Problems

Xiaolong Han

Contents

List of problems

List of symbols

Points and sets

$\mathbb{N}$	the set of natural numbers
$\mathbb{Z}$	the set of integers
$\mathbb{Q}$	the set of rational numbers
$\mathbb{R}^n$	the n-dim Euclidean space
$\mathbb{C}^n$	the n-dim complex space
$\mathbb{S}^n$	the n-dim unit sphere in $\mathbb{R}^{n+1}$
$\mathbb{T}^n$	the n-dim torus $\mathbb{R}^n/\mathbb{Z}^n$ or $\mathbb{R}^n/2\pi\mathbb{Z}^n$
$\mathbb{H}^n$	the n-dim hyperbolic space
$\mathbb{B}^n$	the Poincaré ball model of the n-dim hyperbolic space
$\Re(z)$	the real part of $z \in \mathbb{C}$
$\Im(z)$	the imaginary part of $z \in \mathbb{C}$
$\overline{z}$	the complex conjugate of $z \in \mathbb{C}$
$\partial\Omega$	the boundary of Ω
$\Omega_1 \Subset \Omega$	the closure of Ω_1 is a compact subset of Ω

Matrices and maps

$\mathrm{M}_{m,n}(\mathbb{F})$	the set of $m \times n$ matrices with entries from $\mathbb{F}$, and $\mathrm{M}_n(\mathbb{F}) = \mathrm{M}_{n,n}(\mathbb{F})$
A^{T}	the transpose of a matrix $A \in \mathrm{M}_{m,n}$
$\operatorname{Tr} A$	the trace of a matrix $A \in \mathrm{M}_n$
$\\|A\\|_{\mathrm{HS}}$	the Hilbert-Schimidt norm of $A \in \mathrm{M}_n$
$\mathrm{GL}(n, \mathbb{F})$	the general linear group of degree n over a field $\mathbb{F}$
$\mathrm{SL}(n, \mathbb{F})$	the special linear group of degree n over a field $\mathbb{F}$
$\mathrm{O}(n)$	the orthogonal group of degree n
$\mathrm{SO}(n)$	the special orthogonal group of degree n
$\mathrm{U}(n)$	the unitary group of degree n
$\mathrm{SU}(n)$	the special orthogonal group of degree n
$\mathcal{L}(X, Y)$	the set of bounded linear transformations from X to Y
$\mathcal{L}(X)$	$\mathcal{L}(X, X)$, that is, the set of bounded linear transformations on X
I	the identity map or the identity matrix
$\ker A$	the kernel of $A \in \mathcal{L}(X, Y)$
$\operatorname{im} A$	the image of $A \in \mathcal{L}(X, Y)$
$[A, B]$	the commutator of A and B, that is, $[A, B] = AB - BA$

Analysis

$C(\Omega)$	the space of continuous functions in Ω
$\lvert\alpha\rvert$	the norm of a multi-index $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{N}^n$, that is, $\lvert\alpha\rvert = \alpha_1 + \cdots + \alpha_n$
$C^k(\Omega)$	the space of functions u such that $\partial^\alpha u \in C(\Omega)$ for all α with $\lvert\alpha\rvert \leq k$
$C_0^k(\Omega)$	the space of functions in $C^k(\Omega)$ with compact support in Ω
$C^\gamma(\Omega)$	the space of functions in $C^\gamma(\Omega)$ that are Hölder continuous with exponent $\gamma \in [0, 1]$
$C^{k,\gamma}(\Omega)$	the space of functions in $C^k(\Omega)$ such that $\partial^\alpha u \in C(\Omega)$ for all α with $\lvert\alpha\rvert \leq k$
$L^p_{\mathrm{loc}}(\Omega)$	the space of functions that are in $L^p(\Omega_1)$ for any $\Omega_1 \Subset \Omega$
$\mathcal{D}'(\Omega)$	the space of distributions in Ω, that is, the continuous linear functionals on $C_0^\infty(\Omega)$
$\operatorname{supp} u$	the support of $u \in \mathcal{D}'(\Omega)$
$\mathcal{E}'(\Omega)$	the set of distributions with compact support in Ω
$\mathcal{S}(\mathbb{R}^n)$	the space of Schwartz functions in $\mathbb{R}^n$
$\mathcal{S}'(\mathbb{R}^n)$	the space of Schwartz distributions in $\mathbb{R}^n$
$W^{k,p}(\Omega)$	the Sobolev space in Ω
$\operatorname{esssup} u$	the essential supremum of u
$\mathcal{F}u = \hat{u}$	the Fourier transform of u
$\mathcal{F}_h u$	the semiclassical Fourier transform of u
$\mathrm{Op}(a)$ ($\mathrm{Op}_h(a)$)	the (semiclassical) pseudodifferential operator (SDO) with symbol a
$\mathrm{WF}(u)$ ($\mathrm{WF}_h(u)$)	the (semiclassical) wavefront set of a distribution u
$\{A, B\}$	the Poisson bracket of A and B
p'	the conjugate exponent of $p \in [1, \infty]$, that is, $p = p/(p-1)$

Geometry

$\mathbb{M}$	a differential manifold with dimension n
$\mathcal{D}^r_s(\mathbb{M})$	the space of tensor fields of contravariant degree r and covariant degree s
$C^\infty(\mathbb{M})$	the space of smooth functions on $\mathbb{M}$, that is, $C^\infty(\mathbb{M}) = D^0_0(\mathbb{M})$
$\Lambda_s(\mathbb{M})$	the space of exterior differential forms of degree s on $\mathbb{M}$
$T_x\mathbb{M}$	the tangent space of $\mathbb{M}$ at $x \in \mathbb{M}$
$T\mathbb{M}$	the tangent bundle of $\mathbb{M}$
$T^*_x\mathbb{M}$	the cotangent space of $\mathbb{M}$ at $x \in \mathbb{M}$
$T^*\mathbb{M}$	the cotangent bundle of $\mathbb{M}$
$(\mathbb{M}, g)$	a differential manifold with Riemannian metric $g \in \mathcal{D}^0_2(\mathbb{M})$
$S_x\mathbb{M}$	the sphere space of $\mathbb{M}$ at $x \in \mathbb{M}$
$S\mathbb{M}$	the sphere bundle of $\mathbb{M}$
$S^*_x\mathbb{M}$	the cosphere space of $\mathbb{M}$ at $x \in \mathbb{M}$
$S^*\mathbb{M}$	the cosphere bundle of $\mathbb{M}$
$B(x, r)$	the (open) geodesic ball with center $x \in \mathbb{M}$ and radius $r \geq 0$
$\overline{B}(x, r)$	the closed geodesic ball with center $x \in \mathbb{M}$ and radius $r \geq 0$
Δ_g	the (negative unless specified) Laplace-Beltrami operator (i.e., Laplacian) on $\mathbb{M}$
Vol_g	the Riemannian volume form
$\lvert\Omega\rvert$	the Euclidean volume of $\Omega \subset \mathbb{R}^n$

CHAPTER 1

Classical chaos

By a chaotic system, we refer to a (continuous) dynamical system that is Anosov/hyperbolic: Let X be a Riemannian manifold and $\varphi_t : X \to X$, $t \in \mathbb{R}$, be an Anosov flow that is generated by a vector field V. "Anosov/hyperbolic" means that for each $z \in X$, the tangent space $T_z X$ splits to subspaces

$$E_c(z) \oplus E_s(z) \oplus E_u(z),$$

in which $E_c(z) = \mathbb{R}V$ is the 1-dim subspace spanned by V (i.e., the flow direction), and

$$\begin{cases} |d\varphi_t(v)|_{\varphi_t(z)} \le Ce^{-\alpha t}|v|_z, & \text{if } v \in E_s(z) \text{ and } t \ge 0, \\ |d\varphi_{-t}(v)|_{\varphi_{-t}(z)} \le Ce^{-\alpha t}|v|_z, & \text{if } v \in E_u(z) \text{ and } t \ge 0, \end{cases} \tag{1.1}$$

for some $C, \alpha > 0$. Here, $|\cdot|_z$ is the metric at z and α can be chosen as the maximal Lyapunov exponent. We call $E_s(z)$ (resp. $E_u(z)$) the stable (resp. unstable) subspace and they form the stable (resp. unstable) foliation as z varies in X.

One can similarly define a discrete hyperbolic system that is the iteration of an Anosov diffeomorphism, i.e., $\varphi_t : X \to X$, $t \in \mathbb{Z}$. But notice that there is no E_c for diffeomorphisms.

Example (Hyperbolic toral automorphisms). Let $M \in \mathrm{GL}(2, \mathbb{R})$. Then M induces a map $M : \mathbb{R}^2 \to \mathbb{R}^2$ and is hyperbolic if the eigenvalues of M have module different from one.

Let $M \in \mathrm{SL}(2, \mathbb{Z})$. Then $M : \mathbb{T}^2 \to \mathbb{T}^2$ preserves the Lebesgue measure and is called a hyperbolic toral automorphism if it is hyperbolic. For example, Arnold cat map is given by the matrix

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

We call hyperbolic toral automorphisms "cat maps" after Arnold[i].

The hyperbolic dynamical systems display the typical chaotic behaviors such as exponential sensitivity to the initial condition, divergence of nearby trajectories, mixing, ergodicity, etc. See Katok-Hasselblatt[ii] for a complete treatment.

The most important example of Anosov flows is given by the geodesic flow on negatively curved manifolds.

Example (Geodesic flows). The cotangent bundle of a manifold $\mathbb{M}$ is $T^*\mathbb{M} = \{(x, \xi) : x \in \mathbb{M}, \xi \in T^*_x\mathbb{M}\}$. The geodesic flow G_t in $T^*\mathbb{M}$ is generated by the Hamiltonian vector field

$$V_H(x, \xi) = (\partial_\xi H \cdot \partial_x, -\partial_x H \cdot \partial_\xi) \text{ with Hamiltonian } H = H(x, \xi) = |\xi|_x.$$

---

[i] V. Arnold and A. Avez, *Ergodic problems of classical mechanics.* W. A. Benjamin, Inc., New York-Amsterdam, 1968.

[ii] A. Katok and B. Hasselblatt, *Introduction to the modern theory of dynamical systems.* Cambridge University Press, Cambridge, 1995.

That is, the geodesic flow $G_t : (x(0), \xi(0)) \to (x(t), \xi(t))$ satisfies that

$$\begin{cases} \dot{x} := \frac{dx}{dt} = \partial_\xi H, \\ \dot{\xi} := \frac{d\xi}{dt} = -\partial_x H. \end{cases}$$

One can readily check that the symplectic form $\theta = d\xi \wedge dx$ is preserved under G_t.

Notice that $\partial_\xi H$ has singularity at $\xi = 0$. However, if we restrict G_t to the cosphere bundle $X = S^*\mathbb{M} = \{(x, \xi) \in T^*\mathbb{M} : |\xi|_x = 1\}$, G_t is a smooth flow with unit speed and preserves the canonical contact form on X (which is the restriction of θ to X). The cosphere bundle $S^*\mathbb{M}$ and the sphere bundle $S\mathbb{M}$ can be canonically identified by each other using the Riemannian metric on $\mathbb{M}$. Therefore, one can equivalently consider $S^*\mathbb{M}$ or $S\mathbb{M}$ in dynamical systems.

The geodesic flow G_t on $X = S^*\mathbb{M}$ is Anosov[i] on X if $\mathbb{M}$ has negative sectional curvature. Depending on the geometric structure of $\mathbb{M}$, the foliations E_s and E_u on X have different smoothness properties, which in turn dictate the results and tools available for the system.

(1). If the metric on $\mathbb{M}$ is smooth, then $E_s(z)$ and $E_u(z)$ are in general only Hölder continuous with respect to z. Even if the metric is analytic, this is still true.

(2). If $\mathbb{M}$ is a locally symmetric space, then E_s and E_u are analytic. A (locally) symmetric space is a space such that the reflection through every point is a (local) isometry. For example, the hyperbolic plane $\mathbb{H}^2 = \mathrm{SL}(2, \mathbb{R})/\mathrm{SO}(2)$ is a symmetric space and $\Gamma \backslash \mathbb{H}^2$ is a compact symmetric space if Γ is a cocompact subgroup of the isometry group of $\mathbb{H}^2$. In fact, any compact locally symmetric space can be written in this way: $\Gamma \backslash G/K$, in which G is a Lie group, K is an isometry group that fixes a specified point, and Γ is a cocompact lattice.

(3). Case (2) includes compact hyperbolic manifolds, i.e., compact Riemannian manifolds with constant negative curvature. After normalization, we may assume the curvature is -1. As a consequence, the Lyapunov exponent in (1.1) is uniform:

$$\begin{cases} |dG_t(v)|_{G_t(z)} = e^{-t}|v|_z, & \text{if } v \in E_s(z) \text{ and } t \geq 0, \\ |dG_{-t}(v)|_{G_{-t}(z)} = e^{-t}|v|_z, & \text{if } v \in E_u(z) \text{ and } t \geq 0. \end{cases}$$

Our main concern of the hyperbolic dynamical systems $\varphi_t : X \to X$ are as follows.

(A). Decay of correlation.
(B). Dynamical zeta function.
(C). Classification of the invariant distributions and measures under the flow.
(D). Application of classical chaos to its counterpart, Quantum Chaos, see Chapter 2.

1.1. Decay of correlation

The correlation of $f, g : X \to \mathbb{C}$ (which are often called the observables) with respect to a flow $\varphi_t : X \to X$ (and the Riemannian volume $d\mathrm{Vol}$ on X) is defined by

$$C_t(f, g) := \int_X f \cdot g \circ \varphi_t \, d\mathrm{Vol}.$$

Suppose that φ_t preserves the volume. Then

$$C_t(f, g) = \int_X \varphi_{-t} (f \cdot g \circ \varphi_t) \, d\mathrm{Vol} = \int_X g \cdot f \circ \varphi_{-t} \, d\mathrm{Vol} = C_{-t}(g, f).$$

A typical feature of the hyperbolic system is decay of correlation, i.e., $C_t(f, g) \to 0$ as $|t| \to \infty$ for sufficiently regular observables f, g.

[i] D. Anosov, *Geodesic flows on closed Riemann manifolds with negative curvature.* Proceedings of the Steklov Institute of Mathematics, No. 90 (1967).

For any C^4 hyperbolic flows on a contact manifold, Liverani[i] established exponential decay of correlation of Hölder functions. As a result,

Theorem (Exponential decay of correlations). *Let $\mathbb{M}$ be negatively curved and φ_t be the geodesic flow on $X = S^*\mathbb{M}$. Then for each $\gamma \in (0,1)$ and all $f, g \in C^\gamma(X)$,*

$$\left| C_t(f,g) - \int_X f \int_X g \right| \leq Ce^{-c|t|} \|f\|_{C^\gamma(X)} \|g\|_{C^\gamma(X)}, \tag{1.2}$$

in which $c = c(\gamma, \mathbb{M}), C = C(\gamma, \mathbb{M}) > 0$.

Problem 1.1 (Decay of correlation for distributions in hyperbolic systems). *Let $f \in C^\infty(X)$. Is decay of correlation possible for some $g \in \mathcal{D}'(X)$?*

1.1.1. Decay of correlation for hyperbolic toral automorphisms. Let

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : \mathbb{T}^2 \to \mathbb{T}^2$$

be a hyperbolic toral automorphism. Suppose that the eigenvalues of M are $e^{\pm\lambda}$ with $\lambda > 0$.

We show decay of correlation for smooth functions and investigate potential generalizations to distributions. Consider first the decay of correlation of $f_k(x) = e^{2\pi i k\cdot x}$ and $f_j(x) = e^{2\pi i j\cdot x}$ for $k, j \in \mathbb{Z}^2$ and $x \in \mathbb{T}^2$:

$$\begin{aligned} C_t\left(f_k, f_j\right) &= \int_{\mathbb{T}^2} f_k(x) \cdot f_j \circ M^t(x)\, dx \\ &= \int_{\mathbb{T}^2} e^{2\pi i k\cdot x} \cdot e^{2\pi i j M^t x}\, dx \\ &= \int_{\mathbb{T}^2} e^{2\pi i (k + jM^t)\cdot x}\, dx \\ &= \delta\left(k + jM^t\right). \end{aligned}$$

Let

$$f(x) = \sum_{k\in\mathbb{Z}^2} a_k f_k(x) \in \mathcal{D}'(\mathbb{T}^2).$$

Fix $j \in \mathbb{Z}^2$. Then

$$C_t(f, f_j) = \sum_{k\in\mathbb{Z}^2} a_k \delta\left(k + jM^t\right) = a_{-jM^t}.$$

If $j = (0,0)$, then the decay of correlation of f and f_j is clear. Otherwise, $\int f_j = 0$, and

$$\left|jM^t\right| \approx e^{\lambda|t|}|j| \to \infty \quad \text{as } |t| \to \infty.$$

Furthermore, as $t \to \infty$, jM^t converges toward the eigenaxe E_u of eigenvalue e^λ, while as $t \to -\infty$, jM^t converges toward the eigenaxe E_s of eigenvalue $e^{-\lambda}$.

Example. Design f such that its Fourier coefficients

$$a_k = \begin{cases} = |k|^{-1} & \text{if } k \text{ is near } E_u \text{ or } E_s \\ = 1 & \text{elsewhere.} \end{cases}$$

It is obvious that f is a distribution. In this case,

$$C_t(f, f_j) = a_{-jM^t} = \left|jM^t\right|^{-1} \approx e^{-\lambda|t|},$$

that is, exponential decay of correlation holds for such choices of $f \in \mathcal{D}'(\mathbb{T}^2)$ and $f_j \in C^\infty(\mathbb{T}^2)$.

[i] C. Liverani, *On contact Anosov flows.* Ann. of Math. (2) 159 (2004), no. 3, 1275–1312.

From this example, one sees that the decay of correlation of f and f_j holds, as long as the Fourier coefficients a_k of f have decay for $k \in \mathbb{Z}^2$ along E_u and E_s. In fact, one can design f to have sufficiently decay along E_u and E_s so that $C_t(f, f_j)$ decays arbitrarily fast; in the meanwhile, a_k can be arbitrary elsewhere so that f is a non-smooth distribution.

1.2. Dynamical zeta function

The dynamical zeta function for an Anosov flow $\varphi_t : X \to X$ is defined by

$$\zeta(z) = \prod_{\gamma \text{ prime}} \left(1 - e^{-zT_\gamma}\right), \tag{1.3}$$

in which T_γ is the period of the prime closed orbit γ of φ_t.

One can show that $\zeta(z)$ is analytic when $\Re(z) \gg 1$. Motivated by the investigations of Riemann zeta function, one asks

(1). meromorphic continuation of ζ to $\mathbb{C}$,
(2). location of zeros and poles of ζ,
(3). prime closed orbit counting problem in relation to the zeros/poles of the zeta function,
(4). functional equation of the zeta function,
(5). evaluation of of ζ on $\mathbb{C}$ (such as $\zeta(0)$) and the information these values contain (on geometry and topology of the manifold).

The above questions on $X = S^*\mathbb{M}$ such that $\mathbb{M}$ has constant curvature (and more generally, the locally symmetric spaces) are mostly known. While in the case of variable curvature, the questions remain largely open but there are some recent development, including the meromorphic continuation of ζ for all smooth negatively curved manifolds. See Giulitti-Liverani-Pollicott[i].

1.3. Invariant distribututions

We say that a distribution $u \in \mathcal{D}'(X)$ is invariant under an Anosov flow $\varphi_t : X \to X$ if $\varphi^* u = u$ for all $t \in \mathbb{R}$, or equivalently, $Vu = 0$ in the distributional sense, in which V is the generating vector field of φ_t. There are abundance of invariant distributions. An important class of invariant distributions are measures, on which we focus.

In fact, a distribution $u \in \mathcal{D}'(X)$ is a (positive) measure if and only if $u(\phi) \geq 0$ for all non-negative functions $\phi \in C_0^\infty(X)$, see Theorem 2.1.7 in Hörmander[ii]. (However, it is in general difficult to verify this condition for a given distribution. See Section 1.4.)

If X is compact and μ is a finite measure, then we always normalize μ so that it is a probability measure (i.e., $\mu(X) := \mu(\mathbb{1}_X) = 1$). For example, each closed orbit γ of φ_t supports a delta measure that is invariant:

$$\delta_\gamma := \frac{1}{T_\gamma} \int_{x \in \gamma} \delta_x.$$

In the case of $X = S^*\mathbb{M}$, there is also the canonical Liouville measure $\mu_L = \theta \wedge (d\theta)^{n-1}$ ($n = \dim \mathbb{M}$) which is invariant.

Between these two examples of invariant measures, one can think of δ_γ being the most irregular and μ_L being the most regular. More precisely, δ_γ is smooth along γ and is singular in the transverse directions, which means that the wavefront set $\mathrm{WF}(\delta_\gamma) = N^*\gamma$ (the conormal bundle of γ). While μ_L is smooth so $\mathrm{WF}(\mu_L) = \emptyset$. (The wavefront set of a distribution

[i] P. Giulietti, C. Liverani, and M. Pollicott, *Anosov flows and dynamical zeta functions.* Ann. of Math. (2) 178 (2013), no. 2, 687–773.

[ii] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Second edition. Springer-Verlag, Berlin, 1990.

characterizes where it is singular as well as the direction in which the singularity occurs. In particular, $\mathrm{WF}(\mu) = \emptyset$ iff μ is smooth.)

Remark. Let $\dim X = d$.

- We see that $\dim \operatorname{supp} \delta_\gamma = 0$ and $\dim \operatorname{supp} \mu_L = d$. For any $r \in [0, d]$, there are invariant measures whose supports have dimension r.
- Any invariant measure μ is a limit of the delta measures on closed orbits. That is, there is a sequence of orbits γ_j such that $\delta_{\gamma_j} \to \mu$ weakly as $j \to \infty$. See Sigmund[i].
- Any invariant measure μ can be decomposed to
$$\mu = \int_X \delta_x \, d\mu(x), \quad \text{in which } \delta_x = \lim_{T\to\infty} \frac{1}{T} \int_0^T \delta_{\varphi_t(x)} \, dt. \tag{1.4}$$
This is called the ergodic decomposition of μ.

One of the most important problems in classical chaos is to classify the invariant measures. In particular, our interest of this problem is related to Quantum Chaos in Chapter 2. That is, the Laplacian eigenfunctions on a negatively curved manifold $\mathbb{M}$ induce the “semiclassical measures” (which characterize the macroscopic distribution of eigenfunctions in the phase space $T^*\mathbb{M}$). Semiclassical measures are always probability measures supported on $X = S^*\mathbb{M}$ and are invariant under the geodesic flow.

The Quantum Unique Ergodicity (QUE) conjecture then claims that the semiclassical measure is unique and coincides with the Liouville measure μ_L. Since there are a lot of invariant measures, QUE essentially asks to narrow down the semiclassical measures to one single candidate, μ_L, using further information than the flow invariance.

In the view of δ_γ and μ_L, we can ask for smoothness condition in order to classify the invariant measures.

Definition (Absolutely continuous measures and smooth measures).

- A measure is said to be absolutely continuous if in any smooth local chart it is given by integrating a density.
- An absolutely continuous measure is said to be smooth if the density is a smooth function.

Then there is at most one smooth invariant measure in a compact hyperbolic system, which is μ_L in the case of $X = S^*\mathbb{M}$. See Theorem 20.4.1 in Katok-Hasselblatt.

Being smooth is very restrictive. It turns out that this smoothness condition can be significantly weakened.

Theorem (Dyatlov-Zworski). *Let $\mathbb{M}$ be a compact surface with negative curvature and $X = S^*\mathbb{M}$. Let μ be a (normalized) distribution that is invariant under the geodesic flow on X. Suppose that the wavefront set $\mathrm{WF}(\mu) \subset E_u^*$, which means that μ can only be singular in the unstable directions. Then $\mu = \mu_L$.*

This theorem is a byproduct of the investigation of Dyatlov-Zworski[ii] on the topological information of $\mathbb{M}$ that is encoded in the dynamical zeta function (1.3).

Problem 1.2 (Wavefront of semiclassical measures in Quantum Chaos). *What is $\mathrm{WF}(\mu)$ of a semiclassical measure μ, in particular, is $\mathrm{WF}(\mu) \subset E_u^*$?*

---

[i] K. Sigmund, *On the space of invariant measures for hyperbolic flows.* Amer. J. Math. 94 (1972), 31–37.

[ii] S. Dyatlov and M. Zworski, *Ruelle zeta function at zero for surfaces.* Invent. Math. 210 (2017), no. 1, 211–229.

On manifolds with certain arithmetic structure, a classification of invariant measures was proved by Lindenstrauss[i], using the additional Hecke conditions on such manifolds.

Theorem (Lindenstrauss). *Let $\mathbb{M}$ be an arithmetic hyperbolic surface with finite volume and $X = S^*\mathbb{M}$. (This includes compact arithmetic hyperbolic surfaces and the modular surface $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}^2$.) Let μ be an invariant measure under the geodesic flow on X. Suppose that*

(i). each ergodic component of μ has positive entropy,

(ii). μ is recurrent with respect to a Hecke operator.

Then $\mu = c\mu_L$ for some constant c.

In application to Quantum Chaos by Lindenstrauss and Brooks-Lindenstrauss[ii], one shows that any semiclassical measure μ induced by Hecke eigenfunctions (which are eigenfunctions of the Laplacian and of a Hecke operator) must satisfy the two conditions in the above theorem. As a consequence, $\mu = c\mu_L$. Furthermore, to determine the constant c, we divide into compact and non-compact cases.

(1). If $\mathbb{M}$ is compact, then $c = 1$ in the above theorem since a semiclassical measure is always a probability measure (after normalizing the eigenfunctions in $L^2(\mathbb{M})$).

(2). If $\mathbb{M}$ is the modular surface $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}^2$, which is non-compact and has a cusp, then one has to eliminate the possibility of $c = 0$ in the above theorem. This phenomenon of $c = 0$ is called "escape to infinity in the cusp". It can be roughly understood as follows. A semiclassical measure μ is first defined in compact regions of X, via the test functions (which are symbols of the pseudo-differential operators, which localize the eigenfunctions in the phase space) with compact support. Even though μ is always a probability measure, i.e., $\mu(X) = 1$, μ could potentially concentrates at infinity in the cusp so that $\mu(\Omega) = 0$ for any compactly supported region $\Omega \subset X$, i.e., μ escapes to infinity. Escape to infinity is eliminated by Soundararajan[iii], thus finishing QUE on the modular surface.

1.3.1. Entropy. Let μ be a probability measure on X with a hyperbolic flow φ_t. Suppose that $P = \{P_j\}_{j=1}^J$ is a finite measurable partition of X, i.e., $P_j \subset X$ is μ-measurable and $X = \cup P_j$ (or weaker, $\mu(X \setminus \cup P_j) = 0$). The Shannon entropy of μ with respect to the partition P is defined by

$$h_P(\mu) := -\sum_{j=1}^{J} \mu(P_j) \log \mu(P_j).$$

Here, we agree that $0 \log 0 = 0$.

Assume that μ is invariant under the flow φ_t. For any $n \in \mathbb{N}$, write

$$\begin{aligned} P^{\vee n} &= P \vee \varphi_{-1}(P) \vee \cdots \vee \varphi_{-(n-1)}(P) \\ &= \left\{P_{\alpha_0} \cap \varphi_{-1}(P_{\alpha_1}) \cap \cdots \cap \varphi_{-(n-1)}\left(P_{\alpha_{n-1}}\right) : \alpha_0, \alpha_1, ..., \alpha_{n-1} \in \{1, ..., J\}\right\}. \end{aligned}$$

Denote

$$h_n(\mu, P) = h_{P^{\vee n}}(\mu).$$

---

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

[ii]S. Brooks and E. Lindenstrauss, *Joint quasimodes, positive entropy, and quantum unique ergodicity.* Invent. Math. 198 (2014), no. 1, 219–259.

[iii]K. Soundararajan, *Quantum unique ergodicity for* $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}^2$. Ann. of Math. (2) 172 (2010), no. 2, 1529–1538.

Then the (Kolmogorov-Sinai/metric) entropy of μ with respect to the partition P is defined as

$$h(\mu, P) = \lim_{n\to\infty} \frac{h_n(\mu, P)}{n},$$

and the (Kolmogorov-Sinai/metric) entropy of μ is defined as

$$h(\mu) = \sup_P \frac{h_n(\mu, P)}{n},$$

in which the supremum is taken over all the measurable partitions of X. In fact, because of the hyperbolic nature of the flow, the choice of the partition is not crucial. – The supremum is achieved as soon as $\sup_{P_j\in P} \operatorname{diam}(P_j)$ is small enough.

Remark. The entropy is linear with respect to the ergodic decomposition (1.4):

$$h(\mu) = h\left(\int_X \delta_x \, d\mu(x)\right) = \int_X h\left(\delta_x\right) \, d\mu(x).$$

Since $x_0 \in \varphi_{-k}\left(P_{\alpha_k}\right)$ iff $\varphi_k(x_0) \in P_{\alpha_k}$,

$$\mu\left(P_{\alpha_0} \cap \varphi_{-1}\left(P_{\alpha_1}\right) \cap \cdots \cap \varphi_{-(n-1)}\left(P_{\alpha_{n-1}}\right)\right)$$

measures the μ-probability to visit successively $P_{\alpha_0}, P_{\alpha_1}, ..., P_{\alpha_{n-1}}$ at time $0, 1, ..., n-1$. Hence, the entropy measures the average exponential decay of these probabilities as $n \to \infty$. In particular, if for all sequences $\alpha_0, \alpha_1, ..., \alpha_{n-1}$ chosen from $\{1, ...J\}$,

$$\mu\left(P_{\alpha_0} \cap \varphi_{-1}\left(P_{\alpha_1}\right) \cap \cdots \cap \varphi_{-(n-1)}\left(P_{\alpha_{n-1}}\right)\right) \le Ce^{-h_0 n} \quad \text{for some } C > 0, h_0 \ge 0.$$

Then $h(\mu) \ge h_0$.

Example. Let $\mu = \delta_\gamma$. Then

$$\mu\left(P_{\alpha_0} \cap \varphi_{-1}\left(P_{\alpha_1}\right) \cap \cdots \cap \varphi_{-(n-1)}\left(P_{\alpha_{n-1}}\right)\right) = 0$$

unless $P_{\alpha_0} \cap \gamma \ne \emptyset$ and $P_{\alpha_k} \cap \varphi_1\left(P_{\alpha_{k-1}}\right) \ne \emptyset$ for $k = 1, ..., n-1$. The number of such sequences $P_{\alpha_0}, P_{\alpha_1}, ..., P_{\alpha_{n-1}}$ grows linearly in n, while the total number grows exponentially. Therefore, $h(\mu) = 0$.

Ruelle[i] proved the following upper bound of the entropy, see Theorem S.2.13 in Katok-Hasselblatt for more discussion.

Theorem (Ruelle). *Let μ be an invariant measure on X with respect to a hyperbolic flow φ_t. Then*

$$h(\mu) \le \int_X \left|\log J^u(z)\right| \, d\mu(z),$$

in which $J^u(z)$ is the Jacobian of $E_u(z) \to E_u(\varphi_{-1}(z))$ (i.e., $\log J^u(z)$ is the sum of the positive Lyapunov exponents).

In particular, on $X = S^*\mathbb{M}$ of a hyperbolic manifold $\mathbb{M}$ (i.e., the curvature is -1), the Lyapunov exponents are 1 *and* $\dim E_u = n - 1$. *Hence,* $h(\mu) \le n - 1$.

Furthermore, Ledrappier-Young[ii] proved that μ is absolutely continuous iff the above equality holds. In particular, if $X = S^*\mathbb{M}$, then the equality holds iff $\mu = \mu_L$.

[i] D. Ruelle, *An inequality for the entropy of differentiable maps.* Bol. Soc. Brasil. Mat. 9 (1978), no. 1, 83–87.

[ii] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula.* Ann. of Math. (2) 122 (1985), no. 3, 509–539.

1.4. Invariant measures with uniform projection

In this section, we continue the investigation on the classification of invariant measures in classical chaos: Given an additional condition that the projection to a submanifold is uniform, we hope the invariant measure is uniquely determined (after necessary normalization).

Let $\mathbb{M}$ be a compact manifold with negative curvature and $X = S^*\mathbb{M}$ be the cosphere bundle of $\mathbb{M}$. Denote by $P : X \to \mathbb{M}$ the projection of X onto $\mathbb{M}$.

Problem 1.3 (Invariant measures with uniform projection on hyperbolic manifolds). *Let μ be a probability measure on X which is invariant under the geodesic flow on X. Assume that $P(\mu) = \mathrm{Vol}$, in which $d\mathrm{Vol}$ is the Riemannian volume form on $\mathbb{M}$. Then must μ be (a constant multiple of) the Liouville measure μ_L on X?*

Remark. This problem is motivated by the quantum unique ergodicity (QUE) conjecture in Quantum Chaos, see Chapter 2. That is, let μ be a semiclassical measure induced by a sequence of Laplacian eigenfunctions on $\mathbb{M}$. Then μ is a probability measure which is invariant under the geodesic flow on $X = S^*\mathbb{M}$. The QUE conjecture asserts that $\mu = \mu_L$ (after necessary normalization), which means that the eigenfunctions are equidistributed in the phase space X. As a consequence, $P(\mu) = \mathrm{Vol}$, which means that the eigenfunctions are equidistributed on $\mathbb{M}$.

However, if the answer to the above problem is positive, then we know that equidistribution implies the the QUE conjecture. In particular, the eigenfunctions tend to be equidistributed in the phase space X if and only if they do in the physical space $\mathbb{M}$. Therefore, one only needs to consider symbols depending on the position variable in the QUE conjecture.

1.4.1. Invariant measures with uniform projection for hyperbolic toral automorphisms. Let

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{Z}) : \mathbb{T}^2 \to \mathbb{T}^2$$

be a hyperbolic toral automorphism (i.e., a cat map). Here,

$$\mathbb{T}^2 = \{x = (x_1, x_2) : x_1, x_2 \in \mathbb{T}\}.$$

Denote by P_1 and P_2 the projections onto the x_1-variable and onto the x_2-variable, respectively.

Problem 1.4 (Invariant measures with uniform projection in cat maps). *Let μ be a probability measure on $\mathbb{T}^2$ which is invariant under M. Assume that $P_1(\mu) = dx_1$, in which dx_1 is the Lebesgue measure on $\mathbb{T}$. Then must μ be (a constant multiple of) the Lebesgue measure dx on $\mathbb{T}^2$?*

Expand μ in Fourier series:

$$\mu(x) = \sum_{k \in \mathbb{Z}^2} a_k e^{2\pi i k \cdot x},$$

in which

$$a_k = \int_{\mathbb{T}^2} e^{-ik\cdot x}\, d\mu(x).$$

We gather the conditions on μ via its Fourier coefficients.

(i). μ is a probability measure:

(i.a). $a_k = \overline{a_{-k}}$ for all $k \in \mathbb{Z}^2$,

(i.b). $a_k = 1$ when $k = (0, 0)$,

(i.c). $\{a_k\}_{k\in\mathbb{Z}^2}$ is positive definite.

