Eigenfunctions of the Laplacian of Riemannian manifolds
Updated: August 15, 2017

Steve Zelditch

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL 60208
E-mail address: zelditch@math.university.edu

Key words and phrases. Laplacian, eigenfunction, nodal set, L^p norms, Weyl law, quantum limits

Research partially supported by NSF grant and DMS-1541126 and by the Stefan Bergman trust.

Contents

Preface

These lecture notes are an expanded version of the author's CBMS ten Lectures at the University of Kentucky in June 20-24, 2011. The lectures were devoted to eigenfunctions of the Laplacian and of Schrödinger operators, in particular to their L^p-norms and nodal sets. The lecture notes have undergone extensive revisions in the intervening years, due in part to progress in the field and also to the new publications on related topics, which made some of the original lecture notes obsolete. In particular, the new book [**So2**] of Chris Sogge and the author's 2013 Park City Lecture notes [**Ze7**] are also devoted to eigenfunctions and includes extensive background on pseudo-differential operators and harmonic analysis. (References for the preface can be found at the end of §1.) The book of Maciej Zworski [**Zw**] contains a systematic introduction to semi-classical Fourier integral operators and includes applications to quantum ergodicity of eigenfunctions. The recent book [**GS**] of V. Guillemin and S. Sternberg also gives background on the global theory of Fourier integral operators and in particular on their symbols. Fanghua Lin and Qing Han also have a book in progress on eigenfunctions from viewpoint of local elliptic equations. For this reason, we do not feel it is useful in these lecture notes to provide any systematic background on these techniques, although their properties will be used freely. We do include some background on symplectic geometry, pseudo-differential and Fourier integral operators to establish notation and links to other references. But overall we assume that the reader is willing to consult these other references for the basic techniques.

The purpose of these lecture notes is to convey inter-related themes and results, and so we rarely give detailed proofs. Rather we aim to outline key ideas and how they are related to other results. The lectures concentrate on the following themes:

- Local versus Global analysis of eigenfunctions. The Local analysis of eigenfunctions belongs to the theory of elliptic equations, and pertains to local solutions of the eigenvalue problem $(\Delta + \lambda)\varphi = 0$ on small balls of radius $\frac{C}{\sqrt{\lambda}}$. The global analysis belongs to hyperbolic equations, i.e., studies the eigenfunctions through the wave equation $\cos t\sqrt{-\Delta}\varphi = \cos t\sqrt{\lambda}\varphi$ and their relations to geodesics as $\lambda \to \infty$. One of the aims of these lectures is to survey both local and global methods, and to discuss how they interact. For instance, the main existence theorem that there exists a zero of φ_λ in each ball $B(p, \frac{A_g}{\sqrt{\lambda}})$ whose radius is a certain number C_g of wavelengths is a local result and global methods are not particularly useful in proving it. On the other hand, the basic sup-norm estimates of eigenfunctions are most easily proved using the wave equation. It often seems that researchers on eigenfunctions split into two disjoint groups, exclusively using local or global methods. It is likely that many problems require both types of methods. In §5.3 we review the elliptic methods that have been applied to eigenfunctions by Donnelly-Fefferman, F. H. Lin, Nazarov-Sodin, Colding-Minicozzi, and many others.

- Quantum analogues of classical dynamical methods for ergodic or completely integrable systems. For instance, Birkhoff normal forms are local normal forms on both the classical and quantum level around invariant sets such as closed geodesics, which are useful in study concentration on submanifolds.
- L^p bounds on eigenfunctions and their source in the global dynamics of the geodesic flow.
- Restriction theorems for eigenfunctions under dynamical assumptions mainly in the ergodic setting.
- Nodal geometry in the complex domain. Considerable space is devoted to analytic continuation of eigenfunctions of Laplacians of real analytic Riemannian manifolds to the complexification of the manifold. The rationale for analytic continuation is that the nodal sets are better behaved and easier to study in the complex domain than the real domain. From the viewpoint of quantum mechanics, both the real and complex domains are equally good representations.

0.1. Organization

Let us go over the sequence of events in these lectures and explain what is and what is not contained in them and what is the logic of the presentation.

We introduce the subject of eigenfunctions in terms of vibrating membranes and quantum energy eigenstates. The rich phenomenology of examples developed over the last two hundred years is rapidly surveyed. In Chapter 3 we give an overview of the principal new results that will be discussed in detail. The model surfaces of constant curvature are introduced in §4. Harmonic analysis begins with the Euclidean eigenfunctions $e^{i\langle x, \mathbf{k}\rangle}$ on $\mathbb{R}^n$ or T^n, yet they have very unusual properties compared to eigenfunctions on other Riemannian manifolds. The eigenfunctions of S^2 illustrate virtually the entire range of behavior of eigenfunctions of any Riemannian metric with regard to size and concentration. On the other hand, they are restrictions of harmonic polynomials on $\mathbb{R}^3$ and their nodal sets are potentially tamer than for a general C^∞ metric. Eigenfunctions of hyperbolic surfaces $\mathbb{H}^2/\Gamma$ come next. They are the material of quantum chaos and are the subject of intense investigation over the last 30 years. In §5-5.3 the local elliptic analysis of eigenfunctions is surveyed. This leads §6 on the wave equation on a Riemannian manifold and the Hadamard-Riesz construction of parametrices. This construction parallels the Minakshisundaraman-Pleijel parametrix construction for heat kernels. In some ways, the original presentations of Hadamard and Riesz remain the best expositions, in particular in their presentations of the convergence of the parametrix construction in the real analytic case. It was a precursor to the Fourier integral operator theory, which is rapidly reviewed in §2.1, §2.5, §7. As mentioned above, this material is contained in many other references and is principally used to establish notation. In §8.2 classical results on the pointwise and local Weyl laws are reviewed, and the results presented give the universal sup-norm estimates on eigenfunctions and their gradients. The author is not aware of a proof of such estimates using elliptic estimates. Geometric analysts who are more familiar with elliptic estimates might want to compare their methods to the small time wave equation methods used in the proofs. In §9, the asymptotics and limits of matrix elements $\langle A\varphi_j, \varphi_j\rangle$ of pseudo-differential operators with respect to eigenfunctions are introduced. Matrix elements are the fundamental quantities in quantum mechanics. They are quadratic in the eigenfunctions and thus are related to energy estimates. There exist some results on multilinear eigenfunction estimates but they are not covered in these lectures. In §9.5 the basic facts about quantum ergodic systems are reviewed. At this point in the lecture notes, the global long-time dynamics of the geodesic flow takes over as the dominant player. In §11 some parallel results for quantum integrable systems are presented. At this time there exist only a few results on quantizations of mixed systems, and despite the great interest in mixed systems we do not present these results but only record the existence of several articles devoted to them. L^p norms of eigenfunctions are studied in §10. Sogge's books [**So1, So2**] also concern L^p norms but the material presented here contains both less and more on them. Less, because the universal Sogge estimates are not presented, and more because the more advanced results due to Sogge and the author are given in some detail. In §11.6, L^p norms of eigenfunctions in the quantum integrable case are reviewed. One of the motivations to include this material is the belief that such QCI eigenfunctions are extremals for L^p norms and restrictions of eigenfunctions. Although it is very relevant the restriction theorems of Burq-Gérard-Tzvetkov are not discussed here. Rather we turn to quantum ergodic restriction theorems in §12.21. They have proved useful in the study of nodal sets and that is the main topic for the rest of the lectures. Nodal sets in the real domain are discussed in §13, in particular bounds on hypersurface volumes and counting nodal domains. Starting in §14, the analytic continuation of eigenfunctions to Grauert tubes and their complex zeros are studied. Complex nodal sets and their intersections with complexified geodesics are studied in §14.30. Use of the complexified wave kernel gives a simplified proof of the Donnelly-Fefferman upper bound on the hypersurface measure of nodal sets. The lower bound seems to be disconnected from global methods. In §14.33, Alex Brudnyi has contributed a simplified proof of the Donnelly-Fefferman lower bound. In §14.37, the author's results on equidistribution of complexified nodal sets in the ergodic case are presented. There are parallel results in the completely integrable case which are still in progress. Other results in this section are those of John Toth and the author giving upper bounds on numbers of intersection points of nodal lines with curves in dimension two.

0.2. Topics which are not covered

There are many important topics on eigenfunctions which are not discussed in these lecture notes, due to time and length constraints. A more comprehensive treatment of eigenfunctions would include the following topics:

- Arithmetic quantum chaos. These lecture notes are devoted to PDE methods and therefore we do not get into the special methods available for Hecke-Maass forms on arithmetic quotients. The sharpest results on L^p norms or nodal sets of eigenfunctions are for these special joint eigenfunctions of Δ and of Hecke operators. One might compare their special

properties to those of joint eigenfunctions of a quantum integrable system although they are much more complicated and the dynamics is in the opposite chaotic regime.
- Entropy of quantum limits. The breakthrough results of Anantharaman and the subsequent work of Anantharaman-Nonnenmacher and Rivière are very relevant to the theme of these lectures.
- General L^p bounds on restrictions of eigenfunctions, multilinear estimates and Kakeya-Nikodym bounds.
- Gaussian random spherical harmonics and more general random linear combinations of eigenfunctions.
- Spectral and scattering theory for non-compact complete Riemannian manifolds.

0.3. Topics which are double covered

It is impossible to avoid overlaps with the author's prior expository articles, such as the article on local and global analysis of eigenfunctions [**Ze3**], on nodal sets [**Ze6**] or Park City Lecture notes [**Ze7**] and other expository articles on eigenfunctions and nodal sets.

Another double-coverage is with regard to cited references. Each chapter has a bibliography of the references cited in it. Many references are cited in several chapters. Although this results in duplicated references it seems preferable to only listing hundreds of references at the end of the lecture notes.

0.4. Notation

Notation regarding eigenvalue parameters is given in §1.2 and notation for geometric and dynamical objects is given in §2.

Acknowledgments

Thanks go to Peter Hislop and Peter Perry for organizing the CBMS lecture series at the University of Kentucky. Both author and reader will thank Alex Brudnyi for giving alternative arguments to the Donnelly-Fefferman lower bound in §14.33. Thanks also to Hans Christianson, J. Jung, C.D. Sogge, J.A. Toth for collaboration and for an infinite number of discussions on the topics discussed here. My main thanks go to Robert Chang for reading the text and for suggesting many corrections. Robert also did almost all the technical support in producing the book.

CHAPTER 1

Introduction

In this chapter, we introduce the main objects and themes of this monograph. In particular we introduce the quantum mechanical interpretation of eigenfunctions and their time evolution. At the end we outline the topics emphasized in later chapters.

1.1. What are eigenfunctions and why are they useful

Eigenfunctions of the Laplacian first arose in the study of vibrating plates and membranes. The equations of motion of a vibrating membrane Ω are given by the mixed initial value and Dirichlet problem for the profile $u(t,x)$ on $\mathbb{R} \times \Omega$:

$$\left\{ \begin{array}{ll} \left(\frac{\partial^2}{\partial t^2} - \Delta\right) u(t,x) = 0; & \\ u(0,x) = \varphi_0(x), & \frac{\partial \varphi}{\partial t} u(0,x) = 0; \\ u(t,x) = 0, & x \in \partial\Omega. \end{array} \right. \tag{1.1}$$

If φ_λ is a Dirichlet eigenfunction, i.e., a solution of the *Helmholtz equation*

$$(\Delta + \lambda^2)\varphi_\lambda = 0 \quad \text{and} \quad \varphi_\lambda|_{\partial\Omega} = 0, \tag{1.2}$$

then one obtains a periodic solution of the wave equation on $\mathbb{R} \times \Omega$:

$$u_\lambda(t,x) = (\cos t\lambda)\varphi_\lambda(x). \tag{1.3}$$

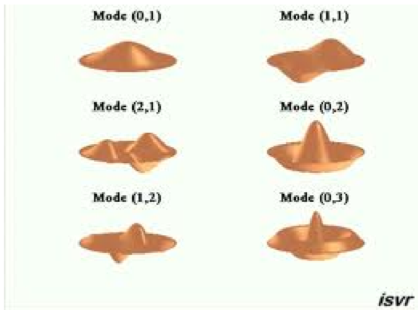
Thus, φ_λ represents the profile of a periodic vibration, i.e., a mode of vibration. In our notation, $-\lambda^2$ is the eigenvalue or energy and λ denotes the frequency.

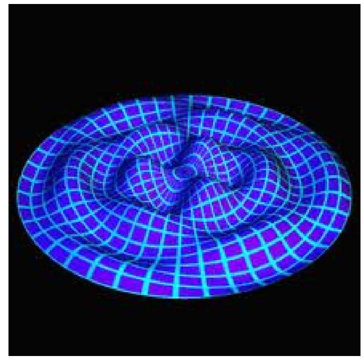
When the domain Ω is compact, the Laplacian has a discrete spectrum with finite multiplicities. We write

$$\lambda_0^2 = 0 < \lambda_1^2 \leq \lambda_2^2 \uparrow \infty \tag{1.4}$$

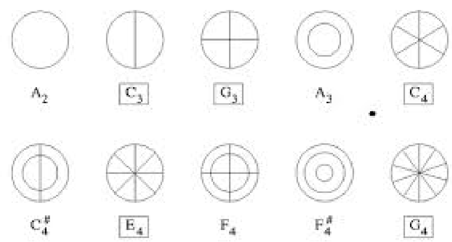
for the ordered sequence of eigenvalues, repeated according to multiplicity. The corresponding set $\{\varphi_j\}$ of eigenfunctions form an orthonormal basis of $L^2(\Omega)$ with respect to the inner product $\langle u, v \rangle = \int_\Omega u\bar{v}\, dV$, where dV is the volume density:

$$(\Delta + \lambda_j^2)\varphi_j = 0 \quad \text{and} \quad \langle \varphi_j, \varphi_k \rangle := \int_M \varphi_j \varphi_k \, dV = \delta_{jk} \tag{1.5}$$





The *nodal set* (or *zero set*) of φ_λ gives the positions at which the vibrating membrane is still. The nodal patterns have been studied since the time of Chladni (ca. 1800).



A natural generalization is to consider eigenfunctions of the Laplacian on Riemannian manifolds (M, g), with or without boundary. The Laplacian of (M, g) is given locally by

$$\Delta_g := \frac{1}{\sqrt{g}} \sum_{i,j=1}^{n} \frac{\partial}{\partial x_i} \left(g^{ij} \sqrt{g} \frac{\partial}{\partial x_j} \right), \tag{1.6}$$

replacing the Euclidean Laplacian Δ above. Here, $g_{ij} = g(\frac{\partial}{\partial x_i}, \frac{\partial}{\partial x_j})$, $[g^{ij}]$ is the inverse matrix to $[g_{ij}]$ and $\sqrt{g} := \sqrt{\det[g_{ij}]}$. Since g is usually understood, in subsequent chapters we suppress the dependency of the metric by writing $\Delta_g = \Delta$.

It follows that on a Riemannian manifold (M, g), the eigenvalue problem (1.2) has the form

$$(\Delta_g + \lambda^2)\varphi_\lambda = 0. \tag{1.7}$$

If M has a non-empty boundary ∂M then we impose the standard Dirichlet or Neumann boundary conditions. If M is compact case, there exists an orthonormal basis $\{\varphi_j\}_{j \geq 0}$ of $L^2(M)$ of eigenfunctions,

$$\Delta_g \varphi_j = -\lambda_j^2 \varphi_j \quad \text{and} \quad \langle \varphi_j, \varphi_k \rangle_{L^2(M)} := \int_M \varphi_j \varphi_k \, dV_g = \delta_{jk} \tag{1.8}$$

and as above the eigenvalues

$$0 = \lambda_0^2 \leq \lambda_1^2 \leq \lambda_2^2 \uparrow \infty \tag{1.9}$$

are repeated according to multiplicity. As in the case of a vibrating membrane, the eigenfunctions φ_λ represent modes of vibration of M.

An equivalent definition of the Laplacian is that it is the operator corresponding to the quadratic form

$$D(f) = \int_M \|df\|^2 \, dV_g = \langle df, df \rangle \tag{1.10}$$

in the sense that

$$D(f) = -\langle \Delta f, f \rangle. \tag{1.11}$$

1.2. Notation for eigenvalues

We often parametrize spectral quantities in terms of the frequencies λ_j, which are eigenvalues of the first order elliptic pseudo-differential operator $\sqrt{-\Delta}$, rather than by the eigenvalues λ_j^2. We warn the reader that many others denote the $(-\Delta)$-eigenvalues by λ_j and the frequencies by $\sqrt{\lambda_j}$.

Regarding eigenfunctions, we write φ_j when the eigenfunction is part of an orthonormal basis as in (1.5) and φ_λ to denote any eigenfunction of eigenvalue λ^2 with $\|\varphi_\lambda\|_{L^2} = 1$. The notation is ambiguous since the eigenvalue λ^2 may not be simple, i.e., the eigenspace may have dimension greater than one, but it is a useful notation when we only care about the dependence on the eigenvalue.

There are two reasons to emphasize frequencies over eigenvalues. One is to simplify the notation by getting rid of square roots. The other is to relate frequencies to Planck's constant. (Planck's constant is also written as $\hbar = \frac{h}{2\pi}$. We use the two notations interchangeably.)

$$h_j = \lambda_j^{-1}, \tag{1.12}$$

which is conceptually important because the high frequency asymptotics of eigenvalues and eigenfunctions is equivalent to the semiclassical asymptotics $h \to 0$. We also use the notation φ_h for φ_λ where it is understood that $h = \lambda^{-1}$ as in (1.12). We sometimes denote an orthonormal basis by φ_{h_j}. Thus we write the Helmholtz equation in semiclassical notation as

$$\Delta \varphi_h = -h^{-2} \varphi_h \iff (h^2 \Delta - 1) \varphi_h = 0. \tag{1.13}$$

As the semiclassical notation suggests, a compelling motivation to study eigenfunctions comes from their role in quantum mechanics.

1.3. Weyl's law for $(-\Delta)$-eigenvalues

When M is a compact manifold, Weyl's law counts the number of eigenvalues of Δ_g. Let

$$N(\lambda) := \{j \colon \lambda_j \leq \lambda\}. \tag{1.14}$$

Weyl's law states the following:

THEOREM 1.1. *If (M, g) is a compact Riemannian manifold of dimension m, then*

$$N(\lambda) = C_m \operatorname{Vol}(M, g) \lambda^m + O(\lambda^{m-1}), \tag{1.15}$$

where $C_m = \operatorname{Vol}(B_1)$ is a dimensional constant, the volume of the unit ball in $\mathbb{R}^m$.

Here, we say $R(\lambda) = O(\lambda^r)$ if there exists a constant C independent of λ so that $R(\lambda) \leq C\lambda^r$ as $\lambda \to \infty$. We also write $R(\lambda) = o(\lambda^r)$ if $R(\lambda) \leq \varepsilon \lambda^r$ as $\lambda \to \infty$ for any $\varepsilon > 0$.

1.4. Quantum Mechanics

Much of quantum mechanics is concerned with the eigenvalue problem for the Schrödinger equation

$$\Big(-\frac{h^2}{2}\Delta + V\Big)\psi = E\psi. \tag{1.16}$$

Here, V stands for multiplication by the potential $V \in C^\infty(M)$, and (as above) h is Planck's constant, a very small parameter. When $V = 0, E = 1$ and $h = \lambda^{-1}$, (1.16) specializes to (1.2). Thus we think of the limit as $\lambda \to \infty$ in (1.2) as the semiclassical limit $h \to 0$.

The Schrödinger eigenvalue problem in quantum mechanics resolves a puzzle about the stability of atoms. Before quantum mechanics, a hydrogen atom was roughly pictured as a 2-body planetary system, i.e., as an electron orbiting the nucleus centered at the origin $0 \in \mathbb{R}^3$ according to Kepler's laws. The orbits are projections to configuration space $\mathbb{R}^3$ of the phase space orbits of the classical Hamiltonian flow defined by Hamilton's equations

$$\begin{cases} \dfrac{dx_j}{dt} = \dfrac{\partial H}{\partial \xi_j}, \\ \dfrac{d\xi_j}{dt} = -\dfrac{\partial H}{\partial x_j}, \end{cases} \tag{1.17}$$

where the Hamiltonian

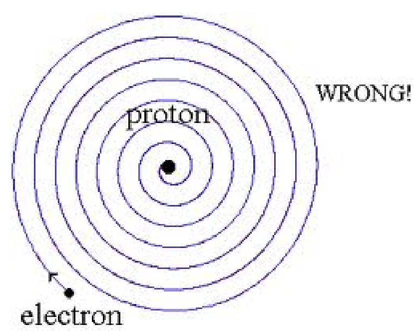
$$H(x,\xi) = \frac{1}{2}|\xi|^2 - \frac{1}{|x|} : T^*\mathbb{R}^3 \to \mathbb{R} \tag{1.18}$$

is the total Newtonian kinetic plus Coulomb potential energy function $V(x) = -\frac{1}{|x|}$ on phase space, the cotangent bundle $T^*\mathbb{R}^3$ of the configuration space $\mathbb{R}^3$. We denote the Hamiltonian (geodesic) flow by

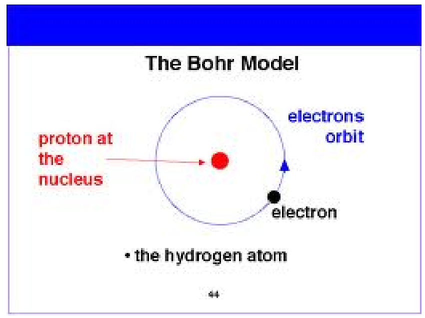
$$G^t(x,\xi) = \exp tX_H, \tag{1.19}$$

where X_H is the Hamilton vector field and $\exp tX$ is the general notation for the flow of X.

But this model cannot be right: the electron would radiate energy and spiral into the nucleus.



The Bohr model (1913) of "old quantum theory" proposed that the electron can only occupy special stable orbits defined by Bohr-Sommerfeld "quantization conditions."



However this theory is too specialized. It relies on the special structure of the orbits of the Coulomb problem, in particular the (hidden) symmetry that makes all of the orbits periodic. It does not extend in any clear way to more complicated atoms such as Helium or even to the hydrogen atom in an electric or magnetic field. In the article *Quantisierung als Eigenwertproblem*, Annalen der Physik (1926), Schrödinger [**Sch1**] proposed to model the electron by a wave function $\psi(x) \in L^2(\mathbb{R}^3)$ with the states of energy $E_j(\hbar)$ solving the eigenvalue problem (as in (1.16))

$$\hat{H}\psi_{\hbar,j} := \left(-\frac{\hbar^2}{2}\Delta + V\right)\psi_{\hbar,j} = E_j(\hbar)\psi_{\hbar,j}, \tag{1.20}$$

for the Schrödinger operator $\hat{H}$, where $\Delta = \sum_j \frac{\partial^2}{\partial x_j^2}$ is the Laplacian and V is the potential, a multiplication operator on $L^2(\mathbb{R}^3)$. Here $\{\psi_{\hbar,j}\}$ denote an orthonormal basis of eigenfunctions with eigenvalues $E_j(\hbar)$ in non-decreasing order.

Classically, the particle evolves according to Hamilton's equations (1.17) with $H(x,\xi) = \frac{1}{2}|\xi|^2 + V(x)$. The Hamiltonian is constant along Hamilton orbits and therefore the orbits lie on level sets $\{H = E\}$ of H. The projection $\{x \colon V(x) \leq E\}$ of these level sets to the configuration space $\mathbb{R}^n$ is known as the *allowed region*; a classical particle cannot enter the *forbidden region*, which is the set $\{x \colon V(x) > E\}$.

Quantum mechanics thus replaces the classical mechanics of Hamilton's equations with linear algebra (an eigenvalue problem). The time evolution of an energy state is given by

$$U_\hbar(t)\psi_{\hbar,j} = e^{-i\frac{t}{\hbar}(-\frac{\hbar^2}{2}\Delta+V)}\psi_{\hbar,j} = e^{-i\frac{tE_j(\hbar)}{\hbar}}\psi_{\hbar,j}. \tag{1.21}$$

Throughout it is assumed that eigenfunctions are L^2 normalized,

$$\int |\psi_{\hbar,j}|^2\, dV = 1, \tag{1.22}$$

so that the state $\psi_{\hbar,j}$ defines a probability amplitude, i.e., its modulus square is a probability measure with

$$|\psi_{\hbar,j}(x)|^2 dx = \text{the probability density of finding the particle at } x\ . \tag{1.23}$$

This probability density is not concentrated in the classically allowed region $\{V \leq E\}$, i.e., a quantum particle has a positive probability of going into the forbidden region $\{x \colon V(x) > E\}$. The only *observable quantities* are the matrix elements

$$\langle A\psi_{\hbar,j}, \psi_{\hbar,j}\rangle = \int \psi_{\hbar,j} A\psi_{\hbar,j}(x)\, dV \tag{1.24}$$

of observables (A is a self adjoint operator). Under the time evolution (1.21), the factors of $e^{-i\frac{tE_j(\hbar)}{\hbar}}$ cancel and so the particle evolves as if "stationary," i.e., observations of the particle are independent of the time t.

Quantum mechanics resolves the puzzle of how the electron can be moving and stationary at the same time. But it also replaces the geometric (classical mechanical) Bohr model of classical orbits with eigenfunctions (1.20), which are not geometric objects and which are difficult to visualize. They are very complicated functions on high dimensional spaces. How does one reconcile the classical picture of orbits with the quantum picture of eigenfunctions, as stationary energy states of atoms? In the semiclassical limit $h \to 0$, the quantum physics should tend to the classical physics, and the eigenfunctions should be related to the classical orbits.

The Bohr model proposed a close relation between the quantum mechanics of a hydrogen atom and the classical mechanics of the corresponding classical Hamiltonian $H(x, \xi) = \frac{1}{2}|\xi|^2 + V(x)$. Can we use classical mechanics to analyze shapes and sizes of quantum eigenstates?

1.5. Dynamics of the geodesic or billiard flow

The (homogeneous) geodesic flow

$$G^t : T^*M\backslash 0 \to T^*M\backslash 0 \tag{1.25}$$

on the punctured cotangent bundle $T^*M\backslash 0 = \{(x, \xi) \in T^*M : \xi \neq 0\}$ is the Hamiltonian flow of the metric norm function

$$H(x, \xi) = \sqrt{\sum_{i,j=1}^{n} g^{ij}(x)\xi_i\xi_j}. \tag{1.26}$$

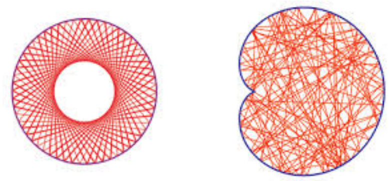
It is free particle motion (with $V = 0$) on M. When $\partial M \neq 0$ the geodesic reflects off the boundary by Snell's law of equal angles. This flow is called the broken geodesic flow or billiard flow.

The Bohr correspondence principle suggests that as $\lambda_j \to \infty$ the asymptotics of eigenfunctions and eigenvalues should be related to dynamics of the geodesic flow. The relations between eigenfunctions and the Hamiltonian flow are best established in two extreme cases: (i) where the Hamiltonian flow is completely integrable on an energy surface, or (ii) where it is ergodic. The hydrogen atom is completely integrable and that is why the special eigenfunctions which are joint eigenfunctions of the Schrödinger operator, the total angular momentum and the z-component of the angular momentum, can be completely understood. These are the eigenfunctions whose images are graphed here. Integrable systems are rare but important in that many of the known results are obtained by perturbing

the integrable case. The quantum integrable case is discussed in some detail in §11.

Ergodic (or more chaotically mixing) dynamical systems are more difficult than integrable systems because they are not explicitly solvable. However, as with the law of large numbers or central limit theorem in probability theory, chaos induces a kind of symmetry or uniformity which makes it possible to prove results by indirect calculations and results.

The extremes are illustrated below in the case of (a) billiards on rotationally invariant annulus, (b) chaotic billiards on a cardioid.

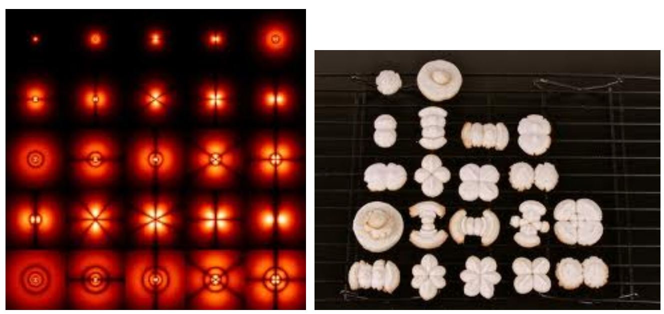


(a) (b)

A typical trajectory in the case of ergodic billiards is uniformly distributed, while all trajectories are quasi-periodic in the integrable case.

1.6. Intensity plots and excursion sets

There are several ways to 'picture' an eigenfunction and the probability density (1.23) that it defines. One vivid kind of picture of a hydrogen atom is an *intensity plot* which darkens in the regions where $|\varphi_j(x)|^2$ is large (most probable locations).



The most probable locations are defined by the *excursion sets*

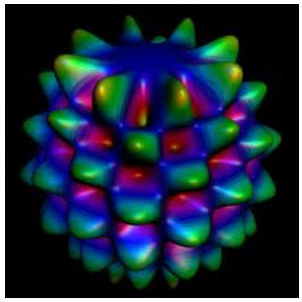
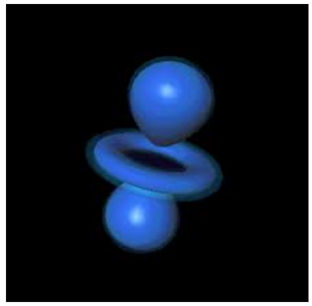
$$\Omega_{j,\hbar,E} = \{x : |\psi_{\hbar,j}(x)|^2 \geq E\}. \tag{1.27}$$

It is particularly interesting to understand the high excursion levels, where $E \simeq A\hbar^{-r}$. One would ideally like to know how the excursion sets are distributed in the semiclassical limit $\hbar \to 0$. How many connected components does it have and what are there shapes and locations? What is the distribution function

$$\mu_{\psi_{\hbar,j}}[E, \infty) := \mathrm{Vol}\{x : |\psi_{\hbar,j}(x)|^2 \geq E\}, \tag{1.28}$$

where Vol denotes the volume measure of E (corresponding to the metric underlying the Laplacian Δ). One could also use (1.23) as the measure to determine the relative proportion of the L^2 mass of the eigenfunction which is concentrated near its top values.

One may also imagine graphing (1.23) over the high-dimensional configuration space and asking for the prominent features of the graph. In the case of a high frequency spherical harmonic on S^2 one may obtain the graph:

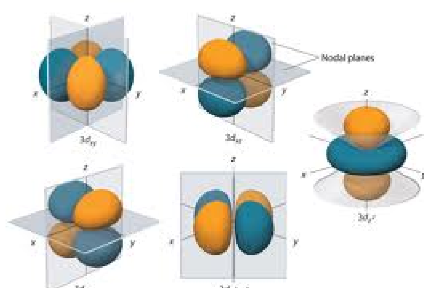
One observes that there are many local maxima near the peak values of the eigenfunction (or its square (1.23)). They appear to be rather uniformly distributed. Can one at least prove that the number of critical points tends to infinity with the eigenvalue (or equivalently as $\hbar \to 0$)? This is known to be false for some eigenfunctions of general Riemannian manifolds. How does the distribution or number of critical points reflect the underlying Hamiltonian dynamics?

These questions are almost completely open and (as in the images) are most accessible for quantum integrable systems. The high excursion sets are the most important sets, but also rather intractable since they involve the distribution function of the eigenfunction. The only case known to the author where the distribution function has been discussed is in the case of toric eigenfunctions on Kähler manifolds [**STZ**]. It is likely that analogous results can be proved for joint eigenfunctions of real integrable systems such as surfaces of revolution or (as in the images) the hydrogen atom eigenfunctions.

1.7. Nodal sets and critical point sets

At the opposite are plots of the nodal hypersurfaces: the zero set

$$\mathcal{N}_{\varphi_\lambda} = \{x \in M : \varphi_\lambda(x) = 0\}. \tag{1.29}$$



These are the points where the probability (density) of the particle's position vanishes. Here, we consider eigenfunctions of the Laplacian Δ_g of a compact Riemannian manifold rather than a general Schrödinger operator.

The *nodal domains* of φ_λ are the connected components Ω_j of $M \backslash \mathcal{N}_{\varphi_\lambda} = \bigcup_{j=1}^{N(\varphi_\lambda)} \Omega_j$. We write

$$N(\varphi_\lambda) := \text{the number of nodal domains of } \varphi_\lambda. \tag{1.30}$$

In [**Br**], J. Brüning (and Yau, unpublished) showed that $\mathcal{H}^1(Z_{\varphi_\lambda}) \geq C_g \lambda$ in the $\dim M = 2$ case, i.e., the length of a nodal line is bounded below by a constant multiple of the frequency for some constant $C_g > 0$. A. Logunov has recently proved the analogous result in all dimensions [**Lo**].

According to Courant's nodal domain theorem [**C**], there exists a universal upper bound for $N(\varphi_j)$:

$$N(\varphi_j) \leq j. \tag{1.31}$$

In terms of order of magnitude, this bound is often obtained: when M is the unit sphere S^2 and φ is a random spherical harmonics, then $N(\varphi_\lambda) \sim c\lambda^2$ holds almost surely for some constant $c > 0$ thanks to Nazarov-Sodin [**NS**] . However, it is known that the bound is not always sharp in terms of order of magnitude. In the chapter on nodal sets ♣ref♣, we will review results of H. Lewy and others that construct sequences of eigenfunctions with a uniform bound on the number of nodal domains. On the other hand, it is very plausible that every compact Riemannian manifold possesses a sequence of eigenfunctions for which the number of nodal domains tends to infinity. In the same chapter, we prove this to be true for almost the entire sequence of eigenfunctions of a non-positively curved surface with concave boundary (for Dirichlet or Neumann boundary conditions) and for negatively curved surfaces possessing an anti-holomorphic isometric involution with dividing fixed point set.

Closely related to nodal sets are the other level sets

$$\mathcal{N}^a_{\varphi_j} = \{x \in M \colon \varphi_j(x) = a\} \tag{1.32}$$

and the sublevel sets

$$\{x \in M \colon |\varphi_j(x)| \leq a\}. \tag{1.33}$$

The zero level is distinguished since the symmetry $\varphi_j \to -\varphi_j$ in the equation preserves the nodal set. A fundamental existence result states that there exists a constant $A > 0$ so that every ball of (M, g) contains a nodal point of any eigenfunction φ_λ if its radius is greater than $\frac{A}{\lambda}$.

Of equal interest is the critical point set

$$\mathcal{C}_{\varphi_j} = \{x \in M \colon \nabla\varphi_j(x) = 0\}. \tag{1.34}$$

The critical point set can be a hypersurface in M. In counting problems it is better to consider the set

$$\mathcal{V}_{\varphi_j} = \{\varphi_j(x) \colon \nabla\varphi_j(x) = 0\} \tag{1.35}$$

of critical values. At this time of writing, there exist (to the author's knowledge) no rigorous upper bounds on the number of critical values except in separation-of-variables situations.

1.8. Local versus global analysis of eigenfunctions

As will be discussed in detail in §5.3, the local study of eigenfunctions uses analysis on small balls of radii $\frac{C}{\lambda}$. One does necessarily assume that the eigenfunctions are global, i.e., that they are eigenfunctions on a global closed manifold without boundary, or that they satisfy Dirichlet or Neumann boundary conditions on a manifold with boundary.

Global harmonic analysis concerns the properties of global eigenfunctions. A key property is that they are eigenfunctions of the evolution operator

$$U(t) = e^{it\sqrt{-\Delta}} \tag{1.36}$$

or propagator

$$U_h(t) = e^{\frac{it}{\hbar}\hat{H}} \tag{1.37}$$

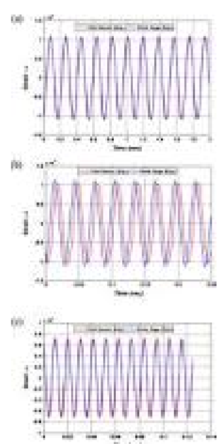
for semiclassical Schrödinger operators.

The goal is then to relate the behavior of eigenfunctions in the semiclassical limit $\lambda_j \to \infty$ or $\hbar \to 0$ to properties of the geodesic flow, or more generally the Hamiltonian flow of $\frac{1}{2}|\xi|^2 + V(x)$ on a fixed energy surface.

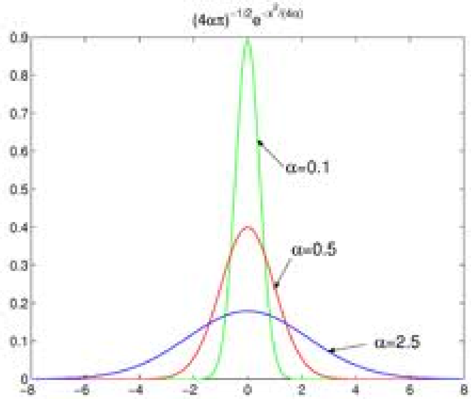
1.9. High frequency limits, oscillation and concentration

The emphasis of these lectures is on high frequency limits $\lambda_j \to \infty$ of zeros sets, norms and mass distribution of sequences of eigenfunctions. For general Schrödinger operators, one studies the semiclassical limit $\hbar \to 0$.

In analogy with polynomials, the degree of a polynomial, resp. the frequency λ of an eigenfunction, is a measure of its "complexity" and the high frequency limit is the large complexity limit. A sequence of eigenfunctions of increasing frequency oscillates more and more rapidly and the problem is to find its "limit shape". In the graph below of $(\sin kx)^2$, the square of the eigenfunction tends in the weak* sense of measures to its mean value. Thus the oscillations smear out to an average value. The eigenfunction sequence itself always tends to zero weakly in L^2.



But it is also possible that a sequence of squares will concentrate on a low-dimensional subset, as in this picture of a sequence of Gaussians tending to the delta function at 0.



In the Riemannian case, there exist sequences of squares of eigenfunctions called *Gaussian beams* which put the two types together: they oscillate more and more rapidly along a geodesic γ and have Gaussian decay in the transverse direction, so that in the limit they tend to a delta function along γ.

1.10. Spectral projections

We now specialize the eigenvalue problem to the setting of Laplacians on compact Riemannian manifold (M, g), i.e. we set $V = 0$. The quantum Hamiltonian is then the Laplacian (1.6).

We denote by $\Pi_{[0,\lambda]} = \Pi_{[0,h^{-1/2}]}$ the spectral projections kernel for the interval $[0, \lambda]$:

$$\Pi_{[0,\lambda]}(x, y) = \Pi_\lambda(x, y) := \sum_{j:\ \lambda_j \leq \lambda} \varphi_j(x)\varphi_j(y). \tag{1.38}$$

It is the Schwartz kernel of the orthogonal projection onto the span of the eigenfunctions with frequencies $\leq \lambda$, and is independent of the choice of orthonormal basis. Sometimes we wish to consider shorter spectral intervals, and then subscript the projection by the relevant interval. An important case is the spectral projections for a short interval:

$$\Pi_{[\lambda,\lambda+1]}(x, y) := \sum_{j:\ \lambda \leq \lambda_j \leq \lambda+1} \varphi_j(x)\varphi_j(y). \tag{1.39}$$

We mainly consider compact Riemannian manifolds in this monograph, although many of the same problems and techniques are valid on non-compact complete Riemannian manifolds. Our main concern is to relate the behavior of eigenvalues/eigenfunctions to the dynamics of the geodesic flow, and the setting of compact Riemannian manifolds is sufficiently rich to illustrate the possible relations. In the non-compact setting, the spectrum is continuous (possibly with embedded eigenvalues) and there is a basis of generalized eigenfunctions. We will briefly consider examples such as the hyperbolic plane and hyperbolic cylinder.

In the non-compact case, it is also natural to study resonances instead of eigenvalues and resonance states instead of eigenfunctions, but the theory is quite different and is not discussed here. See [**DZw**] for a comprehensive exposition. It is also natural to study the scattering phase shifts (eigenvalues) and eigenfunctions of the scattering operator $S(h)$ (see [**GHZ**] for references).

REMARK 1.2. We set the potential V equal to zero for the sake of brevity, but almost everything we do generalizes, often in subtle ways, to Schrödinger operators. The dynamics of geodesic flows is sufficiently rich to exhibit the relations between classical and quantum mechanics. More general Schrödinger operators $-\hbar^2\Delta_g + V$ (or magnetic Schrödinger operators) give rise to significant additional issues that in some cases have barely been explored, such as the behavior of nodal sets in the forbidden region.

Most of the problems stated above are not only unsolved but appear to be completely intractable. They are most accessible when the Schrödinger operator is *completely integrable* on the quantum level in the sense of §1.19, although the problems remain unsolved even in this case. There are many phenomena which show up clearly in numerical plots yet which are far beyond mathematical analysis. Therefore we need to simplify the problems to the point where rigorous results are possible.

1.11. L^p norms

As mentioned above, excursion sets (1.27) are difficult to study in all but the simplest cases (such as the standard spheres or surfaces of revolution). A somewhat more accessible and natural mathematical problem is to study the L^p norms of eigenfunctions (to the pth power),

$$\int_0^\infty t^p d\mu_{\varphi_j}(t) = \int |\varphi_j|^{2p} dV \tag{1.40}$$

as a function of the eigenvalue $E_j(\hbar)$. Here, μ_{φ_j} is the distribution function of $|\varphi_j|^2$ (1.28). Different powers measure different aspects of the intensity plot. Since φ_j is L^2-normalized (1.22), high L^p norms, (e.g., the sup norm $\|\varphi_j\|_\infty = \sup_x |\varphi_j(x)|$) is large when there exist a few very high peaks and are not so large when there exist many relatively shallow peaks. Lower L^p norms are large when the set of rather large values has a large measure. A random spherical harmonic spreads its mass rather evenly around the sphere, and thus has relatively small L^p norms for high p.

General upper bounds on L^p norms of eigenfunctions will be discussed in §10. In §10.4 we discuss the case Riemannian manifolds possessing sequences of eigenfunctions achieving the maximal allowed growth for large p. The case of small p is not understood in general. L^p norms of quantum integrable eigenfunctions are discussed in §11.6.

1.12. Matrix elements and Wigner distributions

L^p norms are 'non-linear' measures of the size of the eigenfunction. The most linear measures are the matrix elements (1.24). For instance, if A is multiplication by the characteristic function $\mathbf{1}_E$ of a Borel set $E \subset M$, one is measuring the L^2-mass

$$\int_E |\varphi_j|^2 \, dV \tag{1.41}$$

of the eigenfunction in the set E, e.g., to determine where it concentrates most. In the case where the boundary ∂E (closure of E minus its interior) has measure zero, there are many results on the semiclassical limits of the L^2 mass. Virtually nothing is known for more general E such as Cantor sets of positive measure. It is possible to allow E to depend on $\hbar$ and to let it shrink at a specified rate as $\hbar \to 0$. Such "small-scale mass" is closely related to L^p norms.

The diagonal matrix elements

$$\rho_j(A) := \langle A\varphi_j, \varphi_j \rangle \tag{1.42}$$

of an observable A (i.e. a bounded operator on $L^2(M)$) are interpreted in quantum mechanics as the expected value of the observable A in the energy state φ_j. The off-diagonal matrix elements

$$\rho_{jk}(A) = \langle A\varphi_i, \varphi_j \rangle, \qquad (j \neq k) \tag{1.43}$$

are interpreted as transition amplitudes. Here, and below, an amplitude is a complex number whose modulus square is a probability.

There is a special class of observables A for which it is possible to study semiclassical limits of matrix elements (1.42), namely (various kinds of) pseudo-differential operators $Op_\hbar(a) = a(x, \hbar D)$. Such operators are understood as 'quantizations' of classical observables, namely functions $a(x, \xi)$ on phase space. For instance, one may let $a(x, \xi) = \mathbf{1}_E(x, \xi)$ where $E \subset T^*M$ is a nice Borel set. Then (1.42) measures the phase space mass of the eigenfunction in E, i.e., the probability

$$\langle \mathbf{1}_E \varphi_j, \varphi_j \rangle \tag{1.44}$$

that its (position, momentum) $= (x, \xi)$ belong to E. Just as φ_j determines the probability measure (1.23) on configuration space, so it also induces probability measures on phase space.

If we fix the quantization $a \to \mathrm{Op}_\hbar(a)$, then the matrix elements can be represented by *Wigner distributions*. In the diagonal case, we define $W_k \in \mathcal{D}'(T^*M)$ by

$$\int_{T^*M} a \, dW_k := \langle \mathrm{Op}_\hbar(a)\varphi_k, \varphi_k \rangle. \tag{1.45}$$

Here, we are using semiclassical pseudo-differential operators (see [**DSj, Zw**]). If we use homogeneous pseudo-differential operators, the Wigner distributions may be defined as distributions on the unit co-sphere bundle S^*M.

The basic compactness theorem regarding the sequence of probability measures (1.23) or their microlocal lifts is simply the compactness of probability measures in the weak* topology. As the name suggests, weak* convergence is a very weak type of convergence and it is difficult to determine many concrete properties of eigenfunctions even from knowledge of the limit measures.

1.13. Egorov's theorem

Egorov's theorem is the precise statement of the correspondence between the Heisenberg time evolution $U_t A U_t^*$ of an observable A and the time evolution of the classical observable (its symbol) $\sigma_A \circ G^t$, where G^t is the corresponding Hamiltonian (geodesic) flow. It states that if $A \in \Psi^0(M)$ (i.e., A is a pseudo-differential operator of order zero), then

$$U_\hbar^t \, \mathrm{Op}_h(a) U_\hbar^{-t} - \mathrm{Op}_h(a \circ G^t) \in \Psi_h^{-1}(M), \tag{1.46}$$

i.e., the difference is a pseudo-differential operator of order -1. In semiclassical notation, order -1 means of order $\mathcal{O}(\hbar)$.

1.14. Eherenfest time

The aim in quantum chaos is to obtain information about the high energy asymptotics as $\lambda_j \to \infty$ of eigenvalues and eigenfunctions by connecting information about U^t and G^t. The connection often comes from Egorov's theorem (Theorem 1.46). But to use the hypothesis that G^t is ergodic or chaotic, one needs to exploit the connection as $\hbar \to 0$ and $t \to \infty$. The difficulty in quantum chaos is that the approximation of U^t by G^t is only a good one for t less than the *Eherenfest time*

$$T_E = \frac{\log |\hbar|}{\lambda_{\max}}, \tag{1.47}$$

where $\lambda_{\max}$ is the so-called maximal Lyapunov exponent.

Roughly speaking, the idea is that the evolution of a well constructed "coherent" quantum state or particle is a moving lump that "tracks along" the trajectory of a classical particle up to time T_E and then slowly falls apart and stops acting like a classical particle. Numerical studies of long time dynamics of wave packets are given in works of E.J. Heller [**He1, He2**] and rigorous treatments are in Bouzouina-Robert [**BoR**], Combescure-Robert [**CR**] and Schubert [**S**].

The basic result expressed in semiclassical notation is that there exists $\Gamma > 0$ such that

$$\|U_\hbar^t \operatorname{Op}(a) U_\hbar^{-t} - \operatorname{Op}(a \circ G^t)\| \leq C\hbar e^{t\Gamma}. \tag{1.48}$$

The exponential growth rate in t has long been known to be the essential stumbling block to precise localization in the spectrum. Thus, one only expects good joint asymptotics as $\hbar \to 0$, $t \to \infty$ for $t \leq T_E$. As a result, one can only exploit the approximation of U^t by G^t for the relatively short time T_E.

1.15. Weak* limit problem

There are two (equivalent) ways to state the weak* limit problem: (i) in terms of quantum statistical mechanical *states* on the algebra of observables, or (ii) in terms of Wigner distributions (or microlocal lifts, microlocal defect measures, etc.). The first is more abstract or at least less PDE oriented but is useful in not requiring any choice of quantization of classical observables. The second is more concrete.

The diagonal matrix elements define linear functionals (1.42) on Ψ^0. We observe that $\rho_j(I) = 1$, that $\rho_j(A) \geq 0$ if $A \geq 0$ and that

$$\rho_k(U^t A U^{-t}) = \rho_k(A). \tag{1.49}$$

Indeed, if $A \geq 0$ then $A = B^* B$ for some $B \in \Psi^0$ and we can move B^* to the right side. Similarly (1.49) is proved by moving U_t to the right side and using the fact that the eigenvalues of U_t are of modulus one. In quantum statistical mechanics, these properties are summarized by saying that ρ_j is an *invariant state* on the algebra Ψ^0, or more precisely, on its closure in the operator norm. An invariant state is the analogue in quantum statistical mechanics of an invariant probability measure.

We denote by $\mathcal{M}_I$ the convex set of invariant probability measures for the geodesic flow. Further, we say that a measure is time-reversal invariant if it is invariant under the anti-symplectic involution $(x, \xi) \to (x, -\xi)$ on T^*M. We denote the time-reversal invariant elements of $\mathcal{M}_I$ by $\mathcal{M}_I^+$.

PROPOSITION 1.3. *Any weak limit of the sequence $\{\rho_j\}$ on Ψ^0 is a time-reversal invariant, G^t invariant probability measure on S^*M, i.e. is an element of $\mathcal{M}_I^+$.*

PROOF. For any compact operator K, $\langle K\varphi_j, \varphi_j\rangle \to 0$. Hence, any limit of $\langle A\varphi_j, \varphi_j\rangle$ is equally a limit of $\langle (A+K)\varphi_j, \varphi_j\rangle$. By the norm estimate, the limit is bounded by $\inf_K \|A+K\|$ (the infimum taken over compact operators). Hence any weak limit is bounded by a constant times the sup norm $\|\sigma_A\|_{L^\infty}$ of the symbol σ_A of A and is therefore continuous on $C(S^*M)$. It is a positive functional since each ρ_j is and hence any limit is a probability measure. By Egorov's theorem and the invariance of ρ_j, any limit of $\rho_j(A)$ is a limit of $\rho_j(\mathrm{Op}(\sigma_A \circ G^t))$ and hence the limit measure is invariant. It is also time-reversal when the eigenfunctions are real-valued, i.e., complex conjugation invariant. □

The Wigner distributions (also called microlocal lifts) dW_k defined by (1.45) of course depend on the choice of $\mathrm{Op}(a)$. If a is chosen to be homogeneous of degree 0 on $T^*M - \{0\}$ (the zero section) then one can arrange that $dw_k \in \mathcal{D}'(S^*M)$. In the semiclassical setting one deals with non-homogeneous symbols. Eigenfunctions localize on the 'energy surface' $\{H = 1\}$, i.e., on the unit co-sphere bundle S_g^*M in the case of the Laplacian, and the corresponding microlocal lifts concentrate there.

PROBLEM 1.4. Determine the set $\mathcal{Q}$ of 'quantum limits,' i.e., weak* limit points of the sequence $\{dW_k\}$.

The set $\mathcal{Q}$ is independent of the definition of quantization $a \to \mathrm{Op}(a)$. The simplest examples are the exponentials on a flat torus $\mathbb{R}^m/\mathbb{Z}^m$. By definition of pseudo-differential operator, $Ae^{2\pi i\langle k,x\rangle} = a(x,k)e^{2\pi i\langle k,x\rangle}$ where $a(x,k)$ is the complete symbol. Thus,

$$\langle Ae^{2\pi i\langle k,x\rangle}, e^{2\pi i\langle k,x\rangle}\rangle = \int_{\mathbb{R}^n/\mathbb{Z}^n} a(x,k)dx \sim \int_{\mathbb{R}^n/\mathbb{Z}^n} \sigma_A\left(x, \frac{k}{|k|}\right) dx. \tag{1.50}$$

A subsequence $e^{2\pi i\langle k_j,x\rangle}$ of eigenfunctions has a weak limit if and only if $\frac{k_j}{|k_j|}$ tends to a limit vector ξ_0 in the unit sphere in $\mathbb{R}^n$. In this case, the associated weak* limit is $\int_{\mathbb{R}^n/\mathbb{Z}^n} \sigma_A(x,\xi_0)dx$, i.e., the delta-function on the invariant torus $T_{\xi_0} \subset S^*M$ for G^t, defined by the constant momentum condition $\xi = \xi_0$. The eigenfunctions are said to localize on this invariant torus. Given ξ_0, we can always define a sequence k_j so that $\frac{k_j}{|k_j|} \to \xi_0$, and thus, every invariant torus measure arises as a quantum limit.

In general, there are many possible limit measures. The most important are:

(1) Normalized Liouville measure μ_L. In fact, the functional ω of integration against normalized Liouville measure is also a state on Ψ^0 for the reason explained above. A subsequence $\{\varphi_{j_k}\}$ of eigenfunctions is considered diffuse if $\rho_{j_k} \to \omega$.
(2) A periodic orbit measure μ_γ defined by $\mu_\gamma(A) = \frac{1}{L_\gamma}\int_\gamma \sigma_A ds$ where L_γ is the length of γ. A sequence of eigenfunctions for which $\rho_{k_j} \to \mu_\gamma$ obviously concentrates (or strongly 'scars') on the closed geodesic.
(3) A finite sum of periodic orbit measures.
(4) A delta-function along an invariant Lagrangian manifold $\Lambda \subset S^*M$. The associated eigenfunctions are viewed as *localizing* along Λ.
(5) A more general measure which is singular with respect to $d\mu$.

All of these possibilities arise as (M, g) varies among Riemannian manifolds. Indeed, the standard sphere provides an extreme example (see [**JZ**])

THEOREM 1.5. *For the standard round sphere* S^n, $\mathcal{Q} = \mathcal{M}_I^+$.

In the case where $\rho_{k_j} \to \omega$, the corresponding eigenfunctions become uniformly distributed on the energy surface S^*M. By testing against multiplication operators, one gets

$$\frac{1}{\mathrm{Vol}(M)} \int_E |\varphi_{k_j}(x)|^2 d\,\mathrm{Vol} \to \frac{\mathrm{Vol}(E)}{\mathrm{Vol}(M)} \tag{1.51}$$

for any measurable set E whose boundary has measure zero. In the interpretation of $|\varphi_{k_j}(x)|^2 d\,\mathrm{Vol}$ as the probability density of finding a particle of energy $\lambda_{k_j}^2$ at x, this says that the sequence of probabilities tends to uniform measure. However, $\rho_{k_j} \to \omega$ is much stronger since it says that the eigenfunctions become uniformly distributed on S^*M and not just on the configuration space M. For instance, on the flat torus $\mathbb{R}^n/\mathbb{Z}^n$, the standard exponentials $e^{2\pi i \langle k, x\rangle}$ satisfy $|e^{2\pi i \langle k, x\rangle}|^2 = 1$, and are thus uniformly distributed in configuration space. On the other hand, as seen above, in phase space they localize on invariant Lagrange tori in S^*M.

The flat torus is a model of a completely integrable system, on both the classical and quantum levels. On the other hand, if the geodesic flow is ergodic one would expect the eigenfunctions to be diffuse in phase space. The statement that the all eigenfunctions are diffuse, i.e., $\mathcal{Q} = \{\omega\}$, is known as *quantum unique ergodicity*. It will be discussed in §1.18.

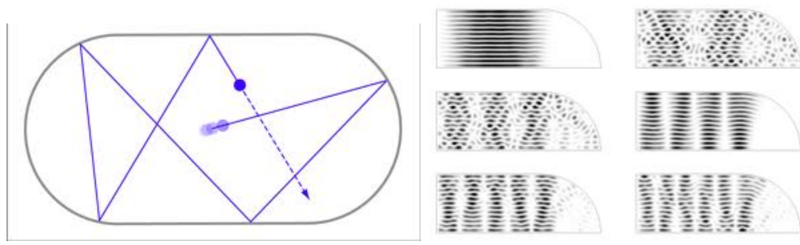
Off-diagonal matrix elements (1.43) are also important as transition amplitudes between states. They no longer define states since $\rho_{jk}(I) = 0$, are no longer positive, and are no longer invariant. Indeed, $\rho_{j,k}(U_t A U_t^*) = e^{it(\lambda_j - \lambda_k)} \rho_{jk}(A)$, so they are eigenvectors of the automorphism $\alpha_t(A) = U_t A U_t^*$. A sequence of such matrix elements cannot have a weak limit unless the spectral gap $\lambda_j - \lambda_k$ tends to a limit $\tau \in \mathbb{R}$. In this case, by the same discussion as above, any weak limit of the functionals ρ_{jk} will be a time-reversal invariant eigenmeasure of the geodesic flow which transforms by $e^{i\tau t}$ under the action of G^t. Examples of such eigenmeasures are orbital Fourier coefficients $\frac{1}{L_\gamma} \int_0^{L_\gamma} e^{-i\tau t} \sigma_A(G^t(x, \xi)) dt$ along a periodic orbit. Here $\tau \in \frac{2\pi}{L_\gamma}\mathbb{Z}$. We denote by $\mathcal{Q}_\tau$ such eigenmeasures of the geodesic flow. Problem 1.4 has the following extension to off-diagonal elements:

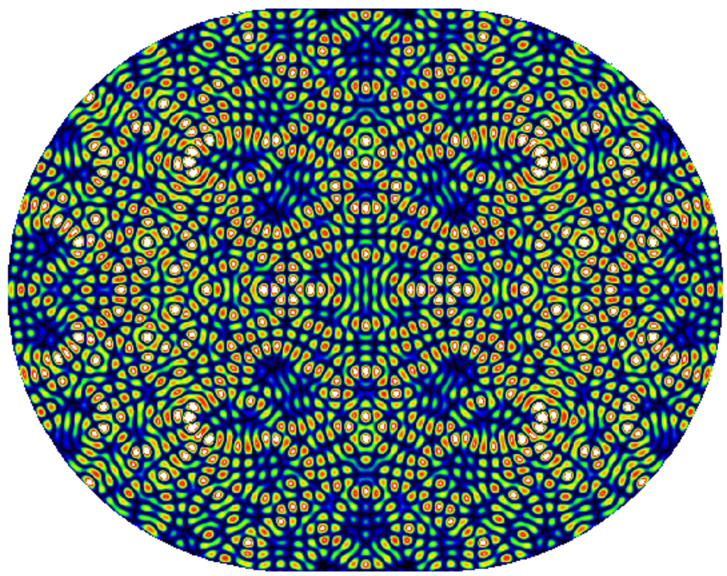
PROBLEM 1.6. Determine the set $\mathcal{Q}_\tau$ of 'quantum limits', i.e., weak* limit points of the sequence $\{\rho_{jk}\}$ on the classical phase space T^*M.

As will be discussed in ♣§9.5.8♣, the asymptotics of off-diagonal elements depends on the weak mixing properties of the geodesic flow and not just its ergodicity.

1.16. Ergodic versus completely integrable geodesic flow

In line with §1.5, one of the principal cases where one can largely control the weak* limits of the Wigner distributions and the complex nodal currents is that of Riemannian manifolds (M, g) with ergodic geodesic flow. When M has a boundary ∂M then the geodesic flow is replace by the billiard flow. In the images below, the left one shows a typical trajectory of the ergodic billiards in a stadium, and the right one gives intensity plots of the ergodic Dirichlet eigenfunctions, as well as several modes of 'bouncing ball type,' corresponding to vertical bouncing ball orbits in the middle rectangle.





1.17. Ergodic eigenfunctions

A subsequence $\{\varphi_{j_k}\}$ of eigenfunctions is called *quantum ergodicquantum ergodic* if the only weak* limit of the sequence of ρ_{j_k} is $d\mu_L$ or equivalently the Liouville state ω.

One can quantize characteristic functions $\mathbf{1}_E$ of open sets in S^*M whose boundaries have measure zero. Then

(1.52)

$$\langle \mathrm{Op}(\mathbf{1}_E)\varphi_j, \varphi_j \rangle = \text{the amplitude that the particle in energy state } \lambda_j^2 \text{ lies in } E.$$

For an ergodic sequence of eigenfunctions,

$$\langle \mathrm{Op}(\mathbf{1}_E)\varphi_{j_k}, \varphi_{j_k} \rangle \to \frac{\mu_L(E)}{\mu_L(S^*M)}, \tag{1.53}$$

so that the particle becomes diffuse, i.e. uniformly distributed on S^*M. This is the quantum analogue of the property of uniform distribution of typical geodesics of ergodic geodesic flows (Birkhoff's ergodic theorem).

On a compact Riemannian manifold (M, g) with ergodic geodesic flow, there exists a subsequence of density one of any orthonormal basis which is quantum ergodic [**Sh1, Ze1, CdV, ZZw**].

1.18. Quantum unique ergodicity (QUE)

The Laplacian Δ or (M, g) is said to be QUE (quantum uniquely ergodic) if $\mathcal{Q} = \{\mu_L\}$, i.e., the only quantum limit measure for any orthonormal basis of eigenfunctions is Liouville measure. An orthonormal basis $\{\varphi_h\}$ of $-h^2\Delta_g$-eigenfunctions is called QUE on M if $\langle a^w(x, hD_x)\varphi_h, \varphi_h\rangle \to \omega(a_0)$ for all pseudodifferential operators $a^w \in \Psi^0(M)$, i.e. if it is not necessary to eliminate a sparse subsequence of eigenfunctions of density zero. Here, $\omega(A) = \int_{S^*M} \sigma_A d\mu_L$ where $d\mu_L$ is normalized, $\sigma_A = a_0$ is the principal symbol of $A \in \Psi^0(M)$. It is conjectured in [**RS**] that QUE should hold for any orthonormal basis of eigenfunctions if the geodesic flow is Anosov, e.g. if the curvature is negative. In the case of the special orthonormal basis of Hecke-Maass forms for arithmetic hyperbolic surfaces, QUE was proved by E. Lindenstrauss [**Li**] (together with the recent final step in the noncompact case by Soundararajan [**Sou**]). However, the known bounds on multiplicities of eigenvalues are too weak for this to imply that QUE holds for all orthonormal bases.

Although we are not discussing quantum cat maps in detail, it should be emphasized that quantizations of hyperbolic (Anosov) symplectic maps of the torus are not QUE. For a sparse sequence of Planck constants $\hbar_k$, there exist eigenfunctions of the quantum cat map which partly scar on a hyperbolic fixed point (see Faure-Nonnenmacher-de Bièvre [**FNB**]). The multiplicities of the corresponding eigenvalues are of order $\hbar_k^{-1}/|\log \hbar|$. It is unknown if anything analogous can occur in the Riemannian setting, but as yet there is nothing to rule it out.

1.19. Completely integrable eigenfunctions

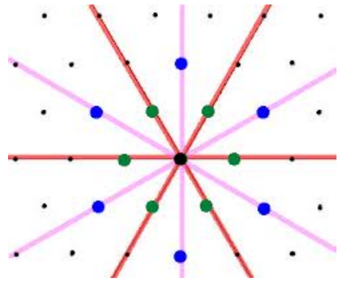
The simplest quantum systems, both on the classical and quantum level, are the completely integrable ones. They will be discussed in §11 and again in §11.6.

- Completely integrable systems: the quantum Hamiltonian

$$\hat{H} := -\frac{\hbar^2}{2}\Delta + V \tag{1.54}$$

  commutes with $n-1$ other observables P_j where $n = \dim M$. The hydrogen atom Hamiltonian and round sphere Laplacian are examples. Trajectories of the classical Hamiltonian system wind around on tori of dimension n.
- $\sqrt{\Delta} = H(P_1, \ldots, P_n)$, where $[P_i, P_j] = 0$.
- The symbols form a moment map $\mathcal{P} : T^*M \to \mathbb{R}^n$.

The joint eigenfunctions φ_α are characterized by $P_j\varphi_\alpha = \alpha_j\varphi_\alpha$ simultaneously for all j. Joint eigenvalues $\alpha = (\alpha_1, \ldots, \alpha_n)$ form a lattice. We take limits along rays in the lattice.



1.20. Heisenberg uncertainty principle

The Heisenberg uncertainty principle is the heuristic principle that one cannot measure things in regions of phase space where the product of the widths in configuration and momentum directions is $\leq \hbar$.

It is useful to microlocalize to sets which shrink as $\hbar \to 0$ using $\hbar$ (or λ)-dependent cutoffs such as $\chi_U(\lambda^\delta(x,\xi))$. The Heisenberg uncertainty principle is manifested in pseudo-differential calculus in the difficulty (or impossibility) of defining pseudo-differential cut-off operators $\mathrm{Op}(\chi_U(\lambda^\delta(x,\xi))$ when $\delta \geq 1$. That is, such small scale cutoffs do not obey the usual rules of semiclassical analysis (behavior of symbols under composition). The uncertainty principle allows one to study eigenfunctions by microlocal methods in configuration space balls $B(x_0, \lambda_j^{-1})$ of radius $\hbar$ (since there is no constraint on the 'height' in the ξ variable) or in a phase space ball of radius λ_j^{-1}.

1.21. Sequences of eigenfunctions and length scales

As mentioned above, most of the semiclassical results concern *sequences* $\{\varphi_{j_k}\}_{k=0}^\infty$ of eigenfunctions with $\lambda_{j_k} \to \infty$. When we consider φ_λ or $\varphi_\hbar$ asymptotically, we implicitly consider sequences. To orient the reader, we provide some terminology and some results on sequences of eigenfunctions.

We say that a subsequence $\mathcal{S} = \{\lambda_{j_k}\}$ has upper asymptotic density M if

$$\limsup_{\lambda\to\infty} \frac{\#\{j\colon \lambda_j \leq \lambda, \lambda_j \in \mathcal{S}\}}{N(\lambda)} \leq M, \tag{1.55}$$

and that $\mathcal{S}$ has lower asymptotic density m if

$$\liminf_{\lambda\to\infty} \frac{\#\{j\colon \lambda_j \leq \lambda, \lambda_j \in \mathcal{S}\}}{N(\lambda)} \geq m. \tag{1.56}$$

If the upper and lower densities agree the sequence is said to have an asymptotic density.

If $\{c_j\}$ is a sequence of positive numbers satisfying

$$\limsup_{\lambda\to\infty} \frac{1}{N} \sum_{j\leq N} c_j \to 0 \tag{1.57}$$

then there is a subsequence $\{c_{j_k}\}$ of upper density one such that $c_{j_k} \to 0$. If the above limit holds with lim sup replaced by lim inf, then there is a subsequence of lower density one such that $c_{j'_k} \to 0$. These statements follow from Chebyshev's inequality

$$\frac{1}{N}\#\{j \leq N : c_j > \varepsilon\} \leq \frac{1}{\varepsilon}\frac{1}{N}\sum_{j \leq N} c_j \tag{1.58}$$

by taking lim sup or lim inf of both sides.

We single out sequences $\{\varphi_{j_k}\}$ of eigenfunctions for which the Wigner distributions of §1.12 have a unique weak* limit. In the case where the geodesic flow is ergodic, and there exists just one weak* limit for any orthonormal basis, the sequence is called QUE. However, this might be taken to imply that the limit is Liouville measure. It does not appear that any general term is standard, so we use our original term:

Definition 1.7. We say that a sequence $\{\varphi_{j_k}\}$ of eigenfunctions is 'coherent' if the associated sequence of Wigner distributions has a unique weak limit. Given a coherent sequence, we define the *microsupport* $\mathrm{MS}\{\varphi_{j_k}\}$ to be the support of the unique weak* limit measure of the sequence, i.e.,

$$\mathrm{MS}\{\varphi_{j_k}\} = \mathrm{supp}(\mu) \subset T^*M. \tag{1.59}$$

It follows from Egorov's theorem that $\mathrm{MS}\{\varphi_{j_k}\} \subset S^*M$ is invariant under the geodesic flow G^t. A priori, the limit measure could be any invariant measure, and as mentioned before, any invariant measure is in fact a weak* limit in the case of the standard sphere. Another characterization of the microsupport is the following:

Definition 1.8. The annihilator $\mathcal{A}_{\{\varphi_{j_k}\}}$ of the sequence $\{\varphi_{j_k}\}$ is the class of semiclassical pseudo-differential operators $\mathrm{Op}_h(a)$ with $h = h_{j_k} = \lambda_{j_k}^{-1}$ such that

$$\lim_{k\to\infty} \langle \mathrm{Op}_h(a)\varphi_{j_k}, \varphi_{j_k}\rangle = 0. \tag{1.60}$$

The connection to the microsupport of $\{\varphi_{j_k}\}$ is that

$$\mathrm{MS}\{\varphi_{j_k}\} = \bigcap_{A \in \mathcal{A}_{\{\varphi_{j_k}\}}} \mathrm{Char}(A), \tag{1.61}$$

where

$$\mathrm{Char}(A) = \{(x,\xi) \in S^*M : \sigma_A(x,\xi) = 0\}. \tag{1.62}$$

Equivalently, if μ is the limit measure then

$$A \in \mathcal{A}_{\{\varphi_{j_k}\}} \iff \mathrm{supp}(\mu) \subset \mathrm{Char}(A). \tag{1.63}$$

The last definition makes sense for any sequence $\{\varphi_{j_k}\}$, whether or not it is coherent. If it is not coherent, then the annihilator of the sequence is the intersection of the annihilators of the coherent subsequences.

1.22. Localization of eigenfunctions on closed geodesics

It follows from the invariance of the weak* limit measures that the smallest possible microsupport a coherent sequence $\{\varphi_{j_k}\}$ can have is a single closed geodesic γ. When the Wigner measures tend to the delta-function $\delta_\gamma(f) = \int_\gamma f ds$, then $\{\varphi_{j_k}\}$ is said to localize or concentrate along γ. One has

$$\langle A\varphi_{j_k}, \varphi_{j_k}\rangle \to \int_\gamma \sigma_A \, ds. \tag{1.64}$$

In configuration space M, this says that

$$\int_M f|\varphi_{j_k}|^2\,dV \to \int_\gamma f\,ds. \tag{1.65}$$

Thus, squares of the eigenfunctions form a delta-sequence on γ.

On special Riemannian manifolds (M,g) there do exist sequences $\{\varphi_{j_k}\}$ of eigenfunctions concentrating on γ. Namely if (M,g) possesses a stable elliptic closed geodesic, then one may construct a Gaussian beam along γ. In general it is only a sequence of approximate eigenfunctions but in special cases they are genuine eigenfunctions.

A measure of the localization of eigenfunctions around a specific closed geodesic γ is given by the L^2 *mass profile*, which studies the restricted L^2 norm-squared $\int_\gamma |\varphi_\lambda|^2\,ds$ for an orthonormal basis. In the case of quantum integrable systems such as the standard sphere or a hyperbolic cylinder, one may study the norm-squares as a function of the joint eigenvalues. The limit mass profile is explicitly computable in model cases. Another measure is to study concentration of L^2 mass in tubes, and this is done for integrable systems in §11.6. If the geodesic is not fixed one may take the supremum over all geodesic arcs of a fixed length. This defines the geodesic maximal function (Definition 3.3). Or one may study the supremum of the mass in tubes using the Kakeya-Nikodym maximal function (Definition 3.9 studied in §3.4).

1.23. Some remarks on the contents and on other texts

There are several very recent references providing a systematic background on eigenfunctions and the wave equation on Riemannian manifolds. In particular, the new book [**So2**] of Chris Sogge covers much of the background and also discusses some relatively recent joint work with the author on L^p norms of eigenfunctions. The earlier book [**So1**] includes a systematic introduction to the relevant theory of Fourier integral operators. The relatively new book of Maciej Zworski [**Zw**] contains a systematic introduction to semiclassical Fourier integral operators and includes applications to quantum ergodicity of eigenfunctions. Other excellent texts covering Fourier integral operators and the wave group are [**Ho1, Ho2, Ho3, Ho4**] and [**SV**]. For this reason, we do not give a detailed treatment here of Fourier integral parametrices for the long time wave kernel, although we do use the results. Instead, we try to limit the foundational material and survey the statements and methods-of-proof of relatively recent results. We do provide background on the Hadamard parametrix taken almost directly from Hadamard's remarkable book [**Had**]. This book is in some ways better than the more recent expositions, in particular because it proves the convergence of Hadamard's parametrix series in the real analytic case. We take some of the details from M. Riesz's classic work [**R**] and from the more recent exposition of P. Bérard [**Be**]. We also review the basic results on eigenfunctions on model spaces of constant curvature ($\mathbb{R}^n$, flat tori $\mathbb{R}^n/\Gamma$, spheres S^n, hyperbolic space $\mathbf{H}^n$ and its quotients $\mathbf{H}^n/\Gamma$).

In addition to this monograph, the author has recently written extensive survey articles on eigenfunctions and their nodal sets [**Ze6**], on recent results on quantum chaos [**Ze4**], and on the use of global wave equation methods in the study of eigenfunctions [**Ze3**]. The main theme of [**Ze3**] was to contrast local ('small ball', elliptic) methods for studying eigenfunctions with global wave equation methods.

We develop that theme here as well. In §5.3 we review the elliptic methods that have been applied to eigenfunctions by Donnelly-Fefferman [**DF**], F.H. Lin (see [**HL**]), Nazarov-Sodin, Colding-Minicozzi, and many others. Also recent are the author's lecture notes [**Ze7**] from the Park City summer program in geometric analysis (2013).

1.24. References

Below are two groups of references. The first are historically important papers in the field, some of which (especially by Hadamard and Riesz) remain among the most lucid and complete expositions of results on the wave equation. The second group consists of research articles relevant to this chapter. Here and henceforth, references are given chapter by chapter to make it easier for the reader to find the relevant reference.

Bibliography

[B] N. Bohr "On the Constitution of Atoms and Molecules, Part I". Philosophical Magazine 26 (151):(1913) 124.

[D] P. A. M. Dirac, *The Principles of Quantum Mechanics.* 3d ed. Oxford, at the Clarendon Press, 1947.

[FH] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, Emended Edition (Dover Books on Physics).

[Had] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations.* Dover Publications, New York, 1953.

[LL] L. D. Landau, and E. M. Lifshitz, *Quantum mechanics: non-relativistic theory.* Course of Theoretical Physics, Vol. 3. Addison-Wesley Series in Advanced Physics. Addison-Wesley Publishing Co., Inc., Reading, Mass; 1958.

[M] A. Messiah, *Quantum mechanics. Vols. I-II.* North-Holland Publishing Co., Amsterdam; Interscience Publishers.

[R] M. Riesz, L'intégrale de Riemann-Liouville et le problème de Cauchy. Acta Math. 81, (1949). 1-223.

[Ra] J. W. Strutt, Baron Rayleigh, The Theory of Sound. 2d ed. Dover Publications, New York, N. Y., 1945.

[Sch1] E. Schrödinger, Collected papers on wave mechanics. Chelsea Publishing Co., New York, 1982.

[Sch2] E. Schrödinger, Quantisierung als Eigenwertproblem, Annalen der Physik (1926).

[Wei] S. Weinberg, *Lectures on quantum mechanics.* Cambridge University Press, Cambridge, 2013.

[Wey] H. Weyl, Gruppentheorie und Quantenmechanik (Wissenschaftliche Buchgesellschaft, Darmstadt), Nachdruck der 2. Aufl., Leipzig 1931. [English translation: Group Theory and Quantum Mechanics, Dover, New York (1950)].

[Wey2] H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung). Math. Ann. 71 (1912), no. 4, 441 - 479.

[Wey3] H. Weyl, Ramifications, old and new, of the eigenvalue problem. Bull. Amer. Math. Soc. 56, (1950). 115 - 139.

[Ag] S. Agmon, Representation theorems for solutions of the Helmholtz equation on Rn. Differential operators and spectral theory, 2743, Amer. Math. Soc. Transl. Ser. 2, 189, Amer. Math. Soc., Providence, RI, 1999

[Be] P. Bérard, On the wave equation on a compact Riemannian manifold without conjugate points. Math. Z. 155 (1977), no. 3, 249-276.

[BoR] A. Bouzouina and D. Robert, Uniform semiclassical estimates for the propagation of quantum observables. Duke Math. J. 111 (2002), no. 2, 223–252.

[Br] J. Bruening, ber Knoten von Eigenfunktionen des Laplace-Beltrami-Operators. Math. Z. 158 (1978), no. 1, 15-21.

[C] I. Chavel, *Eigenvalues in Riemannian geometry.* Including a chapter by Burton Randol. With an appendix by Jozef Dodziuk. Pure and Applied Mathematics, 115. Academic Press, Inc., Orlando, FL, 1984.

[CdV] Y.Colin de Verdière, Ergodicité et fonctions propres du Laplacien, Comm.Math.Phys. 102 (1985), 497-502.

[CR] M. Combescure and D. Robert, Semiclassical spreading of quantum wave packets and applications near unstable fixed points of the classical flow, Asymptot. Anal. 14 (1997), 377–404.

[DSj] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semiclassical limit.* London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.

[DF] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions on Riemannian manifolds, Invent. Math. 93 (1988), 161-183.

[DZw] S. Dyatlov and M. Zworski, *Resonances*, book in progress.

[FNB] F. Faure, S. Nonnenmacher, and S. De Bievre, Scarred eigenstates for quantum cat maps of minimal periods. Comm. Math. Phys. 239 (2003), no. 3, 449-492.

[Fe] C. L. Fefferman, The uncertainty principle. Bull. Amer. Math. Soc. (N.S.) 9 (1983), no. 2, 129-206.

[GS] V. Guillemin and S. Sternberg, *Semi-classical analysis*, International Press, Boston, MA, 2013. MR3157301.

[GHZ] J. Gell-Redman, A. Hassell and S. Zelditch, Equidistribution of phase shifts in semiclassical potential scattering. J. Lond. Math. Soc. (2) 91 (2015), no. 1, 159-179.

[HL] Q. Han and F. H. Lin: *Nodal sets of solutions of Elliptic Differential Equations*, book in preparation (online at http://www.nd.edu/qhan/nodal.pdf).

[He1] E.J. Heller, Bound-state eigenfunctions of classically chaotic Hamiltonian systems: scars of periodic orbits. Phys. Rev. Lett. 53 (1984), no. 16, 1515–1518.

[He2] E.J. Heller, Wavepacket dynamics and quantum chaology. Chaos et physique quantique (Les Houches, 1989), 547–664, North-Holland, Amsterdam, 1991.

[Ho1] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[Ho2] L. Hörmander, *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Classics in Mathematics. Springer-Verlag, Berlin, 2005.

[Ho3] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[Ho4] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[JZ] D. Jakobson and D. Zelditch, Classical limits of eigenfunctions for some completely integrable systems. Emerging applications of number theory (Minneapolis, MN, 1996), 329-354, IMA Vol. Math. Appl., 109, Springer, New York, 1999.

[Li] E. Lindenstrauss, Invariant measures and arithmetic quantum ergodicity, Annals of Math. (2) 163 (2006), no. 1, 165–219.

[Lo] A. Logunov, Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture, arXiv:1605.02589.

[NS] F. Nazarov and M. Sodin, On the number of nodal domains of random spherical harmonics. Amer. J. Math. 131 (2009), no. 5, 1337-1357.

[RS] Z. Rudnick and P. Sarnak, The behaviour of eigenstates of arithmetic hyperbolic manifolds. Comm. Math. Phys. 161 (1994), no. 1, 195–213.

[SV] Yu. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators.* Translations of Mathematical Monographs, 155. American Mathematical Society, Providence, RI, 1997.

[S] R. Schubert, Semiclassical behaviour of expectation values in time evolved Lagrangian states for large times. Comm. Math. Phys. 256 (2005), no. 1, 239–254.

[STZ] B. Shiffman, T. Tate, and S. Zelditch, Distribution laws for integrable eigenfunctions. Ann. Inst. Fourier (Grenoble) 54 (2004), no. 5, 1497-1546, xvxvi, xxi.

[Sh1] A.I. Shnirelman, Ergodic properties of eigenfunctions, Usp.Math.Nauk. 29/6 (1974), 181-182.

[Sh2] A.I. Shnirelman, On the asymptotic properties of eigenfunctions in the region of chaotic motion, addendum to V.F.Lazutkin, *KAM theory and semiclassical approximations to eigenfunctions*, Springer (1993).

[Sou] K. Soundararajan, Quantum unique ergodicity for $SL_2(\mathbb{Z})$ H, Ann. of Math. (2) 172 (2010), no. 2, 1529-1538 (arXiv:0901.4060).

[Sj] J. Sjöstrand, *Singularités analytiques microlocales.* Astérisque, 95, 1166, Astérisque, 95, Soc. Math. France, Paris, 1982.

[So1] C. D. Sogge, *Fourier integrals in classical analysis.* Cambridge Tracts in Mathematics, 105. Cambridge University Press, Cambridge, 1993.

[So2] C. D. Sogge, *Hangzhou lectures on eigenfunctions of the Laplacian*, Annals of Math. Studies 188. Princeton University Press, Princeton, NJ, 2014.

[T1] M. E. Taylor, *Partial differential equations I. Basic theory.* Second edition. Applied Mathematical Sciences, 115. Springer, New York, 2011.

[T2] M. E. Taylor, *Partial differential equations II. Qualitative studies of linear equations.* Second edition. Applied Mathematical Sciences, 116. Springer, New York, 2011.

[TZ] J. A. Toth and S. Zelditch, L^p norms of eigenfunctions in the completely integrable case. Ann. Henri Poincar'e 4 (2003), no. 2, 343-368.

[Ze1] S. Zelditch, Uniform distribution of eigenfunctions on compact hyperbolic surfaces. Duke Math. J. 55 (1987), no. 4, 919-941

[Ze2] S. Zelditch, The inverse spectral problem. With an appendix by Johannes Sjöstrand and Maciej Zworski. Surv. Differ. Geom., IX, Surveys in differential geometry. Vol. IX, 401-467, Int. Press, Somerville, MA, 2004.

[Ze3] S. Zelditch, Local and global analysis of eigenfunctions on Riemannian manifolds. Handbook of geometric analysis. No. 1, 545-658, Adv. Lect. Math. (ALM), 7, Int. Press, Somerville, MA, 2008.

[Ze4] S. Zelditch, Recent developments in mathematical quantum chaos. Current developments in mathematics, 2009, 115-204, Int. Press, Somerville, MA, 2010.

[Ze5] S. Zelditch, Pluri-potential theory on Grauert tubes of real analytic Riemannian manifolds, I. Spectral geometry, 299-339, Proc. Sympos. Pure Math., 84, Amer. Math. Soc., Providence, RI, 2012.

[Ze6] S. Zelditch, Eigenfunctions and Nodal sets, Surveys in Differential Geometry volume 18 (2013), pp. 237-308 (arXiv:1205.2812.)

[Ze7] S. Zelditch, Park City Lectures on Eigenfunctions, Geometric analysis, 111-193, IAS/Park City Math. Ser., 22, Amer. Math. Soc., Providence, RI, 2016.

[ZZw] S.Zelditch and M.Zworski, Ergodicity of eigenfunctions for ergodic billiards, Comm.Math. Phys. 175 (1996), 673-682.

[Zw] M. Zworski, *Semiclassical analysis.* Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012.

CHAPTER 2

Geometric preliminaries

We assume the reader is familiar with basic notions of Riemannian geometry. We now quickly run through the definitions and notations which we will need in analyzing wave invariants. All the notions are discussed in detail in [**Kl**]. We also refer to [**Su**] for background in a context closely related to that of this article.

The following notation is used:

(i)	$r = r(x, y)$	distance function on M;
(ii)	$\lvert\xi\rvert_g : T^*M - 0 \to \mathbb{R}^+$	length of a (co)-vector;
(iii)	$S^*M = \{\lvert\xi\rvert_g = 1\}$	the unit cosphere bundle;
(iv)	$G^t : T^*M - 0 \to T^*M - 0$	the geodesic flow, i.e., the Hamilton flow of $\lvert\xi\rvert_g$;
(v)	γ	closed geodesic, i.e., a closed orbit of G^t in S^*M;
(vi)	$\Lambda(M)$	H^1 loop space of M;
(vii)	$\mathcal{G}(M)$	subset of closed geodesics in $\Lambda(M)$;
(viii)	$\mathcal{G}_{[\gamma]}$	set of closed geodesics in $\mathcal{G}(M)$ whose free homotopy class is $[\gamma]$;
(ix)	$\mathrm{inj}(M, g)$	the injectivity radius.

2.1. Symplectic linear algebra and geometry

Classical Hamiltonian dynamics has been written since Hamilton in the language of symplectic geometry. Quantum mechanics and semiclassical analysis are quantizations of classical Hamiltonian dynamics, and symplectic notions such as Lagrangian submanifolds and canonical relations permeate quantization theory. In many systematic texts on Fourier integral operators [**Du, Ho1, Ho2, Ho3, Ho4, Zw**], the reader will find extensive discussions on symplectic geometry and Hamilton-Jacobi theory in the construction of parametrices for wave equations and more general PDEs. Further use of symplectic geometry is to find canonical transformations to symplectic normal forms, and to quantize the canonical transformations to give quantum normal forms. Estimates of mapping properties of oscillatory integral operators often start from the symplectic geometry of the underlying canonical relations.

The purpose of this section is to introduce notation and terminology that will be used throughout. It is assumed that the reader is familiar with basic symplectic notions. We refer to [**Ho4, DSj, GSj, GuSt1, GuSt2, BW**] for a systematic exposition.

2.1.1. Symplectic linear algebra. We recall that a symplectic vector space (V, σ) equipped with a symplectic form σ is an even dimensional vector space with a non-degenerate skew-symmetric 2-form σ. That is, $\sigma(v,w) = -\sigma(w,v)$ and $v \to \sigma(v,\cdot)$ is an isomorphism from the vector space V to its dual V^*.

The basic symplectic vector space is $(\mathbb{R}^{2n}, \sigma)$ with

$$\sigma((x,\xi),(x',\xi'))) = \langle \xi, x'\rangle - \langle \xi', x\rangle. \tag{2.1}$$

Given any vector space V with dual V^* one can use the same definition to give $V \oplus V^*$ a symplectic structure. The universal construction of symplectic vector spaces is as follows: Let W be a vector space and W^* be its dual. Let $V = W \oplus W^*$ with symplectic form

$$\sigma_0(e \oplus f, e' \oplus f') = f'(e) - f(e'). \tag{2.2}$$

Every symplectic vector space (say of dimension $2n$) has a *symplectic (Darboux) basis* $e_j, f_k (j,k = 1,\dots,n)$ so that

$$\sigma(e^i, e^j) = \sigma(f_i, f_j) = 0 \quad \text{and} \quad \sigma(e^i, f_j) = \delta_{ij}. \tag{2.3}$$

In coordinates relative to this basis,

$$\sigma = \sum_{i=1}^{n} e^j \otimes f_j^*. \tag{2.4}$$

2.1.2. Lagrangian subspaces. Let V be a symplectic vector space of dimension $2n$. A *Lagrangian subspace* of V is a subspace $W \subset V$ of dimension n so that $\sigma|_W = 0$, i.e., $\sigma(v,w) = 0$ for all $v, w \in W$. For instance, $\mathbb{R}\{e_1,\dots,e_n\}$ and $\mathbb{R}\{f_1,\dots,f_n\}$ are Lagrangian subspaces of V, where e_j, f_k are the symplectic basis elements of V introduced in the previous subsection.

Consider the possible Lagrangian subspaces of (V,σ) which are graphs over W so that $f_j = \sum_{k=1}^n A_{jk} e_k$. In other words, let $A : W \to W^*$ be a linear map and let $\Gamma_A = \{(w, Aw) \subset W \oplus W^*\}$ be its graph. Then Γ_A is Lagrangian if and only if A is symmetric. Indeed,

$$\sigma((w,Aw),(w',Aw')) = (Aw')(w) - (Aw)(w'). \tag{2.5}$$

2.1.3. Structure of linear canonical relations. Lemma 25.3.6 of [**Ho4**] states:

PROPOSITION 2.1. *Let $G \subset S_1 \oplus S_2$ be a canonical relation. Then there exist symplectic orthogonal decompositions*

$$S_1 = S_{11} \oplus S_{12} \quad \textit{and} \quad S_2 = S_{21} \oplus S_{22} \tag{2.6}$$

so that

$$G = \lambda_1 \oplus \hat{G} \oplus \lambda_2 \tag{2.7}$$

where $\lambda_j \subset S_{jj}$ is Lagrangian and $\hat{G}$ is the graph of a linear symplectic transformation $S_{21} \to S_{12}$.

PROOF. We copy the proof verbatim from Hörmander. Let

$$\lambda_1 = \{(\gamma \in S_1 : (\gamma, 0) \in G\} \quad \text{and} \quad \lambda_2 = \{(\gamma \in S_2 : (0,\gamma) \in G\}. \tag{2.8}$$

They are isotropic subspaces and $G \subset \lambda_1^\sigma \oplus \lambda_2^\sigma$ since G is Lagrangian. Write $\lambda_1^\sigma = \lambda_1 \oplus S_{12}$. Then $S_{12} = \lambda_1^\sigma/\lambda_1$ is symplectic and $S_1 = S_{11} \oplus S_{12}$ where the symplectic orthogonal complement S_{11} of S_{12} contains λ_1. Define S_{21}, S_{22} analogously. Then $G = \lambda_1 \oplus \hat{G} \oplus \lambda_2$ where $\hat{G} \subset S_{12} \oplus S_{21}$ has bijective projection onto S_{12} and S_{21}. Since $\dim S_{jj} = 2 \dim \lambda_j$, λ_j is symplectic in S_{jj}. □

2.2. Symplectic manifolds and cotangent bundles

A symplectic manifold (X, ω) is a manifold with a closed non-degenerate two form, i.e., $d\sigma = 0$. Non-degeneracy means that $\omega_x : T_x X \to T_x^* X,\ v \to \omega_x(v, \cdot)$, is an isomorphism. One can then define the symplectic gradient Ξ_f of a function $f\colon M \to \mathbb{R}$ by $df = \sigma(\Xi_f, \cdot)$. Given local coordinates $x_1, \ldots, x_n$ on M any covector may be expressed as $\eta = \sum_{j=1}^n \eta_j\, dx_j$ and one defines the local coordinates $\xi_j : T^*M \to \mathbb{R}$ by $\xi_j(\eta) = \eta_j$.

The main example is $X = T^*M$ the cotangent bundle of a Riemannian manifold M. We recall that a cotangent bundle carries a canonical 1-form $\alpha = \xi\, dx$ defined by

$$\alpha_{x,\xi}(v) = \xi(D\pi v). \tag{2.9}$$

It may also be described as follows: A section of the natural projection $\pi : T^*M \to M$ is a co-vector field $\eta : M \to T^*M$. Then α is the unique element of $\Omega^1(T^*M)$ (i.e., unique 1-form on T^*M) satisfying $\eta^*\alpha = \eta$.

Let $x = (x_1, \ldots, x_n)$ be any local coordinate system on M. Then dx_j define local covector fields giving a local trivialization of $T^*M \to M$. One defines the dual coordinates on T^*M of a covector $\eta = \sum_j \eta_j dx_j$ by the coordinates relative to the frame field $dx_1, \ldots, dx_n$, i.e., $\xi_j(\eta) = \eta_j$.

The canonical symplectic form ω of T^*M is defined by $\omega = d\alpha$. Thus, in the above local coordinates,

$$\omega = \sum_j d\xi_j \wedge dx_j. \tag{2.10}$$

Such coordinates are called symplectic. In any coordinate system y_j on T^*M one could write

$$\omega = \sum_{j,k=1}^n \omega_{ij}\, dy_j \wedge dy_k, \tag{2.11}$$

but in the (x, ξ) coordinates ω has constant coefficients

$$\big(\omega_{jk}\big) = J := \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{2.12}$$

where we use the basis $\frac{\partial}{\partial x_j}, \frac{\partial}{\partial \xi_k}$ for $T(T^*M)$. That is, we have

$$\omega\left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial x_k}\right) = 0 = \omega\left(\frac{\partial}{\partial \xi_j}, \frac{\partial}{\partial \xi_k}\right) \quad \text{and} \quad \omega\left(\frac{\partial}{\partial x_j}, \frac{\partial}{\partial \xi_k}\right) = \delta_{jk}. \tag{2.13}$$

Note that for fixed $(x, \xi) \in T^*M$, the tangent space $T_{(x,\xi)}T^*M$ is a symplectic vector space with the symplectic form $\omega_{x,\xi}$.

2.2.1. Compatible complex structure.

A non-degenerate 2-form ω defines an isomorphism

$$J : T_x^*M \to T_x M, \qquad \omega(u, v) = J^{-1}(u)v. \tag{2.14}$$

If $\omega = \frac{1}{2}\sum_{j=1}^n \omega_{ij} dx^i \wedge dx^j$ then

$$J(dx^i) = -\sum_{j=1}^n \omega^{ij}(x) \frac{\partial}{\partial x_j}. \tag{2.15}$$

2.2.2. Hamilton vector fields. The Hamilton vector field Ξ_H of a Hamiltonian $H\colon T^*M \to \mathbb{R}$ is the symplectic gradient of H, i.e., $\omega(\Xi_H \cdot) = dH$. Unlike the Riemannian metric gradient, X_H is tangent to level sets of H. Thus,

$$\Xi_H = \sum_j \frac{\partial H}{\partial \xi_j}\frac{\partial}{\partial x_j} - \frac{\partial H}{\partial x_j}\frac{\partial}{\partial \xi_j}. \tag{2.16}$$

The Poisson bracket of two functions is defined by

$$\{f, g\} = \Xi_f(g) = dg(\Xi_f) = \sum_{j=1}^n \left(\frac{\partial f}{\partial \xi_j}\frac{\partial g}{\partial x_j} - \frac{\partial f}{\partial x_j}\frac{\partial g}{\partial \xi_j} \right). \tag{2.17}$$

2.3. Lagrangian submanifolds

Definition 2.2. A submanifold $\Lambda \subset T^*M$ of dimension $n = \dim M$ is Lagrangian if $\omega|_{T\Lambda} = 0$. That is, $\omega_\lambda(X, Y) = 0$ for all $\lambda \in \Lambda, X, Y \in T_\lambda\Lambda$.

A Lagrangian submanifold is called *projectible* if the natural projection

$$\pi\colon \Lambda \subset T^*M \to M \tag{2.18}$$

is a diffeomorphism. If $U \subset M$ is an open set we say that $\pi\colon \Lambda|_U \to U$ is locally projectible if the projection is a local diffeomorphism.

A section of T^*M is the same as a covector field $\eta\colon M \to T^*M$. The graph of a covector field is the submanifold $\Gamma_\eta = \{(x, \eta_x)\colon x \in M\}$.

Proposition 2.3.

- Γ_η *is a Lagrangian submanifold if and only if* η *is a closed 1-form.*
- $\Lambda \subset T^*M$ *is a globally projectible Lagrangian submanifold if and only if* $\Lambda = \Gamma_\eta$ *where* η *is a closed 1-form, i.e.,* $d\eta = 0$.
- *If* Λ *is locally projectible over a contractible open set* U*, then* $\Lambda = \Gamma_{dS}$ *for some smooth* $S\colon U \to \mathbb{R}$*, i.e., it is the graph of an exact form form* dS*.* S *is called a local generating function of* Λ*. It is determined up to a constant.*

In general, Lagrangian submanifolds are not projectible, even locally. For instance, every curve in $T^*\mathbb{R}$ defines a Lagrangian submanifold. The unit circle $\{x^2 + \xi^2 = 1\}$ is only projectible away from the "turning points" $\xi = 0, x = \pm 1$ and is a two-sheeted cover over $(-1, 1)$. The image $\pi(\Lambda) \subset M$ is called the classically allowed region for Λ. The singular set $\{\lambda \in \Lambda : \ker d\pi_\lambda \neq \{0\}\}$ is called the Maslov singular cycle.

Since Lagrangian submanifolds are rarely projectible, they are often parametrized by Lagrangian immersions

$$\iota\colon \Lambda \to T^*M. \tag{2.19}$$

A key fact is the following

Proposition 2.4. *Suppose that* $\Lambda \subset T^*M$ *is a Lagrangian submanifold and that* $\Lambda \subset \{H = E\}$*. Then* X_H *is tangent to* Λ*.*

Proof. By assumption $\omega|_{T\Lambda} = 0$ and since $T\Lambda$ has dimension n, if X is a vector in $T_\lambda(T^*M)$ such that $\omega_\lambda(X, Y) = 0$ for all $Y \in T_\lambda\Lambda$ then $X \in T_\lambda\Lambda$. Hence it suffices to show that $\omega_\lambda(\Xi_H, Y) = 0$ or all $Y \in T_\lambda\Lambda$. But $\omega_\lambda(\Xi_H, Y) = dH(Y) = 0$ for all $Y \in T_\lambda(\{H = E\})$. □

2.4. Jacobi fields and Poincaré map

♣We say a metric g is *non-degenerate* if the energy functional E on $\Lambda(M)$ is a Bott-Morse function, i.e., $\mathcal{G}(M)$ is a smooth submanifold of $\Lambda(M)$ and $T_c\mathcal{G}(M) = \ker J_c$ where $c \in \mathcal{G}(M)$ and J_c is the Jacobi operator (index form) on $T_c\Lambda(M)$. In other words, $J_c = \nabla^2 + R(\dot{c}, \cdot)\dot{c}$. We also say that g is *bumpy* if, for every $c \in \mathcal{G}(M)$, the orbit $S^1(c)$ of c under the S^1-action of constant reparametrization $c(t+s)$ of $c(t)$ is a non-degenerate critical manifold of E. For the sake of simplicity we will assume that (M, g) is a bumpy Riemannian manifold.♣

Let γ be a closed geodesic of length L_γ of a Riemannian manifold (M, g). Let $\mathcal{J}_\gamma^\perp \otimes \mathbb{C}$ denote the space of complex normal Jacobi fields along γ, a symplectic vector space of (complex) dimension $2n$ (where $n = \dim M - 1$) with respect to the Wronskian

$$\omega(X, Y) = g\left(X, \frac{d}{ds}Y\right) - g\left(\frac{d}{ds}X, Y\right). \tag{2.20}$$

The linear Poincaré map P_γ is then the linear symplectic map on $\mathcal{J}_\gamma^\perp \otimes \mathbb{C}$ defined by

$$P_\gamma Y(t) = Y(t + L_\gamma). \tag{2.21}$$

Recall that, since P_γ is symplectic, its eigenvalues ρ_j come in three types: (i) pairs $\rho, \bar{\rho}$ of conjugate eigenvalues of modulus 1; (ii) pairs ρ, ρ^{-1} of inverse real eigenvalues; and (iii) 4-tuplets $\rho, \bar{\rho}, \rho^{-1}\bar{\rho}^{-1}$ of complex eigenvalues. We will often write them in the forms: (i) $e^{\pm i\alpha_j}$; (ii)$e^{\pm\lambda_j}$; (iii) $e^{\pm\mu_j \pm i\nu_j}$ respectively (with $\alpha_j, \lambda_j, \mu_j, \nu_j \in \mathbb{R}$), although a pair of inverse real eigenvalues $\{-e^{\pm\lambda}\}$ could be negative. Here, and throughout, we make the assumption that P_γ is *non-degenerate* in the sense that $\det(I - P_\gamma) \neq 0$. In constructing the normal form we assume the stronger non-degeneracy assumption that

$$\Pi_{i=1}^{2n} \rho_i^{m_i} \neq 1 \quad \text{for all } \rho_i \in \sigma(P_\gamma) \text{ and } (m_1, \ldots, m_{2n}) \in \mathbf{N}^{2n}. \tag{2.22}$$

A closed geodesic is called *elliptic* if all of its eigenvalues are of modulus one, *hyperbolic* if they are all real, and *loxodromic* if they all come in quadruples as above.

In the elliptic case, namely case (i) where the eigenvalues of P_γ are of the form $e^{\pm i\alpha_j}$, the associated normalized eigenvectors will be denoted $\{Y_j, \overline{Y_j}\}_{j=1,\ldots,n}$:

$$P\gamma Y_j = e^{i\alpha_j} Y_j, \quad P_\gamma \overline{Y}_j = e^{-i\alpha_j} \overline{Y}_j, \quad \omega(Y_j, \overline{Y}_k) = \delta_{jk} \tag{2.23}$$

Relative to a fixed parallel normal frame $e(s) := (e_1(s), ..., e_n(s))$ along γ they will be written in the form

$$Y_j(s) = \sum_{k=1}^{n} y_{jk}(s) e_k(s). \tag{2.24}$$

An elliptic γ is said to be non-degenerate elliptic if $\{\alpha_j\}_{j=1,\ldots,n}$ together with π are independent over $\mathbb{Q}$.

In the case of surfaces, $\mathcal{J}_\gamma^\perp \otimes \mathbb{C}$ has complex dimension two and as mentioned above is spanned by the eigenvectors $\{Y, \bar{Y}\}$. A normal Jacobi field along γ is simply of the form $Y(s) = y(s)\nu(s)$, where $\nu(s)$ is the parallel unit normal vector γ. Jacobi's equation is then a second order scalar equation:

$$y'' + \tau y = 0. \tag{2.25}$$

There is a two dimensional space of solutions: the vertical Jacobi field y_1 with initial conditions $y(0) = 0, y'(0) = 1$ and the horizontal Jacobi field y_2 with initial conditions $y(0) = 1, y'(0) = 0$ with respect to a fixed choice of origin $\gamma(0)$ of γ. We consider the pair (y, y') and form the symplectic Wronskian matrix:

$$a_s := \begin{pmatrix} y_2'(s) & y_1'(s) \\ y_2(s) & y_1(s) \end{pmatrix}. \tag{2.26}$$

We modify the Wronskian matrix so that its columns are given in terms of the normalized eigenvectors (2.23) of the Poincaré map:

$$\mathcal{A}(s) := \begin{pmatrix} \mathrm{Im}\dot{Y} & \mathrm{Re}\dot{Y} \\ \mathrm{Im}Y & \mathrm{Re}Y \end{pmatrix}. \tag{2.27}$$

The somewhat strange positioning of the elements is to maintain consistency with our reference [**F**] on the metaplectic representation.

Of course, geodesics of S^n are degenerate. In the general Zoll case, $P_\gamma = \mathrm{Id}$, i.e., the normal Jacobi fields are periodic (all $\alpha_j = 0$) and the Wronskian matrices a_s resp. $\mathcal{A}(s)$ are periodic. Yet the highest weight spherical harmonics are models of Gaussian beams. Gaussian beams on Zoll surfaces are discussed in detail in [**Z**].

2.4.1. Fermi normal coordinates along a geodesic. Fermi normal coordinates are the normal coordinates defined by the exponential map $\exp\colon N_{\gamma,\varepsilon} \to T_\varepsilon(\gamma)$ from a ball in the normal bundle of γ to a tube of radius ε around γ. Thus we write $(s, y) = \exp_{\gamma(s)} y \cdot \nu_{\gamma(s)}$, where $\nu(s)$ is a choice of unit normal frame along γ. We write the associated metric coefficients as

$$g_{00} = g(\partial_s, \partial_s), \quad g_{0j} = 0, \quad g_{jk} = g(\partial_{y_j}, \partial_{y_k}) = 1. \tag{2.28}$$

In Fermi coordinates along a geodesic, the field $\frac{\partial}{\partial s}$ is a horizontal Jacobi field pointing between nearby normal geodesics to γ and tangent to the wave fronts. Each $\frac{\partial}{\partial y_j}$ is a geodesic vector field. The volume density is given by $j = \sqrt{\det g}$. In dimension two, $j = \|\frac{\partial}{\partial s}\|$.

2.5. Pseudo-differential operators

We briefly review the theory of pseudo-differential operators on manifolds. There are many comprehensive texts and we only go over some very basic definitions, referring to [**GSj, Ho3, Zw**] for extensive background on pseudo-differential operators on manifolds. In the §4.16 we collect notation and background on the Fourier transform on $\mathbb{R}^n$.

Homogeneous pseudo-differential operators on a manifold are defined and discussed in detail in [**Ho3**]. We only consider symbols of type $(\rho, \delta) = (1, 0)$ and denote the space of homogeneous pseudo-differential operators of order m and type $(1, 0)$ by Ψ^m. Symbols are discussed in the next section §2.6. They are operators which are sums of a smoothing operator (i.e., an operator with a smooth kernel) and an oscillatory integral operator

$$Au(x) = (2\pi)^{-n} \int_M \int_{\mathbb{R}^n} e^{i\varphi(x,y,\eta)} a(x, y, \eta) u(y) dy d\eta, \tag{2.29}$$

where $n = \dim M$, $dyd\eta$ is the symplectic volume measure on T^*M, $a \in S^m_{1,0}(M \times M \times \mathbb{R}^n)$ and $\varphi(x, y, \eta)$ is linear in ξ, vanishes on the diagonal and $d\varphi = \xi dx - \xi dy$ at (x, x, ξ) for $\xi \in T^*_x M$.

There are many possible choices of the phase φ and on a Riemannian manifold (M, g). It is natural to choose one which is adapted to the Riemannian metric g. In the case of Euclidean $\mathbb{R}^n$ the canonical choice is $\varphi(x, y, \eta) = \langle x - y, \eta \rangle$, and its generalization to Riemannian manifolds is

$$\varphi(x, y, \eta) = \langle \exp_y^{-1} x, \eta \rangle, \qquad \eta \in T_y^* M. \tag{2.30}$$

It is only defined for when the distance $r(x, y)$ is less than the injectivity radius and so the amplitude $a(x, y, \eta)$ is understood to be cutoff to a sufficiently small neighborhood of the diagonal so that $r(x, y) < \mathrm{inj}(M)$ on the support of a. Articles which give extensive discussions of geometric definitions of pseudo-differential operators on Riemannian manifolds include [**W, Sh, Sa**]. The phase (2.30) is particularly convenient when constructing Fourier integral representations for the wave kernel of (M, g).

In §4.7 we discuss the non-Euclidean Fourier transform on the hyperbolic plane and use it as a basis for the definition of pseudo-differential operators in the hyperbolic setting. This is a special case of pseudo-differential operators on a curved Riemannian manifold where the definition is adapted to the geometry.

The starting idea of pseudo-differential operators is that it is a quantization $a \to \mathrm{Op}(a)$ from classical observables (functions on T^*M) to operators on $L^2(M)$. In the case of $M = \mathbb{R}^n$ the symbols are functions on $T^*\mathbb{R}^n$ and may be identified with functions on the Heisenberg group. The corresponding quantization is the Schrödinger representation of the Heisenberg group, in its integrated form. This viewpoint originates in Hermann Weyl's classic text *Theory of Groups and Quantum Mechanics* and has been developed by E.M. Stein, R. Howe and many others. See for instance [**GrLS, H**].

The main advance in the definition of Kohn-Nirenberg quantization over the definition of Weyl is the restriction of the classical observables to *symbols*, i.e., polyhomogeneous functions on T^*M. They are generalizations of polynomials and thus the Schwartz kernel of the pseudo-differential operator is decomposed into a sum of products of a rapidly oscillating factor (the phase) and a slowly oscillating factor (the symbol). Symbols are constructed so that there exist semiclassical asymptotic expansions.

The theory is rich in applications in classical analysis as well as quantum mechanics. Many of the principal operators in classical analysis, e.g., Green's functions (resolvent kernels) are pseudo-differential operators. The symbol expansion in the homogeneous setting is equivalent to a singularity expansion of the kernel, e.g., along the lines of the Hadamard parametrix for the Green's function (see for instance [**Ho3**]).

Many of the theorems relate conditions on the symbol to conditions on the corresponding operator. For example, the vanishing of the symbol at infinity is related to compactness, integrability of the symbol is related to the trace class property, square integrability is related to the Hilbert-Schmidt property, boundedness of derivatives of the symbol is related to L^2 boundedness of the operator. We refer to the texts in the references for precise statements and proofs.

2.6. Symbols

There are many symbol classes and associated spaces of pseudo-differential operators. The Hörmander style definition defines symbols through the following estimate: On an open set $U \subset \mathbb{R}^n$, we say that $a(x,\xi;\hbar) \in C^\infty(U \times \mathbb{R}^n)$ is in the symbol class $S^{m,k}(U \times \mathbb{R}^n)$, provided

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi;\hbar)| \leq C_{\alpha\beta}\hbar^{-m}(1+|\xi|)^{k-|\beta|}, \tag{2.31}$$

for all multi-indices $\alpha, \beta \in \mathbb{N}^n$. When $m = 0$ the symbol is called polyhomogeneous of order k, i.e., it satisfies estimates similar to a polynomial of degree k. The space of such symbols is denoted Ψ^k_{phg}.

By a classical polyhomogeneous symbol of degree k we mean a symbol satisfying (2.31) which additionally possesses the asymptotic expansion

$$a(x,\xi) \sim \sum_{j=0}^\infty a_j(x,\xi) \text{ with } a_j(x,t\xi) = t^j a_j(x,\xi) \text{ for } |\xi| \geq 1. \tag{2.32}$$

That is, a polyhomogeneous symbol is asymptotic to a sum of homogeneous terms as $|\xi| \to \infty$ which descend in unit steps. The leading term a_k is called the principal symbol. The symbol is said to be *elliptic* if the top term a_k is nowhere vanishing on $T^*M - 0$. Here, $0 = 0_M$ denotes the zero section $\{(x,0)\} \subset T^*M$. Such symbols arise when one constructs Green's functions, i.e., inverses of (elliptic) pseudo-differential operators. The polyhomogeneous expansion of the symbol reflects the polyhomogeneous expansion of the singularity along the diagonal. See Chapter 17 of [**Ho3**] for the construction of a Hadamard parametrix for the Green's function in terms of every more regular homogeneous singular terms. Of course, there are many modifications of such symbol classes which play important roles in partial differential equations.

Quite often the semiclassical parameter $\hbar$ is the key parameter in the symbol and takes the place of order of homogeneity, or is in competition with it. We say that $a \in S^{m,k}_{\text{cl}}(U \times \mathbb{R}^n)$ is a semiclassical symbol of order m of classical type provided there exists an asymptotic expansion:

$$a(x,\xi;\hbar) \sim \hbar^{-m} \sum_{j=0}^\infty a_j(x,\xi)\hbar^j \text{ with } a_j(x,\xi) \in S^{0,k-j}(U \times \mathbb{R}^n). \tag{2.33}$$

Somewhat more generally, let M be a compact manifold. By a semiclassical symbol $a \in S^{m,k}(T^*M \times [0,h_0))$ we mean a smooth function possessing an asymptotic expansion as $h \to 0$ of the form

$$a(x,\xi,h) \sim_{h\to 0^+} \sum_{j=0}^\infty a_{k-j}(x,\xi)h^{m+j} \text{ with } a_{k-j} \in S^k_{1,0}(T^*M). \tag{2.34}$$

Here, $S^k_{1,0}$ is the standard Hörmander class consisting of smooth functions $a(x,\xi)$ satisfying the estimates $|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \leq C_{\alpha,\beta}\langle\xi\rangle^{k-|\beta|}$. Semiclassical symbols and pseudo-differential operators are discussed systematically in [**Zw**].

2.7. Quantization of symbols

The $\hbar$-quantization of a symbol is denoted by $\text{Op}_\hbar(a)$. It is defined locally by the Kohn-Nirenberg formula:

$$\text{Op}_\hbar(a)(x,y) = (2\pi\hbar)^{-n} \int_{\mathbb{R}}^n e^{i(x-y)\xi/\hbar} a(x,\xi,\hbar)\, d\xi. \tag{2.35}$$

By using a partition of unity, one constructs a corresponding class, $Op_\hbar(S^{m,k})$, of properly-supported $\hbar$-pseudodifferential operators acting globally on $C^\infty(M)$; as is well known, it is independent of the choice of partition of unity.

When the symbols are homogeneous then we denote the algebra of pseudo-differential operators by $\Psi^*(M)$ and the space of pseudo-differential operators of order m by $\Psi^m(M)$. Since there are many symbol classes relevant to the problems we study, we often just write Op in front of a symbol class to denote the corresponding class of pseudo-differential operators.

Weyl quantization is often more useful on $\mathbb{R}^n$. The Weyl quantization $\mathrm{Op}_h^w(a)$ (or simply a^w) of a symbol in $a \in S^{m,k}(T^*M \times [0, h_0))$ is defined by the Schwartz kernel

$$\mathrm{Op}_h^w(a)(x,y) = (2\pi h)^{-n} \int_{\mathbb{R}^n} e^{i\langle x-y,\xi\rangle/h} a\Big(\frac{x+y}{2}, \xi; h\Big)\, d\xi \tag{2.36}$$

with $a \in S^{m,k}$. By a symbol of order zero we mean that $a \in S^{0,0}$, and we refer to $a_0(x,\xi)$ as the principal symbol. In the latter case, we simply write

$$S^0_{\mathrm{sc}}(M) := S^{0,0} \quad \text{and} \quad \Psi^0_{\mathrm{sc}}(M) := \mathrm{Op}_h^w(S^{0,0}). \tag{2.37}$$

Weyl quantization is essentially the same as integration of the Schrödinger representation of the Heisenberg group. We refer to [**H, F**] for more details. As this indicates, it is more useful on $\mathbb{R}^n$ than on a general compact manifold and we do not use it in this monograph, but sometimes local computations on a manifold are simpler when using the Weyl formula.

2.8. Action of a pseudo-differential operator on a rapidly oscillating exponential

One of the key calculations is the action of a pseudo-differential operator $P(x,D)$ on a rapidly oscillating WKB Lagrangian state $ae^{i\tau\varphi}$. A semiclassical Lagrangian distribution is defined as an oscillatory integral

$$u(x,\hbar) = \hbar^{-n/2} \int_{\mathbb{R}^n} e^{\frac{i}{\hbar}\varphi(x,\theta)} a(x,\theta,\hbar)\, d\theta, \tag{2.38}$$

where $a(x,\theta,\hbar)$ is a semiclassical symbol. For instance it could be of classical type,

$$a(x,\theta,\hbar) \sim \sum_{k=0}^{\infty} \hbar^{\mu+k} a_k(x,\theta). \tag{2.39}$$

We refer to [**T1, Ho3**] for background.

PROPOSITION 2.5. *Let φ be real-valued and suppose that $d\varphi \neq 0$ (anywhere). Then*

$$P(x,D)(a(x,\xi)e^{i\tau\varphi(x,\xi)}) = a_P(x,\xi)e^{i\tau\varphi(x,\xi)}$$

with

$$a_P(x,\xi) \sim \sum_\alpha \frac{1}{\alpha!} D^\alpha p(x, \tau d\varphi) \mathcal{N}_\alpha(\varphi, \tau, D_x) a(x,\xi), \tag{2.40}$$

where

$$\mathcal{N}_\alpha(\varphi,\tau,D_x)u(x) = D_y^\alpha e^{i\tau\varphi_{(2)}(y,\xi)} u(y)|_{y=x}, \tag{2.41}$$

$\varphi_{(2)}$ being the second order remainder of the Taylor expansion of $\varphi(y,\xi)$ around $y = x$:

$$\varphi(x) - \varphi(y) = \nabla\varphi_x \cdot (x - y) + \varphi_{(2)}(x, y). \tag{2.42}$$

2.8.1. Symbol composition. An application of Proposition 2.5 is the asymptotic formula for the composition of two symbols. Given $a \in S^{m_1,k_1}$ and $b \in S^{m_2,k_2}$, the composition is given by

$$\mathrm{Op}_\hbar(a) \circ \mathrm{Op}_\hbar(b) = \mathrm{Op}_\hbar(c) + \mathcal{O}(\hbar^\infty) \tag{2.43}$$

in $L^2(M)$ where locally,

$$c(x,\xi;\hbar) \sim \hbar^{-(m_1+m_2)} \sum_{|\alpha|=0}^{\infty} \frac{(-i\hbar)^{|\alpha|}}{\alpha!} (\partial_\xi^\alpha a) \cdot (\partial_x^\alpha b). \tag{2.44}$$

The formula shows that pseudo-differential operators of the types mentioned above form algebras.

Bibliography

[BW] S. Bates and A. Weinstein, *Lectures on the geometry of quantization.* Berkeley Mathematics Lecture Notes, 8. American Mathematical Society, Providence, RI; Berkeley Center for Pure and Applied Mathematics, Berkeley, CA, 1997.

[DSj] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semiclassical limit.* London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.

[Du] J. J. Duistermaat, *Fourier integral operators*, reprint of the 1996 edition, based on the original lecture notes published in 1973, Modern Birkhäuser Classics, Birkhäuser/Springer, New York, 2011. MR2741911

[F] G. B. Folland, *Harmonic analysis in phase space.* Annals of Mathematics Studies, 122. Princeton University Press, Princeton, NJ, 1989.

[GSj] A. Grigis and J. Sjöstrand, *Microlocal analysis for differential operators. An introduction.* London Mathematical Society Lecture Note Series, 196. Cambridge University Press, Cambridge, 1994.

[GrLS] A. Grossmann, G. Loupias, G. and E. M. Stein, An algebra of pseudodifferential operators and quantum mechanics in phase space. Ann. Inst. Fourier (Grenoble) 18 1968 fasc. 2, 343-368, viii (1969).

[GuSt1] V. Guillemin and S. Sternberg, Geometric asymptotics. Mathematical Surveys, No. 14. American Mathematical Society, Providence, R.I., 1977.

[GuSt2] V. Guillemin and S. Sternberg, Semiclassical Analysis, International Press, Boston, MA, 2013.

[Ho1] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[Ho2] L. Hörmander, *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Classics in Mathematics. Springer-Verlag, Berlin, 2005.

[Ho3] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[Ho4] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[H] R. Howe, Quantum mechanics and partial differential equations. J. Funct. Anal. 38 (1980), no. 2, 188-254.

[Kl] W. Klingenberg, *Lectures on Closed Geodesics*, Grundlehren der. math. W. **230**, Springer-Verlag (1978).

[Sa] Yu. G. Safarov, Pseudodifferential operators and linear connections, Proc. London Math. Soc. (3), v. 74 (1997) .379-416.

[Sh] V. A. Sharafutdinov, Geometric symbol calculus for pseudodifferential operators. I [Translation of Mat. Tr. 7 (2004), no. 2, 159206; MR2124544]. Siberian Adv. Math. 15 (2005), no. 3, 81125.

[So] C. D. Sogge, *Fourier integrals in classical analysis.* Cambridge Tracts in Mathematics, 105. Cambridge University Press, Cambridge, 1993.

[St] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[Su] T. Sunada, Trace formula and heat equation asymptotics for a nonpositively curved manifold. Amer. J. Math.104 (1982), no. 4, 795?812.

[T1] M.E. Taylor, *Pseudodifferential operators.* Princeton Mathematical Series, 34. Princeton University Press, Princeton, N.J., 1981.

[T2] M. E. Taylor, *Partial differential equations I. Basic theory.* Second edition. Applied Mathematical Sciences, 115. Springer, New York, 2011.

[T3] M. E. Taylor, *Partial differential equations II. Qualitative studies of linear equations.* Second edition. Applied Mathematical Sciences, 116. Springer, New York, 2011.

[W] H. Widom, *A complete symbolic calculus for pseudodifferential operators*, Bull. Sci. Math. (2), v. 104 (1981) 19-63.

[Z] S. Zelditch, *Gaussian beams on Zoll manifolds and maximally degenerate Laplacians*, Spectral theory and partial differential equations, 169–197, Contemp. Math., 640, Amer. Math. Soc., Providence, RI. MR3381021

[Zw] M. Zworski, *Semiclassical analysis.* Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012.

CHAPTER 3

Main results

To clarify the organization of this monograph we give a rapid survey of the relatively new results that form the main content of this monograph. There are many older results, such as norm estimates based on local Weyl laws, or global quantum ergodicity theorems, that we do not discuss in this section, although they are (briefly) reviewed in later sections.

All of the results concern the oscillation and concentration of an orthonormal basis of eigenfunctions. Oscillation is measured by matrix elements or by restricted matrix elements. Concentration is measured by restricted matrix elements or by L^p norms of restrictions to hypersurfaces or thin tubes around hypersurfaces.

An important heuristic (and technical) principle is that results which only make use of the short time behavior of the geodesic flow or wave group are universal. This is because the singularity at $t = 0$ is universal. To break universality one needs to exploit the long time behavior. For instance, curvature assumptions (such as nonpositive curvature) or dynamical assumptions (such as ergodicity or complete integrability) require the use of the wave kernel for long times. The main obstacle to using the long time behavior of the wave kernel and its relation to the geodesic flow is that many estimates blow up exponentially in the time parameter t, i.e., like $e^{C|t|}$, and one can only work up to the Eherenfest time.

Recall that the $(-\Delta)$-eigenvalues are denoted by λ^2 and frequencies by λ and that all eigenfunctions are understood to satisfy the normalization condition $\|\varphi\|_{L^2} = 1$.

3.1. Universal L^p bounds

In the ♣Chapter♣ on L^p norms, we review the universal bounds of Sogge [**S1**] on L^p norms of L^2-normalized eigenfunctions on compact Riemannian manifolds. These bounds are achieved by extremal eigenfunctions on standard spheres but are rarely achieved on general Riemannian manifolds. One of the themes of this monograph is the exploration of converses to extremal L^p norm growth.

In the case of a compact surface (i.e., a Riemannian manifold of dimension $n = 2$), Sogge's universal L^p bounds assert that for $\lambda \geq 1$,

$$\|\varphi_\lambda\|_{L^p(M)} \leq C\lambda^{\frac{1}{2}(\frac{1}{2}-\frac{1}{p})}\|e_\lambda\|_{L^2(M)}, \qquad 2 \leq p \leq 6, \tag{3.1}$$

$$\|\varphi_\lambda\|_{L^p(M)} \leq C\lambda^{(\frac{1}{2}-\frac{1}{p})-\frac{1}{2}}\|\varphi_\lambda\|_{L^2(M)}, \quad 6 \leq p \leq \infty. \tag{3.2}$$

In particular, for a general compact Riemannian manifold of any dimension,

$$\|\varphi_\lambda\|_{L^\infty(M)} \leq C\lambda^{\frac{n-1}{2}}, \tag{3.3}$$

and in dimension 2,

$$\|\varphi_\lambda\|_{L^4(M)} \leq C\lambda^{\frac{1}{8}}\|\varphi_\lambda\|_{L^2(M)}. \tag{3.4}$$

These estimates are sharp for the round sphere S^2. As discussed in ♣§4.4♣, the estimate (3.1) is achieved by the so-called highest weight spherical harmonics (also known as Gaussian beams). Estimate (3.2) is achieved by zonal spherical harmonics (and more generally zonal functions for surfaces of revolution), which concentrate at the poles.

3.2. Self-focal points and extremal L^p bounds for high p

On which manifolds are the above universal sup norm bounds achieved by some sequence of eigenfunctions? To date, the only known examples in dimension two are surfaces of revolution, where the sup norm bounds are achieved by rotationally invariant eigenfunctions at the poles. In higher dimensions the only known examples are compact rank one symmetric spaces (e.g., complex projective space $\mathbb{CP}^n$), or on higher dimensional surfaces of revolution, where the metric is invariant under the action of $SO(n-1)$. The poles have the property that all geodesics emanating from the poles return to the poles at time $t = 2\pi$ and close up smoothly.

In fact, any example that saturates the sup norm bounds has strong similarities to such surfaces of revolution. To state the theorem precisely, we introduce some notation. For a given $x \in M$, let $\mathcal{L}_x \subset S^*_x M$ denote those unit directions ξ for which $G^t(x, \xi) \in S^*_x M$ for some time $t \neq 0$, and let $|\mathcal{L}_x|$ denote its surface measure $d\mu_x$ in $S^*_x M$ induced by the Euclidean metric g_x. Thus, $\mathcal{L}_x$ denotes the initial directions of geodesic loops through x.

In [**SZ2**], it is proved that for a compact Riemannian manifold (M, g), if $|\mathcal{L}_x| = 0$ for all $x \in M$ then

$$\|\varphi_\lambda\|_{L^\infty(M)} = o(\lambda^{\frac{n-1}{2}}), \tag{3.5}$$

A recent stronger result is obtained by the author and Sogge when the manifold is assumed to be real analytic. In this case, there are just two extreme possibilities regarding the nature of the loop directions $\mathcal{L}_x \subset S^*_x M$: either $|\mathcal{L}_x| = 0$ or $\mathcal{L}_x = S^*_x M$. In the second case there is also a minimal time $\ell > 0$ so that $G^\ell(x, \xi) \in S^*_x M$, meaning that all geodesics starting at x loop back at exactly this minimal time ℓ. We call such a point x a *self-focal point*. There may exist more than one self-focal point and the minimal common return time ℓ of the loops may depend on x, but for simplicity of notation we do not embellish ℓ with a subscript. If we then write

$$G^\ell(x, \xi) = (x, \eta_x(\xi)), \quad \xi \in S^*_x M, \tag{3.6}$$

then the *first return map*,

$$\eta_x : S^*_x M \to S^*_x M \tag{3.7}$$

above our self-focal point is real analytic. Following Safarov [**Sa, SaV**], we can associate to this first return map the *Perron-Frobenius operator* $U_x \colon L^2(S^*_x M, d\mu_x) \to L^2(S^*_x M, d\mu_x)$ by setting

$$U_x f(\xi) = f(\eta_x(\xi)) \sqrt{J_x(\xi)} \quad \text{for all } f \in L^2(S^*_x M, d\mu_x), \tag{3.8}$$

where $J_x(\xi)$ denotes the Jacobian of the first return map, that is, $\eta_x^* d\mu_x = J_x(\xi) d\mu_x$. Clearly U_x is a unitary operator and

$$\eta_x^*(f d\mu_x) = U_x(f) d\mu_x. \tag{3.9}$$

The key assumption underlying the result is contained in the following:

Definition 3.1. A self-focal point $x \in M$ is said to be *dissipative* if U_x has no invariant function $f \in L^2(S^*_x M)$. Equivalently, G^ℓ has no invariant L^1 measure with respect to $d\mu_x$.

If the above is satisfied, then the following is proved in [**SZ4, SZ5**]:

Theorem 3.2. *Let (M, g) be a real analytic compact manifold without boundary. In dimensions $n \geq 2$, the L^∞ norm estimate* (3.5) *holds only if every self-focal point is dissipative. In dimension $n = 2$, the estimate* (3.5) *can only hold if there exists a pole p so that all geodesics through p are closed.*

3.3. Low L^p norms and concentration of eigenfunctions around geodesics

For the smaller L^p norms ($p \leq 6$ in dimension 2), the analogue of Theorem 3.2 would be that if the maximal L^p norm is achieved, then there should exist Gaussian beams concentrating on elliptic closed geodesics. An elliptic closed geodesic is one for which the eigenvalues of the linear Poincaré map are of modulus one. A Gaussian beam is similar to a highest weight spherical harmonic. This conjectural picture is however open at the present time.

To test this conjecture it is useful to study integrals of squares of eigenfunctions over unit length geodesic segments γ or over thin tubes around such geodesics. Following Cordoba, Bourgain and Sogge we consider two types of maximal functions which measure concentration of eigenfunctions along geodesics: (i) the geodesic maximal function $\mathcal{G}_p$, which measures L^p norms on the geodesic and (ii) the Kakeya-Nikodym maximal function $\mathcal{M}_p$, which measures L^p norms in λ^{-1}-tubes around geodesics.

Definition 3.3. Let Π denote the set of geodesic segments of length one. The L^p geodesic maximal function is defined by

$$\mathcal{G}_p \varphi_j(x) = \sup_{\gamma \in \Pi :\, x \in \gamma} \left(\int_\gamma |\varphi_j|^p \, ds \right)^{\frac{1}{p}}, \tag{3.10}$$

with ds denoting the arc length measure along γ. When $p = 2$, we simply write $\mathcal{G} = \mathcal{G}_2$.

Universal bounds on $\mathcal{G}_p \varphi_j$ was proved by Burq, Gérard and Tzvetkov [**BuGT**]:

Theorem 3.4. *For any compact surface without boundary, one has*

$$\mathcal{G}_p \varphi_\lambda \leq \begin{cases} C\lambda^{\frac{1}{2}-\frac{1}{p}} \|\varphi_\lambda\|_{L^2(M)} & 4 \leq p \leq \infty, \\ C\lambda^{\frac{1}{4}} \|\varphi_\lambda\|_{L^2(M)} & 2 \leq p \leq 4, \end{cases} \tag{3.11}$$

with C independent of λ and $\gamma \in \Pi$.

For $2 \leq p \leq 4$, these bounds are achieved by the highest weight spherical harmonics on S^2. If one makes additional assumptions on the curvature of the surface, however, C. Sogge and the author proved that when $p = 2$ the bound for $\mathcal{G}\varphi_\lambda$ is not achieved [**SZ1**]:

Theorem 3.5. *For any compact surface of nonpositive curvature, given $\varepsilon > 0$ there is a $\lambda(\varepsilon) < \infty$ so that*

$$\mathcal{G}_2 \varphi_\lambda \leq \varepsilon \lambda^{\frac{1}{4}} \|\varphi_\lambda\|_{L^2(M)} \quad \textit{for all } \lambda > \lambda(\varepsilon). \tag{3.12}$$

In [**CS**] X. Chen and C. Sogge showed that the other bound of Theorem 3.4 also fails to be sharp when the curvature is non-positive:

THEOREM 3.6. *For any compact surface of nonpositive curvature, one has*

$$\limsup_{\lambda_j \to \infty} \lambda_j^{-\frac{1}{4}} \mathcal{G}_4 \varphi_{\lambda_j} = 0 \tag{3.13}$$

One of the principal applications of the geodesic maximal function is to control L^4 norms. Bourgain [**Bo**] proves the following:

THEOREM 3.7. *For any compact surface without boundary, if $p \geq 2$ and $\gamma \in \Pi$, then*

$$\left(\int_\gamma |\varphi_\lambda|^2 \, ds \right)^{\frac{1}{2}} \leq C_\gamma \lambda^{\frac{1}{2p}} \|\varphi_\lambda\|_{L^p(M)}. \tag{3.14}$$

Conversely, for any $\varepsilon > 0$,

$$\|\varphi_{\lambda_j}\|_{L^4(M)} \leq C_\varepsilon \lambda_j^{\frac{1}{16}+\varepsilon} \left(\mathcal{G}\varphi_{\lambda_j} \right)^{\frac{1}{2}}. \tag{3.15}$$

Similar to the sup norm estimates, these bounds are achieved on the standard sphere (by highest weight spherical harmonics), but can be improved under curvature assumptions on the surface. In [**S2**], using in part results from Bourgain [**Bo**], it was shown that for a compact surface without boundary and $2 < p < 6$,

$$\|\varphi_{\lambda_j}\|_{L^p(M)} = o(\lambda_j^{\frac{1}{2}(\frac{1}{2}-\frac{1}{p})}) \quad \text{if and only if} \quad \mathcal{G}\varphi_{\lambda_j} = o(\lambda_j^{\frac{1}{4}}). \tag{3.16}$$

Thus, we have the following corollary to Theorem 3.5.

THEOREM 3.8. *Let M a compact surface with nonpositive curvature. For $\varepsilon > 0$ and $2 < p < 6$ fixed there exists $\lambda(\varepsilon, p) < \infty$ so that*

$$\|\varphi_\lambda\|_{L^p(M)} \leq \varepsilon \lambda^{\frac{1}{2}(\frac{1}{2}-\frac{1}{p})} \|\varphi_\lambda\|_{L^2(M)} \quad \textit{for all } \lambda > \lambda(\varepsilon, p). \tag{3.17}$$

3.4. Kakeya-Nikodym maximal function and extremal L^p bounds for small p

Closely related to the geodesic maximal function is the L^p Kakeya-Nikodym maximal function. It is the maximal function over $\lambda^{-\frac{1}{2}}$-tubes around geodesic arcs. Recall that the frequency of an eigenfunction φ_λ is $\lambda^{-1} = h$, so the radius of the tube is the square root of the frequency. A highest weight spherical harmonic is essentially supported in a tube of this radius.

DEFINITION 3.9. Let Π denote the set of geodesic segments of length one. For $\gamma \in \Pi$, define the tubular neighborhood

$$\mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma) = \{ y \in M : \ d(y, \gamma) \leq \lambda^{-\frac{1}{2}} \} \tag{3.18}$$

of radius $\lambda^{-\frac{1}{2}}$ about γ. The L^p Kakeya-Nikodym maximal function is defined by

$$\mathcal{M}_p \varphi_\lambda = \sup_{\gamma \in \Pi : \, x \in \gamma} \frac{1}{\mathrm{Vol}(\mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma))} \left(\int_{\mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma)} |\varphi_\lambda|^p \, dV \right)^{\frac{1}{p}}. \tag{3.19}$$

When $p = 2$, we simply write $\mathcal{M} = \mathcal{M}_2$.

One of the advantages of the Kakeya-Nikodym tubular average is that automatically $\mathcal{M}\varphi_j \leq 1$. Bourgain and Sogge have proved bounds on L^4 norms of eigenfunctions in terms of the L^∞ norm of $\mathcal{M}\varphi_j$. In [**S2**] C. Sogge proved the following.

THEOREM 3.10. *Given a compact surface without boundary, for all $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that*

$$\|\varphi_\lambda\|_{L^4}^4 \leq \varepsilon\lambda^{\frac{1}{2}}\|\varphi_\lambda\|_{L^2}^4 + C_\varepsilon\lambda^{\frac{1}{2}}|\varphi_\lambda|_{L^2}^2\left(\mathcal{M}\varphi_\lambda\right)^2 + C\|\varphi_\lambda\|_{L^2(M)}^4, \tag{3.20}$$

where the constant C is independent of λ and ε.

(Recall again that the universal bound (3.4) says $\|\varphi_\lambda\|_{L^4(M)} \leq \lambda^{\frac{1}{8}}$.) In [**BS1, BS2**] M. Blair and C. Sogge proved a kind of higher dimensional generalization:

THEOREM 3.11. *Let M be an n-dimensional compact Riemannian manifold without boundary. If $\frac{2(n+2)}{n} < p < \frac{2(n+1)}{n-1}$, then*

$$\|\varphi_\lambda\|_{L^p}^p \leq \varepsilon\lambda^{p(\frac{n-1}{2})(\frac{1}{2}-\frac{1}{p})}\|\varphi_\lambda\|_{L^2}^p + C_\varepsilon\|\varphi_\lambda\|_{L^2}^p + C_\varepsilon\lambda^{p(\frac{n-1}{2})(\frac{1}{2}-\frac{1}{p})}\|\varphi_\lambda\|_{L^2}^2\left(\mathcal{M}\varphi_\lambda\right)^{\frac{n-2}{2}}. \tag{3.21}$$

In general, the sequences of eigenfunctions saturating the L^p bounds are sparse (usually forming an arithmetic progression in the sequence of frequencies, i.e. square roots of eigenvalues). We say that a sequence is sparse if the complementary sequence λ_{j_k} has density 1 in the sense that

$$\lim_{\lambda\to\infty}\frac{\#\{\lambda_{j_k} \leq \lambda\}}{N(\lambda)} = 1. \tag{3.22}$$

(Recall that $N(\lambda) = \#\{\lambda_j \leq \lambda\}$). The following Theorem uses the Kakeya-Nikodym maximal function to relate this sparsity to the measure of the set of periodic geodesics. More precisely, let $G^t : S^*M \to S^*M$ be the geodesic flow on the cosphere bundle, assume that the set of periodic points has measure zero, i.e..

$$\{(x,\xi) \in S^*M \colon G^t(x,\xi) = (x,\xi) \text{ for some } t > 0\} \tag{3.23}$$

has measure zero in S^*M with respect to the volume element. The following is proved in [**SZ1**].

THEOREM 3.12. *Let (M,g) be a compact surface. Assume that the set of periodic points of the geodesic flow has measure zero in S^*M with respect to the volume element. Then if φ_j is any orthonormal basis of eigenfunctions, there is a subsequence λ_{j_k} of density 1 so that*

$$\|\varphi_{\lambda_{j_k}}\|_{L^4(M)} = o(\lambda_{j_k}^{\frac{1}{8}}). \tag{3.24}$$

3.5. Concentration of joint eigenfunctions of quantum integrable Δ around closed geodesics

We recall that a Laplacian or Schrödinger operator is quantum completely integrable (QCI) if it commutes with a full set of independent observables (that is, $n-1$ other observables in dimension n). Independent means that the symbols are functionally independent on an open dense set of T^*M. Joint eigenfunctions are simultaneous eigenfunctions of the commuting observables. In the quantum integrable case, we may explicitly calculate the L^2 norms of restrictions of joint eigenfunctions to geodesics or to tubes around closed geodesics. The norm depends on whether the closed geodesic is elliptic or hyperbolic.

One of the principal reasons for studying quantum integrable systems is that they provide (micro-)local models for all problems involving eigenfunctions. To make this statement precise is the purpose of quantum Birkhoff normal forms. Moreover we expect eigenfunctions of quantum integrable systems to provide extremals for concentration and L^p norm problems. This expectation has led to some rigorous theorems but remains a useful heuristic principle.

The first type of result concerns *mass concentration on small length scales* of eigenfunctions of quantum completely integrable eigenfunctions. These are joint eigenfunctions of Δ and of m commuting pseudo-differential operators of order 1. It will be discussed in the ♣chapter on quantum integrable systems♣. Mass concentration on small scales refers to the local L^2 norm of the eigenfunctions around invariant sets Λ for the geodesic flow or more precisely, level sets of the moment map. We are often interested in (closed) geodesics but in the integrable case is it also natural to work with the projection $\pi(\Lambda)$ to M of an invariant phase space torus under $\pi\colon T^*M \to M$.

We use the following geometric terminology:

- A point (x, ξ) is called a singular point of the moment map $\mathcal{P}$ if $dp_1 \wedge \cdots \wedge dp_n(x, \xi) = 0$.
- A level set $\mathcal{P}^{-1}(c)$ of the moment map is called a singular level if it contains a singular point $(x, \xi) \in \mathcal{P}^{-1}(c)$.
- An orbit $\mathbb{R}^n \cdot (x, \xi)$ of G^t is singular if it is non-Lagrangian, i.e., has dimension strictly less than n;

We denote by $T_\varepsilon(\pi(\Lambda))$ an ε-neighborhood of $\pi(\Lambda)$. For $0 < \delta < \frac{1}{2}$, we introduce a cutoff $\chi_1^\delta(x; \hbar) \in C_0^\infty(M)$ with $0 \leq \chi_1^\delta \leq 1$, satisfying

- (i) supp $\chi_1^\delta \subset T_{\hbar^\delta}(\pi(\Lambda))$
- (ii) $\chi_1^\delta = 1$ on $T_{3/4\hbar^\delta}(\pi(\Lambda))$.
- (iii) $|\partial_x^\alpha \chi_1^\delta(x; \hbar)| \leq C_\alpha \hbar^{-\delta|\alpha|}$.

The following result says that the mass of the joint eigenfunctions φ_μ which microlocally concentrate on Λ must blow up at a logarithmic rate on the projection of Λ. The following is proved in [**TZ2**].

THEOREM 3.13. *Let $\{\varphi_\mu\}$ be joint QCI eigenfunctions of the Laplacian Δ of (M, g), where Δ is quantum completely integrable and* $\dim M = m$. *Then for any $0 \leq \delta < \frac{1}{2}$, we have $\langle \mathrm{Op}_\hbar(\chi_1^\delta)\varphi_\mu, \varphi_\mu\rangle \gg |\log \hbar|^{-m}$.*

Here, $A \gg B$ means that there exists a constant c independent of the eigenvalue so that $A \geq cB$. The main idea of the proof is to conjugate to the model by a semiclassical $\hbar$-Fourier Integral Operator. This reduces the calculation of matrix elements to those $(\mathrm{Op}_\hbar(\chi_2^\delta \circ \kappa)u_\mu, u_\mu)$ in the model. Model eigenfunctions are explicit ones and the calculation gives the result above. A special case is where $\Lambda = \gamma$ is a hyperbolic closed geodesic. We denote by $\pi(\gamma)$ be the image of γ in M under the natural projection map.

THEOREM 3.14. *Let γ be a hyperbolic closed orbit in (M, g) with quantum integrable Δ_g, and let $\{\varphi_\mu\}$ be a sequence of joint eigenfunctions whose joint eigenvalues concentrate on the image of γ under the moment map. Then for any $0 \leq \delta < \frac{1}{2}$, we have*

$$\lim_{\hbar \to 0} \langle \mathrm{Op}_\hbar(\chi_1^\delta)\varphi_\mu, \varphi_\mu\rangle \geq (1 - 2\delta). \tag{3.25}$$

These mass concentration results suggest that quantum completely integrable joint eigenfunctions are extremals for various L^p problems. We now give some applications of the results which may be easier to absorb than the original statements.

One extremal problem is to determine the Riemannian manifolds which possess orthonormal bases of eigenfunctions with uniformly bounded L^∞ norms. The following result shows that flat tori are the unique minimizers in the class of quantum completely integrable systems (cf. [**TZ1**]). Define

$$L^\infty(\lambda, g) = \sup_{\substack{\varphi \in V_\lambda \\ \|\varphi\|_{L^2}=1}} \|\varphi\|_{L^\infty} \quad \text{and} \quad \ell^\infty(\lambda, g) = \inf_{\{\varphi_j\} \subset V_\lambda} (\sup_{j=1,\dots,\dim V_\lambda} \|\varphi_j\|_{L^\infty}). \tag{3.26}$$

Note that ℓ^∞ measures the growth of the smallest possible L^∞-norms for any orthonormal basis, while L^∞ measures the largest growth of any sequence of eigenfunctions. For instance, on a rational flat torus, there exist sequences which tend to infinity at the rate of the square root of the multiplicity of the eigenspace, so L^∞ grows at that rate, but the standard ONB is uniformly bounded by 1 and so $\ell^\infty = 1$.

THEOREM 3.15. *Suppose that Δ is a quantum completely integrable Laplacian on a compact Riemannian manifold (M, g). Let V_λ be the eigenspace of eigenvalue $-\lambda^2$. Then*

(a) *If $L^\infty(\lambda, g) = O(1)$ then (M, g) is flat.*
(b) *If $\ell^\infty(\lambda, g) = O(1)$, then (M, g) is flat.*

There exists a quantitative improvement giving blow-up rates for L^p norms for quantum integrable eigenfunctions concentrating on singular level sets, i.e., level sets which are not regular tori. These eigenfunctions are the extremals for L^p blow-up and mass concentration. The following is proved in [**TZ2**].

THEOREM 3.16. *Suppose that (M, g) is a compact Riemannian manifold of dimension n whose Laplacian Δ is quantum completely integrable. Then, unless (M, g) is a flat torus, the underlying Hamiltonian $\mathbb{R}^n$ action must have a singular orbit of dimension strictly less than n. If the minimal dimension of the singular orbits is ℓ, then for every $\varepsilon > 0$, there exists a sequence of eigenfunctions satisfying:*

$$\begin{cases} \|\varphi_j\|_{L^\infty} \geq C_\varepsilon \lambda_j^{\frac{n-\ell}{4} - \varepsilon} \\ \|\varphi_j\|_{L^p} \geq C_\varepsilon \lambda_j^{\frac{(n-\ell)(p-2)}{4p} - \varepsilon} \quad \textit{for } 2 < p < \infty. \end{cases} \tag{3.27}$$

The idea is to measure local L^2 mass on shrinking tubes around special subsets of M. They are the projections of special singular level sets of the moment map of the underlying integrable system from $S^*M \to M$. Except in the case of a flat torus, singular levels such as closed geodesics always occur.

The general theme of these results is that 'flat eigenfunctions' only occur on flat manifolds. But it is only proved for the very special case of quantum integrable Laplacians. The problem is open in the general setting of compact Riemannian manifolds. Later we will consider weak* limits of matrix elements $\int f \varphi_j^2 \, dV_g$. When the φ_j are uniformly bounded, the limit measures (microlocal defect measures, quantum limits) are invariant measures for the geodesic flow whose projections to M have the form ρdV_g where ρ is uniformly bounded. This is of course the case for

Liouville measure and possibly for other invariant measures. So this characterization does not aid much in characterizing manifolds with flat eigenfunctions.

3.6. Quantum ergodic restriction theorems for Cauchy data

We now consider concentration on hypersurfaces $H \subset M$ for eigenfunctions when the geodesic flow is ergodic. The main result is that for generic hypersurfaces, the restriction of ergodic eigenfunctions on M are ergodic along H. In particular, the L^2 norms of their restrictions to H are bounded below by a positive constant. In general, this is not true for restrictions but is true for Cauchy data of eigenfunctions. The quantum ergodic restriction theorem has numerous applications to the geometry of nodal sets, as will be discussed below.

We are interested in the QER (quantum ergodic restriction) theorem both on manifolds with boundary and manifolds without boundary. The simplest cases occur when $H = \partial M$ (on a manifold with boundary) or when $H \cap \partial M = \emptyset$ (on a manifold with or without boundary).

Following [**CTZ**], we use semiclassical notation for eigenfunctions and eigenvalues: we denote the eigenfunctions in the orthonormal basis by φ_h and the eigenvalues by h^{-2}, and consider the eigenvalue problem taking the semiclassical form

$$\begin{cases} (-h^2\Delta_g - 1)\varphi_h = 0, & \\ B\varphi_h = 0 & \text{on } \partial M, \end{cases} \tag{3.28}$$

where $B = I$ or $B = hD_\nu$ in the Dirichlet or Neumann cases, respectively.

Let $H \subset M$ be a smooth hypersurface which either equals ∂M or does not meet ∂M if ∂M (including the case $\partial M = \emptyset$). The quantum ergodic restriction problem for (semiclassical) Cauchy data

$$\mathrm{CD}(\varphi_h) := \{(\varphi_h|_H,\ hD_\nu\varphi_h|_H)\} \tag{3.29}$$

is to prove that the Cauchy data is quantum ergodic along any hypersurface $H \subset M$ if the eigenfunctions are quantum ergodic on the global manifold M. The Cauchy data is the quantum analogue of the 'cross-section' $S^*_H M$ to the geodesic flow in S^*M. In the case $H = \partial M$ on a manifold with boundary, the Dirichlet (resp. Neumann) boundary condition of course kills one of the two components of the Cauchy data.

In the case where $\partial M = \emptyset$, we assume that H is separating in the sense that

$$M \backslash H = M_+ \cup M_- \tag{3.30}$$

where $M_\pm$ are domains with boundary in M. This is not a restrictive assumption since we can arrange that any hypersurface is part of the boundary of a domain.

We recall that a sequence of functions u_{h_j} on a manifold M indexed by a sequence of Planck constants is said to be quantum ergodic with limit measure $d\mu$ if

$$\langle a^w(x, h_j D_x) u_j, u_j\rangle \to \omega(a_0) := \int_{T^*M} a_0 \, d\mu, \tag{3.31}$$

for all zeroth order semiclassical pseudo-differential operators, where a_0 is the principal symbol of $a^w(x, h_j D_x)$. We also recall that the functional $a \to \langle a^w(x, h_j D_x) u_j, u_j\rangle$ is referred to as a microlocal lift (or a Wigner distribution), and the limit measure or state $\omega(a)$ is called a quantum limit or a semiclassical defect measure. An orthonormal basis of eigenfunctions $\{\varphi_h\}$ of $-h^2\Delta_g$ is called QUE (quantum unique ergodic) on M if $\langle a^w(x, hD_x)\varphi_h, \varphi_h\rangle \to \omega(a_0)$ for all pseudo-differential operators $a^w \in \Psi^0(M)$, i.e., if it is not necessary to eliminate a sparse subsequence of eigenfunctions of density zero.

Given a quantization $a \to \mathrm{Op}_h(a)$ of semiclassical symbols $a \in S^0_{sc}(h)$ of order zero (see §2.7) to semiclassical pseudo-differential operators on $L^2(H)$, we define the microlocal lifts of the Neumann data as the restricted versions of the definition on the global M. Namely, they are linear functionals on $a \in S^0_{sc}(B^*H)$ given by

$$\mu_h^N(a) := \int_{B^*H} a\, d\Phi_h^N := \langle \mathrm{Op}_H(a) hD_\nu \varphi_h|_H, hD_\nu\varphi_h|_H\rangle_{L^2(H)}\,, \tag{3.32}$$

where B^*H is the unit co-ball bundle of H. It is the orthogonal projection to T^*H of $S^*_H M$, the unit convectors with footpoint on H. We also define the *re-normalized microlocal lifts* of the Dirichlet data by

$$\mu_h^{RD}(a) := \int_{B^*H} a\, d\Phi_h^{RD} := \left\langle \mathrm{Op}_H(a)(1 + h^2\Delta_H)\varphi_h|_H, \varphi_h|_H\right\rangle_{L^2(H)}. \tag{3.33}$$

Finally, we define the microlocal lift $d\Phi_h^{CD}$ of the Cauchy data to be the sum

$$d\Phi_h^{CD} := d\Phi_h^N + d\Phi_h^{RD}. \tag{3.34}$$

Here, $h^2\Delta_H$ denotes the negative tangential Laplacian for the induced metric on H, so that the operator $(1 + h^2\Delta_H)$ is characteristic precisely on the glancing set S^*H of H. Intuitively, we have re-normalized the Dirichlet data by damping out the whispering gallery components.

The distributions μ_h^N and μ_h^{RD} are asymptotically positive, but are not normalized to have mass one and may tend to infinity. They depend on the choice of quantization, but their possible weak* limits as $h \to 0$ do not, and subsequent theorems are valid for any choice of quantization. We refer to §2.7 or to [**Zw**] for background on semiclassical microlocal analysis.

The next result from [**CTZ**] states that the Cauchy data of a sequence of quantum ergodic eigenfunctions restricted to H is automatically QER for semiclassical pseudo-differential operators with symbols vanishing on the *glancing set* S^*H. The assumption $H \cap \partial M = \emptyset$ is for simplicity of exposition and because the case $H = \partial M$ was proved earlier and by a different method in [**GeL, HZ, Bu**] at different levels of generality.

THEOREM 3.17. *Suppose $H \subset M$ is a smooth, codimension 1 embedded orientable separating hypersurface and assume $H \cap \partial M = \emptyset$. Assume that $\{\varphi_h\}$ is a quantum ergodic sequence of eigenfunctions* (3.28). *Then the sequence $\{d\Phi_h^{CD}\}$ of microlocal lifts* (3.34) *of the Cauchy data of φ_h is quantum ergodic on H in the sense that for any $a \in S^0_{sc}(H)$,*

$$d\Phi_h^{CD}(a) := d\Phi_h^N(a) + d\Phi_h^{RD}(a) \xrightarrow{h\to 0} \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{\frac{1}{2}}\, d\sigma, \tag{3.35}$$

where $a_0(x', \xi')$ is the principal symbol of $\mathrm{Op}_H(a)$, $h^2\Delta_H$ is the induced tangential (semiclassical) Laplacian with principal symbol $|\xi'|^2$, μ is the Liouville measure on S^*M, and $d\sigma$ is the standard symplectic volume form on B^*H.

The limit along H in Theorem 3.17 holds for the full sequence $\{\varphi_h\}$. Thus, if the full sequence of eigenfunctions is known to be quantum ergodic, i.e., if the sequence is QUE, then the conclusion of the theorem applies to the full sequence of eigenfunctions. The proof simply relates the interior and restricted microlocal lifts and reduces the QER property along H to the QE property of the ambient manifold. If we assume that QUE holds in the ambient manifold, we automatically get QUER, which is our first Corollary:

COROLLARY 3.18. *Suppose that $\{\varphi_h\}$ is QUE on M. Then the distributions $\{d\Phi_h^{CD}\}$ have a unique weak* limit*

$$\omega(a) := \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{\frac{1}{2}}\, d\sigma \tag{3.36}$$

In particular, QUE implies QUER when $H = \partial M$ in the Dirichlet or Neumann case.

We note that $d\Phi_h^{CD}$ involves the microlocal lift $d\Phi_h^{RD}$ rather than the microlocal lift of the Dirichlet data. The next result is a QER theorem for the re-normalized distributions $d\Phi_h^D + d\Phi_h^{RN}$ where the microlocal lift $d\Phi_h^D \in \mathcal{D}'(B^*H)$ of the Dirichlet data of φ_h is defined by

$$\int_{B^*H} a\, d\Phi_h^D := \langle \mathrm{Op}_H(a)\varphi_h|_H, \varphi_h|_H\rangle_{L^2(H)}\,, \tag{3.37}$$

and

$$\int_{B^*H} a\, d\Phi_h^{RN} := \left\langle (1 + h^2\Delta_H + i0)^{-1}\, \mathrm{Op}_H(a) hD_\nu\varphi_h|_H, hD_\nu\varphi_h|_H \right\rangle_{L^2(H)}. \tag{3.38}$$

THEOREM 3.19. *Suppose $H \subset M$ is a smooth, codimension 1 embedded orientable separating hypersurface and assume $H \cap \partial M = \emptyset$. Assume that $\{\varphi_h\}$ is a quantum ergodic sequence. Then, there exists a sub-sequence of density one as $h \to 0^+$ such that for all $a \in S^0_{sc}(H)$,*

(3.39)

$$\langle (1+h^2\Delta_H+i0)^{-1}\, \mathrm{Op}_H(a) hD_\nu\varphi_h|_H, hD_\nu\varphi_h|_H\rangle_{L^2(H)} + \langle \mathrm{Op}_H(a)\varphi_h|_H, \varphi_h|_H\rangle_{L^2(H)} \to_{h\to 0^+} \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{-1/2} d\sigma,$$

where $a_0(x', \xi')$ is the principal symbol of $\mathrm{Op}_H(a)$.

COROLLARY 3.20. *With the same assumptions and notation as in Theorem 3.17, we also have*

$$d\Phi_j^D + d\Phi_j^N \to \omega_D, \tag{3.40}$$

where

$$\omega_D(a^w) = \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{-1/2} d\sigma. \tag{3.41}$$

3.7. Quantum ergodic restriction theorems for Dirichlet data

Above we stated a 'universal' QER theorem for Cauchy data. It is more difficult to determine whether the Dirichlet or Neumann data alone is quantum ergodic on a hypersurface. In fact, there is a rather obvious obstruction illustrated by odd eigenfunctions on a Riemannian manifold with an involutive isometry $\sigma\colon M \to M$. The odd eigenfunctions must vanish on the fixed point set of σ and are obviously not quantum ergodic, but their normal derivatives are. In some sense we wish to prove that this is the only obstruction. The result we describe here does not go that far but the condition is in the same direction.

The QER problem is to determine conditions on a hypersurface H so that the restrictions $\{\gamma_H \varphi_j\}$ to H of an orthonormal basis of eigenfunctions $\{\varphi_j\}$ of (M, g) with ergodic geodesic flow, are quantum ergodic along H. Here, $\gamma_H f = f|_H$ denotes the restriction operator to H. We say that $\{\gamma_H \varphi_j\}$ is quantum ergodic along H if there exists a measure $d\mu_H$ on T^*H and a density one subsequence of eigenfunctions so that, for any zeroth-order pseudo-differential operator $\mathrm{Op}_H(a)$ defined on H,

$$\langle \mathrm{Op}_H(a)\gamma_H\varphi_j, \gamma_H\varphi_j\rangle_{L^2(H)} \to \int_{T^*H} a\, d\mu_H. \tag{3.42}$$

Here, the norm on $L^2(H)$ is $\|f\|^2_{L^2(H)} = \int_H |f|^2\, dS$ where dS is the Riemannian surface measure.

There is a dynamical condition on the position of H relative to the geodesic flow which is sufficient for QER. To introduce it we need some more notation. We denote by

$$T^*_H M = \{(q, \xi) \in T^*_q M : q \in H\} \tag{3.43}$$

the covectors to M with footpoint on H, and by $T^*H = \{(q, \eta) \in T^*_q H : q \in H\}$ the cotangent bundle of H. We further denote by $\pi_H : T^*_H M \to T^*H$ the restriction map

$$\pi_H(x, \xi) = \xi|_{TH}. \tag{3.44}$$

It is a linear map whose kernel is the conormal bundle N^*H to H, i.e., the annihilator of the tangent bundle TH. In the presence of the metric g, we may identify co-vectors in T^*M with vectors in TM and induce a co-metric g on T^*M. The orthogonal decomposition $T_H M = TH \oplus NH$ induces an orthogonal decomposition $T^*_H M = T^*H \oplus N^*H$, and the restriction map (3.44) is equivalent modulo metric identifications to the tangential orthogonal projection (or restriction)

$$\pi_H : T^*_H M \to T^*H. \tag{3.45}$$

For any orientable (embedded) hypersurface $H \subset M$, there exists two unit normal co-vector fields $\nu_\pm$ to H which span half ray bundles $N_\pm = \mathbb{R}_+\nu_\pm \subset N^*H$. Infinitesimally, they define two 'sides' of H – indeed they are the two components of $T^*_H M \backslash T^*H$. We often use Fermi normal coordinates (s, y_n) along H with $s \in H$ and with $x = \exp_x y_n \nu$. We let σ, η_n denote the dual symplectic coordinates.

We also denote by $S^*_H M$ (resp. S^*H) the unit covectors in $T^*_H M$ (resp. T^*H). In general, for any subset $V \subset T^*M$ we denote by $SV = V \cap S^*M$ the subset of unit covectors in V. We may restrict (3.45) to get $\pi_H : S^*_H M \to B^*H$, where B^*H is the unit co-ball bundle of H. Conversely, if $(s, \sigma) \in B^*H$, then there exist two unit covectors $\xi_\pm(s, \sigma) \in S^*_s M$ such that $|\xi_\pm(s, \sigma)| = 1$ and $\xi|_{T_s H} = \sigma$. In the above orthogonal decomposition, they are given by

$$\xi_\pm(s, \sigma) = \sigma \pm \sqrt{1 - |\sigma|^2}\,\nu_+(s). \tag{3.46}$$

We define the reflection involution through T^*H by

$$r_H : T^*_H M \to T^*_H M, \qquad r_H(s, \mu\, \xi_\pm(s, \sigma)) = (s, \mu\, \xi_\mp(s, \sigma)), \qquad \mu \in \mathbb{R}_+. \tag{3.47}$$

Its fixed point set is T^*H.

Definition 3.21. We say that H has a positive measure of microlocal reflection symmetry if

$$\mu_{L,H}\left(\bigcup_{j\neq 0}^{\infty}\{(s,\xi)\in S^*_H M : r_H G^{T^{(j)}(s,\xi)}(s,\xi) = G^{T^{(j)}(s,\xi)} r_H(s,\xi)\}\right) > 0.$$

Otherwise we say that H is asymmetric with respect to the geodesic flow.

For homogeneous pseudo-differential operators, the QER theorem is as follows:

Theorem 3.22. *Let (M,g) be a compact manifold with ergodic geodesic flow, and let $H \subset M$ be a hypersurface. Let φ_{λ_j} denote the L^2-normalized eigenfunctions of Δ_g. If H has a zero measure of microlocal symmetry, then there exists a density-one subset S of $\mathbb{N}$ such that for $\lambda_0 > 0$ and $a(s,\sigma) \in S^0_{cl}(T^*H)$*

$$\lim_{\lambda_j\to\infty; j\in S} \langle \mathrm{Op}_H(a)\gamma_H\varphi_{\lambda_j}, \gamma_H\varphi_{\lambda_j}\rangle_{L^2(H)} = \omega(a), \tag{3.48}$$

where

$$\omega(a) = \frac{2}{\mathrm{Vol}(S^*M)} \int_{B^*H} a_0(s,\sigma)\, \gamma_{B^*H}^{-1}(s,\sigma)\, ds d\sigma. \tag{3.49}$$

Alternatively, one can write $\omega(a) = \frac{1}{vol(S^*M)} \int_{S^*_H M} a_0(s, \pi_H(\xi)) d\mu_{L,H}(\xi)$. Note that $a_0(s,\sigma)$ is bounded but is not defined for $\sigma = 0$, hence $a_0(s,\pi_H(\xi))$ is not defined for $\xi \in N^*H$ if $a_0(s,\sigma)$ is homogeneous of order zero on T^*H. The integral can also be simplified to $\omega(a) = C_{M,n}\int_{S^*H} a_0\, d\mu_L$ where, $C_{M,n} = \frac{2}{\mathrm{Vol}\, S^*M}\left(\int_0^1 (1-r^2)^{-1/2} r^{n-2}\, dr\right)$ and $d\mu_L$ is Liouville measure on S^*H. The analogous result for semiclassical pseudo-differential operators is:

Theorem 3.23. *Let (M,g) be a compact manifold with ergodic geodesic flow, and let $H \subset M$ be a hypersurface. If H has a zero measure of microlocal symmetry, then there exists a density-one subset S of $\mathbb{N}$ such that for $a \in S^{0,0}(T^*H \times [0,h_0))$,*

$$\lim_{h_j\to 0^+; j\in S} \langle \mathrm{Op}_{h_j}(a)\gamma_H\varphi_{h_j}, \gamma_H\varphi_{h_j}\rangle_{L^2(H)} = \omega(a),$$

where

$$\omega(a) = \frac{2}{\mathrm{Vol}(S^*M)} \int_{B^*H} a_0(s,\sigma)\gamma_{B^*H}^{-1}(s,\sigma)\, ds d\sigma.$$

In the special case where $a(s,\sigma) = V(s)$ is a multiplication operator, an application of Theorem 3.22 gives:

Corollary 3.24. *Under the same hypotheses as in Theorem 3.22, with dS the surface measure on H,*

$$\lim_{\lambda_j\to\infty; j\in S} \int_H V(s)\left(\gamma_H\varphi_{\lambda_j}\right)^2 dS = C'_{M,n}\int_H V(s)\, dS,$$

where $C'_{M,n} = \frac{\mathrm{Vol}(S^{n-1})}{\mathrm{Vol}(S^*M)}$.

This gives an asymptotic formula for the L^2-norms of restricted eigenfunctions in the density one subsequence, as opposed to the $O(\lambda^{\frac{1}{2}})$ upper bounds in [**BuGT**]. However, it does not disqualify existence of a zero density subsequence of eigenfunctions whose L^2 norms blow up along H.

3.8. Counting nodal domains and nodal intersections with curves

One of the applications of QER theorems is to counting nodal domains on certain surfaces. In dimension one, the number of nodal points of the nth eigenfunction of a Sturm-Liouville operator on an interval equals $n-1$, and this suggests that the number of nodal domains should tend to infinity with the eigenvalue in any dimension. However, this is not the case and indeed was disproved by Stern [**St**] for squares or flat tori. Later, H. Lewy [**L**] constructed sequences of spherical harmonics on the standard S^2 with degrees tending to infinity for which the number of nodal domains is no bigger than three. But it seems plausible that for any (M, g), there exists *some* orthonormal sequence $\{\varphi_{j_k}\}$ of eigenfunctions for which $N(\varphi_{j_k}) \to \infty$ as $k \to \infty$. We now review results showing that for Riemann surfaces (M, J, σ) with anti-holomorphic involution, and for any negatively curved σ-invariant metric, $N(\varphi_{j_k}) \to \infty$ along an orthonormal sequence of eigenfunctions of density one. The same holds for non-positively curved surfaces with concave boundary.

We state the results on nodal domains on real Riemann surfaces and on non-positively surfaces with concave boundary in parallel. In both cases we have a special curve C given by

(i) $C = \partial M$, resp.
(ii) $C = \mathrm{Fix}(\sigma)$.

The Dirichlet (resp. Neumann) eigenfunctions in the boundary setting (i) correspond to odd (resp. even) eigenfunctions in the boundaryless symmetric setting (ii). Following [**GRS**], the approach is to relate the number of nodal domains of φ_j in M to the number of zeros of φ_j on the curve C. We denote by $n(\varphi_j, 0, C)$ the number of zeros a Neumann eigenfunction on C and by $n(\partial_\nu \varphi_j, 0, C)$ the number of zeros of the normal derivative of the Dirichlet eigenfunction. We also define $N^C(\varphi_j)$ to be the number of nodal domains which touch C. By completing the nodal set into an embedded graph and apply the Euler inequality for embedded graphs, one shows that

$$N^C(\varphi_j) \geq \frac{1}{2} n(\tilde{\varphi}_j, 0, C), \tag{3.50}$$

where $\tilde{\varphi}_j = \varphi_j|_C$ in the Neumann/even case and $\tilde{\varphi}_j = \partial_\nu \varphi_j|_C$ in the Dirichlet/odd case. Here, ∂_ν is a fixed choice of unit normal along C. This relation only holds for very special curves C, because we need arcs along C to count as segments of the nodal set. A relative result holds for any smooth closed curve, i.e., there is a growing number of domains bounded by the union of one arc of a nodal line and at most one arc of C (nodal domains "relative to C"). Hence the main point is to show that Cauchy data of eigenfunctions on C have many zeros.

We first consider real Riemann surfaces without boundary. These Riemann surfaces (M, J) are complexifications of real algebraic curves $M(\mathbb{R})$ which *divide* (or equivalently *separate*) M in the sense that $M \backslash M(\mathbb{R})$ has more than one component. Such a surface possesses an anti-holomorphic involution σ whose fixed point set $\mathrm{Fix}(\sigma)$ is the real curve $M(\mathbb{R})$. The space of real algebraic curves which divide their complexifications has real dimension $3g - 3$.

We define $\mathcal{M}_{M,J,\sigma}$ to be the space of smooth, σ-invariant, negatively curved Riemannian metrics on an orientable Klein surface (M, J, σ). Any negatively curved metric g_1 induces a σ-invariant one by averaging: $g_1 \to g = \frac{1}{2}(g_1 + \sigma^* g_1)$. Hence $\mathcal{M}_{M,J,\sigma}$ is an open set in the space of σ-invariant metrics. The isometry σ commutes with the Laplacian Δ_g and therefore the eigenspaces are spanned by even or odd eigenfunctions with respect to σ. We denote by $\{\varphi_j\}$ of $L^2_{even}(M)$ an orthonormal basis of even eigenfunctions, resp. $\{\psi_j\}$ an orthonormal basis of $L^2_{odd}(M)$.

The nex result is from [**JZ1**]:

THEOREM 3.25. *Let (M, J, σ) be (as above) a compact Riemann surface with anti-holomorphic involution σ for which* Fix(σ) *is dividing. Let $g \in \mathcal{M}_{M,J,\sigma}$ be a negatively curved metric that is invariant under σ. Let $\gamma \subset$* Fix(σ) *be any sub-arc. Then for any orthonormal eigenbasis $\{\varphi_j\}$ of $L^2_{even}(M)$ one can find a density* 1 *subset A of $\mathbb{N}$ such that*

$$\lim_{\substack{j\to\infty \\ j\in A}} \#Z_{\varphi_j} \cap \gamma = \infty. \tag{3.51}$$

It follows that the number of nodal domains in each case tends to infinity.

In fact, we prove that the number of zeros tends to infinity by proving that the number of sign changes tends to infinity. The proof uses the Kuzencov trace formula of [**Z1**] to show that $\int_\gamma \varphi_j \, ds$ is 'small' as $j \to \infty$ for any curve γ and for almost all eigenfunctions. On the other hand the QER theorem shows that $\int_\gamma \varphi_j^2 \, ds$ is large. We then compare $\int_\gamma \varphi_j \, ds$ and $\int_\gamma |\varphi_j| \, ds$ by applying a well known sup norm bound on eigenfunctions in the case of surfaces without conjugate points to replace $\int_\gamma \varphi_j^2 ds$ by $\int_\gamma |\varphi_j| ds$. The comparison just manages to show that for any geodesic arc γ, $\int_\gamma |\varphi_j| ds > \int_\gamma \varphi_j ds$. Hence there must exist sign-changing zeros.

REMARK 3.26. Here and for the remaining of the section, γ always denotes a sub-arc of Fix(σ). For each $g \in \mathcal{M}_{M,J,\sigma}$, it follows from Harnack's theorem that the fixed point set Fix(σ) is a disjoint union

$$\text{Fix}(\sigma) = \gamma_1 \cup \cdots \cup \gamma_k \tag{3.52}$$

of $0 \leq k \leq g + 1$ simple closed geodesics, and by our assumption $k > 0$ and Fix(σ) is dividing. Hence the arcs γ above are geodesic arcs of (M, g).

