [RoTa, Lemma 3.1] for details of the computation. □

4. General assumptions

In this section, we list the general geometric assumptions on the manifold M and analytic assumptions on its Laplacian required for our results to hold. As noted in the introduction, they are satisfied in particular if outside of a compact set, M is isometric to either the Euclidean space (studied in Section 6) or an asymptotically hyperbolic space of constant curvature (studied in Section 7). We also derive some direct consequences of the general assumptions, including averaged estimates on plane waves and the existence of limiting measures μ_ξ.

4.1. **Geometric assumptions.** In this subsection, we specify the geometry of the manifold M at infinity.

Let us introduce some notation and terminology first. On a complete Riemannian manifold (M, g) we denote by g^t the geodesic flow of the metric g, considered as a map on the cotangent bundle T^*M. Any smooth function f on M can be lifted to a function on T^*M; denote by $\dot f, \ddot f \in C^\infty(T^*M)$ the derivatives of f with respect to the geodesic flow:

$$\dot f(m,\nu) := d_t f(g^t(m,\nu))|_{t=0},\ \ \ddot f(m,\nu) := d_t^2 f(g^t(m,\nu))|_{t=0}.$$

We denote by S^*M the unit cotangent bundle $\{(m,\nu) \mid |\nu|_g = 1\} \subset T^*M$.

A *boundary defining function* on a smooth compact manifold $\overline{M}$ with boundary is a smooth function $x : \overline{M} \to [0,\infty)$ such that $x > 0$ on M and x vanishes to first order on $\partial\overline{M}$.

We make the following assumptions:

(G1) (M, g) is a complete Riemannian manifold of dimension $d = n+1$. Moreover, there exists a *compactification* of M, namely a compact manifold with boundary $\overline{M}$ such that M is diffeomorphic to the interior of $\overline{M}$. The boundary $\partial\overline{M}$ is called the *boundary at infinity*;

(G2) There exists a boundary defining function x on M and a constant $\varepsilon_0 > 0$ such that for any point $(m,\nu) \in S^*M$,

$$\text{if } x(m,\nu) \le \varepsilon_0 \text{ and } \dot x(m,\nu) = 0,\ \text{then } \ddot x(m,\nu) < 0; \tag{4.1}$$

(G3) For each $(m,\nu) \in S^*M$ such that $x(m) \le \varepsilon_0$ and $\dot x(m,\nu) \le 0$, the geodesic $g^t(m,\nu)$ (projected onto the base space M) converges as $t \to +\infty$, in the topology of $\overline{M}$, to some point $\xi_{+\infty}(m,\nu) \in \partial\overline{M}$. The function $\xi_{+\infty}$ depends smoothly on (m,ν), and we extend it naturally (as the limit of the corresponding geodesic) to a smooth function on $S^*M \setminus \Gamma_-$, with Γ_- given in Definition 4.1 below;

(G4) There exists an open set $U_\infty \subset M \times \partial\overline{M}$ such that $\overline{U}_\infty$ contains a neighbourhood of $\{(\xi,\xi) \in \overline{M} \times \partial\overline{M} \mid \xi \in \partial\overline{M}\}$ and a smooth real-valued function $\phi(m,\xi) = \phi_\xi(m)$ on U_∞ such that $|\partial_m\phi_\xi(m)|_g = 1$ everywhere and the function

$$\tau(m,\xi) := (m, \partial_m\phi_\xi(m)) \in S^*M,\ \ (m,\xi) \in U_\infty, \tag{4.2}$$

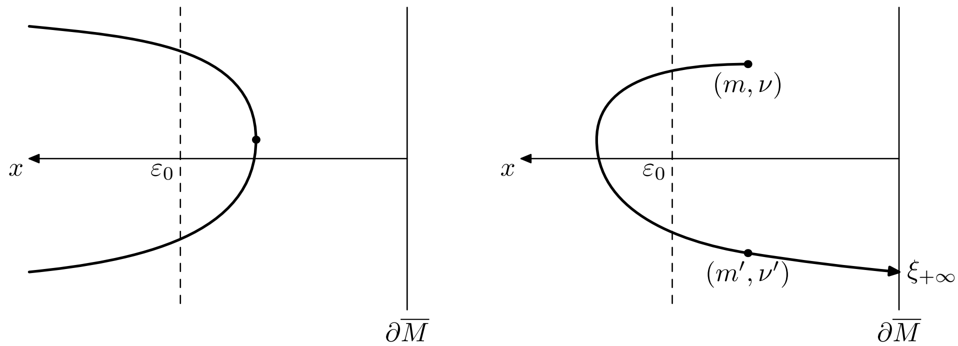

FIGURE 2. Left: an illustration of (G2), showing a forbidden geodesic. Right: an illustration of (G3). The point (m,ν) does not escape directly in the forward direction, but the point (m',ν') does.

is a diffeomorphism from U_∞^+ onto V_∞^+ with inverse given by

$$\tau^{-1}(m,\nu)=(m,\xi_{+\infty}(m,\nu)),\ (m,\nu)\in V_\infty^+$$

where the sets U_∞^+ and V_∞^+ are defined by

$$U_\infty^+ := \{(m,\xi)\in U_\infty \mid x(m)\le\varepsilon_0,\ \dot x(\tau(m,\xi))\le 0\},$$
$$V_\infty^+ := \{(m,\nu)\in S^*M \mid x(m)\le\varepsilon_0,\ \dot x(m,\nu)\le 0,\ (m,\xi_{+\infty}(m,\nu))\in U_\infty\};$$

(G5) if $(m,\nu)\in V_\infty^+$, then $g^t(m,\nu)\in V_\infty^+$ for all $t\ge 0$;

(G6) if $\xi\in\partial\overline{M}$ and $m,m'\in M$ are such that $(m,\xi),(m',\xi)\in U_\infty^+$, then $\partial_\xi\phi_\xi(m)=\partial_\xi\phi_\xi(m')$ if and only if $\tau(m,\xi)$ and $\tau(m',\xi)$ lie on the same geodesic. Moreover, the matrix $\partial_m\partial_\xi\phi_\xi(m)$ has rank n.

Escaping trajectories and the trapped set.. We now define the incoming/outgoing tails $\Gamma_\pm$ and the trapped set K:

Definition 4.1. *Let $\gamma(t)$ be a unit speed geodesic. We say that it* escapes *in the forward, respectively backward, direction, if $\gamma(t)$ goes to infinity in M as $t\to+\infty$, respectively $t\to-\infty$. If $\gamma(t)$ does not escape in some direction, we call it* trapped *in this direction. Denote by $\Gamma_+\subset S^*M$ the union of all geodesics trapped in the backward direction, by Γ_- the union of all geodesics trapped in the forward direction, and put $K=\Gamma_+\cap\Gamma_-$; we call K the* trapped set*.*

An escaping geodesic could potentially spend a long time in the compact part of the manifold. It is helpful to consider geodesics that escape in a straightforward way (with the boundary defining function x decreasing along them); they appeared in assumption (G3) for instance.

Definition 4.2. *We say that $(m,\nu)\in S^*M$* directly escapes *in the forward, respectively backward, direction, if $x(m)\le\varepsilon_0$ and $\dot x(m,\nu)\le 0$, respectively $\dot x(m,\nu)\ge 0$. Here ε_0 is the constant from (G2). Denote by $\mathcal{DE}_+$, respectively $\mathcal{DE}_-$, the set of all points directly escaping in the forward, respectively backward, direction.*

One can verify that $\Gamma_\pm$ are closed sets and the trapped set K is compact (see [GéSj, Appendix]); in fact, since $S^*M \cap \{x \leq \varepsilon_0\} \subset \mathcal{DE}_+ \cup \mathcal{DE}_-$, we have $K \subset \{x > \varepsilon_0\}$.

For the example of $M = \mathbb{R}^{n+1}$ discussed below, we have $\Gamma_\pm = \emptyset$. The point (m, ν) lies in $\mathcal{DE}_+$ if and only if $x(m) \leq \varepsilon_0$ and $m \cdot \nu \geq 0$.

Comments on the geometric assumptions.

A basic example to have in mind for a manifold satisfying our assumptions is $M = \mathbb{R}^{n+1}$ with the radial compactification $\overline{M}$ being a closed ball and the boundary at infinity $\partial\overline{M}$ equal to the sphere $\mathbb{S}^n$. We will often use this example to illustrate the somewhat abstract assumptions of this section. (A more general version will be considered in Section 6.)

An important corollary of the assumption (G2) is that for $\varepsilon \leq \varepsilon_0$, the compact set $\{x \geq \varepsilon\} \subset M$ is *geodesically convex*; i.e., if $\gamma(t)$ is a geodesic and $\gamma(t_1), \gamma(t_2) \in \{x \geq \varepsilon\}$, then $\gamma(t) \in \{x \geq \varepsilon\}$ for $t \in [t_1, t_2]$. For the example of $M = \mathbb{R}^{n+1}$, we can take $x = (1+|m|^{-2})^{-1/2}$, where $|m|$ is the Euclidean length of $m \in \mathbb{R}^d$; the corresponding sets $\{x \geq \varepsilon\}$ are balls centered at zero.

It also follows from (G2) that for $(m, \nu) \in \mathcal{DE}_+$, the function $x(g^t(m, \nu))$ is decreasing for $t \geq 0$. One can show that $x(g^t(m, \nu)) \to 0$ as $t \to +\infty$ and thus $g^t(m, \nu)$ escapes in the forward direction; we do not give a proof of this fact as it follows from the more restrictive assumption (G3). Also, if a geodesic $\gamma(t)$ escapes in the forward direction, then for t large enough we have $\gamma(t) \in \mathcal{DE}_+$. For $M = \mathbb{R}^{n+1}$, we have $\xi_{+\infty}(m, \nu) = \nu \in \mathbb{S}^n$.

Assumption (G4) means that for m sufficiently close to the infinity, the covectors ν such that $(m, \nu) \in \mathcal{DE}_+$ are in one-to-one correspondence with the limit points $\xi_{+\infty}(m, \nu)$, and the inverse correspondence can be described by a phase function. It follows in particular from (G4) that for a fixed $\xi \in \partial\overline{M}$, the set of directly escaping points (m, ν) such that $\xi_{+\infty}(m, \nu) = \xi$ and $(m, \xi) \in U_\infty$ is the intersection of $\mathcal{DE}_+$ with the Lagrangian

$$\Lambda_\xi := \{(m, \partial_m \phi_\xi(m)) \mid (m, \xi) \in U_\infty\}. \tag{4.3}$$

In the model case $M = \mathbb{R}^{n+1}$ we can put for any $R > 0$, $U_\infty = \{(m, \xi) \mid |m| > R\}$, and $\phi_\xi(m) = m \cdot \xi$, so that τ is the canonical map from $\mathbb{R}^{n+1} \times \mathbb{S}^n$ to $S^*\mathbb{R}^{n+1}$. Then $U_\infty^+ = \{(m, \xi) \mid |m| > R,\ m \cdot \xi \geq 0\}$ and $V_\infty^+ = \{(m, \nu) \mid |m| \geq R,\ m \cdot \nu \geq 0\}$; the difference is that U_∞^+ is considered as a subset of $\mathbb{R}^{n+1} \times \mathbb{S}^n$, while V_∞^+ is considered as a subset of $S^*\mathbb{R}^{n+1}$.

The condition (G6) is required in Proposition 5.12. To explain it, note that under the assumption (G4), if $(m, \xi) \in U_\infty^+$ and $(m(t), \nu(t)) = g^t(\tau(m, \xi))$, then

$$\partial_t \phi_\xi(m(t))|_{t=0} = \partial_m \phi_\xi(m) \cdot \partial_t m(t)|_{t=0} = g(\partial_m \phi_\xi(m), \partial_m \phi_\xi(m)) = 1. \tag{4.4}$$

Therefore, $\partial_\xi \phi_\xi(m)$ is constant on the geodesic passing through $\tau(m, \xi)$.

4.2. **Analytic assumptions.** In this subsection, we formulate the analytic assumptions on plane waves. Let M be as in the previous subsection, Δ be the (nonnegative definite) Laplace–Beltrami operator on M, and $h > 0$ be the semiclassical parameter. We make the following assumptions:

(A1) There exists $c_0 \geq 0$ (equal to 0 for the Euclidean and to $n^2/4$ for the hyperbolic case), such that for each $\lambda > 0$, $h > 0$ and $\xi \in \partial\overline{M}$, there exists a function, called *distorted plane wave*, $E_h(\lambda, \xi; m)$, smooth in all variables and solving on M the differential equation (1.4) in m:

$$(h^2\Delta - c_0 h^2 - \lambda^2) E_h(\lambda, \xi; \cdot) = 0.$$

Here ξ gives the direction of the plane wave, while λ corresponds to its semiclassical energy;

(A2) for each $0 < \lambda_1 \leq \lambda_2$, the Schwartz kernel of the semiclassical spectral projector

$$\Pi_{[\lambda_1,\lambda_2]} := \mathbb{1}_{[\lambda_1^2 + c_0 h^2, \lambda_2^2 + c_0 h^2]}(h^2\Delta)$$

can be written in the form

$$\Pi_{[\lambda_1,\lambda_2]}(m, m') = (2\pi h)^{-n-1} \int_{\lambda_1}^{\lambda_2} \lambda^n f_\Pi(\lambda/h) \int_{\partial\overline{M}} E_h(\lambda, \xi; m) \overline{E_h(\lambda, \xi; m')}\, d\xi d\lambda. \tag{4.5}$$

Here integration in ξ is carried with respect to a certain given volume form $d\xi$ on $\partial\overline{M}$ and $f_\Pi(z) > 0$ is a smooth function of z such that $|\partial_z^k f_\Pi(z)| \leq C_k \langle z\rangle^{-k}$ for each k and $f_\Pi(z) \to 1$ as $z \to \infty$.

We now assume that plane waves admit the decomposition

$$E_h(\lambda, \xi; m) = \chi_0(m; \xi) E_h^0(\lambda, \xi; m) + E_h^1(\lambda, \xi; m), \tag{4.6}$$

where χ_0, E_h^0, E_h^1 are respectively a cutoff function, an explicit 'outgoing' part of the wave, and the 'incoming' part, satisfying more precisely the following properties:

(A3) $\chi_0(m; \xi)$ is a function smooth in $m \in \overline{M}$ and $\xi \in \partial\overline{M}$, supported inside the set U_∞ from (G4) and $\chi_0(m, \xi) = 1$ for m sufficiently close to ξ;

(A4) $E_h^0(\lambda, \xi; m)$ is a smooth function of $\lambda \in \mathbb{R}$ and $(m, \xi) \in U_\infty$, of the form

$$E_h^0(\lambda, \xi; m) = e^{\frac{i\lambda}{h}\phi_\xi(m)} b^0(\lambda, \xi, m; h), \tag{4.7}$$

where U_∞ and ϕ_ξ are defined in (G4) and b^0 is a classical symbol in h defined for $\lambda \in \mathbb{R}$ and $(m, \xi) \in U_\infty$; that is, b^0 is smooth in all variables, including h, up to $h = 0$. We also require that $b^0(\lambda, \xi, m; 0)$ is independent of λ;

(A5) for λ in a fixed compact subset of $(0, \infty)$ and ε_0 defined in (G2), the function

$$\widetilde{E}_h^1(\lambda, \xi; m) := \frac{E_h^1(\lambda, \xi; m)}{1 + \|E_h(\lambda, \xi; m)\|_{L^2(\{x \geq \varepsilon_0\})}} \tag{4.8}$$

is h-tempered in the sense of (3.2);

(A6) for λ in a fixed compact subset of $(0, \infty)$, each $\xi \in \partial\overline{M}$, and each $(m, \lambda\nu) \in \mathrm{WF}_h(\widetilde{E}_h^1(\lambda, \xi))$, we have $(m, \nu) \in S^*M$ and either the geodesic $\gamma(t) = g^t(m, \nu)$ does not escape in the forward direction (i.e. $(m, \nu) \in \Gamma_-$) or there exists $t \geq 0$ such that $\gamma(t)$ lies in the set

$$W_\xi := \{(m, \partial_m \phi_\xi(m)) \mid m \in \mathrm{supp}(\partial_m \chi_0)\}. \tag{4.9}$$

The constants in the corresponding estimates (in the definition of the wave front set of a distribution given in Section 3.1) are uniform in λ and ξ;

(A7) there exists $\varepsilon_1 \in (0, \varepsilon_0)$ such that for $(m, \nu) \in S^*M$ directly escaping in the forward direction and $x(m) \leq \varepsilon_1$, the point $(m, \xi_{+\infty}(m, \nu))$ lies in the set U_∞ defined in (G4) and $\chi_0 = 1$ near this point;

(A8) Let $\tau : U_\infty^+ \to V_\infty^+$ be the diffeomorphism from (G4). Then its Jacobian with respect to the volume measure $d\mathrm{vol}(m)d\xi$ on U_∞^+ and the Liouville measure on V_∞^+, is equal to $|b^0(1, \xi, m; 0)|^2$, with b^0 defined in (A4).

For example, for $M = \mathbb{R}^{n+1}$ we put $c_0 = 0$, $E_h(\lambda, \xi; m) = e^{i\lambda\xi\cdot m/h}$ and use the standard volume form on the sphere $\partial\overline{M} = \mathbb{S}^n$. The equation (4.5) then follows from the Fourier inversion formula.

Let us informally explain how the decomposition (4.6) is constructed and provide a justification for assumptions (A3)–(A6), putting for simplicity $\lambda = 1$. First of all, (A4) implies that for any $\chi \in C_0^\infty(M)$, $\chi\chi_0 E_h^0$, as a function of m, is a Lagrangian distribution associated to the Lagrangian Λ_ξ from (4.3). In fact, in the cases considered in the present paper, E_h^0 solves on its domain the equation (1.4); however, we do not make this assumption here, as in more complicated cases (such as asymptotically hyperbolic manifolds of variable curvature) E_h^0 might only be an approximate solution to (1.4) in a certain sense.

If we assume that E_h^0 solves (1.4) on its domain, then the function

$$F_h(\lambda, \xi; m) = (h^2\Delta - \lambda^2 - c_0h^2)(\chi_0(m)E_h^0(\lambda, \xi; m))$$

is equal to $[h^2\Delta, \chi_0]E_h^0$. Since E_h^0 is a Lagrangian distribution associated to Λ_ξ, the wavefront set of F_h is contained in W_ξ. We will now take $E_h^1 = -R_h(\lambda)F_h$, where $R_h(\lambda)$ is the *incoming scattering resolvent*, a certain right inverse of $h^2\Delta - \lambda^2 - c_0h^2$. Moreover, in our cases $R_h(\lambda)$ will be microlocally incoming in the weak sense: if we multiply F_h by a (possibly small) constant to make $R_h(\lambda)F_h$ bounded polynomially in h, then each point in the wavefront set of $R_h(\lambda)F_h$, when propagated forward by the geodesic flow, will either converge to the trapped set or pass through $\mathrm{WF}_h(F_h)$. Thus, the assumption (A6) should be viewed as a direct consequence of the fact that the scattering resolvent is microlocally incoming and of propagation of singularities.

The assumption (A7) looks less natural, but will play an essential role in our proofs, in Propositions 5.2 and 5.5. It holds for both Euclidean and hyperbolic infinities, but for different reasons. For the hyperbolic infinity, $\chi_0(\cdot; \xi)$ is equal to 1 in a small neighborhood of ξ and one can see that for (m, ν) directly escaping in the forward direction and converging to ξ, the distance from m to ξ in $\overline{M}$ is $\mathcal{O}(x(m))$. This is not true in the Euclidean case; however, in that case χ_0 is equal to 1 outside of a compact subset of M (that is, near the whole boundary $\partial\overline{M}$, not just near ξ).

The assumption (A8) is required to relate the natural measure arising from the function E_h^0 to the Liouville measure. If E_h were equal to E_h^0, then this assumption would simply follow by taking the trace in (4.5) with a compactly supported pseudodifferential operator and a smooth cutoff function in λ.

4.3. **Limiting measures.** We now define the family of limiting measures μ_ξ. These measures result from propagating the natural measure arising from the 'outgoing' part E_h^0 of the plane wave, which is supported on the Lagrangian Λ_ξ from (4.3), backwards along the geodesic flow. In contrast with [Dy2], where the exponential decay of the measure along the flow ensured its convergence, our measures will only be defined for almost every ξ.

We first define the measure $\tilde\mu_\xi$ on S^*M, corresponding to E_h^0, as follows: for each compactly supported continuous function a on S^*M, put

$$\int_{S^*M} a\,d\tilde\mu_\xi = \int_{(m,\xi)\in U_\infty^+} |b^0(1,\xi,m;0)|^2 a(\tau(m,\xi))\ \mathrm{dvol}(m). \tag{4.10}$$

The support of $\tilde\mu_\xi$ is contained in the Lagrangian Λ_ξ from (4.3) and the integral (4.10) depends continuously on ξ. We see from (A8) that for any continuous function f on $\partial\overline{M}$,

$$\int_{\partial\overline{M}} f(\xi)\int_{S^*M} a(m,\nu)\,d\tilde\mu_\xi(m,\nu)d\xi = \int_{V_\infty^+} f(\xi_{+\infty}(m,\nu))a(m,\nu)\,d\mu_L(m,\nu). \tag{4.11}$$

We now want to define the measure μ_ξ by

$$\int_{S^*M} a\,d\mu_\xi = \lim_{t\to+\infty}\int_{S^*M} a\circ g^{-t}\,d\tilde\mu_\xi, \tag{4.12}$$

valid for all compactly supported continuous functions a. To show that the limit exists for almost every ξ (chosen independently of a) and for every a, we will use monotonicity. By (4.11), (G5), and using the invariance of the function $\xi_{+\infty}$ and the Liouville measure μ_L under the geodesic flow, we see that if a and f are nonnegative, then

$$\int_{\partial\overline{M}} f(\xi)\int_{S^*M} (a\circ g^{-t})\,d\tilde\mu_\xi d\xi = \int_{g^{-t}(V_\infty^+)} f(\xi_{+\infty}(m,\nu))a(m,\nu)\,d\mu_L(m,\nu)$$

is increasing with t. Therefore, for each ξ the integral

$$I_{a,t}(\xi) = \int_{S^*M}(a\circ g^{-t})\,d\tilde\mu_\xi$$

is increasing in t for any nonnegative a. Moreover, the integral of $I_{a,t}(\xi)$ in ξ is bounded by a t-independent constant, namely by the integral of a by the Liouville measure. Taking a to be an approximation of the characteristic function of each member of a countable family of compact sets exhausting S^*M, and using the monotone convergence theorem, we see that there exists a measure zero set $\mathcal{X}\subset\partial\overline{M}$ such that for $\xi\notin\mathcal{X}$, we have for each j and for any compactly supported continuous function a,

$$\lim_{t\to+\infty}\int_{S^*M}(a\circ g^{-t})\,d\tilde\mu_\xi < \infty.$$

This limit is a continuous functional on the space of continuous compactly supported functions on S^*M; therefore, there exists unique Borel measure μ_ξ such that (4.12) holds. Moreover, we see that the limit (4.12) is uniform in a, as soon as we fix a compact set containing $\operatorname{supp} a$ and impose a bound on $\sup_{S^*M}|a|$. One also sees immediately (1.6), namely that for compactly supported continuous a,

$$\operatorname{supp} a\cap\overline{\xi_{+\infty}^{-1}(\xi)} = \emptyset \Longrightarrow \int_{S^*M} a\,d\mu_\xi = 0,$$

as $\int_{S^*M}(a\circ g^{-t})\,d\tilde\mu_\xi = 0$ for all t.

We can integrate the measure μ_ξ in ξ, getting back the Liouville measure:

Proposition 4.3. *For each $f \in C^\infty(\partial\overline{M})$ and each $a \in C_0^\infty(S^*M)$ we have*

$$\int_{\partial\overline{M}} f(\xi) \int_{S^*M} a(m,\nu)\, d\mu_\xi(m,\nu) d\xi = \int_{S^*M\setminus\Gamma_-} f(\xi_{+\infty}(m,\nu)) a(m,\nu)\, d\mu_L(m,\nu). \tag{4.13}$$

In particular, if $\mu_L(\Gamma_-) = 0$ (which will always be the case in our theorems, see (5.2)), then $\int \mu_\xi\, d\xi$ is the Liouville measure.

Proof. The left-hand side can be written as

$$\lim_{t\to+\infty} \int_{g^{-t}(V_\infty^+)} f(\xi_{+\infty}(m,\nu)) a(m,\nu)\, d\mu_L(m,\nu).$$

It remains to use the dominated convergence theorem; indeed, the function under the integral is bounded and compactly supported, we have $g^{-t_1}(V_\infty^+) \subset g^{-t_2}(V_\infty^+)$ for $t_1 < t_2$, and the union of $g^{-t}(V_\infty^+)$ over all $t \in \mathbb{R}$ is exactly $S^*M \setminus \Gamma_-$, as for every geodesic $\gamma(t)$ escaping in the forward direction and for t large enough, the point $\gamma(t)$ is directly escaping in the forward direction and $(\gamma(t), \xi_{+\infty}(\gamma(t))) \in U^\infty$. □

Finally, the following lemma will be useful to relate our measure μ_ξ to the one obtained from E_h^0 in the proofs of Theorems 1 and 2:

Lemma 4.4. *Let $\xi \notin \mathcal{X}$, so that μ_ξ is well-defined. Let a be a compactly supported continuous function on S^*M.*

1. *μ_ξ is invariant under the geodesic flow: for each $t \in \mathbb{R}$,*

$$\int_{S^*M} a \circ g^t\, d\mu_\xi = \int_{S^*M} a\, d\mu_\xi. \tag{4.14}$$

2. *If $\operatorname{supp} a \subset \mathcal{DE}_+ \cap \{x \le \varepsilon_1\}$, where $\mathcal{DE}_+$ is given by Definition 4.2 and ε_1 is defined in (A7), then*

$$\int_{(m,\xi)\in U_\infty} |b^0(1,\xi,m;0)\chi_0(m;\xi)|^2 a(m, \partial_m\phi_\xi(m))\ \mathrm{dvol}(m) = \int_{S^*M} a\, d\mu_\xi. \tag{4.15}$$

Proof. 1. Follows immediately from the definition (4.12).

2. First of all, note that for m in the support of the function $a(m, \partial_m\phi_\xi(m))$, we have $(m,\xi) \in U_\infty^+$ and $\chi_0(m;\xi) = 1$ by (A7); therefore, the left-hand side of (4.15) becomes the integral of a over the measure $\tilde\mu_\xi$ defined in (4.10). By (4.12), it is enough to show that for $t \ge 0$,

$$\int_{S^*M} a \circ g^{-t}\, d\tilde\mu_\xi = \int_{S^*M} a\, d\tilde\mu_\xi.$$

For that, it is enough to show that for each $f \in C_0^\infty(\partial\overline{M})$,

$$\int_{\partial M} f(\xi) \int_{S^*M} a \circ g^{-t}\, d\tilde\mu_\xi = \int_{\partial M}\int_{S^*M} a\, d\tilde\mu_\xi.$$

Using (4.11), we rewrite this as

$$\int_{g^{-t}(V_\infty^+)} f(\xi_{+\infty}) a\, d\mu_L = \int_{V_\infty^+} f(\xi_{+\infty}) a\, d\mu_L.$$

This is true as $\operatorname{supp} a \subset V_\infty^+ \subset g^{-t}(V_\infty^+)$. □

4.4. **Averaged estimates on plane waves.** One of the principal tools of the present paper are microlocal estimates on the plane waves $E_h(\lambda,\xi)$ *on average* in λ,ξ, where λ takes values in a size h interval. They are direct consequences of (4.5) and the Hilbert–Schmidt norm estimate (3.24). More precisely, restricting to the case $\lambda = 1 + \mathcal{O}(h)$ for simplicity, we have the following

Proposition 4.5. *Let $\chi \in C_0^\infty(M)$. Then:*

1. $\chi\Pi_{[1,1+h]}$ is a Hilbert–Schmidt operator and there exists a global constant C such that for each bounded operator $A : L^2(M) \to L^2(M)$, we have

$$h^{-1}\|A\chi(m)E_h(\lambda,\xi;m)\|^2_{L^2_{m,\xi,\lambda}(M\times\partial\overline{M}\times[1,1+h])} \leq Ch^n\|A\chi\Pi_{[1,1+h]}\|^2_{\mathrm{HS}}. \tag{4.16}$$

2. The functions χE_h are bounded in L^2 on average in the following sense: there exists a constant $C(\chi)$ such that for any h,

$$h^{-1}\|\chi(m)E_h(\lambda,\xi;m)\|^2_{L^2_{m,\xi,\lambda}(M\times\partial\overline{M}\times[1,1+h])} \leq C(\chi). \tag{4.17}$$

The h^{-1} prefactor in both cases is due to the fact that we are integrating over an interval of size h in λ.