Theorem (Herglotz). *$\{a_k\}_{k\in\mathbb{Z}^2}$ is the Fourier coefficients of a probability measure on $\mathbb{T}^2$ if and only if $a_{(0,0)} = 1$ and $\{a_k\}_{k\in\mathbb{Z}^2}$ is positive definite.*

See Section 1.7.6 in Katznelson[i]. Here, a sequence $\{a_k\}_{k\in\mathbb{Z}^2}$ is said to be positive definite if

$$\sum_{m,n\in\mathbb{Z}^2} a_{m-n} z_m \overline{z_n} \geq 0 \quad \text{for each } \{z_n\}_{n\in\mathbb{Z}^2} \subset \mathbb{C}.$$

(ii). μ is invariant under $\mathbb{M}$. This amounts to

$$a_k = a_{M^t k} \quad \text{for all } t \in \mathbb{Z}, k \in \mathbb{Z}^2.$$

(iii). The projection of $P_1(\mu)$,

$$P_1(\mu)(x_1) = \sum_{k_1\in\mathbb{Z}} a_{k_1,0} e^{2\pi i k_1 x_1},$$

equals dx_1. This amounts to

$$a_{k_1,0} = 0 \quad \text{for all } k_1 \in \mathbb{Z} \setminus \{0\}.$$

[i] Y. Katznelson, *An introduction to harmonic analysis*. Cambridge University Press, Cambridge, 2004.

CHAPTER 2

Quantum Chaos

Let $\mathbb{M}$ be a manifold with negative curvature so the geodesic flow G_t is chaotic on the cosphere bundle $X = S^*\mathbb{M}$, see Chapter 1. In Quantum Chaos, our main concern is how the chaotic properties of the classical flow influence its quantum counterpart, i.e., the spectral and scattering properties of the Laplacian.

In the case when $\mathbb{M}$ is compact, the (positive) Laplacian Δ has a discrete set of eigenvalues $\lambda_1^2 \le \lambda_2^2 \le \cdots \to \infty$ and corresponding eigenfunctions $\{u_j\}_{j=1}^{\infty}$, $\Delta u_j = \lambda_j^2 u_j$. One can find an orthonormal basis of eigenfunctions in $L^2(\mathbb{M})$, which is called an eigenbasis. The main intuitions in Quantum Chaos are

(1). the distribution of the eigenvalues is modeled by the eigenvalues of large random matrices,
(2). the properties of the eigenfunctions are modeled by random combinations of (a large number of) plane waves, which are called random waves.

See Sarnak[i]. The precise realization of these two intuitions remain largely open. In addition, the study of random matrices and random waves have their own interests.

In the case when $\mathbb{M}$ is non-compact but has finite volume, for example, the modular surface $\mathrm{SL}(2,\mathbb{Z})\backslash\mathbb{H}^2$, Δ has a discrete set of eigenvalues as well as a continuous spectrum, which correspond to the Eisenstein series. Then the Quantum Chaos questions can be asked about the eigenfunctions as well as the Eisenstein series.

In the case when $\mathbb{M}$ is non-compact and has infinite volume, for example, the convex cocompact hyperbolic manifolds, Δ has a continuous spectrum as well as a discrete set of resonances, which correspond to the resonant states. (One can in fact also define resonant states in the case of finite volume.) Then the Quantum Chaos questions can be asked about the Eisenstein series as well as the resonances and resonant states.

We also consider the quantization of discrete Anosov diffeomorphisms, for example, quantum cat map is the quantized operator of Arnold cat map on the torus $\mathbb{T}^2$,

$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} : \mathbb{T}^2 \to \mathbb{T}^2.$$

Quantum cat map is the simplest model in Quantum Chaos, yet it displays all the important (and extreme) features of the Quantum Chaos phenomenon.

We use "eigenstates" to broadly refer to the eigenfunctions, Eisenstein series, resonant states, and denote by $u(h)$ with $h \to 0$ as the semiclassical parameter. For example, on a compact manifold, $u(h)$ denote the Laplacian eigenfunctions with $h = \lambda^{-1} \to 0$. Our main concern of Quantum Chaos is about the properties of eigenstates in a chaotic system.

(A). The macroscopic distribution of $u(h)$ can be studied via the semiclassical measure that is induced by $u(h)$. A semiclassical measure μ is defined in the phase space $T^*\mathbb{M}$. Its projection to $\mathbb{M}$ is simply the weak limits of $|u(h)|^2\, d\mathrm{Vol}$, in which Vol is the Riemannian

[i] P. Sarnak, *Arithmetic Quantum Chaos.* The Schur lectures (1992) (Tel Aviv), 183–236, Israel Math. Conf. Proc., 8, Bar-Ilan Univ., Ramat Gan, 1995.

volume on $\mathbb{M}$. Suppose that $|u(h_k)|^2\,d\mathrm{Vol} \to d\mu^{\mathbb{M}}$ for a measure $\mu^{\mathbb{M}}$ on $\mathbb{M}$, i.e.,

$$\int_{\mathbb{M}} a(x)|u(h_k)|^2\,d\mathrm{Vol} \to \int_{\mathbb{M}} a(x)\,d\mu^{\mathbb{M}} \quad \text{for all } a \in C_0^\infty(\mathbb{M}).$$

Then the macroscopic distribution of $u(h_k)$ on $\mathbb{M}$ is reflected in $\mu^{\mathbb{M}}$. The eigenstates in chaotic systems are expected to be equidistributed at the macroscopic scale, which requires that $\mu^{\mathbb{M}} = \mathrm{Vol}$. In particular,

$$\int_{\Omega} |u(h_k)|^2\,d\mathrm{Vol} \to \frac{\mathrm{Vol}(\Omega)}{\mathrm{Vol}(\mathbb{M})} \quad \text{for all open sets } \Omega \subset \mathbb{M}.$$

The question of equidistribution of chaotic eigenstates is largely open. To answer this question, one compares $\mu^{\mathbb{M}}$ with Vol in various measure-theoretic or topological considerations.
(a). What is $\operatorname{supp}\mu^{\mathbb{M}}$? Is it the whole manifold?
(b). What is the regularity and wavefront set of $\mu^{\mathbb{M}}$? Is $\mathrm{WF}(\mu^{\mathbb{M}}) = \emptyset$ so $\mu^{\mathbb{M}}$ is smooth?
(c). What is distance between $\mu^{\mathbb{M}}$ and Vol? Is it zero?
(d). What is the weight of the Vol component of $\mu^{\mathbb{M}}$? Is it 1?
...
One should in fact ask these questions for the semiclassical measure μ in the cotangent bundle $T^*\mathbb{M}$, where the classical and quantum chaotic dynamics take place. (Then one is to compare the semiclassical measure to the Liouville measure.) For example, what is the topological or metric entropy of μ? Is it maximal among all the invariant measures with respect to the geodesic flow?

(B). The microscopic distribution of $u(h)$ asks the L^2 mass of $u(h)$ in regions that are at some microscopic scale $r(h) \to 0$ as $h \to 0$. For example, equidistribution of $u(h_k)$ at a scale $r_k = r(h_k) \to 0$ means that

$$\int_{B(x,r_k)} |u(h_k)|^2\,d\mathrm{Vol} = \frac{\mathrm{Vol}(B(x,r_k))}{\mathrm{Vol}(\mathbb{M})} + o\left(\mathrm{Vol}(B(x,r_k))\right) \quad \text{for all } x \in \mathbb{M}.$$

Here, $B(x,r)$ is a geodesic ball with center x and radius r. (The case when r is independent of h corresponds to equidistribution at the macroscopic scale.) It amounts to the study of the limits in (1) for test functions $a(x;h)$ that are supported in small balls of radius $r(h)$. Eigenstates $u(h)$ oscillate at a typical wavelength of h. Therefore, their equidistribution can only be expected above the Planck scale h. On compact manifolds with negative curvature, some equidistribution results at logarithmical scales $r = |\log h|^{-\alpha}$ for some $\alpha > 0$ are known due to Han[i] and Hezari-Rivière[ii]. These scales are much larger than the Planck scale.

(C). The distribution of $u(h)$ on all measurable sets (rather than on only open sets). It would follow from the study of the limits in (A) for test functions $a(x) \in L^p(\mathbb{M})$ for some p.

(D). The L^p norm estimates of $u(h)$ ask the bounds for $p > 2$ that

$$\|u(h)\|_{L^p(\mathbb{M})} \le Ch^{-\delta(p)}\|u\|_{L^2(\mathbb{M})} \quad \text{for some appropriate } \delta(p) > 0.$$

These estimates also reflect the distribution property of $u(h)$. In principle, more equidistributed (at the microscopic scales) $u(h)$ are, smaller $\delta(p)$ that one can expect. See Chapter 3.

(E). The nodal sets of $u(h)$. See Chapter 4.

---

[i] X. Han, *Small scale Quantum Ergodicity in negatively curved manifolds.* Nonlinearity 28 (2015), no. 9, 3263–3288.

[ii] H. Hezari and G. Rivière, *L^p norms, nodal sets, and quantum ergodicity.* Adv. Math. 290 (2016), 938–966.

Problem 2.1 (Quantum Chaos implies classical chaos). *It is almost completely unknown what consequence of Quantum Chaos has on the classical dynamics. For example, suppose that all eigenstates behave like random waves. Then must the classical dynamic be chaotic?*

2.1. Semiclassical analysis

Semiclassical analysis is the mathematical toolbox to connect classical mechanics and quantum mechanics. See Zworski[i] for a complete treatment of this topic.

There are two components of a classical or quantum mechanical system: (1). The kinematics describe the states and observables; (2). the dynamics describe their evolution in time.

2.1.1. Classical mechanics. Consider the classical mechanics of a point particle moving on a manifold $\mathbb{M}$ in the Hamiltonian view. See Section 10.3. Let the Hamiltonian $H(x,\xi) : T^*\mathbb{M} \to \mathbb{R}$.

Theorem (Kinematics in the classical mechanics).

(i). A state is described by a point (x,ξ) in the phase space $T^*\mathbb{M}$. Here, $x \in \mathbb{M}$ is the position variable and $\xi \in T^*_x\mathbb{M}$ is the momentum variable.

(ii). An observable is described by a smooth function $a \in T^*\mathbb{M}$. For example, the kinetic energy observable is given by $a(x,\xi) = |\xi|_x^2/(2m)$, in which m is the mass of the particle.

Theorem (Dynamics in the classical mechanics).

(i). A state evolves under the Hamiltonian flow by $(x,\xi) \to \varphi_t(x,\xi)$. Here, φ_t is generated by the Hamiltonian $H(x,\xi)$, i.e., φ_t is the integral flow of the Hamiltonian vector field $(\partial_\xi H \cdot \partial x, -\partial_x H \cdot \partial_\xi)$ in $T^*\mathbb{M}$.

(ii). An observable evolves under the Hamiltonian flow by the Poisson bracket:

$$\dot{a} = \frac{d}{dt}(a \circ \varphi_t) = \partial_x a \cdot \dot{x} + \partial_\xi a \cdot \dot{\xi} = \partial_x a \cdot \partial_\xi H - \partial_\xi a \cdot \partial_x H := \{H, a\}.$$

In particular, $\{H,H\} = 0$ implies that the Hamiltonian is preserved under the flow.

Example (Geodesic flow). The movement of a free point particle on a manifold $\mathbb{M}$ with unit speed is described by the geodesic flow G_t with Hamiltonian $H(x,\xi) = |\xi|_x$. The Hamiltonian vector field is not smooth at $\xi = 0$. However, since H is preserved under the flow, one can restrict G_t to the cosphere bundle $S^*\mathbb{M} = \{(x,\xi) \in T^*\mathbb{M} : |\xi|_x = 1\}$. Alternatively, one can set $H(x,\xi) = |\xi|_x^2/2$ to define the unit speed geodesic flow on $S^*\mathbb{M}$.

2.1.2. Quantum mechanics. Consider the quantum mechanics of a point particle moving on a manifold $\mathbb{M}$.

Theorem (Kinematics in the quantum mechanics).

(i). A state is described by a function $u \in L^2(\mathbb{M})$, whose density provides the probability distribution of where the particle can be found in $\mathbb{M}$. That is, after normalizing u in $L^2(\mathbb{M})$, the probability of finding the particle in a region $\Omega \subset \mathbb{M}$ is given by $\int_\Omega |u|^2$.

(ii). An observable is described by an operator $\mathrm{Op}_h(a)$ acting on $L^2(\mathbb{M})$, in which $a \in C^\infty(T^*\mathbb{M})$. Here, h is the Planck constant. For example, if $a(x,\xi) = |\xi|_x^2/(2m)$, then $\mathrm{Op}_h(a) = h^2\Delta/(2m)$ and the expected value of this quantum observable of kinetic energy at a state of u is given by $\langle \mathrm{Op}_h(a)u, u\rangle$.

[i] M. Zworksi, *Semiclassical analysis.* American Mathematical Society, Providence, RI, 2012.

Remark (Quantization). Quantization is a procedure that associates every function/symbol $a \in C^\infty(T^*\mathbb{M})$ with a pseudodifferential operator (SDO) $\mathrm{Op}_h(a)$ satisfying the pseudodifferential calculus mentioned below. For example, in $\mathbb{M} = \mathbb{R}^n$, the Weyl quantization is

$$\mathrm{Op}_h^{\mathrm{w}}(a)u(x) = \frac{1}{(2\pi h)^n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi} a\left(\frac{x+y}{2},\xi\right) u(y)\,dyd\xi.$$

- A SDO $\mathrm{Op}_h(a)$ can be understood as a localization by $a(x,\xi)$ in the phase space. That is, the phase portrait of a function u consists of its values $u(x)$ in the physical space and its (semiclassical) Fourier transform $\hat{u}(\xi)$ in the frequency space. Then the phase portrait of the function $\mathrm{Op}_h(a)u$ is obtained by multiplying $a(x,\xi)$ to the phase portrait of u at (x,ξ). For example, if $a(x,\xi) = a(x)$, then $\mathrm{Op}_h(a)u(x) = a(x)u(x)$.
- A SDO $\mathrm{Op}_h(a)$ for $a \in C_0^\infty(T^*\mathbb{M})$ is bounded in $L^2(\mathbb{M})$. In fact,
$$\|\mathrm{Op}_h(a)\|_{L^2(\mathbb{M})\to L^2(\mathbb{M})} \le C \cdot \sup_{(x,\xi)\in T^*\mathbb{M}} |a(x,\xi)|.$$
The upper bound can be achieved by letting $\mathrm{Op}_h(a)$ to act on a function u whose phase portrait is localized near the supremum of $|a|$.

Theorem 2.1 (Pseudodifferential calculus). *Let $a, b \in C_0^\infty(T^*\mathbb{M})$.*

(i). Conjugation:
$$\mathrm{Op}_h(a)^\star = \mathrm{Op}_h(\overline{a}) + O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}(h).$$
(ii). Composition:
$$\mathrm{Op}_h(a)\mathrm{Op}_h(b) = \mathrm{Op}_h(ab) + O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}(h).$$
(iii). Commutation:
$$[\mathrm{Op}_h(a), \mathrm{Op}_h(b)] = \frac{h}{i}\mathrm{Op}_h(\{a,b\}) + O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}\left(h^2\right).$$
Therefore, the commutation can be interpreted as the quantization of the Poisson bracket.

Denote the quantization of the Hamiltonian by $\hat{H} = \mathrm{Op}_h(H)$ and the Schrödinger propagator by $U_t = e^{it\hat{H}/h}$.

Theorem (Dynamics in the quantum mechanics).

(i). A state u evolves under the Schrödinger flow by $u \to U_t(u)$. In particular, if u is an eigenfunction of $\hat{H}$, $\hat{H}u = ku$ for $k \in \mathbb{R}$, then $|U_t(u)|^2 = |u|^2$. This means that the probability distribution defined by the state u stays unchanged in time, i.e., u is a stationary state in the quantum system.

(ii). An observable $\mathrm{Op}_h(a)$ evolves under the Schrödinger flow by $\mathrm{Op}_h(a) \to U_{-t}\mathrm{Op}_h(a)U_t$.

Remark. Semiclassical analysis provides the mathematical tools to the kinematics and the dynamics in quantum systems. In particular,

- the pseudodifferential calculus can be used to describe the quantum kinematics, i.e., states and observables are L^2 functions and operators acting on L^2 functions,
- the Fourier integral calculus can be used to describe the quantum dynamics, i.e., the evolution of states and observables are through the Schrödinger propagator $U_t = e^{it\hat{H}/h}$, which is a Fourier integral operator (FIO).

Bohr's correspondence principle demands that the quantum mechanics reduces to the classical one as $h \to 0$. In semiclassical analysis, the Planck constant h is considered as a dimensionless small parameter. Bohr's principle is then interpreted as the influence of classical mechanics to the asymptotic analysis in quantum mechanics as $h \to 0$. For example, the Egorov's theorem states that the evolution of quantum observables under the Schrödinger flow U_t reduces to the one of the classical observable under the Hamiltonian flow φ_t, as $h \to 0$.

THEOREM 2.2 (Egorov's theorem). *Let $a \in C_0^\infty(T^*\mathbb{M})$ and $0 \le t \le T_0 < \infty$. Then*

$$\|U_{-t}\mathrm{Op}_h(a)U_t - \mathrm{Op}_h(a\circ\varphi_t)\|_{L^2(\mathbb{M})\to L^2(\mathbb{M})} = O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}(h).$$

PROOF. Recall that $U_t = e^{-it\hat{H}/h}$. Compute that

$$\begin{aligned}
&\frac{d}{dt}\left(U_t\mathrm{Op}_h(a\circ\varphi_t)U_{-t}\right)\\
=\ &-\frac{i}{h}\hat{H}U_t\mathrm{Op}_h(a\circ\varphi_t)U_{-t} + U_t\mathrm{Op}_h\left(\frac{d}{dt}(a\circ\varphi_t)\right)U_{-t} + U_t\mathrm{Op}_h(a\circ\varphi_t)\frac{i}{h}\hat{H}U_{-t}\\
=\ &U_t\left(\mathrm{Op}_h(\{H,a\}) - \frac{i}{h}\left[\hat{H},\mathrm{Op}_h(a)\right]\right)U_{-t}\\
=\ &O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}(h).
\end{aligned}$$

This is because U_t is unitary in $L^2(\mathbb{M})$ and according to the commutator rule

$$\left[\hat{H},\mathrm{Op}_h(a)\right] = \frac{h}{i}\mathrm{Op}_h(\{H,a\}) + O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}\left(h^2\right)$$

so

$$\mathrm{Op}_h(\{H,a\}) - \frac{i}{h}\left[\hat{H},\mathrm{Op}_h(a)\right] = O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}(h).$$

Now

$$U_t\mathrm{Op}_h(a\circ\varphi_t)U_{-t} - \mathrm{Op}_h(a) = \int_0^t \frac{d}{ds}\left(U_s\mathrm{Op}_h(a\circ\varphi_s)U_{-s}\right)ds = O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}(h).$$

Therefore, the theorem follows since U_t is unitary. ☐

2.1.3. Symbol classes. The pseudodifferential calculus and Egorov's theorem above involve the symbols in $C_0^\infty(T^*\mathbb{M})$. It is useful to enlarge the class of functions/symbols such that they can also be allowed to depend on the semiclassical parameter $h \to 0$.

Definition (Symbol classes). Let $\rho \in [0,1)$ and $0 < h_0 < 1$. We say that $a \in S_\rho(\mathbb{M})$ if $a(x,\xi;h) \in C_0^\infty(T^*\mathbb{M}\times(0,h_0))$ satisfies that

$$\left|\partial_{x,\xi}^\alpha a(x,\xi;h)\right| \le C_\alpha h^{-\rho|\alpha|}$$

for all multiindex α. In particular, if $a \in C_0^\infty(T^*\mathbb{M})$ is independent of h, then $a \in S_0(\mathbb{M})$.

We denote by $\Psi_\rho(\mathbb{M})$ the corresponding class of SDOs with symbols in $S_\rho(\mathbb{M})$.

Remark.

- Since we require that the symbols in the above definition have compact support, the symbol classes are usually denoted by $S_\rho^{\mathrm{comp}}(\mathbb{M})$ in other texts. One can in fact define more general Kohn-Nirenberg symbols that allow C^∞ functions with certain decay in the frequency variable. See Section 9.3 in Zworski.
- A SDO $\mathrm{Op}_h(a)$ for $a \in S_\rho(\mathbb{M})$ with $\rho \le \frac{1}{2}$ remains bounded in $L^2(\mathbb{M})$ that is independent of h. In fact, $\|\mathrm{Op}_h(a)\|_{L^2(\mathbb{M})\to L^2(\mathbb{M})}$ depends on finite number of seminorms of a (i.e., C_α in the above inequality).
- A SDO $\mathrm{Op}_h(a)$ can still be understood as a localization in the phase space by $a(x,\xi;h)$. If $a \in S_\rho(\mathbb{M})$, then the localization is at scales no smaller than h^ρ in each variable. For example, let $a \in C_0^\infty(\mathbb{R}^n)$. Then $a_\rho(x) := a(x/h^\rho) \in S_\rho(\mathbb{R}^n)$ and $a_\rho(x)u(x)$ is a localization of u at scale h^ρ in the x variables.

- Since the SDOs $\mathrm{Op}_h(a) \in \Psi_\rho(\mathbb{M})$ can localize at scale h^ρ in each variable (x and ξ), $h^\rho \cdot h^\rho = h^{2\rho} \geq h$ according to the uncertainty principle, that is, $\rho \leq \frac{1}{2}$ and one can not simultaneously localize a function in x and in ξ at scales that are smaller than $\sqrt{h}$.
- Let $\rho \in [0, \frac{1}{2})$. Then there is a symbol map $\sigma : \Psi_\rho(\mathbb{M}) \to S_\rho(\mathbb{M})/h^{1-2\rho}S_\rho(\mathbb{M})$ that assigns each $A \in \Psi_\rho(\mathbb{M})$ with a symbol $\sigma(A)$ in $S_\rho(\mathbb{M})$ which is unique module $h^{1-2\rho}S_\rho(\mathbb{M})$. We call $\sigma(A)$ the principle symbol of A.

Theorem (Pseudodifferential calculus). *Let $a \in S_{\rho_1}(\mathbb{M})$ and $b \in S_{\rho_2}(\mathbb{M})$ with $0 \leq \rho_1, \rho_2 < \frac{1}{2}$.*

(i). Composition:

$$\mathrm{Op}_h(a)\mathrm{Op}_h(b) = \mathrm{Op}_h(ab) + O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}\left(h^{1-\rho_1-\rho_2}\right).$$

(ii). Commutation:

$$[\mathrm{Op}_h(a), \mathrm{Op}_h(b)] = \frac{h}{i}\mathrm{Op}_h\left(\{a, b\}\right) + O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}\left(h^{2(1-\rho_1-\rho_2)}\right).$$

2.2. Semiclassical measures

Let $\{u(h)\}_{0<h<h_0}$ be a family of functions for some $0 < h_0 < 1$ such that $\|u(h)\|_{L^2} = 1$. Then there is a sequence $h_k \to 0$ as $k \to \infty$ and a probability measure μ on $T^*\mathbb{M}$ such that

$$\langle \mathrm{Op}_h(a)u(h_k), u(h_k)\rangle_{L^2(\mathbb{M})} \to \int_{T^*\mathbb{M}} a(x, \xi)\, d\mu$$

for all $a \in C_0^\infty(T^*\mathbb{M})$. Here, $\mathrm{Op}_h(a) : L^2(\mathbb{M}) \to L^2(\mathbb{M})$ is its semiclassical quantization. We call μ the semiclassical measure induced by $\{u(h_k)\}_{k=1}^\infty$.

Theorem. *Let $p(x, \xi) : T^*\mathbb{M} \to \mathbb{R}$ and $P(h) = \mathrm{Op}_h(p)$.*

(i). If $\|P(h)u(h)\|_{L^2} = o(1)$, then $\mathrm{supp}\,(\mu) \subset p^{-1}(0) = \{(x, \xi) \in T^*\mathbb{M} : p(x, \xi) = 0\}$. That is, the semiclassical measure is supported on the characteristic set of the symbol p.

(ii). If $\|P(h)u(h)\|_{L^2} = o(h)$, then

$$\int_{T^*\mathbb{M}} \{p, a\}\, d\mu = 0$$

for all $a \in C_0^\infty(T^*\mathbb{M})$. That is, the semiclassical measure is invariant under the Hamiltonian flow H_p generated by the symbol p, $H_p(\mu) = 0$.

Remark.

- The support of the semiclassical measure in (i) follows the elliptic estimates. That is, if $\Omega \cap p^{-1}(0) \neq 0$, then $p \neq 0$ in Ω and $P(h)$ has a local inverse in Ω. Then $\|P(h)u(h)\|_{L^2} = o(1)$ implies that $\|u\|_{L^2(\Omega)} = o(1)$ so the semiclassical measure has no charge on Ω.
- The flow invariance of the semiclassical measure in (ii) is a direct consequence of the commutator rule in Theorem 2.1.

Example (Laplacian eigenfunctions). Let $p(x) = |\xi|^2 - 1$. Then $P(h) = h^2\Delta - 1$, the characteristic set $p^{-1}(0) = S^*\mathbb{M}$, and the Hamiltonian flow H_p is the geodesic flow on $S^*\mathbb{M}$. (One can choose $p(x, \xi) = (|\xi|^2 - 1)/2$ so that that the speed of the geodesic flow $|\dot{x}| = |\xi| = 1$ on $S^*\mathbb{M}$.)

(i). If $\|(h^2\Delta - 1)u(h)\|_{L^2} = o(1)$, i.e., $\|(\Delta - \lambda^2)u\|_{L^2} = o(\lambda^2)$ with $\lambda = h^{-1}$, then $\mathrm{supp}\,(\mu) \subset S^*\mathbb{M}$.

(ii). If $\|(h^2\Delta - 1)u(h)\|_{L^2} = o(h)$, i.e., $\|(\Delta - \lambda^2)u\|_{L^2} = o(\lambda)$, then the semiclassical measure is invariant under the geodesic flow.

In particular, the semiclassical measure induced by Laplacian eigenfunctions are always supported on $S^*\mathbb{M}$ and are invariant under the geodesic flow.

It is very poorly understood how perturbation of the classical geodesic flow can impact on the semiclassical measures. For example, Zelditch asks

Problem 2.2 (Semiclassical measures under perturbation). *Assume that there are a sequence of eigenfunctions concentrating in an invariant torus in $S^*\mathbb{M}$ of a manifold. We perturb the metric of $\mathbb{M}$ such that the invariant torus is not affected. Then what can we say about the semiclassical measure after perturbation? Do they also persist?*

2.3. Quantum Ergodicity

Let $\{u_{j_k}\}$ be a subsequence of Laplacian eigenfunctions that induces a semiclassical measure μ. Then the distribution of $\{u_{j_k}\}$ at the macroscopic scale is reflected by μ. In particular, there are two extreme cases:

(1). $\mu = \delta_\gamma$ for some closed geodesic orbit γ, that is,

$$\langle \mathrm{Op}_h(a) u_{j_k}, u_{j_k} \rangle_{L^2(\mathbb{M})} \to \int_{S^*\mathbb{M}} a(x,\xi)\, d\delta_\gamma$$

for all $a \in C_0^\infty(T^*\mathbb{M})$. In this case, the eigenfunctions become concentrated near γ and are said to be "scarring" on γ.

(2). $\mu = \mu_L$, the Liouville measure, that is,

$$\langle \mathrm{Op}_h(a) u_{j_k}, u_{j_k} \rangle_{L^2(\mathbb{M})} \to \int_{S^*\mathbb{M}} a(x,\xi)\, d\mu_L$$

for all $a \in C_0^\infty(T^*\mathbb{M})$. In this case, the eigenfunctions become diffused on $S^*\mathbb{M}$ and are said to be "equidistributed".

The Quantum Ergodicity (QE) theorem of Shnirelman[i], Zelditch[ii], and Colin de Verdière[iii] (on manifolds without boundary) and Gérard-Leichtnam[iv] and Zelditch-Zworski[v] (on manifolds with boundary) states

Theorem (Quantum Ergodicity). *Let $\{u_j\}$ be a Laplacian eigenbasis on a manifold with ergodic geodesic flow. Then there is a full density subsequence $\{u_{j_k}\}$ such that the corresponding semiclassical measure is μ_L. Here, $\{u_{j_k}\} \subset \{u_j\}$ has full density if*

$$\lim_{N\to\infty} \frac{\#\{j_k \le N\}}{N} = 1.$$

[i]A. Shnirelman, *The asymptotic multiplicity of the spectrum of the Laplace operator.* Uspehi Mat. Nauk 30 (1975), no. 4 (184), 265–266.

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

[iii]Y. Colin de Verdière, *Ergodicité et fonctions propres du laplacien.* Comm. Math. Phys. 102 (1985), no. 3, 497–502.

[iv]P. Gérard and É. Leichtnam, *Ergodic properties of eigenfunctions for the Dirichlet problem.* Duke Math. J. 71 (1993), no. 2, 559–607.

[v]S. Zelditch and M. Zworski, *Ergodicity of eigenfunctions for ergodic billiards.* Comm. Math. Phys. 175 (1996), no. 3, 673–682.

An immediate consequence of the QE theorem is the equidistribution of eigenfunctions on $\mathbb{M}$: By the Portmanteau theorem in Corollary 6.2.4 and Theorem 6.2.5 of Sogge[i], we have that for all Jordan measurable set $\Omega \subset \mathbb{M}$, i.e., Ω is Borel measurable and $\mathrm{Vol}(\partial\Omega) = 0$,

$$\int_\Omega |u_{j_k}|^2 \, d\mathrm{Vol} \to \frac{\mathrm{Vol}(\Omega)}{\mathrm{Vol}(\mathbb{M})} \quad \text{as } k \to \infty. \tag{2.1}$$

Example (QE on the torus). The QE theorem (in the physical space $\mathbb{M}$ rather than in the phase space $S^*\mathbb{M}$) has also been proved on torus $\mathbb{T}^n$ by Marklof-Rudnick[ii], on which the geodesic flow is integrable (and is therefore not ergodic). In this case, the geodesic flow G_t acts ergodic on functions that depend only on the physical variable:

$$\left\| \frac{1}{T} \int_0^T a(x + t\xi) \, dt \to \int_{\mathbb{T}^n} a(x) \, dx \right\|_{L^2(S^*\mathbb{T}^n)} \to 0 \quad \text{as } T \to \infty.$$

Therefore, the QE theorem also applies to symbols that depend only on x:

$$\langle \mathrm{Op}_h(a) u_{j_k}, u_{j_k} \rangle = \int_{\mathbb{T}^n} a(x) |u_{j_k}|^2 \, dx \to \int_{\mathbb{T}^n} a(x) \, dx$$

for all $a \in C^\infty(\mathbb{T}^n)$. Equivalently, the projection of the semiclassical measure (induced by a full density subsequence of eigenfunctions) is the Lebesgue measure dx.

2.4. L^p-Quantum Ergodicity on the torus: Toral eigenfunctions, Take I

L^p-Quantum Ergodicity asks whether one can replace the smooth symbols in Quantum Ergodicity by L^p functions (in the physical space). See Section 4.11 in Zelditch[iii].

Problem 2.3 (L^p-Quantum Ergodicity). *Let $\{u_j\}$ be a Laplacian eigenbasis on a manifold $\mathbb{M}$. Is there a full density subsequence $\{u_{j_k}\}$ such that*

$$\int_{\mathbb{M}} a(x) |u_{j_k}|^2 \, d\mathrm{Vol} \to \int_{\mathbb{M}} a(x) \, d\mathrm{Vol}$$

for all $a \in L^p(\mathbb{M})$?

An immediate consequence of L^p-QE for any p is that (2.1) is valid for all Borel measurable sets $\Omega \subset \mathbb{M}$.

Example. The only known result on L^p-QE is that L^2-QE holds on $\mathbb{T}^2$ by Hezari-Rivière[iv]. In this case, Zygmund[v] proved that

$$\|u_j\|_{L^4(\mathbb{T}^2)} \le C,$$

in which $C > 0$ is absolute. Hence, $|u_j|^2 \, dx$ is compact in $L^2(\mathbb{T}^2)$. But semiclassical measure (or rather its projection to $\mathbb{T}^2$) corresponding to $\{u_{j_k}\}$ is unique, i.e., the Lebesgue measure.

[i] C. Sogge, *Hangzhou lectures on eigenfunctions of the Laplacian.* Princeton University Press, Princeton, NJ, 2014.

[ii] J. Marklof and Z. Rudnick, *Almost all eigenfunctions of a rational polygon are uniformly distributed.* J.Spectr. Theory 2 (2012), no.1, 107–113.

[iii] S. Zelditch, *Eigenfunctions and nodal sets.* Surveys in differential geometry. Geometry and topology, 237–308, Surv. Differ. Geom., 18, Int. Press, Somerville, MA, 2013.

[iv] H. Hezari and G. Rivière, *Quantitative equidistribution properties of toral eigenfunctions.* J. Spectr. Theory 7 (2017), no. 2, 471–485.

[v] A. Zygmund, *On Fourier coefficients and transforms of functions of two variables.* Studia Math. 50 (1974), 189–201.

Remark (Average of the L^p norms of eigenfunctions). Motivated by the above example, L^p-QE would follow from the uniform bound of L^{2p} norm of eigenfunctions. In fact, since L^p-QE only concerns full density eigenfunctions, it would follow if a full density subsequence of eigenfunctions have uniform L^{2p} norm. It in turn requires the average

$$\frac{1}{N(\lambda)}\sum_{\lambda_j\le\lambda}\|u_j\|_{L^{2p}(\mathbb{M})}<\infty.$$

Here, $N(\lambda)=\#\{j:\lambda\le\lambda\}$ is the eigenvalue counting function.

Example. On spheres, one can construct eigenbasis with large average L^p norm. See Han[i]. This phenomenon is not expected on torus or on manifolds with negative curvature.

We next discuss L^p-QE on $\mathbb{T}^n$ in more details. Let $\mathbb{T}^n=\mathbb{R}^n/2\pi\mathbb{Z}^n$ be the n-dim torus. A toral eigenfunction u in the eigenspace E_λ with eigenvalue λ^2 can be written as

$$u(x)=\sum_{|k|^2=\lambda^2}c_ke^{ik\cdot x},\quad \text{in which } k\in\mathbb{Z}^n \text{ and } c_k\in\mathbb{C}.$$

Here, the eigenvalue

$$\lambda^2=k_1^2+\cdots+k_n^2.$$

Denote $N_\lambda=\dim E_\lambda$. For any eigenbasis $\{u_j\}_{j=1}^{N_\lambda}$, the Quantum Ergodicity (QE) states that there is a full density subset $S(\lambda)\subset\{1,...,N_\lambda\}$ such that for all $a\in C^\infty(\mathbb{T}^n)$,

$$\int_{\mathbb{T}^n}a(x)|u_\lambda|^2\,dx\to\int_{\mathbb{T}^n}a(x)\,dx\quad \text{as } \lambda\to\infty \text{ in } S(\lambda).$$

In the following, we present both directions of potential arguments to prove and to disprove L^p-QE on the torus $\mathbb{T}^n$ with $n\ge 3$.