The main ingredient of Theorem 3.25 is the QER (quantum ergodic restriction) theorem for Cauchy data of [**CTZ**]. The QER theorem then says that the Cauchy data is quantum ergodic along $\mathrm{Fix}(\sigma)$. Indeed, it is quantum ergodic along any curve. $\mathrm{Fix}(\sigma)$ is special because the odd eigenfunctions automatically vanish on it and the even eigenfunctions have vanishing normal derivatives. Hence half of the Cauchy data of each eigenfunction automatically vanishes on $\mathrm{Fix}(\sigma)$. Quantum ergodicity forces the sequence of restrictions of eigenfunctions to $\mathrm{Fix}(\sigma)$ to oscillate quickly and thus to have a growing number of zeros as the eigenvalue increases.

Combining with (3.50) gives

THEOREM 3.27. *Let (M, J, σ) be a compact Riemann surface with anti-holomorphic involution σ for which $\mathrm{Fix}(\sigma)$ is dividing. Then for any $g \in \mathcal{M}_{(M,J,\sigma)}$ and any orthonormal Δ_g-eigenbasis $\{\varphi_j\}$ of $L^2_{even}(M)$ and $\{\psi_j\}$ of $L^2_{odd}(M)$, one can find a density 1 subset A of $\mathbb{N}$ such that*

$$\lim_{\substack{j\to\infty \\ j\in A}} N(\varphi_j) = \infty \quad \textit{and} \quad \lim_{\substack{j\to\infty \\ j\in A}} N(\psi_j) = \infty. \tag{3.53}$$

REMARK 3.28. For generic metrics in $\mathcal{M}_{M,J,\sigma}$, the eigenvalues are simple (multiplicity one) and therefore all eigenfunctions are either even or odd. Hence for generic metrics in $\mathcal{M}_{M,J,\sigma}$, Theorem 3.27 says that the number of nodal domains tends to infinity along almost the /entire sequence of eigenfunctions.

For odd eigenfunctions, the same conclusion holds with the assumption $\mathrm{Fix}(\sigma)$ separating replaced by $\mathrm{Fix}(\sigma) \neq \emptyset$, i.e., the conclusion holds for the complexification of any real algebraic curve.

In more recent work [**Z3**], the author has shown that the number of zeros grows like a power of $\log \lambda$.

THEOREM 3.29. *Let (M, J) be a real Riemann surface and let g be a negatively curved invariant metric on M. Then for $j \in A$ (a set of density one),*

$$N(\varphi_j) \geq C_g (\log \lambda_j)^K, \qquad \forall K < \frac{1}{6}. \tag{3.54}$$

resp.

$$N(\psi_j) \geq C_g (\log \lambda_j)^K, \qquad \forall K < \frac{1}{6}. \tag{3.55}$$

It is doubtful that this lower bound is sharp in the negatively curved case. However it is difficult to go beyond logarithms due to the exponential growth of the geodesic flow.

The same methods show that the number of singular points of odd eigenfunctions ψ_j tends to infinity. By singular points of an eigenfunction we mean the set

$$\Sigma_{\varphi_\lambda} = \{x \in Z_{\varphi_\lambda} : d\varphi_\lambda(x) = 0\} \tag{3.56}$$

of critical points φ_λ which lie on the nodal set Z_{φ_j}. For generic metrics, the singular set is empty [**U**]. However for negatively curved surfaces with an isometric involution, odd eigenfunctions ψ always have singular points. Indeed, odd eigenfunctions vanish on γ and they have singular points at $x \in \gamma$ where the normal derivative vanishes, $\partial_\nu \psi_j = 0$.

THEOREM 3.30. *Let (M, J, σ) be a compact Riemann surface with anti-holomorphic involution for which $\mathrm{Fix}(\sigma)$ is dividing. Let $g \in \mathcal{M}_{M,J,\sigma}$. Then for any orthonormal eigenbasis $\{\psi_j\}$ of $L^2_{odd}(M)$, one can find a density 1 subset A of $\mathbb{N}$ such that*

$$\lim_{\substack{j\to\infty \\ j\in A}} \#\Sigma_{\psi_j} \cap \mathrm{Fix}(\sigma) = \infty. \tag{3.57}$$

Furthermore, there are an infinite number of zeros where $\partial_\nu \psi_j|_H$ changes sign.

We now state the parallel results on a surface with non-empty smooth concave boundary $\partial M \neq \emptyset$. The first result states sufficient conditions under which the number of nodal domains tends to infinity along a subsequence of 'almost all' eigenfunctions of any orthonormal basis.

THEOREM 3.31. *Let (M, g) be a surface with non-empty smooth boundary ∂M. Let $\{\varphi_j\}$ be an orthonormal eigenbasis of Dirichlet (resp. Neumann) eigenfunctions. Assume that (M, g) satisfies the following conditions:*

- (i) *The billiard flow G^t is ergodic on S^*M with respect to Liouville measure;*
- (ii) *There does not exist a self-focal point $q \in \partial M$ for the billiard flow, i.e., a point q such that the set $\mathcal{L}_q$ of loop directions $\eta \in B^*_q\partial M$ has positive measure in $B^*_q \partial M$.*
- (iii) *The Cauchy data $(\varphi_j|_{\partial M}, \lambda_j^{-1}\partial_\nu \varphi_j|_{\partial M})$ of the eigenfunctions is strictly sub-maximal in growth, i.e., both components are $o(\lambda_j^{\frac{1}{2}})$ as $\lambda_j \to \infty$.*

Then there exists a subsequence $A \subset \mathbb{N}$ of density one so that

$$\lim_{\substack{j\to\infty \\ j\in A}} N(\varphi_j) = \infty. \tag{3.58}$$

The hypotheses of Theorem 3.31 are satisfied by surfaces of non-positive curvature with concave boundary. By this we mean a non-positively curved surface

$$M = X \backslash \bigcup_{j=1}^{r} \mathcal{O}_j, \tag{3.59}$$

obtained by removing a finite union $\mathcal{O} := \bigcup_{j=1}^{r} \mathcal{O}_j$ of embedded non-intersecting geodesically convex domains (or 'obstacles) $\mathcal{O}_j$ from a closed non-positively curved surface (X, g). We denote the scalar curvature of (X, g) by K and assume $K \leq 0$. In the case where X is a flat torus or a square, such a billiard is called a Sinai billiard.

COROLLARY 3.32. *The conclusion of Theorem 3.27 holds for a non-positively curved surface* (3.59) *with concave boundary.*

3.9. Intersections of nodal lines and general curves on negatively curved surfaces

Although it is not useful for counting nodal domains, we point out a general result on intersections of nodal lines and curves on surfaces.

THEOREM 3.33. *Let (M, g) be a C^∞ compact negatively curved surface, and let H be a closed curve which is asymmetric with respect to the geodesic flow. Then for* *any orthonormal eigenbasis* $\{\varphi_j\}$ *of eigenfunctions, there exists a density* 1 *subset* A *of* $\mathbb{N}$ *such that*

$$\begin{cases} \lim_{\substack{j\to\infty \\ j\in A}} \#Z_{\varphi_j} \cap H = \infty \\ \lim_{\substack{j\to\infty \\ j\in A}} \#\{x \in H : \partial_\nu \varphi_j(x) = 0\} = \infty. \end{cases} \tag{3.60}$$

Furthermore, there are an infinite number of zeros where $\varphi_j|_H$ *(resp.* $\partial_\nu \varphi_j|_H$*) changes sign.*

Theorem 3.33 does not necessarily imply lower bounds on nodal domains because the topological argument used in the case $H = \mathrm{Fix}(\sigma)$ does not necessarily apply. Its proof is essentially the same as that for Theorem 3.25. The main difference is that we use the QER theorem for Cauchy data in Theorem 3.25 and for just the Dirichlet data in Theorem 3.33. The latter requires the asymmetry condition on H.

Finally, we mention an upper bound on the number of nodal intersections in the real analytic case [**TZ3**].

THEOREM 3.34. *Suppose that* $\Omega \subset \mathbb{R}^2$ *is a piecewise real analytic plane domain. Then the number* $n(\lambda_j) = \#Z_{\varphi_{\lambda_j}} \cap \partial\Omega$ *of zeros of the boundary values* $\varphi_{\lambda_j}|_{\partial\Omega}$ *of the* j*th Neumann eigenfunction satisfies* $n(\lambda_j) \leq C_\Omega \lambda_j$*, for some* $C_\Omega > 0$.

The result is surely true for any real analytic Riemannian surface with analytic boundary but has not as yet been generalized that far. To our knowledge, no upper bound on the number of intersections in the general C^∞ case is known.

3.10. Complex zeros of eigenfunctions

As mentioned in the introduction, we have much more control over nodal sets in the complex domain than in the real domain. By the complex domain is meant the complexification of M. When g is real analytic, the eigenfunctions admit simultaneous analytic extensions to a fixed Grauert tube M_ε. The tube radius ε is defined by a tube function $\sqrt{\rho}$ where $\rho(z) = -r^2(z, \bar{z})$. The complex nodal hypersurface of an eigenfunction is defined by

$$Z_{\varphi_\lambda^{\mathbb{C}}} = \{\zeta \in M_{\varepsilon_0} : \varphi_\lambda^{\mathbb{C}}(\zeta) = 0\}. \tag{3.61}$$

As discussed in §14.30.2, there exists a natural current of integration over the nodal hypersurface in any Grauert tube M_ε given by the Poincaré-Lelong formula,

$$\langle [Z_{\varphi_\lambda^{\mathbb{C}}}], f\rangle = \frac{i}{2\pi}\int_{M_\varepsilon} \partial\bar{\partial} \log |\varphi_\lambda^{\mathbb{C}}|^2 \wedge f = \int_{Z_{\varphi_\lambda^{\mathbb{C}}}} f \tag{3.62}$$

for all smooth test $(m-1, m-1)$-forms $f \in \mathcal{D}^{(m-1,m-1)}(M_\varepsilon)$ with support in M_ε. (In the second equality we used the Poincaré-Lelong formula; see §14.30.1.)

The nodal hypersurface $Z_{\varphi_\lambda^{\mathbb{C}}}$, viewed as a complex hypersurface embedded in a Kähler manifold, also carries a natural volume form. By Wirtinger's formula, this volume form equals the restriction of $\frac{\omega_g^{m-1}}{(m-1)!}$ to $Z_{\varphi_\lambda^{\mathbb{C}}}$. Hence, one can regard $Z_{\varphi_\lambda^{\mathbb{C}}}$ as defining the measure

$$\langle |Z_{\varphi_\lambda^{\mathbb{C}}}|, f\rangle = \int_{Z_{\varphi_\lambda^{\mathbb{C}}}} f \frac{\omega_g^{m-1}}{(m-1)!}, \quad \text{where } f \in C(M_\varepsilon). \tag{3.63}$$

We prefer to state results in terms of the current $[Z_{\varphi_\lambda^{\mathbb{C}}}]$ since it carries more information.

THEOREM 3.35. *Let (M, g) be real analytic, and let $\{\varphi_{j_k}\}$ denote a quantum ergodic sequence of eigenfunctions. Let M_{ε_0} be the maximal Grauert tube around M. Let $\varepsilon < \varepsilon_0$. Then,*

$$\frac{1}{\lambda_{j_k}}[Z_{\varphi_{j_k}^{\mathbb{C}}}] \to \frac{i}{\pi}\partial\bar{\partial}\sqrt{\rho} \quad \text{weakly in } \mathcal{D}'^{(m-1,m-1)}(M_\varepsilon) \tag{3.64}$$

in the sense that, for any continuous test form $\psi \in \mathcal{D}^{(m-1,m-1)}(M_\varepsilon)$, we have

$$\frac{1}{\lambda_{j_k}}\int_{Z_{\varphi_{j_k}^{\mathbb{C}}}} \psi \to \frac{i}{\pi}\int_{M_\varepsilon} \psi \wedge \partial\bar{\partial}\sqrt{\rho}. \tag{3.65}$$

Equivalently, for any $f \in C(M_\varepsilon)$,

$$\frac{1}{\lambda_{j_k}}\int_{Z_{\varphi_{j_k}^{\mathbb{C}}}} f\frac{\omega_g^{m-1}}{(m-1)!} \to \frac{i}{\pi}\int_{M_\varepsilon} f\partial\bar{\partial}\sqrt{\rho} \wedge \frac{\omega_g^{m-1}}{(m-1)!}. \tag{3.66}$$

In [**Z2**] we proved a similar result for intersections of the complex nodal curve with complexified geodesics. The complexification of an arc-length parametrized geodesic

$$\gamma_{x,\xi}\colon \mathbb{R} \to M, \qquad \gamma_{x,\xi}(0) = x, \qquad \gamma'_{x,\xi}(0) = \xi \in T_xM \tag{3.67}$$

is defined by analytic continuation

$$\gamma^{\mathbb{C}}_{x,\xi} : S_\varepsilon \to M_\varepsilon \tag{3.68}$$

to the strip

$$S_\varepsilon = \{(t + i\tau \in \mathbb{C} : |\tau| \le \varepsilon\}. \tag{3.69}$$

When we freeze τ we simplify the notation to

$$\gamma^\tau_{x,\xi}(t) := \gamma^{\mathbb{C}}_{x,\xi}(t + i\tau). \tag{3.70}$$

The intersection points of $\gamma^{\mathbb{C}}_{x,\xi}$ and $\mathcal{N}^{\mathbb{C}}_{\varphi_j}$ correspond to the zeros of the pullback $(\gamma^{\mathbb{C}}_{x,\xi})^*\varphi^{\mathbb{C}}_j$. We encode this discrete set by the measure

$$[\mathcal{N}^{\gamma^{\mathbb{C}}_{x,\xi}}_{\lambda_j}] = \sum_{(t+i\tau)\colon\, \varphi^{\mathbb{C}}_j(\gamma^{\mathbb{C}}_{x,\xi}(t+i\tau))=0} \delta_{t+i\tau}. \tag{3.71}$$

Let $\mathcal{S} = \{j_k\} \subset \mathbb{N}$ be a subsequence of the positive integers. We say that the intersection points of the complex nodal sets $\mathcal{N}^{\mathbb{C}}_{\lambda_{j_k}}$ and the complexified geodesic $\gamma^{\mathbb{C}}_{x,\xi}$ for the subsequence $\mathcal{S}$ condense on the real geodesic and become uniformly distributed with respect to arc-length if, for any $f \in C_c(S_\varepsilon)$,

$$\lim_{k\to\infty}\frac{1}{\lambda_{j_k}}\sum_{(t+i\tau)\colon\, \varphi^{\mathbb{C}}_{j_k}(\gamma^{\mathbb{C}}_{x,\xi}(t+i\tau))=0} f(t + i\tau) = \frac{1}{\pi}\int_{\mathbb{R}} f(t)dt. \tag{3.72}$$

That is, $\frac{1}{\lambda_{j_k}}[\mathcal{N}^{\gamma^{\mathbb{C}}_{x,\xi}}_{\lambda_j}] \to \frac{1}{\pi}\delta_0(\tau)dtd\tau$ in the sense of measures.

THEOREM 3.36. *Let (M, g) be a real analytic Riemannian surface with ergodic geodesic flow. Let $\gamma_{x,\xi}$ be a periodic geodesic satisfying the asymmetry QER hypothesis of Definition 3.21. Then there exists a subsequence of eigenvalues λ_{j_k} of density one such that* (3.72) *holds.*

Moreover, we have

PROPOSITION 3.37 (Growth saturation). *If $\gamma_{x,\xi}$ is a periodic geodesic which satisfies the QER asymmetry condition (Definition 3.21) along compact arcs, then there exists a subsequence $\mathcal{S}_{x,\xi}$ of density one so that, for all $\tau < \varepsilon$,*

$$\lim_{k\to\infty} \frac{1}{\lambda_{j_k}} \log \left| \gamma_{x,\xi}^{\tau *} \varphi_{\lambda_{j_k}}^{\mathbb{C}}(t+i\tau) \right|^2 = 2|\tau| \quad \textit{in } L^1_{loc}(S_\tau). \tag{3.73}$$

The subsequence $\mathcal{S}_{x,\xi}$ is the ergodic sequence along $\gamma_{x,\xi}$ given by Theorem 3.42.

Proposition 3.37 immediately implies Theorem 3.36 since we can apply $\partial\bar{\partial}$ to the L^1 convergent sequence $\frac{1}{\lambda_{j_k}} \log \left| \gamma_{x,\xi}^{*} \varphi_{\lambda_{j_k}}^{\mathbb{C}}(t+i\tau) \right|^2$ to obtain a weakly convergence sequence of measures tending to $\partial\bar{\partial}|\tau|$. Proposition 3.37 has an analogue for any real analytic curve but the exact formula is special to geodesics and arises because complex geodesics are isometric embeddings to Grauert tubes. In general, the growth rates of restrictions depend on the curve.

Bibliography

[BS1] M.D. Blair and C.D. Sogge, Refined and microlocal Kakeya-Nikodym bounds for eigenfunctions in two dimensions. Anal. PDE 8 (2015), no. 3, 747-764.

[BS2] M. D. Blair and C.D. Sogge, On Kakeya-Nikodym averages, L^p-norms and lower bounds for nodal sets of eigenfunctions in higher dimensions, J. Eur. Math. Soc. (JEMS) 17 (2015), no. 10, 2513-2543 (arXiv:1301.7468) .

[Bo] J. Bourgain, Geodesic restrictions and Lp-estimates for eigenfunctions of Riemannian surfaces, Linear and complex analysis, 27?35, Amer. Math. Soc. Tranl. Ser. 2, 226, Amer. Math. Soc., Providence, RI, 2009.

[Bu] N. Burq, Quantum ergodicity of boundary values of eigenfunctions: A control theory approach, to appear in Canadian Math. Bull. 48 (2005), 3-15 (math.AP/0301349).

[BuGT] N. Burq, P. Gerard, and N. Tzvetkov, Restrictions of the Laplace-Beltrami eigenfunctions to submanifolds, Duke Math. J. 138 (2007), 445?486. MR 2322684.

[CS] X. Chen and C. D. Sogge, A few endpoint geodesic restriction estimates for eigenfunctions. Comm. Math. Phys. 329 (2014), no. 2, 435-459.

[CTZ] H. Christianson, J. A. Toth and S. Zelditch, Quantum ergodic restriction for Cauchy data: interior que and restricted que. Math. Res. Lett. 20 (2013), no. 3, 465-475.

[GRS] A. Ghosh, A. Reznikov and Peter Sarnak. Nodal domains of Maass forms I. Geom. Funct. Anal., 23(5):1515-1568, 2013.

[GeL] P.Gérard and E.Leichtnam, Ergodic properties of eigenfunctions for the Dirichlet problem, Duke Math J. 71 (1993), 559-607.

[HZ] A. Hassell and S. Zelditch, Quantum ergodicity of boundary values of eigenfunctions. Comm. Math. Phys. 248 (2004), no. 1, 119–168.

[JZ1] J. Jung and S. Zelditch, Number of nodal domains and singular points of eigenfunctions of negatively curved surfaces with an isometric involution. J. Differential Geom. 102 (2016), no. 1, 37-66.

[JZ2] J. Jung and S. Zelditch, Number of nodal domains of eigenfunctions on non-positively curved surfaces with concave boundary. Math. Ann. 364 (2016), no. 3-4, 813-840.

[L] H. Lewy, On the minimum number of domains in which the nodal lines of spherical Comm. Partial Differential Equations 2 (1977), no. 12, 1233-1244.

[Sa] Y. Safarov, Asymptotics of a spectral function of a positive elliptic operator without a nontrapping condition. (Russian) Funktsional. Anal. i Prilozhen. 22 (1988), no. 3,53–65, 96 translation in Funct. Anal. Appl. 22 (1988), no. 3, 213–223 (1989).

[SaV] Y. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators. Translated from the Russian manuscript by the authors.* Translations of Mathematical Monographs, 155. American Mathematical Society, Providence, RI, 1997.

[S1] C. D. Sogge, *Fourier integrals in classical analysis.* Cambridge Tracts in Mathematics, 105. Cambridge University Press, Cambridge, 1993.

[S2] C. D. Sogge, *Kakeya-Nikodym averages and L^p-norms of eigenfunctions*, Tohoku Math. J. (2) 63 (2011), no. 4, 519-538 (arXiv:0907.4827).

[SZ1] C.D. Sogge and S. Zelditch, On eigenfunction restriction estimates and L4-bounds for compact surfaces with nonpositive curvature. *Advances in analysis: the legacy of Elias M. Stein,* 447-461, Princeton Math. Ser., 50, Princeton Univ. Press, Princeton, NJ, 2014.

[SZ2] C. Sogge and S. Zelditch, Riemannian manifolds with maximal eigenfunction growth, Duke Math. J. 114 (2002), 387-437.

[SZ3] C. Sogge and S. Zelditch, Concerning the L^4 norms of typical eigenfunctions on compact surfaces, Recent developments in geometry and analysis, Vol 23, 407-423, Higher Ed. Press and Int. Press, Beijing-Boston (2012) (arXiv:1011.0215).

[SZ4] C.D. Sogge and S. Zelditch, Focal points and sup-norms of eigenfunctions. Rev. Mat. Iberoam. 32 (2016), no. 3, 971-994. (arXiv:1311.3999);

[SZ5] C. D. Sogge and S. Zelditch, Focal points and sup norms of eigenfunctions on real analytic Riemannian manifolds, Rev. Mat. Iberoam. 32 (2016), no. 3, 995-999. (arXiv:1409.2063).

[STZ] C. Sogge, J.A. Toth and S. Zelditch, About the blowup of quasimodes on Riemannian manifolds, issue on Geometric Harmonic Analysis, J. Geom. Anal 21 (2011), 150-173 (arXiv:0908.0688).

[St] A. Stern, Bemerkungen Ueber asymptotisches Verhalten von Eigenwerten und Eigenfunktionen. Diss. Gottingen (30 Juli 1924). Available at http://www-fourier.ujf-grenoble.fr/ pberard/R/stern-1925-thesis-partial-reprod.pdf

[TZ1] J. A. Toth and S. Zelditch, Riemannian manifolds with uniformly bounded eigenfunctions. Duke Math. J. 111 (2002), no. 1, 97-132.

[TZ2] J. A. Toth and S. Zelditch, L^p norms of eigenfunctions in the completely integrable case. Ann. Henri Poincar'e (2003), no. 2, 343-368.

[TZ3] J A. Toth and S. Zelditch, Counting nodal lines which touch the boundary of an analytic domain. J. Differential Geom. 81 (2009), no. 3, 649-686.

[TZ4] J. A. Toth and S. Zelditch, Quantum ergodic restriction theorems, I: interior hypersurfaces in analytic domains,, *Ann. H. Poincaré* 13, Issue 4 (2012), Page 599-670 (arXiv:1005.1636).

[TZ5] J.A. Toth and S. Zelditch, Quantum ergodic restriction theorems: manifolds without boundary. Geom. Funct. Anal. 23 (2013), no. 2, 715-775.

[U] K. Uhlenbeck, Generic properties of eigenfunctions. Amer. J. Math. 98 (1976), no. 4, 1059-1078.

[Z1] S. Zelditch, Kuznecov sum formulae and Szegö limit formulae on manifolds. Comm. Partial Differential Equations 17 (1992), no. 1-2, 221-260.

[Z3] S. Zelditch, Logarithmic lower bound on the number of nodal domains, J. Spectr. Theory 6 (2016), no. 4, 1047-1086 (arXiv:1510.05315).

[Z2] S. Zelditch, Ergodicity and intersections of nodal sets and geodesics on real analytic surfaces. J. Differential Geom. 96 (2014), no. 2, 305-351.

[Zw] M. Zworski, *Semiclassical analysis*. Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012.

CHAPTER 4

Model spaces of constant curvature

In this section we encounter eigenfunctions for the first time. We begin with model spaces of constant curvature, where special eigenfunctions can be explicitly constructed. Aside from giving explicitly computable models, these eigenfunctions are often extremal for various problems, e.g., L^p norms. Hence with begin with the geodesic flow and the eigenfunctions on the model spaces of constant curvature:

- Euclidean $\mathbb{R}^n$ and flat tori $\mathbb{R}^n/\mathbb{Z}^n$ (and other lattices L and quotients $\mathbb{R}^n/L$). This is a quantum integrable system whose joint eigenfunctions are exponential functions. They are extremal in being the 'flattest' eigenfunctions, i.e., in having uniformly bounded sup norms. No eigenfunctions in other settings are known to have such flatness properties.
- Spherical harmonics on the standard sphere S^n. This is another quantum integrable system, but the joint eigenfunctions behave quite differently from the flat case. Certain sequences are extremal in the opposite sense of being of maximal growth in L^p norm. Due to the high multiplicity of the eigenvalues, there are eigenfunctions on S^n exhibiting most known types of behavior that occur on any Riemannian manifold.
- Hyperbolic space $\mathbf{H}^n$ and its quotients $\mathbf{H}^n\backslash\Gamma$, where Γ is a discrete subgroup of the isometry group of $\mathbf{H}^n$. On the universal cover, the Laplacian is quantum integrable and one has explicit eigenfunctions. On compact quotients, the eigenfunctions are a model of quantum chaos and much of their behavior is unknown.

4.1. Euclidean space

Euclidean space is $(\mathbb{R}^n, g)$ with the flat metric $g = \sum_{j=1}^n dx_j^2$. On a sufficiently small length scale, every Riemannian manifold is approximately Euclidean and the theory of harmonic and subharmonic functions on $\mathbb{R}^n$ is approximately true on small balls of any Riemannian manifold. Hence, $\mathbb{R}^n$ is fundamental and we go over some of the main operators used to study harmonic functions on $\mathbb{R}^n$.

The associated Hamiltonian $\sum_{j=1}^n \xi_j^2$ on $T^*\mathbb{R}^n$ is the square of the metric norm, but for applications to the wave equation we take its square root to make it homogeneous of degree one:

$$H(x,\xi) = |\xi| = \left(\sum_{j=1}^n \xi_j^2\right)^{\frac{1}{2}}. \tag{4.1}$$

The square root normalization makes the geodesic flow the same on all energy surfaces $H = E$. The level sets of H are the co-sphere bundles: $S_E^*M = \{(x,\xi)\colon |\xi| = E\}$. The square root also causes a singularity at the 'zero section'

$\xi = 0$, but it is harmless for our purposes because we are interested in asymptotics as $|\xi| \to \infty$.

The Hamiltonian vector field of H is

$$\Xi_H(x,\xi) = \sum_{j=1}^{n} \frac{\xi_j}{|\xi|} \frac{\partial}{\partial x_j}. \tag{4.2}$$

Hence, the geodesic flow is defined by

$$\frac{dx_j}{dt} = \frac{\xi_j}{|\xi|} \quad \text{and} \quad \frac{d\xi_j}{dt} = 0. \tag{4.3}$$

We see that the functions ξ_j are constant along the flow-lines (they are called 'conserved quantities'). Existence of n independent conserved quantities in n dimension is known as complete integrability. The equations of motion can be explicitly solved:

$$x(t) = x(0) + t\frac{\xi}{|\xi|} \quad \text{and} \quad \xi(t) = \xi(0). \tag{4.4}$$

We now consider the quantum picture. The eigenspaces of the Laplacian on $\mathbb{R}^n$ are defined by

$$\mathcal{E}_\lambda = \{\varphi_\lambda \in \mathcal{S}'(\mathbb{R}^n) \colon (\Delta + \lambda^2)\varphi_\lambda = 0\}, \tag{4.5}$$

where $\mathcal{S}'(\mathbb{R}^n)$ is the space of tempered distributions, which rules exponentially growing eigenfunctions such as $e^{\langle x,\xi\rangle}$ with $|\xi| = \lambda$. The Euclidean Laplacian is *quantum integrable* in the sense that it commutes with the infinitesimal translations $\frac{\partial}{\partial x_j}$. A basis of (complex valued) joint eigenfunctions of the differential operators $\{\frac{\partial}{\partial x_j}\}_{j=1}^n$ are the Euclidean plane waves $e^{i\langle x,\xi\rangle}$ for $\xi \in \mathbb{R}^n$.

The spectral theory of the Laplacian is essentially the theory of the Fourier transform. We recall a few facts about the Fourier transform in §4.16 but must assume the reader is familiar with Euclidean Fourier analysis. For purposes of this monograph, we treat Fourier analysis on $\mathbb{R}^n$ as a known object that we wish to generalize to other Riemannian manifolds. We refer to [**DM, StW**] for further background on Fourier analysis and to [**Str**] for more on Riemannian eigenfunctions as a generalization of Fourier analysis.

The eigenspaces $\mathcal{E}_\lambda$ are spanned by the plane waves $e^{i\langle x,\xi\rangle}$. The following theorem is known as the Poisson formula for Euclidean eigenfunctions (see [**Ag, Hel1, Hel2**] for the proof):

THEOREM 4.1. *Let $\varphi_\lambda \in \mathcal{E}_\lambda$. Then there exists a distribution $dT \in \mathcal{D}'(S^{n-1})$ such that*

$$\varphi_\lambda(x) = \int_{S^{n-1}} e^{i\lambda\langle x,\sigma\rangle}\, dT(\sigma). \tag{4.6}$$

This is a global theorem: It applies to global eigenfunctions but not to local eigenfunctions. Indeed, the right side is a global eigenfunction. It is studied in [**Ag**], where the question is raised of proving such formulae on more general Riemannian manifolds. In the next section we will consider the generalization to hyperbolic space.

It is interesting to relate properties of φ_λ to properties of dT. A special class of eigenfunctions occurs when $dT \in L^2(S^{n-1})$. Then one can define a Hilbert space norm on the space $\mathcal{E}_\lambda^{(2)}$ of such eigenfunctions by the $L^2(S^{n-1})$ norm of $dT(\sigma)$. The elements of finite norm define the Hilbert space $\mathcal{E}_\lambda^{(2)}$. The Euclidean motion group E_n (translations and rotations) commutes with the flat Laplacian $\Delta = \Delta_{\mathbb{R}^n}$, and therefore preserves the eigenspaces. Hence $\mathcal{E}_\lambda^{(2)}$ is a representation of E_n. They are in fact irreducible and infinite dimensional, and carry invariant inner products or which the action of E_n is unitary. See for instance [**DM, Hel1, Hel2**].

Let $J_\nu(z)$ be the Bessel function

$$J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+\nu+1)\Gamma(k+1)} \left(\frac{z}{2}\right)^{2k+\nu}, \tag{4.7}$$

then the orthogonal projection onto an individual eigenspace $\mathcal{E}_\lambda^{(2)}$ is given by the Bessel kernel

$$E_\lambda f(x) = \int_{\mathbb{R}^n} J_{\frac{n-2}{2}}(\lambda|x-y|) f(y)\, dy. \tag{4.8}$$

To see this, we first note that by the Fourier inversion formula, the spectral projection for Δ for the spectral interval $[0, \lambda^2]$ is given by

$$e_0(x-y, \lambda^2) = (2\pi)^{-n} \int_{|\xi|<\lambda} e^{i\langle x-y, \xi\rangle}\, d\xi. \tag{4.9}$$

By differentiating, we find that the spectral projection onto $\mathcal{E}_\lambda$ is given by

$$\frac{d}{d\lambda} e_0(x-y, \lambda^2) = (2\pi)^{-n} \int_{|\xi|=\lambda} e^{i\langle x-y, \xi\rangle}\, dS, \tag{4.10}$$

where dS is the standard surface measure. This proves (4.8).

The Euclidean spherical means operator is defined by

$$L_r u(x) = \frac{1}{|S_r(x)|} \int_{S_r(x)} u\, dA, \tag{4.11}$$

where $|S_r(x)|$ is the Riemannian surface area of the sphere $S_r(x)$ of radius r, and the ball means operator is

$$B_r u(x) = \frac{1}{|B_r(x)|} \int_{B_r(x)} u\, dV, \tag{4.12}$$

where $|B_r(x)|$ is the volume of the ball $B_r(x)$ of radius r. These operators have many repercussions for harmonic and subharmonic functions on $\mathbb{R}^n$, and also for the wave equation on $\mathbb{R}^n$. They can also be defined on Riemannian manifolds, a subject we explore in §5.9.3. A word of caution: In flat Euclidean space spherical means and ball means operators agree, but they differ in general on curved Riemannian manifolds, where the surface measure on a geodesic sphere is not the pushforward under the exponential map of the Euclidean surface measure in the tangent space.

4.1.1. Spherical means. On Euclidean $\mathbb{R}^n$, there exists an exact formula for the spherical means operator known as Pizetti's formula:

$$L_r = \sum_{k=0}^{\infty} \frac{\Gamma(\frac{n}{2})}{k!\Gamma(\frac{n}{2}+k)} \left(\frac{r}{2}\right)^{2k} \Delta^k, \tag{4.13}$$

which is valid when applied to real analytic functions u. If we write

$$W_n(z) = \Gamma\left(\frac{n}{2}\right) \left(\frac{2}{z}\right)^{\frac{n}{2}-1} J_{\frac{n-2}{2}}(z), \tag{4.14}$$

then direct computations show that for all analytic functions u on $\mathbb{R}^n$, we have

$$L_r u(x) = W_n(ir\sqrt{\Delta})u(x). \tag{4.15}$$

In dimensions $n = 2$ and $n = 3$, for instance, the spherical means operator is given by

$$J_0(r\sqrt{\Delta}) \quad \text{and} \quad \frac{\sinh(r\sqrt{\Delta})}{r\sqrt{\Delta}}, \tag{4.16}$$

respectively.

There are similar formulae for the ball means operator. It follows by the mean value property that if u is a harmonic function on $\mathbb{R}^n$, then $L_r u = P_r(0)u = u$ and similarly $M_r u = u$. A second identity is that for all u (not necessarily analytic),

$$\int_{B_r(x)} \Delta u(y)\, dy = r^{n-1}\frac{d}{dr}\left(r^{1-n}\int_{S_r(x)} u(y)\, dS(y)\right) = \omega_n r^{n-1}\frac{d}{dr}L_r u(x). \tag{4.17}$$

Here $\omega_n = |S^{n-1}|$ is the volume of the unit $(n-1)$-sphere. If $\Delta u \geq 0$, i.e., if u is subharmonic, then $L_r u(x)$ increases with r. It follows that $u(x) \leq L_r u(x)$ for all r. By integrating the inequality in r, one also has $u(x) \leq M_r u(x)$. Yet a third identity concerns eigenfunctions. If $0 > \lambda > \lambda_0(B_r(x))$, where $\lambda_0(B_r(x))$ is the smallest nonzero Dirichlet eigenvalue of the ball of radius r centered at x, then

$$u(x) = \frac{\tau_{\frac{n}{2}}(r\lambda)}{2^{\frac{n}{2}}\Gamma(\frac{n}{2}+1)}\frac{1}{|B_r(x)|}\int_{B_r(x)} u(y)\, dV(y). \tag{4.18}$$

Here, $\tau_\alpha(z) = z^{-\alpha}J_\alpha(z)$.

Note that the initial expansion of L_r has the form

$$L_r = I + \frac{r^2}{2n}\Delta + \sum_{k=2}^{\infty} P_k(\Delta)r^{2k}, \quad \text{where } P_k(\Delta) = \frac{\Gamma(\frac{n}{2})}{k!\Gamma(\frac{n}{2}+k)}\left(\frac{1}{2}\right)^{2k}\Delta^k \tag{4.19}$$

Comparing with the odd solution operator

$$S(t) := \frac{\sin t\sqrt{\Delta}}{t\sqrt{\Delta}} = I + \frac{\Delta}{2}t^2 + \cdots \tag{4.20}$$

of the initial value problem for the wave equation, we see that even though L_r and $S(r)$ are different functions of Δ, their Taylor expansions agree in the first two terms when $t = \frac{r}{\sqrt{n}}$. This is a first indication of the relation between the wave equation and spherical means on $\mathbb{R}^n$, as discussed at length in [**J**]. A deeper fact is that both L_r (or M_r) and $S(t)$ are Fourier integral operators associated to the same canonical relation, namely the union of the graph of the geodesic flow G^t and of G^{-t}. This is true for small $|t|$ or r on any Riemannian manifold. Therefore there exists an elliptic pseudo-differential operator $A(t, D_t, x, D_x)$ on $\mathbb{R}_t \times \mathbb{R}^n$ so that $S(t) = AL_{t\sqrt{n}}$. The Hadamard parametrix method discussed below gives an explicit construction of $S(t)$ in terms of operations on the spherical means operator on any manifold without conjugate points.

4.1.2. Propagators and fundamental solution of the wave equation.

4.2. Euclidean wave kernels

In this section, we review the exact formulae for the propagators and fundamental solution and the Poisson kernel in Euclidean $\mathbb{R}^n$.

As above, we wish to find exact solution operators for the Cauchy problem

$$\begin{cases} \Box u = 0, \\ u(x,0) = \varphi(x), \\ u_t(x,0) = \psi(x). \end{cases} \tag{4.21}$$

for the homogeneous wave equation. We define the solution operators

$$S(t) = \frac{\sin(t\sqrt{-\Delta})}{\sqrt{-\Delta}}, \quad C(t) = S'(t) = \cos t\sqrt{-\Delta}$$

of the homogeneous wave equation (4.21).

There are several methods to obtain explicit formulae for these propagators.

- Using the spherical means operator L_r.
- Using the Fourier transform.

The spherical means operator is defined by

$$L_r f(x) = \int_{S_x^* M} f(x + r\xi) dS(\xi), \tag{4.22}$$

Intuitively, it should be related to the wave equation because wave fronts are distance spheres $S_t(x) = \partial B(x,t)$ where $B(x,t)$ is the ball of radius $|t|$ around x.

A key point is that $[L_r, \Delta] = 0$ in Euclidean space. This is also true for $\mathbf{H}^n, S^n$ but it is very rarely true on a Riemannian manifold. We will take advantage of this symmetry to express $C(t), S(t)$ in terms of L_t.

It is not necessarily the case that if $[A,B] = 0$ then $A = F(B)$ for some F. But this is the case for L_r: On Euclidean space $\mathbb{R}^n$ there is a classical explicit formula

$$L_r u(x) = W_m(ir\sqrt{\Delta})u(x), \quad (W_m(z) = \Gamma(\frac{m}{2})(\frac{2}{z})^{\frac{m}{2}-1} J_{\frac{m-2}{2}}(z))$$

In the lowest dimensions, they become

$$L_r = \begin{cases} J_0(r\sqrt{-\Delta})u(x), & n = 2 \\ \frac{\sin(r\sqrt{-\Delta})}{r\sqrt{-\Delta}} u(x), & n = 3. \end{cases}$$

In Section 4.2.9 we review the so-called Pizzetti formula giving a Taylor expansion of $L_r \simeq I + \frac{r^2}{2n}\Delta + \cdots$. Note that $J_0(x) = 1 - \frac{x^2}{4} + \cdots$ and $\frac{\sin x}{x} \simeq 1 - x^2/3! + \cdots$.

The general formula on Euclidean $\mathbb{R}^n$ is given by

Proposition 4.2. *Let $u(x,t)$ be the solution of (4.21). Then,*

$$\begin{aligned} u(x,t) &= \frac{1}{\gamma_n}\frac{\partial}{\partial t}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}} \left(t^{n-2} ⨍_{\partial B(x,t)} \varphi(y) dS(y)\right) \\ &\quad + \frac{1}{\gamma_n}(\frac{1}{t}\frac{\partial}{\partial t})^{\frac{n-3}{2}} \left(t^{n-2} ⨍_{\partial B(x,t)} \psi(y) dS(y)\right) \end{aligned}$$

where $\gamma_n = 1 \cdot 3 \cdot 5 \cdot \cdots \cdot (n-2)$, *or in operator terms*

$$\begin{cases} S(t) = \frac{1}{\gamma_n}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-3}{2}} t^{n-2} L_t, \\ C(t) = \frac{1}{\gamma_n}\frac{\partial}{\partial t}\left(\frac{1}{t}\frac{\partial}{\partial t}\right)^{\frac{n-3}{2}} t^{n-2} L_t, \end{cases}$$

The explicit formula for $U(t) = \exp(-it\sqrt{-\Delta})$ in terms of spherical means involves $\sqrt{-\Delta} \cdot L_t$. The calculus of Fourier integral operators allows one to make sense of this and give formulae, but because $\sqrt{-\Delta}$ is non-local we do not expect an averaging operator over the sphere $S_t(x) = \partial B(x,t)$. But we may expect it differs from such an operator by a smoothing operator (an operator with a smooth Schwartz kernel).

4.2.1. Darboux-Euler formula. Let us make the abbreviation

$$\bar{u}(t,r;x) := L_r u(x,t),$$

and consider it for a solution $u(x,t)$ of the homogeneous wave equation. We claim that for each fixed x it is a solution of the Darboux-Euler equation

$$\begin{cases} \bar{u}_{tt} - \bar{u}_{rr} - \frac{n-1}{r}\bar{u}_r = 0, \quad 0 < r < \infty, t \geq 0, \\ \bar{u}(r,0;x) = \bar{\varphi}(x;r), \quad \bar{u}_t(r,0;x) = \bar{\psi}(r;x). \end{cases} \tag{4.23}$$

Proof.

$$\begin{aligned} \bar{u}(r,t;x) &= ⨍_{\partial B_r(x)} u(y,t)dS(y) \\ &= ⨍_{\partial B_1(0)} u(x+ry,t)dS(y). \end{aligned}$$

Hence

$$\begin{aligned} \bar{u}_r(r,t;x) &= ⨍_{\partial B_1(0)} \nabla u(x+ry,t)\cdot y dS(y) \\ &= ⨍_{\partial B_r(x)} \nabla u(y,t)\cdot \frac{y-x}{r} dS(y) \\ &= ⨍_{\partial B_r(x)} \frac{\partial u}{\partial \nu} dS(y) \\ &= \frac{1}{\mathcal{H}^{n-1}(S_r(x)} \int_{\partial B_r(x)} \frac{\partial u}{\partial \nu} dS(y) \\ &= \frac{1}{\mathcal{H}^{n-1}(S_r(x)} \int_{B_r(x)} \Delta u(y,t)dy \\ &= \frac{1}{\mathcal{H}^{n-1}(S_r(x)} \int_{B_r(x)} u_{tt}(y,t)dy. \end{aligned}$$

Then

$$\begin{aligned} \bar{u}_r(r,t;x) &= \frac{1}{\mathcal{H}^{n-1}(S_r(x)} \int_{B_r(x)} u_{tt}(y,t)dy \\ &\Longrightarrow (r^{n-1}\bar{u}_r(r,t;x))_r = \frac{1}{\mathcal{H}^{n-1}(S_1(x)} \int_{\partial B_r(x)} u_{tt}(y,t)dS(y) \\ &= r^{n-1}⨍_{\partial B_r(x)} u_{tt}(y,t)dS(y) = r^{n-1}\bar{u}_{tt}(r,t;x). \end{aligned}$$

It follows that

$$(r^{n-1}\bar{u}_r(r,t;x))_r = r^{n-1}\bar{u}_{tt}(r,t;x)$$

or equivalently

$$(n-1)r^{n-2}\bar{u}_r + r^{n-1}\bar{u}_{rr} = r^{n-1}\bar{u}_{tt}$$

and dividing by r^{n-1} gives the Darboux formula.

□

4.2.2. Proof of Propsition 4.2 in dimension 3. In dimension 3, Proposition 4.2 says:

$$u(x,t) = \frac{1}{\gamma_3}\frac{\partial}{\partial t}\left(\fint_{\partial B(x,t)} \varphi(y)dS(y)\right) + \frac{1}{\gamma_3}\left(t\fint_{\partial B(x,t)} \psi(y)dS(y)\right)$$

The first term is $C(t)$ and the second is $S(t)$.

If we set $n = 3$ in (4.23), we get the equation

$$\bar{u}_{tt} - \bar{u}_{rr} - \frac{2}{r}\bar{u}_r = 0.$$

We now prove the formula in Proposition 4.2 from this.

It is actually HW Exercise 1. Try to do it yourself. The trick is to multiply the spherical means $\bar{u}$ by r and reduce to a 1 D wave equation. The proof given below is the solution to this exercse

The equation is equivalent to

$$\begin{cases} \frac{\partial^2}{\partial t^2}(r\bar{u}) - \frac{\partial^2}{\partial r^2}(r\bar{u}) = 0 \\ r\bar{u}_{t=0} = r\bar{\varphi}, \ \ \partial_t(r\bar{u})|_{t=0} = r\bar{g}. \end{cases} \tag{4.24}$$

This is a 1D wave equation which can be solved by d'Alembert's formula:

$$\begin{aligned} r\bar{u}(x,r,t) &= \tfrac{1}{2}[(r+t)\bar{f}(x,r+t) + (r-t)\bar{f}(x,r-t)] \\ &+ \tfrac{1}{2}\int_{r-t}^{r+t} \tau\bar{g}(x,\tau)d\tau. \end{aligned} \tag{4.25}$$

Now divide by r and take the limit as $r \to 0$ to get

$$\begin{aligned} u(x,t) &= t\bar{g}(x,t) + \partial_t(t\bar{f}(x,t)) \\ &= \tfrac{1}{4\pi t}\int_{|y-x|=t} g(y)dS(y) + \tfrac{\partial}{\partial t}\left(\tfrac{1}{4\pi t}\int_{|y-x|=t} f(y)dS(y)\right). \end{aligned} \tag{4.26}$$

4.2.3. Kirchhoff formula. If $\varphi \in C^1$, then we may perform the differentiation and obtain a simpler formula, known as Kirchoff's formula:

PROPOSITION 4.3. *The solution of* (4.21) *is given by*

$$u(x,t) = \frac{1}{4\pi t^2}\int_{S_t(x)} [\varphi(y) + \nabla\varphi(y)\cdot(y-x) + t\psi(y)]dS(y).$$

HW Exercise 2 is to prove this formula.

4.2.4. Dimension two. The standard method for solving the wave equation on $\mathbb{R}^2$ is to increase the dimension by one to $\mathbb{R}^3$ and pulling back the solution $u(x_1, x_2, t)$ on $\mathbb{R}^2 \times \mathbb{R}$ to a solution $\tilde{u}(x_1, x_2, x_3, t)$ on $\mathbb{R}^2 \times \mathbb{R}$ of the Cauchy problem with pulled back data which is independent of the third coordinate. Thus, the solution is given by

$$\tilde{u}(x_1, x_2, 0, t) = \fint_{\partial B_t(\bar{x})} \left[\tilde{\varphi}(y) + \nabla\tilde{\varphi}(y) \cdot (y - x) + t\tilde{\psi}(y)\right] dS(y).$$

Here $B_t(\bar{x})$ is the ball of $\mathbb{R}^3$ of radius t around $\bar{x} = (x_1, x_2, 0)$. But if F is any function independent of the third coordinate,

$$\begin{aligned} \fint_{\partial B_t(\bar{x})} F(y) dS(y) &= \tfrac{1}{4\pi t^2} \int_{\partial B_t(\bar{x})} F(y) dS(y) \\ &= \tfrac{1}{4\pi t^2} \int_{\partial B_t(x)} F(y)(1 + |\nabla\sqrt{\Gamma}|^2)^{\frac{1}{2}} dy, \end{aligned}$$

where $B_t(x)$ is the ball of radius t around $x \in \mathbb{R}^2$ and $\Gamma(y) = (t^2 - |x-y|^2)$. Some elementary calculations then give

$$u(x,t) = \frac{1}{2\pi t^2} \int_{B_t(x)} \frac{t\varphi(y) + t^2\psi(y) + t\nabla\varphi(y) \cdot (y - x)}{\sqrt{t^2 - |x-y|^2}} dy.$$

The method of descent is universal. Given any even dimensional (M^n, g) we form the product $(M^n \times \mathbb{R}, g \oplus dx_{n+1}^2)$ and solve the wave equation on the product space with data pulled back from M^n.

4.2.5. Poisson kernel formula for $U(t) = \exp it\sqrt{-\Delta}$ in the Euclidean case. The half-wave propagator is constructed on $\mathbb{R}^n$ by the Fourier inversion formula,

$$U(t,x,y) = \int_{\mathbb{R}^n} e^{i\langle x-y,\xi\rangle} e^{it|\xi|} d\xi. \tag{4.27}$$

The Poisson kernel (extending functions on $\mathbb{R}^n$ to harmonic functions on $\mathbb{R}_+ \times \mathbb{R}^n$) is the half-wave propagagor at positive imaginary times $t = i\tau$ ($\tau > 0$),

$$\begin{aligned} U(i\tau, x, y) &= \int_{\mathbb{R}^n} e^{i\langle x-y,\xi\rangle} e^{-\tau|\xi|} d\xi \\ &= \tau^{-n} \left(1 + (\tfrac{x-y}{\tau})^2\right)^{-\frac{n+1}{2}} = \tau\left(\tau^2 + (x-y)^2)\right)^{-\frac{n+1}{2}}. \end{aligned} \tag{4.28}$$

In the case of $\mathbb{R}^n$, the Poisson kernel analytically continues to $t + i\tau, \zeta = x + ip \in \mathbb{C}_+ \times \mathbb{C}^n$ as the integral

$$U(t + i\tau, x + ip, y) = \int_{\mathbb{R}^n} e^{i(t+i\tau)|\xi|} e^{i\langle \xi, x+ip-y\rangle} d\xi, \tag{4.29}$$

which converges absolutely for $|p| < \tau$. If we substitute $\tau \to \tau - it$ and let $\tau \to 0$ we get the formula

$$U(t,x,y) = C_n \lim_{\tau\to 0} it((t+i\tau))^2 - r(x,y)^2)^{-\frac{n+1}{2}}, \tag{4.30}$$

for a constant C_n depending only on the dimension. The limit is taken in the sense of distributions and is then written

$$U(t,x,y) = C_n \; it \; ((t+i0))^2 - r(x,y)^2)^{-\frac{n+1}{2}}, \tag{4.31}$$

Background on distributions is given in the Appendix.

4.2.6. Fourier formula. The wave kernels in $\mathbb{R}^n$ may be expressed as Fourier integrals. We illustrate this only for the half-wave propagator $U(t) = \exp(it\sqrt{-\Delta})$, since we did not give a spherical means formula for it. Since

$$\delta_y(x) = \delta(x-y) = \int_{\mathbb{R}^n} e^{2\pi i \langle x-y,\xi\rangle} d\xi$$

the kernel of $U(t)$ is $U(t)\delta_y(x)$ which is

$$U(t,x,y) = \int_{\mathbb{R}^n} e^{2\pi i\langle x-y,\xi\rangle} e^{it|\xi|} d\xi.$$

If one puts the integral in polar coordinates $\xi = r\omega$, one gets

$$U(t,x,y) = \int_0^\infty \int_{S^{n-1}} e^{2\pi i r\langle x-y,\omega\rangle} e^{itr} r^{n=1} dr d\omega.$$

The spherical integral

$$J_{\frac{n-2}{2}}(r|x-y|) = \int_{S^{n-1}} e^{2\pi i r\langle x-y,\omega\rangle} d\omega$$

is a Bessel function. Hence we get

$$U(t,x,y) = \int_0^\infty J_{\frac{n-2}{2}}(r|x-y|) e^{itr} r^{n-1} dr.$$

One could go further with this calculation, e.g., $r^{n-1}e^{itr} = D_t^{n-1} e^{itr}$ so that

$$U(t,x,y) = D_t^{n-1} \int_0^\infty J_{\frac{n-2}{2}}(r|x-y|) e^{itr} dr.$$

4.2.7. Fundamental solution. An explicit formula for $S(t)$ induces one for the forward fundamental solution and in dimension 3 it says that

$$E^+ * \psi(x,t) = \frac{H(t)}{4\pi} \left(\frac{1}{t} \int_{\partial B(x,t)} \psi(y) dS(y) \right).$$

Another way to write this is that

$$E_+(t,x) = \frac{\delta(t-r)}{4\pi r}.$$

Above, we thought of the propagator as a 1-parameter family of operators on $\mathbb{R}^3$ indexed by t, but now we think of the kernels as distributions on $M \times \mathbb{R}$. In this section, we give another derivation that uses the theory of distributions rather than 'advanced calculus' and the Darboux-Euler formula from [**Ho, GeSh, F**]. It is based on pullbacks of distributions under submersions. The submersion in question is

$$Q(x,t) := t^2 - |x|^2 : \mathbb{R}^{3+1} \backslash \mathcal{N} \to \mathbb{R}\backslash\{0\},$$

where

$$\mathcal{N} = \{(x,t) : Q(x,t) = 0\}$$

is the null cone. Note that 0 is a critical value of Q and that $(0,0)$ is a critical point of Q. Hence $\mathcal{N}$ is a critical level set.

Away from the critical point it makes sense to define the pullback

$$Q^*\delta_0 = \delta_0(t^2 - |x|^2)$$

of the 1D delta function δ_0 at 0. In general, measures and distributions cannot be pulled back under maps. However, the theory of distributions gives it a meaning when the map is a submersion [**Ho, GeSh, F**]. The distribution $\delta(Q)$ is simply the 'Leray measure' on $Q^{-1}(0)$ or conditional measure on this level set. It is the measure supported on $Q^{-1}(0)$ with Gelfand-Leray form

$$Q^*\delta_0 = \frac{dxdt}{dQ}. \tag{4.32}$$

One may define its integral against a test function $\varphi \in C_c^\infty(\mathbb{R}^{3+1})$ by

$$\langle \delta_0(Q), \varphi \rangle = \int_{Q=0} \varphi \frac{dxdt}{dQ} = \int_{Q=0} \varphi \frac{dS}{|\nabla Q|}, \tag{4.33}$$

where dS is the Riemannian surface measure $\iota_\nu dxdt$ where $\nu = \frac{\nabla Q}{|\nabla Q|}$ is the unit normal.

In probability texts, the same formula is derived as follows: Let

$$\varphi_Q(t) := \frac{\partial}{\partial t} \int_{Q<t} \varphi dxdt.$$

Then

$$\langle \delta(Q), \varphi \rangle := \varphi_Q(0).$$

In the case of $Q = t^2 - |x|^2$, (4.33) gives

LEMMA 4.4. *$\delta_0(Q)$ is the following measure:*

$$\langle \delta_0(Q), \varphi \rangle = \frac{1}{2} \int_{\mathbb{R}^3} \varphi(x, |x|) \frac{dx}{|x|} + \frac{1}{2} \int_{\mathbb{R}^3} \varphi(x, -|x|) \frac{dx}{|x|}.$$

We denote the first term by $\delta_+(\varphi)$ and the second by $\delta_-(\varphi)$. Here, the first term corresponds to the upper half of the light cone where $t = |x|$ and the second term correspons to the bottom half. For the first term, we parametrize the light cone by $x \to (x, |x|)$. The Gelfand-Leray form is $\frac{dxdt}{d(t^2-|x|^2)}$ and we eliminate the variable t using $d(t^2 - |x|^2) = 2tdt - 2x \cdot dx$. The Gelfand -Leray form, is the unique form (when restricted to $Q^{-1}(0)$) satisfying $dQ \wedge \frac{dxdt}{dQ} = dxdt$ and clearly this is true $\frac{dxdt}{2tdt} = \frac{1}{2t} dx = \frac{1}{2|x|} dx$ on $Q^{-1}(0)$.

The first term above is therefore

$$\langle E^+, \varphi \rangle := \frac{1}{2} \int_{\mathbb{R}^3} \varphi(x, |x|) \frac{dx}{|x|}.$$

Similarly for the second. QED

PROPOSITION 4.5. *The following distributions on $\mathbb{R}^{3+1}$ are the forward/backward fundamental solutions:*

$$E^+(t, x) = \frac{\delta(t - r)}{4\pi r}, \quad E^-(t, x) = \frac{\delta(t + r)}{4\pi r}.$$

That is,

$$\Box E^+ = 2\pi \delta_0. \tag{4.34}$$

Hence E^+ is a fundamental solution supported in the forward light cone. Similarly for E^- in the backward light cone.

PROOF. The next observation is:

LEMMA 4.6. *$\Box \delta_0(Q) = 0$ on $\mathbb{R}^{3+1} \backslash \{0\}$.*

Here, as usual, $\Box = \frac{\partial^2}{\partial t^2} - \Delta$ is the d'Alembertian.

PROOF. We compute by the chain rule as if δ_0 were a function. Note that

$$\Box f(Q) = \nabla \cdot \nabla f(Q) = \nabla \cdot f'(Q)\nabla Q = f''(Q)\nabla Q \cdot \nabla Q + f'(Q)\Box Q,$$

where the dot product is Lorentzian. Now in dimension $3+1$, $\nabla Q \cdot \nabla Q = 4Q$ and $\Box Q = 8$. In any dimension n, if $f(t)$ homogeneous is of degre a,

$$\Box f(Q) = g(Q), \;\; g(t) := 2nf'(t) + 4tf''(t) = (2n + 4(a-1))f'(t).$$

Here we use that $tf''(t) = (a-1)f'(t)$. When $a = \frac{2-n}{2}$ the right side is zero.

Now suppose $f = \delta_0$. Then all derivatives of f are supported at 0 and f is homogeneous of degree $-1 = \frac{2-4}{2}$, the right side is zero.

□

Since E^+, E^- have disjoint supports it follows that $\Box E^{\pm} = 0$ on $\mathbb{R}^{3+1} \backslash \{0\}$. It follows that

$$\Box E^{\pm} = P(D)\delta_0$$

since all distributions supported at 0 are of this form, where $P(D)$ is a constant coefficient PDO. Now we just consider homogeneities to determine that $P(D)$ must be a constant c: Write $E^+ = \delta_+(Q)$.

- $\delta_+(Q)$ is homogeneous of degree -2.
- $\Box\delta_+(Q)$ is homogenous of degree -4.
- δ_0 is homogeneous of degree -4.
- $D^\alpha\delta_0$ is homogeneous of degree $-4 - |\alpha|$.

It follows that $\alpha = 0$ and $P(D) = c$. By using a test function $\varphi = \rho(t)$ one finds that $c = 2\pi$. (Left to reader). Hence, we proved the Proposition.

□

4.2.8. Higher dimensions. For general $\mathbb{R}^{(n-1)+1}$, one has

$$\Box f(Q) = 4Q(f''(Q)\Box Q + 2nf'(Q)).$$

PROPOSITION 4.7. *The forward fundamental solutions on (spacetime) $\mathbb{R}^n$ are*

$$\begin{cases} E^+ = \frac{1}{2\pi^m}\, \delta_+^{(m-1)}(Q), & n = 2m+2, \\ E^+ = \frac{H(t)}{2\pi^{m-\frac{1}{2}}\Gamma(\frac{3}{2}-m)} Q_+^{-m+\frac{1}{2}}, & n = 2m+1. \end{cases}$$

As in the case $n = 3+1$, the most important step is to prove:

LEMMA 4.8. $\Box f(Q) = 0$ *on* $\mathbb{R}^{(n-1)+1}\backslash\{0\}$ *if*

$$\begin{cases} f(t) = \delta^{(n-\frac{3}{2})}(t), & n \text{ even}, \\ t_+^{-\frac{n}{2}+1}, & n \text{ odd}. \end{cases}$$

We omit the proofs, which may be found in [**GeSh, Ho**].

4.2.9. Pizzetti formula. On Euclidean $\mathbb{R}^n$, there exists an exact formula known as Pizzetti's formula,

$$L_r = P_r(\Delta) := \Gamma(\frac{n}{2}) \sum_{k=0}^{\infty} (\frac{r}{2})^{2k} \frac{1}{k!\Gamma(\frac{n}{2}+k)} \Delta^k \tag{4.35}$$

which is valid on real analytic functions. The initial expansion has the form,

$$L_r = I + \frac{\Delta}{2n} r^2 + \sum_{k=2}^{\infty} P_k(\Delta) r^{2k}.$$

Let $J_\nu(z)$ be the Bessel function

$$J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+\nu+1)\Gamma(k+1)} (\frac{z}{2})^{2k+\nu}.$$

When $0 > \lambda > \lambda_0(B)$ with $B = B_r(x)$ then the eigenfunction of eigenvalue λ can be expressed as the ball mean

$$u(x) = \frac{1}{V_r 2^{\frac{n}{2}} \Gamma(\frac{n}{2}+1)} \tau_{\frac{n}{2}(\sqrt{\lambda} r)} \int_{B_r(x)} u(y) dV(y). \tag{4.36}$$

Here, $\tau_\alpha(z) = z^{-\alpha} J_\alpha(z)$. Also $W_\alpha(z) = \Gamma(\frac{\alpha}{2}) 2^{\frac{\alpha}{2}-1} \tau_{\frac{\alpha}{2}-1}(z)$.

There is a similar related formula for the ball means operator. It follows that if u is a harmonic function on $\mathbb{R}^n$ then $L_r u = P_r(0) u = u$ and similarly $M_r u = u$. A second identity is that

$$\int_{B_r(x)} \Delta u(y) dy = r^{n-1} \frac{d}{dr} \left(r^{1-n} \int_{S_r(x)} u(y) dS(y) \right) = \omega_n r^{n-1} \frac{d}{dr} L_r u(x).$$

Here $\omega_n = |S^{n-1}|$. If $\Delta u \geq 0$, i.e., if u is subharmonic, then $L_r u(x)$ increases with r. It follows that $u(x) \leq L_r u(x)$ for all r. By integrating the inequality in r, one also has $u(x) \leq M_r u(x)$.

The identity (4.35) has many repercussions for harmonic and subharmonic functions on $\mathbb{R}^n$, and also for the wave equation on $\mathbb{R}^n$. Although L_r and $S(r)$ are different functions of Δ, their Taylor expansions agree in the first two terms when $t = \frac{r}{\sqrt{n}}$. A deeper fact is that both L_r (or M_r) and $S(t)$ are Fourier integral operators associated to the same canonical relation, namely the union of the graph of the geodesic flow G^t and of G^{-t}. This is true for small $|t|$ or r on any Riemannian manifold. Therefore there exists an elliptic pseudo-differential operator $A(t, D_t, x, D_x)$ on $\mathbb{R}_t \times \mathbb{R}^n$ so that $S(t) = A L_t$. The Hadamard parametrix method gives an explicit construction of $S(t)$ in terms of operations on the spherical means operator on any manifold without conjugate points.

A classic book on the relations between the wave equation and spherical means is F. John [**J**].

4.2.10. Green's functions. Another important operator is the Green's function of the Helmholtz equation, that is, a distribution G satisfying

$$(\Delta + \lambda^2) G(\lambda, x, y) = \delta_y(x). \tag{4.37}$$

The Green's function is not unique since one may add any homogeneous solution of the Helmholtz equation to a Green's function. In the simplest case of $\mathbb{R}^3$ there are three standard choices of Helmholtz Green's function:

$$
G_0(\lambda, x, y) = -\frac{\cos(\lambda|x-y|)}{4\pi|x-y|} \quad \text{(stationary)}, \tag{4.38}
$$
$$
\begin{cases} G_0(\lambda, x, y) = -\dfrac{\cos(\lambda|x-y|)}{4\pi|x-y|} & \text{(stationary)},\\ G_+(\lambda, x, y) = -\dfrac{e^{i\lambda|x-y|}}{4\pi|x-y|} & \text{(outgoing)},\\ G_-(\lambda, x, y) = -\dfrac{e^{-i\lambda|x-y|}}{4\pi|x-y|} & \text{(incoming)}. \end{cases}
$$

In general dimension n, the free outgoing Helmholtz Green's function is given by

$$
G_+(\lambda, x, y) = \frac{i}{4}\left(\frac{\lambda}{2\pi|x|}\right)^{\frac{n-2}{2}} H^{(1)}_{\frac{n-2}{2}}(\lambda|x|). \tag{4.39}
$$

Here, $H^{(1)}_{\frac{n-2}{2}}$ is the Hankel function of the first kind.

4.3. Flat torus $\mathbb{T}^n$

A flat torus $T^n = \mathbb{R}^n/L$ is a compact quotient of $\mathbb{R}^n$ by a lattice L of full rank such as $L = \mathbb{Z}^n$. Let $x_1, \ldots, x_n$ denote the usual Euclidean coordinates on $\mathbb{R}^n$. They are not globally well-defined on T^n, but the 1-forms $dx_1, \ldots, dx_n$ are. Similarly the vector fields $\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}$ are well-defined independent vector fields on the torus. Define the dual symplectic coordinates ξ_j to the local coordinates x_j by expressing a covector as $\xi = \sum_j \xi_j\, dx_j$. Then (x, ξ) are local symplectic coordinates on T^*T^n, i.e., the canonical symplectic form is $\sigma = \sum_j d\xi_j \wedge dx_j$.

Since the cotangent bundle $T^*T^n \simeq T^n \times \mathbb{R}^n$ is a trivial vector bundle, we may define a Lagrangian torus $T_\xi \subset T^*T^n$ by

$$
T_\xi = \{(x, \xi) \colon x \in T^n\}. \tag{4.40}
$$

(Recall in general a Lagrangian submanifold $\Lambda \subset T^*M$ is a submanifold of dimension $n = \dim M$ for which the restriction of the symplectic form $\sigma|_{T^*\Lambda}$ is zero.) Under the natural projection $\pi : T^*M \to M$ the submanifold projects diffeomorphically to the base, so we have

$$
\pi : T_\xi \simeq T^n = \mathbb{R}^n/L. \tag{4.41}
$$

The symbols $\xi_1, \ldots, \xi_n$ of the differential operators $D_1, \ldots, D_n$ with $D_j = \frac{1}{i}\frac{\partial}{\partial x_j}$ generate a Hamiltonian $\mathbb{R}^n$-action on T^*T^n. The Hamilton vector fields are $\Xi_j = \frac{\partial}{\partial x_j}$ and generate translations in the base. The ξ_j are conserved quantities. In fact the joint Hamiltonian flow of the Ξ_j generates a Hamiltonian torus action $\exp t_1\Xi_1 \circ \cdots \circ \exp t_n\Xi_n$ with periodic lattice L. The tori T_ξ are orbits of this torus action. The geodesic flow is the projection of the geodesic flow of $T^*\mathbb{R}^n$ to T^*T^n and therefore has the same local expression (4.3) as in the case of $\mathbb{R}^n$.

The Lagrangian torus T_ξ is called an 'invariant torus' because it is invariant under the geodesic flow $G^t : T_\xi \to T_\xi$ in the sense that a geodesic $(x(t), \xi(t))$ starting at $(x, \xi) \in T_\xi$ stays on the torus T_ξ. Recall that a geodesic in T^n is the projection to the torus of the straight line $x(0) + t\frac{\xi}{|\xi|}$ in $\mathbb{R}^n$ under the covering map $\mathbb{R}^n \to T^n$. The image of the straight line is a 'winding line' on T^n, and every geodesic in T_ξ is a translate of a single geodesic. If $\xi = \ell \in L$, the winding line is the projection of a line segment from $0 \in \mathbb{R}^n$ to ℓ, and is therefore a closed geodesic. Thus, all geodesics on T_ℓ is periodic, and T_ℓ is sometimes called a 'periodic torus.'

The flat Laplacian commutes with the vector fields $\frac{\partial}{\partial x_j}$ so, on the flat torus (which is compact), there exists an orthonormal basis of joint eigenfunctions given by $e^{2\pi i\langle\lambda,x\rangle}$, where $\lambda\in L^*$ lies in the dual lattice L^* to L. Recall that points of the dual lattice are vectors $\tilde{\lambda}$ such that $\langle\lambda,\ell\rangle\in\mathbb{Z}$ for $\ell\in L$. The corresponding Laplace eigenvalue is $-(2\pi)^2|\lambda|^2$, that is,

$$(\Delta+(2\pi)^2|\lambda|^2)e^{2\pi i\langle\lambda,x\rangle}=0. \tag{4.42}$$

As before, we denote the Laplacian eigenspace of eigenvalue $-(2\pi)^2|\lambda|^2$ by $\mathcal{E}^L_{2\pi\lambda}$. It is a subspace of L-periodic elements of the Euclidean eigenspace $\mathcal{E}_{2\pi\lambda}$ defined in (4.5). For a generic lattice L, the eigenvalue $-(2\pi)^2|\lambda|^2$ has multiplicity two and $\mathcal{E}^L_{2\pi\lambda}$ is spanned by the two eigenfunctions $e^{\pm 2\pi i\langle x,\lambda\rangle}$. In contrast, when L is an arithmetic lattice such as $\mathbb{Z}^n$, then there exists a high multiplicity of eigenvalues. By the exact formula (4.42) for the eigenvalues, the multiplicity is the number of lattice points of the dual lattice L^* lying on the surface of a Euclidean sphere of radius $(2\pi)^2|\lambda|^2$. The multiplicities behave differently depending on the dimension n of the torus T^n. Let $-\mu_1^2<-\mu_2^2<\cdots$ be distinct Laplacian eigenvalues. In dimension $n\geq 5$, the multiplicity of $-\mu_k^2$ grows at the rate k^{n-2}. In dimensions $n\leq 4$ the situation is more complicated. Under certain assumptions, the lattice points tend to become uniformly distributed on the frequency sphere as the eigenvalue tends to infinity. Since the main input is arithmetic rather than analytic we do not review the results in this monograph and refer to [**EH, DSP**] for two relevant research articles.

Recall the integral representation (4.6) for an eigenfunction. In the case of a flat torus, the boundary distribution dT corresponding to $e^{\pm 2\pi i\langle x,\lambda\rangle}$ is simply $\delta_{\pm 2\pi\lambda}$. Hence, these periodic eigenfunctions are not in $\mathcal{E}^{(2)}_{2\pi\lambda}$. General linear combinations $\sum_{\lambda:|\lambda|=R}e^{i\langle x,\lambda\rangle}a_\lambda$ with $\sum|a_\lambda|^2=1$ of plane waves of a fixed eigenvalue for arithmetic tori exhibit complicated behavior that has been studied by D. Jakobson [**J**], N. Anantharaman and F. Macia [**AM**] and Hezari-Riviere [**HR**] (among others). They prove that the projections of the Wigner distributions to the base of any sequence of eigenfunctions is diffuse.

A flat torus is a model of quantum complete integrability, which pertains mainly to the joint eigenfunctions. A key feature of the joint eigenfunctions is that they have the form $a(x)e^{iS(x)/h}$ with $h=|\lambda|^{-1}$. Such functions are known as WKB modes or Lagrangian states. Notice that the 'phase function' $S(x)=\langle x,\frac{\lambda}{|\lambda|}\rangle$ has the property

$$\operatorname{graph}(dS):=\{(x,dS(x))\colon x\in T^n\}=\left\{(x,\xi)\colon \xi=\frac{\lambda}{|\lambda|}\right\}. \tag{4.43}$$

We therefore say that S is the generating function of the Lagrangian torus $\{(x,\xi=\frac{\lambda}{|\lambda|})\}$ in $T^*(T^n)$. By earlier discussion, the Lagrangian tori corresponding to joint eigenfunctions are those of the form T_λ where $\lambda\in L^*$. These tori and their generalizations are known as Bohr-Sommerfeld tori. They are dual in a sense to the periodic tori T_ℓ.

4.4. Spheres S^n

In this section we consider eigenfunctions and geodesic flow on the standard unit sphere S^n. The isometry group of S^n is the orthogonal group $O(n+1)$ of isometries of $\mathbb{R}^{n+1}$ which fix the origin. The orthogonal group is a matrix group with elements

$A \in GL(n+1, \mathbb{R})$ such that $A^t A = I$. Equivalently, such matrices satisfy $\langle Ax, Ay \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{R}^{n+1}$. The special orthogonal group $SO(n+1) \subset O(n+1)$ is the subgroup of matrices with $\det A = 1$, i.e, the group of rotations. The isometry group acts transitively on the sphere and the isotropy subgroup of the north pole is $SO(n-1)$. Hence $S^n = SO(n)/SO(n-1)$. The geodesics of S^n are the great circles centered at $0 \in \mathbb{R}^{n+1}$. All of the geodesics are closed and the geodesic flow is periodic of period 2π.

We now consider the quantum picture, described in quantum mechanics texts as the theory of 'angular momentum'. Under the identification $\mathbb{R}^{n+1} \simeq \mathbb{R}^+ \times S^n$, the Euclidean Laplacian can be written in polar coordinates (r, ω) as

$$\Delta_{\mathbb{R}^{n+1}} = \frac{\partial^2}{\partial r^2} + \frac{n-1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\Delta_{S^n}. \tag{4.44}$$

Here, Δ_{S^n} is the metric Laplacian on the sphere with the usual round metric. It is often written $|L|^2$ in physics texts.

Spherical harmonics are eigenfunctions of Δ_{S^n}. Let $P(x) = P(x_1, \ldots, x_{n+1})$ be a polynomial on $\mathbb{R}^{n+1}$, then recall

- P is a homogeneous of degree k if $P(rx) = r^k P(x)$. We denote the space of such polynomials by $\mathcal{P}_\ell$. A basis is given by the monomials

$$x^\alpha = x_1^{\alpha_1} \cdots x_{n+1}^{\alpha_{n+1}}, \quad \text{where } |\alpha| = \alpha_1 + \cdots + \alpha_{n+1} = \ell. \tag{4.45}$$

- P is a harmonic if $\Delta_{\mathbb{R}^{n+1}} P(x) = 0$. We denote the space of harmonic homogeneous polynomials of degree ℓ by $\mathcal{H}_\ell$.
- The restriction to S^n of a harmonic homogeneous (of degree ℓ) polynomial is a spherical harmonic (of degree ℓ).

Every homogeneous polynomial can be decomposed into a sum of harmonic homogeneous polynomials. Indeed, let $Q \in \mathcal{P}_\ell$, then

$$Q(x) = P_0(x) + |x|^2 P_1(x) + \cdots + |x|^{2\ell} P_\ell(x), \tag{4.46}$$

where P_j is a homogeneous harmonic polynomial of degree $k - 2j$ for $j = 0, \ldots, \ell$. Now suppose that $P \in \mathcal{H}_\ell$ is a homogeneous harmonic polynomial of degree ℓ on $\mathbb{R}^{n+1}$. Then, writing everything in polar coordinates, it follows from the definitions that

$$0 = \Delta_{\mathbb{R}^{n+1}} P = \left(\frac{\partial^2}{\partial r^2} + \frac{n-1}{r}\frac{\partial}{\partial r} \right) r^\ell P(\omega) + \frac{1}{r^2}\Delta_{S^n} P(\omega), \tag{4.47}$$

thereby implying

$$\Delta_{S^n} P|_{S^n} = -(\ell(\ell-1) + (n-1)\ell) P|_{S^n} = -\ell(\ell+n-2) P|_{S^n}. \tag{4.48}$$

This shows that the restriction of $P \in \mathcal{H}_\ell$ to the unit sphere (i.e., a spherical harmonic of degree ℓ by definition) is an eigenfunction of Δ_{S^n} with eigenvalue $-\ell(\ell+n-2)$.

THEOREM 4.9. *Let $\mathcal{H}_\ell$ denote the space of spherical harmonics of degree ℓ, then $L^2(S^n) = \bigoplus_{\ell=0}^\infty \mathcal{H}_\ell$ is a direct sum of orthogonal subspaces of dimensions*

$$\dim \mathcal{H}_\ell = \binom{n+\ell-1}{n-1} + \binom{n+\ell-3}{n-1}. \tag{4.49}$$

We only give a brief sketch of the proof. Orthogonality of $\mathcal{H}_\ell$ is obvious because they are distinct eigenspaces of a self-adjoint operator. Thanks to our earlier observation (4.46), every homogeneous polynomial restricted to S^n can be written as the sum of restrictions of harmonic polynomials. Thus, to establish the spanning property it suffices to show that the restrictions of harmonic polynomials are dense. This follows from the Stone-Weierstrass theorem.

theo In particular when $n = 3$ the eigenvalues are $-\ell(\ell+2) = -(\ell+1)^2 + 1$ and the multiplicity is ℓ^2.

The operator $\mathcal{N}$ whose eigenvalue on $\mathcal{H}_\ell$ is ℓ is known as the degree operator. Consider S^3 where the eigenvalue of Δ is $-(\ell+1)^2 + 1$. Then $\Delta - I$ is a perfect square and we define

$$A = \sqrt{-\Delta + I}.$$

Then the eigenvalues of A are $\ell + 1$ so

$$\mathcal{N} = A - I.$$

In general,

$$A = \sqrt{-\Delta + (\frac{n-1}{2})^2}.$$

A key object in the theory of spherical harmonics is the orthogonal projector

$$\Pi_\ell \colon L^2(S^n) \to \mathcal{H}_\ell \tag{4.50}$$

whose Schwartz kernel $\Pi_\ell(x, y)$ is defined by

$$\Pi_\ell f(x) = \int_{S^n} \Pi_\ell(x, y) f(y)\, dS(y). \tag{4.51}$$

Here, f is any L^2 function on the sphere and dS is the standard surface measure. We note that the along diagonal $\Pi_\ell(x, x) = C_\ell$ is a constant because it is rotationally invariant and $O(n+1)$ acts transitively on S^n. Indeed, by integrating we find

$$\Pi_\ell(x, x) = \frac{\dim \mathcal{H}_\ell}{\mathrm{Vol}(S^n)}. \tag{4.52}$$

4.4.1. Special case: the 2-sphere S^2. On spheres of any dimension there exists a special orthonormal basis of so-called weight vectors, a term from representation theory. It is easiest to explain when $n = 2$. Recall that the Lie algebra of $SO(3)$ is generated by the vector fields

(4.53)

$$L_1 = i\left(x_3 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_3}\right), \quad L_2 = i\left(x_1 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_1}\right), \quad L_3 = i\left(x_2 \frac{\partial}{\partial x_1} - x_1 \frac{\partial}{\partial x_2}\right).$$

Put

$$|L|^2 := L_1^2 + L_2^2 + L_3^2, \tag{4.54}$$

then $|L|^2 = \Delta_{S^2}$ is precisely the Laplacian on the sphere and

$$\Delta_{\mathbb{R}^3} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 \frac{\partial}{\partial r}) - \frac{1}{r^2}|L|^2. \tag{4.55}$$

The subgroup $SO(2) \subset SO(3)$ of rotations around the x_3-axis commutes with the Laplacian. We denote its infinitesimal generator by $L_3 = \frac{1}{i}\frac{\partial}{\partial \theta}$. The 'special' orthonormal basis we alluded to earlier consists of the joint eigenfunctions Y_ℓ^m of $|L|^2$ and of L_3 satisfying

$$\begin{cases} \Delta_{S^2} Y_\ell^m = -\ell(\ell+1) Y_\ell^m \\ \dfrac{1}{i}\dfrac{\partial}{\partial \theta} Y_\ell^m = m Y_\ell^m. \end{cases} \tag{4.56}$$

$$Y_\ell^m(\varphi,\theta) = \sqrt{(2\ell+1)\frac{(\ell-m)!}{(\ell+m)!}}P_\ell^m(\cos\varphi)e^{im\theta}, \tag{4.57}$$

where

$$P_\ell^m(\cos\varphi) = \frac{1}{2\pi}\int_0^{2\pi}(i\sin\varphi\cos\theta+\cos\varphi)^\ell e^{-im\theta}d\theta \tag{4.58}$$

are the Legendre polynomials.

$$P_\ell^m(x) = \sqrt{(2\ell+1)\frac{(\ell-m)!}{(\ell+m)!}}\frac{1}{2^\ell \ell!}(1-x^2)^{-\frac{m}{2}}\frac{d^{\ell-m}}{dx^{\ell-m}}(x^2-1)^\ell. \tag{4.59}$$

Under this orthonormal basis, the kernel of the spectral projection $\Pi\colon L^2(S^2)\to\mathcal{H}_\ell$ is thus

$$\Pi_\ell(x,y) = \sum_m Y_\ell^m(x)\overline{Y_\ell^m}(y). \tag{4.60}$$

4.4.2. From Poisson integral to half-wave kernel. We recall that the Poisson integral formula for the unit ball is:

$$u(x) = \fint_{S^n}\frac{1-|x|^2}{|x-\omega'|^2}f(\omega')dS(\omega').$$

Write $x=r\omega$ with $|\omega|=1$ to get:

$$P(r,\omega,\omega') = \frac{1-r^2}{(1-2r\langle\omega,\omega'\rangle+r^2)^{\frac{n+1}{2}}}.$$

A second formula for $u(r\omega)$ is

$$u(r,\omega) = r^{A-\frac{n-1}{2}}f(\omega) = e^{-t(A-\frac{n-1}{2})}f(\omega),$$

where $A=\sqrt{-\Delta+(\frac{n-1}{2})^2}$ and where $t=\log\frac{1}{r}$. Thus, harmonic extension is written as an evolution equation with generator $A-\frac{n-1}{2}$.[1] This follows from by writing the equation $\Delta_{\mathbb{R}^{n+1}}u=0$ as an Euler equation:

$$\left(r^2\frac{\partial^2}{\partial r^2}+nr\frac{\partial}{\partial r}-\Delta_{S^n}\right)u=0.$$

Another explanation is that on the space $\mathcal{H}_N$ of spherical harmonics of degree N on S^n, the harmonic extension is simply the homogeneous extension as a polynomial of degree N, i.e. by r^N. But $A|_{\mathcal{H}_N}=N+\frac{n-1}{2}$. For instance, in dimension 2, $-\Delta|_{\mathcal{H}_N}=N(N+1)=(N+\frac{1}{2})^2-\frac{1}{4}$, so $A-\frac{1}{2}=N$.

The Poisson operator kernel with $r=e^{-t}$ is given by

$$P(t,\omega,\omega') \quad = \quad C_n\frac{\sinh te^{-(n-1)t}}{(\cosh t-\cos r(\omega,\omega'))^{\frac{n+1}{2}}}.$$

It follows that

$$e^{-tA} = C_n\sinh t(\cosh t-\cos r(\omega,\omega'))^{-\frac{n+1}{2}}.$$

[1] It resembles a heat equation but the generator is roughly $\sqrt{-\Delta}$ rather than $-\Delta$. It is a jump process rather than a continuous one.

Note that $U(t) = e^{itA}$, resp. $P(t) = e^{-tA}$ has the Schwartz kernel

$$\sum_{N=0}^{\infty} e^{it(N+\frac{n-1}{2})}\Pi_N(\omega,\omega'), \text{ resp. } \sum_{\mathrm{N}=0}^{\infty} \mathrm{e}^{-\mathrm{t}(\mathrm{N}+\frac{\mathrm{n}-1}{2})}\Pi_{\mathrm{N}}(\omega,\omega'),$$

where $\Pi_N : L^2(S^n) \to \mathcal{H}_N$. Thus, the $P(t) = U(it)$ for $t > 0$. The Schwartz kernel of $U(t)$ is thus obtained by analytically continuing the Poisson kernel in time. For $t > 0$, $P(t + i\tau)$ is a smoothing operator, but its boundary value at $t = 0$ is the distributional kernel $U(\tau)$. We thus have,

Proposition 4.10.

$$e^{itA} = \lim_{\varepsilon\to 0^+} C_n i \sin t(\cosh\varepsilon\cos t - i\sinh\varepsilon\sin t - \cos r(\omega,\omega')^{-\frac{n+1}{2}}$$

$$= \lim_{\varepsilon\to 0^+} C_n i\sinh(it-\varepsilon)(\cosh(it-\varepsilon) - \cos r(\omega,\omega')^{-\frac{n+1}{2}}.$$

If we formally put $\varepsilon = 0$ we obtain:

$$e^{itA} = C_n i \sin t(\cos t - \cos r)^{-\frac{n+1}{2}}.$$

This expression is singular when $\cos t = \cos r$ and only well-defined if we recall that it is the distribution boundary value above. We note that $r \in [0,\pi]$ and that it is singular on the cut locus $r = \pi$. Also, $\cos : [0,\pi] \to [-1,1]$ is decreasing, so the wave kernel is singular when $t = \pm r$ if $t \in [-\pi,\pi]$.

When n is even, the expression appears to be pure imaginary but that is because we need to regularize it on the set $t = \pm r$. When n is odd, the square root is real if $\cos t \geq \cos r$ and pure imaginary if $\cos t < \cos r$.

We see that the kernels of $\cos tA$, $\frac{\sin tA}{A}$ are supported inside the light cone $|r| \leq |t|$. On the other hand, e^{itA} has no such support property (it has infinite propagation speed). On odd dimensional spheres, the kernels are supported on the distance sphere (sharp Huyghens phenomenon).

4.4.3. Fundamental solution and Propagators on S^3. Given the above formula for e^{itA}, we can read off the formula for the propagators. We only record the formulae in dimension 3.

Proposition 4.11. *On S^3 for $t > 0$,*

$$\frac{\sin t\sqrt{-(\Delta+1)}}{\sqrt{(-(\Delta+1)}}\delta_y(x) = \frac{\delta(t-r)}{4\pi\sin t}, \quad \cos t\sqrt{-(\Delta+1)}\delta_y(x) = \frac{\delta'(t-r)}{4\pi\sin r}$$

In the next section we give a direct proof of the analogous formula on hyperbolic space. By analogy with the Euclidean case, we define

$$Q(t,x,y) = \cos t - \cos r$$

on $S^n \times \mathbb{R}$. The first formula in the Proposition is equivalent to the fact that $E = \delta(Q)$ is a fundamental solution on S^3, the sum $E = E^+ + E^-$ of the forward and backward fundamental solutions. In fact, one may show directly that $(\Box - 1)\delta(Q) = \delta_0$ where $\Box = \frac{\partial^2}{\partial^2} - \Delta$. This is done in the Remark at the end of the next section on Hyperbolic space.

4.4.4. Complete integrability of S^2. Consider the operator $R = (-\Delta + \frac{1}{4})^{\frac{1}{2}}$. It is easy to see that the standard spherical harmonics defined in (4.57) are eigenfunctions of R with eigenvalues $k + \frac{1}{2}$. The operators R and L_3 are first order commuting pseudo-differential operators. Their joint spectrum is the shifted semi-lattice $(m, \ell + \frac{1}{2}) \subset \mathbb{R}_+ \times \mathbb{R}_+$.

The pair R, L_3 is analogous to D_{x_1}, D_{x_2} on the flat torus. Yet there are some interesting differences. Consider the 'symbols' R and L_3. The symbol of R is the metric norm function $|\xi|$ while that of L_3 is the so-called Clairaut integral $p_\theta(x,\xi) = \langle \xi, \frac{\partial}{\partial\theta}\rangle$. The pair $(p_\theta, |\xi|)$ is called the moment map of the completely integrable geodesic flow of S^2. By the Cauchy-Schwarz inequality, $|p_\theta(x,\xi)| \leq |\frac{\partial}{\partial\theta}| \leq 1$ when $|\xi| = 1$. Hence the image of T^*S^2 under the moment map is a triangular cone in $\mathbb{R}^2$ with vertex at the origin, central axis the y-axis and sides $y = \pm x$ in Cartesian coordinates. Compare this to the image of T^*T^2 under the moment map (ξ_1, ξ_2), which is the whole of $\mathbb{R}^2$.

The inverse image of a point $(x_0, y_0) \in \mathbb{R}^2$ of the triangular region under the moment map is the set of points $(x, \xi) \in T^*S^2$ such that $(p_\theta(x,\xi), |\xi|) = (x_0, y_0)$. It is easy to see that this inverse image is invariant under the x_3-axis rotations and under the geodesic flow (i.e., under the Hamiltonian flows of the components of the moment map). Hence, it is a Lagrangian torus. In particular, the image of T^*S^2 under the moment map has a boundary, corresponding singular points of the moment map. The singular points are the unit vectors along the equatorial geodesic, traversed in either of its two orientations. By contrast, the moment map of T^2 is everywhere regular.

It is helpful (and accurate) to imagine Y_ℓ^m and its joint eigenvalue $(m, \ell + \frac{1}{2})$ as corresponding to the torus with $p_\theta(x,\xi) = m$ and $|\xi| = \ell$. If we rescale (by homogeneity) back to S^*S^2 this is $p_\theta = m/k$, which defines a 2-torus. For instance, the central axis is $p_\theta = 0$ and that corresponds to longitudinal great circles, which depart from the north pole, converge at the south pole and then return to the north pole. This is the picture of zonal spherical harmonics.

4.4.5. Special spherical harmonics. We will see that inequalities involving eigenfunctions are often saturated by zonal or highest weight spherical harmonics. Recall the spectral projection kernel $\Pi_\ell(x,y)$. We L^2-normalize this function by dividing by the square root of

$$\|\Pi_\ell(\cdot, y)\|_{L^2}^2 = \int_{S^n} \Pi_\ell(x,y)\Pi_\ell(y,x)\, dS(x) = \Pi_\ell(x,x). \tag{4.61}$$

The resulting function

$$\Phi_\ell^x(y) = \frac{\Pi_\ell(x,y)}{\sqrt{\Pi_\ell(x,x)}}. \tag{4.62}$$

is called the coherent state or zonal spherical harmonic. The 'coherent state' is always the extremal for pointwise norm at y among eigenfunctions $\varphi_\lambda \in V_\lambda$
(4.63)

$$\varphi_\lambda(x) = \int_M \Pi_\lambda(x,y)\varphi_\lambda(y)dy \implies |\varphi_\lambda(x)| \leq \sqrt{\int_M |\Pi_\lambda(x,y)|^2 dy} = \sqrt{\Pi_\lambda(x,x)} = |\Phi_\lambda^x(x)|.$$

The zonal spherical harmonic constructed above is the same as

$$Y_\ell^0(\varphi,\theta) = Y_\ell^0(\varphi) = \sqrt{\frac{(2\ell+1)}{2\pi}} P_\ell(\cos\varphi). \tag{4.64}$$

Another important spherical harmonic is Y_ℓ^ℓ, which is the spherical harmonic in $\mathcal{H}_\ell$ with the largest eigenvalue of L_3, or in other words the highest weight. It corresponds to the boundary points of the triangular image of the moment map, hence to the equatorial great circle.

Indeed, Y_ℓ^ℓ is the restriction of the harmonic polynomial $(x_1 + ix_2)^\ell$ up to normalization. This polynomial is independent of x_3 and is a holomorphic function of x_1 and x_2 so it is certainly a harmonic homogeneous polynomial. Also, by its form it clearly is largest on the unit circle in the (x_1, x_2) plane and tends to zero as $(x_1, x_2) \to (0, 0)$. Hence on the sphere it defines a spherical harmonic which is large on the equator and tends to zero at the poles.

Note that

$$\int_{\mathbb{R}^3} (x_1^2 + x_2^2)^\ell e^{-(x_1^2+x_2^2+x_3^2)}\, dx_1dx_2dx_3 = \|(x_1 + ix_2)^\ell\|^2_{L^2(S^2)} \int_0^\infty r^{2\ell+2} e^{-r^2}\, dr, \tag{4.65}$$

Converting the integral on the left-hand side to polar coordinates, we find

$$\int_{\mathbb{R}^3} (x_1^2 + x_2^2)^\ell e^{-(x_1^2+x_2^2+x_3^2)}\, dx_1dx_2dx_3 = \int_0^\infty r^{2\ell+1} e^{-r^2}\, dr. \tag{4.66}$$

It follows that

$$\|(x_1 + ix_2)^\ell\|^2_{L^2(S^2)} = \frac{\Gamma(\ell + \frac{1}{2})}{\Gamma(\ell + 1)} \sim \ell^{-\frac{1}{2}}. \tag{4.67}$$

We define the normalized highest weight spherical harmonics (also known as Gaussian beams) by the restriction of $\ell^{\frac{1}{4}}(x + iy)^\ell$. It achieves its L^∞ norm at $(1, 0, 0)$ where it has size $\ell^{\frac{1}{4}}$. In general, Gaussian beams on S^n are transverse Gaussian bumps which are concentrated on $\lambda^{-\frac{1}{2}}$ tubes around closed geodesics and have height $\lambda^{\frac{n-1}{4}}$.

4.5. Hyperbolic space and non-Euclidean plane waves

The most natural analogue of the sphere S^n for constant curvature -1 is hyperbolic space $\mathbf{H}^n$, a globally symmetric space. In this section we review a natural basis of eigenfunctions which are analogous to plane waves in Euclidean space. There is also a joint basis of spherical eigenfunctions, i.e., joint eigenfunctions for rotations around the origin and the Laplacian. Hyperbolic is substantially more complicated than S^n due to the fact that it is of infinite volume. We stick to the Hyperbolic plane $n = 2$ because the representation theory is substantially simpler.

4.5.1. Hyperbolic plane $\mathbf{H}^2$. Hyperbolic surfaces are uniformized by the hyperbolic plane $\mathbf{H}^2$ or disc $\mathbf{D}$. In the disc model $\mathbf{D} = \{z \in \mathbb{C} : |z| < 1\}$, the hyperbolic metric has the form

$$ds^2 = \frac{4|dz|^2}{(1 - |z|^2)^2}.$$

The group of orientation-preserving isometries can be identified with $PSU(1, 1)$ acting by Möbius transformations; the stabilizer of 0 is $K \simeq SO(2)$ and thus we will often identify $\mathbf{D}$ with $SU(1, 1)/K$. Computations are sometimes simpler in the

$\mathbf{H}^2$ model, where the isometry group is $PSL(2,\mathbb{R})$. We therefore use the general notation G for the isometry group, and G/K for the hyperbolic plane, leaving it to the reader and the context to decide whether $G = PSU(1,1)$ or $G = PSL(2,\mathbb{R})$.

We denote by $B = \{z \in \mathbb{C} \colon |z| = 1\}$ the boundary at infinity of $\mathbf{D}$. The unit tangent bundle $S\mathbf{D}$ of the hyperbolic disc $\mathbf{D}$ is by definition the manifold of unit vectors in the tangent bundle $T\mathbf{D}$ with respect to the hyperbolic metric. We may, and will, identify $S\mathbf{D}$ with the unit cosphere bundle $S^*\mathbf{D}$ by means of the metric. We will make a number of further identifications:

- $S\mathbf{D} \equiv PSU(1,1)$. This comes from the fact that $PSU(1,1)$ acts freely and transitively on $S\mathbf{D}$. Similarly, if we work with the upper half plane model $\mathbf{H}^2$, we have $S\mathbf{H}^2 \equiv PSL(2,\mathbb{R})$. We identify a unit tangent vector (z,v) with a group element g if $g \cdot (i,(0,1)) = (z,v)$. We identify $S\mathbf{D}$, $S\mathbf{H}^2$, $PSU(1,1)$, and $PSL(2,\mathbb{R})$. In general, we work with the model which simplifies the calculations best. According to a previous remark, $S\mathbf{D}$, $PSU(1,1)$ and $PSL(2,\mathbb{R})$ will often be designated by the letter G.
- $S\mathbf{D} \equiv \mathbf{D} \times B$. Here, we identify $(z,b) \in \mathbf{D} \times B$ with the unit tangent vector (z,v), where $v \in S_z\mathbf{D}$ is the vector tangent to the unique geodesic through z ending at b.

The geodesic flow G^t on $S\mathbf{D}$ is defined by $G^t(z,v) = (\gamma_v(t), \gamma_v'(t))$ where $\gamma_v(t)$ is the unit speed geodesic with initial value (z,v). The space of geodesics is the quotient of $S\mathbf{D}$ by the action of G^t. Each geodesic has a forward endpoint b and a backward endpoint b' in B, hence the space of geodesics of $\mathbf{D}$ may be identified with $B \times B \setminus \Delta$, where Δ denotes the diagonal in $B \times B$: To $(b',b) \in (B \times B) \setminus \Delta$ there corresponds a unique geodesic $\gamma_{b',b}$ whose forward endpoint at infinity equals b and whose backward endpoint equals b'.

We then have the identification

$$S\mathbf{D} \equiv (B \times B \setminus \Delta) \times \mathbb{R}.$$

The choice of time parameter is defined – for instance – as follows: The point $(b',b,0)$ is by definition the closest point to 0 on $\gamma_{b',b}$ and (b',b,t) denotes the point t units from $(b',b,0)$ in signed distance towards b.

4.6. Dynamics and group theory of $G = PSL(2,\mathbb{R})$

We recall the group theoretic point of view towards the geodesic and horocycle flows on the unit cotangent bundle $S^*\mathbf{X}_\Gamma$ of $\mathbf{X}_\Gamma = \Gamma\backslash X$. As stated above, it is equivalent to work with $G = PSU(1,1)$ or $G = PSL(2,\mathbb{R})$; we choose the latter. Our notation follows [**L, AZ**] (see also [**K, Hel2**]) except for the normalization of the metric. The generators of the Lie algebra $\mathfrak{sl}(2,\mathbb{R})$ are denoted by

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad W = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{4.68}$$

The associated one parameter subgroups are denoted by A, A_-, K. Let

$$E^+ = H + iV \quad \text{and} \quad E^- = H - iV \tag{4.69}$$

be the raising/lowering operators for K-weights. The Casimir operator is then given by $4\Omega = H^2 + V^2 - W^2$. On K-invariant functions, the Casimir operator coincides with the metric Laplacian. We also put

$$X_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad X_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{4.70}$$

and denote the associated subgroups by N, N_-.

In the identification $S\mathbf{D} \equiv G$, the geodesic flow is given by the right action of the group A of diagonal matrices, that is, $G^t(g) = ga_t$ where

$$a_t = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}. \tag{4.71}$$

The action of the geodesic flow is closely related to that of the horocycle flow $(h^u)_{u\in\mathbb{R}}$. The horocycle flow is defined by the right action of N, that is, by $h^u(g) = gn_u$ where

$$n_u = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}. \tag{4.72}$$

4.7. The Hyperbolic Laplacian

In hyperbolic polar coordinates centered at the origin, the Laplacian is the operator

$$\Delta = \frac{\partial^2}{\partial r^2} + \coth r \frac{\partial}{\partial r} + \frac{1}{\sinh^2 r}\frac{\partial^2}{\partial \theta^2}. \tag{4.73}$$

The distance on $\mathbf{D}$ induced by the Riemannian metric will be denoted $d_{\mathbf{D}}$. We denote the volume form by $d\,\mathrm{Vol}(z)$.

There are standard notations for writing the eigenvalue of Δ:

$$\lambda^2 = \begin{cases} \lambda_r^2 = \dfrac{1}{4} + r^2 \\ \lambda_s^2 = s(1-s) \quad \text{where } s = \dfrac{1}{2} + ir. \end{cases} \tag{4.74}$$

The reason for introducing λ_r is to shift the bottom of the spectrum $\frac{1}{4}$ of Δ on $\mathbf{D}$ to the origin. The reason for the second is that the parameter s parametrizes the irreducible representations of G (see [**L, K**]). We also put $R = (-\Delta + \frac{1}{4})^{\frac{1}{2}}$ and $U^t = e^{itR}$ the wave operator of $\mathbf{D}$.

As in the Euclidean case we consider the space

$$\mathcal{E}_s = \{f \in \mathcal{S}'(\mathbf{H}^n) : \Delta f = s(1-s)f\} \tag{4.75}$$

of temperate eigenfunctions. The definition of 'temperate' is more involved in the Euclidean case since the volume of balls is exponentially growing. Schwartz functions on G were first defined by Harish-Chandra; the definition was extended to $\mathbf{D} = G/K$ by Eguchi and his collaborators [**Eg**]. Identifying the hyperbolic disc with G/K, a function f belongs to the Schwartz space $\mathcal{C}^p(G/K)$ (for $0 < p \leq 2$) if and only if it is a right-K-invariant smooth function on G satisfying

$$\sup_{g\in G} \varphi_0(gK)^{-2/p}(1 + d(gK, 0))^q |L_1 L_2 f(g)| < +\infty \tag{4.76}$$

for any $q > 0$ and for any differential operators L_1, L_2 on G that are respectively left- and right-invariant. Here $0 \in G$ is the identity element and φ_0 is the spherical function on G/K, which satisfies $\varphi_0(z) \asymp d(z,0)e^{-d(z,0)/2}$ as the hyperbolic distance $d(z,0) \to +\infty$. Functions on $\mathcal{C}^p(G/K)$ are, in particular, in L^p (they are sometimes called Schwartz functions of L^p-type).

4.8. Wave kernel and Poisson kernel on Hyperbolic space $\mathbf{H}^n$

$\mathbf{H}^n$ is the symmetric space G/K where $G = SO(1,n)_0$ and $K = SO(n)$. In geodesic polar coordinates centered at any point y, the meteric has the form

$$g = dr^2 + \sinh^2 r g_{S^{n-1}}$$

and the Riemannian volume form is

$$d\,\mathrm{Vol} = C_n (\sinh r)^{n-1} dr d\omega$$

and the Laplace operator is

$$\Delta = \partial_r^2 + (n-1)\coth r \partial_r + \sinh r^{-2} \Delta_{S^{n-1}}.$$

Also, the gradient is

$$\nabla = \sum_{i,j} g^{ij} \frac{\partial}{\partial x_i} e_j = \frac{\partial|}{\partial r} + G\nabla_\omega.$$

In hyperbolic polar coordinates centered at the origin, the Laplacian is the operator

$$\Delta = \frac{\partial^2}{\partial r^2} + \coth r \frac{\partial}{\partial r} + \frac{1}{\sinh^2 r}\frac{\partial^2}{\partial \theta^2}.$$

It is known that the spectrum of $\Delta + 1$ is $[0, \infty]$.

4.8.1. Sine wave kernel. Let $n = 3$ and consider the sine wave kernel $S(t)$ for $\Delta + 1$ on $\mathbf{H}^3$. As in the Eucliean case, let $E = E^+ + E^-$ be the fundamental solution where $E^\pm$ is the forward/backward fundamental solution.

PROPOSITION 4.12. *Let $Q = \cosh t - \cosh r$. Then:*

- *$E = \delta(Q)$, where*

$$\langle \delta(Q), \varphi \rangle = C_3 [\int_{\mathbf{H}^3} \varphi(x,r) \frac{\sinh^2 r dr dt dd\omega}{\sinh r} + \int_{\mathbf{H}^3} \varphi(x,-r) \frac{\sinh^2 r dr dt dd\omega}{\sinh r}.]$$

- *For $t > 0$,*

$$E^+ = \frac{\sin t\sqrt{-(\Delta+1)}}{\sqrt{(-(\Delta+1)}} \delta_y(x) = \frac{\delta(t-r)}{4\pi \sinh t}, \quad \cos t\sqrt{-(\Delta+1)}\delta_y(x) = \frac{\delta'(t-r)}{4\pi\sinh r}$$

PROOF. The Laplacian of $\mathbf{H}^3$ is

$$\Delta = x_3^2 \Delta_0 - x_3 \frac{\partial}{\partial x_3},$$

and in geodesic normal coordinates it is

$$\Delta = \partial_r^2 + 2\coth r \partial_r + \sinh r^{-2} \Delta_{S^2}.$$

Since $\mathbf{H}^3$ is a symmetric space, the fundamental solutions $E^\pm$ is a function only of (t, r) (verify!) so a fundamental solution must solve

$$[\partial_r^2 + 2\coth r \partial_r - 1] E = \delta_0.$$

Here, δ_0 is the delta-function with respect to the volume form, i.e.

$$\langle \delta_0, \psi \rangle = \psi(0) = \int_{\mathbf{H}^3 \times \mathbb{R}} \delta_0(t,r)\psi(r,t) \sinh^2 r dr d\omega dt.$$

Let $y \in \mathbf{H}^3$. The set $\mathcal{C}_y := \{(x,t) : \cosh r(x,y) - \cosh t = 0\} = \{(x,t) : r(x,y) = |t|\}$ is called the *characteristic conoid* based at y and will appear again later on.

The conoid may be parametrized by points x of $\mathbf{H}^3$ with $t = \pm r(x, y)$ giving the upper and lower sheets of the conoid.

In the case of $Q = \cosh t - \cosh r$, (4.33) (with $d\,\mathrm{Vol} = C_3\ (\sinh r)^2 dr d\omega$ replacing dx) gives the stated expression for $\langle \delta(Q), \varphi \rangle$.
We did not cancel the common factors of $\sinh r$ to clarify the use of Lemma 4.33.

We denote the first term by $\delta_+(\varphi)$ and the second by $\delta_-(\varphi)$. The first term above is therefore

$$\delta_+(\varphi) = \langle E^+, \varphi \rangle := C_3[\int_{\mathbf{H}^3} \varphi(x, r) \frac{\sinh^2 r dr dt dd\omega}{\sinh r}.$$

Similarly for the second. The main point is to show that

$$(\Box - 1)E^+ = 2\pi\delta_0. \tag{4.77}$$

Hence E^+ is a fundamental solution supported in the forward conoid. Similarly for E^- in the backward conoid.

We first show that

$$\Box\delta_0(Q) = C\delta(Q), \text{ on } \mathbf{H}^3\backslash\{\mathrm{y}\}. \tag{4.78}$$

As before, we use that

$$\Box f(Q) = f''(Q)\nabla Q \cdot \nabla Q + f'(Q)\Box Q,$$

where the dot product is Lorentzian. Noting that $(\cosh^2 t - \cosh^2 r) = Q(\cosh t + \cosh r)$, we have

$$\begin{cases} (i)\ \nabla Q \cdot \nabla Q & = & (\sinh t dt, -\sinh r dr) \cdot (\sinh t dt, -\sinh r dr) \\ & = & \sinh^2 t - \sinh^2 r = Q(\cosh t + \cosh r), \\ (ii)\ \Box Q & = & [\frac{\partial^2}{\partial t^2} - \partial_r^2 - 2\coth r \partial_r](\cosh t - \cosh r) \\ & = & \cosh t + \cosh r + 2\coth r \sinh r = \cosh t + 3\cosh r. \end{cases}$$

Hence,

$$\Box\delta(Q) = \delta''(Q)Q(\cosh t + \cosh r) + \delta'(Q)(\cosh t + 3\cosh r).$$

Since $t\delta''(t) = -2\delta'(t)$, we have

$$\delta''(Q)Q(\cosh t + \cosh r) = (\cosh t + \cosh r)(-2\delta'(Q) = -2Q\delta'(Q) - 4\cosh \delta'(Q)).$$

Next, write $(\cosh t + 3\cosh r) = Q + 4\cosh r$. Then,

$$\delta'(Q)(\cosh t + 3\cosh r) = Q\delta'(Q) + 4\cosh r\delta'(Q).$$

It follows that

$$\begin{array}{rcl} \Box\delta(Q) & = & -2Q\delta'(Q) - 4\cosh \delta'(Q)) + Q\delta'(Q) + 4\cosh r\delta'(Q) \\ & = & -Q\delta'(Q) = \delta(Q). \end{array}$$

This concludes the proof that $(\Box - 1)\delta(Q) = 0$ on $\mathbf{H}^3 - \{y\}$.

It follows that $(\Box - 1)\delta(Q)$ is a distribution supported at $\{y\}$ and is therefore a linear combinations of derivatives of δ_y, which we write as $\delta_0(t, r, \omega)$ in normal coordinates. If we Taylor expand the coefficients of $\Box$ around 0 in (t, r) it becomes the Euclidean $\Box$ and the homogeneity calculations in that case also imply that

$(\Box - 1)\delta(Q) = c\delta_0$. The value of c can be calculated from a convenient test function, as in the Euclidean case.

$\Box$

REMARK 4.13. In the case of S^3, one analytically continues the equations above, replacing $\cosh r$ by $\cos r$ and so on. The main change is that second derivatives reverse signs of $\cos t, \cos r$. In this case, we get

$$\begin{cases} (i)\ \nabla Q \cdot \nabla Q &= \sin^2 t - \sin^2 r = -(\cos^2 t - \cos^2 r) = -Q(\cos t + \cos r), \\ (ii)\ \Box Q &= [\frac{\partial^2}{\partial t^2} - \partial_r^2 - 2\cot r \partial_r](\cos t - \cos r) \\ &= -\cos t - \cos r - 2\coth r \sinh r = -\cos t - 3\cosh r. \end{cases}$$

Since all signs reverse, we get $(\Box + 1)\delta(Q) = C_n\delta_0$.

4.8.2. Poisson kernel and wave kernel. We obtain the wave kernel on hyperbolic space by analytic continuation of the wave kernel of $\frac{\sin tA}{A}$ on the sphere:

PROPOSITION 4.14. *The Poisson kernel e^{-tA} on hyperbolic space with*

$$A = \sqrt{\Delta - (\frac{n-1}{2})^2}$$

is

$$U(i\tau, x, y) = \sinh\tau\,(\cosh(\tau + i0) - \cosh r)^{-\frac{n+1}{2}}. \tag{4.79}$$

The right side is by definition

$$\lim_{\varepsilon\to 0^+} -2C_n \mathrm{Im}(\cos(it - \varepsilon) - \cosh r)^{-\frac{n-1}{2}}.$$

Taylor [**T1**] proves this formula by analytic continuation of the standard formula for spheres of all radii R or equivalently, for spheres of all curvatures $K > 0$. To be more precise, we just consider the radial parts of the Laplacians of the various metrics. A ball of radius R is the dilate by R of the unit ball and the metrics are related by the dilation. The radial part of the Laplacian $\Delta_{S^n(R)}$ of radius R is obtained by dilation. One then checks that the radial part of the Laplacian for hyperbolic space $\mathbf{H}^n$ is the analytic continuation in R of the radial part for $S^n(R)$ when $K \to -1$. This implies that the fundamental solutions must also be analytic in the parameter K.

4.8.3. Wave equation and spherical means. On hyperbolic space, the spherical means operator is defined by

$$M_r f(x) = \fint_{S_r(x)} f(y)dS(y),$$

where dS is the Riemannian surface measure on the sphere $S_r(x)$ in the hyperbolic metric.

One then has the following formulae for the solution of the modified wave equation (cf. [**Hel2, GN**])

$$\begin{cases} (\Box + (\frac{n-1}{2})^2)u(x,t) = 0, \\ u(x,0) = \varphi(x), \quad u_t(x,0) = \psi(x). \end{cases}$$

Let

$$N^r_{m,k}f(x) = (\frac{\partial}{\partial \cosh r})^m (M_r f(x) \sinh^k(t)).$$

PROPOSITION 4.15. *When $n \geq 3$ is odd,*

$$u(x,t) = C_n \left(\frac{\partial}{\partial t} N^t_{\frac{n-3}{2}, n-2} \varphi(x) + N^t_{\frac{n-3}{2}, n-2} \psi(x) \right),$$

where $C_n = \frac{1}{(n-2)!}$.

When n is even,

$$\frac{1}{2} \int_0^t \frac{u(x,s) + u(x,-s)}{\sqrt{\cosh s - \cosh r}} ds = C_n N^t_{\frac{n-2}{2}, n-2} \varphi(x).$$

4.9. Poisson kernel

Following [**Hel1, Hel2**], we denote by $\langle z, b \rangle$ the signed distance to 0 of the horocycle through the points $z \in \mathbf{D}$ and $b \in B$. Equivalently,

$$e^{\langle z,b \rangle} = \frac{1 - |z|^2}{|z-b|^2} = P_{\mathbf{D}}(z,b),$$

where $P_{\mathbf{D}}(z,b)$ is the Poisson kernel of the unit disc. (We caution that $e^{\langle z,b\rangle}$ is written $e^{2\langle z,b\rangle}$ in [**Hel1, Hel2**]). We denote Lebesgue measure on B by $|db|$, so that the harmonic measure issued from 0 is given by $P_{\mathbf{D}}(z,b)|db|$. A basic identity (cf. [**Hel1**]) is that

$$\langle g \cdot z, g \cdot b \rangle = \langle z, b \rangle + \langle g \cdot 0, g \cdot b \rangle, \tag{4.80}$$

which implies

$$P_{\mathbf{D}}(gz, gb)\, |d(gb)| = P_{\mathbf{D}}(z,b)\, |db|. \tag{4.81}$$

We represent the elements of $\mathcal{E}_s$ with $s = \frac{1}{2} + ir$ by non-Euclidean Poisson integrals analogous to (4.6),

$$\varphi_r(x) = \int_{S^{n-1}} e^{(\frac{1}{2}+ir)\langle z,b\rangle} dT(b) \in \mathcal{E}_\lambda \tag{4.82}$$

where $dT \in \mathcal{D}'(B)$. (Recall that B denotes the boundary of $\mathbf{D}$.)

As before there exists a subspace $\mathcal{E}_s^{(2)}$ on which one can define a Hilbert inner product, namely those for which $dT \in L^2(B)$. One then defines the norm to be the usual $L^2(S^1)$ norm of $dT(b)$. The elements of finite norm define the Hilbert space $\mathcal{E}_\lambda^{(2)}$. The orthogonal projection onto $\mathcal{E}_s^{(2)}$ is given by the non-Euclidean Bessel kernel,

$$E_s(z,w) = \int_B e^{(\frac{1}{2}+ir)\langle z,b\rangle} e^{-(\frac{1}{2}+ir)\langle w,b\rangle} db, \tag{4.83}$$

i.e. by convolution with the spherical function. Then $\mathcal{E}_s^{(2)}$ are irreducible unitary representations of G.

4.10. Spherical functions on $\mathbf{H}^2$

The spherical functions on $\mathbf{D} = \mathbf{H}^2$ are analogues of the spherical harmonics Y_ℓ^m on the 2-sphere. They are joint eigenfunctions of Δ and of K, the rotations fixing the origin. Thus,

$$\begin{cases} \Delta_{\mathbf{H}^2} \Phi_s^m = s(1-s)\Phi_s^m, \\ \dfrac{1}{i}\dfrac{\partial}{\partial\theta}\Phi_s^m = m\Phi_s^m. \end{cases} \tag{4.84}$$

The spherical functions with $s = \frac{1}{2} + ir$ have the Poisson integral representation

$$\Phi_s^m(z) = \int_B e^{(\frac{1}{2}+ir)\langle z,b\rangle} b^m \, db. \tag{4.85}$$

Here $s = \frac{1}{2} + ir$. In particular the analogue of the zonal spherical harmonic is the spherical function,

$$\varphi_s(z) := \Phi_s^0(z) = \int_B e^{(\frac{1}{2}+ir)\langle z,b\rangle} \, db. \tag{4.86}$$

Since φ_s is the analogue of the zonal spherical harmonic, it is plausible that this sequence (in s) of eigenfunctions should be the extremal for the sup norm functional. The analogy to the compact S^2 is not complete since neither φ_s nor any eigenfunction contributing to the spectrum of Δ on $\mathbf{H}^2$ is L^2-normalized. Studying random 'hyperbolic plane waves' or extremals for norms is in the infinite volume setting remains relatively unexplored.

4.11. The non-Euclidean Fourier transform

The hyperbolic plane waves $e_{(r,b)}(z) = e^{(\frac{1}{2}+ir)\langle z,b\rangle}$ are hyperbolic analogues of Euclidean plane waves $e^{i\langle x,\xi\rangle}$ and give rise to a non-Euclidean Fourier transform [**Hel1, Hel2**]. The non-Euclidean Fourier transform is a unitary operator

$$\mathcal{F}\colon L^2(\mathbf{D}, dV) \to L^2(\mathbb{R}_+ \times \partial\mathbf{D}, dp(\lambda) \otimes db) \tag{4.87}$$

defined by

$$\mathcal{F}u(r,b) = \int_{\mathbf{D}} e^{(\frac{1}{2}-ir)\langle z,b\rangle} u(z) \, d\,\mathrm{Vol}(z). \tag{4.88}$$

The hyperbolic Fourier inversion formula is given by

$$u(z) = \int_{\partial\mathbf{D}} \int_{\mathbb{R}} e^{(\frac{1}{2}+ir)\langle z,b\rangle} \mathcal{F}u(r,b) r \tanh(2\pi r) \, dr|db|. \tag{4.89}$$

The measure $r\tanh(2\pi r)dr|db|$ is the Plancherel measure for $G = PSU(1,1)$, which henceforth we denote by $dp(\lambda)$. In general, for $E \in \mathcal{S}'(B \times \mathbb{R}_+)$, we write

$$\mathcal{F}^{-1}E(z) = \int_{\partial\mathbf{D}} \int_{\mathbb{R}} e^{(\frac{1}{2}+ir)\langle z,b\rangle} E(b,r) r \tanh(2\pi r) \, dr|db|. \tag{4.90}$$

4.12. Hyperbolic cylinders

The simplest hyperbolic quotients are defined by cyclic subgroups of G. If the group is generated by a hyperbolic element one obtains a hyperbolic surface of revolution, $\mathbf{H}^2/\langle\gamma\rangle$ where γ is a hyperbolic element and $\langle\gamma\rangle$ is the cyclic group it generates. This surface is uniformized as the region between two half circles orthogonal to the real axis in the upper half plane. One can also choose γ to be a parabolic element and obtain a different hyperbolic surface of revolution which is uniformized as a vertical strip. These surfaces of revolution are not symmetric spaces but they do have quantum integrable Laplacians and by separating variables one obtains model eigenfunctions that correspond more closely to Gaussian beams on S^2 than to eigenfunctions on $\mathbf{H}^2$. Hence, $\mathbf{H}^2/\langle\gamma\rangle$ is another good analogue of S^2.

4.13. Irreducible representations of G

The non-trivial irreducible representations of G come in three families, the principal series, the complementary series and the discrete series. We briefly review the definitions (see [**K, L**] for background).

The principal series $\mathcal{P}^+_{ir}$ are realized on the Hilbert space $L^2(\mathbb{R})$ by the action

$$\mathcal{P}^{\pm}_{ir}\begin{pmatrix} a & b \\ c & d \end{pmatrix} f(x) = |-bx+d|^{-1-2ir} f\left(\frac{ax-c}{-bx+d}\right). \tag{4.91}$$

The unique normalized K-invariant vector of $\mathcal{P}_{ir_j}$ is a constant multiple of

$$f_{ir,0}(x) = (1+x^2)^{-(\frac{1}{2}+ir)}.$$

The complementary series representations are realized on $L^2(\mathbb{R}, B)$ with inner product

$$\langle Bf, f\rangle = \int_{\mathbb{R}\times\mathbb{R}} \frac{f(x)\overline{f(y)}}{|x-y|^{1-2u}}\, dxdy$$

and with action

$$\mathcal{C}_u\begin{pmatrix} a & b \\ c & d \end{pmatrix} f(x) = |-bx+d|^{-1-2u} f\left(\frac{ax-c}{-bx+d}\right). \tag{4.92}$$

When asymptotics as $|r_j| \to \infty$ are involved, we may ignore the complementary series representations and therefore do not discuss them in detail.

Let $\mathbb{C}_+ = \{z \in \mathbb{C} : \mathrm{Im} z > 0\}$. Then $\mathcal{D}^+_m$ is realized on the Hilbert space

$$\mathcal{H}^+_m = \left\{ f \text{ holomorphic on } \mathbb{C}_+ \text{ and } \int_{\mathbb{C}_+} |f(x,y)|^2 y^{m-2} dxdy < \infty \right\} \tag{4.93}$$

with the action

$$\mathcal{D}^+_m\begin{pmatrix} a & b \\ c & d \end{pmatrix} f(z) = (-bz+d)^{-m} f\left(\frac{az-c}{-bz+d}\right). \tag{4.94}$$

The lowest weight vector of $\mathcal{D}^+_m$ in this realization is $(z+i)^{-m}$.

4.14. Compact hyperbolic quotients $\mathbf{X}_\Gamma = \Gamma\backslash\mathbf{H}^2$

Let $\Gamma \subset G$ be a discrete co-compact subgroup, that is, $\Gamma\backslash G$ is compact. It may be identified with the unit tangent or cotangent bundle $S^*\mathbf{X}_\Gamma$ of $\mathbf{X}_\Gamma = \Gamma\backslash\mathbf{H}^2$.

The closed orbits of the geodesic flow G^t on the quotient $\Gamma\backslash G$ are denoted $\{\gamma\}$ and are in one-to-one correspondence with the conjugacy classes of hyperbolic elements of Γ; see for instance [**McK**]. We denote by G_γ, respectively Γ_γ, the centralizer of γ in G, respectively Γ. The group Γ_γ is generated by an element γ_0 which is called a primitive hyperbolic geodesic. The length of γ is denoted $L_\gamma > 0$ and means that γ is conjugate, in G, to

$$a_\gamma = \begin{pmatrix} e^{L_\gamma/2} & 0 \\ 0 & e^{-L_\gamma/2} \end{pmatrix}. \tag{4.95}$$

If $\gamma = \gamma_0^k$ is a power some some primitive γ_0, then we call L_{γ_0} the primitive length of the closed geodesic γ.

4.15. Representation theory of G and spectral theory of Δ on compact quotients

The representation theory of $L^2(\Gamma\backslash G)$ is intimately related to the spectral theory of Δ. We briefly review the results in the case where the quotient is compact (cf. [**K, L**]).

In the compact case, we have the decomposition into irreducibles,

(4.96)

$$L^2(\Gamma\backslash G) = \bigoplus_{j=1}^{S} \mathcal{C}_{ir_j} \oplus \bigoplus_{j=0}^{\infty} \mathcal{P}_{ir_j} \oplus \bigoplus_{m=2,\ m \text{ even}}^{\infty} \mu_\Gamma(m)\mathcal{D}_m^+ \oplus \bigoplus_{m=2,\ m \text{ even}}^{\infty} \mu_\Gamma(m)\mathcal{D}_m^-,$$

where $\mathcal{C}_{ir_j}$ denotes the complementary series representation, respectively $\mathcal{P}_{ir_j}$ denotes the unitary principal series representation, in which Ω equals $s_j(1-s_j) = \frac{1}{4} + r_j^2$. In the complementary series case, $ir_j \in \mathbb{R}$ while in the principal series case $ir_j \in i\mathbb{R}^+$. The irreducibles are indexed by their K-invariant vectors $\{\varphi_{ir_j}\}$, which is assumed to be the given orthonormal basis of Δ-eigenfunctions. Thus, the multiplicity of $\mathcal{P}_{ir_j}$ is the same as the multiplicity of the corresponding eigenvalue of Δ.

Further, $\mathcal{D}_m^\pm$ denotes the holomorphic (respectively anti-holomorphic) discrete series representation with lowest (respectively highest) weight m, and $\mu_\Gamma(m)$ denotes its multiplicity; it depends only on the genus of $\mathbf{X}_\Gamma$. We denote by $\psi_{m,j}$ ($j = 1, \dots, \mu_\Gamma(m)$) a choice of orthonormal basis of the lowest weight vectors of $\mu_\Gamma(m)\mathcal{D}_m^+$ and write $\mu_\Gamma(m)\mathcal{D}_m^+ = \oplus_{j=1}^{\mu_\Gamma(m)} \mathcal{D}_{m,j}^+$ accordingly.

By an automorphic (τ, m)-eigenfunction, we mean a Γ-invariant joint eigenfunction

$$\begin{cases} \Omega\sigma_{\tau,m} = -\Big(\dfrac{1}{4} + \tau^2\Big)\sigma_{\tau,m}, \\ W\sigma_{\tau,m} = im\sigma_{\tau,m}. \end{cases} \tag{4.97}$$

of the Casimir Ω and the generator W of $K = SO(2)$.

We note that the K-weights in all irreducibles are even. Lowest weight vectors of $\mathcal{D}_m^+$ correspond to (holomorphic) automorphic forms of weight m for Γ in the classical sense of holomorphic functions on $\mathbf{H}$ satisfying

$$f(\gamma \cdot z) = (cz+d)^m f(z), \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma. \tag{4.98}$$

A holomorphic form of weight m defines a holomorphic differential of type $f(z)(dz)^{\frac{m}{2}}$. Forms of weight n in $\mathcal{P}_{ir}, \mathcal{C}_u, \mathcal{D}_{ir}^\pm$ always correspond to differentials of type $(dz)^{\frac{n}{2}}$. Forms of odd weights do not occur in $L^2(\Gamma\backslash PSL(2,\mathbb{R}))$.

4.16. Appendix on the Fourier transform

For $f \in L^1(\mathbb{R}^n, dx)$, the Fourier transform is the bounded continuous function defined by

$$\mathcal{F}f(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} f(x)dx. \tag{4.99}$$

We recall that Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is the class of functions such that, for all α, β,

$$\sup_x |x^\alpha D^\beta f(x)| \leq C_{\alpha,\beta} < \infty.$$

Here, $D_j = \frac{1}{i}\frac{\partial}{\partial x_j}$ and $D^\beta = D_1^{\beta_1} \cdots D_n^{\beta_n}$.

LEMMA 4.16. *$\mathcal{F}\colon \mathcal{S} \to \mathcal{S}$ continuously and $\mathcal{F}D_j\varphi = \xi_j\mathcal{F}\varphi$. Also, $\mathcal{F}x_j\varphi = -D_j\mathcal{F}\varphi$.*

The proof is by explicit calculation and integration by parts. One can integrate by parts and pass derivatives under the integral sign because of the rapid decay and smoothness of the integrand.

LEMMA 4.17 ([**Ho**], Lemma 7.1.4). *Suppose that $T\colon \mathcal{S} \to \mathcal{S}$ is a linear map such that for all $\varphi \in \mathcal{S}$,*

$$TD_j\varphi = D_jT\varphi, \quad Tx_j\varphi = x_jT\varphi, \quad j = 1, \ldots, n, \tag{4.100}$$

then $T\varphi = c\varphi$ for some constant c.

PROOF. If $\varphi \in \mathcal{S}$ and $\varphi(y) = 0$ then (by Taylor's formula) there exist $\varphi_j \in \mathcal{S}$ such that

$$\varphi(x) = \sum_j (x_j - y_j)\varphi_j(x). \tag{4.101}$$

Then

$$T\varphi(x) = \sum_j (x_j - y_j)T\varphi_j(x) = 0 \quad \text{if} \quad x = y. \tag{4.102}$$

It follows that

$$T\varphi(x) = c(x)\varphi(x), \quad c(x) = (T1)(x). \tag{4.103}$$

Indeed, let $\psi_x(y) = e^{-(x-y)^2/2}$. Then $(\varphi(y) - \varphi(x)\psi_x(y)$ vanishes when $y = x$. It follows that $T(\varphi(y) - \varphi(x)\psi_x(y)) = 0$ when $y = x$ or

$$(T\varphi)(x) = \varphi(x)T(\psi_x)(x)$$

. Thus T is multiplication by $T\psi_x(x)$.

Applying to $\varphi(x) = e^{-x^2}$ for instance shows that $c(x) \in C^\infty$. Since $D_jc(x)\varphi = c(x)D_j\varphi$ it follows that $c(x) \equiv c$ is constant. □

THEOREM 4.18. *$\mathcal{F}$ is an isomorphism of $\mathcal{S}$ with inverse given by*

$$\mathcal{F}^{-1}f(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} e^{i\langle x,\xi\rangle}\hat{f}(\xi)d\xi. \tag{4.104}$$

PROOF. We know that $\mathcal{F}^2\colon \mathcal{S} \to \mathcal{S}$ and that it anti-commutes with D_j and x_j. Let $R\varphi(x) = \varphi(-x)$. Then $T = RF^2$ commutes with D_j, x_j hence is a constant multiple of the identity. Thus, $RF^2 = c$. Thus, $F^{-1} = c^{-1}RF$. To determine c one applies $\mathcal{F}$ to $e^{-|x|^2/2}$ and finds that $c = (2\pi)^n$. Indeed, $(x_j + iD_j)\varphi = 0$ and taking the Fourier transform gives $(-D_j + i\xi_j)\hat{\varphi}(\xi) = 0$. Hence $\hat{\varphi} = c_1\varphi$ with $c_1 = \hat{\varphi}(0) = (2\pi)^{n/2}$. Hence $F^2\varphi = c_1^2\varphi$. □

THEOREM 4.19. *For $\varphi, \psi \in \mathcal{S}$, we have*

- $\int \hat{\varphi}\psi dx = \int \psi\hat{\varphi}dx$;
- $\int \hat{\varphi}\overline{\psi}dx = \int \hat{\varphi}\hat{\overline{\psi}}d\xi$.

PROOF. Both sides of the first statement equal

$$\iint \varphi(x)\psi(\xi)e^{-i\langle x,\xi\rangle}\,dxd\xi. \tag{4.105}$$

To prove the second statement, put $\chi = (2\pi)^{-n}\overline{\hat{\psi}}$ and use Fourier inversion to get

$$\overline{\hat{\chi}(\xi)} = (2\pi)^{-n}\int \hat{\psi}(x)e^{i\langle x,\xi\rangle}dx = \psi(\xi). \tag{4.106}$$

Hence the second statement follows from the first if we replace ψ by χ. □

THEOREM 4.20. *$\mathcal{F}$ extends to a unitary isomorphism of $L^2(\mathbb{R}^n)$.*

Bibliography

[Ag] S. Agmon, Representation theorems for solutions of the Helmholtz equation on Rn. Differential operators and spectral theory, 2743, Amer. Math. Soc. Transl. Ser. 2, 189, Amer. Math. Soc., Providence, RI, 1999.

[AM] N. Anantharaman and F. Maci, Semiclassical measures for the Schrdinger equation on the torus. J. Eur. Math. Soc. (JEMS) 16 (2014), no. 6, 1253-1288.

[AZ] N. Anantharaman and S. Zelditch, Patterson-Sullivan distributions and quantum ergodicity. Ann. Henri Poincar 8 (2007), no. 2, 361-426.

[DM] H. Dym and H.P. McKean, *Fourier series and integrals.* Probability and Mathematical Statistics, No. 14. Academic Press, New York-London, 1972.

[DSP] W. Duke and R. Schulze-Pillot, Representation of integers by positive ternary quadratic forms and equidistribution of lattice points on ellipsoids. Invent. Math. 99 (1990), no. 1, 49-57.

[Eg] Eguchi, M. *Asymptotic expansions of Eisenstein integrals and Fourier transform on symmetric spaces.* J. Funct. Anal. 34 (1979), no. 2, 167–216.

[EH] P. Erdös and R. R. P. Hall, On the angular distribution of Gaussian integers with fixed norm. Paul Erdös memorial collection. Discrete Math. 200 (1999), no. 1-3, 87-94.

[F] G. Friedlander, *Introduction to the theory of distributions.* Cambridge University Press, Cambridge, 1998.

[GeSh] I.M. Gel'fand and G.E. Shilov, *Generalized functions. Vol. I: Properties and operations.* Academic Press, New York-London 1964.

[GN] A. Grigor'yan and M. Noguchi, The heat kernel on hyperbolic space, The heat kernel on hyperbolic space. Bull. London Math. Soc. 30 (1998), no. 6, 643-650.

[Hel1] S. Helgason, *Topics in harmonic analysis on homogeneous spaces. Progress in Mathematics, 13. Birkhuser, Boston, Mass., 1981.*

[Hel2] S. Helgason, *Groups and geometric analysis. Integral geometry, invariant differential operators, and spherical functions.* Corrected reprint of the 1984 original. Mathematical Surveys and Monographs, 83. American Mathematical Society, Providence, RI, 2000.

[HR] H. Hezari and G. Riviere, Quantitative equidistribution properties of toral eigenfunctions (arXiv:1503.02794).

[Ho] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[J] D. Jakobson, Quantum limits on flat tori. Ann. of Math. (2) 145 (1997), no. 2, 235-266.

[K] A. W. Knapp,*Representation theory of semisimple groups. An overview based on examples.* Reprint of the 1986 original. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 2001.

[L] S. Lang, $\mathrm{SL}_2(R)$. Graduate Texts in Mathematics, 105. Springer-Verlag, New York, 1985.

[McK] H. P. McKean, Selberg's trace formula as applied to a compact Riemann surface. Comm. Pure Appl. Math. 25 (1972), 225-246.

[So] Sogge, Christopher D. Oscillatory integrals and spherical harmonics. Duke Math. J. 53 (1986), no. 1, 43–65.

[StW] E. M. Stein and G. ; Weiss,*Introduction to Fourier analysis on Euclidean spaces.* Princeton Mathematical Series, No. 32. Princeton University Press, Princeton, N.J., 1971.

[Str] R. S. Strichartz, . Harmonic analysis as spectral theory of Laplacians. J. Funct. Anal. 87 (1989), no. 1, 51-148.

[T1] M. E. Taylor, *Partial Differential Equations, Vol 2*, Springer-Verlag, New York, 1996. MR1395147

[W] J.A. Wolf, *Spaces of constant curvature.* Fifth edition. Publish or Perish, Inc., Houston, Tex., 1984.

CHAPTER 5

Local structure of eigenfunctions

The main purpose of this monograph is to study the asymptotic properties of global eigenfunctions on a compact Riemannian manifold (M, g). However there exists a large literature on local eigenfunctions, and local eigenfunctions may have quite different properties from global ones. The first aim of this section is to explain the 'local-global' distinction and to study the local structure of eigenfunctions. The second aim of this section is to explain the sense in which the local techniques are elliptic while the global techniques are hyperbolic.

5.1. Local versus global eigenfunctions

Some estimates on eigenfunctions are *local* in the sense that they hold for a local solution of $(\Delta + \lambda^2)\varphi_\lambda = 0$ in an open set $U \subset M$, regardless of whether the local eigenfunction extends to a global eigenfunction on M. Others are *global* in that they use the global extension of φ_λ, usually in the form that it is an eigenfunction of the wave group $U(t) = e^{it\sqrt{-\Delta}}$. For instance, the Poisson integral formulae (4.6) on Euclidean space or (4.82) on hyperbolic space are global in the sense that they only hold for global eigenfunctions. Our main interest is in global estimates and their relation to the geodesic flow. In particular, we often pose the problem of determining when the global estimates are achieved by a sequence of eigenfunctions of (M, g).

The distinctions are illustrated by the eigenfunctions

$$\begin{cases} \Phi_{\mu,\nu}(x,y) = \cos\mu x \cos\nu y, & (\lambda^2 = \mu^2 + \nu^2), \\ \Psi_{\mu,\nu}(x,y) = \cosh\mu x \cos\nu y, & (\lambda^2 = \nu^2 - \mu^2), \\ \Psi'_{\mu,\nu}(x,y) = \cosh\mu x \cosh\nu y, & (\lambda^2 = -(\nu^2 + \mu^2)) \end{cases}$$

on $\mathbb{R}^2$. Note that the eigenvalue of the negative operator Δ is positive in the last case (although we still wrote λ^2) and is not in the spectrum of Δ.

Only the oscillatory function $\Phi_{\mu,\nu}$ can be global eigenfunctions on a compact Riemannian manifold. The eigenfunctions $\Psi_{\mu,\nu}$ are not tempered and do not contribute to the L^2 spectrum of Δ on $\mathbb{R}^2$. Depending on the pair (μ, ν), the eigenvalues can have any sign or be of any size, but only $\Phi_{\mu,\nu}$ could be an eigenfunctions on a compact manifold (possibly with boundary). Clearly it is impossible to control the size of the nodal set or the number of critical points of $\Psi_{\mu,\nu}$ by λ, which are controlled by the quantity $\sqrt{\nu^2 + \mu^2}$ and not by the frequency $|\sqrt{\nu^2 - \mu^2}|$. Note that $\sqrt{\nu^2 + \mu^2}$ is of the order of magnitude of the local Dirichlet quotient

$$D_B(\varphi) := \frac{\int_B |\nabla\varphi|^2 \, dV}{\int_B \varphi^2 \, dV}. \tag{5.1}$$

To distinguish these eigenfunctions, we introduce some definitions:

Definition 5.1. Let (M, g) be a connected Riemannian manifold. Define the bottom of the spectrum by

$$\lambda_0^2(M,g) = \inf_{\varphi \in C_0^\infty(M)} \frac{\int_M |\nabla \varphi|^2 \, dV}{\int_M \varphi^2 \, dV}. \tag{5.2}$$

In the case of $\mathbb{R}^n$, $\lambda_0 = 0$. Clearly, $\Psi'_{\mu,\nu}$ is an eigenfunction from the positive spectrum while $\Phi_{\mu,\nu}$ is from the L^2 spectrum. The hybrid $\Psi_{\mu,\nu}$ is from neither spectrum and it is possible that $\nu^2 - \mu^2 = \lambda^2 \gg 0$. Thus, $\Psi_{\mu,\nu}$ is the type of local eigenfunction which does not occur globally.

One of the basis distinctions between $\Phi_{\mu,\nu}$ and $\Psi_{\mu,\nu}$ is their status relative to the maximum principle. If $(\Delta + \lambda^2)u = 0$ in an open domain Ω of a Riemannian manifold (M, g) one says that a non-constant function u satisfies the strong maximum principle if there does not exist $p \in \Omega$ so that $u(p) = \sup_\Omega u$. Equivalently, if $\sup_\Omega u = \sup_{\partial\Omega} u$. The maximum principle holds for the positive spectrum but not globally for the non-constant eigenfunctions of the L^2 spectrum. For instance $\Phi_{\mu,\nu}$ has interior maxima in small balls surrounding its local maxima. If Ω is a nodal domain for φ_λ then we may assume $\varphi_\lambda > 0$ in Ω with $\varphi_\lambda = 0$ on $\partial\Omega$, and obviously φ_λ does not satisfy the maximum principle in Ω or in a ball $B \subset \Omega$ around its maximum point in Ω.

5.2. Small balls and local dilation

As discussed in more detail in the next section §5.9.1, eigenfunctions resemble harmonic functions on balls of radius $r = \varepsilon\lambda^{-1}$ for ε sufficiently small. Since this length scale comes up repeatedly, we pause to give it a name:

Definition 5.2. A "small ball" is a ball of radius $\frac{C}{\lambda}$ where C is a fixed constant.

When $C = \varepsilon$ is sufficiently small, the maximum principle and some mean value inequalities may be apply on the ball $B_{\frac{C}{\lambda}}(p)$.

A useful way to study the local behavior of eigenfunctions around a zero x_0 is to pull back the eigenvalue equation to the tangent space $T_{x_0}M$ and perform a dilation $\varphi_\lambda(u) \to \varphi_\lambda(tu)$ in the tangent space $T_{x_0}M$. When the eigenvalue is λ^2 it is natural to dilate the small ball by $t = \varepsilon^{-1}\lambda$. To simplify notation, we do not distinguish functions on $T_{x_0}M$ with functions expressed in normal coordinates on a small ball $B_r(x_0)$ around x_0 of radius $r = \varepsilon\lambda^{-1}$. In particular, we continue to denote denote the pullback $\exp_{x_0}^* \varphi_\lambda$ by φ_λ. With these notational conventions, we define the dilation operators by

$$D_t^{x_0}\varphi_\lambda(u) = \varphi_\lambda(\exp_{x_0} tu), \qquad u \in T_{x_0}M. \tag{5.3}$$

It is proved in [**Ber, HarW2**] that in a small ball around a zero, an eigenfunction is asymptotic to a homogeneous harmonic polynomial. In a general setting, let

$$L = \sum_{i,j=1}^n g^{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^n \Gamma_j(x) \frac{\partial}{\partial x_j} + V(x) \tag{5.4}$$

be the Laplacian plus a potential. Fix a point p and introduce geodesic normal coordinates x at p. Define the *osculating operator* at p to be the constant coefficient operator

$$L_p f(x) = \sum_{i,j=1}^{n} g^{ij}(0) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{j=1}^{n} \Gamma_j(0) \frac{\partial}{\partial x_j} + V(0). \tag{5.5}$$

The following is Theorem 1 of [**Ber**]:

PROPOSITION 5.3. *Assume that φ_λ vanishes to order k at x_0. Let $\varphi_\lambda(x) = \varphi_k^{x_0}(x) + \varphi_{k+1}^{x_0} + \cdots$ denote the C^∞ Taylor expansion of φ_λ into homogeneous terms in normal coordinates x centered at x_0. Then $\varphi_k^{x_0}(x)$ is a Euclidean harmonic homogeneous polynomial of degree k.*

To prove this, one Taylor expands the equation $\Delta \varphi_\lambda = -\lambda^2 \varphi_\lambda$ and finds the first non-vanishing homogeneous term. The osculating equation has the form $L_0[\varphi]_N = 0$, where $\varphi = [\varphi]_N + O(|x|^{N+1})$. Indeed, conjugation by the dilation operator $D_t^{x_0}$ produces the rescaled Laplacian:

$$\Delta_t^{x_0} := D_t^{x_0} \Delta_g (D_t^{x_0})^{-1} = \sum_{j=1}^{n} \frac{\partial^2}{\partial u_j^2} + \cdots . \tag{5.6}$$

Here, $\cdots$ refer to the lower order terms in the Taylor expansion of the coefficients of Δ around x_0. Collecting terms with like powers of t is equivalent to collecting terms with a given homogeneity order, with the understanding that in the Laplacian each differential expression $\frac{\partial}{\partial x_j}$ is homogeneous of degree -1. We denote the rescaling or dilation $D_t^{x_0} f$ by $f_{[t]}$.

More explicitly, we express Δ in local normal coordinates as

$$\Delta = \sum_{i,j} g^{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_k \Gamma^k \frac{\partial}{\partial x_k}. \tag{5.7}$$

Rescaling gives

$$\Delta_t^{x_0} = t^{-2} \sum_{i,j} g_{[t]}^{ij} \frac{\partial^2}{\partial x_i \partial x_j} + t^{-1} \sum_k \Gamma_{[t]}^k \frac{\partial}{\partial x_k}. \tag{5.8}$$

As $t \to 0$ we have

$$\begin{cases} g_{[t]}^{ij}(u) = \delta^{ij} + t^2 R_{k\ell}^{ij} u_j u_\ell + \cdots , \\ \Gamma_{[t]}^k = t \Gamma_{;j}^k u_j + \cdots . \end{cases} \tag{5.9}$$

Expanding the coefficients in powers of t, dilating each $\frac{\partial}{\partial x_j}$ and collecting like powers of t defines the expansion

$$\Delta_t^{x_0} = t^{-2} \Delta^{(2)} + t^{-1} \Delta^{(1)} + \cdots + t^k \Delta^{(k)} + \cdots , \tag{5.10}$$

with the osculating Laplacian $\Delta^{(2)}$ given by the first term of (5.6).

We then rewrite the eigenvalue equation as

(5.11)

$$D_t^{x_0} \Delta_g (D_t^{x_0})^{-1} \varphi(\exp_{x_0} tu) = \lambda^2 \varphi(\exp_{x_0} tu)$$

$$\implies \left[t^{-2} \Delta^{(2)} + t^{-1} \Delta^{(1)} + \cdots \right] \varphi(\exp_{x_0} tu) = \lambda^2 \varphi(\exp_{x_0} tu). \tag{5.12}$$

The next step is to Taylor expand $\varphi(\exp_{x_0} tu)$ around $t = 0$, i.e., expand

$$\varphi(\exp_{x_0} tu) = p_N(u) + \cdots .$$

in terms of homogeneous polynomials on $T_{x_0}M$. The leading order equation in t as $t \to 0$ becomes

$$t^{-2} t^N \Delta^{(2)} p_N + \cdots = \lambda^2 t^N p_N. \tag{5.13}$$

There is no term of order t^{N-2} on the right side and thus the lowest order homogeneous term is $\Delta^{(2)} p_N = 0$ of degree $N-2$.

One can continue to derive transport type equations for the higher coefficients p_{N+k}. To the author's knowledge, they have never been studied. A problem is that the scaling apparently destroys the global properties of eigenfunctions, in particular that they extend globally to M. The transport equations are not uniquely solvable without some global condition.

There are many applications of this local structure to nodal sets, which will be discussed further in the relevant chapters. For the time being, we only note that in dimension 2, a homogeneous harmonic polynomial of degree N is the real or imaginary part of the unique holomorphic homogeneous polynomial z^N of this degree, i.e., $p_N(r,\theta) = r^N \sin N\theta$. As observed in [**Ch**], there exists a C^1 local diffeomorphism χ in a disc around a zero x_0 so that $\chi(x_0) = 0$ and so that $\varphi_N^{x_0} \circ \chi = p_N$. It follows that the restriction of φ_λ to a curve H is C^1 equivalent around a zero to p_N restricted to $\chi(H)$. The nodal set of p_N around 0 consists of N rays, $\{r(\cos\theta, \sin\theta) \colon r > 0, p_N|_{S^1}(v) = 0\}$. It follows that the local structure of the nodal set in a small disc around a singular point p is C^1 equivalent to N equi-angular rays emanating from p. We refer to [**Ch**] for further details.

5.2.1. Harmonifying eigenfunctions. In the preceding section, it is explained that an eigenfunction is well approximated by a harmonic polynomial in a small ball around a zero. A second relation is that eigenfunctions can be converted into harmonic functions on a space of one higher dimension in various (essentially equivalent) ways. Let $C = M \times \mathbb{R}_+$ be the cone or upper half space over M with metric $dr^2 + r^2 g_M$. Let r denote the distance to the vertex. Then

$$(\Delta + \lambda^2) u = 0 \implies r^d u(x) \text{ is harmonic on } C \text{ (with } \lambda^2 = d(d+n-2)) \tag{5.14}$$

in the sense that

$$\Delta_{C(M)} u = \frac{\partial^2}{\partial r^2} u + \frac{n-1}{r} \frac{\partial}{\partial r} u + \frac{1}{r^2} \Delta_M u = 0. \tag{5.15}$$

Another approach is that

$$(\Delta + \lambda^2) u = 0 \implies e^{-t\lambda} u(x) \text{ is harmonic on } M \times \mathbb{R}_+ \tag{5.16}$$

in the sense that $(\frac{\partial^2}{\partial t^2} + \Delta_M) e^{-t\lambda} u(x) = 0$. For this reason, $e^{-t\sqrt{-\Delta}}$ is called the Poisson semi-group.

5.3. Local elliptic estimates of eigenfunctions

In this section, we state some elliptic estimates on eigenfunctions on Riemannian manifolds. They mainly take the form of mean value inequalities. In the local-global dichotomy of the previous section, they generally belong in the local class. In the next section we move on to semiclassical estimates which treat the Helmholtz operator as hyperbolic. In subsequent sections we develop the theory of the wave equation on compact Riemannian manifolds, which is the main tool in the study of semiclassical asymptotics and estimates.

Classic texts such as [**GT**] on elliptic estimates generally start with elliptic estimates for harmonic functions and extend the estimates to more general elliptic operators of the form $L = \sum_{i,j} \frac{\partial}{\partial x_i} a_{ij} \frac{\partial}{\partial x_j} + c(x)$. In proving maximum principles and mean value inequalities for subsolutions, it is often assumed that $c(x) \leq 0$. This of course excludes the Helmholtz operator and indeed the maximum principle does not apply without some restriction. For instance, it is obviously not true that a Dirichlet eigenfunction on a domain Ω takes its maximum on the boundary. One of the earliest articles to consider mean value inequalities when $c > 0$ is that of [**BNV**] (see also [**Ca, P**]).

We follow the traditional path in beginning with elliptic identities and estimates for harmonic and subharmonic functions on $\mathbb{R}^n$. We then consider eigenfunctions and pay particular attention to estimates which hold for local or global solutions of the Helmholtz equation for large λ.

One of the principal techniques for obtaining mean value inequalities is to use reproducing kernels. For instance, harmonic functions are reproduced by the globally defined spherical means or ball means operators. They are also locally reproduced by the Dirichlet Green's functions of balls, i.e., by a Poisson integral formula, which are only defined on the given balls in general (see Chapter 4 of [**GT**] or Section 1.3 of [**HanL1**]). There also exist local reproducing formula using hyperbolic methods, e.g. by using operators of the form $\rho(\sqrt{\lambda} - \sqrt{\Delta})$ (5.100). One may also use the wave operators $\cos t\sqrt{-\Delta}$ or $\frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}$. At first it seems that these kernels belong to the global class, but they are in a sense local in that for small t they only sample values of the eigenfunction in small balls by the finite propagation speed of the wave equation. Paraphrasing M. Kac, one might call this the principle of not feeling the global manifold for small times.

5.3.1. Mean value inequalities for L-harmonic and subharmonic functions on $\mathbb{R}^n$. In this section we review some classical mean value inequalities for harmonic and subharmonic functions. They only apply to eigenfunctions if one first converts them to harmonic functions as described in §5.2.1.

Let $L = \operatorname{div} A\nabla$ be an elliptic operator, i.e., let

$$Lu = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j} a_{ij} \frac{\partial u}{\partial x_i}$$

be an elliptic operator in divergence form on a ball $B_1 \subset\subset \mathbb{R}^n$. The following inequality for L-harmonic functions is known as the Moser local boundedness theorem (cf. [**HanL1**]).

THEOREM 5.4. *Let a_{ij} be bounded in B_1 and let $Lu = 0$. Then*

$$\sup_{B_{\frac{1}{2}}} u^2 \leq C \fint_{B_1} u^2,$$

where C depends only on L.

Regarding subharmonic functions, one has

THEOREM 5.5. *Let $L = \sum_{i,j=1}^{n} \frac{\partial}{\partial x_j} a_{ij} \frac{\partial u}{\partial x_i}$ be an elliptic operator. Suppose that $m|v|^2 \leq a_{ij} v_i v_j \leq M|v|^2$, and let $k = 2 + \frac{8M^2}{m^2}$. Let $x_0 \in M$. Suppose that $u \geq 0$*

and $Lu \geq 0$. *Then*

$$\sup_{B_1(x_0)} u \leq k \left(\int_{B_2(x_0)} u^2 \right)^{\frac{1}{2}}. \tag{5.17}$$

5.3.2. Mean value inequalities for subharmonic functions on Riemannian manifolds. We now turn to mean value inequalities on Riemannian manifolds. The following is proved in Section 6 of [**SY**]. It compares the value of a subharmonic function at the center of a ball with its mean value in the ball.

THEOREM 5.6. *Let (M, g) be a complete Riemannian manifold, and let $B_R(p)$ be a geodesic ball. Suppose that the sectional curvature $K_M \leq k$ and that $R < \operatorname{inj}(M, g)$. Then for any $u \in C^\infty(M)$ satisfying $\Delta u \geq 0$ and $u \geq 0$ on M,*

$$u(p) \leq \frac{1}{V_k(R)} \int_{B_R(p)} u \, dV. \tag{5.18}$$

Here, $V_k(R) = \operatorname{Vol}(B_R, g_k)$ is the volume of a ball of radius R in the space form of constant curvature k.

Theorem 6.2 of [**SY**] and Theorem 1.2 of [**LiSch**] generalize Theorem 5.4 (Moser's local boundedness estimate) to a mean value inequality for manifolds with a lower Ricci curvature bound:

THEOREM 5.7. *Let (M, g) be a complete Riemannian manifold with $\operatorname{Ric}_M \geq -K$. Let $p \in M$, and $u \geq 0$ and $\Delta u \geq 0$ on $B_R(p)$. Then for any $\tau \in (0, \frac{1}{2})$,*

$$\sup_{B_{(1-\tau)R}(p)} u^2 \leq C_1 \tau^{-C_2(1+\sqrt{K}R)} \fint_{B_R(p)} u^2 \, dV. \tag{5.19}$$

Here, $\fint_K f$ denotes the average value of f on K.

On a general Riemannian manifold one has the following mean value inequality (see Proposition 4.7 of [**CM2**]):

THEOREM 5.8. *Let $\dim M = n$ and suppose that $\operatorname{Ric}_M \geq -(n-1)s^{-2}$. Let $u \geq 0$ and $\Delta_M u \geq -s^{-2} u$. Then*

$$u^2(x) \leq C \fint_{B_s(x)} u^2. \tag{5.20}$$

PROOF. After scaling the metric it suffices to prove the case $s = 1$. Let $N = M \times [-1, 1]$ have the product metric, so that $\operatorname{Ric}_N \geq -(n-1)$ and as in (5.14) let $w(x, t) = u(x) e^t$. Then $\Delta_N = e^t \Delta_M u + e^t u \geq 0$. Hence w is subharmonic on $M \times [-1, 1]$. The mean value theorem for subharmonic functions (Theorem 5.7) gives

$$w^2(x, 0) \leq \frac{C}{\operatorname{Vol}(B_1(x, 0) \subset N)} \int_{B_1(x,0) \subset N} w^2 \tag{5.21}$$

$$\leq 2e^2 \frac{C}{\operatorname{Vol}(B_{\frac{1}{2}}(x) \subset M)} \int_{B_1(x) \subset M} u^2. \tag{5.22}$$

The Proposition follows from the Bishop-Gromov volume comparison theorem bounding the ratio $\operatorname{Vol}(B_1(x)) / \operatorname{Vol}(B_{\frac{1}{2}}(x))$. □

5.3.3. Mean value inequalities for λ-subharmonic functions. The previous results hold for Δ- subharmonic functions. We now consider $(\Delta+\lambda^2)$-harmonic or subharmonic functions, i.e. eigenfunctions or functions satisfying $(\Delta+\lambda^2)u \geq 0$. As mentioned in the introduction to this section, the positivity of the coefficient $c(x)=\lambda^2>0$ changes the theory profoundly unless the balls have size $\varepsilon\lambda^{-1}$.

A good example to keep in mind is an L^2-normalized zonal spherical harmonic $\varphi_\ell = C_\ell \mathrm{Re}(Y_\ell^\ell)$ in a ball of radius $\varepsilon\ell^{-1}$ around the north pole P. The supremum of $|\varphi_\ell| \simeq \sqrt{\ell}$ is taken at P and although it oscillates on the length scale of ℓ^{-1}. The profile decreases towards the equator to size 1. One may choose ε so that $B_\varepsilon(P)$ is the nodal domain which contains P. In this case $|\varphi_\ell|$ oscillates from $\sqrt{\ell}$ to 0 on an interval of size $\frac{1}{\ell}$ and the mean of $|\varphi_\ell|$ is of order $\ell^{\frac{1}{2}}$. But if one chooses a much larger ε so that $B_\varepsilon(P)$ is the northern hemisphere, then the mean of $|\varphi_\ell|^2$ decreases to 1. A mean value inequality must therefore contain a compensating constant of order ℓ. This extreme case shows that we may expect a mean value inequality as in the previous section to be valid on sufficiently small balls of order $\varepsilon\lambda^{-\frac{1}{2}}$ but not on balls of a larger order of magnitude.

COROLLARY 5.9. *Let (M,g) be a complete Riemannian manifold with* Ric $\geq -(n-1)$. *Let $y \in M$ and let $t \leq 1$. Suppose that $v \geq 0$ on $B_{2t}(y)$ and that $\Delta v \geq -\lambda^2 v$ for some $\lambda^2 \geq 0$. Then*

$$\sup_{B_t} v \leq e^{C_n(1+\lambda t)} \fint_{B_{2t}(y)} v. \tag{5.23}$$

PROOF. We rescale the metri c $g \to \lambda^{-2}g$ so that $\Delta \to \lambda^2\Delta$. Then $\mathrm{Ric}_{\lambda^{-1}g} = \lambda^{-1}\mathrm{Ric}_g \geq -\lambda^{-1}(n-1) \geq -(n-1)$. The ball of radius λt around p in the metric $\lambda^{-2}g$ is the same as the ball of radius t around p in the metric g. Applying the mean value inequality of Theorem 5.7 for the metric $\lambda^{-2}g$ with $\tau=e^{-1}$ gives

$$\sup_{B_{(1-\tau)\lambda t}(p)} u^2 \leq e^{C(1+t\lambda^{-1}\lambda)} \fint_{B_{\lambda t}} u^2 \, dV_{\lambda^{-1}g} = e^{C(1+t\lambda)} \fint_{B_t} u^2 \, dV. \tag{5.24}$$

□

REMARK 5.10. Above we are simply rescaling the metric. It is also possible to use a local dilation (rescaling) map on $B_R(p)$ which fixes p and dilates normal coordinates $x \to \lambda x$ centered at p. This dilation map distorts u, whereas rescaling the metric leaves u unchanged.

A related corollary in [**CM3**] is the following:

COROLLARY 5.11. *Let (M,g) be a complete Riemannian manifold and suppose that $(\Delta+\lambda^2)u=0$. Suppose that $u(p)=0$. Then*

$$\sup_{B_{\frac{r}{2}}} |u| \leq C r^{-n} \int_{B_r(p)} |u|.$$

Moreover, if $q \in B_{\frac{r}{3}}(p)$ with $u(q)=0$ and if $r \simeq \lambda^{-\frac{1}{2}}$, then

$$\sup_{B_{\frac{4r}{3}}(q)} u^2 \leq C_0 r^{-n} \int_{B_{2r}(p)} u^2.$$

5.3.4. Bochner identity and mean value inequalities. Another source of subharmonic functions and mean-value inequalities comes from the Bochner identity

$$\Delta|\nabla\varphi_\lambda|^2 = 2|\operatorname{Hess}(\varphi_\lambda)|^2 - 2\lambda^2|\nabla\varphi_\lambda|^2 + 2\operatorname{Ric}(\nabla\varphi_\lambda, \nabla\varphi_\lambda).. \tag{5.25}$$

If the Ricci curvature satisfies a suitable lower bound as in (Proposition 2.2 of [**SY**]), one has

$$\Delta|\nabla\varphi_\lambda|^2 \geq -2(n-1-\lambda^2)|\nabla\varphi_\lambda|^2. \tag{5.26}$$

Hence there exists $C_n > 0$ so that (cf. [**SY**], p. 80)

$$\sup_{B(z_k, \frac{s}{\lambda})} |\varphi_\lambda|^2 \leq s^2 \sup_{B(z_k, \frac{s}{\lambda})} |\nabla\varphi_\lambda|^2 \leq Cs^2 \frac{1}{\operatorname{Vol} B(z_k, \frac{s}{\lambda})} \int_{B(z_k, \frac{s}{\lambda})} |\nabla\varphi_\lambda|^2. \tag{5.27}$$

5.4. λ-Poisson operators

In the standard Euclidean setting, classical results on mean value identities and inequalities are proved using either the spherical means operator, or Poisson integral formula. Unlike the spherical means operators, which are global, the Poisson integral operators are adapted to a domain. They may be used to prove mean value inequalities for λ-subharmonic functions on balls of size $\varepsilon\lambda^{-1}$. There also exist global Poisson integral formulae on Cartan-Hadamard manifolds, i.e. simply connected manifolds of non-positive curvature, in which the ideal (geodesic) boundary at infinity is used.

There are several types of Poisson operator for the eigenvalue problem. By a Poisson operator we mean an extension operator of data along a separating hypersurface H to eigenfunctions in the kernel of $(\Delta + \lambda^2)$ on $M \setminus H = M_+ \cup M_-$.

The first type of Poisson operator is the Poisson operator on two-component Cauchy data. It is always well defined and may be constructed in terms of any Green's function G_λ, i.e., any solution on M_+ of $(\Delta + \lambda)G_\lambda(x, y) = \delta_y(x)$. By Green's formula

$$\int_{M_+} G_\lambda(\Delta + \lambda^2)\varphi_\lambda - \varphi_\lambda(\Delta + \lambda^2)G_\lambda = \int_H (\partial_\nu\varphi_\lambda)G_\lambda - \varphi_\lambda\partial_\nu G_\lambda, \tag{5.28}$$

one gets (for $p \in M_\pm$)

(5.29)

$$\varphi_\lambda(x) = \int_H (\partial_\nu\varphi_\lambda)(q)G_\lambda(x, q)\, dS(q) - \varphi_\lambda(q)\partial_{\nu_q}G_\lambda(x, q)\, dS(q) := \mathcal{P}_\pm(\lambda)\begin{pmatrix}\varphi_\lambda|_H \\ \partial_\nu\varphi_\lambda|_H\end{pmatrix}$$

There is no jump of the right side across H since we assume that φ_λ is a global eigenfunction.

The second type of Poisson operator is with respect to Dirichlet (resp. Neumann) data alone and does not always exist. We consider only Dirichlet data (the discussion for Neumann data is similar). The Poisson operator $P_\pm(\lambda)$ for Dirichlet data on H of the eigenvalue problem $(\Delta + \lambda^2)\varphi = 0$ with respect to $M_\pm$, is the extension operator of Dirichlet data along H as an eigenfunction. That is $P_\pm(\lambda)f$ is the unique eigenfunction on $M_\pm$ whose restriction to H equals f. $P_\pm(\lambda)$ is well-defined unless $-\lambda^2$ is an eigenvalue for the Dirichlet problem on $M_\pm$ for Δ. When this is not the case the Dirichlet Green's function $G_D(\lambda)$ has no pole at λ and we may choose it for the Green's contention G_λ above to get

$$\varphi_\lambda(x) = -\int_H \varphi_\lambda(q)\partial_{\nu_q} G_D(\lambda, x, q)\, dS(q). \tag{5.30}$$

We could exchange the roles of Dirichlet and Neumann data to obtain a formula in terms of Neumann data when $-\lambda^2$ is not an eigenvalue of the Neumann problem.

Bourgain-Rudnick have raised the question whether a sequence of global eigenfunctions exists could vanish on a (separating) curve (or hypersurface) H. By the above, it is equivalent to asking if when the Dirichlet eigenvalue problem on a component $M - H$ can have infinitely many eigenvalues in common with the global eigenvalue problem on M.

5.4.1. Scaling of the Poisson kernel for balls and mean value inequalities. In this section we return to the mean value inequalities of §5.3.3. We use the dilation operator to rescale Green's formula on a small ball $B_{\frac{\varepsilon}{\lambda}}(p)$ of radius $\frac{\varepsilon}{\lambda}$ to obtain a new approach to mean value inequalities. The following is Lemma 7.6 of [**DF1**] (see also Theorem 5.5 and p. 117 of [**HanL2**]):

PROPOSITION 5.12. *Let $(\Delta + \lambda^2)u = 0$, let $p \in M$ satisfy $u(p) = 0$. Let $r = \varepsilon\lambda^{-1}$. Then*

$$\sup_{B_r(p)} |u| \leq C_L \left(\fint_{B_{2r}(p)} |u|^2 \right)^{\frac{1}{2}}. \tag{5.31}$$

We then apply Green's theorem to the ball B to get

$$v(x) - \int_B G_D(\lambda, x, x')((\Delta + \lambda^2)v)\, dV = \int_{\partial B} \partial_{\nu_2} G_D(\lambda, x, q) v(q)\, dS(q). \tag{5.32}$$

We consider a ball $B(p, \frac{\varepsilon}{\lambda})$ with ε to be chosen later and where $\Delta\varphi = -\lambda^2\varphi$. We denote the Dirichlet Green's function of the ball by

$$G_D(\lambda, x, x') = \sum_j \frac{\psi_j(x)\psi_j(x')}{\lambda - \mu_j}, \tag{5.33}$$

where (μ_j, ψ_j) are the Dirichlet eigenfunctions of the ball. Since there exists an orthonormal basis $\{\psi_j\}$ of real valued eigenfunctions, G_D is real valued. Less obviously, for ε sufficiently small, $G_D(\lambda, x, x') < 0$. Here we choose the sign of the Green's function so that

$$(\Delta + \lambda^2) G_D(\lambda, x, y) = \delta_y(x).$$

Thus, $G(\lambda, x, y)$ is "λ-subharmonic" and therefore $-\infty$ at its pole.

LEMMA 5.13. *For ε sufficiently small, so that $\lambda \leq \mu_1(B(p, \frac{\varepsilon}{\lambda}))$, we have $G_D(\lambda, x, x') < 0$.*

PROOF. It is well known that the Dirichlet heat kernel for any metric or domain Ω is positive. But

$$-G_D(\varepsilon, x, x') = \int_0^\infty e^{\varepsilon t} K_D(t, x, x')\, dt \tag{5.34}$$

for any $\varepsilon < \lambda_1^D(\Omega)$, i.e. for ε below the lowest Dirichlet eigenvalue. The integral then converges and the positivity of $K_D(t, x, x') > 0$ implies $-G_D(\varepsilon, x, x') > 0$. □

REMARK 5.14. We could also use the dilation operator $D^p_{\varepsilon^{-1}\lambda}$ to rescale the ball $B = B(p, \frac{\varepsilon}{\lambda})$. Rescaling flattens out the metric so that (for ε sufficiently small) it is close to the Euclidean metric on a ball of radius 1, and changes the eigenvalue from λ to ε. We then observe that the Euclidean Dirichlet Green's function $G_0(\varepsilon, x, y)$ for the unit ball, i.e. the kernel of $(\Delta_0 + \varepsilon^2)^{-1}$, is strictly negative. Indeed, this is well known for $\varepsilon = 0$ where the Dirichlet Green's function can be constructed from the Newtonian potential $-\frac{1}{r^{n-2}}$ by the method of reflections. Hence it must continue to be true for small metric and eigenvalue perturbations. The dilation operator preserves positivity, so the Green's function of the small ball B is also negative.

The following is proved in Section 6 of [**SY**]:

THEOREM 5.15. *Let (M, g) be a complete Riemannian manifold, and let $B = B_{\varepsilon\lambda^{-1}}(p)$ be a geodesic ball of radius $\varepsilon\lambda^{-1}$. Let $B_{\frac{1}{2}} = B_{\frac{1}{2}\varepsilon\lambda^{-1}}(p)$. Suppose that $v \in C^\infty(M)$ satisfies $(\Delta + \lambda^2)v \geq 0$. Then*

$$\sup_{B_{\frac{1}{2}}} |v| \leq C \fint_B v \, dV. \tag{5.35}$$

PROOF. If $(\Delta + \lambda^2)v \geq 0$, then the second term of the left side of (5.32) is positive since Dirichlet Green's function for $(\Delta + \lambda^2)$ is a negative kernel. We then have

$$v(x) \leq \int_{\partial B} \partial_{\nu_2} G_D(\lambda, x, q) v(q) \, dS(q). \tag{5.36}$$

Now let $B_{\frac{1}{2}} = B(p, \frac{1}{2}\frac{\varepsilon}{\lambda})$. Then

$$\sup_{B_{\frac{1}{2}}} |v(x)| \leq \left(\sup_{x \in B_{\frac{1}{2}}, q \in \partial B} |\partial_{\nu_2} G_D(\lambda, x, q)| \right) \int_{\partial B} |v(q)| \, dS(q). \tag{5.37}$$

REMARK 5.16. It is often assumed in addition that $v \geq 0$, and in that case

$$\sup_{B_{\frac{1}{2}}} v(x) \leq \left(\sup_{x \in B_{\frac{1}{2}}, q \in \partial B} |\partial_{\nu_2} G_D(\lambda, x, q)| \right) \int_{\partial B} v(q) \, dS(q). \tag{5.38}$$

Alternatively, by Cauchy-Schwartz, we have

$$\sup_{B_{\frac{1}{2}}} v^2 \leq \left(\sup_{x \in B_{\frac{1}{2}}} \int_{\partial B} |\partial_{\nu_2} G_D(\lambda, x, q)|^2 \, dS(q) \right) \int_{\partial B} v^2(q) \, dS(q). \tag{5.39}$$

To conclude the proof we only need to show

$$\begin{cases} \sup_{x \in B_{\frac{1}{2}}, q \in \partial B} |\partial_{\nu_2} G_D(\lambda, x, q)| \leq C e^{c\lambda r(x,q)}, \\ \sup_{x \in B_{\frac{1}{2}}} \int_{\partial B} |\partial_{\nu_2} G_D(\lambda, x, q)|^2 \, dS(q) \leq e^{c\lambda r(x,q)}. \end{cases} \tag{5.40}$$

This can be proved by rescaling using the dilation operator $D^p_{\varepsilon^{-1}\lambda}$ and comparison to the Euclidean case. $\square$

5.5. Bernstein estimates

In addition to such absolute bounds, there exist Bernstein gradient estimates which compare $\nabla\varphi_\lambda$ to φ_λ either locally or globally. In [**DF3**] the following local Bernstein inequalities are proved:

THEOREM 5.17. *Local eigenfunctions of a Riemannian manifold satisfy:*

(1) L^2 *Bernstein estimate:*

$$\left(\int_{B(p,r)} |\nabla\varphi_\lambda|^2\, dV\right)^{1/2} \leq \frac{C\lambda}{r}\left(\int_{B(p,r)} |\varphi_\lambda|^2\, dV\right)^{1/2}. \tag{5.41}$$

(2) L^∞ *Bernstein estimate: There exists* $K > 0$ *so that*

$$\max_{x\in B(p,r)} |\nabla\varphi_\lambda(x)| \leq \frac{C\lambda^K}{r}\max_{x\in B(p,r)} |\varphi_\lambda(x)|. \tag{5.42}$$

(3) *Dong's improved bound* [**D2**]*:*

$$\max_{B_r(p)} |\nabla\varphi_\lambda| \leq C_1\frac{\lambda^2}{r}\max_{B_r(p)} |\varphi_\lambda| \tag{5.43}$$

for $r \leq C_2\lambda^{-1/2}$.