Proof. 1. It follows immediately from (4.5) that

$$h^{-1}\int_1^{1+h} f_\Pi(\lambda/h)\lambda^n\int_{\partial\overline{M}}(\chi E_h(\lambda,\xi))\otimes(\chi E_h(\lambda,\xi))\,d\xi d\lambda = (2\pi)^{n+1}h^n\chi\Pi_{[1,1+h]}\bar{\chi}. \tag{4.18}$$

Here $\otimes$ denotes the Hilbert tensor product, defined in (2.8). The integral on the left-hand side converges in the trace class norm, as the Schwartz kernels of the integrated operators are smooth and compactly supported. Therefore, $\chi\Pi_{[1,1+h]}\bar{\chi}$ is trace class. Since

$$\chi\Pi_{[1,1+h]}\bar{\chi} = (\chi\Pi_{[1,1+h]})(\chi\Pi_{[1,1+h]})^*,$$

we see that $\chi\Pi_{[1,1+h]}$ is a Hilbert–Schmidt operator. Now, multiplying both sides of (4.18) by A on the left and A^* on the right and taking the trace, we get

$$\begin{aligned} h^{-1}\|\lambda^{n/2}f_\Pi(\lambda/h)^{1/2}A\chi(m)E_h(\lambda,\xi;m)\|^2_{L^2_{m,\xi,\lambda}(M\times\partial\overline{M}\times[1,1+h])}\\ = (2\pi)^{n+1}h^n\operatorname{Tr}((A\chi\Pi_{[1,1+h]})(A\chi\Pi_{[1,1+h]})^*)\\ = (2\pi)^{n+1}h^n\|A\chi\Pi_{[1,1+h]}\|^2_{\mathrm{HS}}. \end{aligned} \tag{4.19}$$

2. We would like to use Lemma 3.11 to estimate $\|\chi\Pi_{[1,1+h]}\|_{\mathrm{HS}}$ (we can put χ on the other side of the projector in (3.24) by taking the adjoint), however this is not directly possible as χ is not compactly microlocalized. We thus use that E_h solve the equation (1.4), writing by the elliptic parametrix construction (same as for the proof of Proposition 3.2)

$$\chi = B + Q_\lambda(h^2\Delta - \lambda^2 - c_0h^2) + R_\lambda \tag{4.20}$$

for $\lambda \in [1,1+h]$, where $B \in \Psi^{\mathrm{comp}}(M)$, $Q_\lambda \in \Psi^{-2}(M)$, and $R_\lambda \in h^\infty\Psi^{-\infty}(M)$ are compactly supported and B is independent of λ and equal to χ microlocally near S^*M. We can also assume that Q_λ and R_λ are smooth in λ. Now, we substitute (4.20) into the left-hand side of (4.17) and use the triangle inequality. By (4.16), the term featuring B is bounded by a constant times $h^n\|B\Pi_{[1,1+h]}\|^2_{\mathrm{HS}}$, which is bounded uniformly in h by Lemma 3.11. The term featuring Q_λ is zero by (1.4).

Finally, we show that the term featuring R_λ is $\mathcal{O}(h^\infty)$. This does not follow immediately from (4.16), as the operator R_λ depends on λ. We use the following variant of (4.19): for $\tilde\lambda \in [1,1+h]$,

$$h^{-1}\|\lambda^{n/2}f_\Pi(\lambda/h)^{1/2}R_{\tilde\lambda}E_h(\lambda)\|^2_{L^2_{m,\xi,\lambda}(M\times\partial\overline{M}\times[1,\tilde\lambda])} = (2\pi)^{n+1}h^n\|R_{\tilde\lambda}\Pi_{[1,\tilde\lambda]}\|^2_{\mathrm{HS}}.$$

Differentiating in $\tilde\lambda$, we get

$$(2\pi)^{n+1}h^n\partial_\lambda\|R_{\tilde\lambda}\Pi_{[1,\tilde\lambda]}\|^2_{\mathrm{HS}} = h^{-1}\|\tilde\lambda^{n/2}f_\Pi(\tilde\lambda/h)^{1/2}R_{\tilde\lambda}E_h(\tilde\lambda)\|^2_{L^2(m,\xi)(M\times\partial M)}+$$
$$2h^{-1}\operatorname{Re}\langle\lambda^{n/2}f_\Pi(\lambda/h)^{1/2}(\partial_{\tilde\lambda}R_{\tilde\lambda})E_h(\lambda),\lambda^{n/2}f_\Pi(\lambda/h)^{1/2}R_{\tilde\lambda}E_h(\lambda)\rangle_{L^2_{m,\xi,\lambda}(M\times\partial M\times[1,\tilde\lambda])}.$$

We now integrate in $\tilde\lambda$ from 1 to $1+h$. The integral of the left-hand side is bounded by a constant times $h^n\|R_{1+h}\|^2_{\mathrm{HS}} = \mathcal{O}(h^\infty)$. The integral of the first term on the right-hand side is the quantity we are estimating. Finally, the second term on the right-hand side is bounded by a constant times $h^n|\operatorname{Tr}((\partial_{\tilde\lambda}R_{\tilde\lambda})\Pi_{[1,1+h]}R^*_{\tilde\lambda})|$, which is $\mathcal{O}(h^\infty)$ uniformly in $\tilde\lambda$, as the Hilbert–Schmidt norms of both $R_{\tilde\lambda}$ and $\partial_{\tilde\lambda}R_{\tilde\lambda}$ are $\mathcal{O}(h^\infty)$. □

5. Proofs

5.1. Proof of Theorem 1. In this section, we prove the convergence Theorem 1 under the following assumption:

$$\mu_L(K) = 0, \tag{5.1}$$

where μ_L denotes the Liouville measure on S^*M and K is the trapped set. First of all, note that (5.1) implies

$$\mu_L(\Gamma_\pm) = 0. \tag{5.2}$$

Indeed, fix $\varepsilon \in (0,\varepsilon_0)$, where ε_0 is defined in (G2), and take the set $\Gamma_+^\varepsilon = \Gamma_+ \cap \{x \geq \varepsilon\}$. For $(m,\nu) \in \Gamma_+ \cap \{x = \varepsilon\}$, we have $\dot x(m,\nu) < 0$; indeed, otherwise (m,ν) directly escapes in the backward direction and thus cannot lie in Γ_+. It follows that $g^{-t}(\Gamma_+^\varepsilon) \subset \Gamma_+^\varepsilon$ for $t \geq 0$. Since Γ_+^ε is bounded, and μ_L is invariant under the geodesic flow, we have

$$\mu_L(\Gamma_+^\varepsilon) = \lim_{t\to+\infty}\mu_L(g^{-t}(\Gamma_+^\varepsilon)) = \mu_L\Big(\bigcap_{t\geq 0} g^{-t}(\Gamma_+^\varepsilon)\Big) = \mu_L(K) = 0.$$

Letting $\varepsilon \to 0$, we get (5.2).

We next note that the averaged L^2 bound (4.17) on E_h on compact sets, together with (1.4) and the elliptic Proposition 3.2, give the following

Proposition 5.1. *Assume that $A \in \Psi^0(M)$ is compactly supported and $\mathrm{WF}_h(A) \cap S^*M = \emptyset$. Then*

$$h^{-1}\|\langle AE_h(\lambda,\xi), E_h(\lambda,\xi)\rangle\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])} = \mathcal{O}(h^\infty). \tag{5.3}$$

Therefore, it is enough to prove (1.5) for a compactly supported $A \in \Psi^{\mathrm{comp}}(M)$ microlocalized in an arbitrarily small neighborhood of S^*M.

Take $t > 0$; we will calculate limits of the form $\lim_{t\to+\infty}\lim_{h\to 0}$, thus $\mathcal{O}_t(h^\infty)$ expressions (that is, expressions that are $\mathcal{O}(h^\infty)$ with the constants depending on t) will be negligible. Take $\chi \in C_0^\infty(M)$ indepedendent of t and such that $A = \chi A \chi$. We first use that E_h is a generalized eigenfunction of the Laplacian (1.4) and apply Lemma 3.10: for each $\lambda \in [1, 1+h]$ and each $\xi \in \partial M$,

$$\chi E_h = \chi e^{-it(\lambda^2+c_0h^2)/2h} U(t)\chi_t E_h + \mathcal{O}_t(h^\infty \|E_h\|_{L^2(K_t)})_{L^2}. \tag{5.4}$$

Here $U(t) = e^{ith\Delta/2}$ is the semiclassical Schrödinger propagator and $\chi_t \in C_0^\infty(M)$ is supported in the interior of the compact set $K_t \subset M$ and satisfies $d_g(\operatorname{supp}\chi, \operatorname{supp}(1-\chi_t)) > t$. We also assume that $|\chi_t| \leq 1$ everywhere and K_t contains $\{x \geq \varepsilon_0\}$, where ε_0 is defined in (G2). By Proposition 3.8, we can write $U(-t)AU(t) = A^{-t} + \mathcal{O}_t(h^\infty)_{L^2\to L^2}$, where $A^{-t} \in \Psi^{\mathrm{comp}}$ is compactly supported. Then

$$\langle AE_h, E_h\rangle = \langle A^{-t}\chi_t E_h, \chi_t E_h\rangle + \mathcal{O}_t(h^\infty)\|E_h\|^2_{L^2(K_t)}. \tag{5.5}$$

We will now write

$$A^{-t} = A_0^{-t} + A_1^{-t},\ A_0^{-t} := A^{-t}\varphi,\ A_1^{-t} := A^{-t}(1-\varphi), \tag{5.6}$$

where the L^2 norm of the principal symbol of A_0^{-t} will decay with t and the operator A_1^{-t} will be negligible on E_h^1. The function $\varphi \in C_0^\infty(M)$ is taken independent of t and such that $\operatorname{supp}\chi \subset \{x > \varepsilon_\chi\}$ for some ε_χ and $\varphi = 1$ near $\{x \geq \varepsilon_\chi\}$. We also require that $\varphi = 1$ near $\{x \geq \varepsilon_1\}$, where ε_1 comes from the assumption (A7).

We first show that the terms in (5.5) featuring both A_1^{-t} and E_h^1 are $\mathcal{O}(h^\infty)$. For that, we need to show that the trajectories in $\mathrm{WF}_h(A_1^{-t}) \subset \operatorname{supp}(1-\varphi)\cap g^t(\operatorname{supp}\chi)$ satisfy the geometric property shown on Figure 1:

Lemma 5.2. *Let $t \geq 0$. Assume that $(m,\nu) \in S^*M$ satisfies $m \in \operatorname{supp}(1-\varphi)$, but $g^{-t}(m,\nu) \in \operatorname{supp}\chi$. Then:*

1. *(m,ν) directly escapes in the forward direction, in the sense of Definition 4.2;*
2. *for each $s \geq 0$, $g^s(m,\nu)$ does not lie in the set W_ξ defined in (4.9), for any $\xi \in \partial\overline{M}$.*

Proof. (1) We have $x(m) < \varepsilon_1 \leq \varepsilon_0$; therefore, if (m,ν) does not directly escape in the forward direction, then it directly escapes in the backward direction; this would imply that $x(g^{-t}(m,\nu))$ is decreasing in $t \geq 0$, which is impossible as $x(m) < \varepsilon_\chi < x(g^{-t}(m,\nu))$.

(2) The point $g^s(m,\nu)$ directly escapes in the forward direction and $x(g^s(m,\nu)) < \varepsilon_1$. If $g^s(m,\nu) \in W_\xi$, then by (G4), $\xi = \xi_{+\infty}(m,\nu)$, but this is impossible as $\chi_0 = 1$ near $(g^s(m,\nu), \xi_{+\infty}(m,\nu))$ by (A7). □

Combining Lemma 5.2 with the microlocal information we have on E_h^1, we get

Proposition 5.3. *If $E_h = \chi_0 E_h^0 + E_h^1$ is the decomposition (4.6), then for each $t \geq 0$,*

$$\begin{aligned}\langle AE_h, E_h\rangle &= \langle A_1^{-t}\chi_t\chi_0 E_h^0, \chi_t\chi_0 E_h^0\rangle \\ &+\langle A_0^{-t}\chi_t E_h, \chi_t E_h\rangle + \mathcal{O}_t(h^\infty(1+\|E_h\|^2_{L^2(K_t)})).\end{aligned} \tag{5.7}$$

where A_0^{-t}, A_1^{-t} *are defined in* (5.6).

Proof. By (5.5), it is enough to show that

$$\langle A_1^{-t}\chi_t E_h, \chi_t E_h\rangle - \langle A_1^{-t}\chi_t\chi_0 E_h^0, \chi_t\chi_0 E_h^0\rangle = \mathcal{O}_t(h^\infty(1+\|E_h\|^2_{L^2(K_t)})).$$

Given that $\|\chi_0 E_h^0\|_{L^2(K_t)} = \mathcal{O}(1)$, it suffices to prove

$$\|B\chi_t E_h^1\|_{L^2} = \mathcal{O}_t(h^\infty(1+\|E_h\|_{L^2(K_t)})),$$

where B is equal to either A_1^{-t} or its adjoint. This in turn follows from

$$\|B\chi_t \widetilde{E}_h^1\|_{L^2} = \mathcal{O}_t(h^\infty), \tag{5.8}$$

with $\widetilde{E}_h^1$ defined in (4.8). Take $(m,\nu)\in \mathrm{WF}_h(B\chi_t\widetilde{E}_h^1)\subset S^*M$. Then by Proposition 3.8,

$$(m,\nu)\in \mathrm{WF}_h(B)\subset \mathrm{WF}_h(A^{-t})\cap \operatorname{supp}(1-\varphi)\subset g^t(\mathrm{WF}_h(A))\cap\operatorname{supp}(1-\varphi).$$

Since $\mathrm{WF}_h(A)\subset\operatorname{supp}\chi$, we see that $m\in\operatorname{supp}(1-\varphi)$ and $g^{-t}(m,\nu)\in\operatorname{supp}\chi$; therefore, by Lemma 5.2, the geodesic $g^s(m,\nu)$ escapes in the forward direction and does not pass through W_ξ for $s\geq 0$. But then by (A6) the point (m,ν) cannot lie in $\mathrm{WF}_h(\widetilde{E}_h^1)$, a contradiction. We showed that the wavefront set of $B\chi_t\widetilde{E}_h^1$ is empty, which implies (5.8). □

We now use the averaged estimate (4.16) and the Hilbert–Schmidt norm estimates from Section 3.3, to estimate the second term on the right-hand side of (5.7):

Proposition 5.4. *There exists a constant* C *independent of* t *such that*

$$h^{-1}\|\langle A_0^{-t}\chi_t E_h,\chi_t E_h\rangle\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])}\leq C\|(\sigma(A)\circ g^{-t})\varphi\|_{L^2(S^*M)}+\mathcal{O}_t(h). \tag{5.9}$$

Here $\|a\|_{L^2(S^*M)}$ *is the* L^2 *norm of the restriction of* a *to* S^*M *with respect to the Liouville measure.*

Proof. Take a real-valued function $\varphi_1\in C_0^\infty(M)$ independent of t such that $\varphi_1=1$ near $\operatorname{supp}\varphi$. Then the left-hand side of (5.9) is bounded by

$$h^{-1}\|\langle A_0^{-t}\chi_t E_h,\varphi_1\chi_t E_h\rangle\|_{L^1_{\xi,\lambda}}+h^{-1}\|\langle(1-\varphi_1)A_0^{-t}\chi_t E_h,\chi_t E_h\rangle\|_{L^1_{\xi,\lambda}},$$

where the L^1, and later L^2, norms in ξ,λ are taken over $\partial\overline{M}\times[1,1+h]$. The second term here is $\mathcal{O}_t(h^\infty)$ by the bound (4.17) and since $(1-\varphi_1)A_0^{-t}=\mathcal{O}_t(h^\infty)_{L^2\to L^2}$ is compactly supported. The first term can be estimated by applying the Cauchy–Schwarz inequality first in m and then in (λ,ξ):

$$\begin{aligned}h^{-1}\|\langle A_0^{-t}\chi_t E_h,\varphi_1\chi_t E_h\rangle\|_{L^1_{\xi,\lambda}}&\leq h^{-1}\|\,\|A_0^{-t}\chi_t E_h\|_{L^2(M)}\cdot\|\varphi_1\chi_t E_h\|_{L^2(M)}\|_{L^1_{\xi,\lambda}}\\&\leq h^{-1/2}\|A_0^{-t}\chi_t E_h\|_{L^2_{m,\xi,\lambda}}\cdot h^{-1/2}\|\varphi_1\chi_t E_h\|_{L^2_{m,\xi,\lambda}}.\end{aligned}$$

Now, $h^{-1/2}\|\varphi_1\chi_t E_h\|_{L^2_{m,\xi,\lambda}}$ is bounded (independently of t) uniformly in h by (4.17). As for $h^{-1/2}\|A_0^{-t}\chi_t E_h\|_{L^2_{m,\xi,\lambda}}$, we can estimate it using (4.16) by a constant times

$$h^{n/2}\|A_0^{-t}\chi_t\Pi_{[1,1+h]}\|_{\mathrm{HS}}.$$

Note that the operator $A_0^{-t}\chi_t\in\Psi^{\mathrm{comp}}$ is compactly supported and it is compactly microlocalized independently of t. It then remains to apply (3.24) (to the adjoint of our operator); by Proposition 3.8, the principal symbol of $A_0^{-t}\chi_t$ is given by $(\sigma(A)\circ g^{-t})\varphi$. □

We now use the dynamical assumption that $\mu_L(K)=0$. The function $(\sigma(A)\circ g^{-t})\varphi$ is supported in a t-independent compact set and bounded uniformly in t. Moreover, it converges to zero pointwise on $S^*M\setminus\Gamma_+$ as $t\to+\infty$. Therefore, by (5.2) and the dominated convergence theorem we have $(\sigma(A)\circ g^{-t})\varphi\to 0$ in $L^2(S^*M)$, as $t\to+\infty$. It then follows from (5.7) together with the bound (4.17) and from (5.9) that

$$\lim_{t\to+\infty}\limsup_{h\to 0} h^{-1}\|\langle AE_h,E_h\rangle-\langle A_1^{-t}\chi_t\chi_0E_h^0,\chi_t\chi_0E_h^0\rangle\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])}=0.$$

To prove Theorem 1, it now remains to show that

$$\lim_{t\to+\infty}\limsup_{h\to 0} h^{-1}\left\|\langle A_1^{-t}\chi_t\chi_0E_h^0,\chi_t\chi_0E_h^0\rangle-\int_{S^*M}\sigma(A)\,d\mu_\xi\right\|_{L^1_\xi(\partial\overline{M})}=0 \tag{5.10}$$

uniformly in $\lambda=1+\mathcal{O}(h)$. We first note that by (4.7) the function

$$\chi_t\chi_0E_h^0(\lambda,\xi;m)=e^{\frac{i\lambda}{h}\phi_\xi(m)}\chi_t(m)\chi_0(m,\xi)b^0(1,\xi,m;0)+\mathcal{O}_t(h)_{L^2}$$

is a compactly supported Lagrangian distribution associated to the Lagrangian Λ_ξ from (4.3). Therefore, by Proposition 3.5, we find

$$A_1^{-t}\chi_t\chi_0E_h^0(\lambda,\xi)=e^{\frac{i\lambda}{h}\phi_\xi}\chi_t\chi_0b^0(1,\xi,m;0)\sigma(A_1^{-t})(m,\partial_m\phi_\xi(m))+\mathcal{O}_t(h)_{L^2}. \tag{5.11}$$

Therefore,

$$\langle A_1^{-t}\chi_t\chi_0E_h^0,\chi_t\chi_0E_h^0\rangle=\int_M\sigma(A_1^{-t})(m,\partial_m\phi_\xi(m))|\chi_t\chi_0b^0(1,\xi,m;0)|^2\,d\mathrm{vol}(m)+\mathcal{O}_t(h).$$

Now, by Proposition 3.8, $\sigma(A_1^{-t})=(\sigma(A)\circ g^{-t})(1-\varphi)$. By Lemma 5.2, this function is supported in $\mathcal{DE}_+\cap\{x<\varepsilon_1\}$, with $\mathcal{DE}_+$ from Definition 4.2. Also, $\chi_t=1$ near $\operatorname{supp}\sigma(A_1^{-t})$. Then by part 2 of Lemma 4.4,

$$\langle A_1^{-t}\chi_t\chi_0E_h^0,\chi_t\chi_0E_h^0\rangle=\int_{S^*M}(\sigma(A)\circ g^{-t})(1-\varphi)\,d\mu_\xi+\mathcal{O}_t(h). \tag{5.12}$$

Therefore, (5.10) reduces to

$$\lim_{t\to+\infty}\left\|\int_{S^*M}(\sigma(A)\circ g^{-t})(1-\varphi)\,d\mu_\xi-\int_{S^*M}\sigma(A)\,d\mu_\xi\right\|_{L^1_\xi(\partial\overline{M})}=0. \tag{5.13}$$

By part 1 of Lemma 4.4 and (4.13), we write the norm on the left-hand side of (5.13) as

$$\left\|\int_{S^*M}\sigma(A)(\varphi\circ g^t)\,d\mu_\xi\right\|_{L^1_\xi(\partial\overline{M})}\leq\int_{S^*M}|\sigma(A)(\varphi\circ g^t)|\,d\mu_L.$$

The expression under the integral on the right-hand side is bounded and compactly supported uniformly in t and converges to zero pointwise on $S^*M\setminus\Gamma_-$; by (5.2) and the dominated convergence theorem, we get (5.13). This finishes the proof of Theorem 1.

The nontrapped case. We briefly discuss the situation when $\mathrm{WF}_h(A)\cap\Gamma_-=\emptyset$. In this case, for t large enough (depending on A), for any $(m,\nu)\in\mathrm{WF}_h(A)$ we have $g^t(m,\nu)\notin\operatorname{supp}\varphi$ and thus

$$A_0^{-t}=\mathcal{O}(h^\infty)_{L^2\to L^2}.$$

Then by (5.7) and the bound (4.17),

$$\langle AE_h,E_h\rangle=\langle A_1^{-t}\chi_t\chi_0E_h^0,\chi_t\chi_0E_h^0\rangle+\mathcal{O}(h^\infty)_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])}.$$

The quantity $\langle A_1^{-t}\chi_t\chi_0E_h^0,\chi_t\chi_0E_h^0\rangle$ is calculated in (5.12) up to $\mathcal{O}(h)$. However, since E_h^0 is a Lagrangian distribution, one can get by Proposition 3.5 a full expansion of this quantity in powers of h; this yields

$$\langle AE_h(\lambda,\xi),E_h(\lambda,\xi)\rangle=\sum_{0\leq j<N}h^j\int_{S^*M}L_ja\,d\mu_\xi+\mathcal{O}(h^{N+1})_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])},\tag{5.14}$$

where $A=\mathrm{Op}_h(a)$ for some symbol a and some quantization procedure Op_h and each L_j is a differential operator of order $2j$ on T^*M, with $L_0=1$.

5.2. **Estimates on the remainder.** In this subsection, we prove (1.12) and establish an approximation fact (Proposition 5.8) used in the proofs of (1.13) and Theorem 4.

Classical escape rate and Ehrenfest time. Let $K_0\subset M$ be a compact geodesically convex set containing a neighborhood of the projection of the trapped set K onto M. As in (1.9), define the set

$$\mathcal{T}(t)=\{(m,\nu)\in S^*M\mid m\in K_0,\ g^t(m,\nu)\in K_0\}.$$

The choice of K_0 does not matter here: if $K_0'\subset M$ is another set with same properties and $\mathcal{T}'(t)$ is defined using K_0' in place of K_0, then there exists a constant $T_0>0$ such that for each $T\geq T_0$ and $t\geq 0$,

$$g^T(\mathcal{T}'(t+2T))\subset\mathcal{T}(t).\tag{5.15}$$

Indeed, assume that (5.15) were false. Then there exists sequences $T_j\to+\infty$, $t_j\geq 0$, and $(m_j,\nu_j)\in S^*M$ such that $g^{-T_j}(m_j,\nu_j)$ and $g^{t_j+T_j}(m_j,\nu_j)$ both lie in K_0', but for each j, either (1) $(m_j,\nu_j)\notin K_0$ or (2) $g^{t_j}(m_j,\nu_j)\notin K_0$. We may assume that case (1) holds for all j; case (2) is handled similarly, reversing the direction of the flow and taking $g^{t_j}(m_j,\nu_j)$ in place of (m_j,ν_j). Take $\varepsilon>0$ such that $K_0'\subset\{x\geq\varepsilon\}$; since $\{x\geq\varepsilon\}$ is geodesically convex for ε small enough, we have $(m_j,\nu_j)\in\{x\geq\varepsilon\}$. Passing to a subsequence, we can assume that $(m_j,\nu_j)\to(m,\nu)\in S^*M$ as $j\to+\infty$. Now, since $g^{-T_j}(m_j,\nu_j)\in K_0'$ and $T_j\to+\infty$, we have $(m,\nu)\in\Gamma_+$ (indeed, otherwise there would exist $s>0$ such that $g^{-s}(m,\nu)\in\{x<\varepsilon\}$ and this would also hold in a neighborhood of (m,ν)). Similarly, since $g^{t_j+T_j}(m_j,\nu_j)\in K_0'$ and $t_j+T_j\to+\infty$, we have $(m,\nu)\in\Gamma_-$. It follows that $(m,\nu)\in K$, which is impossible, as each (m_j,ν_j) does not lie in K_0, which contains a neighborhood of K.

By changing Λ_0 slightly and using (5.15), we see that the choice of K_0 does not matter for the validity of (1.12) and (1.13); more precisely, if $\Lambda_0>\Lambda_0'$, then $r'(h,\Lambda_0')\leq Cr(h,\Lambda_0)$, where r' is defined by (1.10) using $\mathcal{T}'$ in place of $\mathcal{T}$. Also, the maximal expansion rate $\Lambda_{\max}$ defined in (1.11) does not depend on the choice of K_0.

We now choose a geodesically convex K_0 such that its interior contains the supports of all cutoff functions and compactly supported operators used in the argument below. We will rely on Proposition 3.9 (with U equal to the interior of K_0); we let $\Lambda_0 > \Lambda_{\max}$ and fix $\varepsilon_e > 0$ and Λ_0' such that $\Lambda_0 > \Lambda_0' > (1 + 2\varepsilon_e)\Lambda_{\max}$. Define the Ehrenfest time

$$t_e := \log(1/h)/(2\Lambda_0). \tag{5.16}$$

Then when propagating an operator in Ψ^{comp} microlocalized inside

$$\mathcal{E}_{\varepsilon_e} := \{1 - \varepsilon_e \leq |\nu|_g \leq 1 + \varepsilon_e\} \tag{5.17}$$

with cutoffs supported inside K_0, as in Proposition 3.9, for time $t = lt_0 \in [-t_e, t_e]$, we get a mildly exotic pseudodifferential operator in $\Psi^{\mathrm{comp}}_{\rho_e}$, where

$$\rho_e := t_e\Lambda_0' / \log(1/h) = \Lambda_0'/(2\Lambda_0) < 1/2. \tag{5.18}$$

First decomposition of $\langle AE_h, E_h\rangle$. By Proposition 5.1, we may assume that $A \in \Psi^{\mathrm{comp}}(M)$ is compactly supported and microlocalized inside the set $\mathcal{E}_{\varepsilon_e}$ defined in (5.17).

We first establish the following decomposition similar to (5.7):

$$\begin{aligned}\langle AE_h, E_h\rangle &= e^{il\beta}\langle A(\varphi U(t_0))^l \varphi E_h, E_h\rangle \\ &+ \sum_{j=1}^{l} e^{ij\beta}\langle A(\varphi U(t_0))^j (1-\varphi)\varphi_{t_0}\chi_0 E_h^0, E_h\rangle + \mathcal{O}(h^\infty \mathcal{N}(E_h)^2),\end{aligned} \tag{5.19}$$

uniformly in $\xi \in \partial\overline{M}$ and $\lambda \in [1, 1+h]$. Here $l = \mathcal{O}(\log(1/h))$ is a nonnegative integer and $t_0 > 0$ and $\varphi, \varphi_{t_0} \in C_0^\infty(M)$, specified below, are independent of j. The quantity $\mathcal{N}(E_h)$, defined in (5.22), is related to the L^2 norm of E_h on a certain compact set, and is bounded on average by (5.23). The real-valued parameter β is equal to

$$\beta = -t_0(\lambda^2 + c_0 h^2)/(2h) \tag{5.20}$$

and will not play a big role in our argument.

