Commun. Math. Phys. 160, 81–92 (1994)

Communications in
Mathematical
Physics

On the Rate of Quantum Ergodicity I: Upper Bounds

Steven Zelditch*

Department of Mathematics, Johns Hopkins University, Baltimore, MD 21218, USA

Received: 3 May 1993

Abstract: One problem in quantum ergodicity is to estimate the rate of decay of the sums

$$S_k(\lambda; A) = \frac{1}{N(\lambda)} \sum_{\sqrt{\lambda_j} \leqq \lambda} |(A\varphi_j, \varphi_j) - \bar{\sigma}_A|^k$$

on a compact Riemannian manifold (M, g) with ergodic geodesic flow. Here, $\{\lambda_j, \varphi_j\}$ are the spectral data of the Δ of (M, g), A is a 0-th order ψDO, $\bar{\sigma}_A$ is the (Liouville) average of its principal symbol and $N(\lambda) = \#\{j\colon \sqrt{\lambda_j} \leqq \lambda\}$. That $S_k(\lambda; A) = o(1)$ is proved in [S, Z.1, CV.1]. Our purpose here is to show that $S_k(\lambda; A) = O((\log \lambda)^{-k/2})$ on a manifold of (possibly variable) negative curvature. The main new ingredient is the central limit theorem for geodesic flows on such spaces ([R, Si]).

Quantum ergodicity is the study of the spectral properties of Schrödinger operators with ergodic classical flows. In this paper, we will be concerned with a special case: that of a Laplacian on a compact n-dimensional Riemannian manifold M of negative curvature. As is well known, the geodesic flow G^t on S^*M is then ergodic. Δ is also quantum ergodic in the following sense: for any choice of orthonormal basis $\{\phi_j\}$ of eigenfunctions

$$\Delta\phi_j = \lambda_j\phi_j, \quad 0 = \lambda_0 < \lambda_1 \leqq \lambda_2 \ldots \uparrow \infty$$

and any $A \in \Psi^0(M)$, one has

$$\lim_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{\sqrt{\lambda_j} \leqq \lambda} |(A\varphi_j, \varphi_j) - \bar{\sigma}_A| = 0 \, . \tag{0.1}$$

Here, $N(\lambda) = \#\{j\colon \sqrt{\lambda_j} \leqq \lambda\}$, $\Psi^m(M)$ is the space of ψDO's (pseudodifferential operators) of order m, σ_A is the principal symbol and $\bar{\sigma}_A := \frac{1}{\mathrm{vol}(S^*M)} \int_{S^*M} \sigma_A d\mu$ is

* Partially supported by NSF grant #DMS-9243626

the average value of σ_A with respect to Liouville measure $d\mu$. We will refer to the articles [CV, S, Z.1] for further discussion of (0.1) and its interpretation. Let us just note here that it is a somewhat weak version of the limit formula

$$\lim_{\lambda \to \infty} (A\varphi_j, \varphi_j) = \bar{\sigma}_A , \tag{0.2}$$

which has been so far neither proved nor disproved.

A natural problem is to determine the rate at which the sums

$$S_k(\lambda; A) := \frac{1}{N(\lambda)} \sum_{\sqrt{\lambda_j} \leqq \lambda} |(A\varphi_j, \varphi_j) - \bar{\sigma}_A|^k$$

in (0.1) tend to 0 ([CV.2, Sa]). We will call any upper bound of the form

$$S_k(\lambda; A) = O_A(R_k(\lambda))$$

a rate of quantum ergodicity. Our purpose in this note is to estimate the rate of quantum ergodicity for Riemannian manifolds of curvature $K < 0$.

Recently, Sarnak [Sa] has conjectured that (the optimal) rate $R_1(\lambda) = \lambda^{-\frac{1}{4}+\varepsilon}$ for compact (and, with appropriate modifications, non-compact but finite area) hyperbolic surfaces. In fact, he conjectures this rate for the individual terms $|(A\varphi_j, \varphi_j) - \bar{\sigma}_A|$. This rate is suggested by the Lindelöf hypothesis for Rankin–Selberg L functions,

$$L(s, \phi_j \otimes \phi_j) := (E(\circ, s)\phi_j, \phi_j) ,$$

where $E(z, s)$ resp. ϕ_j, is an Eisenstein series, resp. a cusp form, on an arithmetic quotient $\mathbf{H}^2/\Gamma$ (e.g. $\Gamma = SL_2(\mathbf{Z})$). The connection is of course that $E(\circ, s)$ is regarded as a multiplication operator playing the role of A above. We refer to [Sa] for discussion of the grounds for this conjecture, and to [Z.2–3] for some related results.