(4) *The bound of* [**ShXu**]*: There exist* c, C *depending only on* g *so that*

$$c\lambda\|\varphi_\lambda\|_\infty \leq \|\nabla\varphi_\lambda\|_\infty \leq C\lambda\|\varphi_\lambda\|_\infty. \tag{5.44}$$

THEOREM 5.18 (Donnelly-Fefferman). *Let* $\dim M = n$ *and let* $(\Delta + \lambda^2)u = 0$. *Then there exists* $r_0(M)$ *so that, for* $r < r_0$

$$\sup_{B_r(x)} |\nabla u| \leq C\frac{\lambda^{\frac{n+2}{2}}}{r}\sup_{B_r(x)} |u|. \tag{5.45}$$

Donnelly-Fefferman conjecture that $\lambda^{\frac{n+2}{2}}$ can be replaced by λ. This implies the global Bernstein inequality

$$\|\nabla u\|_\infty \leq \lambda\|u\|_\infty. \tag{5.46}$$

This inequality is proved in [**OCP**] as an immediate consequence of the Schoen-Yau estimate

$$\sup_{B_{\frac{a}{2}}} |\nabla h| \leq C_n\left(\frac{1 + a\sqrt{K}}{a}\right)\sup_{B_a} |h|. \tag{5.47}$$

5.6. Frequency function and doubling index

The frequency function $N(a, r)$ of a function u is a local measure of its 'degree' as a polynomial like function in $B_r(a)$. More precisely, it controls the local growth rate of u. In the case of harmonic functions, it is given by

$$N(a, r) = \frac{rD(a, r)}{H(a, r)}, \tag{5.48}$$

where

$$H(a, r) = \int_{\partial B_r(a)} u^2\, d\sigma \quad\text{and}\quad D(a, r) = \int_{B_r(a)} |\nabla u|^2\, dx. \tag{5.49}$$

A well-written detailed treatment of the frequency function and its applications can be found in [**HanL1, Kuk**], following the original treatments in [**GaL1, GaL2, Lin91**].

Frequency functions may also be defined for eigenfunctions. At least two variations have been studied: (i) where the eigenfunctions are converted into harmonic functions on the cone $\mathbb{R}^+ \times M$ as in §5.2.1; (ii) where a frequency function adapted to eigenfunctions on M is defined.

We first consider method (i) in the case of an eigenfunction φ_λ on S^n. The associated harmonic function on the cone is precisely the homogeneous harmonic polynomial on $\mathbb{R}^{n+1}$ whose restriction to the sphere yields φ_λ. By the previous calculation, $N(0, r) \equiv N$ where $\lambda = N(N+n-1)$. We note that on the cone $\mathbb{R}^{n+1}$, the ball of radius r has the form $[0, r] \times S^n$, i.e the frequency function is global on S^n. On a general manifold, the analogous global calculation is cleanest if we define $N(r)$ with $B_r(0)$ everywhere replaced by $[0, r] \times M$. If we 'harmonize' an eigenfunction $\varphi_\lambda \to r^\alpha \varphi_\lambda$ as in §5.2.1, we obtain $N(r) \equiv \alpha$.

The second method is to define a frequency function on balls of M itself. The generalization to eigenfunctions is as follows (see [**GaL1, GaL2, Kuk**]). Fix a point $a \in M$ and choose geodesic normal coordinates centered at a so that $a = 0$. Put

$$\mu(x) = \frac{g_{ij} x_i x_j}{|x|^2}, \tag{5.50}$$

and put

$$D(a, r) := \int_{B_r} \left(g^{ij} \frac{\partial \varphi_\lambda}{\partial x_i} \frac{\partial \varphi_\lambda}{\partial x_j} + \lambda^2 \varphi_\lambda^2 \right) dV \quad \text{and} \quad H(a, r) := \int_{\partial B_r} \mu \varphi_\lambda^2. \tag{5.51}$$

By the divergence theorem, one has

$$D(a, r) = \int_{\partial B_r} \varphi_\lambda \frac{\partial \varphi_\lambda}{\partial \nu} \tag{5.52}$$

Define the frequency function of φ_λ by

$$N(a, r) := \frac{r D(r)}{H(r)}. \tag{5.53}$$

As in the case of harmonic functions, the main properties of the frequency function of an eigenfunction are a certain monotonicity in r in small balls of radius $O(\frac{1}{\lambda})$ (see Theorem 5.19) and the fact that $N(a, r)$ is commensurate with $N(b, r)$ when a and b are close.

Simple examples show that, despite its name, the frequency function measures local growth but not frequency of oscillations of eigenfunctions, and therefore is not necessarily comparable to λ. For instance, the frequency function of $\sin nx$ in a ball or radius $\frac{C}{|n|}$ is bounded. An example considered in [**DF1**] are the global eigenfunctions $e^{sx} \sin ty$ on $\mathbb{R}^2$ of Δ-eigenvalue $s^2 - t^2$ and frequency function of size s. One could let $s, t \to \infty$ with $s^2 - t^2$ bounded and obtain a high frequency function but a low eigenvalue. However, as discussed in [**Lin91**] (p. 291), if one forms the harmonic function from a global eigenfunction as in §5.2.1, then one has $N(0, 2) \leq C\lambda$ where C depends only on the metric. This is a global estimate since a ball centered at 0 in the cone will cover all of M. An application is given in Theorem 5.20.

We now state Theorem 2.3 of [**GaL1**]; see also [**GaL2, Lin91, HanL1**] and [**Kuk**], Theorem 2.3, 2.4.

THEOREM 5.19. *There exists* $C > 0$ *such that* $e^{Cr}(N(r) + \lambda^2 + 1)$ *is a non-decreasing function of* r *in some interval* $[0, r_0(\lambda)]$.

Another basic fact is that the frequency of φ_λ in $B_r(a)$ is comparable to its frequency in $B_R(b)$ if a, b are close and r, R are close. More precisely, there exists $N_0(R) \ll 1$ such that if $N(0,1) \leq N_0(R)$, then φ_λ does not vanish in B_R, while if $N(0,1) \geq N_0(R)$, then

$$N\left(p, \frac{1}{2}(1-R)\right) \leq CN(0,1) \quad \text{for all } p \in B_R. \tag{5.54}$$

5.6.1. Doubling and vanishing order estimates. An important series of estimates are known as *doubling estimates*.

THEOREM 5.20 (Donnelly-Fefferman, Lin and [**HanL1**] Lemma 6.1.1). *Let* φ_λ *be a global eigenfunction of a smooth Riemannian manifold* (M, g). *Then there exists* $C = C(M, g)$ *and* r_0 *such that for* $0 < r < r_0$,

$$\frac{1}{\mathrm{Vol}(B_{2r}(a))} \int_{B_{2r}(a)} |\varphi_\lambda|^2 \, dV_g \leq e^{C\lambda} \frac{1}{\mathrm{Vol}(B_r(a))} \int_{B_r(a)} |\varphi_\lambda|^2 \, dV_g. \tag{5.55}$$

Further,

$$\max_{B(p,r)} |\varphi_\lambda(x)| \leq \left(\frac{r}{r'}\right)^{C\lambda} \max_{x \in B(p,r')} |\varphi_\lambda(x)| \qquad (0 < r' < r). \tag{5.56}$$

The doubling estimates imply the vanishing order estimates. Let $a \in M$ and suppose that $u(a) = 0$. By the vanishing order $\nu(u, a)$ of u at a is meant the largest positive integer such that $D^\alpha u(a) = 0$ for all $|\alpha| \leq \nu$.

THEOREM 5.21. *Suppose that* M *is compact and of dimension* n. *Then there exist constants* $C(n), C_2(n)$ *depending only on the dimension such that the the vanishing order* $\nu(u, a)$ *of* u *at* $a \in M$ *satisfies* $\nu(u, a) \leq C(n)\ N(0,1) + C_2(n)$ *for all* $a \in B_{1/4}(0)$. *In the case of a global eigenfunction,* $\nu(\varphi_\lambda, a) \leq C(M, g)\lambda$.

The bounds are sharp since they are achieved by highest weight spherical harmonics (Gaussian beams), which vanish to maximum order at the poles and saturate doubling estimates in the upper or lower hemisphere at some fixed distance from the equator. As in the case of sup-norm bounds, we may ask for which (M, g) there exist sequences of eigenfunctions saturating the doubling estimates and VO (vanishing order) estimates. Sequences saturating the VO estimates of course saturate doubling estimates, but the converse is not known to be true. In all known examples where the VO estimates are sharp (such as spheres or surfaces of revolution), there exists a point p so that a sequence of eigenfunctions vanishes to maximum order at p, and all geodesics emanating from p return to p at a fixed time. This is the same condition for existence of eigenfunctions saturating sup norm bounds (see the Chapter on L^p norms), but the eigenfunctions are quite different and we do not know any mutual implications between the existence of the two eigenfunction sequences.

5.7. Carleman estimates

A local Carleman estimate is a weighted energy inequality. Given a well-chosen weight function ψ one forms the weighted Hilbert space $L^2(U, e^{\psi/h})$ over an open set $U \subset M$. Given an elliptic operator $P = -\Delta$, a Carleman inequality has the form

$$h\|e^{\psi/h}u\|^2_{L^2(U)} + h^3\|e^{\psi/h}\nabla u\|^2_{L^2(U)} \leq Ch^4\|e^{\psi/h}Pu\|^2_{L^2(U)}. \tag{5.57}$$

The key condition on ψ is that the conjugated operator (a semiclassical pseudo-differential operator)

$$P_\psi = h^2 e^{\psi/h} P e^{-\psi/h} = -h^2\Delta - |d\psi|^2 + \langle \nabla\psi, h\nabla\rangle + h\Delta\psi \tag{5.58}$$

be subelliptic. The principal symbol of P_ψ is

$$p_\psi(x,\xi) = |\xi|^2 - |\nabla\psi|^2 + 2i\langle\nabla\psi,\xi\rangle = \sum_j (\xi_j + i\partial_{x_j}\psi)^2. \tag{5.59}$$

This follows fom the fact that the symbol of $e^{\psi/h}D_j e^{-\psi/h}$ is $\xi_j + i\partial_{x_j}\psi$.

Define

$$Q_1 = \frac{1}{2i}(P_\psi - P_\psi^*) \quad \text{and} \quad Q_2 = \frac{1}{2i}(P_\psi + P_\psi^*). \tag{5.60}$$

Let q_j be the principal symbol of Q_j. Then

$$q_1 = 2\langle\xi,\nabla\psi\rangle \quad \text{and} \quad q_2 = |\xi|^2 - |\nabla\psi(x)|^2. \tag{5.61}$$

The assumption on ψ is that on the characteristic set

$$\text{Char}(p_\psi) = \{(x,\xi) \in \overline{U} \times \mathbb{R}^n : p_\psi(x,\xi) = 0 \iff |\xi| = |\nabla\psi| \text{ and } \langle\xi,\nabla\psi\rangle = 0\}, \tag{5.62}$$

one has a positive Poisson bracket on the ordered pair (q_1, q_2):

$$\text{for all } (x,\xi) \in \overline{U} \times \mathbb{R}^n,\ p_\psi(x,\xi) = 0 \text{ implies } \{q_2, q_1\}(x,\xi) \geq C > 0. \tag{5.63}$$

Note that $\psi = e^{\tau\rho}$ satisfies the condition if $|\nabla\rho| > 0$ in U and τ is sufficiently large.

One then has:

PROPOSITION 5.22. *Let ψ satisfy the assumption above in $\overline{U}$. Then there exists $h_1 > 0$ so that* (5.57) *holds for $u \in C_c^\infty(\overline{U})$ and $0 < h < h_1$.*

We sketch the proof: Let $v = e^{\psi/h}$. Then

$$Pu = f \iff P_\psi v = g := h^2 e^{\psi/h} f \iff Q_2 v + iQ_1 v = g.$$

But $\langle Q_j w_1, w_2\rangle = \langle w_1, Q_j w_2\rangle$ if $w_1, w_2 \in C_c^\infty(\mathbb{R}^n)$. Hence

$$\|g\|^2_{L^2} = \|Q_1 v\|^2_{L^2} + \|Q_2 v\|^2_{L^2} + 2\text{Re}\langle Q_2 v, iQ_1 v\rangle = \langle (Q_1^2 + Q_2^2 + i[Q_2, Q_1])v, v\rangle. \tag{5.64}$$

For sufficiently large $\mu > 0$ and for h so that $h\mu < 1$,

$$h\langle (Q_1^2 + Q_2^2 + i[Q_2, Q_1])v, v\rangle \leq \|g\|^2_{L^2}. \tag{5.65}$$

Then Gårding's inequality gives

$$h\|v\|^2_{H^1} \leq C\|g\|^2_{L^2}. \tag{5.66}$$

Here, H^1 is the Sobolev space. Since $\nabla(e^{\psi/h}u) = h^{-1}e^{\psi/h}(\nabla\psi)y + e^{\psi/h}\nabla u$ and $\|\nabla\psi\| \leq C$, this gives

$$h\|e^{\psi/h}u\|^2_{L^2} + h^3\|\nabla(e^{\psi/h}u)\|^2_{L^2} \leq Ch^4\|e^{\psi/h}f\|^2_{L^2}.$$

We recall that Gårding's inequality asserts that positivity of a symbol implies positivity of the associated semiclassical pseudo-differential operator. More precisely, if $K \subset \mathbb{R}^n$ is compact and $a(x, \xi, h)$ is a semiclassical symbol with principal part a_m, and if there exists $C > 0$ so that $\text{Re} a_m(x, \xi, h) \geq C(1 + |\xi|^2)^{m/2}$ for $x \in K, \xi \in \mathbb{R}^n$ and $h \in (0, h_0)$ then there exists $C' > 0$ so that

$$\text{Re}\langle Op_h(a)u, u\rangle \geq C' \|u\|^2_{H^{m/2}} \quad \text{for all } u \in C_c^\infty(K) \text{ and } 0 < h < h_1. \tag{5.67}$$

We now recall quantitative lower bound estimates. They follow from doubling estimates and also from Carleman inequalities.

Theorem 5.23. *Suppose that M is compact and that φ_λ is a global eigenfunction, $\Delta \varphi_\lambda = -\lambda^2 \varphi_\lambda$. Then for all p, r there exist $C, C' > 0$ so that*

$$\max_{x \in B(p,r)} |\varphi_\lambda(x)| \geq C' e^{-C\lambda}. \tag{5.68}$$

Local lower bounds on $\frac{1}{\lambda} \log |\varphi_\lambda^{\mathbb{C}}|$ follow from doubling estimates. They imply that there exists $A, \delta > 0$ so that, for any $\zeta_0 \in \overline{M_{\tau/2}}$,

$$\sup_{\zeta \in B_\delta(\zeta_0)} |\varphi_\lambda(\zeta)| \geq C e^{-A\lambda}. \tag{5.69}$$

Indeed, there of course exists a point $x_0 \in M$ so that $|\varphi_\lambda(x_0)| \geq 1$. Any point of $\overline{M}_{\tau/2}$ can be linked to this point by a smooth curve uniformly bounded length. We then choose δ sufficiently small so that the δ-tube around the curve lies in M_τ and link $B_\delta(\zeta)$ to $B_\delta(x_0)$ by a chain of δ-balls in M_τ where the number of links in the chain is uniformly bounded above as ζ varies in M_τ. If the balls are denoted B_j we have $\sup_{B_{j+1}} |\varphi_\lambda| \leq e^{\beta\lambda} \sup_{B_j} |\varphi_\lambda|$ since $B_{j+1} \subset 2B_j$. The growth estimate implies that for any ball B, $\sup_{2B} |\varphi_\lambda| \leq e^{C\lambda} \sup_B |\varphi_\lambda|$. Since the number of balls is uniformly bounded,

$$1 \leq \sup_{B_\delta(x_0)} |\varphi_\lambda| \leq e^{A\lambda} \sup_{B_\delta(\zeta)} |\varphi_\lambda|$$

and we get a contradiction if no such A exists.

As an illustration, Gaussian beams such as highest weight spherical harmonics decay at a rate $e^{-C\lambda d(x,\gamma)}$ away from a stable elliptic orbit γ. Hence if the closure of an open set is disjoint from γ, one has a uniform exponential decay rate which saturate the lower bounds.

Example: Let $U = \mathcal{T}_{\lambda^{-1/2}}(\gamma) \backslash \mathcal{T}_{\lambda^{-1}}(\gamma)$. Then we may take $\rho = \log \text{dist}_\gamma(x)$ (the distance to γ) in the previous example. However, it is not clear that the theorem applies since the domain is changing with $h = \lambda^{-1}$.

Let us also try $\psi(x) = \text{dist}^2_\gamma(x)$. Thus, $\psi(s, y) = |y|^2$ in Fermi normal coordinates, and

$$\nabla \psi = 2 \,\text{dist}_\gamma(x) \nabla \,\text{dist}_\gamma(x), \qquad |\nabla \psi| = 2 \,\text{dist}_\gamma(x), \tag{5.70}$$

i.e., $\nabla \psi$ is normal to level sets of the distance to γ. Hence $\langle \xi, \nabla \psi \rangle = 0$ if and only if ξ is tangent to the level sets of dist_γ, and $|\xi| = |\nabla \psi| \iff |\xi| = 2 \,\text{dist}_\gamma(x)$. Since the domain is shrinking we must introduce a cutoff χ_λ supported in the region between the domains.

5.8. Norm square of the Cauchy data

Let $H \subset M$ be a smooth orientable hypersurface and let ∂_ν denote a choice of unit normal. The semiclassical Cauchy data of φ_λ along H is defined by

$$\mathrm{CD}(\varphi_\lambda) := \{(\varphi_\lambda|_H, \lambda^{-1}\partial_\nu \varphi_\lambda|_H)\}. \tag{5.71}$$

The term semiclassical refers to the normalization. The global eigenfunction can be recovered from its Cauchy data along a separating hypersurface.

Following Dong [**D1**], we define the norm-squared of the global Cauchy data,

$$q = q(\varphi) = |\nabla\varphi|^2 + \frac{\lambda^2}{2}\varphi^2. \tag{5.72}$$

In Theorem 3.3 of [**D1**] the following is proved:

THEOREM 5.24. *Let Δ be the negative Laplacian on a surface M^2. Let φ_λ be an eigenfunction of eigenvalue $-\lambda^2$ and let $\mathcal{S} = \{p_1, \ldots, p_\ell\}$ be the set of singular points of φ_λ, i.e. points where $\varphi_\lambda(p) = d\varphi_\lambda(p) = 0$. Let k_j be the vanishing order of φ_λ at p. Then,*

$$\Delta \log q \geq -\lambda^2 + 2\min\{K, 0\} + 4\pi \sum_j (k_j - 1)\delta_{p_j}. \tag{5.73}$$

The exact formula is

$$\int_M f \Delta \log q = 4\pi \sum_j (k_j - 1) f(p_j) \tag{5.74}$$

$$+ \lim_{\varepsilon \to 0} \sum_j \left[\int_{U \setminus B_\varepsilon(p_j)} f \left(\frac{\lambda}{2} \frac{(\varphi_{11} - \varphi_{22})^2}{q^2} + (2K - \lambda^2) \frac{|\nabla\varphi|^2}{q} \right) \right]. \tag{5.75}$$

Here, φ_{jj} are the second derivatives of φ in a local frame field given by the Hessian eigenvectors, so that φ_{jj} are the eigenvalues of the Hessian $\nabla^2\varphi$. Setting $f \equiv 1$ gives

$$4\pi \sum_{p_j \in \Sigma} (k_j - 1) = (\lambda^2 - 2K) \int_M \frac{|\nabla\varphi|^2}{q} + \int_M \left(\frac{\lambda}{2} \frac{(\varphi_{11} - \varphi_{22})^2}{q^2} \right). \tag{5.76}$$

COROLLARY 5.25. *If (M, g) is a Riemannian surface and if $(\Delta + \lambda^2)\varphi_\lambda = 0$ then*

$$\sum_i k_i \leq \frac{1}{4\pi} \left[\lambda^2 \operatorname{Vol}(M) - 2\int_M \min\{K, 0\} \right]. \tag{5.77}$$

Note that the order of magnitude is achieved by separation of variables eigenfunctions on a flat torus or round sphere, e.g., $\sin kx \sin ky$ with eigenvalue $-2k^2$ and with k^2 singular points.

We briefly sketch the proof. The Bochner formula gives an exact formula away from the singular points:

$$\Delta \log q = \frac{1}{q} \left[2|\nabla^2\varphi_\lambda|^2 - \lambda^2 \varphi_\lambda^2 + (2K - \lambda^2)|\nabla\varphi_\lambda|^2 \right] \tag{5.78}$$

$$- \frac{1}{q^2} \left| 2\langle \nabla^2\varphi_\lambda, \nabla\varphi_\lambda \rangle + \lambda^2 \varphi_\lambda \nabla^2 \varphi_\lambda \right|^2. \tag{5.79}$$

Choose an orthogonal coordinate frame in which $\nabla^2\varphi_\lambda$ is diagonal at a fixed point. For simplicity write $\varphi = \varphi_\lambda$. Then

$$\nabla^2\varphi = \begin{pmatrix} \varphi_{11} & 0 \\ 0 & \varphi_{22} \end{pmatrix}, \quad \nabla u = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}. \tag{5.80}$$

Then

$$\Delta \log q = \frac{\lambda^2}{2}\frac{\varphi^2(\varphi_{11}-\varphi_{22})^2}{q^2} + (2K-\lambda^2)\frac{|\nabla\varphi|^2}{q}. \tag{5.81}$$

We briefly review the calculation because the factor φ^2 in the numerator of the first term is missing in [**D1**]. One has

$$\Delta \log q = \frac{1}{q}\left[2(\varphi_{11}^2+\varphi_{22}^2) - (\varphi_{11}+\varphi_{22})^2 + (2K-\lambda^2)|\nabla\varphi|^2\right] \tag{5.82}$$

$$-\frac{1}{q^2}\left(\begin{pmatrix} \varphi_{11}-\varphi_{22} & 0 \\ 0 & \varphi_{22}-\varphi_{11} \end{pmatrix}\begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}\right)^2 \tag{5.83}$$

$$= \frac{1}{q^2}\left(q(\varphi_{11}-\varphi_{22})^2 + q(2K-\lambda^2)|\nabla\varphi|^2 - (\varphi_{11}-\varphi_{22})^2|\nabla\varphi|^2\right) \tag{5.84}$$

$$= \frac{\lambda^2}{2}\frac{\varphi^2(\varphi_{11}-\varphi_{22})^2}{q^2} + (2K-\lambda^2)\frac{|\nabla\varphi|^2}{q}. \tag{5.85}$$

In the last line we use that $q - |\nabla\varphi|^2 = \frac{1}{2}\lambda^2\varphi^2$.

Since $\frac{|\nabla\varphi|}{q} \leq 1$ one gets

$$\Delta \log q \geq -\lambda^2 + 2\min\{K, 0\}. \tag{5.86}$$

Next consider a singular point p. Choose geodesic polar coordinates (r,θ) centered at p. For simplicity denote φ_λ by φ. Then if φ vanishes to order k at $r = 0$,

$$\begin{cases} \varphi = ar^k\cos k\theta + O(r^{k+1}), \\ \nabla\varphi = \nabla(ar^k\cos k\theta) + O(r^k), \\ \nabla^2\varphi = \nabla^2(ar^k\cos k\theta) + O(r^{k-1}). \end{cases} \tag{5.87}$$

$$\begin{cases} \text{(i)} \quad q = |\nabla\varphi|^2 + \lambda^2\dfrac{\varphi^2}{2} = k^2r^{2k-2} + O(r^{2k-1}), \\ \text{(ii)} \quad \log q = \ln k^2 + (2k-2)\log r + O(r), \\ \text{(iii)} \quad \nabla \log q = \dfrac{2k-2}{r} + O(1). \end{cases} \tag{5.88}$$

Let U be ball around p containing no other singular points. Let f be a test function supported in U. Then

$$\int_U f\Delta\log q = \lim_{\varepsilon\to 0}\int_{U-B_\varepsilon(p)} (\Delta f)\log q \tag{5.89}$$

$$= \lim_{\varepsilon\to 0}\left[\int_{\partial B_\varepsilon(p)}\frac{\partial f}{\partial\nu}\log q - \int_{\partial B_\varepsilon(p)} f\frac{\partial}{\partial\nu}\log q + \int_{U\setminus B_\varepsilon(0)} f\Delta\log q\right]. \tag{5.90}$$

It follows from (ii) that the first term tends to zero as $\varepsilon \to 0$. It follows from (iii) that the second term tends to $4\pi(k-1)f(0)$. The third term is an integral over a domain with no singular points where (5.81) is valid. Thus,

$$\int_U f\Delta \log q = 4\pi(k-1)f(0) + \lim_{\varepsilon\to 0}\left[\int_{U\setminus B_\varepsilon(p)} f\left(\frac{\lambda^2}{2}\frac{(\varphi_{11}-\varphi_{22})^2}{q^2} + (2K-\lambda^2)\frac{|\nabla\varphi|^2}{q}\right)\right]. \tag{5.91}$$

It follows that for a non-negative test function f in an open set U containing only one singular point p, and with $k = \operatorname{ord}_p \varphi$,

$$\int_U f\Delta \log q \geq \int_U f\left[-\lambda^2 + 2\min\{K,0\}\right] + 4\pi(k-1)f(p), \tag{5.92}$$

concluding the proof of the theorem.

Note that the limit as $\varepsilon \to 0$ in the integral of (5.91) converges. Indeed, $\frac{|\nabla\varphi|^2}{q} \leq 1$ and $\frac{\varphi(\varphi_{11}-\varphi_{22})}{q} = O(r^k \frac{r^{k-2}}{r^{2k-2}}) = O(1)$. (Note that the factor of φ^2 which is missing in [**D1**]. Its presence makes this ratio bounded.)

5.8.1. The case where $-\lambda^2$ is not an eigenvalue for the $\pm$ Dirichlet problem for $\Delta + \lambda^2$. We may then define the second order Dirichlet-to-Neumann operators by

$$\mathcal{N}_2^{\pm}(\lambda)u(s) = \frac{\partial^2}{\partial y^2}\varphi_\lambda(s,y)|_{y=0}, \tag{5.93}$$

where φ_λ is the extension of $u \in C(H)$ to an eigenfunction of Δ_g of eigenvalue $-\lambda^2$ on $M_\pm$. The first Dirichlet-to-Neumann operator is similar but takes the first derivative. The (first) Dirichlet-to-Neumann operator is well-known to be a first order pseudo-differential operator on H. The second one (5.93) is a second order pseudo-differential operator. But its behavior in λ is more relevant to our problem.

In this case we have

$$-\int_H \left|\frac{\partial\varphi_\lambda}{\partial s}\right|^2 ds + \langle \mathcal{N}_2(\lambda)\varphi_\lambda, \varphi_\lambda\rangle_{L^2(H)} = -\lambda^2 \int_H |\varphi_\lambda|^2 ds. \tag{5.94}$$

We can determine $\mathcal{N}_2(\lambda)$ from Green's formula. We then take two normal derivatives to get

$$\partial_\nu^2 \varphi_\lambda(x)|_H = \int_H (\partial_\nu\varphi_\lambda)\partial_\nu^2 G_\lambda - \varphi_\lambda \partial_{\nu_1}^2\partial_{\nu_2} G_\lambda|_H.$$

When we use the Dirichlet Green's function, the first term vanishes and we obtain

$$\partial_\nu^2 \varphi_\lambda(x)|_H =: \mathcal{N}_2(\lambda)(\varphi_\lambda(x)|_H) = -\int_H \varphi_\lambda \partial_{\nu_1}^2 \partial_{\nu_2} G_\lambda^D|_H. \tag{5.95}$$

Rewriting the restricted eigenvalue equation in terms of (5.93) gives,

$$-\frac{\partial^2}{\partial s^2}\varphi_\lambda(s,0) - \mathcal{N}_2(\lambda)\varphi_\lambda(s,0) = \lambda^2\varphi_\lambda(s,0). \tag{5.96}$$

We view it as a kind of operator Sturm-Liouville equation with λ-dependent 'potential' $V_\lambda = \mathcal{N}_2(\lambda)$.

As in [**TayII**], $\mathcal{N}_2(\lambda)$ is a pseudo-differential operator on H. To see this one may use the ambient Green's kernel $G_M(\lambda, x, y)$ again assuming that $-\lambda \notin \mathrm{Sp}(\Delta)$. We then define the double layer potential on $M \times H$ by $\partial_{\nu_2} G_M(\lambda, x, y)$.

This case is simpler and more amenable to study. Let us suppose that $\varphi_\lambda|_H = 0$. Then for any choice of G_λ in (5.29) we have (as in (5.30)),

$$\varphi_\lambda(x) = \int_H (\partial_\nu \varphi_\lambda)(q) G_\lambda(x, q) dS(q). \tag{5.97}$$

Hence,

$$\partial_\nu^2 \varphi_\lambda(x)|_H = \int_H (\partial_\nu \varphi_\lambda) \partial_\nu^2 G_\lambda dS(q). \tag{5.98}$$

Similarly if $\partial_\nu \varphi_\lambda|_H = 0$ then we obtain (5.95),

$$\partial_\nu^2 \varphi_\lambda(x)|_H = -\int_H \varphi_\lambda(q) \partial_{\nu_x}^2 \partial_{\nu_q} G_\lambda(x, q) dS(q). \tag{5.99}$$

5.9. Hyperbolic aspects

5.9.1. Elliptic on small length scales, hyperbolic on large length scales. In this section we explain the idea that the Helmholtz operator $\Delta + \lambda^2$ is elliptic on small length scales but hyperbolic on large ones. The Helmholtz operator is of course elliptic if we view the constant λ as a lower order term. The symbol of the Laplacian Δ is $|\xi|_g^2 = \sum_{i,j=1}^n g^{ij}(x)\xi_i\xi_j$ and it vanishes only when $\xi = 0$ in T^*M. But when we are interested in the asymptotic behavior as $\lambda \to \infty$, we should view λ as an operator of order 1. More precisely, in semiclassical analysis one writes differential operators as polynomials in $\lambda^{-1}\frac{\partial}{\partial x_j}$, and in particular one rewrites the Helmholtz operator as $\lambda^{-2}\Delta + 1$, whose semiclassical symbol $|\xi|_g^2 - 1$ vanishes on the unit cosphere bundle S^*M. Thus, in semiclassical analysis it is natural to view $\Delta + \lambda^2$ as a hyperbolic operator, and it is not surprising to find that the dynamics of the geodesic flow are intimately related to spectral asymptotics. Indeed, a Fourier transform in time conjugates $\Delta + \lambda^2$ to the wave operator $\Delta + \frac{\partial^2}{\partial t^2}$.

On the other hand, it is natural to view $\Delta + \lambda^2$ as elliptic on length scales of order $\varepsilon\lambda^{-1}$. This is because the lowest eigenvalue of the Dirichlet Laplacian on a ball of radius $r = \varepsilon\lambda^{-1}$ will be larger than λ^2 for ε sufficiently small. As a result, the maximum principle and standard mean value inequalities apply to $\Delta + \lambda^2$ on $B_r(p)$, and the eigenfunctions behave rather like harmonic functions if they are dilated as in §5.2. We summarize with the statement that $\Delta + \lambda^2$ is elliptic on the length scale $r = \varepsilon\lambda^{-1}$ but is hyperbolic on larger length scales.

5.9.2. Wave equation and local reproducing formulae. To study the size of an eigenfunction φ_λ near a point p, it is useful to relate its value at p to its values near p. For this purpose it is useful to have a local reproducing operator, i.e., an operator whose Schwartz kernel is supported in $x \in B_r(p)$ for some small r. For instance, the ball means operator M_r reproduces harmonic functions on $\mathbb{R}^n$ using its values in a ball of radius r around p. It is one of the main sources of estimates on harmonic functions.

There are at least three ways to construct local reproducing operators on a Riemannian manifold. The first involves the wave operators $E(t) = \cos t\sqrt{-\Delta}$ and $S(t) = \frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}$. The key feature of these operators is the finite propagation speed of the wave equation, which means that $E(t,x,p)$ and $S(t,x,p)$ are supported for $x \in B_t(p)$. They do not reproduce the eigenfunctions but rather multiply it by $\cos t\sqrt{-\lambda}$, resp. $\frac{\sin t\sqrt{-\lambda}}{\sqrt{-\lambda}}$. However, if we let ρ be a smooth Schwartz function on $\mathbb{R}$ whose Fourier transform satisfies $\hat{\rho} \subset (-\varepsilon, \varepsilon)$ and $\hat{\rho}(0) = 1$, then

$$\rho(\sqrt{\lambda} - \sqrt{-\Delta})\varphi_\lambda = \varphi_\lambda, \tag{5.100}$$

and the Schwartz kernel $\rho(\sqrt{\lambda} - \sqrt{-\Delta})(x,p)$ is supported for $x \in B_\varepsilon(p)$. Hence $\rho(\sqrt{\lambda} - \sqrt{-\Delta})$ is somewhat like a ball means operator of radius t. However, it is not an averaging operator on all functions, nor is its Schwartz kernel a positive measure. In fact, it is a semiclassical oscillatory (Fourier integral) operator which will be studied in detail below. The rapid oscillation of the kernel is quite distinct from the behavior of a ball means operator.

5.9.3. Spherical means and ball means operators. A second (and somewhat related) method is to relate the ball or spherical means operator to the operators $E(t), S(t)$. The spherical means operator is an averaging operator over the Riemannian sphere $S_r(x)$ of radius r centered at x and the ball means operator averages over the ball $B_r(x)$ centered at x of radius r. There are two natural definitions according to the measure one puts on the spheres, resp. balls. The tangential spherical means operator is defined by

$$L_r^0 f(x) = \int_{S_x^* M} f(\exp_x r\xi)\, d\mu_x(\xi), \tag{5.101}$$

where $d\mu_x$ is the normalized surface measure on the unit co-sphere $S_x^* M$ induced by the metric g. The tangential ball means is defined by

$$M_r^0 f(x) = \int_{rB_x^* M} f(\exp_x r\xi)\, d\mathcal{L}(r\xi), \tag{5.102}$$

where $d\mathcal{L}$ is Lebesgue measure in the tangent space In both cases one uses Euclidean surface area form, resp. volume form, on the tangent space at x.

The (non-tangential) spherical means operator is defined by

$$L_r f(x) = \frac{1}{|S_r(x)|} \int_{S_r(x)} f\, dA_x, \tag{5.103}$$

where $|S_r(x)|$ is the Riemannian surface area of $S_r(x)$, and the ball means operator is

$$M_r f(x) = \frac{1}{|B_r(x)|} \int_{B_r(x)} f\, dV_g \tag{5.104}$$

where $|B_r(x)|$ is the volume of the ball.

In Euclidean $\mathbb{R}^n$ the two types of spherical means (resp. ball means) operators agree, but they differ in general on curved Riemannian manifolds, where the surface measure on a geodesic sphere is not the pushforward under the exponential map of the Euclidean surface measure in the tangent space. An important difference between Euclidean spaces (or more generally harmonic spaces) and general Riemannian manifolds is that in the first case, the spherical means operator or ball means operator reproduce eigenfunctions of Δ up to a scalar. This is a rare property on a Riemannian manifold and on large length scales the wave reproducing kernels are much more useful.

For sufficiently small t (less than the injectivity radius $\text{inj}(M, g)$), the even wave operator $E(t) = \cos t\sqrt{-\Delta}$ and the spherical means operators are both Fourier integral operators with the same canonical relation, i.e., the union of the graphs of the geodesic flow G^t and G^{-t}. Both have nowhere vanishing principal symbols. It follows that for small $|t|$, there exist pseudo-differential operators $A(t, D_t, x, D_x), B(t, D_t, x, D_x)$ so that $E(t) = A(t, D_t, x, D_x)L_t$ and vice-versa, $L_t = B(t, D_t, x, D_x)E(t)$. Applied to an eigenfunction, this gives

$$L_t\varphi_\lambda = B(t, D_t, x, D_x) \cos t\sqrt{\lambda}\varphi_\lambda. \tag{5.105}$$

The change from tangential to local spherical means operator is just a change in B. This is the closest one has to a reproducing formula using spherical means operators. For $|t| > \text{inj}(M, g)$ on a manifold with conjugate points, distance spheres can develop singularities and the spherical means kernel becomes more singular than the wave kernel.

Bibliography

[Ber] L. Bers, Local behavior of solutions of general linear elliptic equations. Comm. Pure Appl. Math. 8 (1955), 473–496.

[BNV] H. Berestycki, L. Nirenberg, and S.R.S. Varadhan, The principal eigenvalue and maximum principle for second-order elliptic operators in general domains. Comm. Pure Appl. Math. 47 (1994), no. 1, 47-92.

[Ca] X. Cabré, On the Alexandroff-Bakelman-Pucci estimate and the reversed Hlder inequality for solutions of elliptic and parabolic equations. Comm. Pure Appl. Math. 48 (1995), no. 5, 539-570.

[Ch] I. Chavel, *Eigenvalues in Riemannian geometry.* Including a chapter by Burton Randol. With an appendix by Jozef Dodziuk. Pure and Applied Mathematics, 115. Academic Press, Inc., Orlando, FL, 1984.

[CM1] T. Colding and W. P. Minicozzi II, Volumes for eigensections. Geom. Dedicata 102 (2003), 19-24.

[CM2] T. Colding and W. P. Minicozzi II, An excursion into geometric analysis. Surveys in differential geometry. Vol. IX, 83-146, Surv. Differ. Geom., IX, Int. Press, Somerville, MA, 2004.

[CM3] T. Colding and W. P. Minicozzi II, Lower bounds for nodal sets of eigenfunctions. Comm. Math. Phys. 306 (2011), no. 3, 777-784.

[D1] R. T. Dong, Nodal sets of eigenfunctions on Riemann surfaces. J. Differential Geom. 36 (1992), no. 2, 493-506.

[D2] R. T. Dong, A Bernstein type of inequality for eigenfunctions. J. Differential Geom. 42 (1995), no. 1, 23-29.

[D3] R. T. Dong, A BMO bound for eigenfunctions on Riemannian manifolds. Degenerate diffusions (Minneapolis, MN, 1991), 71-76, IMA Vol. Math. Appl., 47, Springer, New York, 1993.

[DF1] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions on Riemannian manifolds, Invent. Math. 93 (1988), 161-183.

[DF2] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions: Riemannian manifolds with boundary. Analysis, et cetera, 251–262, Academic Press, Boston, MA, 1990.

[DF3] H. Donnelly and C. Fefferman, Growth and geometry of eigenfunctions of the Laplacian. Analysis and partial differential equations, 635–655, Lecture Notes in Pure and Appl. Math., 122, Dekker, New York, 1990.

[GaL1] N. Garofalo and F. H. Lin, Monotonicity properties of variational integrals, A_p weights and unique continuation. Indiana Univ. Math. J. 35 (1986), no. 2, 245–268.

[GaL2] N. Garofalo and F. H. Lin, Unique continuation for elliptic operators: a geometric-variational approach. Comm. Pure Appl. Math. 40 (1987), no. 3, 347–366.

[GT] D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001.

[HanL1] Q. Han and Fanghua Lin, *Elliptic partial differential equations.* Second edition. Courant Lecture Notes in Mathematics, 1. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2011.

[HanL2] Q. Han and Fanghua Lin, *Nodal sets of Solutions to Elliptic Equations* (preprint).

[HL] X. Han and G. Lu, A geometric covering lemma and nodal sets of eigenfunctions. Math. Res. Lett. 18 (2011), no. 2, 337-352.

[HarS] R. Hardt and L. Simon, Nodal sets for solutions of elliptic equations. J. Differential Geom. 30 (1989), no. 2, 505–522.

[HarW1] P. Hartman and A. Wintner, On the local behavior of solutions of non-parabolic partial differential equations. Amer. J. Math. 75, (1953). 449–476.

[HarW2] P. Hartman and W. Wintner, On the local behavior of solutions of non-parabolic partial differential equations. III. Approximations by spherical harmonics. Amer. J. Math. 77 (1955), 453–474.

[He] E. Hebey, Sobolev Spaces on Riemannian Manifolds, Lecture Notes in Math. 1365, Springer, Berlin, 1996.

[HO12] T. Hoffmann-Ostenhof, Problems for the number of nodal domains in "Geometric Aspects of Spectral Theory xiv" Oberwolfach Reports 33, page 2068-270, 2012.

[J] F. John, *Plane waves and spherical means applied to partial differential equations.* Dover Publications, Inc., Mineola, NY, 2004.

[Kua] I. Kukavica, Nodal volumes for eigenfunctions of analytic regular elliptic problems. J. Anal. Math. 67 (1995), 269–280.

[Kuk] I. Kukavica, Quantitative uniqueness for second-order elliptic operators. Duke Math. J. 91 (1998), no. 2, 225–240.

[LiSch] P. Li and R. Schoen, L^p and mean value properties of subharmonic functions on Riemannian manifolds. Acta Math. 153 (1984), no. 3-4, 279-301.

[Lin91] F. H. Lin, Nodal sets of solutions of elliptic and parabolic equations, Comm. Pure Appl. Math., 45, 287-308, (1991).

[Ni] L. Ni, Mean value theorems on manifolds. Asian J. Math. 11 (2007), no. 2, 277-304.

[OCP] J. Ortega-Cerdà and B. Pridhnani, Carleson measures and Logvinenko-Sereda sets on compact manifolds. Forum Math. 25 (2013), no. 1, 151-172.

[P] P. Padilla, The principal eigenvalue and maximum principle for second order elliptic operators on Riemannian manifolds. J. Math. Anal. Appl. 205 (1997), no. 2, 285-312.

[ShXu] Y. Shi and B. Xu, Gradient estimate of an eigenfunction on a compact Riemannian manifold without boundary. Ann. Global Anal. Geom. 38 (2010), no. 1, 21-26.

[SY] R. Schoen and S. T. Yau, Lectures on differential geometry. . Conference Proceedings and Lecture Notes in Geometry and Topology, I. International Press, Cambridge, MA, 1994.

[Str] R. Strichartz, Mean value properties of the Laplacian via spectral theory. Trans. Amer. Math. Soc. 284 (1984), no. 1, 219-228.

[Su] D. Sullivan, Related aspects of positivity in Riemannian geometry. J. Differential Geom. 25 (1987), no. 3, 327-351.

[TayII] M.E. Taylor, *Partial differential equations II.* Qualitative studies of linear equations. Second edition. Applied Mathematical Sciences, 116. Springer, New York, 2011.

[XBin] Bin Xu, Derivatives of the spectral function and Sobolev norms of eigenfunctions on a closed Riemannian manifold. Ann. Global Anal. Geom. 26 (2004), no. 3, 231- 252.

[Xu1] X. Xu, Gradient estimates for the eigenfunctions on compact manifolds with boundary and Hrmander multiplier theorem. Forum Math. 21 (2009), no. 3, 455-476.

[Xu2] X. Xu, Eigenfunction estimates for Neumann Laplacian and applications to multiplier problems. Proc. Amer. Math. Soc. 139 (2011), no. 10, 3583-3599.

[Z1] S. Zelditch, Real and complex zeros of Riemannian random waves. Spectral analysis in geometry and number theory, 321342, Contemp. Math., 484, Amer. Math. Soc., Providence, RI, 2009.

[Z2] S. Zelditch, Fine structure of Zoll spectra. J. Funct. Anal. 143 (1997), no. 2, 415-460.

CHAPTER 6

Hadamard parametrices on Riemannian manifolds

In this section we construct parametrices for the wave kernels of a general complete Riemannian manifold (M, g). Wave kernels are basic examples of Fourier integral operators, which are discussed in general in §7.

There are several methods to construct Fourier integral representations of wave kernels, the earliest and simplest of which being the Hadamard parametrix. In later sections we also consider the Lax-Hörmander parametrix, which is linear in the time variable.

The Hadamard parametrix is in some sense based on the close relation between the wave kernels and the spherical means operators. We bring out this relation only after discussing general Fourier integral operators in §7.

We largely follow [**Be, H, R, Ga**] in the construction of the Hadamard-Riesz parametrix.

6.1. Hadamard parametrix

The wave group of a Riemannian manifold is the unitary group $U(t) = e^{it\sqrt{-\Delta}}$. As above, we also write $E(t) = \cos t\sqrt{-\Delta}$ and $S(t) = \frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}$. We now review the construction of a Hadamard parametrix for $E(t)$ and $S(t)$. There is a similar parametrix for $U(t)$ but it is somewhat more complicated because $U(t)$ is not a function of Δ.

The basic ansatz ($n = \dim M$) is that

$$S(t, x, y) = \int_0^\infty e^{i\theta(r^2(x,y)-t^2)} \sum_{k=0}^\infty W_k(x, y)\theta^{\frac{n-3}{2}-k} d\theta, \qquad t < \text{inj}(M, g), \tag{6.1}$$

where

$$W_0(x, y) = \Theta^{-\frac{1}{2}}(x, y), \quad \Theta(x, y) \text{ the volume density in normal coordinates.} \tag{6.2}$$

The higher coefficients are determined by transport equations, and θ^r is regularized at 0 (see below). This formula is only valid for times $t < \text{inj}(M, g)$, but using the group property of $U(t)$ it determines the wave kernel for all times. It shows that for fixed (x, t) the kernel $S(t, x, y)$ is singular along the distance sphere $S_t(x)$ of radius t centered at x, with singularities propagating along geodesics. It only represents the singularity and in the analytic case only converges in a neighborhood of the characteristic conoid.

We recall that the even part $E(t)$ of the wave kernel, $\cos t\sqrt{-\Delta}$ which solves the initial value problem

$$\begin{cases} \Big(\dfrac{\partial}{\partial t}^2 - \Delta\Big)u = 0, \\ u|_{t=0} = f, \quad \dfrac{\partial}{\partial t}u|_{t=0} = 0. \end{cases} \tag{6.3}$$

Similarly, the odd part $S(t)$ of the wave kernel, $\frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}$, is the operator solving

$$\begin{cases} \Big(\dfrac{\partial}{\partial t}^2 - \Delta\Big)u = 0, \\ u|_{t=0} = 0, \quad \dfrac{\partial}{\partial t}u|_{t=0} = g. \end{cases} \tag{6.4}$$

These kernels only really involve Δ and may be constructed by the Hadamard-Riesz parametrix method. As above they have the form

$$\int_0^\infty e^{i\theta(r^2-t^2)} \sum_{j=0}^\infty W_j(x,y)\theta_{\text{reg}}^{\frac{n-1}{2}-j} d\theta \quad \text{modulo } C^\infty \text{ functions,} \tag{6.5}$$

where W_j are the Hadamard-Riesz coefficients determined inductively by the transport equations

$$\begin{cases} \dfrac{\Theta'}{2\Theta}W_0 + \dfrac{\partial W_0}{\partial r} = 0, \\ 4ir(x,y)\Big\{\Big(\dfrac{k+1}{r(x,y)} + \dfrac{\Theta'}{2\Theta}\Big)W_{k+1} + \dfrac{\partial W_{k+1}}{\partial r}\Big\} = \Delta_y W_k. \end{cases} \tag{6.6}$$

Here, $r(x,y)$ denotes the distance between $x, y \in M$. The solutions are given by:

$$\begin{cases} W_0(x,y) = \Theta^{-\frac{1}{2}}(x,y), \\ W_{j+1}(x,y) = \Theta^{-\frac{1}{2}}(x,y) \displaystyle\int_0^1 s^k \Theta(x,x_s)^{\frac{1}{2}} \Delta_2 W_j(x,x_s)\, ds, \end{cases} \tag{6.7}$$

where x_s is the geodesic from x to y parametrized proportionately to arc-length and where Δ_2 operates in the second variable.

Going back to the integral (6.5), we use a formula from [**GeSh**, p. 171]:

$$\int_0^\infty e^{i\theta\sigma}\theta_+^\lambda\, d\theta = ie^{i\lambda\pi/2}\Gamma(\lambda+1)(\sigma+i0)^{-\lambda-1}. \tag{6.8}$$

We recall that $(t+i0)^{-n} = e^{-i\pi\frac{n}{2}}\frac{1}{\Gamma(n)}\int_0^\infty e^{itx}x^{n-1}dx$. Therefore,

$$\int_0^\infty e^{i\theta(r^2-t^2)}\theta_+^{\frac{n-3}{2}-j}\, d\theta = ie^{i(\frac{n-1}{2}-j)\pi/2}\Gamma\Big(\frac{n-3}{2}-j+1\Big)(r^2-t^2+i0)^{j-\frac{n-3}{2}-2}. \tag{6.9}$$

Here there is apparently trouble when n is odd since $\Gamma(\frac{n-3}{2}-j+1)$ has poles at the negative integers. Indeed, using

(6.10)
$$\begin{cases}\Gamma(\alpha+1-k)=(-1)^{k+1+[\alpha]}\dfrac{\Gamma(\alpha+1-[\alpha])\Gamma([\alpha]+1-\alpha)}{\alpha+1}\dfrac{1}{\alpha-[\alpha]}\dfrac{1}{\Gamma(k-\alpha)},\\ \Gamma(z)\Gamma(1-z)=\dfrac{\pi}{\sin\pi z},\\ (x+i0)^\lambda=\begin{cases}e^{i\pi\lambda}|x|^\lambda & x<0\\ x_+^\lambda & x>0\end{cases}\quad\text{is entire,}\end{cases}$$

we see that the imaginary part cancels the singularity of $\frac{1}{\alpha-[\alpha]}$ as $\alpha\to\frac{n-3}{2}$ when $n=2m+1$. There is no singularity in even dimensions. In odd dimensions the real part is $\cos\pi\lambda x_-^\lambda+x_+^\lambda$ and we always seem to have a pole in each term!

But in any dimension, the imaginary part is well-defined and we have

$$\frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}(x,y)=C_o\,\mathrm{sgn}(t)\sum_{j=0}^\infty(-1)^jw_j(x,y)\frac{(r^2-t^2)_-^{j-\frac{n-3}{2}-1}}{4^j\Gamma(j-\frac{n-3}{2})}\mod C^\infty. \tag{6.11}$$

By taking the time derivative we also have,

$$\cos t\sqrt{-\Delta}(x,y)=C_o|t|\sum_{j=0}^\infty(-1)^jw_j(x,y)\frac{(r^2-t^2)_-^{j-\frac{n-3}{2}-2}}{4^j\Gamma(j-\frac{n-3}{2}-1)}\mod C^\infty. \tag{6.12}$$

where C_o is a universal constant and where $W_j=\tilde{C}_oe^{-ij\frac{\pi}{2}}4^{-j}w_j(x,y)$.

6.2. Hadamard-Riesz parametrix

We try to construct the kernel as a homogeneous oscillatory integral

$$E(t,x,y)=\int_0^\infty e^{i\theta(r^2-t^2)}A(t,x,y,\theta)\,d\theta, \tag{6.13}$$

where A is a polyhomogeneous symbol in θ,

$$A(t,x,y,\theta)\sim\sum_{j=0}^\infty W_j(t,x,y)\theta_+^{\frac{n-1}{2}-j}d\theta\mod C^\infty. \tag{6.14}$$

Here, θ_+^s is the homogeneous distribution with singularity at $\theta=0$ regularized as in [**Ho1**, Chapter 3.2] or in [**Be**]. The leading term $\theta_+^{\frac{n-1}{2}}$ of the amplitude has the correct power for $\cos t\sqrt{-\Delta}$. It should be $\theta_+^{\frac{n-3}{2}}$ for $\frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}$.

Applying the Fourier transform formula for $\mathcal{F}\theta_+^s$ [**Ho1**, p. 167] gives

(6.15)
$$\int_0^\infty e^{i\theta(r^2-t^2)}\theta_+^{\frac{n-3}{2}-j}d\theta=ie^{i(\frac{n-1}{2}-j)\pi/2}\Gamma\Big(\frac{n-3}{2}-j+1\Big)(r^2-t^2+i0)^{j-\frac{n-3}{2}-2}.$$

When n is odd, $\Gamma(\frac{n-3}{2}-j+1)$ has poles at the negative integers. Thus, this parametrix does not quite work on odd dimensional spaces (= even dimensional space-times). But the correct formulae may be obtained by analytic continuation (cf. [**R, Be**]). Riesz defined a holomorphic family of Riesz kernels $(t^2-r^2)_+^\alpha$ and used analytic continuation to define the value when α is a negative integer. He only studied the imaginary part, where there is no pole. Hadamard used a different regularization procedure (discussed below). In the end,

$$\frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}(x,y) = C_o \operatorname{sgn}(t) \sum_{j=0}^{\infty} (-1)^j W_j(x,y) \frac{(r^2-t^2)_-^{j-\frac{n-3}{2}-1}}{4^j \Gamma(j-\frac{n-3}{2})} \mod C^\infty. \tag{6.16}$$

Here, $\operatorname{sgn}(x) = \frac{x}{|x|}$ for $x \neq 0$ and $= 0$ for $x = 0$.

By taking the time derivative we also have,

$$\cos t\sqrt{-\Delta}(x,y) = C_o|t| \sum_{j=0}^{\infty} (-1)^j W_j(x,y) \frac{(r^2-t^2)_-^{j-\frac{n-3}{2}-2}}{4^j \Gamma(j-\frac{n-3}{2}-1)} \mod C^\infty, \tag{6.17}$$

where C_o is a universal constant and where the Hadamard-Riesz coefficients $W_j(x,y)$ are determined inductively by the transport equations (6.6), whose solutions are given by (6.7).

For $U(t) = \exp it\sqrt{\Delta}$ one may apply $\sqrt{\Delta}$ to the parametrix for $\frac{\sin t\sqrt{\Delta}}{\sqrt{\Delta}}$, resulting in an oscillatory with the same phase and a different amplitude. One may then use Duhamel's formula to construct the exact solution as a Volterra series,

$$U(t,x,y) = U_N(t,x,y) + \int_0^t U_N(t-s)(\partial_t^2 - \Delta)U_N(t-s)ds + \cdots, \tag{6.18}$$

where U_N is an approximate solution obtained by using N terms of a series above.

6.3. The Hadamard-Feynman fundamental solution and Hadamard's parametrix

In his seminal work [**H**], Hadamard constructed a solution of $\Box E = 0$ for $t > 0$ which has the singularity $\Gamma^{\frac{-m+2}{2}}$, $m = n+1 = \dim M \times \mathbb{R}$, where

$$\Gamma = t^2 - r^2 \tag{6.19}$$

in Hadamard's notation. Note the analogy to the elliptic case where the Green's function (the kernel of Δ^{-1}) has the singularity r^{-n+2} if $n > 2$.

The fundamental solution is more complicated in even space-time dimensions (i.e., odd space dimensions). Hadamard found the general solution as follows:

- The elementary solution in odd space-time dimensions has the form

$$U\Gamma^{-\frac{m-2}{2}} \text{ where } U = U_0 + \Gamma U_1 + \cdots + \Gamma^h U_h + \cdots \text{ is holomorphic.} \tag{6.20}$$

  This U is not the half-wave propagator!
- The elementary solution in even space-time dimensions has the form

$$U\Gamma^{-\frac{m-2}{2}} + V\log\Gamma + W, \tag{6.21}$$

  where

$$U = \sum_{j=0}^{m-1} U_j \Gamma^j, \quad V = \sum_{j=0}^{\infty} V_j \Gamma^j, \quad W = \sum_{j=1}^{\infty} W_j \Gamma^j. \tag{6.22}$$

Hadamard's formulae for the fundamental solutions pre-date the Schwartz theory of distributions. We follow his approach of describing the fundamental solutions as branched meromorphic functions (possibly logarithmically branched) on complexified space-time. In modern terms Γ^α (resp. $\log \Gamma$) would be defined as the distributions $(\Gamma + i0)^\alpha$ (resp. $\log(\Gamma + i0)$) as in the constant curvature cases. Hadamard implicitly worked in the complexified setting. For background on $\log(x + i0)$, see [**GeSh**, Chapter III.4.4] or [**Ho1**].

THEOREM 6.1 (Hadamard, 1920). *With the U_j, V_k, W_ℓ defined as above,*

- *In odd space-time dimensions, there exists a formal series U as above so that $E = U\Gamma^{\frac{2-m}{2}}$ solves $\Box E = \delta_0(t)\delta_y(x)$. If (M, g) is real analytic, the series $U = \sum_{j=0}^\infty U_j \Gamma^j$ converges absolutely for $|\Gamma| < \varepsilon$ sufficient small, i.e. near the characteristic conoid and admits a holomorphic continuation to a complex neighborhood of $\mathcal{C}_{\mathbb{C}}$.*
- *In even space-time dimensions, $E = U\Gamma^{\frac{2-m}{2}} + V \log \Gamma + W$ solves $\Box E = \delta_0(t)\delta_y(x)$. If (M, g) is real analytic, all of the series for U, V, W converge for $|\Gamma|$ small enough and admit analytic continuations to a neighborhood of $\mathcal{C}_{\mathbb{C}}$*

In the smooth case, the series do not converge. But if they are truncated at some j_0, the partial sum defines a parametrix, i.e. a fundamental solution modulo functions in C^{j_0}. By the Levi sums method (Duhamel principle) the parametrix differs from a true fundamental solution by a C^{j_0} kernel. We are mainly interested in real analytic (M, g) in this article and do not go into details on the last point. We note that the singularities of the kernel are due to the factors $\Gamma^{\frac{2-m}{2}}, \log \Gamma$, which are branched meromorphic (and logarithmic) kernels. The terms are explicitly evaluated in the case of hyperbolic quotients in [**JL**]. We also refer to Chapter 5.2 of [**Ga**] for a somewhat modern presentation of the proof.

It may be of interest to note that this construction only occupies a third of Hadamard's book [**H**]. The rest is devoted to the use of such kernels to solve the Cauchy problem, using Green's formula applied to a domain obtained by intersecting the backward characteristic conoid from a point (t, x) of space-time with the Cauchy hypersurface. The integrals over the light-like (null) part of the boundary caused serious trouble since the factors $\Gamma^{\frac{2-m}{2}}$ are infinite along them and need to be re-normalized. This was the origin of Hadamard's finite parts of divergent integrals. Riesz used analytic continuation methods instead to define the forward fundamental solution in [**R**].

6.4. Sketch of proof of Hadamard's construction

Let $\Theta = \sqrt{\det(g_{jk})}$ be the volume density in normal coordinates based at y, $dV = \Theta(y, x)dx$. That is,

$$\Theta(x, y) = \left| \det D_{\exp_x^{-1}(y)} \exp_x \right|.$$

Fix $x \in M$ and endow $B_\varepsilon(x)$ with geodesic polar coordinates r, θ. That is, use the chart $\exp_x^{-1} : B_r(x) \to B_{x,r}^* M$ combined with polar coordinates on $T_x^* M$. Then $g^{11} = 1, g^{1j} = 0$ for $j = 2, \ldots, n$. Also, $dV = \Theta(x, y)\, dy = \Theta(x, r, \theta) r^{n-1}\, dr d\theta$. So the volume density J relative to Lebesgue measure $dr d\theta$ in polar coordinates is given by $J = r^{n-1}\Theta$.

In these coordinates,

$$\Delta = \frac{1}{J}\sum_{j,k=1}^{n}\frac{\partial}{\partial x_j}\left(Jg^{jk}\frac{\partial}{\partial x_k}.\right) = \frac{\partial^2}{\partial r^2} + \frac{J'}{J}\frac{\partial}{\partial r} + L, \tag{6.23}$$

where L involves no $\frac{\partial}{\partial r}$ derivatives. Equivalently,

$$\Delta = \frac{\partial^2}{\partial r^2} + (\frac{\Theta'}{\Theta} + \frac{n-1}{r})\frac{\partial}{\partial r} + L. \tag{6.24}$$

The first step in the parametrix construction is to find the phase function. Hadamard chooses to use (6.19). In the Lorentzian metric, Γ satisfies

$$\nabla\Gamma\cdot\nabla\Gamma = 4\Gamma. \tag{6.25}$$

This is not the standard eikonal equation $\sigma_{\Box}(d\varphi) = 0$ of geometric optics, but rather has the form

$$\sigma_{\Box}(d\Gamma) = 4\Gamma. \tag{6.26}$$

But Γ is a good phase, since the Lagrangian submanifold $\{(t, d_t\Gamma, x, d_x\Gamma, y, -d_y\Gamma)\}$ is the graph of the bi-characteristic flow. This is because the $d_x r(x,y)$ is the unit vector pointing along the geodesic joining x to y and $d_y r(x,y)$ is the unit vector pointing along the geodesic pointing from y to x.

To proceed, we introduce the simplifying notation

$$M = \Box\Gamma = -4 - 2r\frac{(n-1)}{r} - 2r\frac{\Theta_r}{\Theta} = 2m + 2r\frac{\Theta_r}{\Theta}. \tag{6.27}$$

We then have

(6.28)

$$\Box\left[f(\Gamma)U_j\right] = \Box\left[f(\Gamma)\right]U_j + 2\nabla\left[f(\Gamma)\right]\nabla U_j + f(\Gamma)\Box U_j$$

$$= \left(f''(\Gamma)\nabla(\Gamma)\cdot\nabla(\Gamma) + f'(\Gamma)\Box(\Gamma)\right)U_j + 2f'(\Gamma)\nabla\Gamma\cdot\nabla U_j + f(\Gamma)\Box U_j. \tag{6.29}$$

In addition to (6.25), we have

$$\begin{cases} \Box\Gamma = 4 + \dfrac{J_r}{J}2r \\ \nabla\Gamma\cdot\nabla = \nabla(t^2 - r^2)\cdot\nabla = 2\Big(t\dfrac{\partial}{\partial t} + r\dfrac{\partial}{\partial r}\Big) = 2s\dfrac{d}{ds}, \end{cases} \tag{6.30}$$

where we recall that that we are using the Lorentz metric of signature $+---$. Here $s^2 = \Gamma$, and the notation $s\frac{d}{ds}$ refers to differentiation along a space-time geodesic.

We then have

(6.31)

$$\Box\left[f(\Gamma)U_j\right] = \left(f''(\Gamma)(4\Gamma) + f'(\Gamma)\Big(4 + \frac{J_r}{J}2r\Big)\right)U_j + 2f'(\Gamma)\Big(-2s\frac{d}{ds}U_j\Big) + f(\Gamma)\Box U_j.$$

We now apply the equation above with $f = x^{\frac{2-m}{2}+j}$ (and later to $f = \log x$), in which case

$$f' = \Big(\frac{2-m}{2}+j\Big)x^{\frac{2-m}{2}+j-1}, \quad f'' = \Big(\frac{2-m}{2}+j\Big)\Big(\frac{2-m}{2}+j-1\Big)x^{\frac{2-m}{2}+j-2}. \tag{6.32}$$

We then attempt to solve

$$\Box\Big(\Gamma^{\frac{2-m}{2}}\sum_{j=0}^{\infty}U_j\Gamma^j\Big) = 0 \tag{6.33}$$

away from the characteristic conoid by setting the coefficient of each power $\Gamma^{\frac{2-m}{2}+j-1}$ of Γ equal to zero. The resulting 'transport equation' is

$$\begin{aligned}(6.34)\quad 0 = \Bigg\{-4\bigg(\Big(\frac{2-m}{2}+j\Big)\Big(\frac{2-m}{2}+j-1\Big) + \Big(\frac{2-m}{2}+j\Big)\Big(-4-\frac{J_r}{J}2r\Big)\bigg)\\ + 2\Big(\frac{2-m}{2}+j\Big)\Big(-2s\frac{d}{ds}\Big)\Bigg\} U_j + \Box U_{j-1}.\end{aligned}$$

They are impossible to solve for all j when m is even because the common factor $(\frac{2-m}{2}+j)$ vanishes when $j=\frac{m-2}{2}$. We thus first assume that m is odd so that it is non-zero for all j. We then recursively solve Hadamard's transport equations in even space dimensions:

$$(6.35)\qquad 4s\frac{dU_k}{ds} + \Big(M - 2m + 2r\frac{J_r}{J}\Big)U_k = -\Box U_{k-1}.$$

When $k=0$ we get

$$(6.36)\qquad 2s\frac{dU_0}{ds} + 2s\frac{\Theta_s}{\Theta} = 0,$$

which is solved by

$$(6.37)\qquad U_0 = \Theta^{-\frac{1}{2}}.$$

The solution of the ℓth transport equation is then,

$$(6.38)\qquad U_\ell = -\frac{U_0}{4\ell s^{m+\ell}}\int_0^s U_0^{-1} s^{\ell+m-1}\Box U_{\ell-1}\, ds.$$

Hence we have a formal solution with the singularity of the Green's function in the elliptic case, and by comparison with the Euclidean case we see that it solves $\Box E = \delta_0$.

We now consider the necessary modifications in the case of even dimensional space-times. In this case, $\Gamma^{\frac{2-m}{2}}\Gamma^j$ is always an integer power. If we could solve the transport equation for $j=\frac{m-2}{2}$, the resulting term would be regular with power Γ^0. The problem is that Γ^0 should actually be a term with a logarithmic singularity $\log\Gamma$.

Thus the parametrix (6.33) is inadequate in even space-time dimensions. Hadamard therefore introduced a logarithmic term $V\log(\Gamma)$. By a similar calculation to the above,

(6.39)

$$\Box\,[(\log\Gamma)V] = \Big(-\Gamma^{-2}(4\Gamma) + \Gamma^{-1}\Big(-4-\frac{J_r}{J}2r\Big)\Big)V + 2\Gamma^{-1}\Big(-2s\frac{d}{ds}V\Big) + \log\Gamma\Box V.$$

Due to (6.25), all terms except the logarithmic term have the same singularity Γ^{-1}. On the other hand, the only way to eliminate the logarithmic term is to insist that $\Box V = 0$. We further assume that

$$V = \sum_{j=0}^{\infty} V_j \Gamma^j.$$

We then return to the unsolvable transport equations for U_j for $j \geq \frac{m-2}{2}$, which now acquires the new V_0 term to become:

(6.40)
$$0 = \Big\{ -4\left((\frac{2-m}{2}+j)(\frac{2-m}{2}+j-1) \; + (\frac{2-m}{2}+j)(-4-\frac{J_r}{J}2r) \right)$$

(6.41)
$$+ 2(\frac{2-m}{2}+j)(-2s\frac{d}{ds}) \Big\} U_j + \Box U_{j-1}$$

(6.42)
$$+ \Gamma^{-1}\left(4 + (-4-\frac{J_r}{J}2r) + 2\frac{d}{ds}\right) V_0.$$

When $j = \frac{m-2}{2}$, everything cancels in the Γ^{-1} term except $\Box U_{m-1}$. Hence, we drop the U_j for $j \geq \frac{m-2}{2}$ and assume the non logarithmic part is just the finite sum $\sum_{j=0}^{m-1} U_j \Gamma^j$. But adding in the V_0 term we get the transport equation

(6.43)
$$-4s\frac{dV_0}{ds} - 2r\frac{J_r}{J}V_0 = -\Box U_{m-1}.$$

Here, U_{m-1} is known and we solve for V_0 to get

(6.44)
$$V_0 = -\frac{U_0}{4s^m}\int_0^s U_0^{-1} s^{m-1} \Box U_{m-1}\, ds.$$

The condition $\Box V = 0$ imposed above then determines the rest of the coefficients V_j,

(6.45)
$$V_\ell = -\frac{U_0}{4\ell s^{m+\ell}}\int_0^s U_0^{-1} s^{\ell+m-1} \Box V_{\ell-1}\, ds.$$

We now have two equations: the original $\Box(U\Gamma^{\frac{2-m}{2}}U + V\log\Gamma) = 0$ and the new $\Box V = 0$. By solving the transport equations for $U_0, \ldots, U_{m-1}, V_0, V_j (j \geq 1)$ we obtain a solution of an inhomogeneous equation of the form,

(6.46)
$$\Box(U\Gamma^{\frac{2-m}{2}} + V\log\Gamma) = \sum_{j=0} w_j \Gamma^j,$$

where the right side is regular. To complete the construction, we add a new term of the form $W = \sum_{\ell=1}^{\infty} W_\ell (r^2-t^2)^\ell$ in order to ensure that

(6.47)
$$\Box\Big(\sum_{j=0}^{m-1} U_j (r^2-t^2)^{-m+j} + V\log(r^2-t^2) + W\Big) = 0$$

away from the characteristic conoid. It then suffices to find W_j so that

(6.48)
$$\Box \sum_{j=1}^{\infty} W_j \Gamma^j = \sum_{j=0} w_j \Gamma^j.$$

This leads to more transport equations which are always solvable (by the Cauchy-Kovalevskaya theorem). This concludes the sketch of the proof of Theorem 6.1.

6.5. Convergence in the real analytic case

The above parametrix construction was formal. However, when the metric is real analytic, Hadamard proved that the formal series converges for $|t|$ and $|\Gamma|$ sufficiently small.

The convergence proof based on the method of majorants.

THEOREM 6.2 ([**H**]; see also [**Ga**]). *Assume that (M, g) is real analytic. Then there exists $K > 0$ so that the Hadamard parametrix converges for any (t, y) such that $t \neq 0$, $r(x, y) < \varepsilon = \mathrm{inj}(x_0)$ and*

$$|t^2 - r^2| \leq \frac{\left((1 - \frac{\|y\|}{\varepsilon}\right)^2}{\left(1 + \frac{m_1}{\varepsilon} + \frac{m_1^2}{\varepsilon^2}\right) K}, \qquad m_1 = \frac{m-2}{2}. \tag{6.49}$$

It follows that the Hadamard fundamental solutions holomorphically extend to a neighborhood of $\mathcal{C}_{\mathbb{C}}$ as branched meromorphic functions with $\mathcal{C}_{\mathbb{C}}$ as branch locus. To obtain single valued distributions, one then needs to restrict the kernels to regions where a unique branch can be defined.

6.6. Away from $\mathcal{C}_{\mathbb{R}}$

A further complication is that the fundamental solution has only been constructed in a neighborhood of $\mathcal{C}_{\mathbb{R}}$. But it is known to be real analytic on $(\mathbb{R} \times M \times M) \backslash \mathcal{C}_{\mathbb{R}}$ and that it extends to a holomorphic kernel away from a neighborhood of $\mathcal{C}_{\mathbb{C}}$. To prove this it suffices to note that the analytic wave front set of the fundamental solution is $\mathcal{C}_{\mathbb{R}}$. A more detailed proof is given by Mizohata [**M1, M2**] for 'elementary solutions', i.e. solutions $E(t, x, y)$ of $\Box E = 0$ such as $\cos t\sqrt{-\Delta}$ or $\frac{\sin t\sqrt{-\Delta}}{\sqrt{-\Delta}}$ whose Cauchy data is either zero or a delta function. Here, $\Box$ operates in the x variable with y as a parameter. To analyze the wave kernels away from the characteristic conoid, Mizohata makes the decomposition

$$E(t, x, y) = E_N(t, x, y) + w_N(t, x, y) + z_N(t, x, y), \tag{6.50}$$

where

- $\Box E_N = f_N$ with $f_N \in C^{N-1}(\mathbb{R} \times M \times M)$;
- $E_N(t, x, y)|_{t=0} = \delta_y + a(x, y)$ with $a(x, y) \in C^{\omega}(M \times M)$ (in fact it is independent of y);
- $\Box w_N(t, x, y) = -f_N(t, x, y)$, with $w(0, x, y) = 0$;
- $\Box z_N = 0, z_N(0, x, y) = -a(x, y)$.

The same kind of decomposition applies to the Hadamard fundamental solution. The term constructed by the parametrix method is E_N. By solving the above equations, it is shown in [**M1**] that the sum is analytic away from $\mathcal{C}_{\mathbb{R}}$. One can see that the Hadamard method is only a branched Laurent type expansion near $\mathcal{C}_{\mathbb{R}}$ by considering kernels for spaces of constant curvature.

6.7. Hadamard parametrix on a manifold without conjugate points

The wave kernels $\cos(t\sqrt{-\Delta})(x, y)$ and $\frac{\sin(t\sqrt{-\Delta})}{\sqrt{-\Delta}}$ can be constructed globally in time on a Riemannian manifold (M, g) without conjugate points, such as a non-positively curved manifold. We refer to §2 for the geometric notions and notations. We denote the universal Riemannian cover of (M, g) by $(\tilde{M}, \tilde{g})$. By definition, there is a unique geodesic (unit speed) between any two points (x, y) of $\tilde{M}$ and the geodesic distance function (squared) is a global smooth function $r^2(x, y)$.

On $\tilde{M}$, the wave operator $\tilde{E}$ can be globally constructed (modulo $C^{\infty}(\mathbb{R} \times M \times M)$) by the Hadamard-Riesz parametrix method (see [**Be**]). That is, the wave kernel $\tilde{E}(t,x,y) = \cos(t\sqrt{-\Delta})$ is given modulo smooth kernels by the Hadamard parametrix,

$$\tilde{E}(t,x,y) \equiv \int_0^\infty e^{i\theta(r^2-t^2)} \sum_{j=0}^\infty W_j(x,y)\theta^{\frac{n-1}{2}-j}\chi(\theta)\,d\theta, \tag{6.51}$$

where χ is (as above) a smooth cutoff near 0 and where the W_j are given recursively by the formulae in (6.7). Note that r^2 and $\Theta^{-\frac{1}{2}}$ are smooth for a metric WCP.

The wave kernel $E(t,x,y)$ on M is obtained by projecting this kernel from $\tilde{M}$, i.e., by summing over the deck transformation group:

$$E(t,x,y) = \sum_{\gamma\in\Gamma} \tilde{E}(t,x,\gamma\cdot y) \tag{6.52}$$

$$\equiv \sum_{\gamma\in\Gamma} \int_0^\infty e^{i\theta(r^2(x,\gamma y)-t^2)} \sum_{j=0}^\infty W_j(x,\gamma y)\theta^{\frac{n-1}{2}-j}\chi(\theta)\,d\theta. \tag{6.53}$$

6.8. Dimension 3

In dimension 3, the Hadamard-Riesz parametrix is relatively elementary, and its relation to spherical means is simpler. The Hadamard parametrix is constructed for the cosine propagator in dimension in [**Don**] and we follow its exposition in the section. The sine-propagator $S(t)$ is one degree smoother but the calculations are equivalent since $S'(t) = C(t)$.

Let $C(t,x,q)$ be the cosine propagator. For each k we construct a parametrix in the first sense so that

$$\begin{cases} \Box C^k \in C^{k-1}(\mathbb{R}\times M), \\ \\ C^k(0,x,q) - \delta_q \in C^{k-1}(M). \end{cases}$$

We then do the same for the sine propagator $S(t,x,q)$.

We follow [**Don**] and start by relating the notation there to the one in the Hadamard parametrix method. Donnelly writes $g = \det(g_{ij})$ in normal coordinates based at x. We denote the same quantity by $\Theta(x,y)$ so that $dV(y) = \Theta(x,y)dy = \sqrt{g(y)}dy$. We then change to geodesic polar coordinates (ρ,ω) so that $dy = \rho^{n-1}d\rho d\omega$. We write $\Theta(x,\rho,\omega)$ for $\Theta(x,\cdot)$ in polar coordinates. Thus, $dV(y) = \Theta\rho^{n-1}d\rho d\omega$.

Let $\Theta = \sqrt{\det(g_{jk})}$ be the volume density in normal coordinates based at y, $dV = \Theta(y,x)dx$. That is,

$$\Theta(x,y) = \left|\det D_{\exp_x^{-1}(y)} \exp_x\right|.$$

Fix $x \in M$ and endow $B_\varepsilon(x)$ with geodesic polar coordinates r,θ. That is, use the chart $\exp_x^{-1} : B_r(x) \to B^*_{x,r}M$ combined with polar coordinates on T^*_xM. Then $g^{11} = 1, g^{1j} = 0$ for $j = 2,\dots,n$. Also, $dV = \Theta(x,y)dy = \Theta(x,r,\theta)r^{n-1}drd\theta$. So the volume density J relative to Lebesgue measure $drd\theta$ in polar coordinates is given by $J = r^{n-1}\Theta$.

In these coordinates,

$$\Delta = \frac{1}{J}\sum_{j,k=1}^n \frac{\partial}{\partial x_j}\left(Jg^{jk}\frac{\partial}{\partial x_k}.\right) = \frac{\partial^2}{\partial r^2} + \frac{J'}{J}\frac{\partial}{\partial r} + L, \tag{6.54}$$

where L involves no $\frac{\partial}{\partial r}$ derivatives. Equivalently,

$$\Delta = \frac{\partial^2}{\partial r^2} + (\frac{\Theta'}{\Theta} + \frac{n-1}{r})\frac{\partial}{\partial r} + L,$$

PROPOSITION 6.3. *For general metrics on Riemannian manifolds of dimension 3, for any k and for $|t| < \mathrm{inj(M, g)}$, there exists a kth order parametrix of the form,*

(6.55)
$$\begin{aligned} C^k(t,x,q) &= a_{-2}(x,q)\delta'(\rho - t) + a_{-1}(x,q)\delta(\rho - t) + a_0(x,q)H^0(\rho - t) \\ &\quad + \cdots + a_k(x,q)H^k(\rho - t). \end{aligned}$$

Here, $H^k(s) = \frac{1}{j!}s_+^j$ and the a_j are constructed so that as $\rho \to 0$,

$$\begin{cases} a_0 = O(1), \\ a_0 + \rho a_1 = O(\rho), \\ \dots \\ a_0 + \rho a_1 + \cdots + a_k\rho^k = O(\rho^k). \end{cases}$$

REMARK 6.4. On $\mathbb{R}^3$ the sine propagator was $\frac{\delta(\rho-t)}{\rho}$ and its t-derivative is the cosine propagator $\frac{\delta'(\rho-t)}{\rho}$. The factor of $\frac{1}{\rho}$ will be absorbed into the amplitudes a_j.

PROOF. Suppose that $f(\rho)$ is a function depending only on ρ. In view of (6.54), we have for $n = 3$,

$$\Delta(f(\rho)\alpha) = \left(f''(\rho) + \frac{J_\rho}{J}f'(\rho)\right)\alpha + 2f'(\rho)\frac{\partial\alpha}{\partial\rho} + f(\rho)\Delta\alpha. \tag{6.56}$$

Define the transport operator

$$\mathcal{R} := J^{-\frac{1}{2}}\frac{\partial}{\partial\rho}J^{\frac{1}{2}}.$$

A straightforward caclulation based on (6.56) gives

$$\begin{aligned} (\tfrac{\partial^2}{\partial t^2} - \Delta)C^k(t,x,q) &= (-2\mathcal{R}a_{-2})\,\delta''(\rho - t) \\ &+ (-2\mathcal{R}a_{-1} - \Delta a_{-2})\,\delta'(\rho - t) \\ &\quad + \cdots + (-\mathcal{R}a_{j-1})\,H^{j-1}(\rho - t) \\ &+ \cdots(-\Delta a_k)H^k(\rho - t). \end{aligned}$$

Note that the coefficient of δ''' cancels due to the argument $\rho - t$. To make the coefficient of δ'' zero, we need to solve

$$\mathcal{R}\,a_{-2} = 0 \iff \frac{\partial}{\partial\rho}J^{\frac{1}{2}}a_{-2} = 0 \iff J^{\frac{1}{2}}a_{-2} = \mathrm{Const}.$$

It follows that

$$a_{-2} = C\frac{\Theta^{\frac{1}{2}}}{\rho}.$$

To obtain the desired initial condition, one makes the constant the same as in the Euclidean case, i.e. $C = 4\pi$. Thus, a_{-2} has a similar form to the Euclidean case where $\Theta \equiv 1$.

The coefficient of $\delta(\rho - t)$ equals

$$-2\mathcal{R}a_{-1} + \Delta a_{-2},$$

and to make it zero we need to define a_{-1} so that

$$\mathcal{R}a_{-1} = -\Delta a_{-2}.$$

Then,

$$a_{-1} = CJ^{-\frac{1}{2}} - \frac{1}{2}J^{-\frac{1}{2}} \int_0^\rho J^{\frac{1}{2}} \Delta J^{-\frac{1}{2}} ds.$$

One must have $C = 0$ if a_{-1} is smooth. Adjusting for constants,

$$a_{-1} = \frac{1}{8\pi} \frac{1}{J^{\frac{1}{2}}} \int_0^\rho J^{\frac{1}{2}} \Delta J^{-\frac{1}{2}} ds.$$

This may be simplified using the special identity in dimension 3,

$$J^{\frac{1}{2}} \Delta J^{-\frac{1}{2}} = \Theta^{-1} \Delta \Theta^{-\frac{1}{2}},$$

to give,

$$a_{-1} = J^{-\frac{1}{2}} \int_0^\rho \Theta^{\frac{1}{2}} \Delta \Theta^{-\frac{1}{2}} ds.$$

In general one has transport equations

$$\mathcal{R}a_j = -\tfrac{1}{2}\Delta a_{j-1},$$

which are solved iteratively as in the case $j = -1$. If we solve the first k transport equations we get

$$\Box C^k = -(\Delta a_k)H^k \in C^{k-1}(\mathbb{R} \times M).$$

This uniquely determines the a_k and implies that

$$C^k(0, x, q) - \delta_q \in C^{k-1}(M).$$

As in [**Don**, Theorem 2.4], we claim:

THEOREM 6.5. $C(t, x, q) - C^k(t, x, q) \in C^{k-1}(\mathbb{R} \times M)$.

PROOF. One has,

$$\begin{cases} \Box(C - C^k) \in C^{k-1}(\mathbb{R} \times M), \\ C(0, x, q) - C^k(0, x, q) \in C^{k-1}(\mathbb{R} \times M), \\ \frac{\partial}{\partial t}(C(t, x, q) - C^k(t, x, q))|_{t=0} = 0. \end{cases}$$

The last equation holds if we extend the solution to be even in t from $t > 0$. Further, by the recursive procedure,

$$C^{k+1} - C^k \in C^k(\mathbb{R} \times M).$$

To prove that $C - C^k \in C^{k-1}$ we use a form of Duhamel's principal for second order equations and the fact that the wave propagator is unitary on Sobolev spaces. It is sufficient to cite [**Ho3**, Lemma 15.5.4]. Let $E(t) = \int_M (|v(x,t)|^2 + |\nabla v(x,t)|^2) dV$.

LEMMA 6.6. *Let $v \in C^\infty([0,T] \times M)$ be the solution of the inhomogeneous initial value problem on $[0,T] \times M$:*

$$\begin{cases} \Box v = h, \\ \\ v(x,0) = v_t(x,0) = 0. \end{cases}$$

Then,

$$E(t) \leq C(\int_0^t \|h(s,x)\| ds.$$

First,

$$\langle \Delta \dot{v}, v \rangle_{L^2} = \langle \dot{v}, \Delta v \rangle_{L^2}.$$

Hence,

$$2\langle h, \dot{v} \rangle_{L^2(M)} = \frac{\partial}{\partial t} E(t).$$

Let

$$M^2 := \sup_{0 \leq s \leq t} \|\dot{v}(s\cdot)\|_{L^2}^2 + \|\nabla v(s,\cdot)\|_{L^2}^2 / C_1 = \sup_{0 \leq s \leq t.} E(s).$$

Therefore,

$$E(t) \leq 2M \int_0^t \|h(s,\cdot)\| ds.$$

Thus,

$$M^2 \leq 2M \int_0^t \|h(s,\cdot)\| ds,$$

proving the Lemma.

□

One can iterate the argument to obtain estimates on higher derivatives of v in terms of higher derivatives of h. The full estimate is [**Ho3**, (17.5.11)]:

$$\sum_{j=0}^{k+1} \|D_t^{K+1-j} v(t,\cdot)\|_{(j)} \leq C_k \left(\int_0^t \|D_s^k h(s,\cdot)\| ds + \sum_{j=0}^{k-1} \|D_t^{k-1-j} h(t,,\cdot)\|_{(j)} \right).$$

It is proved inductively using $\ddot{v} = h - \Delta v$ which equals 0 when $t = 0$. This concludes the proof of Proposition 6.3. □

6.8.1. Sine kernel. The same parametrix construction works for the sine propagator $S(t)$. In Euclidean space it equals $\frac{\delta(\rho - t)}{\rho}$. The factor $\frac{1}{\rho} = a_{-2}$ and $\Theta = 1$.

PROPOSITION 6.7. *For general metrics on Riemannian manifolds of dimension 3, there exists a parametrix of the form,*

(6.57)

$$S^k(t,x,q) = a_{-2}(x,q)\delta(\rho - t) + a_{-1}(x,q)H^0(\rho - t) + a_0(x,q)H^1(\rho - t) + \cdots + a_k(x,q)H^{k-1}(\rho - t).$$

Here, $H^k(s) = \frac{1}{j!} s_+^j$.

One can go through the same steps or simply observe that $S(t) = \int_0^t C(s) ds$. Since the a_j are independent of t one simply integrates up the $H^k(s)$. Note that $\frac{d}{ds} H^k(s) = H^{k-1}(s)$.

6.9. Appendix on Homogeneous distributions

Let $D_t\varphi(x) = \varphi_t(x) = t\varphi(tx)$. A distribution on $\mathbb{R}\backslash\{0\}$ is called homogeneous of degree a if $\langle E, \varphi_t\rangle = t^{-a}\langle E, \varphi\rangle$. One uses the same definition in $\mathbb{R}^n\backslash\{0\}$ with $\varphi_t(x) = t^n\varphi(tx)$. If it extends to $\mathbb{R}^n$ as a distribution with the same property it is called homogeneous of degree a on $\mathbb{R}^n$. The following is from [**Ho1**].

6.9.1. x_+^a**.** For $\mathrm{Re}(a) > -1$ define

$$x_+^a = \begin{cases} x^a & x \geq 0, \\ 0 & x < 0. \end{cases} \tag{6.58}$$

We want to extend the definition to all $a \in \mathbb{C}$ so that

$$\frac{d}{dx}x_+^a = ax_+^{a-1} \quad \text{and} \quad xx_+^{a-1} = x_+^a. \tag{6.59}$$

There is a problem already at $a = 0$ since $\frac{d}{dx}x_+^0 = \delta_0(x)$.

We define

$$I_a(\varphi) = \int_0^\infty x^a\varphi(x)\,dx \tag{6.60}$$

so that

$$I_a(\varphi') = -aI_{a-1}(\varphi), \qquad \mathrm{Re}(a) > 0. \tag{6.61}$$

Then for $\mathrm{Re} a > -1$ and $k \in \mathbb{Z}_+$,

$$I_a(\varphi) = \frac{(-1)^k}{(a+k)\cdots(a+1)}I_{a+k}(\varphi^{(k)}). \tag{6.62}$$

This defines I_a as an analytic family of distributions for $\mathrm{Re} a > -k-1$ except for poles at $a = -1, \ldots, -k$. At $a = -k$ the residue is

$$\lim_{a\to -k}(a+k)I_a(\varphi) = \frac{(-1)^k}{(-1)\cdots(-k+1)}I_0(\varphi^{(k)}) = \frac{\varphi^{(k-1)}}{(k-1)!}. \tag{6.63}$$

Thus,

$$\lim_{a\to -k}(a+k)x_+^a = (-1)^k\frac{\delta_0^{(k-1)}}{(k-1)!}. \tag{6.64}$$

Thus one defines

$$x_+^{-k}(\varphi) = \frac{1}{(k-1)!}\int_0^\infty (\log x)\varphi^{(k)}(x)\,dx + \frac{\varphi^{(k-1)}(0)}{(k-1)!}\left(\sum_{j=1}^k \frac{1}{j}\right). \tag{6.65}$$

6.9.2. x_-^a**.** For $\mathrm{Re}(a) > -1$ define

$$x_+^a = \begin{cases} 0, & x \geq 0, \\ |x|^a & x < 0. \end{cases} \tag{6.66}$$

It is the reflection of x_+^a through the origin.

6.9.3. χ_+^a**.** Also, define

$$\chi_+^\alpha = \frac{x_+^\alpha}{\Gamma(\alpha+1)}. \tag{6.67}$$

This is a holomorphic family of homogeneous distributions and

$$\chi_+^{-k} = \delta_0^{(k-1)}. \tag{6.68}$$

6.9.4. $(x+i0)^a$**.** Define the function z^a on $\mathbb{C}\backslash\mathbb{R}$ defined by $e^{a\log z}$ where $\log z \in \mathbb{R}$, for $z \in \mathbb{R}_+$. Its boundary values are denoted $(x \pm i0)^a$. For $\mathrm{Re}(a) > 0$ one has

$$(x \pm i0)^a = x_+^a + e^{\pm i\pi a} x_-^a. \tag{6.69}$$

We also need that $(x+i0)^\lambda$ is entire and

$$(x+i0)^\lambda = \begin{cases} e^{i\pi\lambda}|x|^\lambda & x < 0 \\ x_+^\lambda & x > 0. \end{cases} \tag{6.70}$$

One has

$$(x\pi \pm i0)^{-k} = x_+^{-k} + (-1)^k x_-^{-k} \pm i\pi(-1)^k \frac{\delta_0^{(k-1)}}{(k-1)!}. \tag{6.71}$$

6.9.5. x^{-n}**.** Here and above t^{-n} is the distribution defined by $t^{-n} = \mathrm{Re}(t + i0)^{-n}$ (see [**Be**] or [**GeSh**, p. 52, 60].)

6.9.6. Fourier transforms of homogeneous distributions. According to [**GeSh**, p. 171],

$$\int_0^\infty e^{i\theta\sigma}\theta_+^\lambda d\lambda = ie^{i\lambda\pi/2}\Gamma(\lambda+1)(\sigma+i0)^{-\lambda-1}. \tag{6.72}$$

Also, $(t+i0)^{-n} = e^{-i\pi\frac{n}{2}}\frac{1}{\Gamma(n)}\int_0^\infty e^{itx}x^{n-1}dx$.

Bibliography

[Be] P. Bérard, On the wave equation without conjugate points, Math. Zeit. **155** (1977), 249–276.

[Don] Donnelly, H. On the wave equation asymptotics of a compact negatively curved surface. Invent. Math. 45 (1978), no. 2, 115-137.

[Du] J.J. Duistermaat, *Classical aspects of partial differential equations.* Communications of the Mathematical Institute, Rijksuniversiteit Utrecht, 13. Rijksuniversiteit Utrecht, Mathematical Institute, Utrecht, 1980.

[Ga] P. R. Garabedian, *Partial differential equations.* Reprint of the 1964 original. AMS Chelsea Publishing, Providence, RI, 1998.

[GeSh] I.M. Gel'fand and G.E. Shilov, *Generalized functions. Vol. I: Properties and operations.* Academic Press, New York-London 1964

[H] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations.* Dover Publications, New York, 1953.

[Ho1] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[Ho2] L. Hörmander, *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Classics in Mathematics. Springer-Verlag, Berlin, 2005.

[Ho3] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[Ho4] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[JL] J. Jorgenson and S. Lang, Analytic continuation and identities involving heat, Poisson, wave and Bessel kernels. Math. Nachr. 258 (2003), 44-70.

[Le] G. Lebeau, The complex Poisson kernel on a compact analytic Riemannian manifold, preprint (2013).

[M1] S. Mizohata, Analyticity of the fundamental solutions of hyperbolic systems. J. Math. Kyoto Univ. 1 (1961/1962), 327-355.

[M2] S. Mizohata, Solutions nulles et solutions non analytiques. J. Math. Kyoto Univ. 1 1961/1962 271-302.

[R] M. Riesz, L'intégrale de Riemann-Liouville et le problème de Cauchy. Acta Math. 81, (1949). 1–223.

[Su] T. Sunada, Trace formula and heat equation asymptotics for a nonpositively curved manifold. Amer. J. Math.104 (1982), no. 4, 795812.

[Z] S. Zelditch, Pluri-potential theory on Grauert tubes of real analytic Riemannian manifolds, I : AMS Proc. Symp. Pure Math. :Spectral Geometry, Alex H. Barnett, Carolyn S. Gordon, Peter A. Perry, and Alejandro Uribe, (eds), (2012), p. 299-339. (arXiv: arXiv:1107.0463).

CHAPTER 7

Lagrangian distributions and Fourier integral operators

7.1. Introduction

This section gives a brief introduction to Fourier integral operators and Lagrangian distributions, in both the homogeneous and semi-classical versions. We are mainly concerned with applications to parametrices for the wave group, but we also consider restriction operators to submanifolds and other Fourier integral operators that arise when studying eigenfunctions. Our aim is to present the theory in an informal and intuitive way, through basic examples and by sketching some of the original ideas and calculations due to Pauli, Van Vleck, Fock and other physicists more or less the way they understood them. Most of the 'modern' material in this section is drawn from [**Ho6, Ho2, Du1, Du2, DSj, GSj, GuSt1, SV**], where the reader can find much more detailed and systematic presentations.

Fourier integral operators are generalizations of pseudo-differential operators in which symplectic transformations are quantized as well as symbols. Fourier integral operators may thus be viewed as quantizations of symplectic maps or (in the homogeneous setting) contact transformations, and in the Japanese literature they are often referred to as quantized contact transformations. They formalize P. Dirac's idea [**Dir**] that unitary operators are the quantum mechanical analogue of symplectic transformations. This analogy was put on a firmer mathematical basis by Van Vleck [**V**], Pauli [**P**] and Fock [**F**].

Fourier integral operators come in two basic types: (i) homogeneous, and (ii) semi-classical. Homogeneous Fourier integral operators are used construct parametrices for 'propagators' of wave equations such as $U(t) = \exp it\sqrt{-\Delta}$ or Green's functions, e.g., Schwarz kernels of Δ^{-1}. They have homogeneous phases of degree 1 and poly-homogeneous amplitudes. Semi-classical Fourier integral operators are used to construct propagators for the Schŕodinger equation with $\hbar$ or many other kernels involving an external parameter $\hbar$. Hadamard's parametrix (6.5) for the wave kernel is a homogeneous Fourier integral operator. Semi-classical propagators were probably introduced by van Vleck and Pauli in the form

$$K_{\rm sc}(x, \tau, y, 0) = (i\hbar)^{-n/2}\sqrt{D}e^{\frac{i}{\hbar}S(t,x,y)}, \tag{7.1}$$

where $S(t, x, y)$ is the generating function of the graph of the associated classical Hamiltonian flow, and where

$$D = D(t, x, y) = (-1)^n \det\left(\frac{\partial S^2}{\partial x_j \partial y_k}\right) \tag{7.2}$$

is now known as the van Vleck determinant. In the symbol calculus of Fourier integral operators, the van Vleck determinant is interpreted in terms of the principal symbol, so that the 'scalar' symbol of (7.1) is equal to one. Pauli gave a formal proof that van Vleck determinants satisfy the cocycle condition $D(2,1)D(1,0) = D(2,0)\det S$ under composition; here, S is the Hessian of $S(2,1)+S(1,0)$ at the critical point 1 (see §7.3.3). Physics reference explaining this include [**V, F, DeW, M, Li**]. Fock [**F**] gave a formal proof that Schwartz kernels of this form are unitary to leading order in $\hbar$ (see §7.3.4). Exponentials of quadratic Hamiltonians, the so-called metaplectic representation [**Fo**], have this form when the canonical relation is projectible §7.3.2. When the canonical relation is not projectible, one needs a more general theory in which propagators are expressed as oscillatory integrals.

This chapter is accordingly split up into a section on homogeneous Fourier integrals and a section on semi-classical Fourier integral operators. As mentioned above, the purpose is only to illustrate some basic notions with concrete examples.

7.1.1. Symbol calculus and symplectic category. In the previous chapter, we constructed the Hadamard parametrix for the wave kernel. In the following section on Weyl laws we use another parametrix for the half-wave group with a different phase function. This raises the question of finding invariants common to the different parametrix constructions. The theory of Fourier integral operators provides such invariants, the main ones being the canonical relation C and principal symbol σ. Use of these invariants simplifies the calculation of the leading order terms in stationary phase expansions of oscillatory integrals obtained by composing many Fourier integral kernels. The symbol calculus gives rules for gluing together stationary phase expansions of the parts. The symbol is a section of the bundle $\Omega_{\frac{1}{2}}(C)$ of half densities on C tensored with the Maslov canonical bundle. In this monograph we omit discussion of the Maslov bundle, referring to the references above for background.

Under certain transversality (or 'clean-ness') assumptions, the composition of Fourier integral operators is a Fourier integral operator and, ignoring technical complications, one may speak of the algebra of Fourier integral operators. The symbol calculus is the induced 'algebra' of pairs (C,σ) of weighed canonical relations, consisting of a canonical relation C and a half-density on C. Thus, given weighted canonical relations (C_1,σ_1) and (C_2,σ_2) we would like to define

$$(C_1 \circ C_2, \sigma_1 \circ \sigma_2). \tag{7.3}$$

The composition $C_1 \circ C_2$ is simply set theoretic composition and the only problem is C_1 must satisfy generalized transversality conditions relative to C_2 so that the composition is a canonical relation, i.e., a Lagrangian submanifold. Weinstein has called the symbol algebra the 'symplectic category' and we follow this reasonably standard terminology. The symplectic category is a geometric model for the 'algebra' of Fourier integral operators and is discussed in detail in [**DuG, GuSt1, GuSt2, Ho6**].

In practice, the symbol calculus is used to simplify lengthy calculations of stationary phase expansions arising from compositions of many Fourier integral operators. The most difficult part to interpret geometrically is the Hessian of the phase on the critical set. The symbol calculus can potentially be used to analyze the Hessian of a phase of many variables as a composition of principal symbols of the factors. Although we do not use the symbol calculus very much in this monograph, it is used in many of the cited references and is conceptually useful.

The principal symbol calculus is formalized as the "symplectic category" whose objects are weighted symplectic manifolds and whose morphisms are weighted canonical relations, where weights are half-densities or forms. Composing symbols is a matter of symplectic linear algebra which is explained in §2 of [**Ho6**] and will be reviewed in §7.1.1.

7.2. Homogeneous Fourier integral operators

The term 'homogeneous' refers to the $\mathbb{R}_+$ action $r \cdot (x, \xi) = (x, r\xi)$ on $T^*M \setminus 0$, with $r \in \mathbb{R}_+$. An oscillatory integral distribution

$$u(x) = \int_{\mathbb{R}^N} e^{i\varphi(x,\theta)} a(x,\theta)\, d\theta \tag{7.4}$$

is said to be a Fourier integral (or Lagrangian) distribution if its phase φ is homogeneous of degree one and if its amplitude $a(x,\theta) \sim \sum_{j=0}^{\infty} a_j(x)|\theta|^{\mu-j}$ is polyhomogeneous. The integral representation is not unique and one would like to associate invariants to u. These are the associated Lagrangian submanifold $\Lambda_\varphi \subset T^*X$ and the symbol.

We first define Λ_φ. The critical set of the phase is given by

$$C_\varphi = \{(x,\theta) \colon d_\theta \varphi = 0\}. \tag{7.5}$$

Under ideal conditions, the map

$$\iota_\varphi \colon C_\varphi \to \iota_\varphi(C_\varphi) =: \Lambda_\varphi \subset T^*X, \quad \iota_\varphi(x,\theta) = (x, d_x\varphi) \tag{7.6}$$

is a Lagrangian embedding, or at least an immersion. In this case the phase is called non-degenerate. Less restrictive, although still an ideal situation, is where the phase is clean. This means that the map

$$\iota_\varphi \colon C_\varphi \to \iota_\varphi(C_\varphi) =: \Lambda_\varphi, \tag{7.7}$$

is locally a fibration with fibers of dimension e. From [**Ho6**, Definition 21.2.5] , the number of linearly independent differentials $d\frac{\partial\varphi}{\partial\theta}$ at a point of C_φ is $N - e$.

We write a tangent vector to $M \times \mathbb{R}^N$ as $(\delta_x, \delta_\theta)$ or $(\delta x, \delta\theta)$. The kernel of

$$D\varphi'_\theta = \begin{pmatrix} \varphi''_{\theta x} & \varphi''_{\theta\theta} \end{pmatrix} \tag{7.8}$$

is $T_{(x,\theta)}C_\varphi$. That is, $(\delta_x, \delta_\theta) \in TC_\varphi$ if and only if $\varphi''_{\theta x}\delta_x + \varphi''_{\theta\theta}\delta_\theta = 0$. Indeed, φ'_θ is the defining function of C_φ and $d\varphi_\theta$ is the defining function of TC_φ. By [**Ho6**, Definition 21.2.5], the number of linearly independent differentials $d\frac{\partial\varphi}{\partial\theta}$ at a point of C_φ is $N - e$ where e is the excess. Then $C \to \Lambda$ is locally a fibration with fibers of dimension e. So to find the excess we need to compute the rank of $\begin{pmatrix} \varphi''_{x\theta} & \varphi''_{\theta\theta} \end{pmatrix}$ on $T_{x,\theta}(\mathbb{R}^N \times M)$. Non-degeneracy is thus the condition that

$$\begin{pmatrix} \varphi''_{\theta x} & \varphi''_{\theta\theta} \end{pmatrix} \text{ is surjective on } C_\varphi \iff \begin{pmatrix} \varphi''_{\theta x} \\ \varphi''_{\theta\theta} \end{pmatrix} \text{ is injective on } C_\varphi. \tag{7.9}$$

Note that

$$d\iota_\varphi(\delta_x, \delta_\theta) = (\delta_x, \varphi''_{xx}\delta_x + \varphi''_{x\theta}\delta_\theta). \tag{7.10}$$

Indeed, if $d\iota_\varphi(\delta_x, \delta_\theta) = 0$ then $\delta_x = 0$ and $\varphi''_{x\theta}\delta_\theta = 0$ but then if $(\delta_x, \delta_\theta) \in TC_\varphi$, then also $\varphi''_{\theta\theta}\delta_\theta = 0$. So if $\begin{pmatrix} \varphi''_{\theta x} \\ \varphi''_{\theta\theta} \end{pmatrix}$ is injective, then $\delta_\theta = 0$.

7.2.1. Fourier integral operators. A homogeneous Fourier integral operator $A\colon C^\infty(X) \to C^\infty(Y)$ is an operator whose Schwartz kernel may be represented by a Fourier integral distribution

$$K_A(x,y) = \int_{\mathbb{R}^N} e^{i\varphi(x,y,\theta)} a(x,y,\theta)\, d\theta, \tag{7.11}$$

where (as above) the phase φ is homogeneous of degree one in θ. The amplitude is understood to be a poly-homogeneous function of θ, i.e., $a(x,y,\theta) \sim \sum_j a_j(x,y)\theta^{\mu-j}$ as $j \to \infty$. There could be any number N of phase variables θ. The critical set of the phase is then given by

$$C_\varphi = \{(x,y,\theta)\colon d_\theta\varphi = 0\}, \tag{7.12}$$

and it is non-degenerate if

$$\iota_\varphi : C_\varphi \to T^*(X \times Y), \quad \iota_\varphi(x,y,\theta) = (x, d_x\varphi, y, -d_y\varphi) \tag{7.13}$$

is a Lagrange embedding, or at least an immersion. As before, the image of ι_φ is denoted by Λ_φ.

The geometry is encoded in the diagram:

$$\begin{array}{ccc} & \Lambda_\phi \subset T^*X \times T^*Y & \\ {}^{\pi}\swarrow & & \searrow^{\rho} \\ T^*X & & T^*Y \end{array}$$

The canonical relation Λ_φ is called a canonical graph if the projections π, ρ are homogeneous diffeomorphisms, in which case it is the graph of a homogeneous canonical (i.e., symplectic) transformation. It is called a local canonical graph if both π, ρ are local diffeomorphisms.

A submanifold $C \subset T^*(X \times Y)$ is called a *canonical relation* if it is Lagrangian for the difference symplectic structure $\omega_X - \omega_Y$ of the canonical symplectic forms. For instance, the graph of a symplectic map $\chi : T^*X \to T^*Y$ is a canonical graph consisting of pairs (x, ξ, y, η) such that $\chi(x,\xi) = (y,\eta)$. One can (and often does) multiply η by -1 to reverse the sign of the symplectic form and make the canonical relation Lagrangian with respect to the standard symplectic structure on the product. Homogeneous canonical relations, diffeomorphisms and phases are understood to be homogeneous of degree 1. Symbols can be homogeneous of any order.

7.2.2. Order and symbol. The second invariant associated to a Lagrangian distribution is its symbol, a section of the bundle of half-densities tensor a Maslov bundle. We omit any discussion of Maslov bundles, but do review half-densities and the half-density part of the symbol.

Half-densities arise when one wishes to endow a manifold with a Hilbert space structure $L^2(X)$. In general there is no preferred volume form and instead of scalar functions one uses half-densities $\Omega^{\frac{1}{2}}$, i.e., square roots of densities, to define the intrinsic Hilbert space $L^2(X, \Omega^{\frac{1}{2}})$. Thus, K_A from (7.11) is a double half-density on $X \times Y$ and its integral against a half-density on Y is well-defined The symbol calculus is based on this half-density formalism. In the Riemannian setting of this monograph, there does exist a preferred half density $\sqrt{dV_g}$, the square root of the volume density, and there is a natural Hilbert space isomorphism $L^2(M, \Omega^{\frac{1}{2}}) \simeq$

$L^2(M, \sqrt{dV_g})$. Whenever possible, we write half-densities as $f\sqrt{dV_g}$. We need to keep track of the half-density aspect to define symbols correctly.

The *order* of a homogeneous Fourier integral operator $A\colon L^2(X) \to L^2(Y)$ in the non-degenerate case is given in terms of a local oscillatory integral formula (7.11) by $m + \frac{N}{2} - \frac{n}{4}$, where $n = \dim X + \dim Y$, m the order of the amplitude, and N the number of phase variables in the local Fourier integral representation (see [**Ho6**, Proposition 25.1.5]); in the general clean case with excess e, the order goes up by $\frac{e}{2}$ (see [**Ho6**, Proposition 25.1.5']).

The space of Fourier integral operators of order m associated to a canonical relation $C \subset T^*(X \times Y) \setminus 0$ is denote by

$$I^m(X \times Y, C). \tag{7.14}$$

The *symbol* of the Fourier integral operator A is a homogeneous section of $\Omega_{\frac{1}{2}} \otimes \mathcal{M}$ of the bundle of half-densities times a section of the Maslov line bundle on C. It involves the delta-density d_{C_φ} along the critical set of the phase. More precisely, d_{C_φ} is the Leray form defined by $\delta(d_\theta \varphi)$, transported to its image in T^*M under ι_φ. If $(\lambda_1, \ldots, \lambda_n)$ are any local coordinates on C_φ, extended as smooth functions in neighborhood, then

$$d_{C_\varphi} := \frac{|d\lambda|}{|D(\lambda, \varphi'_\theta)/D(x,\theta)|}, \tag{7.15}$$

where $d\lambda$ is the Lebesgue density. The symbol

$$\sigma_A := a_0 \sqrt{d_{C_\varphi}} \tag{7.16}$$

is the leading order term $a_0|_C$ of the amplitude on C multiplied by the square-root of the Leray form.

If φ is a non-degenerate phase function of a general Lagrangian distribution (7.4), define the Hessian (cf. [**Ho6**, Proposition 25.1.5])

$$\Phi = \begin{pmatrix} \varphi''_{xx} & \varphi''_{x\theta} \\ \varphi''_{\theta x} & \varphi''_{\theta\theta} \end{pmatrix} \tag{7.17}$$

Then the symbol of (7.4) expressed in local coordinates is equal to

$$\sigma_u = e^{\frac{i\pi}{4}\operatorname{sgn}(\Phi)} a_0 |\det \Phi|^{-\frac{1}{2}} |d\xi|^{\frac{1}{2}}. \tag{7.18}$$

See [**Ho2**, Proposition 4.1.3]. For a clean phase function, the more complicated formula is given in [**Ho6**, p. 15].

The definition of the principal symbol is evidently "extrinsic", i.e., it makes use of a specific representation (7.11), the embedding ι_φ from (7.7), and a density (7.15) that depends on φ and not just on the zero set $\{d_\theta \varphi = 0\}$. Indeed, (7.15) is a Leray form of type $\frac{dV}{df}$ on $\{f = 0\}$, and such a form depends on the choice of defining function f. It is verified in [**Ho2, Du1, GuSt1**] and elsewhere that the principal symbol is independent of these choices. However it is obviously desirable to have an intrinsic formula for the principal symbol as a half-density on Λ_φ.

We now specialize to kernels $K_A \in \mathcal{D}'(X \times Y)$ of Fourier integral operators. In general, if $C \subset T^*Y \times T^*X$ is a canonical relation, one can parametrize C in the neighborhood of a point $c_0 \in C$ by a non-degenerate phase function $\varphi(x,y,\theta) \in C^\infty(X \times Y \times \mathbb{R}^N \backslash 0)$ near (x_0, y_0, θ_0) where $\pi(c_0) = (x_0, y_0)$, i.e., by

$$(x,y,\theta) \in C_\varphi = \{(x,y,\theta)\colon \varphi'_\theta(x,y,\theta) = 0\} \to (x, \varphi'_x, y, -\varphi'_y). \tag{7.19}$$

The map $C \to T^*X$ is bijective if and only if $n_X = n_Y$ and if

$$D(\varphi) := \det \begin{pmatrix} \varphi''_{\theta\theta} & \varphi''_{\theta x} \\ \varphi''_{\theta y} & \varphi''_{yx} \end{pmatrix} \neq 0 \quad \text{at } (x_0, y_0, \theta_0). \tag{7.20}$$

It is easy to check that

$$\frac{D(x, \varphi'_x, \varphi'_\theta)}{D(x, y, \theta)} = \det \begin{pmatrix} \varphi''_{yx} & \varphi''_{\theta x} \\ \varphi''_{\theta y} & \varphi''_{\theta\theta} \end{pmatrix}. \tag{7.21}$$

Since ι_φ is an embedding, $(x, \xi) = (x, \varphi'_x)$ are local coordinates and
(7.22)

$$d_C = |D(\varphi)|^{-1} dx_1 \cdots dx_n d\xi_1 \cdots d\xi_n = \left| \frac{D(x, \varphi'_x, \varphi'_\theta)}{D(x, y, \theta)} \right|^{-1} dx_1 \cdots dx_n d\xi_1 \cdots d\xi_n.$$

See [**Ho2**, Proposition 4.1.3].

A Fourier integral operator $A \in I^m(X \times Y, C)$ is determined up to a lower order term in $I^{m-1}(X \times Y, \mathbb{C})$ by (C, σ_A).

7.2.3. Oscillatory integrals associated to Lagrangian subspaces of a symplectic vector space. The simplest oscillatory integrals or Lagrangian distributions are those associated to Lagrangian subspaces $\lambda \subset S$ of a symplectic vector space (S, σ). Chapter 21.6 of [**Ho5**] is devoted to them, and we review some of the results in this section. In a sense, symbol calculus is simply a combination of this linear case together together with some calculus on manifolds.

The Lagrangian Grassmannian $\Lambda(S)$ is the manifold of all Lagrangian subspaces of (S, σ). One may identify $S \simeq T^*\mathbb{R}^n$ and σ with the standard symplectic form. If $\mu = T_0^*\mathbb{R}^n$ denotes the vertical Lagrangian subspace, then elements $\lambda \in \Lambda_\mu(S)$, the of Lagrangian subspaces transversal to μ, can be represented as graphs of of symmetric matrices, $\lambda = \{(x, Ax)\}$ or as differentials of the quadratic functions $q_A(x) = \langle Ax, x\rangle/2$. Here, μ could be replaced by another Lagrangian plane.

If λ_0, λ_1 are transversal Lagrangian subspaces, then the bilinear map $(X, Y) \in \lambda_0 \times \lambda_1 \to \sigma(X, Y)$ is non-degenerate and defines an isomorphism from $\lambda_1 \to \lambda_0^*$ where λ_0^* is the dual space of λ_0. Introduce linear symplectic coordinates (x, ξ) so that $\lambda_0 = \{x = 0\}$ and $\lambda_1 = \{\xi = 0\}$. Following [**Ho5**, (21.6.4)], define, for $\lambda \in \Lambda(S)$,

$$I(\lambda; \lambda_1) := \Big\{ u \in \mathcal{S}'(\lambda_0^*, \Omega^{\frac{1}{2}}) \colon L(x, D)u = 0 \text{ for all } L(x, \xi) \in S^* \\ \text{such that } L(x, \xi) = 0 \text{ on } \lambda \Big\}. \tag{7.23}$$

Here, $L(x, \xi)$ is a linear functional. The definition is almost independent of λ_1 and the equations $L(x, D) = u$ for $L|_\lambda = 0$ is an involutive system of maximal dimension n, so that $\{u, \lambda) : u \in I(\lambda)\}$ is one-dimensional and $\lambda \to I(\lambda)$ defines a line bundle over $\Lambda(S)$. This may be seen concretely when λ is the graph of $\langle B\xi, \xi\rangle/2$ for a symmetric matrix B. Then $u \in I(\lambda, \lambda_1)$ is the oscillatory integral

$$u(x) = C \int_{\mathbb{R}^n} e^{i(\langle x, \xi\rangle - \langle B\xi, \xi\rangle/2)} \, d\xi. \tag{7.24}$$

Every $u \in I(\lambda, \lambda_1)$, whether or not λ is transversal to λ_1, may be represented in the form

$$u(x) = C \int_{\mathbb{R}^n} e^{iQ(x, \xi)} \, d\xi, \qquad x \in \lambda_1, \tag{7.25}$$

with $Q(x,\xi) = \langle x, \xi\rangle + \langle Ax, x\rangle/2 - \langle B\xi, \xi\rangle/2$ for some symmetric A, B so that

$$\lambda = \left\{ \left(x, \frac{\partial Q}{\partial x} \right) : \frac{\partial Q}{\partial \xi} = 0 \right\}. \tag{7.26}$$

Further, in [**Ho5**, (p. 333)] it is explained that $I(\lambda)$ can be identified with the space of translation invariant half-densities on λ. Then as in (7.15), as half-densities $\Omega^{\frac{1}{2}}(\lambda)$ on λ,

$$d_C^{\frac{1}{2}} = |\det Q''|^{-\frac{1}{2}} |d\xi|^{\frac{1}{2}}. \tag{7.27}$$

The full symbol is

$$\sigma_u = a\, d_C^{\frac{1}{2}} e^{i\pi \operatorname{sgn} Q/4}, \tag{7.28}$$

where a is a normalization of the constant C. Such oscillatory integrals with quadratic phase functions form the metaplectic representation and are discussed further in §7.3.2.

7.2.4. Quantizations of canonical transformations. The simplest homogeneous Fourier integral operators F are those whose canonical relations are graphs of homogeneous symplectic diffeomorphisms (canonical transformations). They may be represented in terms of the Fourier transform by

$$Fu(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} a(x,\xi) e^{i\varphi(x,\xi)} \hat{u}(\xi)\, d\xi. \tag{7.29}$$

This is a local representation in which the manifold is identified with $\mathbb{R}^n$ and the homogeneous phase is assumed to have the form

$$\Phi(x, y, \xi) = \varphi(x, \xi) - \langle y, \xi \rangle \tag{7.30}$$

with phase variable ξ.

The critical set is then given by

$$C_\Phi = \{(x, y, \xi) : y = d_\xi \varphi(x, \xi)\} \subset \mathbb{R}^n \times T^*\mathbb{R}^n \tag{7.31}$$

and the Lagrangian immersion is given by

$$\iota_{\Phi|}(x, y, \xi) = (x, \varphi'_x, y, \xi) = (x, \varphi'_x, \varphi'_\xi, \xi). \tag{7.32}$$

Thus, the independent variables on $C_\Phi \simeq T^*\mathbb{R}^n$ are (x, ξ) and we may view φ as a function on $T^*\mathbb{R}^n$ with $\Lambda_\Phi = \operatorname{graph}(d\varphi)$. As observed in [**Ho6**, Proposition 25.3.3], C_φ is the graph of a canonical transformation if and only if $\det \left(\frac{\partial^2 \varphi}{\partial x \partial \xi} \right) \neq 0$ since this is the condition for the maps $(x, \xi) \to (x, \varphi'_x)$ and $(x, \xi) \to (x, \varphi'_\xi)$ to be local diffeomorphisms, and also for the critical set C_φ to be a local canonical graph.

The principal symbol is the half-density on Λ_Φ given in local (x, ξ) coordinates (see [**Ho6**, p. 27]) by

$$\sigma(F) = \frac{a_0(x,\xi)}{\sqrt{\det \left(\frac{\partial^2 \varphi}{\partial x \partial \xi} \right)}} |dx d\xi|^{\frac{1}{2}}. \tag{7.33}$$

To see that this is consistent with (7.15), we have to show that (7.33) equals the symbol is $a_0 \sqrt{d_{C_\varphi}}$, transported to Λ_Φ. By (7.18), it suffices to observe that (7.33) equals

$a(x,y,\theta)|D(\varphi)|^{-\frac{1}{2}}$ times the symplectic volume half-density, whereas in (7.22),

$$D(\varphi) = \det \begin{pmatrix} \varphi''_{\xi\xi} & \varphi''_{\xi x} \\ \varphi''_{y\xi} & \varphi''_{yx} \end{pmatrix} = \det \begin{pmatrix} \varphi''_{\xi\xi} & \varphi''_{\xi x} \\ -I & 0 \end{pmatrix} = \det \varphi''_{x\xi}. \tag{7.34}$$

One of the principal examples of a homogeneous Fourier integral operator associated to a canonical graph is the half-wave kernel at fixed time, i.e., the Schwartz kernel of the half-wave operator $e^{-it\sqrt{-\Delta}}$. In the case of Euclidean $\mathbb{R}^n$, the half-wave kernel has the above kind of homogeneous Fourier integral representation

$$U(t,x,y) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\langle x-y,\xi\rangle} e^{it|\xi|}\, d\xi. \tag{7.35}$$

For fixed t its homogeneous canonical relation in $T^*(M\times M)$ is given by the graph of the geodesic flow on the characteristic variety $\tau + |y|_g = 0$, i.e.,

$$\mathcal{C} = \{(t,\tau,x,\xi,y,\eta) : \tau + |\eta|_g = 0,\ \ G^t(x,\xi) = (y,\eta)\}. \tag{7.36}$$

For fixed (t,y), the amplitude is 1 and the phase is $\varphi(t,x,y,\xi) = \langle x-y,\xi\rangle + it|\xi|$. Since $|\xi|$ is independent of x, the mixed Hessian $\varphi''_{x,\xi} = 1$ and the symbol is the graph half-density $|dx\wedge d\xi|^{\frac{1}{2}}$.

If one views t, y as variables, then $U(t,x,y) \in \mathcal{D}'(\mathbb{R}\times M\times M)$ is distribution kernel on $\mathbb{R}\times M\times M$ and its canonical relation is a homogeneous canonical relation in $T^*(\mathbb{R}\times M\times M)$ given by the space time graph of the geodesic flow on the characteristic variety $\tau + |y|_g = 0$, i.e.

$$\mathcal{C} = \{(t,\tau,x,\xi,y,\eta)\colon \tau + |\eta|_g = 0,\ G^t(x,\xi) = (y,\eta)\}. \tag{7.37}$$

Its symbol is then $|dt\wedge dx\wedge d\xi|^{\frac{1}{2}}$. We refer to [**Ho6**] and [**DuG**] for background.

In the definition of the symbol (7.33), one divides by the van Vleck determinant. The non-homogeneous analogue of the symbol of (7.1) would then be 1. In [**SV**] (see (2.1.2), (2.2.5) and Definition 2.7.1), homogeneous Fourier integral operators quantizing canonical transformations are represented in the form

$$\mathcal{I}_{\varphi,a}(t,x,y) = \int e^{i\varphi(t,x,y,\eta)} a(t,x,y,\eta)\zeta(t,x,y,\eta)|\det\varphi_{x,\eta}|^{\frac{1}{2}}\, d\eta, \tag{7.38}$$

where ζ is a cutoff. In this representation, the principal symbol is $a_0|_{C_\varphi}|dxd\xi|^{\frac{1}{2}}$. There is an additional Maslov factor $e^{i\pi\arg/4}$ (see [**SV**, (2.2.4)]).

7.2.5. Special parametrization of a conic Lagrangian. In view of the non-uniqueness of the Fourier oscillator integral representations of a Lagrangian distribution, it is desirable to have a kind of canonical parametrization of a Lagrangian submanifold. In the conic (homogeneous) case, where Λ is invariant under the $\mathbb{R}_+$ action, a reasonably canonical local parametrization is to express Λ as the graph of an exact 1-form on the vertical fibers T^*_xM of T^*M. Of course, this is only a local representation when T^*M is not symplectically equivalent to the product $M\times\mathbb{R}^n$.

Theorem 21.2.16 of [**Ho5**] asserts that for any conic Lagrangian submanifold $\Lambda\subset T^*M$ which is locally projectible to the fiber T^*_xM, there exist local symplectic coordinates (x,ξ) and a homogeneous function $H(\xi)$ so that

$$\Lambda = \{(H'(\xi),\xi)\colon \xi\in\mathbb{R}^n\}. \tag{7.39}$$

Here, locally projectible means that there exists a conic neighborhood of (x_0,ξ_0) so that the map $\pi_V\colon \Lambda\to\mathbb{R}^n_\xi$ is non-singular in the cone, i.e., Λ is transverse to

$\xi = C$. By Proposition 25.1.3 of [**Ho6**] a Lagrangian distribution u associated to the Lagrangian Λ may be locally represented in the form

$$Fu(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} a(x,\xi) e^{iH(\xi)} \hat{u}(\xi)\, d\xi. \tag{7.40}$$

The principal symbol is then (up to Maslov factors and dimensional constants) given by

$$\sigma_u = \pi_V^*(a_0 |d\xi|^{\frac{1}{2}}). \tag{7.41}$$

7.2.6. Unitary homogeneous Fourier integral operators. If F is a homogeneous Fourier integral operator associated to the graph of a canonical transformation, then F^*F and FF^* are pseudo-differential operators. This follows from the fact that they are associated to $C \circ C^{-1} = \text{graph(Id)}$. The principal symbol is given by (7.20)

$$\sigma_{FF^*} = \sigma_{F^*F} = |a(x,y,\theta)|^2 ||D(\varphi)|^{-1}. \tag{7.42}$$

For the proof, we refer to [**Tr**, (6.16)].

Suppose that F is a unitary homogeneous Fourier integral operator associated to the graph of a canonical transformation. Then $FF^* \sim F^*F \sim 1$ and it follows from (7.42) that

$$|a(x,y,\theta)||D(\varphi)|^{-\frac{1}{2}} = 1. \tag{7.43}$$

For semi-classical Fourier integral kernels of WKB form, (7.42) and (7.43) were already proved by V. Fock in [**F**]. The proof is sketched in §7.3.4.

7.2.7. Principal symbol of the Hadamard parametrix. The Hadamard parametrix is a Fourier integral representation of the wave kernel for small $|t|$ which has only one phase variable. This reflects the fact that the wave front set of the wave kernel for fixed initial point y is the positive co-normal bundle of the distance sphere of radius t centered at y.

The scalar cosine wave kernel is given by (cf. [**Be**, Section D])

$$\cos t\sqrt{-\Delta}(x,y) = C_n |t| \int_0^\infty e^{i\theta \frac{(t^2 - r^2(x,y)}{2})} A(t,x,y,\theta) \theta^{\frac{n-1}{2}}\, d\theta, \tag{7.44}$$

where C_n is a dimensional constant defined so that the right side is $\delta_x(y)$ at $t=0$. It is a Fourier integral operator with non-degenerate phase function

$$\varphi(t;x,y,\theta) = \frac{1}{2}\theta(t^2 - r^2(x,y)), \tag{7.45}$$

and with leading order term

$$A_0(t,x,y,\theta) = |t| \theta^{\frac{n-1}{2}} \Theta^{-\frac{1}{2}}(x,y), \tag{7.46}$$

where

$$d\,\text{Vol}(y) = \Theta(x,y)\, dy \quad \text{and} \quad \Theta(x,y) = |\det D_y \exp_x(y)|. \tag{7.47}$$

Thus, if $\gamma_{x,y}$ is the minimizing geodesic from x to y, then

$$\Theta(x,y) = \frac{1}{r^{n-1}} |\det g(V_j(r), V_k(r))|^{\frac{1}{2}}, \tag{7.48}$$

where $\{V_j\}_{j=2}^n$ are orthogonal Jacobi fields along $\gamma_{x,y}$ satisfying $V_j(0) = 0, \gamma'(0), V_2'(0), \ldots, V_n'(0)$ is an orthonormal basis of T_xM.

The associated canonical relation for fixed t is the union of the graphs of the geodesic flow G^t and of its inverse G^{-t}, which intersect when $t = 0$ and at points (x, ξ) with $G^{2t}(x, \xi) = (x, \xi)$. Each branch is a canonical graph and so the canonical relation is a local canonical graph. As a distribution in (t, x, y),

$$C_\varphi = \{(t, x, y, \theta) \in \mathbb{R}_t \ \times M \times M \times \mathbb{R}_+ \colon r^2(x, y) = t^2\}.$$

Note that

$$\mathcal{C} = \{(t, x, y) \in \mathbb{R}_t \times M \times M \colon r^2(x, y) = t^2\}$$

is the *characteristic conoid*, and $C_\varphi \to \mathcal{C}$ is an R_+ bundle.

It is sometimes useful to fix a variable and then we subscript C_φ with the frozen variable. If t is fixed,

$$C_{\varphi,t} = \{(x, y, \theta) \in \mathbb{R}_+ \times M \times M \colon r^2(x, y) = t^2\} \tag{7.49}$$

is an $\mathbb{R}_+$ bundle over the family of distance spheres of radius t centered at the moving point x, and the Lagrange map is

$$\iota_t(x, y, \theta) = \left(x, \frac{1}{2}\theta d_x r^2, y, -\frac{1}{2}\theta d_y r^2\right) \colon C_\varphi \cap \{r(x, y) = |t|\} \to T^*(M \times M)\backslash 0. \tag{7.50}$$

For small (non-zero) distances, $(x, d_x r(x, y), y, -d_y r(x, y))$ are the initial resp. terminal vectors of the unique geodesic (of length r) between x, y. To see this, we denote by $\gamma_{x,\xi}(t)$ the unit speed geodesic with initial data (x, ξ), i.e. $\gamma_{x,\xi}(0) = x, \gamma'_{x,\xi}(0) = \xi$. Then $r(x, \gamma_{x,\xi}(t)) = t$. One has $g(\nabla r, \nabla r) = 1$ and

$$\exp_x r(x, y) \nabla_x r(x, y) = y. \tag{7.51}$$

If $\gamma_{x,\xi}(r(x, y)) = y$ then the time reversal of the terminal tangent vector satisfies

$$-\gamma'_{x,\xi}(r(x, y)) = \frac{1}{2} \nabla_y r^2(x, y), \tag{7.52}$$

since $\exp_y(-\gamma'_{x,\xi}(r(x, y))) = x$ and since the right side also satisfies this equation by (7.51). The image of this map is the graph of G^t, so the map $C \to T^*X$ is bijective and $n_X = n_Y$.

Since $e^{-it\sqrt{-\Delta}}$ is a unitary Fourier integral, its symbol has modulus one (7.43). We verify this using (7.15). The symbol we consider is that of the half-density kernel

$$\cos t\sqrt{-\Delta}(x, y) \sqrt{dV_g(x)} \sqrt{dV_g(y)}. \tag{7.53}$$

By (7.15) and (7.20) the symbol is

$$C_n |t| \Theta^{-\frac{1}{2}}(x, y) |D(\varphi)|^{-\frac{1}{2}} \theta^{\frac{n-1}{2}} dx d\xi. \tag{7.54}$$

For $t \neq 0$ and (x, y) close to but not on the diagonal, ι_t is an embedding, $(x, \xi = \theta\varphi'_x = \theta t d_x r(x, y))$ are local coordinates on $T^*\mathbb{R}^n \simeq \text{graph}(G^t)$, and

$$D(\varphi) = \det \begin{pmatrix} 0 & \frac{1}{2} d_x r^2 \\ \frac{1}{2} d_y r^2 & \frac{1}{2}\theta \frac{\partial^2 r^2(x,y)}{\partial x \partial y} \end{pmatrix} = \theta^n \det \left(\tfrac{\partial^2 r^2(x,y)/2}{\partial x \partial y}\right) \left\langle \left(\frac{1}{2}\theta \frac{\partial^2 r^2(x, y)}{\partial x \partial y}\right)^{-1} \frac{d_x r^2}{2}, \frac{d_y r^2}{2} \right\rangle. \tag{7.55}$$

A standard relation (see [**LV, Wi**]) between (7.48) and Hessians of distance functions is that

$$\Theta^{-\frac{1}{2}}(x,y) = \sqrt{\det\left(\frac{\partial^2 r^2(x,y)/2}{\partial x \partial y}\right)}, \quad \Theta(x,y) := \frac{dV_g \otimes dV_g}{(\frac{\partial^2 r^2/2}{\partial x\partial y} dx \wedge dy)^n}, \tag{7.56}$$

so that

(7.57)

$$\Theta^{-\frac{1}{2}}(x,y)\left[\det\left(\tfrac{\partial^2 r^2(x,y)/2}{\partial x\partial y}\right)\right]^{-\frac{1}{2}} = \left(\frac{dV_g \otimes dV_g}{(\frac{\partial^2 r^2/2}{\partial x\partial y} dx\wedge dy)^n}\det\left(\tfrac{\partial^2 r^2(x,y)/2}{\partial x\partial y}\right)\right)^{-\frac{1}{2}} = \left(\frac{dV_g(x)dV_g(y)}{dx\wedge dy)}\right)^{-\frac{1}{2}}.$$

Further, the bilinear form

$$\langle \xi, (d_v \exp_x)^{-1}\eta\rangle_x = \langle (d_w \exp_y)^{-1}\xi, \eta\rangle_y \qquad \text{for all } (\xi,\eta) \in T_xM \times T_yM \tag{7.58}$$

on $T_xM \times T_yM$ (with $v = (\exp_x)^{-1}(y), w = (\exp_y)^{-1}(x)$) coincides up to sign with

$$\langle (-\nabla^2_{x,y} d^2(x,y)/2)\xi, \eta\rangle, \quad \xi \in T_xM, \eta \in T_yM. \tag{7.59}$$

Since

$$(d_v \exp_x)^{-1}(d_y r(x,y)^2/2) = v, \qquad v = d_x r^2(x,y)/2, \tag{7.60}$$

we have, for $v = d_x r^2(x,y)/2$,

(7.61)

$$\left\langle \left(\frac{(\partial^2 r^2(x,y)/2}{\partial x\partial y}\right)^{-1} d_x r^2/2, d_y r^2/2\right\rangle = \langle v, (d_v \exp_x)^{-1} d_y r^2(x,y)/2\rangle_x = r^2.$$

Moreover, $\xi = \varphi'_x = \theta d_x r^2/2(x,y) = \theta t\omega$ when $r(x,y) = t$, and

$$d_C = |D(x, \varphi'_x, \varphi'_\theta)/D(x,y,\theta)|^{-1}\, dx_1 \cdots dx_n d\xi_1 \cdots d\xi_n \tag{7.62}$$

$$= t^2\, dx_1 \cdots dx_n d\xi_1 \cdots d\xi_n. \tag{7.63}$$

Here, $\exp_x t\omega = y$. Hence,

$$D(\varphi)|_C = r^2 \Theta^{\frac{1}{2}} = t^2\Theta^{\frac{1}{2}}, \tag{7.64}$$

and so the principal symbol for $t > 0$ equals

$$|t|(r^2)^{-\frac{1}{2}}\theta^{-\frac{n-1}{2}}\theta^{\frac{n-1}{2}} = 1 \qquad (t = r). \tag{7.65}$$

This corroborates the fact that the principal symbol of the unitary wave propagator $e^{-it\sqrt{-\Delta}}$ equals 1.

As a simple example we recall the Euclidean case where the kernel of $e^{-it\sqrt{-\Delta}}$ is

$$U(t,x,y) = \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}} \frac{it}{(|x-y|^2 - (t-i0)^2)^{\frac{n+1}{2}}} \tag{7.66}$$

$$= \frac{\Gamma(\frac{n+1}{2})}{\pi^{\frac{n+1}{2}}}(it)\int_0^\infty e^{i\theta(|x-y|^2-(t-i0)^2)}\theta^{\frac{n-1}{2}}\, d\theta \tag{7.67}$$

7.2.8. Restriction to a hypersurface. An example of a homogeneous Fourier integral operator that is not associated to a local canonical graph is the pull back operator γ_H (or sometimes r_H) by an embedding $\iota_H : H \to M$ of a submanifold. Thus,

$$\gamma_H f(q) = f|_H = \text{the restriction of } f \text{ to } H. \tag{7.68}$$

In the case of a submanifold $H \subset \mathbb{R}^n$, the restriction operator $\gamma_H \colon C_c(\mathbb{R}^n) \to C_c(H)$ has the Fourier integral representation

$$\gamma_H f(q) = (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i\langle q-y,\eta\rangle} f(y)\, dy d\eta, \qquad q \in H. \tag{7.69}$$

The linear phase $\langle q-y, \eta\rangle$ is the same as for the identity operator but q is restricted to H. Its canonical relation is

$$\Gamma_H = \{(q, \xi|_{TH}, q, \xi) \colon \xi \in T^*_q M, q \in H\} \subset T^*H \times T^*M. \tag{7.70}$$

It is somewhat complicated by the fact that if $\xi \perp T_q H$ is conormal to H, then its restriction to TH is zero, i.e., $\xi|_{TH} = 0$. Thus there are co-vectors of the form $(q, 0, q, \nu)$ in the canonical relation, unlike the canonical relation of the space time geodesic flow.

A combination of the two examples is important in eigenfunction restriction theorems. Let $H \subset M$ be a hypersurface and let $B \in \Psi^0(M)$ be a poly-homogeneous pseudo-differential operator whose symbol vanishes in a conic neighborhood of T^*H. Then define $W \colon C(M) \to C(\mathbb{R} \times H)$ by

$$W \colon f \in C(M) \to \gamma_H B U(t) f \in C(\mathbb{R} \times H). \tag{7.71}$$

If the symbol of B vanishes on covectors T^*H cotangent to H then W is a Fourier integral operator with local canonical graph. (If one does not put in the cutoff B, W is a more complicated degenerate Fourier integral operator with one-sided folds.) W has the canonical relation

(7.72)
$$\Gamma_W = \{(t, \tau, q, \xi|_H, G^t(q, \xi)) \colon (q, \xi) \in T^*_H M, |\xi| = \tau\} \subset T^*(\mathbb{R} \times H)\backslash 0 \times T^*M\backslash 0.$$

Below is the associated diagram

$$\begin{array}{ccccc} & & \Gamma_W \subset T^*(\mathbb{R} \times H) \times T^*M & & \\ & \swarrow^{\pi} & & \searrow^{\rho} & \\ T^*(R \times H) & & & & T^*M \end{array}$$

The left projection is 2-1 except along the set $\{(t, \tau, q, \xi|_H, q, \xi) : |\xi| = \tau = |\xi|_H|\}$ (i.e., $\xi \in T^*H$), where it has a fold singularity. This set is removed from the wave front relation by the cutoff B.

7.3. Semi-classical Fourier integral operators

Semi-classical Lagrangian distributions are defined by oscillatory integrals (see [**Du1, Du2**]),

$$u(x, \hbar) = I_\hbar(a, \varphi) := \hbar^{-N/2} \int_{\mathbb{R}^N} e^{\frac{i}{\hbar}\varphi(x,\theta)} a(x, \theta, \hbar)\, d\theta, \tag{7.73}$$

where the amplitude $a(x,\theta,\hbar)$ is a semi-classical symbol admitting an asymptotic expansion of the form

$$a(x,\theta,\hbar) \sim \sum_{k=0}^{\infty} \hbar^{\mu+k} a_k(x,\theta). \tag{7.74}$$

The order of $u(x,\hbar)$ is μ. There are many possible assumptions on the behavior of a in θ and we refer to [**Zw, GSj, DSj**] for symbol classes of various kinds.

The canonical relation and symbol of a semi-classical Fourier integral operator parallels that in the homogeneous case in §7.2.2. The two main differences are that canonical relations are not homogeneous, and the symbol includes an oscillatory factor. We briefly recapitulate the theory.

Let

$$C_\varphi = \{(x,\theta) : \varphi'_\theta = 0\}.$$

The phase is called non-degenerate if $d(\frac{\partial\varphi}{\partial\theta_1}), \dots, d(\frac{\partial\varphi}{\partial\theta_N})$ are independent on C_φ. In this case,

$$\varphi'_\theta := \left(\left(\frac{\partial\varphi}{\partial\theta_1}\right), \dots, \left(\frac{\partial\varphi}{\partial\theta_N}\right)\right) : X \times \mathbb{R}^N \to \mathbb{R}^N \tag{7.75}$$

is locally a submersion near 0 and C_φ is a manifold of codimension N whose tangent space is $\ker D\varphi'_\theta$. If φ is non-degenerate, then the Lagrange map

$$\iota_\varphi(x,\theta) = (x, \varphi'_x(x,\theta))\colon C_\varphi \to \Lambda_\varphi \subset T^*X \tag{7.76}$$

is an immersion whose image is a Lagrangian submanifold denoted Λ_φ. If the order of (7.73) is μ, we say that $u(x,\hbar)$ is an oscillatory integral associated with Λ_φ and lies in $\mathcal{O}^\mu(X,\Lambda)$, the space of oscillatory integrals of order μ.

As in the case of homogeneous Lagrangian distributions, $u(x,\hbar)$ may be defined in several different ways as an oscillatory integral and one would like to define invariants which are independent of the particular expression. One is the Lagrangian submanifold Λ_φ above. The second is the symbol σ_u, which is a homogeneous section $\sigma_u \in S^\mu(\Lambda, \Omega^{\frac{1}{2}} \otimes \mathcal{M})$ of order μ of the bundle of half-densities (tensor the Maslov bundle) on Λ. We do not intend to give a systematic exposition of the theory here but only to give some heuristic principles and useful methods for calculating symbols of oscillatory integrals that arise when studying eigenfunctions. We refer to [**Du1, Du2, GSj, DSj, GuSt1**] for more systematic and precise expositions.

Analogous to (7.15), the delta function on C_φ is

$$d_{C_\varphi} := \left.\frac{|d\lambda|}{|D(\lambda,\varphi'_\theta)/D(x,\theta)|}\right|_{C_\varphi}. \tag{7.77}$$

The principal symbol of the oscillatory integral (7.73) is defined to be the pushforward to Λ_φ of the leading order part of the amplitude times $\sqrt{d_{C_\varphi}}$, i.e.,

$$\sigma_{u_\hbar} := \iota_{\varphi*} a_0 e^{i\lambda\varphi} \sqrt{d_{C_\varphi}}. \tag{7.78}$$

The only difference to the homogeneous case is that a is a semi-classical symbol rather than a homogeneous one and that the symbol includes the oscillating factor $e^{i\lambda\varphi}$.

Again, as discussed above, this definition of the principal symbol is "extrinsic," i.e., it makes use of a specific representation (7.73) and the embedding ι_φ. It is verified in [**Ho2, Du1, GuSt1**] and elsewhere that the principal symbol is independent of these choices. However it is obviously desirable to have an intrinsic formula for the principal symbol as a half-density on Λ_φ.

7.3.1. Semi-classical kernels associated to projectible Lagrangians. The simplest kernels have the WKB form (7.1) first studied by van Vleck [**V**]:

$$U(q,Q) = (2\pi h)^{-n/2}\sqrt{\det\left(\frac{\partial^2 S}{\partial q \partial Q}\right)} e^{\frac{i}{\hbar}S(q,Q)} \tag{7.79}$$

with no phase variables. Here, we use the physics notation (q,p) for a point in phase space; we also write (q,Q) instead of (x,y). If q and p are classical conjugate variables, then the kernel of the semi-classical Fourier transform $\langle q|p\rangle$ is

$$\langle q|p\rangle = (2\pi i\hbar)^{-n/2} e^{\frac{i}{\hbar}\langle p,q\rangle}. \tag{7.80}$$

More generally, if $(q,p)\to(Q,P)$ is a canonical transformation, then the map is quantized by the following semi-classical unitary operators:

$$\begin{cases} \langle q|P\rangle = \left(\dfrac{1}{2\pi i\hbar}\dfrac{\partial^2 S_2(q,P)}{\partial q\partial P}\right)^{\frac12} \exp\dfrac{iS_1(q,P)}{\hbar}, \\ \langle q|Q\rangle = \left(\dfrac{1}{2\pi i\hbar}\dfrac{\partial^2 S_1(q,Q)}{\partial q\partial Q}\right)^{\frac12} \exp\dfrac{iS_1(q,Q)}{\hbar}, \\ \langle p|P\rangle = \left(\dfrac{1}{2\pi i\hbar}\dfrac{\partial^2 S_3(p,P)}{\partial p\partial P}\right)^{\frac12} \exp\dfrac{iS_3(p,P)}{\hbar}, \end{cases} \tag{7.81}$$

where S_1, S_2, S_3 are the generating functions of the canonical transformation. All of these kernels have symbol 1.

7.3.2. Quadratic phases. In the semi-classical setting, linear Hamiltonians $ax + b\xi$ on $T^*\mathbb{R}^n$ generate 'phase space translations,' known as the Heisenberg group. Quadratic Hamiltonians generate linear Hamiltonian flows fixing the origin. Together, the linear and quadratic functions on $T^*\mathbb{R}^n$ generate a Lie algebra, the associated group being the semi-direct product of the Heisenberg and symplectic groups.

This section is a concerned with quadratic phase functions and the metaplectic representation, a unitary representation of the metaplectic group (the double cover of the symplectic group) as exponentials e^{itH} of quadratic operators H on $\mathbb{R}^n$, which form the elements of the metaplectic representation §7.3.2. For special values of t is impossible to represent them in the WKB form (7.79). The key point is that the Lagrangian submanifold Λ_S associated to the kernel must be projectible to the base, i.e. the natural projection $\pi\colon \Lambda \subset T^*\mathbb{R}^n \to \mathbb{R}^n$ must be a diffeomorphism.

The model semi-classical Fourier integral operators are the quantizations of linear symplectic transformations of $\mathbb{R}^n$ defined by the metaplectic representation. Systematic expositions may be found in [**Fo, GuSt1**].

Let $T^*\mathbb{R}^n \simeq \mathbb{R}^n \oplus \mathbb{R}^n$ with coordinates (x,ξ) and let $\mathcal{J}$ be the standard complex structure. A linear symplectic transformation of $T^*\mathbb{R}^n \simeq \mathbb{R}^n \oplus \mathbb{R}^n$ may be put in block-form:

$$\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{7.82}$$

Then $\mathcal{A}$ is a symplectic transformation, $\mathcal{A} \in Sp(n,\mathbb{R})$, if and only if $\mathcal{A}^*\mathcal{J}A = \mathcal{J}$.

The metaplectic representation $\mu : Sp(n,\mathbb{R}) \to \mathcal{U}(L^2(\mathbb{R}^n))$ quantizes linear symplectic transformations as unitary operators on $L^2(\mathbb{R}^n)$. In the Schrödinger position representation, the metaplectic representation has the form

$$\mu \begin{pmatrix} A & B \\ C & D \end{pmatrix} f(x) = \hbar^{-n/2} i^{n/2} \sqrt{\det\left(\frac{\partial^2 S(x,y)}{\partial x \partial y}\right)} \int_{\mathbb{R}^n} e^{\frac{2\pi i}{\hbar} S(x,y)} f(y)\, dy, \tag{7.83}$$

with
(7.84)
$$S'(x,y) = -\frac{1}{2}\left(xDB^{-1}x + yB^{-1}x - \frac{1}{2} yB^{-1}Ay\right), \qquad \det B^{-1} = \det S''_{x,y}.$$

This is the WKB form of van Vleck (7.79). Here we use [**Fo**, Theorem 4.53] together with

$$(\det B)^{-\frac{1}{2}} = \sqrt{\det\left(\frac{\partial^2 S(x,y)}{\partial x \partial y}\right)}. \tag{7.85}$$

For instance, the quantization $\begin{pmatrix} \cos\theta I & \sin\theta I \\ -\sin\theta I & \cos\theta I \end{pmatrix}$ of the classical isotropic harmonic oscillator flow on $\mathbb{R}^n$ has phase $S(t,x,y) = \cot t(|x|^2 + |y|^2) - \frac{2}{\sin t} x \cdot y$.

In the Schrödinger (momentum) representation, the metaplectic representation with Planck constant $\hbar$ has a form analogous to that of homogeneous Fourier integral operators in §7.2.4:

$$\mu \begin{pmatrix} A & B \\ C & D \end{pmatrix} f(x) = \hbar^{-\frac{n}{2}} \sqrt{\det\left(\frac{\partial^2 S(x,\xi)}{\partial x \partial \xi}\right)} \int_{\mathbb{R}^n} e^{\frac{2\pi i}{\hbar} S(x,\xi)} \hat{f}(\xi)\, d\xi, \tag{7.86}$$

with

$$S(x,\xi) = -\frac{1}{2}\left(xCA^{-1}x + \xi A^{-1}x + \frac{1}{2}\xi A^{-1}B\xi\right), \qquad A^{-1} = \left(\frac{\partial^2 S(x,\xi)}{\partial x \partial \xi}\right). \tag{7.87}$$

Here, we use the formula of [**Fo**, Theorem 4.51] together with

$$(\det A)^{-\frac{1}{2}} = \sqrt{\det\left(\frac{\partial^2 S(x,\xi)}{\partial x \partial \xi}\right)}. \tag{7.88}$$

In the momentum representation, the phase of the isotropic harmonic oscillator is $S(t,x,\xi) = \tan t(|x|^2 + |\xi|^2) - \frac{2}{\cos t} x \cdot \xi$.

7.3.3. Historical remarks on semi-classical propagators. Propagators are unitary groups $U_\hbar(t) = e^{it\hbar \hat{H}_\hbar}$ generated by a pseudo-differential Hamiltonian. If one solves the time-dependent Schrödinger equation by the ansatz $U_\hbar(t) = A_t e^{\frac{i}{\hbar} S_t}$ one gets the transport equation

$$\nabla \cdot (A^2 \nabla S) + \partial_t (A^2) = 0. \tag{7.89}$$

Van Vleck's theorem of 1928 [**V**] is that

$$A^2 = D(q,t;q_0,t_0) = \left(\frac{\partial p_0(q,t;q_0,t_0)}{\partial q}\right) = (-1)^n \det\left(\frac{\partial S^2}{\partial q_j \partial q_k^0}\right). \tag{7.90}$$

In a similar spirit, W. Pauli expressed the propagator for the Schrödinger equation in the form

$$K_c(x,\tau,y,0) = (i\hbar)^{-n/2} \sqrt{D} e^{\frac{i}{\hbar} S(x,\tau;y,0)}, \tag{7.91}$$

where

$$D = D(x,\tau;y,0) = (-1)^n \det\left(\frac{\partial S^2}{\partial x_j \partial y_k}\right). \tag{7.92}$$

De Witte-Morette [**DeW**, Proposition 2] proved that

$$-J^{\alpha\beta}(t_b,t_a)\,\frac{\partial^2 S(a,b)}{\partial a^\alpha \partial b^\alpha} = -\delta^\alpha_\beta. \tag{7.93}$$

Here $(J^{\alpha\beta})$ is the matrix of the Jacobi fields:

$$\begin{cases} J(t_a,t_a) = 0, \\ \nabla_t J(t_a,t_a) = A^{-1}(t_a), \end{cases} \quad \begin{cases} K(t_a,t_a) = A^{-1}(t_a), \\ \nabla_t K(t_a,t_a) = 0, \end{cases} \tag{7.94}$$

where

$$A_{\alpha\beta} = -\frac{\partial^2 L}{\partial \dot q^\alpha \partial \dot q^\beta}.$$

$J(t,t_a)$ resp. $J(t_b,t_a)$ consists of Jacobi fields through stationary paths keeping the initial point a resp. b fixed. Similarly, $K(t,t_a)$, resp. $K(t,t_b)$ are variations keeping v_a resp. v_b fixed. One has

$$J^{\alpha\beta}(t_a,t_b) = -J^{\beta\alpha}(t_b,t_a), \quad J(t_b,t_a)M(t_a,t_b) = -I.$$

Using these relations one can show that the van Vleck determinants propagate as a cocycle,

$$D(q_2,q_1)D(q_1,q_0) = D(q_2,q_0)\det S. \tag{7.95}$$

Here, S is the Hessian of $S(q_2,1_1) + S(q_1,q_0)$ at the critical point q_1.

7.3.4. Fock's unitarity proof. It was observed by Fock [**F**] that (7.79) is unitary to leading order. The discussion is parallel to, and simpler than, the homogeneous case of §7.2.6. We sketch his proof since its heuristics are illuminating and useful. If we define P by

$$S(Q,q) - S(q,Q') = (Q-Q')\cdot P \tag{7.96}$$

and change variables $q \to P$, the volume form changes by

$$dq = \frac{dq}{dP}\,dP, \qquad \frac{dq}{dP} = \left[\det\left(\frac{\partial^2 S}{\partial q \partial Q}\right)\right]^{-1} \tag{7.97}$$

and

(7.98)

$$\int U(Q,q)U^*(q,Q')\,dq = (2\pi h)^{-n}\int \sqrt{\det\left(\tfrac{\partial^2 S}{\partial Q \partial q}\right)}\sqrt{\det\left(\tfrac{\partial^2 S}{\partial q \partial Q'}\right)} e^{\frac{i}{\hbar}(S(Q,q)-S(q,Q')}\,dq$$

$$= (2\pi h)^{-n}\int \sqrt{\det\left(\tfrac{\partial^2 S}{\partial Q \partial q}\right)}\sqrt{\det\left(\tfrac{\partial^2 S}{\partial q \partial Q'}\right)} e^{\frac{i}{\hbar}((Q-Q')\cdot P)}\,dq \tag{7.99}$$

$$= (2\pi h)^{-n}\int \sqrt{\det\left(\tfrac{\partial^2 S}{\partial Q \partial q}\right)}\sqrt{\det\left(\tfrac{\partial^2 S}{\partial q \partial Q'}\right)} \tag{7.100}$$

$$\times \det\left(\tfrac{dq}{dP}\right) e^{\frac{i}{\hbar}((Q-Q')\cdot P)}\,dP \tag{7.101}$$

$$= (2\pi h)^{-n}\int e^{\frac{i}{\hbar}((Q-Q')\cdot P)}\,dP \tag{7.102}$$

$$\simeq \delta(Q-Q'), \tag{7.103}$$

since the integral is equal modulo $\mathcal{O}(\hbar)$ by the value of the amplitude at $Q = Q'$.

7.4. Principal symbol, testing and matrix elements

As mentioned above, it is desirable to have an "intrinsic definition" of the principal symbol (7.16) of homogeneous, resp. (7.78) semi-classical Fourier integral operators as half-densities (tensor Maslov factors) on the associated Lagrangian submanifold Λ_φ. The symbol arises naturally in the leading order term of matrix elements of semi-classical Fourier integral operators relative to oscillatory test functions. It also arises in the leading coefficients of the singularities of the trace of the wave group and related evolution operators as in the foundational paper [**DuG**]. In view of the prominent role of matrix elements in this monograph we provide further background on principal symbols and oscillatory testing, following [**Ho2**] (pages 149-154, especially Theorem 3.2.4), [**Du1, Du2**] and [**GuSt1**, Chapter VII.4].

We begin with a remark on principal symbols in the presence of a Riemannian metric g. The metric endows TM with a volume form dV_g, and T^*M with the dual co-volume form dV_{g^*}. If $\{e_1, \dots, e_n\}$ is an orthonormal frame at T_xM, and if $\theta_1, \dots, \theta_n$ is the dual frame of T_x^*M, then $dV_g = \theta_1 \wedge \cdots \wedge \theta_n$ and $dV_{g^*} = e_1 \wedge \cdots \wedge e_n$, where we identify $V^{**} = V$ for any vector space. The vertical spaces T_x^*M form the "vertical polarization" of T^*M, i.e., a foliation by Lagrangian submanifolds. They are equipped with volume forms. Hence, given any Lagrangian submanifold $\Lambda \subset T^*M$ which is transverse to the vertical, a half-density $\sqrt{V_\Lambda}$ on Λ induces a half-density $\sqrt{V_\Lambda} \otimes \sqrt{dV_{g^*}}$ on $T_\lambda(T^*M)$ for $\lambda \in \Lambda$. Since $T_\lambda\Lambda$ and $T_\lambda T_x^*M$ are transverse (where $x = \pi(\lambda)$), there exist dual bases $e_1, \dots, e_n$ resp. $f_1, \dots, f_n$ of these two vector spaces such that $\omega_\lambda(e_i, f_j) = \delta_{ij}$ where ω is the standard symplectic form. Then $\sqrt{V_\Lambda(e_1, \dots, e_n)} \otimes \sqrt{dV_{g^*}(f_1, \dots, f_n)}$ is a scalar quantity which determines $\sqrt{V_\Lambda}$ in the presence of the metric g. It does not depend on the choice of dual bases $\{e_j, f_k\}$. Calculating this scalar gives an intrinsic formula for the principal symbol (albeit one depending on a choice of metric).

In the next section we consider integrals of (7.73) against oscillatory functions and obtain explicit formulae for the symbol in terms of non-vertical Lagrangian submanifolds transverse to Λ_φ. To pursue the aim of obtaining an intrinsic formula, we record a lot of calculations of symbols and expansions drawn from [**Du1, DSj, GSj**] and elsewhere.

7.4.1. Symbols when the Lagrangian is projectible. Recall that Λ is projectible if $\pi : \Lambda \to M$ is a diffeomorphism. Then Λ is the graph of a closed 1-form and on a simply connected open set $\Lambda = \text{graph}(d\psi)$ for some smooth function ψ. In this case, a semi-classical Lagrangian distribution associated to Λ can be expressed in the form

$$\chi(x) e^{-\frac{i}{\hbar}\psi(x)} |dV_g|^{\frac{1}{2}}, \qquad \psi \in C^\infty(M). \tag{7.104}$$

In this case, the symbol is computed as follows:

LEMMA 7.1. *The symbol* $\sigma_{\chi e^{-i\hbar^{-1}\psi}}(x_0, \xi_0)$ *of* $\chi e^{-i\hbar^{-1}\psi} |dV_g|$ *is the pull back*

$$\pi_\psi^* e^{-i\hbar^{-1}\psi} |dV_g|^{\frac{1}{2}}$$

of this oscillatory half-density to the graph of $d\psi$.

7.4.1.1. *Symbol when Λ is momentum projectible.* We say that Λ is momentum projectible if there exist local coordinates x with dual symplectic coordinates ξ so that $\Lambda = \{(H'(\xi), \xi)\} = \text{graph}(dH(\xi))$ is the graph of a closed (locally exact) 1-form over the vertical axis. Of course, the vertical axis is not intrinsically defined and requires the introduction of the local coordinates. This is especially important for homogeneous Lagrangian submanifolds.

We consider phases $\varphi(x, \xi) = x \cdot \xi - H(\xi)$ satisfying $\varphi'_\xi = 0 \iff x = H'(\xi)$ and $\iota_\varphi(x, \xi) = (x, \varphi'_x) = (H'(\xi), \xi)$. Suppose that the Lagrangian distribution is expressed in the form

$$u_\hbar(x) = (2\pi\hbar)^{-\frac{n}{2}} \int e^{\frac{i}{\hbar}(\langle x,\xi\rangle - H(\xi)} a(\xi, \hbar)\, d\xi |dx|^{\frac{1}{2}}. \tag{7.105}$$

This is a special case of the Fourier transform representation in §7.2.4 with $\varphi(x, \xi) = H(\xi)$.

LEMMA 7.2. *Then the half-density part of the principal symbol is given by*

$$\sigma(u_\hbar)|_{(H'(\xi),\xi)} = e^{\frac{i}{\hbar}(\langle H'(\xi),\xi\rangle - H(\xi))} a_0(\xi) |d\xi|^{\frac{1}{2}} \tag{7.106}$$

on $\Lambda_H = \{(H'(\xi), \xi) \colon \xi \in \mathbb{R}^n\}$.

In general, Λ is neither globally projectible nor momentum projectible. The symbol can still be calculated in a way similar to that in the linear case in §7.2.3 by using a mixed representation with some x and some ξ coordinates, i.e., by using a partial Fourier transform.

In a sense we can reduce the calculation of the principal symbol to the linear case at a point $(x_0, \xi_0) \in \Lambda$ by constructing the "osculating" oscillatory integral at this point, replacing the phase by the quadratic part of its Taylor expansion and the amplitude by its leading term at the critical point. The symbol at $(x_0, \xi_0) \in \Lambda$ then coincides with the symbol of the osculating oscillatory integral. Indeed, the symbol only depends on the osculating data.

7.4.2. Matrix elements and oscillatory testing. As mentioned above, the symbol arises naturally in the leading order term of matrix elements of semi-classical Fourier integral operators relative to oscillatory test functions. Since matrix elements relative to oscillatory functions are simply inner products of a Fourier integral kernel with a tensor product of two oscillatory functions on $M \times M$, we only go through the details of the inner product of an oscillatory integral against an oscillatory test function of the form,

$$\chi(x) e^{-\frac{i}{\hbar}\psi(x)}, \qquad \psi \in C^\infty(M). \tag{7.107}$$

It suffices to consider oscillatory functions locally defined by $\psi(x) = \langle x, \xi_0\rangle$. Then,

$$\langle u(x, \hbar), \chi(x) e^{-\frac{i}{\hbar}\psi(x)}\rangle = \hbar^{-N/2} \int e^{\frac{i}{\hbar}\varphi(x,\theta)} a(x, \theta, \hbar) \chi(x) e^{-\frac{i}{\hbar}\psi(x)}\, d\theta dx. \tag{7.108}$$

The full integral has the phase

$$\varphi(x, \theta) - \psi(x) \in C^\infty(M \times \mathbb{R}^N) \tag{7.109}$$

and we assume at first that it has only non-degenerate critical points as a function of (θ, x). The phase (7.109) has a critical point at (x_0, θ_0) if and only if

$$d_x\varphi(x_0, \theta_0) = d\psi(x_0), \quad d_\theta\varphi(x_0, \theta_0) = 0, \tag{7.110}$$

which means that graph$(d\psi)$ intersects Λ_φ at (x_0, ξ_0), i.e.,

$$d\psi(x_0) \in \Lambda_\varphi \cap T^*_{x_0}M. \tag{7.111}$$

The Hessian of the phase (7.109) in local coordinates is

$$Q_{\varphi,\psi} = d^2_{(x,\theta)}(\varphi - \psi) = \begin{pmatrix} d^2_{xx}(\varphi-\psi) & d^2_{\theta x}\varphi \\ d^2_{x\theta}\varphi & d^2_{\theta\theta}\varphi \end{pmatrix}, \tag{7.112}$$

which is non-degenerate if and only if the rank of the right column $\begin{pmatrix} d^2_{\theta x}\varphi \\ d^2_{\theta\theta}\varphi \end{pmatrix}$ equals N. Indeed, if the rank of the full matrix is $N + \dim M$ then the last N columns must be linearly independent.

LEMMA 7.3. *$Q_{\varphi,\psi}$ is non-degenerate if and only if the intersection of* graph$(d\psi)$ *with Λ_φ is transversal at (x_0, θ_0).*

PROOF. We need to see that

$$T_{(x_0,\xi_0)}\Lambda_\varphi \cap T_{(x_0,\xi_0)}\,\text{graph}(d\psi) = \ker Q_{\varphi,\psi}. \tag{7.113}$$

But for $(\delta x, \delta\theta) \in T_{(x_0,\xi_0)}C_\varphi$, we have

$$Q_\psi \begin{pmatrix} \delta x \\ \delta\theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff (\delta x, d^2_{xx}\varphi\delta x + d_\theta d_x \varphi \delta\theta) = (\delta x, d^2_{xx}\psi\delta x). \tag{7.114}$$

□

Since $\dim \Lambda_\varphi = \dim \text{graph}(d\psi) = \dim M$, transversality is equivalent to

$$T_{(x_0,\xi_0)}\Lambda_\varphi \oplus T_{(x_0,\xi_0)}\,\text{graph}(d\psi) = T_{(x_0,\xi_0)}T^*M. \tag{7.115}$$

It follows that C_φ must be a smooth manifold through (x_0, θ_0) of dimension $\dim M$ and that

$$\iota_\varphi \colon C_\varphi \to \Lambda_\varphi \tag{7.116}$$

is an immersion. The following Proposition is proved in [**Ho2**, (3.2.17) and Theorem 3.2.4] and [**Du1**].

PROPOSITION 7.4. *Assume $\psi(x_0) = 0$, $\chi(x_0) = 1$ and that (x_0, ξ_0) is the only critical point in the support of χ, and that the intersection of* graph$(d\psi)$ *with Λ_φ is transversal at (x_0, θ_0). Then the stationary phase expansion of* (7.108) *is given by* (*cf.* (7.112))

$$\langle u(x,\hbar), \chi(x)e^{-\frac{i}{\hbar}\psi(x)}\rangle \simeq e^{\frac{i}{\hbar}(\varphi(x_0,\theta_0)-\psi(x_0))}\left(\frac{2\pi}{\hbar}\right)^{n/2}|\det Q_{\varphi,\psi}|^{-\frac{1}{2}} \times e^{i\frac{\pi}{4}\,\text{sgn}\,Q_{\varphi,\psi}} a_0(x_0,\xi_0)\hbar^{-\mu}[1 + O(\hbar)]. \tag{7.117}$$

We remark that if graph$(d\psi) \cap \Lambda_\varphi = \emptyset$, then the integral would be rapidly decaying. On the other hand, it increases in order of magnitude in $\hbar^{-1}$ as the dimension of the intersection increases. When the two Lagrangian submanifolds coincide, the inner product of two oscillatory integrals in the same class is calculated asymptotically in terms of the inner product of the symbols; see §7.4.5.

The phase φ in the oscillatory factor has differential equal to the restriction of the action form $\alpha = \xi \cdot dx$ to Λ_φ, since $d_x\varphi(x_0,\xi_0) = \xi_0$ and since $d_\theta\varphi(x_0,\xi_0) = 0$. The leading term above only depends on ψ through $d^2_x\psi(x_0)$, i.e. on the tangent space to graph$(d\psi)$ at x_0.

We now interpret the expansion in terms of the symbol σ_u of $u(x;\hbar)$ and the symbol σ_ψ of $\chi(x)e^{-\frac{i}{\hbar}\psi(x)}$, following [**Du2**] and [**GuSt1**, p. 411-412]. First we develop the linear algebra. The vertical tangent space $V_{x_0,\xi_0} \subset T_{x_0,\xi_0}T^*M$ is invariantly defined as the kernel of $D\pi$ where $\pi : T^*M \to M$ is the natural projection. Since the fibers are vector spaces, we may identify $V_{x_0,\xi_0} \simeq T_{(x_0,\xi_0)}T^*_{x_0}M$. The Riemannian volume form induces a *vertical volume form* Vol_V , which is the quotient

$$\mathrm{Vol}_V\,|_{V_{x_0,\xi_0}} = \frac{d\,\mathrm{Vol}_\omega}{\pi^*\,\mathrm{Vol}_{T_{x_0}M}} = \sqrt{\det(g^{ij}(x_0))}d\xi$$

of the canonical volume form $dvol_\omega$ on $T_{(x_0,\xi_0)}T^*M$ by the pull back of $dV_g = \mathrm{Vol}_{T_{x_0}M}$ under the natural projection $D\pi : T_{(x_0,\xi_0)}T^*M \to T_{(x_0)}M$.

The tangent space to $\mathrm{graph}(d\psi)$ is

$$T\,\mathrm{graph}(d\psi) = \{(\delta x, d_x^2\psi\delta x)\}. \tag{7.118}$$

We denote by π_ψ the natural projection restricted to $\mathrm{graph}(d\psi)$,

$$\pi_\psi\colon\ \mathrm{graph}(d\psi) \to M, \quad \pi_\psi(x,\xi) = x, \tag{7.119}$$

Since a graph is transverse to the vertical, we have the decomposition,

$$T_{(x_0,\xi_0)}T^*M = T_{(x_0,\xi_0)}\,\mathrm{graph}(d\psi) \oplus T_{(x_0,\xi_0)}T^*_{x_0}M. \tag{7.120}$$

of the symplectic vector space into a sum of transverse Lagrangian subspaces along the graph, giving a horizontal complement to the vertical. Let p_ψ be the vertical projection along $T(\mathrm{graph}\,d\psi)$ with respect to (7.120):

$$p_\psi\colon T_{(x_0,\xi_0)}T^*M \to T_{(x_0,\xi_0)}T^*_{x_0}M. \tag{7.121}$$

Let (x,ξ) be local symplectic coordinates induced by local coordinates x on M, and write a tangent vector to T^*M by $(\delta x, \delta\xi)$. The decomposition (7.120) has the form

$$(\delta x, \delta\xi) = (\delta x, d_x^2\psi\delta x) + (0, \delta\xi - d_x^2\psi\delta x), \tag{7.122}$$

and so the vertical projection with respect to the decomposition is give by

$$p_\psi\begin{pmatrix}\delta x\\ \delta\xi\end{pmatrix} = \begin{pmatrix}0\\ \delta\xi - d_x^2\psi\cdot\delta x\end{pmatrix} \in T_{(x_0,\xi_0)}T^*_{x_0}M. \tag{7.123}$$

The complementary projection

$$q_\psi\colon T_{(x_0,\xi_0)}T^*M \to T_{(x_0,\xi_0)}\,\mathrm{graph}(d\psi) \tag{7.124}$$

coming from the splitting (7.120) is defined by

$$q_\psi\begin{pmatrix}\delta x\\ \delta\xi\end{pmatrix} = \begin{pmatrix}\delta x\\ d_x^2\psi\cdot\delta x\end{pmatrix}. \tag{7.125}$$

We note that $p_\psi + q_\psi = \mathrm{Id}$, and that

$$q_\psi\begin{pmatrix}\delta x\\ \delta\xi\end{pmatrix} = D(d\psi)\circ d\pi\begin{pmatrix}\delta x\\ \delta\xi\end{pmatrix}. \tag{7.126}$$

The tangent space to Λ_φ is the image of $D\iota_\varphi$ on TC_φ, i.e., vectors of the form

$$D\iota_\varphi(\delta_x, \delta_\theta) = (\delta_x, \varphi''_{xx}\delta_x + \varphi''_{x\theta}\delta_\theta), \qquad (\delta x, \delta\theta) \in TC_\varphi. \tag{7.127}$$

We recall that

$$TC_\varphi = \begin{pmatrix}\delta x\\ \delta\theta\end{pmatrix} : D_{x,\theta}\varphi'_\theta\begin{pmatrix}\delta x\\ \delta\theta\end{pmatrix} = 0 \iff \varphi''_{\theta x}\delta x + \varphi''_{\theta\theta}\delta\theta = 0. \tag{7.128}$$

Since $T(\operatorname{graph} d\psi)$ is transverse both to the vertical fiber and to Λ_φ,

$$p_{\psi,\varphi} := p_\psi|_{T_{(x_0,\xi_0)}\Lambda_\varphi} : T_{(x_0,\xi_0)}\Lambda_\varphi \to T_{(x_0,\xi_0)}T^*_{x_0}M \tag{7.129}$$

is an isomorphism. The pullback of the vertical volume form Vol_V by $p_{\psi,\varphi}$ defines a volume form on $T_{(x_0,\xi_0)}\Lambda_\varphi$ denoted by

$$\mathrm{Vol}_{\psi,\varphi} := p^*_{\psi,\varphi}\,\mathrm{Vol}_V = (p_\psi|_{T_{(x_0,\xi_0)}\Lambda_\varphi})^*\,\mathrm{Vol}_V\,. \tag{7.130}$$

From (7.112) it follows that

$$Q_{\varphi,\psi} = \begin{pmatrix} p_\psi D\iota_\varphi \\ d_{(x,\theta)}d_\theta\varphi \end{pmatrix} : T_{(x,\theta)}(M\times\mathbb{R}^N) \to T_{\iota_\varphi(x,\theta)}T^*_xM\times\mathbb{R}^N. \tag{7.131}$$

Note that $\mathbb{R}^N$ is equipped with the volume form $d\theta_1\wedge\cdots\wedge d\theta_N := \mathrm{Vol}_{\mathbb{R}^N}$, T_xM has the Riemannian volume form and T^*_xM has the dual Riemannian volume form. Hence it makes sense to take the determinant of $Q_{\varphi,\psi}$.