2.4.1. Proving L^p-QE on the torus.

Unlike on $\mathbb{T}^2$, the L^4 norm of toral eigenfunctions are not uniformly bounded on $\mathbb{T}^n$ for $n\ge 3$. Indeed, Bourgain[ii] conjectured that

Conjecture. *Let $\Delta u=\lambda^2u$ on $\mathbb{T}^n$ with $n\ge 3$. Then*

$$\|u\|_{L^p(\mathbb{T}^n)}\lesssim\begin{cases}\lambda^\varepsilon & \text{if } p\le\frac{2n}{n-2},\\ \lambda^{\frac{n-2}{2}-\frac{n}{p}} & \text{if } p>\frac{2n}{n-2}.\end{cases}$$

While the conjecture is still open, recent advances are achieved by Bourgain-Demeter[iii] using l^2 decoupling argument. However, notice that to prove L^p-QE, we only need the average L^{2p} estimates to be uniformly bounded.

Problem 2.4 (Average L^p norm in an eigenbasis on the torus). *Is the average L^p norm in an eigenbasis uniformly bounded on the torus?*

A positive answer to this problem would lead to $L^{p/2}$-QE on the torus. For example, compute that

$$\|u\|_{L^4(\mathbb{T}^n)}^4 \;=\; \int_{\mathbb{T}^n}\left|u(x)^2\right|^2\,dx$$

[i] X. Han, *Spherical harmonics with maximal L^p $(2<p\le 6)$ norm growth.* J. Geom. Anal. 26 (2016), no. 1, 378–398.

[ii] J. Bourgain, *Eigenfunction bounds for the Laplacian on the n-torus.* Internat. Math. Res. Notices 1993, no. 3, 61–66.

[iii] J. Bourgain and C. Demeter, *The proof of the l^2 decoupling conjecture.* Ann. of Math. (2) 182 (2015), no. 1, 351–389.

$$= \int_{\mathbb{T}^n} \left| \sum_{|k|^2=|j|^2=|\tilde{j}|^2} c_k c_j e^{i(k+j)\cdot x} \right|^2 dx$$
$$= \sum_{\substack{|k|^2=|j|^2=|\tilde{j}|^2=|\tilde{k}|^2=\lambda^2 \\ k+j=\tilde{k}+\tilde{j}}} c_k c_j \overline{c_{\tilde{k}} c_{\tilde{j}}}.$$

Now the average of $\|\cdot\|_{L^4}^4$ in an eigenbasis $\{u\}$ in an eigenspace E_λ would be

$$\frac{1}{N_\lambda} \sum_u \sum_{\substack{|k|^2=|j|^2=|\tilde{j}|^2=|\tilde{k}|^2=\lambda^2 \\ k+j=\tilde{k}+\tilde{j}}} c_k^u c_j^u \overline{c_{\tilde{k}}^u c_{\tilde{j}}^u},$$

which is subject to the normalization and orthogonality conditions:

$$\sum_{|k|^2=\lambda^2} |c_k^u|^2 = 1 \quad \text{and} \quad \sum_{|k|^2=\lambda^2} c_k^u \overline{c_k^v} = 0.$$

2.4.2. Disproving L^p-QE on the torus. It suffices to find a measurable but not Jordan measurable set $\Omega \subset \mathbb{T}^n$ such that for $j \in S(\lambda)$

$$\int_\Omega |u_j|^2 \, dx \not\to \mathrm{Vol}(\Omega) \quad \text{as } \lambda \to \infty.$$

Here, $S(\lambda) \subset \{1, ..., N_\lambda\}$ has positive density.

Notice that it is rather easy to construct semiclassical measures on the torus that is different from the Lebesgue measure. For example, fix $m \in \mathbb{Z}^n$ and define

$$u = \frac{1}{\sqrt{2}} \left(e^{i(k+m)\cdot x} + e^{i(k-m)\cdot x} \right).$$

If $k \cdot m = 0$, then the above function is an eigenfunction with eigenvalue

$$|k+m|^2 = |k-m|^2 = |k|^2 + |m|^2.$$

Now

$$u = \sqrt{2} e^{ik\cdot x} \cos(m \cdot x)$$

clearly fails equidistribution if we let $|k| \to \infty$ (and recall that m is fixed).

However, as mentioned above, the semiclassical measure induced by a full density subset of eigenfunctions has to be the Lebesgue measure. So the above example can only be of zero density in the eigenbasis. Therefore, to disprove L^p-QE, we need to accomplish the following.

(1). Find an irregular set $\Omega \subset \mathbb{T}^n$ (non-Jordan measurable and of positive measure).

(2). Find a positive density subset of eigenfunctions that fails equidistribution on Ω.

We present one potential argument involving the Kakeya set to accomplish the above two conditions. Take Ω as a ε-neighborhood of a Kakeya set. Then Ω contains a tube T_e^ε in each direction $e \in \mathbb{S}^{n-1}$. If we can find $N_\lambda = \dim E_\lambda$ tubes T_k^ε and eigenfunctions w_k that are highly concentrated on T_k^ε, then

$$\int_\Omega |w_k|^2 \geq \int_{T_k^\varepsilon} |w_k|^2 \geq c.$$

But one can construct a Kakeya set such that $|\Omega$ can be arbitrarily small. Therefore, L^p-QE fails at Ω.

Fix $k=(k_1,...,k_n)\in\mathbb{S}_\lambda^{n-1}$. Notice that the eigenfunction

$$u=\sum_{j\in\mathbb{S}_\lambda:|j-k|<\varepsilon}e^{ij\cdot x}$$

concentrates on a tube T_k^ε, *provided* that there are a lot of lattice points $j\in\mathbb{S}_\lambda$ such that $|j-k|<\varepsilon$. In particular, if there are M points j near k, then the L^2-normalized eigenfunction

$$u=\frac{1}{\sqrt{M}}\sum_{j\in\mathbb{S}_\lambda:|j-k|<\varepsilon}e^{ij\cdot x},$$

has lower bound of $\sqrt{M}$ in the tube T_k^ε. This is basically because u does not oscillate in such a tube.

However, the above plan breaks down: On $\mathbb{T}^n$ with $n\geq 5$, there are λ^{n-2} lattice points on $\mathbb{S}_\lambda$. In a ε-neighborhood of a fixed point k, there is only itself as a lattice point. This is because the surface area of $\mathbb{S}_\lambda$ is $O(\lambda^{n-1})$ hence the λ^{n-2} lattice points are very sparse on the sphere $\mathbb{S}_\lambda$. In short, there are not enough Fourier modes for us to use to generate enough eigenfunctions that are concentrated in tubes.

Remark. According to Babich-Lazutkin[i] and Ralston[ii], one can construct approximate eigenfunctions (i.e., quasimodes) that are concentrated near an elliptic closed orbit. They are usually called Gaussian beams with physical concentration $1\times\sqrt{h}$ and frequency uncertainty $h\times\sqrt{h}$. On the torus, all closed geodesics are elliptic so there are abundance of Gaussian beams. It is unclear as how to use them in L^p-QE.

2.5. Quantum scarring in the Bunimovich stadium

Quantum scarring phenomenon concerns the existence of eigenfunctions or approximate eigenfunctions (i.e., quasimodes) that concentrate near proper subset of the domain, for example, near a closed hyperbolic orbit. It has been long observed numerically in the physics circle and is mathematically proved by Hassell[iii] for generic stadiums.

Define the Bunimovich stadium

$$X=R\cup W,\quad \text{in which } R=\left[-\frac{\pi}{2},\frac{\pi}{2}\right]_x\times\left[-\frac{\pi}{2},\frac{\pi}{2}\right]_y \text{ and } R\cap W=R\cap\left\{x=\pm\frac{\pi}{2}\right\}.$$

The billiard flow on S^*X is Anosov and is therefore ergodic. However, it is not uniquely ergodic, since there are vertical trajectories in which the billiard reflects orthogonally off of the two flat edges indefinitely (i.e., the bouncing ball trajectories).

Consider the Dirichlet eigenfunctions $\{u_j\}$. Quantum ergodic theorem states that there exists a full density subsequence of eigenfunctions $\{u_{j_k}\}\subset\{u_j\}$ such that for any measurable subset $\Omega\subset X$ and $\mathrm{Vol}(\partial\Omega)=0$,

$$\lim_{k\to\infty}\int_\Omega|u_{j_k}|^2\to\frac{\mathrm{Vol}(\Omega)}{\mathrm{Vol}(X)}\quad\text{as } k\to\infty.$$

The quantum scarring conjecture (c.f., Conjecture 1.1 in Tao[iv]) states

[i]V. Babich and V. Lazutkin, *Eigenfunctions concentrated near a closed geodesic*, pp. 9–18 in Topics in Mathematical Physics, vol. 2, edited by M. S. Birman, Consultant's Bureau, New York, 1968.

[ii]J. Ralston, *Approximate eigenfunctions of the Laplacian*, J. Differential Geometry 12:1 (1977), 87–100.

[iii]A. Hassell, *Ergodic billiards that are not quantum unique ergodic.* Ann. of Math. (2) 171 (2010), no. 1, 605–619.

[iv]T. Tao, *Structure and randomness.* American Mathematical Society, Providence, RI, 2008.

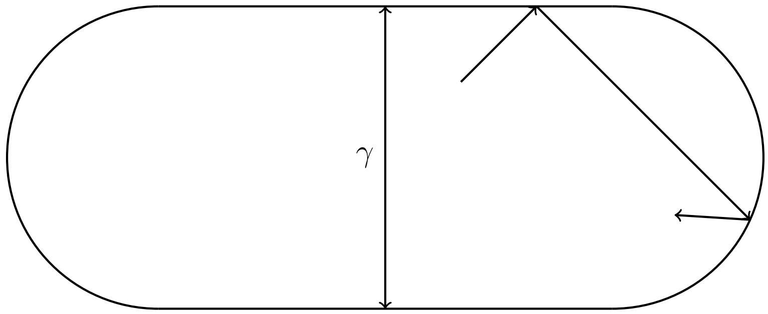

FIGURE 2.1. The billiard flow on a Bunimovich stadium, in which γ is a bouncing ball trajectory.

Problem 2.5 (Quantum scarring in the Bunimovich stadium). *There exists a subset $\Omega \subset X$ and a subsequence of eigenfunctions $\{u_{j_k}\} \subset \{u_j\}$, such that*

$$\lim_{k\to\infty} \int_\Omega |u_{j_k}|^2 \not\to \frac{\mathrm{Vol}(\Omega)}{\mathrm{Vol}(X)} \quad \textit{as } k \to \infty.$$

Informally, the eigenfunctions $\{u_{j_k}\}$ either concentrate (or "scar") in Ω, or on $X \setminus \Omega$.

Consider the one-parameter family of partially rectangular domains $X_t \subset \mathbb{R}^2$:

$$X_t = R_t \cup W, \quad \text{where } R_t = \left[-\frac{t\pi}{2}, \frac{t\pi}{2}\right]_x \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]_y \text{ and } R_t \cap W = R_t \cap \left\{x = \pm\frac{\pi}{2}\right\}.$$

Hassell proved the the phase space version of the quantum scarring conjecture on generic stadiums. The conjecture for all stadiums, as well as the physical space version of the conjecture, remain open.

Theorem (Hassell). *The Laplacian on X_t, with either Dirichlet, Neumann or Robin boundary conditions, is not quantum unique ergodic (non-QUE) for almost everywhere $t \in [1,2]$ (with respect to the Lebesgue measure).*

Consider the (positive) Dirichlet Laplacian $\Delta = \Delta_D$ and its quasimodes (i.e., approximate eigenfunctions)

$$v_n(x,y) = \begin{cases} \chi(x)\sin(ny) & \text{for even } n, \\ \chi(x)\cos(ny) & \text{for odd } n, \end{cases}$$

in which $\operatorname{supp}\chi \in [-\pi/4, \pi/4]$. Hence, v_n is compactly supported in R_t. We can L^2-normalize them by normalizing χ. Then

$$\Delta v_n - n^2 v_n = \begin{cases} \chi''(x)\sin(ny) & \text{for even } n, \\ \chi''(x)\cos(ny) & \text{for odd } n \end{cases}$$

is also compactly supported in the same region within R_t. Since

$$\|(\Delta - n^2)v_n\|_2 \le C$$

with $K > 0$ absolute, v_n is a $O_{L^2}(h^2)$ semiclassical quasimode:

$$\|(h^2\Delta - 1)v_h\|_2 = O(h^2)\|v_h\|_2, \quad \text{in which } h = \frac{1}{n}.$$

Remark. These $O(h^2)$ quasimodes are very special to the stadium domain. In comparison, one can construct $O(h)$ quasimodes and $O(h/|\log h|)$ quasimodes in general domains. It is very rare to find better quasimodes in these general cases.

It is easy to see that v_n are non-QUE since they oscillate in the y-direction and "scar" in the phase space. Because these scarring quasimodes are $O(n^{-2})$, they essentially "lives" in the spectral window $[n^2-C, n^2+C]$ for some constant C. If there are limited number (say bounded by an absolute constant M) of eigenfunctions in such a window, then there must be at least one eigenfunction share the weight of v_n, which results non-QUE of this eigenfunction.

Therefore, the key to prove non-QUE is to show that there exists a subsequence $\{n_j\}$ such that

$$\#\{\text{eigenvalues of } \Delta \text{ in } [n_j^2-2C, n_j^2+2C]\} \leq M \quad \text{uniformly as } j \to \infty.$$

According to the Weyl law, the number of eigenvalues in $[n^2-C, n^2+C]$ is bounded by $O(n)$ so is far from being sufficient for our purpose. However, it also asserts that on average, this number should be uniformly bounded. Hassell's proof contains an important observation that one can find n_j with uniformly bounded number of eigenvalues nearby on generic stadiums.

Indeed, Hassell showed that for almost all $t \in [1,2]$, the uniform boundedness condition holds on the stadiums $\{X_t : t \in [1,2]\}$. It is based on Hadamard variational formula and Quantum Ergodicity on the boundary.

With this in hand, the non-QUE can be finished following an argument by Donnelly[i] and Zelditch[ii].

(1). Let P be the spectral projection of Δ. Then $\|(\Delta-n^2)v_n\|_2 \leq C$ implies that

$$\left\|P_{[n^2-2C,n^2+2C]}v_n\right\|_2^2 \geq \frac{3}{4}.$$

Indeed, since $\{u_j\}$ is an eigenbasis with eigenvalues $\{\lambda_j^2\}$ in $L^2(X_t)$,

$$v_n = \sum_j a_j u_j, \quad \text{in which } \sum_j |a_j|^2 = \|v_n\|_2 = 1.$$

But

$$C^2 \geq \left\|(\Delta-n^2)\sum_{|\lambda_j^2-n^2|>2C} a_j u_j\right\|_2^2 = \sum_{|\lambda_j^2-n^2|>2C} \left\|(\lambda_j^2-n^2)a_j u_j\right\|_2^2 \geq 4C^2 \sum_{|\lambda_j^2-n^2|>2C} |a_j|^2.$$

Hence,

$$\sum_{|\lambda_j^2-n^2|>2C} |a_j|^2 \leq \frac{1}{4} \quad \text{and} \quad \|P_{[n^2-2C,n^2+2C]}v_n\|_2^2 = \sum_{|\lambda_j^2-n^2|\leq 2C} |a_j|^2 \geq \frac{3}{4}.$$

(2). Since the number of eigenvalues in $[n_j^2-2C, n_j^2+2C]$ is uniformly bounded by M as $n_j \to \infty$, there exists an eigenfunction u_{k_j} with eigenvalue n_j^2 such that

$$\langle u_{k_j}, v_{n_j}\rangle \geq \sqrt{\frac{3}{4M}}.$$

(3). To show non-QUE, it remains to find a semiclassical SDO A with $\langle Au_{k_j}, u_{k_j}\rangle$ is uniformly bounded by a positive constant yet $\int \sigma(A) \to 0$. Let A be a self-adjoint semiclassical SDO, properly supported in R_t in both variables x and y, so that

$$\sigma(A) \leq 1 \quad \text{and} \quad \|(\mathrm{I}-A)v_n\|_2 \leq \varepsilon.$$

---

[i] H. Donnelly, *Quantum unique ergodicity.* Proc. Amer. Math. Soc. 131 (2003), no. 9, 2945–2951.

[ii] S. Zelditch, *Note on quantum unique ergodicity.* Proc. Amer. Math. Soc. 132 (2004), no. 6, 1869–1872.

Then

$$\begin{aligned}\langle A^2 u_{k_j}, u_{k_j}\rangle &= \|Au_{k_j}\|_2^2 \geq |\langle Au_{k_j}, v_{n_j}\rangle|^2 \\ &= |\langle u_{k_j}, Av_{n_j}\rangle|^2 \geq \left(|\langle u_{k_j}, v_{n_j}\rangle| - \varepsilon\right)^2 \geq \left(\sqrt{\frac{3}{4M}} - \varepsilon\right)^2.\end{aligned}$$

On the other hand, observe that v_n is semiclassically localized around

$$R_t \times \{(0, \pm 1)\} \subset S^* X_t.$$

Choose $\sigma_A(x, y, \xi, \eta)$ to be compactly supported around $\xi = 0$. Then one can make

$$\int_{S^* X_t} \sigma(A)$$

arbitrarily small.

Notice that Hassell falls short to proving quantum scarring in the physical space, i.e., for some $\Omega \subset X$ such that $|\Omega| < \varepsilon$ (small) while $|u_{j_m}|^2\, dx$ is large on Ω, say

$$\lim_{m\to\infty} \int_\Omega |u_{j_m}|^2 \geq \frac{1}{2} \quad \text{as } m \to \infty.$$

We now take a closed look at the quasimode

$$v_n(x, y) = \frac{1}{\|\varphi\|_2} \varphi(x) \sin(ny),$$

in which $\varphi \in C_0^\infty(-\pi/2, \pi/2)$ and

$$\|\varphi\|_2 = \left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} |\varphi(x)|\, dx\right)^{\frac{1}{2}}.$$

They are L^2-normalized. Moreover,

$$(\Delta - n^2) v_n = \frac{1}{\|\varphi\|_2} \varphi''(x) \sin(ny),$$

and

$$\|(\Delta - n^2) v_n\|_2 \approx \frac{\|\varphi''\|_2}{\|\varphi\|_2}.$$

Example. For $M \gg 1$, take

$$v_n(x, y) = \sqrt{M} \chi(Mx) \sin(ny).$$

Then

$$(\Delta - n^2) v_n = M^{\frac{5}{2}} \varphi''(Mx) \sin(ny),$$

and

$$\|(\Delta - n^2) v_n\|_2 \approx M^2.$$

Hence,

$$\left\| P_{[n^2 - 2M^2, n^2 + 2M^2]} v_n \right\|_2^2 \geq \frac{3}{4}.$$

But there are $O(M^2)$ eigenfunctions in this spectral window, each one gets $O(M^{-2})$ mass weight on average. One can pack M orthogonal quasimodes in R by translation. Hence, each eigenfunction gets $O(M \times M^{-2}) = O(M^{-1})$ mass weight. This can not beat the quantum scarring since each one is concentrated in a $O(M^{-1})$ region.

2.5.1. Burq-Zworski's bouncing ball modes. The quantum scarring in Section 2.5 investigates the possible concentration of eigenfunctions along the bouncing ball trajectories in the rectangular center $R = [0,1] \times [0,a]$ of the Bunimovich stadium $X = R \cup W$. Burq-Zworski[i] proved an interesting result on the opposite direction: The eigenfunction can not concentrate *entirely* in R, that is, the mass has to "spread" over (an open neighborhood of) the wings W. In fact, this result does not depend on the shape of wings.

THEOREM 2.3. *Let $\Delta = -(\partial_x^2 + \partial_y^2)$ be the Laplace operator and V be any open neighborhood of the closure of W in X. Then for any quasimode v in X,*

$$(\Delta - \lambda^2)v = F \text{ in } X \quad \text{and} \quad v = 0 \text{ on } \partial X,$$

we have that

$$\int_X |v(x,y)|^2 \, dxdy \leq C \left(\int_X |F(x,y)|^2 \, dxdy + \int_V |v(x,y)|^2 \, dxdy \right)$$

for some positive constant $C = C(V)$.

Remark. Take $F = 0$. Then v is an eigenfunction with eigenvalue λ^2. In this case, we have that

$$\int_V |v(x)|^2 \, dx \geq C$$

for some positive constant $C = C(V)$.

We first need a proposition.

Proposition. *For any open $\omega \subset R = [0,1]_x \times [0,a]_y$ of the form*

$$\omega = \omega_x \times [0,a]_y,$$

there exists C such that the solution of

$$\begin{cases} (\Delta - \lambda^2)u = f + \partial_x g & \text{in } R, \\ u = 0 & \text{on } \partial R, \end{cases}$$

with an arbitrary $\lambda \geq 0$, we have

$$\int_R |u(x,y)|^2 \, dxdy \leq C \left(\int_R |f(x,y)|^2 \, dxdy + \int_R |g(x,y)|^2 \, dxdy + \int_\omega |u(x,y)|^2 \, dxdy \right).$$

PROOF.

(1). Using the Dirichlet eigenfunctions

$$e_k(y) = \sqrt{\frac{2}{a}} \sin\left(\frac{2k\pi y}{a} \right),$$

we decompose u, f, and g:

$$\begin{cases} u(x,y) = \sum_k u_k(x) e_k(y), \\ f(x,y) = \sum_k f_k(x) e_k(y), \\ g(x,y) = \sum_k g_k(x) e_k(y). \end{cases}$$

Then the equation $(\Delta - \lambda^2)u^2 = f + \partial_x g$ is reduced to

$$\sum_k e_k(y) \left(-u_k''(x) + \left(\frac{2k\pi}{a} \right)^2 u_k(x) - \lambda^2 u_k(x) \right) = \sum_k e_k(y) \left(f_k(x) + g_k'(x) \right).$$

[i] N. Burq and M. Zworski, *Bouncing ball modes and Quantum Chaos.* SIAM Rev. 47 (2005), no. 1, 43–49.

Letting

$$z = \lambda_1^2 = \lambda^2 - \left(\frac{2k\pi}{a}\right)^2,$$

u_k solves the equation

$$\begin{cases} -(d_x^2 + z)u_k = f_k + g_k' & \text{in } [0,1], \\ u_k(0) = u_k(1) = 0. \end{cases}$$

(2). If $\Im\lambda_1 \le 1$, that is, $z = \lambda_1^2 \ge -1$, then

$$\lambda^2 - \left(\frac{2k\pi}{a}\right)^2 = \lambda_1^2 \ge -1.$$

Assume that $\omega_x = (0, \delta)$. Choose $\chi \in C_0^\infty([0,1])$ such that $\chi = 0$ near 0 and equals 1 on $[\frac{\delta}{2}, 1]$. Then

$$\begin{aligned} (d_x^2 + \lambda_1^2)(\chi u_k) &= \chi(d_x^2 + \lambda_1^2)u_k + 2\chi' u_k' + \chi'' u_k \\ &= -\chi(f_k + g_k') + 2\chi' u_k' + \chi'' u_k \\ &:= F_k. \end{aligned}$$

But we have the explicit solution

$$\chi(x)u_k(x) = \frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))F_k(s)\,ds.$$

This is because

$$\begin{cases} d_x\left(\frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))F_k(s)\,ds\right) = \int_0^x \cos(\lambda_1(x-s))F_k(s)\,ds; \\ d_x^2\left(\frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))F_k(s)\,ds\right) = F_k - \lambda_1\int_0^x \sin(\lambda_1(x-s))F_k(s)\,ds; \\ (d_x^2 + \lambda_1^2)\left(\frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))F_k(s)\,ds\right) = F_k(x). \end{cases}$$

By integration by parts,

$$\begin{aligned} & \chi(x)u_k(x) \\ = & \frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))F_k(s)\,ds \\ = & \frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))(-\chi(f_k + g_k') + 2\chi' u_k' + \chi'' u_k)(s)\,ds \\ = & -\frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))\chi(s)f_k(s)\,ds - \frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))\chi(s)g_k'(s)\,ds \\ & +\frac{2}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))\chi'(s)u_k'(s)\,ds + \frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))\chi''(s)u_k(s)\,ds \\ = & -\frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))\chi(s)f_k(s)\,ds + \frac{1}{\lambda_1}\int_0^x \left[\sin(\lambda_1(x-s))\chi(s)\right]' g_k(s)\,ds \\ & -\frac{2}{\lambda_1}\int_0^x \left[\sin(\lambda_1(x-s))\chi'(s)\right]' u_k(s)\,ds + \frac{1}{\lambda_1}\int_0^x \sin(\lambda_1(x-s))\chi''(s)u_k(s)\,ds \\ = & -\int_0^x \frac{\sin(\lambda_1(x-s))}{\lambda_1}\chi(s)f_k(s)\,ds \\ & -\int_0^x \cos(\lambda_1(x-s))\chi(s)g_k(s)\,ds + \int_0^x \frac{\sin(\lambda_1(x-s))}{\lambda_1}\chi'(s)g_k(s)\,ds \end{aligned}$$

$$+2\int_0^x \cos(\lambda_1(x-s))\chi'(s)u_k(s)\,ds - \int_0^x \frac{2\sin(\lambda_1(x-s))}{\lambda_1}\chi''(s)u_k(s)\,ds$$
$$+\int_0^x \frac{\sin(\lambda_1(x-s))}{\lambda_1}\chi''(s)u_k(s)\,ds.$$

Hence,

$$\begin{aligned}
&|\chi(x)u_k(x)|^2\\
\leq\ & C\Big(\int_0^x |f_k(s)|^2\,ds + \int_0^x |g_k(s)|^2\,ds + \int_0^\delta |g_k(s)|^2\,ds\\
&+2\int_0^\delta |u_k(s)|^2\,ds + \int_0^\delta |u_k(s)|^2\,ds + \int_0^\delta |u_k(s)|^2\,ds\Big)\\
\leq\ & C\left(\int_0^1 |f_k(s)|^2\,ds + \int_0^1 |g_k(s)|^2\,ds + \int_0^\delta |u_k(s)|^2\,ds\right).
\end{aligned}$$

Now

$$\begin{aligned}
\int_0^1 |u_k(x)|^2\,dx &= \int_0^{\frac{\delta}{2}} |u_k(x)|^2\,dx + \int_{\frac{\delta}{2}}^1 |u_k(x)|^2\,dx\\
&\leq \int_0^\delta |u_k(x)|^2\,dx + \int_{\frac{\delta}{2}}^1 |\chi(x)u_k(x)|^2\,dx\\
&\leq C\left[\int_0^1 |f_k(x)|^2\,ds + \int_0^1 |g_k(x)|^2\,ds + \int_0^\delta |u_k(x)|^2\,ds\right].
\end{aligned}$$

The proposition then follows by summing up k in the case of $\Im\lambda_1 \leq 1$.

(3). If $\Im\lambda_1 > 1$, that is, $z = \lambda_1^2 \leq -1$, then

$$\lambda^2 - \left(\frac{2k\pi}{a}\right)^2 = \lambda_1^2 \leq -1.$$

A stronger conclusion holds in this case (omitting the subscript k):

$$\begin{aligned}
\int_0^1 \left(f(x)\overline{u(x)} - g(x)\overline{d_x u(x)}\right)\,dx &= \int_0^1 (f(x) + d_x g(x))\,\overline{u(x)}\,dx\\
&= \int_0^1 \left(-d_x^2 - z\right) u(x)\overline{u(x)}\,dx\\
&= \int_0^1 \left(|d_x u(x)|^2 + |z||u(x)|^2\right)\,dx.
\end{aligned}$$

But the left-hand side

$$\begin{aligned}
&\left|\int_0^1 \left(f(x)\overline{u(x)} - g(x)\overline{d_x u(x)}\right)\,dx\right|\\
\leq\ & \left(\int_0^1 \left(|f(x)|^2 + |g(x)|^2\right)\,dx\right)^{\frac{1}{2}} \left(\int_0^1 \left(|u(x)|^2 + |d_x u(x)|^2\right)\,dx\right)^{\frac{1}{2}}.
\end{aligned}$$

Therefore,

$$\int_0^1 \left(|d_x u(x)|^2 + |u(x)|^2\right)\,dx$$

$$\leq \int_0^1 \left(|d_x u(x)|^2 + |z||u(x)|^2\right) dx$$
$$\leq \left(\int_0^1 \left(|f(x)|^2 + |g(x)|^2\right) dx\right)^{\frac{1}{2}} \left(\int_0^1 \left(|u(x)|^2 + |d_x u(x)|^2\right) dx\right)^{\frac{1}{2}}.$$

Then
$$\int_0^1 |u_k(x)|^2 \, dx \leq \int_0^1 \left(|d_x u_k(x)|^2 + |u_k(x)|^2\right) dx \leq \int_0^1 \left(|f_k(x)|^2 + |g_k(x)|^2\right) dx.$$

□

Now we prove Theorem 2.3.

PROOF OF THEOREM 2.3. Consider u and F satisfying $(\Delta - \lambda^2)v = F$ and $v = 0$ on ∂X. Set $\chi(x) \in C_0^\infty(0,1)$ which equals 1 on $[\varepsilon, 1-\varepsilon]$. Then
$$(\Delta - \lambda^2)\chi v = \chi F - 2\chi' \partial_x v - \chi'' v \quad \text{in } R$$
with Dirichlet boundary conditions on ∂R.

Apply the above proposition with $f = \chi F + \chi'' v$ and $g = -2\chi' v$ to obtain
$$\int_R |\chi(x) v(x,y)|^2 \, dxdy \leq C \left(\int_R |\chi(x) f(x,y)|^2 \, dxdy + \int_R |g(x,y)|^2 \, dxdy + \int_{\omega_\varepsilon} |v(x,y)|^2 \, dxdy\right),$$
in which ω_ε is a neighborhood of the support χ'. Since we can choose it to be contained in $R \cap V$, the theorem follows. □

2.6. Quantum Unique Ergodicity

Rudnick-Sarnak[i] conjectured that

Conjecture (Rudnick-Sarnak). *On negatively curved manifolds, the semiclassical measure induced by eigenfunctions is the Liouville measure. That is, all eigenfunctions tend to be equidistributed in the phase space.*

This is of course a stronger statement than the QE theorem, as the latter asserts the conclusion for a full density subsequence of eigenfunctions. Therefore, the above conjecture is called the Quantum Unique Ergodicity (QUE) conjecture. Let μ be a semiclassical measure on a negatively curved manifold $\mathbb{M}$. We know that μ is supported on $S^*\mathbb{M}$ and is invariant under the geodesic flow G_t. Various results on further characterization of μ are known, for example,

(1). Anatharaman[ii] and Anatharaman-Nonnenmacher[iii] proved that μ must have positive entropy. In particular, it implies that the arc-length measures can not be a semiclassical measure, because they have zero entropy, see Section 1.3.1.

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

[ii] N. Anantharaman, *Entropy and the localization of eigenfunctions.* Ann. of Math. (2) 168 (2008), no. 2, 435–475.

[iii] Anantharaman and S. Nonnenmacher, *Half-delocalization of eigenfunctions for the Laplacian on an Anosov manifold.* Festival Yves Colin de Verdière. Ann. Inst. Fourier (Grenoble) 57 (2007), no. 7, 2465–2523.

(2). On surfaces with negative curvature, Dyatlov-Jin[i] and Dyatlov-Jin-Nonnenmacher[ii] proved that μ must have full support on $S^*\mathbb{M}$.

See Dyatlov[iii] for a recent survey on the QUE conjecture.

To attack the QUE conjecture, we ask

Problem 2.6 (Semiclassical measure via quantitative unique continuation). *Can one provide a proof that μ cannot be the arc-length measure on a closed geodesic, based on quantitative unique continuation of eigenfunctions (see Chapter 5) and hyperbolic dynamics?*

Chatterjee-Galkowski[iv] proved that a small perturbation of the Laplacian with Dirichlet boundary condition in a domain satisfies QUE. That is, let $\Omega \subset \mathbb{R}^n$ be a bounded domain. Then for any $\varepsilon > 0$, there is a linear operator

$$\|S_\varepsilon\|_{L^2(\overline{\Omega}) \to L^2(\overline{\Omega})} \leq \varepsilon$$

such that $(\mathrm{I} + S_\varepsilon)\Delta$ is QUE. The idea is based on an observation that a random combination of eigenfunctions is almost surely QUE. Indeed, let $\{u_j\}$ be an eigenbasis with eigenvalues $\lambda_j^2 \to \infty$, $\Delta u_j = \lambda_j^2 u_j$. Divide $[0, \infty)$ into spectral windows $|\lambda_j - \lambda| \leq L$ of width L. Then

$$u = \sum_{|\lambda_j - \lambda| \leq L} a_j u_j$$

is QUE almost surely with respect to random coefficients a_j. Furthermore, choose S_ε so that

$$(\mathrm{I} + S_\varepsilon)\Delta u = \lambda^2 u,$$

that is,

$$\sum_{|\lambda_j - \lambda| \leq L} a_j \lambda_j^2 u + a_j \lambda_j^2 S_\varepsilon u_j = \sum_{|\lambda_j - \lambda| \leq L} a_j \lambda^2 u.$$

This can be satisfied by setting

$$S_\varepsilon u_j = \frac{\lambda^2 - \lambda_j^2}{\lambda_j^2} u_j \approx_{L^2(\overline{\Omega})} u_j,$$

if the width L of the spectral window is chosen small enough. We remark that S_ε is independent of the random coefficients a_j. Therefore, by choosing $\varepsilon = \varepsilon(L)$ properly, $S_\varepsilon : L^2(\overline{\Omega}) \to L^2(\overline{\Omega})$ has small mapping norm, while we also maintain the QUE property of the random eigenfunction $u = \sum a_j u_j$.

This approach does not seem to apply to generic Laplacian or Laplacian with respect to a generic hyperbolic metric. Therefore, we ask

Problem 2.7 (QUE on generic hyperbolic surfaces). *Can one show that QUE is a generic property among all compact hyperbolic surfaces, or find a hyperbolic surface on which QUE holds?*

---

[i] S. Dyatlov and L. Jin, *Semiclassical measures on hyperbolic surfaces have full support.* Acta Math. 220 (2018), no. 2, 297–339.

[ii] S. Dyatlov, L. Jin, and S. Nonnenmacher, *Control of eigenfunctions on surfaces of variable curvature.* J. Amer. Math. Soc. 35 (2021), no. 2, 361–465.

[iii] S. Dyatlov, *Macroscopic limits of chaotic eigenfunctions.* arXiv:2109.09053. To appear in Proceedings of the International Congress of Mathematicians - Moscow 2022.

[iv] S. Chatterjee and J. Galkowski, *Arbitrarily small perturbations of Dirichlet Laplacians are quantum unique ergodic.* J. Spectr. Theory 8 (2018), no. 3, 909–947.