Our estimate of the rate of quantum ergodicity on negatively curved manifolds (M, g) is based on M. Ratner's estimates on the rate of convergence in $L^p(S^*M)$ of the time mean $\bar{a}(z, T) := \frac{1}{T}\int_0^T a(G^t z)dt$ to the space mean $\bar{a}$ of a smooth function a on S^*M. The main result is:

Theorem. *Let (M, g) be a compact Riemannian manifold of (possibly variable) negative curvature. Then $S_n(\lambda; A) = 0((\log \lambda)^{\frac{-n}{2}})$.*

Remarks and Acknowledgements.

(1) This logarithmic improvement over the earlier rate $o(1)$ is somewhat analogous to the logarithmic improvement $0(\lambda^{n-1}/\log \lambda)$ of Randol, Bérard, and others ([B, Rn.1]) on the Duistermaat–Guillemin estimate $o(\lambda^{n-1})$ for the remainder term $R(\lambda)$ in Weyl's law on a negatively curved manifold. In both cases, the logarithm arises from the exponential growth rate of the geodesic flow and length spectrum. Sarnak's conjecture is then somewhat analogous to Randol's conjecture that $R(\lambda) = 0(\lambda^{\frac{1}{2}+\varepsilon})$ on a surface of constant negative curvature [Rn.2].

(2) We thank P. Gerard for correcting some errors in an earlier version and in particular for pointing out an improvement of the main estimate. We also thank D. Hejhal and M. Ratner for helpful comments on the central limit theorem.

1. Preliminaries

We begin by reviewing some relevant background and terminology.

(*1a*) *Friedrichs quantization* ([Z.1, CV, T]). We will fix, once and for all, a quantization map

$$Op^F\colon C^\infty(S^*M) \to \Psi^0(M)$$

with the property that

$$Op^F(a) \geqq 0 \quad \text{if } a \geqq 0\,.$$

Such an Op^F is called a Friedrichs quantization. It is actually defined for all symbols on T^*M; we regard $C^\infty(S^*M)$ as a subspace of symbols of order 0.

(*1b*) *Microlocal Lifts of Eigenfunctions.* We denote by $d\Phi_k \in \mathcal{M}_1^+(S^*M)$ the following positive linear functionals on $C(S^*M)$:

$$\langle a, d\Phi_k\rangle := (Op^F(a)\varphi_k, \varphi_k)\,.$$

Clearly, $d\Phi_k$ is a probability measure on S^*M. For an explicit formula, see ([Z.1, CV]). We will refer to $d\Phi_k$ as the microlocal lift of ϕ_k to S^*M, although of course it projects to φ_k^2.

(*1c*) *Egorov's Theorem* ([T, ch. VIII, CV]). Let $U(t) = \exp it\sqrt{\Delta}$. For each t, $U(t)$ is an FIO (Fourier Integral Operator) in the Hörmander class $I^0(M \times M, C_t)$, where $C_t \subset T^*(M \times M)\backslash 0$ is the graph of the geodesic flow G^t.

Egorov's theorem states:

$$U(-t)Op^F(a)U(t) = Op^F(a \circ G^t) + R_t\,, \tag{1.1}$$

where R_t is a ψDO of order -1.

(*1.d*) *The Central Limit and Moment Estimates for Geodesic Flows on Compact Negatively Curved Manifolds* ([R, Si]). A function $f \in L^\infty(S^*M)$ is said to obey the central limit theorem (CLT) relative to the geodesic flow G^t if

$$\lim_{T\to\infty} \mu\left\{z\colon \frac{\int_0^T (f(G^t(z)) - \bar{f})dt}{\sqrt{D_T(f)}} < \alpha\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2}\,dx\,. \tag{1.2}$$

Here, μ is Liouville measure, $\bar{f}$ is (as above) the Liouville mean of f and D_T is the variance of f:

$$D_T(f) := \int_{S^*M} \left|\int_0^T (f(G^t z) - \bar{f})dt\right|^2 d\mu\,. \tag{1.3}$$

Equivalently, f obeys the CLT if

$$d\mu_T \rightharpoonup \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx \quad (T \to \infty)\,, \tag{1.4}$$

where

$$d\mu_T = \left(\frac{1}{\sqrt{D_T(f)}} \int_0^T (f \circ G^t - \bar{f})dt\right)_* d\mu \tag{1.5}$$

(the pushforward to $\mathbf{R}$ of $d\mu$ under the indicated partial time average of f). Also, $\rightharpoonup$ denotes weak* convergence of measures on $C_b(\mathbf{R})$, the space of bounded continuous functions.