PROPOSITION 7.5. *The principal symbol $\sigma_u(x_0,\xi_0) = e^{\frac{i}{\hbar}\varphi}a_0\sqrt{d_{C_\varphi}}$ is a $\frac{1}{2}$-density on $T_{(x_0,\xi_0)}\Lambda_\varphi$ (depending on the choice of the density dV_g on $T_{x_0}M$) and*

$$\iota_{\varphi*}\sqrt{d_{C_\varphi}} = |\det Q_{\varphi,\psi}|^{-\frac{1}{2}}|Vol_{\psi,\varphi}|^{\frac{1}{2}}. \tag{7.132}$$

It is independent of ψ.

PROOF. Both $\iota_{\varphi*}d_{C_\varphi}$ and $\mathrm{Vol}_{\psi,\varphi}$ are volume forms on Λ_φ, and we would like to show that

$$\det Q_{\varphi,\psi} = \frac{\mathrm{Vol}_{\psi,\varphi}}{d_{C_\varphi}}. \tag{7.133}$$

To prove this, we identify $T_{(x_0,\xi_0)}\Lambda_\varphi = TC_\varphi$ and consider volume forms on $M\times\mathbb{R}^N$. We express the Lebesgue density $dV_g\wedge d\theta$ in two different ways. By definition of C_φ,

$$d\frac{\partial\varphi}{\partial\theta_1}\wedge\cdots\wedge d\frac{\partial\varphi}{\partial\theta_N}\wedge d_{C\varphi} = dV_g\wedge d\theta. \tag{7.134}$$

The second is

$$\mathrm{Vol}_V\otimes\mathrm{Vol}_{\mathbb{R}^N} = dV_g\wedge d\theta. \tag{7.135}$$

Here, we identify the volume form on the vertical space with dV_{g^*} using the symplectic volume form $|\Omega^n|$:

$$|dV_g|\otimes\mathrm{Vol}_V = |\Omega^n|. \tag{7.136}$$

Now let $e_1,\dots,e_n\in\ker(d_{(x,\theta)}d_\theta\varphi) = TC_\varphi$ and let $e_{n+1},\dots,e_{N+n}$ fill out to a basis of $T(M\times\mathbb{R}^N)$. Then

(7.137)

$$(p_\psi D\iota_\varphi)^*\,\mathrm{Vol}_V(e_1,\dots,e_n)(d_{(x,\theta)}d_\theta\varphi)^*\,\mathrm{Vol}_{\mathbb{R}^N}(e_{n+1},\dots,e_{N+n})$$

(7.138)

$$= \mathrm{Vol}_V((p_\psi D\iota_\varphi e_1,\dots,(p_\psi D\iota_\varphi e_n)\,\mathrm{Vol}_{R^N}(d_{(x,\theta)}d_\theta\varphi e_{n+1},\dots,d_{(x,\theta)}d_\theta\varphi e_{n+N})$$

(7.139)

$$= \mathrm{Vol}_V\otimes\mathrm{Vol}_{\mathbb{R}^N}(Q_{\varphi,\psi}e_1,\dots,Q_{\varphi,\psi}e_n,Q_{\varphi,\psi}e_{n+1},\dots,Q_{\varphi,\psi}e_{n+N})$$

(7.140)

$$= \det Q_{\varphi,\psi}d_{C_\varphi}(e_1,\dots,e_n)\,\mathrm{Vol}_{\mathbb{R}^N}(d_{(x,\theta)}d_\theta\varphi e_{n+1},\dots,d_{(x,\theta)}d_\theta\varphi e_{n+N}).$$

In the last line we used (7.134). Canceling $\mathrm{Vol}_{\mathbb{R}^N}(d_{(x,\theta)}d_\theta\varphi e_{n+1},\dots,d_{(x,\theta)}d_\theta\varphi e_{n+N})$ from the second and fourth lines gives

$$\det Q_{\varphi,\psi}d_{C_\varphi}(e_1,\dots,e_n) = \mathrm{Vol}_V((p_\psi D\iota_\varphi e_1,\dots,(p_\psi D\iota_\varphi e_n)$$

, as desired. □

7.4.3. Tensor product. Suppose that Λ_φ and Λ_ψ are transversal Lagrangian submanifolds. Then at an intersection point $\zeta\in\Lambda_\varphi\cap\Lambda_\psi$ one has $T_\zeta\Lambda_\varphi\oplus T_\zeta\Lambda_\psi = T_\zeta T^*M$, and the tensor product of the densities dC_φ on Λ_φ resp. dC_ψ form a density $dC_\varphi\otimes dC_\psi$ on $T^*_\zeta M$. The ratio $\frac{dC_\varphi\otimes dC_\psi}{\Omega}$ with the canonical density gives a scalar invariant of the pair (φ,ψ). By (7.133) and (7.130) it equals

$$\frac{\mathrm{Vol}_{\varphi,\psi}\otimes\mathrm{Vol}_{\psi,\varphi}}{\det Q_{\varphi,\psi}\det Q_{\psi,\varphi}\Omega} = \frac{p^*_{\psi,\varphi}\mathrm{Vol}_V\, p^*_{\varphi,\psi}\mathrm{Vol}_V}{\det Q_{\varphi,\psi}\det Q_{\psi,\varphi}\Omega}. \tag{7.141}$$

We note that $Q_{\varphi,\psi}$ (7.112) is symmetric in φ,ψ except when they are functions of different variables, i.e., (x,θ) resp. x.

7.4.4. Projectible cases. We return to §7.4.1 and reconsider the calculations of $Q_{\varphi,\psi}$ in projectible cases.

7.4.4.1. *Special case (i): Configuration projectible cases.* The first special case occurs when we test an oscillatory integral $u_\hbar(x) = A_\hbar e^{\frac{i}{\hbar}\varphi(x)}\sqrt{dV_g}$ relative to a projectible Lagrangian $\Lambda_\varphi = \mathrm{graph}(\mathrm{d}\varphi(\mathrm{x}))$ with an oscillatory test function relative to a second projectible Lagrangian $\Lambda_\psi = \mathrm{graph}(\mathrm{d}\psi(\mathrm{x}))$ transverse to Λ_φ. The phase has no "phase variables" θ, $C_\varphi = M$, $\iota_\varphi(x) = (x,\varphi'(x))$, $\iota_{\varphi*}d_{C_\varphi} = \pi^*_\varphi dV_g$, and (7.112) takes the form,

$$Q_{\varphi,\psi} = (\varphi''_{xx}-\psi''_{xx}). \tag{7.142}$$

Also

$$p_\psi\circ D\iota_\varphi\delta x = (d^2_x\varphi - d^2_x\psi)\cdot\delta x. \tag{7.143}$$

By definition, $e_j = D(x,\varphi'_x)e^0_j = e^0_j + \varphi''_{xx}e^0_j$ and

$$\mathrm{Vol}_{\psi,\varphi}(e_1,\dots,e_n) = Vol_V(p_\psi(e_1),\dots,p_\psi(e_n)). \tag{7.144}$$

7.4.4.2. *Special case (ii): Momentum projectible case.* The second special case occurs when we test an oscillatory integral

$$u_\hbar(x) = (2\pi\hbar)^{-\frac{n}{2}}\int e^{\frac{i}{\hbar}(\langle x,\xi\rangle - H(\xi)}a(\xi,\hbar)\,d\xi|dx|^{\frac12} \tag{7.145}$$

relative to a momentum projectible Lagrangian $\Lambda_H = \{(H'(\xi),\xi)\colon\xi\in\mathbb{R}^n\}$ with an oscillatory test function relative to a second projectible Lagrangian $\Lambda_\psi = \mathrm{graph}(d\psi(x))$ transverse to Λ_H. This is especially important for homogeneous Lagrangian submanifolds.

We consider phases $\varphi(x,\xi) = x\cdot\xi - H(\xi)$, so that $\varphi'_\xi = 0 \iff x = H'(\xi)$ and $\iota_\varphi(x,\xi) = (x,\varphi'_x) = (H'(\xi),\xi)$. Then

$$Q_{\varphi,\psi} = \begin{pmatrix}(-\psi)''_{xx} & I\\ I & -H''_{\xi\xi}\end{pmatrix}, \tag{7.146}$$

and (by the Schur determinant formula),

$$\det Q = \det(I - \det\psi''_{xx}\circ H''_{\xi\xi}). \tag{7.147}$$

The determinant is non-vanishing as long as the two Lagrangians are transversal. By Proposition 7.4, if the order of u equals zero, then to leading order

(7.148)
$$\langle u(x,\hbar), \chi(x) e^{-\frac{i}{\hbar}\psi(x)}\rangle \simeq e^{\frac{i}{\hbar}(\varphi(x_0,\xi_0)-\psi(x_0)}(\frac{2\pi}{\hbar})^{n/2}|\det(I - \det \psi''_{xx} \circ H''_{\xi\xi})|^{-\frac{1}{2}} e^{i\frac{\pi}{4}\operatorname{sgn} Q_{\varphi,\psi}} a_0(x_0,\xi_0).$$

By Lemma 7.2, the half-density part of the principal symbol is the half-density

$$\sigma(u_\hbar)|_{(H'(\xi),\xi)} = e^{\frac{i}{\hbar}(\langle H'(\xi),\xi\rangle - H(\xi)} a_0(\xi)|d\xi|^{\frac{1}{2}}. \tag{7.149}$$

By Lemma 7.1, the symbol $\sigma_{\chi e^{-i\hbar^{-1}\psi}}(x_0,\xi_0)$ of $\chi e^{-i\hbar^{-1}\psi}|dV_g|$ is the pull back $\pi_\psi^* e^{-i\hbar^{-1}\psi}|dV_g|^{\frac{1}{2}}$ of this oscillatory half-density to the graph of $d\psi$.

7.4.5. Inner product of two oscillatory functions. Above we took the inner product of two oscillatory integrals associated to transverse Lagrangian submanifolds. We now consider the inner product of two oscillatory integrals associated to the same Lagrangian submanifold. For the following, see [**Du1**, (1.3.15)].

LEMMA 7.6. *Suppose that u_1, u_2 are two oscillatory functions of order zero associated to Λ. Denote their symbols by σ_1, σ_2. Then*

$$\int_M u_1 \cdot \overline{u_2} = \int_\Lambda \sigma_1 \overline{\sigma_2} + O(\hbar). \tag{7.150}$$

Note that $\sigma_1\overline{\sigma_2}$ is a density on Λ.

7.5. Composition of half-densities on canonical relations in cotangent bundles

We should at least mention the composition theory of symbols. The composition of half-densities on Lagrangian submanifolds which intersect transversely (or cleanly) is based on a Lemma of symplectic linear algebra, i.e., on the composition of half-densities on Lagrangian subspaces. We follow [**DuG**] and [**GuSt2**].

Let V, W be symplectic vector spaces and let Γ be a Lagrangian subspace of $V \times W$. Let Λ be a Lagrangian subspace of W. Let

$$\Gamma \circ \Lambda = \{v \in V \colon \text{there exists } (v,w) \in \Gamma \text{ with } w \in \Lambda\}. \tag{7.151}$$

Let $\pi\colon \Gamma \to W$ and $\rho\colon \Gamma \to V$ be the coordinate projections. Consider the diagram

$$\begin{array}{ccccc} V & \xleftarrow{\rho} & \Gamma & \longleftarrow & F \subset \Gamma \times \Lambda \\ & & {\scriptstyle\pi}\downarrow & & \downarrow \\ & & W & \xleftarrow{\iota} & \Lambda \end{array}$$

Here, $F = \{(a = (v,w), b = w) \in \Gamma \times \Lambda, \pi(a) = w = \iota(b) \in W\}$ is the fiber product. Let α be the composite map

$$\alpha : F \to \Gamma \xrightarrow{\rho} V, \qquad \alpha(v,w,b) = \rho(v,w) = v.$$

PROPOSITION 7.7. $\Gamma \circ \Lambda$ *is a symplectic subspace of W, and there is a canonical isomorphism,*

$$|\Lambda|^{\frac{1}{2}} \otimes |\Gamma|^{\frac{1}{2}} \simeq |\ker \alpha| \otimes |\Gamma \circ \Lambda|^{\frac{1}{2}}. \tag{7.152}$$

PROOF. First we have the exact sequence

$$0 \to \ker\alpha \to F \xrightarrow{\alpha} \Gamma \circ \Lambda \to 0, \tag{7.153}$$

which implies

$$|F|^{\frac{1}{2}} \simeq |\Gamma \circ \Lambda|^{\frac{1}{2}} \otimes |\ker\alpha|^{\frac{1}{2}}. \tag{7.154}$$

Define

$$\tau\colon \Gamma \times \Lambda \to W, \qquad \tau((v,w),b) = \pi(v,w) - \iota(b) = w - b. \tag{7.155}$$

Associated to the diagram (7.5) is the exact sequence

$$0 \to F \to \Gamma \times \Lambda \xrightarrow{\tau} W \to \operatorname{coker}\tau \to 0, \tag{7.156}$$

which implies

$$|F|^{-\frac{1}{2}} \otimes |\Gamma|^{\frac{1}{2}} \otimes |\Lambda|^{\frac{1}{2}} \otimes |W|^{-\frac{1}{2}} \otimes |\operatorname{coker}\tau|^{\frac{1}{2}} \simeq 1, \tag{7.157}$$

hence

$$|F|^{\frac{1}{2}} \otimes |W|^{\frac{1}{2}} \otimes |\operatorname{coker}\tau|^{-\frac{1}{2}} \simeq |\Gamma|^{\frac{1}{2}} \otimes |\Lambda|^{\frac{1}{2}}. \tag{7.158}$$

Combining (7.154) and (7.158) gives

$$|\Gamma|^{\frac{1}{2}} \otimes |\Lambda|^{\frac{1}{2}} \simeq |\Gamma \circ \Lambda|^{\frac{1}{2}} \otimes |\ker\alpha|^{\frac{1}{2}} \otimes |W|^{\frac{1}{2}} \otimes |\operatorname{coker}\tau|^{-\frac{1}{2}}. \tag{7.159}$$

To complete the proof we need to show that

$$|\operatorname{coker}\tau|^{-\frac{1}{2}} \simeq |\ker\alpha|^{\frac{1}{2}}. \tag{7.160}$$

This follows from the fact that $\ker\alpha$ and $\operatorname{coker}\tau$ are dually paired by the symplectic form on W, so that $(\ker\alpha)^\perp = \operatorname{Im}\tau$. Indeed, $\ker\alpha = \{(a = (v,w'), w) \in F\colon \rho(a) = v = 0\}$ and $(a,w) \in F$ if and only if $w' = w$. Hence $\ker\alpha \simeq \{w \in \Lambda\colon (0,w) \in \Gamma\}$. On the other hand, if $u \in \operatorname{Im}\tau$, then $u = w_2 - w_1$ with $(v_2,w_2) \in \Gamma$ and $w_1 \in \Lambda$. Now suppose that in the identification above $w \in \ker\alpha$, and $u \in \operatorname{Im}\tau$. Then $\Omega_W(w_1,w) = 0$ since $w_1, w \in \Lambda$ and Λ is Lagrangian. Moreover, $\Omega_W(w_2,w) = 0$ since Γ is Lagrangian in $V \times W$ and so $\{w\colon (0,w) \in \Gamma\}$ is isotropic in W. Hence, $\Omega_W(w,u) = 0$. Since Γ and Λ are Lagrangian, it follows that $(\ker\alpha)^\perp = \operatorname{Im}\tau$ in W. This implies (7.160). Since $|W|^{\frac{1}{2}} \simeq 1$ (i.e., there is a canonical choice of half-density), this proves (7.152). □

The linear algebra is used to define a composition law for half-densities on canonical relations. Suppose that X, Y are compact manifolds and $\Gamma \subset T^*(X \times Y)\backslash 0$, $\Lambda \subset T^*Y\backslash 0$ are Lagrangian submanifolds. Let $\Gamma' = \{(x,\xi,y,\eta)\colon (x,\xi,y,-\eta) \in \Gamma\}$. The composition of Γ and Λ is defined by

$$\Gamma' \circ \Lambda = \{(x,\xi)\colon \text{there exists } (x,\xi,y,\eta) \in \Gamma' \text{ where } (y,\eta) \in \Lambda\}. \tag{7.161}$$

It follows from Lemma 7.7 that the composite is a Lagrangian submanifold. Moreover if $F \subset \Gamma' \times \Lambda$ is the fiber product, i.e., set of points $((x,\xi,y,\eta),(y,\eta))$, and if the fibers are compact, then by Lemma 7.7 the half-densities compose on each tangent space to give a density on the fiber with values in half-densities on the composition. That is, one has

COROLLARY 7.8. *Let $q \in \Gamma \circ \Lambda$. Let F_q be the fiber over q and let $m = (m_1, m_2, m_2) \in F_q$. Then,*

$$|T_m F_q| \otimes |T_{m_1,m_2}\Gamma \circ T_{m_2}\Lambda|^{\frac{1}{2}} \simeq |T_{(m_1,m_2)}\Gamma|^{\frac{1}{2}} \otimes |T_{m_2}\Lambda|^{\frac{1}{2}}. \tag{7.162}$$

Following [**Ho6**, Theorem 25.2.3], let us denote the half-density on Γ by σ_2 and the half-density on Λ by σ_1 and let $\sigma_1 \times \sigma_2$ denote the density on $T_f F$ with values in half-densities on on $T_{m_1,m_2}\Gamma \circ T_{m_2}\Lambda$. Then integration over F gives a half-density on the composite. At a point $q \in \Gamma \circ \Lambda$,

$$\sigma_1 \circ \sigma_2|_q = \int_{F_q} \sigma_1 \times \sigma_2. \tag{7.163}$$

Bibliography

[BaW] S. Bates and A. Weinstein, *Lectures on the geometry of quantization.* Berkeley Mathematics Lecture Notes, 8. American Mathematical Society, Providence, RI; Berkeley Center for Pure and Applied Mathematics, Berkeley, CA, 1997.

[Be] P. Bérard, On the wave equation without conjugate points, Math. Zeit. **155** (1977), 249–276.

[DeW] C. DeWitt-Morette, The semiclassical expansion. Ann. Physics 97 (1976), no. 2, 367-399.

[DSj] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit.* London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.

[Dir] P. A. M. Dirac, *The Principles of Quantum Mechanics.* 3d ed. Oxford, at the Clarendon Press, 1947.

[Du1] J.J. Duistermaat, Oscillatory integrals, Lagrange immersions and unfolding of singularities. Comm. Pure Appl. Math. 27 (1974), 207-281.

[Du2] J. J. Duistermaat, Fourier integral operators. Reprint of the 1996 original. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2011.

[DuG] J.J. Duistermaat and V. W. Guillemin, V. W. The spectrum of positive elliptic operators and periodic bicharacteristics. Invent. Math. 29 (1975), no. 1, 39-79.

[F] V. Fock, On the canonical transformation in classical and quantum mechanics, Acta Physical Acad. Scient. Hungaricae, Vol. 27 (1960), 219-224.

[Fo] G. B. Folland, *Harmonic analysis in phase space.* Annals of Mathematics Studies, 122. Princeton University Press, Princeton, NJ, 1989.

[GeSh] I. M. Gel'fand and G.E. Shilov,*Generalized functions. Vol. I: Properties and operations.* Academic Press, New York-London 1964.

[GSj] A. Grigis and J. Sjöstrand, *Microlocal analysis for differential operators. An introduction.* London Mathematical Society Lecture Note Series, 196. Cambridge University Press, Cambridge, 1994.

[GuSt1] V. Guillemin and S. Sternberg, *Geometric asymptotics.* Mathematical Surveys, No. 14. American Mathematical Society, Providence, R.I., 1977.

[GuSt2] V. Guillemin and S. Sternberg, *Semi-classical analysis.* International Press, Boston, MA, 2013.

[Ho1] L. Hörmander, Lars The spectral function of an elliptic operator. Acta Math. 121 (1968), 193-218.

[Ho2] L. Hörmander, Fourier integral operators. I. Acta Math. 127 (1971), no. 1-2, 79183.

[Ho3] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[Ho4] L. Hörmander, *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Classics in Mathematics. Springer-Verlag, Berlin, 2005.

[Ho5] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[Ho6] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[Li] R. G. Littlejohn, The van Vleck formula, Maslov theory and phase space geometry, J. Stat. Phys. 68 (1992), p. 7.

[LV] G. Loeper and C. Villani, Regularity of optimal transport in curved geometry: the nonfocal case. Duke Math. J. 151 (2010), no. 3, 431-485.

[M] W. H. Miller, Semiclassical theory of atom-diatom collisions: Path integrals and the classical S-matrix, J. Chem. Phys. 53 no. 5 (1970), p. 1949.

[P] W. Pauli, *Pauli Lectures on Physics Volume 6: Topics in Field Quantization*, MIT Press, Cambridge, Mass. (1979).

[SV] Safarov, Y. G.; Vassiliev, D., *The asymptotic distribution of eigenvalues of partial differential operators*, Translated from the Russian manuscript by the authors. Translations of Mathematical Monographs, **155** American Mathematical Society, Providence, RI, 1997.

[So] C. D. Sogge, *Fourier integrals in classical analysis.* Cambridge Tracts in Mathematics, 105. Cambridge University Press, Cambridge, 1993.

[T1] M. E. Taylor, *Partial differential equations I. Basic theory.* Second edition. Applied Mathematical Sciences, 115. Springer, New York, 2011.

[T2] M. E. Taylor, *Partial differential equations II. Qualitative studies of linear equations.* Second edition. Applied Mathematical Sciences, 116. Springer, New York, 2011.

[Tr] F. Treves, *Introduction to pseudodifferential and Fourier integral operators. Vol. 2. Fourier integral operators.* The University Series in Mathematics. Plenum Press, New York-London, 1980.

[We] A. Weinstein, Fourier integral operators and partial differential equations (Colloq. Internat., Univ. Nice, Nice, 1974), pp. 34-372, Lecture Notes in Math., Vol. 459, Springer, Berlin, 1975].)

[Wi] T. J. Willmore, *Riemannian geometry* (Oxford Science Publications, Clarendon Press, Oxford 1993).

[V] J. H. Van Vleck, The correspondence principle in the statistical interpretation of quantum mechanics, Proc. Nat. Acad. Sci. 14 (1928), p. 178.

[Y] A. Yoshikawa, On Maslov's canonical operator. Hokkaido Math. J. 4 (1975), 8-38.

[Z] S. Zelditch, Kuznecov sum formulae and Szego limit formulae on manifolds. Comm. Partial Differential Equations 17 (1992), no. 1-2, 221-260.

[Zw] M. Zworski, *Semiclassical analysis.* Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012.

CHAPTER 8

Small time wave group and Weyl asymptotics

In this section we review some estimates on eigenfunctions which use the wave kernel $E(t,x,y) = \cos t\sqrt{-\Delta}$ only for small times. In particular the proofs use the reproducing formula (5.100) of §5.9.2, which we briefly recall here: Let ρ be a smooth Schwartz function on $\mathbb{R}$ whose Fourier transform satisfies $\hat{\rho} \subset (-\varepsilon, \varepsilon)$ and $\hat{\rho}(0) = 1$, then

$$\rho(\sqrt{\lambda} - \sqrt{-\Delta})\varphi_\lambda = \varphi_\lambda, \tag{8.1}$$

where

$$\rho(\sqrt{\lambda} - \sqrt{-\Delta}) = \int_{\mathbb{R}} \hat{\rho}(t) e^{it\lambda} U(t), \tag{8.2}$$

with $U(t) = \exp -it\sqrt{-\Delta}$. The even part is almost the same and replaces $U(t)$ by $E(t) = \cos t\sqrt{-\Delta}$. The resulting kernel $\sum_{\pm} \rho(\sqrt{\lambda} \pm \sqrt{-\Delta})(x,y)$ is supported in $x \in B_\varepsilon(y)$ and is for small $|t|$ a pseudo-differential operator applied to the spherical means kernel.

We use (8.1) to obtain sup norm estimates on eigenfunctions and their derivatives. Since only small $|t|$ behavior of $E(t)$ is used, the results are local and do not use global properties of the metric or geodesic flow. Consequently the results are universal. This might seem contrary to the fact that $\sqrt{-\Delta}$ is only defined for global eigenfunctions, but of course $\cos t\sqrt{-\Delta}$ is a function of Δ and therefore is defined, at least formally, for local eigenfunctions. But it is more relevant to say that small $|t|$ behavior of wave kernels belongs to a kind of intermediate regime between purely local and global behavior of eigenfunctions.

8.1. Hörmander parametrix

To use (8.1), we need to construct small time parametrices for $U(t)$ and $E(t)$. In §6 we constructed the Hadamard parametrix, whose phase is quadratic in t. It is sometimes difficult to use the Hadamard parametrix for calculations at $t = 0$ because the distance squared $r^2(x,y)$ vanishes to order two on the diagonal, resp. t^2 vanishes to order two at $t = 0$. It is often more convenient to use a phase which vanishes only to order one.

Hörmander's parametrix is a parametrix for general unitary one parameter groups $U(t) = e^{itQ}$ generated by first order positive elliptic pseudo-differential operators Q. The parametrix has the form

$$U(t,x,y) \sim \int_{T_x^* M} e^{i\psi(x,y,\eta)} e^{it|\eta|} A(t,x,y,\eta) d\eta, \tag{8.3}$$

where $\sim$ means that the two sides differ by a smooth kernel. The phase ψ solves the Cauchy problem for the Hamilton-Jacobi equation

$$\begin{cases} q(x, d_x\psi(x,y,\eta)) = q(x,\eta), \\ \psi(x,y,\eta) = 0 \iff \langle x-y, \eta\rangle = 0, \\ d_x\psi(x,y,\eta) = \eta \quad \text{for } x = y. \end{cases} \tag{8.4}$$

Here, q is the principal symbol of Q. The proof that there exists a unique solution ψ and the construction of the amplitude $\mathcal{A}$ may be found in [**Ho4, DG, SV**].

When $Q = \sqrt{-\Delta}$ there is a more concrete construction using the phase (2.30). This phase is equivalent to ψ in the sense of [**Ho2**, Theorem 3.1.6]. We thus seek to construct a Fourier integral parametrix of the form

$$U(t,x,y) \sim \int_{T_x^* M} e^{i\langle \exp_y^{-1} x, \eta\rangle} e^{it|\eta|_y} \mathcal{A}(t,x,y,\eta)\, d\eta. \tag{8.5}$$

As in the case of pseudo-differential operators, the amplitude is understood to be cutoff near the diagonal so that the phase is well defined. Hence the right side is only equal to the left side modulo smoothing operators.

For fixed y, the level sets

$$H_{y,\eta,c} := \{(x,\eta) \colon \langle \exp_y^{-1} x, \eta\rangle = c\}$$

are 'distorted plane-wave' hypersurfaces of (M,g) near y with normal η, which generalize Euclidean 'plane waves.' These 'parallel hyperplanes normal to η' are images under $\exp_y$ of level sets of $\langle \xi, \eta\rangle = c$ in $T_y M$.

Unlike ψ, the phase $\langle \exp_y^{-1} x, \eta\rangle$ does not solve the Hamilton-Jacobi equation above for all (x,y,η) but only on the canonical relation. We now verify this fact, which is sufficient for the existence of an amplitude $\mathcal{A}$ so that (8.5) is valid.

LEMMA 8.1. *If* $\exp_y t\eta = x$, *then* $|\nabla_x \langle \exp_y^{-1} x, \eta\rangle|_x = |\eta|$.

PROOF. The radial geodesic in the direction η is of course normal to the exponential image of the hyperplanes $H_{y,\eta,c}$ for all c. If $\exp_y t\eta = x$ then $|\nabla_x \langle \exp_y^{-1} x, \eta\rangle| = \frac{\partial}{\partial t}\langle \exp_y^{-1} \exp_y t\frac{\eta}{|\eta|}, \eta\rangle = t|\eta|_y$. Hence $|\nabla_x \langle \exp_y^{-1} x, \eta\rangle|_x = |\eta|$. □

8.2. Wave group and spectral projections

In this and the next two sections, we introduce wave equation techniques for obtaining asymptotic properties of global eigenfunctions. This section is devoted to spectral projections kernels and their relations to the wave group. In the next section §8.4 the methods are combined with Tauberian theorems to prove several Weyl laws on spectral averages. We then adjust the techniques to apply to individual eigenfunctions.

Differentiating the spectral projections kernel (1.38), we denote by $d\Pi_\lambda$ the spectral measure

$$d\Pi_{[0,\lambda]} := \sum_{j=1}^{\infty} \varphi_j(x)\varphi_j(y)\delta(\lambda_j) \tag{8.6}$$

associated with the spectral projection kernels Π_λ for $\sqrt{-\Delta}$ for the interval $[0,\lambda]$. Here, $\delta(x)$ is the Dirac mass at $x \in \mathbb{R}$. Its Fourier transform is the half-wave group

$$U(t) = \int e^{-it\lambda}\, d\Pi_\lambda = e^{-it\sqrt{-\Delta}}, \tag{8.7}$$

which is the solution operator of the Cauchy problem

$$\left(\frac{1}{i}\frac{\partial}{\partial t} + \sqrt{-\Delta}\right) U(t) = 0, \qquad U(0) = \mathrm{Id}\,. \tag{8.8}$$

The Schwartz kernel $U(t,x,y)$ is a Fourier integral operator in the class $I^{-1/4}(\mathbb{R} \times M \times M, \mathcal{C})$ where $\mathcal{C}$ is the Lagrangian

$$\mathcal{C} = \{(t,x,y;\tau,\xi,\eta) \colon \tau + p(x,\xi) = 0 \text{ and } (x,\xi) = G^t(y,\eta)\}. \tag{8.9}$$

The restriction of the wave kernel to the diagonal in $M \times M$ is the pullback

$$U(t,x,x) = \Delta^* U(t,x,y) \tag{8.10}$$

under the diagonal embedding

$$\Delta \colon \mathbb{R} \times M \to \mathbb{R} \times M \times M, \qquad \Delta(t,x) = (t,x,x). \tag{8.11}$$

By the pullback rule for wave front sets,

$$\mathrm{WF}(U(t,x,x)) \subset \mathcal{C}_\Delta, \tag{8.12}$$

where

$$\mathcal{C}_\Delta = \{(t,\tau,x,\xi-\eta) \colon \tau = -|\xi|,\, G^t(x,\eta) = (x,\xi)\}. \tag{8.13}$$

This follows from the fact that the canonical relation underlying Δ^* is given by [**DG**, (1.20)]:

$$\mathrm{WF}'(\Delta) = \{(t,\tau,x,(\xi+\eta)); (t,\tau,x,\xi,x,\eta))\} \subset T^*((\mathbb{R}\times M)\times(\mathbb{R}\times M\times M)). \tag{8.14}$$

In fact, $\mathcal{C}_\Delta$ is the clean intersection of $\mathcal{C}$ and the second component

$$\mathcal{D} := \{(t,\tau,x,\xi,x,\eta)\} \subset T^*(\mathbb{R} \times M \times M) \tag{8.15}$$

of the canonical relation of Δ^*.

It follows from the composition theorem for Fourier integral operators with cleanly intersecting canonical relations (cf. [**DG, Ho4**]) that $\mathcal{C}_\Delta$ is a Lagrangian submanifold of $T^*(\mathbb{R} \times M)\backslash 0$, and that

$$U(t,x,x) \in I^0(\mathbb{R} \times M, \mathcal{C}_\Delta), \tag{8.16}$$

where $I^0(\mathbb{R}\times M, \mathcal{C}_\Delta)$ is the class of Fourier integral operators of order zero associated to the canonical relation $\mathcal{C}_\Delta$.

8.3. Small-time asymptotics for microlocal wave operators

For various results on spectral asymptotics (see §8.4), it is important to calculate the singularities for small $|t|$ of the restriction $[U(t)Q](x,y)|_{x=y}$ to the diagonal of the right composition

$$U(t)Q \colon C^\infty(M) \to C^\infty(M) \tag{8.17}$$

of $U(t)$ with a polyhomogeneous zero-order pseudodifferential operator $Q(t,x,D_x)$. A somewhat more general class which can be calculated by similar methods is $Q(t,D_t,x,D_x)$. The left composition is also similar. The calculations are done in [**Ho4, SV**].

We simplify the notation by denoting the principal symbol of $\sqrt{-\Delta}$ by $p = \sqrt{\sum g^{jk}(x)\xi_j\xi_k}$ and its subprincipal symbol by $p^s = 0$.

PROPOSITION 8.2. *For any compact Riemannian manifold (M, g), the restriction K of the kernel of $U(t)Q$ to $\mathbb{R} \times \{(x,x) \colon x \in M\}$ is conormal with respect to $\{0\} \times \{(x,x) \colon x \in M\}$ in a neighborhood of this submanifold. Moreover, there is a $\delta > 0$ so that when $|t| < \delta$*

$$K(t,x) = \int_{-\infty}^{\infty} \frac{\partial A(x,\lambda)}{\partial \lambda} e^{-i\lambda t}\, d\lambda, \tag{8.18}$$

where $A \in S^n$, $A(x,0) = 0$ and

$$A(x,\lambda) - (2\pi)^{-n} \int_{p(x,\xi)<\lambda} (b + b^s) d\xi + (2\pi)^{-n} \frac{\partial}{\partial \lambda} \int_{p(x,\xi)<\lambda} \tfrac{i}{2} \{b,p\}\, d\xi \\ + (2\pi)^{-n} \frac{\partial}{\partial \lambda} \int_{p(x,\xi)<\lambda} ir(0,x,\xi)\, d\xi \in S^{n-2}. \tag{8.19}$$

If Q is independent of t then (8.19) coincides with [**Ho4**, Proposition 29.1.2]. Since the kernel $K(t,y)$ in this case is conormal, it is easy to see that this special case also yields the time-dependent case. The constant δ is the injectivity radius of (M, g).

There is a similar formula for compositions

$$V(t) = CU(t)B \tag{8.20}$$

where $C(x,D)$ and $B(x,D)$ are zero-order polyhomogeneous pseudo-differential operators on M that follows from Proposition 8.2 by using Egorov's theorem to move $U(t)$ from the middle to the left position. Let b and c denote the principal symbols of B and C. Note then that

$$\tfrac{i}{2}\{\sigma_{\mathrm{prin}}(CB), p\} - \sigma_{\mathrm{prin}}([\sqrt{-\Delta}, C]B) = \tfrac{i}{2}\big(c\{b,p\} - b\{c,p\}\big). \tag{8.21}$$

We have the following.

PROPOSITION 8.3. *The restriction K of the kernel of $CU(t)B$ to $\mathbb{R} \times \{(x,x) \colon x \in M\}$ is conormal with respect to $\{0\} \times \{(x,x) \colon x \in M\}$ in a neighborhood of this submanifold. Moreover, there is a $\delta > 0$ so that when $|t| < \delta$*

$$K(t,x) = \int_{-\infty}^{\infty} \frac{\partial A_{CB}(x,\lambda)}{\partial \lambda} e^{-i\lambda t}\, d\lambda, \tag{8.22}$$

where $A_{CB} \in S^n$, $A_{CB}(0,\lambda) = 0$ and

$$A_{CB}(x,\lambda) - (2\pi)^{-n} \int_{p(x,\xi)<\lambda} (cb + \sigma_{sub}(CB))\, d\xi \\ + (2\pi)^{-n} \frac{\partial}{\partial \lambda} \int_{p(x,\xi)<\lambda} \tfrac{i}{2}\big(c\{b,p\} - b\{c,p\}\big)\, d\xi \in S^{n-2}. \tag{8.23}$$

To reduce Proposition 8.3 to (8.19), we observe that $V(t)$ solves the Cauchy problem

$$\begin{cases} \big(\frac{1}{i}\frac{\partial}{\partial t} + \sqrt{-\Delta}\big) V(t) = [\sqrt{-\Delta}, C] U(t) B, \\ V(0) = CB. \end{cases} \tag{8.24}$$

Consequently, by Duhamel's formula

$$CU(t)B = U(t)CB + i\int_0^t U(t)\big(U(-s)[\sqrt{-\Delta}, C]U(s)B\big) ds. \tag{8.25}$$

If we change our notation a bit and let $\sigma_{prin}(CB)$ and $\sigma_{sub}(CB)$ denote the principal and subprincipal symbols of CB then Proposition 8.2 tells us that for small $|t|$ we can write the restriction to the diagonal of the kernel of $U(t)CB$ as

$$\int \frac{\partial A_0}{\partial \lambda}(x,\lambda) e^{-i\lambda t}\, d\lambda, \tag{8.26}$$

where

$$A_0(x,\lambda) - (2\pi)^{-n} \int_{p<\lambda} \big(\sigma_{\text{prin}}(CB) + \sigma_{\text{sub}}(CB)\big) d\xi + (2\pi)^{-n} \frac{\partial}{\partial\lambda} \int_{p<\lambda} \tfrac{i}{2}\{\sigma_{\text{prin}}(CB), p\} d\xi \in S^{n-2}. \tag{8.27}$$

To perform the same calculation for the last term in (8.25), we note that by Egorov's theorem

$$Q(t,x,D_x) = i \int_0^t U(-s)[\sqrt{-\Delta}, C] U(s) B\, ds \tag{8.28}$$

is as in Proposition 8.2 with $Q(0,x,D_x) = 0$ and $\partial_t Q(0,x,D_x) = i[\sqrt{-\Delta}, C]B$. Thus, for small $|t|$ the restriction to the diagonal of the kernel of the last term in (8.25) can be written as

$$\int \frac{\partial A_1}{\partial \lambda}(x,\lambda) e^{-i\lambda t}\, d\lambda, \tag{8.29}$$

where

$$A_1(x,\lambda) - (2\pi)^{-n} \frac{\partial}{\partial\lambda} \int_{p<\lambda} \sigma_{prin}([\sqrt{-\Delta}, C]B) d\xi \in S^{n-2}. \tag{8.30}$$

Thus, we can combine the main term for A_1 with the last term for A_0 to complete the proof.

8.4. Weyl law and local Weyl law

The classical Weyl law asymptotically counts the number of eigenvalues less than λ

$$N(\lambda) = \#\{j\colon \lambda_j \le \lambda\} = \frac{|B_n|}{(2\pi)^n} \operatorname{Vol}(M,g) \lambda^n + O(\lambda^{n-1}). \tag{8.31}$$

Here, $|B_n|$ is the Euclidean volume of the unit ball and $\operatorname{Vol}(M,g)$ is the volume of M with respect to the metric g. Equivalently,

$$\operatorname{Tr} \Pi_\lambda = \frac{\operatorname{Vol}(|\xi|_g \le \lambda)}{(2\pi)^n} + O(\lambda^{n-1}), \tag{8.32}$$

where Vol is the symplectic volume measure relative to the natural symplectic form $\sum_{j=1}^n dx_j \wedge d\xi_j$ on T^*M. Thus, the dimension of the space where $H = \sqrt{\Delta}$ is $\le \lambda$ is asymptotically the volume where its symbol $|\xi|_g \le \lambda$.

8.4.1. Two term Weyl laws. An improved, two-term Weyl law has been proved which takes into account the singularities of $\operatorname{Tr}\cos t\sqrt{-\Delta}$ for larger values of t. The singular $t \neq 0$ are the lengths of the closed geodesics γ of G^t. The size of the remainder reflects the measure of closed geodesics. Before stating the result we review some notation concerning closed geodesics.

A periodic geodesic is a periodic orbit of the geodesic flow $G^t : S^*_g M \to S^*_g M$ on the unit cosphere bundle. Its projection to M is a smoothly closed geodesic $\gamma(t)$, i.e., the initial and terminal tangent vectors are the same, $\gamma'(0) = \gamma'(L)$ where L is the length of the geodesic.

We say that geodesic flow of (M, g) is *aperiodic* (or more simply that (M, g) is aperiodic) if the periodic geodesics form a set of measure zero in S^*M. In this case the remainder estimate has been improved by Duistermaat-Guillemin-Ivrii and one can obtain Weyl laws over the shorter interval $[\lambda, \lambda + 1]$. Generic g are aperiodic in this sense.

We say that the geodesic flow of (M, g) periodic if there is a time T so that $G^T = \mathrm{Id}$, i.e., all geodesics are smoothly closed and of period T. Such (M, g) are usually called Zoll manifolds. We also say that (M, g) is *partially periodic* if the set of closed geodesics of (M, g) (i.e., the set of periodic points of G^t in $S^*_g M$) has positive Liouville measure.

THEOREM 8.4 (Two-term Weyl laws).

(1) *In the aperiodic case, the two-term Weyl law of Duistermaat-Guilleimin(-Ivrii) states*

$$N(\lambda) = \#\{j : \lambda_j \leq \lambda\} = c_n \operatorname{Vol}(M, g)\, \lambda^n + o(\lambda^{n-1}) \tag{8.33}$$

where $n = \dim M$ and where c_m is a universal constant.

(2) *In the periodic case, the spectrum of $\sqrt{-\Delta}$ is a union of eigenvalue clusters C_N of the form*

$$C_N = \left\{ \left(\frac{2\pi}{T}\right) \left(N + \frac{\beta}{4}\right) + \mu_{Ni},\ i = 1 \dots d_N \right\} \tag{8.34}$$

with $\mu_{Ni} = O(N^{-1})$. The number d_N of eigenvalues in C_N is a polynomial of degree $n - 1$.

We will sketch the proof of this and related results below but will not give too many details because it is quite standard and not the main point of this monograph. We refer to [**DG, Ho4, SV**] for background and further discussion.

8.4.2. Pointwise Weyl laws. One of the principal methods for relating eigenfunctions and geodesic flow are the *pointwise Weyl laws*, i.e., Weyl asymptotics and remainder estimates for the pointwise sums

$$N(\lambda, x) = \sum_{j : \lambda_j \leq \lambda} |\varphi_{\lambda_j}(x)|^2. \tag{8.35}$$

THEOREM 8.5 (Avakumovich, Levitan, Hörmander, Duistermaat-Guillemin).

$$N(\lambda, x) :\ = \sum_{\lambda_j \leq \lambda} |\varphi_j(x)|^2 = \frac{1}{(2\pi)^n} |B^n| \lambda^n + R(\lambda, x), \tag{8.36}$$

where $R(\lambda, x) = O(\lambda^{n-1})$ uniformly in x. Here, $|B_n|$ is the volume of the unit ball in $\mathbb{R}^n$.

An important consequence is the following estimate on sup norms of L^2 normalized eigenfunctions.

COROLLARY 8.6. *Let (M, g) be a compact C^∞ Riemannian manifold of dimension n, and let φ_λ be an L^2-normalized eigenfunction of eigenvalue λ^2. Then*

$$\sup_{x \in M} |\varphi_\lambda(x)| \leq C_g \lambda^{\frac{n-1}{2}}. \tag{8.37}$$

The proof is that $\sum_{j:\lambda_j = \lambda} \varphi_{\lambda_j}(x)^2 = N(\lambda, x) - N(\lambda - 0, x)$, i.e., it is the jump of the Weyl function at λ. But the asymptotic is continuous and therefore

$$\sum_{j:\lambda_j=\lambda} \varphi_{\lambda_j}(x)^2 = R(\lambda, x) - R(\lambda - 0, x). \tag{8.38}$$

In general we know little about the jump of the remainder but by (8.36) we know that it is $O(\lambda^{n-1})$ and taking square roots gives Corollary 8.6.

In §10.4 we study the geometry of compact (M, g) for which this universal bound is achieved. We show that for C^∞ metrics there must exist a point $x_0 \in M$ such that the set $\mathcal{L}_{x_0} \subset S^*_{x_0} M$ of directions of loops at x_0 has positive measure. Furthermore the first return map on loop directions must have an invariant measure in the class of the standard surface volume form of $S^*_{x_0} M$. In particular, one obtains a $o(\lambda^{n-1})$ remainder if the set of geodesic loops at x has measure zero. Such refinements will be discussed in §10.

8.4.3. Local Weyl law for PsiDO's. The PsiDO local Weyl law concerns the traces $\operatorname{Tr} A\Pi_{[0,\lambda]}$ where $A \in \Psi^m(M)$.

THEOREM 8.7. *Let $A \in \Psi^0(M)$. Then*

$$\sum_{\lambda_j \leq \lambda} \langle A\varphi_j, \varphi_j \rangle = \frac{1}{(2\pi)^n} \left(\int_{B^*M} \sigma_A \, dx d\xi \right) \lambda^n + O(\lambda^{n-1}). \tag{8.39}$$

If (M, g) is aperiodic then the remainder is $o(\lambda^{n-1})$.

When the periodic geodesics form a set of measure zero in S^*M, one obtains an asymptote and remainder if one averages over the shorter interval $[\lambda, \lambda+1]$. The statement about the remainder is quite similar to the one for $A = I$ but was first proved (to the author's knowledge) in [**Z2**].

We introduce the notation,

$$\omega(A) := \frac{1}{\mu(S^*M)} \int_{S^*M} \sigma_A \, d\mu_L \tag{8.40}$$

for the Liouville average (or state), where μ_L is the *Liouville measure* on S^*M, i.e., the surface measure $d\mu_L = \frac{dx d\xi}{dH}$ induced by the Hamiltonian $H = |\xi|_g$ and by the symplectic volume measure $dx d\xi$ on T^*M. When σ_A is homogeneous of degree zero, one evaluate the ball average in polar coordinates in $B^*_x M$ and find that the right side of (8.39) is a universal multiple $C_n = \frac{|B_n|}{|S^{n-1}|}$ (depending on the dimension) of $\omega(A)$.

It follows that

$$\omega(A) = \lim_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} \langle A\varphi_j, \varphi_j \rangle \tag{8.41}$$

8.5. Fourier Tauberian approach

We now sketch the proof of the pointwise (hence global) Weyl asymptotics. The classical method of Carleman, Levitan, Hörmander, Duistermaat-Guillemin and others uses a Fourier Tauberian method to relate asymptotics to the singularity at $t = 0$ of the dual trace

$$S(t,x) = \sum_j e^{it\lambda_j} |\varphi_{\lambda_j}(x)|^2 \quad \text{resp.} \quad S_A(t) = \sum_j e^{it\lambda_j} \langle A\varphi_{\lambda_j}, \varphi_{\lambda_j} \rangle \tag{8.42}$$

Note that $S(t,x) = U(t,x,x)$ where $U(t,x,y)$ is the kernel of $e^{it\sqrt{-\Delta}}$ on the diagonal. It is somewhat simpler to use $\cos t\sqrt{-\Delta}$.

Finding the singularity at $t = 0$ is equivalent to convolving with a test function $\rho \in \mathcal{S}(\mathbb{R})$ with supp $\hat{\rho}$ contained in a sufficiently small neighborhood of 0 and $\hat{\rho} \equiv 1$ in a small neighborhood of 0, and to study the asymptotics as $\lambda \to \infty$ of the smoothed Weyl sums

$$\sum_j \rho(\lambda - \lambda_j) |\varphi_{\lambda_j}(x)|^2 = \rho * d_\lambda N(\lambda, x). \tag{8.43}$$

Proposition 8.8. *Let (M, g) be a C^∞ compact Riemannian manifold of dimension n. Then there exists a sequence $\omega_1, \omega_2, \ldots$ of real valued smooth densities on M such that, for every $\rho \in \mathcal{S}(\mathbb{R})$ with supp $\hat{\rho}$ contained in a sufficiently small neighborhood of 0 and $\hat{\rho} \equiv 1$ in a small neighborhood of 0*

$$\sum_j \rho(\lambda - \lambda_j) |\varphi_{\lambda_j}(x)|^2 \sim \sum_{k=0}^{\infty} \omega_k \lambda^{n-k-1} \tag{8.44}$$

as $\lambda \to \infty$ (and rapidly decaying as $\lambda \to -\infty$) with

$$\omega_0(x) = \mathrm{Vol}(S^*_x M), \quad \omega_1 = 0 = \omega_n, \quad \omega_k = 0 \text{ for odd } k. \tag{8.45}$$

One now uses a short-time parametrix

$$U(t,x,y) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{i\varphi(t,x,y,\eta)} a(t,x,y,\eta)\, d\eta \tag{8.46}$$

where a is a classical symbol of order 0 and where the phase is of the Lax-Hörmander linear form in t, $\varphi(t,x,y,\eta) = \psi(x,y,\eta) - t|\eta|$, with $\psi(x,y,\eta) = 0$ if $\langle x - y, \eta \rangle = 0$. Hence,

$$\rho * dN(\lambda, x) = \int_{\mathbb{R}} e^{i\lambda t} \hat{\rho}(t) U(t,x,x)\, dt \tag{8.47}$$

$$= (2\pi)^{-n} \int_{\mathbb{R}} \int_{\mathbb{R}^n} e^{i\lambda t} \hat{\rho}(t) e^{-it|\eta|} a(t,x,x,\eta)\, d\eta dt. \tag{8.48}$$

One now changes variables $\eta \to \lambda\eta$, puts the $d\xi$ integral into polar coordinates $\xi = r\omega, |\omega| = 1$ and carries out the $dtdr$ integral by the method of stationary phase. One obtains the expansion of Proposition 8.8 with $\alpha_0 = \frac{|B_x|}{(2\pi)^n}$. The calculation of the leading order term is based on the fact that the leading order term of the amplitude equals 1 at $t = 0$, which is forced by the initial condition $U(0) = \mathrm{Id}$.

The same kind of argument applies to $N_A(\lambda)$:

Proposition 8.9. *For $A \in \Psi^m(M)$, let $N_A(\lambda) = \sum_{\lambda_j \leq \lambda} (A\varphi_j, \varphi_j)$. Then for any $\rho \in \mathcal{S}(\mathbb{R})$ with $\hat{\rho} \in C_0^\infty(\mathbb{R})$, $\mathrm{supp}\, \hat{\rho} \cap Lsp(M,g) = \{0\}$ and with $\hat{\rho} \equiv 1$ in some* *interval around 0, we have*

$$\rho * dN_A(\lambda) \sim \sum_{k=0}^{\infty} \alpha_k \lambda^{n+m-k-1} \qquad (\lambda \to +\infty), \tag{8.49}$$

where

$$n = \dim M, \quad \alpha_0 = \int_{S^*M} \sigma_A \, d\mu, \quad \alpha_k = \int_{S^*M} \omega_k \, d\mu \tag{8.50}$$

and ω_k *is determined from the* k*-jet of the complete symbol* a *of* A*.*

Proof. The only new step is to apply A to the parametrix for U_t. By Proposition 2.5, applying a PsiDO to $ae^{i\varphi}$ produces an expression $\alpha e^{i\varphi}$ with the same phase and only a change in the amplitude. Hence,

$$AU(t,x,y) = (2\pi)^{-n} \int_{\mathbb{R}^n} \alpha(t,x,y,\eta) e^{i\varphi(t,x,y,\eta)} d\eta \tag{8.51}$$

where α is a classical symbol of order m. Now proceed as before. □

8.5.1. Local Weyl law for homogeneous Fourier integral operators. In [**Z1**], the local Weyl law for pseudo-differential operators (8.39) was generalized to Fourier integral operators F associated to a local canonical graph. We refer to §7.2 for notation and background, and state the result in the form (8.41).

The canonical relation C is a local canonical graph when both projections in the diagram are (possibly branched) covering maps.

$$\begin{array}{ccccc} & & C \subset T^*M \times T^*M & & \\ & \swarrow^{\pi_X} & & \searrow^{\pi_Y} & \\ T^*M & & \Longleftrightarrow & & T^*M \end{array}$$

If we equip C with the symplectic volume measure pulled back by π_X from T^*M, then we may consider symbols σ_F as functions on C.

Proposition 8.10. *et* $C_F \subset T^*M - 0 \times T^*M - 0$ *be a local canonical graph and* $F \in I^0(M \times M; C_F)$*. Then,*

$$\lim_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{j\colon \lambda_j \leq \lambda} \langle F\varphi_j, \varphi_j \rangle = \int_{S(C_F \cap \Delta_{S^*M \times S^*M})} \sigma_F \, d\mu_L. \tag{8.52}$$

Here, $S(C_F \cap \Delta_{S^*M \times S^*M})$ *is the set of unit vectors in the diagonal part of* C_F*.*

We again take a Fourier Tauberian approach and now study the singularities of $\operatorname{Tr} F e^{it\sqrt{-\Delta}}$ at $t = 0$. Using the composition calculus of FIOs the singularity is determined by the intersection of the canonical relation of F with the diagonal, i.e., with the fixed point set of the canonical relation. The integral is with respect to a canonical volume form on this fixed point set coming from the symbol calculus.

The limit sifts out the pseudo-differential part (i.e., the diagonal branch) of the canonical relation. If the fixed point set has measure zero, then the limit is zero. One can then amplify F by composing it with Δ^s for some s.

8.5.2. Sup norm bounds on φ_λ and $|\nabla\varphi_\lambda|$. The first estimate is a universal sup norm bound. It will be derived in Corollary 8.6 from pointwise Weyl laws. The simplest proofs are by wave equation methods.

THEOREM 8.11. *Let (M,g) be a compact C^∞ Riemannian manifold of dimension n, and let φ_λ be an L^2-normalized eigenfunction of eigenvalue λ^2. Then*

$$\sup_{x\in M}|\varphi_\lambda(x)| \leq C_g \lambda^{\frac{n-1}{4}}. \tag{8.53}$$

The pointwise Weyl law of §8.4 may be differentiated any number of times and produces sup norm estimates on derivatives as well. As an example we consider the sup norm of the gradient, based on twice differentiating the spectral projections kernel

$$d_x \otimes d_y \Pi_{[0,\lambda]}(x,y)|_{x=y} = \sum_{\lambda_j\leq\lambda} |\nabla\varphi_j(x)|_g^2. \tag{8.54}$$

One can run through the proof of the pointwise Weyl law after differentiating the wave kernel and obtain

THEOREM 8.12.

$$N_\nabla(\lambda,x) := \sum_{\lambda_j\leq\lambda}|\nabla\varphi_j(x)|^2 = \frac{1}{(2\pi)^n}\int_{B_x^*M}\xi\otimes\xi d\mu_L(\xi) + R_\nabla(\lambda,x), \tag{8.55}$$

where $R(\lambda,x) = O(\lambda^{\frac{n}{2}})$ uniformly in x.

As a corollary one obtains

COROLLARY 8.13. *Let (M,g) be a compact C^∞ Riemannian manifold of dimension n, and let φ_λ be an L^2-normalized eigenfunction of eigenvalue λ^2. Then*

$$\sup_{x\in M}|\nabla\varphi_\lambda(x)| \leq C_g\lambda^{\frac{n+1}{4}}. \tag{8.56}$$

Such bounds use that φ_λ is a global eigenfunction of (M,g), i.e. an eigenfunction of the wave group. They are apparently superior to Sobolev estimates or Bernstein bounds for local eigenfunctions. These bounds can be found in [**Z4, Z3**] in the boundaryless case and in [**Xu1**] for manifolds with boundary.

In [**SoZ**], a somewhat sharper bound is proved in terms of the L^1-norm of φ_λ:

LEMMA 8.14. *If $\lambda > 0$ then*

$$\|\nabla_g\varphi_\lambda\|_{L^\infty(M)} \lesssim \lambda^{1+\frac{n-1}{2}}\|\varphi_\lambda\|_{L^1(M)}. \tag{8.57}$$

Here, $A(\lambda)\lesssim B(\lambda)$ means that there exists a constant independent of λ so that $A(\lambda)\leq CB(\lambda)$.

8.5.3. Gradient bound in C^0-norm. In this section, we prove the gradient bound

PROPOSITION 8.15. *There exists a constant $C>0$ so that for all $_j$,*

$$\sup_M|\nabla\varphi_j| \leq C\lambda\sup_{x\in M}|\varphi_j|. \tag{8.58}$$

REMARK 8.16. This proof was suggested by C. Sogge. A different one is published in [**ShXu**].

Proof. Choose a Schwartz function ρ whose Fourier transform is supported in $[-1,1]$ and which equals one at the origin. Then $\rho(\lambda^{-1}(\lambda-\sqrt{-\Delta}))\varphi_\lambda=\varphi_\lambda$. Let $K_\lambda(x,y)$ be the Schwartz kernel of this operator. We claim that

$$\nabla K_\lambda \le C\lambda^{n+1}(1+\lambda r(x,y))^{-N}, \tag{8.59}$$

for all N. Indeed, by (8.3),

$$K_\lambda(x,y) = \int_{\mathbb{R}}\int_{T_x^*M} \hat\rho(t\lambda)e^{it\lambda}e^{i\psi(x,y,\eta)}e^{it|\eta|}A(t,x,y,\eta)\,d\eta dt \tag{8.60}$$

$$= \lambda\int_{\mathbb{R}}\int_{T_x^*M} \hat\rho(t)e^{it}e^{i\psi(x,y,\eta)}e^{i\lambda^{-1}t|\eta|}A(\lambda^{-1}t,x,y,\eta)\,d\eta dt \tag{8.61}$$

$$= \lambda^{1+n}\int_{\mathbb{R}}\int_{T_x^*M} \hat\rho(t)e^{it}e^{i\lambda\psi(x,y,\eta)}e^{it|\eta|}A(\lambda^{-1}t,x,y,\lambda\eta)\,d\eta dt. \tag{8.62}$$

Now apply ∇_x, which in one term brings down a power of λ from the phase and changes the amplitude to a sum of new amplitudes with the same bounds. The phase is only stationary in (t,η) if $\exp_x t\xi=\eta$, thus if $|\eta|<r(x,y)$. If we introduce a cutoff $\chi_{B_r(x)}$ to the ball bundle of radius $2r(x,y)$, and break up the integral using $1=\chi_B+(1-\chi_B)$, then repeated integration by parts shows that the $(1-\chi_B)$ is convergent and $O(\lambda^{-N})$ for any N. Henceforth we assume that the amplitude A is supported in $|\eta|\le 2r(x,y)$.

In terms of λ, the phase is $\lambda\psi(x,y,\eta)\sim\lambda\langle x-y,\eta\rangle$. One has $|\nabla_\eta\varphi(x,y,\eta)|\ge C|x-y|\simeq Cr(x,y)$ on the support of the amplitude. We integrate by parts with the operator

$$L=(\|1+\lambda\nabla_\eta\varphi\|^2)^{-1}\Big(I+\frac{1}{i}\nabla_\eta\varphi\cdot\nabla_\eta\Big). \tag{8.63}$$

Using the reproducing formula $Le^{i\lambda\varphi}=e^{i\lambda\varphi}$ and integrating by parts N times gives (8.64)

$$K_\lambda(x,y)=\lambda^{1+n}\int_{\mathbb{R}}\int_{T_x^*M}\hat\rho(t)e^{it}e^{i\lambda\psi(x,y,\eta)}L^{*N}\left(e^{it|\eta|}A(\lambda^{-1}t,x,y,\lambda\eta)\right)\,d\eta dt.$$

To complete the proof it suffices to show that

$$L^{*N}\left(e^{it|\eta|}A(\lambda^{-1}t,x,y,\lambda\eta)\right)\le C(1+\lambda r(x,y))^{-N}, \tag{8.65}$$

which follows by induction, using that λ only appears in the denominator of L. This proves (8.59).

Since

$$\|\nabla\varphi_\lambda\|_{L^\infty}\le C\|\varphi_\lambda\|_{L^\infty}\sup\int_M|K_\lambda(x,y)|\,dV(y), \tag{8.66}$$

we have

$$\|\nabla\varphi_\lambda\|_{L^\infty}\le C\|\varphi_\lambda\|_{L^\infty}\lambda^{n+1}\sup_x\int_M(1+\lambda r(x,y))^{-N}\,dV(y). \tag{8.67}$$

Using geodesic polar coordinates centered at x, the right side is

$$C_n\int_0^1(1+\lambda r)^{-N}r^{n-1}dr=D_{n,N}\lambda^{-n}. \tag{8.68}$$

Here, $C_n, D_{n,N}$ are independent of λ. This completes the proof. □

8.6. Tauberian Lemmas

In this section we shall collect the Tauberian lemmas that we need.

The first one is a special case of [**Ho3**, Lemma 7.5.6]. It is a slight variant of the one used by Ivrii in his proof of the Duistermaat-Guillemin theorem. It requires a monotonicity assumption that will only be fulfilled for the "diagonal" terms of our approximation to $\Pi_{[0,\lambda]}(x,x)$.

LEMMA 8.17. *Suppose that μ is a non-decreasing temperate function satisfying $\mu(0)=0$ and that ν is a function of locally bounded variation such that $\nu(0)=0$. Suppose also that $m \geq 1$ and that $\varphi \in \mathcal{S}(\mathbb{R})$ is a fixed positive function satisfying $\int \varphi(\lambda)d\lambda = 1$ and $\hat{\varphi}(t)=0$, $t \notin [-1,1]$. If $\varphi_\sigma(\lambda) = \sigma^{-1}\varphi(\lambda/\sigma)$, $0<\sigma\leq\sigma_0$, assume that for $\lambda \in \mathbb{R}$*

$$|d\nu(\lambda)| \leq \big(A_0(1+|\lambda|)^m + A_1(1+|\lambda|)^{m-1}\big)\, d\lambda, \tag{8.69}$$

and that

$$|((d\mu - d\nu) * \varphi_\sigma)(\lambda)| \leq B(1+|\lambda|)^{-2}. \tag{8.70}$$

Then

$$|\mu(\lambda)-\nu(\lambda)| \leq C_m\big(\, A_0\sigma(1+|\lambda|)^m + A_1\sigma(1+|\lambda|)^{m-1} + B\,\big), \tag{8.71}$$

where C_m is a uniform constant depending only on σ_0 and our $m \geq 1$.

The other lemma that we require allows us to handle the "off-diagonal" terms in our approximation to $\Pi_{[0,\lambda]}(x,x)$ where the above monotonicity assumption will not be valid. A proof can be found in [**So**, p. 128].

LEMMA 8.18. *Let $g(\lambda)$ be a piecewise continuous function of $\mathbb{R}$. Assume that for some $m \geq 1$ there is a constants A_0 and A_1 so that*

$$|g(\lambda+s)-g(\lambda)| \leq A_0(1+|\lambda|)^m + A_1(1+|\lambda|)^{m-1}, \quad 0<s<1. \tag{8.72}$$

Suppose further that for some fixed $\delta > 0$

$$\hat{g}(t) = 0, \quad |t| < \delta.$$

Then there is a constant $C_{m,\delta}$, depending only on m and δ, so that

$$|g(\lambda)| \leq C_{m,\delta}\big(A_0(1+|\lambda|)^m + A_1(1+|\lambda|)^{m-1}\big). \tag{8.73}$$

We also quote a simple asymptotic result using this Tauberian Lemma and which will be used in §10.20. In the following, SV for any homogeneous $V \subset T^*_x M$ denotes the set of unit vectors in V. For background on Maslov indices, excesses etc. we refer to [**DG, Ho4**] .

LEMMA 8.19. *Let $A \in I^0(M\times M, \Lambda)$ denote a zeroth order Fourier integral operator associated to a homogeneous canonical relation $\Lambda \subset T^*M \times T^*M$, and assume that $\Lambda \cap \Delta_{T^*M}$ is a clean intersection. Then*

$$\sum_{\lambda_\nu \leq \lambda} \langle A\varphi_\nu, \varphi_\nu\rangle = i^m \lambda^{\frac{e-1}{2}} \int_{S\Lambda_\Delta} \sigma_A d\mu_0 + O(\lambda^{\frac{e-1}{2}-1}), \tag{8.74}$$

where $d\mu_0$ is a canonical density on the intersection $\Lambda \cap \Delta_{T^*M}$, σ_A is the symbol of A, m is the Maslov index, and $e = \dim \Lambda \cap \Delta_{T^*M}$.

Here are some additional Fourier cosine Tauberian Lemmas from [**S1, SV**] We denote by F_+ the class of real-valued, monotone nondecreasing functions $N(\lambda)$ of polynomial growth supported on $\mathbb{R}_+$. The following Tauberian theorem uses only the singularity at $t = 0$ of $\widehat{dN}$ to obtain a one term asymptotic of $N(\lambda)$ as $\lambda \to \infty$:

THEOREM 8.20. *Let $N \in F_+$ and let $\psi \in \mathcal{S}(\mathbb{R})$ satisfy the conditions: ψ is even, $\psi(\lambda) > 0$ for all $\lambda \in \mathbb{R}$, $\hat{\psi} \in C_0^\infty$, and $\hat{\psi}(0) = 1$. Then,*

$$\psi * dN(\lambda) \leq A\lambda^\nu \implies |N(\lambda) - N * \psi(\lambda)| \leq CA\lambda^\nu,$$

where C is independent of A, λ.

To obtain a two-term asymptotic formula, one needs to take into account the other singularities of $\widehat{dN}$. We let ψ be as above, and also introduce a second test function $\gamma \in \mathcal{S}$ with $\hat{\gamma} \in C_0^\infty$ and with the supp $\hat{\gamma} \subset (0, \infty)$.

THEOREM 8.21. *Let $N_1, N_2 \in F_+$ and assume:*

(1) *$N_j * \psi(\lambda) = O(\lambda^\nu)$ for $j = 1, 2$;*
(2) *$N_2 * \psi(\lambda) = N_1 * \psi(\lambda) + o(\lambda^\nu)$;*
(3) *$\gamma * dN_2(\lambda) = \gamma * dN_1(\lambda) + o(\lambda^\nu)$.*

Then

$$N_1(\lambda - o(1)) - o(\lambda^\nu) \leq N_2(\lambda) \leq N_1(\lambda + o(1)) + o(\lambda^\nu). \tag{8.75}$$

Bibliography

[DG] J. J. Duistermaat and V. W. Guillemin, The spectrum of positive elliptic operators and periodic bicharacteristics. Invent. Math. 29 (1975), no. 1, 39–79.

[GS] T. E. Gureev and Yu. G. Safarov, G. Exact spectral asymptotics for the Laplace operator on a manifold with periodic geodesics. (Russian) Translated in Proc. Steklov Inst. Mth. 1989, no. 2, 35–53. Boundary value problems of mathematical physics, 13 (Russian). Trudy Mat. Inst. Steklov. 179 (1988), 36–53, 241.

[Ho1] L. Hörmander, The spectral function of an elliptic operator. Acta Math. 121 (1968), 193–218.

[Ho2] L. Hörmander, Fourier integral operators. I. Acta Math. 127 (1971), no. 1-2, 79183.

[Ho3] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[Ho4] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[S1] Y. Safarov, Asymptotics of a spectral function of a positive elliptic operator without a nontrapping condition. (Russian) Funktsional. Anal. i Prilozhen. 22 (1988), no. 3,53–65, 96 translation in Funct. Anal. Appl. 22 (1988), no. 3, 213–223 (1989).

[S2] Y. Safarov, Fourier Tauberian theorems and applications. J. Funct. Anal. 185 (2001), no. 1, 111128.

[SV] Y. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators. Translated from the Russian manuscript by the authors.* Translations of Mathematical Monographs, 155. American Mathematical Society, Providence, RI, 1997.

[ShXu] Y. Shi and B. Xu, Gradient estimate of an eigenfunction on a compact Riemannian manifold without boundary, Ann. Global Anal. Geom. 38 (2010), no. 1, 21-26.

[So] Sogge, C.D., Fourier integrals in classical analysis. Cambridge Tracts in Math., 105. Cambridge Univ. Press, Cambridge, 1993.

[SoZ] C. D. Sogge and S. Zelditch, Lower bounds on the Hausdorff measure of nodal sets II. Math. Res. Lett. 19 (2012), no. 6, 1361-1364.

[XBin] Bin Xu, Derivatives of the spectral function and Sobolev norms of eigenfunctions on a closed Riemannian manifold. Ann. Global Anal. Geom. 26 (2004), no. 3, 231- 252.

[Xu1] X. Xu, Gradient estimates for the eigenfunctions on compact manifolds with boundary and Hrmander multiplier theorem. Forum Math. 21 (2009), no. 3, 455-476.

[Xu2] X. Xu, Eigenfunction estimates for Neumann Laplacian and applications to multiplier problems. Proc. Amer. Math. Soc. 139 (2011), no. 10, 3583-3599.

[Z1] S. Zelditch, Kuznecov sum formulae and Szegö limit formulae on manifolds. Comm. Partial Differential Equations 17 (1992), 221-260.

[Z2] S. Zelditch, On the rate of quantum ergodicity. II. Lower bounds. Comm. Partial Differential Equations 19 (1994), no. 9-10, 1565-1579.

[Z3] S. Zelditch, Fine structure of Zoll spectra. J. Funct. Anal. 143 (1997), no. 2, 415-460.

[Z4] S. Zelditch, Real and complex zeros of Riemannian random waves. Spectral analysis in geometry and number theory, 321342, Contemp. Math., 484, Amer. Math. Soc., Providence, RI, 2009.

CHAPTER 9

Matrix elements

One of the principal techniques for obtaining information on the asymptotics of eigenfunctions is to study matrix elements

$$\rho_{jk}(A) := \langle A\varphi_j, \varphi_k\rangle, \qquad (\rho_j(A) := \langle A\varphi_j, \varphi_j\rangle \text{ when } j = k) \tag{9.1}$$

of pseudo-differential operators $A \in \Psi^0(M)$ with respect to the eigenfunctions. Since the eigenfunctions are normalized, the linear functionals ρ_{jk} are bounded on the space $\mathcal{L}(\mathcal{H})$ of bounded linear operators on the Hilbert space $\mathcal{H} = L^2(M)$ equipped with the operator norm topology. In quantum mechanics, the functional $\rho_j(A)$ is viewed as the 'expected value of the observable A in the energy state φ_j (of energy λ_j^2).

When an orthonormal basis of eigenfunctions is fixed, we refer to

$$\big(\langle A\varphi_j, \varphi_k\rangle\big)_{j,k=0}^{\infty} \tag{9.2}$$

as the matrix of A with respect to the orthonormal basis $\{\varphi_j\}$. The matrix of a pseudo-differential operator has special asymptotic properties distinguishing it from the matrix of a general bounded operator.

When we fix a quantization $a \to \mathrm{Op}(a)$ of symbols as pseudo-differential operators, the matrix elements become linear functionals of the symbol and one has a representation

$$\rho_{jk}(\mathrm{Op}(a)) = \int_{T^*M} a(x,\xi)\, dW_{jk} \tag{9.3}$$

of the linear functional as a distribution dW_{jk} on smooth symbols. The distribution is sometimes called the Wigner distribution of (φ_j, φ_k) and we follow that terminology here. When $j = k$ we denote $W_{j,k}$ by W_j. In the case of homogeneous pseudo-differential operators, we can view

$$\rho_j(\mathrm{Op}(a)) = \int_{S_g^*M} a\, dW_j \tag{9.4}$$

as a distribution on the unit co-sphere bundle (energy surface). In general, we would like to study the asymptotics of the matrix elements or Wigner distributions for as large as possible a class of symbols or operators. The Wigner distribution is (almost) a positive measure and is truly one if $\mathrm{Op}(a)$ is defined in a certain way.

If $A = \mathbf{1}_E$ is multiplication by the characteristic function of a nice open set $E \subset M$, then

$$\rho_j(\mathbf{1}_E) = \int_E |\varphi_j|^2 dV_g \tag{9.5}$$

is the "mass" or the probability that the particle represented by φ_j is located in E. $\mathrm{Op}(\mathbf{1}_E)$ is viewed as the quantization of the characteristic function of a set

$E \subset T^*M$. Then $\langle \mathrm{Op}(\mathbf{1}_E)\varphi_j, \varphi_j\rangle$ is the "probability amplitude that the (position, momentum) of the particle is in E." We may regard it as the measure of the microlocal mass of φ_j in E.

9.1. Invariance properties

We now develop the view that $\rho_j(A) = \langle A\varphi_j, \varphi_j\rangle$ is an *invariant state* on the C^*-algebra $\Psi^0(M)$.

A *state* is a linear functional on $\Psi^0(M)$ such that (i) $\rho_\psi(A^*A) \geq 0$; (ii) $\rho_\psi(I) = 1$; (iii) ρ_ψ is continuous in the norm topology. It is the quantum analogue of a probability measure (a state on $C^0(S^*M)$).

An *invariant state* is a state ρ so that

$$\rho(A) = \rho(\alpha_t(A)) := \rho(U^t A U^{-t}). \tag{9.6}$$

Recall that $U^t = e^{it\sqrt{-\Delta}}$ is the wave group so that $U^t\varphi_k = e^{it\lambda_k}\varphi_k$. Since $|e^{it\lambda_k}| = 1$, the probability measure $|\psi(t,x)|^2 d\,\mathrm{Vol}$ is constant where $\psi(t,x) = U_t\psi(x)$ is the evolving state.

If $\mathrm{Op}(a) \in \Psi^0(M)$ and U^t is as above then Egorov's theorem states that $\alpha_t(\mathrm{Op}(a)) := U^t \,\mathrm{Op}(a) U^{t*} \in \Psi^0(M)$ and the principal symbol of $\alpha_t(\mathrm{Op}(a))$ is $a \circ G^t$. Quantitatively,

$$U^t \,\mathrm{Op}(a) U^{t*} = \mathrm{Op}(a \circ G^t) + R_t, \tag{9.7}$$

where R_t is a pseudo-differential operator of order -1.

Global harmonic analysis exploits the long time behavior of the geodesic flow, e.g., its ergodicity or integrability, to prove results about the high eigenvalue limit of eigenfunctions. The joint asymptotics $t \to \infty$, $\lambda_j \to \infty$ makes the analysis difficult and the geodesic flow is only a good approximation to the quantum dynamics when

$$|t| \leq T_H(\lambda_j) := \kappa \log \lambda_j,$$

for a certain κ.

9.2. Proof of Egorov's theorem

Theorem 9.1. *Let A be a homogenous semiclassical PsiDO on M of order 0 and let $U(t) = e^{-it\sqrt{-\Delta}}$. Let*

$$A(t) = U(t)^* A U(t).$$

Then $A(t)$ is a PsiDO of order 0 with principal symbol

$$\sigma_{A(t)} = (\Phi^t)^*\sigma_A := \sigma_A \circ \Phi^t.$$

Once the theorem is proved for $A \in \Pi^0(M)$ it automatically extends to all $A \in \Psi^*(M)$.

Proof. $A(t) = U(-t)AU(t)$ is the unique solution of the operator ODE,

$$\frac{d}{dt}A(t) = i[\sqrt{-\Delta}, A(t)], \quad A(0) = A. \tag{9.8}$$

The goal is to construct a complete symbol

$$\widetilde{a}(t,x,\xi) \simeq a_0(t,x,\xi) + a_{-1}(t,x,\xi) + \cdots$$

whose quantization $\widetilde{A} = Op(\widetilde{a})$ in some quantization Op satisfies an approximate version of (9.8),

$$\frac{d}{dt}\widetilde{A}(t) = i[\sqrt{-\Delta}, \widetilde{A}(t)] + R_{-\infty} \quad \widetilde{A}(0) = A + R_{-\infty,2}, \tag{9.9}$$

where $R_{-\infty}.R_{-\infty,2} \in \Psi^{-\infty}$ Then we show that

$$A - \widetilde{A} \in \Psi^{-\infty}.$$

Let Φ^t be the Hamiltonian flow of the Hamiltonian vector field X_H of $H(x,\xi) := |\xi|_g$.If we assume that $A(t) \in \Psi^0(M)$ then its principal symbol would solve the ODE

$$\frac{d}{dt}\sigma_{A(t)} = \{H, \sigma_{A(t)}\} = X_H(\sigma_{A(t)}).$$

This suggests we define

$$\begin{cases} a_0(t,x,\xi) = \sigma_A(G^t(x,\xi)), \\ \\ A_0(t) = Op(a_0(t,x,\xi)). \end{cases}$$

Then,

$$\sigma_{i[P,A_0(t)]-A_0'(t)} =: r_{-1}(t,x,\xi) \in S^{-1}.$$

Now construct $a_{-1}(t,x,\xi)$ to solve the inhomogeneous initial value problem,

$$(\frac{\partial}{\partial t} - X_p)a_{-1}(t,x,\xi) = -r_{-1}(t,x,\xi), \;\; a_{-1}(0,x,\xi) = 0.$$

This an inhomogeneous equation, whose solution is

$$a_{-1}(t,x,\xi) = -\int_0^t r_{-1} \circ \Phi^{t-s} ds. \tag{9.10}$$

If $A_{-1}(t) = Op(a_{-1})$ then

$$\frac{d}{dt}[A_0(t) + A_{-1}(t)] + i[\sqrt{-\Delta}, A_0 + A_1] \in \Psi^{-2}$$

since the symbol of order zero vanishes and the symbol of order -1 is $r_{-1} + \dot{a}_{-1} - X_H a_{-1} = 0$. So the leading term is of order -2 and so the complete symbol, denoted r_{-2} is of order -2.

Then construct a_2 in the same way so that

$$(\frac{\partial}{\partial t} - X_H)a_{-2}(t,x,\xi) = -r_{-2}(t,x,\xi), \;\; a_{-2}(0,x,\xi) = 0.$$

Iterating gives a recursive sequence a_{-j} and an asymptotic sum,

$$\widetilde{a}(t,x,\xi) \sim \sum_{j=0}^{\infty} a_j(t,x,\xi),$$

whose quantization satisfies,

$$i[P,\tilde{A}(t)] - \tilde{A}'(t) = R_{-\infty} \in \Psi^{-\infty}, \;\; \tilde{A}(0) = A(0).$$

We now verify that $A(t) - \widetilde{A}$ is a smoothing operator. Let $B(t) = A(t) - \widetilde{A}$. Then,

$$\frac{d}{dt}B(t) - i[\sqrt{-\Delta}, B(t)] = R_{-\infty}, \;\; B(0) = 0. \tag{9.11}$$

By Duhamel's formula (Section 9.7)

$$A(t) - \widetilde{A}(t) = B(t) = \int_0^t U^*(t-s) R_{-\infty}(s) U(t-s) ds. \tag{9.12}$$

The integral has a smooth kernel by the energy estimates

$$U(t) : H^s(M) \to H^s(M), \ \ \forall s, t.$$

This completes the proof of Egorov's theorem.

□

9.3. Weak* limit problem

.

One of the best known problems in semi-classical asymptotics is the following:

PROBLEM 9.2. Determine the set $\mathcal{Q}$ of 'quantum limits', i.e., weak* limit points of the sequence $\{W_k\}$ of Wigner distributions. Equivalently, determine the set of limit states of $\{\rho_k\}$.

Let $\mathcal{M}_I$ denote the compact convex set of G^t-invariant probability measures for the geodesic flow. The following is all one can say in general:

PROPOSITION 9.3. *If M is a compact manifold, then $\mathcal{Q} \subset \mathcal{M}_I$. The limits are time-reversal invariant if the eigenfunctions are real valued.*

Any weak* limit of of $\{\rho_k\}$ is an invariant measure for G^t, i.e., $\mu(E) = \mu(G^t E)$. This is because ρ_k is an invariant state for the automorphism

$$\rho_k(U_t A U_t^*) = \rho_k(A). \tag{9.13}$$

It follows by Egorov's theorem that any limit of $\rho_k(A)$ is a limit of $\rho_k(\mathrm{Op}(\sigma_A \circ G^t))$ and hence the limit measure is G^t invariant.

There are many invariant probability measures and it is difficult to characterize those which arise as quantum limits. Some examples of invariant measures are:

1. Normalized Liouville measure $d\mu_L$.
2. A periodic orbit measure μ_γ defined by $\mu_\gamma(A) = \frac{1}{L_\gamma} \int_\gamma \sigma_A ds$ where L_γ is the length of γ. A finite sum of periodic orbit measures. In this case the eigenfunctions are sometimes said to 'scar' along γ.
3. A delta-function along an invariant Lagrangian manifold $\Lambda \subset S^*M$. The associated eigenfunctions are viewed as *localizing* along Λ.
4. A more general measure which is singular with respect to $d\mu_L$. There are many examples in the hyperbolic case (see e.g. [**Si**]).

On a flat torus for instance,

$$\mathrm{Op}(a) e^{i\langle x, \lambda \rangle} = a(x, \lambda) e^{i\langle x, \lambda \rangle}. \tag{9.14}$$

Hence

$$\langle \mathrm{Op}(a) e^{i\langle x, \lambda \rangle}, e^{i\langle x, \lambda \rangle} \rangle = \int_{T^n} a\left(x, \frac{\lambda}{|\lambda|}\right) dx. \tag{9.15}$$

This is Lebesgue measure on the invariant torus $\xi = \lambda/|\lambda|$. Every Lebesgue measure (i.e., on every invariant torus $\xi = \xi_0$) arises as a weak* limit. For rational tori, eigenvalues are multiple and one may take linear combinations of such exponentials with the same eigenvalue. Some results on the possible weak limits can be found it [**J, HeR1**].

9.4. Matrix elements of spherical harmonics

We now study matrix elements with respect to spherical harmonics, in particular with respect to the standard basis Y_k^m, i.e., $\langle \mathrm{Op}(a) Y_\ell^m, Y_\ell^m \rangle$ as $\ell \to \infty, m/\ell \to c$.

The image of $T^*S^2 - 0$ under the moment map $\mu(x,\xi) = (p_\theta(x,\xi), |\xi|)$ is a vertical triangular wedge. It is a cone, reflecting that $\mu(x, r\xi) = r\mu(x,\xi)$ is homogeneous. We can break the homogeneity by taking a base for the cone with $|\xi| = 1$, i.e. by considering points $(x, 1)$. This corresponds to looking at $p_\theta \colon S^*S^2 \to \mathbb{R}$.

Thus, we consider pairs (m_j, ℓ_j) in the joint spectrum of $D_\theta, A = \sqrt{-\Delta + 1/2} - 1/2$ whose projection to the base of the cone has a limit $(c, 1)$.

THEOREM 9.4. *Suppose that* $m_j/\ell_j \to c$. *Then*

$$\langle \mathrm{Op}(a) Y_\ell^m, Y_\ell^m \rangle \to \int_{\mu^{-1}(c,1)} a_0 dx. \tag{9.16}$$

Thus, the eigenfunctions in this ray localize on the invariant torus $p_\theta^{-1}(c)$.

We define $U(t_1, t_2) = e^{i(t_1 D_\theta + t_2 A)}$ and note that it is a unitary representation of the 2-torus T^2 on $L^2(S^2)$. Further

$$\langle \mathrm{Op}(a) Y_\ell^m, Y_\ell^m \rangle = \langle U(t_1,t_2)^* \,\mathrm{Op}(a) U(t_1,t_2) Y_\ell^m, Y_\ell^m \rangle. \tag{9.17}$$

Indeed, the eigenvalues cancel out. Average this formula over T^2. We note that

$$\langle A \rangle := \int_{T^2} U(t_1,t_2)^* \,\mathrm{Op}(a) U(t_1,t_2)\, dt_1 dt_2 \tag{9.18}$$

commutes with both D_θ and A. Indeed, the commutator with A gives $\frac{d}{dt_2}$ under the integral sign, and the integral of this derivative equals zero.

But D_θ, A have a simple joint spectrum: the dimension of the joint eigenspace equals one. Hence, any operator which commutes with them is a function of them. Thus,

$$\langle A \rangle = F(D_\theta, A). \tag{9.19}$$

The function F must be homogeneous of degree zero. Also, the right side is a ΨDO whose symbol is

$$\langle a_0 \rangle : \int_{T^2} a_0(G^{t_1,t_2}(x,\xi))\, dt_1 dt_2. \tag{9.20}$$

It follows first that

$$\langle \mathrm{Op}(a) Y_\ell^m, Y_\ell^m \rangle = \langle \langle \mathrm{Op}(a) \rangle Y_\ell^m, Y_\ell^m \rangle = F(m,k). \tag{9.21}$$

Secondly, as $(m_j, \ell_j) \to \infty$ with $m_j/\ell_j \to c$, we have

$$\langle \mathrm{Op}(a) Y_\ell^m, Y_\ell^m \rangle = F(m_j, \ell_j) \to F(c,1). \tag{9.22}$$

But also, the limit is the integral of a_0 against an invariant measure. The principal symbol of F is $\langle a_0 \rangle$, which is a function on the image of the moment map. Its value at $(c,1)$ is by definition $\int_{\mu^{-1}(c,1)} a_0 dx$, concluding the proof.

Let us take the 'symbol' of the pair A, L_3. The symbol of A is the metric norm function $|\xi|$ while that of L_3 is the so-called Clairaut integral $p_\theta(x,\xi) = \langle \xi, \frac{\partial}{\partial\theta} \rangle$.

The pair $(p_\theta, |\xi|)$ is called the moment map of the completely integrable geodesic flow of S^2. By the Schwartz inequality, $|p_\theta(x,\xi)| \leq |\frac{\partial}{\partial\theta}| \leq 1$ when $|\xi| = 1$. Hence the image of T^*S^2 under the moment map is a triangular cone in $\mathbb{R}^2$ with vertex at 0 with central axis the y-axis and with sides $y = \pm x$ in the usual $x - y$ coordinates. Compare this to the image of T^*T^2 under the moment map (ξ_1, ξ_2) which is the whole plane.

But it is of intrinsic interest to understand lower bounds on L^p norms as well as upper bounds. At least in the quantum integrable case, there are lower bounds showing that some sequences of eigenfunctions must have power law growth of L^p norms, reflecting the singularities of projections of Lagrangian submanifolds to the base. As mentioned in the introduction, it is proved in [**JZ**] that

PROPOSITION 9.5. *Every invariant measure for the geodesic flow arises as a weak* limit for a sequence of eigenfunctions on the standard* S^2.

SKETCH. It suffices to show that every finite sum of delta functions on closed geodesics arises as a quantum limit. Such measures arise by taking linear combinations of the associated Gaussian beams Y^ℓ_ℓ. □

However, one may hope to constrain the possible limits when the geodesic flow is sufficiently chaotic. To do so, one needs to find properties of Wigner measures which are special and which are preserved to some degree in the semi-classical limit. For the remainder of this section, we consider what kinds of properties of eigenfunctions are measured by matrix elements. We also consider matrix elements with respect to more general kinds of operators.

9.5. Quantum ergodicity and mixing of eigenfunctions

In this section, we assume that the geodesic flow of (M, g) is ergodic. Ergodicity of G^t means that Liouville measure $d\mu_L$ is an ergodic measure for G^t on S^*M, i.e. an extreme point of $\mathcal{M}_I$. Equivalently, any G^t-invariant set has Liouville measure zero or one. Ergodicity is a spectral property of the operator $V^t f(\zeta) = f(G^t(\zeta))$ on $L^2(S^*M, d\mu_L)$, namely that V^t has 1 as an eigenvalue of multiplicity one, i.e. only invariant L^2 functions (with respect to Liouville measure) are the constant functions.

In this case, there is a general result which originated in the work of A.I. Schnirelman and was developed into the following theorem by S. Zelditch, Y. Colin de Verdière on manifolds without boundary and by P. Gérard-E. Leichtnam and S. Zelditch-M. Zworski on manifolds with boundary.

The first result is the following Variance Theorem:

THEOREM 9.6. *Let* (M, g) *be a compact Riemannian manifold (possibly with boundary), and let* $\{\lambda_j, \varphi_j\}$ *be the spectral data of its Laplacian* $-\Delta$. *Then the geodesic flow* G^t *is ergodic on* $(S^*M, d\mu_L)$ *if and only if, for every* $A \in \Psi^0(M)$, *we have:*

(i) $\lim_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\lambda_j \leq \lambda} |(A\varphi_j, \varphi_j) - \omega(A)|^2 = 0$.

(ii) *For all* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that*

$$\limsup_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\substack{j\neq k:\ \lambda_j,\lambda_k\leq\lambda \\ |\lambda_j-\lambda_k|<\delta}} |(A\varphi_j, \varphi_k)|^2 < \varepsilon. \tag{9.23}$$

The diagonal result may be interpreted as a variance result for the local Weyl law. Since all the terms are positive, the asymptotic is equivalent to the existence of a s subsequence $\{\varphi_{j_k}\}$ of eigenfunctions whose indices j_k have counting density one for which $\langle A\varphi_{j_k}, \varphi_{j_k}\rangle \to \omega(A)$ for any $A \in \Psi^0(M)$. As above, such a sequence of eigenfunctions is called ergodic. One can sharpen the results by averaging over eigenvalues in the shorter interval $[\lambda, \lambda + 1]$ rather than in $[0, \lambda]$.

The off-diagonal statement was proved in [**Z3**] and the fact that its proof can be reversed to prove the converse direction was observed by Sunada in [**Su**]. A generalization to finite area hyperbolic surfaces is in [**Z2**].

The first statement (i) is essentially a convexity result. It remains true if one replaces the square by any convex function F on the spectrum of A:

$$\frac{1}{N(E)} \sum_{\lambda_j \leq E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) \to 0. \tag{9.24}$$

The basic QE (quantum ergodicity) theorem is the following:

THEOREM 9.7. *Let (M, g) be a compact Riemannian manifold (possibly with boundary), and let $\{\lambda_j, \varphi_j\}$ be the spectral data of its Laplacian $-\Delta$. Then, if the geodesic flow G^t is ergodic on $(S^*M, d\mu_L)$, there exists a subsequence $\mathcal{S}$ of density one, $D^*(\mathcal{S}) = 1$ such that*

$$\lim_{j\to\infty, j\in\mathcal{S}} \langle A\varphi_j, \varphi_j\rangle \to \omega(A)$$

for all $A \in \Psi^0(M)$.

Density one means that $\frac{1}{N(\lambda)}\#\{j : \lambda_j \leq \lambda, j \in \mathcal{S}\} \to 1$ as $\lambda \to \infty$.

9.5.1. Quantum ergodicity in terms of operator time and space averages. The diagonal variance asymptotics may be interpreted as a relation between operator time and space averages.

Let $A \in \Psi^0$ be an observable and define its time average to be:

$$\langle A\rangle := \lim_{T\to\infty} \langle A\rangle_T, \tag{9.25}$$

where

$$\langle A\rangle_T := \frac{1}{2T}\int_{-T}^{T} U^t A U^{-t} dt \tag{9.26}$$

Further define its space average to be scalar operator

$$\omega(A) \cdot I. \tag{9.27}$$

Then Theorem 9.7 (i) is (almost) equivalent to

$$\langle A\rangle = \omega(A)I + K \text{ where } \lim_{\lambda\to\infty} \omega_\lambda(K^*K) := \lim_{\lambda\to\infty} \mathrm{Tr}\left(\Pi_{[0,\lambda]} K^*K\right) = 0. \tag{9.28}$$

Thus, the time average equals the space average plus a term K which is semi-classically small in the sense that its Hilbert-Schmidt norm square $\|E_\lambda K\|^2_{\mathrm{HS}}$ in the span of the eigenfunctions of eigenvalue $\leq \lambda$ is $o(N(\lambda))$.

This is not exactly equivalent to Theorem 9.7 (i) since it is independent of the choice of orthonormal basis, while the previous result depends on the choice of basis. However, when all eigenvalues have multiplicity one, then the two are equivalent. To see the equivalence, note that $\langle A\rangle$ commutes with $\sqrt{-\Delta}$ and hence is diagonal in the basis $\{\varphi_j\}$ of joint eigenfunctions of $\langle A \rangle$ and of U_t. Hence K is the diagonal matrix with entries $\langle A\varphi_k, \varphi_k\rangle - \omega(A)$. The condition is therefore equivalent to

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \le E} |\langle A\varphi_k, \varphi_k\rangle - \omega(A)|^2 = 0. \tag{9.29}$$

9.5.2. Heuristic proof of Theorem 9.7 (i). There is a simple picture of eigenfunction states which makes Theorem 9.7 seem obvious. Justifying the picture is more difficult than the formal proof below but the reader may find it illuminating and convincing.

We introduce some language from quantum statistical mechanics and C^* algebras. By a *state* on $\Psi^0(M)$ (or more precisely, its norm completion) is meant a bounded iinear functional $\rho : \Psi^0(M) \to R$ satisfying:

- $\rho(I) = 1$;
- $\rho(A^*A) \ge 0$;

We further define the automorphisms $\alpha_t : \Psi^0 \to \Psi^0$ by $\alpha_t(A) = U^{-t}AU^t$. A state ρ is called an invariant state if $\rho(\alpha_t(A)) = \rho(A)$. The set of invariant states is denoted $\mathcal{E}_{\mathbb{R}}$. It is a compact convex state. The Liouville state is denoted by $\omega(A) = \fint \sigma_A d\mu_L$.

One may re-formulate the ergodicity of G^t as a property of the Liouville measure $d\mu_L$: ergodicity is equivalent to the statement $d\mu_L$ is an extreme point of the compact convex set $\mathcal{M}_I$ of G^t-invariant probability measures. Moreover, it implies that the Liouville state ω on $\Psi^0(M)$ is an extreme point of the compact convex set $\mathcal{E}_{\mathbb{R}}$ of invariant states for $\alpha_t(A) = U^{-t}AU^t$; see [**Ru**] for background. But the local Weyl law says that ω is also the limit of the convex combination $\frac{1}{N(E)}\sum_{\lambda_j \le E} \rho_j$. An extreme point cannot be written as a convex combination of other states unless all the states in the combination are equal to it. In our case, ω is only a limit of convex combinations so it need not (and does not) equal each term. However, almost all terms in the sequence must tend to ω, and that is equivalent to (i).

One could make this argument rigorous by considering whether Liouville measure is an *exposed point* of $\mathcal{E}_I$ and $\mathcal{M}_I$. Namely, is there a linear functional Λ which is equal to zero at ω and is < 0 everywhere else on $\mathcal{E}_I$? If so, the fact that $\frac{1}{N(E)}\sum_{\lambda_j \le E} \Lambda(\rho_j) \to 0$ implies that $\Lambda(\rho_j) \to 0$ for a subsequence of density one. For one gets an obvious contradiction if $\Lambda(\rho_{j_k}) \le -\varepsilon < 0$ for some $\varepsilon > 0$ and a subsequence of positive density. But then $\rho_{j_k} \to \omega$ since ω is the unique state with $\Lambda(\rho) = 0$.

In [**Je**] it is proved that Liouville measure (or any ergodic measure) is exposed in $\mathcal{M}_I$. It is stated in the following form: For any ergodic invariant probability measure μ, there exists a continuous function f on S^*M so that μ is the unique f-maximizing measure in the sense that

$$\int f d\mu = \sup\left\{ \int f\, dm \colon m \in \mathcal{M}_I \right\}. \tag{9.30}$$

To complete the proof, one would need to show that the extreme point ω is exposed in $\mathcal{E}_I$ for the C^*-algebra defined by the norm-closure of $\Psi^0(M)$.

9.5.3. Proof of Theorem 9.7. We now sketch the proof of Theorem 9.7 . By time averaging, we have

$$\sum_{\lambda_j \leq E} |\langle A\varphi_k, \varphi_k\rangle - \omega(A)|^2 = \sum_{\lambda_j \leq E} |\langle\langle A\rangle_T - \omega(A)\varphi_k, \varphi_k\rangle|^2 . \tag{9.31}$$

We then apply the Schwartz inequality to get:

$$\begin{aligned} \textstyle\sum_{\lambda_j \leq E} |\langle(\langle A\rangle_T - \omega(A))\,\varphi_k, \varphi_k\rangle|^2 &\leq \textstyle\sum_{\lambda_j \leq E} \langle(\langle A\rangle_T - \omega(A))^2\,\varphi_k, \varphi_k\rangle \\ &= \operatorname{Tr}(\Pi_E[\langle A\rangle_T - \omega(A)]^2 \Pi_E) \\ &=: \omega_E([\langle A\rangle_T - \omega(A)]^2). \end{aligned} \tag{9.32}$$

Above, Π_E is the spectral projection for $\hat{H}$ corresponding to the interval $[0, E]$. By the local Weyl law, $\omega_E \to \omega$. Hence,

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \leq E} |\langle A\varphi_k, \varphi_k\rangle - \omega(A)|^2 \leq \int_{\{H=1\}} |\langle\sigma_A\rangle_T - \omega(A)|^2 d\mu_L.$$

As $T \to \infty$ the right side approaches $\varphi(0)$ by the L^2 von Neumann mean ergodic theorem. Since the left hand side is independent of T, this implies that

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \leq E} |\langle A\varphi_k, \varphi_k\rangle - \omega(A)|^2 = 0.$$

□

It is useful to note that the same result is true if $F(x) = x^2$ is replaced by any convex function. The generalization to convex functions is useful in obtaining rates of quantum ergodicity [**Z4, Schu1, Schu2, AR12**]. The rates were used to improve various results on nodal sets and L^p norms on balls of shrinking radius $r(\lambda) = \frac{1}{(\log\lambda)^\gamma}$ for certain $\gamma > 0$ in [**Han15, HeR2**].

By time averaging, we have

$$\sum_{\lambda_j \leq E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) = \sum_{\lambda_j \leq E} F(\langle\langle A\rangle_T - \omega(A)\varphi_k, \varphi_k\rangle). \tag{9.33}$$

We then apply the Peierls–Bogoliubov inequality

$$\sum_{j=1}^n F((B\varphi_j, \varphi_j)) \leq \operatorname{Tr} F(B)$$

with $B = \Pi_E[\langle A\rangle_T - \omega(A)]\Pi_E$ to get:

$$\sum_{\lambda_j \leq E} F(\langle\langle A\rangle_T - \omega(A)\varphi_k, \varphi_k\rangle) \leq \operatorname{Tr} F(\Pi_E[\langle A\rangle_T - \omega(A)]\Pi_E). \tag{9.34}$$

Above, Π_E is the spectral projection for $\hat{H}$ corresponding to the interval $[0, E]$. By the Berezin inequality (if $F(0) = 0$):

$$\begin{aligned} \frac{1}{N(E)} \operatorname{Tr} F(\Pi_E[\langle A\rangle_T - \omega(A)]\Pi_E) &\leq \frac{1}{N(E)} \operatorname{Tr} \Pi_E F([\langle A\rangle_T - \omega(A)])\Pi_E \\ &= \omega_E(\varphi(\langle A\rangle_T - \omega(A))). \end{aligned}$$

As long as F is smooth, $F(\langle A\rangle_T - \omega(A))$ is a pseudodifferential operator of order zero with principal symbol $F(\langle \sigma_A\rangle_T - \omega(A))$. By the assumption that $\omega_E \to \omega$ we get

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \leq E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) \leq \int_{\{H=1\}} F(\langle \sigma_A\rangle_T - \omega(A)) d\mu_L.$$

As $T \to \infty$ the right side approaches $\varphi(0)$ by the dominated convergence theorem and by Birkhoff's ergodic theorem. Since the left hand side is independent of T, this implies that

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \leq E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) = 0$$

for any smooth convex F on Spec(A) with $F(0) = 0$. □

9.5.4. Proof of Theorem 9.7. The proof is an application of Chebychev's inequality and a diagonal argument to the variance result of Theorem 9.6. The variances can be taken over unit width intervals $[\lambda, \lambda + 1]$ using a slightly stronger local Weyl law. We omit the proof.

First we show that for any $A \in \Psi^0(M)$ there is a density one subsequence $\mathcal{S}_A$ such that $\langle A\varphi_j, \varphi_j\rangle \to \omega(A)$. For fixed λ, we put a probability measure $\mathbf{P}_\lambda$ on $\{j : \lambda_j \leq\in [\lambda, \lambda + 1]\}$ by $\frac{1}{N(\lambda)} \sum_{j:\lambda_j\in[\lambda,\lambda+1]} \delta_{\lambda_j}$. Denote the expected value by $\mathbf{E}_\lambda$ and define the random variable X_λ on $\{j : \lambda_j \in [\lambda, \lambda + 1]\}$ by $X_\lambda(j) = |\langle A\varphi_j, \varphi_j\rangle - \omega(A)|^2$. Then,

$$\mathbf{E}_\lambda X_\lambda = \sum_{j:\lambda_j\in[\lambda\lambda+1]} |\langle A\varphi_j, \varphi_j\rangle - \omega(A)|^2 =: \varepsilon^2(\lambda).$$

By the variance Theorem 9.6 , $\varepsilon(\lambda) \to 0$ as $\lambda \to \infty$. Let

$$\Gamma(\lambda) = \{j : \lambda_j \in [\lambda, \lambda+1], |\langle A\varphi_j, \varphi_j\rangle - \omega(A)|^2 \geq \varepsilon(\lambda)\}, \;\; \Lambda(\lambda) := \{j : \lambda_j \leq \lambda\}\backslash\Gamma(\lambda).$$

By Chebyshev's inequality,

$$\mathbf{P}_\lambda(\Gamma_\lambda) \leq \frac{\varepsilon^2(\lambda)}{\varepsilon(\lambda)} = \varepsilon(\lambda).$$

Then,

$$\mathbf{P}_\lambda(\Lambda(\lambda)) = 1 - \frac{\#\Gamma(\lambda)}{N(\lambda)} \geq 1 - \varepsilon(\lambda).$$

We then dissect $\mathbb{R}_+$ into intervals $[N, N + 1]$ and let $\Lambda_N = \Lambda(\lambda)$ with $\lambda = N$. We then define Λ_∞ by $\Lambda_\infty \cap [N, N + 1] = \Lambda_N \cap [N, N + 1]$.

The sets $\Gamma(\lambda), \Lambda(\lambda)$ depend on A. We now use a diagonalization argument to obtain a subsequence of density one which works for all A.

Since $\Psi^0(M)$ is separable, there exists a countable dense subset $\{A_j\}$ of the unit ball of Ψ^0. For each j, let $\mathcal{S}_j \subset \mathbb{N}$ be a density one subsequence $\mathcal{S}_A$ such that $\langle A\varphi_k, \varphi_k\rangle \to \omega(A)$ (with $k \in \mathcal{S}_A$) for A_j. We may assume $\mathcal{S}_j \subset \mathcal{S}_{j+1}$. Then choose N_j so that

$$\frac{1}{N}\#\{k \in \mathcal{S}_j : k \leq N\} \geq 1 - 2^{-j} \text{ for } N \geq N_j\ .$$

Let $\mathcal{S}_\infty$ be the subsequence defined by

$$\mathcal{S}_\infty \cap [N_j, N_{j+1}] = \mathcal{S}_j \cap [N_j, N_{j+1}].$$

Then $\mathcal{S}_\infty$ is of density one and

$$\lim_{k\to\infty, k\in\mathcal{S}_\infty} \rho_k(A) = \omega(A)$$

for all $A \in \Psi^0$, since it holds for the set $\{A_j\}$ and since $\{A_j\}$ is dense in the unit ball.

9.5.5. Sketch of Proof of Theorem 9.7 (i) for general convex functions. We now sketch the proof of (9.24). Let Π_E be the spectral projection corresponding to the interval $[0, E]$. By time averaging, we have

$$\sum_{\lambda_j \le E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) = \sum_{\lambda_j \le E} F(\langle \langle A\rangle_T - \omega(A)\varphi_k, \varphi_k\rangle). \tag{9.35}$$

We then apply the Peierls–Bogoliubov inequality

$$\sum_{j=1}^{n} F((B\varphi_j, \varphi_j)) \le \operatorname{Tr} F(B) \tag{9.36}$$

with $B = \Pi_E[\langle A\rangle_T - \omega(A)]\Pi_E$ to get

$$\sum_{\lambda_j \le E} F(\langle \langle A\rangle_T - \omega(A)\varphi_k, \varphi_k\rangle) \le \operatorname{Tr} F(\Pi_E[\langle A\rangle_T - \omega(A)]\Pi_E). \tag{9.37}$$

By the Berezin inequality (if $F(0) = 0$),

$$\frac{1}{N(E)} \operatorname{Tr} F(\Pi_E[\langle A\rangle_T - \omega(A)]\Pi_E) \le \frac{1}{N(E)} \operatorname{Tr} \Pi_E F([\langle A\rangle_T - \omega(A)])\Pi_E \tag{9.38}$$

$$= \omega_E(\varphi(\langle A\rangle_T - \omega(A))). \tag{9.39}$$

As long as F is smooth, $F(\langle A\rangle_T - \omega(A))$ is a pseudodifferential operator of order zero with principal symbol $F(\langle \sigma_A\rangle_T - \omega(A))$. By the assumption that $\omega_E \to \omega$ we get

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \le E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) \le \int_{\{H=1\}} F(\langle \sigma_A\rangle_T - \omega(A)) d\mu_L. \tag{9.40}$$

As $T \to \infty$ the right side approaches $\varphi(0)$ by the dominated convergence theorem and by Birkhoff's ergodic theorem. Since the left hand side is independent of T, this implies that

$$\lim_{E\to\infty} \frac{1}{N(E)} \sum_{\lambda_j \le E} F(\langle A\varphi_k, \varphi_k\rangle - \omega(A)) = 0 \tag{9.41}$$

for any smooth convex F on $\operatorname{spec}(A)$ with $F(0) = 0$, completing the proof.

This proof can only be used directly for scalar Laplacians on manifolds without boundary, but it still works as a template in more involved situations. For instance, on manifolds with boundary, conjugation by the wave group is not a true automorphism of the observable algebra. In quantum ergodic restriction theorems, the appropriate conjugation is an endomorphism but not an automorphism. When Δ has continuous spectrum (as in finite area hyperbolic surfaces), one must adapt the proof to states which are not L^2-normalized [**Z2**].

9.5.6. QUE in terms of time and space averages. The quantum unique ergodicity problem (the term is due to Rudnick-Sarnak [**RS**]) is the following:

PROBLEM 9.8. Suppose the geodesic flow G^t of (M, g) is ergodic on S^*M. Is the operator K in the time-average

$$\langle A \rangle := \lim_{T \to \infty} \langle A \rangle_T = \omega(A) + K \tag{9.42}$$

a compact operator? Equivalently is $\mathcal{Q} = \{d\mu_L\}$?

If K is compact, or equivalently $\mathcal{Q} = \{d\mu_L\}$, then $\sqrt{-\Delta}$ is said to be quantum uniquely ergodic (QUE). Compactness of K implies that $\langle K\varphi_k, \varphi_k \rangle \to 0$, hence $\langle A\varphi_k, \varphi_k \rangle \to \omega(A)$ along the entire sequence. Rudnick-Sarnak [**RS**] conjectured that Laplacians of negatively curved manifolds are QUE, i.e., that for any orthonormal basis of eigenfunctions, the Liouville measure is the only quantum limit.

9.5.7. Converse QE. So far we have not mentioned Theorem 9.7 (ii). An interesting open problem is the extent to which (ii) is actually necessary for the equivalence to classical ergodicity.

PROBLEM 9.9. Suppose that $\sqrt{-\Delta}$ is quantum ergodic in the sense that Theorem 9.7 (i) holds. What are the properties of the geodesic flow G^t. Is it ergodic (in the generic case)?

In the larger class of Schrödinger operators, there is a simple example of a Hamiltonian system which is quantum ergodic but not classically ergodic, namely, a Schrödinger operator with a symmetric double well potential W. That is, W is a W shaped potential with two wells and a $\mathbb{Z}_2$ symmetry exchanging the wells. The low energy levels consist of two connected components interchanged by the symmetry, and hence the classical Hamiltonian flow is not ergodic. However, all eigenfunctions of the Schrödinger operator $-\frac{d^2}{dx^2} + W$ are either even or odd and thus have the same mass in both wells. It is easy to see that the quantum Hamiltonian is quantum ergodic.

Recently, B. Gutkin [**G**] has given a two dimensional example of a domain with boundary which is quantum ergodic but not classically ergodic and which is a two dimensional analogue of a double well potential. The domain is a so-called hippodrome (race-track) stadium. Similarly to the double well potential, there are two invariant sets interchanged by a $\mathbb{Z}_2$ symmetry. They correspond to the two orientations with which the race could occur. Hence the classical billiard flow on the domain is not ergodic. After dividing by the $\mathbb{Z}^2$ symmetry the hippodrome has ergodic billiards, hence by Theorem 9.7, the quotient domain is quantum ergodic. But the The eigenfunctions are again either even or odd. Hence the hippodrome is quantum ergodic but not classically ergodic.

Little is known about converse quantum ergodicity in the absence of symmetry. It is known that if there exists an open set in S^*M filled by periodic orbits, then the Laplacian cannot be quantum ergodic (see [**MOZ**] for recent results and references). But it is not even known at this time whether KAM systems, which have Cantor-like invariant sets of positive measure, are not quantum ergodic. It is known that there exist a positive proportion of approximate eigenfunctions (quasimodes) which localize on the invariant tori, but it has not been proved that a positive proportion of actual eigenfunctions have this localization property.

9.5.8. Quantum weak mixing. There are parallel results on quantizations of weak-mixing geodesic flows. We recall that the geodesic flow of (M, g) is weak mixing if the operator V^t has purely continuous spectrum on the orthogonal complement of the constant functions in $L^2(S^*M, d\mu_L)$. The following is proved in [**Z2**]:

THEOREM 9.10. *The geodesic flow G^t of (M, g) is weak mixing if and only if the conditions (i) and (ii) of Theorem 9.7 hold and additionally, for any $A \in \Psi^0(M)$ and any $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$\limsup_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{\substack{j\neq k: \lambda_j, \lambda_k \leq \lambda \\ |\lambda_j - \lambda_k - \tau| < \delta}} |(A\varphi_j, \varphi_k)|^2 < \varepsilon \qquad \text{for all } \tau \in \mathbb{R}. \tag{9.43}$$

The restriction $j \neq k$ is of course redundant unless $\tau = 0$, in which case the statement coincides with quantum ergodicity. This result follows from the general asymptotic formula, valid for any compact Riemannian manifold (M, g), that

$$\begin{aligned} &\frac{1}{N(\lambda)} \sum_{\substack{i\neq j \\ \lambda_i, \lambda_j \leq \lambda}} |\langle A\varphi_i, \varphi_j\rangle|^2 \left| \frac{\sin T(\lambda_i - \lambda_j - \tau)}{T(\lambda_i - \lambda_j - \tau)} \right|^2 \\ &\qquad \sim \left\| \frac{1}{2T} \int_{-T}^{T} e^{it\tau} V_t(\sigma_A) \right\|_2^2 - \left| \frac{\sin T\tau}{T\tau} \right|^2 \omega(A)^2. \end{aligned} \tag{9.44}$$

In the case of weak-mixing geodesic flows, the right hand side approaches 0 as $T \to \infty$. As with diagonal sums, the sharper result is true where one averages over the short intervals $[\lambda, \lambda + 1]$.

Theorem 9.10 is based on expressing the spectral measures of the geodesic flow in terms of matrix elements. The main limit formula is

$$\int_{\tau-\varepsilon}^{\tau+\varepsilon} d\mu_{\sigma_A} := \lim_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{i,j: \lambda_j \leq \lambda,\, |\lambda_i - \lambda_j - \tau| < \varepsilon} |\langle A\varphi_i, \varphi_j\rangle|^2, \tag{9.45}$$

where $d\mu_{\sigma_A}$ is the spectral measure for the geodesic flow corresponding to the principal symbol of A, $\sigma_A \in C^\infty(S^*M, d\mu_L)$. Recall that the spectral measure of V^t corresponding to $f \in L^2$ is the measure $d\mu_f$ defined by

$$\langle V^t f, f\rangle_{L^2(S^*M)} = \int_{\mathbb{R}} e^{it\tau}\, d\mu_f(\tau). \tag{9.46}$$

9.5.9. Evolution of Lagrangian states. In this section, we briefly review results on evolution of Lagrangian states and coherent states. We follow in particular the article of R. Schubert [**Sc3**].

A simple Lagrangian or WKB state has the form $\psi_\hbar(x) = a(\hbar, x) e^{\frac{i}{\hbar} S(x)}$ where $a(\hbar, x)$ is a semi-classical symbol $a \sim \sum_{j=0}^\infty \hbar^j a_j(x)$. The phase S generates the Lagrangian submanifold $(x, dS(x)) \subset T^*M$.

It is proved in [**Sc3**, Theorem 1] that if G^t is Anosov and if Λ is transversal to the stable foliation W^s (except on a set of codimension one), then there exists $C, \tau > 0$ so that for every smooth density on Λ and every smooth function $a \in C^\infty(S^*M)$, the Lagrangian state ψ with symbol σ_ψ satisfies,

$$\left| \langle U^t\psi, AU^t\psi\rangle - \int_{S^*M} \sigma_A d\mu_L \int_\Lambda |\sigma_\psi|^2 \right| \leq Che^{\Gamma|t|} + ce^{-t\tau}. \tag{9.47}$$

In order that the right side tends to zero in the joint limit $\hbar \to 0, t \to \infty$, it is necessary and sufficient that

$$t \leq \frac{1-\varepsilon}{\Gamma} |\log \hbar|. \tag{9.48}$$

9.5.10. Open problems on quantum ergodicity. As a test of how much is known about quantum ergodicity, we offer the following open problems. In each, the motivating case is where the geodesic flow is ergodic or that the eigenfunctions are ergodic.

(1) Suppose that $\varphi_{j_k}^2 \rightharpoonup 1$ in the weak* sense on $C(M)$. Prove that φ_{j_k} does not tend to 1 in the strong sense. For instance show that $\liminf_{k\to\infty} \|\varphi_{j_k}^2 - 1\|_{L^1} > 0$.
(2) Similarly show that $\liminf_{k\to\infty} \|\lambda_{j_k}^{-2} |\nabla \varphi_{j_k}|^2 - 1\|_{L^1} > 0$.
(3) Are the measures φ_j^2 uniformly integrable? Show that $\|\varphi_{j_k}\|_{L^p} \to \infty$ at least along a density one subsequence for each $p > 2$.
(4) Does there exist $\delta > 0$ so that $\|\varphi_{j_k}\|_{L^1} \geq \delta$? This is false for Gaussian beams but might occur in the ergodic case. Find a good lower bound.
(5) Is there a subsequence of density one so that $\int_E \varphi_{j_k}^2 dV \to \frac{|E|}{|M|}$ for every Borel set? Here $|E|$ is its Liouville measure. This is true by the Portmanteau theorem if ∂E has Liouville measure zero. For each Borel set there does exist a subsequence of density one with this property, but L^∞ is not separable and one cannot use a diagonalization argument to find a density one subsequence.
(6) In weak* limit formulae one holds the set fixed. But for quantum ergodic eigenfunctions one might expect the 'mean values' to dominate. Does exist a constant C so that

$$\int_{|\varphi_{j_k}| \leq C} f \varphi_{j_k}^2 \, dV \to \int_M f \, dV? \tag{9.49}$$

Note that if $c(\lambda) \to \infty$ as $\lambda \to \infty$, then

$$\int_{|\varphi_{j_k}| \geq c(\lambda_{j_k})} f \varphi_{j_k}^2 \, dV \to 0. \tag{9.50}$$

(7) Suppose that the eigenfunctions of Δ_g are quantum ergodic in the sense that diagonal matrix elements tend to their means. What can be concluded about the dynamics of the geodesic flow?

9.6. Hassell's scarring result for stadia

This section is an exposition of Hassell's scarring result for the Bunimovich stadium. We follow [**Ha**] and [**Z1**].

A stadium is a domain $X = R \cup W \subset \mathbb{R}^2$ which is formed by a rectangle $R = [-\alpha, \alpha]_x \times [-\beta, \beta]_y$ and where $W = W_{-\beta} \cup W_\beta$ are half-discs of radius β attached at either end. We fix the height $\beta = \pi/2$ and let $\alpha = t\beta$ with $t \in [1, 2]$. The resulting stadium of rectangular width $t\beta$ is denoted X_t.

It has long been suspected that there exist exceptional sequences of eigenfunctions of X which have a singular concentration on the set of "bouncing ball" orbits of R. These are the vertical orbits in the central rectangle that repeatedly bounce orthogonally against the flat part of the boundary. The unit tangent vectors to the orbits define an invariant Lagrangian submanifold with boundary $\Lambda \subset S^*X$.

It is easy to construct approximate eigenfunctions which concentrate microlocally on this Lagrangian submanifold. Namely, let $\chi(x)$ be a smooth cutoff supported in the central rectangle and form $v_n = \chi(x) \sin ny$. Then for any pseudo-differential operator A properly supported in X,

$$\langle Av_n, v_n \rangle \to \int_\Lambda \sigma_A \chi \, d\nu \tag{9.51}$$

where $d\nu$ is the unique normalized invariant measure on Λ.

Numerical studies suggested that there also existed genuine eigenfunctions with the same limit. Recently, A. Hassell has proved this to be correct for almost all stadia.

THEOREM 9.11. *Let X_t be the stadium of rectangular width $t\beta$. Then the Laplacian on X_t is not QUE for almost every $t \in [1, 2]$.*

We now sketch the proof and develop related ideas on quantum unique ergodicity. The main idea is that the existence of the scarring bouncing ball quasimodes implies that either

- There exist actual modes with a similar scarring property; or
- The spectrum has exceptional clustering around the bouncing ball quasi-eigenvalues n^2.

Hassell then proves that the second alternative cannot occur for most stadia. We now explain the ideas in more detail.

We first recall that a quasimode $\{\psi_k\}$ is a sequence of L^2-normalized functions which solve

$$\|(\Delta - \mu_k^2)\psi_k\|_{L^2} = O(1), \tag{9.52}$$

for a sequence of quasi-eigenvalues μ_k^2. By the spectral theorem it follows that there must exist true eigenvalues in the interval $[\mu_k^2 - K, \mu_k^2 + K]$ for some $K > 0$. Moreover, if $\tilde{\Pi}_{[\mu_k^2 - K, \mu_k^2 + K]}$ denotes the spectral projection for $-\Delta$ corresponding to this interval, then

$$\|\tilde{\Pi}_{[\mu_k^2 - K, \mu_k^2 + K]}\psi_k - \psi_k\|_{L^2} = O(K^{-1}). \tag{9.53}$$

To maintain consistency with our use of frequencies μ_k rather than energies μ_k^2, we rephrase this in terms of the projection $\Pi_{[\mu_k^2 - K, \mu_k^2 + K]}$ for $\sqrt{-\Delta}$ in the interval $[\sqrt{\mu_k^2 - K}, \sqrt{\mu_k^2 + K}]$. For fixed K, this latter interval has width $\frac{K}{\mu_k}$.

Given a quasimode $\{\psi_k\}$, the question arises of how many true eigenfunctions it takes to build the quasimode up to a small error.

DEFINITION 9.12. We say that a quasimode $\{\psi_k\}$ of order 0 with $\|\psi_k\|_{L^2} = 1$ has $n(k)$ essential frequencies if

$$\psi_k = \sum_{j=1}^{n(k)} c_{kj} \varphi_j + \eta_k, \qquad \|\eta_k\|_{L^2} = o(1). \tag{9.54}$$

To be a quasimode of order zero, the frequencies λ_j of the φ_j must come from an interval $[\mu_k - \frac{K}{\mu_k}, \mu_k + \frac{K}{\mu_k}]$. Hence the number of essential frequencies is bounded above by the number $n(k) \leq N(k, \frac{K}{k})$ of eigenvalues in the interval. Weyl's law for $\sqrt{-\Delta}$ allows considerable clustering and only gives $N(k, \frac{K}{k}) = o(k)$ in the case where periodic orbits have measure zero. For instance, the quasi-eigenvalue might be a true eigenvalue with multiplicity saturating the Weyl bound. But a typical interval has a uniformly bounded number of Δ-eigenvalues in dimension 2 or equivalently a frequency interval of with $O(\frac{1}{\mu_k})$ has a uniformly bounded number of frequencies. The dichotomy above reflects the dichotomy as to whether exceptional clustering of eigenvalues occurs around the quasi-eigenvalues n^2 of Δ or whether there is a uniform bound on $N(k, \delta)$.

PROPOSITION 9.13. *If there exists a quasimode* $\{\psi_k\}$ *of order* 0 *for* Δ *with the properties:*

- *(i)* $n(k) \leq C$ *for all* k*;*
- *(ii)* $\langle A\psi_k, \psi_k\rangle \to \int_{S^*M} \sigma_A d\mu$ *where* $d\mu \neq d\mu_L$*,*

then Δ *is not QUE.*

The proof is based on the following lemma pertaining to near off-diagonal Wigner distributions. It gives an "everywhere" version of the off-diagonal part of Theorem 9.7 (ii).

LEMMA 9.14. *Suppose that* G^t *is ergodic and* Δ *is QUE. Suppose that* $\{(\lambda_{i_r}, \lambda_{j_r}),\ i_r \neq j_r\}$ *is a sequence of pairs of eigenvalues of* $\sqrt{-\Delta}$ *such that* $\lambda_{i_r} - \lambda_{j_r} \to 0$ *as* $r \to \infty$*. Then* $dW_{i_r, j_r} \to 0$*.*

PROOF. Let $\{\lambda_i, \lambda_j\}$ be any sequence of pairs with the gap $\lambda_i - \lambda_j \to 0$. Then by Egorov's theorem, any weak* limit $d\nu$ of the sequence $\{dW_{i,j}\}$ is a measure invariant under the geodesic flow. The weak limit is defined by the property that

$$\langle A^* A\varphi_i, \varphi_j\rangle \to \int_{S^*M} |\sigma_A|^2\, d\nu. \tag{9.55}$$

If the eigenfunctions are real, then $d\nu$ is a real (signed) measure.

We now observe that any such weak* limit must be a constant multiple of Liouville measure $d\mu_L$. Indeed, we first have:

$$|\langle A^* A\varphi_i, \varphi_j\rangle| \leq |\langle A^* A\varphi_i, \varphi_i\rangle|^{1/2} |\langle A^* A\varphi_j, \varphi_j\rangle|^{1/2}. \tag{9.56}$$

Taking the limit along the sequence of pairs, we obtain

$$\left|\int_{S^*M} |\sigma_A|^2\, d\nu\right| \leq \int_{S^*M} |\sigma_A|^2\, d\mu_L. \tag{9.57}$$

It follows that $d\nu \ll d\mu_L$ (absolutely continuous). But $d\mu_L$ is an ergodic measure, so if $d\nu = f d\mu_L$ is an invariant measure with $f \in L^1(d\mu_L)$, then f is constant. Thus,

$$d\nu = C d\mu_L \quad \text{for some constant } C. \tag{9.58}$$

We now observe that $C = 0$ if $\varphi_i \perp \varphi_j$ (i.e., if $i \neq j$). This follows if we substitute $A = I$ in (9.55), use orthogonality and apply (9.58). □

We now complete the proof of the proposition by arguing by contradiction. The frequencies must come from a shrinking frequency interval, so the hypothesis of the

Proposition is satisfied. If Δ were QUE, we would have in the notation of (9.54):

$$\langle A\psi_k, \psi_k\rangle = \sum_{i,j=1}^{n(k)} c_{kj} c_{ki} \langle A\varphi_i, \varphi_j\rangle + o(1) \tag{9.59}$$

$$= \sum_{j=1}^{n(k)} c_{kj}^2 \langle A\varphi_j, \varphi_j\rangle + \sum_{i\neq j=1}^{n(k)} c_{kj} c_{ki} \langle A\varphi_i, \varphi_j\rangle + o(1) \tag{9.60}$$

$$= \int_{S^*M} \sigma_A \, d\mu_L + o(1), \tag{9.61}$$

by Proposition 9.14. This contradicts (ii). In the last line we used $\sum_{j=1}^{n(k)} |c_{kj}|^2 = 1 + o(1)$, since $\|\psi_k\|_{L^2} = 1$. This completes the proof of Proposition 9.13.

9.6.1. Proof of Hassell's scarring result. We apply and develop this reasoning in the case of the stadium. The quasi-eigenvalues of the Bunimovich stadium corresponding to bouncing ball quasimodes are n^2 independently of the diameter t of the inner rectangle.

By the above, it suffices to show that that there exists a sequence $n_j \to \infty$ and a constant M (independent of j) so that there exist $\leq M$ eigenvalues of Δ in $[n_j^2 - K, n_j^2 + K]$. An somewhat different argument is given in [**Ha**] in this case: For each n_j there exists a normalized eigenfunction u_{k_j} so that $\langle u_{k_j}, v_{k_j}\rangle \geq \sqrt{\frac{3}{4}M}$. It suffices to choose the eigenfunction with eigenvalue in the interval with the largest component in the direction of v_{k_j}. There exists one since

$$\|\tilde{\Pi}_{[n^2-K, n^2+K]} v_n\| \geq \frac{3}{4}. \tag{9.62}$$

The sequence $\{u_{n_k}\}$ cannot be Liouville distributed. Indeed, for any $\varepsilon > 0$, let A be a self-adjoint semi-classical pseudo-differential operator properly supported in the rectangle so that $\sigma_A \leq 1$ and so that $\|(Id - A)v_n\| \leq \varepsilon$ for large enough n. Then

(9.63)

$$\langle A^2 u_{k_j}, u_{k_j}\rangle = \|A u_{k_j}\|^2 \geq \left|\langle A u_{k_j}, v_{k_j}\rangle\right|^2$$

$$= \left|\langle u_{k_j}, A v_{k_j}\rangle\right|^2 \geq \left(|\langle u_{k_j}, v_{k_j}\rangle| - \varepsilon\right)^2 \geq \left(\sqrt{\frac{3}{4}M} - \varepsilon\right)^2. \tag{9.64}$$

Choose a sequence of operators A such that $\|(\mathrm{Id} - A)v_n| \to 0$ and so that the support of σ_A shrinks to the set of bouncing ball covectors. Then the mass of any quantum limit of $\{u_{n_k}\}$ must have mass $\geq \frac{3}{4}M$ on Λ.

Thus, the main point is to eliminate the possibility of exceptional clustering of eigenvalues around the quasi-eigenvalues. In fact, no reason is known why no exceptional clustering should occur. Hassell's idea is that it can however only occur for a measure zero set of diameters of the inner rectangle. The proof is based on Hadamard's variational formula for the variation of Dirichlet or Neumann eigenvalues under a variation of a domain. In the case at hand, the stadium is varied by horizontally (but not vertically) expanding the inner rectangle. In the simplest case of Dirichlet boundary conditions, the eigenvalues are forced to decrease as the rectangle is expanded. The QUE hypothesis forces them to decrease at a uniform rate. But then they can only rarely cluster at the fixed quasi-eigenvalues n^2. If this ever happened, the cluster would move left of n^2 and there would not be time for a new cluster to arrive.

Here is a more detailed sketch. Under the variation of X_t with infinitesimal variation vector field ρ_t, Hadamard's variational formula in the case of Dirichlet boundary conditions gives

$$\frac{dE_j(t)}{dt} = \int_{\partial X_t} \rho_t(s)(\partial_n u_j(t)(s))^2 \, ds, \tag{9.65}$$

where as above ρ_t is the variation of the boundary. Let u^b denote the Cauchy data of u: $u^b = u|_{\partial X_t}$ in the case of Neumann boundary conditions, resp. $u^b = \partial_n|_{\partial X_t}$ in the case of Dirichlet boundary conditions. Then

$$E_j^{-1} \frac{d}{dt} E_j(t) = -\int_{\partial X_t} \rho_t(s) u_j^b(s)^2 \, ds. \tag{9.66}$$

Let $A(t)$ be the area of S_t. By Weyl's law, $E_j(t) \sim c\frac{j}{A(t)}$. Since the area of X_t grows linearly, we have on average $\dot{E}_j \sim -C\frac{E_j}{A(t)}$. Theorem 9.7 gives the asymptotics individually for almost all eigenvalues. Let

$$f_j(t) = \int_{\partial X_t} \rho_t(s) |u_j^b(t;s)|^2 \, ds.$$

Then $\dot{E}_j = -E_j f_j$. Then Theorem 9.7 implies that $|u_j^b|^2 \to \frac{1}{A(t)}$ weakly on the boundary along a subsequence of density one. QUE is the hypothesis that this occurs for the entire sequence, i.e.,

$$f_j(t) \to \frac{k}{A(t)} > 0, \quad k := \int_{\partial S_t} \rho_t(s) ds.$$

Hence,

$$\frac{\dot{E}_j}{E} = -kA(t)(1 + o(1)), \quad j \to \infty.$$

Hence there is a lower bound to the velocity with which eigenvalues decrease as $A(t)$ increases. Eigenvalues can therefore not concentrate in the fixed quasimode intervals $[n^2 - K, n^2 + K]$ for all t. But then there are only a bounded number of eigenvalues in this interval; so Proposition 9.13 implies QUE for the other X_t. A more detailed analysis shows that QUE holds for almost all t.

9.7. Appendix on Duhamel's formula

Suppose that $S(t)$ is the solution operator for the initial value problem

$$\begin{cases} W_t + AW = 0, \\ U(0) = \Phi \end{cases}$$

Then the solution of the inhomogeneous problem

$$\begin{cases} W_t + WU = F, \\ W(0) = \Phi \end{cases}$$

is given by

$$W(t) = S(t)\Phi + \int_0^t S(t-s)F(s)ds. \tag{9.67}$$

Now consider the inhomogeneous initial value problem for the wave equation:

$$\begin{cases} u_{tt} - \Delta u = f(x,t), \\ u(x,0) = \varphi(x), \quad u_t(x,0) = \psi(x). \end{cases} \tag{9.68}$$

PROPOSITION 9.15. *In the case of* $\mathbb{R}^3$, *the solution is*

$$u(x,t) = \frac{1}{4\pi t^2} \int_{\partial B(x,t)} [\varphi(y) + \nabla\varphi(y)\cdot(y-x) + t\psi(y)]\, dS(y), \tag{9.69}$$

$$+ \int_0^t \frac{1}{4\pi(t-s)} \int_{\partial B(x,t-s)} f(y,s)\, dS(y) ds. \tag{9.70}$$

PROOF. Convert this to a first order system:

$$\begin{cases} \begin{pmatrix} u \\ v \end{pmatrix}_t = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}, \\ \begin{pmatrix} u \\ v \end{pmatrix}_{t=0} = \begin{pmatrix} \varphi \\ \psi \end{pmatrix}. \end{cases} \tag{9.71}$$

Let

$$W = \begin{pmatrix} u \\ v \end{pmatrix}, \quad A = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ f \end{pmatrix}, \quad \Phi = \begin{pmatrix} \varphi \\ \psi \end{pmatrix}.$$

Apply Duhamel's formula to complete the proof. □

Bibliography

[AR12] N. Anantharaman and G. Riviere, Dispersion and controllability for the Schrdinger equation on negatively curved manifolds. Anal. PDE 5 (2012), no. 2, 313-338.

[G] B. Gutkin, Note on converse quantum ergodicity. Proc. Amer. Math. Soc. 137 (2009), no. 8, 2795–2800.

[Han15] X. Han, Small scale equidistribution of random eigenbases. Comm. Math. Phys. 349 (2017), no. 1, 425-440.

[Ha] A. Hassell, Ergodic billiards that are not quantum unique ergodic, Appendix by A. Hassell and L. Hillairet, Ann. of Math. (2) 171 (2010), no. 1, 605-619 (arXiv:0807.0666).

[HeR1] H. Hezari and G. Riviere, Quantitative equidistribution properties of toral eigenfunctions, J. Spectr. Theory 7 (2017), no. 2, 471-485 (arXiv:1503.02794).

[HeR2] H. Hezari and G. Riviere, Lp norms, nodal sets, and quantum ergodicity. Adv. Math. 290 (2016), 938-966.

[J] D. Jakobson, Quantum limits on flat tori. Ann. of Math. (2) 145 (1997), no. 2, 235-266.

[JZ] D. Jakobson and D. Zelditch, Classical limits of eigenfunctions for some completely integrable systems. Emerging applications of number theory (Minneapolis, MN, 1996), 329-354, IMA Vol. Math. Appl., 109, Springer, New York, 1999.

[Je] O. Jenkinson, Every ergodic measure is uniquely maximizing, Disc. Cont. Dyn. Syst. 16 (2006).

[MOZ] J. Marklof, S. O'Keefe, Weyl's law and quantum ergodicity for maps with divided phase space. With an appendix "Converse quantum ergodicity" by Steve Zelditch. Nonlinearity 18 (2005), no. 1, 277–304.

[RS] Z. Rudnick and P. Sarnak, The behaviour of eigenstates of arithmetic hyperbolic manifolds. Comm. Math. Phys. 161 (1994), no. 1, 195–213.

[Ru] D. Ruelle, *Statistical mechanics: Rigorous results.* W. A. Benjamin, Inc., New York-Amsterdam 1969.

[Schu1] R. Schubert, On the rate of quantum ergodicity for quantised maps. Ann. Henri Poincaré 9 (2008), no. 8, 1455–1477

[Schu2] R. Schubert, Upper bounds on the rate of quantum ergodicity. Ann. Henri Poincaré 7 (2006), no. 6, 1085–1098.

[Sc3] R. Schubert, Semiclassical behaviour of expectation values in time evolved Lagrangian states for large times. Comm. Math. Phys. 256 (2005), no. 1, 239–254.

[Si] K. Sigmund. On the space of invariant measures for hyperbolic flows, Amer. J.Math. 94 (1972), 31-37.

[Su] T. Sunada, Quantum ergodicity. Progress in inverse spectral geometry, 175–196, Trends Math., Birkhauser, Basel, 1997.

[Z1] S. Zelditch, Note on quantum unique ergodicity, Proc. Amer. Math. Soc. 132 (2004), no. 6, 1869–1872.

[Z2] S. Zelditch, Mean Lindelöf hypothesis and equidistribution of cusp forms and Eisenstein series. J. Funct. Anal. 97 (1991), no. 1, 1–49.

[Z3] S. Zelditch, Quantum transition amplitudes for ergodic and for completely integrable systems. J. Funct. Anal. 94 (1990), no. 2, 415–436.

[Z4] S. Zelditch, On the rate of quantum ergodicity. I. Upper bounds. Comm. Math. Phys. 160 (1994), no. 1, 81-92.

CHAPTER 10

L^p norms

This is the first of several sections devoted to the growth as $\lambda \to \infty$ of L^p norms of L^2-normalized eigenfunctions

$$-\Delta_g \varphi_\lambda(x) = \lambda^2 \varphi_\lambda(x), \quad x \in M. \tag{10.1}$$

As seen earlier, most estimates of eigenfunctions involve quadratic functionals such as matrix elements $\langle A\varphi_\lambda, \varphi_\lambda \rangle$. The L^p norm

$$\|\varphi_\lambda\|_{L^p} = \left(\int_M |\varphi_\lambda|^p \, dV \right)^{\frac{1}{p}} \tag{10.2}$$

of an L^2-normalized eigenfunction is a more 'nonlinear' measure of its concentration. We are interested in both upper and lower bounds on L^p norms, in identifying *extremal sequences* and in comparisons of extremals for different Riemannian manifolds (M, g).

We denote eigenspaces by

$$V_\lambda = \{\varphi \colon \Delta\varphi = -\lambda^2 \varphi\}, \tag{10.3}$$

and measure the upper growth rate of L^p norms by

$$L^p(\lambda, g) = \sup_{\varphi \in V_\lambda \colon \|\varphi\|_{L^2} = 1} \|\varphi\|_{L^p}. \tag{10.4}$$

Although there exist few if any results in the opposite extreme on "flat eigenfunctions," we could also consider

$$\ell^p(\lambda, g) = \inf_{\varphi \in V_\lambda \colon \|\varphi\|_{L^2} = 1} \|\varphi\|_{L^p}. \tag{10.5}$$

Recall $\Pi_{[0,\lambda]} = \Pi_\lambda$ is the spectral projection for $\sqrt{-\Delta}$ to the interval $[0, \lambda]$ and $\chi_\lambda := \Pi_{[\lambda, \lambda+1]}$ is the projection for eigenvalues $\lambda_j \in [\lambda, \lambda + 1]$.

As mentioned in the introduction (for surfaces), C. Sogge [**So1**] has proved universal bounds on L^p norms of eigenfunctions in the sense that they hold for any compact Riemannian manifold (M, g) with O-bounds dependent only on g:

THEOREM 10.1. *For any compact Riemannian manifold (M, g) of dimension n, we have*

$$\sup_{\varphi \in V_\lambda} \frac{\|\varphi\|_{L^p}}{\|\varphi\|_{L^2}} = O(\lambda^{\delta(p)}), \quad 2 \leq p \leq \infty, \tag{10.6}$$

where

$$\delta(p) = \begin{cases} n(\frac{1}{2} - \frac{1}{p}) - \frac{1}{2} & \textit{for } \frac{2(n+1)}{n-1} := p_n \leq p \leq \infty, \\ \frac{n-1}{2}(\frac{1}{2} - \frac{1}{p}) & \textit{for } 2 \leq p \leq \frac{2(n+1)}{n-1}. \end{cases} \tag{10.7}$$

The growth of L^p norms thus has two 'phases' which meet at the critical index $p_n = \frac{2(n+1)}{n-1}$.

Definition 10.2. Given (M, g) we say that a subsequence $\{\varphi_{j_k}\}$ of L^2-normalized eigenfunctions is an *extremal sequence* if $\|\varphi_{j_k}\|_{L^p} = \Omega(\lambda_{j_k}^{\delta(p)})$. We say that (M, g) has maximal L^p eigenfunction growth if it possesses an extremal sequence.

Above, Ω is the lower bound symbol, i.e., $f(\lambda) = \Omega(g(\lambda))$ if there exists a constant $C > 0$ so that $f(\lambda_{j_k}) \geq C g(\lambda_{j_k})$ for some subsequence λ_{j_k}. Since the constant C is not specified, an extremal sequence is only extremal in terms of order of magnitude in λ_{j_k}. In some cases, there is a natural candidate for the 'best constant' $C = C_g$ depending on g. As will be reviewed below, the extremal (optimal) sequences which saturate the inequality have different shapes above and below p_n and at $p = p_n$ there exist a wide variety of extremal shapes.

Two problems motivate this section; only the first problem has been partially solved.

Problem 10.3.

- Characterize (M, g) with maximal L^∞ eigenfunction growth. The same sequence of eigenfunctions should saturate all L^p norms with $p \geq p_n := \frac{2(n+1)}{n-1}$.
- Characterize (M, g) with maximal L^p eigenfunction growth for $2 \leq p \leq \frac{2(n+1)}{n-1}$.

Although the problems seem well out of reach at present, we also mention two related extremal problems:

- Characterize (M, g) possessing an orthonormal basis of eigenfunctions for which $\|\varphi_j\|_{L^1} \geq C > 0$. There are related extremal problems for maximal vanishing order or for nodal volumes in balls. By Hölder's inequality, an upper bound on the L^4 norm (say) gives a lower bound on the L^1 norms, since

$$1 = \|\varphi_\lambda\|_{L^2}^{1/\theta} \leq \|\varphi_\lambda\|_{L^1} \|\varphi_\lambda\|_{L^p}^{\frac{1}{\theta}-1}, \quad \theta = \frac{p}{p-1}\Big(\frac{1}{2} - \frac{1}{p}\Big) = \frac{(p-2)}{2(p-1)}. \tag{10.8}$$

- Characterize (M, g) possessing an orthonormal basis of eigenfunctions which are uniformly bounded in L^∞ (or another L^p norm for $p > 2$). It is an open problem to find geometric or dynamical conditions on (M, g) so that it has an orthonormal basis of eigenfunctions, or even one orthonormal sequence of eigenfunctions, with uniformly bounded L^∞ norms (§10.12).
- It is also interesting to study minimizers of norms, and this problem arises (more often in the holomorphic context) in the theory of Abrikosov lattices. There are almost no results on $\ell^p(\lambda, g)$ or on flat eigenfunctions except in the completely integrable case [**TZ**].

The main themes examined in this section are the following:

- There is a 'phase transition' at $p = p_n = \frac{2(n+1)}{n-1}$. As a rough heuristic, high L^p norms measure concentration around single points. Low L^p norms measure concentration around larger sets such as closed geodesics.
- The Sogge bounds for $p > \frac{2(n+1)}{n-1}$ are achieved by sequences of eigenfunctions on the round sphere S^n. But they are rarely sharp on other (M, g)

(§10.4). Extremal sequences for the functional $L^p(\lambda, g)$ for $p > \frac{2(n+1)}{n-1}$ resemble zonal spherical harmonics: they blow up at a 'pole' with the property that geodesics from the pole return to the pole.

- For $p < \frac{2(n+1)}{n-1}$, the known extremal sequences resembles a highest weight spherical harmonic – a Gaussian beam along a stable elliptic geodesic. It is a conjecture that all extremal sequences are of this kind.

One may pose the same problems for restrictions of eigenfunctions to hypersurfaces or closed geodesics. It is revealing to determine the structure of extremals for the various L^p inequalities.

10.1. Discrete Restriction theorems

The L^p bounds on eigenfunctions are applications of 'discrete L^2 restriction theorems,' which are discussed in detail in [**So2**, §5.1]. We briefly review discrete restriction theorems, both in the classical setting of Euclidean harmonic analysis on $\mathbb{R}^n$ or T^n and on general Riemannian manifolds. Let $\Pi_{[0,\lambda]}$ denote the orthogonal projection onto the span of eigenfunctions φ_j with $\lambda_j \leq \lambda$. Also let Let $\Pi_{[\lambda,\lambda+1]}$ denote the orthogonal projection onto the span of eigenfunctions φ_j with $\lambda \leq \lambda_j \leq \lambda + 1$. A stronger result than (10.1), due to Sogge, is the following.

THEOREM 10.4. *Let (M, g) be a compact Riemannian manifold of dimension n and let $f \in L^p(M)$. Let $\delta(p)$ be as in* (10.7)*. Then*

$$\|\Pi_{[0,\lambda]} f\|_{L^2(M)} \leq \begin{cases} C\lambda^{\delta(p)}\|f\|_p & \text{for } 1 \leq p \leq \frac{2(n+1)}{n+3}, \\ C\lambda^{\frac{(n-1)(2-p)}{4p}} & \text{for } \frac{2(n+1)}{n+3} \leq p \leq 2. \end{cases} \tag{10.9}$$

Dually,

$$\|\Pi_{[\lambda,\lambda+1]} f\|_{L^q(M)} \leq \begin{cases} C\lambda^{\delta(q)}\|f\|_2 & \text{for } \frac{2(n+1)}{n-1} \leq q < \infty, \\ C\lambda^{\frac{(n-1)(2-q')}{4q'}} & \text{for } 2 \leq q \leq \frac{2(n+1)}{n-1}. \end{cases} \tag{10.10}$$

The bounds (10.1) of course follow from the special case $f = \varphi_\lambda$.

As mentioned above, the estimates for $p > p_n = \frac{2(n+1)}{n-1}$ are sharp for the standard round metric on S^n and also for 'surfaces of revolution,' but as we discuss below, they are rarely sharp in the general setting of Riemannian manifolds (M, g). In particular the estimates are far from sharp on a flat torus or on Euclidean $\mathbb{R}^n$.

In the case of $\mathbb{R}^n$, the Stein-Tomas theorem [**Tom**] says that if $f \in L^p(\mathbb{R}^n)$ with $1 \leq p \leq \frac{2n+2}{n+3}$, then

$$\|\hat{f}|_{S^{n-1}}\|_{L^2(S^{n-1})} \leq C_{p,n}\|f\|_{L^p(\mathbb{R}^n)}. \tag{10.11}$$

The spectrum of $\Delta_{\mathbb{R}^n}$ on $\mathbb{R}^n$ is continuous and $\hat{f}(\xi) = \langle f, e_\xi \rangle$ where $e_\xi = e^{2\pi i \langle x, \xi \rangle}$. So this restriction theorem is a special case of the problem of estimating $\sum_j |\langle f, \varphi_j \rangle|^2$ for λ_j in a very short window. Eigenspaces for $\Delta_{\mathbb{R}^n}$ are infinite dimensional, so the restriction to a single eigenspace $\{|\xi|^2 = 1\}$ is an interesting problem. A dual 'extension' problem (by the Plancherel theorem) is to study norms of Fourier transforms $\hat{\mu}(\xi) = \int_{S^{n-1}} e^{-2\pi i \langle x, \xi \rangle} d\mu(\xi)$ of measures μ supported on S^{n-1}, for instance measures $d\mu = F d\sigma$ where $d\sigma$ is the standard volume form and $F \in L^p$. The extension estimate has the form

$$\|\widehat{(F d\sigma)}\|_{L^{p'}} \leq C\|F\|_{L^{q'}(S^{n-1})}. \tag{10.12}$$

See [**Tao**] for a survey.

In the case of the flat torus $\mathbb{R}^n/\mathbb{Z}^n$, there are many studies of L^p norms of the eigenfunctions

$$\varphi_{N,\boldsymbol{a}}(x) := \sum_{k \in S^{n-1}(N)\cap \mathbb{Z}^n} a_k e^{2\pi i \langle k, x\rangle} \tag{10.13}$$

of eigenvalue N^2 in terms of $\|\varphi_{\boldsymbol{a}}\|_{L^2} = \|\boldsymbol{a}\|_{\ell^2}$. Here, $S^{n-1}(N)$ is the sphere of radius $N \in \mathbb{Z}_+$. The restriction conjecture is that, for $p > p_n := \frac{2n}{n-2}$,

$$||\varphi_{N,\boldsymbol{a}}||_{L^p(T^n)} \leq C_\varepsilon N^{\frac{n-2}{2} - \frac{n}{p} + \varepsilon} \|\boldsymbol{a}\|_{\ell^2} \tag{10.14}$$

and that for $2 \leq p < p_n$,

$$\|\varphi_{N,\boldsymbol{a}}\|_{L^p(T^n)} \leq C_\varepsilon N^\varepsilon \|\boldsymbol{a}\|_{\ell^2}. \tag{10.15}$$

This discrete (lattice) restriction theorem has been studied intensively since [**B2**]. For recent results we refer to [**BD**]. These results of course make fundamental use of the properties of the exponentials $e^{2\pi i\langle k,x\rangle}$. The discrete lattice point restriction theorem could be stated for more general quantum integrable systems besides the flat torus, but the joint eigenfunctions do not satisfy the same bounds and one would not expect the same estimates.

10.2. Random spherical harmonics and extremal spherical harmonics

In this section we compute critical L^p norms of extremal spherical harmonics to verify that they are extremal.

10.2.1. Zonal spherical harmonics. The zonal spherical harmonic with pole at w is the L^2-normalized spectral projections kernels $Z^w(z) = \frac{\Pi_N(z,w)}{\sqrt{\Pi_N(w,w)}}$.When $z = w$, $Z^w(w) = \sqrt{\Pi_N(w,w)}$. Note that $\Pi_N(w,w)$ is constant on S^n with value equal to $\frac{\dim V_N}{\text{Vol}(S^n)}$ where V_N is the space of spherical harmonics of degree N and $\Pi_N : L^2(S^n) \to V_N$ is the orthogonal projection. Since $\dim V_N \simeq N^{n-1}$, one has $\|\Phi_N^w\|_{L^\infty} \simeq CN^{\frac{n-1}{2}}$.

10.2.2. Highest weight spherical harmonics on S^n saturate the L^p norm bounds for $p \leq p_n$. Let us prove that for $n = 2$ (and $p_n = 6$), the sequence of L^2-normalized highest weight spherical harmonics has L^6 norm growing like $k^{\frac{1}{6}}$.

We first have to L^2- normalize the highest weight spherical harmonics $(x+iy)^k$. We claim that $\|(x+iy)^k\|_{L^2(S^2)} \sim k^{-1/4}$. This is most easily proved using Gaussian integrals:

(10.16)

$$\begin{cases} \int_{\mathbb{R}^3} (x^2+y^2)^k e^{-(x^2+y^2+z^2)}\, dxdydz = \|(x+iy)^k\|^2_{L^2(S^2)} \int_0^\infty r^{2k} e^{-r^2} r^2\, dr \\ \int_{\mathbb{R}^3} (x^2+y^2)^k e^{-(x^2+y^2+z^2)}\, dxdydz = \int_{\mathbb{R}^2} (x^2+y^2)^k e^{-(x^2+y^2)} = \int_0^\infty r^{2k} e^{-r^2} r\, dr \end{cases}$$

(10.17)

$$\implies \|(x+iy)^k\|^2_{L^2(S^2)} = \frac{\Gamma(k+1)}{\Gamma(k+\frac{3}{2})} \sim k^{-1/2}.$$

Thus, the L^2 normalized highest weight harmonic is essentially $k^{1/4}(x+iy)^k$. It achieves its L^∞ norm at $(1,0,0)$ where it has size $k^{1/4}$.

Next we show that it is an extremal for L^p for $2 \le p \le 6$. We calculate the L^6 norm again by Gaussian integrals:

(10.18)

$$\begin{cases} \int_{\mathbb{R}^3} (x^2+y^2)^{3k} e^{-(x^2+y^2+z^2)}\,dxdydz = \|(x+iy)^k\|^6_{L^6(S^2)} \int_0^\infty r^{6k} e^{-r^2} r^2\,dr \\ \int_{\mathbb{R}^3} (x^2+y^2)^{3k} e^{-(x^2+y^2+z^2)}\,dxdydz = \int_{\mathbb{R}^2} (x^2+y^2)^{3k} e^{-(x^2+y^2)}\,dxdy = \int_0^\infty r^{6k} e^{-r^2} r\,dr \end{cases}$$

(10.19)

$$\Longrightarrow \|(x+iy)^k\|^6_{L^6(S^2)} = \frac{\Gamma(6k+1)}{\Gamma(6k+\frac{3}{2})} \sim k^{-1/2}.$$

Hence, the L^6 norm of $k^{1/4}(x+iy)^k$ is of order

$$k^{1/4}k^{-1/12} = k^{1/6}. \tag{10.20}$$

Since $\lambda_k \sim k^2$ and $\delta(6) = \frac{1}{6}$ in dimension 2, we see that it is an extremal.

Remark 10.5. If we pick a 'random spherical harmonic' of degree k from the unit sphere in $L^2 \cap \mathcal{H}_k$, then the average value of the L^∞ norm is $\sim \sqrt{\log k}$ and the average of the other L^p norms are bounded as $k \to \infty$. See [**CH**] for a proof (related to one in [**FZ**] in the complex domain). Thus, random spherical harmonics are rather flat but do have peaks of order $\sqrt{\log k}$. It is not known if there exists an orthonormal basis of spherical harmonics with uniformly bounded L^∞ norms.

10.3. Sketch of proof of the Sogge L^p estimates

We sketch the proof of the Sogge estimates following [**So2**]. The estimates are proved by interpolation from three estimates: $p = 2, \infty, \frac{2(n+1)}{n-1}$ (for $n \neq 1$). The L^2 estimate is of course trivial. The L^∞ estimate already followed from the pointwise local Weyl law with remainder estimated by the Fourier Tauberian method. Thus the key point is to prove an $L^2 \to L^{\frac{2(n+1)}{n-1}}$ mapping norm estimate on the spectral projections $\chi_\lambda := \Pi_{[\lambda,\lambda+1]}$. To obtain $L^2 \to L^p$ mapping norms of oscillatory integral operators which are bounded on $L^2 \to L^2$, it suffices by the Riesz interpolation theorem to prove the case for ♣Note that $p_n^* < 2$. The constant $\delta(p_n^*)$ has not been defined for such p_n^*.

$$p_n^* = \frac{2(n+1)}{n+3} \text{ the dual exponent to } p_n = \frac{2n+2}{n-1}. \tag{10.21}$$

Thus, the key point is to prove that

$$\|\Pi_{[\lambda,\lambda+1]} f\|_{L^{p_n}} \le C(1+\lambda)^{\frac{n-1}{2n+2}} \|f\|_{L^2}, \tag{10.22}$$

or equivalently,

$$\|\Pi_{[\lambda,\lambda+1]} f\|_{L^2} \le C(1+\lambda)^{\delta(p_n^*)} \|f\|_{p_n^*}, \qquad \delta(p) = n\left|\frac{1}{p} - \frac{1}{2}\right| - \frac{1}{2}. \tag{10.23}$$

♣

We only sketch the proof when $f = \varphi_\lambda$ is an eigenfunction. In this case, the proof involves wave equation techniques. Instead of using a spectral projections kernel, which reproduces all eigenfunctions in a spectral interval, we use a "designer reproducing kernel K_λ" for φ_λ.

We introduce a cutoff $\hat{\rho} \in C_0^\infty(\mathbb{R})$ satisfying $\rho(0) = \int \hat{\rho}\, dt = 1$ and define the operator

$$\rho(\lambda - \sqrt{-\Delta}) : L^2(M) \to L^2(M) \tag{10.24}$$

by

$$\rho(\lambda - \sqrt{-\Delta})f = \int_{\mathbb{R}} \hat{\rho}(t) e^{it\lambda} e^{-it\sqrt{-\Delta}} f\, dt. \tag{10.25}$$

Then (10.24) is a function of Δ and has φ_λ as an eigenfunction with eigenvalue $\rho(\lambda - \lambda) = \rho(0) = 1$. It has the reproducing property $\rho(\lambda - \sqrt{-\Delta})\varphi_\lambda = \varphi_\lambda$. We choose ρ further so that $\hat{\rho}(t) = 0$ for $t \notin [\varepsilon/2, \varepsilon]$. The operator (10.24) is a semi-classical Fourier integral operator T_λ whose kernel is of the form

$$K_\lambda(x,y) = \lambda^{\frac{n-1}{2}} a_\lambda(x,y) e^{i\lambda r(x,y)}, \tag{10.26}$$

where $a_\lambda(x,y)$ is bounded with bounded derivatives in (x,y) and where $r(x,y)$ is the Riemannian distance between points. Using a TT^* argument the estimate (10.22) is equivalent to

$$\|T_\lambda T_\lambda^* g\|_{L^{p_n}} \leq C \lambda^{-\frac{n(n-1)}{n+1}} \|f\|_{L^{p_n^*}}. \tag{10.27}$$

Except for the pre-factor $\lambda^{\frac{n-1}{2}}$, the oscillatory integral operator (10.26) is a semi-classical Fourier integral operator

$$T_\lambda f(x) = \int_{\mathbb{R}^n} e^{i\lambda\psi(x,y)} a(x,y) f(y)\, dy, \tag{10.28}$$

where $a \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^n)$. The mapping norms of such operators on L^p have been studied by Carleson-Sjolin, Hörmander [**Hor1**], Stein [**St, St**], Sogge [**So2**], Bourgain [**B1**] and others. The simplest case occurs for non-degenerate phases like $\psi(x,y) = \langle x, y\rangle$ for which $\det \partial^2\psi/\partial x \partial y \neq 0$. In this case then $\|T_\lambda\|_{L^p \to L^{p^*}} \leq \lambda^{-n/p}$ by an extension of the Hausdorff-Young inequality [**Hor1**].

However, the phase of (10.26) is degenerate in the sense that $x \to \nabla r(x,y)$ locally maps an n-dimensional ball to the surface of a sphere. Thus, $r(x,y)$ satisfies the homogeneous Monge-Ampere equation $\det \partial^2_{x,y} r(x,y) = 0$. However, it is still quite non-degenerate in that the gradient map has rank $n-1$ and its image is a positively curved hypersurface. In the terminology of [**So2**, §2.2], T_λ with kernel (10.26) is said to satisfy the $n \times n$ Carleson-Sjolin condition, namely that the the projections of the critical set

$$C_\psi = \{(x, \varphi'_x, y, -\varphi'_y)\} \subset T^*(\mathbb{R}^n \times T^*\mathbb{R}^n) \tag{10.29}$$

to $T^*\mathbb{R}^n$ have rank $2n-1$ and the maps

$$C_\psi \to S_x = \{(x, \varphi'_x\} \subset T^*\mathbb{R}^n, \quad C_\psi \to S_y = \{(y, -\varphi'_y\} \subset T^*\mathbb{R}^n \tag{10.30}$$

are immersions to hypersurfaces of positive Gaussian curvature. Corollary 2.2.3 of [**So2**] says that for $a \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^n)$, and for such $n \times n$ Carleson-Sjolin phases,

$$\left\| \int_{\mathbb{R}^n} e^{i\lambda\psi(x,y)} a(x,y) f(y)\, dy \right\|_{L^{\frac{2n+2}{n-1}}(\mathbb{R}^n)} \leq \lambda^{-\frac{n(n-1)}{2(n+1)}} \|f\|_{L^2(\mathbb{R}^n)}. \tag{10.31}$$

The representation is local and the estimate holds on manifolds.

The original Carleson-Sjolin oscillatory integral operators in [**So2**, §2.2] have the form

$$T_\lambda f(x) = \int_{\mathbb{R}^{n-1}} e^{i\lambda\varphi(z,y)} a(z,y) f(y)\, dy, \quad (z,y) \in \mathbb{R}^n \times \mathbb{R}^{n-1} \tag{10.32}$$

with $a \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$ and where for fixed z, the projection of the critical set

$$T^*(\mathbb{R}^n \times T^*\mathbb{R}^{n-1}) \supset C_\varphi = \{(z, \varphi_z', y, -\varphi_y')\} \to S_z = \{(z, \varphi_z'\} \subset T^*\mathbb{R}^n \tag{10.33}$$

is a hypersurface of positive Gaussian curvature. Equivalently, the map $y \to \varphi_z'(z_0, y)$ from $\mathbb{R}^{n-1} \to \mathbb{R}^n$ gives a local embedding as a positively curved hypersurface for each z_0. For operators satisfying these conditions, [**So2**, (2.2.9)] asserts that

$$\|T_\lambda f\|_{L^{\frac{2n+2}{n-1}}(\mathbb{R}^n)} \leq C\lambda^{-\frac{n(n-1)}{2n+2}} \|f\|_{L^2(\mathbb{R}^{n-1})}, \tag{10.34}$$

and by (2.2.10) (loc.cit.) this is equivalent to (10.27).

A few words on the proof of (10.27) and (10.34). The estimate (10.31) for $n \times n$ Carleson-Sjolin integral operators follows from the estimate (10.34). One introduces coordinates $y = (u, t)$ so that for fixed t the phase satisfies the Carleson-Sjolin condition and applies (10.34) together with standard estimates (see the proof of [**So2**, Corollary 2.2.3]). Regarding the proof of (10.34), since the rank of the Hessian of $r(x, y)$ is $n - 1$, the idea is to split off a t variable so that the mixed Hessian is non-degenerate in the remaining (x, y) variables. One then defines the operators

$$T_t^\lambda f(x) = \int_{\mathbb{R}^{n-1}} e^{i\lambda\varphi(t,x,y)} a(t,x,y)\, dy.$$

The proof of (10.34) uses the 1-dimensional Hardy-Littlewood-Sobolev inequality to reduce the estimate to

$$\|T_t^\lambda T_{t'}^{\lambda *} f\|_{L^{p_n}(\mathbb{R}^{n-1})} \leq C|t - t'|^{1-(\frac{1}{p_n^*} - \frac{1}{p_n})} \lambda^{-\frac{n(n-1)}{n+1}} \|f\|_{L^{p_n^*}(\mathbb{R}^{n-1})}. \tag{10.35}$$

This is [**So2**, (2.2.10')]. But this follows by interpolation between $L^2 \to L^2$ and $L^1 \to L^\infty$ mapping norm estimates. Of these, only the second is difficult. It uses the curvature hypothesis to prove that the Schwartz kernel of $T_t^\lambda T_{t'}^{\lambda *}$ satisfies

$$|K_{t,t'}^\lambda(x, x')| \leq C \left(\lambda |(x,t) - (x', t')|\right)^{-\frac{n-1}{2}}.$$

This implies that

$$\|T_t^\lambda T_{t'}^{\lambda *} f\|_{L^\infty(\mathbb{R}^{n-1})} \leq C\lambda^{-\frac{n-1}{2}} |t - t'|^{-\frac{n-1}{2}} \|f\|_{L^1(\mathbb{R}^{n-1})}.$$

This implies (10.35); for the details we refer again to [**So2**].

10.4. Maximal eigenfunction growth

The bounds of Theorem 10.1 are achieved by special sequences of eigenfunctions (spherical harmonics) on the standard S^n. However, it is rare that a Riemannian manifold possesses any sequence of eigenfunctions achieving these bounds. In this section, we give geometric conditions on (M, g) which are necessary for the universal L^∞ bound or the Sogge L^p bounds for $p > p_n$ to be achieved. Thus we are concerned with a special case of extremal sequences in Definition 10.2.

DEFINITION 10.6. We say that (M, g) has maximal L^p eigenfunction growth if it possesses a subsequence of eigenfunctions $\varphi_{\lambda_{j_k}}$ which saturates the L^p bound. When $p = \infty$ (so that $\|\varphi_{\lambda_{j_k}}\|_{L^\infty} \geq C_0 \lambda_{j_k}^{(n-1)/2}$ for some $C_0 > 0$ depending only on (M, g)) we say that it has maximal sup norm growth.

The condition that (M, g) does not have maximal sup norm growth is that

$$\|\varphi_\lambda\|_{L^\infty(M)} = o(\lambda^{\frac{n-1}{2}}). \tag{10.36}$$

There is a stronger condition on the $L^2 \to L^\infty$ norm of projection operators onto shrinking spectral bounds, i.e.,

$$\left\|\chi_{[\lambda, \lambda+o(1)]}\right\|_{L^2(M) \to L^\infty(M)} = o(\lambda^{\frac{n-1}{2}}). \tag{10.37}$$

By this we mean that, given $\varepsilon > 0$, we can find a $\delta(\varepsilon) > 0$ and $\Lambda_\varepsilon < \infty$ so that

$$\left\|\chi_{[\lambda, \lambda+\delta(\varepsilon)]} f\right\|_{L^\infty(M)} \leq \varepsilon \lambda^{\frac{n-1}{2}} \|f\|_{L^2(M)}. \quad \lambda \geq \Lambda_\varepsilon. \tag{10.38}$$

In a series of articles [**SoZ1, SoTZ, SoZ2**], ever more stringent characterizations are given of (M, g) with maximal eigenfunction growth. In [**SoZ1**], it was shown that (M, g) of maximal L^p eigenfunction growth for $p \geq p_n$ have self-focal points. (They were also called blow-down points). The main purpose of this section is to give an exposition of the proof in [**SoZ3, SoZ4**] that (M, g) of maximal sup norm growth must possess self-focal points, and that when g is analytic the first return map must preserve an invariant L^1 measure. When $\dim M = 2$ it must possess a pole through which all the geodesics are closed.

The proof is somewhat different from that in [**SoZ3, SoZ4**] and we go further into the geometry of loops. Several purely geometric problems are raised on the possible structure of loops at a point p of a Riemannian manifold (M, g) with $\dim M \geq 3$. Solution of these problems would lead to stronger characterizations of manifolds of maximal eigenfunction ‘.

10.4.1. Geometric and dynamical notions. Given $x \in M$, we let $\mathcal{L}_x$ denote the set of loop directions at x:

$$\mathcal{L}_x = \{\xi \in S^*_x M \colon \text{there exists } T \text{ such that } \exp_x T\xi = x\}. \tag{10.39}$$

We let $T_x \colon S^*_x M \to \mathbb{R}_+ \cup \{\infty\}$ denote the return time function to x:

$$T_x(\xi) = \begin{cases} \inf\{t > 0 \colon \exp_x t\xi = x\} & \text{if } \xi \in \mathcal{L}_x, \\ +\infty & \text{if no such } t \text{ exists.} \end{cases} \tag{10.40}$$

We then define the first return map by

$$\Phi_x = G_x^{T_x} \colon \mathcal{L}_x \to S^*_x M, \tag{10.41}$$

where G^t is the homogeneous geodesic flow (the Hamilton flow of $|\xi|_x$). We also define $T^{(k)}(\xi)$ to be the time of kth return for directions which loop back at least k times.

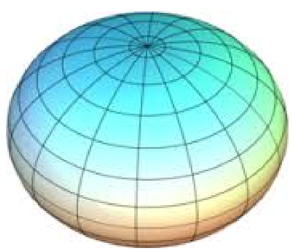
DEFINITION 10.7. We call a point p a *self-focal point* or *blow-down point* if all geodesics leaving p loop back to p at a common time T, i.e., $\exp_p T\xi = p$. (They do not have to be closed geodesics.) We call a point p a partial self-focal point if there exists a positive measure in $S^*_x M$ of directions ξ which loop back to p.

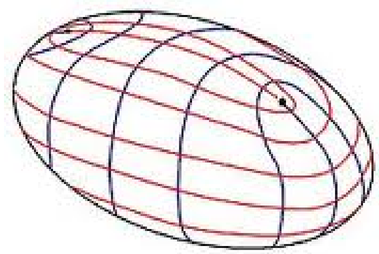
Self-focal points come in two basic kinds, depending on the first return map Φ_x. We say that x is a *pole* if

$$\Phi_x = \mathrm{Id}_x : S^*_x M \to S^*_x M.$$

Equivalently, the set $\mathcal{CL}_x$ of smoothly closed geodesics based at x is all of $S^*_x M$.

The poles of a surface of revolution are self-focal and all geodesics close up smoothly (i.e. are closed geodesics). The umbilic points of an ellipsoid are self-focal but only two directions give smoothly closed geodesics (one up to time reversal). There are topological restrictions on manifolds possessing a self-focal point. In [**Bes**] a manifold with such a point is denoted a $F^{x_0}_\ell$ (or $Y^{x_0}_\ell$-)-manifold; if ℓ is the least common return time for all loops it is denoted by $L^{x_0}_\ell$. If (M, g) has a focal point x_0 from which all geodesics are simple (non-intersecting) loops, then the integral cohomology ring $H^*(M, \mathbb{Z})$ is generated by one element [**Nak**]. For an $F^{x_0}_\ell$ manifold, $H^*(M, \mathbb{Q})$ has a single generator [**Bes**, Theorem 4]. Most results on manifolds with self-focal points consider only the special case of Zoll metrics; see [**Bes**] for classic results.





Lemma 10.8. *Assume that (M, g) is real analytic and that x is a self-focal point. Then the map $\Phi_x \colon S^*_x M \to S^*_x M$ is a real analytic orientation preserving diffeomorphism of $S^*_x M$ which is conjugate to its inverse.*

Proof. Φ_x is the restriction to $S^*_x M$ of a real analytic diffeomorphism $G^{T(x)}$ of S^*M. Hence it is a real analytic diffeomorphism.

Now let $\tau(x, \xi) = (x, -\xi)$ be the time reversal map on S^*M. Then on all of S^*M, we have $\tau G^t \tau = G^{-t}$. To see this, let Ξ_H be the Hamilton vector field of $H(x, \xi) = |\xi|_g$. Then $H \circ \tau = H$, i.e. H is time reversal invariant. We claim that $\tau_* \Xi_H = -\Xi_H$. Written in Darboux coordinates,

$$\Xi_H = \sum_j \frac{\partial H}{\partial \xi_j}\frac{\partial}{\partial x_j} - \frac{\partial H}{\partial x_j}\frac{\partial}{\partial \xi_j}.$$

If we let $(x, \xi) \to (x, -\xi)$ and use invariance of the Hamiltonian we see that the vector field changes sign.