To show (5.19), we start by considering the functions $\varphi, \varphi_1, \varphi_2 \in C_0^\infty(M)$ such that:

- $0 \leq \varphi, \varphi_1, \varphi_2 \leq 1$ everywhere,
- $\varphi = 1$ near $\operatorname{supp}\varphi_2$ and $\varphi_1 = 1$ near $\operatorname{supp}\varphi$, and
- $\varphi_2 = 1$ both near the support of A and near the set $\{x \geq \varepsilon_1\}$, with ε_1 defined in (A7).

The proof of (5.19) only uses the function φ, however the other two functions will be required for the more precise decomposition (5.35) below.

We now have the following analogue of Lemma 5.2:

Lemma 5.5. *There exists $t_0 \geq 0$ such that if $(m, \nu) \in S^*M$ satisfies*

$$m \in \operatorname{supp}(1-\varphi_2) \text{ and } g^{-t}(m,\nu) \in \operatorname{supp}\varphi_1 \text{ for some } t \geq t_0, \tag{5.21}$$

then:

1. *(m, ν) directly escapes in the forward direction;*
2. *for each $s \geq 0$, $g^s(m,\nu)$ does not lie in the set W_ξ defined in (4.9) for any $\xi \in \partial\overline{M}$; and*

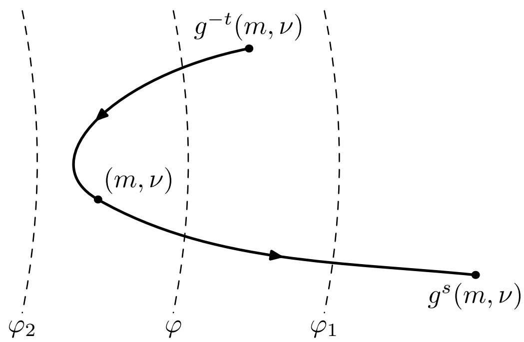

FIGURE 3. An illustration of Lemma 5.5. The functions $\varphi, \varphi_1, \varphi_2$ are supported to the left of the corresponding dashed lines; the right side of the figure represents infinity.

(3) *for each* $s \geq t_0$, $g^s(m,\nu) \notin \operatorname{supp}\varphi_1$.

Proof. (1) Let $\operatorname{supp}\varphi_1 \subset \{x \geq \varepsilon_\varphi\}$. The set $\mathcal{DE}_- \cap \{x \geq \varepsilon_\varphi\}$, where $\mathcal{DE}_-$ is specified in Definition 4.2, is compact; therefore, there exists $t_0 > 0$ such that for $t \geq t_0$ and $(m,\nu) \in \mathcal{DE}_- \cap \{x \geq \varepsilon_\varphi\}$, we have $g^{-t}(m,\nu) \notin \operatorname{supp}\varphi_1$.

Now, assume that (m,ν) satisfies (5.21), but it does not directly escape in the forward direction. Since $(m,\nu) \in \operatorname{supp}(1-\varphi_2)$, we have $x(m) \leq \varepsilon_0$; therefore, $(m,\nu) \in \mathcal{DE}_-$. Then $x(m) \geq x(g^{-t}(m,\nu)) \geq \varepsilon_\varphi$; therefore, $(m,\nu) \in \mathcal{DE}_- \cap \{x \geq \varepsilon_\varphi\}$, a contradiction with the fact that $g^{-t}(m,\nu) \in \operatorname{supp}\varphi_1$ and $t \geq t_0$.

(2) This is proved exactly as part 2 of Lemma 5.2.

(3) It is enough to use part (1), take t_0 large enough, and use that the set $\mathcal{DE}_+ \cap \{x \geq \varepsilon_\varphi\}$ is compact. ∎

Take t_0 from Lemma 5.5. Let $\varphi_{t_0} \in C_0^\infty(M)$ be real-valued and satisfy $d_g(\operatorname{supp}\varphi_1, \operatorname{supp}(1-\varphi_{t_0})) > t_0$. Take a compact set $K_{t_0} \subset M$ whose interior contains $\operatorname{supp}\varphi_{t_0}$. Put

$$\mathcal{N}(E_h) := 1 + \|E_h\|_{L^2(K_{t_0})}; \tag{5.22}$$

this quantity depends on λ and ξ and we know by (4.17) that

$$h^{-1}\|\mathcal{N}(E_h)\|^2_{L^2_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])} = \mathcal{O}(1). \tag{5.23}$$

By (1.4) and Lemma 3.10, we have similarly to (5.4),

$$\varphi E_h = e^{i\beta}\varphi U(t_0)\varphi_{t_0} E_h + \mathcal{O}(h^\infty \mathcal{N}(E_h))_{L^2}. \tag{5.24}$$

Here β is given by (5.20). Iterating (5.24) by writing $\varphi_{t_0} = \varphi + (1-\varphi)\varphi_{t_0}$, we get for $l = \mathcal{O}(\log(1/h))$ (or even for l polynomially bounded in h)

$$\varphi E_h = e^{il\beta}(\varphi U(t_0))^l \varphi E_h + \sum_{j=1}^{l} e^{ij\beta}(\varphi U(t_0))^j (1-\varphi)\varphi_{t_0} E_h + \mathcal{O}(h^\infty \mathcal{N}(E_h))_{L^2}, \tag{5.25}$$

uniformly in $\xi \in \partial\overline{M}$ and $\lambda \in [1, 1+h]$. Same is true if φ is replaced by any function $\varphi' \in C_0^\infty(M)$ such that $d_g(\operatorname{supp}\varphi', \operatorname{supp}(1-\varphi_{t_0})) > t_0$. One can also replace $U(t_0)$ by $U(-t_0)$.

We now use our knowledge of the wavefront set of $\widetilde{E}_h^1$ to prove the following analogue of Proposition 5.3:

Proposition 5.6. *If $E_h = \chi_0 E_h^0 + E_h^1$ is the decomposition (4.6), then*

$$\|\varphi U(t_0)(1-\varphi)\varphi_{t_0} E_h^1\|_{L^2} = \mathcal{O}(h^\infty \mathcal{N}(E_h)), \tag{5.26}$$

uniformly in $\xi \in \partial\overline{M}$ and $\lambda \in [1, 1+h]$. Same is true if we replace each instance of φ by any function in the set $\{\varphi, \varphi_1, \varphi_2\}$.

Proof. Recalling the definition (4.8) of $\widetilde{E}_h^1$, we see that (5.26) follows from

$$\|\varphi U(t_0)(1-\varphi)\varphi_{t_0} \widetilde{E}_h^1\|_{L^2} = \mathcal{O}(h^\infty). \tag{5.27}$$

We now make the following observation: a point $(m, \nu) \in S^*M$ in the wavefront set of $\widetilde{E}_h^1$ will make an $\mathcal{O}(h^\infty)$ contribution to (5.27) unless $m \in \operatorname{supp}(1-\varphi)$, but $g^{-t_0}(m,\nu) \in \operatorname{supp}\varphi$; however, by (A6) and Lemma 5.5, in this case $(m,\nu) \notin \operatorname{WF}_h(\widetilde{E}_h^1)$. To make this argument rigorous, we can write (bearing in mind that $\widetilde{E}_h^1$ is polynomially bounded)

$$\varphi_{t_0}\widetilde{E}_h^1 = B\widetilde{E}_h^1 + \mathcal{O}(h^\infty)_{L^2},$$

where $B \in \Psi^{\mathrm{comp}}$ is compactly supported and such that

$$(m,\nu) \in \operatorname{WF}_h(B) \cap \operatorname{supp}(1-\varphi) \Longrightarrow g^{-t_0}(m,\nu) \notin \operatorname{supp}\varphi.$$

Then the operator $\varphi U(t_0)(1-\varphi)B$ is $\mathcal{O}(h^\infty)_{L^2\to L^2}$ by part 2 of Proposition 3.8, which proves (5.27). □

Using (5.26), we can replace E_h by $\chi_0 E_h^0$ in each term of the sum (5.25):

$$\begin{gathered}\varphi E_h = e^{il\beta}(\varphi U(t_0))^l \varphi E_h + \sum_{j=1}^{l} e^{ij\beta}(\varphi U(t_0))^j (1-\varphi)\varphi_{t_0}\chi_0 E_h^0 \\ +\mathcal{O}(h^\infty \mathcal{N}(E_h))_{L^2}.\end{gathered} \tag{5.28}$$

Applying the operator $A = \varphi A \varphi$, we get (5.19).

Properties of propagators up to Ehrenfest time. We will now establish certain properties of the cut off and iterated propagators up to the Ehrenfest time t_e defined in (5.16), or, in certain cases, up to twice the Ehrenfest time. The need for these properties arises mostly because of the cutoffs present in the argument. Define the Ehrenfest index

$$l_e := \lfloor t_e/t_0 \rfloor + 1 \sim \log(1/h). \tag{5.29}$$

Lemma 5.7. *Assume that $\varphi', \varphi'' \in C_0^\infty(M)$ satisfy $|\varphi'|, |\varphi''| \le 1$ everywhere. Let $B \in \Psi^{\mathrm{comp}}$ be compactly supported and microlocalized inside the set $\mathcal{E}_{\varepsilon_e}$ defined in (5.17). Then:*

1. If $\varphi''=1$ *near* $\operatorname{supp}\varphi'$*, then for* $0\leq j\leq l_e$*,*

$$(\varphi'U(\pm t_0))^j BU(\mp jt_0)=(\varphi'U(\pm t_0))^j B(U(\mp t_0)\varphi'')^j+\mathcal{O}(h^\infty)_{L^2\to L^2},\tag{5.30}$$

$$(U(\pm t_0)\varphi')^j BU(\mp jt_0)=(U(\pm t_0)\varphi')^j B(\varphi''U(\mp t_0))^j+\mathcal{O}(h^\infty)_{L^2\to L^2}.\tag{5.31}$$

2. If $B_1,B_2\in\Psi^{\mathrm{comp}}$ *satisfy same conditions as* B *and moreover* $\mathrm{WF}_h(B_1)\cap\mathrm{WF}_h(B_2)=\emptyset$*, then for* $0\leq j\leq 2l_e$ *(that is, up to twice the Ehrenfest time)*

$$B_1(\varphi'U(\pm t_0))^j B(U(\mp t_0)\varphi'')^j B_2=\mathcal{O}(h^\infty)_{L^2\to L^2}.\tag{5.32}$$

Same is true if we replace $\varphi'U(\pm t_0)$ *by* $U(\pm t_0)\varphi'$ *and/or replace* $U(\mp t_0)\varphi''$ *by* $\varphi''U(\mp t_0)$*.*

3. If $\varphi''=1$ *near* $\operatorname{supp}\varphi'$ *and both* φ'' *and* B *are supported at distance more than* t_0 *from* $\operatorname{supp}(1-\varphi_{t_0})$*, then for* $0\leq j\leq l_e$ *(*β *is defined in (5.20))*

$$e^{\pm ij\beta}(\varphi'U(\pm t_0))^j BE_h=(\varphi'U(\pm t_0))^j B(U(\mp t_0)\varphi'')^j E_h+\mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2},\tag{5.33}$$

$$e^{\pm ij\beta}(U(\pm t_0)\varphi')^j BE_h=(U(\pm t_0)\varphi')^j B(\varphi''U(\mp t_0))^j\varphi_{t_0}E_h+\mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}.\tag{5.34}$$

Proof. We will repeatedly use Propositions 3.8 and 3.9 and omit the $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainders present there.

1. We prove (5.30); (5.31) is proved similarly. Assume that the signs are chosen so that (5.30) features $\varphi'U(t_0)$. We argue by induction in j. The case $j=0$ is obvious. Now, assume that (5.30) is true for $j-1$ in place of j. Then

$$(\varphi'U(t_0))^j BU(-jt_0)=\varphi'B'+\mathcal{O}(h^\infty)_{L^2\to L^2},$$

where

$$B'=U(t_0)(\varphi'U(t_0))^{j-1}B(U(-t_0)\varphi'')^{j-1}U(-t_0)$$

is a compactly supported operator in $\Psi^{\mathrm{comp}}_{\rho_e}$, with ρ_e defined in (5.18). Since $\operatorname{supp}\varphi'\cap\operatorname{supp}(1-\varphi'')=\emptyset$, we have

$$\varphi'B'=\varphi'B'\varphi''+\mathcal{O}(h^\infty)_{L^2\to L^2}$$

and (5.30) follows.

2. We again assume that the signs are chosen so that (5.32) features $\varphi'U(t_0)$. Write $j=j_1+j_2$, where $0\leq j_1,j_2\leq l_e$, and write the left-hand side of (5.32) as $U(j_1t_0)\widetilde{B}_1\widetilde{B}\widetilde{B}_2U(-j_1t_0)$, where

$$\widetilde{B}=(\varphi'U(t_0))^{j_2}B(U(-t_0)\varphi'')^{j_2},$$

$$\widetilde{B}_1=U(-j_1t_0)B_1(\varphi'U(t_0))^{j_1},\ \widetilde{B}_2=(U(-t_0)\varphi'')^{j_1}B_2U(j_1t_0).$$

Now, $\widetilde{B}$ is a compactly supported member of $\Psi^{\mathrm{comp}}_{\rho_e}$. Same can be said about $\widetilde{B}_1$ and $\widetilde{B}_2$, by applying (5.31) and its adjoint (where the role of φ' is played by either φ' or φ'' and the role of φ'', by a suitably chosen cutoff function). Moreover, if U_1,U_2 are bounded open subsets of T^*M such that $\mathrm{WF}_h(B_k)\subset U_k$ and $U_1\cap U_2=\emptyset$, then by Proposition 3.9, $\widetilde{B}_k$ is microsupported, in the sense of Definition 3.1, on the set $g^{j_1t_0}(U_k)$; since these two sets do not intersect, we see that $\widetilde{B}_1\widetilde{B}\widetilde{B}_2=\mathcal{O}(h^\infty)_{L^2\to L^2}$ as needed.

3. We once again fix the sign so that $U(t_0)$ stands next to φ'. Formally, (5.33) and (5.34) follow by applying (5.30) and (5.31), respectively, to the identity $e^{ij\beta}E_h = U(-jt_0)E_h$. To make this observation rigorous, we write by Lemma 3.10

$$e^{i\beta}BE_h = BU(-t_0)\varphi_{t_0}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2},$$
$$e^{i\beta}\varphi''E_h = \varphi''U(-t_0)\varphi_{t_0}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}.$$

We now use induction in j. For $j=0$, both (5.33) and (5.34) are trivial. Now, assume that they both hold for $j-1$ in place of j. We then write

$$\begin{aligned}&e^{ij\beta}(\varphi'U(t_0))^jBE_h\\ &= e^{i\beta}(\varphi'U(t_0))^jB(U(-t_0)\varphi'')^{j-1}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}\\ &= (\varphi'U(t_0))^jB(U(-t_0)\varphi'')^{j-1}U(-t_0)\varphi_{t_0}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}.\end{aligned}$$

The operator $(\varphi'U(t_0))^jB(U(-t_0)\varphi'')^{j-1}U(-t_0)$ is a compactly supported element of $\Psi^{\mathrm{comp}}_\rho$; moreover, as $j\geq 1$, the wavefront set of this operator is contained in $\operatorname{supp}\varphi'$. Since $\varphi''=1$ near $\operatorname{supp}\varphi'$, we can replace φ_{t_0} by φ'' in the last formula, proving (5.33).

We next write

$$e^{ij\beta}(U(t_0)\varphi')^jBE_h = e^{i\beta}(U(t_0)\varphi')^jB(\varphi''U(-t_0))^{j-1}\varphi_{t_0}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}.$$

However, $\varphi'(U(t_0)\varphi')^{j-1}B(\varphi''U(-t_0))^{j-1}$ is a compactly supported element of $\Psi^{\mathrm{comp}}_\rho$ and its wavefront set is contained in $\operatorname{supp}\varphi'$. Since $\varphi''=1$ near $\operatorname{supp}\varphi'$, we can replace φ_{t_0} by φ'', obtaining (5.34):

$$\begin{aligned}&e^{ij\beta}(U(t_0)\varphi')^jBE_h\\ &= e^{i\beta}(U(t_0)\varphi')^jB(\varphi''U(-t_0))^{j-1}\varphi''E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}\\ &= (U(t_0)\varphi')^jB(\varphi''U(-t_0))^{j-1}\varphi''U(-t_0)\varphi_{t_0}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}. \quad \square\end{aligned}$$

Second decomposition of $\langle AE_h, E_h\rangle$. We now analyse the terms of (5.19), reducing $\langle AE_h, E_h\rangle$ to an expression depending on the 'outgoing' part E_h^0 of the plane wave (see (4.6)), with remainder estimated by the classical escape rate for up to twice the Ehrenfest time:

Proposition 5.8. *For $0\leq l\leq 2l_e$,*

$$\begin{gathered}\langle AE_h, E_h\rangle = \sum_{j=1}^{l}\langle \widetilde{A}^j\chi_0E_h^0, \chi_0E_h^0\rangle + \mathcal{O}\big(h\mu_L(\mathcal{T}(lt_0)) + h^\infty\big)_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])},\\ \widetilde{A}^j := \varphi_{t_0}(1-\varphi_2)(U(-t_0)\varphi_1)^jA(\varphi U(t_0))^j(1-\varphi)\varphi_{t_0}.\end{gathered} \tag{5.35}$$

Here l_e is defined in (5.29) and $\mathcal{T}(t)$ in (1.9); we keep the $\mathcal{O}(h^\infty)$ remainder to include the nontrapping case.

We will use Lemma 5.7; since it only applies to pseudodifferential operators microlocalized inside the set $\mathcal{E}_{\varepsilon_e}$ from (5.17), we take an operator

$$X_0\in\Psi^{\mathrm{comp}}(M),\ \ \mathrm{WF}_h(X_0)\subset\mathcal{E}_{\varepsilon_e},\ \ X_0=1 \text{ near } S^*M\cap\operatorname{supp}\varphi_{t_0}, \tag{5.36}$$

compactly supported inside K_{t_0}. By (1.4) and the elliptic estimate (Proposition 3.2), we have

$$\varphi_{t_0}E_h = X_0\varphi_{t_0}E_h + \mathcal{O}(h^\infty \mathcal{N}(E_h))_{L^2} = \varphi_{t_0}X_0E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}. \tag{5.37}$$

Same is true if we replace E_h by $\chi_0 E_h^0$, as by (A4) and the fact that $|\partial_m\phi_\xi|_g = 1$, we have $\mathrm{WF}_h(\chi_0E_h^0) \subset S^*M$. We also recall that $\mathrm{WF}_h(A) \subset \mathcal{E}_{\varepsilon_e}$.

We start the proof of Proposition 5.8 by estimating the first term on the right-hand side of (5.19) for l up to twice the Ehrenfest time, in terms of the classical escape rate:

Proposition 5.9. *There exists a constant C such that for $0 \le l \le 2l_e$, we have*

$$h^{-1}\|\langle A(\varphi U(t_0))^l\varphi E_h, E_h\rangle\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])} \le C\mu_L(\mathcal{T}(lt_0)) + \mathcal{O}(h^\infty). \tag{5.38}$$

Proof. We write $l = l_1 + l_2$, where $0 \le l_1, l_2 \le l_e$; then

$$\langle A(\varphi U(t_0))^l\varphi E_h, E_h\rangle = \langle(\varphi U(t_0))^{l_1}\varphi E_h, (U(-t_0)\varphi)^{l_2}A^*E_h\rangle.$$

Now, by (5.33)

$$\begin{aligned} e^{il_1\beta}(\varphi U(t_0))^{l_1}\varphi E_h &= e^{il_1\beta}(\varphi U(t_0))^{l_1}X_0\varphi E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2} \\ &= B_l^1E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}, \end{aligned}$$

where

$$B_l^1 = (\varphi U(t_0))^{l_1}X_0\varphi(U(-t_0)\varphi_1)^{l_1}.$$

Similarly, by (5.34) (recalling that $A = \varphi A\varphi$)

$$e^{-il_2\beta}(U(-t_0)\varphi)^{l_2}A^*E_h = B_l^2E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2},$$

where

$$B_l^2 = (U(-t_0)\varphi)^{l_2}A^*(\varphi_1U(t_0))^{l_2}\varphi_{t_0}.$$

Put $B_l = (B_l^2)^*B_l^1$; recalling (5.23), it is then enough to show that

$$h^{-1}\|\langle B_lE_h, E_h\rangle\|_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])} \le C\mu_L(\mathcal{T}(lt_0)) + \mathcal{O}(h^\infty). \tag{5.39}$$

Now, by Proposition 3.9, the operator B_l^1 is a compactly supported element of $\Psi^{\mathrm{comp}}_{\rho_e}$ (modulo an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder which we will omit), and it is microsupported, in the sense of Definition 3.1, inside the set $g^{-l_1t_0}(\{\varphi_1 \neq 0\})$ (here we only use that $\operatorname{supp}\varphi \subset \{\varphi_1 \neq 0\}$). Similarly, $B_l^2 \in \Psi^{\mathrm{comp}}_{\rho_e}$ is microsupported inside $g^{l_2t_0}(\{\varphi_1 \neq 0\})$. Therefore, B_l is microsupported on the set

$$\mathcal{S}_l = g^{-l_1t_0}(\{\varphi_1 \neq 0\}) \cap g^{l_2t_0}(\{\varphi_1 \neq 0\}).$$

Note also that B_l is compactly supported independently of l.

Now, by taking the convolution of the indicator function of an h^{ρ_e} sized neighborhood of $\mathcal{S}_l$ with an appropriately rescaled cutoff function, we can construct a compactly supported operator $\widetilde{B}_l \in \Psi^{\mathrm{comp}}_{\rho_e}$ such that $B_l = \widetilde{B}_l^*B_l + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ and $\widetilde{B}_l$ is microsupported inside an $\mathcal{O}(h^{\rho_e})$ sized neighborhood $\widetilde{\mathcal{S}}_l$ of $\mathcal{S}_l$. Using (5.16), (5.18), and the estimate on the Lipschitz constant of the flow given by (3.17), we see that for $(\widetilde{m}, \tilde{\nu}) \in \widetilde{\mathcal{S}}_l \cap S^*M$, there exists $(m, \nu) \in \mathcal{S}_l \cap S^*M$ such that $d((\widetilde{m}, \tilde{\nu}), (m, \nu)) \le Ch^{\rho_e}$ and for $\Lambda_0' > \Lambda_0'' > (1 + 2\varepsilon_e)\Lambda_{\max}$,

$$d(g^{l_1t_0}(\widetilde{m}, \tilde{\nu}), g^{l_1t_0}(m, \nu)) \le Ce^{l_1t_0\Lambda_0''}h^{\rho_e} \le Ce^{t_e\Lambda_0''}h^{\rho_e} \le Ce^{-t_e(\Lambda_0'-\Lambda_0'')}$$

is bounded by some positive power of h. Here d denotes some smooth distance function on T^*M. Same is true if we replace $g^{l_1t_0}$ with $g^{-l_2t_0}$; therefore, if the compact set K_0 used in the definition (1.9) of $\mathcal{T}(t)$ is chosen large enough, we have

$$\widetilde{\mathcal{S}}_l \cap S^*M \subset g^{l_2t_0}(\mathcal{T}(lt_0)). \tag{5.40}$$

Using the Cauchy–Schwartz inequality and (4.16), we bound the left-hand side of (5.39) by

$$\begin{aligned} h^{-1}\|\langle B_lE_h, E_h\rangle\|_{L^1_{\xi,\lambda}} &\leq h^{-1}\|\langle B_lE_h, \widetilde{B}_lE_h\rangle\|_{L^1_{\xi,\lambda}} + \mathcal{O}(h^\infty) \\ &\leq h^{-1}\|B_lE_h\|_{L^2(M)L^2_{\xi,\lambda}} \cdot \|\widetilde{B}_lE_h\|_{L^2(M)L^2_{\xi,\lambda}} + \mathcal{O}(h^\infty) \\ &\leq C(h^{n/2}\|B_l\Pi_{[1,1+h]}\|_{\mathrm{HS}})(h^{n/2}\|\widetilde{B}_l\Pi_{[1,1+h]}\|_{\mathrm{HS}}) + \mathcal{O}(h^\infty). \end{aligned}$$

It remains to use (3.25) (or rather its adjoint). Indeed, both B_l and $\widetilde{B}_l$ are bounded in $\Psi^{\mathrm{comp}}_{\rho_e}$ uniformly in l, and they are microsupported in $\widetilde{\mathcal{S}}_l$; therefore, by (5.40)

$$h^{-1}\|\langle B_lE_h, E_h\rangle\|_{L^1_{\xi,\lambda}} \leq C\mu_L(\widetilde{\mathcal{S}}_l \cap S^*M) + \mathcal{O}(h^\infty) \leq C\mu_L(\mathcal{T}(lt_0)) + \mathcal{O}(h^\infty). \quad \square$$

As for the sum in (5.19), we have the following

Proposition 5.10. *For $1 \leq j \leq 2l_e$, we have*

$$e^{ij\beta}\langle A(\varphi U(t_0))^j(1-\varphi)\varphi_{t_0}\chi_0E^0_h, E_h\rangle = \langle \widetilde{A}^j\chi_0E^0_h, \chi_0E^0_h\rangle + \mathcal{O}(h^\infty\mathcal{N}(E_h)^2), \tag{5.41}$$

uniformly in $\xi \in \partial\overline{M}$ and $\lambda \in [1, 1+h]$, with $\widetilde{A}^j$ defined in (5.35).

Proof. Since $A = \varphi A\varphi$, we can replace E_h by φ_1E_h on the left-hand side of (5.41). Writing down (5.25) for φ_1 in place of φ and using φ_2 in place of φ_1 in the splitting $\varphi_{t_0} = \varphi_1 + (1-\varphi_1)\varphi_{t_0}$ in the last step, we get

$$\begin{aligned} \varphi_1E_h &= e^{ij\beta}(\varphi_1U(t_0))^j\varphi_2E_h + e^{ij\beta}(\varphi_1U(t_0))^j(1-\varphi_2)\varphi_{t_0}E_h \\ &\quad + \sum_{k=1}^{j-1} e^{ik\beta}(\varphi_1U(t_0))^k(1-\varphi_1)\varphi_{t_0}E_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2}. \end{aligned} \tag{5.42}$$

We now substitute (5.42) into the left-hand side of (5.41). The first term gives, after using (5.37) to replace $\varphi_{t_0}\chi_0E^0_h$ by $X_0\varphi_{t_0}\chi_0E^0_h$ and φE_h by $X_0\varphi E_h$

$$\langle A(\varphi U(t_0))^j(1-\varphi)\varphi_{t_0}\chi_0E^0_h, (\varphi_1U(t_0))^j\varphi_2E_h\rangle = \langle B_0\chi_0E^0_h, E_h\rangle + \mathcal{O}(h^\infty\mathcal{N}(E_h)^2),$$

where

$$B_0 = \varphi_2X_0^*(U(-t_0)\varphi_1)^jA(\varphi U(t_0))^j(1-\varphi)X_0\varphi_{t_0} = \mathcal{O}(h^\infty)_{L^2\to L^2}$$

by (5.32), as $\operatorname{supp}\varphi_2 \cap \operatorname{supp}(1-\varphi) = \emptyset$.

Next, we use Proposition 5.6 to write the second term of (5.42) as

$$e^{ij\beta}(\varphi_1U(t_0))^j(1-\varphi_2)\varphi_{t_0}\chi_0E^0_h + \mathcal{O}(h^\infty\mathcal{N}(E_h))_{L^2};$$

therefore, this term gives the right-hand side of (5.41).