2.6.1. Control of eigenfunctions. This section follows Dyatlov-Jin[i].

THEOREM 2.4 (Control of eigenfunctions on compact hyperbolic surfaces). *Suppose that $\mathbb{M}$ is a compact hyperbolic surface. Let $a \in C_0^\infty(T^*\mathbb{M})$ and $a \not\equiv 0$ on $S^*\mathbb{M}$. Then there are positive constants $C = C(a)$ and $h_0 = h_0(a)$ such that*

$$\|u\|_{L^2(\mathbb{M})} \leq C\|\mathrm{Op}_h(a)u\|_{L^2(\mathbb{M})},$$

for all eigenfunctions $u = u(h)$ on $\mathbb{M}$, $(h^2\Delta - 1)u(h) = 0$, and $0 < h < h_0$.

Remark.

- A direct consequence of Theorem 2.4 is that any semiclassical measure μ on a hyperbolic surface $\mathbb{M}$ has full support, i.e., $\operatorname{supp}\mu = S^*\mathbb{M}$.
- Theorem 2.4 is a "control" result, as the eigenfunction u on $\mathbb{M}$ is controlled by $\mathrm{Op}_h(a)u$ localized by a. This theorem and its corollary about the support of semiclassical measures in fact hold for approximate eigenfunctions (i.e., quasimodes) of certain order. Indeed, Dyatlov-Jin proved that
$$\|u\|_{L^2(\mathbb{M})} \leq C\|\mathrm{Op}_h(a)u\|_{L^2(\mathbb{M})} + \frac{C|\log h|}{h}\|(h^2\Delta - 1)u\|_{L^2(\mathbb{M})}.$$
Therefore, the control results hold for quasimodes of order $o(h/|\log h|)$.
- These results have been generalized to surfaces with variable negative curvature by Dyatlov-Jin-Nonnenmacher[ii].

The starting point of the control argument to prove Theorem 2.4 is Egorov's theorem in Theorem 2.2: Denote by $U_t = e^{-it\sqrt{\Delta}}$ the Schrödinger propagator and by G_t the geodesic flow on $S^*\mathbb{M}$. Let $a \in C_0^\infty(T^*\mathbb{M})$. Then by Egorov's theorem,

$$U_{-t}\mathrm{Op}_h(a)U_t \approx \mathrm{Op}_h(a \circ G_t) \quad \text{on } L^2(\mathbb{M}) \text{ for } |t| \leq T.$$

Since u is a Laplacian eigenfunction with eigenvalue h^{-2},

$$U_t(u) = e^{-it\sqrt{\Delta}}u = e^{-it/h}u \quad \text{for any } t \in \mathbb{R}.$$

in which $e^{-ith} \in \mathbb{C}$. Therefore,

$$\|U_{-t}\mathrm{Op}_h(a)U_t u\|_{L^2(\mathbb{M})} = \|\mathrm{Op}_h(a)u\|_{L^2(\mathbb{M})} \quad \text{for any } t \in \mathbb{R}.$$

For such reason, eigenfunctions are called the stationary states in the quantum system.

We now strive to accomplish the following two goals.

(1). Take $|t| < T$ for T as large as possible in Egorov's theorem. Doing so, we hope that
$$\bigcup_{|t|\leq T} G_t(\operatorname{supp} a)$$
covers as much ground as possible on $S^*\mathbb{M}$. Since the eigenfunctions are supported near $S^*\mathbb{M}$, if we can cover the whole $S^*\mathbb{M}$ within some $T < \infty$ that is independent of h, then the control result follows. (In fact, the control result follows under such condition on any manifold.) Unfortunately, this is in general never possible on hyperbolic surfaces. For example, let γ be a closed orbit such that $\operatorname{supp} a \cap \gamma = \emptyset$. Then $G_t(\operatorname{supp} a)$ would never cover γ regardless of how large T is.

[i]S. Dyatlov and L. Jin, *Semiclassical measures on hyperbolic surfaces have full support.* Acta Math. 220 (2018), no. 2, 297–339.

[ii]S. Dyatlov, L. Jin, and S. Nonnenmacher, *Control of eigenfunctions on surfaces of variable curvature.* J. Amer. Math. Soc. 35 (2021), no. 2, 361–465.

(2). Control the L^2 mass of u in regions that do not intersect $\cup_{|t|\le T} G_t(\operatorname{supp} a)$. This requires the "fractal uncertainty principle (FUP)" developed by Dyatlov and his coworkers, see Chapter 7. That is, the regions that do not pass $\operatorname{supp} a$ under the flow have to display certain fractal structure, because of the hyperbolicity of the geodesic flow. As T is sufficiently large, this fractal structure is fine enough such that the L^2 mass of u in these regions is always small, according to the FUP.

However, for the FUP to take effect, we need Egorov's theorem for longer times on hyperbolic surfaces: Let $a \in C_0^\infty(T^*\mathbb{M})$ and $|t| \le T_\rho = \rho|\log h|$ with $0 \le \rho < 1$. (Later, ρ will be chosen close to 1.) Then for all $|t| \le T_\rho$,

$$\|U_{-t}\mathrm{Op}_h(a)U_t - \mathrm{Op}_h(a\circ G_t)\|_{L^2(\mathbb{M})\to L^2(\mathbb{M})} = O_{L^2(\mathbb{M})\to L^2(\mathbb{M})}\left(h^{1-\rho}\right).$$

Remark (Symbol classes). Since the geodesic flow G_t is hyperbolic, $a\circ G_t$ for $t \ge 0$ concentrate near the stable foliation E_s. Moreover, the expansion rate of G_t on hyperbolic surface is uniform and equals 1. This implies that $\operatorname{supp}(a\circ G_t)$ is in the e^{-t} neighborhood of E_s.

As a consequence, beginning from a symbol $a \in C_0^\infty(T^*\mathbb{M}) \subset S_0(\mathbb{M})$, the functions $a\circ G_t$ with $t = t(h)$ may not be in the same symbol class $S_0(\mathbb{M})$ anymore. Indeed, for $t \le T_\rho = \rho|\log h|$,

$$\sup_{z\in S^*\mathbb{M}} |Y_1\cdots Y_m Z_1\cdots Z_k\,(a\circ G_t)(z)| \le Ce^{tk} \le Ch^{-\rho k},$$

in which $Y_1,...,Y_m, Z_1,...,Z_k$ are vector fields in $T^*\mathbb{M}$ such that $Y_1,...,Y_m$ are tangent to the stable foliation E_s. That is, $a\circ G_t$ is smooth in the stable directions and is singular (with respect to h) in the transverse directions. Hence,

- if $\rho < \frac{1}{2}$, then $a\circ G_t \in S_{\frac{1}{2}}(\mathbb{M})$ and the semiclassical analysis in Section 2.1 applies,
- if $\frac{1}{2} \le \rho \le 1$, then Dyatlov et al defined a class of anisotropic symbols $S_{E_s,\rho}(\mathbb{M})$ and develop a calculus $\Psi_{E_s,\rho}(\mathbb{M})$ for the corresponding pseudodifferential operator that is similar to Section 2.1. In particular, two such operators can compose, resulting an operator in the same calculus.

The situation for $t \le 0$ with respect to the unstable foliation E_u is similar.

In the view of the regularity of the symbols $a\circ G_t$, we divide the time frame $-\rho|\log h| \le t \le \rho|\log h|$ into three periods.

(I). $-\rho|\log h| \le t \le -|\log h|/2$. The symbol $a\circ G_t \in S_{E_u,\rho}(\mathbb{M})$ with $\rho > \frac{1}{2}$.
(II). $-|\log h|/2 \le t \le |\log h|/2$. The symbol $a\circ G_t \in S_{\frac{1}{2}}(\mathbb{M})$.
(III). $|\log h|/2 \le t \le \rho|\log h|$. The symbol $a\circ G_t \in S_{E_s,\rho}(\mathbb{M})$ with $\rho > \frac{1}{2}$.

We return to the control argument.

(1). Take $|t| < T_\rho = \rho|\log h|$ for ρ close to 1. Then by the semiclassical analysis for symbols in $S_{\frac{1}{2}}(\mathbb{M})$, one can control the L^2 mass of u in $\cup_{|t|\le T} G_t(\operatorname{supp} a)$.
(2). The region that does not pass through $\operatorname{supp} a$ within T is

$$\bigcap_{|t|\le T_\rho} G_t(S^*\mathbb{M}\setminus \operatorname{supp} a).$$

Let $a_2 = 1 - a$. Then the L^2 mass on such a region can be controlled by (taking t to be discrete)

$$\left\|\prod_{|t|\le T_\rho}\mathrm{Op}_h(a_2\circ G_t)u\right\|_{L^2} = \left\|\prod_{t=-T_\rho}^{-1}\mathrm{Op}_h(a_2\circ G_t)\prod_{t=0}^{T_\rho}\mathrm{Op}_h(a_2\circ G_t)u\right\|.$$

Notice that $a_2 \circ G_t \in S_{E_s,\rho}(\mathbb{M})$ as $t \to \rho|\log h|$ while $a_2 \circ G_t \in S_{E_u,\rho}(\mathbb{M}))$ as $t \to -\rho|\log h|$. The two corresponding anisotropic pseudodifferential operators are in two different calculus $\Psi_{E_s,\rho}(\mathbb{M})$ and $\Psi_{E_s,\rho}(\mathbb{M})$, respectively. They cannot compose according to the semiclassical rules as before.

Notice further that $a_2 \circ G_t$ localizes in an h^ρ-neighborhood of $S_{E_s,\rho}(\mathbb{M})$ as $t \to \rho|\log h|$ while a_2 localizes in an h^ρ-neighborhood of $S_{E_u,\rho}(\mathbb{M}))$ as $t \to -\rho|\log h|$. Interpret E_s and E_u as directions of position and frequency. Then the function u cannot simultaneously concentrate in both neighborhoods. Therefore, the fractal uncertainty principle (FUP) applies to imply that the above term is bounded by h^β for some $\beta > 0$. Here, the FUP takes account for the fractal structure of the neighborhoods of E_s and E_u, as a consequence of the hyperbolicity of the flow.

2.6.1.1. *Control of eigenfunctions with a logarithmical loss.* In this section, we present a weaker control result

$$\|u\|_{L^2(\mathbb{M})} \leq C|\log h|\|\mathrm{Op}_h(a)u\|_{L^2(\mathbb{M})},$$

which has a loss of a log factor comparing with the one in Theorem 2.4.

We set up the control process:

- Let $a_1 = a$ and $a_2 = 1 - a_1$ so $a_1 + a_2 = 1$ near $S^*\mathbb{M}$. Denote $A_1 = \mathrm{Op}_h(a_1)$ and $A_2 = \mathrm{Op}_h(a_2)$. (In practice, one would need to modify the definition to accommodate the smoothness requirement of symbols.) Since eigenfunctions $u(h)$ are concentrated near $S^*\mathbb{M}$, $(A_1 + A_2)u = u + O_{L^2}(h^\infty)$.
- Denote $A_1(t) = U_{-t}A_1U_t$ and $A_2(t) = U_{-t}A_2U_t$. Then by Egorov's theorem, $\sigma(A_1(t)) = a_1 \circ G_t$ and $\sigma(A_2(t)) = a_2 \circ G_t$. Moreover, for eigenfunctions u,
$$(A_1(t) + A_2(t))u = (A_1 + A_2)u = u + O_{L^2}(h^\infty).$$
- Define a word ω of length $|\omega| = N$ to be a sequence $\{w_j\}_{j=0}^{N-1}$ such that $w_j \in \{1, 2\}$. Write $\mathcal{W}_N = \{\omega : |\omega| = N\}$. Then
$$(A_1(N-1) + A_2(N-1)) \cdots (A_1(1) + A_2(1))(A_1 + A_2) = \sum_{\omega \in \mathcal{W}_N} A_\omega,$$
in which $A_\omega = A_{w_{N-1}}(N-1) \cdots A_{w_1}(1)A_{w_0}(0)$ has principle symbol
$$a_\omega := \sigma(A_\omega) = \prod_{j=0}^{N-1} a \circ G_{w_j}.$$
- Let $N_\rho = \lfloor \rho|\log h| \rfloor$. Then
$$u + O_{L^2}(h^\infty) = \sum_{|\omega|=2N_\rho} A_\omega u = \sum_{w_j=1 \text{ for some } j=0,\dots,2N_\rho-1} A_\omega u + A_2(2N_\rho - 1) \cdot A_2(1)A_2u.$$

(1). The summation in the right-hand-side of the above equation can be divided into $2N_\rho$ groups. This is effectively the "controlled" region since they all pass through $\operatorname{supp} a_1$. Here, the j-th group is the collection of words whose j-th terms equal 1. Therefore, the total L^2 mass in each group is always bounded by $\|A_1u\|_{L^2}$. Using the triangle inequality,

$$\left\| \sum_{w_j=1 \text{ for some } j=0,\dots,2N_\rho-1} A_\omega u \right\|_{L^2(\mathbb{M})} \leq 2N_\rho\|A_1u\|_{L^2(\mathbb{M})} \leq C|\log h|\|A_1u\|_{L^2(\mathbb{M})}.$$

(2). The operator $A_2(2N_\rho - 1)\cdot A_2(1)A_2$ has principle symbol

$$\prod_{t=0}^{2N_\rho-1} \mathrm{Op}_h(a_2\circ G_t).$$

So it is not the product of two operators in $\Psi_{E_s,\rho}(\mathbb{M})$ and $\Psi_{E_u,\rho}(\mathbb{M})$ yet and we cannot apply FUP directly to it. However, this can be easily fixed by flowing the whole term backward for time N_ρ:

$$\begin{aligned} & \|A_2(2N_\rho-1)\cdot A_2(1)A_2u\|_{L^2(\mathbb{M})} \\ = & \left\|U_{N_\rho}A_2(2N_\rho-1)\cdot A_2(1)A_2U_{-N_\rho}u\right\|_{L^2(\mathbb{M})} \\ = & \|A_2(N_\rho-1)\cdot A_2(1)A_2A_2(-1)\cdots A_2(-(N_\rho-1))A_2(-N_\rho)u\|_{L^2(\mathbb{M})} \\ = & \left\|\prod_{0}^{t=N_\rho-1}\mathrm{Op}_h(a_2\circ G_t)\prod_{-1}^{t=-N_\rho}\mathrm{Op}_h(a_2\circ G_t)u\right\|_{L^2(\mathbb{M})} \\ \ll & h^\beta, \end{aligned}$$

by the FUP.

Combing the estimates in (1) and (2),

$$\|u\|_{L^2(\mathbb{M})}\leq \sum_{w_j=1 \text{ for some } j=1,\dots,2N_\rho}\|A_\omega u\|_{L^2}+\|A_2(2N_\rho-1)\cdot A_2(1)A_2u\|_{L^2(\mathbb{M})}\leq C|\log h|\|A_1u\|_{L^2(\mathbb{M})}.$$

2.6.1.2. *Control of eigenfunctions without loss.* To improve the coefficient of $\|A_1u\|_{L^2}$, one would need to

(1). Move some terms from the controlled part (1), i.e., $\{|\omega|=2N_\rho : w_j=1 \text{ for some } j\}$, to the FUP part (2). Let $\alpha\in(0,1)$ and define

$$F(\omega)=\frac{\#\{j: w_j=1\}}{N},$$

that is, $F(\omega)$ is the proportion of terms in a word $\omega=\{w_0,\dots,w_{N-1}\}$ which equal 1. Then the words ω such that $F(\omega)\geq\alpha$ are the ones that at least αN of them equal 1. Since $a_1+a_2=1$ on $S^*\mathbb{M}$,

$$1=(a_1+a_2)^N=\sum_{|\omega|=N}a_\omega=\sum_{F(\omega)\geq\alpha}a_\omega+\sum_{F(\omega)<\alpha}a_\omega,$$

in which the second summation can still be controlled by the FUP. (Setting $\alpha=0$ reduces to the discussion in the previous subsection.) While the words in the first summation spend at least αN times around a_1, therefore the L^2 mass of u corresponding to these terms are expected to be controlled by $\|A_1u\|_{L^2}$.

(2). Use better estimate than the triangle inequality to estimate the L^2 mass of u in the first summation above. The crucial idea is to compare the L^2 mass of u that correspond to functions of word collections. More precisely, for any "linear" function $c:\mathcal{W}_N\to\mathbb{C}$, define

$$A_c=\sum_{\omega\in\mathcal{W}_N}c(\omega)A_w.$$

Given two functions c and d such that $|d|\leq c$ on $\mathcal{W}_N$, we have that

$$\|A_du\|_{L^2(\mathbb{M})}\leq\|A_cu\|_{L^2(\mathbb{M})}.$$

Take $c = F$ and $d = 1_{F(\omega)\geq\alpha}$, i.e., the indicator function of the set $\{\omega : F(\omega) \geq \alpha\}$. Then $\alpha d \leq F$ by their definitions. Therefore,

$$\left\| \sum_{F(\omega)\geq\alpha} A_\omega u \right\|_{L^2(\mathbb{M})} \leq C\|A_d u\|_{L^2(\mathbb{M})} \leq \frac{C}{\alpha}\|A_F u\|_{L^2(\mathbb{M})} \leq \frac{C}{\alpha}\|A_1 u\|_{L^2(\mathbb{M})},$$

in which the last inequality follows because there is a $1/N$ factor in F.

Combining the estimates in (1) and (2),

$$\|u\|_{L^2(\mathbb{M})} \leq \left\| \sum_{F(\omega)\geq\alpha} A_\omega u \right\|_{L^2(\mathbb{M})} + \left\| \sum_{F(\omega)<\alpha} a_\omega u \right\|_{L^2(\mathbb{M})} \leq \frac{C}{\alpha}\|A_1 u\|_{L^2(\mathbb{M})} + h^\beta \leq C\|A_1 u\|_{L^2(\mathbb{M})}.$$

2.7. Mixed and KAM systems

The integrable systems and ergodic systems, in some sense, represent the two extreme cases of behaviors in classical dynamics. For example, the phase space in an integrable system is foliated by a family of invariant tori, whereas in an ergodic system, the phase space is filled by an invariant set (except a measure zero region).

In addition to these two extreme cases, there are systems which display the two dynamical behaviors in separate regions of the phase space, for example, the billiard system in a mushroom:

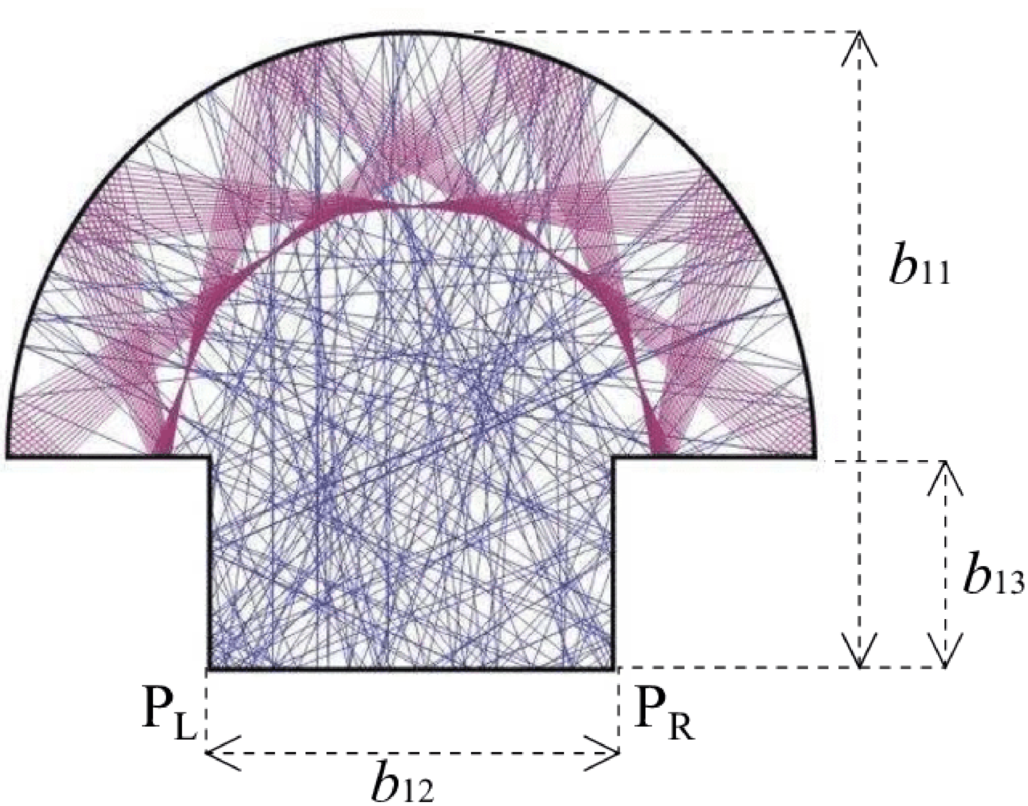

FIGURE 2.2. A mushroom billiard system (Credit: Kryzhevich)

The phase space in a mushroom billiard system splits into two components: The integrable part A when the billiard is confined in the semi-annulus (e.g., the purple trajectories) and the ergodic part B elsewhere (e.g., the blue trajectories). Percival's Conjecture, roughly speaking, predicts that the quantum system follows the same splitting, that is, an eigenbasis $\{u_j\}_{j=1}^{\infty}$ splits into two subsequences S_A and S_B such that

(1). all semiclassical measures induced by S_A are supported on A,
(2). the semiclassical measure induced by S_B is the Liouville measure on B, i.e., S_B tend to be equidistributed on B,
(3). the ratio of the densities of S_A and S_B equals the ratio of the volumes of A and B.

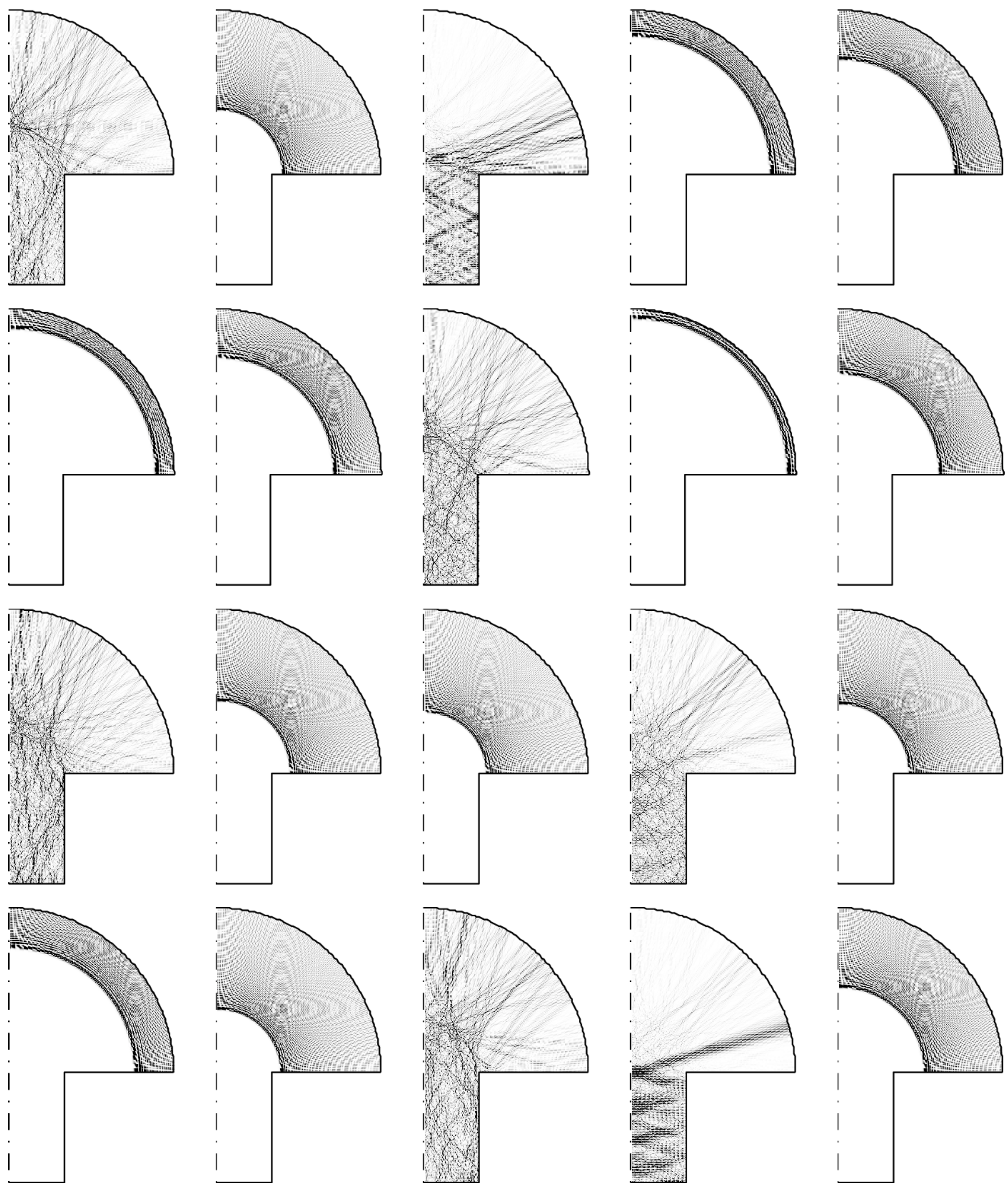

FIGURE 2.3. The density plots of some eigenfunctions in a mushroom billiard system (Credit: Barnett-Betcke)

A positive answer to the weak Percival's Conjecture is given for the generic mushroom billiard system by Gomes[i]. Here, "weak" means that one loses information of a zero density subsequence

[i]S. Gomes, *Percival's Conjecture for the Bunimovich Mushroom Billiard.* Nonlinearity 31 (2018), no. 9, 4108–4136.

of eigenfunctions; while "generic" means that the result holds for almost all mushrooms in a family for which the rectangular region has length, say, $[1, 2]$. The proof uses crucially the "non-clustering" phenonmenon of eigenvalues of Hassell as explained in Section 2.5. In particular, there are quasimodes that are highly localized in the semi-annulus region (given by eigenfunctions of the disc). Their properties are inherited by genuine eigenfunctions, if the eigenvalues do not cluster, which can be guaranteed by varying size of the rectangular region.

The same line of idea also played an important role in Gomes-Hassell[i] about the quantum system of integrable systems under perturbation. In this case, the KAM (Kolmogorov, Arnold and Moser) theory asserts that under small perturbation, the invariant tori which are non-resonant (i.e., the corresponding frequencies are badly approximated by rationals) persist.

Under additional Gevrey assumption of the sytem, for any given non-resonant torus, Popov[ii] constructed quasimodes which are (exponentially) localized nearby. It is natural to ask whether there are eigenfunctions that have the same localization. The obstacle from potential clustering of eigenvalues as before appears here. Hence, Gomes-Hassell introduced 1-parameter family of Gevrey KAM systems and overcame such obstacle. Their result states that for such generic KAM system, there are eigenfunctions which scar on the non-resonant tori. This partly answers the question in Problem 2.2.

[i] S. Gomes and A. Hassell, *Semiclassical scarring on tori in KAM Hamiltonian systems.* J. Eur. Math. Soc. (JEMS) 24 (2022), no. 5, 1769–1790.

[ii] G. Popov, *KAM theorem and quasimodes for Gevrey Hamiltonians*, Mat. Contemp. 26 (2004), 87–107.

CHAPTER 3

L^p norms of Laplacian eigenfunctions

Let Δ be the (positive) Laplacian on a n-dim compact manifold $\mathbb{M}$ without boundary. Suppose that $\Delta u = \lambda^2 u$. Then Sogge's L^p estimates of eigenfunctions state that

$$\|u\|_{L^p(\mathbb{M})} \le C\lambda^{\sigma(p)}\|u\|_{L^2(\mathbb{M})},$$

in which $C = C(p, \mathbb{M}) > 0$ and

$$\sigma(p) = \begin{cases} \frac{n-1}{4} - \frac{n-1}{2p} & \text{if } 2 < p \le \frac{2(n+1)}{n-1} := p_n, \\ \frac{n-1}{2} - \frac{n}{p} & \text{if } p_n \le p \le \infty. \end{cases} \tag{3.1}$$

We are mainly concerned with the extreme (i.e., maximal and minimal) growth rate of $\|u\|_{L^p}$ (after normalization in $L^2(\mathbb{M})$) as $\lambda \to \infty$ in terms of the following questions.

(A). Can one find a (sub)sequence of eigenfunctions $\{u\}$ on some manifold with the maximal L^p norm, i.e.,

$$\|u\|_{L^p(\mathbb{M})} \approx \lambda^{\sigma(p)}\|u\|_{L^2(\mathbb{M})} \quad \text{as } \lambda \to \infty?$$

(B). Can one find a (sub)sequence of eigenfunctions $\{u\}$ on some manifold with the minimal L^p norm, i.e.,

$$\|u\|_{L^p(\mathbb{M})} \approx \|u\|_{L^2(\mathbb{M})} \quad \text{as } \lambda \to \infty?$$

The strongest form would be to find uniformly bounded eigenfunctions.

(C). What is the density of the above subsequence in an eigenbasis? The strongest form would be to find a whole eigenbasis that satisfy the requirement.

(D). What is the semiclassical measure of the above subsequence?

There is a large literature on these topics, see Sogge[i] and Zelditch[ii].

3.1. Eigenfunctions with maximal L^p norms: Spherical harmonics, Take I

Sogge's L^p estimates (3.1) are sharp on the sphere $\mathbb{S}^n$. In this case, Laplacian eigenfunctions are spherical harmonics, that is, harmonic and homogeneous polynomials in $\mathbb{R}^{n+1}$ restricted to $\mathbb{S}^n$. Denote by $\mathbb{SH}_k^n$ the space of spherical harmonics of degree k on $\mathbb{S}^n$. Then

$$\Delta_{\mathbb{S}^n} u = k(k+n-1)u \quad \text{if } u \in \mathbb{SH}_k^n,$$

and

$$\dim \mathbb{SH}_k^n = \frac{2k^{n-1}}{(n-1)!} + O_n\left(k^{n-2}\right),$$

c.f., Section 3.4 in Sogge. See also Section 10.17 for spherical harmonics arising from the representation theory of SO(3).

The maximal L^p norms in (3.1) are saturated by zonal harmonics for $p \ge p_n$ and by Gaussian beams for $p \le p_n$.

[i]C. Sogge, *Hangzhou lectures on eigenfunctions of the Laplacian.* Princeton University Press, Princeton, NJ, 2014.

[ii]S. Zelditch, *Eigenfunctions of the Laplacian on a Riemannian manifold.* American Mathematical Society, Providence, RI, 2017.

Example (Zonal harmonics). Let $\{u_j\}$ be an eigenbasis of $\mathbb{SH}_k^n$. Then the projection kernel of $L^2(\mathbb{S}^n) \to \mathbb{SH}_k^n$ is given by

$$P(x,y) = \sum_{j=1}^{\dim \mathbb{SH}_k^n} u_j(x)\overline{u_j(y)}.$$

Notice that if $R \in \mathrm{O}(n+1)$, then $u \circ R \in \mathbb{SH}_k^n$ since the Laplacian commutes with R. Therefore, $P(x,y)$ must be invariant under $\mathrm{O}(n+1)$, i.e.,

$$P(Rx, Ry) = P(x,y).$$

Taking $y = x$, we have that

$$P(Rx, Rx) = P(x,x) = \sum_{j=1}^{\dim \mathbb{SH}_k^n} |u_j(x)|^2,$$

which is constant on $\mathbb{S}^n$ since $\mathrm{O}(n+1)$ acts on $\mathbb{S}^n$ transitively. As a consequence,

$$S_n \cdot P(x,x) = \int_{\mathbb{S}^n} P(x,x)\,dx = \int_{\mathbb{S}^n} \sum_{j=1}^{\dim \mathbb{SH}_k^n} |u_j(x)|^2\,dx = \dim \mathbb{SH}_k^n,$$

in which S_n is the surface area of $\mathbb{S}^n$. Hence,

$$P(x,x) = \frac{\dim \mathbb{SH}_k^n}{S_n} = ck^{n-1} + O_n\left(k^{n-2}\right),$$

in which $c = c(n) > 0$. Fix $x \in \mathbb{S}^n$. Compute that

$$\int_{\mathbb{S}^n} |P(x,y)|^2\,dy = \sum_{j,k=1}^{\dim \mathbb{SH}_k^n} u_j(x)\overline{u_k(x)} \int_{\mathbb{S}^n} u_k(x)\overline{u_j(y)}\,dy = \sum_{j=1}^{\dim \mathbb{SH}_k^n} |u_j(x)|^2 = P(x,x).$$

The zonal harmonic with pole at x is defined by $P(x,y)$ after normalization in $L^2(\mathbb{S}^n)$:

$$Z_x(y) = \frac{P(x,y)}{\sqrt{P(x,x)}}.$$

Now

$$Z_x(x) = \sqrt{P(x,x)} \approx k^{\frac{n-1}{2}},$$

which saturates the L^∞ norm in (3.1). That is, the zonal harmonic Z_x has a peak at its pole x.

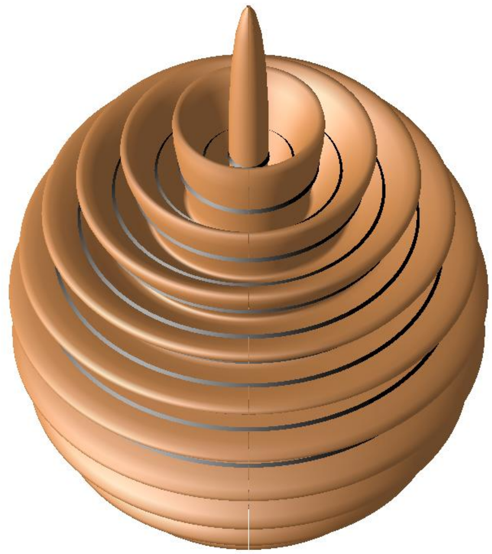

FIGURE 3.1. Zonal harmonic of degree 25 on $\mathbb{S}^2$ with peak at the north pole (and south pole) (Credit: McNamara)

In fact, one can find $r \approx k^{-1}$ such that

$$Z_x(y) \approx k^{\frac{n-1}{2}} \quad \text{for all } y \in B(x,r).$$

That is, the values of $Z_x(y)$ are comparable within a distance of ck^{-1} by choosing c small. It then follows that

$$\int_{\mathbb{S}^n} |Z_x(y)|^p \, dy \geq \int_{B(x,r)} |Z_x(y)|^p \, dy \gtrsim k^{-n} \cdot k^{\frac{p(n-1)}{2}}.$$

So

$$\|Z_x\|_{L^p(\mathbb{S}^n)} \gtrsim k^{\frac{n-1}{2} - \frac{n}{p}}.$$

That is, zonal harmonics saturate the maximal L^p norms for all $p \geq p_n$ in (3.1).