Ya.G. Sinai (in the constant curvature case) and M. Ratner (in the general case) have proved that a broad class of functions, including $C^\infty(S^*M)$, obey the CLT on a negatively curved manifold. Ratner in fact proves the CLT for compact (M, g) with transitive Anosov geodesic flow.

In this paper we will actually need the stronger property:

$$\|\bar{f}(T, \circ) - \bar{f}\|_{2k} = O\left(\frac{1}{\sqrt{T}}\right) \quad (f \in C^\infty(S^*M)) . \tag{1.6}$$

Here $\|\cdot\|_{2k}$ is the L^{2k}-norm on S^*M. In the case $k = 1$, this is contained in Ratner's Variance Theorem:

(1.7) Theorem [R, Theorem 3.1]. *Let V denote the generator of G^t and suppose $f \in C^\infty(S^*M)$. Then*:

(i) *If $f - \bar{f} = Vh$ has no solution in $L^2(S^*M)$, then $D_T f \sim \sigma_f T$ (as $T \to \infty$) for a certain constant $\sigma_f > 0$;*

(ii) *If $f - \bar{f} = Vh$ for some $h \in L^2(S^*M)$, then $D_T f = O(1)$ as $T \to \infty$.* ■

Some comments: First, the h in case (ii) is now known to be C^∞ [LMM, Theorem 2.1], so the $D_T f$ estimates becomes obvious. Second, we may plug the formula for $D_T f$ in case (i) into the CLT to obtain the more familiar version:

$$\text{(CLT)} \qquad d\mu_T \rightharpoonup \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx , \tag{1.8}$$

where

$$d\mu_T = \left(\frac{1}{\sigma\sqrt{T}} \int_0^T (f \circ G^t - \bar{f}) dt\right)_* d\mu .$$

(Here, $\sigma = \sigma_f$.) Third, we see that the L^{2k}-estimates in (1.6) are automatic in case (ii), and in case (i) are equivalent to:

$$\int_{\mathbf{R}} x^{2k} d\mu_T(x) = O(1) \quad \text{as } T \to \infty . \tag{1.9}$$

These moment estimates would follow from (1.8) if the weak convergence were valid against the (unbounded) powers x^{2k}. In fact, this can be proved to be the case. For $k = 1$, (1.9) follows immediately from Ratner's Variance Theorem (1.7). The general case is implicit in [R], but the statements (1.6) and (1.8) are not explicity noted there. For the sake of completeness, we will now explain how to dig the proof of (1.6)–(1.8) out of [R].

(1.9) Theorem. *Let $f \in C^\infty(S^*M)$, and suppose f is in case* (i) *above. Then*

$$\int_{S^*M} |\bar{f}(T, z) - \bar{f}|^{2k} d\mu = \frac{\sigma^{2k}}{T^k} (1 + o(1))$$

(*case* (ii) *is obvious, as mentioned above*).

Proof. The first step is to put this problem in normal form. The normal form is a suspension (S^t, W, v) of a shift $(\phi, X, \boldsymbol{\mu})$ of finite type. We refer to [R] for precise definitions and background. Let us just note that $(\phi, X, \boldsymbol{\mu})$ is constructed from a Markov partition of $(G^t, S^*M, d\mu)$. This Markov partition also determines a positive Hölder continuous function ℓ on X, a suspended flow S^t on

$$W = \{(x, y): x \in X, 0 \leqq y \leqq \ell(x)\}$$

and a Lipschitz continuous, finitely-many-to-one conjugacy ψ: $S^*M \to W$ such that

$$\psi S^t = G^t \psi .$$

ψ takes Liouville measure $d\mu$ on S^*M to an invariant Gibbs measure $v = \psi_* d\mu$ on W and it takes $C^\infty(S^*M)$ to a class (denoted $\Upsilon_{\rho,\kappa}$ in [R]) of Hölder continuous functions on W. The theorem is therefore reduced to the case of suspended shift automorphisms.