It remains to estimate the contribution of each term of the sum in (5.42), which we can write, using (5.37), as $e^{i(j-k)\beta}\langle B_k\chi_0E^0_h, \chi_0E^0_h\rangle + \mathcal{O}(h^\infty\mathcal{N}(E_h)^2)$, with

$$B_k = \varphi_{t_0}X_0^*(1-\varphi_1)(U(-t_0)\varphi_1)^kA(\varphi U(t_0))^j(1-\varphi)X_0\varphi_{t_0}.$$

We need to show that $\|B_k\|_{L^2\to L^2} = \mathcal{O}(h^\infty)$ for $1 \le k < j$. For that, we consider two cases. First, assume that $k \le l_e$. Then we have

$$\varphi_{t_0} X_0^*(1-\varphi_1)(U(-t_0)\varphi_1)^k A(\varphi U(t_0))^k \varphi = \mathcal{O}(h^\infty)_{L^2\to L^2}, \tag{5.43}$$

as the supports of $1-\varphi_1$ and φ do not intersect and the operator in between them is a compactly supported element of $\Psi^{\mathrm{comp}}_{\rho_e}$ (modulo an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder which we will omit). Since B_k is obtained by multiplying the left-hand side of (5.43) on the right by $U(t_0)(\varphi U(t_0))^{j-1-k}(1-\varphi)X_0\varphi_{t_0}$, it is also $\mathcal{O}(h^\infty)_{L^2\to L^2}$.

Now, assume that $k \ge l_e$. Take $\widetilde{\varphi}_1 \in C_0^\infty(M)$ equal to 1 near $\operatorname{supp}\varphi_1$ and such that $|\widetilde{\varphi}_1| \le 1$ everywhere. We write by (5.30) and its adjoint,

$$\begin{aligned}
U((k-l_e)t_0)B_k U(-(j-l_e)t_0) &= B_k^1 B_k^2 B_k^3 + \mathcal{O}(h^\infty)_{L^2\to L^2},\\
B_k^1 &= (\widetilde{\varphi}_1 U(t_0))^{k-l_e}\varphi_{t_0} X_0^*(1-\varphi_1)(U(-t_0)\varphi_1)^{k-l_e},\\
B_k^2 &= (U(-t_0)\varphi_1)^{l_e} A(\varphi U(t_0))^{l_e},\\
B_k^3 &= (\varphi U(t_0))^{j-l_e}(1-\varphi)X_0\varphi_{t_0}(U(-t_0)\varphi_1)^{j-l_e}.
\end{aligned}$$

Now all B_k^i, $i=1,2,3$, are compactly supported members of $\Psi^{\mathrm{comp}}_{\rho_e}$. Let U_1, U_2 be two bounded open sets such that $\operatorname{supp}(\varphi_{t_0}(1-\varphi_1)) \subset U_1$, $\operatorname{supp}\varphi \subset U_2$, and $U_1 \cap U_2 = \emptyset$. Since $k - l_e > j - l_e$ and by Proposition 3.9, the operator B_k^1 is microsupported, in the sense of Definition 3.1, on the set $g^{-(k-l_e)t_0}(U_1)$, while B_k^3 is microsupported on the set $g^{-(k-l_e)t_0}(U_2)$; since these two sets do not intersect, we get $B_k = \mathcal{O}(h^\infty)_{L^2\to L^2}$, finishing the proof. □

Combining (5.19) with (5.38) and (5.41), we finally get (5.35).

Proof of (1.12). Put l equal to the number l_e defined in (5.29). By (5.35), it is enough to approximate the terms $\langle \widetilde{A}^j \chi_0 E_h^0, \chi_0 E_h^0\rangle$. This is done by the following proposition, relying on the Lagrangian structure of E_h^0 and featuring the interpolated escape rate $r(h,\Lambda)$ from (1.10):

Proposition 5.11. *Put $l = l_e$ given by* (5.29)*, and $r(h,\Lambda)$ defined in* (1.10)*. Then the sum on the right-hand side of* (5.35) *is approximated as follows:*

$$\sum_{j=1}^{l} \langle \widetilde{A}^j \chi_0 E_h^0, \chi_0 E_h^0\rangle = \int_{S^*M} \sigma(A)\, d\mu_\xi + \mathcal{O}(hr(h, 2\Lambda_0))_{L^1_{\xi,\lambda}(\partial\overline{M}\times[1,1+h])}. \tag{5.44}$$

Proof. By Proposition 3.9, the operator $\widetilde{A}^j$ is compactly supported and lies in $\Psi^{\mathrm{comp}}_{\rho_j}$, modulo an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder, where

$$\rho_j = \frac{jt_0}{t_e}\rho_e, \qquad \text{with } jt_0/t_e \le 1 + o(1)$$

with t_e and ρ_e defined in (5.16) and (5.18), respectively. Next, $\widetilde{A}^j$ is microsupported, in the sense of Definition 3.1, in the set

$$\mathcal{Q}_j := g^{t_0}(\{\varphi_1 \ne 0\}) \cap g^{jt_0}(\{\varphi_1 \ne 0\}).$$

If the set K_0 from the definition (1.9) of $\mathcal{T}(t)$ is large enough, then $\mathcal{Q}_j\cap S^*M\subset g^{jt_0}(\mathcal{T}(jt_0))$; by the definition (1.10) of $r(h,\Lambda)$, we find

$$h^{1-jt_0/t_e}\mu_L(\mathcal{Q}_j\cap S^*M)\leq r(h,2\Lambda_0). \tag{5.45}$$

By (A4) and Proposition 3.5, we have the following analogue of (5.11):

$$\widetilde{A}^j\chi_0E_h^0=e^{\frac{i\lambda}{h}\phi_\xi}\chi_0b^0(1,\xi,m;0)\sigma(\widetilde{A}^j)(m,\partial_m\phi_\xi(m))+\mathcal{O}(h^{1-2\rho_j})_{L^\infty}.$$

Moreover, by part 2 of the same proposition, we see that $\widetilde{A}^j\chi_0E_h^0$ is $\mathcal{O}(h^\infty)$ outside of the set of points $m\in U_\xi$ such that $(m,\partial_m\phi_\xi(m))\in\mathcal{Q}_j$. By Lemma 5.5, $\sigma(\widetilde{A}^j)$ is supported in $\operatorname{supp}(1-\varphi)\cap g^{t_0}(\operatorname{supp}\varphi_1)\subset\mathcal{DE}_+\cap\{x<\varepsilon_1\}$, with $\mathcal{DE}_+$ from Definition 4.2. Using part 2 of Lemma 4.4, we then get

$$\langle\widetilde{A}^j\chi_0E_h^0,\chi_0E_h^0\rangle=\int_{S^*M}\sigma(\widetilde{A}^j)\,d\mu_\xi+\mathcal{O}(h^{1-2\rho_j}\mu_\xi(\mathcal{Q}_j))+\mathcal{O}(h^\infty), \tag{5.46}$$

uniformly in $\xi\in\partial\overline{M}$ and $\lambda\in[1,1+h]$. Now, we write by (5.45) and Proposition 4.3,

$$\begin{gathered}\sum_{j=1}^l h^{1-2\rho_j}\|\mu_\xi(\mathcal{Q}_j)\|_{L^1_\xi}=\sum_{j=1}^l h^{1-2\rho_j}\mu_L(\mathcal{Q}_j\cap S^*M)\\ \leq r(h,2\Lambda_0)\sum_{j=1}^l h^{(1-2\rho_e)jt_0/t_e}=r(h,2\Lambda_0)\sum_{j=1}^l e^{-2\Lambda_0(1-2\rho_e)jt_0}\leq Cr(h,2\Lambda_0).\end{gathered}$$

It remains to sum up the integrals in (5.46). We have by Proposition 3.9, bearing in mind that $\varphi\varphi_1=\varphi$, $(1-\varphi)(1-\varphi_2)=1-\varphi$, $A=\varphi A\varphi$, and $d_g(\operatorname{supp}(1-\varphi_{t_0}),\operatorname{supp}\varphi)>t_0$,

$$\sigma(\widetilde{A}^j)=(\sigma(A)\circ g^{-jt_0})(1-\varphi)\prod_{k=1}^{j-1}\varphi\circ g^{-kt_0}.$$

By part 1 of Lemma 4.4, we write

$$\begin{aligned}\sum_{j=1}^l\int_{S^*M}\sigma(\widetilde{A}^j)\,d\mu_\xi&=\int_{S^*M}\sigma(A)\sum_{j=1}^l(1-\varphi\circ g^{jt_0})\prod_{k=1}^{j-1}\varphi\circ g^{kt_0}\,d\mu_\xi\\ &=\int_{S^*M}\sigma(A)\Big(1-\prod_{k=1}^l\varphi\circ g^{kt_0}\Big)\,d\mu_\xi.\end{aligned}$$

It remains to note that by Proposition 4.3,

$$\int_{\partial\overline{M}}\int_{S^*M}|\sigma(A)|\prod_{k=1}^l\varphi\circ g^{kt_0}\,d\mu_\xi d\xi=\int_{S^*M}|\sigma(A)|\prod_{k=1}^l\varphi\circ g^{kt_0}\,d\mu_L=\mathcal{O}(\mu_L(\mathcal{T}(lt_0)))$$

since the expression under the last integral is supported in $\mathcal{T}(lt_0)$. □

5.3. **Trace estimates.** In this subsection, we prove a stronger remainder bound (1.13) for the case when $\langle AE_h, E_h\rangle$ is paired with a test function in ξ and obtain an expansion of the trace of spectral projectors with a fractal remainder – Theorem 4.

Expressing $E_h^0 \otimes E_h^0$ via Schrödinger propagators. Our argument will be based on the decomposition (5.35). The remainder in this decomposition is already controlled by the escape rate at twice the Ehrenfest time t_e defined in (5.16). However, in the previous subsection (see Proposition 5.11), we were only able to estimate the sum in (5.35) for l up to the Ehrenfest index $l_e \sim t_e/t_0$ defined in (5.29). We therefore need a better way of writing down the Lagrangian states E_h^0, when coupled with a test function in ξ, and such a way is provided by

Lemma 5.12. *Let $f(\xi) \in C^\infty(\partial\overline{M})$ and define for $\lambda \in (1/2, 2)$,*

$$\Pi_f^0(\lambda) := \int_{\partial\overline{M}} f(\xi)(\chi_0 E_h^0(\lambda,\xi)) \otimes (\chi_0 E_h^0(\lambda,\xi))\, d\xi. \tag{5.47}$$

Here $\otimes$ denotes the Hilbert tensor product, see (2.8). Assume that $\widetilde{X}_1, \widetilde{X}_2 \in \Psi^{\mathrm{comp}}(M)$ are compactly supported and the projections $\pi_S(\mathrm{WF}_h(\widetilde{X}_j))$ of $\mathrm{WF}_h(\widetilde{X}_j)$ onto S^*M along the radial rays in the fibers of T^*M lie inside $\mathcal{DE}_+ \cap \{x \le \varepsilon_1\}$, with $\mathcal{DE}_+$ defined in (4.2) and ε_1 from (A7). Then

$$\widetilde{X}_1 \Pi_f^0(\lambda) \widetilde{X}_2^* = (2\pi h)^n \int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)} U(s) B_s(\lambda)\, ds + \mathcal{O}(h^\infty)_{L^2\to L^2},$$

where $T_0 > 0$ is independent of h and $B_s(\lambda) \in \Psi^{\mathrm{comp}}(M)$ is compactly supported on M, smooth and compactly supported in $s \in (-T_0, T_0)$, and smooth in λ. Moreover, if $\xi_{+\infty}$ is the function defined in (G3), then

$$\sigma(B_0(1))|_{S^*M} = f(\xi_{+\infty})\sigma(\widetilde{X}_1 \widetilde{X}_2^*). \tag{5.48}$$

Proof. We write

$$\widetilde{X}_1 \Pi_f^0(\lambda) \widetilde{X}_2^* = \int_{\partial\overline{M}} f(\xi)(\widetilde{X}_1 \chi_0 E_h^0(\lambda,\xi)) \otimes (\widetilde{X}_2 \chi_0 E_h^0(\lambda,\xi))\, d\xi.$$

By (A4), $\chi_0 E_h^0(\lambda,\xi)$ is a Lagrangian distribution associated to λ times the Lagrangian Λ_ξ from (4.3). By Proposition 3.5, we can write

$$\widetilde{X}_j \chi_0 E_h^0(\lambda,\xi)(m) = e^{\frac{i\lambda}{h}\phi_\xi(m)} b_j(\lambda,\xi,m;h) + \mathcal{O}(h^\infty)_{C_0^\infty},$$

where ϕ_ξ is defined in (G4) and b_j is a classical symbol in h smooth in λ, ξ, m and compactly supported in m. The symbol b_j depends on the operator $\widetilde{X}_j$; in fact, we can make $\operatorname{supp} b_j \subset \tau^{-1}(\pi_S(\mathrm{WF}_h(\widetilde{X}_j)))$, with τ defined in (4.2). We then write the Schwartz kernel of $\widetilde{X}_1 \Pi_f^0(\lambda) \widetilde{X}_2^*$, modulo an $\mathcal{O}(h^\infty)_{C_0^\infty}$ remainder, as

$$\widetilde{\Pi}(m,m';\lambda,h) = \int_{\partial\overline{M}} e^{\frac{i\lambda}{h}(\phi_\xi(m)-\phi_\xi(m'))} f(\xi) b_1(\lambda,\xi,m;h)\overline{b_2(\lambda,\xi;m',h)}\, d\xi. \tag{5.49}$$

Now, the support of each b_j in the (m,ξ) variables lies in the set U_∞^+ defined in (G4). The critical points of the phase $\lambda(\phi_\xi(m) - \phi_\xi(m'))$ are given by $\partial_\xi \phi_\xi(m) = \partial_\xi \phi_\xi(m')$; using (G6), we see that $h^{-n/2}\widetilde{\Pi}(m,m';\lambda,h)$ is a Lagrangian distribution associated to the Lagrangian

$$\widetilde{\Lambda}_\lambda := \{(m,\nu;m',\nu') \mid |\nu|_g=\lambda,\ \exists s\in(-T_0,T_0) : g^s(m,\nu)=(m',\nu')\}. \tag{5.50}$$

Here $T_0>0$ is large, but fixed.

Now, take some family $B_s(\lambda)\in\Psi^{\mathrm{comp}}(M)$ smooth and compactly supported in $s\in(-T_0,T_0)$ and define the operator

$$\Pi_B(\lambda) := (2\pi h)^n \int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)}U(s)B_s(\lambda)\,ds. \tag{5.51}$$

Following the proof of Lemma 3.11, we see that $h^{-n/2}$ times the Schwartz kernel $\Pi_B(m,m';\lambda,h)$ of $\Pi_B(\lambda)$ is, up to an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder, a compactly supported and compactly microlocalized Lagrangian distribution associated to the Lagrangian $\widetilde{\Lambda}_\lambda$. Moreover, the principal symbol of $h^{-n/2}\Pi_B(m,m';\lambda,h)$ at (m,ν,m',ν') such that $g^s(m,\nu)=(m',\nu')$ is a nonvanishing factor times $\sigma(B_s)(m',\nu')$. Arguing as in the proof of part 2 of Proposition 3.3, we see that we can find a family of operators $B_s(\lambda)$ such that

$$\widetilde{\Pi}(m,m';\lambda,h)=\Pi_B(m,m';\lambda,h)+\mathcal{O}(h^\infty)_{C_0^\infty}.$$

It remains to check that the family $B_s(\lambda)$ can be chosen to depend smoothly on λ uniformly in h (this is not automatic, as multiplication by $e^{\frac{i}{h}\psi(\lambda)}$ for some function ψ destroys this property, but does not change the Lagrangians where our kernels are microlocalized for each λ). For that, it is enough to note (by Proposition 3.3) that if we consider $h^{-n/2}\widetilde{\Pi}$ and $h^{-n/2}\Pi_B$ as Lagrangian distributions in m,m',λ, they are associated to the Lagrangian

$$\{(m,\nu,m',\nu',\lambda,q_\lambda) \mid |\nu|_g=\lambda,\ \exists s\in(-T_0,T_0) : g^s(m,\nu)=(m',\nu'),\ q_\lambda=-\lambda s\},$$

where q_λ is the momentum corresponding to λ. For $\widetilde{\Pi}$, this is true as when $\tau(m',\xi)=g^s(\tau(m,\xi))$, we have $\phi_\xi(m)-\phi_\xi(m')=-s$ by (4.4); for Π_B, this is seen directly from the parametrization (3.27), keeping in mind the factor $e^{-i\lambda^2 s/(2h)}$ in the definition of Π_B.

Finally, to show the formula (5.48), put $\lambda=1$, take an arbitrary $Z\in\Psi^{\mathrm{comp}}$, and compute the trace

$$\operatorname{Tr}(\widetilde{X}_1\Pi_f^0(1)\widetilde{X}_2^* Z)=(2\pi h)^n\int_{-T_0}^{T_0} e^{-is/(2h)}\operatorname{Tr}(U(s)B_s(1)Z)\,ds+\mathcal{O}(h^\infty). \tag{5.52}$$

The left-hand side of (5.52) can be computed as at the end of Section 5.1, using the limiting measure μ_ξ; by Proposition 4.3, it is equal to the integral of $f(\xi_{+\infty})\sigma(\widetilde{X}_2^* Z\widetilde{X}_1)$ over the Liouville measure on S^*M, plus an $\mathcal{O}(h)$ remainder. The right-hand side of (5.52) can be computed by the trace formula (3.32), and is equal to the integral of $B_0(1)Z$ over the Liouville measure on S^*M, plus an $\mathcal{O}(h)$ remainder. Therefore,

$$\int_{S^*M}\sigma(Z)f(\xi_{+\infty})\sigma(\widetilde{X}_1\widetilde{X}_2^*)\,d\mu_L=\int_{S^*M}\sigma(Z)\sigma(B_0(1))\,d\mu_L$$

for any Z; (5.48) follows. □

Proof of (1.13). By (5.35), it is enough to approximate the sum in this formula up to twice the Ehrenfest time:

Proposition 5.13. *Fix $f \in C^\infty(\partial\overline{M})$. Put $l = 2l_e$, where l_e is defined in (5.29), and consider*

$$S_f(\lambda) := \sum_{j=1}^{l} \int_{\partial\overline{M}} f(\xi) \langle \widetilde{A}^j \chi_0 E_h^0(\lambda,\xi), \chi_0 E_h^0(\lambda,\xi)\rangle \, d\xi. \tag{5.53}$$

If $\xi_{+\infty}(m,\nu)$ is the limit of $g^t(m,\nu)$ as $t \to +\infty$, for $(m,\nu) \in S^*M \setminus \Gamma_-$ (see (G3)), then for $\lambda \in [1, 1+h]$,

$$S_f(\lambda) = \int_{S^*M} f(\xi_{+\infty}) \sigma(A) \, d\mu_L + \mathcal{O}(r(h,\Lambda_0)). \tag{5.54}$$

Here $r(h,\Lambda)$ is defined in (1.10). Moreover, for each k

$$\sup_{\lambda\in[1,1+h]} |\partial_\lambda^k S_f(\lambda)| \leq C_k h^{-k\rho_e}, \tag{5.55}$$

where ρ_e is defined in (5.18).

Proof. First of all, take a compactly supported operator $\widetilde{X} \in \Psi^{\mathrm{comp}}$ such that $\mathrm{WF}_h(\widetilde{X}) \cap S^*M$ lies inside the set $\mathcal{DE}_+ \cap \{x \leq \varepsilon_1\}$ and for X_0 defined in (5.36),

$$\varphi U(t_0)(1-\varphi)\varphi_{t_0} X_0 (1-\widetilde{X}) = \mathcal{O}(h^\infty)_{L^2\to L^2},$$
$$\varphi_1 U(t_0)(1-\varphi_2)\varphi_{t_0} X_0 (1-\widetilde{X}) = \mathcal{O}(h^\infty)_{L^2\to L^2}.$$

Such an operator exists by Lemma 5.5 (it can be easily seen that in part 1 of this lemma, (m,ν) actually lies in the interior of $\mathcal{DE}_+$). Then by (5.37), the definition (5.35) of $\widetilde{A}^j$, and Lemma 5.12,

$$\begin{aligned} S_f(\lambda) &= \sum_{j=1}^{l} \int_{\partial\overline{M}} f(\xi) \langle \widetilde{A}^j X_0 \widetilde{X} \chi_0 E_h^0(\lambda,\xi), \widetilde{X}\chi_0 E_h^0(\lambda,\xi)\rangle \, d\xi + \mathcal{O}(h^\infty) \\ &= \sum_{j=1}^{l} \mathrm{Tr}(\widetilde{A}^j X_0 \widetilde{X} \Pi_f^0(\lambda) \widetilde{X}^*) + \mathcal{O}(h^\infty) \\ &= (2\pi h)^n \sum_{j=1}^{l} \int_{-T_0}^{T_0} e^{-i\lambda^2 s/(2h)} \, \mathrm{Tr}(\widetilde{A}^j X_0 U(s) B_s(\lambda)) \, ds + \mathcal{O}(h^\infty) \end{aligned}$$

for some fixed $T_0 > 0$ and some family $B_s(\lambda) \in \Psi^{\mathrm{comp}}$ smooth in s and λ and compactly supported in s; we can make B_s microlocalized inside the set $\mathcal{E}_{\varepsilon_e}$ defined in (5.17). We will henceforth ignore the $\mathcal{O}(h^\infty)$ term.

Now, take $1 \leq j \leq l$ and put $j = j_1 + j_2$, where $0 \leq j_1, j_2 \leq l_e$, $j_2 \geq 1$, and $|j_1 - j_2| \leq 1$. Using the cyclicity of the trace, we find

$$\mathrm{Tr}(\widetilde{A}^j X_0 U(s) B_s(\lambda)) = \mathrm{Tr}(U(s) B_1^j B_{2,s}^j(\lambda)),$$
$$B_1^j := (U(-t_0)\varphi_1)^{j_1} A (\varphi U(t_0))^{j_1},$$
$$B_{2,s}^j(\lambda) := (\varphi U(t_0))^{j_2} (1-\varphi)\varphi_{t_0} X_0 U(s) B_s(\lambda) \varphi_{t_0} (1-\varphi_2)(U(-t_0)\varphi_1)^{j_2} U(-s).$$

Put $\rho_j = (jt_0/t_e)\rho_e$; since $j_1, j_2 \leq j/2 + 1$, by Proposition 3.9 the operator B_1^j is a compactly supported element of $\Psi^{\mathrm{comp}}_{\rho_j/2}$ (modulo an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder which we will ignore). Same can be said about the operator

$$B^j_{2,s}(\lambda) = (\varphi U(t_0))^{j_2} \cdot (1-\varphi)\varphi_{t_0} X_0 U(s) B_s(\lambda) \varphi_{t_0}(1-\varphi_2) U(-s) \cdot (U(-t_0)\cdot U(s)\varphi_1 U(-s))^{j_2}.$$

(The operator $U(s)\varphi_1 U(-s)$ is not pseudodifferential because φ_1 is not compactly microlocalized, but this can be easily fixed by taking $\widetilde{X}_0 \in \Psi^{\mathrm{comp}}$ equal to the identity on a sufficiently large compact set and replacing $U(s)\varphi_1 U(-s)$ by $U(s)\varphi_1 \widetilde{X}_0 U(-s)$ in $B_{2,s}(\lambda)$, with an $\mathcal{O}(h^\infty)$ error.) Therefore, $B^j_1 B^j_{2,s}(\lambda)$ also lies in $\Psi^{\mathrm{comp}}_{\rho_j/2}$; moreover, it depends smoothly on s and λ, uniformly in this operator class. (In principle, we get powers of $l \sim \log(1/h)$ when differentiating in s, due to the $(U(-t_0)\cdot U(s)\varphi_1 U(s))^{j_2}$ term, but they can be absorbed into the powers of h in the expansion (3.32).)

We now use the trace formula of Lemma 3.12, writing

$$S_f(\lambda) = (2\pi h)^n \sum_{j=1}^{l} \int_{-T_0}^{T_0} e^{-i\lambda s^2/(2h)} \operatorname{Tr}(U(s) B^j_1 B^j_{2,s}(\lambda))\, ds. \tag{5.56}$$

The operator $B_{2,s}(\lambda)^j$ is microsupported, in the sense of Definition 3.1, inside the set $g^{-j_2 t_0}(\{\varphi_2 \neq 1\}) \cap g^{-(j_2-1)t_0}(\{\varphi_1 \neq 0\})$; by Lemma 5.5, this set lies inside $g^{-j_2 t_0}(\mathcal{DE}_+)$ and in particular does not intersect any closed geodesics, therefore (3.31) holds. The estimate (5.55) now follows immediately from (3.32). The power $h^{-k\rho_e}$ arises because we integrate over the energy surface $\{|\nu|_g = \lambda\}$ depending on λ; therefore, $\partial^k_\lambda S_f(\lambda)$ will involve kth derivatives of the full symbol of $B^j_1 B^j_{2,s}(\lambda)$ in the direction transversal to the energy surface, which are bounded by $h^{-k\rho_j}$. The sum (5.56) has $l \sim \log(1/h)$ terms; however, our estimate is not multiplied by $\log(1/h)$ because one can see that the sum of Liouville measures of the sets where these terms are microsupported is bounded.

As for the approximation (5.54), we write (note that we take $s = 0$ in $B_{2,s}$)

$$\sigma(B^j_1) = (\sigma(A)\circ g^{-j_1 t_0}) \prod_{k=1}^{j_1-1} \varphi \circ g^{-k t_0},$$

$$\sigma(B^j_{2,0}(\lambda))|_{S^*M} = \big((1-\varphi)\sigma(B_0(\lambda))\big)\circ g^{j_2 t_0} \prod_{k=0}^{j_2-1} \varphi\circ g^{k t_0}.$$

Since the Liouville measure is invariant under the geodesic flow, the contribution of the principal term of (3.32) to $S_f(\lambda)$ for $\lambda = 1$ is

$$\int_{S^*M} \sigma(A) \sum_{j=1}^{l} \big((1-\varphi)\sigma(B_0(\lambda))\big)\circ g^{j t_0} \prod_{k=1}^{j-1} \varphi \circ g^{k t_0}\, d\mu_L.$$

Now, by (5.48), $\sigma(B_0(1)) = f(\xi_{+\infty})$ on the support of the integrated expression; recombining the terms as in the proof of Proposition 5.11, we get the right-hand side of (5.54), with an $\mu_L(\mathcal{T}(l t_0))$ remainder. The subprincipal terms (and also the difference $S_f(\lambda) - S_f(1)$ for $\lambda \in [1, 1+h]$) are estimated using the bound on the Liouville measure of the set where $B^j_1 B^j_{2,s}$ is microsupported; arguing as in the proof of Proposition 5.11, we see that they are bounded by a constant times $r(h, \Lambda_0)$. □

Expansion of the trace of spectral projectors in powers of h. We now use the results obtained so far to derive an asymptotic expansion for the trace of the product of the spectral projector $\mathbb{1}_{[0,\lambda^2]}(P(h))$ with a compactly supported pseudodifferential operator, with the remainder depending on the classical escape rate for up to twice the Ehrenfest time. Here we denote

$$P(h) := h^2(\Delta - c_0), \tag{5.57}$$

with the constant c_0 from (A1). It will also be more convenient for us to use the spectral parameter $s = \lambda^2$ in the following corollary and theorem (not to be confused with the time variable s used in Lemma 5.12).