Now, G^{-t} is the Hamilton flow of $-\Xi_H$ and that is $\tau_* \Xi_H$. But the Hamilton flow of the latter is $\tau G^t \tau$. Since $S^*_x M$ is invariant under τ, so we just restrict the identity $\tau G^t \tau = G^{-t}$ to $S^*_x M$ to see that Φ_x is conjugate to its inverse. □

10.4.2. The Lagrangian submanifold Λ_x at a self-focal point. Associated to a self-focal point x is the flowout manifold

$$\Lambda_x = \bigcup_{0 \le t \le \ell} G^t S^*_x M, \tag{10.42}$$

where ℓ is the minimal common return time. It is an immersed Lagrangian submanifold of S^*M whose projection

$$\pi\colon \Lambda_x \to M. \tag{10.43}$$

has a "blow-down singularity" at $t = 0, \ell$. For this reason self-focal points were called blow-down points in [**SoTZ**]. We may view Λ_{x_0} as the embedding of the mapping cylinder $\mathcal{C}_x$ of Φ_x, i.e., as

$$\mathcal{C}_x = S^*_x M \times [0, \ell]/\sim \text{ where } (\xi, \ell) \sim (\Phi_\ell(x, \xi), 0). \tag{10.44}$$

It is easily seen that the map

$$\iota_x(\xi, t) = G^t(x, \xi)\colon \mathcal{C}_x \to \Lambda_x \subset S^*_x M$$

is a Lagrange immersion whose image is $\Lambda_x \subset S^*_x M$. If ℓ is the minimal period of all loops (i.e., if there are no exceptionally short loops) then $\iota_z|_{S^*_z \times (0,\ell)}$ is an embedding [**SoTZ**].

It is helpful to keep in mind the following general theorem of Serre (see [**NR**] for quantitative results and background):

Theorem 10.9. *Let x, y be two (not necessarily distinct) points of a Riemannian manifold (M, g). Then there exist infinitely many geodesics between x and y.*

Corollary 10.10. *Let $\pi\colon T^*M \to M$ be the natural projection. If x is a self-focal point, then $\pi\colon \Lambda_x \to M$ is surjective.*

Thus,

$$M = \Lambda_x/\sim \text{ where } G^t(x, \xi) \sim G^s(x, \eta) \iff \exp_x t\xi = \exp_x s\eta.$$

For instance, on a surface of revolution where the distance between the two poles $\{x, \exp_x T\xi\}$ is T, $\exp_x T\xi = \exp_x T\eta$ for all $\xi, \eta \in S^*_x M$ and π is the standard map from $T^2 \to S^2$ which blows down two circles to the two poles, and is a double cover everywhere else. Since $\Lambda_x \simeq S^1 \times S^{n-1}$, existence of Λ_x puts topological restrictions on M [**Bes**].

The singularities of (10.43) have been studied in model cases in a variety of articles. For small values of t, the projection of $G^t S^*_m$ is a distance sphere $S_t(x)$ centered at x. As it evolves by its outward unit normal field, it develops singularities at the focal points to x. The blow-down singularities of a surface of revolution are very non-generic. Cusp singularities develop generically as a distance circle evolves by its normals in two dimensions.

10.4.3. Perron-Frobenius operators. In this section we consider ergodic properties of the first return map Φ_x. Recall that $\mu_x = |d\omega|$ denotes the surface area measure on $S^*_x M$ induced by the metric g_x on T^*M. We associate to the first return map (10.41) at a self-focal point the *Perron-Frobenius operator* $U_x\colon L^2(S^*_x M, |d\omega|) \to L^2(S^*_x M, |d\omega|)$ by setting (cf. [**Saf, SaV**])

$$U_x f(\xi) = f(\Phi_x(\xi))\sqrt{J_x(\xi)}, \quad f \in L^2(S^*_x M, |d\omega|), \tag{10.45}$$

where $J_x(\xi)$ denotes the Jacobian of the first return map, i.e., $\Phi^*_x|d\omega| = J_x(\xi)|d\omega|$.. Clearly U_x is a unitary operator and

$$\Phi^*_x(f\, d\mu_x) = U_x(f)\, d\mu_x. \tag{10.46}$$

Also define

$$U_x(\lambda)f(\xi) = e^{i\lambda T_x(\xi)} U_x f(\xi). \tag{10.47}$$

DEFINITION 10.11. We say that a self-focal point $x_0 \in M$ is

- *dissipative* if U_x has no invariant function $f \in L^2(S^*_x M)$. Equivalently, Φ_ℓ has no invariant measure absolutely continuous with respect to $d\mu_x$ and whose density lies in L^2.
- is a σ_0 point if U_{x_0} has a non-zero invariant L^2 function. As discussed in [**D**], it is equivalent that there exists a finite Φ_{x_0}-invariant measure on $S^*_{x_0} M$.
- is a dissipative point if, in the Hopf decomposition of Φ_{x_0} or U_{x_0}, the set of conservative points has measure zero. In this case, the spectrum of U_{x_0} is absolutely continuous; hence a dissipative point is never a σ_0 point.

The dissipative condition is a spectral condition on U_x. If U_x has any L^2 eigenfunction g then $U_x g = e^{i\theta} g$ form some $e^{i\theta} \in L^2$ and then $U_x|g| = |U_x g| = |g|$. Hence the dissipative condition is the condition that the spectrum of U_x is purely continuous. For this reason, one might prefer the term 'weak mixing'; but that might create the wrong impression that Φ_ℓ is weak mixing with respect to some given invariant measure. The term 'dissipative' refers to the Hopf decomposition of Φ_ℓ on $S^*_x M$ into conservative and dissipative parts [**Kl**].

DEFINITION 10.12. Let (X, μ) be a measure space and let $T\colon L^1(X) \to L^1(X)$ be a positive contraction, i.e., $f \geq 0$ implies $Tf \geq 0$ and $\|f\| \leq \|f\|$. Let $u \in L^1(X)$ satisfy $u(x) > 0$ a.e. Then $X = C \cup D$ where

$$C = \left\{ x : \sum_{n=0}^{\infty} T^k u(x) = \infty \right\}, \quad D = \left\{ x \colon \sum_{n=0}^{\infty} T^k u(x) < \infty \right\}. \tag{10.48}$$

T is called completely dissipative if $C = \emptyset$.

In the case of a triaxial ellipsoid $E \subset \mathbb{R}^3$, the first return map Φ_x is a totally dissipative expanding map of the circle with two fixed points, one a source and one a sink. It has invariant δ-measures at the fixed points and an infinite locally L^1 invariant measure on each component of the complement.

In the case of C^∞ metrics, we can only prove existence of a partial blow-down point. In this case we define (as in [**Saf, SaV**]) the positive partially unitary operator (the Perron-Frobenius operator, compare (10.45))

$$U_x \colon L^2(\mathcal{L}_x, |d\omega|) \to L^2(S^*_x, |d\omega|), \quad U_x f(\xi) = \begin{cases} f((\Phi_x)(\xi))\sqrt{J_x(\xi)}, & \xi \in \mathcal{L}_x, \\ 0, & \xi \notin \mathcal{L}_x. \end{cases} \tag{10.49}$$

Here, as above, J_x is the Jacobian of the map Φ_x. We have

$$\ker U_x = \{ f \in L^2(S^*_x M) \colon \operatorname{supp} f \cap \Phi_x(\mathcal{L}_x) = \emptyset \}, \tag{10.50}$$

$$\operatorname{Image} U_x = \{ f \in L^2(S^*_x M) : \operatorname{supp} f \subset \mathcal{L}_x \}. \tag{10.51}$$

Note that U_x (10.45) is a positive contraction and $U_x(\lambda)$ (10.47) is a contraction which is not positive. The following Proposition exemplifies the difference between a conservative and dissipative first return map.

PROPOSITION 10.13. *Suppose that Φ_x is completely dissipative. Let $U_x(\lambda)$ be as defined in* (10.47) *and let $f \equiv \mathbf{1}$. Then*

$$F(\lambda, x) := \int_{\mathcal{L}_x} \sum_{k=1}^{\infty} U_x(\lambda)^k \mathbf{1} \, |d\xi|$$

is uniformly continuous in λ.

PROOF. By definition of 'dissipative' (Definition 10.12), the sum is absolutely convergent. We want to show that for all δ there exists ε such that $|F(\lambda+\varepsilon) - F(\lambda)| \leq \delta$ for any λ. The difference is

$$\sum_{k \in \mathbb{Z} \setminus 0} \int_{\mathcal{L}_x} \left(1 - e^{i\varepsilon \sum_{j=1}^k T_x(\Phi_x^j \xi)}\right) e^{i\lambda \sum_{j=1}^k T_x(\Phi_x^j \xi)} \overline{U_x^k \mathbf{1}} \, d\xi$$

and it is obviously bounded by

$$\sum_{k \in \mathbb{Z} \setminus 0} \int_{\mathcal{L}_x} \left|1 - e^{i\varepsilon \sum_{j=1}^k T_x(\Phi_x^j \xi)}\right| |\overline{U_x(\lambda)^k \mathbf{1}}| \, d\xi.$$

Since there is a positive minimal first return time T_0 we have $\sum_{j=1}^k T_x(\Phi_x^j \xi) \geq kT_0$. Since $|1 - e^{i\varepsilon k}| < \delta$ if $|k| \leq \frac{\delta}{\varepsilon}$ the latter sum is bounded for any $\varepsilon > 0$ by

(10.52)

$$\delta \sum_{\substack{0 \neq k \\ |k| \leq \frac{\delta}{\varepsilon}}} \left| \int_{\mathcal{L}_x} \overline{U_x(\lambda)^k \mathbf{1}} \right| d\xi + \sum_{|k| \geq \frac{\delta}{\varepsilon}} \left| \int_{\mathcal{L}_x} \overline{U_x(\lambda)^k \mathbf{1}} \right| d\xi \leq C\delta + \sum_{\substack{0 \neq k \\ |k| \geq \frac{\delta}{\varepsilon}}} \left| \int_{\mathcal{L}_x} \overline{U_x(\lambda)^k \mathbf{1}} \right| d\xi.$$

Given δ we now choose ε so that

$$\sum_{|k| \geq \frac{\delta}{\varepsilon}} \left| \int_{\mathcal{L}_x} \overline{U_x(\lambda)^k \mathbf{1}} \right| d\xi \leq C\delta,$$

which is clearly possible since the left side tends to zero as $\varepsilon \to 0$. □

10.4.4. Statement of results. Recall Definition 10.6 of maximal eigenfunction growth. In [**SoZ1**] the following is proved:

THEOREM 10.14. *Suppose (M, g) is a C^∞ Riemannian manifold with maximal eigenfunction sup-norm growth. Then there must exist a point $x \in M$ for which the loopset $\mathcal{L}_x$ at x has positive measure in $S_x^* M$ in $S^* M$.*

The theorem, as well as the results of [**SoZ1, SoTZ**], are proved by studying the remainder term $R(\lambda, x)$ in the pointwise Weyl law,

$$N(\lambda, x) = \sum_{j:\, \lambda_j \leq \lambda} |\varphi_j(x)|^2 = C_m \lambda^m + R(\lambda, x). \tag{10.53}$$

The first term $N_W(\lambda) = C_m \lambda^m$ is called the Weyl term. It is classical that the remainder is of one lower order, $R(\lambda, x) = O(\lambda^{m-1})$.

The relevance of the remainder to maximal eigenfunction growth is through the following well-known Lemma (see e.g. [**SoZ1**]):

LEMMA 10.15. *Fix $x \in M$. If $\lambda \in \operatorname{spec}(\sqrt{-\Delta})$, then*

$$\sup_{\varphi \in V_\lambda} \frac{|\varphi(x)|}{\|\varphi\|_2} = \sqrt{R(\lambda, x) - R(\lambda - 0, x)}. \tag{10.54}$$

Here, for a right continuous function $f(x)$ we denote by $f(x+0) - f(x-0)$ the jump of f at x.

Thus, Theorem 10.14 follows from:

THEOREM 10.16. *Let $R(\lambda, x)$ denote the remainder for the local Weyl law at x. Then*

$$R(\lambda, x) = o(\lambda^{n-1}) \text{ if } |\mathcal{L}_x| = 0. \tag{10.55}$$

Additionally, if $|\mathcal{L}_x| = 0$, then given $\varepsilon > 0$, there is a neighborhood $\mathcal{N}$ of x and a $\Lambda < \infty$, both depending on ε so that

$$|R(\lambda, y)| \leq \varepsilon \lambda^{n-1}, \quad y \in \mathcal{N}, \quad \lambda \geq \Lambda. \tag{10.56}$$

Theorem 10.14 is not sharp: on a tri-axial ellipsoid (three distinct axes), the umbilic points are self-focal points. But the eigenfunctions which maximize the sup-norm only have L^∞ norms of order $\lambda^{\frac{n-1}{2}} / \log \lambda$.[1] An improvement is given in [**SoTZ**], and more recently in [**SoZ3**], Sogge and the author have further improved the result in the case of real analytic (M, g). In this case $|\mathcal{L}_x| > 0$ implies that $\mathcal{L}_x = S^*_x M$ and the geometry simplifies.

THEOREM 10.17. *Suppose that (M, g) is a compact real analytic manifold without boundary with maximal eigenfunction sup-norm growth. Then (M, g) possesses a self-focal point p whose first return map Φ_p is conservative, i.e., has an invariant measure in the class of $|d\omega|$ on $S^*_x M$.*

In two dimensions, we can give a rather definitive result [**SoZ4**]:

THEOREM 10.18. *Let (M, g) be a compact real analytic compact surface without boundary with maximal eigenfunction sup-norm growth, then (M, g) possesses a pole, i.e., a point p so that every geodesic starting at p returns to p at time $2T_p$ as a smoothly closed geodesic.*

Thus, (M, g) is a $C^p_{2T_p}$-manifold in the terminology of [**Bes**, Definition 7.7(e)]. It follows by combining Theorem 10.17 with the following simple dynamical result.

LEMMA 10.19. *Let (S^2, g) be a two-dimensional real analytic Riemannian surface. Suppose that $p \in S^2$ is a self-focal point and that the first return map $\Phi_p \colon S^*_p S^2 \to S^*_p S^2$ preserves a probability measure which is in $L^1(S^*_p S^2)$. Then Φ_p^2 is the identity map, and in particular all geodesics through p are smoothly closed with the common period $2T_p$.*

10.4.5. Open problems.

- The only known examples of maximal eigenfunction growth in dimension 2 are surfaces of revolution. Although every point of a Zoll surface is a pole, it is doubtful that many have maximal eigenfunction growth. Must s real analytic Riemannian surface with maximal eigenfunction growth be a surface of revolution?
- In higher dimensions, does maximal eigenfunction growth imply existence of poles?

---

[1] It does not appear that a proof of this result has appeared in the literature.

- In the real analytic case, the Lagrangian submanifold Λ_x (10.42) exists. The existence proof does not relate it directly to the sequence of maximally growing eigenfunctions. Must they be 'quasi-modes' associated to Λ_x, i.e., do they microlocally concentrate on Λ_x? Do the microlocal defect measures of the sequence charge Λ_x?

By taking linear combinations of zonal spherical harmonics with different poles, it is easy to see that that without a more quantitative condition on the L^∞ growth one cannot conclude that the sequence has a unique microlocal limit or that it is concentrated on a single Λ_x. If there exists a self-focal point p with a smooth invariant function, then one can construct a *quasi-mode* of order zero concentrated on Λ_x.

PROPOSITION 10.20. *Suppose that (M, g) has a point p which is a self-focal point whose first return map Φ_x at the return time T is the identity map of S^*_pM. Then there exists a quasi-mode of order zero associated to the sequence $\{\frac{2}{\pi}Tk + \frac{\beta}{2} : k = 1, 2, 3, \dots\}$ that concentrates microlocally on Λ_x (10.42).*

Above, β is the common Morse index of the periodic orbits of periodic T. The 'symbol' is the smooth invariant density on the flow-out Λ_x. Theorem 10.18 is valid for quasi-modes as well as eigenfunctions. Indeed, most microlocal methods cannot distinguish modes and quasi-modes. It is not clear how the quasi-modes of Proposition 10.20 are related to the extremal sequence of eigenfunctions.

In this chapter we give a proof of a weaker version Theorem 10.17 which is more 'geometric' than the one in [**SoZ3**].

THEOREM 10.21. *Suppose that (M, g) is a compact real analytic manifold without boundary with maximal eigenfunction growth. Suppose further that the set $\mathcal{TL}$ of twisted self-focal points is finite. Then there must exist a conservative self-focal point, i.e., (M, g) possesses a self-focal point p whose first return map Φ_p is conservative, i.e., has an invariant measure in the class of $|d\omega|$ on S^*_xM.*

We prove this weaker version because there are no known examples of (M, g) in dimensions ≥ 3 which have twisted self-focal points and no known examples in any dimension where there are infinitely many self-focal points. Hence, the technical steps necessary to deal with such points may turn out to be vacuous. The assumption is only used to simplify some uniformity issues in remainder estimates.

10.5. Geometry of loops and return maps.

In this section we make some further remarks and conjectures on loops that are studied further in [**Z**].

Let $\mu_x = |d\omega|$ be the standard surface area on S^*_xM induced by the metric g_x. We summarize the following geometric notions.

DEFINITION 10.22. If (M, g) is C^∞, we say that $x \in M$

- is an $\mathcal{L}$ point ($x \in \mathcal{L}$) if $\mu_x(\mathcal{L}_x) > 0$. That is, there exists a positive measure of closed loops at x. In the real analytic setting, this implies $\mathcal{L}_x = \pi^{-1}(x) \simeq S^*_xM$. We call a point such that $\mathcal{L}_x = S^*_xM$ 'self-focal'.
- is a $\mathcal{CL}$ point ($x \in \mathcal{CL}$) if $\mu_x\{\xi \in \mathcal{L}_x : \Phi_x(\xi) = \xi\} > 0$. These are points where there exists a positive measure of smoothly closed loops. In the real analytic setting, it means that all geodesics emanating from x are

smoothly closed geodesic loops so that $\Phi_x = \mathrm{Id}$. Such a point is called a 'pole,' by analogy with a pole of a surface of revolution.

- is a *twisted self-focal point* if x is self-focal but $\Phi_x \neq \mathrm{Id}$ [2]. We write that x is a $\mathcal{TL}$ point ($x \in \mathcal{TL}$) if $\mu(\mathcal{L}_x) > 0$ but $\mu_x\{\xi \in \mathcal{L}_x : \Phi_x(\xi) = \xi\} = 0$, and denote the set of such points by $\mathcal{TL}$. Later we give a more quantitative 'twistedness' condition.
- x_0 is a $\mathcal{TL}_T$ point ($x \in \mathcal{TL}_T$) if $x \in \mathcal{TL}$ and if $T(x, \xi) \leq T$ for all $\xi \in \pi^{-1}(x)$. We denote the set of such points by $\mathcal{TL}_T$.

The notion of 'twisted' will be discussed further in §10.5.

CONJECTURE 10.23. If (M, g) is real analytic, and $\mathcal{CL} = \emptyset$, then $\mathcal{TL}_T$ is a finite set.

The motivation is that there do not seem to exist any examples with an infinite number of twisted self-focal points.

We think of this question as the possible 'rigidity' of self-focal points. On a sphere (or any Zoll surface), every point is a $\mathcal{CL}$ point, so non-twisted self-focal points can be 'moved around'. On the other hand, the umbilic points of an ellipsoid are isolated $\mathcal{TL}$ points. Must such points always be isolated, e.g. finite in number? At this time of writing, we cannot exclude the possibility, even in the real analytic setting, that there are an infinite number of $\mathcal{TL}$ points with twisted return maps. We let $\overline{\mathcal{L}}$ denote the closure of the set of self-focal points. At this time of writing, we do not know even how to exclude that $\overline{\mathcal{L}} = M$, i.e. that the set of self-focal points is dense.

This may also be seen in terms of the Lagrangian submanifolds Λ_x (10.42). If x_t is a smooth curve of self-focal points (in the real analytic case), then one has a curve of G^t-invariant Lagrangian submanifolds Λ_{x_t}. The idea is that there should exist an obstruction to deforming Λ_x through $\Lambda_{x(t)} \subset S^*M$ (that is, through G^t-invariant Lagrangian submanifolds) if Φ_x is twisted. There is evidently no obstruction when $\Phi_x = \mathrm{Id}$, so the main problem is to exploit the 'twisted condition' in the symplectic setting.

Our knowledge is so primitive that we do not even know whether there exist real analytic metrics with twisted self-focal points in higher dimensions.

PROBLEM 10.24. Do there exist (M, g) with $\dim M \geq 3$ possessing self-focal points with $\Phi_x \neq \mathrm{Id}$. That is, do there exist generalizations of umbilic points of ellipsoids in dimension two.

There do not seem to exist any known examples. Experts on geodesic flows on ellipsoids do not seem to have asked whether higher dimensional ellipsoids have self-focal points such as umbilic points in dimension 2. Recall that the geodesic flow of an ellipsoid E is integrable and S^*E is foliated by invariant Lagrangian tori (and cylinders at singular orbits). On the other hand, the flow-out Lagrangian Λ_x is $S^1 \times S^{n-1}$. In dimension 2, and only in dimension 2, Λ_x is a torus. In higher dimensions, existence of Λ_x might conflict with the Lagrangian torus fibration.

10.5.1. Reversibility and orientation. In this section, we assume p is a self-focal point. We note that Φ_p is the restriction of the geodesic flow G^T to

[2]The term is adopted from 'Dehn Twist' rather than 'twist map'; Φ_x does not resemble a twist map.

the invariant set S_p^*M. This sphere is contained in a symplectic transversal S_p to the geodesic flow. On the symplectic transversal G^T is a symplectic map which is invertible. Hence Φ_p is invertible. Thus, $(D\Phi_p)_\omega$ is non-zero for all $\omega \in S_p^*M$.

Next we use time reversal invariance of the geodesic flow to show that Φ_p is conjugate to its inverse. Let $\tau(x,\xi) = (x, -\xi)$ on S^*M.

LEMMA 10.25. $\Phi_p = \tau \Phi_p^{-1} \tau$, *i.e.*, Φ_p *is reversible (conjugate to its inverse).*

PROOF. On all of S^*M, we have $\tau G^t \tau = G^{-t}$. Indeed, let Ξ_H be the Hamilton vector field of $H(x,\xi) = |\xi|_g$. Then $H \circ \tau = H$, i.e., H is time reversal invariant. We claim that $\tau_* \Xi_H = -\Xi_H$. Written in Darboux coordinates,

$$\Xi_H = \sum_j \frac{\partial H}{\partial \xi_j} \frac{\partial}{\partial x_j} - \frac{\partial H}{\partial x_j} \frac{\partial}{\partial \xi_j}.$$

If we let $\tau(x,\xi) = (x, -\xi)$ and use invariance of the Hamiltonian we see that the vector field changes sign. Now, G^{-t} is the Hamilton flow of $-\Xi_H$ and that is $\tau_* \Xi_H$. But the Hamilton flow of the latter is $\tau G^t \tau$.

Since S_p^*M is invariant under τ, we just restrict the identity $\tau G^t \tau = G^{-t}$ to S_p^*M to see that Φ_p is reversible. □

10.5.2. Jacobi fields and the first return map. The first return map Φ_x may be expressed in terms of normal vertical Jacobi fields along a loop. The space $\mathcal{J}_\gamma^\perp$ of (real) orthogonal Jacobi fields along a geodesic $\gamma_{x,\xi}$ is the real symplectic vector space of dimension $2(m-1)$ of solutions of the Jacobi equation

(10.57)
$$\ddot{Y} + R(T,Y)T = 0, \quad g(Y(t), \dot{\gamma}_{x,\xi}(t)) = 0, \quad (T \text{ the unit tangent vector along } \gamma).$$

Here, $\dot{Y}$ and $\ddot{Y}$ are short for $\frac{DY}{dt}$ and $\frac{D^2Y}{dt}$. The symplectic structure is given by the Wronskian

$$\omega(X,Y) = g\left(X, \frac{D}{ds}Y\right) - g\left(\frac{D}{ds}X, Y\right).$$

We note that

$$0 = \frac{d}{dt} g(Y(t), \dot{\gamma}_{x,\xi}(t)) = g(\dot{Y}(t), \dot{\gamma}(t)),$$

so $\dot{Y}(t)$ is also a normal vector field along γ.

Jacobi fields arise from the derivative of the geodesic flow. Let ∇ denote the Riemannian connection, and recall that it determines a horizontal subbundle of $T(S^*M)$ complementary to the vertical subbundle of the projection $\pi : S^*M \to M$. Together with the symplectic structure, we get a splitting

$$T(S^*M) = \bar{H} \oplus \bar{V} \oplus \bar{T}$$

where $\bar{T}$ is the real span of $\dot{\gamma}$, and $\bar{H} \oplus \bar{V}$ is the horizontal plus vertical decomposition of the kernel of the contact form $\alpha = \xi \cdot dx$ (or equivalently, of the symplectic orthogonal of T and the cone axis). The subspaces $\bar{H}, \bar{V}$ are symplectically paired Lagrangian subspaces of $T(T^*M)$. Given a vector $X \in N_{\gamma(t)}$, we denote by X^h the horizontal lift of X to $\bar{H}_{\gamma(t)}$ and by X^v the vertical lift to $\bar{V}_{\gamma(t)}$. The correspondence

$$Y(t) \to (Y(t)^h, \dot{Y}(t)^v) \tag{10.58}$$

then defines an isomorphism between the spaces of Jacobi fields and geodesic flow invariant vector fields along $(\gamma(t), \dot{\gamma}(t)$ (cf [Kl, Lemma 3.1.6]). That is,

$$dG^s_{(\gamma(0),\dot{\gamma}(0))} \colon (Y(0)^h, \dot{Y}(0)^v) \to (Y(s)^h, \dot{Y}(s)^v)$$

where $Y(s)$ is the Jacobi field with the given initial conditions. Moreover, since G^t is a Hamiltonian flow dG^s is a linear symplectic mapping from $(\bar{H} \oplus \bar{V})_{\gamma(0),\dot{\gamma}(0)}$ to $(\bar{H} \oplus \bar{V})_{\gamma(s),\dot{\gamma}(s)}$.

Let U_i, V_i be a basis of the normal Jacobi fields along a loop with initial condition

$$\begin{pmatrix} U_i(0) = \nu_i & V_i(0) = 0 \\ \dot{V}_i(0) = 0 & \dot{V}_i(0) = \nu_i \end{pmatrix}, \tag{10.59}$$

where $\nu_1, \ldots, \nu_{m-1}$ is a parallel orthonormal basis for the normal bundle along the loop. The V_i are the vertical Jacobi fields. Along a geodesic loop the solution of the Jacobi equation with initial matrix (10.59) evolves by

$$M_{x,\xi}(t) = \begin{pmatrix} \langle \nu_i, U_j(t) \rangle & \langle \nu_i, V_j(t) \rangle \\ \langle \nu_i, \dot{U}_j(t) \rangle & \langle \nu_i, \dot{V}_j(t) \rangle \end{pmatrix} \tag{10.60}$$

Thus, $DG^t_{x,\xi}$ is a linear symplectic map on a symplectic transversal of rank $m-1$ and $M_{x,\xi}(t)$ is its matrix relative to a parallel normal frame and is therefore symplectic. It is expressed in block form,

$$\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad A^*C = C^*A, \quad B^*D = D^*B, \quad A^*D - C^*B = \mathrm{Id}\,. \tag{10.61}$$

10.5.3. $D_\xi \Phi_x$ as a block of $\mathcal{M}_{x,\xi}$. Now consider a self-focal point x and a loop $\gamma_{x,\xi}$ and $D_\xi \Phi_x \colon T_\xi S^*_x M \to T_{\Phi_x(\xi)} S^*_x M$.

First we observe that $D_\xi \Phi_x \colon T_\xi S^*_x M \to T_{\Phi_x(\xi)} S^*_x M$ is (roughly speaking) the D component of $D_\xi G^T_{x,\xi}$. More precisely, the D component is the 'block' of $D_\xi \Phi_x$ transversal to $\gamma'_{x,\xi}(t)$. When ξ is a conjugate direction, $D_\xi \Phi_x$ is a linear map on time derivatives of vertical Jacobi fields along $\gamma_{x,\xi}$, i.e., Jacobi fields obtained by varying the initial direction of geodesics at x.

We denote by $C_x \subset T^*_x M$ the set of tangential conjugate points ξ where $D_\xi \exp_x$ is singular. By SC_x we also denote the unit vectors $\frac{\xi}{|\xi|}$ for $\xi \in C_x$, and refer to these unit vectors as 'conjugate directions.'

LEMMA 10.26. *Let x be a self-focal point length $T = T(x)$. Then:*

- *For all $\xi \in S^*_x M$, $\eta = T(x)\xi \in SC_x$ and the matrix of $D_\eta \exp_x$ restricted to tangent directions in SC_x is the vertical-horizontal block $\langle \nu_i, V_j(T) \rangle = 0$. Thus, $\ker D_\eta \exp_x = T_\eta SC_x$.*
- *The matrix of $D_\xi \Phi_x$ is the vertical-vertical D block of (10.60) for $(x, \xi) \in S^*M$. More precisely, $D_\xi \Phi_x(\dot{V}(0)) = \dot{V}(T)$, or equivalently,*

$$D_\xi \Phi_x = \left(\langle \nu_i, \dot{V}^h_j(T) \rangle \right)_{i,j=1}^{m-1} : T_\xi S^*_x M \to T_{\Phi_x(\xi)} S^*_x M$$

.

PROOF. If $Y(0) = 0, \nabla Y(0) = V$ then $D_\xi \Phi_x = \nabla Y(T(x))$, i.e., the value of the covariant derivative of the Jacobi along $\gamma_{x,\xi}$ at the first return time.

Let $\xi(s)$ be a curve in $S^*_x M$ with $\xi(0) = \xi$ and $\frac{d}{ds}\xi(s) = V$. Then $D_\xi \Phi_x(V) = \frac{d}{ds}|_{s=0} \Phi_x(\xi(s))$. Since $S^*_x M = \mathcal{L}_x$, the one-parameter family

$$F(s,t) := \exp_x t\xi(s) \colon [0,1] \times [0, T(x)] \to M$$

is a variation of geodesic loops through x. By definition,

$$\Phi_x(\xi(s)) = \frac{d}{dt}\bigg|_{t=T(x)} \exp_x t\xi(s).$$

Hence,

$$D_\xi D\Phi_x(V) = \frac{d}{ds}\Big|_{s=0} \frac{d}{dt}\Big|_{t=T(x)} \exp_x t\xi(s).$$

Note that the $\frac{d}{ds}$-derivative is the usual Euclidean derivative of a curve in T_xM. We note that

$$\frac{d}{ds}\Big|_{s=0} \exp_x t\xi(s) = Y_V(t)$$

is the vertical Jacobi field along $\gamma_{x,\xi}$ with $Y_V(0) = 0, \dot{Y}_V(0) = V$. The last statement follows because we may commute the derivatives (see e.g. [**M**, Lemma 8.7]). For the same reason,

$$D_\xi D\Phi_x(V) = \frac{D}{dt}\Big|_{t=T(x)} Y_V(t) = \dot{Y}_V(T(x)).$$

□

LEMMA 10.27. *When (M, g) is real analytic and when x is a self-focal point (an $\mathcal{L}$ point in Definition 10.22) of length T then the vertical Jacobi fields along each loop direction ξ satisfy $V_j(T) = 0$ and*

$$M_{x,\xi}(T) = \begin{pmatrix} \langle \nu_i, U_j(T)\rangle, & 0 \\ \langle \nu_i, \dot{U}_j(T)\rangle, & \langle \nu_i, \dot{V}_j(T)\rangle \end{pmatrix} \tag{10.62}$$

is a lower triangular symplectic matrix. Φ_x is orientation preserving if and only if $\det D > 0$ *if and only if* $\det A > 0$.

PROOF. Since $B = 0$ the last equation above forces $A^*D = \mathrm{Id}$. Φ_x is orientation preserving if and only if $\Phi_x^* \, dV_{S_x^*M} = J_x \, dV_{S_x^*M}$ with $J_x > 0$ as in (10.45) and $J_x(\xi) = \det D$. Clearly $\det D > 0$ if and only if $\det A > 0$. □

10.5.4. Periodic geodesics and fixed points of Φ_x. When $\gamma_{x,\xi}$ is a periodic orbit of period T, then $dG^T_{x,\xi}$ is the linear Poincare map P_γ of γ. For a periodic geodesic, P_γ acts as the translation map by a period

$$P_\gamma \colon Y(t) \to Y(t + T_\gamma)$$

on Jacobi fields, where T_γ is the period. It decomposes into the two-dimensional symplectic plane of tangential Jacobi fields and the $2m-2$ dimensional space of normal Jacobi fields. A Lagrangian subspace of the latter is the $(m-1)$ dimensional subspace of vertical normal Jacobi fields. P_γ may be expressed as the matrix (10.60) acting on the tangent space $T_{x,\xi} \subset T_{x,\xi}S^*M$ to a surface of section $S \subset S^*M$, i.e., a symplectic transversal.

Fixed points of Φ_x correspond to closed geodesics through x and are evidently important in the study of loops. We record a number of facts relevant to $\mathrm{Fix}(\Phi_x)$. If $\xi \in \mathcal{CL}_x$, so that $\gamma_{x,\xi}(t) = \exp_x t\xi$ is a closed geodesic through x, then $\Phi_x\xi = \xi$. Conversely if $\Phi_x\xi = \xi$ then $\gamma_{x,\xi}(t)$ is a closed geodesic.

We recall that a diffeomorphism F is said to have *clean* fixed point sets if $\mathrm{Fix}(F)$ is a manifold and if $T\,\mathrm{Fix}(F) = \mathrm{Fix}(DF)$. Cleanliness of Φ_x thus means that $\mathrm{Fix}(\Phi_x) \subset S^*_xM = \mathcal{CL}_x$ is a manifold and $T_\xi\mathcal{CL}_x = \ker(D_\xi\Phi_x - I)$.

The Lefschetz fixed point theorem implies that a diffeomorphism $f \colon S^n \to S^n$ has a fixed point as long as $\deg f \neq (-1)^{n+1}$. Thus, an orientation preserving diffeomorphism has a fixed point when n is even, and an orientation reversing diffeomorphism has a fixed point when n is odd. Hence, either Φ_p or Φ_p^2 has a fixed point when $\dim S_p^*M$ is even.

In Lemma 10.61 we show that when $\dim M = 2$, Φ_p^2 always has a fixed point. In the case of the umbilic points p of a 2-dimensional tri-axial ellipsoid, $\mathrm{Fix}(\Phi_p)$ consists of 2 directions.

10.5.5. The tangential loopset. We have been focusing on directions $\xi \in \mathcal{L}_x \subset S_x^*M$ of loops. In this section, we consider the magnitudes as well as directions, that is $(x, T(\xi)\xi) \in T_x^*M$ where $T(\xi)$ is the first return time. In the analytic case, the two sets are closely related since, if all $\xi \in \pi^{-1}(x)$ are loop directions, then the time $T(x,\xi)$ of first return is constant on $\pi^{-1}(x)$. This is because an analytic function is constant on its critical point set.

One has $\Phi_x(\xi) = \xi$ with $\xi \in S_x^*M$ if and only if $\exp_x T(\xi)\xi = x$. Define the function

$$E(x,\xi)\colon TM \to M, \quad E(x,\xi) = \exp_x \xi - x = \pi G^1(x,\xi) - \pi(x,\xi). \tag{10.63}$$

Here we are using local coordinates $U \simeq \mathbb{R}^n$ on M so that subtraction makes sense. One could define $E(x,\xi)$ invariantly but it is simpler to use local coordinate subtraction.

When g is real analytic, E is a real analytic function. The set of initial data of loops is the analytic set

$$E^{-1}(0) := \mathcal{E} = \{(x,\xi) \in T^*M\colon \exp_x \xi = x\},$$

where as usual we identify vectors and co-vectors with the metric. Under the natural projection $\pi\colon T^*M \to M$, $\pi(\mathcal{E})$ is the set of points through which there exists a loop. If we fix x we obtain

$$E_x\colon T_xM \to M$$

and its zero set $\mathcal{E}_x$ is the set of loop vectors through x. The set $\mathrm{Fix}(\Phi_x) = \mathcal{L}_x$ of loop directions at x is the spherical projection $\xi \to \frac{\xi}{|\xi|}$ of $\mathcal{E}_x$. Moreover, if $E(x,\xi) = 0$, then (using the homogeneous geodesic flow rather than the one defining $\exp_x \xi$),

$$G^{|\xi|}\left(x, \frac{\xi}{|\xi|}\right) = \left(x, \Phi_x\left(\frac{\xi}{|\xi|}\right)\right).$$

At a self-focal point p, $\mathcal{E}_p = \mathcal{E} \cap T_pM$ contains a union of spheres of radii $kT(p)$, $k = 1, 2, 3, \ldots$.

If p is self-focal and $\Phi_p = \mathrm{Id}$ on S_p^*M, then the entire flowout Lagrangian Λ_p (10.42) lies in $\mathcal{E}$. This is because $g^t(p,\xi)$ is also a loop since $G^TG^t(p,\xi) = G^tG^T(p,\xi) = G^t\Phi_p(p,\xi) = G^t(p,\xi)$. The same calculation shows that the flowout of $\mathrm{Fix}(\Phi_p)$ lies in $\mathcal{E}_p$. In general, Λ_p is not contained in $\mathcal{E}$, since the geodesic through $G^t(p,\xi)$ generally does not loop back to $\exp_p t\xi$.

10.5.5.1. *Examples.* As mentioned above, by Theorem 10.9 it follows that for any point p of any compact Riemannian manifold, there exists an infinite set of distinct geodesic loops through p. It follows that

Corollary 10.28. *$\dim \mathcal{E} \geq m = \dim M$.*

The dimension of $\mathcal{E}$ is closely related to the rank of $D_{x,\xi}E$ at points $(x,\xi) \in \mathcal{E}$. Recall that if $f\colon M \to N$ is a smooth map, then the rank of the derivative D_pf is lower semi-continuous: If $\mathrm{rank}(D_pf) = r$, then there exists a neighborhood of p so that $\mathrm{rank}(D_qf) \geq r$ for $q \in U$. Hence if $\mathrm{rank}(D_{x,\xi}E) \geq m$ (the maximal rank) at a point $(x,\xi) \in \mathcal{E}$ then there is a neighborhood of (x,ξ) where the rank is maximal and $\mathcal{E}$ is locally a manifold of dimension m through (x,ξ).

Some examples: The minimum dimension is achieved on a negatively curved manifold, for instance. At the opposite extreme, in the case of the standard S^m or a Zoll manifold, at each $p \in M$ the set $E(p,\xi) = 0$ is a union of spheres $2\pi k S^*_x M$ of maximal dimension $m-1$, so that $\mathcal{E} = E^{-1}(0)$ is of dimension $2m-1$. This reflects the degeneracy of DE, which has rank one at every loop (x,ξ). For a generic surface of revolution, if p is the north or south pole, $E_p^{-1}(0)$ is again a union of circles in $T^*_p M$ of radii kT_x where T_x is the common primitive period. Moreover $\mathrm{Fix}(\Phi_p) = S^*_p M$ so $\Lambda_p \subset \mathcal{E}$. Away from the poles, $E_x^{-1}(0)$ is a countable set which at least contains the meridian covector along the geodesic through x from the poles. Hence $\dim \mathcal{E} = m$.

In the case of a tri-axial ellipsoid $E \subset \mathbb{R}^3$, the first return map at the umbilic points p is fully twisted, and $\mathrm{Fix}(\Phi_p)$ consists of two collinear vectors $(p,\xi),(p,-\xi)$ lying along the closed geodesic through p. Hence, $\mathcal{E}$ contains $S^*_p E$ and also the covectors along these two closed geodesics.

10.6. Proof of Theorem 10.21. Step 1: Safarov's pre-trace formula

The purpose of the rest of this section is to prove a slightly weaker version of Theorem 10.17 in which we make the additional assumption that (M,g) has only finitely many twisted self-focal points. By Proposition 10.23, there are only finitely many with return times less than a given T.

THEOREM 10.29. *Suppose that (M,g) is a compact real analytic manifold without boundary which possesses a sequence $\{\varphi_{j_k}\}$ of eigenfunctions satisfying*

$$\|\varphi_{j_k}\|_{L^\infty(M)} \geq C_g \lambda_{j_k}^{\frac{n-1}{2}}. \tag{10.64}$$

Suppose that the number of twisted self-focal points of (M,g) is finite. Then (M,g) possesses a self-focal point p whose first return map Φ_p is conservative, i.e., has an invariant measure in the class of $|d\omega|$ on $S^*_x M$.

The proof is somewhat different from the one in [**SoZ2**]. The motivation to prove the weaker result is that it involves relations between dynamics and spectral theory beyond those of [**SoZ2**].

The first step in the proof is a pre-trace formula for a smoothing of the cosine wave kernel $E(t,x,x)$ on the diagonal. We recall some notation: $E(t,x,y)$ denotes the kernel of $\cos t\sqrt{-\Delta}$. Let $\rho \in \mathcal{S}(\mathbb{R})$ with $\hat{\rho} \in C_0^\infty(\mathbb{R})$, and denote its dilation by $\hat{\rho_T}(\tau) = \hat{\rho}(\frac{\tau}{T})$. Then,

$$\int_{\mathbb{R}} \hat{\rho}(t) e^{it\lambda} E(t,x,x) = \sum_j (\rho(\lambda - \lambda_j) + \rho(\lambda + \lambda_j)) \varphi_j^2(x).$$

The second term is negligible since $\rho \in \mathcal{S}(\mathbb{R})$ and $\lambda + \lambda_j \to \infty$. We denote the pointwise Weyl function by

$$N(\lambda, x) = \sum_{j:\ \lambda_j \leq \lambda} \varphi_j^2(x),$$

so that

$$\rho * dN(\lambda, x) = \sum_j \rho(\lambda - \lambda_j) \varphi_j^2(x). \tag{10.65}$$

The Safarov pre-trace formula expresses (10.65) in terms of iterates of the return map Φ_x and the Perron-Frobenius operator U_x (10.45) on the loopset $\mathcal{L}_x$ (10.39); see [**SaV**, Theorem 4.4.10] or [**Saf**, (4.4)]. As in (10.47), let

$$U_x^{\pm}(\lambda) = e^{i\lambda T_x^{\pm}} U_x^{\pm}. \tag{10.66}$$

Recall that $T_x^{(k)}(\xi)$ is the kth return time of ξ for Φ_x.

PROPOSITION 10.30. *Let $x \in M$. Then*

$$\rho_T * dN(\lambda, x) = \lambda^{n-1} + \lambda^{n-1} \int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\left(\frac{T_x^{(k)}(\xi)}{T}\right) U_x(\lambda)^k \mathbf{1} |d\xi| + \mathcal{R}_{T,x}(\lambda), \tag{10.67}$$

where

$$\mathcal{R}_{T,x}(\lambda) = o_{T,x}(\lambda^{n-1}). \tag{10.68}$$

The first term is due to the singularity at $t = 0$. Once Proposition 10.30 is proved, the remainder of the proof of Theorem 10.17 consists in studying the two remainder terms. In Proposition 10.42, the L^2 ergodic theorem is applied to the term

$$\int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}(\frac{T_x^{(k)}(\xi)}{T}) U_x(\lambda)^k \mathbf{1} |d\xi| \tag{10.69}$$

for reach x to show that the average tends to zero. But additionally one needs uniformity of the convergence to zero in x to prove the Theorem. We also prove uniformity of $o_{T,x}(\lambda^{n-1})$ in x. The proof of uniformity is done separately for x near to a self-focal point and for x far from the self-focal locus. Of course, if there are only a finite number of self-focal points, proofs of uniformity simplify a great deal.

REMARK 10.31. The principal term λ^{n-1} is due to the singularity of $E(t, x, x)$ at $t = 0$. The other terms are due to singularities caused by loops at x for $t \neq 0$, and vanish if x is $|\mathcal{L}_x| = 0$. Duistermaat-Guillemin [**DG**] proved the existence of an asymptotic expansion for $\rho_T * dN(\lambda, x)$ and showed that the leading coefficient equals 1. Safarov calculated the second term of order λ^{n-1} in terms of Φ_x. It vanishes if x is not a partial self-focal point. The uniformity in x is not obvious. Uniformity is discussed in §10.8.3.

We break up the proof of Proposition 10.30 into two steps. The first step only involves the construction of the wave kernel and the application of stationary phase to reduce to an integral over $S_x^* M$. The second step is to reduce the latter integral to one over $\mathcal{L}_x$. Even in simple examples, the second step cannot be done by stationary phase.

To state the result of the first step we introduce some notation from [**SaV**]. Let $\varphi_j(t, x, y, \xi)$ be a local generating function for the graph C of the geodesic flow, in the sense that

$$C = \{(t, \varphi_t, x, d_x\varphi, y, -d_y\varphi \colon d_\xi \varphi = 0\}.$$

Also let

$$\tilde{t}_j(x, \xi) = T_x(\xi), \quad \text{and} \quad r_* = r_*(x, \xi) := -(\varphi_t(T_x^{(k_j)}(\xi)), x, \xi)^{-1}. \tag{10.70}$$

DEFINITION 10.32. Define

$$R_j(x, \lambda) := \int_{S_x^* M} e^{i\lambda \tilde{t}_j(x,\xi)} r_* \left((\hat{\rho} a_0)|_{\tilde{t}_j = T_x(\xi)} \right) |d\xi|. \tag{10.71}$$

When ρ is fixed and we dilate $\rho \to \rho_T$ then we denote the remainder by

$$R_j(\lambda, x; T) := \int_{S^*_x M} e^{i\lambda \tilde{t}_j(x,\xi)} r_* \left((\hat{\rho}_T a_0)|_{\tilde{t}_j = T_x(\xi)} \right) |d\xi|.$$

We note that $R_j(\lambda, x, T)$ depends on λ only through the oscillatory factor $e^{i\lambda \tilde{t}_j(x,\xi)}$.

LEMMA 10.33.

$$\rho_T * dN(\lambda, x) = \lambda^{n-1} + \lambda^{n-1} \sum_{j=1}^{\infty} R_j(\lambda, x, T) + O_T(\lambda^{n-2}), \tag{10.72}$$

where $O_T(\lambda^{n-2})$ is uniform in x.

REMARK 10.34. Referring to (10.67), this Lemma states that

$$\lambda^{n-1} \int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\left(\frac{T_x^{(k)}(\xi)}{T} \right) U_x(\lambda)^k 1 |d\xi| + \mathcal{R}_{T,x}(\lambda) = \lambda^{n-1} \sum_{j=1}^{\infty} R_j(\lambda, x, T) + O_T(\lambda^{n-2}). \tag{10.73}$$

PROOF OF LEMMA 10.33. For t in the support of $\hat{\rho}$ and near a point x we construct a microlocal parametrix for $E(t, x, x) := \cos t\sqrt{-\Delta}$ as a sum of a finite number of local parametrices $E_j(t, x, y)$ (depending on T), each of which has an amplitude a_j and a phase function $\varphi_j(t, x, y, \xi)$ locally parametrizing C as above. One expresses each $E_j(t, x, x)$ as an integral over $T^*_x M$, changes to polar coordinates $(r, \omega) \in \mathbb{R}_+ \times S^*_x M$ and performs the *dtdr* integral by stationary phase. This gives a smoothed spectral expansion [**DG**] and [**SaV**, Theorem 4.1.2]:

$$\begin{aligned} \rho_T * dN(\lambda, x) &= \int_{\mathbb{R}} \hat{\rho}\left(\frac{t}{T} \right) e^{i\lambda t} E(t, x, x)\, dt \\ &= a_0 \lambda^{n-1} + a_1 \lambda^{n-2} + \lambda^{n-1} \sum_{j=1}^{\infty} R_j(\lambda, x, T) + O_T(\lambda^{n-2}), \end{aligned} \tag{10.74}$$

with uniform remainder in x. The sum over j is a sum over charts needed to parametrize the graph of the geodesic flow. The critical times $t \le T$ are the lengths of loops at x. Let

$$\begin{cases} \tilde{t}_j(x, \xi) = T_x(\xi) = \varphi_j(t, x, x, \xi) & \text{at the stationary phase point,} \\ r_*(x, \xi) := -(\varphi_t(T_x^{(k_j)}(\xi)), x, \xi)^{-1}. \end{cases} \tag{10.75}$$

That is, $\tilde{t}_j(x, \xi)$ is the local t-solution of $\varphi_j(t, x, x, \xi) = 0$. Summing over the charts indexed by j gives a sum of stationary phase expansions,

(10.76)

$$\rho_T * dN(\lambda, x) \sim \lambda^{n-1} \sum_j \int_{S^*_x M} e^{i\lambda \tilde{t}_j(x,\xi)} r_* \left(\sum_{m,k=0}^{\infty} \lambda^{-k-m} \mathcal{L}_{m,k}(\hat{\rho} a_{-k})|_{\tilde{t}_j = T_x(\xi)} \right) |d\xi|,$$

for certain operators $\mathcal{L}_{m,k}$.

It is obvious that except for the term with $m = k = 0$ all other terms are uniformly of order λ^{n-2}. Since we can drop all terms except those for which $m = k = 0$, we reduce to

(10.77)
$$\rho_T * dN(\lambda, x) \sim \lambda^{n-1} \sum_j \int_{S^*_x M} e^{i\lambda \tilde{t}_j(x,\xi)} r_* \Big((\hat{\rho} a_0)|_{\tilde{t}_j = T_x(\xi)} \Big) |d\xi| + O_T(\lambda^{n-2}),$$

where the remainder is uniform in x. □

We refer to [**DG, SaV, Hor2, V**] for further details and background. To illustrate the notation, we consider a flat torus $\mathbb{R}^n/\Gamma$ with $\Gamma \subset \mathbb{R}^n$ a full rank lattice. As is well-known, the wave kernel then has the form

$$U(t,x,y) = \sum_{\gamma \in \Gamma} \int_{\mathbb{R}^n} e^{i\langle x-y-\gamma, \xi\rangle} e^{it|\xi|}\, d\xi.$$

Thus, the indices j may be taken to be the lattice points $\gamma \in \Gamma$, and

$$\rho_T * dN(\lambda, x) = \sum_{\gamma \in \Gamma} \int_{\mathbb{R}} \int_0^\infty \int_{S^{n-1}} \hat{\rho}\Big(\frac{t}{T}\Big) e^{ir\langle \gamma, \omega\rangle} e^{itr} e^{-it\lambda} r^{n-1}\, dr dt d\omega.$$

We change variables $r \to \lambda r$ to get a full phase $\lambda(r\langle \gamma, \omega\rangle + tr - t)$. The stationary phase points in (r,t) are $\langle \gamma, \omega \rangle = t$ and $r = 1$. Thus,

$$\tilde{t}_\gamma(x, \omega) = \langle \gamma, \omega \rangle.$$

The geometric interpretation of $t^*_\gamma(x,\omega)$ is that it is the value of t for which the geodesic $\exp_x t\omega = x + t\omega$ comes closest to the representative $x + \gamma$ of x in the γth chart. Indeed, the line $x + t\omega$ is 'closest' to $x + \gamma$ when $t\omega$ closest to γ, since $|\gamma - t\omega|^2 = |\gamma|^2 - 2t\langle \gamma, \omega\rangle + t^2$. On a general (M, g) without conjugate points,

$$\tilde{t}_\gamma(x,\omega) = \langle \exp_x^{-1} \gamma x, \omega \rangle.$$

10.6.1. Decomposition of R_j into almost loops and non-loops. In this section, we study the remainders R_j (10.71) and relate them to the two terms of (10.73).

To relate R_j to left side of (10.71), we use a kind of Lemma of non-stationary phase to reduce the integral over $S^*_x M$ in (10.76) to an integral over $\mathcal{L}_x$. It is based on the fact that the phase of (10.76) is rapidly oscillating away from its critical set, i.e., directions $\xi \in \mathcal{L}_x$ such that $\nabla_\xi \tilde{t}_j = 0$. It is impossible to apply stationary phase to the integral over $S^*_x M$ without some 'cleanliness assumption' on the phase (i.e., without assuming that $\tilde{t}_j$ is a Bott-Morse function). But one can apply a weak version of the Lemma of stationary phase, which shows that the oscillatory integral is decaying in λ away from its stationary phase set.

As in [**SaV**], we pick a non-negative $f \in C^\infty_0(\mathbb{R})$ which equals 1 on $|s| \leq 1$ and zero for $|s| \geq 2$ and split up the jth term of (10.77) into two terms using R_{j1} resp. R_{j2} by using $f(\varepsilon^{-2} |\nabla_\xi \tilde{t}_j|^2)$ and $1 - f(\varepsilon^{-2} |\nabla_\xi \tilde{t}_j|^2)$. We also multiply by $\lambda^{-(n-1)}$ for notational simplicity. Thus, referring to (10.71),

$$R_j(\lambda, x, T) = R_{j1}(\lambda, x, T) + R_{j2}(\lambda, x, T), \tag{10.78}$$

where
(10.79)
$$R_{j1}(\lambda, x; \varepsilon) := \int_{S^*_x M} e^{i\lambda \tilde{t}_j} f(\varepsilon^{-1} |\nabla_\xi \tilde{t}_j(x,\xi)|^2) r_* (\hat{\rho}(T_x(\xi))) a_0(T_x(\xi), x, \xi) r^n_* d\xi|_{\tilde{t} = T^{(k)}}.$$

The second term R_{j2} comes from the $1-f(\varepsilon^{-2}|\nabla_\xi T_x(\xi)|^2)$ term. Since it is simpler, we consider it first. We emphasize that the remainder (10.79) is the coefficient of λ^{n-1}.

10.6.2. R_{j2} is uniformly $o_T(\lambda^{n-1})$ in x. The next Lemma shows that R_{j2} can be absorbed into the the $\mathcal{R}_{T,x}(\lambda) = o_{T,x}(\lambda^{n-1})$ term of (10.67), and in fact the estimate is uniform in x. For all $T > 0$ and $\varepsilon \geq \lambda^{-\frac{1}{2}} \log \lambda$ we have

$$\sup_{x \in M} |R_{j2}(\lambda, x, T, \varepsilon)| \leq C(\varepsilon^2 \lambda)^{-1}.$$

LEMMA 10.35. *For $\varepsilon \geq \lambda^{-\frac{1}{2}} \log \lambda$ we have*

$$\sup_{x \in M} |R_{j2}(\lambda, x, \varepsilon)| \leq C(\log \lambda)^{-1}.$$

PROOF OF LEMMA 10.35. We integrate the R_{j2} term by parts once with

$$L_x := \frac{1}{i\lambda} |\nabla_\xi \tilde{t}_j|^{-2} \nabla_\xi \tilde{t}_j \cdot \nabla_\xi.$$

This operator reproduces $e^{i\lambda \tilde{t}_j}$ and after one partial integration we have

$$R_{j2}(\lambda, x, \varepsilon) = \frac{1}{i\lambda} \int_{S_x^* M} e^{i\lambda \tilde{t}_j} L_x^t (1 - f(\varepsilon^{-2} |\nabla_\xi \tilde{t}_j|^2) r_*(\hat{\rho}(\tilde{t}_j) a_0(\tilde{t}_j), x\xi) r_*^n d\xi \tag{10.80}$$

$$= O\left(\frac{1}{\lambda \varepsilon^2}\right). \tag{10.81}$$

In the last line we use that

$$L_x^t f = \frac{1}{|\nabla_\xi \tilde{t}_j|^2} \nabla_\xi \tilde{t}_j \cdot \big(f'(\varepsilon^{-2} |\nabla_\xi \tilde{t}_j|^2) \varepsilon^{-2} \nabla_\xi \cdot |\nabla_\xi \tilde{t}_j|^2\big) + \operatorname{div}\left(\frac{\nabla_\xi \tilde{t}_j}{|\nabla_\xi \tilde{t}_j|^2}\right)(1-f) = O(\varepsilon^{-2}).$$

Indeed, the final expression is bounded by $|\nabla_\xi \tilde{t}_j(\xi)|^{-2}$ and on the support of the integrand this is bounded by ε^{-2}. These estimates are uniform in x and again can be summed over the the finite number of charts indexed by j. Thus, as long as $\varepsilon \geq \lambda^{-1/2} \log \lambda$, we have $\lambda^{-1} \varepsilon^{-2} \leq C(\log \lambda)^{-1}$. □

We therefore see that the sum of the R_{j2} terms of the $o_x(\lambda^{n-1})$ in Proposition 10.30 is therefore uniform in x.

COROLLARY 10.36.

$$\lambda^{n-1} \sum_{j=1}^{\infty} R_{j2}(\lambda, x, T) = o_T(\lambda^{n-1}).$$

10.6.3. Decomposition of $R_{j1}(\lambda, x, \varepsilon)$. We now decompose the R_{j1} terms (10.79) into terms corresponding to the $\lambda^{n-1} \int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}(\frac{T_x^{(k)}(\xi)}{T}) U_x(\lambda)^k 1 |d\xi|$ and $\mathcal{R}_{T,x}(\lambda)$ terms in (10.73).

We observe the critical set consists of loop directions at x, i.e., $\{\nabla_\xi \tilde{t}_j(x,\xi) = 0\} = \mathcal{L}_x$. We then decompose R_{j1} into an integral over the stationary phase set $\mathcal{L}_x$ and its complement:

$$R_{j1}(x, \varepsilon, T) = \int_{\mathcal{L}_x} e^{i\lambda T_x(\xi)} \hat{\rho}(T_x(\xi)) a_0(T_x(\xi), x, \xi) r_*^n |d\xi| |_{\tilde{t} = T^{(k)}} + \tilde{R}_{j1}, \tag{10.82}$$

where
(10.83)

$$\tilde{R}_{j1}(x,\varepsilon,T)=\int_{|\nabla\tilde{t}_j|>0} e^{i\lambda T_x(\xi)}f(\varepsilon^{-2}|\nabla_\xi\tilde{t}_j(x,\xi)|^2)r_*(\hat{\rho}(T_x(\xi)))a_0(T_x(\xi),x,\xi)r_*^n d\xi.$$

LEMMA 10.37. *$R_{j1}(x,\varepsilon,T)$ is independent of ε and when $\rho=\rho_T$,*

$$\sum_j R_{j1}(x,T)=\int_{\mathcal{L}_x}\sum_{k=1}^{\infty}\hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big)U_x(\lambda)^k\mathbf{1}|d\xi|.$$

PROOF. It is only necessary to compare the formulae (10.49)-(10.66) for $U_x(\lambda)$ with $a_0(T_x(\xi),x,\xi)r_*^n d\xi$. The calculations are done in [**SaV**, Proposition 4.1.10] and [**SaV**, Proposition 4.1.16]. On $\mathcal{L}_x$, $r_*=1$ and the surface measure is the Jacobian in (10.49). Also, since the amplitudes solve transport equations along the Hamilton orbits (see e.g. [**DG, SaV, So3**]) at time $T_x(\xi)$ they are given by iterates of the first return map. We refer to [**SaV**] for further details. □

Since the remainders only involve λ via $e^{i\lambda T}$, it is convenient to introduce notations for upper bounds in which we take absolute values. For future reference, we define λ-independent functions as follows.

DEFINITION 10.38. Set
(10.84)

$$|R|_{j1}(x;\varepsilon)=\int_{S_x^*M} f(\varepsilon^{-2}|\nabla_\xi\tilde{t}_j(x,\xi)|^2)r_*(\hat{\rho}(T_x(\xi)))a_0(T_x(\xi),x,\xi)r_*^n d\xi|_{\tilde{t}=T^{(k)}}$$

(10.85)

$$|R_j|(x,T)=\int_{S_x^*M} r_*\left((\hat{\rho}_T a_0)|_{\tilde{t}_j=T_x(\xi)}\right)|d\xi|$$

(10.86)

$$|\tilde{R}_{j1}|(x,\varepsilon,T)=\int_{|\nabla\tilde{t}_j|>0} f(\varepsilon^{-2}|\nabla_\xi\tilde{t}_j(x,\xi)|^2)r_*(\hat{\rho}(T_x(\xi)))a_0(T_x(\xi),x,\xi)r_*^n d\xi.$$

We note that $|R_j(\lambda,x,T)|\le|R_j|(x,T)$.

10.6.4. Perturbation theory of the remainder. We now compare the absolute remainders at nearby points. The integrands of the remainders vary smoothly with the base point and only involve integrations over different fibers S_x^*M of $S^*M\to M$.

LEMMA 10.39. *We have,*

$$||R|(x,T)-|R|(y,T)|\le Ce^{aT}\operatorname{dist}(x,y).$$

Indeed, we write the difference as the integral of its derivative. The derivative involves the change in Φ_x^n as x varies over iterates up to time T and therefore is estimated by the sup norm e^{aT} of the first derivative of the geodesic flow up to time T. If we choose a ball of radius δe^{-aT} around a focal point, we obtain'

COROLLARY 10.40. *For any $\eta>0$, $T>0$ and any focal point $p\in\mathcal{TL}$ there exists $r(p,\eta)$ so that*

$$\sup_{y\in B(p,r(p,\eta))}|R(\lambda,y,T)|\le\eta.$$

It will be important later on to observe that the absolute remainder varies continuously in x. We choose local coordinates so that $S^*_x M$ is identified with S^{n-1}.

PROPOSITION 10.41. *Then for all η there exists $\delta > 0$ so that for $y \in B_\delta(x)$.*

$$| \, |R|_{j1}(x, \varepsilon, T) - |R_{j1}|(y, \varepsilon, T)| \leq C_T \eta.$$

PROOF. We are just asserting the continuity of $|R_{j1}|(y, \delta, T)$ on $B_\delta(x)$. Indeed, the integrand is uniformly bounded and continuous in all variables. □

10.7. Proof of Theorem 10.29. Step 2: Estimates of remainders at $\mathcal{L}$-points

In view of Lemma 10.37, the next step is to study

$$\int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big) U_x(\lambda)^k \mathbf{1} |d\xi|$$

as x varies over M. The next step is to show that this sum is small at self-focal points if there do not exist invariant L^2 functions.

PROPOSITION 10.42. *Assume that x is a self-focal point and that U_x has no invariant L^2 function. Then, for all $\eta > 0$, there exists T_x so that for $T \geq T_x$,*

$$\frac{1}{T} \left| \int_{S^*_x M} \sum_{k=0}^{\infty} \hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big) U_x(\lambda)^k \mathbf{1} |d\xi| \right| \leq \eta. \tag{10.87}$$

REMARK 10.43. In the proof of Theorem 10.29 it is only necessary to study twisted self-focal points. If there are only finitely many twisted self-focal points, then the time $T = T_x$ above is uniform in x.

This is a simple application of the von Neumann mean ergodic theorem to the unitary operator U_x. Recall that the mean ergodic theorem for unitary operators on a Hilbert space H states that $A_N f = \frac{1}{N} \sum_{n=0}^{N} U^n f \to Pf$ in H where $P \colon H \to H^0$ is the orthogonal projection onto the space H^0 of invariant functions satisfying $Uf = f$. See for instance [**Kr**, Theorem 1.4].

The mean ergodic theorem applies to our problem as follows:

PROPOSITION 10.44. *Let $H = L^2(S^*_x M, |d\xi|)$ and let $U = U_x$ (10.45). Then $H^0 = \{0\}$ if and only if*

$$\frac{1}{T} \int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big) U_x(\lambda)^k \mathbf{1} |d\xi| \to 0, \quad T \to \infty.$$

We only use the 'only if' direction in the main result.

PROOF. Suppose $H^0 = \{0\}$. Since $T_x(\xi) \geq \tau_x > 0$, $T^{(k)} \geq k\tau_x$. Then $T_x^{(k)}(\xi) \leq T$ implies $k \leq \frac{T}{\tau_x}$ and

$$\sum_{k=1}^{\infty} \int_{\mathcal{L}_x M} \hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big) U_x(\lambda)^k \mathbf{1} |d\xi| \leq \frac{1}{T} \sum_{k=1}^{\frac{T}{\tau_x}} \langle U_x(\lambda)^k \mathbf{1}, \mathbf{1} \rangle. \tag{10.88}$$

Assuming $H^0 = \{0\}$ the right side tends to zero as $T \to \infty$, proving the 'only if' side.

We may assume $\hat{\rho} \geq 0$. Using a lower bound for $\hat{\rho}$ on a small interval around 0, we see that there also exists $\tau_x' > 0$ so that

$$\frac{1}{T}\int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big) U_x(\lambda)^k \mathbf{1} |d\xi| \geq \frac{1}{T}\sum_{k=1}^{\frac{T}{\tau_x'}} \langle U_x(\lambda)^k \mathbf{1}, \mathbf{1}\rangle.$$

If the left side tends to zero then so does the right side. □

When there exists an invariant L^2 function for U_x it can be used to define an invariant measure for Φ_x in the class of $d\mu_x$. If $f \in L^2(S_x^*M, d\mu_x)$ satisfies $U_x f = f \iff f(\Phi_x \xi)\sqrt{J_x}(\xi) = f(\xi)$, then $|f|^2 d\mu_x$ is an invariant measure which is absolutely continuous with respect to $d\mu_x$.

REMARK 10.45. Consider the special case

$$\Psi_x^*(\xi) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \sqrt{J_x^{(k)}}(\xi) = \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} U_x(\lambda)^k \mathbf{1}$$

where the limit is taken in $L^2(S_x^*M, d\mu_x)$. By the mean ergodic theorem, $\Psi_x^* = P\cdot 1$, and a finite invariant measure is given by $(\Psi_x^*)^2 d\mu_x$. We must verify that $\Psi_x^* \neq 0$. However, $(\Psi_x^*)^2 d\mu_x$ is a maximal finite invariant measure in the sense that any other finite invariant measure must be absolutely continuous with respect to it. So if $\Psi_x^* = 0$ a.e., then there cannot exist any finite invariant measure. This contradiction proves that $\Psi_x^* \neq 0$.

We recall that we assume all $\mathcal{L}$ points are $\mathcal{TL}$ (see Definition 10.22).

LEMMA 10.46. *Suppose that there exist only a finite number of $\mathcal{TL}$ points. Then the statement of Proposition 10.42 is correct.*

PROOF. In this case, $U_x(\lambda) = 0$ except for a finite number of points $x_1, \dots, x_M$. Since there are only a finite number of twisted self-focal points we can take the maximum of T_{x_j} for each x_j so that all inequality (10.87) holds for the N $\mathcal{TL}$ points. □

10.8. Completion of the proof of Proposition 10.30 and Theorem 10.29: study of $\tilde{R}_{j1}$

The remainder of the proof of Proposition 10.30 and Theorem 10.29 is mainly a study of the term $\tilde{R}_{j1}$ (10.83). The following Lemma turns out to be useful:

LEMMA 10.47. *The sum*

$$\int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}(\frac{T_x^{(k)}(\xi)}{T}) U_x(\lambda)^k \mathbf{1} |d\xi| + \tilde{R}_{j1}$$

is a continuous function of x.

PROOF. This follows from Proposition 10.41. It is evident that the original remainder is a smooth function of x, and the absolute value (10.86) is continuous (in fact smooth). This sum was then broken up into the critical point integral over $\mathcal{L}_x$ and its complement (10.86). The critical point integral was identified with the term $\int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}(\frac{T_x^{(k)}(\xi)}{T}) U_x(\lambda)^k 1 |d\xi|$. The second is $\tilde{R}_{j1}$. □

10.8.1. Estimate of $\tilde{R}_{j1}$. To complete the proof of Proposition 10.30 it suffices to show that, for any $\eta > 0$ there exists ε so that $|\tilde{R}_{j1}(x;\varepsilon)| \leq \eta$. It is helpful to observe the following simplification:

LEMMA 10.48. *There exists a uniform positive constant C so that for all (x, ε),*

$$|\tilde{R}_{j1}(x;\varepsilon)| \leq C\mu_x \left(\{\xi : 0 < |\nabla_\xi \tilde{t}_j(x,\xi)|^2 < \varepsilon\}\right). \tag{10.89}$$

PROOF. To prove the Lemma, we make a dyadic decomposition of this set into the subsets

$$E_k := \{\xi\colon \varepsilon 2^{-k-1} \leq |\nabla_\xi \tilde{t}_j|^2 \leq \varepsilon 2^{-k}\}.$$

Since we punctured out the critical set from $\tilde{R}_{j1}$ we have

(10.90)

$$\tilde{R}_{j1}(x;\varepsilon) = \int_{S_x^*M} \left\{ \int_0^1 \frac{d}{ds} f((s\varepsilon)^{-1} |\nabla_\xi \tilde{t}_j(x,\xi)|^2) ds \right\} r_*(\hat{\rho}(\tilde{t}_j(x,\xi)) a_0(\tilde{t}_j(x,\xi), x, \xi) r_*^n d\xi$$

(10.91)

$$= \int_{S_x^*M} \left\{ \int_0^1 f'((s\varepsilon)^{-1} |\nabla_\xi \tilde{t}_j(x,\xi)|^2) s^{-2} ds \right\} \left\{ \varepsilon^{-1} |\nabla_\xi \tilde{t}_j(x,\xi)|^2 \right\} r_*(\hat{\rho}(T_x(\xi)))$$

(10.92)

$$\times\, a_0(T_x(\xi), x, \xi) r_*^n d\xi.$$

Recalling the properties of f from §10.6.1, the function $f'((s\varepsilon)^{-1}|\nabla_\xi \tilde{t}_j|^2)$ vanishes unless $s \leq \varepsilon^{-1}|\nabla_\xi \tilde{t}_j|^2 \leq 2s$, we can bound

$$\left| \int_0^1 f'((s\varepsilon)^{-1}|\nabla_\xi \tilde{t}_j|^2) s^{-2}\, ds \right| \varepsilon^{-1} |\nabla_\xi \tilde{t}_j|^2 \leq 2 \left| \int_0^1 |f'((s\varepsilon)^{-1}|\nabla_\xi \tilde{t}_j|^2) s^{-1}\, ds \right|.$$

We have

$$\frac{1}{2}\varepsilon^{-1}|\nabla_\xi \tilde{t}_j|^2 \leq s \leq \varepsilon^{-1}|\nabla_\xi \tilde{t}_j|^2,$$

and hence in the set E_k we have $2^{-k-1} \leq s \leq 2^{-k}$. Therefore for $\xi \in E_k$, we have

$$\left| \int_0^1 |f'((s\varepsilon)^{-1}|\nabla_\xi \tilde{t}_j|^2) s^{-1}\, ds \right| \leq C \int_{2^{-k-1}}^{2^{-k}} \frac{ds}{s} = C \log 2,$$

and so

$$|\tilde{R}_{j1}(x;\varepsilon)| \leq C \sum_{k \in \mathbb{Z}} \mu_x(E_k) = C\mu_x \{\xi : 0 < |\nabla_\xi \tilde{t}_j|^2 < \varepsilon\}.$$

Thus the claimed bound on $\tilde{R}_{j1}$ in Lemma 10.48 is proved. □

REMARK 10.49. Lemma 10.48 is only useful away from self-focal points. In small balls around self-focal points, the measure of the almost loop points in the Lemma might not tend to zero uniformly. Indeed, at the center one uses ergodic theory of the first return map to show that the integral over $\mathcal{L}_x$ tends to zero. It would be interesting to have a dynamical interpretation of the integral over $\{\xi : |\nabla_\xi \tilde{t}_j| < \varepsilon(\lambda)\}$ where $\varepsilon(\lambda) \to 0$ as $\lambda \to \infty$. Since almost-loops are not loops, it is not clear how to obtain a dynamical system on S_x^*M. But it seems reasonable to expect that the conservative-dissipative duality exists in some form for the behavior of almost-loops.

10.8.2. Completion of the proof of Proposition 10.30. To summarize the progress so far towards the proof Proposition 10.30, we combine Lemma 10.33, Corollary 10.36, Lemma 10.37 and Lemma 10.48 to obtain:

COROLLARY 10.50.

$$\rho_T * dN(\lambda, x) = \lambda^{n-1} + \lambda^{n-1} \int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\Big((\frac{T_x^{(k)}(\xi)}{T} \Big) U_x(\lambda)^k \mathbf{1} |d\xi| \\ + \lambda^{n-1} \mu_x \left(\{ \xi : 0 < |\nabla_\xi \tilde{t}_j(x, \xi)|^2 < \varepsilon \} \right) + o_T(\lambda^{n-1}) + O(\lambda^{n-1}(\varepsilon^2 \lambda)^{-1}), \tag{10.93}$$

where $o_T(\lambda^{n-1})$ is uniform in x.

Thus, we see that

$$\mathcal{R}_{T,x}(\lambda) = \lambda^{n-1} \mu_x \left(\{ \xi : 0 < |\nabla_\xi \tilde{t}_j(x, \xi)|^2 < \varepsilon \} \right) + o_{T,\varepsilon}(\lambda^{n-1}). \tag{10.94}$$

Since $\mu_x\{\xi : 0 < |\nabla_\xi \tilde{t}_j|^2 < \varepsilon\} \to 0$ as $\varepsilon \to 0$, the combination of Lemma 10.35 and Lemma 10.48 implies Proposition 10.30. As in Lemma 10.35, we obtain a uniform remainder in ε for the $o_{T,\varepsilon}(\lambda^{n-1})$ term if $\varepsilon \geq \lambda^{-\frac{1}{2}} \log \lambda$.

By Corollary 10.50 and Proposition 10.42, the only remaining step in the proof is to estimate $\mu_x \left(\{ \xi : 0 < |\nabla_\xi \tilde{t}_j(x, \xi)|^2 < \varepsilon \} \right)$. One may regard this term as the measure of 'almost-loops' at x, i.e., almost critical points of $\tilde{t}(x, \xi)$.

For $\xi \in S_x^* M$, define

$$V(x, \varepsilon) := \mu_x\{\xi : |\nabla_\omega \tilde{t}_j(x, \xi)| \leq \varepsilon\} \quad \text{and} \quad \mu(x, \varepsilon) := \mu_x\{\xi : 0 < |\nabla_\xi \tilde{t}_j| < \varepsilon\} \tag{10.95}$$

We note that $V(x, \varepsilon)$ is the integral of the characteristic function of closed set, hence is USC (upper semi-continuous) while $\mu(x, \varepsilon)$ is the same for an open set and is LSC (lower semi-continuous); by Fatou's Lemma, the integral of an LSC function is also LSC. If $|\mathcal{L}_x| = 0$ they are equal and therefore are continuous at x. Thus, the only possible points of discontinuity of $\mu(x_j, \varepsilon)$ are the self-focal points x_j. At such a point, $V(x, \varepsilon) - \mu(x, \varepsilon) = |\mathcal{L}_x|$. We have,

$$\text{(i) } \mu(x, \varepsilon) \downarrow 0 \quad \text{and} \quad \text{(ii) } V(x, \varepsilon) \downarrow \begin{cases} 0, & x \neq x_j, \\ |\mathcal{L}_{x_j}|, & x = x_j. \end{cases} \tag{10.96}$$

Obviously, the limit (ii) is not uniform. It is possible that (i) is uniform but this depends on the properties of $|\nabla_\xi \tilde{t}_j(x, \xi)|$ and does not generally hold for functions on M of the type $\alpha(x, \varepsilon) = \mu_x\{0 < q(x, \xi) \leq \varepsilon\}$ where $q(x, \xi)$ is smooth. For instance, if $q(x, \xi)$ is independent of ξ and equals $d(x, x_0)$ where x_0 is a perfect self-focal point ($\mathcal{L}_x = S_x^* M$), then $\alpha(x, \varepsilon) = |S_x^* M|$ if $0 < d(x, x_0) \leq \varepsilon$ and $\alpha(x_0, \varepsilon) = 0$. In our situation, where $q(x, \xi) = |\nabla_\xi \tilde{t}_j(x, \xi)|$ this kind of behavior could occur if the set of ε-'almost critical points' of $\tilde{t}_j(x, \xi)$ had measure bounded below by some $C_0 > 0$ for x along a sequence tending to the focal point x_0 and if the almost-critical set tends to the (punctured out) critical point set when $\varepsilon = 0$.

Since the possible geometric scenarios appears complicated, we work instead on the complement of small balls around the self-focal points and apply the perturbation estimate of Proposition 10.47 inside the small balls.

10.8.3. Uniformity of the remainder $o_{x,T}(\lambda^{n-1})$ in x. In this section we prove that $\mathcal{R}_{x,T}(\lambda) = o_T(\lambda^{n-1})$ uniformly in x as long as there are only a finite number N of self-focal points $\{x_j\}_{j=1}^N$. We may assume they are twisted.

PROPOSITION 10.51. *Suppose that* $|\mathcal{L}_x| = 0$ *except at a finite number of twisted self-focal points* $\{x_j\}_{j=1}^N$. *Then, for any* $\eta > 0$ *there exists* ε, T *so that* $\mathcal{R}_{x,T}(\lambda^{n-1})$ *is* $\leq \eta$ *uniformly in* x.

PROOF. For $\delta > 0$ let $B(x, \delta)$ denote the open ball of radius δ around x and decompose

$$M = \left[\bigcup_{j=1}^{N} B(x_j, \delta)\right] \bigcup \left[M \backslash \bigcup_{j=1}^{N} B(x_j, \delta)\right]$$

with $\delta > 0$ to be chosen later. We refer to $\bigcup_{j=1}^N B(x_j, \delta)$ as points 'near' the self-focal points and its complement as points far from self-focal points, and work on each set separately.

First consider the points 'far' from the self-focal points x_j:

$$x \in M \backslash \bigcup_{j=1}^{M} B(x_j, \delta). \tag{10.97}$$

For x in this set, $\mu(x, \varepsilon) = V(x, \varepsilon)$ (see (10.95)), so they are continuous functions on the closed set (10.97).

LEMMA 10.52. *For any* $\delta > 0$, $\mu(x, \varepsilon) = V(x, \varepsilon) \to 0$ *uniformly in* $x \in M \backslash \bigcup_{j=1}^M B_\delta(x_j)$ *as* $\varepsilon \to 0$.

PROOF. This follows from Dini's theorem that a sequence of continuous (or just usc functions) which decrease to a continuous (or just lsc function) do so uniformly. $V(x, \varepsilon) = \mu(x, \varepsilon)$ is continuous on $M \backslash \{x_j\}_{j=1}^M$ and it is clearly decreasing as $\varepsilon \to 0$ to 0 on $M \backslash \bigcup_{j=1}^M B_\delta(x_j)$. □

To complete the proof of Proposition 10.51 we need uniformity on $\bigcup_{j=1}^N B(x_j, \delta)$.

LEMMA 10.53. *For any* $\eta > 0$ *there exists* T, δ, ε *so that*

$$\int_{\mathcal{L}_x} \sum_{k=1}^{\infty} \hat{\rho}\left(\frac{T_x^{(k)}(\xi)}{T}\right) U_x(\lambda)^k \mathbf{1} |d\xi| + \mathcal{R}_{T,x}(\lambda) \leq \eta$$

uniformly on $\bigcup_{j=1}^N \overline{B(x_j, \delta)}$.

PROOF. It suffices to prove the statement for each ball $B(x_j, \delta)$. But then this follows from Proposition 10.42 and from Lemma 10.47. Indeed, the sum is uniformly continuous on $\overline{B(x_j, \delta)}$, and by Proposition 10.42 the first term is $\leq \eta$ for any given η at the center for sufficiently large T and zero elsewhere in the ball. The value at any other point in the ball is therefore $\leq 2\eta$ by Proposition 10.47. □

This completes the proof of Proposition 10.51. □

REMARK 10.54. Dini's theorem does not apply to $V(x, \varepsilon)$ on all of M. It would apply to $\mu(x, \varepsilon)$ when this function is continuous, but we do not know if this occurs. An interesting example might be a tri-axial ellipsoid, when x_0 is an umbilic point. The almost- return time $\tilde{t}_j(x, \xi)$ could probably be evaluated for x near x_0 using action-angle variables.

This completes the proof of Theorem 10.29 in the case where there are only finitely many twisted self-focal points.

10.9. Infinitely many twisted self-focal points

In this section, we modify the proof of Theorem 10.29 so that it applies to the case of infinitely many self-focal points as long as there are only finitely many with return time $\leq T$ for every T. According to Proposition 10.23, this occurs if all focal points are twisted, as we may assume.

10.9.1. Case 3: For each T_0 there are only finitely many $\mathcal{TL}_{T_0}$ points. In this case, we allow an infinite number of $\mathcal{TL}$ points but assume that, for any $T_0 > 0$ there are only finitely many, $M(T_0)$, $\mathcal{TL}_T$-points for $T \leq T_0$.

LEMMA 10.55. *If for all $T_0 > 0$, the number $\#\mathcal{TL}_{T_0} = M(T_0) < \infty$, then the statement of Proposition 10.42 is correct: for any $\eta > 0$ there exists $T(\eta)$ such that for $T \geq T(\eta)$,*

$$\frac{1}{T}\sum_{k=1}^{\infty}\int_{\mathcal{L}_x M}\hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big)U_x(\lambda)^k\mathbf{1}|d\xi| \leq \eta.$$

PROOF. We first consider the $M(T_0)$ points in $\mathcal{TL}_{T_0}$ Since T_0 are fixed but T is free to vary, we can use the hypothesis that the ergodic means tend for each x to zero to choose T large enough so that

$$\max_{x\in\mathcal{TL}_{T_0},}\left|\frac{1}{T}\int_{\mathcal{L}_x}\sum_{k=1}^{T}\hat{\rho}(\frac{T_x^{(k)}(\xi)}{T})U_x^k\cdot 1|d\xi|\right| \leq \eta. \tag{10.98}$$

We then consider $x \in \mathcal{TL}\backslash\mathcal{TL}_{T_0}$.

Given any small η we first pick T_0 so that $\frac{1}{T_0} \leq \eta$. Since $T_x(\xi) \geq T_0 > 0$, when $x \notin \mathcal{TL}_{T_0}$ lies in the second set, $T^{(k)} \geq kT_0$ and $T_x^{(k)}(\xi) \leq T \implies k \leq \frac{T}{T_0}$. If follows that for all $T > 0$ and $x \in \mathcal{TL}\backslash\mathcal{TL}_T$, we have

$$\frac{1}{T}\sum_{k=1}^{\infty}\int_{\mathcal{L}_x M}\hat{\rho}\Big(\frac{T_x^{(k)}(\xi)}{T}\Big)U_x(\lambda)^k\mathbf{1}|d\xi| \leq \frac{1}{T}\sum_{k=1}^{\frac{T}{T_0}}\langle U_x^k 1, 1\rangle \leq \frac{1}{T_0} \leq \eta. \tag{10.99}$$

Combining the two inequalities (10.98) and (10.99), it follows that for any $\eta > 0$ we can pick T sufficiently large so that for all x, the inequality of (10.99) is correct, proving the Lemma. □

We further need to generalize Proposition 10.51 to this case.

PROPOSITION 10.56. *Suppose that for each $T > 0$ the number of self-focal points with return time $\leq T$ is finite. Then, for any $\eta > 0$ there exists ε, T so that $\mathcal{R}_{x,T}(\lambda^{n-1})$ is $\leq \eta$ uniformly in x.*

For each T, we split the $\{x_j\}$ into two families, the finite set $\{x_j\}_{j=1}^{M(T)} \subset \mathcal{L}_T$ and the infinite number $\{y_j\} \in \mathcal{L}\backslash\mathcal{L}_T$. We then break up $M = \bigcup_{j=1}^{M(T)} B_\delta(x_j)$ and its complement.

By Proposition 10.42, we have:

LEMMA 10.57. *If $r(x, x_\lambda) \leq C\delta(\lambda)$ with $\delta(\lambda) = o(1)$ and if x is a $\mathcal{TL}$ point then for any $\eta > 0$, there exists T, ε so that*

$$|R(\lambda, x, T) \leq \eta + o(1)$$

LEMMA 10.58. *Assume that*

- *U_x has no invariant L^2 function for any x.*
- *For any T there exist a finite number $M(T)$ of self-focal points with first return time $\leq T$.*

Then for any $\eta > 0$, there exists T, ε so that

$$|R(\lambda, x, T, \varepsilon)| \leq \frac{C}{T} + o_T(1) + \eta$$

as $\lambda \to \infty$, where C and $o_T(1)$ are independent of x and λ.

PROOF. By Lemma 10.55, the second hypothesis implies that, for any $\delta > 0$, there exists ε_0 and a uniform constant $C > 0$ so that so that for $\varepsilon \leq \varepsilon_0$

$$\frac{1}{T}\sum_k |R_k(\lambda, x, T)| \leq \left(\frac{1}{T}\sum_k \hat{\rho}\left(\frac{k\tau_0}{T}\right)\right) V(x, \varepsilon) \leq C\eta$$

uniformly in $x \in M \backslash \bigcup_{j=1}^M B_\delta(x_j)$ as $\varepsilon \to 0$. Also, $\lim_{\varepsilon\to 0} \sup_x \tilde{T}_\varepsilon(x) = 0$. The same estimate holds at self-focal points with return time $\geq T$ by §10.9.1. This is stronger than the stated conclusion on the designated set. We therefore may assume in the rest of the proof that $x \in \bigcup_{j=1}^M B_\delta(x_j)$.

But then let $\delta_\lambda = d(x_\lambda, x_k) \to 0$ and apply Lemma 10.57 and Proposition 10.42 to conclude the proof. □

To complete the proof of Theorem 10.29 we prove

LEMMA 10.59. *Let $x \in \overline{\mathcal{TL}} \backslash \mathcal{TL}$. Then for any $\eta > 0$ there exists $r(x, \eta) > 0$ so that*

$$\sup_{y \in \overline{B(x, r(x,\eta))}} |R(\lambda, y, T)| \leq \eta.$$

Indeed, let $p_j \to x$ with $T(p_j) \to \infty$. Up to a uniform negligible term, the remainder is given at each p_j by (10.69). But for any fixed T, the first term of (10.69) has at most one term for j sufficiently large. Since the remainder is continuous, the remainder at x is the limit of the remainders at p_j and is therefore $O(T^{-1}) + O(\lambda^{-1})$.

By the perturbation estimate, one has the same remainder estimate in a sufficiently small ball around x. The first term (10.82) is $U_x(\lambda)\mathbf{1}$. The sum in Lemma 10.37 is studied in §10.7.

10.10. Dynamics of the first return map at a self-focal point

This section is purely dynamical. We consider the geometry of first return maps fixing an L^1 invariant measure in the real analytic case.

10.10.1. Real analytic surfaces: Proof of Proposition 10.19.

We may assume with no loss of generality that M is diffeomorphic to S^2. We also assume that Φ_p is real analytic, since that is the case in our setting; most of the statements below are true for smooth circle maps.

10.10.1.1. *Orientation preserving case.* First let us assume that Φ_p is orientation preserving. Then it has a rotation number. We recall that the rotation number of a circle homeomorphism is defined by

$$r(f) = \left(\lim_{n \to \infty} \frac{F^n(x) - x}{n} \right) \mod 1.$$

Here, $F\colon \mathbb{R} \to \mathbb{R}$ is a lift of f, i.e., a map satisfying $F(x+1) = F(x)$ and $f = \pi \circ F$ where $\pi : \mathbb{R} \to \mathbb{R}/\mathbb{Z}$ is the standard projection. The rotation number is independent of the choice of F or of x. It is rational if and only if f has a periodic orbit.

LEMMA 10.60. *The rotation number of Φ_p is either 0 or π.*

PROOF. For a circle homeomorphism, the rotation number $\tau(f^{-1})$ is always $-\tau(f)$. Since Φ_p is reversible, $\tau(\Phi_p) = -\tau(\Phi_p)$, i.e. its rotation number can only be $0, \pi$. □

LEMMA 10.61. *Φ_p^2 has fixed points.*

PROOF. The rotation number of Φ_p^2 is 0. But it is known that $\tau(f) = 0$ if and only if f has a fixed point. See [**Fr**, Theorem 2.4]. □

We now complete the proof that $\Phi_p^2 = Id$ if Φ_p is orientable. Since Φ_p is real analytic, this is the case if Φ_p^2 has infinitely many fixed points, so we may assume that $\mathrm{Fix}(\Phi_p^2)$ is finite (and non-empty). We write $\#\,\mathrm{Fix}(\Phi_p^2) = N$ and denote the fixed points by p_j.

If $N = 1$, i.e. Φ_p^2 has one fixed point Q, then $S^1 \backslash \{Q\}$ is an interval and Φ_p^2 is a monotone map of this interval. So every orbit is asymptotic to the fixed point of Φ_p^2.

Let μ be the L^1 invariant measure for Φ_p^2 and let $K = \operatorname{supp} \mu$. We can decompose K into N subsets K_j such that $\Phi_p^2(K_j) \to p_j$. K_j is the basin of attraction of p_j.

Then

$$\mu(K_j) = \mu(\Phi^{2p}(K)_j) \to \mu(\{p_j\}).$$

But $p_j \in K_j$ so it must be that $K_j = \{p_j\}$. This shows that μ cannot be L^1, concluding the proof.

10.10.1.2. *Orientation reversing case.* The square η_p^2 of an orientation reversing diffeomorphism of $S_p^* S^2$ is an orientation preserving diffeomorphism. If η_p preserves the measure $d\mu$ then so does η_p^2. Thus we reduce to the orientation preserving case.

10.11. Proof of Proposition 10.20

In this section we consider the case of a perfect self-focal point z, e.g. in the real analytic case. The main point is that the flow-out of $S_z^* M$ is an embedded Lagrangian submanifold $\mathcal{C}_T$ invariant under the geodesic flow. We will explain how to associate quasi-modes to this Lagrangian submanifold. It is an open problem to relate these quasi-modest to the sequence of eigenfunctions of maximal sup norm growth. For instance the 'micro-local' defect measures (or quantum limits) of such a sequence 'should' have the natural invariant Lebesgue measure on $\mathcal{C}_T$ as a component. We only sketch the proof because our motivation is to pose a problem rather than to solve one.

Since z is a perfect self-focal point, i.e., $\mathcal{L}_z = S^*zM$ the geodesic flow induces a smooth first return map. Let $\mathcal{C}_T$ denote the mapping cylinder of G_z^T, namely

$$\mathcal{C}_T = S_z^*M \times [0,T]/\cong \text{ where } (\xi, T) \cong (G_z^T(\xi), 0). \tag{10.100}$$

The $\mathcal{C}_T$ is a smooth manifold. It naturally fibers over S^1 by the map

$$\pi\colon \mathcal{C}_T \to S^1, \quad \pi(\xi, t) = t \mod 2\pi\mathbb{Z}.$$

Proposition 10.62. *Let (M, g) be an n-dimensional Riemannian manifold, and assume that it possesses a blow down point z. Let $\iota\colon \mathcal{C}_T \to T^*M$ be the map*

$$\iota_z(\xi, t) = G^t(z, \xi).$$

Then ι is a Lagrange embedding whose image is a geodesic-flow invariant Lagrangian manifold, Λ_z, diffeomorphic to $S^1 \times S^{n-1} \simeq \mathcal{C}_T$.

Proof. We let ω denote the canonical symplectic form on T^*M. Then, under the map

$$\iota_x : S^1 \times S_x^*M \to T^*M, \quad \iota(t, x, \xi) \to G^t(x, \xi),$$

we have

$$\iota^*\omega = \omega - dH \wedge dt, \quad H(x, \xi) = |\xi|_g.$$

The map ι_z is the restriction of ι to $\mathbb{R} \times S_z^*M$. Since $dH = 0$ on S^*M and $\omega = 0$ on S_x^*M, the right side equals zero.

Thus, ι_x is a Lagrange immersion. To see that it is an embedding, it suffices to prove that it is injective, but this is clear from the fact that G^t has no fixed points. □

Let α_Λ denote the action form $\alpha = \xi \cdot dx$ restricted to Λ. Also, let m_Λ denote the Maslov class of Λ. A Lagrangian Λ satisfies the Bohr-Sommerfeld quantization condition if

$$\frac{r_k}{2\pi}[\alpha_\Lambda] \equiv \frac{m_\Lambda}{4} \mod H^1(\Lambda, \mathbb{Z}), \tag{10.101}$$

where

$$r_k = \frac{2\pi}{T}(k + \frac{\beta}{4}),$$

with β equal to the common Morse index of the geodesics $G^t(z, \xi), \xi \in S_z^*M$.

Proposition 10.63. *Λ_z satisfies the Bohr-Sommerfeld quantization condition.*

Proof. We need to identify the action form and Maslov class.

Lemma 10.64. *We have:*

1. $\iota_z^*\alpha_\Lambda = dt$.
2. $\iota_z^* m_{\Lambda_z} = \frac{\beta}{T}[dt]$.

Proof. For (1), let ξ_H denote the Hamiltonian vector field of H. Since $(G^t)^*\alpha = \alpha$ for all t, we may restrict to $t = T$ and to S_z^*M to obtain $(G_z^T)^*\alpha|_{S_z^*M} = \alpha|_{S_z^*M}$. But clearly, $\xi \cdot dx|_{S_z^*M} = 0$.

For (2), recall that $m_{\Lambda_z} \in H^1(\Lambda_z, \mathbb{Z})$ gives the oriented intersection class with the singular cycle $\Sigma \subset \Lambda_z$ of the projection $\pi : \Lambda_z \to M$. Given a closed curve α on Λ_z, we deform it to intersect Σ transversely and then $\int_\alpha m_{\Lambda_z}$ is the oriented intersection number of the curve with Σ. Our claim is that $\int_\alpha m_{\Lambda_z} = \beta$ where β is the common Morse index of the (not necessarily smoothly) closed geodesic loops $\gamma_\xi(t) = G^t(z, \xi), \xi \in S_z^+M$.

The inverse image of the singular cycle of Λ_z under ι_z consists of the following components:

$$\iota_z^{-1}\Sigma = S_z^* M \cup \mathrm{Conj}(z),$$

where

$$\mathrm{Conj}(z) = \{(t, \xi)\colon 0 < t < T,\ \xi \in S_z^*,\ |\det d_z \exp t\xi| = 0\}$$

is the tangential conjugate locus of z. All of $S_z^* M$ consists of self-conjugate vectors at the time T.

If $\dim M \geq 3$, then $H^1(\mathcal{C}_T, \mathbb{Z}) = \mathbb{Z}$ is generated by the homology class of a closed geodesic loop at z and in this case $\int_\alpha m_{\Lambda_z} = \beta$ by definition of the Morse index. If $\dim M = 2$, then $H^1(\mathcal{C}_T, \mathbb{Z})$ has two generators, that of a closed geodesic loop and that of $S_z^* M$. The value of m_{Λ_z} on the former is the same as for $\dim M \geq 3$, so it suffices to determine $\int_{S_z^* M} m_{\Lambda_z}$. To calculate the intersection number, we deform $S_z^* M$ so that it intersects $\iota_z^{-1}\Sigma$ transversely. We can use $G^\varepsilon S_z^* M$ as the small deformation, and observe that it has empty intersection with $\iota_z^{-1}\Sigma$ for small ε since the set of conjugate times and return times have non-zero lower bounds. □

The Lemma immediately implies (10.101), completing the proof. □

We now complete the proof of Proposition 10.20. By assumption there exists an invariant density ν for Φ_x on $S_x^* M$. Let us further assume that it is C^1. Then we can construct an invariant C^1 density on $\mathcal{C}_T$ in the parametrizing $S_z^* \times [0, T]$ as $\nu \otimes dt$. It is well-defined and smooth on the quotient since the identification $(\xi, T) \simeq (\Phi_x \xi, 0)$ preserves ν.

Since $\mathcal{C}_T$ satisfies the Bohr-Sommerfeld conditions, one can 'quantize' the invariant density as a semi-classical sequence $\{u_k\}$ with microsupport on $\mathcal{C}^T$, i.e construct an oscillatory integral whose phase parametrizes $\mathcal{C}_T$ locally and whose amplitude equals the invariant density. The construction is explained in detail in [**Dui74, CdV**] and will not be repeated here. Only the eikonal and first transport equation are satisfied, and then one has

$$\left(\Delta + \left(\frac{2\pi}{T}k + \frac{\beta}{4}\right)^2\right) u_k = O(1), \quad k \to \infty.$$

10.12. Uniformly bounded orthonormal basis

At the opposite extreme from seeking eigenfunctions which optimize the L^p-inequalities, we consider (M, g) possessing sequences of eigenfunctions which minimize the functionals ℓ^p (10.5). In particular, we ask for which (M, g) does there exist an orthonormal basis of eigenfunctions with uniformly bounded L^∞ norms? For which does there exist any orthonormal sequence of eigenfunctions (possibly sparse) with uniformly bounded L^∞ norms.

There are almost no results on this problem. It would also be interesting to add the condition that $\lambda_j^{-1}\nabla\varphi_j$ are uniformly bounded. To our knowledge the only result pertains to the quantum integrable case. In the following section we will review a quantitative solution to this problem for integrable systems. In this section we only point out an obvious necessary dynamical condition. At the present time there do not exist results on the dynamical problem that can be used for the eigenfunction problem.

LEMMA 10.65. *Suppose that $\{\varphi_j\}$ is an orthonormal basis of eigenfunctions for (M, g) with $\|\varphi_j\|_{L^\infty} \leq C_g$ for all k. Then $\{\varphi_j^2\}$ is weakly pre-compact in L^1 and each converging sequence $\varphi_{j_k}^2 \to g$ (weakly) has the property that $\int_E \varphi_{j_k}^2 dV_g \to \int_E g\, dV_g$ for any Borel set E. Moreover, any weak* limit (defect measure) of its microlocal lifts is an invariant measure for the geodesic flow on S^*M whose projection to M has a bounded density relative to dV_g.*

PROOF. The sequence φ_j^2 is uniformly integrable and has bounded L^1 norms, hence is weakly compact in L^1 by the Dunford-Pettis theorem.

The assumption implies that $\int_M f|\varphi_j|^2\, dV \leq C \int_M |f|\, dV$ for any $f \in L^\infty(M)$. Any weak* limit μ of the microlocal lifts satisfies $\int |f| d\mu \leq C \int |f|\, dV$ for $f \in L^\infty(M)$. But this implies that $\pi_* d\mu \leq C dV$. □

The conclusion implies that

$$\int_{B(r,p)} |\varphi_j|^2\, dV \leq C r^n, \quad n = \dim M$$

and similarly for the projection of any weak* limit.

The Lemma gives a very strong condition on the sequence of eigenfunctions, but there do not seem to exist results constraining Riemannian manifolds whose invariant measures project to measures on M with bounded densities except in completely integrable cases, where M can only be a flat torus. Furthermore, we do not see how to rule out that the microlocal lifts of the orthonormal basis is QUE, i.e. has one limit measure, nor that the limit is Liouville measure (whose projection is dV_g). As mentioned earlier, as far as we know at this time of writing, QE might be a generic property among Riemannian manifolds.

10.13. Appendix: Integrated Weyl laws in the real domain

The geodesic flow G^t of (M, g) of a real analytic Riemannian manifold is of one of the following two types:

(1) *aperiodic:* The Liouville measure of the closed orbits of G^t, i.e. the set of vectors lying on closed geodesics, is zero; or
(2) *periodic = Zoll:* $G^T = id$ for some $T > 0$; henceforth T denotes the minimal period. The common Morse index of the T-periodic geodesics will be denoted by β.

In the real domain, the two-term Weyl laws counting eigenvalues of $\sqrt{\Delta}$ are very different in these two cases.

(1) Let $I_\lambda = [0, \lambda]$ and let $N(\lambda) = \int_M \Pi_{I_\lambda}(x, x) dV(x)$. In the *aperiodic* case, the Duistermaat-Guillemin-Ivrii two term Weyl law states
$$N(\lambda) = \#\{j : \lambda_j \leq \lambda\} = c_m\, Vol(M, g)\, \lambda^m + o(\lambda^{m-1})$$
where $m = \dim M$ and where c_m is a universal constant.
(2) In the *periodic* case, the spectrum of $\sqrt{\Delta}$ is a union of eigenvalue clusters C_N of the form
$$C_N = \{(\frac{2\pi}{T})(N + \frac{\beta}{4}) + \mu_{Ni},\ i = 1 \ldots d_N\} \tag{10.102}$$
with $\mu_{Ni} = 0(N^{-1})$. The number d_N of eigenvalues in C_N is a polynomial of degree $m - 1$.

In the aperiodic case, we can choose the center of the spectral interval I_λ arbitrarily. In the Zoll case we center it along the arithmetic progression $\{\frac{2\pi}{T})(N + \frac{\beta}{4})\}$. We refer to [**Bes, Hor2, SaV**] for background and further discussion.

Bibliography

[ABB1] S. B. Alexander, I. D. Berg and R. L. Bishop, The Riemannian obstacle problem. Illinois J. Math. 31 (1987), no. 1, 167–184.

[ABB2] S. B. Alexander, I. D. Berg and R. L. Bishop, Geometric curvature bounds in Riemannian manifolds with boundary. Trans. Amer. Math. Soc. 339 (1993), no. 2, 703–716.

[BB] L. Béard-Bergery, Quelques exemples de variétés riemanniennes o toutes les gódésiques issues d'un point sont fermées et de meme longueur, suivis de quelques rsultats sur leur topologie. Ann. Inst. Fourier (Grenoble) 27 (1977), no. 1, xi, 23124.

[Bes] A. L. Besse, *Manifolds all of whose geodesics are closed.* Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], 1978.

[B1] J. Bourgain, L^p estimates for oscillatory integrals in several variables, GAFA, Geom. funct. anal. 1 (1991), 321-374.

[B2] J. Bourgain, Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. Part I: Schrdinger equations. Geom. Funct. Anal. 3 (1993), no. 2, 107-156.

[B3] J. Bourgain, Geodesic restrictions and Lp-estimates for eigenfunctions of Riemannian surfaces. Linear and complex analysis, 27-35, Amer. Math. Soc. Transl. Ser. 2, 226, Amer. Math. Soc., Providence, RI, 2009.

[BD] J. Bourgain and C. Demeter, New bounds for the discrete Fourier restriction to the sphere in four and five dimensions, Int. Math. Res. Not. IMRN 2015, no. 11, 3150-3184 (arXiv:1310.5244).

[Bo] L. Boutet de Monvel, Convergence dans le domaine complexe des séries de fonctions propres. C. R. Acad. Sci. Paris Sér. A-B 287 (1978), no. 13, A855–A856.

[BoGu] L. Boutet de Monvel and V. Guillemin, *The Spectral Theory of Toeplitz Operators*, Ann. Math. Studies 99, Princeton Univ. Press, Princeton, 1981.

[BoSj] L. Boutet de Monvel and J. Sjöstrand, Sur la singularité des noyaux de Bergman et de Szegö, *Asterisque* 34–35 (1976), 123–164.

[CH] Y. Canzan and B. Hanin, Fixed frequency eigenfunction immersions and supremum norms of random waves. Electron. Res. Announc. Math. Sci. 22 (2015), 76-86.

[CdV] Y. Colin de Verdiere, Quasi-modes sur les varietes Riemanniennes. Invent. Math. 43 (1977), no. 1, 15-52.

[D] Y. N. Dowker, On measurable transformations in finite measure spaces. Ann. of Math. (2) 62 (1955), 504–516.

[Dui74] J. J. Duistermaat, Oscillatory integrals, Lagrange immersions and unfolding of singularities. Comm. Pure Appl. Math. 27 (1974), 207-281.

[DG] J. J. Duistermaat and V. W. Guillemin, The spectrum of positive elliptic operators and periodic bicharacteristics. Invent. Math. 29 (1975), no. 1, 39–79.

[FZ] R. Feng and S. Zelditch, Median and mean of the supremum of L2 normalized random holomorphic fields, J. Func. Anal., 266 (2014), 5085-5107.

[Fr] J. Franks, Rotation numbers and instability sets. Bull. Amer. Math. Soc. (N.S.) 40 (2003), no. 3, 263–279.

[G] Jeffrey Galkowski, Defect measures of eigenfunctions with maximal L^∞ growth, (arXiv:1704.01452).

[GT] Jeffrey Galkowski and John A. Toth, Eigenfunction scarring and improvements in L^∞ bounds, arXiv:1703.10248.

[GSaf] T. E. Gureev and Yu. G. Safarov, G. Exact spectral asymptotics for the Laplace operator on a manifold with periodic geodesics. (Russian) Translated in Proc. Steklov Inst. Mth.

1989, no. 2, 35–53. Boundary value problems of mathematical physics, 13 (Russian). Trudy Mat. Inst. Steklov. 179 (1988), 36–53, 241.

[GLS] F. Golse, E. Leichtnam, and M. Stenzel, the heat equation on compact real-analytic Riemannian manifolds. Ann. Sci. École Norm. Sup. (4) 29 (1996), no. 6, 669–736.

[H] E. Hopf, "Ergodentheorie" , Springer (1937)

[Hor1] L. Hörmander, Oscillatory integrals and multipliers on FLp. Ark. Mat. 11, 1-11. (1973).

[Hor2] L. Hörmander, *The Analysis of Linear Partial Differential Operators*, Volumes III-IV, Springer-Verlag Berlin Heidelberg, 1983.

[Kl] W. Klingenberg, *Riemannian geometry.* de Gruyter Studies in Mathematics, 1. Walter de Gruyter & Co., Berlin-New York, 1982.

[Kr] U. Krengel, *Ergodic Theorems*, de Gruyter Studies in Mathematics 6.

[Lo] S. Lojasiewicz, Ensemble semi-analytiques, IHES Lecture Notes, 1965.

[M] J. Milnor,*Morse theory.* Annals of Mathematics Studies, No. 51 Princeton University Press, Princeton, N.J. 1963.

[NR] A. Nabutovsky and R. Rotman, Length of geodesics and quantitative Morse theory on loop spaces. Geom. Funct. Anal. 23 (2013), no. 1, 367-414.

[Nak] H. Nakagawa, A note on theorems of Bott and Samelson. J. Math. Kyoto Univ. 7 1967, 205-220.

[Saf] Y. Safarov, Asymptotics of a spectral function of a positive elliptic operator without a nontrapping condition. (Russian) Funktsional. Anal. i Prilozhen. 22 (1988), no. 3,53–65, 96 translation in Funct. Anal. Appl. 22 (1988), no. 3, 213–223 (1989).

[SaV] Y. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators. Translated from the Russian manuscript by the authors.* Translations of Mathematical Monographs, 155. American Mathematical Society, Providence, RI, 1997.

[Sj] J. Sjöstrand, *Singularités analytiques microlocales.* Astérisque, 95, 1–166, Astérisque, 95, Soc. Math. France, Paris, 1982

[So1] C. D. Sogge, Concerning the Lp norm of spectral clusters for second-order elliptic operators on compact manifolds. J. Funct. Anal. 77 (1988), no. 1, 123-138.

[So2] C. D. Sogge, *Fourier integrals in classical analysis.* Cambridge Tracts in Mathematics, 105. Cambridge University Press, Cambridge, 1993.

[So3] C. D. Sogge,*Hangzhou lectures on eigenfunctions of the Laplacian.* Annals of Mathematics Studies, 188. Princeton University Press, Princeton, NJ, 2014.

[SoZ1] C. Sogge and S. Zelditch, Riemannian manifolds with maximal eigenfunction growth, Duke Math. J. 114 (2002), 387-437.

[SoZ2] C. Sogge and S. Zelditch, Concerning the L^4 norms of typical eigenfunctions on compact surfaces, Recent developments in geometry and analysis, Vol 23, 407-423, Higher Ed. Press and Int. Press, Beijing-Boston (arXiv:1011.0215).

[SoZ3] C.D. Sogge and S. Zelditch, Focal points and sup-norms of eigenfunctions. Rev. Mat. Iberoam. 32 (2016), no. 3, 971-994. (arXiv:1311.3999);

[SoZ4] C. D. Sogge and S. Zelditch, Focal points and sup norms of eigenfunctions on real analytic Riemannian manifolds, Rev. Mat. Iberoam. 32 (2016), no. 3, 995-999. (arXiv:1409.2063).

[SoTZ] C. Sogge, J.A. Toth and S. Zelditch, About the blowup of quasimodes on Riemannian manifolds, issue on Geometric Harmonic Analysis, J. Geom. Anal 21 (2011), 150-173 (arXiv:0908.0688).

[St] E. M. Stein, , Oscillatory integrals in Fourier analysis, in "Beijing Lectures in Harmonic Analysis (Beijing, 1984)" (E.M. Stein, ed.), Annals of Mathematics Studies 112, Princeton University Press, New Jersey (1986). 307-355.

[St] E.M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[Tao] T. Tao, Some recent progress on the restriction conjecture. Fourier analysis and convexity, 217-243, Appl. Numer. Harmon. Anal., Birkhuser Boston, Boston, MA, 2004.

[Tay] M.E. Taylor, *Noncommutative harmonic analysis.* Mathematical Surveys and Monographs, 22. American Mathematical Society, Providence, RI, 1986.

[TZ] J. A. Toth and S. Zelditch, Riemannian manifolds with uniformly bounded eigenfunctions. Duke Math. J. 111 (2002), no. 1, 97-132.

[Tom] P.A. Tomas, A restriction theorem for the Fourier transform. Bull. Amer. Math. Soc. 81 (1975), 477-478.

[V] A. V. Volovoy, Improved two-term asymptotics for the eigenvalue distribution function of an elliptic operator on a compact manifold. Comm. Partial Differential Equations 15 (1990), no. 11, 1509-1563.

[Wa] F. Warner, The conjugate locus of a Riemannian manifold, Amer. J. Math.

[We] A. Weinstein, Symplectic manifolds and their Lagrangian submanifolds. Advances in Math. 6 (1971), 329-346.

[Z] S. Zelditch, Geometry of loop-sets (in preparation).

CHAPTER 11

Quantum Integrable systems

Roughly speaking, there are only two types of dynamical systems whose eigenfunctions are well understood: the ergodic ones and the integrable ones. The eigenfunctions of the two types exhibit almost opposite behavior: eigenfunctions of ergodic systems are diffuse, while eigenfunctions of integrable systems localize in phase space. In this section we study eigenfunctions in the integrable case. Although very special, integrable systems supply microlocal approximations for any dynamical system near an invariant submanifold (known as Birkhoff normal forms).

We recall that Δ is *quantum completely integrable* (QCI) if there exist $n = \dim M$ number of first-order analytic pseudo-differential operators $P_1 = \sqrt{-\Delta}, P_2 \dots, P_n$ of order one such that

$$[P_i, P_j] = 0 \quad \text{for all } i, j = 1, \dots, n \tag{11.1}$$

and whose symbols $p_1, \dots, p_n$ satisfy the non-degeneracy condition

$$dp_1 \wedge dp_2 \wedge \dots \wedge dp_n \neq 0 \quad \text{on a open dense set } \Omega \subset T^*M - 0. \tag{11.2}$$

We are assuming that $P_1 = \sqrt{-\Delta}$, but it is often simpler to assume that $\sqrt{-\Delta}$ is some other function $\hat{H}(P_1, \dots, P_m)$. Note that the symbols must Poisson commute, $\{p_i, p_j\} = 0$, i.e., the associated geodesic flow is completely integrable in the classical sense. Simple examples of QCI Laplacians in dimension two include flat tori, surfaces of revolution, ellipsoids, and Liouville tori (for background and references see [**T2, TZ1**]). There are many further examples if one considers quantum integrable Schrödinger operators (see e.g. [**He, T2, T3, T1**]), but for the sake of brevity we only consider Laplacians here.

11.1. Classical integrable systems

A completely integrable system is defined by an abelian subalgebra

$$\boldsymbol{p} = \mathbb{R}\{p_1, \dots, p_n\} \subset (C^\infty(T^*M - 0), \{,\}). \tag{11.3}$$

Here, $\{,\}$ is the standard Poisson bracket. We assemble the generators into the moment map

$$\mathcal{P} = (p_1, \dots, p_m) \colon T^*M \to B \subset \mathbb{R}^n. \tag{11.4}$$

The Hamiltonians p_j generate the $\mathbb{R}^n$-action

$$\Phi_{\boldsymbol{t}}(x, \xi) := \exp(t_1 \Xi_{p_1}) \circ \dots \circ \exp(t_1 \Xi_{p_m})(x, \xi), \quad \boldsymbol{t} = (t_1, \dots, t_m), \tag{11.5}$$

We denote Φ_t-orbits by $\mathbb{R}^n \cdot (x, \xi)$. By the Liouville-Arnold theorem [**AM**], the orbits of the joint flow Φ_t are diffeomorphic to $\mathbb{R}^k \times T^m$ for some (k, m), $k + m \leq n$. We now consider level sets $\mathcal{P}^{-1}(b)$ of the moment map and their decompositions into orbits.

First, we suppose that b is a regular value. Since $\mathcal{P}$ is proper, a regular level has the form

$$\mathcal{P}^{-1}(b) = \Lambda^{(1)}(b) \cup \cdots \cup \Lambda^{(m_{cl})}(b), \quad (b \in B_{\text{reg}}) \tag{11.6}$$

where each $\Lambda^{(l)}(b) \simeq T^n$ is an n-dimensional Lagrangian torus. Here, $m_{cl}(b) = \#\mathcal{P}^{-1}(b)$ is the number of orbits on the level set $\mathcal{P}^{-1}(b)$. In sufficiently small neighborhoods $\Omega^{(l)}(b)$ of each component torus, $\Lambda^{(l)}(b)$, the Liouville-Arnold theorem also gives the existence of local action-angle variables $(I_1^{(l)}, ..., I_m^{(l)}, \theta_1^{(l)}, ..., \theta_m^{(l)})$ in terms of which the joint flow of $\Xi_{p_1}, ..., \Xi_{p_m}$ is linearized [**AM**].

Now let us consider singular levels of the moment map and singular orbits of the $\mathbb{R}^n$-action. We use the following notation:

Definition 11.1.

- A point (x, ξ) is called a singular point of $\mathcal{P}$ if $dp_1 \wedge \cdots \wedge dp_n(x, \xi) = 0$.
- A level set $\mathcal{P}^{-1}(c)$ of the moment map is called a singular level if it contains a singular point $(x, \xi) \in \mathcal{P}^{-1}(c)$. (We then say c is a singular value and write $c \in B_{\text{sing}}$.)
- A connected component of $\mathcal{P}^{-1}(c)$ is a singular component if it contains a singular point.
- An orbit $\mathbb{R}^n \cdot (x, \xi)$ of Φ_t is singular if it is non-Lagrangian, i.e., has dimension $< n$;

Suppose that $c \in B_{\text{sing}}$. We first decompose the singular level

$$\mathcal{P}^{-1}(c) = \bigcup_{j=1}^{r} \Gamma_{\text{sing}}^{(j)}(c) \tag{11.7}$$

into connected components $\Gamma_{\text{sing}}^{(j)}(b)$ and then decompose each component into orbits:

$$\Gamma_{\text{sing}}^{(j)}(c) = \cup_{k=1}^{p} \mathbb{R}^n \cdot (x_k, \xi_k). \tag{11.8}$$

Both decompositions can take a variety of forms. The regular components $\Gamma_{\text{sing}}^{(j)}(b)$ must be Lagrangian tori. Under a non-degeneracy assumption (see Definition 11.2), the singular component consists of finitely many orbits. The orbit $\mathbb{R}^n \cdot (x, \xi)$ of a singular point is necessarily singular, hence has the form $\mathbb{R}^k \times T^m$ for some (k, m) with $k + m < n$. Regular points may of course also occur on a singular component; their orbits are Lagrangian and can take any one of the forms $\mathbb{R}^k \times T^m$ for some (k, m) with $k + m = n$.

Now let $v \in \mathcal{P}^{-1}(c)$ and assume the orbit $\mathbb{R}^n \cdot (v) := \{\exp t_1 \Xi_{p_1} \circ \cdots \circ \exp t_k \Xi_{p_k}(v) \colon t = (t_1, ..., t_k) \in \mathbb{R}^n\}$ is compact and of rank k in the sense that

$$\text{rank}(dp_1, ..., dp_n)|_v = \text{rank}(dp_1, ..., dp_k) = k < n. \tag{11.9}$$

Following [**E1**, p. 9], we observe that the Hessians $d_v^2 p_j$ determine an abelian subalgebra

$$d_v^2 \boldsymbol{p} \subset S^2(K/L, \omega_v)^* \tag{11.10}$$

of quadratic forms on the reduced symplectic subspace K/L, where we put

$$K = \bigcap_{i=1}^{n} \ker dp_i(v), \qquad L = \text{span}\{\Xi_{p_1}(v), \dots, \Xi_{p_n}(v)\}. \tag{11.11}$$

DEFINITION 11.2 ([**E1**]). The orbit $\mathbb{R}^n \cdot v$ is said to be non-degenerate of rank k if $d_v^2\mathbf{p}$ is a Cartan (i.e., maximally abelian) subalgebra of $S^2(K/L, \omega_v)^*$.

Actually, this definition is (superficially) more general than the one in [**E1**, p. 6], since Eliasson assumes through most of [**E1**] that the subalgebra is elliptic (in a sense we describe below). However, most of Eiliasson's ideas apply to generic integrable systems where the Cartan subalgebra is of mixed type, with real or complex hyperbolic generators as well as elliptic ones, as discussed in the last section of [**E1**] and in [**E2**]. Also, our assumption that (11.10) is a CSA is somewhat stronger than in [**E1**].

The definition can be rephrased in terms of reduced Hamiltonian systems, as follows. First, there is a singular Liouville-Arnold theorem which produces action variables conjugate to the angle variables on the singular orbit. As in (11.9), we choose indices so that $dp_1, ..., dp_k$ are linearly independent everywhere on $\mathbb{R}^n \cdot (v_0)$. The singular Liouville-Arnold theorem [**AM**] states that there exists local canonical transformation

$$\psi = \psi(I, \theta, x, y) \colon \mathbb{R}^{2n} \to T^*M - 0, \tag{11.12}$$

where

(11.13)
$$I = (I_1, ..., I_k), \theta = (\theta_1, ..., \theta_k) \in \mathbb{R}^k, \quad x = (x_1, ..., x_{n-k}), y = (y_1, ..., y_{n-k}) \in \mathbb{R}^{n-k}$$

defined in an invariant neighborhood of $\mathbb{R}^n \cdot (v)$ such that

$$p_i \circ \psi = I_i \qquad (i = 1, ..., k) \tag{11.14}$$

and such that the symplectic form ω on T^*M takes the form

$$\psi^*\omega = \sum_{j=1}^{k} dI_j \wedge d\theta_j + \sum_{j=1}^{n-k} dx_j \wedge dy_j. \tag{11.15}$$

As for the remaining Hamiltonians p_j, there exist constants c_{ij} with $i = k+1, ..., n$ and $j = 1, ..., k$, such that at each point of the orbit, $\mathbb{R}^n \cdot (v)$,

$$dp_i = \sum_{j=1}^{k} c_{ij} dp_j. \tag{11.16}$$

Since $dp_1, ..., dp_k$ are linearly independent in a sufficiently neighborhood U of $v \in \mathcal{P}^{-1}(c)$, the action of the flows corresponding to the Hamilton vector fields, $\Xi_{p_1}, ..., \Xi_{p_k}$ generates a symplectic $\mathbb{R}^k$ action on $\mathcal{P}^{-1}(c_0) \cap U$. We reduce U with respect to the partial moment map $\mathcal{I} := (I_1, \ldots, I_k)(= (p_1, \ldots, p_k))$, i.e., we take $\{\mathcal{I} = 0\}$ and divide by the Hamiltonian flow. This produces a $2(n-k)$-dimensional symplectic manifold

$$\Sigma_k := \mathcal{P}^{-1}(c_0) \cap U/\mathbb{R}^k, \tag{11.17}$$

with the induced symplectic form σ. We will denote the canonical projection map by:

$$\pi_k \colon \mathcal{P}^{-1}(c_0) \cap U \longrightarrow \Sigma_k. \tag{11.18}$$

Since $\{p_i, p_j\} = 0$ for all $i, j = 1, ..., n$, it follows that $p_{k+1}, ..., p_n$ induce C^∞ functions on Σ_k, which we will, with some abuse of notation, continue to write as $p_{k+1}, ..., p_n$. From (11.16), it follows that

$$dp_i(\pi_k(v)) = 0 \qquad (i = k+1, ..., n). \tag{11.19}$$

Here, we denote the single point $\pi_k(\mathbb{R}^n \cdot (v))$ by $\pi_k(v)$. We thus obtain an abelian subalgebra $\boldsymbol{p}_{\rm red} = \mathbb{R}\{p_{k+1}, ..., p_n\}$ of $(C^\infty(\Sigma_k), \{,\})$ equipped with the Poisson bracket defined by σ, consisting of functions with a critical point at $\pi_k(v)$. Equivalent to Definition 11.2 is:

DEFINITION 11.3. The orbit $\mathbb{R}^n \cdot v$ is non-degenerate of rank k if $d_v^2 \boldsymbol{p}_{\rm red}$ is a maximally abelian subalgebra of $(C^\infty(\Sigma_k), \{,\})$.

11.2. Normal forms of integrable Hamiltonians near non-degenerate singular orbits

Eliasson's normal form theorem for completely integrable systems near a compact non-degenerate singular orbit $\Lambda \subset \mathcal{P}^{-1}(c)$ of rank k expresses the Hamiltonians p_j in terms of the linear action variables I_k of (11.14) and of additional action variables in the symplectic transversal (or reduced space). Before stating the normal form theorem, we recall the definitions of the action variables.

Let $\mathcal{Q}(2m)$ denote the Lie algebra of quadratic forms on $\mathbb{R}^{2m}$ equipped with its standard Poisson bracket. It contains the following special elements (action variables):

- (i) Real hyperbolic: $I_i^h = x_i \xi_i$;
- (ii) Elliptic: $I_i^e = x_i^2 + \xi_i^2$;
- (iii) Complex hyperbolic: $I_i^{ch} = x_i \xi_{i+1} - x_{i+1} \xi_i + \sqrt{-1}(x_i \xi_i + x_{i+1} \xi_{i+1})$.

Let us call the reduced (or transversal) Hamiltonian system around the equilibrium point (or singular orbit) non-degenerate elliptic, if it is non-degenerate in the sense of Definitions 11.2 and 11.3 and if the generators of $d_v^2 \boldsymbol{p}_{\rm red}$ are elliptic as in (ii). Elliason's elliptic normal form theorem states that in this non-degenerate elliptic case, there exists a local symplectic diffeomorphism

$$\kappa \colon V \to U, \qquad \kappa(\boldsymbol{T}^k \times \{0\}) = \mathbb{R}^n \cdot v \tag{11.20}$$

from a neighborhood V of $\mathbf{T}^k \times 0$ in $T^*(\mathbf{T}^k \times \mathbb{R}^{n-k})$ to a neighborhood U of the orbit and a locally defined C^∞ function $f \colon (\mathbb{R}^n, 0) \to (\mathbb{R}, 0)$ such that

$$p_i \circ \kappa^{-1} - c_i = f(I_1^e, ..., I_{n-k}^e, I_1, ..., I_k). \tag{11.21}$$

Here, $\mathbf{T^k}$ is the standard k-dimensional torus.

There is a corresponding normal form theorem in the hyperbolic case or in the case of mixed elliptic-hyperbolic systems. The statement and proof are alluded to in [**E1**] and discussed in detail in [**E2**]. We let $2m = 2(n-k) = \dim K/L$ as above. By our assumption, the sub-algebra $d_v^2 \mathbf{p}$ is a Cartan subalgebra of $\mathcal{Q}(2m)$. By simultaneously diagonalizing the quadratic forms, we can find a basis of $d_v^2 \boldsymbol{p}$ consisting of generators of the above types. The normal form theorem on the reduced (or transversal) space now states that there exists a locally-defined canonical mapping $\kappa \colon U \to U_0$ from a small neighborhood U of $\pi_k(v) \in \Sigma_k$ to a neighborhood U_0 of $0 \in \mathbb{R}^{2m}$, with the property that

$$\{p_i \circ \kappa^{-1}, I_j^e\} = \{p_i \circ \kappa^{-1}, I_j^h\} = \{p_i \circ \kappa^{-1}, I_j^{ch}\} = 0 \quad \text{for all } i, j = 1, \ldots, n. \tag{11.22}$$

Here, p_j are actually the functions induced by $p_{k+1}, \ldots, p_n$ on Σ_k. By making a second-order Taylor expansion about $I^e = I^{ch} = I^h = 0$, it follows from (11.22) that for all $i = 1, ..., n$, there locally exist $M_{ij} \in C^\infty(U_0)$ with $\{I_i^e, M_{ij}\} = \{I_i^h, M_{ij}\} = \{I_i^{ch}, M_{ij}\} = 0$ such that

$$p_i \circ \kappa^{-1} - c_i = \sum_{j=1}^{H} M_{ij} \cdot I_j^h + \sum_{j=H+1}^{H+L+1} M_{ij} \cdot I_j^{ch} + \sum_{j=H+L+1}^{n} M_{ij} \cdot I_j^e. \tag{11.23}$$

Non-degeneracy is easily seen to be equivalent to

$$(M_{ij})(0) \in Gl(n; \mathbb{R}). \tag{11.24}$$

Here $I^h := (I_1^h, ..., I_H^h)$, $I^{ch} := (I_{H+1}^{ch}, ..., I_{H+L+1}^{ch})$ and $I^e := (I_{H+L+2}^e, ..., I_{n-k}^e)$ denote the elements defined above and $I := (I_1, ..., I_k)$ are momentum coordinates of $T^*(\mathbf{T}^k)$. The M_{ij} Poisson-commute with all the action functions.

The proof of (11.23) is similar to the elliptic case in [**E1**]; for discussion of how the results can be extended to mixed elliptic-hyperbolic systems we refer to [**E1, E2, CdVP, NV**].

11.3. Joint eigenfunctions

We now return to quantum integrable systems defined by commuting operators (11.1). We denote by $\{\varphi_\alpha\}$ an orthonormal basis of *joint eigenfunctions*,

$$P_j \varphi_\alpha = \alpha_j \varphi_\alpha, \quad \langle \varphi_\alpha, \varphi_{\alpha'} \rangle = \delta_{\alpha, \alpha'} \tag{11.25}$$

of the P_j and the joint spectrum of $(P_1, \dots, P_m)$ by

$$\operatorname{spec}(P_1, \dots, P_m) = \Sigma := \{\boldsymbol{\alpha} := (\alpha_1, ..., \alpha_m)\} \subset \mathbb{R}^m. \tag{11.26}$$

The eigenvalues of $\sqrt{-\Delta}$ are thus of the form $H(\boldsymbol{\mu})$ with $\boldsymbol{\mu} \in \Sigma$ and the multiplicity of an eigenvalue is the number of $\boldsymbol{\mu}$ with a given value of $H(\boldsymbol{\mu})$. We refer to the special joint eigenfunctions (11.25) as the QCI eigenfunctions. The QI eigenfunctions are complex-valued and we consider the nodal sets

$$\{\operatorname{Re}(\varphi_\alpha) = 0\}, \quad \{\operatorname{Im}(\varphi_\alpha) = 0\} \tag{11.27}$$

of their real or imaginary parts. Note that the complex QI eigenfunctions might have empty nodal sets, e.g., $e^{i\langle x, k\rangle}$ have no zeros on a flat torus (and its complexification). We partially characterize the QI systems with this property. It is also both interesting and difficult to consider nodal sets of all $(-\Delta)$-eigenfunctions in cases where the spectrum has high multiplicity.

The nodal hypersurfaces of the the QI eigenfunctions can sometimes be determined by separation of variables, i.e., if there exists a coordinate system in which one can separate variables in $-\Delta$ and express φ_α as a product of one dimensional functions satisfying a one dimensional differential equation. The nodal sets in both the real and complex domain are then unions of the nodal sets of the factors and the nodal distribution problem becomes a much simpler and classical problem of locating the zeros of eigenfunctions of a one dimensional equation. But it is unknown whether QCI systems always admit separation of variables and our methods do not involve reductions to one dimensional problems.

We further restrict to the "quantum toric integrable case". Although such systems are rare, the techniques and results in these model cases indicate the results for more general QCI systems. Our main results give the limit distribution of the normalized complex nodal currents along "ladders" or "rays" of joint eigenvalues. Before stating the results, we recall the definition of ladders.

11.4. Quantum toral integrable systems

DEFINITION 11.4. The Laplacian Δ of a compact, Riemannian n-manifold (M, g) is *quantum torus integrable* if there exist generators $\hat{I}_j$ of the ring of commuting first order pseudo-differential operators with the property that $e^{2\pi i \hat{I}_j} = \nu_j Id$ for some constant ν_j of modulus one.

Such generators are known as quantum action operators. We refer to [**CdV1, TZ1**] for background on the definition and also for the following

Toric integrable systems are always toric on the quantum level in the following sense: One can choose generators $\hat{I}_1, \ldots, \hat{I}_m$ of the algebra of pseudo-differential operators commuting with Δ whose exponentials generate a unitary representation of T^m on $L^2(M)$, at least up to scalars. That is, the joint spectrum is contained in an off-set of a conic subset Λ of a lattice,

$$\text{spec}(\hat{I}_1, \ldots, \hat{I}_m) = \Lambda + \nu \subset \mathbb{Z}^m + \nu, \tag{11.28}$$

where $\nu \in (\mathbb{Z}/4)^m$ is a Maslov index. For instance in the case of the standard S^2 one can choose generators whose spectrum is the set $\{(m, n + \frac{1}{2}) \colon -n \leq m \leq n,\, n \geq 0\}$.

Semiclassical limits are taken along ladders in the joint spectrum. In the case of general $\mathbb{R}^m$ action, we define for a fixed $b = (b_1, b_2, ..., b_m) \in \mathbb{R}^m$, a ladder of joint eigenvalues of $P_1 = \sqrt{-\Delta}, P_2, ..., P_m$ by

(11.29)

$$L_b := \{(\lambda_{1k}, ..., \lambda_{mk}) \in \text{spec}(P_1, ..., P_m) \colon \lim_{k \to \infty} \frac{\lambda_{jk}}{|\lambda_k|} = b_j \text{ for all } j = 1, \ldots, m\},$$

where $|\lambda_k| := \sqrt{\lambda_{1k}^2 + ... + \lambda_{mk}^2}$.

In the case of quantum torus actions, we define *rational ladders* by

$$L_\alpha = \mathbb{R}\alpha + \nu, \quad (\alpha \in \Lambda). \tag{11.30}$$

Thus, rational rays consist of multiples of a given lattice point. The definition extends to any point α in $B + \nu$. We only prove limit formulae for rational rays but the same proof works with no essential change for all ladders; see [**STZ**] for the necessary modifications.

We refer to a ladder as a *regular ladder* if $\mathcal{P}^{-1}(\alpha)$ is a regular level, and as a *singular ladder* if $\mathcal{P}^{-1}(\alpha)$ is a singular level. In this article we only consider limit distribution along ladders for regular levels. In §13.7.6 we show that the results are in fact different for singular ladders (see the example of highest weight spherical harmonics).

Restricting this Fourier integral T^m action to a ladder $L = \mathbb{N}\alpha$ is a special case of homogeneous quantization and ladder representations of Guillemin-Sternberg [**GS**]. It follows from the abelian T^m case of of [**GS**, Theorem 6.7] that the ladder projector for the ladder $L = \mathbb{N}\alpha$

$$\pi_L(x, y) = \sum_{k=1}^{\infty} \varphi_{k\alpha}(x) \varphi_{k\alpha}(y)$$

is a homogeneous Fourier integral operator associated to the canonical relation

$$Z \times_\pi Z = \Gamma' \circ \Lambda_L,$$

where $Z = \mathcal{I}^{-1}(\mathbb{R}\alpha)$ is a co-isotropic submanifold of T^*M and $Z \times_\pi Z$ is the set of pairs (z, z') with $z \sim z'$, i.e., which lie on the same leaf of its null-foliation, given in the Abelian case by the full T^m action. Also, Λ_L is the character Lagrangian, i.e., the wave front set of the ladder character $\chi_L(\boldsymbol{t}) = \sum_{k=1}^{\infty} e^{ik\langle\alpha,\boldsymbol{t}\rangle}$ (see [**GS**, Theorem 6.3]). Further it follows by [**GS**, Corollary 6.10] that $W(t)\pi_L$ is a ladder representation of T^m by Fourier integral operators associated to the composition

$$\Gamma' \circ (Z \times_\pi Z) = \{(\boldsymbol{t}, r\alpha, z, \Phi_{\boldsymbol{t}}(z')) \colon z \sim z', \mathcal{I}(z) = r\alpha\}.$$

To determine growth rates of $\varphi_{k\alpha}$ in the real and complex domains, we employ semi-classical versions of this result. Concretely, we construct oscillatory integrals representations of the L^2-normalized joint eigenfunctions by Fourier analysis on the torus:

$$\varphi_{k\alpha}(x)\varphi_{k\alpha}(y) = \int_{T^m} W(\boldsymbol{t}, x, y) e^{-2\pi ik\langle(\alpha+\nu),\boldsymbol{t}\rangle}\, d\boldsymbol{t}. \tag{11.31}$$

Proposition 11.5. *If Δ is quantum torus integrable, then there exists a unitary Fourier integral representation* (11.33)

$$W(t_1, \ldots, t_n) \colon T^n \to U(L^2(M)), \quad W(t_1, \ldots, t_n) = e^{i(t_1\hat{I}_1 + \ldots t_n\hat{I}_n)}$$

of the n-torus and a symbol $\hat{H} \in S^1(\mathbb{R}^n - 0)$ such that $\sqrt{-\Delta} = \hat{H}(I_1, \ldots, I_n)$.

The Schwartz kernel of the quantum torus action $W(\boldsymbol{t}) = \prod_{j=1}^m e^{it_j\hat{I}_j}$ is a Fourier integral representation of T^m on $L^2(M)$. It has the eigenfunction expansion,

$$W(\boldsymbol{t}, x, y) = \sum_{\alpha\in\Lambda} e^{2\pi i\langle(\alpha+\nu),\boldsymbol{t}\rangle}\varphi_\alpha(x)\overline{\varphi_\alpha(y)}, \qquad \boldsymbol{t} = (t_1, \ldots, t_m) \tag{11.32}$$

and also an oscillatory integral representation

$$W(\boldsymbol{t}, x, y) = \int_{\mathbb{R}^m} e^{iS(\boldsymbol{t},x,y,\xi)} A(\boldsymbol{t}, x, y, \xi) d\xi, \tag{11.33}$$

where S is homogeneous of degree one in ξ and A is a classical symbol in ξ and the 'moment Lagrangian'

$$\Gamma := \{(\boldsymbol{t}, \nabla_t S, x, \nabla_x S, y, -\nabla_x S) \colon \nabla_\xi S(\boldsymbol{t}, x, y, \xi) = 0\} \subset T^*(T^m \times M \times M) \tag{11.34}$$

is the 'space-time' graph of the Hamiltonian torus action.

Combining (11.33) and (11.31), and changing variables $\xi \to k\xi$ in (11.33), we have

Lemma 11.6. *For fixed y, the sequence $\{\varphi_{k\alpha}(x)\varphi_{k\alpha}(y)\}_{k=1}^\infty$ is a semi-classical Lagrangian distribution defined by an oscillatory integral with real phase $S - \langle\alpha, \boldsymbol{t}\rangle$. The critical point $\boldsymbol{t}$ occurs when $\nabla_{\boldsymbol{t}} S(\boldsymbol{t}, x, y, \xi) = \alpha$ and $\nabla_\xi S = 0$, that is, if $(x, \nabla_x S(t, x, y)) \in \Lambda_\alpha$.*

The semi-classical Lagrangian distribution of Lemma 11.6 is the semi-classical de-homogenization of the ladder projector π_L. In effect, we compose π_L with $W(\boldsymbol{t})$ to obtain the ladder representation and then integrate in $\boldsymbol{t}$ against $e^{-ik\langle\boldsymbol{t},\alpha\rangle}$. This operation can be viewed as the pushforward $\pi_{\boldsymbol{t}*}\pi_L W(\boldsymbol{t})$. and

$$\{\varphi_{k\alpha}(x)\varphi_{k\alpha}(y)\}_{k=1}^\infty \in \mathcal{O}^0(M \times M, Z_\alpha \times Z_\alpha, kH(\alpha)), \tag{11.35}$$

where

$$Z_\alpha \times_\pi Z_\alpha = \{(z, z') \colon z, z' \in \mathcal{I}^{-1}(\alpha)\}.$$

Here, $\mathcal{O}$ denotes the space of oscillatory integrals with respect to the given Lagrangian submanifold.

The complication in applying this formalism is that it only provides asymptotics for (x, y) which are real points in the classically allowed region, i.e. in image of the natural projection $\pi \colon \Lambda_\alpha \to M \times M$. In effect, we need to analytically continue homogeneous quantization into the complex domain. The generalization is carried out in §14.

11.5. Lagrangian torus fibration and classical moment map

The inverse image of a point (x_0, y_0) of the triangular region under the moment map is the set of points (x, ξ) such that $(p_\theta(x, \xi), |\xi|) = (x_0, y_0)$. It is easy to see that the inverse image is invariant under the x_2-axis rotations and under the geodesic flow, i.e., under the Hamiltonian flows of the components of the moment map. This is a Lagrangian torus.

The boundary of the image is quite special: it corresponds to singular points of the moment map, namely the equatorial geodesic, traversed in either of its two orientations.

It is helpful (and accurate) to imagine Y^k_m and its joint eigenvalue $(m, k + \frac{1}{2})$ as corresponding to the torus with $p_\theta(x, \xi) = m$ and $|\xi| = k$. If we rescale back to S^*S^2 this is $p_\theta = m/k$ which defines a 2-torus.

For instance, the central axis is $p_\theta = 0$ and that corresponds to longitudinal great circles, which depart from the north pole, converge at the south pole and then return to the north pole. This is the picture of zonal spherical harmonics.

11.6. L^p norms of Quantum integrable eigenfunctions

In this section we continue the study of (mainly joint) eigenfunctions of quantum integrable Laplacians begun in §11. The very special property of the joint eigenfunctions is that they concentrate microlocally on level sets of the classical moment map, and usually on one component (i.e., orbit of the joint Hamiltonian flow). The geometry of the foliation by orbits has implications for the L^p norms of the joint eigenfunctions. Roughly speaking, the L^p norm blows up along a sequence of joint eigenfunctions if they localize along a level set which has a singular projection to the base. There are many possible types of singularities, some of which must occur for any (or almost any) integrable system and others which are special. The geometry of the level sets and the singularities of their projections accounts for all known extremal phenomena regarding L^p norms of eigenfunctions, and this motivates our attention to QI systems. Moreover, locally (or microlocally) around a regular invariant set of the geodesic flow on any (M, g) the classical and quantum systems are well approximated by a QI systems (its Birkhoff normal form).

11.6.1. Mass concentration on small length scales. All of the examples we know where eigenfunctions saturate L^p bounds are QCI systems. These generalize the example of flat tori, round spheres and ellipsoids. The importance of these rare but special (M, g) is that they are computable and are most likely extremal for L^p problems.

One extremal problem is to determine the Riemannian manifolds which possess orthonormal bases of eigenfunctions with uniformly bounded L^∞ norms. The following result from [**TZ1**] shows that flat tori are the unique minimizers in the class of QCI systems. In the following, let $L^\infty(\lambda, g)$, resp. $\ell^\infty(\lambda, g)$ denote the maximum (resp. minimum) L^∞ norm among L^2-normalized eigenfunctions of eigenvalue λ^2.

Theorem 11.7. *Suppose that Δ is a quantum completely integrable Laplacian on a compact Riemannian manifold (M, g). Then*

(a) If $L^\infty(\lambda, g) = O(1)$ then (M, g) is flat.
(b) If $\ell^\infty(\lambda, g) = O(1)$, then (M, g) is flat.

There exists a quantitative improvement giving blow-up rates for L^p norms for QI eigenfunctions concentrating on singular level sets, i.e. level sets which are not regular tori. These eigenfunctions are the extremals for L^p blow-up and mass concentration. The following is proved in from [**TZ2**]:

Theorem 11.8. *Suppose that (M, g) is a compact Riemannian manifold whose Laplacian Δ is quantum completely integrable. Then, unless (M, g) is a flat torus, this action must have a singular orbit of dimension $< n$. If the minimal dimension of the singular orbits is ℓ, then for every $\varepsilon > 0$, there exists a sequence of eigenfunctions satisfying:*

$$\begin{cases} \|\varphi_k\|_{L^\infty} \geq C(\varepsilon)\lambda_k^{\frac{n-\ell}{2}-\varepsilon}, \\ \|\varphi_k\|_{L^p} \geq C(\varepsilon)\lambda_k^{\frac{(n-\ell)(p-2)}{2p}-\varepsilon}, \quad 2 < p < \infty. \end{cases}$$

- A point (x, ξ) is called a singular point of the moment map $\mathcal{P}$ if $dp_1 \wedge \cdots \wedge dp_n(x, \xi) = 0$;
- A level set $\mathcal{P}^{-1}(c)$ of the moment map is called a singular level if it contains a singular point $(x, \xi) \in \mathcal{P}^{-1}(c)$;
- An orbit $\mathbb{R}^n \cdot (x, \xi)$ of Φ_t is singular if it is non-Lagrangian, i.e., has dimension $< n$.

The idea is to measure local L^2 mass on shrinking tubes around special subsets of M. They are the projections of special singular level sets of the moment map of the underlying integrable system from $S^*M \to M$. Except in the case of a flat torus, singular levels such as closed geodesics always occur.

The mass in a shrinking tube can be calculating by using microlocal FIO's to conjugate to a quantum Birkhoff normal form around the level. This technique is powerful and should have other applications to L^p norms, e.g. for proving upper bounds.

11.7. Sketch of proof of Theorem 11.8

We sketch the mass estimates in shrinking tubes. We use semi-classical notation $\hbar = \lambda^{-1}$.

Let $\Lambda := \mathbb{R}^n \cdot v$ be a compact, $k < m$-dimensional singular orbit of the Hamiltonian $\mathbb{R}^n$-action generated by $(p_1, ..., p_n)$. In this section, we study mass concentration of modes in shrinking tubes of radius $\sim \hbar^\delta$ for $0 < \delta < 1/2$ around $\pi(\Lambda)$ in M, where $\pi\colon T^*M \longrightarrow M$ denotes the canonical projection map.

We denote by $T_\varepsilon(\pi(\Lambda))$ the set of points of distance $< \varepsilon$ from $\pi(\Lambda)$. For $0 < \delta < 1/2$, we introduce a cutoff $\chi_1^\delta(x; \hbar) \in C_0^\infty(M)$ with $0 \leq \chi_1^\delta \leq 1$, satisfying

- (i) $\operatorname{supp} \chi_1^\delta \subset T_{\hbar^\delta}(\pi(\Lambda))$;
- (ii) $\chi_1^\delta = 1$ on $T_{3/4\hbar^\delta}(\pi(\Lambda))$;
- (iii) $|\partial_x^\alpha \chi_1^\delta(x; \hbar)| \leq C_\alpha h^{-\delta|\alpha|}$.

Under the assumption that Λ is an embedded submanifold of M, the functions

$$\chi_1^\delta(x; \hbar) = \zeta_1(\hbar^{-2\delta} d^2(x, \pi(\Lambda))) \tag{11.36}$$

are smooth on $T_\varepsilon(\pi(\Lambda))$ and satisfy the conditions. Here, $d(.,.)$ is the Riemannian distance function. Also, $\zeta_1 \in C_0^\infty(\mathbb{R})$ with $0 \leq \zeta_1 \leq 1, \zeta_1(x) = 1$ for $|x| \leq 3/4$ and $\operatorname{supp} \zeta_1 \subset (-1,1)$.

THEOREM 11.9. *Let $\{\varphi_\mu\}$ edit must be associated to the level set Λ be a joint eigenfunction of the QCI system. Then for any $0 \leq \delta < 1/2$, $(\mathrm{Op}_\hbar(\chi_1^\delta)\varphi_\mu, \varphi_\mu) \gg |\log \hbar|^{-m}$.*

We briefly sketch the proof. Let $\chi_2^\delta(x,\xi;\hbar) \in C_0^\infty(T^*M;[0,1])$ be a second cutoff supported in a radius $\hbar^\delta$ tube, $\Omega(\hbar)$, around Λ with $\Omega(\hbar) \subset \operatorname{supp} \chi_1^\delta$ and such that $\chi_1^\delta = 1$ on $\operatorname{supp} \chi_2^\delta$. Then, clearly

$$\chi_1^\delta(x,\xi) \geq \chi_2^\delta(x,\xi), \tag{11.37}$$

for any $(x,\xi) \in T^*M$. By Gårding's inequality, (11.37) implies

$$(\mathrm{Op}_\hbar(\chi_1^\delta)\varphi_\mu, \varphi_\mu) \gg (\mathrm{Op}_\hbar(\chi_2^\delta)\varphi_\mu, \varphi_\mu). \tag{11.38}$$

We now conjugate to the model by $\hbar$-FIO. By Egorov's theorem

$$(\mathrm{Op}_\hbar(\chi_2^\delta)\varphi_\mu, \varphi_\mu) = |c(\hbar)|^2 (\mathrm{Op}_\hbar(\chi_2^\delta \circ \kappa)u_\mu, u_\mu) - C_3 \hbar^{1-2\delta}, \tag{11.39}$$

where $c(\hbar)u_\mu(y,\theta;\hbar)$ is the microlocal normal form for the eigenfunction φ_μ. Under reasonable geometric assumptions, $|c(\hbar)|^2 \gg |\log \hbar|^{-m}$.

This reduces things to calculating the matrix elements $(\mathrm{Op}_\hbar(\chi_2^\delta \circ \kappa)u_\mu, u_\mu)$ from below in the model. But the matrix elements are now in terms of elementary model eigenfunctions and the calculation has become easy. The normal form eigenfunctions separate into a product of factors and one only has to calculate one (or two) dimensional integrals. As an example, in the hyperbolic case the integral has the form

$$\begin{aligned} M_h &= \frac{1}{\log \hbar} \left(\int_0^\infty \chi(\hbar\xi/\hbar^\delta) \left| \int_0^\infty e^{-ix} x^{-1/2+i\lambda/\hbar} \chi(x/\hbar^\delta \xi) dx \right|^2 \frac{d\xi}{\xi} \right) \\ &\geq \frac{1}{C_0} (\log \hbar)^{-1} \int_0^{\hbar^{\delta-1}} \frac{d\xi}{\xi} \left| \int_0^{\hbar^\delta \xi} e^{-ix} x^{-1/2+i\lambda/\hbar} dx \right|^2 + \mathcal{O}(|\log \hbar|^{-1}) \\ &\gg |\Gamma(1/2 + i\lambda/\hbar)|^2 (1-2\delta) + \mathcal{O}(|\log \hbar|^{-1}) \geq C(\varepsilon) > 0 \end{aligned} \tag{11.40}$$

uniformly for $\hbar \in (0, \hbar_0(\varepsilon)]$.

11.7.1. Completion of the proof of Theorem 11.8. The small scale mass estimates immediately imply lower bounds on L^∞ norms and L^p norms due to the shrinking volumes of the tubes. For instance

$$\begin{aligned} \int_M |\varphi_\mu(x)|^2 \chi_1^\delta(x;\hbar)\, dV(x) &\leq \sup_{x \in T_{\hbar^{2\delta}}(\pi(\Lambda))} |\varphi_\mu(x)|^2 \int_M \chi_1^\delta(x;\hbar)\, dV(x) \\ &\leq \|\varphi_\mu\|_{L^\infty}^2 \cdot \int_M \chi_1^\delta(x;\hbar)\, dV(x) \end{aligned} \tag{11.41}$$

and it follows from Lemma 11.9 that

$$\|\varphi_\mu\|_{L^\infty}^2 \left(\int_M \chi_1^\delta(x;\hbar)\, d\,\mathrm{Vol}(x) \right) \geq C(\varepsilon) |\log \hbar|^{-m}, \tag{11.42}$$

uniformly for $\hbar \in (0, \hbar_0(\varepsilon)]$. Since

$$\int_M \chi_1^\delta(x;\hbar)\, dV(x) = \mathcal{O}(\hbar^{\delta(n-\ell)}), \tag{11.43}$$

(11.42) implies

$$\|\varphi_\mu\|^2_{L^\infty} \geq C(\varepsilon)\hbar^{-\frac{1}{2}(n-\ell)+\varepsilon}|\log\hbar|^{-m}.$$

Recalling that $\hbar^{-1} \in \{\lambda_j : \lambda_j \in \operatorname{spec}(\sqrt{-\Delta})\}$, this gives:

$$\|\varphi_{\lambda_j}\|_{L^\infty} \geq C(\varepsilon)\lambda_j^{\frac{n-\ell}{4}-\varepsilon}.$$

11.8. Mass concentration of special eigenfunctions on hyperbolic orbits in the quantum integrable case

The mass profile of scarring eigenfunctions near a hyperbolic in the completely integrable case is studied in [**CdVP**] on tubes of fixed radius and in [**NV, TZ2**] on tubes of shrinking radius. Let $\gamma \subset S^*M$ be a closed hyperbolic geodesic of an (M, g) with completely integrable geodesic flow and for which Δ_g is quantum integrable (i.e., commutes with a maximal set of pseudo-differential operators; see [**TZ2**] for background). We then consider joint eigenfunctions Δ_g and of these operators. It is known (see [**TZ2**, Lemma 6]) that there exists a special sequence of eigenfunctions concentrating on the momentum level set of γ. We will call them the *γ-sequence.*

Assume for simplicity that the moment level set of γ just consists of the orbit together with its stable/unstable manifolds. Then it is proved in [**TZ2**] that the mass of φ_μ in the shrinking tube of radius h^δ around γ with $\delta < \frac{1}{2}$ is $\simeq (1-2\delta)$ (see also [**NV**] for a closely related result in two dimensions). Thus, the mass profile of such scarring integrable eigenfunctions only differs by the numerical factor $(1-2\delta)$ from the mass profile of Gaussian beams. The difference is that the 'tails' in the hyperbolic case are longer. Also the peak is logarithmically smaller than in the elliptic case (a somewhat weaker statement is proved in [**TZ2**]).

Let us state the result precisely and briefly sketch the argument. It makes an interesting comparison to the situation discussed later on of possible scarring in the Anosov case.

We denote by $\pi\colon S^*M \to M$ the standard projection and let $\pi(\gamma)$ be the image of γ in M. We denote by $T_\varepsilon(\pi(\Lambda))$ the tube of radius ε around $\pi(\Lambda)$. For $0 < \delta < 1/2$, we introduce a cutoff $\chi_1^\delta(x;\hbar) \in C_0^\infty(M)$ with $0 \leq \chi_1^\delta \leq 1$, satisfying

- (i) supp $\chi_1^{\delta_c 11} \subset T_{\hbar^\delta}(\pi(\gamma))$
- (ii) $\chi_1^\delta = 1$ on $T_{3/4\hbar^\delta}(\pi(\gamma))$.

THEOREM 11.10. *Let γ be a hyperbolic closed orbit in (M, g) with quantum integrable Δ_g, and let $\{\varphi_\mu\}$ be an L^2 normalized γ-sequence of joint eigenfunctions Then for any $0 \leq \delta < 1/2$, $\lim_{\hbar\to 0}(\mathrm{Op}_\hbar(\chi_1^\delta)\varphi_\mu, \varphi_\mu) \geq (1-2\delta)$.*

11.8.1. Outline of proof of Theorem 11.10. For simplicity we assume $\dim M = 2$. Let $\chi_2^\delta(x,\xi;\hbar) \in C_0^\infty(T^*M;[0,1])$ be a second cutoff supported in a radius $\hbar^\delta$ tube, $\Omega(\hbar)$, around γ with $\Omega(\hbar) \subset \operatorname{supp}\chi_1^\delta$ and such that $\chi_1^\delta = 1$ on $\operatorname{supp}\chi_2^\delta$. Thus, $\chi_1^\delta(x,\xi) \geq \chi_2^\delta(x,\xi)$, for any $(x,\xi) \in T^*M$. By the Gårding inequality, there exists a constant $C_1 > 0$ such that:

$$(\mathrm{Op}_\hbar(\chi_1^\delta)\varphi_\mu, \varphi_{\mu)}) \geq (\mathrm{Op}_\hbar(\chi_2^\delta)\varphi_\mu, \varphi_\mu) - C_1\hbar^{1-2\delta}. \tag{11.44}$$

We now conjugate the right side to the model setting of $S^1 \times \mathbb{R}^1$, i.e., the normal bundle N_γ to γ. The conjugation is done by $\hbar$ Fourier integral operators and is known as conjugation to quantum Birkhoff normal form. In the model space, the conjugate of Δ_g is a function of $D_s = \frac{\partial}{i\partial s}$ along S^1 and the dilation operator $\hat{I}^h := \hbar(D_y y + y D_y)$ along $\mathbb{R}$. By Egorov's theorem

$$(\mathrm{Op}_\hbar(\chi_2^\delta)\varphi_\mu, \varphi_\mu) = |c(\hbar)|^2(\mathrm{Op}_\hbar(\chi_2^\delta \circ \kappa)u_\mu, u_\mu) - C_3\hbar^{1-2\delta} \tag{11.45}$$

where $u_\mu(y, s; \hbar)$ is the model joint eigenfunction of $D_s, \hat{I}^h$, and $c(\hbar)$ is a normalizing constant. This reduces our problem to estimating the explicit matrix elements $(Op_\hbar(\chi_2^\delta \circ \kappa)u_\mu, u_\mu)$ of the special eigenfunctions in the model setting. The operator $\hat{I}^h$ has a continuous spectrum with generalized eigenfunctions $y^{-1/2+i\lambda/\hbar}$. The eigenfunctions on the 'singular' level γ correspond to $\lambda \sim E\hbar$. A calculation shows that the mass in the model setting is given by

$$M_h = \frac{1}{\log \hbar}\left(\int_0^\infty \chi(\hbar\xi/\hbar^\delta)\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\hbar^\delta\xi)dx\right|^2 \frac{d\xi}{\xi}\right). \tag{11.46}$$

Analysis of (11.46) shows that the right side tends to $1 - 2\delta$ as $\hbar \to 0$ if $\lambda \sim E\hbar$ (see §11.9 for a detailed discussion).

11.9. Details on M_h

To estimate M_h, we assume $\lambda \sim E\hbar$ and let $\chi \in C_0^\infty(\mathbb{R}; [0, 1])$ with $\chi(x) = 1$ for $|x| \leq 1$ and $\chi(x) = 0$ for $|x| \geq 2$. Note that we divide by $|\log \hbar|$ in the model distribution $u_h(x) = \sqrt{|\log \hbar|}^{-1/2} x^{-1/2+iE/\hbar}Y(x)$, so that $\|\mathrm{Op}_\hbar(\chi)u_h\|_{L^2} \sim 1$. To estimate the mass on shrinking tubes of size $\hbar^\delta$ we write

$$\begin{aligned} M_h &= |\log \hbar|^{-1}\int_0^\infty \chi(\hbar^{1-\delta}\xi)\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\hbar^\delta\xi)dx\right|^2\frac{d\xi}{\xi} \\ &= |\log \hbar|^{-1}\int_0^\infty \chi(\hbar^{1-\delta}\xi)\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\hbar^\delta\xi)dx\right|^2\frac{d\xi}{\xi} \\ &= |\log \hbar|^{-1}\int_0^{\hbar^{\delta-1}} \frac{d\xi}{\xi}\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\hbar^\delta\xi)dx\right|^2 + \mathcal{O}(|\log \hbar|^{-1}). \end{aligned} \tag{11.47}$$

The last step follows since

$$\begin{aligned} &|\log \hbar|^{-1}\int_{\hbar^{\delta-1}}^{2\hbar^{\delta-1}} \frac{d\xi}{\xi}\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\hbar^\delta\xi)dx\right|^2 \\ &= |\log \hbar|^{-1}\int_{\hbar^{\delta-1}}^{2\hbar^{\delta-1}} \frac{d\xi}{\xi}\left|\int_1^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\hbar^\delta\xi)dx + \mathcal{O}(1)\right|^2 = \mathcal{O}(|\log \hbar|^{-1}). \end{aligned}$$

The last estimate follows by integration by parts, since when $\xi \geq \hbar^{\delta-1}$ we have $\hbar^\delta\xi \geq \hbar^{2\delta-1}$ and we assume that $2\delta - 1 < 0$ so $D_x(\chi(x/\hbar^\delta\xi)) = \mathcal{O}(\hbar^{-\delta}\xi^{-1}) \to 0$. Also, $|\Gamma(1/2 + i\lambda/\hbar)|^2 = |\Gamma(1/2 + iE)|^2 = \frac{\pi}{\cosh(\pi E)} = O(1)$. To simplify (11.47) we first make a change of variables $\xi \mapsto \hbar^\delta\xi$ and get

$$M_h = |\log \hbar|^{-1}\int_0^{\hbar^{2\delta-1}} \frac{d\eta}{\eta}\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\eta)dx\right|^2 + \mathcal{O}(|\log \hbar|^{-1}) \tag{11.48}$$

Next, we get rid of the interval $0 \leq \eta \leq 1$ by observing that when $\eta \in [0, 1]$,

$$\left|\int_0^\infty e^{-ix}x^{-1/2+i\lambda/\hbar}\chi(x/\eta)dx\right| \leq \int_0^{2\eta} x^{-1/2}dx = \mathcal{O}(\eta^{1/2})$$

and so,

$$|\log \hbar|^{-1} \int_0^1 \frac{d\eta}{\eta} \left| \int_0^\infty e^{-ix} x^{-1/2+i\lambda/\hbar} \chi(x/\eta) dx \right|^2 \ll |\log \hbar|^{-1} \int_0^1 \eta^{-1} \eta d\eta \ll |\log \hbar|^{-1}.$$

Thus,

$$M_h = |\log \hbar|^{-1} \int_1^{\hbar^{2\delta-1}} \frac{d\eta}{\eta} \left| \int_0^\infty e^{-ix} x^{-1/2+i\lambda/\hbar} \chi(x/\eta) dx \right|^2 + \mathcal{O}(|\log \hbar|^{-1}). \tag{11.49}$$

Next, one gets rid of the cutoff $\chi(x/\eta)$ by integrating by parts. When $\eta \geq 1$,

$$\int_\eta^\infty e^{-ix} x^{-1/2+i\lambda/\hbar} \chi(x/\eta) dx = \int_\eta^\infty D_x(e^{-ix}) x^{-1/2+i\lambda/\hbar} \chi(x/\eta) dx = \mathcal{O}(\eta^{-1/2}).$$

So the $[\eta, \infty]$-range of integration in x gives a contribution to M_h that is $\ll |\log \hbar|^{-1}$. The end result is the formula

$$M_h = |\log \hbar|^{-1} \int_1^{\hbar^{2\delta-1}} \frac{d\eta}{\eta} \left| \int_0^\eta e^{-ix} x^{-1/2+i\lambda/\hbar} dx \right|^2 + \mathcal{O}(|\log \hbar|^{-1}). \tag{11.50}$$

By contour deformation, for $\eta \geq 1$,

$$\left| \int_0^\eta e^{-ix} x^{-1/2+i\lambda/\hbar} dx \right|^2 = \left| \Gamma(1/2 + i\lambda/\hbar) + \mathcal{O}(\eta^{-1/2}) \right|^2. \tag{11.51}$$

Substitution into (11.50) gives

$$M_h = |\log \hbar|^{-1} |\Gamma(1/2 + i\lambda/\hbar)|^2 \int_1^{\hbar^{2\delta-1}} \frac{d\eta}{\eta} + \mathcal{O}(|\log \hbar|^{-1}). \tag{11.52}$$

This last step follows since the $\mathcal{O}(\eta^{-1/2})$-terms in (11.51) give a contribution to $M_\hbar$ that is $\ll |\log \hbar|^{-1} \int_1^{\hbar^{2\delta-1}} \frac{d\eta}{\eta} \eta^{-1/2} = \mathcal{O}(|\log \hbar|^{-1})$.

The final formula follows immediately from (11.52). With $\lambda \sim E\hbar$, one gets

$$M_h \sim |\Gamma(1/2 + iE)|^2 (1 - 2\delta). \tag{11.53}$$

Since $|\Gamma(1/2 + iE)|^2 = \frac{\pi}{\cosh \pi E}$ is bounded away from zero, it follows that (11.40) tends to $1 - 2\delta$ as $\hbar \to 0$.

11.10. Concentration of quantum integrable eigenfunctions on submanifolds

Similar methods were used in [**T3**] to obtain sharp bounds on L^2 norms for restrictions to submanifolds in the quantum integrable case, making more precise the results of [**BGT**] in this special case. For simplicity, let us consider curves on surfaces. First is the generic upper bound from [**T3**]:

THEOREM 11.11. *Let* $\varphi_{\lambda_j}; j = 1, 2, 3, ...$ *be the* L^2*-normalized joint Laplace eigenfunctions of the commuting operators* $P_1 = -\Delta$ *and* P_2 *on a Riemannian surface* (M^2, g). *Then for a generic curve* γ *such that* $\iota^* p_2|_{S^*_\gamma M}$ *is Morse, we have*

$$\int_\gamma |\varphi_{\lambda_j}|^2 ds = \mathcal{O}_{|\gamma|} (\log \lambda_j).$$

It is also proved in [**T3**] that if the curve is a geodesic, the bounds depend on the type of level set the geodesic lies on:

THEOREM 11.12. *Let $P_j(\hbar); j = 1,2$ be a non-degenerate quantum integrable system system on a surface, (M, g). Then,*

- *(i) When γ is the projection of a geodesic segment contained in $\mathcal{P}^{-1}(\mathcal{B}_{\text{reg}})$,*

$$\int_\gamma |\varphi_{\lambda_j}(s)|^2 ds = \mathcal{O}_{|\gamma|}(1),$$

- *(ii) When γ is the projection of a singular joint orbit in $\mathcal{P}^{-1}(\mathcal{B}_{\text{sing}})$,*

$$\int_\gamma |\varphi_{\lambda_j}(s)|^2\, ds = \mathcal{O}_{|\gamma|}(\lambda_j^{1/2}).$$

  Moreover, there exists a constant $c_\gamma > 0$ depending only on the curve γ, and a subsequence of joint eigenfunctions, $\varphi_{\lambda_{j_k}=}$, $k = 1, 2, \ldots$ such that

$$\int_\gamma |\varphi_{\lambda_{j_k}}(s)|^2\, ds \geq c_\gamma \lambda_{j_k}^{1/2} \quad \textit{when } \gamma \textit{ is stable,}$$

$$\int_\gamma |\varphi_{\lambda_{j_k}}(s)|^2 ds \geq c_\gamma \lambda_{j_k}^{1/2} |\log \lambda_{j_k}|^{-1} \quad \textit{when } \gamma \textit{ is unstable.}$$

Thus the exact bound depends on the nature of the geodesic. In the general quantum integrable case, most geodesics lie on regular Lagrangian tori in $\mathcal{P}^{-1}(B_{\text{reg}})$ and these geodesics do not support large L^2-bounds. But as in Theorem 11.8, there always exists a subsequence of joint eigenfunctions of P_1 and P_2 with mass concentrated along (singular) orbits contained in $\mathcal{P}^{-1}(\mathcal{B}_{\text{sing}})$, and the associated eigenfunctions saturate the upper bounds. For instance in the case of a simple surface of revolution, the equator is the projection of a singular orbit of the $\mathbb{R}^2$ action generated by geodesic flow and rotation. The corresponding joint eigenfunctions (the analogs of highest weight spherical harmonics) satisfy $\int_\gamma |\varphi_{\lambda_j}|^2\, ds \sim \lambda_j^{1/2}$ along the equator, γ. The equatorial geodesic is singular and the L^2 norms along it had singular blowup. In the case of the meridian great circles, the closed geodesic lies in the base space projection of a maximal Lagrangian torus. The zonal harmonics have $\hbar$-microsupport on this torus and have L^2-restriction bound $\sim \log \lambda$ along any meridian great circle.

11.10.1. Appendix on semi-classical pseudo-differential operators. For the reader's convenience, we review the definition of semiclassical pseudo-differential operators. Given an open $U \subset \mathbb{R}^n$, we say that $a(x, \xi; \hbar) \in C^\infty(U \times \mathbb{R}^n)$ is in the symbol class $S^{m,k}(U \times \mathbb{R}^n)$, provided

$$|\partial_x^\alpha \partial_\xi^\beta a(x, \xi; \hbar)| \leq C_{\alpha\beta} \hbar^{-m} (1 + |\xi|)^{k-|\beta|}. \tag{11.54}$$

We say that $a \in S^{m,k}_{\text{cl}}(U \times \mathbb{R}^n)$ provided there exists an asymptotic expansion

$$a(x, \xi; \hbar) \sim \hbar^{-m} \sum_{j=0}^{\infty} a_j(x, \xi) \hbar^j, \tag{11.55}$$

valid for $|\xi| \geq \frac{1}{C} > 0$ with $a_j(x, \xi) \in S^{0,k-j}(U \times \mathbb{R}^n)$ on this set.

We denote the associated $\hbar$ Kohn-Nirenberg quantization by $\text{Op}_\hbar(a)$, where this operator has Schwartz kernel given locally by the formula

$$\text{Op}_\hbar(a)(x, y) = (2\pi\hbar)^{-n} \int_{\mathbb{R}^n} e^{i(x-y)\xi/\hbar} a(x, \xi; \hbar)\, d\xi. \tag{11.56}$$

By using a partition of unity, one constructs a corresponding class, $\mathrm{Op}_\hbar(S^{m,k})$, of properly-supported $\hbar$-pseudodifferential operators acting on $C^\infty(M)$.

Bibliography

[AM] R.Abraham and J.E.Marsden, *Foundations of mechanics*, second edition, Benjamin/Cummings (1978).

[BGT] N. Burq, P. Gerard, N. Tzvetkov, Restrictions of the Laplace-Beltrami eigenfunctions to submanifolds. Duke Math. J. 138(3) (2007), 445-486.

[CdV1] Y. Colin de Verdière, Spectre conjoint d'opérateurs pseudo-différentiels qui commutent II: Le cas intégrable, Math.Zeit. 171 (1980), 51-75.

[CdV2] Y. Colin de Verdière, Quasi-modes sur les variétés riemanniennes, Invent.Math. 43 (1977), 15-52.

[CdV3] Y. Colin de Verdière, Sur le spectre des opérateurs elliptiques à bicaracteristiques toutes périodiques. (French) Comment. Math. Helv. 54 (1979), 508–522.

[CdVP] Y. Colin de Verdière and B. Parisse, Équilibre instable en régime semi-classique I: concentration microlocale, Comm. in P.D.E. 19 (1994), 1535-1563.

[D] J. J. Duistermaat, On global action-angle coordinates, Comm. Pure Appl. Math. 33 (1980), 687–706.

[E1] L.H.Eliasson, Normal forms for Hamiltonian systems with Poisson commuting integrals-elliptic case, Comment.Math.Helv. 65 (1990), 4-35.

[E2] L.H. Eliasson, Hamiltonian systems with Poisson commuting integrals. PhD thesis, University of Stockholm (1984).

[GS] V.Guillemin and S.Sternberg, Homogeneous quantization and multiplicities of group representations, J.Fun.Anal. 47(1982), 344- 380.

[He] G.J. Heckman, Quantum integrability for the Kovalevsky top. Indag. Math. (N.S.) 9 (1998), 359–365.

[NV] S. Nonnenmacher and A. Voros, Eigenstate structures around a hyperbolic point. J. Phys. A 30 (1997), no. 1, 295–315.

[STZ] C.D. Sogge, J. A. Toth and S. Zelditch, About the blowup of quasimodes on Riemannian manifolds. J. Geom. Anal. 21 (2011), no. 1, 150-173.