The proof of (1.9) for Hölder f, relative to (S^t, W, v) is implicitly contained in the proof of [R, Theorem 3.1]. However, only the case $k = 1$ is explicitly considered. Let us indicate, briefly, the modifications necessary for general k. We will assume the reader is familiar with the notation and material in [R]. In particular, we will need to use:

(i) $a(T, w) = \frac{1}{\sqrt{T}} \left(\int_0^T (f(S^{-u}w) - \bar{f}) du \right);$

(ii) $B(T, x) = \frac{1}{\sqrt{T}} \left(\int_0^T (f(S^{-u}(x, 0)) - \bar{f}) du \right);$

(iii) $F(x) = \int_0^{\ell(x)} f(x, y) dy;$

(iv) $\tilde{F}(x) = F(x) - \left(\frac{\bar{F}}{\bar{\ell}} \right) \ell(x);$

(v) $D(T, x) = \sum_{i=0}^{n(T,x)} \tilde{F}(\phi^{-i}x);$ (1.10)

where

(vi) $n(T, x) =$ the number of times an S^t-trajectory, starting from x, hits X in time T;

(vii) $S_k(T, x) = \frac{1}{\sqrt{T}} \sum_{i=0}^{[T/\bar{\ell} + k\varepsilon\sqrt{T}]} \tilde{F}(\phi^{-i}x)$

(where $\varepsilon > 0$ is given),

(viii) $A_{kT} = \{x \in X: \frac{T}{\bar{\ell}} + k\varepsilon\sqrt{T} \leqq n(T, x) \leqq \frac{T}{\bar{\ell}} + (k+1)\varepsilon\sqrt{T}\}$,

and finally

(ix) $S(T, x) = S_k(T, x)$ for $x \in A_{kT}$.

One has [R, p. 188]:

$$|a(T, w) - B(T, x)| < \frac{C_1}{\sqrt{T}}, \tag{1.11}$$

where $w = (x, y)$ and C_1 is independent of w and x. Therefore,

$$a^{2k}(T, w) = B^{2k}(T, x) + O\left(\sum_{j=1}^{2k} T^{-j/2} B^{2k-j}(T, x) \right). \tag{1.12}$$

Integrating both sides of (1.12) against dv on W, using that $dv = \frac{1}{\ell}(d\mu \times dy)$ ([R, p. 182]), and applying Hölder's inequality, we get:

$$\left| \int_W a^{2k}(T, w) dv - \int_X B^{2k}(T, x) d\mu_\ell \right| \leqq C \sum_{j=1}^{2k} T^{-j/2} \left(\int_X B^{2k}(T, x) d\mu_\ell \right)^{\frac{2k-j}{2k}} . \tag{1.13}$$

Here, $d\mu_\ell = \frac{\ell(x)}{\ell} d\mu$ ([R, (19)]). Further, we have ([R, p. 194–5]):

$$\left| B(T, x) - \frac{1}{\sqrt{T}} D(T, x) \right| < \frac{C_2}{\sqrt{T}}, \tag{1.14i}$$

$$\left| \frac{D(T, x)}{\sqrt{T}} - S_k(T, x) \right| \leqq R\varepsilon \quad \text{on } A_{kT} \quad (\text{for a certain } R > 0) . \tag{1.14ii}$$

Hence,

$$\left| \int_X B^{2k}(T, x) d\mu_\ell - \int_X \frac{D^{2k}(T, x)}{T^k} d\mu_\ell \right| \leqq C \sum_{j=1}^{2k} T^{-j/2} \left(\int_X D^{2k}(T, x) d\mu_\ell \right)^{\frac{2k-j}{2k}} \tag{1.15i}$$

and

$$\left| \int_X \frac{D^{2k}(T, x)}{T^k} d\mu_\ell - \int_X S^{2k}(T, x) d\mu_\ell \right| \leqq \sum_{j=1}^{2k} \varepsilon^j \left(\int_X S^{2k}(T, x) d\mu_\ell \right)^{\frac{2k-j}{2k}} \tag{1.15ii}$$

(compare [R, (19), (20)].

By [R, (23)] one has:

$$E_{\mu_\ell}(S^{2k}(T, x)) \leqq C_k \frac{1}{L^k}, \tag{1.16}$$

where $0 < L < \inf_X \ell(x)$, and where E_{μ_ℓ} is expectation relative to $d\mu_\ell$. Hence,

$$\begin{aligned} E_{\mu_\ell}(S^{2k}(T, x)) &= O(1) \quad \text{as } T \to \infty \\ &\Rightarrow E_{\mu_\ell}(D^{2k}(T, x)) = O(1) \quad \text{by (1.15ii)} \\ &\Rightarrow E_{\mu_\ell}(B^{2k}(T, x)) = O(1) \quad \text{by (1.15i)} \\ &\Rightarrow \int_W a^{2k}(T, w) dv = O(1) \quad \text{by (1.13)} \\ &\Rightarrow \int_W |\bar{f}(T, w) - \bar{f}|^{2k} dv = O\left(\frac{1}{T^k}\right), \\ &\Rightarrow \int_{S^*M} |\bar{f}(T, z) - \bar{f}|^{2k} dv = O\left(\frac{1}{T^k}\right) \end{aligned} \tag{1.17}$$

(where f denotes any Hölder continuous function in either case). This latter estimate is all we need for the rate of quantum ergodicity, but we will go on with the proof of the sharper asymptotic formula (1.9) since we have nearly completed it anyway.