We start with the following consequence of the decomposition (5.35), the bound (5.54), and the spectral formula (4.5):

Corollary 5.14. *Take $\Lambda_0 > \Lambda_{\max}$, with $\Lambda_{\max}$ defined in (1.11), and let $\mathcal{T}(t)$ be defined in (1.9). For $\varepsilon > 0$, let $\varphi \in C_0^\infty((1-\varepsilon, 1+\varepsilon))$ equal to 1 near $[1-\varepsilon/2, 1+\varepsilon/2]$ and for $s \in \mathbb{R}$, define $\varphi_s := \varphi \cdot \mathbb{1}_{(-\infty,s]}$. If $\varepsilon > 0$ is small enough, then for each compactly supported $A \in \Psi^0(M)$, there exist some functions $S_h(s), Q_h(s)$ and some constants $C > 0$, $C_k > 0$ such that for all $s \in \mathbb{R}$ and all $k \in \mathbb{N}$*

$$\begin{gathered}\operatorname{Tr}\big(A\varphi_s(P(h))\big) = S_h(s) + Q_h(s),\ |\partial_s^k S_h(s)| \leq C_k h^{-n-1-k/2},\\ |Q_h(s+u) - Q_h(s)| \leq C h^{-n} \mu_L\bigg(\mathcal{T}\bigg(\frac{|\log h|}{\Lambda_0}\bigg)\bigg) + \mathcal{O}(h^\infty) \text{ for } u \in [0,h].\end{gathered} \tag{5.58}$$

Proof. By (4.5),

$$\operatorname{Tr}\big(A\varphi_s(P(h))\big) = (2\pi h)^{-n-1} \int_{\sqrt{1-\varepsilon}}^{\sqrt{1+\varepsilon}} \lambda^n f_\Pi(\lambda/h)\varphi_s(\lambda^2) \int_{\partial \overline{M}} \langle AE_h(\lambda,\xi), E_h(\lambda,\xi)\rangle \, d\xi d\lambda.$$

Now, note that the decomposition (5.35) (with $l = 2l_e$) is still valid in any $\mathcal{O}(h)$ sized interval inside $(\sqrt{1-\varepsilon}, \sqrt{1+\varepsilon})$, if ε is small enough. More precisely, if we write

$$S_h(s) := (2\pi h)^{-n-1} \int_{\sqrt{1-\varepsilon}}^{\sqrt{1+\varepsilon}} \lambda^n f_\Pi(\lambda/h)\varphi_s(\lambda^2) S_1(\lambda)\, d\lambda,$$

where $S_1(\lambda)$ is defined by (5.53) with $f(\xi) \equiv 1$, then we have the expansion (5.58) with $Q_h(s)$ satisfying the required bound. To estimate the derivatives of $S_h(s)$, it now suffices to use the bound (5.54), noting that it is valid for $|\lambda^2 - 1| < \varepsilon$ if ε is small enough. □

Using the last corollary, we can show the following trace decomposition with a fractal remainder:

Theorem 4. *Let $P(h)$ be defined in (5.57), let $A = \operatorname{Op}_h(a) \in \Psi^0(M)$ be a compactly supported operator, then there exist some smooth differential operators L_j of order $2j$ on*

T^*M, depending on the quantization procedure Op_h, with $L_0 = 1$, such that for any compact interval $I \subset (0,\infty)$, all $s \in I$, all $h > 0$ small, and all $N \in \mathbb{N}$

$$\mathrm{Tr}\big(A\,\mathbb{1}_{[0,s]}(P(h))\big) = (2\pi h)^{-n-1}\sum_{j=0}^{N} h^j \int\limits_{|\nu|_g^2\leq s} L_j a\, d\mu_\omega + h^{-n}\mathcal{O}\big(\mu_L(\mathcal{T}(\Lambda_0^{-1}|\log h|)) + h^N\big)$$

where the remainder is uniform in s. Here μ_ω is the standard volume form on T^*M; we have $\mu_\omega = \omega_S^{n+1}/(n+1)!$, where ω_S is the symplectic form.

Proof. By rescaling h, it suffices to prove the result for $|s-1| \leq \varepsilon/2$ where $\varepsilon > 0$ is obtained in Corollary 5.14, we can thus assume $|s-1| \leq \varepsilon/2$. Let φ_s be defined as in Corollary 5.14, and $\psi \in C_0^\infty((-1+\varepsilon/2, 1-\varepsilon/2))$ such that $\psi + \varphi = 1$ on $[0, 1+\varepsilon/2]$. For $s \in (1-\varepsilon/2, 1+\varepsilon/2)$, one has $\mathbb{1}_{[0,s]}(P(h)) = \varphi_s(P(h)) + \psi(P(h))$ and it suffices to study the expansion in h of $\sigma_{A,h}(s)$ and $\mathrm{Tr}(A\psi(P(h)))$ where

$$\sigma_{A,h}(s) := \mathrm{Tr}\,\big(A\varphi_s(P(h))\big) = \mathrm{Tr}\,\big(A\varphi_s(P(h))\chi\big). \tag{5.59}$$

if $\chi \in C_0^\infty(M)$ is such that $A = \chi A \chi$. Since A is compactly supported, one can use the functional calculus of Helffer–Sjöstrand [DiSj, Chapters 8-9] to deduce that $A\psi(P(h))\chi \in \Psi^{\mathrm{comp}}(M)$ is a compactly supported and microsupported pseudodifferential operator with a classical symbol[2]. Its trace thus has a complete expansion in powers of h (see [DiSj, Th 9.6]):

$$\mathrm{Tr}(A\psi(P(h))\chi) = (2\pi h)^{-n-1}\sum_{j=0}^{N} h^j \int_{T^*M} L_j'' a\, d\mu_\omega + \mathcal{O}(h^{-n+N}) \tag{5.60}$$

where L_j'' are some differential operators of order $2j$ and $L_0'' a(m,\nu) = a(m,\nu)\psi(|\nu|_g^2)$.

Let us now analyze the function $\sigma_{A,h}$. This is a smooth function of $s > 0$ by the smoothness assumption on the $E_h(\lambda,\xi)$ in λ, it is constant in s for $|1-s| > \varepsilon$, and we know that $\sigma_{A,h}(s) = \mathcal{O}(h^{-n-1})$ uniformly in s by Lemma 3.11. Let $\theta(s) \in \mathscr{S}(\mathbb{R})$ be a Schwartz function such that $\hat\theta \in C_0^\infty(-\eta,\eta)$ for some small $\eta > 0$ and $\hat\theta(t) = 1$ near $t = 0$, and let $\theta_h(s) = h^{-1}\theta(s/h)$. We write

$$\sigma'_{A,h}(s) := \partial_s \sigma_{A,h}(s) = \mathrm{Tr}(A\varphi(P(h))d\Pi_s(P(h))\chi) \in C_0^\infty((0,\infty))$$

where $d\Pi_s(P(h))$ is the spectral measure of $P(h)$. The operator $A\varphi(P(h))d\Pi_s(P(h))\chi$ has a smooth compactly supported kernel and is trace class. We clearly have $\sigma'_{A,h} \star \theta_h \in \mathscr{S}(\mathbb{R})$ and by a simple computation, its semi-classical Fourier transform is given by

$$\int_{\mathbb{R}} e^{-i\frac{t}{h}s}\sigma'_{A,h} \star \theta_h(s)ds = \mathrm{Tr}(A\varphi(P(h))e^{-i\frac{t}{h}P(h)})\hat\theta(t)$$

and thus

$$\sigma'_{A,h} \star \theta_h(s) = \frac{1}{2\pi h}\int_{\mathbb{R}} e^{i\frac{s}{h}t}\,\mathrm{Tr}(A\varphi(P(h))e^{-i\frac{t}{h}P(h)})\hat\theta(t)dt.$$

[2] An alternative method in the Euclidean near infinity setting is the functional calculus of Helffer–Robert [HeRo].

Now we can apply Lemma 3.12 with $B_u = \frac{1}{2}e^{ic_0hu/2}\hat{\theta}(-u/2)A\varphi(P(h))$; the condition (3.31) is satisfied because $\hat{\theta}$ is supported in a small neighborhood of zero. This shows that, as $h \to 0$, we have the expansion (locally uniformly in s)

$$\sigma'_{A,h} \star \theta_h(s) = (2\pi h)^{-n-1}\Big(\sum_{j=0}^{N} h^j \int_{S^*M} \widetilde{L}_j b(m, \sqrt{s}\nu) d\mu_L(m,\nu) + \mathcal{O}(h^{N+1})\Big)$$

for all $N \in \mathbb{N}$, where b is a symbol such that $\mathrm{Op}_h(b) = \frac{1}{2}A\varphi(P(h)) + \mathcal{O}(h^\infty)$, $\widetilde{L}_j$ are differential operators of order $2j$ on T^*M, smooth in $\sqrt{s}$, with $\widetilde{L}_0 = s^{\frac{n-1}{2}}$. In particular, one has $\widetilde{L}_0 b(m,\sqrt{s}\nu) = \frac{1}{2}s^{\frac{n-1}{2}}a(m,\sqrt{s}\nu)\varphi(s|\nu|_g^2)$. Notice that b is supported in $\{(m,\nu) \in T^*M \mid |\nu|_g^2 \in \operatorname{supp}\varphi\}$ thus $\widetilde{L}_j b(m,\sqrt{s}\nu)$ is smooth in $s \in \mathbb{R}$ when $(m,\nu) \in S^*M$. Since $\sigma_{A,h}(s)$ is bounded in s, the convolution $\sigma_{A,h} \star \theta_h(s)$ is well defined (as an element in $L^\infty(\mathbb{R})$) and we have for all N and for all $s \leq 2$

$$\sigma_{A,h} \star \theta_h(s) = (2\pi h)^{-n-1}\Big(\sum_{j=0}^{N} h^j \int_0^s \int_{S^*M} \widetilde{L}_j b(m, \sqrt{u}\nu) d\mu_L(m,\nu) du + \mathcal{O}(h^{N+1})\Big).$$

We are going to show that, uniformly in $s \in \mathbb{R}$,

$$\sigma_{A,h}(s) - \sigma_{A,h} \star \theta_h(s) = \mathcal{O}(h^\infty) + \mathcal{O}\big(h^{-n}\mu_L(\mathcal{T}(\Lambda_0^{-1}|\log h|))\big), \tag{5.61}$$

using the decomposition

$$\sigma_{A,h}(s) = S_h(s) + Q_h(s)$$

defined in (5.58). Since $\partial_s S_h(s)$ is a compactly supported symbol we get by integrating by parts N times

$$(1-\hat{\theta}(t))\int e^{-i\frac{t}{h}s}\partial_s S_h(s) ds = (1-\hat{\theta}(t))\int e^{-i\frac{t}{h}s}\Big(\frac{h}{it}\Big)^N \partial_s^{N+1}(S_h(s)) ds = \mathcal{O}\Big(\frac{h^{N/4}}{(1+|t|)^N}\Big)$$

for all $t \in \mathbb{R}$ and all $N \gg 1$. Thus, taking the Fourier transform we deduce that $S_h(s) - S_h \star \theta_h(s) = \mathcal{O}(h^\infty)$ uniformly in s. From (5.58), we obtain by induction that for all s, u

$$|Q_h(s+u) - Q_h(s)| \leq Ch^{-n}\Big(1 + \frac{|u|}{h}\Big)\mu_L(\mathcal{T}(\Lambda_0^{-1}|\log h|)) + \mathcal{O}(h^\infty)$$

then multiplying by $\theta_h(-u)$ and integrating in u, we obtain (5.61).

Given (5.61), we have

$$\begin{aligned}\sigma_{A,h}(s) = (2\pi h)^{-n-1}&\sum_{j=0}^{N} h^j \int_0^s \int_{S^*M} \widetilde{L}_j b(m, \sqrt{u}\nu) d\mu_L(m,\nu) du \\ &+\mathcal{O}\big(h^{-n}\mu_L(\mathcal{T}(\Lambda_0^{-1}|\log h|))\big) + \mathcal{O}(h^{-n+N}).\end{aligned} \tag{5.62}$$

Since the symbol of b is explicitly obtained from a using Moyal product, we can rewrite this expression with a instead of b and with some new differential operators with the same properties as $\widetilde{L}_j$ but supported in $\{|\nu|_g^2 \in \operatorname{supp}\varphi\}$; using polar coordinates $S^*M \times \mathbb{R}^+_{\sqrt{u}}$ on T^*M, we deduce that there exist some differential operators L'_j of order $2j$ on T^*M such that

$$\int_0^s \int_{S^*M} \widetilde{L}_j b(m, \sqrt{u}\nu) d\mu_L(m,\nu) du = \int_{|\nu|_g^2 \leq s} L'_j a(m,\nu) d\mu_\omega(m,\nu)$$

and $L_0'a(m,\nu) = \varphi(|\nu|_g^2)a(m,\nu)$. Combining this with (5.62) and (5.60), we obtain the desired result where L_j in the statement of the Theorem corresponds now to $L_j' + L_j''$. □

6. Euclidean near infinity manifolds

In this section, we assume that (M,g) is a complete Riemannian manifold such that there exists a compact set $K_0 \subset M$ such that for $\mathcal{E} := M \setminus K_0$,

$$(\mathcal{E}, g) \text{ is isometric to } (\mathbb{R}^{n+1} \setminus B(0,R), g_{\mathrm{eucl}})$$

where $R > 0$, $B(0,R)$ is the Euclidean ball of center 0 and radius R and g_{eucl} is the Euclidean metric. We will check that all the assumptions of Section 4 are satisfied.

6.1. **Geometric assumptions.** We let $x \in C^\infty(M)$ be an everywhere positive function equal to $x(m) = |m|^{-1}$ in $\mathcal{E}$ identified with $\mathbb{R}^{n+1} \setminus B(0,R)$, and such that $x \geq R^{-1}$ in K_0. (We take it instead of the function $(1+|m|^2)^{-1/2}$ used in Section 4 for the model case of $\mathbb{R}^{n+1}$, to simplify the calculations and since we no longer need smoothness at zero.) We shall use the polar coordinates $m = \omega/x$ in $\mathcal{E}$, where $\omega \in \mathbb{S}^n$. Assumption (G1) is satisfied by taking the radial compactification of M, i.e. adding the sphere at infinity: the map $\psi : \mathbb{R}^{n+1} \setminus B(0,R) \to (0, 1/R) \times \mathbb{S}^n$ defined by $\psi(m) = (x(m), x(m)m)$ is a diffeomorphism and the radial compactification of M is obtained by setting $\overline{M} = M \sqcup \partial\overline{M}$ where $\partial\overline{M} := \mathbb{S}^n$, the smooth structure on $\overline{M}$ is the same as before on M but we extend it to $\overline{M}$ by asking that ψ extends smoothly to the boundary $\partial\overline{M}$ and $\psi(\xi) = (0,\xi)$ if $\xi \in \partial\overline{M} = \mathbb{S}^n$ (see for instance [Me] for more details). In other words, smooth functions on $\overline{M}$ are smooth functions on M with an asymptotic expansion in integer powers of $1/|m|$ to any order near infinity.

Assumption (G2) is clearly satisfied for $\varepsilon_0 := 1/2R$ since the trajectories of the geodesic flow in $x \leq \varepsilon_0$ are simply $g^t(m,\nu) = (m + t\nu, \nu)$. A point $(m,\nu) \in S^*M$ is directly escaping in the forward direction in the sense of Definition 4.2 if and only if $x(m) \leq \varepsilon_0$ and $m \cdot \nu \geq 0$. Now, (G3) is satisfied with $\xi_{+\infty}(m,\nu) = \nu$ for $(m,\nu) \in \mathcal{DE}_+$.

For the assumption (G4), we define

$$\widetilde{U}_\infty = \{x < \varepsilon_0\} \times \partial\overline{M} \subset \overline{M} \times \partial\overline{M},$$
$$\phi_\xi(m) = m \cdot \xi, \ (m,\xi) \in U_\infty.$$

Then $\tau : U_\infty \to S^*M$ from (4.2) maps each $(m,\xi) \in (\mathbb{R}^n \setminus B(0,2R)) \times \mathbb{S}^n$ to itself as an element of $S^*(\mathbb{R}^n \setminus B(0,2R))$. Assumptions (G4) and (G5) follow immediately. To see assumption (G6), we note that for $x(m), x(m') < \varepsilon_0$ and some $\xi \in \mathbb{S}^n$, we have $\partial_\xi \phi_\xi(m) = \partial_\xi \phi_\xi(m')$ if and only if $m - m'$ is a multiple of ξ.

6.2. **Distorted plane waves and analytic assumptions.** We recall a few well known facts about scattering theory on perturbations of $\mathbb{R}^n$, we refer to [Me] for a geometric approach and to [MeZw, HaVa] in a more general setting (asymptotically Euclidean case). A plane wave for the flat Laplacian on $\mathbb{R}^{n+1}$ is the function, for $\lambda \in (1/2, 2)$,

$$u(\lambda; m, \xi) := ce^{\frac{i\lambda}{h} m\cdot\xi}, \quad \xi \in \mathbb{S}^n, \ m \in \mathbb{R}^{n+1}, \ c \in \mathbb{C}. \tag{6.1}$$

This is a semiclassical Lagrangian distribution, its oscillating phase has level sets given by planes orthogonal to ξ. The continuous spectrum of the Laplacian Δ associated to the metric g is the half-line $[0,\infty)$. We will take the resolvent of $h^2\Delta$ to be the L^2-bounded operator

$$R_h(\lambda) := (h^2\Delta - \lambda^2)^{-1} \text{ in } \mathrm{Im}(\lambda) < 0.$$

This admits a continuous extension to $\{\lambda \neq 0, \mathrm{Im}(\lambda) \leq 0\}$ as a bounded operator from L^2_{comp} to L^2_{loc}. For $\lambda > 0$ we call $R_h(\lambda)$ the incoming resolvent and $R_h(-\lambda)$ the outgoing resolvent

The *distorted plane wave* is defined for $\xi \in \mathbb{S}^n$ (see [MeZw, HaVa]) by

$$E_h(\lambda,\xi;m) := 2i\lambda h\Big(\frac{2\pi h}{i\lambda}\Big)^{\frac{n}{2}} \lim_{x'\to 0}[(x')^{-n/2}e^{\frac{i\lambda}{hx'}}R_h(\lambda;m,\xi/x')], \tag{6.2}$$

where $R_h(\lambda;m,m')$ is the Schwartz kernel of $R_h(\lambda)$ and $\xi/x' \in \mathcal{E}$. This is a smooth function of $(m,\xi) \in M\times\mathbb{S}^n$, and in the case of $M = \mathbb{R}^{n+1}$ it is given by (6.1) with $c = 1$ (see [Me, Chapter 1]). We shall use the notation $E_h(\lambda,\xi)$ for the $C^\infty(M)$ function defined by $m \mapsto E_h(\lambda,\xi;m)$ and we notice that $(h^2\Delta - \lambda^2)E_h(\lambda,\xi) = 0$ in M. One has $\overline{E_h(\lambda;m,\xi)} = E_h(-\lambda;m,\xi)$ since $R_h(\lambda)^* = R(-\lambda)$ for $\lambda\in\mathbb{R}$, and the decomposition of the spectral measure in terms of these functions in given as follows: by Stone's formula, the semiclassical spectral measure is given by

$$d\Pi_h(\lambda) = \frac{i\lambda}{\pi}(R_h(\lambda) - R_h(-\lambda))\,d\lambda \quad \text{for } \lambda\in(0,\infty) \tag{6.3}$$

in the sense that $F(h^2\Delta) = \int_0^\infty F(\lambda^2)d\Pi_h(\lambda)$ for any bounded function F; by combining this with the Green's type formula of [HaVa, Lemma 5.2], we deduce that

$$d\Pi_h(\lambda;m,m') = \lambda^n(2\pi h)^{-n-1}\int_{\mathbb{S}^n} E_h(\lambda;m,\xi)\overline{E_h(\lambda,m',\xi)}\,d\xi d\lambda.$$

Here $d\xi$ corresponds to the standard volume form on the sphere $\mathbb{S}^n$. The assumptions (A1) and (A2) are then satisfied. In fact, using [HaVa], one can define distorted plane waves and verify assumptions (A1) and (A2) for the more general case of scattering manifolds.

Outgoing/incoming decomposition. We now construct the decomposition (4.6) of E_h into the outgoing and incoming parts and verify assumptions (A3)–(A8). Take $\chi_0 \in C^\infty(\overline{M})$ (thus constant in ξ) supported in $\{x < \varepsilon_0\}$ and equal to 1 near $\{x \leq \varepsilon_0/2\}$, so that assumptions (A3) and (A7) hold, where we put $\varepsilon_1 := \varepsilon_0/2$. We next put

$$E_h^0(\lambda,\xi;m) := e^{\frac{i\lambda}{h}m\cdot\xi},\ x(m) < \varepsilon_0,$$

so that (A4) holds with $b^0 \equiv 1$ and (A8) follows. We then claim that

$$E_h = \chi_0 E_h^0 + E_h^1, \tag{6.4}$$

where

$$E_h^1 := -R_h(\lambda)F_h,\ F_h(\lambda,\xi) = (h^2\Delta - \lambda^2)(\chi_0 E_h^0(\lambda,\xi)) = [h^2\Delta,\chi_0]E_h^0(\lambda,\xi).$$

We can apply $R_h(\lambda)$ to $F_h(\lambda,\xi)$ as the latter lies in $C_0^\infty(M)$; in fact, $\operatorname{supp} F_h \subset \{\varepsilon_0/2 < x < \varepsilon_0\}$. To show (6.4), note that the incoming resolvent $R_h(-\lambda)$ satisfies

$$R_h(\lambda)\chi_1 = \chi_0 R_h^0(\lambda)\chi_1 - R_h(\lambda)[h^2\Delta,\chi_0]R_h^0(\lambda)\chi_1$$

if $\chi_1 \in C^\infty(\overline{M})$ is such that $\chi_0\chi_1 = \chi_1$ and $R_h^0(\lambda)$ is the incoming scattering resolvent of the free semiclassical Laplacian $h^2\Delta$ on $\mathbb{R}^{n+1}$ (we use again the isometry $\mathcal{E} \simeq \mathbb{R}^{n+1} \setminus B(0,R)$). Multiplying the last equation on the right by $x^{-\frac{n}{2}} e^{\frac{i\lambda}{hx}}$, taking the Schwartz kernel of the obtained operators and considering the limit in the second variable along a line with tangent vector ξ as in (6.2) (using the formula (6.1) with $c = 1$ for the expression (6.2) for the free resolvent $R_h^0(\lambda)$), we obtain (6.4).

Microlocalization of E_h^1. It remains to verify assumptions (A5) and (A6). By rescaling h and using that $E_h(\lambda,\cdot)$ depends only on λ/h, we may assume that $\lambda = 1$. Fix ξ and take $\chi_2 \in C^\infty(\overline{M})$ equal to 1 near $\{x \le \varepsilon_0\}$, but supported inside $\mathcal{E}$. Then

$$\chi_2 E_h^1 = R_h^0(\lambda) F_h^0,\ F_h^0 := (h^2\Delta - \lambda^2)(\chi_2 E_h^1) = F_h + [h^2\Delta, \chi_2] E_h^1. \tag{6.5}$$

The function F_h^0 is supported inside $\{x > \varepsilon_0/2\}$ and

$$\|F_h^0\|_{H_h^{-1}} \le Ch(1 + \|E_h\|_{L^2(\{x \ge \varepsilon_0\})}).$$

The free resolvent $R_h^0(\lambda)$ is bounded $H^{-1}_{h,\mathrm{comp}} \to L^2_{\mathrm{loc}}$ with norm $\mathcal{O}(h^{-1})$ by [Bu02, Prop 2.1]; therefore, for each compact set $K \subset M$, there exists a constant C_K such that

$$\|E_h^1\|_{L^2(K)} \le C_K(1 + \|E_h\|_{L^2(\{x \ge \varepsilon_0\})}).$$

This shows (A5), namely that the function

$$\widetilde{E}_h^1 := \frac{E_h^1}{1 + \|E_h\|_{L^2(\{x \ge \varepsilon_0\})}}$$

is h-tempered. To prove (A6), we use semiclassical elliptic estimate and propagation of singularities (see for example [Va11, Section 4.1]). We have

$$(h^2\Delta - \lambda^2)\widetilde{E}_h^1 = -\widetilde{F}_h,\ \widetilde{F}_h := \frac{F_h}{1 + \|E_h\|_{L^2(\{x \ge \varepsilon_0\})}}.$$

Now, F_h is a Lagrangian distribution associated to $\{(m,\xi) \mid m \in \mathrm{supp}(d\chi_0)\}$; therefore,

$$\mathrm{WF}_h(\widetilde{F}_h) \subset \mathrm{WF}_h(F_h) \subset W_\xi,$$

with $W_\xi \subset S^*M$ defined in (4.9).

Take $(m,\nu) \in \mathrm{WF}_h(\widetilde{E}_h^1)$. By the elliptic estimate, $(m,\nu) \in S^*M$. Next, if $\gamma(t) = g^t(m,\nu)$, then by propagation of singularities, either $\gamma(t) \in \mathrm{WF}_h(\widetilde{F}_h) \subset W_\xi$ for some $t \ge 0$ or $\gamma(t) \in \mathrm{WF}_h(\widetilde{E}_h^1)$ for all $t \ge 0$. Now, the free resolvent $R_h^0(\lambda)$ is semiclassically incoming in the following sense: if f is a compactly supported h-tempered family of distributions, then for each $(m',\nu') \in \mathrm{WF}_h(R_h^0(\lambda)f)$, there exists $t \ge 0$ such that $g^t(m',\nu') \in \mathrm{supp}\, f$. This can be seen for example from the explicit formulas for $R_h^0(\lambda)$, see [Me]. By (6.5) and since $\mathrm{supp}(F_h^0) \subset \{x > \varepsilon_0/2\}$, we see that for $(m',\nu') \in \mathrm{WF}_h(\widetilde{E}_h^1)$, we cannot have $x(m') < \varepsilon_0/2$ and $m' \cdot \nu' \ge 0$. Therefore, if $\gamma(t) \notin W_\xi$ for all $t \ge 0$, then $\gamma(t)$ is trapped as $t \to +\infty$; this proves (A6).

7. Hyperbolic near infinity manifolds

In this section, we verify the assumptions of Section 4 for certain asymptotically hyperbolic manifolds. Let (M, g) be an $(n+1)$-dimensional asymptotically hyperbolic manifold as defined in the introduction. It has a compactification $\overline{M} = M \cup \partial\overline{M}$ and the metric can be written in the product form (1.3):

$$g = \frac{dx^2 + h(x)}{x^2}$$

where x is a boundary defining function and $h(x)$ a smooth family of metrics on $\partial\overline{M}$ defined near $x = 0$. The function x putting the metric in the form (1.3) is not unique, and those functions (thus satisfying $|d\log(x)|_g = 1$ near $\partial\overline{M}$) are called *geodesic boundary defining functions*. The set of such functions parametrizes the conformal class of $h(0)$, as shown in [GrLe, Lemma 5.2]. The metric is called *even* if $h(x)$ is an even function of x, this condition is independent of the choice of geodesic boundary defining function. A choice of geodesic boundary defining function induces a metric on $\partial\overline{M}$ by taking $h_0 = h(0) = x^2 g|_{T\partial\overline{M}}$, and therefore one has a Riemannian volume form, denoted $d\xi$, on $\partial\overline{M}$ induced by the choice of x. Any other choice $\hat{x} = e^{\omega}x$ of boundary defining function induces a volume form

$$\widehat{d\xi} = e^{n\omega_0} d\xi \qquad \text{where } \omega_0 = \omega|_{\partial\overline{M}}. \tag{7.1}$$

We will further assume that M has constant sectional curvature -1 outside of some compact set, even though some of the assumptions of Section 4 hold for general asymptotically hyperbolic manifolds.

7.1. **Geometric assumptions.** Let (M, g) be an asymptotically hyperbolic manifold. The assumption (G1) is satisfied. We are now going to prove a Lemma which implies directly that the assumptions (G2) and (G3) are satisfied, except that this only proves continuous dependence of $\xi_{+\infty}$ in (m, ν) in (G3). To prove C^1 dependence in a general setting, a bit more analysis would be required, but we shall later concentrate only on cases with constant curvature near infinity, in which case the dependence is smooth (see below).

Lemma 7.1. *Let (M, g) an asymptotically hyperbolic manifold. Then there exists $\varepsilon_0 > 0$ such that the function x satisfies (4.1) and for any unit speed geodesic $\gamma(t) = (m(t), \nu(t))$ with $x(m(0)) \le \varepsilon_0$ and $\partial_t x(m(t))|_{t=0} \le 0$, one has the following: $\partial_t x(m(t)) \le 0$ for all $t \ge 0$ and $m(t)$ converges in the topology of $\overline{M}$ to some point $\xi_{+\infty} \in \partial\overline{M}$. More precisely, the distance with respect to the compactified metric $\bar{g} = x^2 g$ between $m(t)$ and $\xi_{+\infty}$ is bounded by*

$$d_{\bar{g}}(m(t), \xi_{+\infty}) \le Ct^{-1}.$$

Proof. Consider coordinates $(m, \nu) = (x, y; \rho dx + \theta \cdot dy)$ on T^*M near the boundary $\partial\overline{M} = \{x = 0\}$. The geodesic flow is the Hamiltonian flow of $p/2$, where $p(m, \nu) = x^2(\rho^2 + |\theta|^2_{h_m})$; if dots denote time derivatives with respect to the geodesic flow, we get

$$\dot{x} = \rho x^2, \quad \dot{\rho} = -x^{-1}p(m, \nu) - x^2 \partial_x h_{(x,y)}(\theta, \theta)/2. \tag{7.2}$$

Since $\partial_x h_{(x,y)}$ is smooth up to $x = 0$, there exists a constant C such that

$$|x^2 \partial_x h_{(x,y)}(\theta, \theta)/2| \le Cx^2 h_{(x,y)}(\theta, \theta) \le Cp(m, \nu).$$

Therefore, there exists $\varepsilon_0 > 0$ such that along any unit speed geodesic, we have

$$x \leq \varepsilon_0 \Longrightarrow \dot{\rho} = -x^{-1} + \mathcal{O}(1) \leq -x^{-1}/2 < 0. \tag{7.3}$$

This in particular implies (4.1).