Example (Gaussian beams). Let $x = (x_1, ..., x_{n+1}) \in \mathbb{R}^{n+1}$. Define

$$u(x) = (x_1 + ix_2)^k.$$

Then u is harmonic and homogeneous of degree k. Hence, the restriction of u to $\mathbb{S}^n$, still denoted by u, is in $\mathbb{SH}_k^n$. Denote by G the great circle as the intersection of $\mathbb{S}^n$ and the x_1x_2-plane.

Choose a spherical coordinate system so

$$x_1 = \sin\phi\cos\theta \quad \text{and} \quad x_2 = \sin\phi\sin\theta,$$

in which $\phi \in [0, \pi]$ and $\theta \in [0, 2\pi)$. Then $G = \{\phi = \pi/2\}$ and

$$|u| = \left|\cos\left(\frac{\pi}{2} - \phi\right)\right|^k \gtrsim e^{-k\left(\frac{\pi}{2}-\phi\right)^2}.$$

That is, $|u| \approx 1$ in a tube around G with width $k^{-\frac{1}{2}}$ and decays exponentially fast in the transverse directions of the tube on $\mathbb{S}^n$. This tube has size $k^{-(n-1)/2}$ so

$$\|u\|_{L^p(\mathbb{S}^n)} \approx \left(k^{-\frac{n-1}{2}}\right)^{\frac{1}{p}} = k^{-\frac{n-1}{2p}}.$$

The Gaussian beam u_G is defined by u after normalization in $L^2(\mathbb{S}^n)$:

$$u_G(x) = k^{\frac{n-1}{4}}(x_1 + ix_2)^k.$$

Then $|u_G| \approx k^{(n-1)/4}$ in the $k^{-\frac{1}{2}}$-neighborhood of G. Noticing that $u_G = |u_G|e^{ik\theta}$, we know that u_G has Gaussian concentration in the transverse directions of G and propagates along G with frequency k. For these reasons, u_G is called a Gaussian beam.

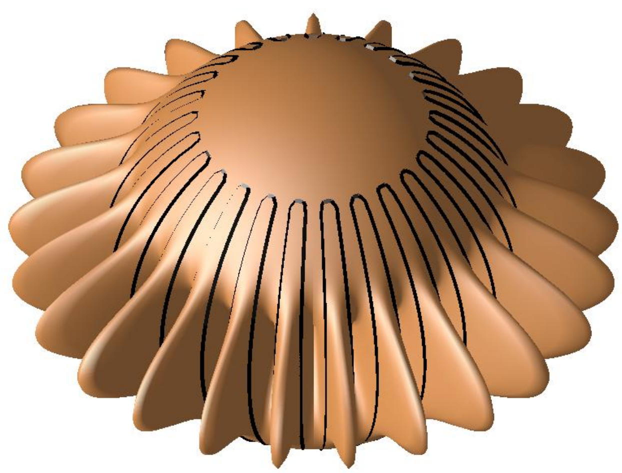

FIGURE 3.2. Gaussian beam of degree 25 on $\mathbb{S}^2$ that is concentrated around and propagates along the equator (Credit: McNamara)

We have that

$$\|u_G\|_{L^p(\mathbb{S}^n)} \approx k^{\frac{n-1}{4}-\frac{n-1}{2p}}.$$

That is, Gaussian beams saturate the maximal L^p norms for all $p \leq p_n$ in (3.1).

An important question in the study of L^p norms of eigenfunctions is that, do all eigenfunctions (on other manifolds) with maximal L^p norms resemble zonal harmonics for large p and Gaussian beams for small p? In addition, if a manifold supports such eigenfunctions, then must the geometry resemble the pole for zonal harmonic for large p and the closed geodesic (i.e., great circle on the sphere) for small p? See Sogge and Zelditch.

The other interesting question is that how many are these eigenfunctions with maximal L^p norms in an eigenbasis? In particular, zonal harmonics and Gaussian beams are both very rare in the standard eigenbasis of spherical harmonics, indeed, their density are zero. So one asks

Problem 3.1 (Eigenbasis with maximal L^p norms). *Find a positive density subsequence of eigenfunctions with maximal L^p norms. In addition, determine if there exists eigenbasis with maximal L^p norms.*

Han[i] constructed a positive density subsequence of spherical harmonics that saturate the maximal L^p estimates for $p \le p_n$. These eigenfunctions are small perturbations of Gaussian beams. The case for $p > p_n$ remains unknown. Moreover, the existence of eigenbasis with maximal L^p norms for either range of p values is also unknown.

3.2. Eigenfunctions with minimal L^p norms: Toral eigenfunctions, Take II

The strongest form of the problem to find eigenfunction with minimal L^p norms is

Problem 3.2 (Uniformly bounded eigenfunctions). *Find and characterize the eigenfunctions that are uniformly bounded.*

The only known examples of manifolds that support uniformly bounded eigenfunctions are the torus (or the fundamental domain of a torus with suitable boundary conditions).

Example (Toral eigenfunctions). The basic eigenmodes $e^{ik\cdot x}$ are uniformly bounded on $\mathbb{T}^n$. They are equidistributed in the physical space but are highly concentrated in the frequency space. Indeed, consider the eigenfunctions

$$u_k(x) = e^{ik\cdot x} = e^{i\frac{k}{|k|}\cdot x/h_k},$$

in which the semiclassical parameter $h_k = |k|^{-1}$. If $k/|k| \to \xi_0$ as $|k| \to \infty$ for some fixed $\xi_0 \in \mathbb{S}^{n-1}$, then

$$\left\langle \mathrm{Op}_{h_k}(a)u_{h_k}, u_{h_k}\right\rangle_{L^2(\mathbb{T}^n)} \to \int_{\mathbb{T}^n} a(x,\xi_0)\,dx \quad \text{as } h_k \to 0,$$

that is, the corresponding semiclassical measure is $dx\delta_{\xi=\xi_0}$, which is supported on the invariant set $\mathbb{T}^n \times \{\xi_0\} \subset S^*\mathbb{T}^n$. This shows that while these toral eigenfunctions $e^{ik\cdot x}$ are equidistributed on $\mathbb{T}^n$, they are highly localized in the frequency space so are not equidistributed in the phase space.

Toth-Zelditch[ii] proved that, under a certain condition of "quantum integrability", a manifold on which all eigenfunctions are uniformly bounded must be a torus. It is, however, unclear whether this condition can be replaced by a weaker "classical integrability". Furthermore, it is unclear whether there can exist eigenfunctions (rather than all) which are uniformly bounded on a given manifold.

We shall point out that uniformly boundedness barely misses being a typical property of eigenfunctions. For example, even though there are spherical harmonics with maximal L^p norms, Burq-Lebeau[iii] proved that the random spherical harmonics u (normalized in $L^2(\mathbb{S}^d)$) have the following estimates.

$$\|u\|_{L^p(\mathbb{S}^n)} \approx \sqrt{p} \quad \text{and} \quad \|u\|_{L^\infty(\mathbb{S}^n)} \approx \log\lambda.$$

Therefore, uniformly bounded spherical harmonics, as well as the ones with maximal L^p norm growth such as zonal harmonics and Gaussian beams, are very atypical. In particular, finding

[i] X. Han, *Spherical harmonics with maximal L^p ($2 < p \le 6$) norm growth.* J. Geom. Anal. 26 (2016), no. 1, 378–398.

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

[iii] N. Burq and G. Lebeau, *Probabilistic Sobolev embeddings, applications to eigenfunctions estimates.* Geometric and spectral analysis, 307–318, Contemp. Math., 630, Centre Rech. Math. Proc., Amer. Math. Soc., Providence, RI, 2014.

uniformly bounded ones amounts to "beating" probabilisitic arguments and therefore is difficult. In the next sections, we visit these results, as well as relevant background in the random setting.

3.3. Rudin-Shapiro polynomials

Beating the log bound in the random setting is a rather common theme in analysis and probability. For example, Rudin-Shapiro polynomials, roughly speaking, are a family of trigonometric polynomials with coefficients ± 1 that have better supremum bound than the random trigonometric polynomials. So we begin from the simplest random setting.

Let $\{\epsilon_n\}_{n\in\mathbb{Z}}$ be i.i.d. random variables with $P(\epsilon_n = \pm 1) = \frac{1}{2}$ for every $n \in \mathbb{Z}$, i.e., a sequence with Rademacher distribution. Consider the trigonometric polynomials with these random coefficients

$$\sum_{n=-\infty}^{\infty} \epsilon_n a_n e^{int} \quad \text{on the torus } \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}.$$

Then Paley-Zygmund[i] theorem applying to such random trigonometric polynomials asserts that

Theorem (Paley-Zygmund).

(i). If $\sum_n |a_n|^2 < \infty$, then a.s. (almost surely)

$$f(t) = \sum_{n=-\infty}^{\infty} \epsilon_n a_n e^{int} \in \bigcap_{1\le p<\infty} L^p(\mathbb{T}).$$

(ii). If $\sum_n |a_n|^2 = \infty$, then a.s.

$$\sum_{n=-\infty}^{\infty} \epsilon_n a_n e^{int} \quad \textit{diverges a.e. (almost everywhere).}$$

See Chapter V in Kahane[ii] for a full account, as well as the history, of the Paley-Zygmund theorem.

Example. If $\sum |a_n|^2 < \infty$, then (ii) says that almost surely $f(t)$ represent a function which is in $L^p(\mathbb{T})$ for all $1 \le p < \infty$. In general, it is not possible to go further and conclude that $f(t)$ is uniformly bounded on $\mathbb{T}$. For example, consider a lacunary series

$$f(t) = \sum_{j=1}^{\infty} \epsilon_j a_j e^{in_j t}, \qquad \text{in which } \liminf_{j\to\infty} \frac{n_{j+1}}{n_j} > 1.$$

Set $\sum |a_j|^2 < \infty$ and $\sum |a_j| = \infty$. Then one can prove that $f(t) \notin L^\infty(\mathbb{T})$ for any choices of $\epsilon_j = \pm 1$.

To estimate the L^∞ growth of the above random trigonometric polynomials, we consider the finite series with L^2-normalization

$$f_N(t) = \frac{1}{\sqrt{\sum_{n=0}^{N} |a_n|^2}} \sum_{n=0}^{N} \epsilon_n a_n e^{int}.$$

---

[i] R. Paley and A. Zygmund, *On some series of functions.* (1), (2), and (3). Proc. Camb. Phil. Soc. 26 (1930) 337-357, 26 (1930) 458-474, and 28 (1932) 190-205.

[ii] J-P. Kahane, *Some random series of functions.* Second edition. Cambridge University Press, Cambridge, 1985.

Then Salem-Zygmund[i] proved that

Theorem (Salem-Zygmund). *There exist positive constants c and C such that*

$$c \le \liminf_{N\to\infty} \frac{\|f_N\|_{L^\infty(\mathbb{T})}}{\sqrt{\log N}} \le \limsup_{N\to\infty} \frac{\|f_N\|_{L^\infty(\mathbb{T})}}{\sqrt{\log N}} \le C, \quad a.s.$$

That is, almost surely $\|f_N\|_{L^\infty(\mathbb{T})}$ grows at the order of $\sqrt{\log N}$. See Chapter VI in Kahane for more details. Then a question arises naturally:

Question 3.1 (Uniformly bounded trigonometric polynomials). *Can one find a trigonometric polynomial with coefficients* ± 1,

$$f_N(t) = \sum_{n=0}^{N} \epsilon_n e^{int} \text{ such that } \|f_N\|_{L^\infty(\mathbb{T})} \approx \|f_N\|_{L^2(\mathbb{T})} \approx \sqrt{N}?$$

Answer. This question was answered by Rudin[ii] and Shapiro and their example is now called the Rudin-Shapiro polynomials.

Notice that if the uni-modular coefficients are allowed to be complex, then the example was known by choosing $\epsilon_n = e^{in\log n}$, according to Hardy-Littlewood, see Rudin.

We now revisit the construction of the Rudin-Shapiro polynomials.

Definition (Rudin-Shapiro polynomials). Let

$$P_0 = Q_0 = 1,$$

and

$$\begin{cases} P_{m+1}(t) = P_m(t) + e^{i2^m t} Q_m(t), \\ Q_{m+1}(t) = P_m(t) - e^{i2^m t} Q_m(t). \end{cases}$$

The resulting polynomials are written as

$$P(t) = \sum_{n=0}^{N} \sigma_n^P e^{int} \quad \text{and} \quad Q_N(t) = \sum_{n=0}^{N} \sigma_n^Q e^{int},$$

in which $\{\sigma_n^P\}$ and $\{\sigma^Q\}$ are called the Rudin-Shapiro sequences.

Example. Compute the first several polynomials:

$$\begin{cases} P_1(t) = 1 + e^{it}, \\ Q_1(t) = 1 - e^{it}; \end{cases}$$

$$\begin{cases} P_2(t) = 1 + e^{it} + e^{2it}(1 - e^{it}) = 1 + e^{it} + e^{2it} - e^{3it}, \\ Q_2(t) = 1 + e^{it} - e^{2it}(1 - e^{it}) = 1 + e^{it} - e^{2it} + e^{3it}; \end{cases}$$

$$\begin{cases} P_3(t) = 1 + e^{it} + e^{2it} - e^{3it} + e^{4it}(1 + e^{it} - e^{2it} + e^{3it}) = 1 + e^{it} + e^{2it} - e^{3it} + e^{4it} + e^{5it} - e^{6it} + e^{7it}, \\ Q_3(t) = 1 + e^{it} + e^{2it} - e^{3it} - e^{4it}(1 + e^{it} - e^{2it} + e^{3it}) = 1 + e^{it} + e^{2it} - e^{3it} - e^{4it} - e^{5it} + e^{6it} - e^{7it}. \end{cases}$$

Therefore,

$$\sigma_n^P = 1, 1, 1, -1, 1, 1, -1, 1, \ldots \quad \text{and} \quad \sigma_n^Q = 1, 1, 1, -1, -1, -1, 1, -1, \ldots$$

[i] R. Salem and A. Zygmund, *Some properties of trigonometric series whose terms have random signs.* Acta Math. 91 (1954), 245–301.

[ii] W. Rudin, *Some theorems on Fourier coefficients.* Proc. Amer. Math. Soc. 10 (1959) 855–859, in which he remarked that "After I found this construction I learned that the problem had been solved earlier, by essentially the same method, in the 1951 Master's Thesis of H. S. Shapiro [in MIT]."

We next prove that the Rudin-Shapiro polynomials satisfy the uniform bounded condition.

PROOF.

(1). From the definition, we have that

$$|P_{m+1}(t)|^2 + |Q_{m+1}(t)|^2 = 2\big(|P_m(t)|^2 + |Q_m(t)|^2\big).$$

Then

$$|P_m(t)|^2 + |Q_m(t)|^2 = 2^{m+1}.$$

Hence,

$$\|P_m\|_{L^\infty(\mathbb{T})}, \|Q_m\|_{L^\infty(\mathbb{T})} \le 2^{\frac{m+1}{2}}.$$

Notice that

$$\|P_m\|_{L^2(\mathbb{T})} = \|Q_m\|_{L^2(\mathbb{T})} = \sqrt{2\pi 2^m} = \sqrt{2\pi}2^{\frac{m}{2}},$$

which immediately shows the uniformly bounded condition of P and Q when $N = 2^m$.

(2). We inductively prove that for $\{\sigma_n\} = \{\sigma_n^P\}$ or $\{\sigma^Q\}$,

$$\left\|\sum_{n=0}^{N} \sigma_n e^{int}\right\|_{L^\infty(\mathbb{T})} \le (\sqrt{2}+2)2^{\frac{m}{2}}, \quad \text{when } 1 \le N < 2^m.$$

It is obviously true for $m = 1$. Suppose that it is true for $0 \le N < 2^m$. Let $2^m \le N < 2^{m+1}$. Then

$$\begin{aligned}
&\left\|\sum_{n=0}^{N} \sigma_n e^{int}\right\|_{L^\infty(\mathbb{T})} \\
\le\ & \left\|\sum_{n=0}^{2^m-1} \sigma_n e^{int}\right\|_{L^\infty(\mathbb{T})} + \left\|\sum_{n=2^m}^{N} \sigma_n e^{int}\right\|_{L^\infty(\mathbb{T})} \\
\le\ & 2^{\frac{m+1}{2}} + \left\|e^{i2^m t}\sum_{n=0}^{N-2^m} \sigma_n^Q e^{int}\right\|_{L^\infty(\mathbb{T})} \qquad \text{from the construction of } P_m \text{ and } Q_m \\
\le\ & 2^{\frac{m+1}{2}} + (\sqrt{2}+2)2^{\frac{m}{2}} \quad \text{by induction since } N - 2^m < 2^m \\
=\ & (\sqrt{2}+2)2^{\frac{m+1}{2}}.
\end{aligned}$$

Now given any $N \ge 1$, suppose that $2^{m-1} \le N \le 2^m$. Then

$$\left\|\sum_{n=0}^{N} \sigma_n e^{int}\right\|_{L^\infty(\mathbb{T})} \le (\sqrt{2}+2)2^{\frac{m}{2}} \le (2+2\sqrt{2})\sqrt{N} \le 5\sqrt{N}.$$

$\square$

In some sense, the Rudin-Shapiro polynomials possess the "square-root cancellation" of $\sigma_n e^{int}$, $n = 0, ..., N$, in the L^∞ norm. This concept was improved further by Bourgain[i] that the "square-root cancellation" persists to any interval $a \le n \le b$ with $0 \le a < b < \infty$.

[i] J. Bourgain, *Applications of the spaces of homogeneous polynomials to some problems on the ball algebra.* Proc. Amer. Math. Soc. 93 (1985), no. 2, 277–283.

Theorem (Bourgain). *There exists an absolute constant $c > 0$ such that*

$$\left\| \sum_{n=a}^{b} \sigma_n e^{int} \right\|_{L^\infty(\mathbb{T})} \leq c\sqrt{b-a},$$

for all $0 \leq a < b < \infty$.

Proof.

(1). If $[a, b]$ is a dyadic interval, i.e.,

$$[a, b] = [j2^k, (j+1)2^k - 1] \text{ with } j, k \in \mathbb{N},$$

then

$$\left\| \sum_{n=a}^{b} \sigma_n e^{int} \right\|_{L^2\infty(\mathbb{T})} = \left\| \sum_{n=0}^{2^k-1} \sigma_n e^{int} \right\|_{L^2\infty(\mathbb{T})} = \|P_k(t)\|_{L^\infty(\mathbb{T})} \text{ or } \|Q_k(t)\|_{L^\infty(\mathbb{T})}.$$

Hence,

$$\left\| \sum_{n=a}^{b} \sigma_n e^{int} \right\|_{L^\infty(\mathbb{T})} \leq 2^{\frac{k+1}{2}} \leq \sqrt{2}\sqrt{b-a}.$$

(2). For a general interval $[a, b]$, we perform a dyadic decomposition as follows.

- Let $I_1 = [a_1, b_1] \subset [a, b]$ as the largest dyadic interval in $[a, b]$. Then $(b-a)/2 \leq |I_1| \leq (b-a)$. This is because if $|I_1| < (b-a)/2$ then there would be a larger dyadic interval in $[a, b]$.
- Let $I_2^{(1)} = [a_2, a_1] \subset [a, a_1]$ and $I_2^{(2)} = [b_1, b_2] \subset [b_1, b]$ be the largest dyadic intervals in $[a, a_1]$ and in $[b_1, b]$. Then $(a_1 - a)/2 \leq |I_2^{(1)}| \leq |I_1|/2$ and $(b - b_1)/2 \leq |I_2^{(1)}| \leq |I_1|/2$. Hence,
$$|I_2^{(1)}| + |I_2^{(2)}| \leq |I_1|.$$
- Let $I_3^{(1)} = [a_3, a_2] \subset [a, a_2]$ and $I_3^{(2)} = [b_2, b_3] \subset [b_2, b]$ be the largest dyadic intervals in $[a, a_2]$ and in $[b_2, b]$. Then $(a_2 - a)/2 \leq |I_3^{(1)}| \leq |I_2|/2$ and $(b - b_2)/2 \leq |I_3^{(1)}| \leq |I_2|/2$. Hence,
$$|I_3^{(1)}| + |I_3^{(2)}| \leq |I_2| \leq |I_1|/2.$$
- Choosing the dyadic intervals inductively, in m-th stage (other than $m = 1$), there are two dyadic intervals $I_m^{(1)}$ and $I_m^{(2)}$ satisfying
$$|I_m^{(1)}| + |I_m^{(2)}| \leq 2^{2-m}|I_1|.$$

Now

$$\begin{aligned}
& \left\| \sum_{n=a}^{b} \sigma_n e^{int} \right\|_{L^\infty(\mathbb{T})} \\
\leq & \left\| \sum_{m} \sum_{n \in I_m} \sigma_n e^{int} \right\|_{L^\infty(\mathbb{T})} \\
\leq & \sum_{m} 5\sqrt{|I_m^{(1)}|} + 5\sqrt{|I_m^{(2)}|} \\
\leq & 5\sqrt{2} \sum_{m} \sqrt{|I_m^{(1)}| + |I_m^{(2)}|}
\end{aligned}$$

$$\begin{aligned} &\leq 10\sqrt{2}\sqrt{b-a}\sum_{m} 2^{-\frac{m}{2}} \\ &\leq \frac{20}{\sqrt{2}-1}\sqrt{b-a}. \end{aligned}$$

□

As an immediately consequence, we have that

THEOREM 3.1. *There is an absolute constant $c > 0$ such that if $\{\lambda_j\}_{j=a}^{b} \subset \mathbb{R}$ is a monotone sequence and $0 \leq a < b < \infty$, then*

$$\left\| \sum_{j=a}^{b} \lambda_j \sigma_j e^{ijt} \right\|_{L^\infty(\mathbb{T})} \leq c \sup_{a \leq j \leq b} |\lambda_j| \sqrt{b-a}.$$

PROOF. Write

$$S_k(t) = \sum_{j=a}^{k} \sigma_j e^{ijt}$$

if $a < k \leq b$. Thus, $\|S_k\|_{L^\infty(\mathbb{T})} \leq c\sqrt{b-a}$ for all $a < k \leq b$. Put $S_{a-1}(t) = S_a(t) = 0$. Using partial summation formula, see Section 3.41 in Rudin[i],

$$\sum_{j=a}^{b} \lambda_j \sigma_j e^{ijt} = \sum_{j=a}^{b-1} S_j(\lambda_j - \lambda_{j+1}) + S_b \lambda_b - S_{a-1}\lambda_a.$$

Hence,

$$\begin{aligned} \left\| \sum_{j=a}^{b} \lambda_j \sigma_j e^{ijt} \right\|_{L^\infty(\mathbb{T})} &\leq \left\| \sum_{j=a}^{b-1} S_j(\lambda_j - \lambda_{j+1}) \right\|_{L^\infty(\mathbb{T})} + \|S_b\lambda_b\|_{L^\infty(\mathbb{T})} + \|S_{a-1}\lambda_a\|_{L^\infty(\mathbb{T})} \\ &\leq c\sqrt{b-a}\left(\sum_{j=a}^{b-1} |\lambda_j - \lambda_{j+1}| + |\lambda_b| + |\lambda_a| \right) \\ &\leq c \sup_{a \leq j \leq b} |\lambda_j| \sqrt{b-a}, \end{aligned}$$

because $\{\lambda_j\}$ is monotone. □

An interesting question is the higher-dimensional analogue of Rudin-Shapiro polynomials.

Problem 3.3 (Rudin-Shapiro polynomials in higher dimensions). *Let $k = (k_1, ..., k_n) \in \mathbb{Z}^n$ and $t = (t_1, ..., t_n) \in \mathbb{T}^n$. Can one find a trigonometric polynomial with coefficients $\sigma_k = \pm 1$,*

$$f_N(t) = \sum_{k_1,\ldots,k_n=0}^{N} \sigma_k e^{ik\cdot t} \text{ such that } \|f_N\|_{L^\infty(\mathbb{T}^n)} \approx \|f_N\|_{L^2(\mathbb{T}^n)} \approx N^{\frac{n}{2}}?$$

Lemma 4 in Bourgain[ii] provides a relevant answer in 2-dim: Set

$$\sigma_k = e^{2\pi i\sqrt{2}|k|^2} \quad \text{for } k = (k_1, k_2) \in \mathbb{Z}^2.$$

[i] W. Rudin, *Principles of mathematical analysis.* Third edition. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.

[ii] J. Bourgain, *On uniformly bounded bases in spaces of holomorphic functions.* Amer. J. Math. 138 (2016), no. 2, 571–584.

THEOREM 3.2. *Let $\rho : [0,1] \to [0,1]$ be smooth and compactly supported. Then there exists an absolute constant $C > 0$ such that for any $c_1, c_2 \in \mathbb{R}$ and $M_1, M_2 > 0$, we have that*

$$\sup_{t\in\mathbb{T}^2}\left|\sum_{k\in\mathbb{Z}^2}\rho\left(\frac{k_1-c_1}{M_1}\right)\rho\left(\frac{k_2-c_2}{M_2}\right)\sigma_k e^{ik\cdot t}\right| \le C\sqrt{M_1M_2}.$$

Remark. This theorem is the smooth version of a fact that the "square-root cancellation" of sum of $\sigma_k e^{ikt}$ holds when summing in any rectangle with sides parallel to the axes.

3.4. Bourgain polynomials: Spherical harmonics, Take II

Let $\mathbb{B}^m_{\mathbb{C}} = \{\zeta = (z_1, ..., z_m) \in \mathbb{C}^m : |\zeta| < 1\}$ be the unit ball in $\mathbb{C}^m$ and $\mathbb{S}^m_{\mathbb{C}} = \partial\mathbb{B}^m_{\mathbb{C}}$ be the unit sphere in $\mathbb{C}^m$. A question raised by Rudin is whether there exists a basis of holomorphic and homogeneous polynomials that are uniformly bounded on $\mathbb{S}^m_{\mathbb{C}}$ (after normalization in $L^2(\mathbb{S}^m_{\mathbb{C}})$). This question was answered by Bourgain when $m = 2$[i] and more recently $m = 3$[ii]. Such results were partially extended to more general Kähler manifolds by Shiffman[iii] and Marzo-Ortega-Cerdà[iv]. In particular, for any $\varepsilon > 0$, there is an orthonormal set of holomorphic and homogeneous sections which has density at least $1 - \varepsilon$ in an orthonormal basis. It is unclear whether these results can be improved to the existence of an orthonormal basis of holomorphic sections.

In this section, we explain Bourgain's construction on $\mathbb{S}^2_{\mathbb{C}}$. However, our motivation is somewhat different from Rudin's and is related to spherical harmonics. Denote by $\mathcal{P}^m_N$ as the space of holomorphic polynomials of homogeneous degree N on $\mathbb{S}^m_{\mathbb{C}}$, i.e.,

$$\mathcal{P}^m_N = \mathrm{span}\{z_1^{\alpha_1}\cdots z_m^{\alpha_m} : \alpha_1 + \cdots + \alpha_m = N\}.$$

Since $\mathbb{S}^m_{\mathbb{C}} = \mathbb{S}^{2m-1}_{\mathbb{R}}$, if $u \in \mathcal{P}^m_N$, then u is also a spherical harmonic on $\mathbb{S}^{2m-1}_{\mathbb{R}}$. In particular, Bourgain's construction of uniformly bounded holomorphic and homogeneous polynomials on $\mathbb{S}^2_{\mathbb{C}}$ and on $\mathbb{S}^3_{\mathbb{C}}$ provide uniformly bounded spherical harmonics of arbitrary degree on $\mathbb{S}^3_{\mathbb{R}}$ and on $\mathbb{S}^5_{\mathbb{R}}$, respectively. Moreover, the results of Shiffman and Marzo-Ortega-Cerdà conclude the same for all odd-dimensional spheres. Therefore, an interesting question is

Problem 3.4 (Uniformly bounded spherical harmonics). *Are there uniformly bounded spherical harmonics on $\mathbb{S}^n_{\mathbb{R}}$ of arbitrary degree for n even?*

Furthermore, since $\dim(\mathcal{P}^m_N) = O(N^{m-1})$ while $\dim(\mathbb{SH}^n_N) = O(N^{n-1})$, the density of holomorphic polynomials in the space of spherical harmonics when $n = 2m-1$ is odd is $O(N^{-(m-1)}) \to 0$ as $N \to \infty$. Therefore, a more challenging question than the above one is that, can one find a uniformly bounded orthonormal basis of spherical harmonics? This is open in any dimension.

We next present Bourgain polynomials on $\mathbb{S}^2_{\mathbb{C}}$. For notational simplicity, we drop the subscript $\mathbb{C}$.

[i]J. Bourgain, *Applications of the spaces of homogeneous polynomials to some problems on the ball algebra.* Proc. Amer. Math. Soc. 93 (1985), no. 2, 277–283.

[ii]J. Bourgain, *On uniformly bounded bases in spaces of holomorphic functions.* Amer. J. Math. 138 (2016), no. 2, 571–584.

[iii]B. Shiffman, *Uniformly bounded orthonormal sections of positive line bundles on complex manifolds.* Analysis, complex geometry, and mathematical physics: in honor of Duong H. Phong, 227–240, Contemp. Math., 644, Amer. Math. Soc., Providence, RI, 2015.

[iv]J. Marzo and J. Ortega-Cerda, *Uniformly bounded orthonormal polynomials on the sphere.* Bull. Lond. Math. Soc. 47 (2015), no. 5, 883–891.

Theorem (Bourgain). *Let $\{\sigma_j\}_{j=0}^N$ be the Rudin-Shapiro sequence. Then*

$$P_k(\zeta) = \frac{1}{\sqrt{N+1}} \sum_{j=0}^{N} \sigma_j e^{\frac{2\pi ijk}{N+1}} \frac{z^j w^{N-j}}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)}}, \quad k = 0, 1, ..., N,$$

form an orthonormal basis in $\mathcal{P}_N$ that are uniformly bounded.

It is unclear what these Bourgain polynomials look like, in addition to their uniformly bounded condition. We therefore ask the characterization of these functions as Laplacian eigenfunctions on the spheres.

Problem 3.5 (Semiclassical measures induced by Bourgain polynomials). *Characterize the semiclassical measures induced by Bourgain polynomials (viewed as spherical harmonics on $\mathbb{S}^3_{\mathbb{R}}$).*

Remark. The coefficients $e^{\frac{2\pi ijk}{N+1}}$ in the construction is to ensure that P_k, $k = 0, ..., N$, are mutually orthogonal. Indeed,

$$\begin{aligned}
& \langle P_k, P_{k'} \rangle_{L^2(\mathbb{S}^2)} \\
= \;& \frac{1}{4\pi^2(N+1)} \int_0^{2\pi}\int_0^{2\pi}\int_0^1 \sum_{j=0}^{N} \sigma_j e^{\frac{2\pi ijk}{N+1}} \frac{z^j w^{N-j}}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)}} \overline{\sum_{l=0}^{N} \sigma_l e^{\frac{2\pi ilk'}{N+1}} \frac{z^l w^{N-l}}{\|z^l w^{N-l}\|_{L^2(\mathbb{S}^2)}}} \, d\rho d\theta d\psi \\
= \;& \frac{1}{4\pi^2(N+1)} \sum_{j,l=0}^{N} \sigma_j \sigma_l e^{\frac{2\pi(ijk - ilk')}{N+1}} \frac{1}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)} \|z^l w^{N-l}\|_{L^2(\mathbb{S}^2)}} \\
& \int_0^{2\pi}\int_0^{2\pi}\int_0^1 z^j w^{N-j} \overline{z}^l \overline{w}^{N-l} \, d\rho d\theta d\psi \\
= \;& \frac{1}{4\pi^2(N+1)} \sum_{j,l=0}^{N} \sigma_j \sigma_l e^{\frac{2\pi(ijk - ilk')}{N+1}} \frac{1}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)} \|z^l w^{N-l}\|_{L^2(\mathbb{S}^2)}} \\
& \int_0^{2\pi}\int_0^{2\pi}\int_0^1 \sqrt{\rho}^j e^{ij\theta} \sqrt{1-\rho}^{N-j} e^{i(N-j)\psi} \sqrt{\rho}^l e^{-il\theta} \sqrt{1-\rho}^{N-l} e^{-i(N-l)\psi} \, d\rho d\theta d\psi \\
= \;& \frac{1}{(N+1)} \sum_{j,l=0}^{N} |\sigma_j|^2 \sigma_l e^{\frac{2\pi(ijk - ilk')}{N+1}} \delta_{jl} \\
= \;& \frac{1}{(N+1)} \sum_{j=0}^{N} e^{\frac{2\pi ij(k-k')}{N+1}} \\
= \;& \delta(k - k').
\end{aligned}$$

By a change of variables $z \to z e^{2\pi ik/(N+1)}$, the L^∞ estimate of P_k can be reduced to one of

$$\sup_{\rho \in [0,1], \theta \in \mathbb{T}} \frac{1}{\sqrt{N+1}} \left| \sum_{j=0}^{N} \frac{\rho^{\frac{j}{2}} (1-\rho)^{\frac{N-j}{2}}}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)}} e^{ij\theta} \right| = \left\| \sum_{j=0}^{N} \sigma_j \alpha_j e^{ij\theta} \right\|_{L^\infty(\mathbb{T})},$$

in which

- $$(z, w) = (\sqrt{\rho} e^{i\theta}, \sqrt{1-\rho} e^{i\psi}) \quad \text{with } \rho \in [0,1] \text{ and } \theta, \psi \in [0, 2\pi],$$
- $$\alpha_j = \frac{1}{\sqrt{N+1}} \frac{\rho^{\frac{j}{2}} (1-\rho)^{\frac{N-j}{2}}}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)}}.$$

The proof of uniform boundedness is based on two observations: Fix $\rho \in [0,1]$.

(1). α_j increases as j increases to ρN and then decreases as j passes ρN.

(2). α_j is concentrated in a $\sqrt{N}$-neighborhood of $j = \rho N$. In such a neighborhood, $|\alpha_j| \approx N^{-1/4}$ and decays exponentially fast outside of this neighborhood.

Assuming these properties of α_j, we then apply Theorem 3.1 involving the square-root cancellation of trigonometric polynomials with Rudin-Shapiro sequence as the coefficients:

$$\left\|\sum_{j=0}^{N} \sigma_j \alpha_j e^{ij\theta}\right\|_{L^\infty(\mathbb{T})} \lesssim \left\|\sum_{|j-\rho N|\lesssim\sqrt{N}} \sigma_j \alpha_j e^{ij\theta}\right\|_{L^\infty(\mathbb{T})} \lesssim N^{-\frac{1}{4}}\sqrt{\sqrt{N}} \leq C.$$

The theorem therefore is done.

We next establish the two observations concerning α_j.

Proof.