[T1] J. A. Toth, Various quantum mechanical aspects of quadratic forms, J. Func. Anal. 130 (1995), 1-42.

[T2] J. A. Toth, Eigenfunction localization in the quantized rigid body, J.Diff.Geom. 43(4)(1996), 844-858.

[T3] J. A. Toth, On the quantum expected values of integrable metric forms, J.Diff.Geom. 52(2)(1999), 327-374.

[T4] J. A. Toth, L^2-restriction bounds for eigenfunctions along curves in the quantum completely integrable case. Comm. Math. Phys. 288 (2009), no. 1, 379-401.

[TZ1] J. A. Toth and S. Zelditch, Riemannian manifolds with uniformly bounded eigenfunctions. Duke Math. J. 111 (2002), no. 1, 97-132.

[TZ2] J. A. Toth and S. Zelditch, L^p norms of eigenfunctions in the completely integrable case. Ann. Henri Poincaré 4 (2003), no. 2, 343-368.

[TZ3] J. A. Toth and S. Zelditch, Norms of modes and quasi-modes revisited. Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001), 435-458, Contemp. Math., 320, Amer. Math. Soc., Providence, RI, 2003.

[Z] S. Zelditch, Quantum transition amplitudes for classically ergodic or completely integrable systems, J. Fun. Anal. 94 (1990), 415-436.

[Zu] N.T. Zung, Symplectic topology of integrable Hamiltonian systems. I. Arnold-Liouville with singularities, Comp. Math. 101(1996) no. 2, 179-215.

CHAPTER 12

Restriction theorems

This chapter is devoted to restriction theorems for eigenfunctions. The term 'restriction theorem' often refers to Fourier restriction theorems, i.e., theorems about restrictions $f \to \hat{f}|_H$ of Fourier transforms of L^p functions to curved hypersurfaces H. Discrete restriction theorems in [**So1**] refers to the $L^p \to L^q$ mapping norms of spectral projections for short spectral intervals ('clusters'). The restriction in these theorems is in the frequency domain and some of the results in the setting of general Riemannian manifolds were surveyed in the section on L^p norms of eigenfunctions. In this section, 'restriction' refers to the 'space domain', i.e., refers to restrictions $\varphi_j|_H$ of eigenfunctions to a hypersurface (or other submanifold) H. This has become a large subject and we only survey a few of the many results, focusing on two types of theorems: (i) L^2 norms of restrictions, and (ii) quantum ergodic restriction theorems. We also briefly discuss integrals $\int_H f\varphi_j \, dS$ of restrictions (Kuznecov sum formulae in the sense of [**Z**]) and a result from [**HHHZ**] saying that restrictions of eigenfunctions with eigenvalues in a short interval provide an 'asymptotic orthonormal basis' for $L^2(H)$. There is also a large theory of L^p norms of restrictions and Kakeya-Nikodym norms but we do not review it here; see [**So3**] for a recent survey.

We recall the notation for the eigenvalue problem on M

$$\begin{cases} -\Delta_g \varphi_j = \lambda_j^2 \varphi_j, \quad \langle \varphi_j, \varphi_k \rangle = \delta_{jk}, \\ B\varphi_j = 0 \text{ on } \partial M, \end{cases} \tag{12.1}$$

where where B is the boundary operator, i.e., $B\varphi = \varphi|_{\partial M}$ in the Dirichlet case or $B\varphi = \partial_\nu \varphi|_{\partial M}$ in the Neumann case. We also allow $\partial M = \emptyset$. We introduce the Planck constant $h_j = \lambda_j^{-1}$; for notational simplicity we often drop the subscript j. We then denote the eigenfunctions in the orthonormal basis by φ_h and the eigenvalues by h^{-2}, so that the eigenvalue problem takes the semi-classical form

$$\begin{cases} (-h^2 \Delta_g - 1)\varphi_h = 0, \\ B\varphi_h = 0 \text{ on } \partial M. \end{cases} \tag{12.2}$$

Above we defined restrictions to be 'Dirichlet data' of eigenfunctions on a hypersurface or lower dimensional submanifold H, i.e., restrictions $\varphi_j|_H$. It is also interesting to study Neumann data $\partial_\nu \varphi_j|_H$ where ∂_ν is the unit normal derivative with respect to a given orientation of H. Further, one may combine the Dirichlet and Neumann data to form the Cauchy data of the eigenfunction on H. For this we use the semi-classical notation

$$\mathrm{CD}(\varphi_h) := (\varphi_h|_H, \; hD_\nu \varphi_h|_H) \tag{12.3}$$

as in [**CTZ**]. Here $D_\nu = \frac{1}{i}\partial_\nu$.

We often use metric Fermi normal coordinates along H, i.e., exponentiate the normal bundle to H. We denote by s the coordinates along H and y_n the normal coordinate, so that $y_n = 0$ is a local defining function for H. We also let σ, ξ_n be the dual symplectic Darboux coordinates. Thus the canonical symplectic form is $\omega_{T^*M} = ds \wedge d\sigma + dy_n \wedge d\xi_n$.

Before getting started we point out two of the possible extremes of restriction of a sequence of eigenfunctions: (i) when there is no positive lower bound on L^2-norms of restrictions of a sequence of eigenfunctions to a hypersurface H, and (ii) where a sequence of eigenfunctions blows up when restricted to a hypersurface.

12.1. Null restrictions, degenerate restrictions and 'goodness'

Regarding the first scenario, it is possible that $\varphi_{j_k}|_H \equiv 0$ for a subsequence of eigenfunctions. The question arises how to characterize hypersurfaces (or lower dimensional submanifolds or other sets) on which an infinite sequence of eigenfunctions vanishes. This question was raised by Bourgain-Rudnick [**BouR1**] and studied by them on flat tori. When the hypersurface bounds a domain it is also the question when the Dirichlet spectrum of the domain can have an infinite number of eigenvalues in common with the global spectrum of (M, g).

The same question can be posed for Neumann data. A simple example is given by the eigenfunctions $\sin(2\pi n_1 x_1)$ on a flat torus all vanish on the totally geodesic submanifold $\{x_1 = 0\}$. Related examples occur in other completely integrable settings where one can separate variables. In the case of the unit disc, a 'ray' or spoke of angle $\frac{2\pi}{m}$ is the common zero set of separation-of-variables eigenfunctions $J_m(\rho_{m,j} r) \sin m\theta$. Here, $\rho_{m,j}$ is a zero of J_m. Such eigenfunctions also vanish on certain concentric circles when $r\rho_{m,j}$ is also a zero of J_m. But it was proved by C.L. Siegel that no J_m and J_n have no common zeros for $m \neq n$. On the standard S^2, or on any surface of revolution, one has a similar situation where an infinite sequence of separation-of-variables eigenfunctions can vanish on a fixed meridian, but (as pointed out by Z. Rudnick) it is an unresolved classical conjecture that only a finite sequence can vanish on a fixed latitude circle. Another obvious (but important) kind of example, which can occur on negatively curved manifolds, is where $H \subset \mathrm{Fix}(\sigma)$ is a component of the fixed point set of an isometric involution. Then any odd eigenfunction $\varphi_j \circ \sigma = -\varphi_j$ will vanish on H. Hence it is possible for 'half' of the eigenfunctions to vanish on a single hypersurface.

Obviously, the Neumann data of odd eigenfunctions on (M, g) with an isometric involution σ do not vanish on components $H \subset \mathrm{Fix}(\sigma)$, but the Neumann data of even eigenfunctions do. Each component of the fixed point set of an isometric involution is a totally geodesic submanifold as in the integrable examples above. To our knowledge it is unknown if there exist hypersurfaces which are not totally geodesic on which a sequence of eigenfunctions vanishes. In [**BouR1**] Bourgain-Rudnick proved that on a two- or three-dimensional flat torus T^n,

$$\|\varphi_\lambda\|_{L^2(\Sigma)} \geq C \|\varphi_\lambda\|_{L^2} \tag{12.4}$$

for any positively curved smooth hypersurface. In [**BouR2**] they prove that a necessary and sufficient condition for a smooth curve $\Sigma \subset T^2$ of a flat two-torus to lie in the nodal set of an infinite sequence is that it is an arc of a closed geodesic. They further generalize this result to real analytic hypersurfaces with nowhere vanishing principal curvatures of a flat torus T^d of any dimension. In [**BouR4**] they prove positive lower bounds (12.4) for such hypersurfaces in two and three dimensional flat tori.

A related question arises how fast a sequence of restricted eigenfunctions can tend to zero in L^2 when the sequence does not vanish identically. The question is illustrated by L^2-normalized Gaussian beams φ_j^γ centered on a closed geodesic γ of a surfaces of revolution (S^2, g). On the one hand, they peak on γ and satisfy

$$\int_\gamma |\varphi_j^\gamma|^2 ds \simeq C_g \lambda_j^{\frac{1}{4}}. \tag{12.5}$$

On the other hand, they decay at a Gaussian rate in the normal direction to γ and so on a latitude circle C they satisfy

$$\int_C |\varphi_j^\gamma|^2 \, ds \simeq A_g(C) e^{-a(C)\lambda_j}. \tag{12.6}$$

Here, C_g, A_g, a_g are constants that may depend on g or the curve C but do not depend on the eigenvalue. This raises the question whether L^2 norms of non-zero restrictions can decay at a faster rate than (12.6).

J. Toth and the author have made this question precise with the notion of a *good* submanifold, specifically a good curve [**ToZ1, ToZ4**]. It is assumed that the metric is real analytic but the definition makes sense for C^∞ data as well. For instance, a real analytic curve $H \subset M$ of a real analytic Riemannian manifold (of any dimension) is said to be ‘good’ if the full sequence of eigenfunctions has a lower bound of the type (12.6). More precisely, one considers the sequence of normalized logarithms

$$u_j := \frac{1}{\lambda_j} \log |\varphi_j|^2 \tag{12.7}$$

and define the restricted sequence

$$u_j^C := \frac{1}{\lambda_j} \log |\varphi_j^{\mathbb{C}}|_C|^2. \tag{12.8}$$

Definition 12.1. Given a subsequence $\mathcal{S} := \{\varphi_{j_k}\}$ of eigenfunctions, we say that a real analytic curve $C \subset M$ is *$\mathcal{S}$-good* if there exists a constant $M_{\mathcal{S}} > 0$ so that $u_j^C \geq -M_{\mathcal{S}}$ for all $j \in \mathcal{S}$. If C is $\mathcal{S}$-good for every subsequence of every orthonormal basis sequence $\{\varphi_j\}$, we say that C is *good*.

If C fails to be good, then there must exist a sequence $\mathcal{S}$ so that C fails to be $\mathcal{S}$-good, and we then say it is $\mathcal{S}$-bad. We refer to $\mathcal{S}$ as a *bad* sequence for C. This definition makes sense for a submanifold C of any dimension, such as a hypersurface. It is very difficult to determine when a submanifold is good and it is mainly in the ergodic setting where curves have been proved to be good, at least for a density one subsequence of eigenfunctions [**JJ2**]. A more general criterion for goodness is given in [**ToZ4**].

Obviously, if a infinite subsequence $\mathcal{S}$ of eigenfunctions vanishes on C then it is $\mathcal{S}$-bad. We call such a C a *nodal curve* (or hypersurface) We do not know of any other cases of bad submanifolds. A good case study would be to check if there are any ‘bad’ latitude circles of a surface of revolution. By separation of variables, this amounts to checking whether the radial parts of a sequence of eigenfunctions can decay faster than $e^{-C\lambda_j}$ at one single point of the unit interval. Of course, the sequence would correspond to a sequence of distinct radial Sturm Liouville problems.

In the case of Gaussian beams, the Cauchy data satisfies the bounds

$$\|CD(\varphi_h)|_H\|_{L^2(H)} \leq C_g e^{-c_g(H)\lambda_j} \tag{12.9}$$

along a latitude circle away from the center of the beam. That is, the Neumann data is just as small as the Dirichlet data. On the other hand, it is impossible for a closed hypersurface which bounds a domain to be 'bad' for the Cauchy data. This follows from boundary Carleman estimates for Cauchy data [**Bu1, V**].

Having discussed lower bound and degeneracy issues, we now turn to upper bounds.

12.2. L^2 upper bounds on Dirichlet or Neumann data of eigenfunctions

The first result is due to Burq-Gerard-Tzvetkov [**BuGT**]:

THEOREM 12.2. *Let M be a compact manifold without boundary and assume* $\dim M = 2$. *If γ is a unit-length geodesic, then*

$$\int_\gamma |\varphi_j(s)|^2 \, ds = \mathcal{O}(\lambda_j^{\frac{1}{2}}), \tag{12.10}$$

with ds denoting arc-length measure on γ.

If γ is a curve with strictly positive geodesic curvature, then

$$\int_\gamma |\varphi_j(s)|^2 \, ds = \mathcal{O}(\lambda_j^{\frac{1}{3}}). \tag{12.11}$$

As noted above, the L^2 restriction bound for closed geodesics is saturated by Gaussian beams around elliptic closed geodesics on surfaces of revolution. Non-positively curved surfaces do not possess such elliptic closed geodesics or Gaussian beams, and that suggests the estimate can be improved for them. Generally speaking, the improvements will not be better than by log factors due to the exponential growth of the geodesic flow. Sogge and the author improved the result on surfaces of non-positive curvature [**SoZ2**].

THEOREM 12.3. *Let M be a compact manifold without boundary and assume* $\dim M = 2$. *Let Π denote the space of unit length geodesics. Then given $\varepsilon > 0$ there is a $\lambda(\varepsilon) < \infty$ so that*

$$\sup_{\gamma \in \Pi} \left(\int_\gamma |\varphi_j|^2 \, ds \right)^{1/2} \leq \varepsilon \lambda_j^{\frac{1}{2}}, \qquad \lambda > \lambda(\varepsilon). \tag{12.12}$$

J. Toth [**To**] proved that for quantum integrable eigenfunctions and generic curves,

$$\int_\gamma |\varphi_j(s)|^2 \, ds = \mathcal{O}(\log \lambda_j). \tag{12.13}$$

For instance, zonal spherical harmonics have such bounds along meridians. Marshall gave a power law improvement in arithmetic cases [**Ma**].

The next result pertains to a manifolds of general dimension:

THEOREM 12.4. *Let M be a compact manifold without boundary. Let $H \subset M$ be a smooth codimension* 1 *submanifold. Then*

$$\|\varphi_j|_H\|_{L^2(H)} = \mathcal{O}(\lambda_j^{1/4})$$

and

$$\|hD_\nu \varphi_j|_H\|_{L^2(H)} = \mathcal{O}(\lambda_j^{1/4}).$$

The first estimate is from [**BuGT**], while the second estimate follows from the Rellich-type argument in [**HT**].

12.3. Cauchy data of Dirichlet eigenfunctions for manifolds with boundary

Let M be a compact Riemannian manifold with boundary ∂M. In this section we consider the same L^2 restriction bounds for Dirichlet or Neumann eigenfunctions on a manifold with boundary with $H = \partial M$. Since the eigenfunctions satisfy boundary conditions, these are actually restriction bounds on Cauchy data rather than just on Dirichlet or Neumann data.

A key hypothesis in the results is a 'no-trapping' hypothesis. The no-trapping hypothesis is that every geodesic intersects ∂M. An example is the upper hemisphere of a sphere or a convex domain in $\mathbb{R}^n$. An example with trapped geodesics is a Lorentz-Sinai billiard of higher genus, in which a disc is removed from a non-positively curved surface, in which case there exist closed geodesics which do not touch the boundary. Another is an equatorial annulus on a standard sphere, in which the equator does not touch the boundary. As mentioned above, Gaussian beams concentrate at the middle geodesic and are therefore not so large on the boundary. Denote by $\{u_j\}$ an orthonormal basis of Dirichlet eigenfunctions, and by $\partial u_j/\partial \nu$ its normal derivative on ∂M. In [**HT**] the following is proved:

THEOREM 12.5. *For any (M, g) with C^∞ boundary, the Dirichlet eigenfunctions satisfy*

$$\int_{\partial M} \|\partial u_j/\partial \nu\|^2 \, dS \leq C\lambda_j^2$$

and for any (M, g) satisfying the no-trapping hypothesis,

$$\int_{\partial M} \|\partial u_j/\partial \nu\|^2 \, dS \geq C\lambda_j^2.$$

The upper bound is not necessarily sharp and the lower bound is not necessarily true in general. In [**HT**], Example 6 of a hyperbolic cylinder with boundary,

$$\frac{c\lambda^2}{\log \lambda} \leq \|\partial u_j/\partial \nu\|^2_{L^2(\partial M)} \leq \frac{C\lambda^2}{\log \lambda}.$$

One may also study L^p norms of Cauchy data. Sup norms of Cauchy data are studied in [**SoZ1**] on manifolds with concave boundary.

12.3.1. Rellich identities.

We will not sketch the proofs of these results but mention that in [**HT**], the results follow from studying Rellich's identities

$$\int_M \langle u, [\Delta, A]u\rangle \, dV = \int_{\partial M} \frac{\partial u}{\partial \nu}(Au) - u\frac{\partial (Au)}{\partial \nu} \, dS \tag{12.14}$$

for eigenfunctions $-\Delta u = \lambda^2 u$. Here, $[A, B] = AB - BA$ is the commutator. To prove (12.14) one uses that $[\Delta, A] = [\Delta + \lambda^2, A]$ and applies Green's identity to integrate by parts the right side of

$$\int_M \langle u, [\Delta, A]u\rangle \, dV = \int_M \langle (\Delta + \lambda^2)u, Au\rangle - \langle u, (\Delta + \lambda^2)Au\rangle \, dV \tag{12.15}$$

$$= -\int_M \langle u, (\Delta + \lambda^2)Au\rangle \, dV. \tag{12.16}$$

To obtain bounds from Rellich's identity one chooses A so that the right side (resp. the left side) is a positive form in $\partial_\nu u$. In [**HT**] A is chosen to be $A = \chi(y)\partial_y$ (writing $r = y$ as the distance to the boundary) and χ is a cutoff that vanishes for $r \geq \delta$. Some of the results originally were proven in [**BLR**].

12.4. Restriction bounds for Neumann eigenfunctions

In [**BaHT**] some of the techniques and results of [**HT**] are extended to Neumann eigenfunctions. For convex Euclidean domains Ω, the use of Rellich's Lemma gives $\|w_j\|_{L^2(\partial\Omega)} \geq C$. In [**HT**] the authors show that it is not true on general manifolds with boundary that $\|w_j\|_{L^2(\partial M)} \geq C$ for Neumann eigenfunctions if there are trapped geodesics. In [**BaHT**] the authors show that indeed $\|w_j\|_{L^2(\partial M)} \geq C$ if there are no trapped geodesics. The complementary sharp upper bound is $\|w_j\|_{L^2(\partial M)} \leq C\mu_j^{1/3}$, as proved by Tataru [**T**]. Here, μ_j^2 is the Neumann eigenvalue.

The following theorem of [**CHT**] generalizes [**HT**] for boundary traces of Dirichlet eigenfunctions to general interior hypersurfaces.

Theorem 12.6. *Let (M, g) be a compact smooth manifold with boundary ∂M. If $H \subset M$ is a smooth, embedded, orientable separating hypersurface and $H \cap \partial M = \emptyset$, then for any orthonormal basis of eigenfunctions,*

$$\|\lambda_j^{-1}\partial_\nu \varphi_{\lambda_j}\|_{L^2(H)} = \mathcal{O}(1).$$

In [**HZ, BaH**] the authors provide some intuition towards the results. First, the analogue for Neumann eigenfunctions of Neumann data $\partial_\nu \varphi_j|_{\partial M}$ for Dirichlet eigenfunctions is not the Dirichlet data $|\varphi_j|_{\partial M}$ but $(1 - \lambda_j^{-2}\Delta_{\partial M})_+^{\frac{1}{2}}\varphi_j|_{\partial M}$ where $\Delta_{\partial M}$ is the *positive* boundary Laplacian. In [**BaH**], the authors explain that the w_j do not behave as uniformly as the Dirichlet traces. In [**BaH**] the authors prove

Proposition 12.7. *Let (M, g) be a compact smooth manifold with boundary ∂M, and let $\{v_j\}$ be an orthonormal basis of Neumann eigenfunctions with eigenvalues $-\mu_j^2$, and let $w_j = v_j|_{\partial M}$. Then,*

$$\|(1 - \lambda_j^{-2}\Delta_{\partial M})_+^{\frac{1}{2}} w_j\|_{L^2(\partial M)} \leq C.$$

A key ingredient in the proof is the estimate,

$$\int_{\partial\Omega} v_j d_n^2 v_j \leq C\mu_j^2.$$

12.5. Periods and Fourier coefficients of eigenfunctions on a closed geodesic

Instead of considering norms we may consider 'periods' or integrals of eigenfunctions over submanifolds. More generally we may consider 'Fourier coefficients' of restriction and Fourier series expansions (in a generalized sense) along submanifolds. In the theory of automorphic forms, Weyl sums of periods are called Kuznecov sum formulae or local trace formulae (usually in arithmetic situations). See for instance

[**K, Bru1, H1, H2, I, KMW, R, P2, Ts**] The author adopted this terminology for general submanifolds of Riemannian manifolds in [**Z**]. Among other applications, comparisons of periods and L^1 norms on submanifolds are used to prove existence of 'many' zeros of $\varphi_j|_H$.

In [**Z**] the following was proved:

THEOREM 12.8. *Let $H \subset M$ be a hypersurface of M. Let $f \in C^\infty(H)$. Then there exists a constant $c > 0$ such that,*

$$\sum_{\lambda_j < \lambda} \left| \int_H f \varphi_j \, dS \right|^2 = \pi \left| \int_H f \, dS \right|^2 \lambda + O_f(1).$$

Here, dS is any density on H.

By Weyl's law, it follows that the average size of the period is $\frac{1}{\sqrt{\lambda_j}}$ and that a density one subsequence of eigenfunctions satisfies period bounds $\frac{\log \lambda_j}{\sqrt{\lambda_j}}$. Moreover, one has unconditionally the estimate

COROLLARY 12.9. *There exists a constant $C = C_{M,H,f}$ so that*

$$\left| \int_Y \varphi_j \, d\nu \right| \leq C.$$

The leading order term and a remainder estimate of $\mathcal{O}(1)$ are given in [**Z**, Corollary 3.3] for any compact Riemannian manifold M^n of dimension n. For a general submanifold $Y \subset M$ of dimension d, and for any density dS on Y, the Corollary states that

$$\sum_{j:\lambda_j \leq \lambda} \left| \int_Y \varphi_j \, dS \right|^2 \simeq C_n \operatorname{Vol}(SN^*Y) \lambda^{n-d} + \mathcal{O}(\lambda^{n-d-1}). \tag{12.17}$$

The same result is valid for the Neumann data.

THEOREM 12.10. *Let $H \subset M$ be a hypersurface and let $f \in C^\infty(H)$. Then there exists a constant $c > 0$ such that,*

$$\sum_{\lambda_j < \lambda} \left| \lambda_j^{-1} \int_H f \partial_\nu \varphi_j \, dS \right|^2 = \pi \left| \int_H f dS \right|^2 \lambda + O_f(1).$$

It is quite useful to improve the remainder estimate. Experience with remainder estimates in Weyl laws suggests that one cannot do better than logarithmic improvements. In a recent series of articles, Sogge et al have provided logarithmic improvements. In [**SoXZ**] the following is proved:

THEOREM 12.11. *Let (M, g) be a negatively curved compact Riemannian surface. Let γ be a closed geodesic. Then for any smooth f on γ,*

$$\left| \int_\gamma f(s) \varphi_j(s) \, ds \right| \leq C_M L_\gamma \frac{1}{\sqrt{\log \lambda}}.$$

CONJECTURE 12.12. Let (M, g) be a negatively curved compact Riemannian manifold of any dimension n. Let γ be a closed geodesic. Then,

$$\sum_{j:\lambda_j \leq \lambda} \left| \int_\gamma f \varphi_j \, ds \right|^2 = \pi \left| \int_\gamma f \, ds \right|^2 \lambda^{n-1} + \mathcal{O}\left(\frac{\lambda^{n-2}}{\log \lambda} \right).$$

12.6. Kuznecov sum formula: Proofs of Theorems 12.8 and 12.10

To prove the results one studies the singularity expansion for the distribution

$$S_H(t) = \int_H \int_H E(t,q,q')\, dS(q') dS(q),$$

where where $H \subset M$ is a smooth submanifold and where $E(t) = \cos t\sqrt{-\Delta}$ is the even wave kernel The singularities of $S_H(t)$ correspond to trajectories of the geodesic flow which intersect H orthogonally at two distinct times, and to be singular at the length T of the trajectory between the two intersections. We refer to such trajectories as H-*orthogonal* geodesics.

For simplicity we assume $H \subset M$ is an interior hypersurface, but the proofs for $H = \partial M$ are not much more difficult. We let dS denote the standard surface area form on H and let $f \in C^\infty(\partial H)$. Define

$$S_f(t) = \int_H \int_H E(t,q,q') f(q) f(q')\, dS(q) dS(q') \tag{12.18}$$

$$= \sum_j \cos t\sqrt{\lambda_j} \left| \int_H f(q)\varphi_j(q)\, dS(q) \right|^2. \tag{12.19}$$

We then introduce a smooth cutoff $\rho \in \mathcal{S}(\mathbb{R})$ with $\operatorname{supp} \hat{\rho} \subset (-\varepsilon, \varepsilon)$, where $\hat{\rho}$ is the Fourier transform of ρ, and consider

$$S_f(\lambda, \rho) = \int_{\mathbb{R}} \hat{\rho}(t) S_f(t) e^{it\lambda}\, dt.$$

Let $H \Subset M$ be a smooth $(n-1)$-dimensional orientable hypersurface, we denote the Cauchy data of u_j on H by

$$\begin{cases} \text{Dirichlet data: } \omega_j = \varphi_j|_H, \\ \text{Neumann data: } \psi_j = \partial_\nu \varphi_j|_H. \end{cases}$$

PROPOSITION 12.13. *If* $\operatorname{supp} \hat{\rho}$ *is contained in a sufficiently small interval around* 0*, with* $\hat{\rho} \equiv 1$ *in a smaller interval, then* $S_f(\lambda, \rho)$ *is a semi-classical Lagrangian distribution whose asymptotic expansion for Neumann data is given by*

$$S_f(\lambda, \rho) = \frac{\pi}{2} \sum_j \rho(\lambda - \lambda_j) \lambda_j^{-2} |\langle \psi_j, f \rangle|^2 = \|f\|^2_{L^2(\partial M)} + o(1), \tag{12.20}$$

and for Dirichlet data by

$$S_f(\lambda, \rho) = \frac{\pi}{2} \sum_j \rho(\lambda - \lambda_j) |\langle \omega_j, f \rangle|^2 = \|f\|^2_{L^2(\partial M)} + o(1). \tag{12.21}$$

PROOF. There exists $\varepsilon_0 > 0$ so that the

$$\operatorname{sing\,supp} S_f(t) \cap (-\varepsilon_0, \varepsilon_0) = \{0\}. \tag{12.22}$$

This follows from propagation of singularities for the wave kernel and its restriction to the boundary. It is known that $\mathrm{WF}(E(t,x,y))$ on a smooth domain consists of geodesic trajectories. The pullback $E_H(t,q,q')$ to $H \times H$ forces the trajectories contributing to $\mathrm{WF}(E^b)$ to begin and end on H and integration over ∂M forces them to be orthogonal to H at both endpoints. Hence there exists $\varepsilon_0 > 0$ so that no trajectory starting orthogonally from a regular point of H can hit H again at any point. Thus the only singularity in this time interval is $t = 0$.

For $\varepsilon < \varepsilon_0$, we only need to determine the contribution of the main singularity of $S_f(t)$ at $t = 0$. We then express $S_f(t)$ and $S_f(\lambda, \rho)$ in terms of pushforwards under the submersion

$$\pi \colon \mathbb{R} \times H \times H \to \mathbb{R}, \quad \pi(t, q, q') = t.$$

By Lemma 12.22, for $t \in (-\varepsilon, \varepsilon)$

$$\text{WF}(S_f^{\varepsilon}(t)) = \{(0, \tau) \colon \pi^*(0, \tau) = (0, \tau, 0, 0) \in \text{WF}(\cos t\sqrt{-\Delta}(t, q, q'))\}. \tag{12.23}$$

These wave front elements correspond to the points $(0, \tau, \tau\nu_q, \tau\nu_q) \in T_0^*\mathbb{R} \times N_q^* H \times N_{q'}^* H$, i.e., where both covectors are co-normal to H. Indeed, the wave front set of $S_f(t)$ is the set

$$\{(t, \tau) \in T^*\mathbb{R} \colon \exists (x, \xi, y, \eta) \in C_t' \cap N^*(H) \times N^*(H)\}$$

in the support of the symbol. Thus, we may microlocalize to the normal directions. The wave kernel has a geometric optics Fourier integral representation, which implies that $S_f(t)$ is classical co-normal at $t = 0$. The order of the singularity and the principal symbol can be read off using the symbol calculus of Fourier integral operators. Since the terms of the Weyl-Kuznecov sums are positive, one can apply a Fourier Tauberian theorem to deduce the expansion and remainder estimate. □

Thus, the geodesic geometry controlling the remainder is that of geodesic orthogonal to H.

It would also useful to study high frequency Fourier coefficients

$$a_n(\lambda) := \int_0^L e^{-\frac{2\pi i n s}{L}} \varphi_\lambda(\gamma(s))\, ds$$

of an eigenfunction along a closed geodesic of length L. By the principle that angular momentum (tangential frequency) is smaller than energy (or total frequency), the Fourier coefficients are small if $|n| >> \lambda_j$. When n is bounded, the Kuznecov bounds above apply. But one may ask how the Fourier coefficients behave when $\frac{n}{\lambda} \to \tau$. Thus, the frequency of the Fourier coefficient is comparable with the frequency of the eigenfunction. Simple examples from surfaces of revolution show that these high frequency Fourier coefficients can be much larger than those where n is fixed as $\lambda \to \infty$, in effect because multiplication by $e^{-in\theta}$ cancels the oscillations of the restricted eigenfunction if the Fourier series of the restriction happens to be concentrated in one frequency. Obviously, $|a_n(\lambda)| \leq \int_\gamma |\varphi_\lambda| ds$ and true cancellation would reverse the inequality. However this frequency concentration happens only very rarely. In ergodic cases, one expects all Fourier coefficients with $|n| \leq \lambda_j$ to have the same size $\frac{1}{\sqrt{n}}$. For Fourier coefficients of frequency $n \simeq \alpha\lambda$, the normal bundle gets 'tilted' to angle α and in the extreme case $m \simeq \lambda_{m,0}$ it becomes the tangent bundle TY. Then the condition is $G^t TY = TY$ for some t, essentially that Y is totally geodesic.

12.7. Restricted Weyl laws

In preparation for QER theorems, we need to understand Weyl sums of restricted eigenfunctions.

We first review some notation and terminology concerning pseudo-differential operators. Homogeneous (or Kohn-Nirenberg) pseudo-differential operators are those with classical poly-homogeneous symbols $a(s,\sigma) \in C^\infty(T^*H)$,

$$a(s,\sigma) \sim \sum_{k=0}^{\infty} a_{-k}(s,\sigma) \text{ positive homogeneous of order } -k$$

as $|\sigma| \to \infty$ on T^*H. On the other hand, semi-classical pseudo-differential operators we mean h-quantizations of semi-classical symbols $a \in S^{0,0}(T^*H \times (0,h_0])$ of the form

$$a_h(s,\sigma) \sim \sum_{k=0}^{\infty} h^k a_{-k}(s,\sigma), \quad a_{-k} \in S^0_{1,0}(T^*H)$$

as in [**Zw, HZ, ToZ2**].

The following is from [**ToZ2, ToZ3**].

PROPOSITION 12.14. *Let H be a smooth interior hypersurface, and let $\mathrm{Op}_H(a) \in \Psi^0(H)$. Then,*

$$\lim_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{j:\lambda_j\le\lambda} \langle \mathrm{Op}_H(a)\gamma_H\varphi_j, \gamma_H\varphi_j\rangle = \frac{2}{\mathrm{Vol}(S^*M)} \int_{B^*H} a_0 \gamma_{B^*H}^{-1} \, |ds \wedge d\sigma|,$$

where $|ds \wedge d\sigma|$ is symplectic volume measure on B^*H, and a_0 is the principal symbol of $\mathrm{Op}_H(a)$.

Following a standard cosine Tauberian approach, the asymptotics arises from a singularity analysis of the dual sums

$$\begin{aligned}\sum_{j:\lambda_j\le\lambda} \langle \mathrm{Op}_H(a)\gamma_H\varphi_j, \gamma_H\varphi_j\rangle_{L^2(H)} e^{it\lambda_j} &= \sum_{j:\lambda_j\le\lambda} \langle \gamma_H^* \mathrm{Op}_H(a)\gamma_H\varphi_j, \varphi_j\rangle_{L^2(M)} e^{it\lambda_j} \\ &= \mathrm{Tr}\, U(t)\gamma_H^* \mathrm{Op}_H(a)\gamma_H, \end{aligned} \tag{12.24}$$

where $U(t) = \exp(it\sqrt{-\Delta})$. Modulo some technical complications, $\gamma_H^* \mathrm{Op}_H(a)\gamma_H$ is a homogeneous Fourier integral operator, and the singularities of (12.24) can be determined by a study of the canonical relations and symbols of the trace. This is not literally true due to the presence of 0 in its wave front relation for conormal vectors to H. Tangential covectors in T^*H also cause problems when we conjugate by wave group. We therefore introduce cutoff operators to cutoff away from T^*H and from $N^*H \times 0_{T^*M} \cup 0_{T^*M} \times N^*H$.

We then begin by proving the local Weyl law for a cutoff of $\gamma_H^* \mathrm{Op}_H(a)\gamma_H$ denoted by $(\gamma_H^* \mathrm{Op}_H(a)\gamma_H)_{\ge\varepsilon}$ on M, that is, we prove

$$\begin{aligned}\lim_{\lambda\to\infty} \frac{1}{N(\lambda)} \sum_{j:\lambda_j\le\lambda} &\langle (\gamma_H^* \mathrm{Op}_H(a)\gamma_H)_{\ge\varepsilon}\varphi_j, \varphi_j\rangle \\ &= \frac{2}{\mathrm{Vol}(S^*M)} \int_{B^*H} a_0(1-\chi_\varepsilon)\gamma_{B^*H}^{-1} \, |ds \wedge d\sigma|,\end{aligned} \tag{12.25}$$

which we study via the trace

$$\mathrm{Tr}\, U(t)(\gamma_H^* \mathrm{Op}_H(a)\gamma_H)_{\ge\varepsilon}. \tag{12.26}$$

In the following section we define this cutoff and make a systematic study of the composition in (12.26).

REMARK 12.15. When $a = V$ is a multiplication operator by a smooth function (extended smoothly to all of M), the local Weyl law on H follows from the pointwise Weyl asymptotic,

$$\frac{1}{N(\lambda)} \sum_{j:\lambda_j \leq \lambda} |\varphi_j(x)|^2 = (2\pi)^{-n} + O(\lambda^{-1}).$$

The pointwise asymptotics imply that the L^2-norm squares of $\gamma_H \varphi_j$ are bounded on average,

$$\frac{1}{N(\lambda)} \sum_{j:\lambda_j \leq \lambda} \|\gamma_H \varphi_j\|^2_{L^2(H)} = (2\pi)^{-n} \operatorname{Vol}^{n-1}(H) + O(\lambda^{-1}). \tag{12.27}$$

More generally, by [**Ho2**, Proposition 29.1.2], for any pseudo-differential operator B of order zero on M, the Schwartz kernel $K_B(t,x)$ of $U(t)B$ or $BU(t)$ on the diagonal $\Delta_{M\times M}$ is conormal to $\{t = 0\}$ with respect to $\mathbb{R} \times \Delta_{M\times M}$ and if

$$\frac{\partial A(\lambda, x)}{\partial \lambda} = \mathcal{F}_{t\to\lambda} K_B(t,x)$$

the $A(\lambda, x)$ is a symbol of order 0 on a manifold of dimension n with

$$A(\lambda, x) = \sum_{j:\lambda_j \leq \lambda} \varphi_j(x) A\varphi_j(x) \sim (2\pi)^{-n} \int_{|\xi|<\lambda} a_0 \, d\xi + O(\lambda^{n-1}) \tag{12.28}$$

in the case where $A = A^*$. There is an analogous statement for $AU(t)B$. Integrating (12.28) over H gives

$$\sum_{j:\lambda_j \leq \lambda} \langle \gamma_H A\varphi_j, \gamma_H B\varphi_j \rangle_{L^2(H)} \sim C_n \lambda^n \int_{B^*_H M} a_0 b_0 \gamma^{-1}_{B^*H} \, |ds \wedge d\sigma|. \tag{12.29}$$

This is almost the statement of the local Weyl law along H except that we wish to use $\operatorname{Op}_H(a)$ rather than a global pseudo-differential operator on M. It is not difficult to extend $\operatorname{Op}_H(a)$ to M and to prove the local Weyl law as above, but instead we give a longer and more complicated argument because the same techniques will be needed in the proof of the QER theorem where we will need the Weyl law for Fourier integral operators.

It should be said that the local Weyl law and the QER theorem are valid for semi-classical pseudo-differential operators on H [**ToZ2, ToZ3, DZ**]. The results for semi-classical pseudo-differential operators are simpler and more general. But we work with homogeneous pseudo-differential and Fourier integral operators to avoid introducing more machinery and because such techniques are somewhat more familiar to the PDE community. The semi-classical approach is used in [**HZ, ToZ2, DZ**] and the appendix to [**ToZ3**].

12.8. Relating matrix elements of restrictions to global matrix elements

Let (M, g) be a compact Riemannian manifold and let H be a compact embedded C^∞ submanifold. We denote by γ_H the restriction operator $\gamma_H f = f|_H : C(M) \to C(H)$ and by γ_H^* the adjoint of γ_H with respect to the inner product on $L^2(M, dV)$ where dV is the Riemannian volume form. Thus,

$$\gamma_H^* f = f\delta_H, \text{ since } \langle \gamma_H^* f, g\rangle = \int_H f g dS,$$

where dS is the surface measure on H induced by the ambient Riemannian metric. The fact that γ_H^* does not preserve smooth functions is due to the fact that $WF'_M(\gamma_H) = N^*H$. Thus, $\gamma_H^* Op_H(a)\gamma_H$ is not a Fourier integral operator with a homogeneous canonical relations because its wave front relation contains $N^*H \times 0_{T^*M} \cup 0_{T^*M} \times N^*H$ (where 0_{T^*M} is the zero section of T^*M). For this reason we need to introduce microlocal cutoffs. In the following, $\chi \in C_0^\infty(\mathbb{R};[0,1])$ is a cutoff function with $\chi(t) = 1$ for $|t| \leq 1$ and supp $\chi \subset [-2,2]$.

Define:

$$\begin{cases} V(t;a) := U(-t)\gamma_H^* \operatorname{Op}_H(a)\gamma_H U(t), \\ \bar{V}_T(a) := \dfrac{1}{T}\displaystyle\int_{-\infty}^{\infty} \chi(T^{-1}t)V(t;a)\,dt, \end{cases} \tag{12.30}$$

LEMMA 12.16. *For any* $a \in C_0^\infty(T^*H)$,

$$\langle \operatorname{Op}_H(a)\varphi_j|_H, \varphi_j|_H\rangle_{L^2(H)} = \langle \bar{V}_T(a)\varphi_j, \varphi_j\rangle_{L^2(M)}. \tag{12.31}$$

PROOF. This follows from the sequence of identities

$$\langle \operatorname{Op}_H(a)\varphi_j|_H, \varphi_j|_H\rangle_{L^2(H)} = \langle \operatorname{Op}_H(a)\gamma_H\varphi_j, \gamma_H\varphi_j\rangle_{L^2(H)} \tag{12.32}$$

$$= \langle \gamma_H^* \operatorname{Op}_H(a)\gamma_H U(t)\varphi_j, U(t)\varphi_j\rangle_{L^2(M)} \tag{12.33}$$

$$= \langle V(t;a)\varphi_j, \varphi_j\rangle_{L^2(M)} \tag{12.34}$$

$$= \langle \bar{V}_T(a)\varphi_j, \varphi_j\rangle_{L^2(M)}.$$

□

After cutting off from the tangential singular set $\Sigma_T \subset T^*M \times T^*M$ and the the conormal sets $N^*H \times 0_{T^*M}, 0_{T^*M} \times N^*H$, $\bar{V}_T(a)$ becomes a Fourier integral operator $\overline{V}_{T,\varepsilon}(a)$ with canonical relation given by

$$\begin{aligned} \operatorname{WF}(\overline{V}_{T,\varepsilon}(a)) := \{&(x,\xi,x',\xi') \in T^*M \times T^*M : \exists t \in (-T,T), \\ &\exp_x t\xi = \exp_{x'} t\xi' = s \in H,\ G^t(x,\xi)|_{T_sH} = G^t(x',\xi')|_{T_sH},\ |\xi| = |\xi'|\}. \end{aligned} \tag{12.35}$$

We now discuss the wave front aspects of these operators. We discuss the operator aspects and cutoffs more thoroughly in Section 12.25.

12.9. Geodesic geometry of hypersurfaces

Before proving Proposition 12.14 we review the symplectic and Riemannian geometric issues and introduce the cutoffs defining $(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon}$.

Let $H \subset M$ be a hypersurface in a Riemannian manifold (M,g). We consider two hypersurfaces of T^*M, the set T_H^*M of covectors with footpoint on H and the unit cotangent bundle S^*M of g.

Let $\pi\colon T^*M \to M$ be the natural projection. We identity $\pi^*y_n = y_n$ as functions on T^*M. Then $f = y_n = 0$ is the defining function of T_H^*M. The hypersurface S^*M is defined by $g = |\xi| = 1$, the metric norm function. It is clear that df, dg are linearly independent, so that T_H^*M, S^*M are a pair of transversal hypersurfaces in T^*M.

In general, let $F, G \subset T^*M$ be two transversely intersecting hypersurfaces, and let f, g denote defining function of F, G, so that $f = 0$ on F, $g = 0$ on G and df, dg are linearly independent. Then their intersection $J = F \cap G$ is a submanifold of codimension two. The intersection fails to be symplectic along the set $K = \{x \in J : \{f,g\}(x) = 0\}$.

When $F = T^*_H M$, $G = S^*M$, $J = S^*_H M$, $K = S^*H$, the Hamilton vector field of y_n equals $\frac{\partial}{\partial \xi_n}$ and its orbits are vertical curves of the form $(s, 0, \sigma, \xi_{n0} + t)$; they define the characteristic foliation of $T^*_H M$. The hypersurface S^*M is defined by $g = |\xi| = 1$, the metric norm function, and its characteristic foliation is given by orbits of the homogeneous geodesic flow G^t. Evidently,

$$\{x_n, |\xi|_g\} = \frac{\partial}{\partial \xi_n} |\xi|_g = |\xi|_g^{-1} \sum_j g^{jn}(x)\xi_j = \xi_n \text{ on } S^*_H M,$$

so $\{x_n, |\xi|_g\} = 0$ defines S^*H. Equivalently,

LEMMA 12.17. *S^*H is the set of points of $S^*_H M$ where $S^*_H M$ fails to be transverse to G^t, i.e. where the Hamilton vector field H_g of $g = |\xi|$ is tangent to $S^*_H M$.*

Indeed, this happens when $H_g(f) = df(H_g) = 0$. One may also see it in Riemannian terms as follows: the generator H_g is the horizontal lift η^h of η to (q, η) with respect to the Riemannian connection on S^*M, where we freely identify covectors and vectors by the metric. Lack of transversality occurs when η^h is tangent to $T_{(q,\eta)}(S^*_H M)$. The latter is the kernel of dy_n. But $dy_n(\eta^h) = dy_n(\eta) = 0$ if and only if $\eta \in TH$. We also note that for any hypersurface H, $dy_n, d\xi_n, d|\xi|_g$ are linearly independent.

Two closely related restriction maps will be important. The first is the linear restriction map $\pi_H : T^*_H M \to T^*H$. If we orthogonally decompose $T^*_H M = T^*H \oplus N^*H$, then π_H is the orthogonal projection with respect to this decomposition. It is a fiber bundle with fiber $N^*_s H$. On the other hand, we consider the restriction map on $S^*_H M \to B^*H$. For $s \in H$, the orthogonal projection map $\gamma_{T^*_H M} : S^*_s M \to B^*_s H$ is the standard projection of a sphere to a ball, which has a fold singularity along the 'equator.'

In our setting, the full restriction map $\gamma_H : S^*_H M \to B^*H$ is a folding map with fixed point set S^*H and involution given by the reflection map r_H (12.111). When H is orientable, S^*H divides $S^*_H M$ into two connected components, and the involution on $\mathbb{R} \times S^*_H M$ is given by $r(t, x, \xi) = (t, r_H(x, \xi))$. Indeed, as observed above, this is true for each $x \in H$, and $D\gamma_H$ is the identity in the directions tangent to H. The reflection bundle at $(s, \sigma) \in S^*H$ is spanned by the Hamilton vector field $H_{y_n} = \frac{\partial}{\partial \xi_n}$.

We also need the following variant.

LEMMA 12.18. *The maps $G : \mathbb{R} \times S^*_H M \to S^*M$ defined by $(t, x, \xi) \to G^t(x, \xi)$ (resp. $G : \mathbb{R} \times T^*_H M - 0 \to T^*M - 0$ defined by $(t, x, \xi) \to G^t(x, \xi)$) are folding maps with folds along $\mathbb{R} \times S^*H$ (resp. $\mathbb{R} \times T^*H$).*

PROOF. In both cases, the spaces are of equal dimension, so the maps are local diffeomorphisms whenever the derivatives are injective. By Lemma 12.17, $DG(\frac{\partial}{\partial t} - H_g) = 0$ on $T_{(t,x,\xi)}(\mathbb{R} \times S^*_H M)$ if $(x, \xi) \in S^*H$, and these are the only vectors in its kernel. Indeed, suppose $X \in T_{x,\xi} S^*_H M$. We note that $DG_{(t,x,\xi)} \frac{\partial}{\partial t} = H_g(G^t(x, \xi))$ and $DG_{t,x,\xi} X$ (as t varies) is a Jacobi field along the geodesic $\gamma_{x,\xi}(t) = \pi G^t(x, \xi)$. Since G^t is a diffeomorphism, the only possible elements of the kernel have the form $\frac{\partial}{\partial t} + X$. If $H_g + D_{x,\xi} G^t X = 0$, then $X = -H_g$, i.e. it is the tangential Jacobi field $\dot{\gamma}$. But by Lemma 12.17, this implies $(x, \xi) \in S^*H$ and $X \in T(S^*H)$.

Since G^t is homogeneous on $T^*M = 0$ the same statements are true on $\mathbb{R} \times T^*_H M$. $\square$

12.10. Tangential cutoffs

To cut off tangential directions to H, we define

$$\begin{cases} (S^*_H M)_{\leq \varepsilon} = \{(x,\xi) \in S^*_H M \colon |\langle \xi, \eta \rangle| \leq \varepsilon, \forall \eta \in S^*_x H\} \\ (S^*_H M)_{\geq \varepsilon} = \{(x,\xi) \in S^*_H M : |\langle \xi, \eta \rangle| \geq \varepsilon, \forall \eta \in S^*_x H\} \end{cases} \tag{12.36}$$

i.e., the covectors which make an angle $\leq \varepsilon$ (resp. $\geq \varepsilon$) with H. We homogenize by defining

$$\begin{cases} (T^*_H M)_{\leq \varepsilon} = \left\{ (x,\xi) \in T^*_H M : \dfrac{\xi}{|\xi|} \in (S^*_H M)_{\leq \varepsilon} \right\}, \\ (T^*_H M)_{\geq \varepsilon} = \left\{ (x,\xi) \in T^*_H M : \dfrac{\xi}{|\xi|} \in (S^*_H M)_{\geq \varepsilon} \right\}. \end{cases} \tag{12.37}$$

Let $x = (s, x_n)$ be Fermi normal coordinates along H, i.e., let $x = \exp_{q_H(s)} x_n \nu_s$ where $s \mapsto q_H(s) \in H$ denotes a local parametrization of H. Then $H = \{x_n = 0\}$. Let $\xi = (\sigma, \xi_n) \in T^*M$ denote the corresponding symplectically dual fiber coordinates.

Let $\psi_\varepsilon \in C_0^\infty(\mathbb{R})$, $\psi_\varepsilon \equiv 1$ on $[-\varepsilon/2, \varepsilon/2]$ and $\psi_\varepsilon \equiv 0$ on $(-\infty, -\varepsilon] \cup [\varepsilon, \infty)$. In Fermi normal coordinates, we may take the cutoff $\chi_\varepsilon^{(tan)} \in C^\infty(T^*M)$ (see also (i)-(iii) in the Introduction) to be

$$\chi_\varepsilon^{(tan)}(s, y_n, \sigma, \eta_n) = \psi_\varepsilon \left(\frac{|\eta_n|^2}{|\sigma|^2 + |\eta_n|^2} \right) \cdot \psi_\varepsilon(y_n), \tag{12.38}$$

which is equal to one in a conic neighborhood of $T^*H = \{y_n = \eta_n = 0\}$. We further introduce a homogeneous cutoff $\chi_\varepsilon^{(n)} \in C^\infty(T^*M)$ given by

$$\chi_\varepsilon^{(n)}(s, y_n, \sigma, \eta_n) = \psi_\varepsilon \left(\frac{|\sigma|^2}{|\sigma|^2 + |\eta_n|^2} \right) \cdot \psi_\varepsilon(y_n) \tag{12.39}$$

which equals one on a conic neighborhood of $N^*H = \{y_n = \sigma = 0\}$. More precisely, we multiply (12.38) and (12.39) by a bump function $\psi(\xi)$ which vanishes identically near the zero section.

We also put

$$\chi_\varepsilon := \chi_\varepsilon^{(tan)} + \chi_\varepsilon^{(n)} \tag{12.40}$$

and denote the corresponding pseudo-differential operator by $\chi_\varepsilon(x, D)$ or by $\mathrm{Op}(\chi_\varepsilon)$.

12.11. Canonical relation of γ_H

As mentioned above, γ_H fails to be a homogeneous Fourier integral operator due to 0-components in its wave front set. The In this section we go through the calculation.

We define

$$\begin{cases} C_H = \{(s, \xi, s, \xi') \in T^*_H M \times T^*_H M : s \in H, \, \xi|_{TH} = \xi'|_{TH}\}, \\ \hat{C}_H = \{(s, \xi, s, \xi') \in T^*_H M \times T^*_H M : s \in H, \, \xi|_{TH} = \xi'|_{TH}, \quad |\xi| = |\xi'|\}, \\ S\hat{C}_H = \{(s, \xi, s, \xi') \in T^*_H M \times T^*_H M : s \in H, \, \xi|_{TH} = \xi'|_{TH}, \quad |\xi| = |\xi'| = 1\} \end{cases} \tag{12.41}$$

As above, SF denotes the unit vectors in any set F. Thus, $\hat{C}_H = \mathbb{R}_+ S\hat{C}_H$. As will be seen below, C_H is the canonical relation of $\gamma_H^* \operatorname{Op}_H(a)\gamma_H$, and $\hat{C}_H$ arises in the canonical relation of $\overline{V}_{T,\varepsilon}(a)$.

We recall that the fiber product of two fiber bundles $\pi : X \to Z$ and $\rho : Y \to Z$ is the submanifold $X \times_Z Y \subset X \times Y$ equal to $(\{(x, y) : \pi(x) = \rho(y)\}$. We apply the same terminology with $X = Y = S_H^* M$, $Z = B^* H$ and $\pi, \rho = \gamma_H$, but as just observed, the restriction map is not a fiber bundle projection but a folding map.

Lemma 12.19. *We have:*

- $C_H \simeq T_H^* M \times_{T^*H} T_H^* M$ *is the fiber square of* $T_H^* M$ *with respect to the restriction map* $\gamma_H : T_H^* M \to T^* H$. *It is an embedded Lagrangian submanifold of* $T^*M \times T^*M$.
- $\hat{C}_H := \mathbb{R} S\hat{C}_H \simeq T_H^* M \times_{S^*H} T_H^* M$ *is an immersed homogeneous isotropic submanifold of dimension* $2n - 1$ *with transverse crossings on the self-intersection locus* $\mathbb{R}_+ \Delta_{S^*H \times S^*H} = \Delta_{T^*H \times T^*H}$. *Also,* $\hat{C}_H \cap (T^*H \times T^*H) = \Delta_{T^*H \cap T^*H}$.
- $S\hat{C}_H \simeq S_H^* M \times_{S^*H} S_H^* M$ *is the 'fiber square' of* $S_H^* M$ *with respect to the (folding) restriction map* $\gamma_H : S_H^* M \to S^* H$. *It is an immersed isotropic submanifold of dimension* $2n - 2$ *with transversal crossings on the self-intersection locus* $\Delta_{S^*H \times S^*H}$.

Proof. The defining equations of $C_H \subset T_H^* M \times T_H^* M$ are given by equating the map $(v, w) \to v|_{TH} - w|_{TH} \in T^*H$ to zero. This map is a submersion. Suppressing the $s \in H$ variable, it is just the map $(\sigma, y_n, \sigma', y_n') \to \sigma - \sigma'$ with $\sigma, \sigma' \in \mathbb{R}^{n-1}, y_n \in \mathbb{R}$. Thus, the zero set is a regular level set, hence an embedded submanifold of codimension $n = \dim M$.

We observe that $S\hat{C}_H$ is the union $S\hat{C}_H = \operatorname{graph}(\operatorname{Id}) \cup \operatorname{graph}(r_H)$ of the identity and reflection maps. The graphs intersect transversely along the diagonal $\Delta_{S^*H \times S^*H} \subset S^*H \times S^*H$, since the tangent space to the identity graph is the diagonal and the tangent space to the reflection map is the 'anti-diagonal' $(v, -v) \in T(S^*H \times S^*H)$. That is, the equation $\pi_H(\zeta, \zeta') := \gamma_H(\zeta) - \gamma_H(\zeta') = 0$ in $S_H^* M \times S_H^* M$ defines a submanifold of codimension $n - 1$ on the dense open set where $D_\zeta \gamma_H, D_{\zeta'} \gamma_H$ spans TB^*H. Suppressing the variable along H, the singularities at each $x \in H$ are those of the map $\pi : S^{n-1} \times S^{n-1} \to \mathbb{R}^{n-1}, \pi(\sigma, y_n; \sigma', y_n') = \sigma - \sigma'$, where $(\sigma, y_n) \in \mathbb{R}^{n-1} \times \mathbb{R}, |\sigma|^2 + y_n^2 = 1$. Thus, $y_n = \pm\sqrt{1 - |\sigma|^2}$ and $\pi^{-1}(0) = \{\sigma, y_n, \sigma, y_n)\} \cup \{(\sigma, y_n, \sigma, -y_n)\}$. Here, we fix $s \in H$ and identify $T_s^* M \simeq \mathbb{R}^n, T_s^* H \simeq \mathbb{R}^{n-1}$.

Since $\mathbb{R}_+ S\hat{C}_H$ is the homogenization, we only need to homogenize the results for $S\hat{C}_H$. In more detail, we again fix x and consider the map $\pi(r, s, y_n, s', y_n') = r(s - s')$ from $\mathbb{R}_+ \times S^{n-1} \times S^{n-1} \to \mathbb{R}^{n-1}$. The zero set is again defined by $s = s'$. The radial tangent direction is in the kernel of $D\pi$ along $\pi^{-1}(0)$. Finally, we note that if $(x, \xi, x, \xi') \in \hat{C}_H \cap (T^*H \times T^*H)$, then $\xi = \xi'$. □

12.12. The canonical relation of $\gamma_H^* \operatorname{Op}_H(a)\gamma_H$

In Fermi normal coordinates,

(12.42)
$$\operatorname{Op}_H(a)\gamma_H(s; x_n, s') = C_n \int e^{i\langle s - s', \sigma\rangle - i x_n \xi_n} a(s, \sigma)(1 - \chi_\varepsilon(s', x_n, \sigma', \xi_n))\, d\xi_n d\sigma.$$

The phase $\varphi(s, x_n, s', \xi_n, \sigma) = \langle s - s', \sigma\rangle - x_n\xi_n$ is linear and non-degenerate. The number of phase variables is $N = d$ and $n = 2d - 1$, where $d = \dim M$, so $\frac{N}{2} - \frac{n}{4} = \frac{1}{4}$. Then $C_\varphi = \{(s, x_n, s', \sigma, \xi_n) : s = s', x_n = 0\}$ and $\iota_\varphi(s, 0, s, \sigma, \xi_n) \to (s, \sigma, s, \sigma, 0, \xi_n)$.

We recall ([**Ho1**, Theorem 8.2.14]) that the general composition of wave front sets has the following form. Let $A\colon C_0^\infty(Y) \to \mathcal{D}'(X)$, $B\colon C_0^\infty(Z) \to \mathcal{D}'(Y)$. Then if $\mathrm{WF}'_Y(A) \cap \mathrm{WF}'_Y(B) = \emptyset$, then $A \circ B\colon C_0^\infty(Z) \to \mathcal{D}'(X)$ and

$$\mathrm{WF}'(A \circ B) \subset \mathrm{WF}'(A) \circ \mathrm{WF}'(B) \cup (\mathrm{WF}'_X(A) \otimes 0_{T^*Z}) \cup (0_{T^*X} \times \mathrm{WF}'_Z(B)).$$

Thus,

$$\begin{aligned} &\mathrm{WF}'(\gamma_H^* \operatorname{Op}_H(a)\gamma_H) \subset \{(q, \xi, q, \xi') : \xi|_H = \xi'|_H), (q, \xi), (q, \xi') \in T_H^*M - 0\} && (12.43)\\ &\qquad \cup \{(q, \nu, q, 0) : (q, \nu) \in N^*H - 0\} && (12.44)\\ &\qquad \cup \{(q, 0, q, \nu) : \nu \in N^*H - 0\}. && (12.45)\end{aligned}$$

With the cutoff $(1 - \chi_{\frac{\varepsilon}{2}})$ on the left and $(1 - \chi_\varepsilon)$ on the right of $\gamma_H^* \operatorname{Op}_H(a)\gamma_H$, the last two sets are erased. Observing that $(1 - \chi_{\frac{\varepsilon}{2}})(1 - \chi_\varepsilon) = 1 - \chi_\varepsilon$, we have proved

LEMMA 12.20. *If $a \in S^0_{cl}(T^*H)$, then $(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon} \in I^{\frac{1}{2}}(M \times M, C_H)$. In the Fermi normal coordinates the symbol is given by*

$$\sigma_{(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon}}(s, \sigma, \eta_n, \eta'_n) = (1 - \chi_\varepsilon) a(s, \xi|_{TH}) |\Omega|^{\frac{1}{2}},$$

where $\Omega = |ds \wedge d\sigma \wedge d\eta_n \wedge d\eta'_n|$.

PROOF. In Lemma 12.19, we showed that C_H is an embedded Lagrangian submanifold of $T^*M \times T^*M$. The proof shows that the composition of $\Lambda_H^* \circ \Lambda_H$ is transversal. Since the order of γ_H^* equals that of γ_H and the orders add under transversal composition, the order of $(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon}$ is $\frac{1}{2}$. Hence, for any homogeneous pseudo-differential operator $Op_H(a)$ on H,

$$(\gamma_H^* \operatorname{Op}(a)\gamma_H)_{\geq\varepsilon} \in I^{\frac{1}{2}}(M \times M, C_H). \tag{12.46}$$

Next we compute its principal symbol. By Lemma 12.19, C_H is the fiber product $T_H^*M \times_{T^*H} T_H^*M$, hence it carries a canonical half-density (associated to the fiber map). As discussed in [**GSt**, p.350], on any fiber product $A \times_B C$, half-densities on A, C together with a negative density on B induce a half density on the fiber product. In our setting, the canonical half-density on T_H^*M is given by the square root of the quotient $\frac{\Omega_{T^*M}}{dy_n} = ds \wedge d\sigma \wedge d\eta_n$ of the symplectic volume density on T^*M by the differential of the defining function y_n of T_H^*M. We also have a canonical density $|ds \wedge d\sigma|$ on T^*H, which induces a canonical -1-density. The induced half-density on C_H is then $|ds \wedge d\sigma \wedge d\eta_n \wedge d\eta'_n|^{\frac{1}{2}}$.

We compute the principal symbol and order using the special oscillatory integral formula,

$$\begin{aligned}\gamma_H^* \operatorname{Op}(a)\gamma_H(s, x_n; s', x'_n) &= C_n \delta_0(x_n) \int e^{i\langle s - s', \sigma\rangle - ix'_n\xi'_n} a(s, \sigma)\, d\xi'_n d\sigma\\ &= C_n \int_{\mathbb{R}^n \times \mathbb{R} \times \mathbb{R}} e^{i\langle s - s', \sigma\rangle + ix_n\xi_n - ix'_n\xi'_n} a(s, \sigma)\, d\xi_n d\sigma d\xi'_n. \qquad (12.47)\end{aligned}$$

If we compose on left and right by $(1 - \chi_\varepsilon)$ and $(1 - \chi_{\frac{\varepsilon}{2}})$ respectively then we further obtain factor of $(1 - \chi_\varepsilon(s, x_n, \sigma, \xi_n))$ under the integral. The phase is

$\varphi(s, x_n, s', x_n', \xi_n, \xi_n', \sigma) = \langle s - s', \sigma \rangle + x_n \xi_n - x_n' \xi_n'$ with phase variables (ξ_n, ξ_n', σ), and

$$C_\varphi = \{(s, x_n, s', x_n', \sigma, \xi_n, \xi_n') : s = s', x_n = 0, x_n' = 0\}.$$

Also,

$$\iota_\varphi(s, 0, s, 0, \sigma, \xi_n, \xi_n') = (s, \sigma, s, \sigma, 0, \xi_n, 0, \xi_n') \in T^*M \times T^*M.$$

Thus, $(s, \sigma, \xi_n, \xi_n')$ define coordinates on C_φ.

The delta-function on C_φ is given by

$$d_{C_\varphi} = \frac{|ds \wedge d\sigma \wedge d\xi_n \wedge d\xi_n'|}{|D(s, \sigma, \xi_n, \xi_n', \varphi'_{\xi_n}, \varphi'_{\xi_n'} \varphi'_\sigma)/D(s, s', \sigma, x_n, x_n', \xi_n, \xi_n')|}.$$

Since

$$|D(\varphi'_{\xi_n}, \varphi'_{\xi_n'}, \varphi'_\sigma)/D(s', x_n, x_n')| = 1,$$

the lemma follows. □

12.13. The canonical relation $\Gamma^* \circ C_H \circ \Gamma$

It is well known (see [**Ho3**]) that $U(t) \in I^{-\frac{1}{4}}(\mathbb{R} \times M \times M, \Gamma)$, with $\Gamma = \{(t, \sigma, x, \xi, G^t(x, \xi)) : \sigma + |\xi| = 0\}$. As in [**DG**], the half density symbol of $U(t, x, y)$ is the canonical volume half density $\sigma_{U(t,x,y)} = |dt \otimes dx \wedge d\xi|^{\frac{1}{2}}$ on Γ.

Here,

$$\text{(12.48)} \quad \Gamma^* \circ C_H \circ \Gamma = \{(t', -|\xi'|, t, |\xi|, G^{t'}(s, \xi'), G^t(s, \xi)) \\ \in T^*\mathbb{R} \times T^*\mathbb{R} \times T^*M \times T^*M \text{ such that } \xi|_{TH} = \xi'|_{TH}\}.$$

Lemma 12.21. *The (set-theoretic) composition $\Gamma^* \circ C_H \circ \Gamma$ is transversal, and $\Gamma^* \circ C_H \circ \Gamma \subset T^*\mathbb{R} \times T^*\mathbb{R} \times T^*M \times T^*M$ is the Lagrangian submanifold parametrized by the embedding*

$$\text{(12.49)} \quad \iota_{\Gamma^* C_H \Gamma} : \mathbb{R} \times \mathbb{R} \times T^*_H M \to T^*(\mathbb{R} \times \mathbb{R} \times M \times M),$$

$$\text{(12.50)} \quad \iota_{\Gamma^* C_H \Gamma}(t', t, s, \xi, \xi') = (t', -|\xi'|, t, |\xi|, G^{t'}(s, \xi), G^t(s, \xi')), \quad \xi|_{TH} = \xi'|_{TH}.$$

Proof. This follows from the following observation: if $\chi : T^*M - 0 \to T^*M - 0$ is a homogeneous canonical transformation and $\Gamma_\chi \subset T^*M \times T^*M$ is its graph, and if $\Lambda \subset T^*M \times T^*M$ is any homogeneous Lagrangian submanifold with no elements of the form $(0, \lambda_2)$, the $\Gamma_\chi \circ \Lambda$ is a transversal composition with composed relation $\{(\chi(\lambda_1), \lambda_2) : (\lambda_1, \lambda_2) \in \Lambda\}$. The condition that $\lambda_1 \neq 0$ is so that $\chi(\lambda_1)$ is well-defined.

We recall that transversality refers to the intersection

$$\Gamma_\chi \times \Lambda \cap T^*M \times \Delta_{T^*M \times T^*M} \times T^*M.$$

Now, the tangent space at any intersection point to $T^*M \times \Delta_{T^*M \times T^*M} \times T^*M$ contains all vectors of the form $(v, 0, 0, 0)$ and $(0, 0, 0, v')$ with $v, v' \in T(T^*M)$. Hence to prove transversality it suffices to fill in the middle two components. The diagonal $T\Delta_{T^*M \times T^*M}$ contributes all tangent vectors of the form (w, w). On the other hand, the middle components of $\Gamma_\chi \times \Lambda$ have the form $(w, \delta\lambda_1)$ where $w \in T(T^*M)$ is arbitrary. The sum of such vectors with the diagonal contains all vectors of the form $(w + v, \delta\lambda + v')$ and therefore clearly spans the middle $T(T^*M \times T^*M)$.

We apply this observation in two steps. First, we compose

$$C_H \circ \Gamma = \{(s, \xi', G^t(s,\xi), t, -|\xi|) : (s,\xi) \in T^*_H M, \xi|_{TH} = \xi'|_{TH}\} \tag{12.51}$$

$$\subset T^*M \times T^*M \times T^*\mathbb{R}\backslash 0. \tag{12.52}$$

By the first part of Lemma 12.19, C_H is a Lagrangian submanifold, so the argument about graphs applies to show that this composition is transversal (including the innocuous $T^*\mathbb{R}$ factor.) We then apply the same argument to the left composition with Γ. It is straightforward to determine the composite as stated above. □

12.14. The pullback $\Gamma_H := \Delta_t^* \Gamma^* \circ C_H \circ \Gamma$

We now consider the pullback of $\Gamma^* \circ C_H \circ \Gamma$ under the time diagonal embedding $\Delta_t(t, x, y) = (t, t, x, y) \colon \mathbb{R} \times M \times M \to \mathbb{R} \times \mathbb{R} \times M \times M$. We define

$$(G^t \times G^t)(C_H) = \{(G^t(s,\xi), G^t(s,\xi')) : (s, \xi, s, \xi') \in C_H\}, \tag{12.53}$$

and

$$\Gamma_H := \{t, |\xi| - |\xi'|, (G^t(s,\xi), G^t(s,\xi')) \in T^*\mathbb{R} \times (G^t \times G^t)(C_H))\}. \tag{12.54}$$

Lemma 12.22. *The map Δ_t is transversal to $(\Gamma^* \circ C_H \circ \Gamma)$, hence*

$$\Delta_t^*(\Gamma^* \circ C_H \circ \Gamma) = \Gamma_H$$

is a smoothly embedded canonical relation under the Lagrange embedding

$$\iota_{\Gamma_H} \colon \mathbb{R} \times T^*_H M \to T^*(\mathbb{R} \times M \times M), \tag{12.55}$$

$$\iota_{\Gamma_H}(t, s, \xi, \xi') = (t, |\xi| - |\xi'|, G^t(s,\xi), G^t(s,\xi')), \quad \xi|_{TH} = \xi'|_{TH}. \tag{12.56}$$

Proof. The explicit formula for the composition is simple to verify. We recall that a map $f : X \to Y$ is said to be transversal to $W \subset T^*Y$ if $df^*\eta \neq 0$ for any $\eta \in W$. By [**GSt**, Proposition 4.1], if $f \colon X \to Y$ is smooth and $\Lambda \subset T^*Y$ is Lagrangian, and if $f : X \to Y$ and $\pi|_\Lambda \colon \Lambda \to Y$ are transverse then $f^*\Lambda \subset T^*X$ is Lagrangian. It is clear from the explicit formula for the pullback that transversality holds.

Since $G^t \times G^t$ is a homogeneous diffeomorphism, $G^t \times G^t(C_H)$ is a smooth embedded manifold, and the map $\iota_{t,C_H} : T^*_H M \times_{T^*H} T^*_H M \to G^t \times G^t(C_H) \subset T^*M \times T^*M$ is a smooth embedding. □

12.15. The pushforward $\pi_{t*}\Delta_t^* \Gamma^* \circ C_H \circ \Gamma$

We now consider the map $\pi_t \colon \mathbb{R} \times M \times M \to M \times M$ and push forward the canonical relation $\Delta_t^* \Gamma^* \circ C_H \circ \Gamma$. We recall that $\overline{V}_{T,\varepsilon}(a)$ is cutoff in time (by χ_T) to $|t| \leq T$ and thus (12.125),

$$\Delta_{T^*M \times T^*M} \cup \Gamma_T = \bigcup_{|t| \leq T} \{(G^t(s,\xi), G^t(s,\xi')) : (s, \xi, s, \xi') \in C_H, |\xi| = |\xi'|\} \tag{12.57}$$

$$= \bigcup_{|t| \leq T} (G^t \times G^t)\hat{C}_H. \tag{12.58}$$

is the proper pushforward

$$\Gamma_T = \pi_{t*}\Delta_t^* \Gamma^* \circ C_H \circ \Gamma, \quad \pi_t : [-T, T] \times M \times M \to M \times M. \tag{12.59}$$

Of course, the sharp cutoff to $[-T,T]$ puts a boundary in Γ_T, but it causes no problems since all of our operators are smooth in a neighborhood of the boundary and since we use the smooth cutoff $\chi(\frac{t}{T})$ in the definition of $\bar{V}_T$.

We recall that the pushforward of $\Lambda \subset T^*X$ under a map $f\colon X \to Y$ is defined by $f_*\Lambda = \{(y,\eta) : y = f(x),\, (x, f^*\eta) \in \Lambda)\}$. As discussed in [**GSt**, Proposition 4.2], if $f\colon X \to Y$ is a smooth map of constant rank and $H^*(X)$ is the bundle of horizontal covectors, and if $\Lambda \cap H^*(X)$ is transversal then $f^*(\Lambda)$ is a Lagrangian submanifold. Here, $H^*(X) = f^*T^*Y$ is the set of covectors which annihilate the tangent space to the fibers.

In our setting, $\pi_t^* T^*(M\times M)$ is the co-horizontal space $H^* \subset T^*(\mathbb{R}\times M\times M)$ which is co-normal to the fibers of π_t, i.e., its elements have the form $(t,0,x,\xi,y,\eta)$. Let $\tau\colon T^*\mathbb{R}\times T^*M\times T^*M \to \mathbb{R}$ be the projection onto the second component of $T^*\mathbb{R} = \{(t,\tau)\}$. Thus,

$$\begin{aligned}\Gamma_H \cap H^*(M\times M) &= \Delta_t^*\Gamma^* \circ C_H \circ \Gamma \cap H^*(M\times M)\\ &= \{z \in \Delta_t^*\Gamma^* \circ C_H \circ \Gamma : \tau(z) = 0\},\end{aligned} \tag{12.60}$$

and the pushforward relation is

$$\begin{aligned}\Gamma_T &= \bigcup_{|t|\le T}\{(G^t(s,\xi), G^t(s,\xi')) : (s,\xi,s,\xi') \in C_H, |\xi| = |\xi'|\}\\ &= \bigcup_{|t|\le T}(G^t\times G^t)\hat{C}_H.\end{aligned} \tag{12.61}$$

Note that $\bigcup_{|t|\le T} G^t(T_x^*M)$ projects (for small t) to M to the ball of radius t around x.

Proposition 12.23. *For any $\varepsilon > 0$, $\Delta_{T^*M\times T^*M} \cup \Gamma_{T,\varepsilon} \subset T^*M\times T^*M$ is smoothly immersed homogeneous canonical relation.*

By Lemma 12.19, (12.57) is the flow-out of an immersed Lagrangian submanifold with transversal crossings on $\mathbb{R}_+\Delta_{S^*H\times S^*H}$. Equivalently, the pushforward relation is parametrized by the Lagrange mapping

$$\iota : \mathbb{R}\times \hat{C}_H \to T^*M\times T^*M\colon (t,s,\xi,\xi') \mapsto (G^t(s,\xi), G^t(s,\xi')). \tag{12.62}$$

The following Lemma is the final step in the proof of Proposition 12.23, and indeed is more precise than necessary for the proof.

Lemma 12.24. *We have*

- $d\tau \neq 0$ *on* (12.60) *except on the set of points of* $\mathbb{R}\times\Delta_{S^*H\times S^*H}$. *Consequently,* (12.60) *is a smooth manifold except at these points and the pushforward*
$$\pi_{t*}\Delta_t^*\left(\Gamma^*\circ C_H\circ\Gamma \backslash T^*\mathbb{R}\times\mathbb{R}^+(\Delta_{S^*H\times S^*H})\right)$$
is an (immersed) Lagrangian submanifold.
- $\iota|_{\mathbb{R}\times(\hat{C}_H\backslash\mathbb{R}\Delta_{S^*H\times S^*H})}$ *is a Lagrange immersion, with self-intersections corresponding to 'return times.'*

Proof. As noted above, if $\Gamma_H = \Delta_t^*\Gamma^*\circ C_H\circ\Gamma$ intersects $0_{\mathbb{R}}\times T^*M\times T^*M$ transversely, then $\pi_{t*}\Delta_t^*\Gamma^*\circ C_H\circ\Gamma$ is Lagrangian. Since $H^*(M\times M)$ is of co-dimension one, $\Delta^*\Gamma^*\circ C_H\circ\Gamma$ fails to be transverse at an intersection point only if its tangent space is contained in $T(H^*(M\times M))$. Thus, it fails to be transverse only at points where $d\tau = \tau = 0$. Since $\tau(t,s,\xi,\xi') = |\xi| - |\xi'| = \sqrt{\sigma^2+\eta_n^2} - \sqrt{\sigma^2+(\eta_n')^2}$, we see that $\tau = 0$ if and only if $\eta_n = \pm\eta_n'$ and $d\tau = 0$ on this set if and only if

$\eta_n = \eta'_n = 0$. This proves that the intersection (12.60) is transversal except on the set $0_{\mathbb{R}} \times \Delta_{T^*H \times T^*H}$ and that it fails to be transversal there. Consequently, the pushforward is a smoothly immersed Lagrangian submanifold away from this singular set.

We now consider ι and first restrict it to $\mathbb{R} \times (\hat{C}_H \backslash \Delta_{T^*H \times T^*H})$ since $\hat{C}_H$ does not have a well-defined tangent plane on the critical locus. The map ι is then an immersion as long as $(G^t \times G^t)(\hat{C}_H)$ is transverse to the orbits of $G^t \times G^t$. Note that S^*H is the set of points of $S^*_H M$ where the Hamilton vector field H_g of $g = |\xi|$ is tangent to $S^*_H M$. Hence, $\iota|_{\mathbb{R} \times (\hat{C}_H \backslash \mathbb{R}\Delta_{S^*H \times S^*H})}$ is a Lagrange immersion. It follows that $\pi_{t*}\Delta_t^* \Gamma^* \circ C_H \circ \Gamma$ is an immersed canonical relation away from the set $\mathbb{R}_+ \bigcup_{|t| \leq T} (G^t \times G^t)(S^*H \times S^*H)$.

We next consider self-intersection set of this immersion. The fiber

$$\iota^{-1}(x_0, \xi_0, y_0, \eta_0) = \{(t, s, \xi, \xi') \in \mathbb{R} \times \hat{C}_H : (G^{-t}(s, \xi), G^{-t}(s, \xi')) = (x_0, \xi_0, y_0, \eta_0)\}, \tag{12.63}$$

of ι over a point in the image corresponds to simultaneous hitting times of (x_0, ξ_0) and (y_0, η_0) on $T^*_H M$. Thus, the self-intersection locus of $\Gamma_{T,\varepsilon}$ consists of the image of pairs $(t, s, \xi, \xi'), (t', s', \eta, \eta')$ such that

$$G^t(s, \xi') = G^{t'}(s', \eta') \iff G^{t-t'}(s, \xi), G^{t-t'}(s, \xi') \in T^*_H M.$$

If $\xi = \xi'$ then $(s, \eta) = (s', \eta')$ and the self-intersection points correspond to the return times and positions of (s, ξ) to $T^*_H M$. If $\xi' = r_H \xi$, then the self-intersection points correspond to the times where the left and right times are the same. Away from $T^*H \times T^*H$ the set of return times is discrete.

This concludes the proof of the Lemma and hence of Proposition 12.23. □

12.16. The symbol of $U(t_1)^*(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon} U(t_2)$

The canonical relation of $U(t_1)^*(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon} U(t_2)$ was determined in Lemma 12.21. We now work out its symbol.

LEMMA 12.25. *If $a \in S^0_{cl}(T^*H)$, then*

$$U(-t_1)(\gamma_H^* \operatorname{Op}(a)\gamma_H)_{\geq\varepsilon} U(t_2) \in I^0(\mathbb{R} \times M \times \mathbb{R} \times M, \Gamma^* \circ C_H \circ \Gamma).$$

Under the embedding $\iota_{\Gamma^* C_H \Gamma}$ of Lemma 12.21, the principal symbol pulls back to the homogeneous function on $\mathbb{R} \times \mathbb{R} \times T^*_H M$ given by

$$(1 - \chi_\varepsilon) a_H(s, \xi) \, |dt \wedge dt_1 \wedge \Omega|^{\frac{1}{2}},$$

where $|dt \wedge dt_1 \wedge \Omega|^{\frac{1}{2}}$ is the canonical volume half-density on $\Gamma^* \circ C_H \circ \Gamma$ (defined in the proof).

PROOF. It is well known (see [**Ho3**]) that $U(t) \in I^{-\frac{1}{4}}(\mathbb{R} \times M \times M, \Gamma)$, with $\Gamma = \{(t, \tau, x, \xi, G^t(x, \xi)) : \tau + |\xi| = 0\}$. As in [**DG**], the half density symbol of $U(t, x, y)$ is the canonical volume half density $\sigma_{U(t,x,y)} = |dt \otimes dx \wedge d\xi|^{\frac{1}{2}}$ on Γ. By Proposition 12.21, the composition is $\Gamma^* \circ C_H \circ \Gamma$ is transversal for any hypersurface H, hence $U(-t_1)(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq\varepsilon} U(t_2)$ is a Fourier integral operator with the stated canonical relation. Under transversal composition the orders add, and the stated order follows from Lemma 12.20 together with the fact that $U(t) \in I^{-\frac{1}{4}}(\mathbb{R} \times M \times M, \Gamma)$.

To prove the formula for the symbol, we observe that $\Gamma^* \circ C_H \circ \Gamma$ is parametrized by

$$\iota_1(t, t', s, \sigma, \eta_n, \eta_n') = (t, |\xi|, t', -|\xi'|, G^t(s, \xi), G^{t'}(s, \xi')) : s \in H,\ \xi, \xi' \in T_x M,\ \xi|_{TH} = \xi'|_{TH}, \tag{12.64}$$

where $\xi = (\sigma, \eta_n), \xi' = (\sigma, \eta_n'))$ are dual Fermi coordinates in the orthogonal decomposition of $T_H^* M = T^* H \oplus N^* H$. The natural volume half density on parameter domain of $\Gamma \circ C_H \circ \Gamma$ is $|dt_1 \wedge dt_2 \wedge ds \wedge d\sigma \wedge d\eta_n \wedge d\eta_n'|^{\frac{1}{2}}$ where $ds \wedge d\sigma$ is the symplectic volume form on $T^* H$, where (η_n, η_n') are the normal components of (ξ, ξ') and where $d\eta_n$ is the Riemannian density on $N_s^* H$. The stated symbol then follows by transversal composition from the symbols of $U(t)$ and of $\gamma_H^* \operatorname{Op}_H(a) \gamma_H$ determined in Lemma 12.21). □

By standard wave front calculus, it follows that

$$\gamma_H^* \operatorname{Op}_H(a) \gamma_H = (\gamma_H^* Op_H(a) \gamma_H))_{\geq \varepsilon} + (\gamma_H^* \operatorname{Op}_H(a) \gamma_H))_{\leq \varepsilon} + K_\varepsilon, \tag{12.65}$$

with $\langle K_\varepsilon \varphi_j, \varphi_j \rangle_{L^2(M)} = \mathcal{O}(\lambda_j^{-\infty})$.

12.17. Proof of the restricted local Weyl law: Proposition 12.14

By the analysis of the canonical relations above, $U(t)(\gamma_H^* \operatorname{Op}_H(a) \gamma_H)_{\geq \varepsilon}$ is a Fourier integral operator of order $\frac{1}{4}$ associated to the (clean) composition of the canonical relation Γ of $U(t)$ and C_H. The trace (12.26) is the further composition with $\pi_* \Delta^*$ as in [**DG**], where $\Delta \colon \mathbb{R} \times M \to \mathbb{R} \times M \times M$ is the embedding $(t, x) \to (t, x, x)$ and $\pi \colon \mathbb{R} \times M \to \mathbb{R}$ is the natural projection. Then $\pi_* \Delta^* U(t) \circ (\gamma_H^* Op_H(a) \gamma_H)_{\geq \varepsilon}$ has singularities at times t so that $G^t(x, \xi) = (x, \xi)$ with $(x, \xi) \in S_H^* M$. By the standard Fourier Tauberian theorem (see [**Ho2**]) the growth rate of the sums above are determined by the singularity at $t = 0$ of the trace, where of course all of $S_H^* M$ is fixed. Hence the fixed point set is a codimension one submanifold of $S^* M$. If $n = \dim M$, $\pi_* \Delta^* U(t) \circ (\gamma_H^* \operatorname{Op}_H(a) \gamma_H)_{\geq \varepsilon} \in I^{\frac{1}{4} + n - 1 - \frac{1}{4}}(T_0^* \mathbb{R})$.

Note that, due to the drop of one in codimension, the singularity of the trace (12.26) loses a degree of $\frac{1}{2}$, but due to the extra $\frac{1}{2}$ in the order of $(\gamma_H^* \operatorname{Op}_H(a) \gamma_H)_{\geq \varepsilon}$ (compared to a pseudo-differential operator), it gains it back again. Hence the order of the singularity of (12.26) is the same as for pseudo-differential operators, and so the spectral asymptotics have the same order in λ. The principal symbol of the trace is determined by the symbol composition and Lemma 12.20. Except for the factor of $\gamma_H^* a$, the half-density symbol is the canonical Liouville volume form on $S_H^* M$. Since $\gamma_H^* a$ is a pullback from $B^* H$, we can project the measure to $B^* H$ and then we obtain the stated formula.

It remains to show that

$$\lim_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{j : \lambda_j \leq \lambda} \langle \operatorname{Op}_H(a) \gamma_H \varphi_j, \gamma_H \varphi_j \rangle = (12.25) + o(1) \text{ as } \varepsilon \to 0.$$

In view of (12.65), it is enough to prove

$$\lim_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{j : \lambda_j \leq \lambda} \langle \chi_{2\varepsilon} \gamma_H^* \operatorname{Op}_H(a) \gamma_H \chi_\varepsilon \varphi_j, \varphi_j \rangle = o(1) \text{ as } \varepsilon \to 0.$$

There are three types of terms: one with the tangential cutoff in both cutoff positions, one with the normal cutoff in both positions and two mixed ones with one tangential and one normal cutoff. Successive applications of the inequality $ab \leq \frac{1}{2}(a^2+b^2)$, Cauchy-Schwarz and L^2-boundedness of $\mathrm{Op}_H(a)$ implies that

$$\left|\frac{1}{N(\lambda)}\sum_{\lambda_j\leq\lambda}\langle(\gamma_H^*\,\mathrm{Op}_H(a)\gamma_H)_{\leq\varepsilon}\varphi_j,\varphi_j\rangle\right| \leq \frac{C}{N(\lambda)}\sum_{\lambda_j\leq\lambda}\Big(\|\gamma_H\chi_\varepsilon^{(tan)}\varphi_j\|^2_{L^2(H)} + \|\gamma_H\chi_{2\varepsilon}^{(tan)}\varphi_j\|^2_{L^2(H)} + \|\gamma_H\chi_\varepsilon^{(n)}\varphi_j\|^2_{L^2(H)} + \|\gamma_H\chi_{2\varepsilon}^{(n)}\varphi_j\|^2_{L^2(H)}\Big). \tag{12.66}$$

Finally, one applies the pointwise local Weyl law (12.29) on M to estimate the right side in (12.66). It follows that

$$\frac{1}{N(\lambda)}\sum_{\lambda_j\leq\lambda}\|\gamma_H\chi^{(tan)}_{\varepsilon,2\varepsilon}\varphi_j\|^2_{L^2(H)} = \mathcal{O}(\varepsilon), \tag{12.67}$$

and the same is true for the other cutoff operators $\chi^{(n)}_{\varepsilon,2\varepsilon}$.

12.18. Asymptotic completeness and orthogonality of Cauchy data

In this section we consider convergence of eigenfunction expansions and, in a new direction, convergence of expansions in terms of Dirichlet, Neumann and Cauchy data of eigenfunctions on a hypersurface H, including the boundary when the Laplacian has boundary conditions. The first topic is very classical and we only briefly discuss it. The second is new and is still in its infancy. We review several aspects of the classical (interior) case before discussing the analogues for Cauchy data.

12.18.1. Convergence and localization of eigenfunction expansions and their Riesz means. Two classical problems on eigenfunctions are convergence and localization of eigenfunction expansions of functions belonging to various Sobolev or Holder spaces. Classical results include the almost everywhere convergence of partial sums of Fourier series of continuous or L^2 functions on the circle, or partial Fourier integrals of L^2 functions on $\mathbb{R}$. Another type of classical result is the Riemann localization principle for Fourier series on the circle, which states if an integrable function vanishes in an open set U then the partial sums of its Fourier series converge uniformly to zero on compact subsets of U. In dimensions ≥ 2, there are many ways to define partial sums or integrals in Fourier inversion formulae and that quickly leads to many well known open problems. For instance one may define rectangular partial sums or spherical partial sums of Fourier series and integrals. The Gibbs phenomenon in dimension one concerns the lack of uniformity in the convergence of partial sums of Fourier series of piecewise smooth functions with a jump discontinuity.

Tonelli showed that localization does not hold on higher dimensional tori $\mathbb{T}^d$, $d>1$. Bochner put this in a quantitative form for spherical summation of multiple Fourier series or integrals. He introduced the Bochner-Riesz means of order δ,

$$S^\delta_\lambda f(x) = \sum_j\left(1-\frac{\lambda_j}{\lambda}\right)^\delta E_j f, \tag{12.68}$$

where $E_j \colon L^2(T^d) \to \mathcal{E}(\lambda_j)$ is the eigenspace projection. Localization may fail if $\delta < \frac{d-1}{2}$ while it holds if $\delta \geq \frac{d-1}{2}$ in the sense

$$S^\delta_\lambda f(x) \to 0, \qquad \delta \geq \frac{d-1}{2} \tag{12.69}$$

at all points of an open set U where $f(x) = 0$.

The set on which the partial sums/integrals of a function diverge is known as the set of divergence or an exceptional set. When f belongs to certain Sobolev spaces, its set of divergence is nicer than a general set of measure zero, and has capacity zero. We refer to [**CGV**] for a recent discussion of localization and "sets of non-localization" for convergence of Riesz means is given in [**CGV**] in terms of Riesz kernels. Let $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ and consider the norms $\|S^\delta_\lambda\|_{L^q(r(x,y) \geq \varepsilon)}$. of the linear functionals

$$\Lambda_{x,\delta,\lambda}(f) = S^\delta_\lambda f(x) \text{ on } L^p(M)$$

restricted to the set

$$\{f \in L^p : \ f = 0 \text{ on } \mathrm{B}_\varepsilon(\mathrm{x})\}. \tag{12.70}$$

The uniform boundedness principle implies that $S^\delta_\lambda(f)(x) \to 0$ for all $f \in L^p$ with $f = 0$ on $B_\varepsilon(x)$ only if these restricted functionals are uniformly bounded. In the case of $\mathbb{R}^d$, the Riesz kernel at $(0, y)$ has size $\lambda^{(d-1)/2-\delta}|y|^{-\frac{(d+1)}{2}-\delta}$ and are therefore uniformly bounded if and only if $\delta \geq \frac{d-1}{2}$.

A natural question for a geometric analyst is to determine the set of divergence for functions with controlled singularities, e.g., functions with a jump discontinuity along a hypersurface. Such questions have been studied by M. Pinsky and M. Taylor. The first issue is to generalize the Gibbs phenomenon to Euclidean spaces $\mathbb{R}^d$ of higher dimensions. In the case $\delta = 0$, Pinsky observed that pointwise localization fails for the characteristic function χ_B of the unit ball $B \subset \mathbb{R}^d$ and that at the center 0, $S^0_R \chi_B(0)$ oscillates between two distinct positive limits (see the Example of [**P2**, p.657]) even though χ_B is smooth in a neighborhood of 0. The discontinuity at the boundary ∂B propagates in a sense to 0 where it causes divergence of the partial sums of the eigenfunction expansion. More general results on this 'Pinsky phenomenon' and a wave equation analysis of it are given in [**P1, PT, Tay2**]. Taylor studies partial sums and Bochner-Riesz means of Fourier series or integrals for general conormal distributions, a vast generalizations of functions which are smooth except for a jump discontinuity across a hypersurface.

Geometric analysis of the wave equation enters when one seeks to generalize convergence or localization results to eigenfunction expansions on Riemannian manifolds or to Schrödinger operators.. On any Riemannian manifold one may consider the Bochner-Riesz means of order δ (12.68). The Schwartz kernel of the Bochner-Riesz means is

$$S^\delta_\lambda(x, y) = \sum_{j : \lambda_j \leq \lambda} (1 - \frac{\lambda_j}{\lambda})^\delta \varphi_j(x) \varphi_j(y). \tag{12.71}$$

The exponentials $e^{i\langle x, \xi \rangle}$ have many special features discussed in the section on L^p norms, such as uniform boundedness, and partial sums of eigenfunctions on Riemannian manifolds may be expected to reflect curvature properties of the metric or dynamics of the geodesic flow. To the author's knowledge, the only articles devoted to the effect of geodesic dynamics on convergence, localization and Gibbs or Pinsky effects are those of [**Tay2, BraC1, BraC2**].

Closely related to the study of partial sums or Bochner-Riesz means of conormal distributions is that of pointwise asymptotics for spectral projections kernels and for Bochner-Riesz kernels (12.71). In effect this amounts to studying partial sums for the conormal distribution $\delta_x(y)$. There are a variety of results on the $L^p - L^q$ mapping properties of such kernels [**So1**], but few if any results on pointwise asymptotics. The asymptotics for (12.71) simplify if $x = y$ and in particular for the integrated kernel $\int_M S_\lambda^\delta(x,x)\,dV(x)$, which depends only on the eigenvalues. Some results on asymptotics of integrated Bochner-Riesz means are given in [**Brun, Sa2**]. In [**Sa2**] the following is stated without proof for $\delta = 1, 2, \dots$ and is proved in [**Brun**] for general $\delta > 0$.

THEOREM 12.26. *There exists a bounded function $Q(\lambda)$ and constants $C_{j,\delta}$ such that, if $\dim M = d$,*

$$\operatorname{Tr} S_\lambda^\delta = \sum_{i=0}^{[\delta]_+} C_{j,\delta}\lambda^{d-j} + \lambda^{d-1-\delta} Q^\delta(\lambda) + o(\lambda^{d-1-\delta}).$$

Here, $[\delta]_+$ is the smallest integer $> \delta$.

Bruneau works with semi-classical Schrodinger operators at a fixed energy level, so that his results are more general. In fact $Q(\lambda)$ is given explicitly in [**Sa2, Brun**] and is the same Q that arose as the middle term in the Safarov trace or pre-trace formula and two-term Weyl law discussed in the section on L^p norms. There exists a pointwise analogue for the full kernel (12.71) but to the author's knowledge it has not been discussed in the literature.

12.19. Expansions in Cauchy data of eigenfunctions

In this section we consider analogous problems for convergence and localization of Cauchy data of eigenfunctions on a hypersurface.

The Dirichlet or Neumann data of global eigenfunctions on a hypersurface H may be expected to provide an 'over-complete' basis for $L^2(H)$ and one may consider 'eigenfunction expansions' of certain $f \in C^2(H)$ with respect to restrictions of global eigenfunctions. By 'over-complete' we mean that there are "too many basis elements' in the sense that the growth rate of the interior eigenvalues is one degree higher than for the boundary eigenvalues. But if one restricts the frequencies to a narrow window, then the Cauchy data does resemble an asymptotically orthonormal basis. This was conjectured in [**BFS, BFSS**] and proved in [**HHHZ**] if one uses a particular type of Schwartz weight to define the partial sums of the eigenfunction expansions.

First we consider Neumann data of Dirichlet eigenfunctions, resp. Dirichlet data of Neumann eigenfunctions, on $H = \partial M$ and then consider a general interior hypersurface (possibly on a manifold without boundary). To simplify notation, we denote the Cauchy data of global eigenfunctions u_j on H by

$$\begin{cases} \text{Dirichlet data: } \omega_j = \varphi_j|_H, \\ \text{Neumann data: } \psi_j = \partial_\nu \varphi_j|_H. \end{cases}$$

THEOREM 12.27 (Completeness of boundary traces of Dirichlet eigenfunctions). *Let $\rho \in \mathcal{S}(\mathbb{R})$ be such that $\hat{\rho}$ is identically* 1 *near* 0*, and has sufficiently small support. Then for any $f \in C^\infty(\partial M)$, we have*

$$f(x) = \lim_{\lambda\to\infty} \frac{\pi}{2} \sum_j \rho(\lambda - \lambda_j)\lambda_j^{-2} \langle \psi_j, f\rangle \psi_j(x), \tag{12.72}$$

where $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\partial M}$ denotes the inner product in $L^2(\partial M)$.

The expansion on the right side thus exhibits the localization property: if one modifies f away from x it changes the Fourier coefficients but not the value of the expansion at x. Localization essentially involves finite propagation speed of the wave equation and is ensured here by the support assumption on $\hat{\rho}$. In the Neumann case,

THEOREM 12.28 (Completeness of boundary traces of Neumann eigenfunctions). *Let ρ be as in Theorem 12.27. Then for any $\varphi \in C^\infty(\partial M)$, we have as $\lambda \to \infty$,*

$$\frac{\pi}{2} \sum_j \rho(\lambda - \lambda_j) \langle \omega_j, f\rangle \omega_j(y) = f(y) + \frac{1}{2}\lambda^{-2} \left[\Delta_{\partial M} - \frac{3}{4}(n-1)^2 H_y^2 + \frac{1}{2}(n-1)(n-2)K_y\right] f(y) + O(\lambda^{-3}). \tag{12.73}$$

This completeness result also holds for any interior hypersurface H.

THEOREM 12.29 (Completeness of Cauchy data on interior hypersurfaces). *Let $\rho \in \mathcal{S}(\mathbb{R})$ be as in Theorem 12.27. Then for any $\varphi \in C^\infty(H)$, we have*

$$\varphi(x) = \lim_{\lambda\to\infty} \pi \sum_j \rho(\lambda - \lambda_j) \langle \omega_j, \varphi\rangle \omega_j(x),$$

and

$$\varphi(x) = \lim_{\lambda\to\infty} \pi \sum_j \rho(\lambda - \lambda_j)\lambda_j^{-2} \langle \psi_j, \varphi\rangle \psi_j(x),$$

where $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\partial M}$ denotes the inner product in $L^2(H)$.

The theorems are proved by studying the operator K_λ^D of (12.72):

$$K_\lambda^D = \frac{\pi}{2} \sum_j \rho(\lambda - \lambda_j)\lambda_j^{-2} \psi_j \langle \psi_j, \cdot \rangle, \tag{12.74}$$

resp. the Neumann analogue K_λ^N:

$$K_\lambda^N = \frac{\pi}{2} \sum_j \rho(\lambda - \lambda_j) \omega_j \langle \omega_j, \cdot \rangle. \tag{12.75}$$

The key property of these operators is:

THEOREM 12.30. *Let A_h be a semiclassical pseudo-differential operator on ∂M, microsupported in $\{(y,\eta) \in T^*(\partial M) : |\eta| < 1 - \varepsilon_1\}$ for some $\varepsilon_1 > 0$. Let ρ be such that $\hat{\rho}$ is supported sufficiently close to* 0 *(depending on ε_1). Then*

(1) *$A_h K_{h^{-1}}^D$ and $K_{h^{-1}}^D A_h$ are semiclassical pseudo-differential operators with principal symbol*

$$\sigma(A)(1 - |\eta|^2)^{1/2}; \tag{12.76}$$

(2) *$A_h K^N_{h^{-1}}$ and $K^N_{h^{-1}} A_h$ are semiclassical pseudo-differential operators with principal symbol*

$$\sigma(A)(1-|\eta|^2)^{-1/2}. \tag{12.77}$$

This is proved by studying the canonical relations and principal symbols of the Fourier transform

$$2\pi\hat{\rho}(t) R_y R_{y'} \cos(t\sqrt{-\Delta}) \quad \text{and} \quad 2\pi\hat{\rho}(t) R_y R_{y'} d_{n_y} d_{n_{y'}} \frac{\cos(t\sqrt{-\Delta})}{-\Delta}$$

of the operators K^D_λ and K^N_λ in λ (t is the dual variable of λ). Here R_y denotes the restriction operator. One finds that

(1) the kernels of

$$\hat{\rho}(t)\chi(y, D_t, D_y) \circ R_y R_{y'} \cos(t\sqrt{-\Delta}), \quad \hat{\rho}(t) R_y R_{y'} \cos(t\sqrt{-\Delta}) \circ \chi(y, D_t, D_y)$$

are distributions conormal to $\{y = y',\, t = 0\}$ with principal symbol

$$\chi(y,\tau,\eta)\left(1 - \frac{|\eta|^2_{\tilde{g}}}{\tau^2}\right)^{-\frac{1}{2}};$$

(2) the kernels of

$$\hat{\rho}(t)\chi(y, D_t, D_y) \circ R_y R_{y'} d_{n_y} d_{n_{y'}} \frac{\cos(t\sqrt{-\Delta})}{-\Delta}, \quad \hat{\rho}(t) R_y R_{y'} d_{n_y} d_{n_{y'}} \frac{\cos(t\sqrt{-\Delta})}{-\Delta} \circ \chi(y, D_t, D_y)$$

are distributions conormal to $\{y = y',\, t = 0\}$ with principal symbol

$$\chi(y,\tau,\eta)\left(1 - \frac{|\eta|^2_{\tilde{g}}}{\tau^2}\right)^{\frac{1}{2}},$$

in which $\tilde{g}$ is the induced metric on H.

12.20. Bochner-Riesz means for Cauchy data

Comparing results on summation and localization of eigenfunction expansions on M and on H suggests a refinement in which one replaces (12.74) or (12.10) by Bochner-Riesz means. In this section we briefly mention a conjectural sharpening of Theorem 12.27 to give what might be called Bochner-Riesz means for Cauchy data. We sharpen the Schwartz-weighted operators (12.74) and (12.10) by taking the Cauchy data of the interior Bochner-Riesz means on H. For simplicity of notation we use the superscript K^b to be the relevant component of the Cauchy data of a kernel on H. As above the hypersurface could be an interior hypersurface or the boundary (when $\partial M \neq \emptyset$).

Definition 12.31. We define the Cauchy data Bochner-Riesz kernels of order δ for Cauchy data by

$$S^{\delta,b}_\lambda(q,q') = \sum_j \left(1 - \frac{\lambda_j}{\lambda}\right)^\delta E^b_j(q,q') = \sum_{j:\lambda_j\le\lambda} \left(1 - \frac{\lambda_j}{\lambda}\right)^\delta \varphi^b_j(q)\varphi^b_j(q'). \tag{12.78}$$

The Riesz mean of the Dirichlet data expansion of a function $f \in C^\infty(\partial M)$ is defined by

$$S^{\delta,b}_\lambda f(x) = \sum_{j:\lambda_j\le\lambda} \left(1 - \frac{\lambda_j}{\lambda}\right)^\delta_+ \left(\int_H f(y)\varphi^b_j(y)\, dS(y)\right) \varphi^b_j(x). \tag{12.79}$$

The question is whether it converges pointwise to f for $f \in C^\infty(M)$ or for f in some given Sobolev class. By comparing to Theorem 12.26 one might state the conjecture:

CONJECTURE 12.32. Let $(M, g, \partial M)$ be a C^∞ Riemannian manifold with boundary. Assume that $Q^b(x, \lambda) = 0$. Then for either Dirichlet or Neumann boundary conditions,

$$S_\lambda^{\delta,b} f(x) = \sum_{j:\lambda_j \leq \lambda} (1 - \frac{\lambda_j}{\lambda})_+^{n-1} \left(\int_Y f(y) \varphi_j^b(y) \, dS(y) \right) \varphi_j^b(x) \tag{12.80}$$

$$\simeq \lambda^{n-1-\delta} f(x) + o(\lambda^{n-1-\delta}). \tag{12.81}$$

12.21. Quantum ergodic restriction theorems

In this section, we review a series of QER (quantum ergodic restriction) theorems. Roughly speaking, such theorems say that if the geodesic flow is ergodic, then Cauchy data of global ergodic eigenfunctions remain ergodic when restricted to a hypersurface. There are two types of QER theorems: (i) one for the Cauchy data of eigenfunctions, which is universal in that it does not require any restrictions on the hypersurface H; (ii) a subtler one for Dirichlet data (or for Neumann data), which does require an 'asymmetry' condition on H. For generic hypersurfaces, the Dirichlet and Neumann data are separately ergodic.

To be more precise, let $H \subset M$ be a hypersurface and consider the Cauchy data $(\varphi_j|_H, \lambda_j^{-1} \partial_\nu \varphi_j|_H)$ of eigenfunctions along H. A QER theorem seeks to find limits of matrix elements of this data along H with respect to pseudo-differential operators $\mathrm{Op}_H(a)$ on H. The main idea is that $S_H^* M$, the set of unit covectors with footpoints on H, is a cross-section to the geodesic flow and the first return map of the geodesic flow for $S_H^* M$ is ergodic. The Cauchy data should be the quantum analogue of such a cross section and therefore should be quantum ergodic on H. QER theorems give asymptotics for $\int_H f \varphi_j^2 dS$ for any $f \in C^\infty(H)$ and for more general matrix elements $\langle Op_h(a) \varphi_j|_H \varphi_j|_H \rangle_{L^2(H)}$ relative to semi-classical pseudo-differential operators on H. As is usual in quantum ergodic theory, we must discard a possible sparse subsequence of eigenfunctions of a given orthonormal basis to obtain limits.

12.21.1. Quantum ergodicity of Cauchy data of restrictions. Let $H \subset M$ be a smooth hypersurface which does not meet ∂M if $\partial M \neq \emptyset$. The main result of this section (Theorem 12.33) is that the semiclassical Cauchy data (12.3) of eigenfunctions is *always* quantum ergodic along any hypersurface $H \subset M$ if the eigenfunctions are quantum ergodic on the global manifold M. The theorem is due to Christianson, Toth and the author [**CTZ**] but is a generalization of the boundary case where $H = \partial M$, which was proved in [**HZ**] and in [**Bu2**] and for Dirichlet eigenfunctions of $C^{1,1}$ plane domains in [**GL**].

To state the results precisely, we introduce some notation. We work with the semiclassical calculus of pseudo-differential operators on both M and H. We fix (Weyl) quantizations $a \to a^w$ of semi-classical symbols to semi-classical pseudo-differential operators. When it is necessary to indicate which manifold is involved, we either write $\mathrm{Op}_H(a)$ for pseudo-differential operators on H or we use capital letters $A^w(x, hD)$ to indicate operators on M and small letters $a^w(y, hD_y)$ to indicate operators on H. With no loss of generality, we assume that H is orientable, embedded, and separating in the sense that

$$M \backslash H = M_+ \cup M_-$$

where $M_\pm$ are domains with boundary in M. Since the results are local on H, this is not a restriction on the hypersurface.

Given a quantization $a \to \mathrm{Op}_H(a)$ of semi-classical symbols $a \in S^0_{sc}(H)$ of order zero to semi-classical pseudo-differential operators on $L^2(H)$, we define the microlocal lifts of the Neumann data as the linear functionals on $a \in S^0_{sc}(H)$ given by

$$\mu_h^N(a) := \int_{B^*H} a \, d\Phi_h^N := \langle \mathrm{Op}_H(a) hD_\nu \varphi_h|_H, hD_\nu \varphi_h|_H \rangle_{L^2(H)}.$$

We also define the *renormalized microlocal lifts* of the Dirichlet data by

$$\mu_h^D(a) := \int_{B^*H} a \, d\Phi_h^{RD} := \langle \mathrm{Op}_H(a)(1 + h^2 \Delta_H) \varphi_h|_H, \varphi_h|_H \rangle_{L^2(H)}.$$

Finally, we define the microlocal lift $d\Phi_h^{CD}$ of the Cauchy data to be the sum

$$d\Phi_h^{CD} := d\Phi_h^N + d\Phi_h^{RD}. \tag{12.82}$$

Here, $h^2 \Delta_H$ denotes the negative tangential Laplacian for the induced metric on H, so that the operator $(1 + h^2 \Delta_H)$ is characteristic precisely on the glancing set S^*H of H. Intuitively, we have renormalized the Dirichlet data by damping out the whispering gallery components.

The distributions μ_h^N, μ_h^D are asymptotically positive, but are not normalized to have mass one and may tend to infinity. They depend on the choice of quantization, but their possible weak* limits as $h \to 0$ do not. We refer to [**Zw**] for background on semi-classical microlocal analysis.

Our first result is that the Cauchy data of a sequence of quantum ergodic eigenfunctions restricted to H is automatically QER for semiclassical pseudodifferential operators with symbols vanishing on the glancing set S^*H, i.e., that $d\Phi_h^{CD} \to \omega$, where

$$\omega(a) = \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{1/2} \, d\sigma$$

is the limit state of Theorem 12.33. This was proved in a different way in [**ToZ2**] in the case of piecewise smooth Euclidean domains. The assumption $H \cap \partial M = \emptyset$ is for simplicity of exposition and because the case $H = \partial M$ is already known.

THEOREM 12.33. *Suppose $H \subset M$ is a smooth, codimension 1 embedded orientable separating hypersurface and assume $H \cap \partial M = \emptyset$. Assume that $\{\varphi_h\}$ is a quantum ergodic sequence of eigenfunctions* (12.2). *Then the sequence $\{d\Phi_h^{CD}\}$* (12.82) *of microlocal lifts of the Cauchy data of φ_h is quantum ergodic on H in the sense that for any $a \in S^0_{sc}(H)$,*

(12.83)

$$\langle \mathrm{Op}_H(a) hD_\nu \varphi_h|_H, hD_\nu \varphi_h|_H \rangle_{L^2(H)} + \left\langle \mathrm{Op}_H(a)(1 + h^2 \Delta_H) \varphi_h|_H, \varphi_h|_H \right\rangle_{L^2(H)}$$
$$\to_{h \to 0^+} \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{1/2} \, d\sigma,$$

where $a_0(x', \xi')$ is the principal symbol of $\mathrm{Op}_H(a)$, $-h^2 \Delta_H$ is the induced tangential (semiclassical) Laplacian with principal symbol $|\xi'|^2$, μ is the Liouville measure on S^*M, and $d\sigma$ is the standard symplectic volume form on B^*H.

REMARK 12.34. We emphasize that the limit along H in Theorem 12.33 holds for the full sequence $\{\varphi_h\}$. Thus, if the full sequence of eigenfunctions is known to be quantum ergodic, i.e., if the sequence is QUE, then the conclusion of the theorem applies to the full sequence of eigenfunctions.

The proof simply relates the interior and restricted microlocal lifts and reduces the QER property along H to the QE property of the ambient manifold. If we assume that QUE holds in the ambient manifold, we automatically get QUER, which is our first Corollary:

COROLLARY 12.35. *Suppose that $\{\varphi_h\}$ is QUE on M. Then the distributions $\{d\Phi_h^{CD}\}$ have a unique weak* limit*

$$\omega(a) := \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{1/2} d\sigma.$$

We note that $d\Phi_h^{CD}$ involves the microlocal lift $d\Phi_h^{RD}$ rather than the microlocal lift of the Dirichlet data. However, in Theorem 12.36, we see that the analogue of Theorem 12.33 holds for a *density one subsequence* if we use the further renormalized distributions $d\Phi_h^D + d\Phi_h^{RN}$ where the microlocal lift $d\Phi_h^D \in \mathcal{D}'(B^*H)$ of the Dirichlet data of φ_h is defined by

$$\int_{B^*H} a\, d\Phi_h^D := \langle \mathrm{Op}_H(a) \varphi_h|_H, \varphi_h|_H \rangle_{L^2(H)},$$

and

$$\int_{B^*H} a\, d\Phi_h^{RN} := \langle (1 + h^2 \Delta_H + i0)^{-1} \mathrm{Op}_H(a) h D_\nu \varphi_h|_H, h D_\nu \varphi_h|_H \rangle_{L^2(H)}.$$

THEOREM 12.36. *Suppose $H \subset M$ is a smooth, codimension 1 embedded orientable separating hypersurface and assume $H \cap \partial M = \emptyset$. Assume that $\{\varphi_h\}$ is a quantum ergodic sequence. Then, there exists a sub-sequence of density one as $h \to 0^+$ such that for all $a \in S^0_{sc}(H)$,*

$$\begin{aligned} (12.84) \quad & \big\langle (1 + h^2 \Delta_H + i0)^{-1} \mathrm{Op}_H(a) h D_\nu \varphi_h|_H, h D_\nu \varphi_h|_H \big\rangle_{L^2(H)} \\ & + \langle \mathrm{Op}_H(a) \varphi_h|_H, \varphi_h|_H \rangle_{L^2(H)} \to_{h \to 0^+} \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{-1/2}\, d\sigma, \end{aligned}$$

where $a_0(x', \xi')$ is the principal symbol of $\mathrm{Op}_H(a)$.

The result relies on the local Weyl law along H (Proposition 12.14; see §12.17) showing that only a sparse set of eigenfunctions could scar on the glancing set S^*H. Such eigenfunctions, if they exist, would not be quantum ergodic and are deleted. This is why we obtain a QER result but not a QUER result (i.e., a quantum uniquely ergodic restriction theorem). However, the following is a direct consequence of Theorem 12.36.

COROLLARY 12.37. *Suppose that $\{\varphi_h\}$ is QUE on M. Then the distributions $\{d\Phi_h^D + d\Phi_h^{RN}\}$ have a unique weak* limit*

$$\omega(a) := \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x', \xi')(1 - |\xi'|^2)^{-1/2}\, d\sigma$$

with respect to the subclass of symbols which vanish on S^*H.

12.22. Rellich approach to QER: Proof of Theorem 12.33

Let $x = (x_1, ..., x_{n-1}, x_n) = (x', x_n)$ be Fermi normal coordinates in a small tubular neighborhood $H(\varepsilon)$ of H defined near a point $x_0 \in H$. In these coordinates we can locally write

$$H(\varepsilon) := \{(x', x_n) \in U \times \mathbb{R},\, |x_n| < \varepsilon\}.$$

Here $U \subset \mathbb{R}^{n-1}$ is a coordinate chart containing $x_0 \in H$ and $\varepsilon > 0$ is arbitrarily small but for the moment, fixed. We let $\chi \in C_0^\infty(\mathbb{R})$ be a cutoff with $\chi(x) = 0$ for $|x| \geq 1$ and $\chi(x) = 1$ for $|x| \leq 1/2$. In terms of the normal coordinates,

$$-h^2\Delta_g = \frac{1}{g(x)} hD_{x_n} g(x) hD_{x_n} + R(x_n, x', hD_{x'})$$

, where R is a second-order h-differential operator along H with coefficients that depend on x_n, and $R(0, x', hD_{x'}) = -h^2\Delta_H$ is the induced tangential semiclassical Laplacian on H.

The Rellich identity (see §12.3.1) relates matrix elements on H to matrix elements on M, which are known by global quantum ergodicity. Here, $D_{x_j} = \frac{1}{i}\frac{\partial}{\partial x_j}$, $D_{x'} = (D_{x_1}, ..., D_{x_{n-1}})$, $D_{x_n} = D_\nu = \frac{1}{i}\partial_\nu$ where ∂_ν is the interior unit normal to M_+. It is proved in the same way as (12.14).

Given $a \in S^{0,0}(T^*H \times (0, h_0])$, we define

$$A(x', x_n, hD_x) = \chi\left(\frac{x_n}{\varepsilon}\right) hD_{x_n} a^w(x', hD').$$

By the Rellich identity

$$\frac{i}{h}\int_{M_+} \left([-h^2\Delta_g, A(x, hD_x)]\varphi_h(x)\right)\overline{\varphi_h(x)}\, dx \tag{12.85}$$

$$= \int_H (hD_\nu A(x', x_n, hD_x)\varphi_h|_H)\,\overline{\varphi_h}|_H\, d\sigma_H + \int_H (A(x', x_n, hD_x)\varphi_h|_H)\,\overline{hD_\nu\varphi_h}|_H\, d\sigma_H.$$

Since $\chi(0) = 1$ it follows that the second term on the right side of (12.85) is just

$$\langle a^w(x', hD')hD_{x_n}\varphi_h|_H, hD_{x_n}\varphi_h|_H\rangle\,. \tag{12.86}$$

The first term on right hand side of (12.85) equals

$$\int_H hD_n(\chi(x_n/\varepsilon)hD_n a^w(x', hD')\varphi_h)\Big|_{x_n=0}\overline{\varphi_h}\Big|_{x_n=0} d\sigma_H \tag{12.87}$$

$$= \int_H \Big(\chi(x_n/\varepsilon)a^w(x', hD')(hD_n)^2\varphi_h \tag{12.88}$$

$$+ \frac{h}{i\varepsilon}\chi'(x_n/\varepsilon)hD_n a^w(x', hD')\varphi_h\Big)\Big|_{x_n=0}\overline{\varphi_h}\Big|_{x_n=0} d\sigma_H \tag{12.89}$$

$$= \int_H (\chi(x_n/\varepsilon)a^w(x', hD')(1 - R(x_n, x', hD'))\varphi_h)\Big|_{x_n=0}\overline{\varphi_h}\Big|_{x_n=0} d\sigma_H, \tag{12.90}$$

since $\chi'(0) = 0$ and $((hD_n)^2 + R + O(h))\varphi_h = \varphi_h$ in these coordinates.

It follows from (12.85)-(12.87) that

(12.91)
$$\langle \mathrm{Op}_H(a)hD_\nu\varphi_h|_H, hD_\nu\varphi_h|_H\rangle_{L^2(H)} + \big\langle \mathrm{Op}_H(a)(1+h^2\Delta_H)\varphi_h|_H, \varphi_h|_H\big\rangle_{L^2(H)}$$
$$= \left\langle \mathrm{Op}\left(\left\{\xi_n^2 + R(x_n,x',\xi'), \chi(\frac{x_n}{\varepsilon})\xi_n a(x',\xi')\right\}\right)\varphi_h, \varphi_h\right\rangle_{L^2(M_+)} + \mathcal{O}_\varepsilon(h).$$

We now assume that φ_h is a sequence of quantum ergodic eigenfunctions, and take the $h \to 0^+$ limit on both sides of (12.91). We apply interior quantum ergodicity to the term on the right side of (12.91). We compute

$$\left\{\xi_n^2 + R(x_n,x',\xi'),\, \chi(\frac{x_n}{\varepsilon})\xi_n a(x',\xi')\right\} = \frac{2}{\varepsilon}\chi'\left(\frac{x_n}{\varepsilon}\right)\xi_n^2 a(x',\xi') + \chi\left(\frac{x_n}{\varepsilon}\right)R_2(x',x_n,\xi'), \tag{12.92}$$

where R_2 is a zero order symbol. Let $\chi_2 \in \mathcal{C}^\infty$ satisfy $\chi_2(t) = 0$ for $t \le -1/2$, $\chi_2(t) = 1$ for $t \ge 0$, and $\chi_2'(t) > 0$ for $-1/2 < t < 0$, and let ρ be a boundary defining function for M_+. Then $\chi_2(\rho/\delta)$ is 1 on M_+ and 0 outside a $\delta/2$ neighborhood. Now the assumptions that the sequence φ_h is quantum ergodic implies that the matrix element of the second term on the right side of (12.92) is bounded by

$$\begin{aligned}&\Big|\langle(\,\chi(x_n/\varepsilon)R_2(x,\xi'))^w\varphi_h,\varphi_h\rangle_{L^2(M_+)}\Big| \\ &\quad\le \|\chi_2(\rho/\delta)\chi(x_n/\varepsilon)\varphi_h\|_{L^2(M)}\|\tilde\chi_2(\rho/\delta)\tilde\chi(x_n/\varepsilon)\varphi_h\|_{L^2(M)} = \mathcal{O}_\delta(\varepsilon) + o_{\delta,\varepsilon}(1),\end{aligned} \tag{12.93}$$

where $\tilde\chi$ and $\tilde\chi_2$ are smooth, compactly supported functions which are one on the support of χ and χ_2 respectively. Here, the last line follows from interior quantum ergodicity of the φ_h since the volume of the supports of $\chi(x_n/\varepsilon)$ and $\tilde\chi(x_n/\varepsilon)$ is comparable to ε.

To handle the matrix element of the first term on the right side of (12.92), we note that $\chi'(x_n/\varepsilon)|_{M_+} = \tilde\chi'(x_n/\varepsilon)$ for a smooth function $\tilde\chi \in \mathcal{C}^\infty(M)$ satisfying $\tilde\chi = 1$ in a neighborhood of $M \setminus M_+$ and zero inside a neighborhood of H. Then, again by interior quantum ergodicity, we have

(12.94)
$$2\left\langle\left(\frac{1}{\varepsilon}\chi'\left(\frac{x_n}{\varepsilon}\right)\xi_n^2 a(x',\xi')\right)^w\varphi_h,\varphi_h\right\rangle_{L^2(M_+)}$$

(12.95)
$$= 2\left\langle\left(\frac{1}{\varepsilon}\tilde\chi'\left(\frac{x_n}{\varepsilon}\right)\xi_n^2 a(x',\xi')\right)^w\varphi_h,\varphi_h\right\rangle_{L^2(M)}$$

(12.96)
$$= \frac{2}{\mu(S^*M)}\int_{S^*M}\frac{1}{\varepsilon}\tilde\chi'\left(\frac{x_n}{\varepsilon}\right)(1 - R(x',x_n,\xi'))a(x',\xi')\,d\mu + O(\varepsilon) + o_\varepsilon(1)$$

(12.97)
$$= \frac{2}{\mu(S^*M)}\int_{S^*M_+}\frac{1}{\varepsilon}\chi'\left(\frac{x_n}{\varepsilon}\right)(1 - R(x',x_n,\xi'))a(x',\xi')\,d\mu + O(\varepsilon) + o_\varepsilon(1),$$

since $\tilde{\chi}'$ and χ' are supported inside M_+. Combining the above calculations yields

(12.98)
$$\langle \mathrm{Op}_H(a) hD_\nu \varphi_h|_H, hD_\nu \varphi_h|_H\rangle_{L^2(H)} + \big\langle \mathrm{Op}_H(a)(1+h^2\Delta_H)\varphi_h|_H, \varphi_h|_H\big\rangle_{L^2(H)}$$
$$= \frac{2}{\mu(S^*M)} \int_{S^*M_+} \frac{1}{\varepsilon}\chi'\left(\frac{x_n}{\varepsilon}\right)(1-R(x',x_n,\xi'))a(x',\xi')\, d\mu + O_\delta(\varepsilon) + o_{\delta,\varepsilon}(1).$$

Finally, we take the $h \to 0^+$-limit in (12.98) followed by the $\varepsilon \to 0^+$-limit, and finally the $\delta \to 0^+$ limit. The result is that, since the left-hand side in (12.98) is independent of ε and δ,

$$\lim_{h\to 0^+} \langle \mathrm{Op}_H(a) hD_\nu \varphi_h|_H, hD_\nu \varphi_h|_H\rangle_{L^2(H)} + \big\langle \mathrm{Op}_H(a)(1+h^2\Delta_H)\varphi_h|_H, \varphi_h|_H\big\rangle_{L^2(H)}$$
$$= \frac{2}{\mu(S^*M)} \int_{S^*_H M} (1-R(x', x_n=0, \xi'))\, d\tilde{\sigma} = \frac{4}{\mu(S^*M)} \int_{B^*H} (1-|\xi'|^2)^{1/2} a(x',\xi'))\, d\sigma,$$

where $d\tilde{\sigma}$ is the symplectic volume form on $S^*_H M$, and $d\sigma$ is the symplectic volume form on B^*H.

12.23. Proof of Theorem 12.36 and Corollary 12.37

The proof follows as in Theorem 12.33 with a few modifications. For fixed $\varepsilon_1 > 0$ we choose the test operator

$$A(x', x_n, hD_x) := (I + h^2\Delta_H(x', hD') + i\varepsilon_1)^{-1} \chi(\frac{x_n}{\varepsilon}) hD_{x_n} a^w(x', hD') \tag{12.99}$$

and since $WF_h'(\varphi_h|_H) \subset B^*H$ (see [**ToZ2**, §11]) it suffices to assume that $a \in C_0^\infty(T^*H)$ with
$$\operatorname{supp} a \subset B^*_{1+\varepsilon_1^2}(H).$$

Let $\chi_{\varepsilon_1}(x',\xi') \in C_0^\infty(B^*_{1+\varepsilon_1^2}(H) \setminus B^*_{1-2\varepsilon_1^2}(H); [0,1])$ be a cutoff near the glancing set S^*H with $\chi_{\varepsilon_1}(x',\xi') = 1$ when $(x',\xi') \in B^*_{1+\varepsilon_1^2}(H) \setminus B^*_{1-\varepsilon_1^2}(H)$. Then, with $A(x, hD_x)$ in (12.99), the same Rellich commutator argument as in Theorem 12.33 gives

(12.100)
$$\big\langle (1+h^2\Delta_H + i\varepsilon_1)^{-1} a^w(x', hD')(1-\chi_{\varepsilon_1})^w hD_\nu \varphi_h|_H, hD_\nu\varphi_h|_H\big\rangle_{L^2(H)}$$
$$+ \left\langle a^w(x',hD')(1-\chi_{\varepsilon_1})^w \left(\frac{1-|\xi'|^2}{1-|\xi'|^2+i\varepsilon_1}\right)^w \varphi_h|_H, \varphi_h|_H\right\rangle_{L^2(H)}$$
$$\to \frac{4}{\mu(S^*M)} \int_{B^*H} a_0(x',\xi')(1-\chi_{\varepsilon_1}(x',\xi')) \left(\frac{(1-|\xi'|^2)^{1/2}}{1-|\xi'|^2+i\varepsilon_1}\right) d\sigma.$$

It remains to determine the contribution of the glancing set S^*H. As in [**Bu2, DZ, HZ, ToZ2**] we use a local Weyl law to do this. Because of the additional normal derivative term the argument is slightly different than in the cited articles and so we give some details. For the rest of this proof, we need to recall that $h \in \{\lambda_j^{-1}\}$, and we write h_j for this sequence to emphasize that it is a discrete sequence of values $h_j \to 0$. Since $\|a^w(x',hD')\|_{L^2\to L^2} = O(1)$, it follows that for

$h \in (0, h_0(\varepsilon_1)]$ with $h_0(\varepsilon_1) > 0$ sufficiently small,

$$\begin{aligned}
(12.101)\quad & \frac{1}{N(h)} \sum_{h_j \geq h} \left| \langle a^w(x', hD')\chi^w_{\varepsilon_1} \varphi_{h_j}|_H, \varphi_{h_j}|_H \rangle_{L^2(H)} \right| \\
& \leq C \frac{1}{N(h)} \sum_{h_j \geq h} \big(\|\chi^w_{\varepsilon_1} \varphi_{h_j}|_H\|_{L^2(H)} \, \|\chi^w_{2\varepsilon_1} \varphi_{h_j}|_H\|_{L^2(H)} + \mathcal{O}(h_j^\infty) \big) \\
& \leq \frac{C}{2} \frac{1}{N(h)} \sum_{h_j \geq h} \Big(\|\chi^w_{\varepsilon_1} \varphi_{h_j}|_H\|^2_{L^2(H)} + \|\chi^w_{2\varepsilon_1} \varphi_{h_j}|_H\|^2_{L^2(H)} + \mathcal{O}(h_j^\infty) \Big) \\
(12.102)\quad & = \mathcal{O}(\varepsilon_1^2).
\end{aligned}$$

By a Fourier Tauberian argument, it follows that for $h \in (0, h_0(\varepsilon_1)]$

$$(12.103)\qquad \frac{1}{N(h)} \sum_{h_j \geq h} |\chi^w_{\varepsilon_1, 2\varepsilon_1} \varphi_{h_j}|_H(x')|^2 = \mathcal{O}(\varepsilon_1^2)$$

uniformly for $x' \in H$. The last estimate in (12.101) follows from (12.103) by integration over H.

To estimate the normal derivative terms, we first recall the standard resolvent estimate

$$\|(1 + h^2 \Delta_H + i\varepsilon_1)^{-1} u\|_{H^2_h(H)} \leq C \varepsilon_1^{-1} \|u\|_{L^2(H)},$$

where H^2_h is the semiclassical Sobolev space of order 2 (see [**Zw**, Lemma 13.6]). Applying the obvious embedding $H^2_h(H) \subset L^2(H)$, we recover

$$\begin{aligned}
(12.104)\quad & \|(1 + h^2 \Delta_H + i\varepsilon_1)^{-1} u\|_{L^2(H)} \leq C \|(1 + h^2 \Delta_H + i\varepsilon_1)^{-1} u\|_{H^2_h(H)} \\
(12.105)\quad & \qquad \leq C \varepsilon_1^{-1} \|u\|_{L^2(H)}
\end{aligned}$$

to get that

(12.106)

$$\begin{aligned}
& \frac{1}{N(h)} \sum_{h_j \geq h} |\langle (1 + h^2 \Delta_H + i\varepsilon_1)^{-1} a^w(x', hD') \chi^w_{\varepsilon_1} h_j D_{x_n} \varphi_{h_j}|_H, h_j D_{x_n} \varphi_{h_j}|_H \rangle_{L^2(H)}| \\
& \leq C' \varepsilon_1^{-1} \frac{1}{N(h)} \sum_{h_j \geq h} \big(\|\chi^w_{\varepsilon_1} h D_{x_n} \varphi_{h_j}|_H\|_{L^2(H)} \, \|\chi^w_{2\varepsilon_1} h D_{x_n} \varphi_{h_j}|_H\|_{L^2(H)} + \mathcal{O}(h_j^\infty) \big) \\
& \leq \frac{C' \varepsilon_1^{-1}}{2} \frac{1}{N(h)} \sum_{h_j \geq h} \Big(\|\chi^w_{\varepsilon_1} h D_{x_n} \varphi_{h_j}|_H\|^2_{L^2(H)} + \|\chi^w_{2\varepsilon_1} h D_{x_n} \varphi_{h_j}|_H\|^2_{L^2(H)} + \mathcal{O}(h_j^\infty) \Big) \\
& = \mathcal{O}(\varepsilon_1^{-1} \varepsilon_1^2) \\
& = \mathcal{O}(\varepsilon_1).
\end{aligned}$$

The last estimate follows again from the Fourier Tauberian argument in [**ToZ2**, §8.4], which gives

$$(12.107)\qquad \frac{1}{N(h)} \sum_{h_j \geq h} |\chi^w_{\varepsilon_1, 2\varepsilon_1} h_j D_{x_n} \varphi_{h_j}|_H(x')|^2 = \mathcal{O}(\varepsilon_1^2)$$

uniformly for $x' \in H$. Since $\varepsilon_1 > 0$ is arbitrary, Theorem 12.36 follows from (12.101) and (12.106) by letting $\varepsilon_1 \to 0^+$ in (12.100).

12.24. Quantum ergodic restriction (QER) theorems for Dirichlet data

For applications to nodal sets and other problems, it is important to know if the Dirichlet data alone satisfies a QER theorem. The answer is obviously 'no' in general. For instance if (M, g) has an isometric involution and with a hypersurface H of fixed points, then any eigenfunction which is odd with respect to the involution vanishes on H. But in [**ToZ2, ToZ3**] a sufficient condition is given for quantum ergodic restriction, which rules out this and more general situations. The symmetry condition is that geodesics emanating from the 'left side' of H have a different return map from geodesics on the 'right side' when the initial conditions are reflections of each other through TH. To take the simplest example of the circle, the restriction of $\sin kx$ to a point is never quantum ergodic but the full Cauchy data $(\cos kx, \sin kx)$ of course satisfies $\cos^2 kx + \sin^2 kx = 1$. In [**CTZ**] it is proved that Cauchy data always satisfies QER for any hypersurface. This has implications for (at least complex) zeros of even or odd eigenfunctions along an axis of symmetry, e.g., for the case of Maass forms for the modular domain $SL(2, \mathbb{Z})/\mathbb{H}^2$.

Let

$$T^*_H M = \{(q, \xi) \in T^*_q M,\ q \in H\}, \quad T^* H = \{(q, \eta) \in T^*_q H,\ q \in H\}. \tag{12.108}$$

We further denote by $\pi_H : T^*_H M \to T^* H$ the restriction map,

$$\pi_H(x, \xi) = \xi|_{TH}. \tag{12.109}$$

For any orientable (embedded) hypersurface $H \subset M$, there exists two unit normal co-vector fields $\nu_\pm$ to H which span half ray bundles $N_\pm = \mathbb{R}_+ \nu_\pm \subset N^* H$. Infinitesimally, they define two 'sides' of H, indeed they are the two components of $T^*_H M \backslash T^* H$. We use Fermi normal coordinates (s, y_n) along H with $s \in H$ and with $x = \exp_x y_n \nu$ and let σ, η_n denote the dual symplectic coordinates. For $(s, \sigma) \in B^* H$ (the co-ball bundle), there exist two unit covectors $\xi_\pm(s, \sigma) \in S^*_s M$ such that $|\xi_\pm(s, \sigma)| = 1$ and $\xi|_{T_s H} = \sigma$. In the above orthogonal decomposition, they are given by

$$\xi_\pm(s, \sigma) = \sigma \pm \sqrt{1 - |\sigma|^2}\, \nu_+(s). \tag{12.110}$$

We define the reflection involution through $T^* H$ by

$$r_H : T^*_H M \to T^*_H M, \quad r_H(s, \mu \xi_\pm(s, \sigma)) = (s, \mu \xi_\mp(s, \sigma)), \quad \mu \in \mathbb{R}_+. \tag{12.111}$$

Its fixed point set is $T^* H$. We denote by G^t the homogeneous geodesic flow of (M, g), i.e., Hamiltonian flow on $T^* M \setminus 0$ generated by $|\xi|_g$. We define the *first return time* $T(s, \xi)$ on $S^*_H M$ by,

$$T(s, \xi) = \inf\{t > 0 : G^t(s, \xi) \in S^*_H M,\ (s, \xi) \in S^*_H M)\}. \tag{12.112}$$

By definition $T(s, \xi) = +\infty$ if the trajectory through (s, ξ) fails to return to H. Inductively, we define the jth return time $T^{(j)}(s, \xi)$ to $S^*_H M$ and the jth return map Φ^j when the return times are finite.

We define the first return map on the same domain by

$$\Phi : S^*_H M \to S^*_H M, \quad \Phi(s, \xi) = G^{T(s, \xi)}(s, \xi) \tag{12.113}$$

When G^t is ergodic, Φ is defined almost everywhere and is also ergodic with respect to Liouville measure $\mu_{L,H}$ on $S^*_H M$.

DEFINITION 12.38. We say that H has a positive measure of microlocal reflection symmetry if

$$\mu_{L,H}\left(\bigcup_{j\neq 0}^{\infty}\{(s,\xi)\in S^*_HM : r_HG^{T^{(j)}(s,\xi)}(s,\xi)=G^{T^{(j)}(s,\xi)}r_H(s,\xi)\}\right)>0.$$

Otherwise we say that H is asymmetric with respect to the geodesic flow.

The QER theorem we state below holds for both poly-homogeneous (Kohn-Nirenberg) pseudo-differential operators as in [**Ho2**] and also for semi-classical pseudo-differential operators on H [**Zw**] with essentially the same proof. To avoid confusion between pseudodifferential operators on the ambient manifold M and those on H, we denote the latter by $\mathrm{Op}_H(a)$ where $a\in S^0_{cl}(T^*H)$.

We further introduce the zeroth order homogeneous function

$$\gamma(s,y_n,\sigma,\eta_n)=\frac{|\eta_n|}{\sqrt{|\sigma|^2+|\eta_n|^2}}=\left(1-\frac{|\sigma|^2}{r^2}\right)^{\frac12},\quad (r^2=|\sigma|^2+|\eta_n|^2) \tag{12.114}$$

on T^*_HM and also denote by

$$\gamma_{B^*H}=(1-|\sigma|^2)^{\frac12} \tag{12.115}$$

its restriction to $S^*_HM=\{r=1\}$.

For homogeneous pseudo-differential operators, the QER theorem is as follows [**ToZ2, ToZ3, DZ**]:

THEOREM 12.39. *Let (M,g) be a compact manifold with ergodic geodesic flow, and let $H\subset M$ be a hypersurface. Let φ_{λ_j} denote the L^2-normalized eigenfunctions of Δ_g. If H has a zero measure of microlocal symmetry, then there exists a density-one subset S of $\mathbb{N}$ such that for $\lambda_0>0$ and $a(s,\sigma)\in S^0_{cl}(T^*H)$*

$$\lim_{\substack{\lambda_j\to\infty\\ j\in S}}\langle \mathrm{Op}_H(a)\gamma_H\varphi_{\lambda_j},\gamma_H\varphi_{\lambda_j}\rangle_{L^2(H)}=\omega(a),$$

where

$$\omega(a)=\frac{2}{\mathrm{Vol}(S^*M)}\int_{B^*H}a_0(s,\sigma)\gamma^{-1}_{B^*H}(s,\sigma)\,dsd\sigma.$$

Alternatively, one can write $\omega(a)=\frac{1}{\mathrm{Vol}(S^*M)}\int_{S^*_HM}a_0(s,\pi_H(\xi))\,d\mu_{L,H}(\xi)$. Note that $a_0(s,\sigma)$ is bounded but is not defined for $\sigma=0$, hence $a_0(s,\pi_H(\xi))$ is not defined for $\xi\in N^*H$ if $a_0(s,\sigma)$ is homogeneous of order zero on T^*H. The analogous result for semi-classical pseudo-differential operators is the following [**ToZ2, ToZ3, DZ**]:

THEOREM 12.40. *Let (M,g) be a compact manifold with ergodic geodesic flow, and let $H\subset M$ be a hypersurface. If H has a zero measure of microlocal symmetry, then there exists a density-one subset S of $\mathbb{N}$ such that for $a\in S^{0,0}(T^*H\times[0,h_0))$,*

$$\lim_{\substack{h_j\to 0^+\\ j\in S}}\langle \mathrm{Op}_{h_j}(a)\gamma_H\varphi_{h_j},\gamma_H\varphi_{h_j}\rangle_{L^2(H)}=\omega(a),$$

where

$$\omega(a)=\frac{2}{\mathrm{Vol}(S^*M)}\int_{B^*H}a_0(s,\sigma)\,\gamma^{-1}_{B^*H}(s,\sigma)\,dsd\sigma.$$

Examples of asymmetric curves on surfaces in the case where (M, g) is a finite area hyperbolic surface are the following:

- H is a geodesic circle;
- H is a closed horocycle of radius $r < \mathrm{inj}(M, g)$, the injectivity radius.
- H is a generic closed geodesic or an arc of a generic non-closed geodesic.

12.25. Time averaging

We begin the proofs of Theorems 12.39 and 12.40 by lifting the matrix elements back to M and time-averaging, as in Section 12.8. We use the identity (12.31) where as in (12.30),

$$\begin{cases} V(t;a) := U(-t)\gamma_H^* \operatorname{Op}_H(a)\gamma_H U(t), \\ \bar{V}_T(a) := \dfrac{1}{T}\displaystyle\int_{-\infty}^{\infty} \chi(T^{-1}t)\, V(t;a)\, dt, \\ \bar{V}_{T,R}(a) := \dfrac{1}{2R}\displaystyle\int_{-R}^{R} U(r)^* \bar{V}_T(a) U(r)\, dr. \end{cases} \tag{12.116}$$

A technical complication is that $\overline{V}_T(a)$ is a Fourier integral operator with fold singularities. To define a Fourier integral operator $\overline{V}_{T,\varepsilon}(a)$, we need to introduce cutoff operators to cutoff away from T^*H and from $N^*H \times 0_{T^*M} \cup 0_{T^*M} \times N^*H$. We let $\chi \in C_0^\infty(\mathbb{R}), [0,1]$ be a cutoff supported in $(-1-\delta, 1+\delta)$ with $\chi(t) = 1$ for $t \in [-1+\delta, 1-\delta]$, $\int_{-\infty}^{\infty} \chi(t) dt = 1$. For fixed $\varepsilon > 0$, we introduce two cutoff pseudo-differential. The first, $\chi_\varepsilon^{(tan)}(x, D) = \operatorname{Op}(\chi_\varepsilon^{(tan)}) \in \operatorname{Op}(S^0_{cl}(T^*M))$, has homogeneous symbol $\chi_\varepsilon^{(tan)}(x,\xi)$ supported in an ε-aperture conic neighborhood of $T^*H \subset T^*M$ with $\chi_\varepsilon^{(tan)} \equiv 1$ in an $\frac{\varepsilon}{2}$-aperture subcone. The second cutoff operator $\chi_\varepsilon^{(n)}(x, D) = \operatorname{Op}(\chi_\varepsilon^{(n)}) \in \operatorname{Op}(S^0_{cl}(T^*M))$ has its homogeneous symbol $\chi_\varepsilon^{(n)}(x,\xi)$ supported in an ε-conic neighborhood of N^*H with $\chi_\varepsilon^{(n)} \equiv 1$ in an $\frac{\varepsilon}{2}$ subcone. To simplify notation, define the total cutoff operator

$$\chi_\varepsilon(x, D) := \chi_\varepsilon^{(tan)}(x, D) + \chi_\varepsilon^{(n)}(x, D), \tag{12.117}$$

and put

$$(\gamma_H^* Op_H(a)\gamma_H)_{\geq \varepsilon} = (I - \chi_{\frac{\varepsilon}{2}})\gamma_H^* \operatorname{Op}_H(a)\gamma_H(I - \chi_\varepsilon), \tag{12.118}$$

and

$$(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\leq \varepsilon} = \chi_{2\varepsilon}\gamma_H^* \operatorname{Op}_H(a)\gamma_H \chi_\varepsilon. \tag{12.119}$$

By standard wave front calculus, it follows that

$$\gamma_H^* \operatorname{Op}_H(a)\gamma_H = (\gamma_H^* \operatorname{Op}_H(a)\gamma_H))_{\geq \varepsilon} + (\gamma_H^* \operatorname{Op}_H(a)\gamma_H))_{\leq \varepsilon} + K_\varepsilon, \tag{12.120}$$

where, $\langle K_\varepsilon \varphi_j, \varphi_j \rangle_{L^2(M)} = \mathcal{O}(\lambda_j^{-\infty})$. We then define

$$V_\varepsilon(t;a) := U(-t)(\gamma_H^* \operatorname{Op}_H(a)\gamma_H)_{\geq \varepsilon} U(t), \tag{12.121}$$

and

$$\overline{V}_{T,\varepsilon}(a) := \frac{1}{T}\int_{-\infty}^{\infty} \chi(T^{-1}t) V_\varepsilon(t;a)\, dt. \tag{12.122}$$

The next proposition provides a detailed description of $\overline{V}_{T,\varepsilon}(a)$ as a Fourier integral operator with local canonical graph away from its fold set and computes its principal symbol. After cutting off from the tangential singular set $\Sigma_T \subset T^*M \times$

T^*M and the the conormal sets $N^*H \times 0_{T^*M}, 0_{T^*M} \times N^*H$, $\bar{V}_T(a)$ becomes a Fourier integral operator $\overline{V}_{T,\varepsilon}(a)$ with canonical relation given by

(12.123)
$$\mathrm{WF}(\overline{V}_{T,\varepsilon}(a)) = \Big\{(x,\xi,x',\xi') \in T^*M \times T^*M : \text{there exists } t \in (-T,T) \text{ such that } \exp_x t\xi = \exp_{x'} t\xi' = s \in H,\ G^t(x,\xi)|_{T_sH} = G^t(x',\xi')|_{T_sH},\ |\xi| = |\xi'|\Big\}.$$

We decompose $\overline{V}_{T,\varepsilon}(a)$ into a pseudo-differential and a Fourier integral part according to the dichotomy that (x,ξ,x',ξ') in (12.123) satisfy either

$$\text{(i) } G^t(x,\xi) = G^t(x',\xi') \text{ or (ii) } G^t(x',\xi') = r_H G^t(x,\xi), \tag{12.124}$$

where r_H is the reflection map of T^*H in ((12.111). Thus,

$$\mathrm{WF}(\overline{V}_{T,\varepsilon}(a)) = \Delta_{T^*M \times T^*M} \cup \Gamma_T, \tag{12.125}$$

where

$$\begin{cases} \Delta_{T^*M \times T^*M} := \{(x,\xi,x,\xi) \in T^*M \times T^*M\}, \\ \Gamma_T = \displaystyle\bigcup_{(s,\xi) \in T^*_H M} \bigcup_{|t|<T} \{(G^t(s,\xi), G^t(r_H(s,\xi)))\}. \end{cases} \tag{12.126}$$

The two 'branches' or components intersect along the singular set

$$\Sigma_T := \bigcup_{|t|<T} (G^t \times G^t)\Delta_{T^*H \times T^*H}. \tag{12.127}$$

We further subscript Γ_T with ε to indicate the points $\Gamma_{T,\varepsilon}$ outside the support of the tangential cutoff.

Since $G^t(r_H(s,\xi)) = G^t r_H G^{-t} G^t(s,\xi)$, $\Gamma_{T,\varepsilon} \subset \Gamma_T \backslash \Sigma_T$ is the graph of a symplectic correspondence. That is, for any $\varepsilon > 0$, $\Gamma_{T,\varepsilon}$ is the union of a finite number $N_{T,\varepsilon}$ of graphs of partially defined canonical transformations

$$\mathcal{R}_j(x,\xi) = G^{t_j(x,\xi)} r_H G^{-t_j(x,\xi)}(x,\xi). \tag{12.128}$$

which we term H-reflection maps. Here $t_j(x,\xi)$ is the jth 'impact time', i.e. the time to the jth impact with H. We denote its domain (up to time T) by $\mathcal{D}^{(j)}_{T,\varepsilon}$. By homogeneity of $G^t \colon T^*M \to T^*M$, for all $j \in \mathbb{Z}$,

$$t_j(x,\xi) = t_j\left(x, \frac{\xi}{|\xi|}\right), \quad \xi \neq 0. \tag{12.129}$$

Proposition 12.41. *Fix $T, \varepsilon > 0$ and let $a \in S^0_{cl}(T^*H)$ with $a_H(s,\xi) = a(s,\xi|_H) \in S^0(T^*_H M))$. Then $\overline{V}_{T,\varepsilon}(a)$ is a Fourier integral operator with local canonical graph and possesses the decomposition*

$$\overline{V}_{T,\varepsilon}(a) = P_{T,\varepsilon}(a) + F_{T,\varepsilon}(a) + R_{T,\varepsilon}(a),$$

where

(i) *$P_{T,\varepsilon}(a) \in \mathrm{Op}_{cl}(S^0(T^*M))$ is a pseudo-differential operator of order zero with principal symbol*

$$a_{T,\varepsilon}(x,\xi) := \sigma(P_{T,\varepsilon}(a))(x,\xi) \tag{12.130}$$

$$= \frac{1}{T}\sum_{j \in \mathbb{Z}} (1-\chi_\varepsilon)(\gamma^{-1}a_H)(G^{t_j(x,\xi)}(x,\xi))\chi(T^{-1}t_j(x,\xi)), \tag{12.131}$$

where $t_j(x,\xi) \in C^\infty(T^*M)$ *are the impact times of the geodesic* $\exp_x(t\xi)$ *with* H*, and* γ *is defined by* (12.114).

(ii) $F_{T,\varepsilon}(a)$ *is a Fourier integral operator of order zero with canonical relation* $\Gamma_{T,\varepsilon}$.

$$F_{T,\varepsilon}(a) = \sum_{j=1}^{N_{T,\varepsilon}} F^{(j)}_{T,\varepsilon}(a), \tag{12.132}$$

where the $F^{(j)}_{T,\varepsilon}(a); j = 1, ..., N_{T,\varepsilon}$ *are zeroth-order homogeneous Fourier integral operators with*

$$\mathrm{WF}'(F^{(j)}_{T,\varepsilon}(a)) = \mathrm{graph}(\mathcal{R}_j) \cap \Gamma_{T,\varepsilon},$$

and symbol

$$\sigma(F^{(j)}_{T,\varepsilon})(x,\xi) = \frac{1}{T}(\gamma^{-1}a_H)(G^{t_j(x,\xi)}(x,\xi))\chi(T^{-1}t_j(x,\xi))\,|dxd\xi|^{\frac{1}{2}}.$$

(iii) $R_{T,\varepsilon}(a)$ *is a smoothing operator.*

12.26. Completion of the proofs of Theorems 12.39 and 12.40

By (12.31) and by Proposition 12.41, the weak* limits of the restricted matrix elements are those of

$$\langle \overline{V}_{T,\varepsilon}(a)\varphi_j,\varphi_j\rangle_{L^2(M)} = \langle P_{T,\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)} + \langle F_{T,\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)} + \langle R_{T,\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)}. \tag{12.133}$$

It is clear that $\langle R_{T,\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)} \to 0$ for the entire sequence of eigenfunctions. Since

$$\langle(\gamma_H^* \,\mathrm{Op}_H(a)\gamma_H))_{\geq\varepsilon}\varphi_j|_H,\varphi_j|_H\rangle_{L^2(H)} = \langle \overline{V}_{T,\varepsilon}(a)\varphi_j,\varphi_j\rangle_{L^2(M)}$$

it follows from (12.133) that

$$\limsup_{\lambda\to\infty}\frac{1}{N(\lambda)}\sum_{j:\lambda_j\leq\lambda}\Big|\langle(\gamma_H^*\,\mathrm{Op}_H(a)\gamma_H))_{\geq\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)} - \langle[P_{T,\varepsilon}(a)+F_{T,\varepsilon}(a)]\varphi_j,\varphi_j\rangle_{L^2(M)}\Big|^2 = 0. \tag{12.134}$$

12.26.1. Removing the $F_{T,\varepsilon}$ term. We now consider the Fourier integral matrix elements $\langle F_{T,\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)}$.

Lemma 12.42. *Suppose that* H *is an asymmetric hypersurface. Then there exists a subsequence of the eigenfunctions of density one for which* $\langle F_{T,\varepsilon}\varphi_j,\varphi_j\rangle_{L^2(M)}$ *tends to zero.*

Proof. It suffices to show that

$$\limsup_{\lambda\to\infty}\frac{1}{N(\lambda)}\sum_{j:\lambda_j\leq\lambda}\big|\langle F_{T,\varepsilon}(a)\varphi_j,\varphi_j\rangle_{L^2(M)}\big|^2 = 0. \tag{12.135}$$

This does not use ergodicity of the geodesic flow.

To prove (12.135) we first use the Schwartz inequality

$$\frac{1}{N(\lambda)}\sum_{j:\lambda_j\leq\lambda}\big|\langle F_{T,\varepsilon}(a)\varphi_j,\varphi_j\rangle_{L^2(M)}\big|^2 \leq \frac{1}{N(\lambda)}\sum_{j:\lambda_j\leq\lambda}\langle F_{T,\varepsilon}(a)^*F_{T,\varepsilon}(a)\varphi_j,\varphi_j\rangle_{L^2(M)} \tag{12.136}$$

to bound the variance sum by a trace. We then use the local Weyl law for Fourier integral operators associated to local canonical graphs,

$$\frac{1}{N(\lambda)} \sum_{j:\lambda_j \leq \lambda} \langle F\varphi_{\lambda_j}, \varphi_{\lambda_j} \rangle \to \int_{S\Gamma_F \cap \Delta_{T^*M}} \sigma_\Delta(F)\, d\mu_L, \tag{12.137}$$

where Γ_F is the canonical relation of F, $S\Gamma_F$ is the set of vectors of norm one, and $S\Gamma_F \cap \Delta_{T^*M}$ is its intersection with the diagonal of $T^*M \times T^*M$. Also, $\sigma_\Delta(F)$ is the (scalar) symbol in this set and $d\mu_L$ is Liouville measure. Thus, if Γ_F is a local canonical graph, the right side is zero unless the intersection has dimension $m = \dim M$. The microlocal asymmetry condition is precisely that the intersection has measure zero, i.e.,

$$\int_{S\Gamma_F \cap \Delta_{T^*M}} \sigma_\Delta(F) d\mu_L = 0.$$

□

12.26.2. Contribution of the pseudo-differential term $P_{T,\varepsilon}(a)$. In view of (12.135), it follows from (12.134) that

$$\limsup_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{j:\lambda_j \leq \lambda} \Big| \langle (\gamma_H^* \operatorname{Op}_H(a) \gamma_H))_{\geq \varepsilon} \varphi_j, \varphi_j \rangle_{L^2(M)} - \langle P_{T,\varepsilon}(a) \varphi_j, \varphi_j \rangle_{L^2(M)} \Big|^2 = 0. \tag{12.138}$$

Hence to complete the proof it suffices to show that

$$\limsup_{\lambda \to \infty} \frac{1}{N(\lambda)} \sum_{j:\lambda_j \leq \lambda} \big| \langle P_{T,\varepsilon}(a) \varphi_j, \varphi_j \rangle_{L^2(M)} - \omega(a) \big|^2 = 0. \tag{12.139}$$

Since $P_{T,\varepsilon}(a)$ is a pseudo-differential operator, this follows from the quantum ergodicity of the eigenfunctions on M. The limit state is precisely the Liouville average of the principal symbol of $P_{T,\varepsilon}(a)$.

This completes the proof of Theorem 12.39. The proof of the semi-classical case is similar.

Bibliography

[BFS] A. Bäcker, S. Fürstberger and R. Schubert, Poincaré Husimi representation of eigenstates in quantum billiards, Phys. Rev. E 70 (2004), 036204.

[BFSS] A. Bäcker, S. Fürstberger, R. Schubert, and F. Steiner, Behaviour of boundary functions for quantum billiards. J. Phys. A 35 (2002), no. 48, 10293–10310.

[BLR] C. Bardos, G. Lebeau, J. Rauch, Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary, SIAM J. Control Optim. 30 (1992), no. 5, 10241065.

[Ba] A. Barnett, *Asymptotic rate of quantum ergodicity in chaotic Euclidean billiards.* Comm. Pure Appl. Math. 59 (2006), no. 10, 1457–1488.

[BaH] A. Barnett and A. Hassell, *Boundary quasi-orthogonality and sharp inclusion bounds for large Dirichlet eigenvalues.* SIAM J. Numer. Anal. 49 (2011), no. 3, 1046–1063.

[BaH] A. H. Barnett and A. Hassel, Estimates on Neumann eigenfunctions at the boundary, and the "Method of Particular Solutions" for computing them, *Spectral geometry,* 195208, Proc. Sympos. Pure Math., 84, Amer. Math. Soc., Providence, RI, 2012. l arXiv:1107.2172.

[BaHT] A. Barnett, A. Hassell and M. Tacy, Comparable upper and lower bounds for boundary values of Neumann eigenfunctions and tight inclusion of eigenvalues (arXiv:1512.04165).

[Be] P. Berard, On the wave equation without conjugate points, Math. Zeit. **155** (1977), 249–276.

[Bou] J. Bourgain, *Geodesic restrictions and L^p-estimates for eigenfunctions of Riemannian surfaces*, Linear and complex analysis, 27–35, Amer. Math. Soc. Tranl. Ser. 2, 226, Amer. Math. Soc., Providence, RI, 2009.

[BraC1] L. Brandolini and L. Colzani, Localization and convergence of eigenfunction expansions. J. Fourier Anal. Appl. 5 (1999), no. 5, 431-447.

[BraC2] L. Brandolini and L. Colzani, Decay of Fourier transforms and summability of eigenfunction expansions. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 29 (2000), no. 3, 611-638.

[Bru1] Bruggeman, Roelof W. Fourier coefficients of automorphic forms. Lecture Notes in Mathematics, 865. Mathematische Lehrbcher und Monographien, II. Abteilung: Mathematische Monographien [Mathematical Textbooks and Monographs, Part II: Mathematical Monographs], 48. Springer-Verlag, Berlin-New York, 1981

[Bru2] R. W. Bruggeman, Fourier coefficients of cusp forms, Invent. Math. 45(1) (1978) 118.

[Brun] V. Bruneau, Semi-classical asymptotics of Riesz means. J. London Math. Soc. (2) 62 (2000), no. 2, 613-624.

[BouR1] J. Bourgain and Z. Rudnick, Restriction of toral eigenfunctions to hypersurfaces. C. R. Math. Acad. Sci. Paris 347 (2009), no. 21-22, 1249-1253.

[BouR2] J. Bourgain and Z. Rudnick, On the nodal sets of toral eigenfunctions. Invent. Math. 185 (2011), no. 1, 199237.

[BouR3] J. Bourgain and Z. Rudnick, On the geometry of the nodal lines of eigenfunctions of the two-dimensional torus. Ann. Henri Poincar 12 (2011), no. 6, 10271053.

[BouR4] J. Bourgain and Z. Rudnick, Restriction of toral eigenfunctions to hypersurfaces and nodal sets. Geom. Funct. Anal. 22 (2012), no. 4, 878937.

[Bu1] N. Burq, Lower bounds for shape resonances widths of long range Schrdinger operators, Amer. J. Math. 124 (2002), no. 4, 677-735.

[Bu2] N. Burq, Quantum ergodicity of boundary values of eigenfunctions: A control theory approach, to appear in Canadian Math. Bull. 48 (2005), 3-15 (math.AP/0301349).

[BuGT] N. Burq, P. Gerard and N. Tzvetkov, Restrictions of the Laplace-Beltrami eigenfunctions to submanifolds. Duke Math. J. 138 (2007), no. 3, 445-486.

[CSo] X. Chen and C.D. Sogge, A few endpoint geodesic restriction estimates for eigenfunctions. Comm. Math. Phys. 329 (2014), no. 2, 435-459.

[CHT] H, Christianson, A.Hassell and J. A. Toth, Exterior mass estimates and L2-restriction bounds for Neumann data along hypersurfaces. Int. Math. Res. Not. IMRN 2015, no. 6, 1638-1665.

[CTZ] H. Christianson, J. A. Toth and S. Zelditch, Quantum ergodic restriction for Cauchy Data: Interior QUE and restricted QUE, Math. Res. Lett. 20 (2013), no. 3, 465-475 (arXiv:1205.0286).

[CGV] L. Colzani, G. Gigante, and A. Vargas, Localization for Riesz means of Fourier expansions. Trans. Amer. Math. Soc. 366 (2014), no. 12, 6229-6245.

[C] R. Cooper, The extremal values of Legendre polynomials and of certain related functions. Proc. Cambridge Philos. Soc. 46, (1950). 54955.

[DG] J.J. Duistermaat and V. Guillemin, The spectrum of positive elliptic operators and periodic bicharacteristics, Invent. Math. 29 (1975), 39-79.

[DZ] S. Dyatlov, and M. Zworski, Quantum ergodicity for restrictions to hypersurfaces (arXiv:1204.0284).

[GL] P.Gérard and E.Leichtnam, Ergodic properties of eigenfunctions for the Dirichlet problem, Duke Math J. 71 (1993), 559-607.

[Gi] V. M. Gichev, A Note on the Common Zeros of Laplace Beltrami Eigenfunctions. Ann. Global Anal. Geome. 26, 201208 (2004).

[GLS] F. Golse, E. Leichtnam, and M. Stenzel, Intrinsic microlocal analysis and inversion formulae for the heat equation on compact real-analytic Riemannian manifolds. Ann. Sci. École Norm. Sup. (4) 29 (1996), no. 6, 669–736.

[G] A. Good, Local Analysis of Selberg's Trace Formula, Lecture Notes in Mathematics, Vol. 1040 (Springer-Verlag, 1983).

[GSt] V. Guillemin and S. Sternberg, *Geometric asymptotics.* Mathematical Surveys, No. 14. American Mathematical Society, Providence, R.I., 1977.

[HHHZ] X. Han, A. Hassell, H. Hezari, and S. Zelditch, Completeness of boundary traces of eigenfunctions, Proc. Lond. Math. Soc. (3) 111 (2015), no. 3, 749-773 (arXiv:1311.0935).

[HT] A. Hassell and T. Tao, Upper and lower bounds for normal derivatives of Dirichlet eigenfunctions. Math. Res. Lett. 9 (2002), no. 2-3, 289-305; Erratum for "Upper and lower bounds for normal derivatives of Dirichlet eigenfunctions" Math. Res. Lett. 17 (2010), no. 4, 793-794.

[HZ] A. Hassell and S. Zelditch, Quantum ergodicity of boundary values of eigenfunctions. Comm. Math. Phys. 248 (2004), no. 1, 119–168.

[H1] D.A. Hejhal, Dennis A. Sur certaines sries de Dirichlet associes aux godsiques fermes d'une surface de Riemann compacte. C. R. Acad. Sci. Paris Sr. I Math. 294 (1982), no. 8, 273-276; Sur quelques proprits asymptotiques des priodes hyperboliques et des invariants algbriques d'un sous-groupe discret de PSL(2,R). C. R. Acad. Sci. Paris Sr. I Math. 294 (1982), no. 15, 509-512.

[H2] D. A. Hejhal, Some Dirichlet series with coefficients related to periods of automorphic eigenforms. II. Proc. Japan Acad. Ser. A Math. Sci. 59 (1983), no. 7, 335338; Hejhal, Dennis A. Some Dirichlet series with coefficients related to periods of automorphic eigenforms. Proc. Japan Acad. Ser. A Math. Sci. 58 (1982), no. 9, 413417.

[Ho1] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[Ho2] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[Ho3] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[Hu] R. Hu, Lp norm estimates of eigenfunctions restricted to submanifolds. Forum Math. 21 (2009), no. 6, 1021-1052.

[I] H. Iwaniec, The spectral growth of automorphic L-functions. J. Reine Angew. Math. 428 (1992), 139159.

[JJ1] J. Jung, Quantitative quantum ergodicity and the nodal domains of Hecke-Maass cusp forms, Comm. Math. Phys. 348 (2016), no. 2, 603-653. (arXiv: 1301.6211).

[JJ2] J. Jung, Sharp bounds for the intersection of nodal lines with certain curves. J. Eur. Math. Soc. (JEMS) 16 (2014), no. 2, 273-288.

[KMW] M. Kimball, M. McKee and E. Wambach, A relative trace formula for a compact Riemann surface. Int. J. Number Theory 7 (2011), no. 2, 389429

[K] N.V. Kuznecov, The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture. Sums of Kloosterman sums. (Russian) Mat. Sb. (N.S.) 111(153) (1980), no. 3, 334-383, 479.

[Ma] S. Marshall, Geodesic restrictions of arithmetic eigenfunctions. Duke Math. J. 165 (2016), no. 3, 463-508.

[Mi] Y. Miyazaki, Asymptotic behavior of normal derivatives of eigenfunctions for the Dirichlet Laplacian. J. Math. Anal. Appl. 388 (2012), no. 1, 205-218.

[O] S. Ozawa, Asymptotic property of eigenfunction of the Laplacian at the boundary, Osaka J. Math. 30 (1993), no. 2, 303–314.

[P1] Mark A. Pinsky, Fourier inversion for piecewise smooth functions in several variables, Proc. Amer. Math. Soc. 118 (1993), no. 3, 903-910.

[P2] M. Pinsky, Pointwise Fourier inversion and related eigenfunction expansions. Comm. Pure Appl. Math. 47 (1994), no. 5, 653-681.

[PT] Mark A. Pinsky and Michael E. Taylor, Pointwise Fourier inversion: a wave equation approach, J. Fourier Anal. Appl. 3 (1997), no. 6, 647-703.

[Pit] N. J.E. Pitt, A sum formula for a pair of closed geodesics on a hyperbolic surface. Duke Math. J. 143 (2008), no. 3, 407-435.

[R] A. Reznikov, A uniform bound for geodesic periods of eigenfunctions on hyperbolic surfaces. Forum Math. 27 (2015), no. 3, 1569-1590.

[Sa1] Yu. G. Safarov, Asymptotics of a spectral function of a positive elliptic operator without a nontrapping condition. (Russian) Funktsional. Anal. i Prilozhen. 22 (1988), no. 3, 53-65, 96; translation in Funct. Anal. Appl. 22 (1988), no. 3, 213-223 (1989).

[Sa2] Yu. G. Safarov, Riesz means of a function of the distribution of the eigenvalues of an elliptic operator. (Russian. English summary) Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 163 (1987), Kraev. Zadachi Mat. Fiz. i Smezhn. Vopr. Teor. Funktsii 19, 143–145, 189–190; translation in J. Soviet Math. 49 (1990), no. 5, 1210-1212.

[SaV] Yu.G. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators*, Translated from the Russian manuscript by the authors. Translations of Mathematical Monographs, **155** American Mathematical Society, Providence, RI, 1997.

[So1] C. D. Sogge, *Fourier integrals in classical analysis.* Cambridge Tracts in Mathematics, 105. Cambridge University Press, Cambridge, 1993.

[So2] C.D. Sogge, *Kakeya-Nikodym averages and L^p-norms of eigenfunctions.* Tohoku Math. J. **63** (2011), 519-538.

[So3] C. D. Sogge, Problems related to the concentration of eigenfunctions (arXiv:1510.07723)

[SoXZ] C. D. Sogge, Y. Xi, and C. Zhang Geodesic period integrals of eigenfunctions on Riemann surfaces and the Gauss-Bonnet Theorem (arXiv:1604.03189).

[SoZ1] C. D. Sogge and S. Zelditch, Sup norms of Cauchy data of eigenfunctions on manifolds with concave boundary (arXiv:1411.1035).

[SoZ2] C.D. Sogge and S. Zelditch, On eigenfunction restriction estimates and L4-bounds for compact surfaces with nonpositive curvature. Advances in analysis: the legacy of Elias M. Stein, 447-461, Princeton Math. Ser., 50, Princeton Univ. Press, Princeton, NJ, 2014.

[T] D. Tataru, On the regularity of boundary traces for the wave equation. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 26 (1998), no. 1, 185206.

[Tay1] M.E. Taylor, *Pseudodifferential operators.* Princeton Mathematical Series, 34. Princeton University Press, Princeton, N.J., 1981.

[Tay2] Taylor, Michael E. The Gibbs phenomenon, the Pinsky phenomenon, and variants for eigenfunction expansions. Comm. Partial Differential Equations 27 (2002), no. 3-4, 565605.

[To] J.A. Toth, L^2-restriction bounds for eigenfunctions along curves in the quantum completely integrable case, Comm. Math. Phys. 288 (2009), no. 1, 379-401 (arXiv:0803.0978).

[ToZ1] J.A. Toth and S. Zelditch, Counting nodal lines which touch the boundary of an analytic domain. J. Differential Geom. 81 (2009), no. 3, 649-686.

[ToZ2] J. A. Toth and S. Zelditch, Quantum ergodic restriction theorems, I: interior hypersurfaces in analytic domains,, *Ann. H. Poincaré* 13, Issue 4 (2012), Page 599-670 (arXiv:1005.1636).

[ToZ3] J. A. Toth and S. Zelditch, Quantum ergodic restriction theorems, II: manifolds without boundary, GAFA: Volume 23, Issue 2 (2013), Page 715-775 (arXiv:1104.4531).

[ToZ4] J. A. Toth and S. Zelditch, Upper bound on the number of nodal points on curves, (preprint, 2017).

[Ts] M. Tsuzuki, Spectral square means for period integrals of wave functions on real hyperbolic spaces. J. Number Theory 129 (2009), no. 10, 23872438.

[U] K. Uhlenbeck, Generic properties of eigenfunctions. Amer. J. Math. 98 (1976), no. 4, 1059–1078.

[V] M. VanValkenburgh, Exponential lower bounds for quasimodes of semiclassical Schrdinger operators. Math. Res. Lett. 16 (2009), no. 4, 721-734.

[Y] M. P. Young, The quantum unique ergodicity conjecture for thin sets, arXiv: 1306.1554.

[Z] S. Zelditch, Kuznecov sum formulae and Szegö limit formulae on manifolds. Comm. Partial Differential Equations 17 (1992), no. 1-2, 221–260.

[Zw] M. Zworski, *Semiclassical analysis*, Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012.

CHAPTER 13

Nodal sets: Real domain

We now begin the study of nodal sets of eigenfunctions. The theory is somewhat analogous to the study of zero sets of polynomials in algebraic geometry with the frequency λ playing the role of the degree. The analogy is strongest when (M, g) is real analytic and is incomplete at this time for general C^∞ metrics, although the recent breakthrough work of Logunov (and Logunov-Malinnikova) has changed the situation at this time of writing [**L1, L2, LM**]. As in algebraic geometry, when g is real analytic the nodal sets can be studied in the complex domain (i.e., in the complexification $M_{\mathbb{C}}$ of M) and are simpler to study there. However, there does not exist an algebraic theory of nodal sets of eigenfunctions, and we can only rely on the equation $(\Delta + \lambda^2)\varphi_\lambda = 0$ to obtain information.

One of the principal problems is to estimate the $(n-1)$-dimensional Hausdorff (hypersurface volume) measure of the nodal set of φ_λ, which is denoted by

$$\mathcal{H}^{m-1}(Z_{\varphi_\lambda}). \tag{13.1}$$

A more elaborate problem is to determine the equidistribution of nodal sets. Other important problems are to determine the number of connected components, the number of nodal domains, the sizes and shapes of nodal domains or other local properties such as the integral curvature of components of the nodal set. These problems are analogues of well-studied problems for real algebraic varieties but are difficult to study by PDE methods alone.

Roughly speaking, there are three approaches to estimating nodal volumes, either from above or below.

- Local method of doubling estimates and frequency functions, using monotonicity properties of frequency functions of harmonic functions. Relation of growth of nodal sets to local growth of eigenfunctions or harmonic functions as measured by doubling exponents (or indices);
- Global construction of the delta-function on the nodal set and integral estimates;
- Estimate from above: Crofton formula and counting nodal points on curves.