As in [R, p. 195], we introduce a continuous cut-off function $h_N \in C_c(\mathbf{R})$, and write:

$$E_{\mu_\ell}[S^{2k}] = E_{\mu_\ell}[S^{2k} h_N(S)] + E_{\mu_\ell}[S^{2k}(1 - h_N)(S)] \, . \tag{1.18}$$

Since $x^{2k} h_N(x) \in C_b(\mathbf{R})$, the CLT applies to the first term and gives the limit σ^{2k} as $T \to \infty$. We claim that the second term is $O\left(\frac{1}{N}\right)$, uniformly as $T \to \infty$. To see this, we only need to make the following changes in [R, p. 195–6]:

(i) On the first line of (22), change S^2 to S^{2k} everywhere;
(ii) On the second line, change S^4 to S^{4k};
(iii) On the third line, change N^{2m-4} to N^{2m-4k};
(iv) In the middle of p. 196, set $m \geqq 2k + 1$.

Ratner's argument is stable under these modifications, and shows that

$$\overline{\lim_{T \to \infty}} \left| E_{\mu_\ell}(S^{2k}(T, x)) - \int_{\mathbf{R}} x^2 h_N(x) e^{-x^2/2\sigma^2} dx \right| \leqq \frac{C}{N} ,$$

hence

$$\lim_{T \to \infty} E_{\mu_\ell}(S^{2k}(T, x)) = \sigma^{2k} ,$$

which implies, as above, (1.9). ■

2. Proof of the Theorem

The proof will closely follow the pattern of our proof of (0.1) in [Z.1]. The new feature is the control over the time dependence of all the relevant terms.

Our first step is, predictably, to invoke the identity

$$(A\varphi_j, \varphi_j) = (U(-t)AU(t)\varphi_j, \varphi_j) \tag{2.1}$$

and Egorov's theorem

$$U(-t)Op^F(a)U(t) = Op^F(a \circ G^t) + R_t \, , \tag{2.2}$$

where ord $R_t = -1$ (cf. (1.c)).

Averaging in t for $t \in [0, T]$ and substituting in $S_k(\lambda; A)^{\frac{1}{k}}$, we easily obtain:

$$S_k(\lambda; A)^{\frac{1}{k}} \leqq S_k\left(\lambda; \frac{1}{T} \int_0^T Op^F(\sigma_A \circ G^t) dt\right)^{\frac{1}{k}}$$
$$+ S_k\left(\lambda; \frac{1}{T} \int_0^T R_t dt\right)^{\frac{1}{k}} . \tag{2.3}$$

We will estimate these two terms in Lemmas 1–2 below. For notational simplicity, we will abbreviate $\frac{1}{T} \int_0^T \sigma_A \circ G^t dt$ by $\bar{\sigma}_A(T)$.

(2.4) Lemma 1. $S_{2k}(\lambda; Op^F(\bar{\sigma}_A(T))) = O_A\left(T^{-k} + \frac{e^{cT}}{\lambda}\right)$ *for some* $c \in \mathbf{R}^+$.

Proof. We may (and will) assume $\bar{\sigma}_A = 0$.
Then

$$S_{2k}(\lambda, Op^F(\bar{\sigma}_A(T))) = \frac{1}{N(\lambda)} \sum_{\sqrt{\lambda_j} \leqq \lambda} |\langle \bar{\sigma}_A(T), d\Phi_j \rangle|^{2k} .$$

By Hölder's inequality on $L^{2k}(S^*M, d\Phi_j)$,

$$|\langle \bar{\sigma}_A(T), d\Phi_j \rangle|^{2k} \leqq \langle |\bar{\sigma}_A(T)|^{2k}, d\Phi_j \rangle .$$

Hence,

$$S_{2k}(\lambda; Op^F(\bar{\sigma}_A(T))) \leqq \frac{1}{N(\lambda)} \sum_{\sqrt{\lambda_j} \leqq \lambda} \langle |\bar{\sigma}_A(T)|^{2k}, d\Phi_j \rangle$$
$$= \frac{1}{N(\lambda)} \operatorname{Tr} \pi_\lambda Op^F(|\bar{\sigma}_A(T)|^{2k}) , \tag{2.5}$$

where π_λ is the orthogonal projection onto the span of the φ_j's with $\sqrt{\lambda_j} \leqq \lambda$.
We now study the asymptotics of the traces in (2.5):