(1). Evaluate that

$$\|z^j w^{N-j}\|^2_{L^2(\mathbb{S}^2)} = \int_0^1 u^j (1-u)^{N-j}\, du = \frac{j!(N-j)!}{(N+1)!}.$$

Here, we use the Beta function (c.f., Exercise 7 in Stein-Shakarchi[i])

$$B(\alpha,\beta) = \int_0^1 (1-t)^{\alpha-1} t^{\beta-1}\, dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)},$$

where

$$\Gamma(s) = \int_0^\infty e^{-t} t^{s-1}\, dt. \quad (\Gamma(n+1) = n! \text{ if } n \in \mathbb{N}.)$$

Then,

$$\alpha_j = \frac{1}{\sqrt{N+1}} \frac{\rho^{\frac{j}{2}}(1-\rho)^{\frac{N-j}{2}}}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)}} = \frac{1}{\sqrt{N+1}} \left[\int_0^1 \left(\frac{u}{\rho}\right)^j \left(\frac{1-u}{1-\rho}\right)^{N-j} du\right]^{-\frac{1}{2}}.$$

Taking the derivative with respect to j:

$$\begin{aligned}
&\frac{d\alpha_j}{dj}\\
= \;& -\frac{1}{2\sqrt{N+1}} \left[\int_0^1 \left(\frac{u}{\rho}\right)^j \left(\frac{1-u}{1-\rho}\right)^{N-j} du\right]^{-\frac{3}{2}} \times \\
& \left[\int_0^1 \left(\frac{u}{\rho}\right)^{j-1} \frac{j}{\rho} \left(\frac{1-u}{1-\rho}\right)^{N-j} du - \int_0^1 \left(\frac{u}{\rho}\right)^j \left(\frac{1-u}{1-\rho}\right)^{N-j-1} \frac{N-j}{1-\rho}\, du\right]\\
= \;& -\frac{1}{2\sqrt{N+1}} \left[\int_0^1 \left(\frac{u}{\rho}\right)^j \left(\frac{1-u}{1-\rho}\right)^{N-j} du\right]^{-\frac{3}{2}} \times \int_0^1 \left(\frac{u}{\rho}\right)^{j-1} \left(\frac{1-u}{1-\rho}\right)^{N-j-1} \frac{j-\rho N}{\rho(1-\rho)}\, du
\end{aligned}$$

We can see that α_j increases until $j_0 \approx \rho N$; then it decreases.

[i]E. Stein and R. Shakarchi, *Complex analysis.* Princeton University Press, Princeton, NJ, 2003.

(2). Recall that Sterling's formula yields (c.f., Theorem 2.3 in Stein-Shakarchi)

$$\Gamma(s) \approx \sqrt{2\pi} s^{s-\frac{1}{2}} e^{-s}.$$

That is,

$$n! \approx (n+1)^{n+\frac{1}{2}} e^{-n} \approx n^{n+\frac{1}{2}} e^{-n}.$$

Then

$$\begin{aligned}
\alpha_j &= \frac{1}{\sqrt{N+1}} \frac{\rho^{\frac{j}{2}}(1-\rho)^{\frac{N-j}{2}}}{\|z^j w^{N-j}\|_{L^2(\mathbb{S}^2)}} \\
&= \left[\frac{\rho^j(1-\rho)^{N-j}}{N+1} \frac{(N+1)!}{j!(N-j)!}\right]^{\frac{1}{2}} \\
&\approx \left[\rho^j(1-\rho)^{N-j} \frac{N^{N+\frac{1}{2}} e^{-N}}{j^{j+\frac{1}{2}} e^{-j} (N-j)^{N-j+\frac{1}{2}} e^{-(N-j)}}\right]^{\frac{1}{2}} \\
&\approx \frac{N^{\frac{1}{4}}}{j^{\frac{1}{4}}(N-j)^{\frac{1}{4}}} \left[\frac{\rho^j(1-\rho)^{N-j} N^N}{j^j (N-j)^{N-j}}\right]^{\frac{1}{2}} \\
&\lesssim \min\{j, N-j\}^{-\frac{1}{4}} \left[\left(\frac{\rho N}{j}\right)^j \left(\frac{(1-\rho)N}{N-j}\right)^{N-j}\right]^{\frac{1}{2}} \quad \text{let } x = j - \rho N \\
&= \min\{j, N-j\}^{-\frac{1}{4}} \left[\left(\frac{\rho N}{x+\rho N}\right)^j \left(\frac{(1-\rho)N}{(1-\rho)N - x}\right)^{N-j}\right]^{\frac{1}{2}} \\
&= \min\{j, N-j\}^{-\frac{1}{4}} \left[\left(1+\frac{x}{\rho N}\right)^{-j} \left(1-\frac{x}{(1-\rho)N}\right)^{-(N-j)}\right]^{\frac{1}{2}}.
\end{aligned}$$

We have that

$$\begin{aligned}
&\log\left[\left(1+\frac{x}{\rho N}\right)^{-j} \left(1-\frac{x}{(1-\rho)N}\right)^{-(N-j)}\right] \\
= \;& -j\log\left(1+\frac{x}{\rho N}\right) - (N-j)\log\left(1-\frac{x}{(1-\rho)N}\right) \quad \text{using } \log(1+x) \approx x. \\
\approx \;& -\frac{jx}{\rho N} + \frac{(N-j)x}{(1-\rho)N} \\
= \;& \frac{(-j(1-\rho) + (N-j)\rho)x}{N\rho(1-\rho)} \\
= \;& \frac{(-j+N\rho)x}{N\rho(1-\rho)} \\
= \;& -\frac{(j-N\rho)^2}{N\rho(1-\rho)}.
\end{aligned}$$

Putting the above bounds together,

$$\alpha_j \le C \min\{j, N-j\}^{-\frac{1}{4}} e^{-\frac{c(j-\rho N)^2}{N\rho(1-\rho)}}.$$

$\square$

3.5. Strichartz estimates on compact manifolds

Let $u \in C_0^\infty(\mathbb{R}^n)$. Then

$$u(t,x) = e^{it\Delta}u_0(x)$$

solves the linear Schrödinger equation

$$\begin{cases} i\partial_t u(t,x) + \Delta_x u(t,x) = 0 & \text{for } (t,x) \in \mathbb{R}\times\mathbb{R}^n, \\ u(0,x) = u_0(x). \end{cases}$$

the Strichartz[i] estimates provide the $L_t^q L_x^p$ bounds of $u(t,x)$ in terms of the L_x^2 norm of the initial data $u_0(x)$:

Theorem (Strichartz).

$$\|u(t,x)\|_{L_t^q(\mathbb{R})L_x^p(\mathbb{R}^n)} \lesssim \|u_0(x)\|_{L_x^2(\mathbb{R}^n)},$$

in which $2 \le p, q \le \infty$, $(p,q) \neq (2,\infty)$, *and*

$$\frac{n}{p} + \frac{2}{q} = \frac{n}{2}.$$

For a proof, see Keel-Tao[ii]. It is built upon the following estimates. On one hand, tt is obvious that the L^2 norm of u is conserved under the Schrödinger flow:

$$\|e^{it\Delta}u\|_{L_x^2(\mathbb{R}^n)} = \|u\|_{L_x^2(\mathbb{R}^n)}.$$

On the other hand,

$$e^{it\Delta}u(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i\xi\cdot x}e^{it|\xi|^2}\hat{u}(\xi)\,d\xi = \frac{1}{(2\pi)^n}\iint_{\mathbb{R}^{2n}} e^{i(\xi\cdot(x-y)+t|\xi|^2)}u(y)\,dyd\xi,$$

the kernel of $e^{it\Delta}$ is given by

$$K(x,y) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} e^{i(\xi\cdot(x-y)+t|\xi|^2)}d\xi = \frac{1}{(4\pi i t)^{n/2}}e^{\frac{i|x-y|^2}{4t}}.$$

It then follows that

$$\|e^{it\Delta}u\|_{L_x^\infty(\mathbb{R}^n)} \lesssim \frac{1}{t^{n/2}}\|u\|_{L_x^1(\mathbb{R}^n)}.$$

We are mainly concerned with the Strichartz estimates on Riemannian manifold $\mathbb{M}$: We say that the (local-in-time) Strichartz estimates hold if

$$\|e^{it\Delta}u(x)\|_{L_t^q([0,T])L_x^p(\mathbb{M})} \le C\|u(x)\|_{L_x^2(\mathbb{M})},$$

for some $C = C(T) > 0$. If

$$\sup_{T\in(0,\infty)} C < \infty,$$

then we say that global-in-time Strichartz estimates hold on $\mathbb{M}$.

Remark. Notice that on a compact manifold, taking $u = 1$ implies that global-in-time Strichartz estimates cannot hold.

Burq-Gérard-Tzvetkov[iii] proved that

---

[i] R. Strichartz, *Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations*. Duke Math. J. 44 (1977), no. 3, 705–714.

[ii] M. Keel and T. Tao, *Endpoint Strichartz estimates*. Amer. J. Math. 120 (1998), no. 5, 955–980.

[iii] N. Burq, P. Gérard, and N. Tzvetkov, *Strichartz inequalities and the nonlinear Schrödinger equation on compact manifolds*. Amer. J. Math. 126 (2004), no. 3, 569–605.

Theorem (Burq-Gérard-Tzvetkov). *Let $\mathbb{M}$ be compact. Then*

$$\|e^{it\Delta}u(x)\|_{L^q_t([0,T])L^p_x(\mathbb{M})} \le C\|u(x)\|_{H_x^{\frac{1}{q}}(\mathbb{M})},$$

in which $C = C(T) > 0$.

That is, Strichartz estimates hold with a loss of $\frac{1}{q}$ derivatives. See also earlier result by Staffilani-Tataru[i].

Remark (Strichartz estimates and L^p estimates of eigenfunctions). Set u be a Laplacian eigenfunction, $-\Delta u = \lambda^2 u$. Then

$$\|u\|_{H^{\frac{1}{q}}(\mathbb{M})} \approx \lambda^{\frac{1}{q}}\|u\|_{L^2(\mathbb{M})},$$

and

$$\|e^{it\Delta}u(x)\|_{L^q_t([0,T])L^p_x(\mathbb{M})} = \|e^{-it\lambda^2}u(x)\|_{L^q_t([0,T])L^p_x(\mathbb{M})} = T^{\frac{1}{q}}\|u\|_{L^p(\mathbb{M})}.$$

Therefore, the Strichartz estimates reduce to the L^p norm estimates of u:

$$\|u\|_{L^p(\mathbb{M})} \le C\lambda^{\frac{1}{q}}\|u\|_{L^2(\mathbb{M})},$$

in which

$$\frac{n}{p} + \frac{2}{q} = \frac{n}{2} \quad \text{so } \frac{1}{q} = \frac{n}{4} - \frac{n}{2p}.$$

In the case when $q = 2$,

$$p = \frac{2n}{n-2}$$

falls into the second range of p values in Sogge's L^p esimates of eigenfunctions that

$$\|u\|_{L^p(\mathbb{M})} \le C\lambda^{\sigma(p)}\|u\|_{L^2(\mathbb{M})},$$

in which $C = C(p, \mathbb{M}) > 0$ and

$$\sigma(p) = \begin{cases} \frac{n-1}{4} - \frac{n-1}{2p} & \text{if } 2 < p \le \frac{2(n+1)}{n-1} := p_n, \\ \frac{n-1}{2} - \frac{n}{p} & \text{if } p_n \le p \le \infty. \end{cases}$$

See Chapter 3. Hence,

$$\sigma\left(\frac{2n}{n-2}\right) = \frac{n-1}{2} - \frac{n}{\frac{2n}{n-2}} = \frac{1}{2},$$

which agrees with the one from the Strichartz estimate. This means that the Strichartz estimate on the sphere is sharp when $q = 2$ in the sense that the loss of $\frac{1}{2}$ derivatives cannot be improved by a loss of fewer derivatives. Moreover, the sharp example is to put the initial data as zonal harmonics (which achieves the maximal L^p norm when $p = 2n/(n-2)$).

One sees that the Strichartz estimates on compact manifolds implies the L^p estimates of eigenfunction (in certain range of p values). On the other hand, the latter provides strong evidence on how much the loss of derivatives one should expect in the former. For example, on the torus $\mathbb{T}^2 = \mathbb{R}^2/2\pi\mathbb{Z}^2$, Zygmund[ii] proved that the L^4 norm of eigenfunctions are uniformly bounded. For the Strichartz estimates with $p = q = 4$, Bourgain[iii] proved that

[i]G. Staffilani and D. Tataru, *Strichartz estimates for a Schrödinger operator with nonsmooth coefficients*, Comm. Partial Differential Equations 27 (2002), 1337–1372.

[ii]A. Zygmund, *On Fourier coefficients and transforms of functions of two variables.* Studia Math. 50 (1974), 189–201.

[iii]J. Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrdinger equations.* Geom. Funct. Anal. 3 (1993), no. 2, 107–156.

Theorem (Bourgain). *For any $\varepsilon > 0$, there is $C = C(\varepsilon) > 0$ such that*

$$\|e^{it\Delta}u(x)\|_{L^4_t(\mathbb{T})L^4_x(\mathbb{T}^2)} \le C\|u(x)\|_{H^\varepsilon(\mathbb{T}^2)}.$$

PROOF. It suffices to show

$$\left\|e^{it\Delta}u(x)\right\|_{L^4_t(\mathbb{T})L^4_x(\mathbb{T}^2)} \lesssim_\varepsilon N^\varepsilon\|u(x)\|_{L^2_x(\mathbb{T})},$$

in which $\operatorname{supp}\hat{u} \subset \{k : N \le |k| \le 2N\}$, that is,

$$u(x) = \sum_{k\in\mathbb{Z}^n, N\le|k|\le 2N} a_k e^{ik\cdot x}.$$

Notice also that it suffices to prove the bilinear version

$$\left\|e^{it\Delta}u(x)e^{it\Delta}v(x)\right\|_{L^2_t(\mathbb{T})L^2_x(\mathbb{T})} \lesssim_\varepsilon N^\varepsilon\|u(x)\|_{L^2_x(\mathbb{T})}\|v(x)\|_{L^2_x(\mathbb{T})}.$$

Write

$$u(x) = \sum_{k\in\mathbb{Z}^2} a_k e^{ik\cdot x} \quad \text{and} \quad v(x) = \sum_{j\in\mathbb{Z}^2} b_j e^{ij\cdot x},$$

in which $a_k = 0$ if $|k| > N$ and $b_j = 0$ if $|j| > N$. Then

$$\begin{aligned}
e^{it\Delta}u(x)e^{it\Delta}v(x) &= \sum_{k,j\in\mathbb{Z}^2} e^{it(|k|^2+|j|^2)}e^{i(k+j)\cdot x}a_k b_j \\
&= \sum_{m\in\mathbb{N},n\in\mathbb{Z}^2}\ \sum_{m=|k|^2+|j|^2,n=k+j} e^{itm}e^{in\cdot x}a_k b_j \\
&= \sum_{m\in\mathbb{N},n\in\mathbb{Z}^2} e^{itm}e^{in\cdot x}\sum_{(k,j)\in\Lambda_{m,n}} a_k b_j,
\end{aligned}$$

in which

$$\Lambda_{m,n} = \{(k,l) : |k|^2 + |j|^2 = m, k + j = n\}.$$

So Plancherel's theorem for a Fourier series with indices m, n implies that

$$\begin{aligned}
\left\|e^{it\Delta}u(x)e^{it\Delta}v(x)\right\|^2_{L^2_{(t,x)}(\mathbb{T}^3)} &= \sum_{m\in\mathbb{N},n\in\mathbb{Z}^2}\left(\sum_{(k,j)\in\Lambda_{m,n}} a_k b_j\right)^2 \\
&\le \sum_{m\in\mathbb{N},n\in\mathbb{Z}^2} \#\Lambda_{m,n}\sum_{(k,j)\in\Lambda_{m,n}} |a_k b_j|^2 \\
&\lesssim_\varepsilon N^\varepsilon \sum_{m\in\mathbb{N},n\in\mathbb{Z}^2}\ \sum_{(k,j)\in\Lambda_{m,n}} |a_k b_j|^2 \\
&\lesssim_\varepsilon N^\varepsilon \sum_{k,j\in\mathbb{Z}^2} |a_k b_j|^2 \\
&\lesssim_\varepsilon N^\varepsilon \sum_{k\in\mathbb{Z}^2} |a_k|^2 \sum_{j\in\mathbb{Z}^2} |b_j|^2 \\
&\lesssim_\varepsilon N^\varepsilon \|u\|^2_{L^2_x}\|v\|^2_{L^2_x}.
\end{aligned}$$

Here, we use the crucial fact that

$$\#\{(k,l) : |k|, |j| \le N, |k|^2 + |j|^2 = m, k + j = n\} \lesssim_\varepsilon N^\varepsilon.$$

$\square$

The contrast of the Strichartz estimates on the sphere and on the torus indicates a strong connection of these estimates and the geometric background, particularly, the dynamics of the geodesic flow. It is then interesting to ask

Problem 3.6 (Strichartz estimates and dynamics of the geodesic flow). *Can one prove the Strichartz estimates on the torus based on the dynamics of its geodesic flow?*

Sarnak[i] conjectured that

$$\|u\|_{L^\infty(\mathbb{M})} \leq_\varepsilon \lambda^\varepsilon \|u\|_{L^2(\mathbb{M})},$$

for eigenfunctions, $-\Delta u = \lambda^2 u$, on compact hyperbolic surfaces. Then correspondingly, one would ask

Problem 3.7 (Strichartz estimates on compact hyperbolic surfaces). *Can on prove the Strichartz estimates on compact hyperbolic surfaces with arbitrarily loss of derivatives?*

[i]P. Sarnak, *Arithmetic Quantum Chaos.* The Schur lectures (1992) (Tel Aviv), 183–236, Israel Math. Conf. Proc., 8, Bar-Ilan Univ., Ramat Gan, 1995.

CHAPTER 4

Nodal geometry of Laplacian eigenfunctions

Let Δ be the Laplacian on a n-dim compact manifold $\mathbb{M}$. Suppose that $-\Delta u = \lambda^2 u$. Then the nodal set of u, denoted by

$$\mathcal{N}(u) = \{x \in \mathbb{M} : u(x) = 0\},$$

is a union of $(n-1)$-dim smooth submanifolds (except possibly at a singular set of lower dimension), see Hardt-Simon[i]. Our main concern of the nodal geometry of Laplacian eigenfunctions are

(A). Yau's nodal size conjecture, i.e., the upper and lower bounds of the $(n-1)$-dim Hausdorff measure of $\mathcal{N}(u)$ in terms of λ,
(B). distribution of nodal sets in the manifold, i.e., classification of the weak limits of $d\mathcal{H}^{n-1}|_{\mathcal{N}(u)}$ (after appropriate normalization by their sizes according to Yau's conjecture),
(C). nodal sets and volume spectrum,
(D). number of connected components of $\mathbb{M} \setminus \mathcal{N}(u)$, i.e., the nodal domain counting problems,
(E). overlapping nodal sets (i.e., a submanifold $H \subset \mathbb{M}$ that is contained in the nodal sets of multiple different eigenfunctions) and nodal intersections (i.e., the size of $\mathcal{N}(u) \cap H$),
(F). vanishing orders of eigenfunctions at nodal points.

On manifolds such as sphere and torus, Yau's nodal size conjecture is implied by Donnelly-Fefferman[ii] since the metric is analytic. However, the other problems on the above list remain largely open. On these special manifolds, eigenfunctions enjoy some additional properties, which can potentially be useful to answer the problems. For example, eigenfunctions on the torus $\mathbb{T}^n = \mathbb{R}^n / 2\pi\mathbb{Z}^n$ are trigonometric polynomials

$$u(x) = \sum_{k \in \mathbb{Z}^n : |k|^2 = \lambda^2} c_\xi e^{ik\cdot x},$$

whose nodal set behaviors are closely related to the distribution of lattice points $\{k \in \mathbb{Z}^n : |k|^2 = \lambda^2\}$ on the sphere $\mathbb{S}^{n-1}_\lambda$.

One can ask the above problems for the singular set of u,

$$\mathcal{S}(u) = \{x \in \mathbb{M} : u(x) = 0 \text{ and } \nabla u(x) = 0\},$$

for which they are mostly open.

4.1. Nodal set problems

4.1.1. Nodal size estimates. The most influential problem in nodal geometry is the nodal size conjecture of Yau[iii]: There are positive constants $c = c(\mathbb{M}), C = C(\mathbb{M})$ such that if $-\Delta u =$

[i] R. Hardt and L. Simon, *Nodal sets for solutions of elliptic equations.* J. Differential Geom. 30 (1989), no. 2, 505–522.

[ii] H. Donnelly and C. Fefferman, *Nodal sets of eigenfunctions on Riemannian manifolds.* Invent. Math. 93 (1988), no. 1, 161–183.

[iii] S.-T. Yau, *Open problems in geometry.* Proc. Sympos. Pure Math. Vol. 54, Part 1, Providence RI: Amer. Math. Soc. 1993, pp. 1–28.

$\lambda^2 u$, then

$$c\lambda \le \mathcal{H}^{n-1}(\mathcal{N}(u)) \le C\lambda.$$

Here, $\mathcal{H}^{n-1}$ is the $(n-1)$-dim Hausdorff measure. The following (incomplete) list is a brief history of attacking this conjecture.

(i). Upper bounds (omitting the constant):
- $\mathbb{M}$ is analytic: λ by Donnelly-Fefferman[i] and Lin[ii],
- 2-dim: λ^3 by Donnelly-Fefferman[iii] and Dong[iv],
- n-dim, $n \ge 3$: $\lambda^{2\lambda}$ by Hardt-Simon[v],
- n-dim: $n \ge 2$: λ^{α} for some $\alpha = \alpha(n) > 1$ by Logunov[vi].

(ii). Lower bounds (omitting the constant):
- $\mathbb{M}$ is analytic: λ by Donnelly-Fefferman,
- 2-dim: λ by Brüning[vii] and Yau,
- n-dim, $n \ge 3$:
  - $\lambda^{\frac{7-3n}{4}}$ by Sogge-Zelditch[viii];
  - $\lambda^{\frac{3-n}{2}}$ by Colding-Minicozzi[ix] and Sogge-Zelditch[x].
  - λ by Logunov[xi].

The remaining problem to Yau's conjecture is to prove the sharp upper bound $O(\lambda)$ of the nodal size on smooth manifolds.

Remark (Doubling index and nodal size). Several approaches on the nodal size estimates mentioned above are based on the study of the doubling index of eigenfunctions: Let

$$N_u(x,r) = \log_2\left(\frac{\max_{\overline{B}(x,2r)}|u|}{\max_{\overline{B}(x,r)}|u|}\right).$$

• The lower bounds of nodal size: In a ball with radius $r_0 \approx \lambda^{-1}$, i.e., the typical wavelength of u, it is well known that the nodal size of u in $B(x,r_0)$ is bounded below by $r_0^{n-1} = \lambda^{-(n-1)}$, provided that the doubling index $N_u(x,r_0) \le C$ for some uniform constant C. Such balls with

---

[i] H. Donnelly and C. Fefferman, *Nodal sets of eigenfunctions on Riemannian manifolds.* Invent. Math. 93 (1988), no. 1, 161–183.

[ii] F. Lin, *Nodal sets of solutions of elliptic and parabolic equations.* Comm. Pure Appl. Math. 44 (1991), no. 3, 287–308.

[iii] H. Donnelly and C. Fefferman, *Nodal sets for eigenfunctions of the Laplacian on surfaces.* J. Amer. Math. Soc. 3 (1990), 333–353.

[iv] R.-T. Dong, *Nodal sets of eigenfunctions on Riemann surfaces.* J. Differential Geom. 36 (1992), no. 2, 493–506.

[v] R. Hardt and L. Simon, *Nodal sets for solutions of elliptic equations.* J. Differential Geom. 30 (1989), no. 2, 505–522.

[vi] A. Logunov, *Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure.* Ann. of Math. (2) 187 (2018), no. 1, 221–239.

[vii] J. Brüning, *Über Knoten von Eigenfunktionen des Laplace-Beltrami-Operators.* Math. Z. 158 (1978), no. 1, 15–21.

[viii] C. Sogge and S. Zelditch, *Lower bounds on the Hausdorff measure of nodal sets.* Math. Res. Lett. 18 (2011), no. 1, 25–37.

[ix] T. Colding and W. Minicozzi, *Lower bounds for nodal sets of eigenfunctions.* Comm. Math. Phys. 306 (2011), no. 3, 777–784.

[x] C. Sogge and S. Zelditch, *Lower bounds on the Hausdorff measure of nodal sets II.* Math. Res. Lett. 19 (2012), no. 6, 1361–1364.

[xi] A. Logunov, *Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture.* Ann. of Math. (2) 187 (2018), no. 1, 241–262.

bounded doubling index are called "good" balls. Then the global lower bound of the nodal size is achieved by finding (disjoint) good balls. More good balls one finds, better the lower bound is. For example, if there are $1/r_0^n \approx \lambda^n$ good balls, then the global lower bound $\lambda^{-(n-1)} \cdot \lambda^n = \lambda$ follows. This was accomplished by Donnelly-Fefferman on analytic manifolds.

- The upper bounds of nodal size: On Riemannian surfaces (i.e., $n = 2$),

$$\mathcal{H}^1(\mathcal{N}(u) \cap B(x, r_0)) \lesssim r_0 N_u(x, r_0).$$

See Roy-Fortin[i]. In particular, in a good ball, the nodal size $\approx r_0 \approx \lambda^{-1}$, which agrees with the lower bound mentioned above.

Cover the manifold with $\approx r_0^{-2} \approx \lambda^2$ balls with radius r_0. Then we have that

$$\mathcal{H}^1(\mathcal{N}(u)) \lesssim \sum_{B(x,r_0)} r_0 N_u(x, r_0) \approx \lambda \sum_{B(x,r_0)} N_u(x, r_0) \cdot \mathrm{Vol}(B(x, r_0)).$$

In fact, Roy-Fortin proved that

$$\mathcal{H}^1(\mathcal{N}(u)) \approx \lambda \int_{\mathbb{M}} N_u(x, r)\, dx.$$

That is, Yau's nodal size conjecture is equivalent to proving that the "average" doubling index of u is uniformly bounded as $\lambda \to \infty$.

There is a clear link between the doubling index and the distribution of $|u|^2$ at the scale r_0. That is, if u is equidistributed at such a scale, then $N_u(x, r_0) \lesssim C$. See Chapter 1 for the background of small scale equidistribution of eigenfunctions. However, the wave-length scale $r_0 \approx \lambda^{-1}$ is too small to expect equidistribution phenomenon. On the other hand, the nodal size estimate only requires an average estimate of the doubling index. So one might ask

Problem 4.1 (Small scale equidistribution and average doubling index). *On a Riemannian surface, suppose that there is a sequence of eigenfunctions u which are equidistributed at certain small scale $r_1 \geq r_0$. Can one prove that*

$$\int_{\mathbb{M}} N_u(x, r_1)\, dx \leq C?$$

If so, what consequence it has on the nodal size of u?

We also mention the (asymptotically) optimal constants in Yau's nodal size conjecture: Let

$$H_l(\mathbb{M}) = \liminf_{\lambda \to \infty} \frac{\mathcal{H}^{n-1}(\mathcal{N}(u))}{\lambda} \quad \text{and} \quad H_u(\mathbb{M}) = \limsup_{\lambda \to \infty} \frac{\mathcal{H}^{n-1}(\mathcal{N}(u))}{\lambda}. \tag{4.1}$$

Problem 4.2 (Sharp constants in Yau's nodal size conjecture). *Find $H_l(\mathbb{M})$ and $H_u(\mathbb{M})$. Moreover, haracterize the eigenfunctions which achieve these limits.*

These constants are known for very few examples of manifolds, for example, Brüning-Gromes[ii] on irrational torus and Han-Murray-Tran[iii] on discs (see Section 4.2). In these examples, the eigenfunctions are explicit and the eigenvalues have bounded multiplicity. So the sharp nodal size can be computed directly. These examples, however, are very rare.

---

[i] G. Roy-Fortin, *Nodal sets and growth exponents of Laplace eigenfunctions on surfaces.* Anal. PDE 8 (2015), no. 1, 223–255.

[ii] J. Brüning and D. Gromes, *Ïber die Länge der Knotenlinien schwingender Membranen.* Math. Z. 124 (1972), 79–82.

[iii] X. Han, M. Murray, and C. Tran, *Nodal lengths of eigenfunctions in the disc.* Proc. Amer. Math. Soc. 147 (2019), no. 4, 1817–1824.

4.1.2. Distribution of nodal sets. In the view of the Yau's nodal size conjecture, we normalize by the size and ask

Problem 4.3 (Classification of the weak limits of nodal measures). *Classify all the weak limits of*

$$\left\{ \frac{1}{\lambda} d\mathcal{H}^{n-1} \Big|_{\mathcal{N}(u)} \right\}.$$

If u is quantum ergodic, then the L^2 mass of u tend to be equidistributed, it is likely that the nodal set also tend to be equidistributed:

Problem 4.4 (Weak limits of nodal measures of quantum ergodic eigenfunctions). *Prove that for quantum ergodic eigenfunctions* $\{u_k\}$,

$$\lim_{k\to\infty} \frac{\mathcal{H}^{n-1}(\mathcal{N}(u_k))}{\lambda_k} = c$$

for some $c > 0$. *Moreover, prove that*

$$\frac{1}{\lambda_k} d\mathcal{H}^{n-1}|_{\mathcal{N}(u_k)} \to c\, d\mathrm{Vol} \quad \textit{as } k \to \infty,$$

in which $d\mathrm{Vol}$ *is the Riemannian volume of* $\mathbb{M}$.

4.1.3. Number of nodal domains. The nodal set $\mathcal{N}(u)$ of an eigenfunction u is a union of $(n-1)$-dim hypersurfaces so divides $\mathbb{M}$ into connected components. In each connected component, u does not change sign and we call it a nodal domain. The nodal domain counting problems are concerned with the lower and upper bounds of the number of nodal domains, denoted by $\mathcal{N}(u)$, particularly as the eigenvalues tend to infinity.

4.1.3.1. *Upper bounds.* The most famous nodal counting result is the following upper bound due to Courant[i].

Theorem (Courant). *Let* $\{u_j\}_{j=1}^{\infty}$ *be a Dirichlet Laplacian eigenbasis on a compact manifold* $\mathbb{M}$ *with boundary. Then* $\mathcal{N}(u_j) \le j$ *for all* $j \in \mathbb{N}$.

Proof. It follows the Rayleigh quotient characterization of eigenvalues. That is,

$$\lambda_1 = \min\left\{ \frac{\int_{\mathbb{M}} |\nabla u|^2}{\int_{\mathbb{M}} |u|^2} : u \in H^2(\mathbb{M}) \right\}.$$

Furthermore, for $j \ge 2$,

$$\lambda_j = \min\left\{ \frac{\int_{\mathbb{M}} |\nabla u|^2}{\int_{\mathbb{M}} |u|^2} : u \in H^2(\mathbb{M}) \text{ and } u \perp u_k \text{ for all } k = 1, ..., j-1 \right\}.$$

Now fix j and denote $u_j = u$ and $\mathcal{N}(u_j) = N$ for simplicity. Let $D_1, ..., D_N$ be the nodal domains of u. Write $w_m = u \cdot \chi_{D_m}$ for $m = 1, ..., N$.

Suppose that $N > j$. Then there are constants $a_1, ..., a_j$ such that

$$w := a_1 w_1 + \cdots a_j w_j \perp u_k$$

for all $k = 1, ..., j-1$. According to the Rayleigh quotient,

$$\frac{\int_{\mathbb{M}} |\nabla w|^2}{\int_{\mathbb{M}} |w|^2} \ge \lambda_j.$$

[i] R. Courant and D. Hilbert, *Methods of mathematical physics.* Vol. I. Interscience Publishers, Inc., New York, NY, 1953.

Apply Green's formula in D_m:

$$\int_{D_m} |\nabla w|^2 = -a_m^2 \int_{D_m} u_j \Delta u_j = \lambda_j a_m^2 \int_{D_m} |u_j|^2,$$

which implies that

$$\frac{\int_{\mathbb{M}} |\nabla w|^2}{\int_{\mathbb{M}} |w|^2} = \frac{\sum_{m=1}^{j} \int_{D_m} |\nabla w|^2}{\sum_{m=1}^{j} \int_{D_m} |w|^2} = \frac{\lambda_j \sum_{m=1}^{j} a_m^2 \int_{D_m} |u_j|^2}{\sum_{m=1}^{j} a_m^2 \int_{D_m} |u_j|^2} = \lambda_j.$$

It then follows that w is an eigenfunction with eigenvalue λ_j. But $w = a_1 w_1 + \cdots a_j w_j = 0$ on D_N since $N > j$. This is impossible because w can not vanish on a non-empty open set. □

Remark. If $\mathbb{M}$ has no boundary, then $\mathcal{N}(u_j) \leq j+1$. In particular, u_1 has two nodal domains and u_j for $j \geq 2$ has at most $j+1$ nodal domains.

Courant's nodal domain theorem in not sharp. Pleijel[i] proved that

Theorem (Pleijel). *Let $\{u_j\}_{j=1}^{\infty}$ be a Dirichlet Laplacian eigenbasis on a compact domain $\mathbb{M} \subset \mathbb{R}^2$ with boundary. Then*

$$\limsup_{j \to \infty} \frac{\mathcal{N}(u_j)}{j} \leq \frac{4}{j_0^2} \approx 0.692,$$

in which $j_0 \approx 2.405$ is the first zero of the zeroth Bessel function (so j_0^2 is the first Dirichlet Laplacian eigenvalue on the unit disc).

Proof. It follows the Faber[ii]-Krahn[iii] inequality that the first eigenvalue in a domain Ω is no smaller than the one on the disc D with the same area:

$$\lambda_1(\Omega) \geq \lambda_1(D) = \frac{\pi j_0^2}{\mathrm{Vol}(D)} = \frac{\pi j_0^2}{\mathrm{Vol}(\Omega)}. \tag{4.2}$$

Hence,

$$\mathrm{Vol}(\Omega) \geq \frac{\pi j_0^2}{\lambda_1(\Omega)}.$$

Suppose that u_j has $\mathcal{N}(u_j) = N_j$ nodal domains $D_1, ..., D_{N_j}$. Then on each nodal domain, λ_j is the first Laplacian eigenvalue. Therefore, for each $k = 1, ..., N_j$,

$$\mathrm{Vol}(D_k) \geq \frac{\pi j_0^2}{\lambda_j},$$

which implies that

$$\mathrm{Vol}(\mathbb{M}) = \sum_{k=1}^{N_j} \mathrm{Vol}(D_k) \geq \frac{\pi j_0^2 N_j}{\lambda_j}.$$

So

$$\frac{N_j}{\lambda_j} \leq \frac{\mathrm{Vol}(\mathbb{M})}{\pi j_0^2}.$$

[i] Å, Pleijel, *Remarks on Courant's nodal line theorem.* Comm. Pure Appl. Math. 9 (1956), 543–550.

[ii] G. Faber, *Beweis, dass unter allen homogenen Membranen von gleicher Fläche und gleicher Spannung die kreisförmige den tiefsten Grundton gibt.* Sitzungsber. Bayer. Akad. Wiss. München, Math.-Phys. Kl. (1923) pp. 169–172.