Now, let $\gamma(t) = (x(t), y(t); \rho(t), \theta(t))$ be a unit speed geodesic and assume that $x(0) \leq \varepsilon_0$ and $\dot{x}(0) \leq 0$. It follows from (4.1) that for $t \geq 0$, we have $\dot{x}(t) \leq 0$ and thus $x(t) \leq \varepsilon_0$. (Indeed, for each $s \geq 0$ the minimal value of $x(t)$ on the interval $[0, s]$ has to be achieved at $t = s$.) It remains to show that as $t \to +\infty$, $x(t)$ converges to 0 and $y(t)$ converges to some $\xi_{+\infty} \in \partial M$. For that, note that by (7.3), $\dot{\rho}(t) \leq -\varepsilon_0^{-1}/2$ for $t \geq 0$; since $\dot{x}(0) \leq 0$, we have $\rho(0) \leq 0$ and thus

$$\rho(t) \leq -\frac{\varepsilon_0^{-1}}{2} t.$$

Setting $u(t) := x(t)^{-1}$, we find $\dot{u}(t) = -\rho(t) \geq (\varepsilon_0^{-1}/2)t$; therefore,

$$x(t) \leq \frac{\varepsilon_0}{1 + t^2/4}.$$

In particular, $x(t) \to 0$ as $t \to +\infty$. Now the equation for $\dot{y}(s)$ tells us that

$$\dot{y}_i(t) = x \sum_{k=1}^{n} h^{ki}_{(x,y)} x \theta_k = \mathcal{O}(x(t)) = \mathcal{O}(t^{-2})$$

and therefore $|y(t) - y(t')| \leq C/t'$ for any $t > t' > 0$, which implies $\lim_{t \to \infty} m(t) = \xi_\infty$ for some $\xi_\infty \in \partial \overline{M}$ and $|m(t) - \xi_\infty| = \mathcal{O}(1/t)$. □

The geometric assumption (G4) is a more complicated one, and we will restrict ourselves to asymptotically hyperbolic manifolds with constant curvature -1 in a neighbourhood of ∂M and x a geodesic boundary defining function. Let $\xi \in \partial \overline{M}$, then there exists a neighborhood V_ξ of ξ in $\overline{M}$, and an isometric diffeomorphism ψ_ξ from $V_\xi \cap M$ into the following neighbourhood V_{q_0} of the north pole q_0 in the unit ball $\mathbb{B} := \{m \in \mathbb{R}^{n+1}; |m| < 1\}$ equipped with the hyperbolic metric $g_{\mathbb{H}^{n+1}}$ (the hyperbolic space $\mathbb{H}^{n+1}$ is $\mathbb{B}$ equipped with $g_{\mathbb{H}^{n+1}}$):

$$V_{q_0} := \{q \in \mathbb{B} \mid |q - q_0| < 1/4\}, \quad g_{\mathbb{H}^{n+1}} = 4 \frac{dq^2}{(1 - |q|^2)^2} \tag{7.4}$$

where $\psi_\xi(\xi) = q_0$ and $|\cdot|$ denotes the Euclidean length. This statement is proved for instance in [GuZw, Lemma 3.1]. We shall choose the boundary defining function

$$x_0 = 2\frac{(1 - |q|)}{(1 + |q|)} \tag{7.5}$$

on $\mathbb{B}$. Note that we might not be able to find ψ_ξ such that $x = \psi_\xi^* x_0$. This would be possible only if the boundary is globally conformally flat and x is chosen so that the metric on ∂M induced by $x^2 g$ is flat.

We define for each $p \in \mathbb{S}^n = \partial \overline{\mathbb{B}}$ the Busemann function on $\mathbb{B}$

$$\phi_p^{\mathbb{B}}(q) = \log \Big(\frac{1 - |q|^2}{|q - p|^2} \Big).$$

The geodesic trajectory $g^t(q, d\phi_p^{\mathbb{B}}(q))$ generated by the differential $d\phi_p^{\mathbb{B}}$ converges (in the Euclidean ball topology) to p and the Lagrangian manifold

$$\Lambda_p^{\mathbb{B}} := \{(q, d\phi_p^{\mathbb{B}}(q)) \in S^*\mathbb{H}^{n+1} \mid q \in \mathbb{B}\}$$

is the stable manifold of the geodesic flow associated to p on $\mathbb{B}$. The level sets of $\phi_p^{\mathbb{B}}$ are horospheres based at p. We cover a neighbourhood of $\partial\overline{M}$ by finitely many V_{ξ_j} for some $\xi_j \in \partial\overline{M}$ and take a partition of unity $\chi_j \in C^\infty(\partial\overline{M})$ on $\partial\overline{M}$ with χ_j supported in $V_{\xi_j} \cap \partial\overline{M}$. Then there exists $\varepsilon > 0$ such that for all j and all $\xi \in \operatorname{supp}\chi_j$, the set

$$U_\xi := \{m \in \overline{M} \mid d_{\bar{g}}(m, \xi) < \varepsilon\} \tag{7.6}$$

lies inside V_{ξ_j}, where $\bar{g} = x^2 g$ is the compactified metric. Put

$$U_\infty := \{(m, \xi) \in M \times \partial\overline{M} \mid m \in U_\xi\}.$$

Define the function

$$\phi_\xi(m) := \sum_j \chi_j(\xi)\phi^{\mathbb{B}}_{\psi_{\xi_j}(\xi)}(\psi_{\xi_j}(m)), \ (m, \xi) \in U_\infty. \tag{7.7}$$

Since ψ_{ξ_j} are isometries, each function $\phi_\xi^j(m) := \phi^{\mathbb{B}}_{\psi_{\xi_j}(\xi)}(\psi_{\xi_j}(m))$ is such that $d\phi_\xi^j(m)$ is the unit covector which generates the unique geodesic in M starting at m, staying in U_ξ for positive times, and converging to ξ (therefore, the difference of any two functions ϕ_ξ^j for different j is a function of ξ only). Therefore $\partial_m\phi_\xi(m) = \sum_j \chi_j(\xi)\partial_m\phi_\xi^j(m)$ is also equal to this unit covector; (G4) and (G5) follow. The dependence of all objects in m, ξ is smooth here. Finally, (G6) can be reduced via ψ_{ξ_j} to the following statement that can be verified by a direct computation: if $p \in \mathbb{S}^n$ and $q, q' \in \mathbb{B}$, then $\partial_p\phi_p^{\mathbb{B}}(q) = \partial_p\phi_p^{\mathbb{B}}(q')$ if and only if q and q' lie on a geodesic converging to p, and the matrix $\partial^2_{pq}\phi_p^{\mathbb{B}}(q)$ has rank n.

7.2. **Eisenstein functions and analytic assumptions.** Let (M, g) be asymptotically hyperbolic. The Laplacian Δ on (M, g) has absolutely continuous spectrum on $[n^2/4, \infty)$ and a possibly non-empty finite set of eigenvalues in $(0, n^2/4)$. By [MaMe, Gu], if g is an even metric[3], the resolvent of the Laplacian

$$R(s) := (\Delta - s(n-s))^{-1} \quad \text{defined in the half plane } \operatorname{Re}(s) > n/2$$

admits a meromorphic continuation to the whole complex plane $\mathbb{C}$, with poles of finite rank (i.e. the Laurent expansion at each pole consists of finite rank operators), as a family of bounded operators

$$R(s) : x^N L^2(M) \to x^{-N} L^2(M), \quad \text{if} \quad \operatorname{Re}(s) - n/2 + N > 0,$$

moreover it has no poles on the line $\operatorname{Re}(s) = \frac{n}{2}$ except possibly $s = \frac{n}{2}$, as shown by Mazzeo [Ma]. Let us fix a geodesic boundary defining function x on $\overline{M}$. By [MaMe], the resolvent integral kernel $R(s; m, m')$ near the boundary $\partial\overline{M}$ has an asymptotic expansion given as follows: for any $m \in M$ fixed

$$m' \mapsto R(s; m, m')x(m')^{-s} \in C^\infty(\overline{M})$$

[3]There is a simpler proof by Guillopé-Zworski [GuZw] when the curvature is constant outside a compact set.

and similarly for $m' \in M$ fixed and $m \to \partial\overline{M}$. Since we are interested in high frequency asymptotics, we will consider the semiclassical rescaled versions

$$R_h(\lambda) := h^{-2}R(n/2 + i\lambda/h),$$

Notice that the physical region $\operatorname{Re} s > n/2$, in which the resolvent is bounded on L^2, corresponds to $\operatorname{Im}\lambda > 0$, which agrees with our convention for Euclidean case.

Definition 7.2. *Let $1/2 \leq |\lambda| \leq 2$ and $h > 0$, then Eisenstein functions are the functions in $C^\infty(M \times \partial\overline{M})$ defined for any fixed $\xi \in \partial\overline{M}$ by the following limit of the resolvent kernel at infinity*

$$\begin{aligned} E_h(\lambda, \xi; m) &:= \frac{2i\lambda h}{C(\lambda/h)} \lim_{m' \to \xi} x(m')^{-n/2 - i\lambda/h} R_h(\lambda; m, m'), \\ C(z) &:= 2^{-iz}(2\pi)^{-\frac{n}{2}} \frac{\Gamma(\frac{n}{2} + iz)}{\Gamma(iz)}. \end{aligned} \tag{7.8}$$

The normalisation constant in (7.8) is like the constant in (6.2) so that in $\mathbb{B}$, $E_h(\lambda)$ is a *horospherical wave* as described below in (7.11). For any $\xi \in \partial\overline{M}$, we will denote by $E_h(\lambda, \xi)$ the function $m \mapsto E_h(\lambda, \xi; m)$, and we observe that they solve (1.4):

$$(h^2(\Delta - n^2/4) - \lambda^2)E_h(\lambda, \xi) = 0.$$

One also has $\overline{E_h(\lambda, \xi; m)} = E_h(-\lambda, \xi; m)$ as an easy consequence of $R_h(\lambda)^* = R(-\lambda)$ for $\lambda \in \mathbb{R}$. From its definition, $E_h(\lambda, \xi)$ depends on the choice of the boundary defining function x, but considering such a change we easily see from (7.1) that the density on $\partial\overline{M}$

$$\langle AE_h(\lambda, \xi), E_h(\lambda, \xi)\rangle_{L^2(M)}\, d\xi \quad \text{for } A \in \Psi^{\mathrm{comp}}(M) \tag{7.9}$$

is *independent* of x.

Let us recall the decomposition of the spectral measure in terms of these functions. By Stone's formula, the semiclassical spectral measure is given by

$$d\Pi_h(\lambda) = \frac{i\lambda}{\pi}(R_h(\lambda) - R_h(-\lambda))\, d\lambda \quad \text{for } \lambda \in (0, \infty)$$

in the sense that $F(h^2(\Delta - n^2/4)) = \int_0^\infty F(\lambda^2) d\Pi_h(\lambda)$ for any bounded function F supported in $(0, \infty)$. Now we can write (see [Gu]) for any m, m'

$$d\Pi_h(\lambda; m, m') = \frac{|C(\lambda/h)|^2}{2\pi h} \int_{\partial\overline{M}} E_h(\lambda, \xi; m) E_h(-\lambda, \xi; m') d\xi \; d\lambda. \tag{7.10}$$

where $(2\pi h)^n |C(\lambda/h)|^2 \to \lambda^n$ as $h \to 0$ uniformly in $\lambda \in [1/2, 2]$. The assumptions (A1) and (A2) are then satisfied in the general asymptotically hyperbolic case (without asking the constant curvature near infinity).

Outgoing/incoming decomposition. To check assumptions (A3)–(A8), we give a representation of the Eisenstein functions as sums of the 'outgoing' part E_h^0 and the 'incoming' part E_h^1. We assume constant curvature near infinity in what follows. The expression for

$E_h^{\mathbb{B}}(\lambda)$ on hyperbolic space $\mathbb{H}^{n+1}$ viewed as a unit ball $\mathbb{B}$, defined using the boundary defining function x_0 of (7.5), is given by [GuNa, Section 2.2]

$$E_h^{\mathbb{B}}(\lambda,p;q)=\Big(\frac{1-|q|^2}{|q-p|^2}\Big)^{n/2+i\lambda/h},\ p\in\mathbb{S}^n,\ q\in\mathbb{B} \tag{7.11}$$

We thus set $E_h^0(\lambda,\xi;m)$ to be

$$E_h^0(\lambda,\xi;m):=e^{(n/2+i\lambda/h)\phi_\xi(m)}, \tag{7.12}$$

where ϕ_ξ is the Busemann function defined in (7.7). Viewing the neighbourhood U_ξ as a subset of one of the $V_{\xi_j}\simeq_{\psi_{\xi_j}}V_{q_0}$ where $V_{q_0}\subset\mathbb{B}$ is defined in (7.4), the Laplacian in this hyperbolic chart pulls back to $\Delta_{\mathbb{H}^{n+1}}$. Since $\phi_\xi(m)=\phi^{\mathbb{B}}_{\psi_{\xi_j}(\xi)}(\psi_{\xi_j}(m))+c_j(\xi)$ for some function $c_j(\xi)$ independent of m, we directly have in U_ξ (U_ξ is defined in (7.6))

$$(h^2(\Delta-n^2/4)-\lambda^2)E_h^0(\lambda;m,\xi)=0.$$

We let $\chi_0\in C^\infty(\partial\overline{M}\times\overline{M})$ be a function such that $\chi(\xi,\cdot)$ is supported in U_ξ, equal to 1 near ξ and smooth in x^2. Therefore we obtain

$$F_h(\lambda,\xi):=(h^2(\Delta-n^2/4)-\lambda^2)\chi_0E_h^0(\lambda,\xi)=[h^2\Delta,\chi_0]E_h^0(\lambda,\xi) \tag{7.13}$$

and we claim that

$$F_h(\lambda,\xi)\in x^{\frac{n}{2}+2+\frac{i\lambda}{h}}C^\infty(\overline{M})\text{ and }\|x^{-1}F_h(\lambda,\xi)\|_{L^2(M)}=\mathcal{O}(h)$$

uniformly in ξ. Indeed, this is an elementary calculation since from (7.11) we see that $E_h^0(\lambda,\xi)\in x^{\frac{n}{2}+i\frac{\lambda}{h}}C^\infty(\overline{M}\setminus\{\xi\})$ and in geodesic normal coordinates near the boundary

$$[\Delta,\chi_0]=-x^2(\partial_x^2\chi_0)-2x(\partial_x\chi_0)x\partial_x+x^2[\Delta_{h(x)},\chi_0]+n(x\partial_x\chi_0)-\frac{1}{2}\operatorname{Tr}_{h(x)}(\partial_xh(x))x^2(\partial_x\chi_0)$$

is a first order operator with coefficients vanishing in a neighbourhood of ξ. We thus correct the error by the incoming resolvent $R_h(\lambda)$ by setting

$$E_h(\lambda,\xi):=\chi_0E_h^0(\lambda,\xi)+E_h^1(\lambda,\xi),\quad\text{with}\quad E_h^1(\lambda,\xi):=-R_h(\lambda)F_h(\lambda,\xi) \tag{7.14}$$

and this makes sense since for $\lambda\in\mathbb{R}$, $R_h(\lambda):x^\alpha L^2(M)\to x^{-\alpha}L^2(M)$ for any $\alpha>0$ and $F_h\in xL^2(M)$. We claim that

Proposition 7.3. *The function $E_h(\lambda,\xi)$ of (7.14) is the Eisenstein function defined in (7.8) for a certain boundary defining function x.*

Proof. Let $R_h^{\mathbb{B}}(\lambda)$ be the resolvent of the hyperbolic space (that is, the incoming right inverse to $h^2(\Delta_{\mathbb{H}^{n+1}}-n^2/4)-\lambda^2)$ in the ball model and let $\chi_1\in C^\infty(\partial\overline{M}\times\overline{M})$ be such that $\chi_1(\xi,\cdot)$ is supported in U_ξ and $\chi_1\chi_0=\chi_1$. Through the pull-back by ψ_{ξ_j} (for each j), the operator $R_h^{\mathbb{B}}(\lambda)$ induces an operator $R_h^j(\lambda)$ on V_{ξ_j}; if $U_\xi\subset V_{\xi_j}$, then we have the resolvent identity

$$R_h(\lambda)\chi_1=\chi_0R_h^j(\lambda)\chi_1-R_h(\lambda)[h^2\Delta,\chi_0]R_h^j(\lambda)\chi_1 \tag{7.15}$$

for $\lambda\in\mathbb{R}$, the composition makes sense as a map $x^\alpha L^2\to x^{-\alpha}L^2$ for any $\alpha>0$. Let x be a boundary defining function, so in V_{ξ_j}, one has $\psi^*_{\xi_j}x_0=xe^{\omega_j}$ for some function $\omega_j\in$

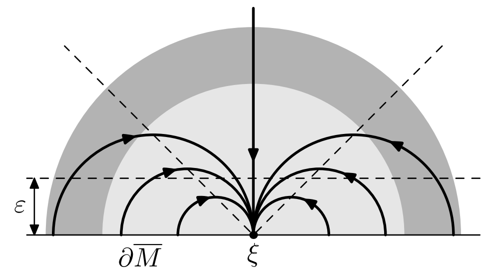

FIGURE 4. Illustration of (A7) for the half-plane model of $\mathbb{H}^{n+1}$: the set of points on trajectories converging to $\xi \in \partial\mathbb{H}^{n+1}$ with $\dot{x} < 0$ and $x < \varepsilon$ is the triangle formed by dashed lines, lying $O(\varepsilon)$ close to ξ. For ε small enough, this triangle lies inside the lighter shaded region, denoting the set $\{\chi_0 = 1\}$.

$C^\infty(\partial\overline{M}\cap V_{\xi_j})$. Then multiplying (7.15) by $x^{-n/2-i\lambda/h}$ on the right, and taking the restriction of the Schwartz kernels on $M \times \partial\overline{M}$, we have

$$E_h(\lambda,\xi) = \chi_0\widetilde{E}_h^0(\lambda,\xi) - R_h(\lambda)[h^2\Delta,\chi_0]\widetilde{E}_h^0(\lambda,\xi)$$

with $\widetilde{E}_h^0(\lambda;m,\xi) = \frac{2i\lambda h}{C(\lambda/h)}\lim_{m'\to\xi}(x(m')^{-\frac{n}{2}-i\frac{\lambda}{h}}R_h^j(\lambda;m,m'))$ a smooth function of $m \in U_\xi$ and $C(\lambda/h)$ the constant in (7.8). Note that the Schwartz kernels of R_h^j and R_h^k are the same on the intersection of their domains, therefore $\widetilde{E}_h^0$ does not depend on the choice of j. Now, since $E_h^{\mathbb{B}}(\lambda;m,\xi)$ in (7.11) is the Eisenstein function on $\mathbb{B}$ for the defining function x_0, we deduce that in $U_\xi \subset V_{\xi_j}$, one has $\widetilde{E}_h^0(\lambda,\xi;m) = E_h^0(\lambda,\xi;m)e^{(\frac{n}{2}+i\frac{\lambda}{h})(\omega_j(\xi)-c_j(\xi))}$. Here $c_j(\xi) = \phi_\xi(m) - \phi^{\mathbb{B}}_{\psi_{\xi_j}(\xi)}(\psi_{\xi_j}(m))$. Since $E_h^0(\lambda;m,\xi)$ does not vanish, this shows that on any intersection $\partial\overline{M}\cap V_{\xi_j}\cap V_{\xi_k}$ of the cover of $\partial\overline{M}$ by the open sets $V_{\xi_j}\cap\partial\overline{M}$, we get $\omega_j(\xi)-c_j(\xi) = \omega_k(\xi)-c_k(\xi)$ and therefore this defines a global smooth function θ on $\partial\overline{M}$. In its definition, $E_h(\lambda,\xi)$ only depends on the first jet of x at $\partial\overline{M}$ and thus modifying x to be xe^θ, this shows the claim. □

It follows that (A3) and (A4) are satisfied, with $b^0 = e^{\frac{n}{2}\phi_\xi(m)}$. Assumption (A8) is then checked by a direct calculation, with the measure $d\xi$ on $\partial\overline{M}$ corresponding to the choice of the function x in Proposition 7.3.

Assumption (A7) can be reduced, using the isometries ψ_{ξ_j}, to the following statement: if $(q,\nu)\in S^*\mathbb{H}^{n+1}$ is directly escaping in the forward direction and converging to some $p\in\mathbb{S}^n$, then $|q-p|\le Cx_0(q)$ for some global constant C; the latter statement is verified directly, see Figure 4.

Microlocalization of E_h^1. Finally, assumptions (A5) and (A6) follow, by rescaling h and using that $E_h(\lambda,\cdot)$ is a function of λ/h, from

Proposition 7.4. *Let $K_0\subset M$ be a compact set containing a neighborhood of the trapped set. Assume that $\lambda = 1$ and define*

$$\widetilde{E}_h^1(\lambda,\xi) = \frac{E_h^1(\lambda,\xi)}{1+\|E_h(\lambda,\xi)\|_{L^2(K_0)}}. \tag{7.16}$$

Then:

1. $\widetilde{E}_h^1(\lambda,\xi)$ *is h-tempered in the sense of* (3.2).

2. The wavefront set $\mathrm{WF}_h(\widetilde{E}_h^1)$ *is contained in* S^*M.

3. If $(m,\nu)\in S^*M$ *and* $g^t(m,\nu)$ *escapes to infinity as* $t\to+\infty$ *and never passes through the set*

$$W_\xi := \{(m,\partial_m\phi_\xi(m)) \mid m\in\operatorname{supp}(\partial_m\chi_0)\}$$

for $t\geq 0$, *then* $(m,\nu)\notin \mathrm{WF}_h(\widetilde{E}_h^1)$.

Moreover, the corresponding estimates are uniform in $\lambda\in[1/2,2]$ *and* $\xi\in\partial M$.

Proof. We will use the construction of [Va11]. (See also [Va10]; note however that in that paper L_+ and L_- switch places compared to the notation of [Va11] that we are using.) Let $\overline{M}_{\mathrm{even}}$ (called $X_{0,\mathrm{even}}$ in [Va11]) be the space $\overline{M}$ with the smooth structure at the boundary $\partial\overline{M}$ changed so that x^2 is the new boundary defining function. As in [Va11, (3.5)], introduce the modified Laplacian

$$P_1(\lambda) := x^{-2}x^{-s}(1+x^2)^{s/4-n/8}(h^2\Delta - s(n-s))(1+x^2)^{n/8-s/4}x^s,\ \ s:=n/2+i\lambda/h.$$

(The conjugation by $(1+x^2)^{s/4-n/8}$ is irrelevant in our case, as $s/4-n/8=i\lambda/(4h)$ is purely imaginary. In [Va11], it is needed to show estimates far away in the physical plane, that is for $\operatorname{Re} s\gg 1$.) Note that we change the sign of λ in the conjugation (in the notation of [Va11], $P_1(\lambda)=P_\sigma$ with $\sigma=-\lambda/h$); therefore, our resolvent will be semiclassically incoming, instead of semiclassically outgoing, for $\lambda>0$. The operator P_1 is smooth up to the boundary of $\overline{M}_{\mathrm{even}}$; as in [Va11, Section 3.5], we embed $\overline{M}_{\mathrm{even}}$ as an open set in a certain compact manifold without boundary X, and extend P_1 to a differential operator in $\Psi^2(X)$. We also consider the semiclassical complex absorbing operator $Q(\lambda)\in\Psi^2(X)$ satisfying the assumptions of [Va11, Section 3.5]; in particular, $Q(\lambda)$ is supported outside of $\overline{M}_{\mathrm{even}}\subset X$. Then $(P_1(\lambda)-iQ(\lambda))^{-1}: C^\infty(X)\to C^\infty(X)$ is a meromorphic family of operators in λ, and for $f\in C^\infty(X)$, we have (see the proof of [Va11, Theorem 5.1])

$$x^s(1+x^2)^{n/8-s/4}(P_1(\lambda)-iQ(\lambda))^{-1}f|_M = R_h(\lambda)(1+x^2)^{n/8-s/4}x^sx^2(f|_M).$$

Here $R_h(\lambda)$ is the incoming scattering resolvent on M. In principle, depending on the choice of $Q(\lambda)$, the operator $(P_1(\lambda)-iQ(\lambda))^{-1}$ could have a pole at λ. However, as $R(\lambda)$ does not have a pole for $\lambda\in[1/2,2]$, the terms in the Laurent expansion of $(P_1(\lambda)-iQ(\lambda))^{-1}$ have to be supported outside of $\overline{M}_{\mathrm{even}}$ and we can ignore them in the analysis.

Let $\widehat{F}_h\in C^\infty(X)$ be any function such that $\widehat{F}_h=O(h)_{H_h^N}$ for all N, and

$$F_h=(1+x^2)^{n/8-s/4}x^{s+2}(\widehat{F}_h|_M).$$

Such a function exists as $x^{-s}\chi_0E_h^0\in C^\infty(\overline{M}_{\mathrm{even}}\setminus\xi)$, $\chi_0\in C^\infty(\overline{M}_{\mathrm{even}})$, and

$$F_h=x^{2+s}(1+x^2)^{n/8-s/4}[P_1(s),\chi_0](1+x^2)^{s/4-n/8}x^{-s}E_h^0$$

is supported away from ξ. Define the function $\widehat{E}_h^1\in C^\infty(X)$ by

$$\widehat{E}_h^1=-\frac{(P_1(\lambda)-iQ(\lambda))^{-1}\widehat{F}_h}{1+\|E_h(\lambda,\xi)\|_{L^2(K_0)}}.$$

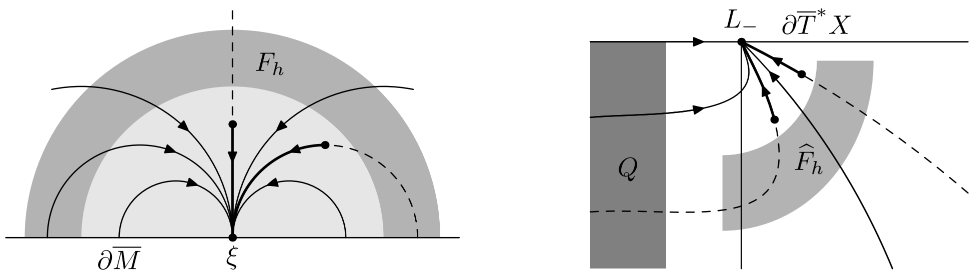

FIGURE 5. Left: physical space picture of geodesics converging to ξ. The darker shaded region is the support of $d\chi_0$, and thus of F_h. In the lighter shaded region, $\chi_0 = 1$. Right: phase space picture near ξ after the conjugation of [Va11]. L_- is the sink consisting of radial points, Q is the complex absorbing operator, and the shaded region corresponds to the wavefront set of $\widehat{F}_h$. The vertical line hitting L_- is the boundary of $\overline{M}_{\mathrm{even}}$, while the horizontal line is the fiber infinity. In both pictures, we mark two points (m,ν) satisfying the assumption of part 3 of Proposition 7.4 and the forward geodesics starting at these points.