(2.6) Proposition. *Let* $b \in C^\infty(S^*M)$, *and let* $B = Op^F(b)$. *Then*:

$$\left| \frac{1}{N(\lambda)} \operatorname{Tr} \pi_\lambda B - \bar{b} \right| = O(\|b\|_{C^m} \lambda^{-1}) ,$$

where $\|b\|_{C^m}$ *is the* C^m *norm and* m *is some number* $\geqq 2(3n+4)$; *the O-symbol is independent of b.*

Proof. Except for the b-dependence, this estimate has been proved by Guillemin in [G, §3 Theorem 1′]. We therefore have only to keep track of the b-aspect in this proof. Actually, it is convenient to modify the proof a little to bring this aspect out more clearly.
Let

$$\phi_B(\lambda) = \frac{1}{N(\lambda)} \operatorname{Tr} \pi_\lambda B .$$

Following [G, §3 Lemmas 2–3], we note that

$$\Phi_B(\lambda) = \rho * \phi_B(\lambda) + O(\|B\| \lambda^{-1}) \tag{2.7}$$

for any $\rho \in C^\infty(\mathbf{R})$ with $\hat{\rho} \in C_0^\infty(\mathbf{R})$, $\int \rho = 1$. The B-dependence is easily read off from [G, loc. cit]. The O-symbols are understood to be independent of B unless indicated otherwise.

It follows from the Calderon–Vaillancourt theorem [T, ch. XIII] that $\|B\| \leqq \|b\|_{C^m}$ for any $m \geqq 2(3n+4)$. The remainder of the proposition therefore amounts to showing that

(2.8) Proposition. $\rho * \phi_B(\lambda) = \bar{b} \frac{1}{n(2\pi)^n} + O(\|b\|_{C^m} \lambda^{-1})$ *if* $\hat{\rho} \in C_0^\infty(\mathbf{R})$, $\int \rho = 1$.

Proof. $\rho * \phi_B(\lambda) = \int_0^\lambda \rho * \frac{d\Phi_B}{d\lambda} d\lambda$, and

$$\rho * \frac{d\Phi_B}{d\lambda}(\lambda) = \int e^{i\lambda t} \hat{\rho}(t) \operatorname{Tr} BU(t) dt , \tag{2.9}$$

where $U(t) = \exp it\sqrt{\Delta}$. Now for small t, say $|t| < \varepsilon$, we can represent the Schwartz kernel $U(t, x, y)$ in the form [G, loc. cit.]:

$$U(t, x, y) = \int_{\mathbf{R}^n} (1 + a_{-1}(y, \eta))e^{i(t|\eta|) + q(x,y,\eta)}\, d\eta + T_t(x, y) ,$$

where:

(i) a_{-1} is a symbol of order -1.
(ii) q is homogeneous of degree 1 in η and $q(x, x, \eta) = 0$.
(iii) T_t is a smoothing operator.

$BU(t)$ can be represented in a similar form, but it seems more convenient just to compose the kernels directly. Writing B in the usual form

$$B(x, y) = \int_{\mathbf{R}^n} b^F(x, \xi)e^{i\langle \xi, x-y\rangle}\, d\xi$$

(cf. [T, VII]), and composing, we get

$$\operatorname{Tr} BU(t) = \iiiint b^F(x, \xi)(1 + a_{-1}(y, \eta))e^{i[\langle x-y, \xi\rangle + t|\eta| + q(x, y, \eta)]}\, dx\, dy\, d\xi\, d\eta + \operatorname{Tr} BT_t\, . \tag{2.10}$$

Plugging (2.10) into (2.9), and replacing (ξ, η) by $(\lambda\xi, \lambda\eta)$, we end up with an oscillatory integral as $\lambda \to \infty$ plus a remainder $T_\lambda = \int e^{i\lambda t}\hat{\rho}(t)\operatorname{Tr} BT_t dt$. We can integrate T_λ by parts in λ to see that it is $O(\|B\|\lambda^{-N})$ for all N, so again the Calderon–Vaillancourt estimate shows that this term is as it should be.