[iii] E. Krahn, *Über eine von Rayleigh formulierte Minimaleigenschaft des Kreises.* Math. Ann. 94 (1925), no. 1, 97–100.

From the Weyl Law

$$\#\{j : \lambda_j \leq \lambda\} = \frac{1}{4\pi}\mathrm{Vol}(\mathbb{M})\lambda + o(\lambda),$$

we have that

$$\lambda_j = \frac{4\pi}{\mathrm{Vol}(\mathbb{M})} j + o(j).$$

Finally,

$$\frac{\mathrm{Vol}(\mathbb{M})}{\pi j_0^2} \geq \frac{N_j}{\lambda_j} = \frac{N_j}{\frac{4\pi}{\mathrm{Vol}(\mathbb{M})} j + o(j)} \gtrsim \frac{\mathrm{Vol}(\mathbb{M}) N_j}{4\pi j},$$

which implies that

$$\limsup_{j\to\infty} \frac{N_j}{j} \leq \frac{4}{j_0^2}.$$

□

Remark. Notice that the Faber-Krahn inequality becomes equality only if the domain is a disc. Whereas the nodal domains, covering an planar region (except the nodal points), can not be all discs. Therefore, the application of the Faber-Krahn inequality to each nodal domain in the above estimate is rarely sharp. As a result, Pleijel's improvement of Courant's theorem is not sharp either.

Bourgain[i] obtained a small improvement over Pleijel:

$$\limsup_{j\to\infty} \frac{\mathcal{N}(u_j)}{j} \leq \frac{4}{j_0^2} - 10^{-9}.$$

The proof is a combination of two ingredients:

(1). The Faber–Krahn inequality can be improved if the domain is not a disc,

(2). The surface is divided into nodal domains, some of which are not discs.

It is widely believed that the sharp constant in the upper bound of number of nodal domains is achieved by the following example on the rectangle. See, e.g., Polterovich[ii].

Example (The sharp upper bound of number of nodal domains). Let $\mathbb{M} = [0,\pi] \times [0,\pi]$. Then there are Dirichlet eigenfunctions as $u_{l,m} = \sin(lx)\sin(my)$ with eigenvalue $l^2 + m^2$. The number of nodal domains of $u_{l,m}$ is obviously $\mathcal{N}(u_{l,m}) = lm$. The Weyl law implies that the eigenvalue is the j-th eigenvalue such that

$$l^2 + m^2 = \frac{4\pi}{\mathrm{Vol}(\mathbb{M})} j + o(j) = \frac{4}{\pi} j + o(j).$$

So

$$j \approx \frac{\pi(l^2 + m^2)}{4}.$$

Hence,

$$\limsup \frac{\mathcal{N}(u_{l,m})}{j} \geq \frac{4}{\pi} \limsup \frac{lm}{l^2 + m^2} \geq \frac{2}{\pi} \approx 0.637.$$

Note that the conjectured sharp constant $frac2\pi$ is in fact very close to $\frac{4}{j_0^2} \approx 0.692$ in Pleijel's theorem, albeit very difficult to prove.

[i]J. Bourgain, *On Pleijel's nodal domain theorem.* Int. Math. Res. Not. IMRN 2015, no. 6, 1601–1612.

[ii]I. Polterovich, *Pleijel's nodal domain theorem for free membranes.* Proc. Amer. Math. Soc. 137 (2009), no. 3, 1021–1024.

Problem 4.5 (Sharp upper bound of number of nodal domains). *Prove that on any surface*
$$\limsup_{j\to\infty} \frac{\mathcal{N}(u_j)}{j} \leq \frac{2}{\pi}.$$

4.1.3.2. *Lower bounds.* In general, there is no lower bound of the number of nodal domains. Indeed, Lewy[i] constructed spherical harmonics (i.e., Laplacian eigenfunctions on the sphere) that have at most two or three nodal domain as the eigenvalues go to infinity. (In fact, such result was already known earlier due to Stern in 1920s, see Bérard-Helffer[ii].)

However, it is widely believed that there is always eigenfunctions with growing number of nodal domains.

Problem 4.6 (Eigenfunctions with number of nodal domains tending infinity). *Prove that on any surface there is a subsequence of eigenfunctions whose numbers of nodal domains go to infinity. That is,*
$$\limsup_{\lambda\to\infty} \mathcal{N}(u) = \infty.$$

Some results are known due to Ghosh-Reznikov-Sarnak[iii], Jung-Zelditch[iv], Jang-Jung[v], etc on surfaces with negative curvature. Otherwise the problem is open.

4.1.4. Nodal sets and volume spectrum. Let $\mathbb{M}$ be a n-dim compact manifold and $p \in \mathbb{N}$. Then we say a family F of hypersurfaces is a p sweep-out of $\mathbb{M}$, if for any p points $x_1, ..., x_p$ in $\mathbb{M}$, there is $c \in F$ such that $x_i \in c$ for all $i = 1, ..., p$. Set
$$\omega_p(\mathbb{M}) = \inf_F \sup_{c\in F} \mathcal{H}^{n-1}(c).$$

Then Gromov[vi] and Guth[vii] proved that
$$c\mathrm{Vol}(\mathbb{M})^{\frac{n-1}{n}} p^{\frac{1}{n}} \leq \omega_p(\mathbb{M}) \leq C\mathrm{Vol}(\mathbb{M})^{\frac{n-1}{n}} p^{\frac{1}{n}}$$
for some positive constants c and C depending on $\mathbb{M}$. Liokumovich-Marques-Neves[viii] proved an asymptotic for this volume spectrum
$$\omega_p(\mathbb{M}) = c_n \mathrm{Vol}(\mathbb{M})^{\frac{n-1}{n}} p^{\frac{1}{n}} + o\left(p^{\frac{1}{n}}\right).$$

---

[i] 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.

[ii] P. Bérard and B Helffer, *A. Stern's analysis of the nodal sets of some families of spherical harmonics revisited.* Monatsh. Math. 180 (2016), no. 3, 435–468.

[iii] A. Ghosh,A. Reznikov, and P. Sarnak, Peter, *Nodal domains of Maass forms, I.* Geom. Funct. Anal. 23 (2013), no. 5, 1515–1568, and *Nodal domains of Maass forms, II.* Amer. J. Math. 139 (2017), no. 5, 1395–1447.

[iv] 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, and *Number of nodal domains of eigenfunctions on non-positively curved surfaces with concave boundary.* Math. Ann. 364 (2016), no. 3-4, 813–840.

[v] S. Jang and J. Jung, *Quantum unique ergodicity and the number of nodal domains of eigenfunctions.* J. Amer. Math. Soc. 31 (2018), no. 2, 303–318.

[vi] M. Gromov, *Systoles and intersystolic inequalities.* Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992), 291–362, Sémin. Congr., 1, Soc. Math. France, Paris, 1996.

[vii] L. Guth, *Minimax problems related to cup powers and Steenrod squares.* Geom. Funct. Anal. 18 (2009), no. 6, 1917–1987.

[viii] Y. Liokumovich, A. Neves, and F. Marques, *Weyl law for the volume spectrum.* Ann. of Math. (2) 187 (2018), no. 3, 933–961.

It can be called the Weyl law for the volume spectrum, i.e., the analogue of the Weyl Law for the Laplacian eigenfunctions, see Section 9.1: Let $\lambda_j^2 \to \infty$ be the Laplacian eigenvalues on a Riemannian surface. Then

$$\lambda_j = \frac{2\sqrt{\pi}}{\mathrm{Vol}(\mathbb{M})^{\frac{1}{2}}} j^{\frac{1}{2}} + o\left(j^{\frac{1}{2}}\right).$$

In particular, when $n = 2$,

$$\omega_p(\mathbb{M}) = c_2 \mathrm{Vol}(\mathbb{M})^{\frac{1}{2}} p^{\frac{1}{2}} + o\left(p^{\frac{1}{2}}\right).$$

Comparing with the Weyl Law for eigenvalues, the explicit constant in the volume spectrum is not known. In fact, the explicit volume spectrum is not known in any manifold.

Marques-Neves[i] suggests that c_n may be computed via the maximal nodal sets of the linear combinations of eigenfunctions, that is,

$$\omega_p(\mathbb{M}) = \max_{\|u\|_{L^2(\mathbb{M})}=1} \left\{\mathcal{H}^{n-1}\left(\mathcal{N}(u)\right), u = c_0 u_0 + c_1 u_1 + \cdots c_p u_p\right\}.$$

According to this intuition, we can compute the explicit volume spectrum on $\mathbb{S}^2$: Let $\lambda^2 = k(k+1)$. Then there are

$$p = \frac{1}{4\pi} \mathrm{Vol}(\mathbb{S}^2)\lambda^2 + o\left(\lambda^2\right) \approx k^2$$

eigenfunctions in the linear combination. Observe that

$$\omega_p(\mathbb{S}^2) = \max_{\|u\|_{L^2(\mathbb{M})}=1} \{\mathcal{H}^{n-1}(\mathcal{N}(u)), u = c_0 u_0 + c_1 u_1 + \cdots c_p u_p\} = 2k\pi.$$

The maximum is obtained by Gaussian beams, see Section 4.3. Hence,

$$c_2 = \frac{\omega_p(\mathbb{S}^2)}{\mathrm{Vol}(\mathbb{M})^{\frac{1}{2}} p^{\frac{1}{2}}} = \sqrt{\pi}.$$

But of course one would have to prove Marques-Neves' intuition that the volume spectrum on $\mathbb{M} = \mathbb{S}^2$ is achieved by spherical harmonics.

Problem 4.7 (Volume spectrum and nodal sets). *What is the constant c_n in the Weyl Law for the volume spectrum?*

4.1.5. Overlapping nodal curves and nodal intersections. The nodal intersection problem is about estimating the size of the intersection between nodal sets $\mathcal{N}(u)$ and a fixed submanifold $H \subset \mathbb{M}$, as the eigenvalues tend to infinity. Such problems are in general difficult so we begin from the case on surfaces when H is a curve and $\mathcal{N}(u)$ is a union of nodal curves.

Problem 4.8 (Nodal intersections). *Let $H \subset \mathbb{M}$ be a fixed smooth curve on a surface $\mathbb{M}$. Then estimate the number of intersections*

$$\#\{H \cap \mathcal{N}(u)\}.$$

But before answering this problem, one needs to rule out the case when H is contained in the nodal sets of infinitely many eigenfunctions (so the intersection becomes 1-dim). That is, infinitely many eigenfunctions vanish on the same curve. Result on the torus, as well as examples on the disc and on the sphere, suggest that

Problem 4.9 (Overlapping nodal curves must be geodesics). *Suppose that a curve H is contained in the nodal sets of infinitely many eigenfunctions on a surface. Then H must be a piece of a geodesic.*

[i]F. Marques and A. Neves, *Existence of infinitely many minimal hypersurfaces in positive Ricci curvature.* Invent. Math. 209 (2017), no. 2, 577–616.

This overlapping nodal curve problem is only answered on the torus, which is due to Bourgain-Rudnick[i]. The basic toral eigenmodes of sine or cosine functions of course can have overlapping nodal curves as a piece of straight lines. On the other hand, a fixed curve H is contained in the nodal sets of infinitely many toral eigenfunctions iff H is straight. It is interesting to see if this result generalizes to other surfaces, such as the disc in Section 4.2 and the sphere in Section 4.3.

4.1.6. Vanishing orders. Donnelly-Fefferman[ii] proved that the vanishing order of an eigenfunction u with eigenvalue λ^2 is at most $C\lambda$, in which $C > 0$ depends only on the manifold.

This vanishing order bound is sharp on the sphere and is achieved by the Gaussian beams at their poles, for example, the Gaussian beam $(x_1 + ix_2)^k$ has eigenvalue $k(k+n-1)$ on $\mathbb{S}^n$ and has vanishing order k at the points $x = (x_1, x_2, ..., x_{n+1})$ with $x_1 = x_2 = 0$.

Zelditch asked whether the maximal vanishing order must be achieved by surfaces and eigenfunctions that are similar to the spheres and the Gaussian beams, for example, the points at which the maximal vanishing order is obtained on $\mathbb{S}^2$ are "focal", i.e., all geodesics from this point return at the same time.

Problem 4.10 (Maximal vanishing order and focal points). *Suppose that there is a sequence of eigenfunctions which achieve the maximal vanishing order at a point on a surface. Then must this point be focal?*

4.2. Nodal sets on the disc

Let $B = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ be the unit disc. Then the Dirichlet eigenfunctions on B are the nonzero solutions to

$$\begin{cases} \Delta\varphi(x,y) = \mu\varphi(x,y), & \text{if } (x,y) \in B, \\ \varphi(x,y) = 0, & \text{if } (x,y) \in \partial B. \end{cases}$$

Here,

$$\Delta = \Delta_{\mathbb{R}^2} = \begin{cases} \partial_x^2 + \partial_y^2, & \text{in the rectangular coordinates } (x,y), \\ \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2, & \text{in the polar coordinates } (r,\theta). \end{cases}$$

The Dirichlet eigenfunctions are

$$J_k\left(\sqrt{\lambda_{k,s}} \cdot r\right) e^{\pm ik(\theta+\theta_0)}, \quad \text{in which } k = 0,1,2,3,... \text{ and } s = 1,2,3,...$$

Here, J_k is the k-th Bessel function and $\sqrt{\lambda_{k,s}} = j_{k,s}$, where $j_{k,s}$ is the s-th nonnegative zero of J_k. That is, J_k solves the differential equation

$$u''(z) + \frac{1}{z}u'(z) + \left(1 - \frac{k^2}{z^2}\right)u(z) = 0,$$

So all the Dirichlet eigenvalues are the squares of zeros of Bessel functions.

[i]J. Bourgain and Z. Rudnick, *On the nodal sets of toral eigenfunctions.* Invent. Math. 185 (2011), no. 1, 199–237.

[ii]H. Donnelly and C. Fefferman, *Nodal sets of eigenfunctions on Riemannian manifolds.* Invent. Math. 93 (1988), no. 1, 161–183.

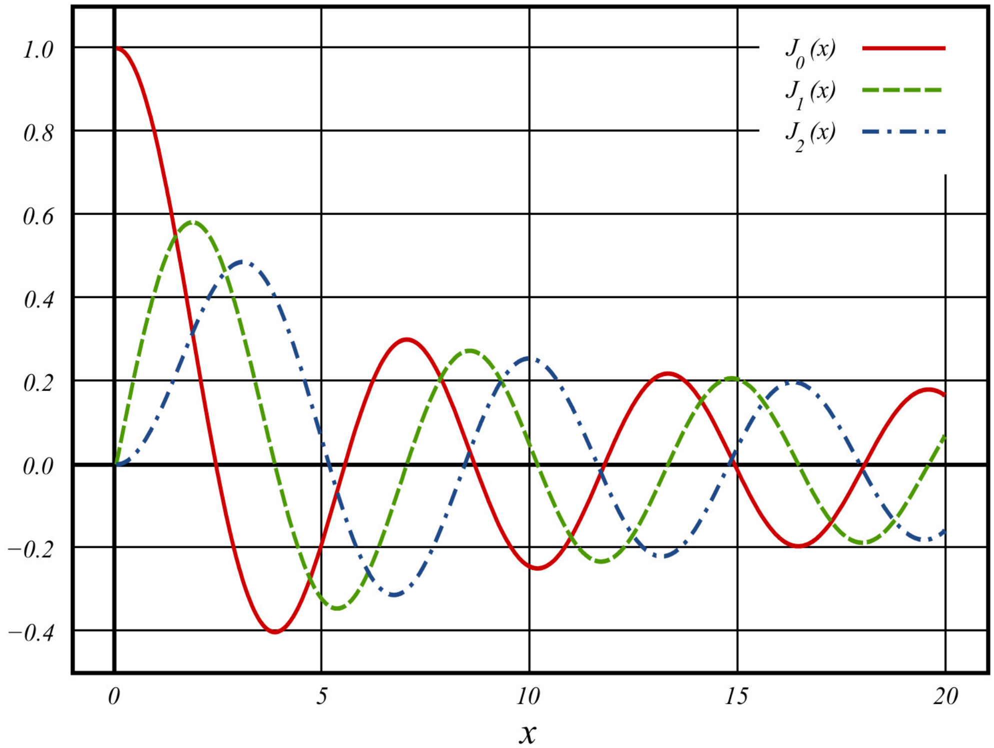

FIGURE 4.1. Bessel polynomials J_k with $0 \leq k \leq 2$ (Credit: Wikipedia)

From Table 9.5 in Abramowitz-Stegun[i],

$$j_{0,1} \approx 2.40483.$$

The Neumann eigenfunctions on B are the nonzero solutions to

$$\begin{cases} \Delta\psi(x,y) = \mu\psi(x,y), & \text{if } (x,y) \in B, \\ \partial_\nu\psi(x,y) = 0, & \text{if } (x,y) \in \partial B. \end{cases}$$

Here, at $(x,y) \in \partial B$, $r = 1$ in the polar coordinates so

$$\partial_\nu = \begin{cases} x\partial_x + y\partial_y & \text{in the rectangular coordinates } (x,y), \\ \partial_r & \text{in the polar coordinates } (r,\theta). \end{cases}$$

The Neumann eigenfunctions are

$$J_k\left(\sqrt{\mu_{k,s}} \cdot r\right) e^{\pm ik(\theta+\theta_0)}, \quad \text{in which } k = 0,1,2,3,\ldots \text{ and } s = 1,2,3,\ldots$$

Here, J_k is the k-th Bessel function and $\sqrt{\mu_{k,s}} = j'_{k,s}$, where $j'_{k,s}$ is the s-th nonnegative zero of J'_k. So all the Neumann eigenvalues are the squares of zeros of the derivatives of Bessel functions.

From Table 9.5 in Abramowitz-Stegun, $j'_{0,1} = 0$, which corresponding to the 0-th Neumann eigenfunction, a constant function. Moreover, $j_{0,2} \approx 3.83171$. According to Section 9.5 in Abramowitz-Stegun, we can list $j_{k,j}$ and $j'_{k,j}$ such that

$$k \leq j'_{k,1} < j_{k,1} < j'_{k,2} < j_{k,2} < \cdots.$$

[i]M. Abramowitz and I. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* National Bureau of Standards. U.S. Government Printing Office, Washington, DC, 1964.

However, this ordering for a fixed k does not answer the following problem.

Problem 4.11 (Interlacing of Dirichlet and Neumann eigenvalues on the disc). *Do the Dirichlet and Neumann eigenvalues on the disc interlace?*

Remark (Nodal curves of real Dirichlet eigenfunctions on the disc). Consider the real Dirichlet eigenfunctions on the disc

$$\varphi_{k,s}(r,\theta) = J_k\left(\sqrt{\lambda_{k,s}} \cdot r\right) \sin(k\theta).$$

- The nodal set $\mathcal{N}(\varphi_{k,s})$ consists of s concentric circles with radii $j_{k,l}/j_{k,s}$, $l = 1, ..., s-1$, and $2k$ radials. So the total length of its nodal curves:
$$\mathcal{H}^1(\mathcal{N}(\varphi_{k,s})) = 2\pi \sum_{l=1}^{s} \frac{j_{k,l}}{j_{k,s}} + 2k.$$
- The critical set $\mathcal{C}(\varphi_{k,s})$ consists of all the intersecting points of the concentric circles and the radials. So the total number of its critical points:
$$\mathcal{H}^0(\mathcal{C}(\varphi_{k,s})) = 2ks.$$
- The nodal curves of $\varphi_{k,s}$ divide B into $2ks$ nodal domains.

Han-Murray-Tran[i] proved that

$$H_l(B) = \liminf_{\lambda\to\infty} \frac{\mathcal{H}^1(\mathcal{N}(\varphi))}{\sqrt{\lambda}} = 1,$$

which are saturated by the eigenfunctions $\varphi_{0,s}$ as $s \to \infty$, and

$$H_u(B) = \limsup_{\lambda\to\infty} \frac{\mathcal{H}^1(\mathcal{N}(\varphi))}{\sqrt{\lambda}} = 2,$$

which are saturated by the eigenfunctions $\varphi_{k,1}$ as $k \to \infty$.

Next we discuss overlapping nodal curves in Problem 4.9. That is, for a fixed curve $H \subset B$, can H be contained in the nodal sets of infinitely many eigenfunctions? In this case, nodal curves are either diameters or circles centered at the origin with radii

$$\frac{j_{k,l}}{j_{k,s}}.$$

Clearly, if H is a piece of diameter, then one can find infinitely many eigenfunctions that vanish on H. On the other hand, two nodal circles overlap iff

$$\frac{j_{k,l}}{j_{k,s}} \neq \frac{j_{\tilde{k},\tilde{l}}}{j_{\tilde{k},\tilde{s}}}.$$

for $(k,l,s) \neq (\tilde{k},\tilde{l},\tilde{s})$, equivalently, $(j_{k,l}, j_{k,s}, j_{\tilde{k},\tilde{l}}, j_{\tilde{k},\tilde{s}})$ solve the polynomial equation with integral coefficients $xw - zy = 0$, and therefore are algebraically dependent. It is plausible that this never happens, i.e., the nodal curves as circles never overlap. Indeed, a stronger condition might be true:

Problem 4.12 (Algebraic independence of zeros of Bessel functions). *The zeros of Bessel functions are algebraically independent.*

[i]X. Han, M. Murray, and C. Tran, *Nodal lengths of eigenfunctions in the disc.* Proc. Amer. Math. Soc. 147 (2019), no. 4, 1817–1824.

Remark. Notice that
$$j_{k,s} \sim \left(s + \frac{k}{2} - \frac{1}{4}\right)\pi \quad \text{as } s \to \infty.$$
So
$$\frac{j_{k,l}}{j_{k,s}} \sim \frac{l + \frac{k}{2} - \frac{1}{4}}{s + \frac{k}{2} - \frac{1}{4}}.$$
Hence, one can easily arrange (k, l, s) and $(\tilde{k}, \tilde{l}, \tilde{s})$ such that $j_{k,l}/j_{k,s}$ equals $j_{\tilde{k},\tilde{l}}/j_{\tilde{k},\tilde{s}}$ asymptotically. This essentially asks for extra help from algebraic number theory, because the asymptotic formulas are the best one can get using analysis only.

Using the Siegel-Shidlovskii theorem (see Section 10.19), we show a weaker fact that all the zeros $j_{k,s}$ and $j'_{k,s}$ of Bessel functions J_k and their derivatives J'_k are transcendental (except 0). This was first proved by Siegel[i].

Theorem (Siegel). *Let $\alpha \neq 0$ be a zero of J_k or J'_k. Then α is transcendental.*

PROOF. Notice that J_k solves the differential equation
$$u''(z) + \frac{1}{z}u'(z) + \left(1 - \frac{k^2}{z^2}\right)u(z) = 0,$$
Set $f_1 = u$ and $f_2 = u'$. Then (f_1, f_2) solves the first order differential system
$$\begin{pmatrix} f_1' \\ f_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 + \frac{k^2}{z^2} & \frac{1}{z} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}.$$
The common denominator of the entries in the matrix is $T(z) = z^2$.

Because f_1 and f_2 are linearly independent over $\overline{\mathbb{Q}}[z]$, for any $\alpha \in \overline{\mathbb{Q}}$ and $\alpha T(\alpha) \neq 0$ (i.e., $\alpha \neq 0$), $f_1(\alpha)$ and $f_2(\alpha)$ are linearly independent over $\overline{\mathbb{Q}}$, by the Siegel-Shidlovskii theorem in Section 10.19. But this implies that α can not be a zero of f_1 or f_2, that is, the zeros of f_1 and f_2 are all transcendental (except 0). $\square$

Similar argument shows that the non-zero zeros of the sine or cosine functions are transcendental, i.e., π is transcendental. See Section 10.19. However, notice that the zeros of the sine function, $k\pi$ with $k \in \mathbb{Z}$, are clearly linearly dependent so are also algebraically dependent. To answer Problem 4.12, one will have to distinguish Bessel functions and sine function, maybe using the fact that the differential equation for Bessel function has a regular singularity at 0, while the differential equation for sine function $u'' - u = 0$ has no singularity.

As a consequence of their transcendental property, we show that there are no common zeros of Bessel functions J_k's with different k. This is in fact the Bourget's hypothesis, see Section 15.28 in Watson[ii]. Indeed, we first notice that J_k and J_{k+1} never share any common zeros. In fact, we can list the zeros of J_k and J_{k+1} such that
$$j_{k,1} < j_{k+1,1} < j_{k,2} < j_{k+1,2} < j_{k,3} < \cdots.$$
See Theorem 9.5.2 in Abramowitz-Stegun.

To pass to J_{k+m} and J_k, we need a recurrence relation that
$$J_{k+1}(z) = \frac{2k}{z}J_k(z) - J_{k-1}(z).$$

---

[i]C. Siegel, *Über einige Anwendungen diophantischer Approximationen.* Abhandlungen der Preußischen Akademie der Wissenschaften. Physikalisch-mathematische Klasse (1929) Nr. 1.

[ii]G. Watson, *A Treatise on the Theory of Bessel Functions.* Second edition. Cambridge University Press, Cambridge, 1944.

See Theorem 9.1.27 in Abramowitz-Stegun. Inductively, we have that

$$J_{k+m}(z) = R_{m,k}(z)J_k(z) - R_{m-1,k+1}(z)J_{k-1}(z).$$

in which $R_{m,k}(z)$ is a rational function such that $R_{m,k}(1/z)$ is a polynomial of degree m (which are called the Lommel polynomials, see Section 9.6 in Watson). Therefore, any common zero of J_{k+m} and J_k must be a zero of $R_{m-1,k+1}$, that is, it is algebraic. But this is impossible by the theorem above.

4.3. Nodal sets on the sphere: Spherical harmonics, Take III

We first present the nodal sets of spherical harmonics in the standard basis on $\mathbb{S}^2$, in which case the pictures are clearest. We denote by $\mathbb{SH}_k = \mathbb{SH}_k^2$ the eigenspace of spherical harmonics with eigenvalue $k(k+1)$. Then $\dim \mathbb{SH}_k = 2k+1$. In the spherical coordinates, write the longitudinal coordinate $\phi \in [0,\pi]$ and latitudinal coordinates $\theta \in [0, 2\pi)$ so that

$$\mathbb{S}^2 \ni x = (\sin\phi\cos\theta, \sin\phi\sin\theta, \cos\phi).$$

Then one can write the standard eigenbasis of $\mathbb{SH}_k$ as $\{Y_m^k\}_{m=-k}^k$:

$$Y_k^m(\phi,\theta) = cP_k^{|m|}(\cos\phi)e^{im\theta},$$

in which $c = c(k,m) > 0$ is the L^2 normalization factor and P_k^m is the associated Legendre polynomial. See Section 2.1 in Han-Lin[i].

Remark (Associated Legendre polynomials). Let $k \geq 1$ and $m = 0, ..., k$.

- P_k^m is the solution to the second order linear differential equation
$$(1-z^2)u''(z) - 2zu'(z) + \left[k(k+1) - \frac{m^2}{1-z^2}\right]u(z) = 0.$$
- We have that
$$P_k^m(z) = (1-z^2)^{\frac{m}{2}}\frac{d^{k+m}}{dz^{k+m}}(1-z^2)^k.$$
is a polynomial of degree k.
- When $m = 0$,
$$P_k(z) := P_k^0(z) = \frac{d^k}{dx^k}(1-z^2)^k$$
is called the Legendre polynomial of degree k. For example,
$$P_0(z) = 1,$$
$$P_1(z) = -2z,$$
$$P_2(z) = -4 + 12z^2.$$

[i]Q. Han and F.-H. Lin, *Nodal sets of elliptic partial differential equations*. Lecture notes.

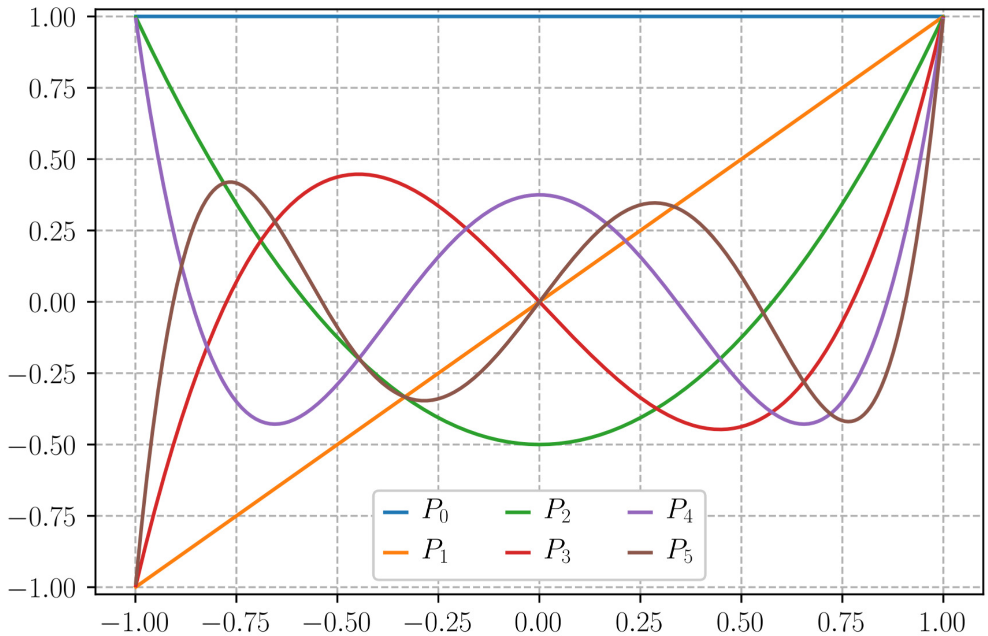

FIGURE 4.2. Legendre polynomials P_k with $0 \leq k \leq 5$ (Credit: Wikipedia)

- When $m = k$,
$$P_k^k(z) = (1 - z^2)^{\frac{k}{2}}.$$
- $P_k^m(z)$ has $k - m$ zeros of in $z \in [-1, 1]$ which are symmetric with respect to $z = 0$. In particular, if $k - m$ is odd, then $P_k^m(0) = 0$.

Consider the real-valued spherical harmonics in the standard basis:
$$Y_k^m(\phi, \theta) = cP_k^{|m|}(\cos\phi)\sin(m\theta).$$
The nodal set as a union of m great circles (from $\sin(m\theta) = 0$) and $k - |m|$ latitudinal circles (from $P_k^{|m|}(\cos\phi) = 0$). Accordingly, Y_k^m can be categorized as follows.

(1). $m = 0$: $Z_k = Y_k^0$ are called zonal harmonics, whose nodal set is a union of k latitudinal circles.

(2). $-k < m < k$: Y_m^k are called tesseral harmonics, whose nodal set is a union of $|m|$ great circles (that passes through the north and south pole) and $k - |m|$ latitudinal circles.

(3). $m = \pm k$: $Q_{\pm k} = Y_{\pm k}^k$ are called sectoral harmonics (i.e., Gaussian beams), whose nodal set is a union of k great circles.

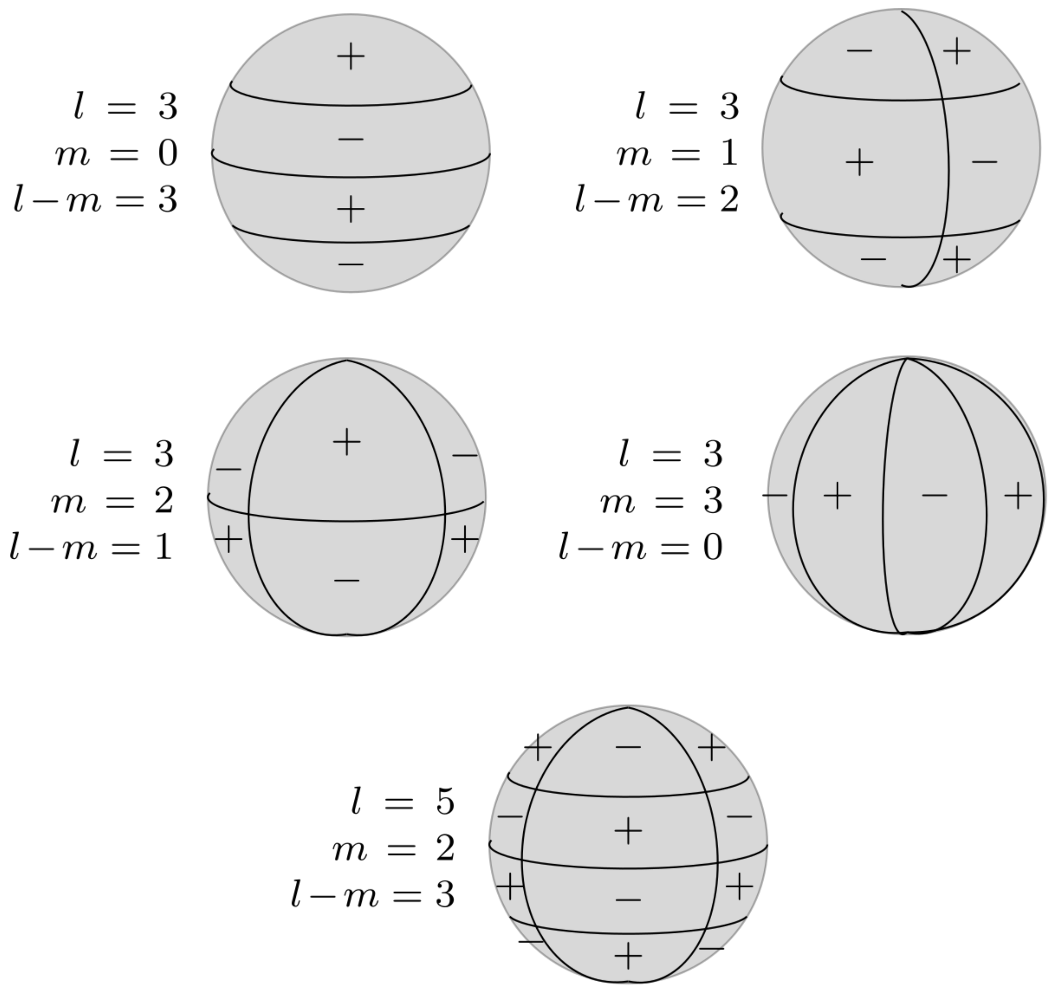

FIGURE 4.3. Nodal portraits of zonal, tesseral, or sectoral harmonics on $\mathbb{S}^2$ (Credit: Wikipedia)

Notice that each spherical harmonics Y_m^k in the standard eigenbasis always has k total nodal curves (great circles and latitudinal circles). Therefore, it is obvious that the Gaussian beam has the longest total nodal length (since all nodal curves are great circles).