Then

$$\widetilde{E}_h^1 = x^s(1+x^2)^{n/8-s/4}\widehat{E}_h^1|_M.$$

Consider the map $\iota : T^*M \to T^*X$ given by

$$\iota(m,\nu) = \bigg(m, \nu - d\Big(\ln x(m) - \frac{1}{4}\ln(1+x(m)^2)\Big)\bigg),\ m\in M,\ \nu\in T_m^*M;$$

then for an h-tempered $u\in C^\infty(X)$,

$$\mathrm{WF}_h(x^s(1+x^2)^{n/8-s/4}u|_M) = \iota^{-1}(\mathrm{WF}_h(u)).$$

Then

$$\mathrm{WF}_h((P_1(\lambda)-iQ(\lambda))\widehat{E}_h^1)\cap T^*M \subset \iota(\mathrm{WF}_h(F_h)) \subset \iota(W_\xi). \tag{7.17}$$

Now, as $\|E_h^0\|_{L^2(K_0)} \leq C$ and thus $\|E_h^1\|_{L^2(K_0)} \leq C + \|E_h\|_{L^2(K_0)}$, we have

$$\|\widehat{E}_h^1\|_{L^2(K_0)} \leq C.$$

Consider an operator $Q_K \in \Psi^{\mathrm{comp}}(X)$ supported in K_0 such that $\sigma(Q_K)\leq 0$ everywhere and each unit speed geodesic $\gamma(t)$ either escapes as $t\to+\infty$ or passes through the region $\{\sigma(Q_K)<0\}$ at some positive time. This is possible since K_0 contains a neighborhood of the trapped set. Then the operator $P_1(\lambda)-iQ(\lambda)-iQ_K$ satisfies the semiclassical nontrapping assumptions [Va11, Section 3.5]; therefore, by the nontrapping estimate [Va11, Theorem 4.8],

$$\begin{aligned}\|\widehat{E}_h^1\|_{L^2(X)} &\leq Ch^{-1}\|(P_1(\lambda)-iQ(\lambda)-iQ_K)\widehat{E}_h^1\|_{L^2(X)}\\ &\leq Ch^{-1}\|\widehat{F}_h\|_{L^2(X)} + Ch^{-1}\|Q_K\widehat{E}_h^1\|_{L^2(X)}.\end{aligned}$$

However, $\|Q_K\widehat{E}_h^1\|$ is bounded by $\|\widehat{E}_h^1\|_{L^2(K_0)}$; therefore, $\|\widehat{E}_h^1\|_{L^2(X)} = O(h^{-1})$ and in particular $\widehat{E}_h^1$ is tempered; it follows that $\widetilde{E}_h^1$ is also tempered. This proves part 1 of the proposition; part 2 follows by ellipticity (note that $\mathrm{WF}_h(F_h)\subset W_\xi\subset S^*M$).

Now, assume that $(m,\nu)\in S^*M$ satisfies the assumption of part 3 of this proposition. Then it follows directly from (7.17), the analysis of [Va11, Section 2.2], and the definition of ι, that the Hamiltonian flow line of $\sigma(P_1)$ starting at $\iota(m,\nu)$ converges to the set L_- of radial points as $t\to+\infty$ and does not intersect $\mathrm{WF}_h((P_1(\lambda)-iQ(\lambda))\widehat{E}^1_h)$ for $t\geq 0$. In a fashion similar to the global argument of [Va11, Section 4.4], we combine elliptic regularity and propagation of singularities (see [Va11, Section 4.1]) with the radial points lemma [Va11, Proposition 4.5] for L_-, to get $\iota(m,\nu)\notin \mathrm{WF}_h(\widehat{E}^1_h)$. Therefore, $(m,\nu)\notin\mathrm{WF}_h(\widetilde{E}^1_h)$ as required. □

Appendix A. Limiting measures for hyperbolic quotients

In this appendix, we give an explicit description of the limiting measures μ_ξ in case when M is a hyperbolic quotient $\Gamma\backslash\mathbb{H}^{n+1}$, in terms of the group Γ. This is a particular case of asymptotically hyperbolic manifolds discussed in Section 7.

A.1. **Convex co-compact groups.** Let $\mathbb{B}$ be the unit ball in $\mathbb{R}^{n+1}$, and $\mathbb{H}^{n+1}$ the $(n+1)$-dimensional hyperbolic space, which we view as $\mathbb{B}$ equipped with the constant negative curvature metric $g_{\mathbb{H}^{n+1}}:=4|dm|^2/(1-|m|^2)^2$. The boundary $\mathbb{S}^n=\partial\overline{\mathbb{B}}$ is the sphere of radius 1, which is also the conformal boundary of $\mathbb{H}^{n+1}$. A *convex co-compact* group Γ of isometries of $\mathbb{H}^{n+1}$ is a discrete group of hyperbolic transformations (i.e., transformations having 2 disjoint fixed points on $\overline{\mathbb{B}}$) with a compact convex core, and Γ is not co-compact. The convex core is the smallest convex subset in $\Gamma\backslash\mathbb{H}^{n+1}$, which can be obtained as follows. The limit set Λ_Γ of the group and the discontinuity set Ω_Γ are defined by

$$\Lambda_\Gamma:=\overline{\{\gamma(m)\in\mathbb{B};\gamma\in\Gamma\}}\cap\mathbb{S}^n\,,\quad \Omega_\Gamma:=\mathbb{S}^n\setminus\Lambda_\Gamma \tag{A.1}$$

where the closure is taken in the closed unit ball $\overline{\mathbb{B}}$ and $m\in\mathbb{B}$ is any point (the set Λ_Γ does not depend on the choice of m). The group Γ acts on the convex hull of Λ_Γ (with respect to hyperbolic geodesics) and the convex core is the quotient space.

An important quantity is the Hausdorff dimension of Λ_Γ

$$\delta:=\dim_H\Lambda_\Gamma<n \tag{A.2}$$

which in turn is, by Patterson [Pa] and Sullivan [Su79], the exponent of convergence of Poincaré series: for any $m\in\mathbb{B}$,

$$\sum_{\gamma\in\Gamma}e^{-sd(m,\gamma m)}<\infty\iff s>\delta; \tag{A.3}$$

we henceforth denote by $d(\cdot,\cdot)$ the distance function of the hyperbolic metric on $\mathbb{B}$. Notice that the series (A.3) is locally uniformly bounded in $m\in\mathbb{B}$.

The group Γ acts properly discontinuously on Ω_Γ as conformal transformations of the sphere and the quotient space $\Gamma\backslash\Omega_\Gamma$ is a smooth compact manifold of dimension n. The quotient

$$M=\Gamma\backslash\mathbb{H}^{n+1}$$

is a smooth non-compact manifold equipped with the hyperbolic metric g induced by $g_{\mathbb{H}^{n+1}}$, and it admits a smooth compactification by setting $\overline{M}=M\cup(\Gamma\backslash\Omega_\Gamma)$, i.e. with $\partial\overline{M}=\Gamma\backslash\Omega_\Gamma$.

Then M is an asymptotically hyperbolic manifold in the sense of Section 7, of constant sectional curvature -1. We shall denote the covering map by

$$\pi : \mathbb{B} \cup \Omega_\Gamma \to \overline{M}.$$

We refer the reader to [Ni] for more details and properties of convex co-compact groups.

A.2. **Limiting measures in this setting.** In constant curvature, it turns out that the limiting measure μ_ξ exists for all ξ (rather than for Lebesgue almost every ξ as in Section 4.3), and can be described as a converging sum over the group. We give an expression below, which is the same as the one obtained in [GuNa] when $\delta < n/2$.

For $\xi \in \mathbb{S}^n$, we let ϕ_ξ be the Busemann function[4] on the unit ball $\mathbb{B}$ defined by

$$\phi_\xi(m) = \log\Big(\frac{1-|m|^2}{|m-\xi|^2}\Big).$$

The map Φ defined by

$$\Phi : \mathbb{B} \times \mathbb{S}^n \to S^*\mathbb{H}^{n+1}, \quad \Phi : (m, \xi) \mapsto (m, \partial_m \phi_\xi(m)) \tag{A.4}$$

gives a diffeomorphism between the unit cotangent bundle $S^*\mathbb{H}^{n+1}$ and $\mathbb{B} \times \mathbb{S}^n$, and satisfies

$$\Phi^* d\mu_L = e^{n\phi_\xi(m)}\,\mathrm{dvol}_{\mathbb{H}^{n+1}}(m) \wedge d\xi, \qquad \text{with} \quad e^{n\phi_\xi(m)} = \Big(\frac{1-|m|^2}{|m-\xi|^2}\Big)^n,$$

if $d\mu_L$ is the Liouville measure (viewed as a volume form on the unit cotangent bundle) and $d\xi$ the canonical measure on $\mathbb{S}^n$. (This is a more general version of (A8) for the considered case.) Any isometry γ of $\mathbb{H}^{n+1}$ acts on both spaces by

$$\gamma.(m, \nu) = (\gamma m, (d\gamma(m)\nu^*)^*), \text{ for } (m, \nu) \in S^*\mathbb{H}^{n+1};$$
$$\gamma.(m, \xi) = (\gamma m, \gamma\xi), \quad \text{for } (m, \xi) \in \mathbb{B} \times \mathbb{S}^n,$$

where * denotes the map identifying $T^*\mathbb{H}^{n+1}$ with $T\mathbb{H}^{n+1}$ through the metric. We have $\Phi(\gamma.(m, \xi)) = \gamma.\Phi(m, \xi)$ and thus Φ descends to a map $\Gamma\backslash(\mathbb{H}^{n+1} \times \mathbb{S}^n) \to S^*(\Gamma\backslash\mathbb{H}^{n+1})$, which we also denote by Φ.

The limiting measure μ_ξ in the considered case is given by

Lemma A.1. *Let $M = \Gamma\backslash\mathbb{H}^{n+1}$ be a quotient of $\mathbb{H}^{n+1}$ by a convex co-compact group Γ of isometries, let $\mathcal{F}$ be a fundamental domain. Then the measure $\mu_{\pi(\xi)}$ of (4.12) exists for all $\xi \in \Omega_\Gamma$ and is described as a converging series by the following expression: if $\xi \in \Omega_\Gamma \cap \overline{\mathcal{F}}$ and $a \in C_0^\infty(S^*M)$, then*

$$\int_M a\, d\mu_{\pi(\xi)} = \int_{\mathcal{F}} \sum_{\gamma \in \Gamma} a(m, d\phi_{\gamma\xi}(m)) e^{n(\phi_{\gamma\xi}(m) + \log|d\gamma(\xi)|)}\,\mathrm{dvol}_{\mathbb{H}^{n+1}}(m)$$

where $\phi_\xi(m)$ is the Busemann function on $\mathbb{B}$ associated to $\xi \in \mathbb{S}^n$ and $|d\gamma(\xi)|$ is the Euclidean norm of $d\gamma(\xi)$.

[4] In Section 7, we used the coordinate $q \in \mathbb{B}$, $p \in \mathbb{S}^n$ for certain charts near infinity of M, and the notation $\phi_p^{\mathbb{B}}(q)$ for the Busemann function on the ball. This was to avoid confusion with the coordinate m, ξ on $M, \partial\overline{M}$. We keep in this appendix the notation $\phi_\xi(m)$ to match the notation of the general setting of the article.

Proof. We can view a as a compactly supported function on the unit cotangent bundle $S^*\mathcal{F}$ over a fundamental domain $\mathcal{F} \subset \mathbb{B}$ and we extend a by 0 in $S^*\mathbb{H}^{n+1}\backslash S^*\mathcal{F}$ (the resulting function might not be smooth, but it does not matter here). The flow g^t on S^*M is obtained by projecting down the geodesic flow $\widetilde{g}^t$ of the cover $S^*\mathbb{H}^{n+1}$. Let $\xi \in \Omega_\Gamma \cap \overline{\mathcal{F}}$, then small neighbourhoods of $\pi(\xi)$ in M are isometric through π to small neighbourhoods of ξ in the unit ball $\mathbb{B}$. By the construction of the decomposition (4.6) for the asymptotically hyperbolic case in Section 7.2, the function $E_h^0(\lambda, \pi(\xi); \pi(m))$ is equal to $e^{(n/2+i\lambda/h)\phi_\xi(m)}$ for m near ξ (ξ being fixed) and thus $|b^0|^2 = e^{n\phi_\xi(m)}$. One has

$$\int_M a(g^{-t}(m, d\phi_\xi(m)))e^{n\phi_\xi(m)}\, \mathrm{dvol}_M(m) = \int_{\mathcal{F}} \widetilde{a}(\widetilde{g}^{-t}\Phi(m,\xi))e^{n\phi_\xi(m)}\, \mathrm{dvol}_{\mathbb{H}^{n+1}}(m)$$

where $\widetilde{a}(m,\nu) := \sum_{\gamma\in\Gamma} a(\gamma.(m,\nu))$ is the lift to $S^*\mathbb{H}^{n+1}$ of the function a on S^*M and $\mathrm{dvol}_{\mathbb{H}^{n+1}}(m)$ is the Riemannian measure on $\mathbb{H}^{n+1}$. Using the map Φ of (A.4), one can define a map $\widetilde{g}^t_\xi : \mathbb{B} \to \mathbb{B}$ by

$$\widetilde{g}^t\Phi(m,\xi) = \Phi(\widetilde{g}^t_\xi(m), \xi),$$

this is a diffeomorphism which preserves the measure $e^{n\phi_\xi(m)}\, \mathrm{dvol}_{\mathbb{H}^{n+1}}$. By [GuNa, Lemma 4], we have $e^{n\phi_\xi(\gamma^{-1}m)} = e^{n\phi_{\gamma\xi}(m)}|d\gamma(\xi)|^n$, but we also have $\gamma.\Phi(m,\xi) = \Phi(\gamma m, \gamma\xi)$. Let U_∞^+ be defined in (G4) and put $U := \{m \mid (m, \pi(\xi)) \in U_\infty^+\}$, then U lies in a small neighborhood of $\pi(\xi)$ in M. We can identify U with a small neighbourhood $\widetilde{U}$ of ξ in $\mathcal{F}$ and we get for $\tilde{\mu}_{\pi(\xi)}$ defined in (4.10),

$$\begin{aligned}
\int_{S^*M} (a\circ g^{-t})\, d\tilde{\mu}_{\pi(\xi)} &= \int_U a(g^{-t}(m, d\phi_\xi(m)))e^{n\phi_\xi(m)}\, \mathrm{dvol}_M(m) \\
&= \int_{\widetilde{U}} \sum_{\gamma\in\Gamma} a(\gamma.\widetilde{g}^{-t}\Phi(m,\xi))e^{n\phi_\xi(m)}\, \mathrm{dvol}_{\mathbb{H}^{n+1}}(m) \\
&= \sum_{\gamma\in\Gamma} \int_{\widetilde{U}} a(\Phi(\gamma\widetilde{g}^{-t}_\xi m, \gamma\xi))e^{n\phi_\xi(m)}\, \mathrm{dvol}_{\mathbb{H}^{n+1}}(m) \\
&= \sum_{\gamma\in\Gamma} \int_{\gamma\widetilde{g}^{-t}_\xi(\widetilde{U})} a(m, d\phi_{\gamma\xi}(m))e^{n\phi_\xi(\gamma^{-1}m)}\, \mathrm{dvol}_{\mathbb{H}^{n+1}}(m).
\end{aligned} \tag{A.5}$$

We now observe that for all $\gamma \in \Gamma$, $\lim_{t\to+\infty} \mathbb{1}_{\gamma\widetilde{g}^{-t}_\xi\widetilde{U}} = 1$, since $\widetilde{U}$ is a neighbourhood of ξ in $\mathbb{B}$ containing all points directly escaping to ξ. This achieves the proof by recalling the definition (4.12) of $\mu_{\pi(\xi)}$ and taking the limit in (A.5) and using the dominated convergence theorem, as there exists $C, C' > 0$ such that for all m in the compact set $\mathrm{supp}(a)$

$$\sum_{\gamma\in\Gamma} e^{n\phi_\xi(\gamma^{-1}m)} = \sum_{\gamma\in\Gamma}\Big(\frac{1-|\gamma^{-1}m|}{|\gamma^{-1}m-\xi|^2}\Big)^n \le C \sup_{m\in\mathrm{supp}(a)} e^{-n\,d(\gamma^{-1}m,0)} \le C'$$

by locally uniform (in m) convergence of Poincaré series (A.3) at $s = n$. □

Appendix B. The escape rate

Let us discuss the classical escape rate in some particular cases, following the work of Bowen-Ruelle [BoRu], Young [Yo], and Kifer [Ki].

B.1. **Escape rate and the pressure of the unstable Jacobian.** We consider (M, g) a complete non-compact Riemannian manifold which has a compact set $K_0 \subset S^*M$ which is geodesically convex, that is any geodesic trajectory in S^*M which leaves K_0 never comes back:

$$\exists t_0 < t_1,\ g^{t_0}(m,\nu) \in K_0 \text{ and } g^{t_1}(m,\nu) \in M \setminus K_0 \Longrightarrow \forall t \geq t_1,\ g^t(m,\nu) \in M \setminus K_0.$$

We assume that K_0 contains a neighborhood of the trapped set K. The examples we consider are (M, g) which are hyperbolic or Euclidean near infinity, and $K_0 = S^*M \cap \{x \geq \varepsilon_0\}$ with x, ε_0 given in (G2). The trapped set from Definition 4.1 can be written as

$$K = \bigcap_{t\in\mathbb{R}} g^t(K_0) = \bigcap_{j\in\mathbb{Z}} g^j(K_0)$$

This is a compact maximal invariant set for the flow g^t. We define the escape rate as in [Yo, Ki] by

$$Q := \limsup_{t\to\infty} \frac{1}{t} \log \mu_L(\mathcal{T}(t)),$$

with μ_L the Liouville measure and $\mathcal{T}(t)$ defined in (1.9). Note that, since K_0 is geodesically convex, we have $\mathcal{T}(t_2) \subset \mathcal{T}(t_1)$ for $0 \leq t_1 \leq t_2$. The escape rate is clearly non-positive.

In this section, we assume that $\mu_L(K) = 0$ and write Q in terms of the topological pressure of the flow, under certain dynamical assumptions. More precisely, we assume that the trapped set K is *uniformly partially hyperbolic*, in the following sense: there exists $\varepsilon_f > 0$ and a splitting of $T(S^*M)$ over K into continuous subbundles invariant under the flow

$$T_zS^*M = E_z^{cs} \oplus E_z^u, \quad \forall z \in K$$

such that the dimensions of E^u and E^{cs} are constant on K and for all $\varepsilon > 0$, there is $t_0 \in \mathbb{R}$ such that

$$\forall z \in K,\ \forall t \geq t_0, \begin{cases} \forall v \in E_z^u,\ |dg_z^t v| \geq e^{\varepsilon_f t}|v|, \\ \forall v \in E_z^{cs},\ |dg_z^t v| \leq e^{\varepsilon t}|v|. \end{cases}$$

Let J^u be the unstable Jacobian of the flow, defined by

$$J^u(z) := -\partial_t(\det dg_z^t|_{E_z^u})|_{t=0}$$

where $dg^t : E_z^u \to E_{g_z^t}^u$ and the determinant is defined using the Sasaki metric for choosing orthornormal bases in E^u. If μ is a g^t-invariant measure on K, one has

$$\int_K J^u d\mu = -\int_K \sum_j \Lambda_j^+ d\mu$$

where $\Lambda_j^+(z)$ are the positive Lyapunov exponents at a regular point $z \in K$ counted with multiplicity (regular points are points where the exponents are well defined, and this is set of full μ-measure by the Oseledec theorem). It is also direct to see that $\int_K J^u d\mu = -\int_K \log\det(dg^1|_{E^u})d\mu$.

The *topological pressure* of a continuous function $\varphi : K \to \mathbb{R}$ with respect to the flow can be defined by the variational formula

$$P(\varphi) := \sup_{\mu\in\mathcal{M}(K)} \Big(h_\mu(g^1) + \int_K \varphi\, d\mu \Big) \tag{B.1}$$

where $\mathcal{M}(K)$ is the set of g^t-invariant Borel probability measures and $h_\mu(g^1)$ is the measure theoretic entropy of the flow at time 1 with respect to μ. In particular $P(0)$ is the topological entropy of the flow.

A particular case of uniformly partially hyperbolic dynamics is when K is *uniformly hyperbolic*, that is when there is a continuous splitting $E^{cs} = \mathbb{R}H_p \oplus E^s$ into flow direction (H_p is the vector field generating the geodesic flow) and stable directions E^s where for $t \geq t_0$

$$\forall v \in E^s_z,\ |dg^t_z v| \leq e^{-\varepsilon_f t}|v|.$$

The set K is called a basic hyperbolic set and the flow is said to be *Axiom A* when the periodic orbits of g^t on K are dense in K and $g^t|_K$ is topologically transitive.

It is proved by Young [Yo, Theorem 4] that if K is uniformly partially hyperbolic, then

$$Q = \lim_{t\to\infty} \frac{1}{t} \log \mu_L(\mathcal{T}(t)) = P(J^u). \tag{B.2}$$

In the Axiom A case, the same formula was essentially contained in the work of Bowen-Ruelle (using the volume lemma [BoRu, Lemma 4.2 and 4.3]). Moreover by [BoRu, Theorem 5], if the incoming tail Γ_- (which is the union of stable manifolds over the trapped set) has Liouville measure 0, then $P(J^u) < 0$. Thus we deduce by (5.2)

$$\mu_L(K) = 0 \text{ and } g^t \text{ is Axiom A } \Longrightarrow P(J^u) < 0.$$

Young [Yo, Theorem 4] gives a lower bound $Q \geq P(-\sum_j \Lambda_j^+)$ which applies without any assumption on K (but we are more interested in an upper bound).

B.2. **Relation with fractal dimensions in particular cases.** Assume first that the metric has constant curvature -1 in a small neighbourhood of the trapped set K (this includes the case of convex co-compact hyperbolic quotients studied in Appendix A). Then the geodesic flow on S^*M is uniformly hyperbolic on K and has Lyapunov exponents 0 (with multiplicity 1) and ± 1 (each with multiplicity n). Therefore, the maximal expansion rate $\Lambda_{\max}$ from (1.11) is equal to 1, one has $J^u(z) = -n$ for all $z \in K$, and (see for example [Fa, Theorem 4])

$$P(J^u) = h_{\rm top}(K) - n = (\dim_H(K) - 1)/2 - n \tag{B.3}$$

where $h_{\rm top}$ is the topological entropy of the flow on K, and $\dim_H(K) \in (0, n)$ is the Hausdorff dimension of K (which is equal to the Minkowski box dimension in this case). For convex co-compact hyperbolic quotients $\Gamma\backslash\mathbb{H}^{n+1}$ (see Section A for definition), one has by Sullivan [Su84]

$$\delta := \dim_H(\Lambda_\Gamma) = h_{\rm top}(K) \tag{B.4}$$

where Λ_Γ is the limit set of the group Γ defined in (A.1).

If g has negative pinched curvature near the trapped set, then one still has upper and lower bounds on $P(J^u)$ in terms of $h_{\rm top}(K)$ and the pinching constant. If the trapped set K is uniformly hyperbolic, it is also shown in [Fa] that $\dim_H(K) \leq 1 + 2h_{\rm top}(K)/\Lambda_{\max}$. In dimension 2 there is an explicit relation between the Hausdorff dimension $\dim_H(K)$ and entropies of certain measures for Axiom A cases: if

$$a^u(z) = \lim_{t\to 0} \frac{1}{t} \log \|dg^t|_{E^u}\| > 0, \quad a^s(z) = \lim_{t\to 0} \frac{1}{t} \log \|Dg^t|_{E^u}\| < 0$$

then Pesin–Sadovskaya [PeSa] show the following formula

$$\dim_H(K) = 1 + t^u + t^s, \qquad \text{with } P(-t^u a^u) = P(-t^s a^s) = 0.$$

Appendix C. Egorov's theorem until Ehrenfest time

In this section, we prove Proposition 3.9, following the methods of [BoRo], [AnNo, Section 5.2], and [Zw, Theorem 11.12]. Without lack of generality, we assume that $t_0 > 0$.

C.1. **Estimating higher derivatives of the flow.** First of all, we need to estimate the derivatives of symbols under propagation for long times. Consider the open set

$$U_1 = \{(m,\nu) \in T^*M \mid m \in U,\ 1 - 2\varepsilon_e < |\nu|_g < 1 + 2\varepsilon_e\}.$$

For each k, we fix a norm $\|\cdot\|_{C^k(\overline{U_1})}$ for the space $C^k(\overline{U_1})$ of k times differentiable functions on $\overline{U_1}$. (The particular choice of the norm does not matter, as long as it does not depend on t.) The following estimate is an analogue of [AnNo, (5.6)]; we include the proof for the case of manifolds for the reader's convenience.

Lemma C.1. *Take $\Lambda_1 > (1+2\varepsilon_e)\Lambda_{\max}$. Then for each k, there exists a constant $C(k)$ such that for each $a \in C_0^\infty(\overline{U_1})$ and each $t \in \mathbb{R}$,*

$$\|a \circ g^t\|_{C^k(\overline{U_1})} \leq C(k) e^{k\Lambda_1 |t|} \|a\|_{C^k(\overline{U_1})}. \tag{C.1}$$

Proof. Without loss of generality, we assume that $t > 0$. We first recall the formula for derivatives of the composition $b \circ \psi$ of a function $b \in C^\infty(\mathbb{R}^d)$ with a map $\psi : \mathbb{R}^d \to \mathbb{R}^d$:

$$\partial^\alpha (b \circ \psi) = \sum_{\alpha, j} c_{\alpha,j} (\partial_{j_1\dots j_m} b) \circ \psi \cdot \prod_{l=1}^m \partial^{\alpha_l} \psi_{j_l}, \tag{C.2}$$

where $c_{\alpha,j}$ are constants, $j_1,\dots,j_m \in \{1,\dots,d\}$, and $\alpha_1,\dots,\alpha_m$ are nonzero multiindices whose sum equals α. We see from (C.2) that (C.1) is implied by the following estimate on the derivatives of the flow g^t (required to hold in any coordinate system):

$$|\alpha| \leq k \Longrightarrow \sup_{U_1 \cap g^{-t}(U_1)} |\partial^\alpha g^t| \leq C_\alpha e^{|\alpha|\Lambda_1 t}. \tag{C.3}$$

The converse is also true, which can be seen by substituting cooordinate functions in place of a in (C.1).

To estimate higher derivatives of the flow, we will need several definitions from differential geometry. For a vector field X on $\overline{U_1}$, define its pushforward $g^t_* X$ by

$$X(a \circ g^t) = ((g^t_* X)a) \circ g^t,\ a \in C^\infty(g^t(\overline{U_1})).$$

Then $g^t_* X$ is a vector field on $g^t(\overline{U_1})$. In local coordinates, we have

$$(g^t_* X)^j = \sum_l (X^l \partial_l g^t_j) \circ g^{-t}.$$

Note that since $g^t = \exp(tH_p/2)$, where H_p is the Hamiltonian vector field of p, and since $g^t_* H_p = H_p$, we have

$$\partial_t g^t_* X = -\frac{1}{2}[H_p, g^t_* X] = -\frac{1}{2} g^t_* [H_p, X]. \tag{C.4}$$

We fix a symmetric affine connection ∇ on T^*M. For vector fields X and Y, consider the differential operator ∇^2_{XY}, acting on functions or on vector fields, defined as follows: for a function f and a vector field Z,

$$\nabla^2_{XY}f = XYf - (\nabla_X Y)f,\ \nabla^2_{XY}Z = \nabla_X\nabla_Y Z - \nabla_{\nabla_X Y}Z. \tag{C.5}$$

In local coordinates, we have (using Einstein's summation convention)

$$\begin{aligned}\nabla^2_{XY}f &= X^iY^j(\partial^2_{ij}f - \Gamma^l_{ij}\partial_l f),\\ (\nabla^2_{XY}Z)^m &= X^iY^j(\partial^2_{ij}Z^m + \Gamma^m_{j\alpha}\partial_i Z^\alpha + \Gamma^m_{i\alpha}\partial_j Z^\alpha - \Gamma^\alpha_{ij}\partial_\alpha Z^m\\ &\quad +(\partial_i\Gamma^m_{j\alpha} + \Gamma^m_{i\beta}\Gamma^\beta_{j\alpha} - \Gamma^\beta_{ij}\Gamma^m_{\alpha\beta})Z^\alpha).\end{aligned}$$

Here Γ^l_{ij} are the Christoffel symbols of the connection ∇. The advantage of ∇^2_{XY} over XY is that the coefficients of this differential operator at any point depend (bilinearly) only on the values of X and Y at this point, but not on their derivatives.

We now return to the proof of (C.1). The estimate (C.3) for $k=1$ follows directly from the definition (3.17) of $\Lambda_{\max}$. It is then enough to assume that (C.3) holds for some $k\geq 1$ and prove the estimate (C.1) for $k+1$. It suffices to show that for any two vector fields X, Y on T^*M and any $a\in C_0^\infty(U_1)$, we have the estimate

$$\|XY(a\circ g^t)\|_{C^{k-1}(\overline{U_1})} \leq Ce^{(k+1)\Lambda_1 t}\|a\|_{C^{k+1}(\overline{U_1})}. \tag{C.6}$$

The left-hand side of (C.6) is equal to $\|(g^t_*Xg^t_*Ya)\circ g^t\|_{C^{k-1}(\overline{U}_1)}$. We first claim that

$$\|(\nabla^2_{g^t_*Xg^t_*Y}a)\circ g^t\|_{C^{k-1}(\overline{U_1})} \leq Ce^{(k+1)\Lambda_1 t}\|a\|_{C^{k+1}(\overline{U_1})}. \tag{C.7}$$

Indeed, in local coordinates

$$(\nabla^2_{g^t_*Xg^t_*Y}a)\circ g^t = (X^\alpha\partial_\alpha g^t_i)(Y^\beta\partial_\beta g^t_j)\big((\partial^2_{ij}a - \Gamma^l_{ij}\partial_l a)\circ g^t\big). \tag{C.8}$$

We can now apply (C.2) to get an expression for any derivative of order no more than $k-1$ of (C.8). The result will involve derivatives of orders $1,\dots,k$ of g^t, but not its $k+1$'st derivative; therefore, we can apply (C.3) to get (C.7).