We are left with the principal term,

$$\beta_\lambda = \int e^{i\lambda t}\hat{\rho}(t)I(\lambda, t)dt ,$$

where $I(\lambda, t)$ is the integral in (2.10). We can apply the method of stationary phase to determine the asymptotics of β_λ. The principal term has to agree with the result in [G, loc. cit.], so the only question is the b-dependence of the remainder estimate. As is well-known, the remainder is of the form $O(\lambda^{-1}\|b^F\|_{C^2})$, where the O-symbol depends on the C^2 norms of a_{-1}, q, etc., which are irrelevant to our purposes [Hö I, Theorem 7.7.5]). Thus, the proposition reduces to showing that $\|b^F\|_{C^2}$ is bounded by $\|b\|_{C^m}$ for some $m \geqq 2(3n + 4)$. This is straightforward, and we will only sketch the proof. First, one has an explicit integral formula relating b^F to b [T, VII (1.5), (2.1)]:

$$b^F(x, \xi) = \iint F(\xi_2, \zeta)b(y, \zeta)F(\xi, \zeta)e^{i(y-x)\cdot(\xi_2-\xi)}\, dy\, d\xi_2\, d\zeta ,$$

where F is a certain amplitude [T, loc. cit.]. We can pass any derivatives on b^F under the integral sign and estimate the resulting integral. The only problem is the absolute convergence of the $d\zeta d\xi_2$ integrals, which due to the form of F, can be reduced to the convergence of the $d\xi_2$ integral as $|\xi_2 - \xi| \to \infty$. However, repeated partial integration with $\dfrac{(\xi_2 - \xi)\cdot Dy}{|\xi_2 - \xi|^2}$ introduces sufficient decay in ξ_2 to render the integral absolutely convergent. The result is evidently bounded by a C^m-norm of b, concluding the proof of both Propositions (2.8) and (2.9). ■

From Proposition 2.6, it follows that:

(2.11) Corollary.

$$S_{2k}(\lambda; Op^F(\bar{\sigma}_A(T))) = \int_{S^*M} |\bar{\sigma}_A(T)|^{2k}\, d\mu + O(\| |\bar{\sigma}_A(T)|^{2k} \|_{C^m})$$

(*for some* $m \in \mathbf{N}$).

Note that $|\bar{\sigma}_A(T)|^{2k}$ is in fact smooth, since the power is even.

The next step is to estimate these two terms. The first term was estimated in Theorem (1.9):

$$\int_{S^*M} |\bar{\sigma}_A(T)|^{2k}\, d\mu = O\left(\frac{1}{T^k}\right). \tag{2.12}$$

This leaves us to estimate the C^k-norm of $|\bar{\sigma}_A(T)|^{2k}$.

(2.13) Proposition. *Let* $f \in C^\infty(S^*M)$. *Then there exists* $L > 0$ *so that*

$$\| f \circ G^t \|_{C^k} = O(e^{Lt}) \quad \text{as } t \to \infty\,.$$

Proof. This follows immediately from the well-known exponential growth of the derivatives of G^t (i.e. of Jacobi fields and their derivatives). We refer the reader to [Be, Appendix] or to [V], for further details. ■

Lemma 1 now follows from Propositions 2.12 and 2.13. ■

We now turn to the second term in (2.3):

(2.14) Lemma 2. $S_k\left(\lambda; \frac{1}{T}\int_0^T R_t dt\right)^{\frac{1}{k}} = O\left(\frac{e^{LT}}{\lambda}\right)$ *for some* $L > 0$.

Proof. It suffices to show that $\|\sqrt{\Delta} R_t\| = O(e^{Lt})$.

Essentially, this requires a series of estimates on amplitudes, phases and remainder terms of the wave kernel as $t \to \infty$. A very careful and detailed study of these estimates has recently been published by Volovoy [V]. For the sake of brevity, we will only summarize the main points in the estimate of $\|\sqrt{\Delta} R_t\|$ and refer the interested reader to [V] for further discussion.

First, we recall [V, Theorem 0.2] that for any t_0, there is an interval $(t_0 - \varepsilon, t_0 + \varepsilon)$ on which $U(t)$ can be represented in the form

$$U(t) = Q(t) + T(t)\,,$$

where $Q(t)$ is a local FIO constructed by the method of geometric optics and $T(t)$ is a smoothing operator satisfying

$$\| T(t)(x, y) \|_{C^k} = O(e^{Lt}) \quad \text{for some } L > 0\,.$$

Further, the amplitudes and phases of $Q(t)$ satisfy a similar exponential bound.

We can (locally) break up $Q(t)$

$$Q(t) = Q_0(t) + Q_{-1}(t)$$

into its leading order part plus a (-1-st) order remainder. Then

$$U(-t)AU(t) = Q_0(-t)AQ_0(t) + S_t\,,$$

where S_t is a ψDO of order -1. Hence,

$$\|\sqrt{\Delta}R_t\| \leqq \|\sqrt{\Delta}(Q_0(-t)Op^F(a)Q_0(t) - Op^F(a \circ G^t))\| + \|\sqrt{\Delta}S_t\| .$$

$\sqrt{\Delta}S_t$ is a ψDO of order 0 and its norm can, as above, be estimated by some C^k norm of its complete symbol. By Volovoy's estimates, all such norms have exponential growth.