There is a pronounced difference in techniques and results between C^∞ metrics and real analytic C^ω metrics, since many more methods are available in the latter case. In this section we mainly consider techniques available for general C^∞ metrics. In the next section we analytically continue eigenfunctions to the complexification of M for real analytic Riemannian manifolds and study the complex nodal set. The real nodal set is the real 'slice' of the complex nodal set and the slice might lack good transversality properties. The complexification $M_{\mathbb{C}}$ in the real analytic case is essentially the 'phase space' B^*M (a co-ball bundle) and the phase space zeros are simpler to relate to the dynamics of the geodesic flow than the configuration space zeros given by the real nodal set. In any case, results on the asymptotics of nodal sets in the real domain are rather few and rather weak when $\dim M > 2$. The purpose of this section is to explain some of the main results.

Just as this exposition was being finished, A. Logunov posted three articles [**L1, L2, LM**] which prove the Yau conjectured lower bound on nodal sets in the smooth setting and prove a polynomial upper bound. Time does not permit a detailed exposition of the new results but obviously they render many of the old ones obsolete, although the proofs may still have interest.

13.1. Fundamental existence theorem for nodal sets

A real algebraic polynomial need not have any real zeros. Since $\int_M \varphi_\lambda \, dV = 0$ it is clear than any non-constant eigenfunction must have zeros. But it is not clear how 'many' zeros it has or how dense the zero set is. The fundamental existence theorem of the title asserts that the zero set is $\lambda^{-\frac{1}{2}}$ dense. The proofs involve rescaling the eigenvalue problem in small balls, and thus the existence theorem is an elliptic result in the sense of §5.3.

THEOREM 13.1. *For any (M, g) there exists a constant $A > 0$ so that every ball of (M, g) of radius greater than $\frac{A}{\lambda}$ contains a nodal point of any eigenfunction φ_λ.*

PROOF. Let $D_{\lambda,j}$ be a nodal domain of φ_λ. Since $\varphi_\lambda > 0$ in a nodal domain (after multiplication by -1 if necessary), $\lambda^2 = \lambda_1^2(D_{\lambda,j})$ where $\lambda_1^2(D_{\lambda,j})$ is the smallest Dirichlet eigenvalue for the nodal domain.

Now fix x_0, r and consider a ball $B(x_0, r)$. If φ_λ has no zeros in $B(x_0, r)$, then $B(x_0, r) \subset D_{\lambda,j}$ for some j, i.e. the ball must be contained in the interior of a nodal domain $D_{\lambda,j}$ of φ_λ. But $\lambda_1^2(D_{\lambda,j})$ is domain monotonic: i.e., the lowest Dirichlet eigenvalue $\lambda_1(\Omega)$ of a domain decreases as Ω increases. Hence

$$\lambda^2 = \lambda_1^2(D_{\lambda,j}) \leq \lambda_1(B(x_0, r)). \tag{13.2}$$

We now show that

$$\lambda_1^2(B(x_0, r)) \leq \frac{C_g}{r^2}, \tag{13.3}$$

where C_g depends only on the metric. Granted (13.3), it follows from (13.2) and (13.3) that

$$\lambda^2 \leq \sup\left\{\frac{C_g}{r^2} : \varphi_\lambda > 0 \text{ in } B(x_0, r)\right\}. \tag{13.4}$$

Clearly, $r \leq \frac{C_g}{\lambda}$ for any r satisfying $\varphi_\lambda > 0$ in $B(x_0, r)$.

To complete the proof, we must prove the inequality (13.3). It is proved by comparing $\lambda_1^2(B(x_0, r))$ for the metric g with the lowest Dirichlet Eigenvalue $\lambda_1^2(B(x_0, cr); g_0)$ of the Euclidean ball $B(x_0, cr; g_0)$ centered at x_0 of radius cr. Here, g_0 is the Euclidean metric defined by freezing the coefficients of g at x_0; c is chosen so that $B(x_0, cr; g_0) \subset B(x_0, r, g)$. Again by domain monotonicity, $\lambda_1^2(B(x_0, r, g)) \leq \lambda_1^2(B(x_0, cr; g))$ for $c < 1$. By comparing the Rayleigh quotients $\frac{\int_\Omega |df|^2 dV_g}{\int_\Omega f^2 dV_g}$ one easily sees that

$$\lambda_1^2(B(x_0, cr; g)) \leq C \lambda_1^2(B(x_0, cr; g_0))$$

for some C depending only on the metric. But by explicit calculation with Bessel functions, $\lambda_1^2(B(x_0, cr; g_0)) \leq \frac{C}{r^2}$, and (13.3) follows. □

13.1.1. A second proof. Another proof is given in [**HL**]. Let u_r denote the ground state Dirichlet eigenfunction for $B(x_0, r)$. Then $u_r > 0$ on the interior of $B(x_0, r)$. If $B(x_0, r) \subset D_{\lambda,j}$ then also $\varphi_\lambda > 0$ in $B(x_0, r)$. Hence the ratio $\frac{u_r}{\varphi_\lambda}$ is smooth and non-negative, vanishes only on $\partial B(x_0, r)$, and must have its maximum at a point y in the interior of $B(x_0, r)$. It follows that

$$\nabla \left(\frac{u_r}{\varphi_\lambda}\right)(y) = 0, \;\; \Delta \left(\frac{u_r}{\varphi_\lambda}\right)(y) \leq 0.$$

But at a critical point y,

$$\Delta \left(\frac{u_r}{\varphi_\lambda}\right)(y) = \frac{\varphi_\lambda \Delta u_r(y) - u_r(y)\Delta\varphi_\lambda(y)}{\varphi_\lambda^2(y)} = -\frac{(\lambda_1^2(B(x_0, r)) - \lambda^2)\varphi_\lambda u_r}{\varphi_\lambda^2} \leq 0.$$

Since $\frac{\varphi_\lambda u_r}{\varphi_\lambda^2} > 0$, this is possible only if $\lambda_1^2(B(x_0, r)) \geq \lambda$.

To complete the proof we need to show that $\lambda_1(B(x_0, r)) \geq \lambda$ implies $r \leq \frac{A}{\lambda}$ for some A depending only on g. As before, we rescale the ball by $x \to \sqrt{\lambda}x$ (with normal coordinates centered at x_0) and obtain an essentially Euclidean ball of radius r. Then $\lambda_1(B(x_0, \frac{r}{\lambda})) = \lambda\lambda_1 B_{g_0}(x_0, r)$. Therefore we only need to choose r so that $\lambda_1 B_{g_0}(x_0, r) = 1$.

13.2. Curvature of nodal lines and level lines

Continuing the local discussion of nodal sets we present a few formulae for their geodesic curvature in dimension 2. Let $Q = |\nabla\varphi|^2$. We claim that in dimension two, the geodesic curvature η of the nodal line $\partial\Omega$ bounding a nodal domain of an eigenfunction φ is given by

$$\eta = \frac{1}{2}\nabla_\nu \log Q. \tag{13.5}$$

Here ν is the inward normal, $\nu = \frac{\nabla\varphi}{|\nabla\varphi|}$ is the unit normal pointing into $\{\varphi > c\}$.

To see this, we recall that the geodesic curvature of a contour line of a function u on a surface is given by

$$k = \frac{u_x^2 u_{xx} + u_y^2 u_{yy} - 2u_x u_y u_{xy}}{|\nabla u|^3}.$$

As in [**Sa, Ta**], it follows that the curvature k of the level lines of u and the curvature h of the gradient lines of u are given respectively by

$$\begin{cases} k = -\operatorname{div} \dfrac{\nabla\varphi}{|\nabla\varphi|}, \\ h = -\operatorname{div} J\dfrac{\nabla\varphi}{|\nabla\varphi|}, \end{cases} \tag{13.6}$$

where $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Following [**Sa**] we verify the formulae as follows: Let u be any smooth function on M and let $N = u^{-1}(c)$ be a level set. If $\nabla u(p) \neq 0$ then the trace of the second fundamental form of N at p is given by

$$\eta = \frac{\Delta u}{|\nabla u|} + \frac{1}{|\nabla u|}\nabla(|\nabla u|) \cdot \nu = \frac{\Delta u}{|\nabla u|} + \frac{1}{2}\nabla \log(|\nabla u|^2) \cdot \nu. \tag{13.7}$$

Recall that the second fundamental form of N is defined by

$$\ell(X, Y) = (\nabla_X Y - \nabla_Y^N Y) \cdot \nu.$$

Since $\nabla \perp N$, $\nabla^2 u|_{T_pN} = -|\nabla u|\ell$. Since $\mathrm{Tr}\,\nabla^2 f = -\Delta f$, and $\nabla_\nu \nu \cdot \nu = 0$, we have

$$\begin{cases}\mathrm{Tr}(\nabla^2 u|_{T_pN}) = \mathrm{Tr}\,\nabla^2 f u - \nabla^2 u(\nu,\nu), \\ \nabla^2 u(\nu,\nu) = \nabla(|\nabla u|)\cdot\nu.\end{cases}$$

We now specialize to the case where M is a surface and $u = \varphi$ is an eigenfunction. Let Ω be a nodal domain. Then on $\partial\Omega$, $\frac{\Delta u}{|\nabla u|} = 0$, so that (13.7) implies (13.5).

Remark 13.2. Green's formula implies that

$$\int_{\partial\Omega} \eta = \frac{1}{2}\int_\Omega \Delta \log Q\, dA. \tag{13.8}$$

It would be interesting to bound the right hand side above or below. For the related Cauchy data square, $q = |\nabla\varphi|^2 + \frac{1}{2}\lambda^2\varphi^2$ there is a distribution lower bound for $\Delta \log q$ by Dong (§13.7.4 or [**D**, Theorem 3.3]).

13.3. Sub-level sets of eigenfunctions

Colding-Minicozzi have used the $\frac{1}{\lambda}$-density of nodal sets together with small scale mean value inequalities to obtain lower bounds on sublevel sets of eigenfunctions [**CM2**].

Theorem 13.3. *Let (M^n, g) be a compact Riemannian manifold without boundary such that $\mathrm{Ric}(g) \geq -(n-1)$. Suppose that φ_λ is an L^2 normalized eigenfunction. Let $V = \mathrm{Vol}(M,g)$. Then there there exists $C_n > 0$ and $\Lambda(n) > 0$ so that*

$$\mathrm{Vol}\left(\left\{x \in M : |\varphi_\lambda|^2 < \frac{\varepsilon}{V}\right\}\right) \geq C\varepsilon^n\,\mathrm{Vol}(M,g).$$

Proof. We sketch the proof, leaving out some of the details from [**CM2**]. It was just proved above that there exists $C(M,g) > 0$ so that $\mathcal{N}_{\varphi_\lambda} \cap B(x_0, \frac{C}{\lambda}) \neq \emptyset$. Let $\{B_{\frac{C}{\lambda}}(x_k)\}_k$ be a maximal collection of disjoint balls of radius $\frac{C}{\lambda}$. By maximality, double the balls covers M. By the volume comparison theorem, the multiplicity of the covering is bounded by a constant C_n depending only on the dimension of M. Here, the volume comparison theorem states that if (M^n, g) complete and $\mathrm{Ric} \geq (n-1)K$ then for any $p \in M$,

$$\frac{\mathrm{Vol}(B(p,r))}{\mathrm{Vol}_K(B(p_K,r))} \text{ is a non-increasing function of r.}$$

Here, p_K is a point in the simply connected space form of constant curvature K and Vol_K denotes the volume in the space form. Thus,

$$\mathrm{Vol}(B(p,r)) \leq \mathrm{Vol}_K(B(p_K,r)), \quad r > 0.$$

Let $z_k \in \mathcal{N}_{\varphi_\lambda} \cap B_{\frac{C}{\lambda}}(x_k)$. Under the assumption $\mathrm{Ric} \geq -(n-1)$, one has the mean value inequality that if $\Delta v \geq -cv$ for some $c \geq 0$, then

$$\sup_{B_t} v \leq e^{C_3 + t\sqrt{c}}\mathrm{Ave}_{B_{2t}} v,$$

where

$$\mathrm{Ave}_B v = \frac{1}{\mathrm{Vol}(B)}\int_B v\, dV_g.$$

By the Bochner formula and the Ricci curvature assumption that

$$\Delta|\nabla\varphi_\lambda|^2 = 2|\,\mathrm{Hess}(\varphi_\lambda)|^2 + 2\lambda^2|\nabla\varphi_\lambda|^2 + 2\,\mathrm{Ric}(\nabla\varphi_\lambda, \nabla\varphi_\lambda) \geq -2(n-1-\lambda^2)|\nabla\varphi_\lambda|^2.$$

Hence there exists $C_n > 0$ so that (cf. [**SY**, p.80])

$$\sup_{B(z_k, \frac{s}{\lambda})} |\varphi_\lambda|^2 \leq s^2 \sup_{B(z_k, \frac{s}{\lambda})} |\nabla\varphi_\lambda|^2 \leq Cs^2 \frac{1}{\operatorname{Vol} B(z_k, \frac{s}{\lambda})} \int_{B(z_k, \frac{s}{\lambda})} |\nabla\varphi_\lambda|^2.$$

Here, we use that

$$f(x) - f(z_k) = \int \frac{d}{dt} f(\gamma(t))\, dt = \int \nabla f \cdot \gamma'(t)\, dt$$

where γ is a curve from z_k to x.

Since $z_k \in B(x_k, \frac{C}{\lambda}), B(z_k, \frac{2s}{\lambda}) \subset B(x_k, \frac{4s}{\lambda})$. Combining the above with the volume comparison theorem gives

(13.9)
$$\sup_{B(z_k, \frac{s}{\lambda})} |\varphi_\lambda|^2 \leq s^2 \sup_{B(x_k, \frac{4s}{\lambda})} |\nabla\varphi_\lambda|^2 = \text{const} \cdot \lambda^2 \frac{1}{\operatorname{Vol} B(x_k, \frac{4s}{\lambda})} \int_{B(x_k, \frac{4s}{\lambda})} |\nabla\varphi_\lambda|^2.$$

From the multiplicity bound for the cover it follows that

$$\sum_k \int_{B(x_k, \frac{4s}{\lambda})} |\nabla\varphi_\lambda|^2 \leq C\lambda^2.$$

Let $\{j\}$ be the subset of $\{k\}$ where

$$\frac{1}{\operatorname{Vol} B(x_k, \frac{4s}{\lambda})} \int_{B(x_k, \frac{4s}{\lambda})} |\nabla\varphi_\lambda|^2 \leq 2C\lambda^2. \tag{13.10}$$

Since the balls with twice the radius cover M,

$$\operatorname{Vol}(M) \leq 2 \sum_j \operatorname{Vol}(B(x_j, 4\frac{s}{\lambda}).$$

By the volume comparison theorem and since the original balls are disjoint, and $z_j \in B(x_j, \frac{s}{\lambda})$ one has

$$\operatorname{Vol}(M) \leq C \operatorname{Vol}\left(\bigcup_j B(z_j, \frac{s}{\lambda})\right). \tag{13.11}$$

The theorem follows by combining (13.9), (13.10) and (13.11). □

REMARK 13.4. This proof is not quite rigorous; one has to introduce a new parameter $\ell \geq 1$ with $s\ell = 4C_0/\lambda$ and let $\{B_{2\ell s}(x_k)\}$ be a maximal collection of balls. We refer to [**CM2**].

13.3.1. Inradius. It is known that in dimension two, the minimal possible area of a nodal domain of a Euclidean eigenfunction is $\pi(\frac{j_1}{\lambda})^2$. This follows from the two-dimensional Faber-Krahn inequality,

$$\lambda_k(\Omega)\text{Area}(D) = \lambda_1(D)\text{Area}(D) \geq \pi j_1^2,$$

where D is a nodal domain in Ω. In higher dimensions, the Faber-Krahn inequality shows that on any Riemannian manifold the volume of any nodal domain is $\geq C\lambda^{-n}$.

Another size measure of a nodal domain is its inradius r_λ, i.e., the radius of the largest ball contained inside the nodal domain. As can be seen from computer graphics, there are a variety of 'types' of nodal components. In [**M1**], Mangoubi proves that

$$\frac{C_1}{\lambda} \geq r_\lambda \geq \frac{C_2}{\lambda^{\frac{1}{2}k(n)}(\log \lambda)^{2n-4}}, \tag{13.12}$$

where $k(n) = n^2 - 15n/8 + 1/4$; note that eigenvalues in [**M4**] are denoted λ while here we denote them by λ^2. Recently, B. Georgiev [**Ge**] has improved the inradius estimate to

$$\frac{C_1}{\lambda} \geq r_\lambda \geq \frac{C_2}{\lambda^2}, \tag{13.13}$$

improving the result in dimensions $n \geq 5$.

In dimension 2, it is known (loc.cit.) that

$$\frac{C_1}{\lambda} \geq r_\lambda \geq \frac{C_2}{\lambda}. \tag{13.14}$$

13.4. Nodal sets of real homogeneous polynomials

One of the standard local techniques in studying nodal and singular sets of eigenfunctions is to approximate the eigenfunction in a neighborhood of a zero by a homogeneous harmonic polynomial q_d of degree d on $\mathbb{R}^n$ and to use knowledge from real algebraic geometry about nodal sets of homogeneous polynomials. Eigenfunctions are approximated by harmonic functions on small balls and the harmonic functions are approximated by polynomials. In this section we cite a few relevant and classical results.

The following Bezout type bound is proved in [**HS**, Theorem 2.1].

THEOREM 13.5. *Let $q\colon \mathbb{R}^n \to \mathbb{R}$ be a polynomial of degree $\leq d$ and suppose that* $\dim q^{-1}(0) \leq k$. *Then*

$$\mathcal{H}^k(q^{-1}(0) \cap B_1) \leq C d^{n-k}.$$

The proof uses the coordinate projections

$$p_\lambda\colon \mathbb{R}^n \to \mathbb{R}^k, \quad p_\lambda(x^1, \dots, x^n) = (x^{\lambda_1}, \dots, x^{\lambda_k})$$

for

$$\lambda \in \Lambda_{n,k} := \{(i_1, \dots, i_k) \in \mathbb{Z}^k,\ 1 \leq i_1 < \cdots < i_k \leq n\}.$$

If $\dim q^{-1}(0) \leq k$ then $q^{-1}(0) \cap p_\lambda^{-1}(0)$ is finite for $\lambda \in \Lambda_{n,k}$, and

$$\operatorname{card}(p_\lambda^{-1}(y) \cap q^{-1}(0)) \leq C d^{n-k}$$

for almost all y since this set is defined by the vanishing of a polynomial of degree $\leq d$ on a Euclidean space of dimension $n-k$. Indeed, the number of components of $q^{-1}(0)$ is $\leq Cd^n$. By Crofton's formula,

$$\mathcal{H}^{n-1}(q^{-1}(0) \cap B_1) \leq \sum_{\lambda \in \Lambda(n,k)} \int_{p_\lambda(B_1)} \operatorname{card}(p_\lambda^{-1}(y) \cap q^{-1}(0)) d\mathcal{L}^k(y) \leq d^{n-k}$$

Now add the assumption that φ is a non-constant harmonic homogenous polynomial of degree d on $\mathbb{R}^n$. By Theorem 13.5, with $k = n-1$,

$$\mathcal{H}^{n-1}(\varphi^{-1}(0) \cap B_1) \leq Cd.$$

Moreover, $\dim\{x : |\nabla \varphi(x)| = 0\} \leq n-2$. Indeed, real algebraic varieties are stratified manifolds, and if the critical set were of dimension $n-1$ there would exist an $(n-1)$-dimensional submanifold Y on which $\nabla \varphi = 0$. Then φ would be constant on Y with zero Cauchy data, contradicting unique continuation for harmonic functions. It also follows from Theorem 13.5 with $k = n-2$ that

LEMMA 13.6. *Let φ be a non-constant harmonic polynomial of degree d on $\mathbb{R}^n$. Then*

$$\mathcal{H}^{n-2}(\{x : \nabla\varphi(x) = 0\} \cap B_1) \leq Cd^2.$$

Hardt-Simon further prove an upper bound for the volume of a tube around the almost-critical set of a harmonic polynomial. It depends on two constants $(\theta = \theta(n), c = c(n))$ depending only on the dimension. Define the $(\theta\varepsilon)$-almost critical set of φ by

$$\mathcal{AC}(\varphi, \theta\varepsilon) := \{x : |\nabla\varphi| \leq (\theta\varepsilon)^{d-1}\}.$$

In [**HS**, Theorem 3.3] they prove

THEOREM 13.7. *If φ is a harmonic polynomial of degree d on $\mathbb{R}^n$ with $\sup_{B_1}|\varphi - \varphi(0)| = 1$ and $|\nabla\varphi(0)| \leq (\theta\varepsilon)^{d-1}$ then*

$$\mathcal{L}^n(B_1 \cap \{x : \mathrm{dist}(x, \mathcal{AC}(\varphi, \theta\varepsilon) < \varepsilon\}) \leq Cd^{2n+2}\varepsilon^2 \log \varepsilon^{-1}.$$

We refer to [**CNV**] for more recent results on polynomial approximation.

13.5. Rectifiability of the nodal set

We recall that the nodal set of an eigenfunction φ_λ is its zero set. When zero is a regular value of φ_λ the nodal set is a smooth hypersurface. This is a generic property of eigenfunctions [**U**]. It is pointed out in [**Bae**] that eigenfunctions can always be locally represented in the form

$$\varphi_\lambda(x) = v(x)\left(x_1^k + \sum_{j=0}^{k-1} x_1^j u_j(x')\right),$$

in suitable coordinates (x_1, x') near p, where φ_λ vanishes to order k at p, where $u_j(x')$ vanishes to order $k-j$ at $x' = 0$, and where $v(x) \neq 0$ in a ball around p. It follows that the nodal set is always countably $n-1$ rectifiable when $\dim M = n$.

13.5.1. Quantitatively transversal analytic functions.

In [**Do**], Donaldson studied zero sets of quantitatively transversal holomorphic sections of powers L^k of positive Hermitian line bundles $L \to M$ over Kähler manifolds. The analogous notion of quantitative transversality for eigenfunctions is the following:

DEFINITION 13.8. An eigenfunction is η-quantitatively transversal if there exists $\eta > 0$ such that

$$|\varphi_j(x)| \leq \eta \implies |\nabla\varphi_j(x)| \geq \lambda_j\eta. \tag{13.15}$$

Equivalently if $q(x) := \lambda^{-2}|\nabla\varphi(x)|^2 + \frac{1}{2}|\varphi|^2 \geq \eta^2$.

It is difficult to prove existence of η-QT (quantitatively transversal) holomorphic sections of line bundles and one may expect QT eigenfunctions to be rare. It is not clear that a generic Δ has any infinite sequence of η-QT eigenfunctions for any $\eta > 0$. If it does, they may form a sparse subsequence. On the standard S^2 one may expect existence of such eigenfunctions but the set of QT eigenfunctions to have a very low probability. Donaldson used ideas of Yomdin that were later simplified by Auroux [**Au**].

In [**Do**] Donaldson proved that zero sets of holomorphic quantitatively transversal sections of L^k are uniformly distributed in the sense that, for any $\psi \in \mathcal{D}^{n-1,n-1}(M)$,

$$\left| \int_{Z_{s_k}} \psi - \frac{1}{2\pi} \int_X \omega_k \wedge \psi \right| \leq C\sqrt{k} \|d\psi\|_{L^\infty}.$$

In [**SD**] a sharper bound was obtained for $\mathbb{CP}^1$, based on the fact that a sequence of quantitatively transversal sections with uniformly bounded sup-norms satisfies,

$$\left| \int_{\mathbb{CP}^1} \log \|p_k\|^2 \right| \leq C_0. \tag{13.16}$$

It is not known how the sup norm and L^2 norms are related for η-QT sections or eigenfunctions. An analogous result is proved in [**DF1**, Lemma 6.4] for general polynomials:

PROPOSITION 13.9. *Let P be a polynomial of degree d on $\mathbb{R}^n$ with* $\max_{S^{n-1}} |P| = 1$*. Then*

$$\int_{|\omega|=1} |\log |P(\omega)| d\omega| \leq Cd. \tag{13.17}$$

There is no QT assumption on P and we observe that the estimate for a QT section is better by an order of magnitude. This raises the question of generalizing (13.16) to QT eigenfunctions and of generalizing (13.17) to general eigenfunctions. The proof of (13.17) is based on facts about polynomials and their zeros and we reproduce it here:

PROOF. Use spherical coordinates (θ, φ) where θ is spherical coordinates on the hemisphere S^{n-1}_+ and $\varphi \in [-\pi, \pi]$ where $\varphi = 0$ is the north pole. For fixed θ write P on S^{n-1} as $P_1(\cos\varphi) + P_2(\cos\varphi)\sin\varphi$ and let

$$\bar{P} = P_1(\cos\theta) - \sin\varphi P_2(\cos\varphi), \quad Q = P\bar{P}.$$

For fixed θ,

$$Q_\theta(\cos\varphi) = \pm \prod_{j=1}^{d_\theta} \frac{(\cos\varphi - \alpha_\nu)}{1 - \alpha_\nu}$$

with $d_\theta \leq 2d$. Here $Q_\theta(1) = 1$. Then

$$\int_0^\pi \log |Q_\theta(\cos\varphi)| \sin^{n-2}\varphi \, d\varphi = \int_0^\pi \sum_{\nu=1}^{d_\theta} \log \left| \frac{\cos\varphi - \alpha_\nu}{1 - \alpha_\nu|} \right| \sin^{n-2}\varphi \, d\varphi \geq Cd.$$

Integrating over $\theta \in S^{n-2}_+$ proves the Lemma. □

Donaldson's results on holomorphic sections probably have more immediate analogues for analytic continuations of eigenfunctions to the complexification of M, which is the topic of the next chapter. If we analytically continue to the complexification of M then the quantitatively transversality condition becomes: there exists $\eta > 0$ such that

$$|\varphi_j^{\mathbb{C}} z)| \leq \eta \implies |\partial s_k(z)| \geq \lambda_j \eta. \tag{13.18}$$

13.5.2. Vanishing order. By the vanishing order $\nu(u,a)$ of u at a is meant the largest positive integer such that $D^\alpha u(a) = 0$ for all $|\alpha| \leq \nu$. A unique continuation theorem shows that the vanishing order of an eigenfunction at each zero is finite. The following estimate is a quantitative version of this fact. (See [**DF1**], [**Lin**, Proposition 1.2 and Corollary 1.4].)

THEOREM 13.10. *Suppose that M is compact and of dimension n. Then there exist constants $C(n), C_2(n)$ depending only on the dimension such that the the vanishing order $\nu(u,a)$ of u at $a \in M$ satisfies $\nu(u,a) \leq C(n)N(0,1) + C_2(n)$ for all $a \in B_{1/4}(0)$. In the case of a global eigenfunction, $\nu(\varphi_\lambda, a) \leq C(M,g)\lambda$.*

Highest weight spherical harmonics $C_N(x_1 + ix_2)^N$ on S^2 are examples which vanish at the maximal order of vanishing at the poles $x_1 = x_2 = 0, x_3 = \pm 1$. Gaussian beams on surfaces of revolution also do. It is curious that the points at which maximal vanishing order occurs are also poles at which maximal sup norm growth occurs for the very different zonal eigenfunctions. it is not clear if this coincidence holds more generally, and that suggests the

PROBLEM 13.11. for which (M,g) does there exist a subsequence of eigenfunctions φ_{j_k} and a sequence of points x_{j_k} so that φ_{j_k} achieves the maximal vanishing order estimate at φ_{j_k}?

A recent result due to [**Hez16**, Theorem 1.2] shows that quantum ergodic eigenfunctions on a negatively curved manifold never achieve maximal vanishing order. More significantly, he gives a quantitative improvement that, for all $\varepsilon > 0$,

$$\nu(\varphi_{\lambda_j}, p) \leq C_g (\log \lambda_j)^{-\frac{1}{2m}+\varepsilon} \lambda_j,$$

where $m = \dim M$. The result does not rule out an exceptional subsequence which vanishes to maximal order.

13.6. Doubling estimates

In this section we continue to discuss *doubling estimates*. They play a fundamental role in estimates of $\mathcal{H}^{n-1}(\mathcal{N}_\lambda)$.

THEOREM 13.12 (See [**DF1**], [**Lin**]). *Let φ_λ be a global eigenfunction of a C^∞ (M,g) there exists $C = C(M,g)$ and r_0 such that for $0 < r < r_0$,*

$$\frac{1}{\mathrm{Vol}(B_{2r}(a))} \int_{B_{2r}(a)} |\varphi_\lambda|^2 \, dV_g \leq e^{C\lambda} \frac{1}{\mathrm{Vol}(B_r(a))} \int_{B_r(a)} |\varphi_\lambda|^2], dV_g.$$

Further,

$$\max_{B(p,r)} |\varphi_\lambda(x)| \leq \left(\frac{r}{r'}\right)^{C\lambda} \max_{x \in B(p,r')} |\varphi_\lambda(x)|, \qquad 0 < r' < r. \tag{13.19}$$

The doubling estimates imply the vanishing order estimates. Let $a \in M$ and suppose that $u(a) = 0$. By the vanishing order $\nu(u,a)$ of u at a is meant the largest positive integer such that $D^\alpha u(a) = 0$ for all $|\alpha| \leq \nu$.

THEOREM 13.13. *Suppose that M is compact and of dimension n. Then there exist constants $C(n), C_2(n)$ depending only on the dimension such that the the vanishing order $\nu(u,a)$ of u at $a \in M$ satisfies $\nu(u,a) \leq C(n)\, N(0,1) + C_2(n)$ for all $a \in B_{1/4}(0)$. In the case of a global eigenfunction, $\nu(\varphi_\lambda, a) \leq C(M,g)\lambda$.*

Following [**NPS**] and [**Ro**], define the supnorm doubling exponent $\beta(\varphi, B)$ for a ball B by

$$\beta(\varphi, B) = \log \frac{\sup_B |\varphi|}{\sup_{\frac{1}{2}B} |\varphi|}.$$

More generally,

$$\beta(\varphi, B; \alpha) = \log \frac{\sup_B |\varphi|}{\sup_{\alpha B} |\varphi|}.$$

Donnelly-Fefferman proved that for any C^∞ metric,

$$\beta(\varphi_\lambda, B) \leq C\lambda.$$

Let

$$B_1(\lambda) = \int_M \beta(B_{a\lambda^{-1}}(x))\, dA(x).$$

The following was conjectured in [**NPS**] and proved by Roy-Fortin [**Ro**]:

Theorem 13.14. *For a C^∞ surface (M, g),*

$$C\lambda B_1(\lambda) \leq \mathcal{H}^1(Z_\lambda) \leq C'\lambda(B_1(\lambda) + 1).$$

Theorem 13.14 indicates that the set of points with maximal doubling indices on wave-length balls is of measure zero, and that the average value is a constant.

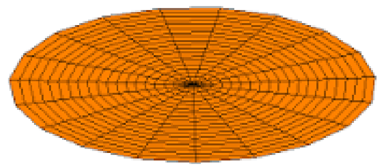
To get some intuition on the dependence of doubling indices on the center and radius, it is useful to consider the simplest example of r^N on $[0, T]$ for some T. This function vanishes to order N at $r = 0$ and models the (normalized) highest weight spherical harmonic $(x + iy)^N$ near the poles. Obviously, the doubling index $\beta(B_1(0))$ is $N \log 2$ for the interval $[0, 1]$. On the interval $[0, \frac{1}{N}]$ the doubling index is $\log \frac{(2/N)^N}{(1/N)^N} = N \log 2$, i.e. in balls centered at a point of maximal vanishing order the doubling index is scale invariant.

Now re-center the interval at some point a and let the radius be ρ and compare $(a + \rho)^N$ to $(a + 2\rho)^N$. Obviously,

$$\log \left(\frac{2\rho + a}{\rho + a} \right)^N = N \log 2 + N \log \frac{(1 + \frac{a}{2\rho})}{(1 + \frac{a}{\rho})}.$$

Let $t = \frac{a}{\rho}$ and consider $\log(1 + \frac{t}{2}) - \log(1 + t)$. This function is decreasing from 0 at $t = 0$ to $-\log 2$ as $t \to \infty$, canceling the $\log 2$ term. Hence the doubling exponent is non-uniform as the radius and center change. In order that $t \to \infty$ either $a \to \infty$ (impossible on a compact manifold) or $\rho \to 0$. On wavelength scales $\rho = \frac{1}{N}$ the doubling index is $N \log 2 + N \log \frac{(1+\frac{Na}{2})}{(1+(aN)}$ Again we observe that if $a > 0$ and $N \to \infty$, the doubling index decreases to 0 at a rate depending on a.

13.6.1. Doubling index to lower bound. Intuition suggests that the highest 'concentration' of the nodal set occurs at singular points where φ vanishes to order $\simeq \lambda$, or more generally where the doubling index is of order λ. In dimension 2, the first statement is provable, but the higher dimensional case is complicated. One does not expect the latter statement to be exactly true. By the local structure of eigenfunctions around a point of vanishing order λ in dimension two, one sees there are $\simeq \lambda$ 'spokes' in the nodal set emanating from the singular point, and the density of the nodal set is λ times the generic density.



13.7. Lower bounds for $\mathcal{H}^{m-1}(\mathcal{N}_\lambda)$ for C^∞ metrics

S. T. Yau has conjectured that for any C^∞ metric, there exist $c_1, C_2 > 0$ (depending only on g) so that

$$c_1\lambda \lesssim \mathcal{H}^{m-1}(Z_{\varphi_\lambda}) \lesssim C_2\lambda.$$

In dimension 2, the lower bound was proved by Brüning [**Br**] and by Yau (unpublished). In 1988, Donnelly-Fefferman [**DF1**] proved the conjectured upper and lower bounds bounds for real analytic Riemannian manifolds (possibly with boundary). We re-state the result as the following

THEOREM 13.15. *Let (M, g) be a compact real analytic Riemannian manifold, with or without boundary. Then*

$$c_1\lambda \lesssim \mathcal{H}^{m-1}(Z_{\varphi_\lambda}) \lesssim \lambda.$$

We give a new proof of this theorem from [**Z3**] in the §14.30. The proof of the upper bound is based on complexifying the nodal set and using plurisubharmonic theory in Grauert tubes. The lower bound is based on the fundamental existence Theorem 13.1, and otherwise uses almost no facts about eigenfunctions per se. Rather it is based on pure complex analysis. We present a new proof due to A. Brudnyi.

Logunov has proved the conjectured lower bound in the general C^∞ case [**L1**]. In this section, we review an older and non-sharp lower bound. Although it is now obsolete, we have retained the proof from the original lectures since it is based on a global integral geometry formula that as yet has not been connected to the local analysis in [**L1**].

In this section we review the lower bounds on $\mathcal{H}^{n-1}(Z_{\varphi_\lambda})$ from [**CM1, SoZ1, SoZ2, HS, HeW**]. Here

$$\mathcal{H}^{n-1}(Z_{\varphi_\lambda}) = \int_{Z_{\varphi_\lambda}} dS$$

is the Riemannian surface measure, where dS denotes the Riemannian volume element on the nodal set, i.e., the insert $\iota_n dV_g$ of the unit normal into the volume form of (M, g). As mentioned above, Logunov has recently proved the Yau lower bound, so the lower bound presented here is now obsolete.

THEOREM 13.16. *Let (M, g) be a C^∞ Riemannian manifold. Then there exists a constant C independent of λ such that*

$$C\lambda^{1-\frac{n-1}{2}} \leq \mathcal{H}^{n-1}(Z_{\varphi_\lambda}).$$

We sketch the proof of Theorem 13.16 from [**SoZ1, SoZ2**]. The starting point is an identity from [**SoZ1**] (inspired by an identity in [**D**]):

PROPOSITION 13.17. *For any* $f \in C^2(M)$,

$$\int_M |\varphi_\lambda| (\Delta_g + \lambda^2) f \, dV_g = 2 \int_{Z_{\varphi_\lambda}} |\nabla_g \varphi_\lambda| f \, dS, \tag{13.20}$$

When $f \equiv 1$ we obtain

COROLLARY 13.18.

$$\lambda^2 \int_M |\varphi_\lambda| \, dV_g = 2 \int_{Z_{\varphi_\lambda}} |\nabla_g \varphi_\lambda| f \, dS, \tag{13.21}$$

PROOF OF PROPOSITION 13.17. The nodal domains form a partition of M, i.e.,

$$M = \bigcup_{j=1}^{N_+(\lambda)} D_j^+ \cup \bigcup_{k=1}^{N_-(\lambda)} D_k^- \cup \mathcal{N}_\lambda,$$

where the D_j^+ and D_k^- are the positive and negative nodal domains of φ_λ, i.e., the connected components of the sets $\{\varphi_\lambda > 0\}$ and $\{\varphi_\lambda < 0\}$.

Let us assume for the moment that 0 is a regular value for φ_λ. Then each D_j^+ has smooth boundary ∂D_j^+, and so if ∂_ν is the Riemann outward normal derivative on this set, by the Gauss-Green formula we have

$$\int_{D_j^+} ((\Delta + \lambda^2) f) |\varphi_\lambda| \, dV = \int_{D_j^+} ((\Delta + \lambda^2) f) \varphi_\lambda \, dV \tag{13.22}$$

$$= \int_{D_j^+} f \, (\Delta + \lambda^2) \varphi_\lambda dV - \int_{\partial D_j^+} f \partial_\nu \varphi_\lambda \, dS \tag{13.23}$$

$$= \int_{\partial D_j^+} f |\nabla \varphi_\lambda| \, dS. \tag{13.24}$$

We use that $-\partial_\nu \varphi_\lambda = |\nabla \varphi_\lambda|$ since $\varphi_\lambda = 0$ on ∂D_j^+ and φ_λ decreases as it crosses ∂D_j^+ from D_j^+. A similar argument shows that

$$\int_{D_k^-} ((\Delta + \lambda^2) f) |\varphi_\lambda| \, dV = \int_{\partial D_k^-} f \, |\nabla \varphi_\lambda| \, dS, \tag{13.25}$$

using that φ_λ increases as it crosses ∂D_k^- from D_k^-.

If we sum these two identities over j and k, we get

$$\int_M ((\Delta + \lambda^2) f) |\varphi_\lambda| \, dV = \sum_j \int_{D_j^+} ((\Delta + \lambda^2) f) |\varphi_\lambda| \, dV + \sum_k \int_{D_k^-} ((\Delta + \lambda^2) f) |\varphi_\lambda| \, dV \tag{13.26}$$

$$= \sum_j \int_{\partial D_j^+} f |\nabla \varphi_\lambda| \, dS + \sum_k \int_{\partial D_k^-} f |\nabla \varphi_\lambda| \, dS \tag{13.27}$$

$$= 2 \int_{\mathcal{N}_\lambda} f |\nabla \varphi_\lambda| \, dS, \tag{13.28}$$

using the fact that $\mathcal{N}_\lambda$ is the disjoint union of the ∂D_j^+ and the disjoint union of the ∂D_k^-. □

Corollary 13.18 implies

(13.29)
$$\lambda^2 \int_M |\varphi_\lambda| \, dV = 2 \int_{Z_\lambda} |\nabla_g \varphi_\lambda|_g \, dS \leq 2|Z_\lambda| \, \|\nabla_g \varphi_\lambda\|_{L^\infty(M)} \lesssim 2|Z_\lambda| \lambda^{1+\frac{n-1}{2}} \|\varphi_\lambda\|_{L^1(M)}.$$

Thus Theorem 13.16 follows from the somewhat curious cancellation of $\|\varphi_\lambda\|_{L^1}$ from the two sides of the inequality.

13.7.1. More on L^1 norms and nodal sets. Hezari-Sogge modified the proof Proposition 13.17 in [**HS**] to prove

THEOREM 13.19. *For any C^∞ compact Riemannian manifold, the L^2-normalized eigenfunctions satisfy*
$$\mathcal{H}^{n-1}(Z_{\varphi_\lambda}) \geq C\lambda \|\varphi_\lambda\|^2_{L^1}.$$

They first apply the Schwarz inequality to get

$$\lambda^2 \int_M |\varphi_\lambda| \, dV_g \leq 2(\mathcal{H}^{n-1}(Z_{\varphi_\lambda}))^{1/2} \left(\int_{Z_{\varphi_\lambda}} |\nabla_g \varphi_\lambda|^2 \, dS \right)^{1/2}. \tag{13.30}$$

They then use the test function

$$f = \left(1 + \lambda^2 \varphi_\lambda^2 + |\nabla_g \varphi_\lambda|^2_g \right)^{\frac{1}{2}} \tag{13.31}$$

in Proposition 13.17 to show that

$$\int_{Z_{\varphi_\lambda}} |\nabla_g \varphi_\lambda|^2 \, dS \leq \lambda^3. \tag{13.32}$$

See also [**Ar**] for the generalization to the nodal bounds to Dirichlet and Neumann eigenfunctions of bounded domains.

Theorem 13.19 shows that Yau's conjectured lower bound would follow for a sequence of eigenfunctions satisfying $\|\varphi_\lambda\|_{L^1} \geq C > 0$.

13.7.2. Lower bounds on L^1 norms of eigenfunctions. The following universal lower bound is optimal as (M, g) ranges over all compact Riemannian manifolds.

PROPOSITION 13.20. *For any (M, g) and any L^2-normalized eigenfunction,*
$$\|\varphi_\lambda\|_{L^1} \geq C_g \lambda^{-\frac{n-1}{2}}.$$

REMARK 13.21. There are few results on L^1 norms of eigenfunctions. The reason is probably that $|\varphi_\lambda|^2 \, dV$ is the natural probability measure associated to eigenfunctions. It is straightforward to show that the expected L^1 norm of random L^2-normalized spherical harmonics of degree N and their generalizations to any (M, g) is a positive constant C_N with a uniform positive lower bound. One expects eigenfunctions in the ergodic case to have the same behavior.

PROBLEM 13.22. A difficult but interesting problem would be to show that $\|\varphi_\lambda\|_{L^1} \geq C > 0$ on a compact hyperbolic manifold. A partial result in this direction would be useful.

13.7.3. Nodal sets of solutions of more general equations. Nodal problems are of interest for many other elliptic equations, and we briefly discuss some recent results.

One equation is the Schrödinger equation

$$\Big(-\frac{\hbar^2}{2}\Delta + V\Big)\psi_{\hbar,E} = E(\hbar)\psi_{\hbar,E}.$$

At this time of writing, none of the techniques for $V = 0$ in the C^∞ setting have been generalized to non-zero V.

Another setting is the Steklov eigenvalue problem

$$\begin{cases} \Delta u = 0, & \text{in } \Omega, \\ \frac{\partial u}{\partial \nu} = \lambda u, & \text{on } \partial\Omega. \end{cases}$$

The Steklov eigenvalue problem is equivalent to the eigenvalue problem for the Dirichlet-to-Neumann operator, and the nodal set problem concerns the zero sets of $u|_{\partial\Omega}$. In [**Bel**], Bellova proves that in the case of a real analytic domain $\Omega \subset \mathbb{R}^n$,

$$\mathcal{H}^{n-2}(\{x \in \partial\Omega : u(x) = 0\}) \leq C\lambda^6$$

but points out that the expected upper bound is $C\lambda$. The Dirichlet to Neumann operator is a non-local pseudo-differential operator of order 1 and so it is not simple to construct the delta-function on the nodal set.

13.7.4. Dong's upper bound on number of singular points. Let (M, g) be a compact C^∞ Riemannian manifold of dimension n. Let

$$q = |\nabla\varphi|^2 + \lambda^2\varphi^2. \tag{13.33}$$

In [**D**, Theorem 2.2], R. T. Dong proves the bound

$$\mathcal{H}^{n-1}(\mathcal{N} \cap \Omega) \leq \frac{1}{2}\int_\Omega |\nabla \log q| + \sqrt{n}\,\mathrm{Vol}(\Omega)\lambda + \mathrm{Vol}(\partial\Omega). \tag{13.34}$$

He also proves (Theorem 3.3) that on a surface,

$$\Delta \log q \geq -\lambda + 2\min(K, 0) + 4\pi\sum_i (k_i - 1)\delta_{p_i}, \tag{13.35}$$

where $\{p_i\}$ are the singular points and k_i is the order of p_i. In Dong's notation, $\lambda > 0$. Using a weak Harnack inequality, Dong shows (Theorem 4.2) how (13.35) and (13.34) combine to produce the upper bound $\mathcal{H}^1(\mathcal{N} \cap \Omega) \leq \lambda^{3/2}$ in dimension 2.

Problem 13.23. To what extent can one generalize these estimates to higher dimensions?

13.7.5. Colding-Minicozzi estimate. The Colding-Minicozzi argument (as in Donnelly-Fefferman) is based on covering M by a good cover of balls B_i of some radius r so that each point is contained in some fixed number C_M of the double balls. Given $d > 1$, the balls are grouped into *d-good* and *d-bad* classes. A ball is d-good if

$$\int_{2B_i} \varphi_\lambda^2 \leq 2^d \int_{B_i} \varphi_\lambda^2.$$

Let G_d be the union of the d-good balls.

LEMMA 13.24. *There exists d_M depending only on C_M so that if $d \geq d_M$,*

$$\int_{G_d} \varphi_\lambda^2 \geq \frac{3}{4}.$$

To prove this it suffices to get a small upper bound on $\int_{B_d} \varphi_\lambda^2$ where B_d is the union of the bad balls. But

$$\int_{B_d} \varphi_\lambda^2 \leq \sum_{B_i \text{ bad}} \int_{B_i} \varphi_\lambda^2 \leq \sum_i 2^{-d} \int_{2B_i} \varphi_\lambda^2 \leq 2^{-d} C_M.$$

Choose d_M so that $2^{-d_M} C_M = \frac{1}{4}$.

LEMMA 13.25. *There exists C_M so that there exist at least $C_M \lambda^{\frac{n+1}{4}}$ balls which are d_M-good.*

Let N be the number of d_M-good balls. Given any $p > 2$, we have

$$\int_G \varphi_\lambda^2 \leq (\mathrm{Vol}(G))^{\frac{p-2}{2}} \|\varphi_\lambda\|_{L^p}^2.$$

Raising both sides to the power $\frac{p}{p-2}$ gives

$$\left(\frac{3}{4}\right)^{\frac{p}{p-2}} \|\varphi_\lambda\|_{L^p}^{-\frac{2p}{p-2}} \leq \left(\int_G \varphi_\lambda^2\right)^{\frac{p}{p-2}} \|\varphi_\lambda\|_{L^p}^{-\frac{2p}{p-2}} \leq \mathrm{Vol}(G).$$

The Sogge eigenfunction bound for $p \leq \frac{2(n+1)}{n-1}$ gives

$$C_p \lambda^{\frac{1-n}{4}} = C_p \left(\lambda^{\frac{(n-1)(p-2)}{8p}}\right)^{-\frac{2p}{p-2}} \leq \mathrm{Vol}(G).$$

Since $\mathrm{Vol}(B_i) \leq C'_M r^n = C' \lambda^{-\frac{n}{2}}$, this gives

$$C \lambda^{\frac{1-n}{4}} \leq \mathrm{Vol}(G) \leq N C \lambda^{-\frac{n}{2}}.$$

The sharpest bound comes from $p = \frac{2(n+1)}{n-1}$.

A local lower bound for nodal volumes is

PROPOSITION 13.26. *Let $d > 1$ and $\rho > 1$. Then there exist $\mu > 0$ and $\bar{\lambda}$ so that if $\Delta \varphi_\lambda = -\lambda \varphi_\lambda$ on $B_r(p)$ with $r \leq \rho \lambda^{-\frac{1}{2}}, \lambda \geq \bar{\lambda}\mu$, if φ_λ vanishes somewhere in $B_{\frac{r}{3}}(p)$ and if*

$$\int_{B_{2r}(p)} \varphi_\lambda^2 \leq 2^d \int_{B_r(p)} \varphi_\lambda^2,$$

then

$$\mathcal{H}^{n-1}(B_r(p)) \cap \{\varphi_\lambda = 0\}) \geq \mu r^{n-1}.$$

The Proposition implies the Theorem: if B_i is any of the balls in the cover, then there exists a nodal point in $\frac{1}{3} B_i$. If B_i is d_M-good, then the Proposition above (with $d = d_M, r = a\lambda^{-\frac{1}{2}}$) gives

$$\mathcal{H}^{n-1}(B_i \cap \{\varphi_\lambda = 0\}) \geq C_1 \lambda^{-\frac{n-1}{2}}.$$

Then

$$\mathcal{H}^{n-1}\{\varphi_\lambda = 0\} \geq C_m^{-1} \sum_{B_i \text{ good}} \mathcal{H}^{n-1}(B_i \cap \{\varphi_\lambda = 0\}) \geq C \lambda^{\frac{n+1}{4}} \lambda^{-\frac{n-1}{2}}.$$

It remains to prove the Proposition. The first step is to prove the mean value inequality: If $\varphi_\lambda(p) = 0$ then

$$\left|\int_{B_r(p)} \varphi_\lambda\right| \leq \frac{1}{3}\int_{B_r(p)} |\varphi_\lambda|. \tag{13.36}$$

We omit the proof and refer to [**CM1**] for this step. Granted (13.36), we complete the proof. Let $q \in B_{\frac{r}{3}}$ with $\varphi_\lambda(q) = 0$. Then

$$B_r(p) \subset B_{\frac{r}{3}}(q), \quad B_{\frac{5r}{3}}(q) \subset B_{2r}(p).$$

Since r is of order $\lambda^{-\frac{1}{2}}$, apply the mean value inequality to φ_λ^2 to get

$$\sup_{B_{\frac{4r}{3}}} \varphi_\lambda^2 \leq C_0 r^{-n}\int_{B_{2r}(p)} \varphi_\lambda^2 \leq C_0 2^d r^{-n}\int_{B_r(p)} \varphi_\lambda^2 \leq C_0 2^d r^{-n}\int_{B_{\frac{4r}{3}}(q)} \varphi_\lambda^2. \tag{13.37}$$

Then use the reverse Hölder inequality for integrals over $B_{\frac{4r}{3}}(q)$:

$$\left(\int \varphi_\lambda^2\right)^2 \leq \sup \varphi_\lambda^2 \left(\int |\varphi_\lambda|\right)^2 \leq C_0 2^d r^{-n}\left(\int \varphi_\lambda^2\right)\left(\int |\varphi_\lambda|\right)^2.$$

This simplifies to

$$\int \varphi_\lambda^2 \leq C_0 2^d r^{-n}\left(\int |\varphi_\lambda|\right)^2.$$

To deal with the integral of $|\varphi_\lambda|$ let $\varphi_\lambda^\pm$ be the positive/negative parts of φ_λ. Then

$$\int \varphi_\lambda^\pm \geq \frac{1}{3}\int |\varphi_\lambda|.$$

Let $B^+ \subset B_{\frac{4r}{3}}(q) \cap \{\varphi_\lambda > 0\}$ and let B_- be the corresponding negative part. Apply Cauchy-Schwartz to φ_λ^+ gives

$$\frac{1}{9}\left(\int |\varphi_\lambda|\right)^2 \leq \left(\int \varphi_\lambda^+\right)^2 \leq \mathrm{Vol}(B^+)\int \varphi_\lambda^2 \leq \mathrm{Vol}(B^+) C_0 2^d r^{-n}\left(\int |\varphi_\lambda|\right)^2.$$

Dividing by the square of $\|\varphi_\lambda\|_{L^1}^2$ gives

$$\frac{r^n}{9C_0 2^d} \leq \mathrm{Vol}(B^+).$$

The same argument applies to φ_λ^- to give the same lower bound for $\mathrm{Vol}(B^-)$. Then apply the isoperimetric inequality to get the lower bound for the measure of the nodal set in B, completing the proof of the Proposition.

13.7.6. Examples.

13.7.6.1. *Flat tori.* We have $|\nabla \sin\langle k, x\rangle|^2 = \cos^2\langle k, x\rangle |k|^2$. Since $\cos\langle k, x\rangle = 1$ when $\sin\langle k, x\rangle = 0$ the integral is simply $|k|$ times the surface volume of the nodal set, which is known to be of size $|k|$. Also, we have $\int_{\mathbf{T}} |\sin\langle k, x\rangle| dx \geq C$. Thus, our method gives the sharp lower bound $\mathcal{H}^{n-1}(Z_{\varphi_\lambda}) \geq C\lambda^1$ in this example.

So the upper bound is achieved in this example. Also, we have $\int_T |\sin\langle k, x\rangle|\, dx \geq C$. Thus, our method gives the sharp lower bound $\mathcal{H}^{n-1}(Z_{\varphi_\lambda}) \geq C\lambda^1$ in this example. Since $\cos\langle k, x\rangle = 1$ when $\sin\langle k, x\rangle = 0$ the integral is simply $|k|$ times the surface volume of the nodal set, which is known to be of size $|k|$.

13.7.6.2. *Spherical harmonics on S^2.* The L^1 of Y_0^N norm can be derived from the asymptotics of Legendre polynomials

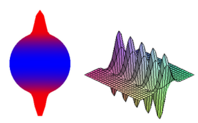
$$P_N(\cos\theta) = \sqrt{2}(\pi N \sin\theta)^{-\frac{1}{2}} \cos\left((N+\frac{1}{2})\theta - \frac{\pi}{4}\right) + O(N^{-3/2})$$

where the remainder is uniform on any interval $\varepsilon < \theta < \pi - \varepsilon$. We have

$$\|Y_0^N\|_{L^1} = 4\pi\sqrt{\frac{(2N+1)}{2\pi}} \int_0^{\pi/2} |P_N(\cos r)|\, dv(r) \sim C_0 > 0,$$

i.e., the L^1 norm is asymptotically a positive constant. Hence $\int_{Z_{Y_0^N}} |\nabla Y_0^N|\, ds \simeq C_0 N^2$. In this example $|\nabla Y_0^N|_{L^\infty} = N^{\frac{3}{2}}$ saturates the sup norm bound. The length of the nodal line of Y_0^N is of order λ, as one sees from the rotational invariance and by the fact that P_N has N zeros. The defect in the argument is that the bound $|\nabla Y_0^N|_{L^\infty} = N^{\frac{3}{2}}$ is only obtained on the nodal components near the poles, where each component has length $\simeq \frac{1}{N}$.

The left image is a zonal spherical harmonic of degree N on S^2: it has high peaks of height $\sqrt{N}$ at the north and south poles. The right image is a Gaussian beam: its height along the equator is $N^{1/4}$ and then it has Gaussian decay transverse to the equator.



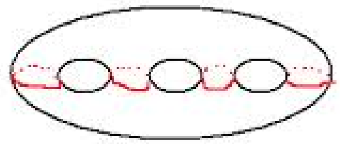
Gaussian beams are Gaussian shaped lumps which are concentrated on $\lambda^{-\frac{1}{2}}$ tubes $\mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma)$ around closed geodesics and have height $\lambda^{\frac{n-1}{4}}$. We note that their L^1 norms decrease like $\lambda^{-\frac{(n-1)}{4}}$, i.e., they saturate the Sogge L^p bounds for small p. In such cases we have $\int_{Z_{\varphi_\lambda}} |\nabla\varphi_\lambda|\, dS \simeq \lambda^2 \|\varphi_\lambda\|_{L^1} \simeq \lambda^{2-\frac{n-1}{4}}$. It is likely that Gaussian beams are minimizers of the L^1 norm among L^2-normalized eigenfunctions of Riemannian manifolds. Also, the gradient bound $\|\nabla\varphi_\lambda\|_{L^\infty} = O(\lambda^{\frac{n+1}{2}})$ is far off for Gaussian beams, the correct upper bound being $\lambda^{1+\frac{n-1}{4}}$. If we use these estimates on $\|\varphi_\lambda\|_{L^1}$ and $\|\nabla\varphi_\lambda\|_{L^\infty}$, our method gives $\mathcal{H}^{n-1}(Z_{\varphi_\lambda}) \geq C\lambda^{1-\frac{n-1}{2}}$, while λ is the correct lower bound for Gaussian beams in the case of surfaces of revolution (or any real analytic case). The defect is again that the gradient estimate is achieved only very close to the closed geodesic of the Gaussian beam. Outside of the tube $\mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma)$ of radius $\lambda^{-\frac{1}{2}}$ around the geodesic, the Gaussian beam and all of its derivatives decay like $e^{-\lambda d^2}$ where d is the distance to the geodesic. Hence $\int_{Z_{\varphi_\lambda}} |\nabla\varphi_\lambda|\, dS \simeq \int_{Z_{\varphi_\lambda} \cap \mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma)} |\nabla\varphi_\lambda|\, dS$. Applying the gradient bound for Gaussian beams to the latter integral gives $\mathcal{H}^{n-1}(Z_{\varphi_\lambda} \cap \mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma)) \geq C\lambda^{1-\frac{n-1}{2}}$, which is sharp since the intersection $Z_{\varphi_\lambda} \cap \mathcal{T}_{\lambda^{-\frac{1}{2}}}(\gamma)$ cuts across γ in $\simeq \lambda$ equally spaced points (as one sees from the Gaussian beam approximation).

13.8. Counting nodal domains

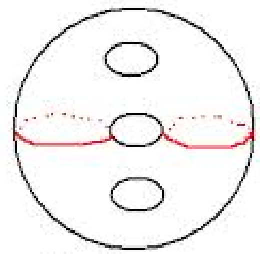
In this section, we review two recent results which count nodal domains on surfaces. One concerns Riemann surfaces corresponding to real algebraic curves which divide their complexifications; the second concerns non-positively curved surfaces with concave boundary. In both cases, the quantum ergodic restriction theorems are used to prove that there exist 'many' zeros on the relevant curve, and then a topological argument is used to prove that the nodal lines through these zeros must bound a growing number of nodal domains. The latter argument owes a good deal to prior work of Ghosh-Reznikov-Sarnk [**GRS**].

We recall that $N(\varphi_\lambda)$ denote the number of nodal domains of φ_λ. In dimension 1, the number of nodal points of the nth eigenfunction of a Sturm-Liouville operator on an interval equals $n-1$, and this suggests that the number of nodal domains should tend to infinity with the eigenvalue in any dimension. However, this is not the case and indeed was disproved by a student of Courant for squares or flat tori. Later, H. Lewy constructed sequences of spherical harmonics on the standard S^2 with degrees tending to infinity for which the number of nodal domains is ≤ 3 [**Lew**]. But it seems plausible that for any (M, g), there exists *some* orthonormal sequence $\{\varphi_{j_k}\}$ of eigenfunctions for which $N(\varphi_{j_k}) \to \infty$ as $k \to \infty$. In this section, we prove that for certain Riemann surfaces (M, J, σ) with anti-holomorphic involution, and for any negatively curved σ-invariant metric, $N(\varphi_{j_k}) \to \infty$ along an orthonormal sequence of eigenfunctions of density one. We also prove a similar result for surfaces with boundary.

13.8.1. Real algebraic curves which divide their complexifications. The relevant Riemann surfaces (M, J) are complexifications of real algebraic curves $M(\mathbb{R})$ which *divide* (equivalently, *separate*) M in the sense that $M \backslash M(\mathbb{R})$ has more than one component. Such a surface possesses an anti-holomorphic involution σ whose fixed point set $\mathrm{Fix}(\sigma)$ is the real curve $M(\mathbb{R})$. It is a classical result of F. Klein [**Kl**] and G. Weichold [**W**] that Riemann surfaces (M, J, σ) with anti-holomorphic involution and with dividing fixed point set $\mathrm{Fix}(\sigma) \neq \emptyset$ exist in any genus, and that the number of connected components equals 2. If $M(\mathbb{R}) = \mathrm{Fix}(\sigma)$ is dividing, then M/σ is orientable, while in the non-dividing case it is non-orientable. Theorem 3.3 of [**Na1**] (see also [**Na1**, Theorem 6.1]) and Corollary 2.1 of [**Na2**] express the moduli space of Klein surfaces of type (g, n, a) in terms of a Teichmüller space modulo a discrete group action, and state that the moduli space of dividing real algebraic curves to be diffeomorphic to $\mathbb{R}^{3g-3}$. Hence the space of real algebraic curves which divide their complexifications has real dimension $3g-3$. Further background and references are given in §13.8.6. Some images taken from the web are given below.



M-curve



dividing curve

We define $\mathcal{M}_{M,J,\sigma}$ to be the space of C^∞ σ-invariant negatively curved Riemannian metrics on an orientable Klein surface (M, J, σ). Any negatively curved metric g_1 induces a σ-invariant one by averaging, $g_1 \to g = \frac{1}{2}(g_1 + \sigma^* g_1)$. Hence $\mathcal{M}_{M,J,\sigma}$ is an open set in the space of σ-invariant metrics. The isometry σ commutes with the Laplacian Δ_g and therefore the eigenspaces are spanned by even or odd eigenfunctions with respect to σ. We denote by $\{\varphi_j\}$ of $L^2_{even}(M)$ an orthonormal basis of even eigenfunctions, resp. $\{\psi_j\}$ an orthonormal basis of $L^2_{odd}(M)$.

Theorem 13.27. *Let (M, J, σ) be a compact Riemann surface with anti-holomorphic involution σ such that* Fix(σ) *divides M into two connected components. Let $\mathcal{M}_{M,J,\sigma}$ be the space of σ-invariant negatively curved C^∞ Riemannian metrics on M.Then for any $g \in \mathcal{M}_{(M,J,\sigma)}$ and any orthonormal Δ_g-eigenbasis $\{\varphi_j\}$ of $L^2_{even}(M)$, resp. $\{\psi_j\}$ of $L^2_{odd}(M)$, one can find a density* 1 *subset A of* $\mathbb{N}$ *such that*

$$\lim_{\substack{j\to\infty \\ j\in A}} N(\varphi_j) = \infty \quad \textit{resp.} \quad \lim_{\substack{j\to\infty \\ j\in A}} N(\psi_j) = \infty. \tag{13.38}$$

Theorem 13.27 applies to all metrics in $\mathcal{M}(M, J, \sigma)$ for any real algebraic curve with $M_{\mathbb{R}}$ dividing $M_{\mathbb{C}}$.

Remark 13.28. For generic metrics in $\mathcal{M}_{M,J,\sigma}$, the eigenvalues are simple (multiplicity one) and therefore all eigenfunctions are either even or odd. Hence for generic metrics in $\mathcal{M}_{M,J,\sigma}$, Theorem 13.27 says that the number of nodal domains tends to infinity along almost the entire sequence of eigenfunctions. See Proposition 13.43 for the proof.

For odd eigenfunctions, the same conclusion holds with the assumption Fix(σ) separating replaced by Fix(σ) $\neq \emptyset$, i.e., for the complexification of any real algebraic curve.

There are two key ingredients in the proof of Theorem 13.27. The first is a proof that the the number of intersections points of the nodal line with Fix(σ) of the even eigenfunctions φ_j, resp. the number of singular points of odd eigenfunctions ψ_j, tends to infinity along a density one subsequence of eigenfunctions. This is a statement of independent interest and we discuss the relevant result in more detail in §13.8.2. We further discuss a significant generalization in §13.8.4.

The second ingredient (Lemma 13.42) is a topological argument. Using the Euler inequality for embedded graphs, we show that the growing number of nodal intersections with Fix(σ) in Theorem 13.29 implies a growing number of nodal domains. This topological argument uses simplicity that Fix(σ) is the common boundary of the two components of $M \backslash$ Fix(σ).

13.8.2. Number of intersection points of nodal lines and curves. The analytical part of Theorem 13.27 is the following:

THEOREM 13.29. *Let (M, J, σ) be (as above) a compact Riemann surface with anti-holomorphic involution for which* $\text{Fix}(\sigma)$ *is dividing. Let g be a negatively curved metric, $g \in \mathcal{M}_{M,J,\sigma}$, invariant under σ as in Theorem 13.27. Let $\gamma \subset \text{Fix}(\sigma)$ be any sub-arc. Then for any orthonormal eigenbasis $\{\varphi_j\}$ of $L^2_{even}(M)$, resp. $\{\psi_j\}$ of $L^2_{odd}(M)$, one can find a density* 1 *subset A of $\mathbb{N}$ such that*

$$\lim_{\substack{j\to\infty \\ j\in A}} \#Z_{\varphi_j} \cap \gamma = \infty.$$

Furthermore, there are an infinite number of zeros where $\varphi_j|_H$ changes sign.

In fact, we prove that the number of zeros tends to infinity by proving that the number of sign changes tends to infinity. The proof uses the Kuzencov trace formula of [**Z2**] to show that $\int_\gamma \varphi_j \, ds$ is 'small' as $j \to \infty$ for any curve γ and for almost all eigenfunctions. On the other hand the QER theorem shows that $\int_\gamma \varphi_j^2 \, ds$ is large. We then compare $\int_\gamma \varphi_j \, ds$ and $\int_\gamma |\varphi_j| \, ds$ by applying a well known sup norm bound on eigenfunctions in the case of surfaces without conjugate points to replace $\int_\gamma \varphi_j^2 \, ds$ by $\int_\gamma |\varphi_j| \, ds$. The comparison just manages to show that for any geodesic arc γ, $\int_\gamma |\varphi_j| \, ds > \int_\gamma \varphi_j \, ds$. Hence there must exist sign-changing zeros.

REMARK 13.30. Here and henceforth, γ always denotes a sub-arc of $\text{Fix}(\sigma)$. For each $g \in \mathcal{M}_{M,J,\sigma}$, it follows from Harnack's theorem that the fixed point set $\text{Fix}(\sigma)$ is a disjoint union

$$\text{Fix}(\sigma) = \gamma_1 \cup \cdots \cup \gamma_k \tag{13.39}$$

of $0 \le k \le g + 1$ simple closed geodesics, and by our assumption $k > 0$ and $\text{Fix}(\sigma)$ is dividing. Hence the arcs γ above are geodesic arcs of (M, g).

REMARK 13.31. $Z_{\varphi_j} \cap \gamma$ must be a finite set of points. For, if $Z_{\varphi_j} \cap \gamma$ contains a curve, then tangential derivative of φ_j along the curve vanishes. Hence together with $\partial_\nu \varphi_j = 0$, we have $d\varphi_j(x) = 0$ along the curve, contradicting the upper bound in [**D**] on the number of singular points.

The main ingredient of Theorem 13.29 is the QER (quantum ergodic restriction) theorem for Cauchy data of [**CTZ**]. This is the 'easy' QER theorem which holds without any conditions on the hypersurface (i.e., curve) in M. We recall its statement in §13.8.9 (see in particular Theorem 13.35). Roughly speaking, the negatively curvature of g is used to guarantee that the geodesic flow is ergodic. The QER theorem then says that the Cauchy data is quantum ergodic along $\text{Fix}(\sigma)$. Indeed, it is quantum ergodic along any curve. $\text{Fix}(\sigma)$ is special because the odd eigenfunctions automatically vanish on it and the even eigenfunctions have vanishing normal derivatives. Hence half of the Cauchy data of each eigenfunction automatically vanishes on $\text{Fix}(\sigma)$. Quantum ergodicity forces the sequence of restrictions of eigenfunctions to $\text{Fix}(\sigma)$ to oscillate quickly and thus to have a growing number of zeros as the eigenvalue increases. This lower bound on the number of zeros in the presence of ergodicity is a kind of converse to the upper bound of [**TZ2**]. There it is shown that the number of nodal intersections in the real analytic case is bounded above by $\sqrt{\lambda}$.

13.8.3. Number of singular points. Our methods also show that the number of singular points of odd eigenfunctions ψ_j tends to infinity. By singular points of an eigenfunction we mean the set

$$\Sigma_{\varphi_\lambda} = \{x \in Z_{\varphi_\lambda} : d\varphi_\lambda(x) = 0\}$$

of critical points φ_λ which lie on the nodal set Z_{φ_j}. It is proved in [**D**] that the number of singular points of φ_λ is bounded by $C\sqrt{\lambda}$ on any surface. For generic metrics, the singular set is empty [**U**]. However for negatively curved surfaces with an isometric involution, odd eigenfunctions ψ always have singular points. Indeed, odd eigenfunctions vanish on γ and they have singular points at $x \in \gamma$ where the normal derivative vanishes, $\partial_\nu \psi_j = 0$.

THEOREM 13.32. *Let (M, J, σ) be (as above) a compact Riemann surface with anti-holomorphic involution for which* $\text{Fix}(\sigma)$ *is dividing. Let g be a negatively curved metric, $g \in \mathcal{M}_{M,J,\sigma}$, invariant under σ as in Theorem 13.27. Then for any orthonormal eigenbasis $\{\psi_j\}$ of $L^2_{odd}(M)$, one can find a density* 1 *subset A of* $\mathbb{N}$ *such that*

$$\lim_{\substack{j\to\infty \\ j\in A}} \# \, \Sigma_{\psi_j} \cap \text{Fix}(\sigma) = \infty.$$

Furthermore, there are an infinite number of zeros where $\partial_\nu \psi_j|_H$ changes sign.

13.8.4. Quantum ergodic restriction and intersections of nodal lines and generic curves. In this section, we state a quite general result on intersections of nodal lines and geodesics, somewhat analogous to Theorem 13.29 but involving different assumptions. At present, we do not know if the generalization does leads to lower bounds on numbers of nodal domains. It is based on the more difficult QER theorem of [**TZ2**] for Dirichlet data, which says that if (M, g) is a surface with ergodic geodesic flow and $H \subset M$ is a curve satisfying a generic asymmetry condition, then the restriction of a density one subsequence of eigenfunctions to H is quantum ergodic. The asymmetry condition is essentially that the two geodesics with mirror image initial velocities emanating from a point of H almost never return to H at the same time to the same place.

THEOREM 13.33. *Let (M, g) be a C^∞ compact negatively curved surface, and let H be a closed curve which is asymmetric with respect to the geodesic flow. Then for any orthonormal eigenbasis $\{\varphi_j\}$ of Δ-eigenfunctions of (M, g), there exists a density* 1 *subset A of* $\mathbb{N}$ *such that*

$$\begin{cases} \lim\limits_{\substack{j\to\infty \\ j\in A}} \# Z_{\varphi_j} \cap H = \infty, \\ \lim\limits_{\substack{j\to\infty \\ j\in A}} \# \{x \in H : \partial_\nu \varphi_j(x) = 0\} = \infty. \end{cases}$$

Furthermore, there are an infinite number of zeros where $\varphi_j|_H$ (resp. $\partial_\nu \varphi_j|_H$) changes sign.

Theorem 13.33 does not necessarily imply lower bounds on nodal domains because the topological argument used in the case $H = \text{Fix}(\sigma)$ does not necessarily apply. We include the result here because the proof of Theorem 13.33 is essentially the same as for Theorem 13.29. The main difference is that we use the QER theorem for Cauchy data in Theorem 13.29 and for just the Dirichlet data in Theorem 13.33. The latter requires the asymmetry condition on H, which was shown to be generic in [**TZ2**] and related articles (we refer there for references).

13.8.5. Curvature assumption. We assume the surfaces are negatively curved for two reasons. First we need that the geodesic flow is ergodic. Ergodicity is assumed so that the Quantum Ergodic Restriction (QER) results of [**CTZ**] apply. In fact, this theorem generalizes to all dimensions and all hypersurfaces but since our main results pertain to surfaces we only state the results in this case. Non-positivity of the curvature is also used to ensure that (M, g) has no conjugate points and that the estimates on sup-norms of eigenfunctions in [**Ber**] apply.

13.8.6. Background on real algebraic curves. The assumption that $\text{Fix}(\sigma) \neq \emptyset$ is equivalent to the fixed point set being a real algebraic curve. There is a moduli space of real dimension $3g - 3$ of real algebraic curves. One defines $n(M)$ to be the number of connected components of $M(\mathbb{R})$, the real locus. If M has genus g then $M(\mathbb{R})$ consists of $n(M)$ disjoint circles. The complement $M(\mathbb{C}) - M(\mathbb{R})$ of the real locus in the complex locus has either one or two connected components. Put $a(M) = 0$ if $M(\mathbb{R})$ divides the complex locus and $a(M) = 1$ if $M(\mathbb{C}) - M(\mathbb{R})$ is connected. The triple (g, n, a) is a complete set of topological invariants of a real algebraic curve and is called the topological type of (M, J, σ). Weichold [**W**] proved that this data determines (M, J, σ) up to an equivariant homeomorphism (see [**Na1, Jaf**] for modern presentations). This classification result is based on the identification of $M(\mathbb{C})$ as the *complex double* of the quotient $M(\mathbb{C})/\sigma$ (see [**Na1**, §1] for background on doubles).

When $\text{Fix}(\sigma) = M(\mathbb{R})$ is dividing, then $M \setminus \text{Fix}(\sigma) = M_+ \cup M_-$ where $M_\pm$ are connected, where $M_+^0 \cap M_-^0 = \emptyset$ (the interiors are disjoint), where $\sigma(M_+) = M_-$ and where $\partial M_+ = \partial M_- = \text{Fix}(\sigma)$.

The quotient of $M(\mathbb{C})$ by complex conjugation is a connected 2-manifold X with $n(M)$ boundary components. X has Euler characteristic $1 - g$ and is orientable if and only if $a(M) = 0$. One has the following constraints:

(1) $0 \leq n(X) \leq g + 1$;
(2) If $n(X) = 0$ then $a(X) = 1$. If $n(X) = g + 1$ then $a(X) = 0$;
(3) If $a(X) = 0$ then $n(X) = g + 1 \mod 2$.

Klein [**Kl**] proved that any pair $(n(M), a(M))$ which satisfies these constraints is realized by some real curve of genus g. We refer to [**Jaf**] for a modern proof.

As mentioned in the introduction, the moduli space of real algebraic curves of a given topological type (g, n, a) is diffeomorphic to the quotient of $\mathbb{R}^{3g-3}$ by a discrete group action. We refer to Corollary 2.1 of [**Na2**] for the precise statement and for references.

13.8.7. Kuznecov sum formula on surfaces. The second ingredient is the Kuznecov asymptotics, which have the following implication:

Proposition 13.34. *There exists a subsequence of eigenfunctions φ_j of natural density one so that, for all $f \in C^\infty(\gamma)$,*

$$\begin{cases} \left|\int_\gamma f\varphi_j \, ds\right| = O_f(\lambda_j^{-1/2}(\log \lambda_j)^{1/2}), \\ \lambda_j^{-\frac{1}{2}} \left|\int_\gamma f\partial_\nu \varphi_j \, ds\right| = O_f(\lambda_j^{-1/2}(\log \lambda_j)^{1/2}). \end{cases} \tag{13.40}$$

The proof makes use of the general Kuznecov asymptotics on a compact Riemannian manifold. Let (M, g) be a Riemannian surface, let $\{u_j\}$ be an orthonormal basis of Δ_g eigenfunctions, and let $C \subset M$ be a closed curve of a surface M. Let $f \in C^\infty(C)$. Then,

$$\sum_{\lambda_j < \lambda} \left|\int_C f u_j ds\right|^2 = \frac{1}{\pi}\left|\int_C f ds\right|^2 \lambda + O_f(1). \tag{13.41}$$

Proof. Denote by $N(\lambda)$ the number of eigenfunctions in $\{j \mid \lambda < \lambda_j < 2\lambda\}$.

For each f, we have by Kuznecov asymptotics and by Chebyshev's inequality,

$$\frac{1}{N(\lambda)}\left\{j \mid \lambda < \lambda_j < 2\lambda, \left|\int_{\gamma_i} f\varphi_j ds\right| \geq \lambda_j^{-1/2}(\log \lambda_j)^{\frac{1}{2}}\right\} = O_f\left(\frac{1}{\log \lambda}\right).$$

It follows that the upper density of exceptions to (13.40) tends to zero. We then choose a countable dense set $\{f_n\}$ and apply the diagonalization argument (see [**Zw**, Theorem 15.5 step (2)]) to conclude that there exists a density one subsequence for which (13.40) holds for all $f \in C^\infty(\gamma)$. The same holds for the normal derivative. □

13.8.8. Proof of Theorem 13.29. The proof is based on the QER theorem for Cauchy data of [**CTZ**]. Our application is to the curve $H = \text{Fix}(\sigma)$, i.e. curve in the Riemann surface M, given by the fixed point set (13.39) of the isometric involution σ.

It is interesting to note that such a hypersurface (i.e., curve) is precisely the kind ruled out in the hypothesis of the Dirichlet data (or Neumann data) QER theorem [**TZ2**]. However the quantum ergodic restriction theorem for Cauchy data in [**CTZ**] does apply and shows that the even eigenfunctions are quantum ergodic along H, hence along each component γ_j and any subarc γ of some γ_j.

13.8.9. Quantum ergodic restriction theorems for Cauchy data. The statement we use is the following:

Theorem 13.35. *Assume that (M, g) has an orientation reversing isometric involution with separating fixed point set H (13.39). Let γ be a component of H, and let $Op_\gamma(a)$ be a semi-classical pseudo-differential operator on $L^2(\gamma, ds)$. Let φ_h be the sequence of even ergodic eigenfunctions. Then,*

$$\left\langle \text{Op}_\gamma(a)\varphi_h|_\gamma, \varphi_h|_\gamma \right\rangle_{L^2(\gamma)} \to_{h\to 0^+} \frac{4}{2\pi \text{Area}(M)} \int_{B^*\gamma} a_0(s, \sigma)(1 - |\sigma|^2)^{-1/2} \, ds d\sigma. \tag{13.42}$$

In particular, this holds when $\text{Op}_\gamma(a)$ is multiplication by a smooth function f.

We only use the special case of the QER theorem in which we test $\varphi_j|_H$ against functions supported on H. Since the symbol a may be assumed to be supported on one component γ_j, there is no essential difference in stating the result for operators

$\mathrm{Op}_H(a)$ along H and for ones supported on a component of H. We state it in the latter form because we plan to use the result on small sub-arcs. We have dropped the subscript γ_j for notational simplicity and also because the result is valid for a sub-arc of a γ_j. It also follows that normal derivatives of odd eigenfunctions are quantum ergodic along γ, but we do not use this result here.

We briefly indicate how to derive Theorem 13.35 from the Cauchy-data QER theorem:

THEOREM 13.36. *Assume that $\{\varphi_h\}$ is a quantum ergodic sequence of eigenfunctions on M. Then,*

$$\left\langle \mathrm{Op}_\gamma(a) hD_\nu \varphi_h|_\gamma, hD_\nu \varphi_h|_\gamma \right\rangle_{L^2(\gamma)} + \left\langle \mathrm{Op}_\gamma(a)(1 + h^2\Delta_\gamma)\varphi_h|_\gamma, \varphi_h|_\gamma \right\rangle_{L^2(\gamma)} \to_{h\to 0^+} \frac{4}{\mu(S^*M)} \int_{B^*\gamma} a_0(s,\sigma)(1-|\sigma|^2)^{1/2}\, ds d\sigma, \tag{13.43}$$

where a_0 is the principal symbol of $\mathrm{Op}_\gamma(a)$.

When applied to even eigenfunctions under an orientation-reversing isometric involution with separating fixed point set, the Neumann data drops out and we get

COROLLARY 13.37. *Let (M, g) have an orientation-reversing isometric involution with separating fixed point set H and let γ be one of its components. Then for any sequence of even quantum ergodic eigenfunctions of (M, g),*

$$\left\langle \mathrm{Op}_\gamma(a)(1 + h^2\Delta_\gamma)\varphi_h|_\gamma, \varphi_h|_\gamma \right\rangle_{L^2(\gamma)} \to_{h\to 0^+} \frac{4}{\mu(S^*M)} \int_{B^*\gamma} a_0(s,\sigma)(1-|\sigma|^2)^{1/2}\, ds d\sigma.$$

This is not the result we wish to apply since we would like to have a limit formula for the integrals $\int_\gamma f\varphi_h^2 ds$. Thus we need a more complicated application involving the the microlocal lift $d\Phi_h^D \in \mathcal{D}'(B^*\gamma)$ of the Dirichlet data of φ_h,

$$\int_{B^*\gamma} a\, d\Phi_h^D := \langle \mathrm{Op}_\gamma(a)\varphi_h|_\gamma, \varphi_h|_\gamma \rangle_{L^2(\gamma)}.$$

In order to obtain a quantum ergodicity result for the Dirichlet data we need to introduce the renormalized microlocal lift of the Neumann data,

$$\int_{B^*\gamma} a\, d\Phi_h^{RN} := \langle (1 + h^2\Delta_\gamma + i0)^{-1}\, \mathrm{Op}_\gamma(a) hD_\nu\varphi_h|_\gamma, hD_\nu\varphi_h|_\gamma \rangle_{L^2(\gamma)}.$$

THEOREM 13.38. *Assume that $\{\varphi_h\}$ is a quantum ergodic sequence on M. Then, there exists a sub-sequence of density one as $h \to 0^+$ such that for all $a \in S^0_{sc}(\gamma)$,*

$$\left\langle (1 + h^2\Delta_\gamma + i0)^{-1}\, \mathrm{Op}_\gamma(a) hD_\nu\varphi_h|_H, hD_\nu\varphi_h|_\gamma \right\rangle_{L^2(\gamma)} + \left\langle \mathrm{Op}_\gamma(a)\varphi_h|_\gamma, \varphi_h|_\gamma \right\rangle_{L^2(\gamma)} \to_{h\to 0^+} \frac{4}{2\pi \mathrm{Area}(M)} \int_{B^*\gamma} a_0(s,\sigma)(1-|\sigma|^2)^{-1/2}\, ds d\sigma. \tag{13.44}$$

Theorem 13.35 follows from Theorem 13.38 since the Neumann term drops out (as before) under the hypothesis of Corollary 13.37.

13.8.10. Proof of Theorem 13.29.

Proof. We first consider the even eigenfunctions. Then the first term of Theorem 13.35 vanishes.

Fix $R \in \mathbb{N}$. Let $\gamma_1, \cdots, \gamma_R$ be a partition of the closed curve H and let $\beta_i \subset \gamma_i$ be proper subsegments. Let $f_1, \cdots, f_R \in C_0^\infty(H)$ be given such that

$$\operatorname{supp}\{f_i\} = \gamma_i, \quad f_i \geq 0 \text{ on } H, \quad f_i = 1 \text{ on } \beta_i.$$

We may assume that the sequence $\{\varphi_j\}$ has the quantum restriction property of Theorem 13.35 which implies that

$$\lim_{j\to\infty} \|\varphi_j\|_{L^2(\beta_i)} = B \cdot \text{length}(\beta_j)$$

for all $j = 1, \cdots, R$ for some constant $B > 0$. Namely, $B = \int_{-1}^1 (1-\sigma^2)^{\frac{1}{2}}\, d\sigma$. Then

$$\int_{\beta_i} |\varphi_j|\, ds \geq \|\varphi_j\|^2_{L^2(\beta_i)} \|\varphi_j\|^{-1}_{L^\infty(M)} \gg \lambda_j^{-1/2} \log \lambda_j.$$

Here we use the well-known inequality $\|\varphi_j\|_{L^\infty(M)} \ll \lambda_j^{1/4}/\log\lambda_j$ which follows from the remainder estimate in the pointwise Weyl law of [**Ber**].

By Proposition 13.34,

$$\left|\int_{\gamma_i} f_i \varphi_j\, ds\right| = O_R(\lambda_j^{-1/2}(\log\lambda_j)^{1/2})$$

is satisfied for any $i = 1, \cdots, R$ for almost all φ_j. Therefore for all sufficiently large j, such φ_j has at least one sign change on each segment γ_i proving that $\#Z_{\varphi_j} \cap H \geq R$ is satisfied for every $R > 0$ by almost all φ_j. Now we apply Lemma 13.39 with $a_j = \#Z_{\varphi_j} \cap H$ to conclude Theorem 13.29.

The proof for Neumann data is essentially the same, since for odd eigenfunctions, the second term of Theorem 13.35 vanishes. □

13.8.10.1. *Appendix on density.* Define the natural density of a set $A \in \mathbb{N}$ by

$$\lim_{X\to\infty} \frac{1}{X}\Big|\{x \in A \mid x < X\}\Big|$$

whenever the limit exists. We say "almost all" when corresponding set $A \in \mathbb{N}$ has the natural density 1. Note that intersection of finitely many density 1 set is a density 1 set. When the limit does not exist we refer to the lim sup as the upper density and the lim inf as the lower density.

Lemma 13.39. *Let a_n be a sequence of real numbers such that for any fixed $R > 0$, $a_n > R$ is satisfied for almost all n. Then there exists a density 1 subsequence $\{a_n\}_{n\in A}$ such that*

$$\lim_{\substack{n\to\infty\\ n\in A}} a_n = +\infty.$$

Proof. Let n_k be the least number such that for any $n \geq n_k$,

$$\frac{1}{n}|\{j \leq n : a_j > k\}| > 1 - \frac{1}{2^k}.$$

Note that n_k is nondecreasing, and $\lim_{k\to\infty} n_k = +\infty$.

Define $A_k \subset \mathbb{N}$ by

$$A_k = \{n_k \leq j < n_{k+1} : a_j > k\}.$$

Then for any $n_k \leq m < n_{k+1}$,

$$\{j \leq m : a_j > k\} \subset \bigcup_{l=1}^{k} A_l \cap [1, m],$$

which implies by the choice of n_k that

$$\frac{1}{m}\left|\bigcup_{l=1}^{k} A_l \cap [1, m]\right| > 1 - \frac{1}{2^k}.$$

This proves $A = \cup_{k=1}^{\infty} A_k$ is a density 1 subset of $\mathbb{N}$, and by the construction we have the statement of the Lemma. $\square$

13.8.11. Sign changing zeros and singular points. As mentioned above, $\mathrm{Fix}(\sigma)$ consists of a union of closed geodesics. Let $\gamma \subset \mathrm{Fix}(\sigma)$ be any component geodesic.

We recall that a singular point $x_0 \in M$ for and eigenfunction φ_λ is a point where $\varphi_\lambda(x_0) = d\varphi_j(x_0) = 0$. A non-singular zero is called a regular zero.

LEMMA 13.40. *Let φ_λ be an even eigenfunction, and let $x_0 = \gamma(s_0)$ be a zero of $\varphi_\lambda|_\gamma$. Then if x_0 is a regular zero, then $\varphi_\lambda|_\gamma$ changes sign. That is, if the even eigenfunction does not change sign at the zero x_0 along γ, x_0 must be a singular point.*

Indeed, since φ is even, its normal derivative vanishes everywhere on γ. If φ does not change sign at x_0, then γ is tangent to Z_{φ_j} at x_0, i.e. $\frac{d}{ds}\varphi_j(\gamma(s)) = 0$, so that x_0 is a singular point.

Next we consider odd eigenfunctions and let ψ_λ be an odd eigenfunction. As above, let γ be a component of $\mathrm{Fix}(\sigma)$. Then $\psi_\lambda \equiv 0$ on γ and the zeros of $\partial_\nu \psi_\lambda$ on γ are singular points of ψ_λ.

LEMMA 13.41. *Let ψ_λ be an odd eigenfunction. Then the zeros of the normal derivative $\partial_\nu \psi_\lambda$ on γ are intersection points of the nodal set of ψ_λ in $M \backslash \gamma$ with γ, i.e. point where at least two nodal branches cross.*

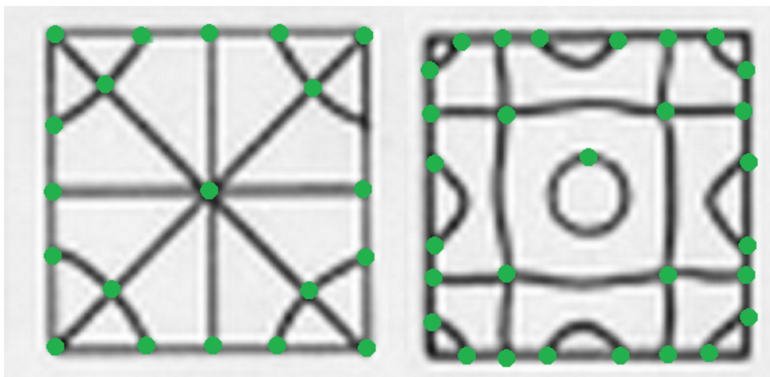
PROOF. If x_0 is a singular point, then $\varphi_j(x_0) = d\varphi_j(x_j) = 0$, so the zero set of φ_λ is similar to that of a spherical harmonic of degree $k \geq 2$, which consists of $k \geq 2$ arcs meeting at equal angles at 0. It follows that at least two transverse branches of the nodal set of an odd eigenfunction meet at each singular point on γ. $\square$

13.8.12. Graph structure of the nodal set and completion of proof of Theorem 13.27. Using the local structure of the nodal set, on can give a graph structure (i.e. the structure of a one-dimensional CW complex) to Z_{φ_λ} as follows.

(1) For each embedded circle which does not intersect γ, we add a vertex.
(2) Each singular point is a vertex.
(3) If $\gamma \not\subset Z_{\varphi_\lambda}$, then each intersection point in $\gamma \cap Z_{\varphi_\lambda}$ is a vertex.
(4) Edges are the arcs of Z_{φ_λ} ($Z_{\varphi_\lambda} \cup \gamma$, when φ_λ is even) which join the vertices listed above.

This way, we obtain a graph embedded into the surface M. We recall that an embedded graph G in a surface M is a finite set $V(G)$ of vertices and a finite set $E(G)$ of edges which are simple (non-self-intersecting) curves in M such that any two distinct edges have at most one endpoint and no interior points in common.

The *faces* f of G are the connected components of $M \backslash V(G) \cup \bigcup_{e \in E(G)} e$. The set of faces is denoted $F(G)$. An edge $e \in E(G)$ is *incident* to f if the boundary of f contains an interior point of e. Every edge is incident to at least one and to at most two faces; if e is incident to f then $e \subset \partial f$. The faces are not assumed to be cells and the sets $V(G), E(G), F(G)$ are not assumed to form a CW complex. Indeed the faces of the nodal graph of odd eigenfunctions are nodal domains, which do not have to be simply connected. In the even case, the faces which do not intersect γ are nodal domains and the ones which do are inert nodal domains which are cut in two by γ.



Now let $v(\varphi_\lambda)$ be the number of vertices, $e(\varphi_\lambda)$ be the number of edges, $f(\varphi_\lambda)$ be the number of faces, and $m(\varphi_\lambda)$ be the number of connected components of the graph. Then by Euler's formula [**Gr**, Appendix F],

$$v(\varphi_\lambda) - e(\varphi_\lambda) + f(\varphi_\lambda) - m(\varphi_\lambda) \geq 1 - 2g_M, \tag{13.45}$$

where g_M is the genus of the surface.

We use this inequality to give a lower bound for the number of nodal domains for even and odd eigenfunctions.

Lemma 13.42. *For an odd eigenfunction* ψ_j,

$$N(\psi_j) \geq \# \left(\Sigma_{\psi_j} \cap \gamma\right) + 2 - 2g_M,$$

and for an even eigenfunction φ_j,

$$N(\varphi_j) \geq \frac{1}{2} \# \left(Z_{\varphi_j} \cap \gamma\right) + 1 - g_M.$$

Proof. **Odd case.** For an odd eigenfunction ψ_j, $\gamma \subset Z_{\psi_j}$. Therefore $f(\psi_j) = N(\psi_j)$. Let $n(\psi_j) = \# \Sigma_{\psi_j} \cap \gamma$ be the number of singular points on γ. These points correspond to vertices having degree at least 4 on the graph, hence

$$0 = \sum_{x \text{ vertices}} \deg(x) - 2e(\psi_j) \geq 2\left(v(\psi_j) - n(\psi_j)\right) + 4n(\psi_j) - 2e(\psi_j).$$

Therefore $e(\psi_j) - v(\psi_j) \geq n(\psi_j)$, and plugging into (13.45) with $m(\psi_j) \geq 1$, we obtain $N(\psi_j) \geq n(\psi_j) + 2 - 2g_M$.

Even case. For an even eigenfunction φ_j, let $N_{in}(\varphi_j)$ be the number of nodal domain U which satisfies $\sigma U = U$ (inert nodal domains). Let $N_{sp}(\varphi_j)$ be the number of the rest (split nodal domains). From the assumption that $Fix(\sigma)$ is separating, inert nodal domains intersect $\text{Fix}(\sigma)$ on simple segments, and $\text{Fix}(\sigma)$ divides each nodal domain into two connected components. This implies that, because $\gamma \subset \mathrm{Fix}(\sigma)$ is added when giving the graph structure, the inert nodal domain may correspond to two faces on the graph, depending on whether the nodal domain intersects γ or not. Therefore $f(\varphi_j) \leq 2N_{in}(\varphi_j) + N_{sp}(\varphi_j)$.

Observe that each point in $Z_{\varphi_j} \cap \gamma$ corresponds to a vertex having degree at least 4 on the graph. Hence by the same reasoning as the odd case, we have

$$N(\varphi_j) \geq N_{in} + \frac{1}{2}N_{sp}(\varphi_j) \geq \frac{f(\varphi_j)}{2} \geq \frac{n(\varphi_j)}{2} + 1 - g_M,$$

where $n(\varphi_j) = \# Z_{\varphi_j} \cap \gamma$. □

Now Theorem 13.27 follows from Theorem 13.29 and Lemma 13.42.

13.8.13. Generic simplicity of eigenvalues. In this section we prove the genericity result stated in Remark 13.28:

PROPOSITION 13.43. *Let (M, J, σ) be a Riemann surface with an anti-holomorphic involution with separating fixed point set* $\mathrm{Fix}(\sigma)$. *Then for generic negatively σ-invariant curved metrics, the Laplace eigenfunctions are either even or odd.*

PROOF. Any eigenfunction may be decomposed as a sum of its even part and its odd part. To prove the Proposition it suffices to show that for a residual set of non-positively σ-invariant curved metrics, the multiplicity of each eigenvalue is equal to one. The eigenfunction is then unique up to scalar multiple and must be either even or odd. In a standard way, it suffices to show that for each j there exists an open dense set of such metrics for which the jth eigenvalue is simple.

Openness is simple since a sufficiently small perturbation of a metric for which the jth eigenvalue is simple also has a simple jth eigenvalue. Regarding density, assume that one cannot split the eigenvalue at some non-positively curved metric g_0. One issue is that a small perturbation of g need not be non-positively curved, so density at g_0 is problematic. Hence we consider strictly negatively curved metrics. If we cannot separate the eigenvalue then for any infinitesimal area preserving σ-invariant perturbation we have $\int_M \dot{\rho}|\varphi_j^1|^2 = \int_M \dot{\rho}|\varphi_j^2|^2$, where φ_j^1 and φ_j^2 are two distinct σ-invariant eigenfunctions corresponding to the same eigenvalue.

But this says that $|\varphi_j^1|^2 - |\varphi_j^2|^2$ is orthogonal to all σ-invariant functions $\dot{\rho}$ so that $\int_M \dot{\rho}\, dV_g = 0$. Since $|\varphi_j^1|^2, |\varphi_j^2|^2$ are also σ-invariant, we make take the quotient by the $\mathbb{Z}_2$ action defined by σ and find that $\int_{M/\mathbb{Z}_2} \dot{\rho}(|\varphi_j^1|^2 - \dot{\rho}|\varphi_j^2|^2)\, dV = 0$ for all smooth $\dot{\rho}$ on $M/\mathbb{Z}_2$ such that $\int \dot{\rho} = 0$. That is, $|\varphi_j^1|^2 - |\varphi_j^2|^2 = C$ for some constant C on $M/\mathbb{Z}_2$. Integrating over M shows that $C = 0$ and therefore $\varphi_j^1 = \varepsilon \varphi_j^2$ where $\varepsilon = \pm 1$. The sign must be constant by regularity, and we then get a contradiction. □

Bibliography

[Ar] S. Ariturk, Lower bounds for nodal sets of Dirichlet and Neumann eigenfunctions, to appear in Comm. Math. Phys. (arXiv:1110.6885).

[Au] Auroux, D. A remark about Donaldson's construction of symplectic submanifolds. J. Symplectic Geom. 1 (2002), no. 3, 647658.

[Bae] C. Baer, On nodal sets for Dirac and Laplace operators, Comm. Math. Phys. 188(3) (1997), 709-72.

[Bak] L. Bakri, Critical set of eigenfunctions of the Laplacian, arXiv:1008.1699.

[Bel] K. Bellova and F.H. Lin, Nodal sets of Steklov eigenfunctions. Calc. Var. Partial Differential Equations 54 (2015), no. 2, 2239-2268.

[Ber] P. Bérard, On the wave equation on a compact Riemannian manifold without conjugate points. Math. Z. 155 (1977), no. 3, 249-276.

[Bers] L. Bers, Local behavior of solutions of general linear elliptic equations. Comm. Pure Appl. Math. 8 (1955), 473–496.

[BR1] J. Bourgain and Z. Rudnick, Nodal intersections and L^p restriction theorems on the torus (arXiv:1308.4247).

[BR2] J. Bourgain and Z. Rudnick, On the nodal sets of toral eigenfunctions. Invent. Math. 185 (2011), no. 1, 19923.

[Br] J. Brüning, Über Knoten Eigenfunktionen des Laplace-Beltrami Operators, Math. Z. 158 (1978), 15-21.

[CNV] J. Cheeger, A. Naber and D. Valtorta, Critical Sets of Elliptic Equations, Comm. Pure Appl. Math. 68 (2015), no. 2, 173-209 (arXiv:1207.4236).

[Ch1] S. Y. Cheng, Eigenfunctions and eigenvalues of Laplacian. Differential geometry (Proc. Sympos. Pure Math., Vol. XXVII, Stanford Univ., Stanford, Calif., 1973), Part 2, pp. 185–193. Amer. Math. Soc., Providence, R.I., 1975.

[Ch2] S. Y. Cheng, Eigenfunctions and nodal sets. Comment. Math. Helv. 51 (1976), no. 1, 43–55.

[CTZ] H. Christianson, J. A. Toth and S. Zelditch, Quantum ergodic restriction for Cauchy data: Interior QUE and restricted QUE, Math. Res. Lett. 20 (2013), no. 3, 465475 (arXiv:1205.0286).

[CM1] T. H. Colding and W. P. Minicozzi II, Lower bounds for nodal sets of eigenfunctions. Comm. Math. Phys. 306 (2011), no. 3, 777 - 784.

[CM2] T. Colding and W.P. Minicozzi, Volumes for eigensections. Geom. Dedicata 102 (2003), 19-24.

[CH] R. Courant and D. Hilbert, *Methods of mathematical physics.* Vol. I. Interscience Publishers, Inc., New York, N.Y., 1953.

[D] R-T. Dong, Nodal sets of eigenfunctions on Riemann surfaces. J. Differential Geom. 36 (1992), no. 2, 493–506.

[DF1] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions on Riemannian manifolds, Invent. Math. 93 (1988), 161-183.

[DF2] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions: Riemannian manifolds with boundary. Analysis, et cetera, 251–262, Academic Press, Boston, MA, 1990.

[DF3] H. Donnelly and C. Fefferman, Growth and geometry of eigenfunctions of the Laplacian. Analysis and partial differential equations, 635–655, Lecture Notes in Pure and Appl. Math., 122, Dekker, New York, 1990.

[DF4] H. Donnelly and C. Fefferman, Nodal sets for eigenfunctions of the Laplacian on surfaces. J. Amer. Math. Soc. 3 (1990), no. 2, 333–353.

[Do] S.K. Donaldson, Symplectic submanifolds and almost-complex geometry, J. Diff. Geom. 44 (1996), 666-705.

[EJN] A. Eremenko, D. Jakobson and N. Nadirashvili, On nodal sets and nodal domains on S2 and R2. Festival Yves Colin de Verdire. Ann. Inst. Fourier (Grenoble) 57 (2007), no. 7, 23452360.

[F] H. Federer, *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York 1969.

[GL1] N. Garofalo and F. H. Lin, Monotonicity properties of variational integrals, A_p weights and unique continuation. Indiana Univ. Math. J. 35 (1986), no. 2, 245–268.

[GL2] N. Garofalo and F. H. Lin, Unique continuation for elliptic operators: a geometric-variational approach. Comm. Pure Appl. Math. 40 (1987), no. 3, 347–366.

[Ge] B. Georgiev, On the lower bound of the inner radius of nodal domains, arXiv:1607.03816.

[GRS] A. Ghosh, A. Reznikov and P. Sarnak Nodal domains of Maass forms I: Geom. Funct. Anal. 23 (2013), no. 5, 15151568. (arXiv:1207.6625.)

[Gi] V. M. Gichev, A Note on the Common Zeros of Laplace Beltrami Eigenfunctions. Ann. Global Anal. Geome. 26, 201208 (2004).

[Gr] M. Grohe, Fixed-point definability and polynomial time on graphs with excluded minors. J. ACM 59 (2012), no. 5, Art. 27.

[HHL] Q. Han, R. Hardt, and F.H. Lin, Geometric measure of singular sets of elliptic equations. Comm. Pure Appl. Math. 51 (1998), no. 11-12, 1425-1443.

[HL] Q. Han and F. H. Lin, *Nodal sets of solutions of Elliptic Differential Equations*, book in preparation (online at http://www.nd.edu/qhan/nodal.pdf).

[HHOHON] R. Hardt, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and N. Nadirashvili, Critical sets of solutions to elliptic equations. J. Differential Geom. 51 (1999), no. 2, 359-373.

[HS] R. Hardt and L. Simon, Nodal sets for solutions of elliptic equations. J. Differential Geom. 30 (1989), no. 2, 505-522.

[HW] P. Hartman and W. Wintner, On the local behavior of solutions of non-parabolic partial differential equations. III. Approximations by spherical harmonics. Amer. J. Math. 77 (1955), 453–474.

[Hez16] H. Hezari, Applications of small scale quantum ergodicity in nodal sets, arXiv:1606.02057.

[HS] H. Hezari and C. D. Sogge, A natural lower bound for the size of nodal sets, Analysis and PDE vol. 5 (2012), 1133-1137 (arXiv:1107.3440).

[HeW] H. Hezari and Z. Wang, Lower bounds for volumes of nodal sets: an improvement of a result of Sogge-Zelditch, Spectral geometry, 229235, Proc. Sympos. Pure Math., 84, Amer. Math. Soc., Providence, RI, 2012 (arXiv:1107.0092.)

[H] T. Hoffmann-Ostenhof, Problems for the number of nodal domains in "Geometric Aspects of Spectral Theory xiv" Oberwolfach Reports 33, page 2068-270, 2012.

[Jaf] H. Jaffee, Real algebraic curves. Topology 19 (1980), no. 1, 81-87.

[JM] D. Jakobson and D. Mangoubi, Tubular Neighborhoods of Nodal Sets and Diophantine Approximation, Amer. J. Math. 131 (2009), no. 4, 1109–1135 (arXiv:0707.4045).

[JN1] D. Jakobson and N. Nadirashvili, Eigenfunctions with few critical points. J. Differential Geom. 53 (1999), no. 1, 177–182.

[JN2] D. Jakobson and N. Nadirashvili, Quasi-symmetry of L^p norms of eigenfunctions. Comm. Anal. Geom. 10 (2002), no. 2, 397–408.

[JL] D. Jerison and G. Lebeau, Nodal sets of sums of eigenfunctions. Harmonic analysis and partial differential equations (Chicago, IL, 1996), 223–239, Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 1999.

[JJ] J. Jung, Quantitative Quantum Ergodicity and the nodal lines of Maass-Hecke cusp forms, arXiv:1301.6211 [math.NT] (2013).

[JZ] J. Jung and S. Zelditch, Number of nodal domains of eigenfunctions on non-positively curved surfaces with concave boundary, (arXiv:1401.4520).

[Kl] F. Klein, Über Realitätsverhltnisse bei einem beliebigen Geschlechte zugehrigen Normalkurve der φ. Math. Ann. 42 (1893) 129

[Ku1] I. Kukavica, Nodal volumes for eigenfunctions of analytic regular elliptic problems. J. Anal. Math. 67 (1995), 269–280.

[Ku2] I. Kukavica, Quantitative uniqueness for second-order elliptic operators. Duke Math. J. 91 (1998), no. 2, 225–240.

[Lew] H. Lewy, On the minimum number of domains in which the nodal lines of spherical harmonics divide the sphere. Comm. Partial Differential Equations 2 (1977), no. 12, 1233-1244.

[Ley] J. Leydold, On the number of nodal domains of spherical harmonics. Topology 35 (1996), no. 2, 301321.

[Lin] F.H. Lin, Nodal sets of solutions of elliptic and parabolic equations. Comm. Pure Appl. Math. 44 (1991), no. 3, 287–308.

[LM] A. Logunov and E. Malinnikova, Nodal sets of Laplace eigenfunctions: estimates of the Hausdorff measure in dimension two and three, arXiv:1605.02595.

[L1] A. Logunov, Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture, arXiv:1605.02589.

[L2] A. Logunov, Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure, arXiv:1605.02587.

[M1] D. Mangoubi, On the inner radius of a nodal domain. Canad. Math. Bull. 51 (2008), no. 2, 249-260.

[M2] D. Mangoubi, Local asymmetry and the inner radius of nodal domains. Comm. Partial Differential Equations 33 (2008), no. 7-9, 1611-1621.

[M3] D. Mangoubi, The volume of a local nodal domain. J. Topol. Anal. 2 (2010), no. 2, 259-275.

[M4] D. Mangoubi, A Remark on Recent Lower Bounds for Nodal Sets, Comm. Partial Differential Equations 36 (2011), no. 12, 2208–2212 (arXiv:1010.4579.)

[Na1] S. M. Natanzon, Klein surfaces. (Russian) Uspekhi Mat. Nauk 45 (1990), no. 6(276), 47–90, 189; translation in Russian Math. Surveys 45 (1990), no. 6, 53-108.

[Na2] S. M. Natanzon, *Moduli of Riemann surfaces, real algebraic curves, and their superanalogs*. Translated from the 2003 Russian edition by Sergei Lando. Translations of Mathematical Monographs, 225. American Mathematical Society, Providence, RI, 2004.

[NPS] F. Nazarov, L. Polterovich and M. Sodin, Sign and area in nodal geometry of Laplace eigenfunctions. Amer. J. Math. 127 (2005), no. 4, 879–910.

[NS] F. Nazarov and M. Sodin, On the number of nodal domains of random spherical harmonics. Amer. J. Math. 131 (2009), no. 5, 1337-1357.

[P] A. Pleijel, Remarks on Courant's nodal line theorem, Comm. Pure Appl. Math., 9, 543-550 (1956).

[PS] L. Polterovich and M. Sodin, Nodal inequalities on surfaces. Math. Proc. Cambridge Philos. Soc. 143 (2007), no. 2, 459–467 (arXiv:math/0604493).

[Ro] Guillaume Roy-Fortin, Nodal sets and growth exponents of Laplace eigenfunctions on surfaces, Anal. PDE 8 (2015), no. 1, 223-255 (arXiv:1408.7101.)

[Sa] A. Savo, Lower bounds for the nodal length of eigenfunctions of the Laplacian. Ann. Global Anal. Geom. 19 (2001), no. 2, 133-151.

[SD] R. Sena-Dias, Estimated transversality and rational maps. J. Symplectic Geom. 4 (2006), no. 2, 199-236.

[SY] R. Schoen and S.T. Yau, *Lectures on differential geometry.* Conference Proceedings and Lecture Notes in Geometry and Topology, I. International Press, Cambridge, MA, 1994.

[SoZ1] C. D. Sogge and S. Zelditch, Lower bounds on the hypersurface measure of nodal sets, Math. Research Letters 18 (2011), 27-39 (arXiv:1009.3573).

[SoZ2] C. D. Sogge and S. Zelditch, Lower bounds on the Hausdorff measure of nodal sets II, Math. Res. Lett. Volume 19 (2012) Number 6, pp. 1361-1364. (arXiv:1208.2045).

[St] A. Stern, Bemerkungen uber asymptotisches Verhalten von Eigenwerten und Eigenfunktionen, Göttingen (1925).

[Ta] G. Talenti, On functions, whose lines of steepest descent bend proportionally to level lines. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 10 (1983), no. 4, 587-605.

[TZ1] J. A. Toth and S. Zelditch, Counting nodal lines which touch the boundary of an analytic domain. J. Differential Geom. 81 (2009), no. 3, 649686.

[TZ2] J. A. Toth and S. Zelditch, Quantum ergodic restriction theorems: manifolds without boundary. Geom. Funct. Anal. 23 (2013), no. 2, 715775.

[U] K. Uhlenbeck, Generic properties of eigenfunctions. Amer. J. Math. 98 (1976), no. 4, 1059–1078

[W] G. Weichold, Über symmetrische Riemannsche Flächen und die Periodizitätsmodulen der zugerhörigen Abelschen Normalintegrale ersters Gattung. 1883. Leipziger dissertation.

[Y] S. T. Yau, Problem section, Seminar on Differential Geom., (ed. by S. T. Yau) Ann. of Math. Stud. 102 (1982) 669706; , Open problems in geometry, Differential geometry: partial differential equations on manifolds, Proc. Symp. Pure Math. 54 (1993) 128.

[Z1] S. Zelditch, Uniform distribution of eigenfunctions on compact hyperbolic surfaces. Duke Math. J. 55 (1987), no. 4, 919-941.

[Z2] S. Zelditch, Kuznecov sum formulae and Szegö limit formulae on manifolds. Comm. Partial Differential Equations 17 (1992), no. 1-2, 221-260.

[Z3] S. Zelditch, Hausdorff measure of nodal sets of analytic Steklov eigenfunctions. Math. Res. Lett. 22 (2015), no. 6, 1821-1842.

[Zw] M. Zworski, *Semiclassical analysis*. Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI (2012). MR2952218

CHAPTER 14

Eigenfunctions in the complex domain

In this section we consider eigenfunctions of real analytic Riemannian manifolds. On a real analytic Riemannian manifold (M, g) of dimension m, we analytically continue an orthonormal basis $\{\varphi_{\lambda_j}\}$ of eigenfunctions into the complexification $M_{\mathbb{C}}$ of M. As recalled in §14.1, eigenfunctions admit analytic continuations $\varphi^{\mathbb{C}}_{\lambda_j}$ to a maximal uniform 'Grauert tube'

$$M_\tau = \{\zeta \in M_{\mathbb{C}}, \sqrt{\rho}(\zeta) < \tau\} \tag{14.1}$$

independent of λ_j, where the radius is measured by the Grauert tube function $\sqrt{\rho}(\zeta)$ corresponding to g (see §14.1 and [**LS, GS1**]). As discussed below, given the metric g there is a relatively canonical identification of M_ε with a ball bundle $B^*_\varepsilon \subset T^*M$, so that one may view M_ε as phase space with a complex structure. The modulus squares

$$|\varphi_j^{\mathbb{C}}(\zeta)|^2 \colon M_\varepsilon \to \mathbb{R}_+ \tag{14.2}$$

are sometimes known as Husimi functions. They are holomorphic extensions of L^2-normalized functions but are not themselves L^2 normalized on M_ε. However, as will be discussed below, their L^2 norms may on the Grauert tubes (and their boundaries) can be determined. One can then ask how the mass of the normalized Husimi function is distributed in phase space, or how the L^p norms behave.

The first motivation to analytically continue eigenfunctions is that it enables us to give a relatively simple proof of the Donnelly-Fefferman theorem on nodal hypersurface volumes. Let Z_{φ_λ} be the nodal set of an eigenfunction φ_λ of eigenvalue $-\lambda^2$.

THEOREM 14.1. *Let (M, g) be a real analytic Riemannian manifold. Then, there exists constants $C, c > 0$ depending only on (M, g) so that*

$$c\lambda \leq \mathcal{H}^{n-1}(Z_{\varphi_\lambda}) \leq C\lambda.$$

The upper bound is based on proofs of nodal upper bounds in [**Z5, Z8**]. The key tool is the analytic continuation of the Poisson wave kernel to Grauert tubes and its description as a Fourier integral operator with complex phase. The Hausdorff measure of the complex nodal set

$$Z^{\mathbb{C}}_{\varphi_\lambda} = \{\zeta \in (\partial\Omega)_{\mathbb{C}} : \psi^{\mathbb{C}}_{\lambda_j}(\zeta) = 0\} \tag{14.3}$$

gives an upper bound for the Hausdorff measure of the real nodal set. In the complex domain one may use the Poincare-Lelong formula and a global Jensen type argument to give the upper bound (§14.31.1). The proof of the lower bound is also based on analytic continuation of eigenfunctions and is a new proof due to A. Brudnyi (§14.33).

Analytic continuation to the complex domain gives strong compactness properties to the sequence $u_j = \frac{1}{\lambda_j} \log |\varphi_j^{\mathbb{C}}|^2$ of pluri-subharmonic functions. This gives rise to a new weak* limit problem for eigenfunctions discussed in §14.30.2. It is possible to solve the problem in ergodic cases (§14.37) and in integrable cases [**Z10**]. This allows one to determine the equidistribution of complex nodal sets in these settings, something which seems out of reach in the real domain.

The relation between the distribution of zeros of analytic eigenfunctions in the real domain and complex domain is almost completely open, but we do give a few results of the relation. They mostly involve 'nodal restriction' or 'nodal intersection' theorems. There is enough control over complex nodal sets to prove results about their intersection with complexified curves such as geodesics. For instance, it is possible to determine the equidistribution of intersection points of complexified geodesics and nodal sets in the ergodic case ([**Z7**] and §14.45). One result is that if a curve H satisfies a non-degeneracy condition called 'goodness' relative to a sequence $\{\varphi_j\}$ of eigenfunctions, then the number of real zeros satisfies the bound $\#\{Z_{\varphi_j} \cap H\} \leq C_H \lambda_j$. This was proved in [**TZ1**] for Euclidean plane domains with analytic (or piecewise analytic) boundary and more recently in [**TZ4**] for any analytic Riemannian manifold without boundary by reversing the argument in [**Z7**]. One often has enough control over the complex nodal sets to pin down the equidistribution of intersections of complex nodal sets and complex geodesics.

An obvious omission in this section (and in the literature) is the analytic continuation of Poisson wave or heat kernels on analytic Riemannian manifolds (M, g) with non-empty analytic boundary ∂M. This is mainly due to the fact that we use Hadamard or Hörmander type parametrix constructions for the Dirichlet or Neumann wave or heat kernels to prove the analytic continuation, and such parametrices do not exist even for short times on general manifolds with boundary. It is very likely that one can use parametrix methods to analytically the Poisson wave kernel on analytic manifolds with analytic concave boundary using the Melrose-Taylor parametrix, but this has never been attempted as yet. It may be technically more convenient to analytically continue Dirichlet or Neumann heat kernels. At this time, there are no rigorous results in this direction.

We now give background on Grauert tubes, Szegő and Poisson kernels, on the analytic continuation of eigenfunctions and the wave group following [**Z5, Z7, Z8**].

14.1. Grauert tubes and complex geodesic flow

By a theorem of Bruhat-Whitney, a real analytic Riemannian manifold M admits a complexification $M_{\mathbb{C}}$, i.e. a complex manifold into which M embeds as a totally real submanifold. Corresponding to a real analytic metric g is a unique plurisubharmonic exhaustion function $\sqrt{\rho}$ on $M_{\mathbb{C}}$ (known as the Grauert tube function) satisfying two conditions (i) It satisfies the Monge-Ampère equation $(i\partial\bar{\partial}\sqrt{\rho})^n = \delta_{M,g}$ where $\delta_{M,g}$ is the delta function on M with density dV_g equal to the volume density of g; (ii) the Kähler metric $\omega_g = i\partial\bar{\partial}\rho$ on $M_{\mathbb{C}}$ agrees with g along M.

In fact,

$$\sqrt{\rho}(\zeta) = \frac{1}{2i}\sqrt{r^2_{\mathbb{C}}(\zeta, \bar{\zeta})}, \tag{14.4}$$

where $r^2(x, y)$ is the square of the distance function and $r^2_{\mathbb{C}}$ is its holomorphic extension to a small neighborhood of the anti-diagonal $(\zeta, \bar{\zeta})$ in $M_{\mathbb{C}} \times M_{\mathbb{C}}$. In the case of flat $\mathbb{R}^n$, $\sqrt{\rho}(x+i\xi) = 2|\xi|$ and in general $\sqrt{\rho}(\zeta)$ measures how far ζ reaches into the complexification of M. The open Grauert tube of radius τ is defined by $M_\tau = \{\zeta \in M_{\mathbb{C}}, \sqrt{\rho}(\zeta) < \tau\}$. We use the imprecise notation $M_{\mathbb{C}}$ to denote the open complexificaiton when it is not important to specify the radius.

The $(1,1)$ form $\omega = \omega_\rho := i\partial\bar{\partial}\rho$ defines a Kähler metric on $M_{\mathbb{C}}$. The Grauert tubes M_τ are strictly pseudo-convex domains in $M_{\mathbb{C}}$, whose boundaries ∂M_τ are strictly pseudo-convex CR manifolds. The boundary is endowed with the contact form

$$\alpha = \frac{1}{i}\partial\rho|_{\partial M_\tau} = d^c\sqrt{\rho}. \tag{14.5}$$

14.2. Analytic continuation of the exponential map

The geodesic flow is a Hamiltonian flow on T^*M. In fact, there are two standard choices of the Hamiltonian. In PDE it is most common to define the (real) homogeneous geodesic flow g^t of (M,g) as the Hamiltonian flow on T^*M generated by the Hamiltonian $|\xi|_g$ with respect to the standard Hamiltonian form ω. This Hamiltonian is real analytic on $T^*M\backslash 0$. In Riemannian geometry it is standard to let the time of travel equal $|\xi|_g$; this corresponds to the Hamiltonian flow of $|\xi|_g^2$, which is real analytic on all of T^*M. We denote its Hamiltonian flow by G^t. In general, we denote by Ξ_H the Hamiltonian vector field of a Hamiltonian H and its flow by $\exp t\Xi_H$. Both of the Hamiltonian flows

$$g^t = \exp t\Xi_{|\xi|_g} \quad \text{and} \quad G^t = \exp t\Xi_{|\xi|_g^2} \tag{14.6}$$

are important in analytic continuation of the wave kernel. The exponential map is the map $\exp_x : T^*M \to M$ defined by $\exp_x \xi = \pi G^t(x,\xi)$ where π is the standard projection.

We denote by $\text{inj}(x)$ the injectivity radius of (M,g) at x, i.e. the radius r of the largest ball on which $\exp_x : B_rM \to M$ is a diffeomorphism to its image. Since (M,g) is real analytic, $\exp_x t\xi$ admits an analytic continuation in t and the imaginary time exponential map

$$E\colon B_\varepsilon^*M \to M_{\mathbb{C}}, \quad E(x,\xi) = \exp_x i\xi \tag{14.7}$$

is, for small enough ε, a diffeomorphism from the ball bundle B_ε^*M of radius ε in T^*M to the Grauert tube M_ε in $M_{\mathbb{C}}$. We have $E^*\omega = \omega_{T^*M}$ where $\omega = i\partial\bar{\partial}\rho$ and where ω_{T^*M} is the canonical symplectic form; and also $E^*\sqrt{\rho} = |\xi|$ [**GS1, LS**]. It follows that E^* conjugates the geodesic flow on B^*M to the Hamiltonian flow $\exp t\Xi_{\sqrt{\rho}}$ of $\sqrt{\rho}$ with respect to ω, i.e.

$$E(g^t(x,\xi)) = \exp t\Xi_{\sqrt{\rho}}(\exp_x i\xi).$$

14.3. Maximal Grauert tubes

A natural definition of *maximal Grauert tube* is the maximum value of ε so that (14.7) is a diffeomorphism. We refer to this radius as the *maximal geometric tube radius*. But for purposes of this paper, another definition of maximality is relevant: the maximal tube on which all eigenfunctions extend holomorphically. A closely related definition is the maximal tube to which the Poisson kernel (14.42) extends holomorphically. We refer to the radius as the *maximal analytic tube radius*.

A natural question is to relate these notions of maximal Grauert tube has not been explored. We therefore define the radii more precisely:

DEFINITION 14.2. (1) The maximal geometric tube radius τ_g is the largest radius ε for which E (14.7) is a diffeomorphism.

(2) The maximal analytic tube radius τ_{an} $M_{\tau_{an}} \subset M_{\mathbb{C}}$ is the maximal tube to which all eigenfunctions extend holomorphically and to which the anti-diagonal $U(2i\tau, \zeta, \bar{\zeta})$ of the Poisson kernel admits an analytic continuation.

It is possible to prove that $\tau_g = \tau_{an}$. In §14.6 we sketch the proof that τ_{an} is the maximal radius for which the coefficients of Δ_g have holomorphic extensions. This radius is similar to the geometric radius, since the leading coefficients are geometric. But the coefficients of the first degree terms are not quite geometric in the same sense and at this time of writing the geometric radius has not been related to the maximal domain in which Δ_g extends holomorphically. The proof which is based on holomorphic extensions of solutions of analytic PDE across non-characteristic hypersurfaces. We found a similar argument in [**KS**] in the case of locally symmetric spaces but employing additional arguments.

14.4. Model examples

We consider some standard examples to clarify these analytic continuations.

(i) **Complex tori:** The complexification of the torus $M = \mathbb{R}^m/\mathbb{Z}^m$ is $M_{\mathbb{C}} = \mathbb{C}^m/\mathbb{Z}^m$. The adapted complex structure to the flat metric on M is the standard (unique) complex structure on $\mathbb{C}^m$. The complexified exponential map is $\exp_x^{\mathbb{C}}(i\xi) = z := x + i\xi$, while the distance function $r(x,y) = |x-y|$ extends to $r_{\mathbb{C}}(z,w) = \sqrt{(z-w)^2}$. Then $\sqrt{\rho}(z,\bar{z}) = \sqrt{(z-\bar{z})^2} = \pm 2i|\mathrm{Im} z| = \pm 2i|\xi|$.

The complexified cotangent bundle is $T^*M_{\mathbb{C}} = \mathbb{C}^m/\mathbb{Z}^m \times \mathbb{C}^m$, and the holomorphic geodesic flow is the entire holomorphic map

$$G^t(\zeta, p_\zeta) = (\zeta + tp_\zeta, p_\zeta).$$

(ii) ***n*-sphere:** (See [**GS1**].) The unit sphere $x_1^2 + \cdots + x_{n+1}^2 = 1$ in $\mathbb{R}^{n+1}$ is complexified as the complex quadric

$$S^n_{\mathbb{C}} = \{(z_1, \ldots, z_n) \in \mathbb{C}^{n+1} : z_1^2 + \cdots + z_{n+1}^2 = 1\}.$$

If we write $z_j = x_j + i\xi_j$, the equations become $|x|^2 - |\xi|^2 = 1, \langle x, \xi\rangle = 0$. The geodesic flow $G^t(x,\xi) = (\cos t|\xi|)x + (\sin t|\xi|)\frac{\xi}{|\xi|}, -|\xi|(\sin t|\xi|)x + (\cos t|\xi|)\xi)$ on T^*S^n complexifies to

$$\begin{aligned} G^t(Z,W) = (\cos t\sqrt{W\cdot W})Z + (\sin t\sqrt{W\cdot W}))\frac{W}{\sqrt{W\cdot W}} \\ - \sqrt{W\cdot W}(\sin t\sqrt{W\cdot W}))Z + (\cos t\sqrt{W\cdot W})W, \qquad (Z,W) \in T^*S^m_{\mathbb{C}}. \end{aligned} \tag{14.8}$$

Here, the real cotangent bundle is the subset of $T^*\mathbb{R}^{n+1}$ of (x,ξ) such that $x \in S^n, x\cdot\xi = 0$ and the complexified cotangent bundle $T^*S^n_{\mathbb{C}} \subset T^*\mathbb{C}^{n+1}$ is the set of vectors $(Z,W) : Z\cdot W = 0$. We note that although $\sqrt{W\cdot W}$ is singular at $W = 0$, both $\cos\sqrt{W\cdot W}$ and $\sqrt{W\cdot W}\sin t\sqrt{W\cdot W})$ are holomorphic. The Grauert tube function equals

$$\sqrt{\rho}(z) = i\cosh^{-1}|z|^2, \qquad z \in S^n_{\mathbb{C}}.$$

It is globally well defined on $S^n_{\mathbb{C}}$. The characteristic conoid is defined by $\cosh\frac{1}{i}\sqrt{\rho} = \cosh\tau$.

(iii) **Hyperbolic space:** The hyperboloid model of hyperbolic space is the hypersurface in $\mathbb{R}^{n+1}$ defined by

$$\mathbf{H}^n = \{x_1^2 + \cdots x_n^2 - x_{n+1}^2 = -1, \;\; x_n > 0\}.$$

Then,

$$H_{\mathbb{C}}^n = \{(z_1, \dots, z_{n+1}) \in \mathbb{C}^{n+1} : z_1^2 + \cdots z_n^2 - z_{n+1}^2 = -1\}.$$

In real coordinates $z_j = x_j + i\xi_j$, this is:

$$\langle x, x\rangle_L - \langle \xi, \xi\rangle_L = -1, \;\; \langle x, \xi\rangle_L = 0$$

where $\langle , \rangle_L$ is the Lorentz inner product of signature $(n, 1)$. Hence the complexified hyperbolic space is the hypersurface in $\mathbb{C}^{n+1}$ given by the same equations.

We obtain $\mathbf{H}_{\mathbb{C}}^n$ from $S_{\mathbb{C}}^n$ by the map $(z', z_{n+1}) \to (iz', z_{n+1})$. The complexified geodesic flow is given for $(Z, W) \in T^*\mathbf{H}^m$ by

$$\begin{aligned}(14.9) \quad G^t(Z, W) = (\cosh t\sqrt{\langle W, W\rangle_L} Z + (\sinh t\sqrt{\langle W, W\rangle_L}))\frac{W}{\sqrt{\langle W, W\rangle_L})} \\ - \sqrt{\langle W, W\rangle_L})(\sinh t\sqrt{\langle W, W\rangle_L}))Z + (\cosh t\sqrt{\langle W, W\rangle_L}))W).\end{aligned}$$

The Grauert tube function is

$$\sqrt{\rho}(z) = \cos^{-1}(\|x\|_L^2 + \|\xi\|_L^2 - \pi)/\sqrt{2}.$$

The radius of maximal Grauert tube is $\varepsilon = 1$ or $r = \pi/\sqrt{2}$.

14.5. Analytic continuation of eigenfunctions

A function f on a real analytic manifold M is real analytic, $f \in C^\omega(M)$, if and only if it satisfies the Cauchy estimates

$$(14.10) \qquad |D^\alpha f(x)| \leq K L^{|\alpha|} \alpha! \quad \text{for some } K, L > 0.$$

In place of all derivatives it is sufficient to use powers of Δ. In the language of Baouendi-Goulaouic [**BG1, BG2, BG3**], the Laplacian of a compact real analytic Riemannian manifold has the property of iterates, i.e., the real analytic functions are precisely the functions satisfying Cauchy estimates relative to Δ:

$$(14.11) \quad C^\omega(M) = \{u \in C^\infty(M) : \exists L > 0, \; \forall k \in \mathbf{N}, \; \|\Delta^k u\|_{L^2(M)} \leq L^{k+1}(2k)!\}.$$

It is classical that all of the eigenfunctions extend holomorphic to a fixed Grauert tube.

Theorem 14.3 (Morrey-Nirenberg Theorem). *Let $P(x, D)$ be an elliptic differential operator in Ω with coefficients which are analytic in Ω. If $u \in \mathcal{D}'(\Omega)$ and $P(x, D)u = f$ with $f \in C^\omega(\Omega)$, then $u \in C^\omega(\Omega)$.*

The proof shows that the radius of convergence of the solution is determined by the radius of convergence of the coefficients.

Let us consider examples of holomorphic continuations of eigenfunctions:

- On the flat torus $\mathbb{R}^m/\mathbb{Z}^m$, the real eigenfunctions are $\cos\langle k, x\rangle, \sin\langle k, x\rangle$ with $k \in 2\pi\mathbb{Z}^m$. The complexified torus is $\mathbb{C}^m/\mathbb{Z}^m$ and the complexified eigenfunctions are $\cos\langle k, \zeta\rangle, \sin\langle k, \zeta\rangle$ with $\zeta = x + i\xi$.

- On the unit sphere S^m, eigenfunctions are restrictions of homogeneous harmonic functions on $\mathbb{R}^{m+1}$. The latter extend holomorphically to holomorphic harmonic polynomials on $\mathbb{C}^{m+1}$ and restrict to holomorphic function on $S^m_{\mathbb{C}}$.
- On $\mathbf{H}^m$, one may use the hyperbolic plane waves $e^{(i\lambda+1)\langle z,b\rangle}$, where $\langle z,b\rangle$ is the (signed) hyperbolic distance of the horocycle passing through z and b to 0. They may be holomorphically extended to the maximal tube of radius $\pi/4$.
- On compact hyperbolic quotients $\mathbf{H}^m/\Gamma$, eigenfunctions can be then represented by Helgason's generalized Poisson integral formula [**H**],

$$\varphi_\lambda(z) = \int_B e^{(i\lambda+1)\langle z,b\rangle}\, dT_\lambda(b).$$

Here, $z \in D$ (the unit disc), $B = \partial D$, and $dT_\lambda \in \mathcal{D}'(B)$ is the boundary value of φ_λ, taken in a weak sense along circles centered at the origin 0. To analytically continue φ_λ it suffices to analytically continue $\langle z,b\rangle$. Writing the latter as $\langle \zeta,b\rangle$, we have:

$$\varphi^{\mathbb{C}}_\lambda(\zeta) = \int_B e^{(i\lambda+1)\langle \zeta,b\rangle}\, dT_\lambda(b). \tag{14.12}$$

In [**BG2**, Theorem 2] and [**BGH**, Theorem 1.2] it is proved that the operator Δ has the iterate property if and only if, for all $b > 1$, each eigenfunction extends holomorphically to some Grauert tube M_τ and satisfies

$$\sup_{z\in M_\tau} |\varphi^{\mathbb{C}}_{\lambda_j}(z)| \le b^{\lambda_j} \sup_{x\in M} |\varphi_{\lambda_j}(x)|. \tag{14.13}$$

The concept of Grauert was not actually used in these articles, so the relation between the growth rate and the Grauert tube function was not stated. But it again shows that all eigenfunctions extend to some fixed Grauert tube.

14.6. Maximal holomorphic extension

The question then arises if all eigenfunctions extend to the maximal Grauert tube allowed by the geometry as in Definition 14.2. This indeed is true and follows from theorems on extensions of holomorphic solutions of holomorphic PDE across non-characteristic hypersurfaces.

THEOREM 14.4. [**Zer, HoIII, KS, L, Z5**] *Let f be analytic in the open set $Z \subset \mathbb{C}^n$ and suppose that $P(x,D)u = f$ in the open set $Z_0 \subset Z$. If $z_0 \in Z \cap \partial Z_0$ and if Z_0 has a C^1 non-characteristic boundary at z_0, then u can be analytically continued as a solution of $P(x,D)u = f$ in a neighborhood of z_0.*

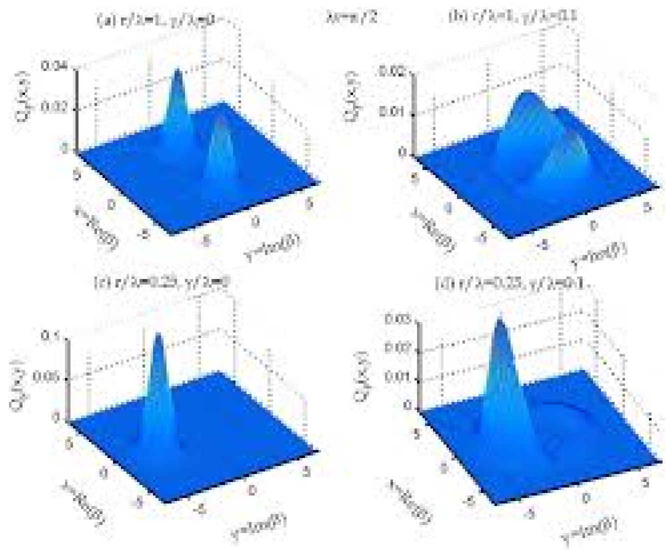
The idea of the proof is to rewrite the equation as a Cauchy problem with respect to the non-characteristic hypersurface and to apply the Cauchy Kowaleskaya theorem. To employ the theorem we need to verify that the hypersurfaces ∂M_τ are non-characteristic for the complexified Laplacian $\Delta_{\mathbb{C}}$, i.e., that $\sum_{i,j} g^{ij}(\zeta)\frac{\partial\sqrt{\rho}}{\partial\zeta_i}\frac{\partial\sqrt{\rho}}{\partial\zeta_j} \ne 0$. To prove this, we observe that in the real domain $g(\nabla r^2, \nabla r^2) = 4r^2$. In this formula $r^2 = r^2(x,y)$ and we differentiate in x. We now analytically continue the identity in $x \to \zeta, y, \to \bar\zeta$ and differentiate only with the holomorphic derivatives $\frac{\partial}{\partial\zeta_j}$. From (14.4), we get

$$g_{\mathbb{C}}(\partial r^2_{\mathbb{C}}(\zeta,\bar\zeta), \partial r^2_{\mathbb{C}}(\zeta,\bar\zeta)) = -4r^2_{\mathbb{C}}(\zeta,\bar\zeta) = \rho(\zeta,\bar\zeta) > 0.$$

Hence Zerner's theorem applies to the maximal domain to which Δ_g extends holomorphically, and we can analytically continue eigenfunctions across any point of any ∂M_τ for $\tau < \tau_g$, the maximal radius of a Grauert tube in which the coefficients of $\Delta_{\mathbb{C}}$ are defined and holomorphic. We can take the union of the open sets where $\varphi_j^{\mathbb{C}}$ has a holomorphic extension to obtain a maximal domain of holomorphy. If it fails to be M_{τ_g} there exists a point ζ with $\sqrt{\rho}(\zeta) < \tau_g$ so that $\varphi_j^{\mathbb{C}}$ cannot be holomorphically extended across ∂M_τ at ζ. This contradicts Zerner's theorem and shows that the maximal domain must be the maximal M_τ to which Δ_g extends holomorphically.

14.7. Husimi functions

The (L^2-normalizations of the) modulus squares (14.2) are sometimes known as Husimi functions (after [**Hu**]). They are holomorphic extensions of L^2-normalized functions but are not themselves L^2 normalized on M_ε. However, as will be discussed below, their L^2 norms may on the Grauert tubes (and their boundaries) can be determined. One can then ask how the mass of the normalized Husimi function is distributed in phase space, or how the L^p norms behave.



One of the general problems of quantum dynamics is to determine all of the weak* limits of the sequence,

$$\{\frac{|\varphi_j^{\mathbb{C}}(z)|^2}{\|\varphi_j^{\mathbb{C}}\|_{L^2(\partial M_\varepsilon)}} d\mu_\varepsilon\}_{j=1}^\infty.$$

Here, $d\mu_\varepsilon$ is the natural measure on ∂M_ε corresponding to the contact volume form on $S_\varepsilon^* M$. Recall that a sequence μ_n of probability measures on a compact space X is said to converge weak* to a measure μ if $\int_X f\, d\mu_n \to \int_X f\, d\mu$ for all $f \in C(X)$. We refer to Theorem 14.43 for the ergodic case. In the integrable case one has localization results showing that complex zeros lie on hypersurfaces (see [**Z10**]).

14.8. Poisson wave operator and Szegő projector on Grauert tubes

In this section, we introduce the Poisson wave operator, the Szegő projector, and complexified spectral projections and state some basic results on analytic continuation and growth (Theorem 14.5 and Theorem 14.7). The theorems on analytic continuation of the Poisson wave kernel are proved in §14.14 following [**Z5, L**]. The theorems on growth of complexified and tempered spectral projections are proved in §14.26 with refinements sketched in §14.28.

14.9. Poisson operator and analytic Continuation of eigenfunctions

The half-wave group of (M, g) is the unitary group $U(t) = e^{it\sqrt{\Delta}}$ generated by the square root of the positive Laplacian. Its Schwartz kernel is a distribution on $\mathbb{R} \times M \times M$ with the eigenfunction expansion

$$U(t, x, y) = \sum_{j=0}^{\infty} e^{it\lambda_j} \varphi_j(x)\varphi_j(y). \tag{14.14}$$

By the Poisson operator we mean the analytic continuation of$U(t)$ to positive imaginary time:

$$e^{-\tau\sqrt{-\Delta}} = U(i\tau). \tag{14.15}$$

The eigenfunction expansion then converges absolutely to a real analytic function on $\mathbb{R}_+ \times M \times M$.

Let $A(\tau)$ denote the operator of analytic continuation of a function on M to the Grauert tube M_τ. Since

$$U_{\mathbb{C}}(i\tau)\varphi_\lambda = e^{-\tau\lambda}\varphi_\lambda^{\mathbb{C}}, \tag{14.16}$$

it is simple to see that

$$A(\tau) = U_{\mathbb{C}}(i\tau)e^{\tau\sqrt{-\Delta}} \tag{14.17}$$

where $U_{\mathbb{C}}(i\tau, \zeta, y)$ is the analytic continuation of the Poisson kernel in x to M_τ. In terms of the eigenfunction expansion, one has

$$U(i\tau, \zeta, y) = \sum_{j=0}^{\infty} e^{-\tau\lambda_j} \varphi_j^{\mathbb{C}}(\zeta)\varphi_j(y), \quad (\zeta, y) \in M_\varepsilon \times M. \tag{14.18}$$

This is a very useful observation because $U_{\mathbb{C}}(i\tau)e^{\tau\sqrt{-\Delta}}$ is a Fourier integral operator with complex phase and can be related to the geodesic flow. The analytic continuability of the Poisson operator to M_τ implies that every eigenfunction analytically continues to the same Grauert tube.

14.10. Analytic continuation of the Poisson wave group

The analytic continuation of the Possion wave kernel to M_τ in the x variable is discussed in detail in [**Z5**] and ultimately derives from the analysis by Hadamard of his parametrix construction. We only briefly discuss it here and refer to [**Z5**] for further details. In the case of Euclidean $\mathbb{R}^n$ and its wave kernel $U(t, x, y) = \int_{\mathbb{R}^n} e^{it|\xi|}e^{i\langle \xi, x-y\rangle}\, d\xi$, which analytically continues to $t + i\tau, \zeta = x + ip \in \mathbb{C}_+ \times \mathbb{C}^n$ as the integral

$$U_{\mathbb{C}}(t + i\tau, x + ip, y) = \int_{\mathbb{R}^n} e^{i(t+i\tau)|\xi|}e^{i\langle \xi, x+ip-y\rangle}\, d\xi.$$

The integral clearly converges absolutely for $|p| < \tau$.

Exact formulae of this kind exist for S^m and $\mathbf{H}^m$. For a general real analytic Riemannian manifold, there exists an oscillatory integral expression for the wave kernel of the form

$$U(t,x,y) = \int_{T_y^*M} e^{it|\xi|_{g_y}} e^{i\langle \xi, \exp_y^{-1}(x)\rangle} A(t,x,y,\xi)\, d\xi, \tag{14.19}$$

where $A(t,x,y,\xi)$ is a polyhomogeneous amplitude of order 0. The holomorphic extension of (14.19) to the Grauert tube $|\zeta| < \tau$ in x at time $t = i\tau$ then has the form

$$U_{\mathbb{C}}(i\tau,\zeta,y) = \int_{T_y^*} e^{-\tau|\xi|_{g_y}} e^{i\langle \xi, \exp_y^{-1}(\zeta)\rangle} A(t,\zeta,y,\xi)\, d\xi, \qquad \zeta = x+ip. \tag{14.20}$$

14.11. Complexified spectral projections

The next step is to holomorphically extend the spectral projectors $d\Pi_{[0,\lambda]}(x,y) = \sum_j \delta(\lambda-\lambda_j)\varphi_j(x)\varphi_j(y)$ of $\sqrt{\Delta}$. The complexified diagonal spectral projections measure is defined by

$$d_\lambda \Pi^{\mathbb{C}}_{[0,\lambda]}(\zeta,\bar\zeta) = \sum_j \delta(\lambda-\lambda_j)|\varphi_j^{\mathbb{C}}(\zeta)|^2. \tag{14.21}$$

Henceforth, we generally omit the superscript and write the kernel as $\Pi^{\mathbb{C}}_{[0,\lambda]}(\zeta,\bar\zeta)$. This kernel is not a tempered distribution due to the exponential growth of $|\varphi_j^{\mathbb{C}}(\zeta)|^2$. Since many asymptotic techniques assume spectral functions are of polynomial growth, we simultaneously consider the damped spectral projections measure

$$d_\lambda P^\tau_{[0,\lambda]}(\zeta,\bar\zeta) = \sum_j \delta(\lambda-\lambda_j) e^{-2\tau\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2, \tag{14.22}$$

which is a temperate distribution as long as $\sqrt{\rho}(\zeta) \le \tau$. When we set $\tau = \sqrt{\rho}(\zeta)$ we omit the τ and put

$$d_\lambda P_{[0,\lambda]}(\zeta,\bar\zeta) = \sum_j \delta(\lambda-\lambda_j) e^{-2\sqrt{\rho}(\zeta)\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2. \tag{14.23}$$

The integral of the spectral measure over an interval I gives

$$\Pi_I(x,y) = \sum_{j:\lambda_j\in I} \varphi_j(x)\varphi_j(y).$$

Its complexification gives the spectral projections kernel along the anti-diagonal

$$\Pi_I(\zeta,\bar\zeta) = \sum_{j:\lambda_j\in I} |\varphi_j^{\mathbb{C}}(\zeta)|^2, \tag{14.24}$$

and the integral of (14.22) gives its temperate version

$$P_I^\tau(\zeta,\bar\zeta) = \sum_{j:\lambda_j\in I} e^{-2\tau\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2, \tag{14.25}$$

or in the crucial case of $\tau = \sqrt{\rho}(\zeta)$,

$$P_I(\zeta,\bar\zeta) = \sum_{j:\lambda_j\in I} e^{-2\sqrt{\rho}(\zeta)\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2, \tag{14.26}$$

14.12. Poisson operator as a complex Fourier integral operator

The damped spectral projection measure $d_\lambda\ P^\tau_{[0,\lambda]}(\zeta,\bar\zeta)$ (14.22) is dual under the real Fourier transform in the t variable to the restriction

$$U(t+2i\tau,\zeta,\bar\zeta)=\sum_j e^{(-2\tau+it)\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2 \tag{14.27}$$

to the anti-diagonal of the mixed Poisson wave group. The adjoint of the Poisson kernel $U(i\tau,x,y)$ also admits an anti-holomorphic extension in the y variable. The sum (14.27) are the diagonal values of the complexified wave kernel

$$U(t+2i\tau,\zeta,\bar\zeta')=\int_M U(t+i\tau,\zeta,y)E(i\tau,y,\bar\zeta')\,dV_g(x) \tag{14.28}$$

$$=\sum_j e^{(-2\tau+it)\lambda_j}\varphi_j^{\mathbb{C}}(\zeta)\overline{\varphi_j^{\mathbb{C}}(\zeta')}. \tag{14.29}$$

We obtain (14.28) by orthogonality of the real eigenfunctions on M.

Since $U(t+2i\tau,\zeta,y)$ takes its values in the CR holomorphic functions on ∂M_τ, we consider the Sobolev spaces $\mathcal{O}^{s+\frac{n-1}{4}}(\partial M_\tau)$ of CR holomorphic functions on the boundaries of the strictly pseudo-convex domains M_ε, i.e.,

$$\mathcal{O}^{s+\frac{m-1}{4}}(\partial M_\tau)=W^{s+\frac{m-1}{4}}(\partial M_\tau)\cap\mathcal{O}(\partial M_\tau),$$

where W_s is the sth Sobolev space and where $\mathcal{O}(\partial M_\varepsilon)$ is the space of boundary values of holomorphic functions. The inner product on $\mathcal{O}^0(\partial M_\tau)$ is with respect to the Liouville measure

$$d\mu_\tau=(i\partial\bar\partial\sqrt{\rho})^{m-1}\wedge d^c\sqrt{\rho}. \tag{14.30}$$

We then regard $U(t+i\tau,\zeta,y)$ as the kernel of an operator from $L^2(M)\to\mathcal{O}^0(\partial M_\tau)$. It equals its composition $\Pi_\tau\circ U(t+i\tau)$ with the Szegő projector

$$\Pi_\tau: L^2(\partial M_\tau)\to\mathcal{O}^0(\partial M_\tau)$$

for the tube M_τ, i.e., the orthogonal projection onto boundary values of holomorphic functions in the tube.

This is a useful expression for the complexified wave kernel, because $\tilde\Pi_\tau$ is a complex Fourier integral operator with a small wave front relation. More precisely, the real points of its canonical relation form the graph Δ_Σ of the identity map on the symplectic one $\Sigma_\tau\subset T^*\partial M_\tau$ spanned by the real one-form $d^c\rho$, i.e.,

$$\Sigma_\tau=\{(\zeta;rd^c\rho(\zeta)):\zeta\in\partial M_\tau,\ r>0\}\subset T^*(\partial M_\tau). \tag{14.31}$$

We note that for each τ, there exists a symplectic equivalence $\Sigma_\tau\simeq T^*M$ by the map $(\zeta,rd^c\rho(\zeta))\to(E_{\mathbb{C}}^{-1}(\zeta),r\alpha)$, where $\alpha=\xi\cdot dx$ is the action form (cf. [**GS2**]).

The following result was first stated by Boutet de Monvel [**Bo**] and has been proved in detail in [**Z5, L, St**].

THEOREM 14.5. *$\Pi_\varepsilon\circ U(i\varepsilon)\colon L^2(M)\to\mathcal{O}(\partial M_\varepsilon)$ is a complex Fourier integral operator of order $-\frac{m-1}{4}$ associated to the canonical relation*

$$\Gamma=\{(y,\eta,\iota_\varepsilon(y,\eta)\}\subset T^*M\times\Sigma_\varepsilon.$$

Moreover, for any s,

$$\Pi_\varepsilon\circ U(i\varepsilon)\colon W^s(M)\to\mathcal{O}^{s+\frac{m-1}{4}}(\partial M_\varepsilon)$$

is a continuous isomorphism.

14.13. Complexified Poisson kernel as a complex Fourier integral operator

The following theorem is stated in [**Bo**]. For proofs, see [**Z5, L**].

THEOREM 14.6. *(See* [**Bo, GS2, GLS**]*.) For sufficiently small* $\tau > 0$, $U_{\mathbb{C}}(i\tau)$: $L^2(M) \to \mathcal{O}(\partial M_\tau)$ *is a Fourier integral operator of order* $-\frac{m-1}{4}$ *with complex phase associated to the canonical relation*

$$\Lambda = \{(y, \eta, \iota_\tau(y, \eta)\} \subset T^*M \times \Sigma_\tau.$$

Moreover, for any s,

$$U_{\mathbb{C}}(i\tau)\colon W^s(M) \to \mathcal{O}^{s+\frac{m-1}{4}}(\partial M_\tau)$$

is a continuous isomorphism.

The proof of Theorem 14.6 is barely sketched in [**Bo**]. However, the theorem follows almost immediately from the construction of the branched meromorphic Hadamard parametrix in Corollary 14.8, or alternatively from the analytic continuation of the Hörmander parametrix of §14.19. It suffices to show that either is a parametrix for $U_{\mathbb{C}}(i\tau, \zeta, y)$, i.e., differs from it by an analytic kernel (smooth would be sufficient by analytic wave front set considerations). But the Hadamard parametrix construction is an exact formula and actually gives a more precise description of the singularities of $U_{\mathbb{C}}(i\tau, \zeta, y)$ than is stated in Theorem 14.6. We briefly explain how either the Hadamard or Hörmander parametrix can be used to complete the proof.

Using the complexified Poisson wave kernel, one can prove the following sup-norm estimate:

PROPOSITION 14.7. *Suppose* (M, g) *is real analytic. Then*

$$\sup_{\zeta \in M_\tau} |\varphi_\lambda^{\mathbb{C}}(\zeta)| \leq C\lambda^{\frac{m+1}{2}} e^{\tau\lambda} \quad \textit{and} \quad \sup_{\zeta \in M_\tau} \left|\frac{\partial \varphi_\lambda^{\mathbb{C}}(\zeta)}{\partial \zeta_j}\right| \leq C\lambda^{\frac{m+3}{2}} e^{\tau\lambda}.$$

14.14. Analytic continuation of the Poisson wave kernel

In this section we prove Theorem 14.5 and Theorem 14.6, closely following [**Z5**]. Other closely related proofs can be found in [**L, St**].

14.15. Hörmander parametrix for the Poisson wave kernel

A more familiar construction of $U(t, x, y)$ and its analytic continuation which is particularly useful for small $|t|$ is the one based on the Fourier inversion formula. Its generalization to Riemannian manifolds is given by

$$U(t, x, y) = \int_{T_y^*M} e^{it|\xi|_{g_y}} e^{i\langle \xi, \exp_y^{-1}(x)\rangle} A(t, x, y, \xi)\, d\xi, \tag{14.32}$$

for (x, y) sufficiently close to the diagonal. We use this parametrix to prove Theorem 14.15 (2).

The amplitude is a polyhomognous symbol of the form

$$A(t, x, y, \xi) \sim \sum_{j=}^{\infty} A_j(t, x, y, \xi), \tag{14.33}$$

where the asymptotics are in the sense of the symbol topology and where

$$A_j(t,x,y,\tau\xi) = \tau^{-j} A_j(t,x,y,\xi), \qquad |\xi| \geq 1.$$

The principal term $A_0(t,x,y,\xi)$ equals 1 when $t=0$ on the diagonal, and the higher A_j are determined by transport equations discussed in [**DuG**].

It can be verified that in the case of real analytic (M,g), the amplitude is a classical formal analytic symbol (see §14.23). Hence if $\mathcal{A}(t,x,y,\xi)$ is a realization of the amplitude $A(t,x,y,\xi)$, then one obtains an analytic parametrix

$$U(t,x,y) = \int_{T_y^* M} e^{it|\xi|_{g_y}} e^{i\langle \xi, \exp_y^{-1}(x)\rangle} \mathcal{A}(t,x,y,\xi)\, d\xi, \tag{14.34}$$

which approximates the wave kernel for small $|t|$ and (x,y) near the diagonal up to a holomorphic error, whose amplitude is exponentially decaying in $|\xi|$.

14.16. Subordination to the heat kernel

The parametrix (14.19) can also be obtained by subordinating the Poisson wave kernel to the heat kernel. To make use of it, one needs to analytically the heat kernel to M_τ. This analytic continuation was studied by Golse-Leichtnam-Stenzel in [**GLS**], who proved the following: For any $x_0 \in M$ there exists $\varepsilon, \rho > 0$ and an open neighborhood W of x_0 in M_ε such that for $0 < t < 1$ and $(x,y) \in W \times W$,

$$E(t,x,y) = N(t,x,y) e^{-\frac{r^2(x,y)}{4t}} + R(t,x,y),$$

where

$$N(t,x,y) = \sum_{0 \leq j \leq \frac{1}{Ct}} W_j(x,y) t^j$$

as $t \downarrow 0^+$ where $W_j(x,y)$ are the Hadamard-Minakshisundaram-Pleijel heat kernel coefficients. is an analytic symbol of order $n/2$ with respect to t^{-1} in the sense of [**Sj**]. As above, the remainder is exponentially small,

$$|R(t,x,y)| \leq C e^{-\frac{\rho}{8t}}$$

with a uniform C in (x,y) as $t \downarrow 0^+$. The heat kernel itself obviously admits a holomorphic extension in the open subset $\mathrm{Re} r^2_{\mathbb{C}}(x,y) > 0$ ot $M_{\mathbb{C}} \times M_{\mathbb{C}}$.

14.17. Fourier integral distributions with complex phase

First, we review the relevant definitions (see [**HoIV**, §25.5] or [**MeSj**]). A Fourier integral distribution with complex phase on a manifold X is a distribution that can locally be represented by an oscillatory integral

$$A(x) = \int_{\mathbb{R}^N} e^{i\varphi(x,\theta)} a(x,\theta)\, d\theta,$$

where $a(x,\theta) \in S^m(X \times V)$ is a symbol of order m in a cone $V \subset \mathbb{R}^N$ and where the phase φ is a positive regular phase function, i.e., it satisfies

- $\mathrm{Im}\varphi \geq 0$;
- $d\frac{\partial \varphi}{\partial \theta_1}, \ldots, d\frac{\partial \varphi}{\partial \theta_N}$ are linearly independent complex vectors on
$$C_{\varphi\mathbb{R}} = \{(x,\theta) : d_\theta(x,\theta) = 0\};$$
- In the analytic setting (which is assumed in this article), φ admits an analytic continuation $\varphi_{\mathbb{C}}$ to an open cone in $X_{\mathbb{C}} \times V_{\mathbb{C}}$.

Define

$$C_{\varphi_{\mathbb{C}}} = \{(x,\theta) \in X_{\mathbb{C}} \times V_{\mathbb{C}} : \nabla_\theta \varphi_{\mathbb{C}}(x,\theta) = 0\}.$$

Then $C_{\varphi_{\mathbb{C}}}$ is a manifold near the real domain. One defines the Lagrangian submanifold $\Lambda_{\varphi_{\mathbb{C}}} \subset T^* X_{\mathbb{C}}$ as the image

$$(x,\theta) \in C_{\varphi_{\mathbb{C}}} \to (x, \nabla_x \varphi_{\mathbb{C}}(x,\theta)).$$

14.18. Analytic continuation of the Hadamard parametrix

As in §14.23, we can express $U_{\mathbb{C}}(i\tau, \zeta, y)$ as a local Fourier integral distribution with complex phase by rewriting the Hadamard series in Corollary 14.8 as oscillatory integrals. Here we assume that $\tau > 0, t \geq 0$.

A complication is that we can only use the complexified phase $\Gamma = t^2 - r^2$ in regions of complexified $\mathbb{R} \times M \times M$ where its imaginary part is ≥ 0. We could also use the phase $t - r$ (resp. $t + r$) in regions where $t + r \neq 0$ (resp. $t - r \neq 0$) and where the contour $\mathbb{R}_+$ can be deformed back to itself after the the change of variables $\theta \to (t + r)\theta$.

14.19. Analytic continuation of the Hörmander parametrix

As was the case in $\mathbb{R}^n$, the parametrix (14.34) admits an analytic continuation in time to a strip $\{t + i\tau : \tau < \tau_{an}, |t| < 1\}$. In the space variables, the parametrix then admits an analytic continuation to complex x, y satisfying $|r_{\mathbb{C}}(x,y)| \leq \tau$.

The analytically continued parametrix (14.19) approximates the true analytically continued Poisson kernel up to a holomorphic kernel. More precisely, for any $x_0 \in M$ and $\tau > 0$, there exists $\varepsilon, \rho > 0$ and an open neighborhood W of x_0 in M_τ such that for $|t| < 1$ and $(x,y) \in W \times W$,

$$U(t + i\tau, x, y) = \int_{T_y^* M} e^{-\tau |\xi|_{g_y}} e^{i\langle \xi, \exp_y^{-1}(x)\rangle} \mathcal{A}(t + i\tau, x, y, \xi) d\xi + R(t, x, y), \tag{14.35}$$

where $R(t,x,y)$ is holomorphic for small $|t|$ and for (x,y) near the diagonal.

The parametrix is only defined near the diagonal where $\exp_y^{-1}$ is defined. However one can extend it to a global holomorphic kernel away from $\mathcal{C}_{\mathbb{C}}$ by cutting off the first term of (14.19) with a smooth cutoff $\chi(x,y)$ supported near the diagonal in $M_\tau \times M_\tau$ and then solving a $\bar{\partial}$ problem on the Grauert tube (or a $\bar{\partial}_b$ problem on its boundary) to extend the kernel to be globally holomorphic (resp. CR). We refer to [**Z5**] for a more detailed discussion. This gives an alternative to the Hadamard parametrix construction of Corollary 14.8.

This concludes the sketch of proof of Theorem 14.6.

14.20. $\Delta_g, \Box_g$ and characteristics

In the real domain, Δ is an elliptic operator with principal symbol $\sigma_\Delta(x,\xi) = \sum_{i,j=1}^n g^{ij}(x)\xi_i\xi_j$. Hence its characteristic set (the zero set of its symbol) consists only of the zero section $\xi = 0$ in T^*M. But when we continue it to the complex domain it develops a complex characteristic set

$$\mathrm{Char}(\Delta_{\mathbb{C}}) = \left\{(\zeta,\xi) \in T^* M_{\mathbb{C}} : \sum_{i,j=1}^n g^{ij}(\zeta)\xi_i\xi_j = 0\right\}. \tag{14.36}$$

The wave operator on the product spacetime $(\mathbb{R} \times M, dt^2 - g_x)$ is given by

$$\Box_g = \frac{\partial^2}{\partial t^2} + \Delta_g.$$

The unusual sign in front of Δ_g is due to the sign normalization above making the Laplacian non-negative. Again we omit the subscript when the metric is fixed. The characteristic variety of $\Box$ is the zero set of its symbol

$$\sigma_\Box(t, \tau, x, \xi) = \tau^2 - |\xi|_x^2,$$

that is,

$$\text{Char}(\Box) = \{(t, \tau, x, \xi) \in T^*(\mathbb{R} \times M) : \tau^2 - |\xi|_x^2 = 0\}. \tag{14.37}$$

The null-bicharacteristic flow of $\Box$ is the Hamiltonian flow of $\tau^2 - |\xi|_x^2$ on $\text{Ch}(\Box)$. Its graph is thus

$$\Lambda = \{(t, \tau, x, \xi, y, \eta) : \tau^2 - |\xi|_x^2 = 0,\ G^t(x, \xi) = (y, \eta)\} \subset T^*(\mathbb{R} \times M \times M).$$

14.21. Characteristic variety and characteristic conoid

Following [**H**], we put

$$\Gamma(t, x, y) = t^2 - r^2(x, y). \tag{14.38}$$

Here, $r(x, y)$ is the distance between x, y. It is singular at $r = 0$ and also when y is in the "cut locus" of x. In this article we only consider (x, y) so that $r(x, y) < \text{inj}(x)$, where $\text{inj}(x)$ is the injectivity radius at x, i.e., is the largest ε so that

$$\exp_x : B^*_{x,\varepsilon} M \to M$$

is a diffeomorphism to its image. The injectivity radius $inj(M, g)$ is the maximum of $\text{inj}(x)$ for $x \in M$. Thus, we work in a sufficiently small neighborhood of the diagonal so that cut points do not occur.

The squared distance $r^2(x, y)$ is smooth in a neighborhood of the diagonal. On a simply connected manifold $(\tilde{M}, g)$ without conjugate points, it is globally smooth on $\tilde{M} \times \tilde{M}$. We recall that 'without conjugate points' means that $\exp_x : T_x M \to M$ is non-singular for all x.

The characteristic conoid is the set

$$\mathcal{C} = \{(t, x, y) : r(x, y) < \text{inj}(x),\ r^2(x, y) = t^2\} \subset \mathbb{R} \times M \times M. \tag{14.39}$$

It separates $\mathbb{R} \times M \times M$ into the forward/backward semi-cones

$$\mathcal{C}_\pm = \{(t, x, y) : t^2 - r^2(x, y) > 0,\ \pm t > 0\}.$$

The complexification of $\mathcal{C}$ is the complex characteristic conoid

$$\mathcal{C}_\mathbb{C} = \{(t, x, y) : r_\mathbb{C}^2(x, y) = t^2\} \subset \mathbb{C} \times M_\mathbb{C} \times M_\mathbb{C}. \tag{14.40}$$

We note that $\mathcal{C}_\mathbb{R} \subset \mathcal{C}_\mathbb{C}$ is a totally real submanifold. Another totally real submanifold of central importance in this article is the 'diagonal' (or anti-diagonal) conoid,

$$\mathcal{C}_\Delta = \{(2i\tau, \zeta, \bar{\zeta}) : \tau \in \mathbb{R}_+,\ \zeta, \bar{\zeta} \in \partial M_\tau\}. \tag{14.41}$$

By definition, $r_\mathbb{C}^2(\zeta, \bar{\zeta}) = -4\tau^2$ if $\zeta \in \partial M_\tau$.

14.22. Hadamard parametrix for the Poisson wave kernel

We are most interested in the Hadamard parametrix for the half-wave kernel, which does not seem to have been discussed in the literature. We are more generally interested in the Poisson wave semi-group $e^{i(t+i\tau)\sqrt{\Delta}}$ for $\tau > 0$. The Poisson wave kernel

$$U(t+i\tau, x, y) = \sum_j e^{i(t+i\tau)\lambda_j} \varphi_j(x)\varphi_j(y) \tag{14.42}$$

is a real analytic kernel which possesses an analytic extension to a Grauert tube. Thus, there exists a non-zero analytic radius $\tau_{an} > 0$ so that the Poisson kernel admits a holomorphic extension $U(t+i\tau, \zeta, y)$ to $M_\tau \times M$ for $\tau \leq \tau_{an}$. Since

$$U(i\tau)\varphi_\lambda = e^{-\tau\lambda}\varphi_\lambda^{\mathbb{C}}, \tag{14.43}$$

the eigenfunctions analytically extend to the same maximal tube as does $U(i\tau)$.

We would like to construct a Hadamard type parametrix for (14.42). We may derive it from the Feynman-Hadamard fundamental solution using that

$$\frac{d}{dt}\frac{e^{i|t|\sqrt{-\Delta}}}{\sqrt{-\Delta}} = i\,\mathrm{sgn}(t) e^{i|t|\sqrt{-\Delta}} \tag{14.44}$$

and

$$e^{it\sqrt{-\Delta}} = \frac{1}{i}H(t)\frac{d}{dt}\frac{e^{i|t|\sqrt{-\Delta}}}{\sqrt{-\Delta}} - \frac{1}{i}H(-t)\frac{d}{dt}\frac{e^{-i|t|\sqrt{-\Delta}}}{\sqrt{-\Delta}}. \tag{14.45}$$

Hence,

$$\frac{1}{i}\frac{d}{dt}U_F(t) = e^{it\sqrt{-\Delta}}, \qquad t > 0. \tag{14.46}$$

The restriction to $t > 0$ is consistent with the fact that $e^{it\sqrt{-\Delta}}(x,y)$ has the singularity $((t+i0)^2 - r^2)^{\frac{-m}{2}}$ (in odd spacetime dimensions) while $U_F(t)$ has the singularity $(t^2 - r^2 + i0)^{\frac{2-m}{2}}$. We note (again) that $((t+i0)^2 - r^2)^\alpha = (t^2 - r^2 + i0)^\alpha$ for $t > 0$.

From Theorem 6.1 we conclude:

Corollary 14.8. *Let (M, g) be real analytic. Then with the U_j, V_k, W_ℓ defined as in Theorem 6.1, we have:*

- *In odd spacetime dimensions, for $t > 0$ the Poisson wave kernel $U(t + i\tau, x, y)$ for $\tau > 0$ has the form $A\Gamma^{\frac{-m}{2}}$ where $A = \sum_{j=0}^\infty A_j \Gamma^j$ with A_j holomorphic. The series converges absolutely to a holomorphic function for $|\Gamma| < \varepsilon$ sufficient small, i.e., near the characteristic conoid.*
- *In even spacetime dimensions, for $t > 0$, the Poisson wave kernel has the form $B\Gamma^{\frac{-m}{2}} + C\log\Gamma + D$ where the coefficients B, C, D are holomorphic in a neighborhood of $\mathcal{C}_{\mathbb{C}}$, and have the same Γ expansions as A.*

We use this parametrix to prove Theorem 14.15 (1).

14.23. Hadamard parametrix as an oscillatory integral with complex phase

Corollary 14.8 gives a precise description of the singularities of the Poisson wave propagator. It implicitly describes the kernel as a Fourier integral kernel. We now make this description explicit in the real domain. In the following sections, we extend the description to the complex domain.

We first express $\Gamma^{\frac{-m}{2}+j}$ as an oscillatory integral with one phase variable using the well-known identity

$$\int_0^\infty e^{i\theta\sigma}\theta_+^\lambda\, d\lambda = ie^{i\lambda\pi/2}\Gamma(\lambda+1)(\sigma+i0)^{-\lambda-1}. \tag{14.47}$$

At least formally, this leads to the representation

$$\int_0^\infty e^{i\theta(t^2-r^2)}\theta_+^{\frac{n-1}{2}-j}\, d\theta = ie^{i(\frac{n-1}{2}-j)\pi/2}\Gamma\left(\frac{n-1}{2}-j+1\right)(t^2-r^2+i0)^{j-\frac{n-1}{2}-1}$$

for the principal term of the Poisson wave. Here, the notation $\Gamma = t^2 - r^2$ unfortunately clashes with that for the Gamma function, and we temporarily write out its definition.

In even space dimensions, the Hadamard parametrix for the Hadamard-Feynman fundamental solution thus has the form

$$\begin{aligned} &\sum_{j=0}^\infty U_j(t,x,y)\Gamma^{\frac{1-n}{2}+j} \\ &\quad = \int_0^\infty e^{i\theta(t^2-r^2)}\left(\sum_{j=0}^\infty U_j(t,x,y)(ie^{i(\frac{n-1}{2}-j)\pi/2})^{-1}\frac{\theta_+^{\frac{n-3}{2}-j}}{\Gamma(\frac{n-3}{2}-j+1)}\right)d\theta. \end{aligned} \tag{14.48}$$

Here we follow Hadamard's notation, but it is simpler to re-define the coefficients U_j so that the Γ-factors appear on the left side as in [**Be**] (7). We thus define

$$\mathcal{U}_j(t,x,y) = \left((ie^{i(\frac{n-1}{2}-j)\pi/2})^{-1}\frac{1}{\Gamma(\frac{n-3}{2}-j+1)}\right)U_j(t,x,y).$$

By the duplication formula $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin\pi z}$ with $z = \frac{m}{2} - k - \frac{\alpha}{2}$, i.e.,

$$\Gamma\left(\frac{m}{2}-j-\frac{\alpha}{2}\right) = (-1)^j\frac{\pi}{\sin\pi(\frac{m}{2}-\frac{\alpha}{2})}\frac{1}{\Gamma(-\frac{m}{2}+1+j+\frac{\alpha}{2})},$$

it follows that

$$U_j(t,x,y) = \left((-1)^j\frac{\pi}{\sin\pi(\frac{m}{2}-\frac{\alpha}{2})}\frac{1}{\Gamma(-\frac{m}{2}+1+j+\frac{\alpha}{2})}\right)\mathcal{U}_j(t,x,y),$$

so that the formula in odd spacetime dimensions becomes

$$\begin{aligned} &C_n\frac{1}{\sin\pi(\frac{m}{2}-\frac{\alpha}{2})}\sum_{j=0}^\infty(-1)^j\frac{\mathcal{U}_j(t,x,y)}{\Gamma(-\frac{m}{2}+1+j+\frac{\alpha}{2})}(t^2-r^2)^{-\frac{m-2}{2}+j} \\ &\quad = \int_0^\infty e^{i\theta(t^2-r^2)}\left(\sum_{j=0}^\infty \mathcal{U}_j(t,x,y)\theta_+^{\frac{n-3}{2}-j}\right)d\theta. \end{aligned} \tag{14.49}$$

The amplitude in the right side of (14.49) is then a formal analytic symbol,

$$A(t,x,y,\theta) = \sum_{j=0}^\infty \mathcal{U}_j(t,x,y)\theta_+^{\frac{n-3}{2}-j}, \tag{14.50}$$

Due to the Gamma-factors appearing in the identity (14.47), convergence of the series on the left side of (14.49) does not imply convergence of the series (14.50).