Given (C.7) and (C.5), it is enough to show

$$\|((\nabla_{g^t_*X}g^t_*Y)a)\circ g^t\|_{C^{k-1}(\overline{U_1})} \leq Ce^{(k+1)\Lambda_1 t}\|a\|_{C^k(\overline{U_1})}. \tag{C.9}$$

The vector field $\nabla_{g^t_*X}g^t_*Y$ involves the second derivatives of g^t, therefore the left-hand side of (C.9) depends on the $k+1$'st derivatives of g^t and we cannot apply (C.3) directly. We will instead use the method of the proof of [BoRo, Lemma 2.2], computing by (C.4)

$$\begin{aligned}\partial_t(g^{-t}_*(\nabla_{g^t_*X}g^t_*Y)) &= \frac{1}{2}g^{-t}_*([H_p,\nabla_{g^t_*X}g^t_*Y]\\ &-\nabla_{[H_p,g^t_*X]}g^t_*Y - \nabla_{g^t_*X}[H_p,g^t_*Y]) = \frac{1}{2}g^{-t}_*Z_t,\end{aligned}$$

where Z_t is the vector field given by

$$Z_t = \nabla^2_{g^t_*Xg^t_*Y}H_p + R_\nabla(H_p,g^t_*X)(g^t_*Y).$$

Here R_∇ is the curvature tensor of the connection ∇. Then

$$\nabla_{g_*^t X} g_*^t Y = g_*^t(\nabla_X Y) + \frac{1}{2}\int_0^t g_*^{t-s} Z_s\,ds. \tag{C.10}$$

We have

$$\begin{aligned}\|(g_*^t(\nabla_X Y)a)\circ g^t\|_{C^{k-1}(\overline{U_1})} &= \|\nabla_X Y(a\circ g^t)\|_{C^{k-1}(\overline{U_1})}\\ &\leq C\|a\circ g^t\|_{C^k(\overline{U_1})} \leq Ce^{k\Lambda_1 t}\|a\|_{C^k(\overline{U_1})}.\end{aligned}$$

It is then enough to handle the integral part of (C.10). The field Z_s depends quadratically on the first derivatives of g^s, but does not depend on its higher derivatives; therefore, writing an expression for Z_s in local coordinates similar to (C.8), we get for $a \in C_0^\infty(U_1)$,

$$\|(Z_s a)\circ g^s\|_{C^{k-1}(\overline{U_1})} \leq Ce^{(k+1)\Lambda_1 s}\|a\|_{C^k(\overline{U_1})}.$$

Applying (C.1) for the C^k norm (given by the induction hypothesis) and using the geodesic convexity of U, we get

$$\begin{aligned}\int_0^t \|((g_*^{t-s}Z_s)a)\circ g^t\|_{C^{k-1}(\overline{U_1})}\,ds &= \int_0^t \|(Z_s(a\circ g^{t-s}))\circ g^s\|_{C^{k-1}(\overline{U_1})}\,ds\\ \leq C\int_0^t e^{(k+1)\Lambda_1 s}\|a\circ g^{t-s}\|_{C^k(\overline{U_1})}\,ds &\leq C\int_0^t e^{(k+1)\Lambda_1 s}e^{k\Lambda_1(t-s)}\|a\|_{C^k(\overline{U_1})}\,ds\\ &\leq Ce^{(k+1)\Lambda_1 t}\|a\|_{C^k(\overline{U_1})}\end{aligned}$$

and the proof is finished. □

C.2. **Proof of Proposition 3.9.** The proof of Proposition 3.9 is based on repeatedly applying the following corollary of Lemma C.1. The functions $b^{(j)}$ below will be the remainders in the formula for the commutator $[h^2\Delta, A^{(j)}(t)]$, while the functions $c^{(j)}$ will be the errors arising from multiplying our operators by X_1 and X_2.

Proposition C.2. *Take $\Lambda_1 > (1+2\varepsilon_e)\Lambda_{\max}$. Fix $t_0 > 0$ and let $\varphi \in C_0^\infty(U_1)$ satisfy $|\varphi| \leq 1$. Assume that $a_0 \in C^\infty(T^*M)$ and for each $j \geq 0$, $b^{(j)}(t) \in C^\infty([0,t_0]\times T^*M)$, and $c^{(j)} \in C^\infty(T^*M)$, with support contained in some j-independent compact set. For $j \geq 0$, define $a^{(j)} \in C^\infty([0,t_0]\times T^*M)$ inductively as the solutions to the equations*

$$\begin{gathered}a^{(0)}(0) = a_0,\ \ a^{(j+1)}(0) = \varphi\cdot a^{(j)}(t_0) + c^{(j+1)};\\ \partial_t a^{(j)}(t) = \frac{1}{2}H_p a^{(j)}(t) + b^{(j)}(t).\end{gathered}$$

Then for each k, and each j, we have (bearing in mind that each $a^{(j)}$ is supported inside some j-independent compact set and thus its C^k norm is well-defined up to a constant)

$$\begin{aligned}\sup_{t\in[0,t_0]}\|a^{(j)}(t)\|_{C^k(T^*M)} &\leq C(k)\big(e^{jk\Lambda_1 t_0}\|a_0\|_{C^k}\\ &+\max_{0\leq i\leq j} e^{(j-i)k\Lambda_1 t_0}\big(\sup_{t\in[0,t_0]}\|b^{(i)}(t)\|_{C^k} + \|c^{(i)}\|_{C^k}\big)\big),\end{aligned}$$

where $C(k)$ is a constant independent of j.

Proof. We can write

$$a^{(j)}(t)=a^{(j)}(0)\circ g^t+\int_0^t b^{(j)}(s)\circ g^{t-s}\,ds.$$

Since t_0 is fixed, it is enough to estimate the derivatives of $a^{(j)}(0)$. Define

$$\varphi^{(j)}=\prod_{0\leq m<j}(\varphi\circ g^{mt_0});$$

applying the Leibniz rule to $\varphi^{(j)}$, estimating each nontrivial derivative of $\varphi\circ g^{mt_0}$ by Lemma C.1, using that $|\varphi|\leq 1$ and absorbing the (polynomial in l) number of different terms in the Leibniz formula into the exponential by increasing Λ_1 slightly, we get $\|\varphi^{(j)}\|_{C^k}=\mathcal{O}(e^{jk\Lambda_1 t_0})$. Now,

$$\begin{aligned}a^{(j)}(0)=\varphi^{(j)}\cdot(a_0\circ g^{jt_0})&+\sum_{i=0}^{j-1}\varphi^{(j-i)}\int_0^{t_0} b^{(i)}(s)\circ g^{(j-i)t_0-s}\,ds\\&+\sum_{i=1}^{j}\varphi^{(j-i)}\cdot(c^{(i)}\circ g^{(j-i)t_0}).\end{aligned}$$

Here we put $\varphi^{(0)}=1$. We can now apply Lemma C.1 again to get the required estimate. □

We are now ready to prove Proposition 3.9. Fix a quantization procedure Op_h on M; our symbols will be supported in a certain compact set (in fact, no more than distance t_0 to the set U) and we require that the corresponding operators be compactly supported. Put $\Lambda_1=\Lambda_0'$.

Let l satisfy (3.18). We will construct the operators

$$A_m^{(j)}(t)=\mathrm{Op}_h\Big(\sum_{0\leq m'\leq m}a_{m'}^{(j)}(t)\Big),\ 0\leq t\leq t_0,\ 0\leq j\leq l,\ m\geq 0,$$

Here the symbols $a_m^{(j)}$ will be supported in a fixed compact subset of T^*M and satisfy the derivative bounds

$$\sup_{t\in[0,t_0]}\|a_m^{(j)}(t)\|_{C^k}\leq C(k,m)h^{(1-2\rho_j)m-\rho_j k}. \tag{C.11}$$

with the constants $C(k,m)$ independent on j and ρ_j defined by (3.19). The operators $A_m^{(j)}(t)$ will satisfy the relations

$$\begin{aligned}A_m^{(0)}(0)&=A+\mathcal{O}(h^\infty)_{\Psi^{-\infty}},\\A_m^{(j+1)}(0)&=X_2A_m^{(j)}(t_0)X_1+\mathrm{Op}_h(c_m^{(j)})+\mathcal{O}(h^\infty)_{\Psi^{-\infty}},\\hD_tA_m^{(j)}(t)&=\frac{1}{2}[h^2\Delta,A_m^{(j)}(t)]+\frac{h}{i}\mathrm{Op}_h(b_m^{(j)}(t))+\mathcal{O}(h^\infty)_{\Psi^{-\infty}},\end{aligned} \tag{C.12}$$

where the symbols $b_m^{(j)}(t)$ and $c_m^{(j)}$ are supported in some fixed compact set and satisfy bounds

$$\sup_{t\in[0,t_0]}\|b_m^{(j)}(t)\|_{C^k},\|c_m^{(j)}\|_{C^k}\leq C(k,m)h^{(1-2\rho_j)(m+1)-\rho_j k}, \tag{C.13}$$

with the constants $C(k,m)$ again independent on j.

We construct the symbols $a_m^{(j)}$ iteratively, by requiring that they solve the equations

$$a_m^{(0)}(0) = \delta_{m0} \cdot a_0,\ a_m^{(j+1)}(0) = \varphi a_m^{(j)}(t_0) - c_{m-1}^{(j)},$$
$$\partial_t a_m^{(j)}(t) = \frac{1}{2} H_p a_m^{(j)}(t) - b_{m-1}^{(j)}(t).$$

Here $A = \mathrm{Op}_h(a_0) + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}$ and we put $b_{-1}^{(j)} = c_{-1}^{(j)} = 0$. The function $\varphi \in C_0^\infty(U_1)$ is equal to $\sigma(X_1)\sigma(X_2)\psi(|\nu|)$, where $\psi \in C_0^\infty(1-2\varepsilon_e, 1+2\varepsilon_e)$ is such that $\psi(|\nu|) = 1$ near $\mathrm{WF}_h(A)$. We use the fact that the function $|\nu|$ is invariant under the geodesic flow. The estimate (C.11) follows immediately from (C.13) and Proposition C.2. As for the equations (C.12) and the bounds (C.13), they follow from (C.11) and the following commutator formula:

$$[h^2\Delta, \mathrm{Op}_h(a)] = \frac{h}{i} \mathrm{Op}_h(H_p a) + \mathrm{Op}_h(b) + \mathcal{O}(h^\infty)_{\Psi^{-\infty}},\ b = \mathcal{O}(h^{2-2\rho}\|a\|_{S_\rho})_{S_\rho},$$

true for any $\rho < 1/2$ and any $a \in S_\rho^{\mathrm{comp}}$.

Now, consider the asymptotic sums

$$a^{(j)}(t) \sim \sum_{m\geq 0} a_m^{(j)}(t)$$

and define the operators $A^{(j)}(t) = \mathrm{Op}_h(a^{(j)}(t))$. By (C.12), these operators satisfy

$$A^{(0)}(0) = A + \mathcal{O}(h^\infty)_{\Psi^{-\infty}},\ A^{(j+1)}(0) = X_2 A^{(j)}(t_0) X_1 + \mathcal{O}(h^\infty)_{\Psi^{-\infty}},$$
$$hD_t A^{(j)}(t) = \frac{1}{2}[h^2\Delta, A^{(j)}(t)] + \mathcal{O}(h^\infty)_{\Psi^{-\infty}}.$$

We then have

$$(X_2 U(t_0))^l A (U(-t_0) X_1)^l = A^{(l)}(0) + \mathcal{O}(h^\infty)_{L^2\to L^2}.$$

It remains to recall that $a^{(l)}(0) \in S_{\rho_l}^{\mathrm{comp}}$ uniformly in l. The principal symbol and microlocal vanishing statements follow directly from the procedure we used to construct the symbols $a_m^{(j)}$.

Appendix D. Proof of quantum ergodicity in the semiclassical setting

In this section, we illustrate how our methods yield a proof of the following integrated quantum ergodicity statement in the semiclassical setting:

Theorem 5. *Let (M, g) be a compact Riemannian manifold of dimension d and assume that the geodesic flow g^t on M is ergodic with respect to the Liouville measure μ_L on the unit cotangent bundle S^*M. For each $h > 0$, let $(e_j)_{j\in\mathbb{N}}$ be an orthonormal basis of eigenfunctions of $h^2\Delta$ with eigenvalues λ_j^2. Then for each semiclassical pseudodifferential operator $A \in \Psi^0(M)$, we have*

$$h^{d-1} \sum_{\lambda_j\in[1,1+h]} \left| \langle Ae_j, e_j\rangle_{L^2(M)} - \frac{1}{\mu_L(S^*M)} \int_{S^*M} \sigma(A)\, d\mu_L \right| \to 0 \text{ as } h \to 0. \tag{D.1}$$

A more general version of Theorem 5 was proved in [HeMaRo], in particular relying on the result of [DuGu, PeRo] on $o(h)$ remainders for the Weyl law when the closed geodesics form a set of measure zero. The purpose of this Appendix is to provide a shorter proof. Theorem 5 is formulated here for the semiclassical Laplacian for simplicity of notation, but it applies to any self-adjoint semiclassical pseudodifferential operator $P(h)$ with compact resolvent on a compact manifold, if the Hamiltonian flow of the principal symbol p of $P(h)$ has no fixed points and is ergodic on the energy surface $p^{-1}(0)$ and we take eigenvalues in the interval $[0,h]$.

The key component of our proof is the following estimate:

Lemma D.1. *Let M be as in Theorem 5. Then for each $A\in\Psi^0(M)$, we have*

$$h^{d-1}\sum_{\lambda_j\in[1,1+h]}\|Ae_j\|^2_{L^2(M)}\leq (C\|\sigma(A)\|_{L^2(S^*M)}+\mathcal{O}(h))^2. \tag{D.2}$$

Here $\|\sigma(A)\|_{L^2(S^*M)}$ is the L^2 norm of the restriction of $\sigma(A)$ to S^*M with respect to the Liouville measure. The constant in $\mathcal{O}(h)$ depends on A, but the constant C does not.

Proof. Assume first that A is compactly microlocalized. We can rewrite the left-hand side of (D.2) as the square of the Hilbert–Schmidt norm of $h^{(d-1)/2}A\Pi_{[1,1+h]}$, where $\Pi_{[1,1+h]}=\mathbb{1}_{[1,(1+h)^2]}(h^2\Delta)$ is a spectral projector. It can then be estimated using the local theory of semiclassical Fourier integral operators, by (3.24) (applied to the adjoint of the operator in interest).

To handle the case of a general A, it remains to note that if $\mathrm{WF}_h(A)\cap S^*M=\emptyset$, then the left-hand side of (D.2) is $\mathcal{O}(h^\infty)$, as each Ae_j is $\mathcal{O}(h^\infty)$ by the elliptic estimate (Proposition 3.2; see also the proof of Proposition 4.5). ☐

Putting A equal to the identity in (D.2), we get the following upper Weyl bound:

$$\#\{j\mid \lambda_j\in[1,1+h]\}\leq Ch^{1-d}. \tag{D.3}$$

We can now prove Theorem 5. Take $A\in\Psi^0(M)$; by subtracting a multiple of the identity operator and applying the ellipticity estimate, we may assume that A is compactly microlocalized and

$$\int_{S^*M}\sigma(A)\,d\mu_L=0. \tag{D.4}$$

Define the quantum average

$$\langle A\rangle_T=\frac{1}{T}\int_0^T U(t)AU(-t)\,dt.$$

Here $U(t)=e^{ith\Delta/2}$ is the semiclassical Schrödinger propagator. By Egorov's theorem (Proposition 3.8), for any fixed T the operator $\langle A\rangle_T$ lies in Ψ^0, modulo an $\mathcal{O}(h^\infty)_{L^2\to L^2}$ remainder, and its principal symbol is

$$\sigma(\langle A\rangle_T)=\langle\sigma(A)\rangle_T:=\frac{1}{T}\int_0^T\sigma(A)\circ g^t\,dt.$$

Note that for each j, we have $U(t)e_j = e^{it\lambda_j/(2h)}$ and thus $\langle\langle A\rangle_T e_j, e_j\rangle = \langle Ae_j, e_j\rangle$. Using Cauchy–Schwarz inequality in j and the bounds (D.2) and (D.3), we get

$$h^{d-1}\sum_{\lambda_j\in[1,1+h]}|\langle Ae_j,e_j\rangle| = h^{d-1}\sum_{\lambda_j\in[1,1+h]}|\langle\langle A\rangle_T e_j,e_j\rangle|$$

$$\leq h^{d-1}\sum_{\lambda_j\in[1,1+h]}\|\langle A\rangle_T e_j\|_{L^2} \leq C\Big(h^{d-1}\sum_{\lambda_j\in[1,1+h]}\|\langle A\rangle_T e_j\|_{L^2}^2\Big)^{1/2}$$

$$\leq C\|\langle\sigma(A)\rangle_T\|_{L^2(S^*M)} + \mathcal{O}_T(h).$$

However, by (D.4) and the von Neumann ergodic theorem [Zw, Theorem 15.1], we have $\|\langle\sigma(A)\rangle_T\|_{L^2(S^*M)} \to 0$ as $T\to\infty$. Therefore, for each $\varepsilon > 0$ we can choose T large enough so that the left-hand side of (D.1) is bounded by $\varepsilon/2 + \mathcal{O}(h)$. Then for h small enough, it is bounded by ε; since the latter was chosen arbitrarily small, we get (D.1).

Acknowledgements. We would like to thank Viviane Baladi, Dima Jakobson, Frederic Naud, Stéphane Nonnenmacher, Steve Zelditch, and Maciej Zworski for useful discussions and providing references on the subject. We additionally thank Stéphane Nonnenmacher for explaining some estimates on higher derivatives of the flow (forming the basis for Lemma C.1). S.D. would also like to thank the DMA of Ecole Normale Supérieure where part of this work was done. S.D. was partially supported by NSF grant DMS-0654436. C.G. is supported by ANR grant ANR-09-JCJC-0099-01.

References

[AnNo] N. Anantharaman and S. Nonnenmacher, *Half-delocalization of eigenfunctions for the Laplacian on an Anosov manifold,* Ann. Inst. Fourier **57**(2007), no. 7, 2465-2523.

[BoRo] A. Bouzouina and D. Robert, *Uniform semiclassical estimates for the propagation of quantum observables,* Duke Math. J. **111**(2002), no. 2, 223–252.

[BoRu] R. Bowen and D. Ruelle, *The ergodic theory of Axiom A flows,* Invent. Math. **29**(1975), no. 3, 181–202.

[Bu02] N. Burq, *Semi-classical estimates for the resolvent in nontrapping geometries,* Int. Math. Res. Not. 2002, no. 5, 221–241.

[Bu05] N. Burq, *Quantum ergodicity of boundary values of eigenfunctions: a control theory approach,* Canad. Math. Bull. **48**(2005), no. 1, 3–15.

[Ch] P.R. Chernoff, *Essential self-adjointness of powers of generators of hyperbolic equations,* J. Funct. Anal. **12**(1973), 401–414.

[CdV] Y. Colin de Verdière, *Ergodicité et fonctions propres du Laplacien,* Comm. Math. Phys. **102**(1985), no. 3, 497–502.

[DiSj] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit,* Cambridge University Press, 1999.

[DuGu] H. Duistermaat and V. Guillemin, *The spectrum of positive elliptic operators and periodic bicharacteristics,* Invent. Math. **29**(1975), no. 1, 39–79.

[Dy1] S. Dyatlov, *Asymptotic distribution of quasi-normal modes for Kerr–de Sitter black holes,* to appear in Ann. Henri Poincaré, arXiv:1101.1260.

[Dy2] S. Dyatlov, *Microlocal limits of Eisenstein functions away from the unitarity axis,* to appear in Journal of Spectral Theory, arXiv:1109.3338.

[DyGu] S. Dyatlov and C. Guillarmou, *Scattering phase asymptotics with fractal remainders,* in preparation.

[DyZw] S. Dyatlov and M. Zworski, *Quantum ergodicity for restrictions to hypersurfaces,* preprint, arXiv:1204.0284.

[Fa] A. Fathi, *Expansiveness, hyperbolicity, and Hausdorff dimension,* Comm. Math. Phys. **126**(1989), no. 2, 249–262.

[GéSj] C. Gérard and J. Sjöstrand, *Semiclassical resonances generated by a closed trajectory of hyperbolic type,* Comm. Math. Phys. **108**(1987), no. 3, 391–421.

[GéLe] P. Gérard and É. Leichtnam, *Ergodic properties of eigenfunctions for the Dirichlet problem,* Duke Math. J. **71**(1993), no. 2, 559–607.

[GrLe] C.R. Graham and J.M. Lee, *Einstein metrics with prescribed conformal infinity on the ball,* Adv. Math. **87**(1991), no. 2, 186–225.

[GrSj] A. Grigis and J. Sjöstrand, *Microlocal analysis for differential operators: an introduction,* Cambridge University Press, 1994.

[Gu] C. Guillarmou, *Meromorphic properties of the resolvent on asymptotically hyperbolic manifolds,* Duke Math. J. **129**(2005), no. 1, 1–37.

[GuNa] C. Guillarmou and F. Naud, *Equidistribution of Eisenstein series on convex co-compact hyperbolic manifolds,* preprint, arXiv:1107.2655.

[GuSt] V. Guillemin and S. Sternberg, *Geometric asymptotics,* AMS, 1990.

[GuZw] L. Guillopé and M. Zworski, *Polynomial bounds on the number of resonances for some complete spaces of constant negative curvature near infinity,* Asymp. Anal. **11**(1995), no. 1, 1–22.

[HaVa] A. Hassell and A. Vasy, *The spectral projections and the resolvent for scattering metrics,* J. Anal. Math. **79**(1999), 241–298.

[HaZe] A. Hassell and S. Zelditch, *Quantum ergodicity for boundary values of eigenfunctions,* Comm. Math. Phys. **248**(2004), no. 1, 119–168.

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

[HeRo] B. Helffer and D. Robert, *Calcul fonctionnel par la transformation de Mellin et opérateurs admissibles,* J. Funct. Anal. **53**(1983), no. 3, 246–268.

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

[HöIV] L. Hörmander, *The Analysis of Linear Partial Differential Operators IV. Fourier Integral Operators,* Springer, 1985.

[Ja] D. Jakobson, *Quantum unique ergodicity for Eisenstein series on $PSL_2(\mathbb{Z})\backslash PSL_2(\mathbb{R})$,* Ann. Inst. Fourier **44**(1994), no. 5, 1477–1504.

[Ki] Y. Kifer, *Large deviations in dynamical systems and stochastic processes,* Trans. Amer. Math. Soc. **321**(1990), no. 2, 505–524.

[Li] E. Lindenstrauss, *Invariant measures and arithmetic quantum unique ergodicity,* Ann. of Math. (2), **163**(2006), no. 1, 165–219.

[LuSa] W.Z. Luo and P. Sarnak, *Quantum ergodicity of eigenfunctions on* $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathbb{H}^2$, Publ. Math. de l'IHES, **81**(1995), 207–237.

[Ma] R. Mazzeo, *Unique continuation at infinity and embedded eigenvalues for asymptotically hyperbolic manifolds,* Amer. J. Math. **113**(1991), no. 1, 25–45.

[MaMe] R.R. Mazzeo and R.B. Melrose, *Meromorphic extension of the resolvent on complete spaces with asymptotically constant negative curvature,* J. Funct. Anal. **75**(1987), no. 2, 260–310.

[Me] R.B. Melrose, *Geometric scattering theory,* Lectures at Stanford, Cambridge University Press.

[MeSBVa] R.B. Melrose, A. Sá Barreto, and A. Vasy, *Analytic continuation and semiclassical resolvent estimates on asymptotically hyperbolic spaces,* preprint, arXiv:1103.3507.

[MeZw] R.B. Melrose and M. Zworski, *Scattering metrics and geodesic flow at infinity,* Invent. Math. **124**(1996), no. 1–3, 389–436.

[Ni] P.J. Nicholls, *The ergodic theory of discrete groups,* Lecture Note series 143, London Mathematical Society, Cambridge University Press.

[No] S. Nonnenmacher, *Spectral problems in open quantum chaos,* preprint, arXiv:1105.2457.

[NoZw] S. Nonnenmacher and M. Zworski, *Quantum decay rates in chaotic scattering,* Acta Math. **203**(2009), no. 2, 149–233.

[Pa] S.J. Patterson, *The limit set of a Fuchsian group,* Acta Math. **136**(1976), no. 3–4, 241–273.

[PeSa] Ya.B. Pesin and V. Sadovskaya, *Multifractal analysis of conformal Axiom A flows,* Comm. Math. Phys. **216**(2001), no. 2, 277–312.

[PeRo] V. Petkov and D. Robert, *Asymptotique semi-classique du spectre d'hamiltoniens quantiques et trajectoires classiques périodiques,* Comm. PDE **10**(1985), no. 4, 365–390.

[Ro] D. Robert, *Autour de l'approximation semi-classique,* Progress in Mathematics 68, Birkhäuser.

[RoTa] D. Robert and H. Tamura, *Semi-classical asymptotics for local spectral densities and time delay problems in scattering processes,* J. Funct. Anal. **80**(1988), no. 1, 124–147.

[Sa] P. Sarnak, *Recent progress on the quantum unique ergodicity conjecture,* Bull. Amer. Math. Soc. **48**(2011), no. 2, 211–228.

[Sh] A.I. Shnirelman, *Ergodic properties of eigenfunctions,* Usp. Mat. Nauk. **29**(1974), 181–182.

[So] K. Soundararajan, *Quantum unique ergodicity for* $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$*,* Ann. of Math. (2) **172**(2010), no. 2, 1529–1538.

[Su79] D. Sullivan, *The density at infinity of a discrete group of hyperbolic motions,* Publ. Math. de l'IHES **50**(1979), 171–202.

[Su84] D. Sullivan, *Entropy, Hausdorff measures old and new, and limit sets of geometrically finite Kleinian groups,* Acta Math. **153**(1984), no. 3–4, 259–277.

[ToZe10] J. Toth and S. Zelditch, *Quantum ergodic restriction theorems, I: interior hypersurfaces in domains with ergodic billiards,* preprint, arXiv:1005.1636.

[ToZe11] J. Toth and S. Zelditch, *Quantum ergodic restriction theorems, II: manifolds without boundary,* preprint, arXiv:1104.4531.

[Va10] A. Vasy, *Microlocal analysis of asymptotically hyperbolic and Kerr–de Sitter spaces,* with an appendix by S. Dyatlov, preprint, arXiv:1012.4391.

[Va11] A. Vasy, *Microlocal analysis of asymptotically hyperbolic spaces and high energy resolvent estimates,* preprint, arXiv:1104.1376.

[VũNg] San Vũ Ngọc, *Systèmes intégrables semi-classiques: du local au global,* Panoramas et Synthèses 22, 2006.

[Yo] L-S. Young, *Large deviations in dynamical systems,* Trans. Amer. Math. Soc. **318**(1990), no. 2, 525–543.

[Ze87] S. Zelditch, *Uniform distribution of eigenfunctions on compact hyperbolic surfaces,* Duke Math. J. **55**(1987), no. 4, 919–941.

[Ze91] S. Zelditch, *Mean Lindelöf hypothesis and equidistribution of cusp forms and Eisenstein series,* J. Funct. Anal. **97**(1991), no. 1, 1–49.

[Ze09] S. Zelditch, *Recent developments in mathematical quantum chaos,* Curr. Dev. Math., 2009, 115–204.

[ZeZw] S. Zelditch and M. Zworski, *Ergodicity of eigenfunctions for ergodic billiards,* Comm. Math. Phys. **175**(1996), no. 3, 673–682.

[Zw] 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`

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

E-mail address: `cguillar@dma.ens.fr`

DMA, U.M.R. 8553 CNRS, École Normale Superieure, 45 rue d'Ulm, 75230 Paris cedex 05, France