This leaves the first term, which we know from Egorov's theorem is also a ψDO of order 0. It thus suffices to estimate the C^k norm of its complete symbol.

To determine this complete symbol, we can explicitly write out the composition

$$\sqrt{\Delta}(Q_0(-t)Op^F(a)Q_0(t) - Op^F(a \circ G^t))$$

as a Fourier integral. The phase in the first term can be made to match that in the second by an application of the method of stationary phase. The leading term in this expansion is of course cancelled by the second term above, $Op^F(a \circ G^t)$, leaving a 0-th order symbol, which is determined from the amplitude and phases of $Q_0(t)$ and of $Op^F(a \circ G^t)$, together with some time independent data from $Op^F(a)$. By Volovoy's estimates and the exponential growth of G^t, it is a straightforward observation that the C^k norms of this complete symbol have exponential growth. ■

We now conclude from Lemmas 1 and 2 that

$$S_{2k}(\lambda; A) = O\left(T^{-k} + \frac{e^{cT}}{\lambda} + \frac{e^{LT}}{\lambda^{2k}}\right) \tag{2.15}$$

for any $T > 0$. Setting $T = \frac{1}{M}\log\lambda$ for sufficiently large M, we get

$$S_{2k}(\lambda; A) = O\left(\frac{1}{(\log\lambda)^k}\right).$$

If n is odd, we also have $S_n(\lambda; A) \leqq S_{2n}(\lambda, A)^{\frac{1}{2}} = O\left(\frac{1}{(\log\lambda)^{\frac{n}{2}}}\right)$. ■

References

[B] Berard, P.: On the wave equation of a compact Riemannian manifold without conjugate points. Math. Zeit **155**, 249–276 (1977)

[CV 1] Colin de Verdiere, Y: Ergodicité et functions propres. Commun Math Phys **102**, 497–502 (1985)

[CV 2] Colin de Verdiere, Y.: Hyperbolic geometry in two dimensions and trace formulas. In: Chaos et Physique Quantique Les Houches 1989 (Session LII), Giannoni, M , Voros, A., Zinn-Justin, J (eds), Amsterdam: North Holland Pr, 1991

[G] Guillemin, V: Some classical theorems in spectral theory revisited In: Seminar on Singularities Hörmander, L. (ed.) Princeton, NJ: Princeton Univ Press, 1978, pp 219–259

[LMM] de La Llave, R., Marco, J M , Moriyon, R.: Canonical perturbation theory of Anosov systems and regularity results for the Livsic cohomology equation Ann. Math **123**, 537–611 (1986)

[Rn.1] Randol, B: The Riemann hypothesis for Selberg's zeta function and the asymptotic behaviour of eigenvalues of the Laplace operator. Trans A.M S **263**, 209–223 (1978)

[Rn.2] Randol, B.: A Dirichlet Series of eigenvalue type with applications to asymptotic estimates Bull. Lond. Math. Soc. **13**, 309–315 (1981)

[R] Ratner, M.: The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. Israel J. Math **16**, 181–197 (1973)

[Sa] Sarnak, P.: Arithmetic Quantum Chaos. Schur Lectures, Tel Aviv preprint, 1992

[S] Schnirelman, A I.: Ergodic properties of eigenfunctions Usp. Math. Nauk. **29**, 181–182 (1974)

[Si] Sinai,Ya.G.: The central limit theorem for geodesic flows on manifolds of constant negative curvature Sov. Math Doklady **1**, 983–987 (1960)

[T] Taylor, M.: Pseduo Differential Operators. Princeton, NJ: Princeton Univ. Press, 1981

[V] Volovoy, A V : Improved two term asymptotics for the eigenvalue distribution function Comm. P.D.E. **15**, 1509–1563 (1990)

[Z 1] Zelditch, S.: Uniform distribution of eigenfunctions on compact hyperbolic surfaces. Duke Math. J. **55**, 919–941 (1987)

[Z.2] Zelditch, S.: Mean Lindelöff hypothesis and equidistribution of cusp forms and Eisenstein series. J. Funct. Anal **97**, 1–49 (1991)

[Z.3] Zelditch, S : Selberg Trace Formulae and equidistribution theorems. A.M S. Memoir, no **465** (1992)

Communicated by Ya.G. Sinai