However, there exists a realization of the formal symbol (14.50) by a holomorphic symbol

$$\mathcal{A}(t,x,y,\theta) = \sum_{0 \le j \le \frac{\theta}{eC}} \mathcal{U}_j(t,x,y)\theta_+^{\frac{n-3}{2}-j},$$

and one obtains an analytic parametrix

$$U(t,x,y) = \int_0^\infty e^{i\theta\Gamma}\mathcal{A}(t,x,y,\theta)\,d\theta \tag{14.51}$$

that approximates the wave kernel for small $|t|$ and (x,y) near the diagonal up to a holomorphic error, whose amplitude is exponentially decaying in θ. Here, we recall (see [**Sj**, p.3 and §9]) that a *classical formal analytic symbol* ([**Sj**, p.3]) on a domain $\Omega \subset \mathbb{C}^n$ is a formal semi-classical series

$$a(z,\lambda) = \sum_{k=0}^{\infty} a_k(z)\lambda^{-k},$$

where $a_k(z,\lambda) \in \mathcal{O}(\Omega)$ for all $\lambda > 0$. Then for some $C > 0$, the $a_k(z) \in \mathcal{O}(\Omega)$ satisfy

$$|a_k(z)| \le C^{k+1}k^k, \quad k = 0,1,2,\dots.$$

A realization of the formal symbol is a genuine holomorphic symbol of the form,

$$a(z,\lambda) = \sum_{0 \le k \le \frac{\lambda}{eC}} a_k(z)\lambda^{-k}.$$

It is an analytic symbol since, with the index restriction,

$$|a_k(z)\lambda^{-k}| \le C_\Omega \left(\frac{Ck}{\lambda}\right)^k \le Ce^{-k}.$$

Hence the series converges uniformly on Ω to a holomorphic function of z for each λ.

Returning to (14.50), the Hadamard-Riesz coefficients $\mathcal{U}_j$ are determined inductively by the transport equations

$$\begin{cases} \dfrac{\Theta'}{2\Theta}\mathcal{U}_0 + \dfrac{\partial \mathcal{U}_0}{\partial r} = 0, \\ 4ir(x,y)\{(\dfrac{k+1}{r(x,y)} + \dfrac{\Theta'}{2\Theta})\mathcal{U}_{j+1} + \dfrac{\partial \mathcal{U}_{j+1}}{\partial r}\} = \Delta_y \mathcal{U}_j. \end{cases} \tag{14.52}$$

whose solutions are given by:

$$\begin{cases} \mathcal{U}_0(x,y) = \Theta^{-\frac{1}{2}}(x,y), \\ \mathcal{U}_{j+1}(x,y) = \Theta^{-\frac{1}{2}}(x,y)\displaystyle\int_0^1 s^k \Theta(x,x_s)^{\frac{1}{2}}\Delta_2 \mathcal{U}_j(x,x_s)\,ds \end{cases} \tag{14.53}$$

where x_s is the geodesic from x to y parametrized proportionately to arc-length and where Δ_2 operates in the second variable.

As discussed above, the representation (14.49) does not suffice when n is odd, since $\Gamma(z)$ and θ_+^z have poles at the negative integers. To rescue the representation when n is odd, we need to use the distributions θ_+^{-n} with $n = 1, 2, \dots$, defined as follows (see [**HoI**]):

$$\theta_+^{-k}(\varphi) = \int_0^\infty (\log\theta)\varphi^{(k)}(\theta)d\theta/(k-1)! + \varphi^{(k-1)}(0)(\sum_{j=1}^{k} 1/j)/(k-1)!.$$

This family behaves in an unusual way under derivation:

$$\frac{d}{d\theta}\theta_+^{-k} = -k\theta_+^{-k-1} + \frac{(-1)^k}{k!}\delta_0^{(k)}$$

(see [**HoI**, (3.2.2)"]) and is therefore sometimes avoided in the Hadamard-Riesz parametrix construction (as in [**Be**]).

However, we have already constructed the parametrices and only want to express them in terms of the above oscillatory integrals to make contact with Fourier integral operator theory. In odd space dimensions, the Hadamard parametrices can be written in the form

(14.54)
$$\int_0^\infty e^{i\theta\Gamma}\left(U_0(t,x,y)\theta_+^m + \cdots + U_m\theta_+^0\right) d\theta + \int_0^\infty e^{i\theta\Gamma}\left(U_{m+1}\theta_+^{-1} + U_{m+2}\theta_+^{-2} + \cdots\right) d\theta.$$

Again the amplitude is a formal symbol. To produce a genuine amplitude it needs to be replace by a realization which approximates it modulo a holomorphic symbol which is exponentially decaying in θ.

We are paying close attention to the regularization of the integral at $\theta = 0$, but only the behavior of the amplitude as $\theta \to \infty$ is relevant to the singularity. The terms with θ_+^{-k} for $k > 0$ produce logarithmic terms in the kernel. If we use a smooth cutoff at $\theta = 0$, we obtain distributions of the form

$$u_\mu(\Gamma) = \int_{\mathbb{R}} e^{i\theta\Gamma}\chi(\theta)\theta^\mu \, d\theta$$

where $\chi(\theta) = 1$ for $\theta \geq 1$ and $\chi(\theta) = 0$ for $\theta \leq \frac{1}{2}$. Then

$$u_{-k}(\Gamma) = i^{k+1}\Gamma^{k-1}\log\Gamma \mod C^\infty$$

Hence the terms with negative powers of θ_+ in (14.54) produce the logarithmic terms and the holomorphic terms.

Above, we discussed the Hadamard-Feynman fundamental solution, but for $t > 0$ we only need to differentiate it in t (according to Proposition 14.8) to obtain the parametrices for the Poisson wave group. Away from the characteristic conoid the Schwartz kernels of the Poisson wave group and Hadamard-Feynman fundamental solution are holomorphic by the theorem on propagation of analytic wave front sets [**Sj**]. The Fourier integral structure and mapping properties follow immediately from the order of the amplitude and from the exact formula for the phase.

14.24. Tempered spectral projector and Poisson semi-group as complex Fourier integral operators

To study the tempered spectral projection kernels (14.25), we further need to continue $U_{\mathbb{C}}(t, \zeta, y)$ anti-holomorphically in the y variable. The discussion is similar to the holomorphic case except that we need to double the Grauert tube radius to obtain convergence. We thus have (cf. (14.27))

$$U_{\mathbb{C}}(t+2i\tau,\zeta,\bar{\zeta}) = \sum_j e^{(-2\tau+it)\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2 = \int_{\mathbb{R}} e^{it\lambda} d_\lambda P^\tau_{[0,\lambda]}(\zeta,\bar{\zeta}). \tag{14.55}$$

Properties of these kernels may be obtained from kernels which are analytically continued in one variable only from the formula (14.28)

$$U_{\mathbb{C}}(t+2i\tau,\zeta,\bar{\zeta}') = \int_M U(t+i\tau,\zeta,y)U_{\mathbb{C}}(i\tau,y,\bar{\zeta}')\,dV_g(x) \tag{14.56}$$

$$= \sum_j e^{(-2\tau+it)\lambda_j}\varphi_j^{\mathbb{C}}(\zeta)\overline{\varphi_j^{\mathbb{C}}(\zeta')}. \tag{14.57}$$

We have

Proposition 14.9. *For small $t,\tau>0$ and for sufficiently small $\tau \geq \sqrt{\rho}(\zeta) > 0$, there exists a realization $\mathcal{B}(t,\zeta,\bar{\zeta},\theta)$ of a formal analytic symbol $B(t,\zeta,\bar{\zeta},\theta)$ so that as tempered distributions on $\mathbb{R}\times M_\tau$,*

$$U_{\mathbb{C}}(t+2i\tau,\zeta,\bar{\zeta}) = \int_0^\infty e^{i\theta((t+2i\tau)-2i\sqrt{\rho}(\zeta))}\mathcal{B}(t,\zeta,\bar{\zeta},\theta)\,d\theta + R(t+2i\tau,\zeta,\bar{\zeta}), \tag{14.58}$$

where $R(t+2i\tau,\zeta,\bar{\zeta})$ is the restriction to the anti-diagonal of a holomorphic kernel. Moreover

- *$\theta((t+2i\tau)-2i\sqrt{\rho}(\zeta))$ is a phase of positive type;*
- *If $\sqrt{\rho}(\zeta)<\tau$ the entire kernel is locally holomorphic;*
- *If $\sqrt{\rho}(\zeta)=\tau$ then*

$$U_{\mathbb{C}}(t+2i\tau,\zeta,\bar{\zeta}) = \int_0^\infty e^{i\theta t}\mathcal{B}(t,\zeta,\bar{\zeta},\theta)d\theta + R(t+2i\tau,\zeta,\bar{\zeta}). \tag{14.59}$$

Proof. We use the Hadamard parametrix (Corollary 14.8) for $U(t+2i\tau,\zeta,\bar{\zeta})$ and use (14.4) to simplify the phase, i.e., we write

$$\Gamma(t+2i\tau,\zeta,\bar{\zeta}) = (t+2i\tau-2i\sqrt{\rho})(t+2i\tau+2i\sqrt{\rho})$$

in the Hadamard parametrix in Corollary 14.8. The factors of $(t+2i\tau+2i\sqrt{\rho})$ are non-zero when $\tau>0$ and can be absorbed into the Hadamard coefficients. We denote the new amplitude by $\mathcal{B}$ to distinguish it from the amplitude in Corollary 14.8. We then express each term as a Fourier integral distribution of complex type with phase $t+2i\tau-2i\sqrt{\rho}$. It is manifestly of positive type. On ∂M_τ, $t+2i\tau-2i\sqrt{\rho}$ simplifies to t. □

14.25. Complexified wave group and Szegő kernels

As in [**Z2**] it will also be necessary for us to understand the composition $U_{\mathbb{C}}(i\tau)^*U_{\mathbb{C}}(i\tau)$. In this regard, it is useful to introduce the Szegő kernels Π_τ of M_τ, i.e., the orthogonal projections

$$\Pi_\tau\colon L^2(\partial M_\tau,d\mu_\tau)\to H^2(\partial M_\tau,d\mu_\tau), \tag{14.60}$$

where $d\mu_\tau$ is the Liouville volume form. Here, $H^2(\partial M_\tau,d\mu_\tau)$ is the Hardy space of boundary values of holomorphic functions in M_τ which belong to $L^2(\partial M_\tau,d\mu_\tau)$. It is simple to prove that the restrictions of $\{\varphi^{\mathbb{C}}_{\lambda_j}\}$ to ∂M_τ is a basis of $H^2(\partial M_\tau,d\mu_\tau)$. The Szegő projector Π_τ is a complex Fourier integral operator with a positive complex canonical relation. The real points of its canonical relation form the graph Δ_Σ of the identity map on the symplectic cone $\Sigma_\tau \subset T^*\partial M_\tau$ (14.31). We refer to [**Z5**] for further background. We only need the first statement in the following:

LEMMA 14.10. *Let $\Psi^s(X)$ denote the class of pseudo-differential operators of order s on X. Then,*

- $U_{\mathbb{C}}(i\tau)^* U_{\mathbb{C}}(i\tau) \in \Psi^{-\frac{m-1}{2}}(M)$ *with principal symbol* $|\xi|_g^{-(\frac{m-1}{2})}$.
- $U_{\mathbb{C}}(i\tau) \circ U_{\mathbb{C}}(i\tau)^* = \Pi_\tau A_\tau \Pi_\tau$ *where* $A_\tau \in \Psi^{\frac{m-1}{2}}(\partial M_\tau)$ *has principal symbol* $|\sigma|_g^{\frac{m-1}{2}}$ *as a function on* Σ_τ.

PROOF. This follows from Proposition 14.6. The first statement is a special case of the following Lemma from [**Z2**]: Let $a \in S^0(T^*M - 0)$. Then for all $0 < \tau < \tau_{\max}(g)$, we have:

$$U(i\tau)^* \Pi_\tau a \Pi_\tau U(i\tau) \in \Psi^{-\frac{m-1}{2}}(M),$$

with principal symbol equal to $a(x,\xi)\ |\xi|_g^{-(\frac{m-1}{2})}$.

The second statement follows from Theorem 14.6 and the composition theorem for complex Fourier integral operators. We note that

$$U_{\mathbb{C}}(i\tau) \circ U_{\mathbb{C}}(i\tau)^*(\zeta, \zeta') = \sum_j e^{-2\tau\lambda_j} \varphi^{\mathbb{C}}_{\lambda_j}(\zeta) \overline{\varphi^{\mathbb{C}}_{\lambda_j}(\zeta')}. \tag{14.61}$$

□

14.26. Growth of complexified eigenfunctions

14.26.1. A Siciak-Zaharjuta extremal function for Grauert tubes. Before defining the analogues, let us first recall the definitions of relative maximal or extremal pluri-subharmonic functions satisfying bounds on a pair $E \subset \Omega \subset \mathbb{C}^m$ where Ω is a bounded open set. There are two definitions:

- The pluri-complex Green's function relative to a subset $E \subset \Omega$, defined [**Si**] as the upper semi-continuous regularization $V^*_{E,\Omega}$ of
$$V_{E,\Omega}(z) = \sup\{u(z) : u \in \mathrm{PSH}(\Omega),\ u|_E \le 0, u|_{\partial\Omega} \le 1\}$$
- Let $\|f\|_E := \sup_{z\in E} |f(z)|$ and $\mathcal{P}^N$ be the space of all complex analytic polynomials of degree N. The Siciak-Zaharjuta extremal function relative to $E \subset \Omega$ is defined by
$$\log \Phi_E := \limsup_{N\to\infty} \log \Phi^N_E, \tag{14.62}$$
$$\log \Phi^N_E(\zeta) := \sup\left\{\frac{1}{N} \log |p_N(\zeta)| : p \in \mathcal{P}^N_E\right\}, \tag{14.63}$$
where $\mathcal{P}^N_E = \{p \in \mathcal{P}^N : \|p\|_E \le 1,\ \|p\|_\Omega \le e^N\}$.

Siciak proved that $\log \Phi_E = V_E$ (see [**Si**, Theorem 1]). Intuitively, there are enough polynomials that one can obtain the sup by restricting to polynomials.

There are analogous definitions in the case of unit co-disc bundles in the dual of a positive holomorphic Hermitian line bundle $L \to M$ over a Kähler manifold. In the case of $\mathbb{CP}^n$, one defines

$$V_K(z) = \sup\{u(z) : u \in \mathcal{L},\ u \le 0 \text{ on } K\}$$

where $\mathcal{L}$ denotes the Lelong class of all global PSH functions u on $\mathbb{C}^n$ with $u(z) \leq c_u + \log(1+|z|)$.

We now define an analogue of the Siciak-Zaharjuta extremal function for Grauert tubes in the special case where $E = M$, the underlying real manifold. The Riemannian analogue of $\mathcal{P}^N$ is the space

$$\mathcal{H}^\lambda = \left\{ p = \sum_{j:\lambda_j \in I_\lambda} a_j \varphi_{\lambda_j}^{\mathbb{C}},\ a_1, \ldots,\ a_{N(\lambda)} \in \mathbb{R} \right\}$$

spanned by the eigenfunctions with 'degree' $\lambda_j \leq \lambda$. Here, $N(\lambda) = \#\{j : \lambda_j \in I_\lambda\}$. As above, we could let $I_\lambda = [0, \lambda]$ or $I_\lambda = [\lambda, \lambda + c]$ for some $c > 0$. It is simpler to work with L^2 based norms than sup norms, and so we define

$$S\mathcal{H}_M^\lambda = \left\{ \psi = \sum_{j:\lambda_j \leq \lambda} a_j \varphi_{\lambda_j}^{\mathbb{C}},\ \sum_{j=1}^{N(\lambda)} |a_j|^2 = 1 \right\}.$$

DEFINITION 14.11. The Riemannian Siciak-Zaharjuta extremal function (with respect to the real locus M) is defined by:

$$\begin{cases} \log \Phi_M^\lambda(\zeta) = \sup\{\frac{1}{\lambda} \log |\psi(\zeta)| : \psi \in S\mathcal{H}_M^\lambda\}, \\ \log \Phi_M = \limsup_{\lambda \to \infty} \log \Phi_M^\lambda. \end{cases} \tag{14.64}$$

REMARK 14.12. One could define the analogous notion for any set $E \subset M_\tau$ with

$$S\mathcal{H}_E^\lambda = \{p \in \mathcal{H}^\lambda,\ \|p\|_{L^2(E)} \leq 1\}.$$

But we only discuss the results for $E = M$.

One could also define the pluri-complex Green's function of M_τ as follows:

DEFINITION 14.13. Let (M, g) be a real analytic Riemannian manifold, let M_τ be an open Grauert tube, and let $E \subset M_\tau$. The Riemannian pluri-complex Green's function with respect to (E, M_τ, g) is defined by

$$V_{g,E,\tau}(\zeta) = \sup\{u(z) : u \in \mathrm{PSH}(M_\tau),\ u|_E \leq 0,\ u|_{\partial M_\tau} \leq \tau\}.$$

It is obvious that $V_{g,M,\tau}(\zeta) \geq \sqrt{\rho}(\zeta)$ and it is almost standard that $V_{g,M,\tau}(\zeta) = \sqrt{\rho}(\zeta)$. See [**GZ**, Proposition 4.1] or [**BT**, Corollary 9]. The set $M = (\sqrt{\rho})^{-1}(0)$ is often called the center. As proved in [**LS**, Theorem 1.1], there are no smooth exhaustion functions solving the exact homogeneous complex Monge'Ampère equation. Hence u must be singular on its minimum set. In [**HW1**] it is proved that the minimum set of strictly PSH function is totally real.

14.26.2. Statement of results. Our first results concern the logarithmic asymptotics of the complexified spectral projections.

THEOREM 14.14 (See also [**Z5**]). *Let* $I_\lambda = [0, \lambda]$. *Then*

(1) $\log \Phi_M^\lambda(\zeta) = \frac{1}{\lambda} \log \Pi_{I_\lambda}^{\mathbb{C}}(\zeta, \bar{\zeta})$.
(2) $\log \Phi_M = \lim_{\lambda \to \infty} \log \Phi_M^\lambda = \sqrt{\rho}$.

To prove the Theorem, it is convenient to study the tempered spectral projection measures (14.26), or in differentiated form (14.22):

$$d_\lambda P_{[0,\lambda]}^\tau(\zeta, \bar{\zeta}) = \sum_j \delta(\lambda - \lambda_j) e^{-2\tau\lambda_j} |\varphi_j^{\mathbb{C}}(\zeta)|^2, \tag{14.65}$$

which is a temperate distribution on $\mathbb{R}$ for each ζ satisfying $\sqrt{\rho}(\zeta) \leq \tau$. When we set $\tau = \sqrt{\rho}(\zeta)$ we omit the τ and write as in (14.23):

$$d_\lambda P_{[0,\lambda]}(\zeta, \bar{\zeta}) = \sum_j \delta(\lambda - \lambda_j) e^{-2\sqrt{\rho}(\zeta)\lambda_j} |\varphi_j^{\mathbb{C}}(\zeta)|^2. \tag{14.66}$$

The advantage of the tempered projections is that they have polynomial asymptotics and one can use standard Tauberian theorems to analyze their growth.

We prove the following one-term local Weyl law for complexified spectral projections:

THEOREM 14.15. *On any compact real analytic Riemannian manifold (M, g) of dimension n, we have, with remainders uniform in ζ,*

(1) *For $\sqrt{\rho}(\zeta) \geq \frac{C}{\lambda}$,*

$$P_{[0,\lambda]}(\zeta, \bar{\zeta}) = (2\pi)^{-n} \left(\frac{\lambda}{\sqrt{\rho}}\right)^{\frac{n-1}{2}} \left(\frac{\lambda}{(n-1)/2+1} + O(1)\right);$$

(2) *For $\sqrt{\rho}(\zeta) \leq \frac{C}{\lambda}$,*

$$P_{[0,\lambda]}(\zeta, \bar{\zeta}) = (2\pi)^{-n} \lambda^n \left(1 + O(\lambda^{-1})\right).$$

This implies new bounds on pointwise norms on complexified eigenfunctions, improving those of [**GLS**]. inequality gives

COROLLARY 14.16. *Suppose (M, g) is real analytic of dimension n, and that $I_\lambda = [0, \lambda]$. Then*

(1) *For $\tau \geq \frac{C}{\lambda}$ and $\sqrt{\rho}(\zeta) = \tau$, there exists $C > 0$ so that*

$$C\lambda_j^{-\frac{n-1}{2}} e^{\tau\lambda} \leq \sup_{\zeta \in M_\tau} |\varphi_\lambda^{\mathbb{C}}(\zeta)| \leq C\lambda^{\frac{n-1}{4}+\frac{1}{2}} e^{\tau\lambda}.$$

(2) *For $\tau \leq \frac{C}{\lambda}$, and $\sqrt{\rho}(\zeta) = \tau$, there exists $C > 0$ so that*

$$|\varphi_\lambda^{\mathbb{C}}(\zeta)| \leq \lambda^{\frac{n-1}{2}}.$$

The lower bound of Corollary 14.16 (1) combines Theorem 14.15 with Gårding's inequality. The upper bound sharpens the estimates claimed in [**Bo, GLS**],

$$\sup_{\zeta \in M_\tau} |\varphi_\lambda^{\mathbb{C}}(\zeta)| \leq C_\tau \lambda^{n+1} e^{\tau\lambda}. \tag{14.67}$$

The improvement is due to using spectral asymptotics rather than a crude Sobolev inequality.

14.27. Siciak extremal functions: Proof of Theorem 14.14 (1)

In this section we prove Theorem 14.14. First we prove a pointwise local Weyl law in the complex domain.

14.27.1. Proof of Theorem 14.14 (2). This follows from Theorem 14.15 together with the following:

LEMMA 14.17 ([**Z5**]). *For any* $\tau = \sqrt{\rho}(\zeta) > 0$*, and for any* $\delta > 0$*,*

$$2\sqrt{\rho}(\zeta) - \frac{\log|\delta|}{\lambda} + O\left(\frac{\log\lambda}{\lambda}\right) \leq \frac{1}{\lambda}\log\Pi_{[0,\lambda]}(\zeta,\bar{\zeta}) \leq 2\sqrt{\rho}(\zeta) + O(\frac{\log\lambda}{\lambda}),$$

hence

$$\lim_{\lambda\to\infty}\frac{1}{\lambda}\log\Pi_{[0,\lambda]}(\zeta,\bar{\zeta}) = 2\sqrt{\rho}(\zeta).$$

PROOF. For the upper bound, we use

$$\Pi_{[0,\lambda]}(\zeta,\bar{\zeta}) \leq e^{2\lambda\sqrt{\rho(\zeta)}} \sum_{j:\lambda_j\in[0,\lambda]} e^{-2\sqrt{\rho(\zeta)}\lambda_j}|\varphi^{\mathbb{C}}_{\lambda_j}(\zeta)|^2 = e^{2\lambda\sqrt{\rho(\zeta)}}\, P_{]0,\lambda]}(\zeta,\bar{\zeta}). \tag{14.68}$$

We then take $\frac{1}{\lambda}\log$ of both sides and apply Theorem 14.15 to conclude the proof.

The lower bound is subtler for reasons having to do with the distribution of eigenvalues (see the Remark below). It is most natural to prove two-term Weyl asymptotics for $P_{[0,\lambda]}(\zeta,\bar{\zeta})$ and to deduce Weyl asymptotics for short spectral intervals $[\lambda,\lambda+1]$. But that requires an analysis of the singularity of the trace of the complexified wave group for longer times than a short interval around $t = 0$ and we postpone the more refined analysis until [**Z9**].

Instead we use the longer intervals $[(1-\delta)\lambda,\lambda]$ for some $\delta > 0$. We clearly have

$$e^{2(1-\delta)\lambda\sqrt{\rho(\zeta)}} \sum_{j:\lambda_j\in[(1-\delta)\lambda,\lambda]} e^{-2\sqrt{\rho(\zeta)}\lambda_j}|\varphi^{\mathbb{C}}_{\lambda_j}(\zeta)|^2 \leq \Pi_{[0,\lambda]}(\zeta,\bar{\zeta}). \tag{14.69}$$

By Theorem 14.15,

$$\sum_{j:\lambda_j\in[(1-\delta)\lambda,\lambda]} e^{-2\sqrt{\rho(\zeta)}\lambda_j}|\varphi^{\mathbb{C}}_{\lambda_j}(\zeta)|^2 = P_{[0,\lambda]}(\zeta,\bar{\zeta}) - P_{[0,(1-\delta)\lambda]}(\zeta,\bar{\zeta}) \tag{14.70}$$

$$= C_n(\tau)[1-(1-\delta)^n]\lambda^{\frac{n+1}{2}} + O(\lambda^{\frac{n-1}{2}}) \tag{14.71}$$

Taking $\frac{1}{\lambda}\log$ then gives

$$\frac{1}{\lambda}\log\Pi_{[0,\lambda]}(\zeta,\bar{\zeta}) \geq 2(1-\delta)\sqrt{\rho}(\zeta) - \frac{|\log\delta|}{\lambda} + O(\frac{\log\lambda}{\lambda}).$$

It follows that for all $\delta > 0$,

$$\liminf_{\lambda\to\infty}\frac{1}{\lambda}\log\Pi_{[0,\lambda]}(\zeta,\bar{\zeta}) \geq 2(1-\delta)\sqrt{\rho}(\zeta).$$

The conclusion of the Lemma follows from the fact that the left side is independent of δ. □

REMARK 14.18. The problematic issue in the lower bound is the width of I_λ. If (M,g) is a Zoll manifold, the eigenvalues of $\sqrt{\Delta}$ form clusters of width $O(\lambda^{-1})$ around an arithmetic progression $\{k+\frac{\beta}{4}\}$ for a certain Morse index β. Unless the intervals I_λ are carefully centered around this progression, P_{I_λ} could be zero. Hence we must use long spectral intervals if we do not analyze the long time behavior of the geodesic flow; for short ones no general lower bound exists.

14.27.2. Proof of Theorem 14.14 (1). We need to show that

$$\Pi^{\mathbb{C}}_{I_\lambda}(\zeta, \bar{\zeta}) = \sup \left\{ |\varphi(\zeta)|^2 : \varphi = \sum_{j: \lambda_j \in I} a_j \varphi^{\mathbb{C}}_{\lambda_j}, \ \|a\| = 1 \right\}.$$

We define the 'coherent state'

$$\Phi^z_\lambda(w) = \frac{\Pi^{\mathbb{C}}_{I_\lambda}(w, \bar{z})}{\sqrt{\Pi^{\mathbb{C}}_{I_\lambda}(z, \bar{z})}},$$

satisfying

$$\Phi^z_\lambda(w) = \sum_{j: I_\lambda} a_j \varphi^{\mathbb{C}}_j(w), \quad a_j = \frac{\overline{\varphi^{\mathbb{C}}_j(\zeta)}}{\sqrt{\Pi^{\mathbb{C}}_{I_\lambda}(z, \bar{z})}}, \quad \sum_j |a_j|^2 = 1.$$

ence, $\Phi^\zeta_{I_\lambda}$ is a competitor for the sup and since $|\Phi^\zeta_{I_\lambda}(\zeta)|^2 = \Pi_{I_\lambda}(\zeta, \bar{\zeta})$ one has

$$\Pi^{\mathbb{C}}_{I_\lambda}(\zeta, \bar{\zeta}) \leq \sup \left\{ |\psi(\zeta)|^2 : \psi = \sum_{j: \lambda_j \in I} a_j \varphi^{\mathbb{C}}_j, \ \|a\| = 1 \right\}.$$

By the Schwartz inequality for ℓ^2, for any $\psi = \sum_{j: \lambda_j \in I} a_j \varphi^{\mathbb{C}}_j$ one has

$$\left| \sum_{j: \lambda_j \in I} a_j \varphi^{\mathbb{C}}_j \right|^2 = |\langle a, \psi \rangle|^2 \leq \|a\|^2 \sum |\varphi^{\mathbb{C}}_j|^2 = \Pi_{I_\lambda}(\zeta, \bar{\zeta})$$

and one has

$$\Pi^{\mathbb{C}}_I(\zeta, \bar{\zeta}) \geq \sup \left\{ |\psi(\zeta)|^2 : \psi = \sum_{j: \lambda_j \in I} a_j \varphi^{\mathbb{C}}_j, \ \|a\| = 1 \right\}.$$

□

REMARK 14.19. Since $N(I_\lambda) \sim \lambda^{m-1}$,

$$\frac{1}{\lambda} \log \Pi_{I_\lambda}(\zeta, \bar{\zeta}) = \frac{1}{\lambda} \log \left(\sum_{j: \lambda_j \in I_\lambda} |\varphi^{\mathbb{C}}_{\lambda_j}(\zeta)|^2 \right) \tag{14.72}$$

$$= \max_{j: \lambda_j \in I_\lambda} \left\{ \frac{1}{\lambda} \log |\varphi^{\mathbb{C}}_{\lambda_j}(\zeta)|^2 \right\} + O(\frac{\log \lambda}{\lambda}). \tag{14.73}$$

Recall that a sequence of eigenfunctions is quantum ergodic if $\langle A\varphi_j, \varphi_j \rangle \to \frac{1}{\mu(S^*_g M)} \int_{S^*_g M} \sigma_A \, d\mu$. The complexified eigenfunctions then satisfy $\frac{1}{\lambda_j} \log |\varphi_j(\zeta)| \to \sqrt{\rho}(\zeta)$. It follows that ergodic eigenfunctions are asymptotically maximal, i.e., have the same logarithmic asymptotics as Φ^λ_M.

14.28. Pointwise phase space Weyl laws on Grauert tubes

14.28.1. Two term pointwise Weyl laws in Grauert tubes. The asymptotics of the complexified spectral projection kernels (14.26) are complex analogues of those of the diagonal spectral projections in the real domain and reflect the structure of complex geodesics from ζ to $\bar{\zeta}$. As in the real domain, one can obtain more refined asymptotics of $P_{[\lambda, \lambda+1]}(\zeta, \bar{\zeta})$ by using the structure of geodesic segments from ζ to $\bar{\zeta}$. This is the subject of the work in progress [**Z9**]. For the sake of completeness, we state the results here: There exists an explicit complex oscillatory factor $Q_\zeta(\lambda)$ depending on the geodesic arc from ζ to $\bar\zeta$ such that

(1) For $\sqrt{\rho}(\zeta) \geq \frac{C}{\lambda}$,

$$P^\tau_{[0,\lambda]}(\zeta,\bar\zeta) = (2\pi)^{-n}\lambda\left(\frac{\lambda}{\sqrt{\rho}}\right)^{\frac{n-1}{2}}\left(1+Q_\zeta(\lambda)\lambda^{-1}+o(\lambda^{-1})\right);$$

(2) For $\sqrt{\rho}(\zeta) \leq \frac{C}{\lambda}$,

$$P^\tau_{[0,\lambda]}(\zeta,\bar\zeta) = (2\pi)^{-n}\ \lambda^n + Q_\zeta(\lambda)\lambda^{n-1} + o(\lambda^{n-1}),$$

In this section, we prove Theorem 14.15 (1). To prove the local Weyl law we employ parametrices for the Poisson wave kernel adapted to $e^{i(t+i\tau)\sqrt{\Delta}}$ for $\tau > 0$ which are best adapted to the complex geometry.

Proof. As in the real domain, we obtain asymptotics of $P^\tau_{[0,\lambda]}(\zeta,\bar\zeta)$ by the Fourier-Tauberian method of relating their asymptotics to the singularities in the real time t of the Fourier transform (14.27). We refer to [**SV**] for background on Tauberian theorems. We follow the classical argument of [**DuG**, Proposition 2.1], , to obtain the local Weyl law with remainder one degree below the main term.

The proof is based on the oscillatory integral representation of Proposition 14.9. We are working in the case where $\sqrt{\rho}(\zeta) = \tau$ and hence can simplify it to (14.59).

We then introduce a cutoff function $\psi \in \mathcal{S}(\mathbb{R})$ with $\hat\psi \in C_0^\infty$ supported in sufficiently small neighborhood of 0 so that no other singularities of $U_{\mathbb{C}}(t+2i\tau,\zeta,\bar\zeta)$ lie in its support. We also assume $\hat\psi \equiv 1$ in a smaller neighborhood of 0. We then change variables $\theta \to \lambda\theta$ and apply the complex stationary phase to the integral

$$\text{(14.74)}\qquad \int_{\mathbb{R}} \hat\psi(t)e^{-i\lambda t}U_{\mathbb{C}}(t+2i\tau,\zeta,\bar\zeta)\,dt$$

$$= \int_{\mathbb{R}}\int_0^\infty \hat\psi(t)e^{-i\lambda t}e^{i\theta t}\left(\mathcal{B}(t,\zeta,\bar\zeta,\theta)\,d\theta + R(t+2i\tau,\zeta,\bar\zeta)\right)\,dt.$$

The second R term can be dropped since it is of order λ^{-M} for all $M > 0$. In the first we change variables $\theta \to \lambda\theta$ to obtain a semi-classical Fourier integral distribution of real type with phase $e^{i\lambda t(\theta-1)}$. The critical set consists of $\theta = 1, t = 0$. The phase is clearly non-degenerate with Hessian determinant one and inverse Hessian operator $D^2_{\theta,t}$. Taking into account the factor of λ^{-1} from the change of variables, the stationary phase expansion gives

$$\text{(14.75)}\qquad \sum_j \psi(\lambda-\lambda_j)e^{-2\tau\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2 \sim \sum_{k=0}^\infty \lambda^{\frac{n-1}{2}-k}\omega_k(\tau;\zeta)$$

where the coefficients $\omega_k(\tau,]\zeta)$ are smooth for $\zeta \in \partial M_\tau$. However the coefficients are not uniform as $\tau \to 0^+$ due to the factors of $(t+2i\tau+2i\sqrt{\rho}(\zeta))$ which were left in the denominators of the modified Hadamard parametrix. Since $t = 0$ at the stationary phase point, the resulting expansion is equivalent to one with the large parameter $\tau\lambda$ (or $\sqrt{\rho}(\zeta)\lambda$). The uniform expansion is then

$$\text{(14.76)}\qquad \sum_j \psi(\lambda-\lambda_j)e^{-2\tau\lambda_j}|\varphi_j^{\mathbb{C}}(\zeta)|^2 \sim \sum_{k=0}^\infty \left(\frac{\lambda}{\tau}\right)^{\frac{n-1}{2}-k}\omega_k(\zeta,\bar\zeta),$$

where ω_j are smooth in ζ, and $\omega_0 = 1$. The remainder has the same form.

To complete the proof, we apply the Fourier Tauberian theorem. Let $N \in F_+$ and let $\psi \in \mathcal{S}(\mathbb{R})$ satisfy the conditions: ψ is even, $\psi(\lambda) > 0$ for all $\lambda \in \mathbb{R}$, $\hat{\psi} \in C_0^\infty$, and $\hat{\psi}(0) = 1$. Then,

$$\psi * dN(\lambda) \le A\lambda^\nu \implies |N(\lambda) - N * \psi(\lambda)| \le CA\lambda^\nu,$$

where C is independent of A, λ. We apply it twice, first in the region $\sqrt{\rho}(\zeta) \ge C\lambda^{-1}$ and second in the complementary region.

In the first region, we let $N_{\tau,\zeta}(\lambda) = P_{\tau,\lambda}(\zeta, \bar{\zeta})$. It is clear that for $\sqrt{\rho} = \tau$, $N_{\tau,\zeta}(\lambda)$ is a monotone non-decreasing function of λ of polynomial growth which vanishes for $\lambda \le 0$. For $\psi \in \mathcal{S}$ positive, even and with $\hat{\psi} \in C_0^\infty(\mathbb{R})$ and $\hat{\psi}(0) = 1$, we have by (14.76) that

$$\psi * dN_{\tau,\zeta}(\lambda) \le C \left(\frac{\lambda}{\tau}\right)^{\frac{n-1}{2}}, \tag{14.77}$$

where C is independent of ζ, λ. It follows by the Fourier Tauberian theorem that

$$N_{\tau,\zeta}(\lambda) = N_{\tau,\zeta}(\lambda) * \psi(\lambda) + O\left(\frac{\lambda}{\tau}\right)^{\frac{n-1}{2}}.$$

Further, by integrating (14.76) from 0 to λ we have

$$N_{\tau,\zeta}(\lambda) * \psi(\lambda) = \left(\frac{\lambda}{\tau}\right)^{\frac{n-1}{2}} \left(\frac{\lambda}{\frac{n-1}{2} + 1} + O(1)\right),$$

proving (1).

To obtain uniform asymptotics in τ down to $\tau = 0$, we use instead the analytic continuation of the Hörmander parametrix (14.19). We choose local coordinates near x and write $\exp_x^{-1}(y) = \Psi(x, y)$ in these local coordinates for y near x, and write the integral T_y^*M as an integral over $\mathbb{R}^m$ in these coordinates. The holomorphic extension of the parametrix to the Grauert tube $|\zeta| < \tau$ at time $t + 2i\tau$ has the form (14.20)–(14.59), i.e.,

$$U_{\mathbb{C}}(t + 2i\tau, \zeta, \bar{\zeta}) = \int_{\mathbb{R}^n} e^{(it-2\tau)|\xi|_{g_y}} e^{i\langle \xi, \Psi(\zeta,\bar{\zeta})\rangle} A(t, \zeta, \bar{\zeta}, \xi)\, d\xi. \tag{14.78}$$

Again, we use a cutoff function $\psi \in \mathcal{S}(\mathbb{R})$ with $\hat{\psi} \in C_0^\infty$ supported in sufficiently small neighborhood of 0 so that no other singularities of $E(t + 2i\tau, \zeta, \bar{\zeta})$ lie in its support and so that $\hat{\psi} \equiv 1$ in a smaller neighborhood of 0. We write the integral in polar coordinates and obtain

$$\begin{aligned} &\int_{\mathbb{R}} \hat{\psi}(t) e^{-i\lambda t} U_{\mathbb{C}}(t + 2i\tau, \zeta, \bar{\zeta})\, dt \\ &\quad = \lambda^m \int_0^\infty \int_{\mathbb{R}} \hat{\psi}(t) e^{-i\lambda t} \int_{S^{n-1}} e^{(it-2\tau)\lambda r} e^{ir\lambda\langle \omega, \Psi(\zeta,\bar{\zeta})\rangle} A(t, \zeta, \bar{\zeta}, \lambda r\omega) r^{n-1}\, dr d\omega. \end{aligned} \tag{14.79}$$

We then apply complex stationary phase to the $drdt$ integral, regarding

$$\int_{S^{n-1}} e^{ir\lambda\langle \omega, \Psi(\zeta,\bar{\zeta})\rangle} A(t, \zeta, \bar{\zeta}, \lambda r\omega) r^{m-1}\, d\omega$$

as the amplitude. When $\sqrt{\rho}(\zeta) \leq \frac{C}{\lambda}$ the exponent is bounded in λ and the integral defines a symbol. Applying stationary phase again to the $dtd\theta$ integral now gives

$$\sum_j \psi(\lambda - \lambda_j) e^{-2\tau\lambda_j} |\varphi_j^{\mathbb{C}}(\zeta)|^2 \sim \sum_{k=0}^{\infty} \lambda^{n-1-k} \omega_k(\zeta, \bar{\zeta}), \tag{14.80}$$

where $\omega_k(\zeta, \bar{\zeta})$ is smooth down to the zero section.

We apply the Fourier Tauberian theorem again, but this time with the estimates

$$\psi * dN_{\tau,\zeta}(\lambda) \leq C\lambda^{n-1},$$

where C is independent of ζ. We conclude that

$$N_{\tau,\zeta}(\lambda) = C\lambda^n + O(\lambda^{n-1}),$$

proving (2). □

COROLLARY 14.20. *For all $\zeta \in M_{\mathbb{C}}$, and with $\tau = \sqrt{\rho}(\zeta)$,*

$$c\lambda^{\frac{n+1}{2}} \leq P^{\tau}_{[0,\lambda]}(\zeta, \bar{\zeta}) \leq C\lambda^n.$$

14.29. Proof of Corollary 14.16

For the upper bound, we use that

$$\sup_{\zeta \in \partial M_\tau} |\varphi_\lambda^{\mathbb{C}}(\zeta)|^2 \leq \sup_{\zeta \in \partial M_\tau} |\Pi_{I_\lambda}(\zeta, \bar{\zeta})| \leq \sup_{\zeta \in \partial M_\tau} e^{\lambda\sqrt{\rho}(\zeta)} |P_{I_\lambda}(\zeta)|.$$

The upper bound stated in Corollary 14.16 then follows from Corollary 14.20 to Theorem 14.15.

For the lower bound in (2) of Corollary 14.16, we use that

$$\|\varphi_j^{\mathbb{C}}\|_{L^2(\partial M_\tau)} = e^{2\tau_j} \langle U(i\tau)^* U(i\tau)\varphi_j, \varphi_j\rangle_{L^2(M)}.$$

By Lemma 14.10, the operator $U(i\tau)^*U(i\tau)$ is an elliptic pseudodifferential operator of order $\mu = -\frac{n-1}{2}$ (or so). Let $C > 0$ be a lower bound for its symbol times $\langle\xi\rangle^\mu$. Then by Gårding's inequality,

$$\langle U(i\tau)^* U(i\tau)\varphi_j, \varphi_j\rangle_{L^2(M)} \geq C\lambda_j^{-\mu},$$

and so

$$\|\varphi_j^{\mathbb{C}}\|_{L^2(\partial M_\tau)} \geq C\lambda_j^{-\mu} e^{2\tau\lambda_j}. \tag{14.81}$$

□

14.30. Complex nodal sets and sequences of logarithms

We regard the zero set $[Z_f]$ as a *current of integration*, i.e., as a linear functional on $(m-1, m-1)$ forms ψ

$$\langle [Z_{\varphi_j}], \psi\rangle = \int_{Z_{\varphi_j}} \psi.$$

Recall that a *current* is a linear functional (distribution) on smooth forms. We refer to [**GH**] for background. On a complex manifold one has (p,q) forms with p dz_j and q $d\bar{z}_k$'s. In (14.121) we use the Kähler hypersurface volume form ω_g^{m-1} (where $\omega_g = i\partial\bar{\partial}\rho$) to make Z_{φ_j} into a measure:

$$\langle [Z_{\varphi_j}], f\rangle = \int_{Z_{\varphi_j}} f\omega_g^{m-1}, \qquad f \in C(M).$$

14.30.1. Poincaré-Lelong formula. One of the two key reasons for the gain in simplicity is that there exists a simple analytical formula for the delta-function on the nodal set. The *Poincaré-Lelong formula* gives an exact formula for the delta-function on the zero set of φ_j

$$\frac{i}{2\pi}\partial\bar{\partial}\log|\varphi_j^{\mathbb{C}}(z)|^2 = [\mathcal{N}_{\varphi_j^{\mathbb{C}}}]. \tag{14.82}$$

Thus, if ψ is an $(n-1, n-1)$ form, then

$$\int_{\mathcal{N}_{\varphi_j^{\mathbb{C}}}} \psi = \frac{1}{2\pi}\int_{M_\varepsilon} \psi \wedge i\partial\bar{\partial}\log|\varphi_j^{\mathbb{C}}(z)|^2.$$

14.30.2. Sequences of PSH functions and a weak* limit problem. We next consider logarithms of Husimi functions, which are PSH functions on M_ε. A function f on a domain in a complex manifold is PSH if $i\partial\bar{\partial}f$ is a *positive (1,1) current.* That is, $i\partial\bar{\partial}f$ is a singular form of type $\sum_{i\bar{j}} a_{i\bar{j}} dz^i \wedge d\bar{z}^k$ with $(a_{j\bar{k}})$ positive definite Hermitian. If f is a local holomorphic function, then $\log|f(z)|$ is PSH and $i\partial\bar{\partial}\log|f(z)| = [Z_f]$. General references are [**GH, HoC**].

A sequence of $(1,1)$ currents E_k converges weak* to a current E if $\langle E_k, \psi\rangle \to \langle E, \psi\rangle$ for all smooth $(m-1, m-1)$ forms. Thus, for all f

$$[Z_{\varphi_j}] \to i\partial\bar{\partial}\sqrt{\rho} \iff \int_{Z_{\varphi_j}} f\omega^{m-1} \to i\int_{M_\varepsilon} f\partial\bar{\partial}\sqrt{\rho}\wedge\omega^{m-1,m-1}.$$

14.30.3. Pluri-subharmonic functions and compactness. In the real domain, we have emphasized the problem of finding quantum limits (or microlocal defect measures). The same problem exists in the complex domain for the sequence of Husimi functions (14.2). However, there also exists a new problem involving the sequence of normalized logarithms

$$\{u_j := \frac{1}{\lambda_j}\log|\varphi_j^{\mathbb{C}}(z)|^2\}_{j=1}^{\infty}. \tag{14.83}$$

A key fact is that this sequence is pre-compact in $L^p(M_\varepsilon)$ for all $p < \infty$ and even that

$$\left\{\frac{1}{\lambda_j}\nabla\log|\varphi_j^{\mathbb{C}}(z)|^2\right\}_{j=1}^{\infty}. \tag{14.84}$$

is pre-compact in $L^1(M_\varepsilon)$.

LEMMA 14.21 (Hartog's Lemma; see Theorem 4.1.9 of [**HoI**]). *Let $\{v_j\}$ be a sequence of subharmonic functions in an open set $X \subset \mathbb{R}^m$ which have a uniform upper bound on any compact set. Then either $v_j \to -\infty$ uniformly on every compact set, or else there exists a subsequence v_{j_k} which is convergent to some $u \in L^1_{loc}(X)$. Further, $\limsup_n u_n(x) \le u(x)$ with equality almost everywhere. For every compact subset $K \subset X$ and every continuous function f,*

$$\limsup_{n\to\infty}\sup_K(u_n - f) \le \sup_K(u - f).$$

In particular, if $f \ge u$ and $\varepsilon > 0$, then $u_n \le f + \varepsilon$ on K for n large enough.

14.30.4. A general weak* limit problem for logarithms of Husimi functions. The study of exponential growth rates gives rise to a new kind new weak* limit problem for complexified eigenfunctions.

PROBLEM 14.22. Find the weak* (in fact, L^1) limits G on M_ε of sequences

$$\frac{1}{\lambda_{j_k}} \log |\varphi_{j_k}^{\mathbb{C}}(z)|^2 \to G.$$

See Theorem 14.44, 14.45 and 14.47 for the solution to this problem (modulo sparse subsequences) in the ergodic case.

Here is a general Heuristic principle to pin down the possible G: If $\frac{1}{\lambda_{j_k}} \log |\varphi_{j_k}^{\mathbb{C}}(z)|^2 \to G(z)$ then

$$|\varphi_{j_k}^{\mathbb{C}}(z)|^2 \simeq e^{\lambda_{j_k} G(z)}(1 + o(1)) \qquad \lambda_j \to \infty.$$

But $\Delta_{\mathbb{C}} |\varphi_{j_k}^{\mathbb{C}}(z)|^2 = \lambda_{j_k}^2 |\varphi_{j_k}^{\mathbb{C}}(z)|^2$, so we should have

CONJECTURE 14.23. Any limit G as above solves the Hamilton-Jacobi equation $(\nabla_{\mathbb{C}} G)^2 = 1$.

REMARK 14.24. The weak* limits of $\frac{\|\varphi_j^{\mathbb{C}}(z)\|^2}{\|\varphi_j^{\mathbb{C}}\|_{L^2(\partial M_\varepsilon)}} d\mu_\varepsilon$ must be supported in $\{G = G_{\max}\}$ (i.e., in the set of maximum values).

14.31. Real zeros and complex analysis

A natural but rather intractable problem to obtain the distribution of real zeros from knowledge of the complex nodal distribution. There exist few if any general results on this problem. In the next section we explain how to get upper bounds on real zeros using complex zeros.

It is possible to obtain results on complex zeros which are within λ^{-1} of the real domain by rescaling the nodal set by a factor of λ^{-1} in M_τ. But we cannot distinguish such 'almost real zeros' from real zeros.

It would be interesting to understand (at least in real dimension 2) how the complex nodal set 'sprouts' from the real nodal set. How do the connected components of the real nodal set fit together in the complex nodal set?

14.31.1. Proof of the Donnelly-Fefferman upper bound. To prove Theorem 14.1, we use Crofton's formula and a multi-dimensional Jensen's formula to give an upper bound for $\mathcal{H}^{n-2}(\mathcal{N}_\lambda)$ in terms of the integral geometry of $\mathcal{N}_\lambda^{\mathbb{C}}$. The integral geometric approach to the upper bound is inspired by the classic paper of Donnelly-Fefferman [**DuG**] (see also [**Lin**]). But, instead of doubling estimates or frequency function estimates, we use the Poisson wave kernel to obtain growth estimates on eigenfunctions, and then use results on pluri-subharmonic functions rather than functions of one complex variable to relate growth of zeros to growth of eigenfunctions. This approach was used in [**Z2**] to prove equidistribution theorems for complex nodal sets when the geodesic flow is ergodic. The Poisson wave kernel approach works for Steklov eigenfunctions as well as Laplace eigenfunctions, and in fact for eigenfunctions of any positive elliptic analytic pseudo-differential operator.

We first use the Poisson wave group (14.28) to analytically continue eigenfunctions in the form (14.16),

$$U_{\mathbb{C}}(i\tau)\psi_j(\zeta) = e^{-\tau\lambda_j}\psi_j^{\mathbb{C}}(\zeta). \tag{14.85}$$

We then use (14.85) to determine the growth properties of $\psi_j^{\mathbb{C}}(\zeta)$ in Grauert tubes of the complexification of $\partial\Omega$. The relevant notion of Grauert tube is the standard Grauert tube for $\partial\Omega$ with the metric $g_{\partial\Omega}$ induced by the ambient metric g on M. This is because the principal symbol of Λ is the same as the principal symbol of $\sqrt{\Delta_{\partial\Omega}}$.

14.31.2. Proof of Theorem 14.1. We start with the integral geometric approach of [**DF**, Lemma 6.3] (see also [**Lin**, (3.21)]). There exists a "Crofton formula" in the real domain which bounds the local nodal hypersurface volume above:

$$\mathcal{H}^{m-1}(\mathcal{N}_{\varphi_\lambda} \cap U) \leq C_L \int_{\mathcal{L}} \#\{\mathcal{N}_{\varphi_\lambda} \cap \ell\} d\mu(\ell). \tag{14.86}$$

Thus, $\mathcal{H}^{m-1}(\mathcal{N}_{\varphi_\lambda} \cap U)$ is bounded above by a constant C_L times the average over all line segments of length L in a local coordinate patch U of the number of intersection points of the line with the nodal hypersurface. The measure $d\mu_L$ is known as the 'kinematic measure' in the Euclidean setting [**F1**, Chapter 3]; see also of [**AP1**, Theorem 5.5]. We will be using geodesic segments of fixed length L rather than line segments, and parametrize them by $S^*M \times [0, L]$, i.e., by their initial data and time. Then $d\mu_\ell$ is essentially Liouville measure $d\mu_L$ on S^*M times dt.

The complexification of a real line $\ell = x + \mathbb{R}v$ with $x, v \in \mathbb{R}^m$ is $\ell_{\mathbb{C}} = x + \mathbb{C}v$. Since the number of intersection points (or zeros) only increases if we count complex intersections, we have

$$\int_{\mathcal{L}} \#(\mathcal{N}_{\varphi_\lambda} \cap \ell)\, d\mu(\ell) \leq \int_{\mathcal{L}} \#(\mathcal{N}_{\varphi_\lambda}^{\mathbb{C}} \cap \ell_{\mathbb{C}})\, d\mu(\ell). \tag{14.87}$$

Note that this complexification is quite different from using intersections with all complex lines to measure complex nodal volumes. If we did that, we would obtain a similar upper bound on the complex hypersurface volume of the complex nodal set. But it would not give an upper bound on the real nodal volume and indeed would the complex volume tends to zero as one shrinks the Grauert tube radius to zero, while (14.87) stays bounded below.

Hence to prove Theorem 14.1 it suffices to show

LEMMA 14.25. *We have,*

$$\mathcal{H}^{m-1}(\mathcal{N}_{\varphi_\lambda}) \leq C_L \int_{\mathcal{L}} \#(\mathcal{N}_{\varphi_\lambda})^{\mathbb{C}} \cap \ell_{\mathbb{C}})\, d\mu(\ell) \leq C\lambda.$$

We now sketch the proofs of these results using a somewhat novel approach to the integral geometry and complex analysis.

14.32. Background on hypersurfaces and geodesics

The proof of the Crofton formula given below in Lemma 14.29 involves the geometry of geodesics and hypersurfaces. To prepare for it we provide the relevant background.

As above, we denote by $d\mu_L$ the Liouville measure on S^*M. We also denote by ω the standard symplectic form on T^*M and by α the canonical one form. Then $d\mu_L = \omega^{n-1} \wedge \alpha$ on S^*M. Indeed, $d\mu_L$ is characterized by the formula $d\mu_L \wedge dH = \omega^m$, where $H(x, \xi) = |\xi|_g$. So it suffices to verify that $\alpha \wedge dH = \omega$ on S^*M. We take the interior product ι_{Ξ_H} with the Hamilton vector field Ξ_H on both sides, and the identity follows from the fact that $\alpha(\Xi_H) = \sum_j \xi_j \frac{\partial H}{\partial \xi_j} = H = 1$

on S^*M, since H is homogeneous of degree one. Henceforth we denote by $\Xi = \Xi_H$ the generator of the geodesic flow.

Let $N \subset M$ be a smooth hypersurface in a Riemannian manifold (M, g). We denote by $T^*_N M$ the of covectors with footpoint on N and $S^*_N M$ the unit covectors along N. We introduce Fermi normal coordinates (s, y_n) along N, where s are coordinates on N and y_n is the normal coordinate, so that $y_m = 0$ is a local defining function for N. We also let σ, ξ_m be the dual symplectic Darboux coordinates. Thus the canonical symplectic form is $\omega_{T^*M} = ds \wedge d\sigma + dy_m \wedge d\xi_m$. Let $\pi : T^*M \to M$ be the natural projection. For notational simplicity we denote $\pi^* y_m$ by y_m as functions on T^*M. Then y_m is a defining function of $T^*_N M$.

The hypersurface $S^*_N M \subset S^*M$ is a kind of Poincaré section or symplectic transversal to the orbits of G^t, i.e. is a symplectic transversal away from the (at most codimension one) set of $(y, \eta) \in S^*_N M$ for which $\Xi_{y,\eta} \in T_{y,\eta} S^*_N M$, where as above Ξ is the generator of the geodesic flow. More precisely,

Lemma 14.26. *The restriction $\omega|_{S^*_N M}$ is symplectic on $S^*_N M \backslash S^*N$.*

Indeed, $\omega|_{S^*_N M}$ is symplectic on $T_{y,\eta} S^*N$ as long as $T_{y,\eta} S^*_N M$ is transverse to $\Xi_{y,\eta}$, since $\ker(\omega|_{S^*M}) = \mathbb{R}\Xi$. But S^*N is the set of points of $S^*_N M$ where $\Xi \in T S^*_N M$, i.e. where $S^*_N M$ fails to be transverse to G^t. Indeed, transversality fails when $\Xi(y_m) = dy_m(\Xi) = 0$, and $\ker dy_m \cap \ker dH = T S^*_N M$. One may also see it in Riemannian terms as follows: the generator $\Xi_{y,\eta}$ is the horizontal lift η^h of η to (y, η) with respect to the Riemannian connection on S^*M, where we freely identify covectors and vectors by the metric. Lack of transversality occurs when η^h is tangent to $T_{(y,\eta)}(S^*_N M)$. The latter is the kernel of dy_n. But $dy_m(\eta^h) = dy_m(\eta) = 0$ if and only if $\eta \in TN$.

It follows from Lemma 14.26 that the symplectic volume form of $S^*_N M \backslash S^*N$ is $\omega^{n-1}|_{S^*_N M}$. The following Lemma gives a useful alternative formula:

Lemma 14.27. *Define*

$$d\mu_{L,N} = \iota_{\Xi} d\mu_L \;|_{S^*_N M},$$

where as above, $d\mu_L$ is Liouville measure on S^*M. Then

$$d\mu_{L,N} = \omega^{m-1}|_{S^*_N M}.$$

Indeed, $d\mu_L = \omega^{m-1} \wedge \alpha$, and $\iota_{\Xi} d\mu_L = \omega^{m-1}$.

Corollary 14.28. $\mathcal{H}^{m-1}(N) = \frac{1}{\beta_m} \int_{S^*_N M} |\omega^{m-1}|$.

14.32.1. Hausdorff measure and Crofton formula for real geodesic arcs. First we sketch a proof of the integral geometry estimate using geodesic arcs rather than local coordinate line segments. For background on integral geometry and Crofton type formulae we refer to [**AP1, AP2**]. As explained there, a Crofton formula arises from a double fibration

$$\begin{array}{ccc} & \mathcal{I} & \\ {}^{\pi_1}\swarrow & & \searrow^{\pi_2} \\ \Gamma & & B \end{array}$$

where Γ parametrizes a family of submanifolds B_γ of B. The points $b \in B$ then parametrize a family of submanifolds $\Gamma_b = \{\gamma \in \Gamma : b \in B_\gamma\}$ and the top space is the incidence relation in $B \times \Gamma$ that $b \in B_\gamma$.

We would like to define Γ as the space of geodesics of (M, g), i.e. the space of orbits of the geodesic flow on S^*M. Heuristically, the space of geodesics is the quotient space $S^*M/\mathbb{R}$ where $\mathbb{R}$ acts by the geodesic flow G^t (i.e. the Hamiltonian flow of H). Of course, for a general (i.e., non-Zoll) (M, g) the 'space of geodesics' is not a Hausdorff space and so we do not have a simple analogue of the space of lines in $\mathbb{R}^n$. Instead we consider the space $\mathcal{G}_T$ of geodesic arcs of length T. If we only use partial orbits of length T, no two partial orbits are equivalent and the space of geodesic arcs $\gamma^T_{x,\xi}$ of length T is simply parametrized by S^*M. Hence we let $B = S^*M$ and also $\mathcal{G}_T \simeq S^*M$. The fact that different arcs of length T of the same geodesic are distinguished leads to some redundancy.

In the following, let L_1 denote the length of the shortest closed geodesic of (M, g).

Proposition 14.29. *Let $N \subset M$ be any smooth hypersurface*[1]*, and let $S^*_N M$ denote the unit covers to M with footpoint on N. Then for $0 < T < L_1$,*

$$\mathcal{H}^{m-1}(N) = \frac{1}{\beta_m T} \int_{S^*M} \#\{t \in [-T, T] : G^t(x, \omega) \in S^*_N M\} \, d\mu_L(x, \omega),$$

where β_m is $2(m-1)!$ times the volume of the unit ball in $\mathbb{R}^{m-2}$.

Proof. By Corollary 14.28, the Hausdorff measure of N is given by

$$\mathcal{H}^{m-1}(N) = \frac{1}{\beta_m} \int_{S^*_N M} |\omega^{m-1}|. \tag{14.88}$$

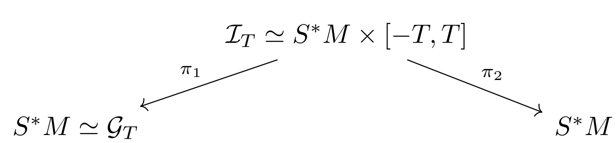
We use the Lagrange (or more accurately, Legendre) immersion

$$\iota \colon S^*M \times \mathbb{R} \to S^*M \times S^*M, \quad \iota(x, \omega, t) = (x, \omega, G^t(x, \omega)).$$

We also let $\pi \colon T^*M \to M$ be the standard projection. We restrict ι to $S^*M \times [-T, T]$ and define the incidence relation

$$\mathcal{I}_T = \{((y, \eta), (x, \omega), t) \subset S^*M \times S^*M \times [-T, T] \colon (y, \eta) = G^t(x, \omega)\},$$

which is isomorphic to $[-T, T] \times S^*M$ under ι. We form the diagram



using the two natural projections, which in the local parametrization take the form

$$\pi_1(t, x, \xi) = G^t(x, \xi), \quad \pi_2(t, x, \xi) = (x, \xi).$$

As noted above, the bottom left S^*M should be thought of as the space of geodesic arcs. The fiber

$$\pi_1^{-1}(y, \eta) = \{(t, x, \xi) \in [-T, T] \times S^*M : G^t(x, \xi) = (y, \eta)\} \simeq \gamma^T_{(y,\eta)}$$

[1] The same formula is true if N has a singular set Σ with $\mathcal{H}^{m-2}(\Sigma) < \infty$

may be identified with the geodesic segment through (y,η) and the fiber $\pi_2^{-1}(x,\omega) \simeq [-T,T]$.

We 'restrict' the diagram above to $S^*_N M$:

$$\begin{array}{ccc} & \mathcal{I}_T \simeq S^*_N M \times [-T,T] & \\ \swarrow^{\pi_1} & & \searrow^{\pi_2} \\ (S^*_N M)_T & & S^*_N M \end{array}$$

where
$$(S^*_N M)_T = \pi_1 \pi_2^{-1}(S^*_N M) = \bigcup_{|t|<T} G^t(S^*_N M).$$

We define the Crofton density φ_T on $S^*_N M$ corresponding to the diagram (14.32.1) [**AP1**, §4] by

$$\varphi_T = (\pi_2)_* \pi_1^* \, d\mu_L. \tag{14.89}$$

Since the fibers of π_2 are 1-dimensional, φ_T is a differential form of dimension $2\dim M - 2$ on S^*M. To make it smoother, we can introduce a smooth cutoff χ to $(-1,1)$, equal to 1 on $(-\frac{1}{2},\frac{1}{2})$, and use $\chi_T(t) = \chi(\frac{t}{T})$. Then $\pi_1^*(d\mu_L \otimes \chi_T \, dt)$ is a smooth density on $\mathcal{I}_T$.

Lemma 14.30. *The Crofton density* (14.89) *is given by* $\varphi_T = T d\mu_{L,N}$.

Proof. In (14.32.1) we defined the map $\pi_1 \colon (y,\eta,t) \in S^*_N M \times [-T,T] \to G^t(y,\eta) \in (S^*M)_\varepsilon$. We first claim that $\pi_1^* d\mu_L = d\mu_{L,N} \otimes dt$. This is essentially the same as Lemma 14.27. Indeed, $d\pi_1(\frac{\partial}{\partial t}) = \Xi$, hence $\iota_{\frac{\partial}{\partial t}} \pi_1^* d\mu_L|_{(t,y,\eta)} = (G^t)^* \omega^{m-1} = \omega^{m-1}|_{T_{y,\eta} S^*_N M}$.

Combining Lemma 14.30 with (14.88) gives

$$\int_{S^*_N M} \varphi_T = \int_{\pi_2^{-1}(S^*_N M)} d\mu_L = T \beta_m \mathcal{H}^{m-1}(N). \tag{14.90}$$

□

We then relate the integral on the left side to numbers of intersections of geodesic arcs with N. The relation is given by the co-area formula: if $f : X \to Y$ is a smooth map of manifolds of the same dimension and if Φ is a smooth density on Y, and if $\#\{f^{-1}(y)\} < \infty$ for every regular value y, then

$$\int_X f^* \Phi = \int_Y \#\{f^{-1}(y)\} \, \Phi.$$

If we set set $X = \pi_2^{-1}(S^*_N M)$, $Y = S^*M$, and $f = \pi_1|_{\pi_2^{-1}(S^*_N M)}$ then the co-area formula gives,

$$\int_{\pi_2^{-1}(S^*_N M)} \pi_1^* d\mu_L = \int_{S^*M} \#\{t \in [-T,T] : G^t(x,\omega) \in S^*_N M\} d\mu_L(x,\omega). \tag{14.91}$$

Combining (14.90) and (14.91) gives the result stated in Proposition 14.29

$$T \beta_m \mathcal{H}^{m-1}(N) = \int_{S^*M} \#\{t \in [-T,T] : G^t(x,\omega) \in S^*_N M\} \, d\mu_L(x,\omega). \tag{14.92}$$

□

14.32.2. Proof of Lemma 14.25. The next step is to complexify.

Proof. We complexify the Lagrange immersion ι from a line (segment) to a strip in $\mathbb{C}$: Define

$$F\colon S_\varepsilon \times S^*M \to M_{\mathbb{C}}, \quad F(t+i\tau, x, v) = \exp_x(t+i\tau)v, \quad (|\tau| \leq \varepsilon)$$

By definition of the Grauert tube, ψ is surjective onto M_ε. For each $(x,v) \in S^*M$,

$$F_{x,v}(t+i\tau) = \exp_x(t+i\tau)v$$

is a holomorphic strip. Here, $S_\varepsilon = \{t+i\tau \in \mathbb{C} : |\tau| \leq \varepsilon\}$. We also denote by $S_{\varepsilon,L} = \{t+i\tau \in \mathbb{C} : |\tau| \leq \varepsilon, |t| \leq L\}$.

Since $F_{x,v}$ is a holomorphic strip,

$$F_{x,v}^*\left(\frac{1}{\lambda} dd^c \log |\psi_j^{\mathbb{C}}|^2\right) = \frac{1}{\lambda} dd^c_{t+i\tau} \log |\psi_j^{\mathbb{C}}|^2 (\exp_x(t+i\tau)v) \tag{14.93}$$

$$= \frac{1}{\lambda} \sum_{t+i\tau : \psi_j^{\mathbb{C}}(\exp_x(t+i\tau)v)=0} \delta_{t+i\tau}. \tag{14.94}$$

Put

$$\mathcal{A}_{L,\varepsilon}\left(\frac{1}{\lambda} dd^c \log |\psi_j^{\mathbb{C}}|^2\right) = \frac{1}{\lambda} \int_{S^*M} \int_{S_{\varepsilon,L}} dd^c_{t+i\tau} \log |\psi_j^{\mathbb{C}}|^2 (\exp_x(t+i\tau)v)\, d\mu_L(x,v). \tag{14.95}$$

A key observation of [**DF, Lin**] is that

$$\#\{\mathcal{N}_\lambda^{\mathbb{C}} \cap F_{x,v}(S_{\varepsilon,L})\} \geq \#\{\mathcal{N}_\lambda^{\mathbb{R}} \cap F_{x,v}(S_{0,L})\}, \tag{14.96}$$

since every real zero is a complex zero. It follows then from Proposition 14.29 (with $N = \mathcal{N}_\lambda$) that

$$\mathcal{A}_{L,\varepsilon}\left(\frac{1}{\lambda} dd^c \log |\psi_j^{\mathbb{C}}|^2\right) = \frac{1}{\lambda} \int_{S^*M} \#\{\mathcal{N}_\lambda^{\mathbb{C}} \cap F_{x,v}(S_{\varepsilon,L})\}\, d\mu(x,v) \geq \frac{1}{\lambda} \mathcal{H}^{m-1}(\mathcal{N}_{\psi_\lambda}).$$

Hence to obtain an upper bound on $\frac{1}{\lambda}\mathcal{H}^{m-1}(\mathcal{N}_{\psi_\lambda})$ it suffices to prove that there exists $M < \infty$ so that

$$\mathcal{A}_{L,\varepsilon}(\frac{1}{\lambda} dd^c \log |\psi_j^{\mathbb{C}}|^2) \leq M. \tag{14.97}$$

To prove (14.97), we observe that since $dd^c_{t+i\tau} \log |\psi_j^{\mathbb{C}}|^2 (\exp_x(t+i\tau)v)$ is a positive $(1,1)$ form on the strip, the integral over S_ε is only increased if we integrate against a positive smooth test function $\chi_\varepsilon \in C_c^\infty(\mathbb{C})$ which equals one on $S_{\varepsilon,L}$ and vanishes off $S_{2\varepsilon,L}$. Integrating by parts the dd^c onto χ_ε, we have

$$\mathcal{A}_{L,\varepsilon}\left(\frac{1}{\lambda} dd^c \log |\psi_j^{\mathbb{C}}|^2\right) \leq \frac{1}{\lambda} \int_{S^*M} \int_{\mathbb{C}} dd^c_{t+i\tau} \log |\psi_j^{\mathbb{C}}|^2 (\exp_x(t+i\tau)v) \tag{14.98}$$

$$\times\, \chi_\varepsilon(t+i\tau)\, d\mu_L(x,v) \tag{14.99}$$

$$= \frac{1}{\lambda} \int_{S^*M} \int_{\mathbb{C}} \log |\psi_j^{\mathbb{C}}|^2 (\exp_x(t+i\tau)v) \tag{14.100}$$

$$\times\, dd^c_{t+i\tau} \chi_\varepsilon(t+i\tau)\, d\mu_L(x,v). \tag{14.101}$$

Now write $\log|x| = \log_+|x| - \log_-|x|$. Here $\log_+|x| = \max\{0, \log|x|\}$ and $\log_| x| = \max\{0, -\log|x|\}$. Then we need upper bounds for

$$\frac{1}{\lambda}\int_{S^*M}\int_{\mathbb{C}} \log_\pm|\psi_j^{\mathbb{C}}|^2(\exp_x(t+i\tau)v) dd^c_{t+i\tau}\chi_\varepsilon(t+i\tau)\, d\mu_L(x,v).$$

For $\log_+$ the upper bound is an immediate consequence of Proposition 14.7. For $\log_-$ the bound is subtler: we need to show that $|\varphi_\lambda(z)|$ cannot be too small on too large a set. As we know from Gaussian beams, it is possible that $|\varphi_\lambda(x)| \leq Ce^{-\delta\lambda}$ on sets of almost full measure in the real domain; we need to show that nothing worse can happen.

The map (14.7) is a diffeomorphism and since $B^*_\varepsilon M = \bigcup_{0\leq\tau\leq\varepsilon} S^*_\tau M$ we also have that

$$E\colon S_{\varepsilon,L}\times S^*M \to M_\tau, \quad E(t+i\tau, x, v) = \exp_x(t+i\tau)v$$

is a diffeomorphism for each fixed t. Hence by letting t vary, E is a smooth fibration with fibers given by geodesic arcs. Over a point $\zeta \in M_\tau$ the fiber of the map is a geodesic arc

$$\{(t+i\tau, x, v) : \exp_x(t+i\tau)v = \zeta, \quad \tau = \sqrt{\rho}(\zeta)\}.$$

Pushing forward the measure $dd^c_{t+i\tau}\chi_\varepsilon(t+i\tau)d\mu_L(x,v)$ under E gives a positive measure $d\mu$ on M_τ. We claim that

$$\mu := E_*\, dd^c_{t+i\tau}\chi_\varepsilon(t+i\tau)d\mu_L(x,v) = \left(\int_{\gamma_{x,v}} \Delta_{t+i\tau}\chi_\varepsilon\, ds\right) dV_\omega, \tag{14.102}$$

where dV_ω is the Kähler volume form $\frac{\omega^m}{m!}$.

In fact, $d\mu_L$ is equivalent under E to the contact volume form $\alpha\wedge\omega_\rho^{m-1}$ where $\alpha = d^c\sqrt{\rho}$. Hence the claim amounts to saying that the Kähler volume form is $d\tau$ times the contact volume form. In particular it is a smooth (and of course signed) multiple J of the Kähler volume form dV_ω, and we do not need to know the coefficient function J beyond that it is bounded above and below by constants independent of λ. We then have

(14.103)
$$\int_{S^*M}\int_{\mathbb{C}} \log|\psi_j^{\mathbb{C}}|^2(\exp_x(t+i\tau)v)dd^c_{t+i\tau}\chi_\varepsilon(t+i\tau)\, d\mu_L(x,v) = \int_{M_\tau} \log|\psi_j^{\mathbb{C}}|^2\, JdV.$$

To complete the proof of (14.97) it suffices to prove that the right side is $\geq -C\lambda$ for some $C > 0$.

We use the well-known Lemma 14.21. This Lemma implies the desired lower bound on (14.103): there exists $C > 0$ so that

$$\frac{1}{\lambda}\int_{M_\tau} \log|\psi_\lambda|\, JdV \geq -C. \tag{14.104}$$

For if not, there exists a subsequence of eigenvalues λ_{j_k} so that $\frac{1}{\lambda_{j_k}}\int_{M_\tau}\log|\psi_{\lambda_{j_k}}|JdV \to -\infty$. By Proposition 14.7, $\{\frac{1}{\lambda_{j_k}}\log|\psi_{\lambda_{j_k}}|\}$ has a uniform upper bound. Moreover the sequence does not tend uniformly to $-\infty$ since $\|\psi_\lambda\|_{L^2(M)} = 1$. It follows that a further subsequence tends in L^1 to a limit u and by the dominated convergence theorem the limit of (14.104) along the sequence equals $\int_{M_\tau} u\, JdV \neq -\infty$. This contradiction concludes the proof of (14.104), hence (14.97), and thus the theorem. □

14.33. Proof of the Donnelly-Fefferman lower bound (A. Brudnyi)

The proof of the lower bound on nodal volumes in [**DF**] involves complex analytic arguments but is not purely complex analytic. The role of complex analysis is subtler than in the upper bound because there is no general lower bound in algebraic geometry, i.e. no reason why a polynomial should have real zeros. The fact that eigenfunctions have many real zeros is due to the fundamental existence theorem that the real nodal set is $\frac{1}{\lambda}$ dense in the real manifold M. This existence theorem does not use analytic extensions. But from the existence theorem one may employ complex analytic arguments to obtain lower bounds. The purpose of this section is to give A. Brudnyi's proof of the lower bound, which starts from some theorems in [**DF**] but brings out the complex analytic component in a simpler way. It is published here for the first time as a response to the author's question whether the Donnelly-Fefferman lower bound could be based on a higher dimensional Cartan Lemma.

14.33.1. Results from Donnelly-Fefferman. We begin with some background results from [**DF**], mainly from Sections 5 and 7.

The argument is local. To prove the lower bound, it suffices to prove it in one coordinate chart U. By the existence theorem, there exists $a_1 > 0$ so that every ball or cube of radius $\frac{a_1}{\lambda}$) contains a zero. Cover U by cubes Q_ν of side $a_2\lambda^{-1}$ with $a_2 > a_1$ so that x_ν lies in the middle tenth of Q_ν. Choose a_3 so that $B_\nu = B(x_\nu, a_3\lambda^{-1})$ lies in the middle $\frac{1}{2}$ of Q_ν.

Definition 14.31. Given Q_ν, and $C_5 > 0$ define the following subset:

$$R_\nu(C_5) := \{x \in Q_\nu : \left|\log \varphi_\lambda^2(x) - \log \mathrm{Ave}_{Q_\nu} \varphi_\lambda^2\right| \leq C_5\}.$$

In [**DF**] (Proposition 5.11 and Lemmas 7.3-7.4) is proved:

Proposition 14.32. *Given $\varepsilon > 0$ there exists $C_5(\varepsilon), a_4 > 0$ so that for at least half of the Q_ν,* $\mathrm{Vol}(R_\nu(C_5(\varepsilon))) \geq (1 - a_4\varepsilon)\,\mathrm{Vol}(Q_\nu)$.

Here, $\mathrm{Vol}(E) = |E|$ is the Riemannian volume measure.

Definition 14.33. Say that Q_ν (or ν or B_ν) is 'preferred' (or 'good') if the inequality of Proposition 14.32 is satisfied. Denote the set of preferred ν by $\mathcal{S}$.

In §14.34, we review results of [**DF**] showing that eigenfunctions in good B_ν are rather flat in that that different L^p norms are equivalent on them.

The next key Proposition of [**DF**] is:

Proposition 14.34. *Let*

$$G_\nu^+ = \{x \in B_\nu : \varphi_\lambda > 0\}, \quad G_\nu^- = \{x \in B_\nu : \varphi_\lambda < 0\}.$$

For preferred $\nu \in \mathcal{S}$, i.e. for 'good balls' B_ν,

$$\min\{|G_\nu^+|, |G_\nu^-|\} \geq E_8|B_\nu|.$$

The lower bound on $\mathcal{H}^{n-1}(Z_{\varphi_\lambda})$ is derived from Propositions 14.32 and 14.34. The idea is the nodal hypersurface cuts 'good' cubes (or balls) roughly into halves. One may then use the isoperimetric inequality for analytic sets ([**F1**, p.476]) to show that the hypersurface volume of the nodal set in Q_ν is bounded below by that of a hyperplane dissecting the cube,

$$\mathcal{H}^{n-1}(Z_{\varphi_\lambda} \cap B_\nu) \geq C\lambda^{-(n-1)}.$$

Summing over $\simeq \frac{1}{2}\lambda^n$ preferred cubes then gives the lower bound of λ.

14.33.2. A simpler statement. A simpler statement than Proposition 14.34 that illustrates some of the ideas in the proof is the following:

PROPOSITION 14.35. *For preferred $\nu \in \mathcal{S}$,*

$$\min\left\{\int_{G_\nu^+} |\varphi_\lambda|\,dx,\ \int_{G_\nu^-} |\varphi_\lambda|\,dx\right\} \geq C_8 \int_{B_\nu} |\varphi_\lambda|\,dx.$$

PROOF. For sufficiently small a_3, rescale φ in the ball $B(x_\nu, a_3\lambda^{-1})$ to the ball of radius 1. Then express the scaled φ using the Dirichlet Green's function $G_D(\lambda, x, y)$ of the scaled ball for the scaled operator. It is almost the flat Green's function. The function $G_D(\lambda, x, x_\nu)$ is denoted φ in [**DF**, p.182]. Then write in unscaled coordinates

$$\varphi_\lambda(x_\nu) = \int_{|x-x_\nu|=r} G_D(\lambda, x, x_\nu)\varphi_\lambda(x)d\theta.$$

Here $0 < r < a_3\lambda^{-1}$. □

We then claim that

$$E_8 > G_D(\lambda, x, x_\nu) > C_7 > 0.$$

This is just a property of the Dirichlet Poisson kernel of the scaled problem, scaled back again.

Since $\varphi_\lambda(x_\nu) = 0$ we have

$$\int_{|x-x_\nu|=r} G_D(\lambda, x, x_\nu)\varphi_\lambda(x)\,d\theta = 0.$$

Proposition 14.35 follows from this and the bounds in the last Lemma.

14.34. Properties of eigenfunctions in good balls

In this section, we review the properties of eigenfunctions in good balls B_ν. The main theme is that they are rather flat and all their L^pnorms are equivalent on a good ball.

This following Lemma says that the L^1 mass of eigenfunctions is rather uniform in good balls.

LEMMA 14.36. *Let (B_ν, Q_ν) be good. Then for any measurable $G_\nu \subset B_\nu$,*

$$\int_{G_\nu} |\varphi_\lambda| \leq E_6 \left(\frac{|G_\nu|}{|B_\nu|}\right)^{\frac{1}{2}} \int_{B_\nu} |\varphi_\lambda|.$$

Moreover,

$$\frac{1}{|B_\nu|}\int_{B_\nu} |\varphi_\lambda|^2 \geq e^{-C_6}\frac{1}{|Q_\nu|}\int_{Q_\nu} |\varphi_\lambda|^2.$$

The Lemma implies that, on good balls, the L^∞ norm and normalized L^2 norm and normalized L^1 norm are equivalent.

PROPOSITION 14.37. *For the good balls,*

$$\|\varphi_\lambda\|_{L^\infty(B_\nu)} \leq E_5 \left(\frac{1}{|B_\nu|}\int_{B_\nu} |\varphi_\lambda|^2\right)^{\frac{1}{2}}.$$

Also

$$\left(\frac{1}{|B_\nu|}\int_{B_\nu}|\varphi_\lambda|^2\right)^{\frac{1}{2}} \leq E_4\frac{1}{|B_\nu|}\int_{B_\nu}|\varphi_\lambda|.$$

14.35. Background on good-ness

Definition 14.31 and Proposition 14.32 arise from the following Proposition about holomorphic functions of one complex variable.

PROPOSITION 14.38 (See Proposition 5.1 of [**DF**]). *Let* $F \in \mathcal{O}(D(0,3))$ *and assume*

$$\max_{D(0,2)}|F(z)| \leq |F(0)|e^{Cd}.$$

Assume $F(x)$ *is real and* ≥ 0 *for* $|x| \leq 1$. *For* d *sufficiently large consider a cover of* $|x| \leq 1$ *by disjoint subintervals* Q_ν *of length* C_2/d. *Then, for all* $\varepsilon > 0$, *there exists* E *so that*

$$|\log F(x) - \log \mathrm{Ave}_{Q_\nu} F| \leq C, \quad x \in Q_\nu \backslash E$$

where meas(E) $< \varepsilon$. *Equivalently, for all* ε *there exists* C *so that for all* ν,

$$m\{|\log F(x) - \log \mathrm{Ave}_{Q_\nu} F| \geq C\} \leq \varepsilon.$$

Thus, if we delete a set E of small length, (which depends on F) any 'singularity' of $\log F$ (i.e. a near zero of F) is canceled by the singularity of $\log \mathrm{Ave}_{Q_\nu} F$ except for points in E.

14.36. A. Brudnyi's proof of Proposition 14.38

To prove Proposition 14.38, the authors write

$$F(z) = e^{G(z)}\prod_\alpha B_r(z,\alpha), \quad |z| \leq r,$$

where

$$B_r(z) = \frac{(z-\alpha)/r}{1-\bar{\alpha}z/r^2}.$$

The e^G factor is relatively harmless and one reduces to proving $|f(x) - \log \mathrm{Ave}(e^f)| \leq C$ outside E if $f(z) = \sum_\alpha \log|z-\alpha|$.

Let $Z_F = \{\alpha \in D(0,\frac{3}{2}) : F(\alpha) = 0\}$ and let $N_F = \#Z_F$ (counted with multiplicity). Then $N_F \leq C_1'd$ by Jensen's formula with $C_1' = \frac{1}{\log\frac{4}{3}}C_1$.

We then use the Cartan Lemma.

LEMMA 14.39. *There exists a set of at most* N_F *discs* $D_1,\ldots,D_k$ *covering the* Z_F *such that the sum of the radii is* $\leq 2\varepsilon$ *and for* $x \notin \bigcup_{j=1}^k D_j$ *the number of elements in* $D(x,R)\cap Z_F$ *is* $\leq \frac{N_F}{\varepsilon}R$.

Let E_1 be as in [**DF**, Lemma 5.4] and let

$$E_2 = \bigcup_j D_j \cap [-1,1].$$

Then $|E_2| \leq 2\varepsilon$. Suppose that $Q_\nu \not\subset E_1 \cup E_2$ and let $x_\nu \notin E_1\cup E_2$. Let A_ν be the set of zeros $z_\nu \in Z_F$ of distance $\leq \frac{2C_2}{d} = 2m(Q_\nu)$ from x_ν. Write

$$f(x) = \sum_{\alpha\in A_\nu}\log|x-\alpha| + \sum_{\alpha\notin A_\nu}\log|x-\alpha| =: b_\nu(x) + g_\nu(x).$$

LEMMA 14.40. *We have*

(1) Card $A_\nu \le \frac{2C_2 N_F}{d\varepsilon} \le \frac{9C_1C_2}{\varepsilon}$;
(2) $|g'_\nu(x_\nu)| \le C_{10} d$;
(3) $|g''_\nu(x)| \le C_{11} d^2$ *for all* $x \in Q_\nu$.

PROOF. (1) follows from the choice of covering, the Cartan Lemma and Lemma (14.39). For (2) we define

$$d_\nu = \inf_{\alpha \in Z_F} |x_\nu - \alpha|.$$

We claim that $d_\nu \ge \frac{\varepsilon}{N_F}$. Suppose that $d_\nu = |x_\nu - \alpha_0|$. By the Cartan Lemma the number of zeros in the disc $|x - \alpha_0| \le d_\nu$ is ≥ 1, hence $1 \le \frac{N_F}{\varepsilon} d_\nu$ and that gives the required estimate.

Let $\mu = \sum_{z_\nu \in Z_F} \delta_{z_\nu}$. We have $n_\nu(r) = 0$ for $r < \frac{\varepsilon}{N_F}$. By integration by parts, by (1) and by the Cartan Lemma,

$$|b'_\nu(x_\nu)| \le \int_{|z - x_\nu| \le 2C_2/d} \frac{1}{|z - x_\nu|}\, d\mu(z) = \int_0^{2C_2/d} \frac{1}{t}\, dn_\nu(t) = \frac{n_\nu(t)}{t}\Big|_0^{2C_2/d}.$$

Moreover,

$$\int_0^{2C_2/d} \frac{1}{t^2}\, dn_\nu(t) \le \frac{9C_1C_2}{\varepsilon} + \int_{\varepsilon/d}^{2C_2/d} \frac{N_F}{\varepsilon}\frac{t}{t^2}\, dn_\nu(t) \tag{14.105}$$

$$\le \frac{9C_1C_2}{\varepsilon} + \frac{9C_1 d}{\varepsilon} \log \frac{2C_2}{\varepsilon} \tag{14.106}$$

$$:= C_9 d. \tag{14.107}$$

Since $x_\nu \notin E_1$

$$|g'_\nu(x_\nu)| \le |f'_\nu(x_\nu)| + |b'_\nu(x_\nu)| \le (C_7 + C_9)d = C_{10} d.$$

For (3): For every $x \in Q_\nu$ and $\alpha \notin A_\nu$,

$$|x - \alpha| \ge |x_\nu - \alpha| - |x - x_\nu| \ge \frac{2C - 2}{d} - \frac{C_2}{d} = \frac{C_2}{d}.$$

Thus,

$$|g''_\nu(x)| \le \sum_{\alpha \notin A_\nu} \frac{1}{|x - \alpha|^2} \tag{14.108}$$

$$= \sum_{\alpha \notin A_\nu} \frac{1}{|x_\nu - \alpha|^2} \frac{|x_\nu - \alpha|^2}{|x - \alpha|^2} \tag{14.109}$$

$$\le \sum_{\alpha \notin A_\nu} \frac{1}{|x_\nu - \alpha|^2} \frac{(|x - \alpha| + |x_\nu - x|)^2}{|x - \alpha|^2} \le 3 \sum_{\alpha \notin A_\nu} \frac{1}{|x_\nu - \alpha|^2}. \tag{14.110}$$

Again using integration by parts and Cartan's Lemma,

$$\sum_{\alpha\notin A_\nu} \frac{1}{|x_\nu-\alpha|^2} = \int_{|z-x_\nu|\geq 2C_2/d} \frac{1}{|z-x_\nu|^2}\,d\mu(z) \tag{14.111}$$

$$= \int_{2C_2/d}^{\infty} \frac{1}{t^2}\,dn_\nu(t) \tag{14.112}$$

$$= \frac{n_\nu(t)}{t^2}\Big|_{2C_2/d}^{\infty} + \int_{2C_2/d}^{\infty} \frac{2n_\nu(t)}{t^3}\,dt \tag{14.113}$$

$$\leq \int_{2C_2/d}^{\infty} \frac{2N_F t}{\varepsilon t^3}\,dt \tag{14.114}$$

$$\leq \frac{18C_1 d^2}{C_2\varepsilon}. \tag{14.115}$$

Together with the previous inequality, this implies (3). □

Now set $S = \bigcup_\nu Q_\nu$ where the union is taken over all $Q_\nu \not\subset E_1 \cup E_2$. We have

$$m([-1,1]\backslash S) \leq m(E_1\cup E_2) \leq a_2\varepsilon.$$

For $x_\nu \in Q_\nu \subset S$, by Taylor's formula with remainder,

$$g_\nu(x) - g_\nu(x_\nu) = g'_\nu(x_\nu(x-x_\nu) + \frac{1}{2}g''_\nu(y)(x-x_\nu)^2 \quad \text{for some } y \in Q_\nu.$$

Together with Lemma 5.5 (2)(3) this gives

$$|g_\nu(x) - g_\nu(x_\nu)| \leq C_{10}C_2 + C_{11}C_2^2/2.$$

Since $\log \operatorname{Ave}_{Q_\nu} e^{g_\nu} = g_\nu(z)$ for some $z \in Q_\nu$ we have

$$|g_\nu(x) - \log \operatorname{Ave}_{Q_\nu} e^{g_\nu}| \leq C_{12} \quad \text{for all } x \in Q_\nu. \tag{14.116}$$

Now consider b_ν for $Q_\nu \subset S$. By Lemma 5.5 (1), the number of terms in the expression for b_ν is at most C_{13}. The classical Remez polynomial inequality for e^{b_ν} then implies

$$\sup_{Q_\nu} e^{b_\nu} \leq \left(\frac{4m(Q_\nu)}{m(S_\nu)}\right)^{C_{13}} \sup_{S_\nu} e^{b_\nu}, \tag{14.117}$$

for any measurable subset $S_\nu \subset Q_\nu$. Put

$$S_\nu = \{x \in Q_\nu : b_\nu(x) - \sup_{Q_\nu} b_\nu \leq -C_{14} = C_{13}\log\varepsilon\}.$$

Then $\sup_{S_\nu} e^{b_\nu} = e^{-C_{14}} \sup_{Q_\nu} e^{b_\nu}$, hence

$$m(S_\nu) \leq 4m(Q_\nu)e^{-C_{14}/C_{13}} = 4C_2\frac{\varepsilon}{d}.$$

Define $E_3 = \bigcup_{Q_\nu\subset S} S_\nu$. Then

$$m(E_3) \leq \frac{2d}{C_2}4C_2\frac{\varepsilon}{d} = 8\varepsilon.$$

Further use the well-known inequality relating L^1 and L^∞ norms of polynomials

$$\sup_{Q_\nu} e^{b_\nu} \leq a_4 C_{13}^2 \operatorname{Ave}_{Q_\nu} e^{b_\nu}.$$

(See [**Ti**, §4.9.6].) Then by the previous inequalities, for $x \in Q_\nu \backslash E_3$,

$$|b_\nu(x) - \mathrm{Av}_{Q_\nu} e^{b_\nu}| \leq |b_\nu(x) - \sup_{Q_\nu} b_\nu| + |\sup_{Q_\nu} b_\nu - \log \mathrm{Ave}_{Q_\nu} e^{b_\nu}| \leq C_{15}. \tag{14.118}$$

The required statement follows from (14.116), (14.118) for all $x \in Q_\nu \backslash (E_1 \cup E_2 \cup E_3)$. Note that $m(E_1 \cup E_2 \cup E_3) \leq a_5 \varepsilon$.

14.37. Equidistribution of complex nodal sets of real ergodic eigenfunctions

We now consider global results when hypotheses are made on the dynamics of the geodesic flow. Use of the global wave operator brings into play the relation between the geodesic flow and the complexified eigenfunctions, and this allows one to prove global results on nodal hypersurfaces that reflect the dynamics of the geodesic flow. In some cases, one can determine not just the volume, but the limit distribution of complex nodal hypersurfaces. The complex nodal hypersurface of an eigenfunction is defined by

$$Z_{\varphi_\lambda^{\mathbb{C}}} = \{\zeta \in M_{\varepsilon_0} : \varphi_\lambda^{\mathbb{C}}(\zeta) = 0\}. \tag{14.119}$$

The Poincaré-Lelong formula (§14.30.1) gives an explicit form to the natural current of integration over the nodal hypersurface in any Grauert tube M_ε with $\varepsilon < \varepsilon_0$, given by

$$\langle [Z_{\varphi_\lambda^{\mathbb{C}}}], \varphi \rangle = \frac{i}{2\pi} \int_{M_\varepsilon} \partial\bar{\partial} \log |\varphi_\lambda^{\mathbb{C}}|^2 \wedge \varphi = \int_{Z_{\varphi_\lambda^{\mathbb{C}}}} \varphi, \qquad \varphi \in \mathcal{D}^{(m-1,m-1)}(M_\varepsilon). \tag{14.120}$$

We recall that $\mathcal{D}^{(m-1,m-1)}(M_\varepsilon)$ stands for smooth test $(m-1, m-1)$-forms with support in $M*_\varepsilon$.

The nodal hypersurface $Z_{\varphi_\lambda^{\mathbb{C}}}$ also carries a natural volume form $|Z_{\varphi_\lambda^{\mathbb{C}}}|$ as a complex hypersurface in a Kähler manifold. By Wirtinger's formula, it equals the restriction of $\frac{\omega_g^{m-1}}{(m-1)!}$ to $Z_{\varphi_\lambda^{\mathbb{C}}}$. Hence, one can regard $Z_{\varphi_\lambda^{\mathbb{C}}}$ as defining the measure

$$\langle |Z_{\varphi_\lambda^{\mathbb{C}}}|, \varphi \rangle = \int_{Z_{\varphi_\lambda^{\mathbb{C}}}} \varphi \frac{\omega_g^{m-1}}{(m-1)!}, \qquad \varphi \in C(B_\varepsilon^* M). \tag{14.121}$$

We prefer to state results in terms of the current $[Z_{\varphi_\lambda^{\mathbb{C}}}]$ since it carries more information.

THEOREM 14.41. *Let (M, g) be real analytic, and let $\{\varphi_{j_k}\}$ denote a quantum ergodic sequence of eigenfunctions of its Laplacian Δ. Let M_{ε_0} be the maximal Grauert tube around M. Let $\varepsilon < \varepsilon_0$. Then:*

$$\frac{1}{\lambda_{j_k}} [Z_{\varphi_{j_k}^{\mathbb{C}}}] \to \frac{i}{\pi} \partial\bar{\partial}\sqrt{\rho} \quad \textit{weakly in } \mathcal{D}'^{(1,1)}(M_\varepsilon)$$

in the sense that, for any continuous test form $\psi \in \mathcal{D}^{(m-1,m-1)}(M_\varepsilon)$, we have

$$\frac{1}{\lambda_{j_k}} \int_{Z_{\varphi_{j_k}^{\mathbb{C}}}} \psi \to \frac{i}{\pi} \int_{M_\varepsilon} \psi \wedge \partial\bar{\partial}\sqrt{\rho}.$$

Equivalently, for any $\varphi \in C(M_\varepsilon)$,

$$\frac{1}{\lambda_{j_k}} \int_{Z_{\varphi^{\mathbb{C}}_{j_k}}} \varphi \frac{\omega_g^{m-1}}{(m-1)!} \to \frac{i}{\pi} \int_{M_\varepsilon} \varphi \partial\bar{\partial}\sqrt{\rho} \wedge \frac{\omega_g^{m-1}}{(m-1)!}.$$

14.38. Sketch of the proof

The first step is to find a nice way to express $\varphi_j^{\mathbb{C}}$ on $M_{\mathbb{C}}$. Very often, when we analytically continue a function, we lose control over its behavior. The trick is to observe that the complexified wave group analytically continues the eigenfunctions, i.e., we use

$$U(i\tau, \zeta, y) = \sum_{j=0}^{\infty} e^{-\tau\lambda_j} \varphi_j^{\mathbb{C}}(\zeta)\varphi_j(y).$$

It is holomorphic in $\zeta \in M_\tau$, i.e., when $\sqrt{\rho}(\zeta) < \tau$. But the main point is that it remains a Fourier integral operator after analytic continuation:

THEOREM 14.42 (Hadamard, Boutet de Monvel, Z, M. Stenzel, G. Lebeau). $U(i\varepsilon, z, y)\colon L^2(M) \to H^2(\partial M_\varepsilon)$ *is a complex Fourier integral operator of order* $-\frac{m-1}{4}$ *quantizing the complexified exponential map* $\exp_y i\varepsilon\frac{\eta}{|\eta|}\colon S^*_\varepsilon \to \partial M_\varepsilon$.

We first observe that $U(i\tau)\varphi_{\lambda_j} = e^{-\tau\lambda_j}\varphi^{\mathbb{C}}_{\lambda_j}$. This follows immediately by integrating $U(i\tau,\zeta,y) = \sum_{k=0}^{\infty} e^{-\tau\lambda_k}\varphi_k^{\mathbb{C}}(\zeta)\varphi_k(y)$ against φ_j and using orthogonality.

But we know that $U(i\tau)\varphi_{\lambda_j}$ is a Fourier integral operator. It is a fact that such an operator can only change L^2 norms by powers of λ_j. So $\|U(i\tau)\varphi_{\lambda_j}\|^2_{L^2(\partial M_\varepsilon)}$ has polynomial growth in λ_j and therefore we have

$$\|\varphi_{\lambda_j}\|^2_{L^2(\partial M_\varepsilon)} = \lambda_j^{\alpha} e^{\tau\lambda_j}.$$

The power α is irrelevant because we are taking the normalized logarithm.

The first step is to prove quantum ergodicity of the complexified eigenfunctions:

THEOREM 14.43. *Assume the geodesic flow of* (M, g) *is ergodic. Then*

$$\frac{|\varphi^{\varepsilon}_{j_k}(z)|^2}{\|\varphi^{\varepsilon}_{j_k}\|^2_{L^2(\partial M_\varepsilon)}} \to 1 \quad \textit{weak* in } \partial M_\varepsilon$$

along a density one subsequence of λ_j. *That is, for any continuous* V,

$$\int_{\partial M_\varepsilon} V \frac{|\varphi^{\varepsilon}_{j_k}(z)|^2}{\|\varphi^{\varepsilon}_{j_k}\|^2_{L^2(\partial M_\varepsilon)}}\, dV \to \int_{\partial M_\varepsilon} V\, dV.$$

Thus, Husimi measures tend to 1 weakly as measures. We then apply Hartogs' Lemma (Lemma 14.21) to obtain

LEMMA 14.44. *We have: For all but a sparse subsequence of eigenvalues,*

$$\frac{1}{\lambda_{j_k}} \log \frac{|\varphi^{\varepsilon}_{j_k}(z)|^2}{\|\varphi^{\varepsilon}_{j_k}\|^2_{L^2(\partial M_\varepsilon)}} \to 0 \quad \textit{in } L^1(M_\varepsilon).$$

This is almost obvious from the QE theorem. The limit is ≤ 0 and it were < 0 on a set of positive measure it would contradict

$$\frac{|\varphi^{\varepsilon}_{j_k}(z)|^2}{\|\varphi^{\varepsilon}_{j_k}\|^2_{L^2(\partial M_\varepsilon)}} \to 1.$$

Combine Lemma 14.44 with Poincaré-Lelong:

$$\frac{1}{\lambda_{j_k}}[Z_{j_k}] = i\partial\bar{\partial}\log|\varphi_{j_k}^{\mathbb{C}}|^2.$$

We get

$$\frac{1}{\lambda_{j_k}}\partial\bar{\partial}\log|\varphi_{j_k}^{\mathbb{C}}|^2 \sim \frac{1}{\lambda_{j_k}}\partial\bar{\partial}\log|\varphi_{j_k}^{\mathbb{C}}|^2_{L^2(\partial M_\varepsilon)} \quad \text{weak* on } M_\varepsilon.$$

To complete proof we need to prove:

$$\frac{1}{\lambda_j}\log\|\varphi_j^{\mathbb{C}}\|^2_{\partial M_\varepsilon} \to 2\varepsilon. \tag{14.122}$$

But $U(i\varepsilon) = e^{-\varepsilon\lambda_j}\varphi_j^{\mathbb{C}}$, hence $\|\varphi_\lambda^{\mathbb{C}}\|^2_{L^2(\partial M_\varepsilon)}$ equals $e^{2\varepsilon\lambda_j}$ times

$$\langle U(i\varepsilon)\varphi_\lambda, U(i\varepsilon)\varphi_\lambda\rangle = \langle U(i\varepsilon)^*U(i\varepsilon)\varphi_\lambda, \varphi_\lambda\rangle.$$

But $U(i\varepsilon)^*U(i\varepsilon)$ is a pseudodifferential operator of order $\frac{n-1}{2}$. Its symbol $|\xi|^{-\frac{n-1}{2}}$ doesn't contribute to the logarithm.

We now provide more details at each step.

14.39. Growth properties of complexified eigenfunctions

In this section we prove Lemma 14.44 in more detail. We state it in combination with (14.122).

THEOREM 14.45. *If the geodesic flow is ergodic, then for all but a sparse subsequence of λ_j,*

$$\frac{1}{\lambda_{j_k}}\log|\varphi_{j_k}^{\mathbb{C}}(z)|^2 \to \sqrt{\rho} \quad \textit{in } L^1(M_\varepsilon).$$

The Grauert tube function is a maximal PSH function with bound $\le \varepsilon$ on M_ε. Hence Theorem 14.45 says that ergodic eigenfunctions have the maximum exponential growth rate possible for any eigenfunctions.

A key object in the proof is the sequence of functions $U_\lambda(x,\xi) \in C^\infty(M_\varepsilon)$ defined by

(14.123)

$$U_\lambda(x,\xi) := \frac{\varphi_\lambda^{\mathbb{C}}(x,\xi)}{\rho_\lambda(x,\xi)} \quad \text{where } (x,\xi)\in M_\varepsilon \text{ and } \rho_\lambda(x,\xi) := \|\varphi_\lambda^{\mathbb{C}}|_{\partial M_{|\xi|_g}}\|_{L^2(\partial M_{|\xi|_g})}.$$

Thus, $\rho_\lambda(x,\xi)$ is the L^2-norm of the restriction of $\varphi_\lambda^{\mathbb{C}}$ to the sphere bundle ∂M_ε where $\varepsilon = |\xi|_g$. U_λ is of course not holomorphic, but its restriction to each sphere bundle is CR holomorphic there, i.e.,

$$u_\lambda^\varepsilon = U_\lambda|_{\partial M_\varepsilon} \in \mathcal{O}^0(\partial M_\varepsilon). \tag{14.124}$$

Our first result gives an ergodicity property of holomorphic continuations of ergodic eigenfunctions.

LEMMA 14.46. *Assume that $\{\varphi_{j_k}\}$ is a quantum ergodic sequence of Δ-eigenfunctions on M. Then for each $0<\varepsilon<\varepsilon_0$,*

$$|U_{j_k}|^2 \to \frac{1}{\mu_1(S^*M)}\sqrt{\rho}^{-m+1} \quad \textit{weakly in } L^1(M_\varepsilon,\omega^m).$$

We note that $\omega^m = r^{m-1} dr d\omega d\,\mathrm{Vol}(x)$ in polar coordinates, so the right side indeed lies in L^1. The actual limit function is otherwise irrelevant. The next step is to use a compactness argument to obtain strong convergence of the normalized logarithms of the sequence $\{|U_\lambda|^2\}$. The first statement of the following lemma immediately implies the second.

LEMMA 14.47. *Assume that* $|U_{j_k}|^2 \to \frac{1}{\mu_\varepsilon(\partial M_\varepsilon)} \sqrt{\rho}^{-m+1}$ *weakly in* $L^1(M_\varepsilon, \omega^m)$. *Then*

(1) $\frac{1}{\lambda_{j_k}} \log |U_{j_k}|^2 \to 0$ *strongly in* $L^1(M_\varepsilon)$;
(2) $\frac{1}{\lambda_{j_k}} \partial\bar{\partial} \log |U_{j_k}|^2 \to 0$ *weakly in* $\mathcal{D}'^{(1,1)}(M_\varepsilon)$.

Separating out the numerator and denominator of $|U_j|^2$, we obtain that

$$\frac{1}{\lambda_{j_k}} \partial\bar{\partial} \log |\varphi_{j_k}^{\mathbb{C}}|^2 - \frac{2}{\lambda_{j_k}} \partial\bar{\partial} \log \rho_{\lambda_{j_k}} \to 0. \tag{14.125}$$

The next lemma shows that the second term has a weak limit:

LEMMA 14.48. *For* $0 < \varepsilon < \varepsilon_0$,

$$\frac{1}{\lambda j_k} \log \rho_{\lambda j_k}(x, \xi) \to \sqrt{\rho}, \quad \textit{in } L^1(M_\varepsilon).$$

Hence,

$$\frac{1}{\lambda_{j_k}} \partial\bar{\partial} \log \rho_{\lambda_{j_k}} \to \partial\bar{\partial}\sqrt{\rho} \quad \textit{weakly in } \mathcal{D}'(M_\varepsilon).$$

It follows that the left side of (14.125) has the same limit, and that will complete the proof of Theorem 14.41.

We begin by proving a weak limit formula for the CR holomorphic functions u_λ^ε defined in (14.124) for fixed ε.

LEMMA 14.49. *Assume that* $\{\varphi_{j_k}\}$ *is a quantum ergodic sequence. Then for each* $0 < \varepsilon < \varepsilon_0$,

$$|u_{j_k}^\varepsilon|^2 \to \frac{1}{\mu_\varepsilon(\partial M_\varepsilon)}, \quad \textit{weakly in } L^1(\partial M_\varepsilon, d\mu_\varepsilon).$$

That is, for any $a \in C(\partial M_\varepsilon)$,

$$\int_{\partial M_\varepsilon} a(x,\xi) |u_{j_k}^\varepsilon((x,\xi)|^2 \, d\mu_\varepsilon \to \frac{1}{\mu_\varepsilon(\partial M_\varepsilon)} \int_{\partial M_\varepsilon} a(x,\xi) \, d\mu_\varepsilon.$$

PROOF. It suffices to consider $a \in C^\infty(\partial M_\varepsilon)$. We then consider the Toeplitz operator $\Pi_\varepsilon a \Pi_\varepsilon$ on $\mathcal{O}^0(\partial M_\varepsilon)$. We have

$$\langle \Pi_\varepsilon a \Pi_\varepsilon u_j^\varepsilon, u_j^\varepsilon \rangle = e^{2\varepsilon\lambda_j} \, \|\varphi_\lambda^{\mathbb{C}}\|_{L^2(\partial M_\varepsilon)}^{-2} \langle \Pi_\varepsilon a \Pi_\varepsilon U(i\varepsilon)\varphi_j, U(i\varepsilon)\varphi_j \rangle_{L^2(\partial M_\varepsilon)} \tag{14.126}$$

$$= e^{2\varepsilon\lambda_j} \|\varphi_\lambda^{\mathbb{C}}\|_{L^2(\partial M_\varepsilon)}^{-2} \langle U(i\varepsilon)^* \Pi_\varepsilon a \Pi_\varepsilon U(i\varepsilon)\varphi_j, \varphi_j \rangle_{L^2(M)}. \tag{14.127}$$

It is not hard to see that $U(i\varepsilon)^* \Pi_\varepsilon a \Pi_\varepsilon U(i\varepsilon)$ is a pseudodifferential operator on M of order $-\frac{m-1}{2}$ with principal symbol $\tilde{a}|\xi|_g^{-\frac{m-1}{2}}$, where $\tilde{a}$ is the (degree 0) homogeneous extension of a to $T^*M \setminus 0$. The normalizing factor $e^{2\varepsilon\lambda_j} \|\varphi_\lambda^{\mathbb{C}}\|_{L^2(\partial B_\varepsilon^* M)}^{-2}$ has the same form with $a = 1$. Hence, the expression on the right side of (14.126) may be written as

$$\frac{\langle U(i\varepsilon)^* \Pi_\varepsilon a \Pi_\varepsilon U(i\varepsilon)\varphi_j, \varphi_j \rangle_{L^2(M)}}{\langle U(i\varepsilon)^* \Pi_\varepsilon U(i\varepsilon)\varphi_j, \varphi_j \rangle_{L^2(M)}}. \tag{14.128}$$

By the standard quantum ergodicity result on compact Riemannian manifolds with ergodic geodesic flow, we have

$$\frac{\langle U(i\varepsilon)^*\Pi_\varepsilon a\Pi_\varepsilon U(i\varepsilon)\varphi_{j_k},\varphi_{j_k}\rangle_{L^2(M)}}{\langle U(i\varepsilon)^*\Pi_\varepsilon U(i\varepsilon)\varphi_{j_k},\varphi_{j_k}\rangle_{L^2(M)}} \to \frac{1}{\mu_\varepsilon(\partial M_\varepsilon)}\int_{\partial M_\varepsilon} a\,d\mu_\varepsilon. \tag{14.129}$$

More precisely, the numerator is asymptotic to the right side times $\lambda^{-\frac{m-1}{2}}$, while the denominator has the same asymptotics when a is replaced by 1. We also use that $\frac{1}{\mu_\varepsilon(\partial M_\varepsilon)}\int_{\partial M_\varepsilon} a d\mu_\varepsilon$ equals the analogous average of $\tilde{a}$ over ∂M_ε. Taking the ratio produces (14.129).

Combining (14.126), (14.129) and the fact that

$$\langle \Pi_\varepsilon a \Pi_\varepsilon u_j^\varepsilon, u_j^\varepsilon\rangle = \int_{\partial B^*_\varepsilon M} a|u_j^\varepsilon|^2\,d\mu_\varepsilon$$

completes the proof of the lemma. □

We now complete the proof of Lemma 14.46, i.e., we prove that

$$\int_{M_\varepsilon} a|U_{j_k}|^2\omega^m \to \frac{1}{\mu_\varepsilon(\partial M_\varepsilon)}\int_{M_\varepsilon} a\sqrt{\rho}^{-m+1}\omega^m \tag{14.130}$$

for any $a \in C(M_\varepsilon)$. It is only necessary to relate the surface Liouville measures $d\mu_r$ (14.30) to the Kähler volume measure. One may write $d\mu_r = \frac{d}{dt}|_{t=r}\chi_t\omega^m$, where χ_t is the characteristic function of $M_t = \{\sqrt{\rho} \le t\}$. By homogeneity of $|\xi|_g$, $\mu_r(\partial M_r) = r^{m-1}\mu_\varepsilon(\partial M_\varepsilon)$. If $a \in C(M_\varepsilon)$, then $\int_{M_\varepsilon} a\omega^m = \int_0^\varepsilon\{\int_{\partial M_r} a\,d\mu_r\}dr$. By Lemma 14.49, we have

$$\int_{M_\varepsilon} a|U_{j_k}|^2\omega^m = \int_0^\varepsilon \left(\int_{\partial M_r} a|u_{j_k}^r|^2\,d\mu_r\right)dr \tag{14.131}$$

$$\to \int_0^\varepsilon \left(\frac{1}{\mu_r(\partial B_r^*)}\int_{\partial M_r} a\,d\mu_r\right)dr \tag{14.132}$$

$$= \frac{1}{\mu_\varepsilon(\partial M_\varepsilon)}\int_{M_\varepsilon} ar^{-m+1}\omega^m, \tag{14.133}$$

which implies $w^* - \lim_{\lambda\to\infty}|U_{j_k}|^2 = \frac{1}{\mu_1(\partial M_\varepsilon)}\sqrt{\rho}^{-m+1}$.

14.40. Proof of Lemma 14.48

In fact, one has

$$\frac{1}{\lambda}\log\rho_\lambda(x,\xi) \to \sqrt{\rho} \quad \text{uniformly on } M_\varepsilon.$$

Again using $U(i\varepsilon)\varphi_\lambda = e^{-\lambda\varepsilon}\varphi_\lambda^{\mathbb{C}}$, we have for $\varepsilon = |\xi|_{g_x}$,

$$\rho_\lambda^2(x,\xi) = \langle \Pi_\varepsilon\varphi_\lambda^{\mathbb{C}}, \Pi_\varepsilon\varphi_\lambda^{\mathbb{C}}\rangle_{L^2(\partial B^*_\varepsilon M)} \tag{14.134}$$

$$= e^{2\lambda\varepsilon}\langle \Pi_\varepsilon U(i\varepsilon)\varphi_\lambda, \Pi_\varepsilon U(i\varepsilon)\varphi_\lambda\rangle_{L^2(\partial B^*_\varepsilon M)} \tag{14.135}$$

$$= e^{2\lambda\varepsilon}\langle U(i\varepsilon)^*\Pi_\varepsilon U(i\varepsilon)\varphi_\lambda, \varphi_\lambda\rangle_{L^2(M)}. \tag{14.136}$$

Hence,

$$\frac{2}{\lambda}\log\rho_\lambda(x,\xi) = 2|\xi|_{g_x} + \frac{1}{\lambda}\log\langle U(i\varepsilon)^*\Pi_\varepsilon U(i\varepsilon)\varphi_\lambda,\varphi_\lambda\rangle. \tag{14.137}$$

The second term on the right side is the matrix element of a pseudo-differential operator, so is bounded by some power of λ. Taking the logarithm gives a remainder of order $\frac{\log \lambda}{\lambda}$. □

14.41. Proof of Lemma 14.47

PROOF. We wish to prove that $\psi_j := \frac{1}{\lambda_j} \log |U_j|^2 \to 0$ in $L^1(M_\varepsilon)$. As we have said, this is almost obvious from Lemma 14.46 and 14.49. If the conclusion is not true, then there exists a subsequence ψ_{j_k} satisfying $\|\psi_{j_k}\|_{L^1(B^*_\varepsilon M)} \geq \delta > 0$. To obtain a contradiction, we use Lemma 14.21.

To see that the hypotheses are satisfied in our example, it suffices to prove these statements on each surface ∂M_ε with uniform constants independent of ε. On the surface ∂M_ε, $U_j = u_j^\varepsilon$. By the Sobolev inequality in $\mathcal{O}^{\frac{m-1}{4}}(\partial M_\varepsilon)$, we have

$$\sup_{(x,\xi) \in \partial M_\varepsilon} |u_j^\varepsilon(x,\xi)| \leq \lambda_j^m \|u_j^\varepsilon(x,\xi)\|_{L^2(\partial M_\varepsilon)} \leq \lambda_j^m.$$

Taking the logarithm, dividing by λ_j, and combining with the limit formula of Lemma 14.48 proves (i) – (ii).

We now settle the dichotomy above by proving that the sequence $\{\psi_j\}$ does not tend uniformly to $-\infty$ on compact sets. That would imply that $\psi_j \to -\infty$ uniformly on the spheres ∂M_ε for each $\varepsilon < \varepsilon_0$. Hence, for each ε, there would exist $K > 0$ such that for $k \geq K$,

$$\frac{1}{\lambda_{j_k}} \log |u_{j_k}^\varepsilon(z)| \leq -1. \tag{14.138}$$

However, (14.138) implies that

$$|u_{j_k}(z)| \leq e^{-2\lambda_{j_k}} \quad \text{for all } z \in \partial M_\varepsilon,$$

which is inconsistent with the hypothesis that $|u_{j_k}^\varepsilon(z)| \to 1$ in $\mathcal{D}'(\partial M_\varepsilon)$.

Therefore, there must exist a subsequence, which we continue to denote by $\{\psi_{j_k}\}$, that converges in $L^1(M_{\varepsilon_0})$ to some $\psi \in L^1(M_{\varepsilon_0})$. Then,

$$\psi(z) = \limsup_{k \to \infty} \psi_{j_k} \leq 2|\xi|_g \qquad \text{a.e.}$$

Now let

$$\psi^*(z) := \limsup_{w \to z} \psi(w) \leq 0$$

be the upper semi-continuous regularization of ψ. Then ψ^* is plurisubharmonic on M_ε and $\psi^* = \psi$ almost everywhere.

If $\psi^* \leq 2|\xi|_g - \delta$ on a set U_δ of positive measure, then $\psi_{j_k}(\zeta) \leq -\delta/2$ for $\zeta \in U_\delta,\ k \geq K$; i.e.,

$$|\psi_{j_k}(\zeta)| \leq e^{-\delta \lambda_{j_k}} \quad \text{for all } \zeta \in U_\delta \text{ and } k \geq K. \tag{14.139}$$

This contradicts the weak convergence to 1 and concludes the proof. □

14.42. Intersections of nodal sets and analytic curves on real analytic surfaces

It is often possible to obtain more refined results on nodal sets by studying their intersections with some fixed curve C or with a hypersurface H. In dimension two, curves are hypersurfaces and the results are most complete in this dimension. On the one hand, one would like to prove generic upper bounds on the number of intersection points when the curve is non-degenerate in a precise sense called 'good'. On the other hand, in some cases one can prove lower bounds on the number of intersection points. To date, this has only been done under ergodicity assumptions but it is possible that one could find other dynamical assumptions which are sufficient.

Upper bounds on nodal intersections with curves were prove in dimension two for analytic Euclidean plane domains in [**TZ1**]. In the boundary case, it has not been proved at this time that the Poisson wave kernel admits an analytic continuation as a Fourier integral operator with complex phase in a suitable sense [2] In place of the Poisson wave kernel, double layer potentials were used. In §14.43 we discuss upper bounds on the number of intersection points of the nodal set with the boundary of a real analytic plane domain and more general 'good' analytic curves. It is likely that this method can be generalized to higher dimensions, but this has so far not been done. In the boundaryless case, upper bounds on numbers of nodal intersections with 'good curves' C can be proved using the techniques of [**TZ4**].

To obtain lower bounds or asymptotics, we need to add some dynamical hypotheses. In case of ergodic geodesic flow, we can obtain equidistribution theorems for intersections of nodal sets and geodesics on surfaces [**Z7**] (see §14.45). The dimensional restriction is due to the fact that the results are partly based on the quantum ergodic restriction theorems of [**TZ3, TZ2**], which concern restrictions of eigenfunctions to hypersurfaces. Nodal sets and geodesics have complementary dimensions and intersect in points, and therefore it makes sense to count the number of intersections. But we do not yet have a mechanism for studying lower bounds on numbers of nodal zeros for restrictions to geodesics when $\dim M \geq 3$.

In all the results we discuss in this section, there is a significant assumption which as yet is poorly understood: Namely, we assume that the real analytic curve C is 'good,' a kind of Carleman estimate type assumption for a curve. There are few results on the existence or genericity of good curves at this time (one is given in [**JJ1**]).

14.43. Counting nodal lines which touch the boundary in analytic plane domains

In this section, we review the results of [**TZ1**] giving upper bounds on the number of intersections of the nodal set with the boundary of an analytic (or more generally piecewise analytic) plane domain. It would be interesting to generalize the results to higher dimensions, either counting nodal intersections with curves or by measuring codimension two nodal hypersurface volumes within the boundary.

Thus we would like to count the number of nodal lines (i.e., components of the nodal set) which touch the boundary. Here we assume that 0 is a regular value so that components of the nodal set are either loops in the interior (closed nodal loops) or curves which touch the boundary in two points (open nodal lines). It is known that for generic piecewise analytic plane domains, zero is a regular value of all the eigenfunctions φ_{λ_j}, i.e., $\nabla \varphi_{\lambda_j} \neq 0$ on $Z_{\varphi_{\lambda_j}}$ [**U**]; we then call the nodal set regular. Since the boundary lies in the nodal set for Dirichlet boundary conditions, we remove it from the nodal set before counting components. Henceforth, the number of components of the nodal set in the Dirichlet case means the number of components of $Z_{\varphi_{\lambda_j}} \backslash \partial\Omega$.

[2]The real wave kernel is at best a Fourier-Airy integral operator so it is not to be expected that the analytic continuation is any classical kind of Fourier integral operator.

It is important to consider piecewise analytic domains because they are the only plane domains which are know to have ergodic billiards. By a piecewise analytic domain $\Omega^2 \subset \mathbb{R}^2$, we mean a compact domain with piecewise analytic boundary, i.e., $\partial\Omega$ is a union of a finite number of piecewise analytic curves which intersect only at their common endpoints. Such domains are often studied as archetypes of domains with ergodic billiards and quantum chaotic eigenfunctions, in particular the Bunimovich stadium or Sinai billiard.

For the Neumann problem, the boundary nodal points are the same as the zeros of the boundary values $\varphi_{\lambda_j}|_{\partial\Omega}$ of the eigenfunctions. The number of boundary nodal points is thus twice the number of open nodal lines. Hence in the Neumann case, the Theorem follows from:

THEOREM 14.50. *Suppose that $\Omega \subset \mathbb{R}^2$ is a piecewise real analytic plane domain. Then the number $n(\lambda_j) = \#Z_{\varphi_{\lambda_j}} \cap \partial\Omega$ of zeros of the boundary values $\varphi_{\lambda_j}|_{\partial\Omega}$ of the jth Neumann eigenfunction satisfies $n(\lambda_j) \leq C_\Omega \lambda_j$, for some $C_\Omega > 0$.*

We prove Theorem 14.50 by analytically continuing the boundary values of the eigenfunctions and counting *complex zeros and critical points* of analytic continuations of Cauchy data of eigenfunctions. When $\partial\Omega \in C^\omega$, the eigenfunctions can be holomorphically continued to an open tube domain in $\mathbb{C}^2$ projecting over an open neighborhood W in $\mathbb{R}^2$ of Ω which is independent of the eigenvalue. We denote by $\Omega_{\mathbb{C}} \subset \mathbb{C}^2$ the points $\zeta = x + i\xi \in \mathbb{C}^2$ with $x \in \Omega$. Then $\varphi_{\lambda_j}(x)$ extends to a holomorphic function $\varphi^{\mathbb{C}}_{\lambda_j}(\zeta)$ where $x \in W$ and where $|\xi| \leq \varepsilon_0$ for some $\varepsilon_0 > 0$.

Assuming $\partial\Omega$ real analytic, we define the (interior) complex nodal set by

$$Z^{\mathbb{C}}_{\varphi_{\lambda_j}} = \{\zeta \in \Omega_{\mathbb{C}} : \varphi^{\mathbb{C}}_{\lambda_j}(\zeta) = 0\}.$$

THEOREM 14.51. *Suppose that $\Omega \subset \mathbb{R}^2$ is a piecewise real analytic plane domain, and denote by $(\partial\Omega)_{\mathbb{C}}$ the union of the complexifications of its real analytic boundary components. Let $n(\lambda_j, \partial\Omega_{\mathbb{C}}) = \#Z^{\partial\Omega_{\mathbb{C}}}_{\varphi_{\lambda_j}}$ be the number of complex zeros on the complex boundary. Then there exists a constant $C_\Omega > 0$ independent of the radius of $(\partial\Omega)_{\mathbb{C}}$ such that $n(\lambda_j, \partial\Omega_{\mathbb{C}}) \leq C_\Omega \lambda_j$.*

The theorems on real nodal lines and critical points follow from the fact that real zeros and critical points are also complex zeros and critical points, hence

$$n(\lambda_j) \leq n(\lambda_j, \partial\Omega_{\mathbb{C}}). \tag{14.140}$$

All of the results are sharp, and are already obtained for certain sequences of eigenfunctions on a disc (see §13.7.6).

To prove Theorem 14.51, we represent the analytic continuations of the boundary values of the eigenfunctions in terms of layer potentials. Let $G(\lambda_j, x_1, x_2)$ be any 'Green's function' for the Helmholtz equation on Ω, i.e. a solution of $(-\Delta - \lambda_j^2)G(\lambda_j, x_1, x_2) = \delta_{x_1}(x_2)$ with $x_1, x_2 \in \bar{\Omega}$. By Green's formula,

$$\varphi_{\lambda_j}(x, y) = \int_{\partial\Omega} \big(\partial_\nu G(\lambda_j, q, (x, y))\varphi_{\lambda_j}(q) - G(\lambda_j, q, (x, y))\partial_\nu \varphi_{\lambda_j}(q)\big)\, d\sigma(q), \tag{14.141}$$

where $(x, y) \in \mathbb{R}^2$, where $d\sigma$ is arc-length measure on $\partial\Omega$ and where ∂_ν is the normal derivative by the interior unit normal. Our aim is to analytically continue this formula. In the case of Neumann eigenfunctions φ_λ in Ω,

$$\varphi_{\lambda_j}(x, y) = \int_{\partial\Omega} \frac{\partial}{\partial\nu_q} G(\lambda_j, q, (x, y))\varphi_{\lambda_j}(q) d\sigma(q). \tag{14.142}$$

To obtain concrete representations we need to choose G. We choose the real ambient Euclidean Green's function S

$$S(\lambda_j, \xi, \eta; x, y) = -Y_0(\lambda_j r((x, y); (\xi, \eta))), \tag{14.143}$$

where $r = \sqrt{zz^*}$ is the distance function (the square root of r^2 above) and where Y_0 is the Bessel function of order zero of the second kind. The Euclidean Green's function has the form

$$S(\lambda_j, \xi, \eta; x, y) = A(\lambda_j, \xi, \eta; x, y) \log \frac{1}{r} + B(\lambda_j, \xi, \eta; x, y), \tag{14.144}$$

where A and B are entire functions of r^2. The coefficient $A = J_0(\lambda_j r)$ is known as the Riemann function.

By the 'jumps' formulae, the double layer potential $\frac{\partial}{\partial \nu_{\tilde{q}}} S(\lambda_j, \tilde{q}, (x, y))$ on $\partial\Omega \times \bar{\Omega}$ restricts to $\partial\Omega \times \partial\Omega$ as $\frac{1}{2}\delta_q(\tilde{q}) + \frac{\partial}{\partial \nu_{\tilde{q}}} S(\lambda_j, \tilde{q}, q)$ (see e.g. [**T1, T2**]). Hence in the Neumann case the boundary values v_{λ_j} of φ_{λ_j} satisfy

$$v_{\lambda_j}(q) = 2 \int_{\partial\Omega} \frac{\partial}{\partial \nu_{\tilde{q}}} S(\lambda_j, \tilde{q}, q) v_{\lambda_j}(\tilde{q}) \, d\sigma(\tilde{q}). \tag{14.145}$$

We have

$$\frac{\partial}{\partial \nu_{\tilde{q}}} S(\lambda_j, \tilde{q}, q) = -\lambda_j Y_1(\lambda_j r) \cos \angle(q - \tilde{q}, \nu_{\tilde{q}}). \tag{14.146}$$

It is equivalent, and sometimes more convenient, to use the (complex valued) Euclidean outgoing Green's function ${\rm Ha}^{(1)}{}_0(kz)$, where ${\rm Ha}^{(1)}{}_0 = J_0 + iY_0$ is the Hankel function of order zero. It has the same form as (14.144) and only differs by the addition of the even entire function J_0 to the B term. If we use the Hankel free outgoing Green's function, then in place of (14.146) we have the kernel

(14.147)
$$N(\lambda_j, q(s), q(s')) = \frac{i}{2} \partial_{\nu_y} {\rm Ha}^{(1)}{}_0(\lambda_j |q(s) - y|)|_{y = q(s')}$$

$$= -\frac{i}{2} \lambda_j {\rm Ha}^{(1)}{}_1(\lambda_j |q(s) - q(s')|) \cos \angle(q(s') - q(s), \nu_{q(s')}), \tag{14.148}$$

and in place of (14.145) we have the formula

$$v_{\lambda_j}(q(t)) = \int_0^{2\pi} N(\lambda_j, q(s), q(t)) \, v_{\lambda_j}(q(s)) \, ds. \tag{14.149}$$

The next step is to analytically continue the layer potential representations (14.145) and (14.149). The main point is to express the analytic continuations of Cauchy data of Neumann and Dirichlet eigenfunctions in terms of the real Cauchy data. For brevity, we only consider (14.145) but essentially the same arguments apply to the free outgoing representation (14.149).

As mentioned above, both $A(\lambda_j, \xi, \eta, x, y)$ and $B(\lambda_j, \xi, \eta, x, y)$ admit analytic continuations. In the case of A, we use a traditional notation $R(\zeta, \zeta^*, z, z^*)$ for the analytic continuation and for simplicity of notation we omit the dependence on λ_j.

The details of the analytic continuation are complicated when the curve is the boundary, and they simplify when the curve is interior. So we only continue the sketch of the proof in the interior case.

As above, the arc-length parametrization of C is denoted by by $q_C \colon [0, 2\pi] \to C$ and the corresponding arc-length parametrization of the boundary, $\partial\Omega$, by q :

$[0, 2\pi] \to \partial\Omega$. Since the boundary and C do not intersect, the logarithm $\log r^2(q(s); q_C^C(t))$ is well defined and the holomorphic continuation of equation (14.149) is given by:

$$\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t)) = \int_0^{2\pi} N(\lambda_j, q(s), q_C^{\mathbb{C}}(t))\varphi_{\lambda_j}(q(s))\, d\sigma(s), \tag{14.150}$$

From the basic formula (14.147) for $N(\lambda_j, q, q_C)$ and the standard integral formula for the Hankel function $\mathrm{Ha}^{(1)}{}_1(z)$, one easily gets an asymptotic expansion in λ_j of the form:

$$\begin{aligned} N(\lambda_j, q(s), q_C^{\mathbb{C}}(t)) = e^{i\lambda_j r(q(s); q_C^{\mathbb{C}}(t))} \sum_{m=0}^{k} a_m(q(s), q_C^{\mathbb{C}}(t))\lambda_j^{1/2-m} \\ + O(e^{i\lambda_j r(q(s); q_C^{\mathbb{C}}(t))}\lambda_j^{1/2-k-1}). \end{aligned} \tag{14.151}$$

Note that the expansion in (14.151) is valid since for interior curves,

$$C_0 := \min_{(q_C(t), q(s)) \in C \times \partial\Omega} |q_C(t) - q(s)|^2 > 0.$$

Then, $\mathrm{Re} r^2(q(s); q_C^C(t)) > 0$ as long as

$$|\mathrm{Im} q_C^{\mathbb{C}}(t)|^2 < C_0. \tag{14.152}$$

So, the principal square root of r^2 has a well-defined holomorphic extension to the tube (14.152) containing C. We have denoted this square root by r in (14.151).

Substituting (14.151) in the analytically continued single layer potential integral formula (14.150) proves that for $t \in A(\varepsilon)$ and $\lambda_j > 0$ sufficiently large,

(14.153)

$$\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t)) = 2\pi\lambda_j^{\frac{1}{2}} \int_0^{2\pi} e^{i\lambda_j r(q(s):q_C^{\mathbb{C}}(t))} a_0(q(s), q_C^{\mathbb{C}}(t))(1 + O(\lambda_j^{-1})\,)\varphi_{\lambda_j}(q(s))\, d\sigma(s).$$

Remark 14.52. We remark that up to this point, the approach works almost the same way in all dimensions, although the integral is over a higher dimensional boundary and the outgoing Green's function depends on the dimension.

Taking absolute values of the integral on the right-hand side in (14.153) and applying the Cauchy-Schwartz inequality proves

Lemma 14.53. *For $t \in [0, 2\pi] + i[-\varepsilon, \varepsilon]$ and $\lambda_j > 0$ sufficiently large*

$$|\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))| \leq C_1 \lambda_j^{1/2} \exp \lambda_j \left(\max_{q(s) \in \partial\Omega} \mathrm{Re} ir(q(s); q_C^{\mathbb{C}}(t)) \right) \cdot \|\varphi_{\lambda_j}\|_{L^2(\partial\Omega)}.$$

From the pointwise upper bounds in Lemma 14.53, it is immediate that

$$\log \max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon))} |\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))| \leq C_{\max}\lambda_j + C_2 \log \lambda_j + \log \|\varphi_{\lambda_j}\|_{L^2(\partial\Omega)}, \tag{14.154}$$

where

$$C_{\max} = \max_{(q(s), q_C^{\mathbb{C}}(t)) \in \partial\Omega \times Q_C^{\mathbb{C}}(A(\varepsilon))} \mathrm{Re}\, ir(q(s); q_C^{\mathbb{C}}(t)).$$

Finally, we use that $\log \|u_{\lambda_j}\|_{L^2(\partial\Omega)} = O(\lambda_j)$ by the assumption that C is a good curve and apply Proposition 14.55 to get that $n(\lambda_j, C) = O(\lambda_j)$.

The following estimate, suggested by [**DF**, Lemma 6.1], gives an upper bound on the number of zeros in terms of the growth of the family.

LEMMA 14.54. *Let $F(z)$ be holomorphic in an open neighborhood of $|z| \leq 1$. Assume that $|F(z)| \leq 1$. Let $\ell = \#\{\alpha : |\alpha| \leq \frac{1}{2}, F(\alpha) = 0\}$. Then:*

$$\ell \leq C_1 \left| \log \max_{|z| \leq \frac{1}{2}} |F(z)| \right|.$$

PROOF. Represent F as a Blashcke product

$$F(z) = G(z) \prod_{j=1}^{\ell} \frac{z - \alpha_j}{1 - \bar{\alpha}_j z}, \qquad |z| \leq 1.$$

By the maximum principle $|G(z)| \leq 1$. Hence for $|z| \leq \frac{1}{2}$ there exists $C_2 < 1$ so that $|F(z)| \leq C_2^{\ell}$. □

In the above proof, $\max_{|z| \leq \frac{1}{2}} |F(z)| \leq C_2^{\ell}$ and $\log \max_{|z| \leq \frac{1}{2}} |F(z)| < 0$. In the following modification, the $\log \max$ is necessarily ≥ 0. To state the next result we need some new notation. We define a pair $(C, \mathcal{S})$ consisting of an irreducible real analytic curve C and a subsequence $\mathcal{S} \subset \mathbb{N}$ to be *good* if the sequence of normalized logarithms

$$u_{\lambda_j} := \frac{1}{\lambda_j} \log |\varphi_{\lambda_j}|^2 \tag{14.155}$$

restricted to C,

$$u_j^C := \gamma_C u_j := \frac{1}{\lambda_j} \log |\varphi_j^C|^2, \tag{14.156}$$

does *not* tend to $-\infty$ uniformly on C as $j \to \infty$. If $\mathcal{S} = \mathbb{N}$ we say more simply that the curve is good.

PROPOSITION 14.55. *Suppose that C is a good real analytic curve. Normalize u_{λ_j} so that $\|\varphi_{\lambda_j}\|_{L^2(C)} = 1$. Then, there exists a constant $C(\varepsilon) > 0$ such that for any $\varepsilon > 0$,*

$$n(\lambda_j, Q_C^{\mathbb{C}}(A(\varepsilon/2))) \leq C(\varepsilon) \max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon))} \log |\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))|.$$

PROOF. Let G_ε denote the Dirichlet Green's function of the 'annulus' $Q_C^{\mathbb{C}}(A(\varepsilon))$. Also, let $\{a_k\}_{k=1}^{n(\lambda_j, Q_C^{\mathbb{C}}(A(\varepsilon/2)))}$ denote the zeros of $u_{\lambda_j}^{\mathbb{C}}$ in the sub-annulus $Q_C^{\mathbb{C}}(A(\varepsilon/2))$. Let

$$\Phi_{\lambda_j} = \frac{\varphi_{\lambda_j}^{\mathbb{C}}}{\|\varphi_{\lambda_j}^{\mathbb{C}}\|_{Q_C^{\mathbb{C}}(A(\varepsilon))}} \text{ where } \|f\|_{Q_C^{\mathbb{C}}(A(\varepsilon))} = \max_{\zeta \in Q_C^{\mathbb{C}}(A(\varepsilon))} |f(\zeta)|.$$

Then,

(14.157)

$$\log |\Phi_{\lambda_j}(q_C^{\mathbb{C}}(t))| = \int_{Q_C^{\mathbb{C}}((A(\varepsilon/2)))} G_\varepsilon(q_C^{\mathbb{C}}(t), w) \partial\bar{\partial} \log |\varphi_{\lambda_j}^{\mathbb{C}}(w)| + H_{\lambda_j}(q_C^{\mathbb{C}}(t))$$

$$= \sum_{a_k \in Q_C^{\mathbb{C}}(A(\varepsilon/2)) : \varphi_{\lambda_j}^{\mathbb{C}}(a_k) = 0} G_\varepsilon(q_C^{\mathbb{C}}(t), a_k) + H_{\lambda_j}(q_C^{\mathbb{C}}(t)), \tag{14.158}$$

since $\partial\bar\partial \log|\varphi_{\lambda_j}^{\mathbb{C}}(w)| = \sum_{a_k \in C_{\mathbb{C}}: \varphi_{\lambda_j}^{\mathbb{C}}(a_k)=0} \delta_{a_k}$. Moreover, the function H_{λ_j} is subharmonic on $Q_C^{\mathbb{C}}(A(\varepsilon))$ since

$$\partial\bar\partial H_{\lambda_j} = \partial\bar\partial \log|\Phi_{\lambda_j}(q_C^{\mathbb{C}}(t))| - \sum_{a_k \in Q_C^{\mathbb{C}}(A(\varepsilon/2)): \varphi_{\lambda_j}^{\mathbb{C}}(a_k)=0} \partial\bar\partial G_\varepsilon(q_C^{\mathbb{C}}(t), a_k) \tag{14.159}$$

$$= \sum_{a_k \in Q_C^{\mathbb{C}}(A(\varepsilon)) \setminus Q_C^{\mathbb{C}}(A(\varepsilon/2))} \delta_{a_k} > 0. \tag{14.160}$$

So, by the maximum principle for subharmonic functions,

$$\max_{Q_C^{\mathbb{C}}(A(\varepsilon))} H_{\lambda_j}(q_C^{\mathbb{C}}(t)) \leq \max_{\partial Q_C^{\mathbb{C}}(A(\varepsilon))} H_{\lambda_j}(q_C^{\mathbb{C}}(t)) = \max_{\partial Q_C^{\mathbb{C}}(A(\varepsilon))} \log|\Phi_{\lambda_j}(q_C^{\mathbb{C}}(t))| = 0.$$

It follows that

$$\log|\Phi_{\lambda_j}(q_C^{\mathbb{C}}(t))| \leq \sum_{a_k \in Q_C^{\mathbb{C}}(A(\varepsilon/2)): \varphi_{\lambda_j}^{\mathbb{C}}(a_k)=0} G_\varepsilon(q_C^{\mathbb{C}}(t), a_k), \tag{14.161}$$

hence that

$$\max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} \log|\Phi_{\lambda_j}(q_C^{\mathbb{C}}(t))| \leq \left(\max_{z,w \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} G_\varepsilon(z,w)\right) n(\lambda_j, Q_C^{\mathbb{C}}(A(\varepsilon/2))). \tag{14.162}$$

Now

$$\begin{cases} G_\varepsilon(z,w) \leq \max_{w \in Q_C^{\mathbb{C}}(\partial A(\varepsilon))} G_\varepsilon(z,w) = 0, \\ G_\varepsilon(z,w) < 0 \text{ for } z,w \in Q_C^{\mathbb{C}}(A(\varepsilon/2)). \end{cases} \tag{14.163}$$

It follows that there exists a constant $\nu(\varepsilon) < 0$ so that $\max_{z,w \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} G_\varepsilon(z,w) \leq \nu(\varepsilon)$. Hence,

$$\max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} \log|\Phi_{\lambda_j}(Q_C^{\mathbb{C}}(t))| \leq \nu(\varepsilon)\ \ n(\lambda_j, Q_C^{\mathbb{C}}(A(\varepsilon/2))). \tag{14.164}$$

Since both sides are negative, we obtain

$$n(\lambda_j, Q_C^{\mathbb{C}}(A(\varepsilon/2))) \leq \frac{1}{|\nu(\varepsilon)|} \left| \max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} \log|\Phi_{\lambda_j}(q_C^{\mathbb{C}}(t))| \right| \tag{14.165}$$

$$\leq \frac{1}{|\nu(\varepsilon)|} \left(\max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon))} \log|\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))| \right. \tag{14.166}$$

$$\left. - \max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} \log|\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))| \right) \tag{14.167}$$

$$\leq \frac{1}{|\nu(\varepsilon)|} \max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon))} \log|\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))|, \tag{14.168}$$

where in the last step we use that $\max_{q_C^{\mathbb{C}}(t) \in Q_C^{\mathbb{C}}(A(\varepsilon/2))} \log|\varphi_{\lambda_j}^{\mathbb{C}}(q_C^{\mathbb{C}}(t))| \geq 0$, which holds since $|\varphi_{\lambda_j}^{\mathbb{C}}| \geq 1$ at some point in $Q_C^{\mathbb{C}}(A(\varepsilon/2))$. Indeed, by our normalization, $\|\varphi_{\lambda_j}\|_{L^2(C)} = 1$, and so there must already exist points on the real curve C with $|\varphi_{\lambda_j}| \geq 1$. Putting $C(\varepsilon) = \frac{1}{|\nu(\varepsilon)|}$ finishes the proof. □

This completes the proof of Theorem 14.51.

14.44. Application to Pleijel's conjecture

I. Polterovich [**Po**] observed that Theorem 14.50 can be used to prove an old conjecture of A. Pleijel regarding Courant's nodal domain theorem, which says that the number n_k of nodal domains (components of $\Omega \backslash Z_{\varphi_{\lambda_k}}$) of the kth eigenfunction satisfies $n_k \leq k$. Pleijel improved this result for Dirichlet eigenfunctions of plane domains: For any plane domain with Dirichlet boundary conditions, $\limsup_{k\to\infty} \frac{n_k}{k} \leq \frac{4}{j_1^2} \simeq 0.691...$, where j_1 is the first zero of the J_0 Bessel function. He conjectured that the same result should be true for a free membrane, i.e. for Neumann boundary conditions. This was recently proved in the real analytic case by I. Polterovich [**Po**]. His argument is roughly the following: Pleijel's original argument applies to all nodal domains which do not touch the boundary, since the eigenfunction is a Dirichlet eigenfunction in such a nodal domain. The argument does not apply to nodal domains which touch the boundary, but by the Theorem above the number of such domains is negligible for the Pleijel bound.

14.45. Equidistribution of intersections of nodal lines and geodesics on surfaces

To understand the relation between real and complex zeros, we intersect nodal lines and geodesics on surfaces $\dim M = 2$. This section is based on [**Z7**].

We fix $(x, \xi) \in S^*M$ and let

$$\gamma_{x,\xi} \colon \mathbb{R} \to M, \quad \gamma_{x,\xi}(0) = x, \quad \gamma_{x,\xi}'(0) = \xi \in T_xM \tag{14.169}$$

denote the corresponding parametrized geodesic. Our goal is to determine the asymptotic distribution of intersection points of $\gamma_{x,\xi}$ with the nodal set of a highly eigenfunction. As usual, we cannot cope with this problem in the real domain and therefore analytically continue it to the complex domain. Thus, we consider the intersections

$$\mathcal{N}_{\lambda_j}^{\gamma_{x,\xi}^{\mathbb{C}}} = Z_{\varphi_j^{\mathbb{C}}} \cap \gamma_{x,\xi}^{\mathbb{C}}$$

of the complex nodal set with the (image of the) complexification of a generic geodesic If

$$S_\varepsilon = \{(t + i\tau \in \mathbb{C} : |\tau| \leq \varepsilon\} \tag{14.170}$$

then $\gamma_{x,\xi}$ admits an analytic continuation

$$\gamma_{x,\xi}^{\mathbb{C}} : S_\varepsilon \to M_\varepsilon. \tag{14.171}$$

In other words, we consider the zeros of the pullback,

$$\{\gamma_{x,\xi}^* \varphi_\lambda^{\mathbb{C}} = 0\} \subset S_\varepsilon.$$

We encode the discrete set by the measure

$$[\mathcal{N}_{\lambda_j}^{\gamma_{x,\xi}^{\mathbb{C}}}] = \sum_{(t+i\tau) : \varphi_j^{\mathbb{C}}(\gamma_{x,\xi}^{\mathbb{C}}(t+i\tau))=0} \delta_{t+i\tau}. \tag{14.172}$$

We would like to show that for generic geodesics, the complex zeros on the complexified geodesic condense on the real points and become uniformly distributed with respect to arc-length. This does not always occur: as in our discussion of QER theorems, if $\gamma_{x,\xi}$ is the fixed point set of an isometric involution, then "odd" eigenfunctions under the involution will vanish on the geodesic. The additional hypothesis is that QER holds for $\gamma_{x,\xi}$. The following is proved in [**Z7**]:

THEOREM 14.56. *Let (M^2, g) be a real analytic Riemannian surface with ergodic geodesic flow. Let $\gamma_{x,\xi}$ satisfy the QER hypothesis. Then there exists a subsequence of eigenvalues λ_{j_k} of density one such that for any $f \in C_c(S_\varepsilon)$,*

$$\lim_{k\to\infty} \sum_{(t+i\tau):\varphi_j^{\mathbb{C}}(\gamma_{x,\xi}^{\mathbb{C}}(t+i\tau))=0} f(t+i\tau) = \int_{\mathbb{R}} f(t)\,dt.$$

In other words,

$$\text{weak}^* \lim_{k\to\infty} \frac{i}{\pi\lambda_{j_k}}[\mathcal{N}_{\lambda_j}^{\gamma_{x,\xi}^{\mathbb{C}}}] = \delta_{\tau=0},$$

in the sense of weak* convergence on $C_c(S_\varepsilon)$. Thus, the complex nodal set intersects the (parametrized) complexified geodesic in a discrete set which is asymptotically (as $\lambda \to \infty$) concentrated along the real geodesic with respect to its arc length.

This concentration- equidistribution result is a 'restricted' version of the result of §14.37. As noted there, the limit distribution of complex nodal sets in the ergodic case is a singular current $dd^c\sqrt{\rho}$. The motivation for restricting to geodesics is that restriction magnifies the singularity of this current. In the case of a geodesic, the singularity is magnified to a delta-function; for other curves there is additionally a smooth background measure.

The assumption of ergodicity is crucial. For instance, in the case of a flat torus, say $\mathbb{R}^2/L$ where $L \subset \mathbb{R}^2$ is a generic lattice, the real eigenfunctions are $\cos\langle\lambda, x\rangle, \sin\langle\lambda, x\rangle$ where $\lambda \in L^*$, the dual lattice, with eigenvalue $-|\lambda|^2$. Consider a geodesic $\gamma_{x,\xi}(t) = x + t\xi$. Due to the flatness, the restriction $\sin\langle\lambda, x_0 + t\xi_0\rangle$ of the eigenfunction to a geodesic is an eigenfunction of the Laplacian $-\frac{d^2}{dt^2}$ of submanifold metric along the geodesic with eigenvalue $-\langle\lambda, \xi_0\rangle^2$. The complexification of the restricted eigenfunction is $\sin\langle\lambda, x_0 + (t+i\tau)\xi_0\rangle|$ and its exponent of its growth is $\tau|\langle\frac{\lambda}{|\lambda|}, \xi_0\rangle|$, which can have a wide range of values as the eigenvalue moves along different rays in L^*. The limit current is $i\partial\bar{\partial}$ applied to the limit and thus also has many limits

The proof involves several new principles which played no role in the global result of §14.37 and which are specific to geodesics. However, the first steps in the proof are the same as in the global case. By the Poincaré-Lelong formula (14.82), we may express the current of summation over the intersection points in (14.172) in the form,

$$[\mathcal{N}_{\lambda_j}^{\gamma_{x,\xi}^{\mathbb{C}}}] = i\partial\bar{\partial}_{t+i\tau} \log\left|\gamma_{x,\xi}^*\varphi_{\lambda_j}^{\mathbb{C}}(t+i\tau)\right|^2. \tag{14.173}$$

Thus, the main point of the proof is to determine the asymptotics of $\frac{1}{\lambda_j}\log\left|\gamma_{x,\xi}^*\varphi_{\lambda_j}^{\mathbb{C}}(t+i\tau)\right|^2$. When we freeze τ we put

$$\gamma_{x,\xi}^{\tau}(t) = \gamma_{x,\xi}^{\mathbb{C}}(t+i\tau). \tag{14.174}$$

PROPOSITION 14.57. *(Growth saturation) If $\{\varphi_{j_k}\}$ satisfies QER along any arcs of $\gamma_{x,\xi}$, then in $L^1_{loc}(S_\tau)$, we have*

$$\lim_{k\to\infty}\frac{1}{\lambda_{j_k}}\log\left|\gamma_{x,\xi}^{\tau*}\varphi_{\lambda_{j_k}}^{\mathbb{C}}(t+i\tau)\right|^2 = |\tau|.$$

Proposition 14.57 immediately implies Theorem 14.56 since we can apply $\partial\bar{\partial}$ to the L^1 convergent sequence $\frac{1}{\lambda_{j_k}}\log\left|\gamma_{x,\xi}^*\varphi_{\lambda_{j_k}}^{\mathbb{C}}(t+i\tau)\right|^2$ to obtain $\partial\bar{\partial}|\tau|$.

The upper bound in Proposition 14.57 follows immediately from the known global estimate

$$\lim_{k\to\infty} \frac{1}{\lambda_j} \log |\varphi_{j_k}(\gamma_{x,\xi}^{\mathbb{C}}(\zeta)| \le |\tau|$$

on all of ∂M_τ. Hence the difficult point is to prove that this growth rate is actually obtained upon restriction to $\gamma_{x,\xi}^{\mathbb{C}}$. This requires new kinds of arguments related to the QER theorem.

- Complexifications of restrictions of eigenfunctions to geodesics have incommensurate Fourier modes, i.e. higher modes are exponentially larger than lower modes.
- The quantum ergodic restriction theorem in the real domain shows that the Fourier coefficients of the top allowed modes are 'large' (i.e., as large as the lower modes). Consequently, the L^2 norms of the complexified eigenfunctions along arcs of $\gamma_{x,\xi}^{\mathbb{C}}$ achieve the lower bound of Proposition 14.57.
- Invariance of Wigner measures along the geodesic flow implies that the Wigner measures of restrictions of complexified eigenfunctions to complexified geodesics should tend to constant multiples of Lebesgue measures dt for each $\tau > 0$. Hence the eigenfunctions everywhere on $\gamma_{x,\xi}^{\mathbb{C}}$ achieve the growth rate of the L^2 norms.

These principles are most easily understood in the case of periodic geodesics. We let $\gamma_{x,\xi} : S^1 \to M$ parametrize the geodesic with arc-length (where $S^1 = \mathbb{R}/L\mathbb{Z}$ where L is the length of $\gamma_{x,\xi}$).

LEMMA 14.58. *Assume that $\{\varphi_j\}$ satisfies QER along the periodic geodesic $\gamma_{x,\xi}$. Let $\|\gamma_{x,\xi}^{\tau*}\varphi_j^{\mathbb{C}}\|^2_{L^2(S^1)}$ be the L^2-norm of the complexified restriction of φ_j along $\gamma_{x,\xi}^{\tau}$. Then,*

$$\lim_{\lambda_j\to\infty} \frac{1}{\lambda_j} \log \|\gamma_{x,\xi}^{\tau*}\varphi_j^{\mathbb{C}}\|^2_{L^2(S^1)} = |\tau|.$$

To prove Lemma 14.58, we study the orbital Fourier series of $\gamma_{x,\xi}^{\tau*}\varphi_j$ and of its complexification. The orbital Fourier coefficients are

$$\nu_{\lambda_j}^{x,\xi}(n) = \frac{1}{L_\gamma} \int_0^{L_\gamma} \varphi_{\lambda_j}(\gamma_{x,\xi}(t)) e^{-\frac{2\pi i n t}{L_\gamma}}\, dt,$$

and the orbital Fourier series is

$$\varphi_{\lambda_j}(\gamma_{x,\xi}(t)) = \sum_{n\in\mathbb{Z}} \nu_{\lambda_j}^{x,\xi}(n) e^{\frac{2\pi i n t}{L_\gamma}}. \tag{14.175}$$

Hence the analytic continuation of $\gamma_{x,\xi}^{\tau*}\varphi_j$ is given by

$$\varphi_{\lambda_j}^{\mathbb{C}}(\gamma_{x,\xi}(t+i\tau)) = \sum_{n\in\mathbb{Z}} \nu_{\lambda_j}^{x,\xi}(n) e^{\frac{2\pi i n (t+i\tau)}{L_\gamma}}. \tag{14.176}$$

By the Paley-Wiener theorem for Fourier series, the series converges absolutely and uniformly for $|\tau| \le \varepsilon_0$. By "energy localization" only the modes with $|n| \le \lambda_j$ contribute substantially to the L^2 norm. We then observe that the Fourier modes decouple, since they have different exponential growth rates. We use the QER hypothesis in the following way:

Lemma 14.59. *Suppose that $\{\varphi_{\lambda_j}\}$ is QER along the periodic geodesic $\gamma_{x,\xi}$. Then for all $\varepsilon > 0$, there exists $C_\varepsilon > 0$ so that*

$$\sum_{n:|n|\geq(1-\varepsilon)\lambda_j} |\nu^{x,\xi}_{\lambda_j}(n)|^2 \geq C_\varepsilon.$$

Lemma 14.59 implies Lemma 14.58 since it implies that for any $\varepsilon > 0$,

$$\sum_{n:|n|\geq(1-\varepsilon)\lambda_j} |\nu^{x,\xi}_{\lambda_j}(n)|^2 e^{-2n\tau} \geq C_\varepsilon e^{2\tau(1-\varepsilon)\lambda_j}.$$

To go from asymptotics of L^2 norms of restrictions to Proposition 14.57 we then use the third principle:

Proposition 14.60. *(Lebesgue limits) If $\gamma^*_{x,\xi}\varphi_j \neq 0$ (identically), then for all $\tau > 0$ the sequence*

$$U_j^{x,\xi,\tau} = \frac{\gamma^{\tau *}_{x,\xi}\varphi_j^{\mathbb{C}}}{\|\gamma^{\tau *}_{x,\xi}\varphi_j^{\mathbb{C}}\|_{L^2(S^1)}}$$

is QUE with limit measure given by normalized Lebesgue measure on S^1.

The proof of Proposition 14.57 is completed by combining Lemma 14.58 and Proposition 14.60. Theorem 14.56 follows easily from Proposition 14.57.

The proof for non-periodic geodesics is considerably more involved, since one cannot use Fourier analysis in quite the same way.

Bibliography

[AP1] J. C. Alvarez Paiva and E. Fernandes, Gelfand transforms and Crofton formulas. Selecta Math. (N.S.) 13 (2007), no. 3, 369–390.

[AP2] J. C. Alvarez Paiva and G. Berck, What is wrong with the Hausdorff measure in Finsler spaces. Adv. Math. 204 (2006), no. 2, 647663.

[BG1] M.S. Baouendi and C. Goulaouic, Régularité analytique et itérés d'opérateurs elliptiques dégénérés; applications. J. Functional Analysis 9 (1972), 208–248.

[BG2] M.S. Baouendi and C. Goulaouic, Itérés d'opérateurs elliptiques et prolongement de fonctions propres. Hommage au Professeur Miron Nicolescu pour son 70e anniversaire, I. Rev. Roumaine Math. Pures Appl. 18 (1973), 1495–1501.

[BG3] M.S. Baouendi and C. Goulaouic, Approximation of analytic functions on compact sets and Bernstein's inequality. Trans. Amer. Math. Soc. 189 (1974), 251–261.

[BG4] M.S. Baouendi and C. Goulaouic, Cauchy problem for analytic pseudo-differential operators. Comm. Partial Differential Equations 1 (1976), no. 2, 135-189.

[BGH] M. S. Baouendi, C. Goulaouic, and B. Hanouzet, Caractérisation de classes de fonctions C^∞ et analytiques sur une variété irrégulière à l'aide d'un opérateur différentiel. J. Math. Pures Appl. (9) 52 (1973), 115–144.

[Bae] C. Bär, On nodal sets for Dirac and Laplace operators. Comm. Math. Phys. 188 (1997), no. 3, 709–721.

[BT] E. Bedford and B. A. Taylor, A new capacity for plurisubharmonic functions. Acta Math. 149 (1982), no. 1-2, 1-40.

[Be] P. Bérard, On the wave equation without conjugate points, Math. Zeit. **155** (1977), 249–276.

[Bers] L. Bers, Local behavior of solutions of general linear elliptic equations. Comm. Pure Appl. Math. 8 (1955), 473–496.

[Bo] L. Boutet de Monvel, Convergence dans le domaine complexe des séries de fonctions propres. C. R. Acad. Sci. Paris Sér. A-B 287 (1978), no. 13, A855–A856.

[BoGu] L. Boutet de Monvel and V. Guillemin, *The Spectral Theory of Toeplitz Operators*, Ann. Math. Studies 99, Princeton Univ. Press, Princeton, 1981.

[BoSj] L. Boutet de Monvel and J. Sjöstrand, Sur la singularité des noyaux de Bergman et de Szegö, *Asterisque* 34–35 (1976), 123–164.

[BW] F. Bruhat and H. Whitney, Quelques propriétés fondamentales des ensembles analytiques-réels. Comment. Math. Helv. 33 1959 132–160.

[CV1] Y. Colin de Verdiére, Ergodicité et fonctions propres du laplacien. Comm. Math. Phys. 102 (1985), no. 3, 497–502.

[CV2] Y.Colin de Verdiere, Quasi-modes sur les varietes Riemanniennes compactes, Invent.Math. 43 (1977), 15-52.

[DSj] M. Dimassi and J. Sjöstrand, *Spectral asymptotics in the semi-classical limit.* London Mathematical Society Lecture Note Series, 268. Cambridge University Press, Cambridge, 1999.

[DF] H. Donnelly and C. Fefferman, Nodal sets of eigenfunctions on Riemannian manifolds. Invent. Math. 93 (1988), no. 1, 161-183.

[Du] J. J. Duistermaat, Fourier integral operators. Reprint of the 1996 original. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2011.

[DuG] J.J.Duistermaat and V.Guillemin, The spectrum of positive elliptic operators and periodic bicharacteristics, Inv.Math. 24 (1975), 39-80.

[DuH] J. J. Duistermaat and L. Hörmander, Fourier integral operators. II. Acta Math. 128 (1972), no. 3-4, 183-269.

[ElT] L. El-Hajj, J. A. Toth Intersection bounds for nodal sets of planar Neumann eigenfunctions with interior analytic curves, J. Diff. Geom. 100 (2015), no. 1, 1-53 (arXiv:1211.3395).

[F1] H. Federer, Geometric measure theory. Die Grundlehren der mathematischen Wissenschaften, Band 153 Springer-Verlag New York Inc., New York 1969.

[F2] H. Federer, Slices and potentials. Indiana Univ. Math. J. 21 (1971/72), 373-382.

[GeSh] I. M. Gel'fand and G.E. Shilov,*Generalized functions. Vol. I: Properties and operations.* Academic Press, New York-London 1964.

[GeS] I. M. Gelfand and M. Smirnov, Lagrangians satisfying Crofton formulas, Radon transforms, and nonlocal differentials. Adv. Math. 109 (1994), 188–227.

[GLS] F. Golse, E. Leichtnam, and M. Stenzel, Intrinsic microlocal analysis and inversion formulae for the heat equation on compact real-analytic Riemannian manifolds. Ann. Sci. École Norm. Sup. (4) 29 (1996), no. 6, 669–736.

[GH] P. Griffiths and J. Harris, *Principles of Algebraic Geometry*, Wiley-Interscience, New York, 1978.

[GrSj] A. Grigis and J. Sjöstrand, *Microlocal analysis for differential operators. An introduction.* London Mathematical Society Lecture Note Series, 196. Cambridge University Press, Cambridge, 1994.

[GZ] V. Guedj and A. Zeriahi, Intrinsic capacities on compact Kähler manifolds. J. Geom. Anal. 15 (2005), no. 4, 607–639.

[G] V. Guillemin, Toeplitz operators in n dimensions. Integral Equations Operator Theory 7 (1984), no. 2, 145205.

[GS1] V. Guillemin and M. Stenzel, Grauert tubes and the homogeneous Monge-Ampère equation. J. Differential Geom. 34 (1991), no. 2, 561–570.

[GS2] V. Guillemin and M. Stenzel, Grauert tubes and the homogeneous Monge-Ampère equation. II. J. Differential Geom. 35 (1992), no. 3, 627–641.

[GSt1] V. Guillemin and S. Sternberg, Geometric asymptotics. Mathematical Surveys, No. 14. American Mathematical Society, Providence, R.I., 1977.

[GSt2] V. Guillemin and S. Sternberg, *Semi-classical Analysis* International Press, Boston, MA, 2013.

[H] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations.* Dover Publications, New York, 1953.

[HL] Q. Han and F. H. Lin: *Nodal sets of solutions of Elliptic Differential Equations*, book in preparation (online at http://www.nd.edu/qhan/nodal.pdf).

[HaS] R. Hardt and L. Simon, Nodal sets for solutions of elliptic equations. J. Differential Geom. 30 (1989), no. 2, 505–522.

[HW1] P. Hartman and A. Wintner, On the local behavior of solutions of non-parabolic partial differential equations. Amer. J. Math. 75, (1953). 449–476.

[HW2] P. Hartman and W. Wintner, On the local behavior of solutions of non-parabolic partial differential equations. III. Approximations by spherical harmonics. Amer. J. Math. 77 (1955), 453–474.

[HK] B. Hasselblatt and A. Katok, *A first course in dynamics. With a panorama of recent developments.* Cambridge University Press, New York, 2003

[HoI] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Classics in Mathematics. Springer-Verlag, Berlin, 2003.

[HoII] L. Hörmander, *The analysis of linear partial differential operators. II. Differential operators with constant coefficients.* Classics in Mathematics. Springer-Verlag, Berlin, 2005.

[HoIII] L. Hörmander, *The analysis of linear partial differential operators. III. Pseudo-differential operators.* Classics in Mathematics. Springer, Berlin, 2007.

[HoIV] L. Hörmander, *The analysis of linear partial differential operators.* IV. Fourier integral operators. Reprint of the 1994 edition. Classics in Mathematics. Springer-Verlag, Berlin, 2009.

[HoC] L. Hörmander, *Notions of convexity.* . Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA, 2007.

[Hu] K. Husimi, Some formal properties of the density matrix Proc. Phys. Math. Soc. Japan 22 (1940), 264314.

[JZ] D. Jakobson and S. Zelditch, Classical limits of eigenfunctions for some completely integrable systems. Emerging applications of number theory (Minneapolis, MN, 1996), 329-354, IMA Vol. Math. Appl., 109, Springer, New York, 1999.

[Jin] Long Jin, Semiclassical Cauchy Estimates and Applications, Trans. Amer. Math. Soc. 369 (2017), no. 2, 975-995 (arXiv:1302.5363).

[JJ1] J. Jung, Sharp bounds for the intersection of nodal lines with certain curves. J. Eur. Math. Soc. (JEMS) 16 (2014), no. 2, 273-288 (arXiv: 1108.2335).

[JJ2] J. Jung, Quantitative QUE and the nodal domains of Maass-Hecke eigenforms, Comm. Math. Phys. 348 (2016), no. 2, 603-653 (arXiv:1301.6211).

[KS] B. Kroetz and H. Schlichtkrull, Holomorphic extension of eigenfunctions. Math. Ann. 345 (2009), no. 4, 835-841.

[Kua] I. Kukavica, Nodal volumes for eigenfunctions of analytic regular elliptic problems. J. Anal. Math. 67 (1995), 269–280.

[L] G. Lebeau, The complex Poisson kernel on a compact analytic Riemannian manifold, preprint (2013).

[LS] L. Lempert and R. Szöke, Global solutions of the homogeneous complex Monge-Ampère equation and complex structures on the tangent bundle of Riemannian manifolds. Math. Ann. 290 (1991), no. 4, 689–712.

[Lin] F.H. Lin, Nodal sets of solutions of elliptic and parabolic equations. Comm. Pure Appl. Math. 44 (1991), no. 3, 287–308.

[LM] J. L. Lions and E. Magenes, Non-homogeneous boundary value problems and applications. Vol. III. Translated from the French by P. Kenneth. Die Grundlehren der mathematischen Wissenschaften, Band 183. Springer-Verlag, New York-Heidelberg, 1973.

[Mar] S. Marshall, Geodesic restrictions of eigenfunctions on arithmetic surfaces, Duke Math. J. 165 (2016), no. 3, 463-508 (arXiv:1204.0781).

[MeSj] A. Melin and J. Sjöstrand, Fourier integral operators with complex-valued phase functions.*Fourier integral operators and partial differential equations* (Colloq. Internat., Univ. Nice, Nice, 1974), pp. 120-223. Lecture Notes in Math., Vol. 459, Springer, Berlin, 1975.

[Po] I. Polterovich, Pleijel's nodal domain theorem for free membranes, Proc. Amer. Math. Soc. 137 (2009), no. 3, 10211024 (arXiv:0805.1553).

[S] Y. Safarov, Asymptotics of a spectral function of a positive elliptic operator without a nontrapping condition. (Russian) Funktsional. Anal. i Prilozhen. 22 (1988), no. 3,53–65, 96 translation in Funct. Anal. Appl. 22 (1988), no. 3, 213–223 (1989).

[SV] Y. Safarov and D. Vassiliev, *The asymptotic distribution of eigenvalues of partial differential operators. Translated from the Russian manuscript by the authors.* Translations of Mathematical Monographs, 155. American Mathematical Society, Providence, RI, 1997.

[Si] J. Siciak, On some extremal functions and their applications in the theory of analytic functions of several complex variables. Trans. Amer. Math. Soc. 105 1962 322-357.

[Sj] J. Sjöstrand, *Singularités analytiques microlocales.* Astérisque, 95, 1166, Astérisque, 95, Soc. Math. France, Paris, 1982.

[St] M. Stenzel, On the analytic continuation of the Poisson kernel. Manuscripta Math. 144 (2014), no. 1-2, 253-276.

[SHM] RM Stratt, NC Handy, WH Miller, On the quantum mechanical implications of classical ergodicity, The Journal of Chemical Physics, 1979.

[T1] M. E. Taylor, *Partial differential equations I. Basic theory.* Second edition. Applied Mathematical Sciences, 115. Springer, New York, 2011.

[T2] M. E. Taylor, *Partial differential equations II. Qualitative studies of linear equations.* Second edition. Applied Mathematical Sciences, 116. Springer, New York, 2011.

[Ti] A. F. Timan, *Theory of Approximation of Functions of Real Variable*, Pergamon Press, Oxford (1963).

[TZ1] J. A. Toth and S. Zelditch, Counting Nodal Lines Which Touch the Boundary of an Analytic Domain, Jour. Diff. Geom. 81 (2009), 649- 686 (arXiv:0710.0101).

[TZ2] J. A. Toth and S. Zelditch, Quantum ergodic restriction theorems, I: interior hypersurfaces in analytic domains,, *Ann. H. Poincaré* 13, Issue 4 (2012), Page 599-670 (arXiv:1005.1636).

[TZ3] J. A. Toth and S. Zelditch, Quantum ergodic restriction theorems, II: manifolds without boundary, Geom. Funct. Anal. 23 (2013), no. 2, 715-775 (arXiv:1104.4531).

[TZ4] J. A. Toth and S. Zelditch, Upper bound on the number of nodal points on curves, in preparation.

[U] K. Uhlenbeck, Generic properties of eigenfunctions. Amer. J. Math. 98 (1976), no. 4, 1059–1078.

[Z1] S. Zelditch, Uniform distribution of eigenfunctions on compact hyperbolic surfaces. Duke Math. J. 55 (1987), no. 4, 919–941.

[Z2] S. Zelditch,, Complex zeros of real ergodic eigenfunctions. Invent. Math. 167 (2007), no. 2, 419–443.

[Z3] S. Zelditch, Local and global analysis of eigenfunctions on Riemannian manifolds. Handbook of geometric analysis. No. 1, 545-658, Adv. Lect. Math. (ALM), 7, Int. Press, Somerville, MA, 2008.

[Z4] S. Zelditch, Recent developments in mathematical quantum chaos. Current developments in mathematics, 2009, 115-204, Int. Press, Somerville, MA, 2010.

[Z5] S. Zelditch, Pluri-potential theory on Grauert tubes of real analytic Riemannian manifolds, I : AMS Proc. Symp. Pure Math. :Spectral Geometry, Alex H. Barnett, Carolyn S. Gordon, Peter A. Perry, and Alejandro Uribe, (eds), (2012), p. 299-339. (arXiv: arXiv:1107.0463).

[Z6] S. Zelditch, Eigenfunctions and Nodal sets, Surveys in Differential Geometry volume 18 (2013), pp. 237-308 (arXiv:1205.2812.)

[Z7] S. Zelditch, Ergodicity and intersections of nodal sets and eigenfunctions on real analytic surfaces, J. Diff. Geom. 96 (2014), no. 2, 305-351 (arXiv:1210.0834).

[Z8] S. Zelditch, Hausdorff measure of nodal sets of analytic Steklov eigenfunctions. Math. Res. Lett. 22 (2015), no. 6, 1821-1842.

[Z9] S. Zelditch, Pointwise phase space Weyl laws on Grauert tubes (in preparation).

[Z10] S. Zelditch, Complex zeros of quantum integrable eigenfunctions, (in preparation).

[Zer] M. Zerner, Domain d'holomorphie des fonctions vérifiant une équation aux dérivées partielles, C. R. Acad. Sci. Paris Ser. I Math. 272 (1971) 1646-1648.

[Zw] M. Zworski, *Semiclassical analysis*. Graduate Studies in Mathematics, 138. American Mathematical Society, Providence, RI, 2012.

Index