Indeed, Gichev[i] proved a stronger result that Gaussian beams have the longest total nodal length among all spherical harmonics in $\mathbb{SH}_k$ (not only the ones from the standard eigenbasis). In fact, his result holds on $\mathbb{S}^n$ for any n. Consider a Gaussian beam in $\mathbb{SH}_k^n$ that is given by

$$u_k(x) = \Re(x_1 + ix_2)^k \in \mathbb{SH}_k^n.$$

Notice that u_k is not normalized but it does not affect the nodal set estimates. Indeed, the nodal set of u_k is a union of k copies of $\mathbb{S}^{n-1}$. So

$$\mathcal{H}^{n-1}(\mathcal{N}(u_k)) = kS_{n-1},$$

in which S_{n-1} is the area of $\mathbb{S}^{n-1}$. Recall Theorem 3.2.48 in Federer[ii]:

[i] V. Gichev, *Some remarks on spherical harmonics*. St. Petersburg Math. J. 20 (2009), no. 4, 553–567.

[ii] H. Federer, *Geometric measure theory*. Springer-Verlag, New York, NY, 1969.

Theorem. *Let $A, B \subset \mathbb{S}^n$ be compact such that A is k-rectifiable and B is l-rectifiable. Set $r = k + l - n$. Suppose that $r \geq 0$. Then*

$$\int_{\mathrm{O}(n+1)} \mathcal{H}^r(A \cap gB)\, d\mu_g = C\mathcal{H}^k(A)\mathcal{H}^l(B),$$

in which $C = S_r/(S_k S_l)$ and μ_g is normalized Haar measure on $\int_{\mathrm{O}(n+1)}$ so that $\int_{\mathrm{O}(n+1)} 1\, d\mu_g = 1$.

To apply the above theorem to the nodal size estimate of $u \in \mathbb{SH}_k$, take $A = \mathcal{N}(u)$ and $B = \mathbb{S}^1$. Then gB is also a great circle on $\mathbb{S}^n$. Since u is a homogeneous polynomial of degree k, the restriction of u to the plane that contains $g\mathbb{S}^1$ is a homogeneous polynomial with two variables and has degree k. Then restricting to $g\mathbb{S}^1$, we have that

$$\mathcal{H}^0(\mathcal{N}(u) \cap g\mathbb{S}^1) \leq 2k.$$

The equation holds as long as $\mathcal{N}(u) \cap g\mathbb{S}^1$ is a discrete set of points, which is true almost everywhere for $g \in \mathrm{O}(n+1)$.

With $k = n-1$, $l = 1$, and $r = 0$, we compute that

$$C\mathcal{H}^{n-1}(\mathcal{N}(u))\mathcal{H}^1(\mathbb{S}^1) = \int_{\mathrm{O}(n+1)} \mathcal{H}^0(\mathcal{N}(u) \cap g\mathbb{S}^1)\, dg \leq 2k,$$

in which

$$C = \frac{S_0}{S_{n-1}S_1} = \frac{1}{\pi S_{n-1}}.$$

Hence,

$$\mathcal{H}^{n-1}(\mathcal{N}(u)) \leq kS_{n-1}.$$

Remark. This approach on $\mathbb{S}^n$ is similar in spirit to one that uses the integral geometric formula, Theorem 4.1, on $\mathbb{T}^n$ (or in $\mathbb{R}^n$). That is, to estimate the size of a $(n-1)$-dim set E in a n-dim space, one uses the "test lines" l to intersect E and count the intersecting points, i.e., $\mathcal{H}^0(E \cap l)$. In $\mathbb{R}^n$, the test lines are chosen as n families of straight lines that are parallel to the axes. On $\mathbb{S}^n$, the test lines are chosen as one family of great circles.

To estimate the largest nodal size, the argument of "straight lines" in $\mathbb{R}^n$ lead to certain loss in the integral geometric formula, when E is not parallel to any axis. Whereas the argument of "great circles" on $\mathbb{S}^n$ avoids such loss. However, the latter is highly dependent on the symmetry of $\mathbb{S}^n$ and seems difficult to generalize.

In particular, the longest total nodal length of spherical harmonics in $\mathbb{SH}_k$ is always achieved by Gaussian beams, whose nodal set is a union of k great circles. Recall that the same result for nodal sets on the disc in Section 4.2. One therefore asks

Problem 4.13 (Maximal nodal length and geodesics). *On a surface, the maximal nodal length is saturated by eigenfunctions whose nodal curves are all geodesics.*

On the other hand, we estimate the total nodal length of zonal harmonics in $\mathbb{SH}_k$. In this case,

$$Z_k(\phi, \theta) = c_{k,m} P_k(\cos\theta),$$

in which $P_k(\cos\theta)$ has zeros $\theta_1, ..., \theta_k$ in $(0, \pi)$ satisfying that

$$\frac{2j-1}{2k+1}\pi \leq \theta_j \leq \frac{2j}{2k+1}\pi \quad \text{for } j = 1, ..., k.$$

See Section 6.21 in Szegö[i]. Hence,

$$\mathcal{H}^1(\mathcal{N}(Z_k)) = \sum_{j=1}^{k} 2\pi \sin(\theta_j) \approx 2\pi \cdot \frac{k}{\pi} \int_0^{\pi} \sin\theta \, d\theta = 4k.$$

which is of course smaller than $2\pi k$, the longest total nodal curves for any spherical harmonics in $\mathbb{SH}_k$. It is plausible, but not known, that zonal harmonics have the shortest total nodal length:

Problem 4.14 (Nodal size of zonal harmonics). *Zonal harmonics have the shortest total nodal length among all the spherical harmonics of the same degree on* $\mathbb{S}^2$.

Next we discuss the problem of overlapping nodal curves on $\mathbb{S}^2$, again from the spherical harmonics in the standard eigenbasis of $\mathbb{SH}_k$:

$$Y_m^k(\phi, \theta) = cP_k^{|m|}(\cos\phi)e^{im\theta},$$

That is, when can the nodal curves of different spherical harmonics overlap?

Example (Gaussian beams). Since Gaussian beams have nodal curves as great circles, it is straightforward to find examples of overlapping nodal curves. For example, let $u_k(x, y, z) = \Re(x + iy)^k \in \mathbb{SH}_k$. Then u_k vanishes on the great circle as the intersection of the xz-plane and $\mathbb{S}^2$ for all $k \in \mathbb{N}$.

Example (Zonal harmonics). Since zonal harmonics have nodal curves as latitudinal circles, two zonal harmonics $Z_k \in \mathbb{SH}_k$ and $Z_j \in \mathbb{SH}_j$ share a nodal curve iff the Legendre polynomials Y_k and Y_j have a common zero. Indeed, Stieltjes conjectured that Legendre polynomials do not share common zeros except at 0 (in which case the corresponding nodal curve is a great circle). See Bourgain-Rudnick.

Unlike Bessel functions in Section 4.2, Legendre polynomials $P_k \in \mathbb{Z}[z]$ have degree k and their zeros are automatically algebraic. While we use the Siegel-Shidlovskii theorem to show that the zeros of Bessel functions are all transcendental (except 0), it is unclear how to prove linear or algebraic independence among the zeros of Legendre polynomials, such as in the Stieltjes conjecture.

4.4. Nodal sets on the torus: Toral eigenfunctions, Take III

Let $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ be the torus. An eigenfunction on $\mathbb{T}^2$ can be written as

$$u(x) = \sum_{k \in \mathbb{Z}^2, |k| = \lambda} a_k e^{2\pi i k \cdot x},$$

in which $x = (x_1, x_2) \in \mathbb{T}^2$ and $k = (k_1, k_2) \in \mathbb{Z}^2$. We assume that u is real-valued. Let

$$\mathcal{N}(u) = \{x \in \mathbb{T}^2 : u(x) = 0\}$$

be the nodal set of u. We want to estimate the maximal and minimal nodal length of u in the eigenspace with eigenvalue $4\pi^2|k|^2 = 4\pi^2\lambda^2$.

Example. Let $u_k(x_1, x_2) = \cos(2\pi k_1 x_1)\cos(2\pi k_2 x_2)$ with $k_1, k_2 \in \mathbb{N}$. (The choice of cos or sin here is irrelevant.) Then u has $2k_1$ nodal lines that are parallel to x_2-axis (when $2\pi k_1 x_1 \in \pi\mathbb{Z} + \pi/2$) and $2k_2$ nodal lines that are parallel to x_1-axis (when $2\pi k_2 x_2 \in \pi\mathbb{Z} + \pi/2$). Hence,

$$\mathcal{H}^1(\mathcal{N}(u_k)) = 2k_1 + 2k_2.$$

[i] G. Szegö, *Orthogonal polynomials*. Fourth edition. American Mathematical Society, Providence, RI, 1975.

Notice that
$$k_1^2 + k_2^2 = |k|^2 = \lambda^2.$$
We have that
$$\lambda = \sqrt{k_1^2 + k_2^2} \le k_1 + k_2 \le \sqrt{2}\sqrt{k_1^2 + k_2^2} = \sqrt{2}\lambda.$$
Furthermore,

(1). the equality on the left-hand-side is attained when $k_1 = 0$ or $k_2 = 0$, i.e., $u_k(x_1, x_2) = \cos(2\pi k_1 x_1)$ or $\cos(2\pi k_2 x_2)$. In this case, the nodal length of u_k is 2λ;

(2). the equality on the right-hand-side is attained (asymptotically) when $k_1 = k_2 = \sqrt{2}\lambda/2$. In this case, the nodal length of u_k is $2\sqrt{2}\sqrt{\lambda}$.

It is very likely that these two examples provide maximal and minimal nodal lengths among all the eigenfunctions in the eigenspace.

Problem 4.15 (Sharp nodal size estimates on the torus). *Prove that*
$$H_l(\mathbb{T}^2) = \liminf_{\lambda\to\infty} \frac{\mathcal{H}^1(\mathcal{N}(u))}{2\pi\lambda} = \frac{1}{\pi} \quad \textit{and} \quad H_u(\mathbb{T}^2) = \limsup_{\lambda\to\infty} \frac{\mathcal{H}^1(\mathcal{N}(u))}{2\pi\lambda} = \frac{\sqrt{2}}{\pi}.$$

The integral geometric formula is a vital tool for the nodal set estimate.

THEOREM 4.1 (Integral geometric formula). *Let* $\mathrm{P}_j : \mathbb{R}^n \to \mathbb{R}$ *be the projection from* $\mathbb{R}^n$ *to the one deleting the* j*-th variable, i.e.,*
$$\mathrm{P}_j(x_1, ..., x_n) = (x_1, ..., x_{j-1}, x_{j+1}, ..., x_n).$$
Suppose that E *is a smooth* $(n-1)$-dim *hypersurface in* $\mathbb{R}^n$*. Write*
$$a_j = \int_{\mathbb{R}^{n-1}} \mathcal{H}^0(E \cap \mathrm{P}_j^{-1}(y))\, dy.$$
Then
$$\left(\sum_{j=1}^n a_j^2\right)^{\frac{1}{2}} \le \mathcal{H}^{n-1}(E) \le \sum_{j=1}^n a_j.$$

See Theorem 1.2.10 in Han-Lin for more background of the integral geometric formula. In particular, it has a clear geometric interpretation: To estimate the size of a $(n-1)$-dim object in $\mathbb{R}^n$, we only need to examine the intercepts of this set with all straight lines parallel to axis.

Inspired by Appendix B in Bourgain-Rudnick[i], we complexify the toral eigenfunction and write the corresponding Laurent polynomial
$$u^{\mathbb{C}}(z_1, z_2) = \sum_{|k|=\lambda} a_k z_1^{k_1} z_2^{k_2},$$
in which $z = (z_1, z_2) \in \mathbb{S}^1 \times \mathbb{S}^1$ is defined by $z_1 = e^{2\pi i x_1}$ and $z_1 = e^{2\pi i x_2}$. Note that $|dz|/|dx| = 2\pi$. It is then obvious that the nodal length of u on $\mathbb{T}^2$ and $u^{\mathbb{C}}$ in $\mathbb{C}^2$ can be compared:
$$\mathcal{H}^1(\mathcal{N}(u)) = \frac{\mathcal{H}^1(\mathcal{N}(u^{\mathbb{C}}))}{2\pi}. \tag{4.3}$$
For each toral eigenfunction u, denote
$$m_1 = \min_{a_k \ne 0}\{k_1\} \quad \text{and} \quad m_2 = \min_{a_k \ne 0}\{k_2\}.$$

[i] J. Bourgain and Z. Rudnick, *On the nodal sets of toral eigenfunctions.* Invent. Math. 185 (2011), no. 1, 199–237.

Then

$$P_u(z_1, z_2) = \frac{u^{\mathbb{C}}(z_1, z_2)}{z_1^{m_1} z_2^{m_2}} = \sum_{4\pi^2|k|^2=\lambda} a_k z_1^{k_1-m_1} z_2^{k_2-m_2}$$

is the unique polynomial determined by u which has the same nodal set as $u^{\mathbb{C}}$.

We know that $|k_1|, |k_2|, |m_1|, |m_2| \leq \lambda$ because $|k_1|^2 + |k_2|^2 = |k|^2 = \lambda^2$. Hence, the degree of $P_u(z_1, t)$ for fixed z_1 is bounded by $|k_2| + |m_2| \leq 2\lambda$ and same for $P_u(t, z_2)$ for fixed z_2.

By the fundamental theorem of algebra,

$$N_1 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(z_1, t) = 0\}\, dz_1 \leq 4\pi\lambda,$$

and

$$N_2 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(t, z_2) = 0\}\, dz_2 \leq 4\pi\lambda.$$

By the integral geometric formula,

$$\mathcal{H}^1(\mathcal{N}(P_u)) \leq N_1 + N_2 \leq 8\pi\lambda. \tag{4.4}$$

Therefore,

$$\mathcal{H}^1(\mathcal{N}(u)) = \frac{\mathcal{H}^1(\mathcal{N}(P_u))}{2\pi} \leq 4\lambda, \tag{4.5}$$

which is larger than our predicted value $2\sqrt{2}\lambda$ in Example (2). Why does it happen? We take a closer look at the examples:

Example.

(1). $u(x_1, x_2) = \cos(2\pi k_1 x_1)$ with $k_1 = \lambda$. Then

$$u^{\mathbb{C}}(z_1, z_2) = \frac{1}{2i}\left(z_1^{k_1} + z_1^{-k_1}\right),$$

and

$$P_u(z_1, z_2) = \frac{1}{2i}\left(z_1^{2k_1} + 1\right).$$

Now

$$N_1 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(z_1, t) = 0\}\, dz_1 = 0,$$

and

$$N_2 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(t, z_2) = 0\}\, dz_2 = 4\pi k_1.$$

Then

$$\mathcal{H}^1(\mathcal{N}(P_u)) \leq N_1 + N_2 = 4\pi k_1.$$

In this case, the integral geometric formula (4.4) is an equation since there are $2k_1$ lines parallel to z_2-axis. We have that $\mathcal{H}^1(\mathcal{N}(u)) = \mathcal{H}^1(\mathcal{N}(P_u))/2\pi = 2k_1 = 2\lambda$.

(2). $u(x_1, x_2) = \cos(2\pi k_1 x_1)\cos(2\pi k_2 x_2)$ with $k_1 = k_2 = \sqrt{2}\lambda/2$. Then

$$u^{\mathbb{C}}(z_1, z_2) = \frac{1}{4}\left(z_1^{k_1} z_2^{k_2} + z_1^{k_1} z_2^{-k_2} + z_1^{-k_1} z_2^{k_2} + z_1^{-k_1} z_2^{-k_2}\right),$$

and

$$P_u = \frac{1}{4}\left(z_1^{2k_1} z_2^{2k_2} + z_1^{2k_1} + z_2^{2k_2} + 1\right) = \frac{1}{4}\left(z_1^{2k_1} + 1\right)\left(z_2^{2k_2} + 1\right).$$

Now

$$N_1 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(z_1, t) = 0\}\, dz_1 = 4k_2\pi,$$

and

$$N_2 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(t, z_2) = 0\}\, dz_2 = 4k_1\pi.$$

Then

$$\mathcal{H}^1(\mathcal{N}(P_u)) \leq N_1 + N_2 = 4\pi(k_1 + k_2).$$

In this case, the integral geometric formula (4.4) is an equation since there are $2k_1$ lines parallel to z_2-axis and $2k_2$ lines parallel to z_1-axis. We have that $\mathcal{H}^1(\mathcal{N}(u)) = \mathcal{H}^1(\mathcal{N}(P_u))/2\pi = 2(k_1 + k_2) = \sqrt{2}\lambda$.

(3). $u(x_1, x_2) = \cos(2\pi k x_1) + \cos(2\pi k x_2)$ with $k = \lambda$. Then

$$u^{\mathbb{C}}(z_1, z_2) = \frac{1}{2}\left(z_1^k + z_1^{-k} + z_2^k + z_2^{-k}\right),$$

and

$$P_u(z_1, z_2) = \frac{1}{2}\left(z_1^{2k} z_2^k + z_2^k + z_1^k z_2^{2k} + z_1^k\right).$$

Now

$$N_1 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(z_1, t) = 0\}\, dz_1 = 4k\pi,$$

and

$$N_2 = \int_{\mathbb{S}^1} \#\{t \in \mathbb{S}^1 : P_u(t, z_2) = 0\}\, dz_2 = 4k\pi.$$

Then

$$\mathcal{H}^1(\mathcal{N}(P_u)) \leq N_1 + N_2 = 8k\pi \quad \text{and} \quad \mathcal{H}^1(\mathcal{N}(u)) \leq 4k = 4\lambda.$$

However, $\mathcal{N}(u)$ consists of $2k$ straight lines and each with length $\sqrt{2}$ (the k translations of the two hypotenuse in the square representation of the torus.) So $\mathcal{H}^1(\mathcal{N}(u)) = 2\sqrt{2}k$. in this case, the integral geometric formula (4.5) is not sharp.

Essentially, in the integral geometric formula, one chooses the two families of straight lines (Family 1 parallel to x_1-axis and Family 2 parallel to x_2-axis) to measure the nodal size by testing the intersections.

- When $u(x_1, x_2) = \cos(2\pi k_1 x_1)\cos(2\pi k_2 x_2)$, the two families of testing lines in the integral geometric formula are "doing the work individually". That is, Family 1 "sees" the nodal lines parallel to x_2-axis (so there are intersections with lines from Family 1); while Family 2 "sees" the nodal lines parallel to x_1-axis. So the application of the integral geometric formula is sharp.
- When $u(x_1, x_2) = \cos(2\pi k_1 x_1) + \cos(2\pi k_2 x_2)$, the two families of testing lines in the integral geometric formula are "doing the work together". That is, Families 1 and 2 both "see" the nodal lines which are parallel to the hypotenuse. The infinitesimal nodal length

$$dl^2 = dx_1^2 + dx_2^2.$$

  While the integral geometric formula uses the triangle inequality $dl \leq dx_1 + dx_2$, which precisely accounts for the loss of $\sqrt{2}$.

One potential improvement is to use different integral geometric formula. That is, one can design different families of 1-dim objects to test the intersection, and try to avoid the above loss. Take $u(x_1, x_2) = \cos(2\pi k_1 x_1) + \cos(2\pi k_2 x_2)$ for example. The integral geometric formula with the lines parallel to axis leads to a loss. While using lines parallel to two hypotenuse, the application of the formula becomes sharp, for the same reason as before. However, it seems difficult to design integral geometric formula that is sharp to estimate all nodal curves. So it is challenging here.

CHAPTER 5

Multiscale structure of solutions to Schrödinger equations

Definition (Vanishing orders). Let $u \in C^\infty(B(x,r)) \subset \mathbb{R}^n$. The vanishing order of u at x, denoted by $V_u(x)$, is the smallest integer d such that $\partial^\alpha u(x) = 0$ for all $|\alpha| < d$. (If $\partial^\alpha u(x) = 0$ for all α, then we set $V_u(x) = \infty$; if $u(x) \neq 0$, then we set $V_u(x) = 0$.)

One can in fact define the vanishing order of u at $x \in \Omega$ as long as $u \in L^2_{\mathrm{loc}}(\Omega)$, that is,

$$V_u(x) = \sup_{k \in \mathbb{R}} \left\{ \lim_{r \to 0} \frac{\|u\|_{L^2(B(x,r))}}{r^{k+n}} = 0 \right\}.$$

Noticing the normalizing factor r^n in the denominator, this vanishing order therefore agrees with the original definition for smooth functions.

Proof. If u is C^{k+1} in a neighborhood of the line segment from x to $x+y$, then the Taylor expansion

$$u(x+y) = \sum_{|\alpha| \le k} \frac{\partial^\alpha u(x)}{\alpha!} y^\alpha + \frac{1}{k!} \sum_{|\alpha| = k+1} \left(\int_0^1 \partial^\alpha (x+ty)(1-t)^k \, dt \right) y^\alpha.$$

In particular,

$$|u(x+y) - P_k(y)| = o\left(|y|^k\right) \quad \text{as } y \to 0,$$

in which

$$P_k(y) = \sum_{|\alpha| \le k} \frac{\partial^\alpha u(x)}{\alpha!} y^\alpha$$

is a polynomial of degree k. See Section 1.1 in Hörmander[i].

Therefore, if d is the smallest integer such that $\partial^\alpha u(x) = 0$ for all $|\alpha| < d$, then $u(x+y) \approx P_k(y)$ when $y \to 0$, in which $P_k(y)$ is a non-zero homogeneous polynomial of degree d. In this case,

$$\|u\|_{L^2(B(x,r))} \approx \|P\|_{L^2(B(0,r))} \approx r^{d+n}.$$

□

Consider the Schrödinger equation

$$\Delta u = Vu \quad \text{with potential } V \in C^\infty(\Omega)$$

for a sufficiently large domain $\Omega \subset \mathbb{R}^n$. Then $u \in C^\infty(\Omega)$ and satisfies strong unique continuation property, i.e., the vanishing order of u is finite everywhere. In addition, the nodal set of u, denoted by $\mathcal{N}(u)$, has dimension at most $(n-1)$. See Hardt-Simon[ii].

Example.
(1). If $V = 0$, then u is harmonic.

[i]L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Second edition. Springer-Verlag, Berlin, 1990.

[ii]R. Hardt and L. Simon, *Nodal sets for solutions of elliptic equations.* J. Differential Geom. 30 (1989), no. 2, 505–522.

(2). If $V = -\lambda^2$, then u is a Laplacian eigenfunction with eigenvalue λ^2.

We view the harmonic functions and the Laplacian eigenfunctions in one family of solutions to Schrödinger equations. The functions in this family appear to have a multiscale structure that is related to the mass distribution in Chapter 2, the L^p norm estimates in Chapter 3, and the nodal geometry in Chapter 4. Such structure is probably most apparent through the study of the doubling index.

Definition (Doubling index).

- For $\overline{B}(x, 2r) \subset \Omega$, define the doubling index of u in $B(x, r)$ as
$$N_u(x, r) = \log_2 \left(\frac{\max_{\overline{B}(x,2r)} |u|}{\max_{\overline{B}(x,r)} |u|} \right).$$
- For a cube Q such that $3\sqrt{n}Q \subset \Omega$ (which is to ensure that $\overline{B}(x, 2\mathrm{diam}(Q)) \subset \Omega$ for all $x \in Q$), define the doubling index of u in Q as
$$N_u(Q) = \sup_{p \in Q, r \in (0, \mathrm{diam}(Q))} N_u(x, r).$$

Theorem.

(i).
$$\lim_{r \to 0} N_u(x, r) = V_u(x).$$
In particular, $N_u(x, r) \to 0$ as $r \to 0$ unless $u(x) = 0$, i.e., $x \in \mathcal{N}(u)$.

(ii). $N_u(Q)$ is monotone in Q, that is, if $Q_1 \subset Q_2$, then
$$N_u(Q_1) \leq N_u(Q_2).$$

PROOF.

(i). It can first be verified for homogeneous polynomials. Then smooth functions can be approximated by homogeneous polynomials as before.

(ii). It follows immediately from the definition.

□

Let $Q_0 = [0, 1]^n$ be the unit cube and $3\sqrt{n}Q_0 \subset \Omega$. Since $\mathcal{N}(u)$ is at most $(n-1)$-dim,
$$\lim_{r \to 0} N_u(x, r) = V_u(x) = 0$$
almost everywhere in Q_0. In addition, $N_u(x, r) \leq N_u(Q_0) < \infty$ for all $x \in Q_0$ and $r < \mathrm{diam}(Q_0)$. By the dominated convergence theorem, we have that
$$\lim_{r \to 0} \int_{Q_0} N_u(x, r)\, dx = \int_{Q_0} \lim_{r \to 0} N_u(x, r)\, dx = 0.$$
Set $M = M(n) > 0$ to be an absolute constant. We are concerned with the optimal (i.e., largest) scale $r_0 = r_0(u, M)$ such that
$$\int_{Q_0} N_u(x, r_0)\, dx \leq M.$$
It is more convenient to work with cubes since they divide into smaller cubes.

Definition (Average doubling index). Divide Q_0 into subcubes q of side-length $l(q) = r$ for some $0 < r \leq 1$. Define the average doubling index of u at the scale r by
$$A_u(r) := \frac{1}{\#\{q : l(q) = r\}} \sum_{l(q)=r} N_u(q) = r^n \sum_{l(q)=r} N_u(q).$$

It is straightforward to show that $A_u(r)$ is monotone in r: Suppose that $0 < r \leq R \leq 1$ and assume that r, R are dyadic, i.e., they are of the form 2^{-m} with $m \geq 0$. (The dyadic assumption is generally harmless in application.) Then

$$\begin{aligned} A_u(r) &= \frac{1}{\#\{q : l(q) = r\}} \sum_{l(q)=r} N_u(q) \\ &= r^n \sum_{l(Q)=R} \left(\sum_{q \subset Q, l(q)=r} N_u(q) \right) \\ &\leq r^n \sum_{l(Q)=R} \frac{|Q|}{|q|} \cdot N_u(Q) \\ &= R^n \sum_{l(Q)=R} N_u(Q) \\ &= A_u(R). \end{aligned}$$

Therefore, $A_u(r)$ is uniformly bounded by $N_u(Q_0)$ and converges to 0 as $r \to 0$ for the same reason as before. (Indeed, $N_u(Q) \to V_u(x)$ if the cubes $Q \to \{x\}$.) Then the question arises

Problem 5.1 (Optimal scale for the average doubling index to be bounded). *What is the optimal (i.e., largest) scale $r_0 = r_0(u, M)$ at which $A_u(r_0) \leq M$?*

We shall mention two examples of functions which guide our intuition toward the problems in this chapter.

Example.

(1). Let

$$u_p(x, y) = \prod_{j=1}^{m} (z - z_j)^{k_j},$$

in which $z = (x + iy) \in \mathbb{C}$ and $z_1, ..., z_m \in Q_0/2$. Then $V_{u_p}(z_j) = k_j$ and $V_{u_p}(z) = 0$ at other points. One can put many z_j's with large k_j's, so that

$$A_{u_p}(r) \to k_1 + \cdots + k_m$$

is arbitrarily large, as $r \to 0$. Notice that $k_1 + \cdots + k_m \approx N_{u_p}(Q_0)$, which controls $A_{u_p}(r)$ in this example.

(2). Let

$$u_e(x, y) = e^{ax} \sin(by),$$

for which $\Delta u_e = V u_e$ with $V = a^2 - b^2$. Then $V_{u_e}(x, y) = 1$ along the lines $y = j\pi/b$, $j = 0, ..., b/\pi$. Hence,

$$A_{u_e}(r) \to \frac{rb}{\pi} \quad \text{if } r < \frac{\pi}{b}$$

Notice that $N_{u_e}(Q_0) \approx a$, which together with $V = a^2 - b^2$, control b and thus $A_{u_e}(r)$.

The above two examples indicate the appropriate parameters to control $A_u(r)$ are $N_u(Q_0)$ and $\|V\|_{L^\infty}$. Accordingly, we define

$$D(L, N, r) := \sup_{\Delta u = Vu, \|V\|_{L^\infty(\Omega)} \leq L, N_u(Q_0) \leq N} A_u(r),$$

and the special case for harmonic functions:

$$D(N,r) := D(0,N,r) = \sup_{\Delta u=0, N_u(Q_0)\le N} A_u(r).$$

That is, $D(L,N,r)$ depends on the following three parameters.

(1). L quantifies how close u is to a harmonic function at the global scale. In particular, if $L=0$, then u is harmonic and $D(0,N,r)=D(N,r)$.

(2). $N = N_u(Q_0)$ is the "global" doubling index of u.

(3). r is the scale of the side-length of the subcubes.

It is obvious that $D(L,N,r)$ is monotone in L,N,r: $D(L_1,N_1,r_1)\le D(L_2,N_2,r_2)$ if $L_1\le L_2$, if $N_1\le N_2$, or if $r_1\le r_2$. We are concerned with

(A). the multiscale relation of $D(L,N,r)$ when the scale $r=r(L,N)$ varies,

(B). the optimal scale $r_0 = r_0(L,N,M)$ at which $D(L,N,r_0)\le M$ for an absolute constant M,

(C). the application of the multiscale results to the quantitative unique continuation and the nodal size estimate of u.

Example. The two examples of u_p and u_e in $\mathbb{R}^2$ generalize to $\mathbb{R}^n$, however, with much more diversity. For example,

(1). The harmonic function

$$u_p(x_1,x_2,x_3) = (x_1+ix_2)^k \quad \text{in } \mathbb{R}^3$$

vanishes to order k along the x_3-axis, while the harmonic function

$$u_p(x_1,x_2,x_3,x_4) = (x_1+ix_2)^{k_1}(x_3+ix_4)^{k_2} \quad \text{in } \mathbb{R}^4$$

vanishes to order k_1 on the hyperplane $\{x_3=x_4=0\}$ and to order k_2 on the hyperplane $\{x_1=x_2=0\}$.

(2). The function

$$u_e(x_1,x_2,x_3) = e^{ax_1}\sin(bx_2)\sin(cx_3) \quad \text{in } \mathbb{R}^3$$

solves the equation $\Delta u_e = (a^2-b^2-c^2)u_e$ and vanishes to order 1 on the hyperplanes $\{x_2 = j\pi/b\}$ and $\{x_3 = k\pi/c\}$, while the function

$$u_e(x_1,x_2,x_3) = e^{ax_1}e^{bx_2}\sin(cx_3) \quad \text{in } \mathbb{R}^3$$

solves the equation $\Delta u_e = (a^2+b^2-c^2)u_e$ and vanishes to order 1 on the hyperplanes $\{x_3 = k\pi/c\}$,

Therefore, the problems in dimensions higher than two are more challenging. While we mention the cases in higher dimensions, our primary focus is on dimension two.

5.1. Harmonic functions

We again consider the two examples of harmonic functions.

Example. Denote $x=(x_1,...,x_n)\in\mathbb{R}^n$. Let $u_p(x_1,...,x_n) = (x_1+ix_2)^k$ in $\mathbb{R}^n$. Then $\Delta u_p = 0$ and $N_{u_p}(Q_0)\approx k$. Furthermore,

$$N_{u_p}(x,r) = \log_2\left(\frac{\max_{\overline{B}(x,2r)}|u_p|}{\max_{\overline{B}(x,r)}|u_p|}\right) = \log_2\left(\frac{(|(x_1,x_2)|+2r)^k}{(|(x_1,x_2)|+r)^k}\right) \le \frac{ckr}{|(x_1+x_2)|+r}.$$

That is, the doubling index of u_p at $B(x,r)$ is determined by the distance between (x_1, x_2) and the origin on the x_1x_2-plane. Therefore, we sum up the doubling index of subcubes q according to the distance between its projection to the x_1x_2-plane and the origin:

$$A_{u_p}(r) = \frac{1}{\#\{q : l(1) = r\}} \sum_{l(q)=r} N_{u_p}(q) \lesssim r^n \sum_{j=1}^{r^{-1}} \frac{ckr}{jr + r} \cdot jr \cdot r^{-(n-2)} \lesssim r^2 k \approx r^2 N_{u_p}(Q_0).$$

Example. Let $u_e(x_1, ..., x_n) = e^{ax_1} \sin(bx_2)$ in $\mathbb{R}^n$ with $a, b > 0$. Then $N_{u_e}(Q_0) \approx a$ and $\Delta u_e = (a^2 - b^2)u_e = 0$ if $a = b$. Furthermore,

$$N_{u_e}(x,r) = \log_2 \left(\frac{\max_{\overline{B}(x,2r)} |u_e|}{\max_{\overline{B}(x,r)} |u_e|} \right) \leq \log_2 \left(\frac{e^{a(|x_1|+2r)}}{e^{a(|x_1|+r)}} \right) = \log_2 e \cdot ra.$$

That is, the doubling index is $\approx ra$ for all subcubes. Hence,

$$A_{u_e}(r) \approx ra \approx rN_{u_e}(Q_0).$$

It is plausible that the second example is the "worse" case in the sense that the average doubling index is largest.

Problem 5.2 (Average doubling index of harmonic functions). *There exists $C = C(n) > 0$ such that if we partition Q_0 into subcubes q of side-length r, then*

$$A_u(r) \leq CrN_u(Q_0)$$

for all harmonic functions u. As a consequence,

$$D(N, N^{-1}) \lesssim 1.$$

We now deduce a weaker result in 2-dim using a result of Roy-Fortin[i], which is based on an approach developed by Donnelly-Fefferman[ii]. Indeed, the following theorem is extracted from Lemma 4.2.3 in Roy-Fortin and Lemma 6.1 in Donnelly-Fefferman.

Theorem (Roy-Fortin, Donnelly-Fefferman). *In $\mathbb{R}^2$, there exist $C, M > 0$ such that if we partition Q_0 into subcubes q of side-length $r \leq N_u(Q)^{-1}$, then*

$$\#\{q : N_u(q) \geq M\} \leq CN_u(Q_0)^2$$

for all harmonic functions u.

Remark. The estimate on the number of "bad" cubes, i.e., the ones with doubling index not uniformly bounded, is only meaningful at scales $r \leq N^{-1}$. That is, in the case when $r \geq N^{-1}$, the total number of cubes is $\leq N^2$ so the above estimate is meaningless.

The key information of the above theorem is that as r gets smaller than $N_u(Q_0)^{-1}$, the total number of subcubes r^{-2} increases, however, the number of the ones with doubling index larger than M stays bounded by $CN_u(Q_0)^2$ which is independent of r. Thus,

$$\begin{aligned} A_u(r) &= \frac{1}{\#\{q : l(q) = r\}} \sum_{l(q)=r} N_u(q) \\ &\leq r^2 \left[M \cdot \#\{q : N_u(q) < M\} + N_u(Q_0) \cdot \#\{q : N_u(q) \geq M\} \right] \end{aligned}$$

---

[i]G. Roy-Fortin, *Nodal sets and growth exponents of Laplace eigenfunctions on surfaces.* Anal. PDE 8 (2015), no. 1, 223–255.

[ii]H. Donnelly and C. Fefferman, *Nodal sets for eigenfunctions of the Laplacian on surfaces.* J. Amer. Math. Soc. 3 (1990), no. 2, 333–353.

$$
\begin{aligned}
&\leq r^2\left[M \cdot r^{-2} + N_u(Q_0)\cdot CN_u(Q_0)^2\right]\\
&\leq M + Cr^2 N_u(Q_0)^3\\
&\lesssim Cr^2 N_u(Q_0)^3.
\end{aligned}
$$

In particular, if $r \leq N_u(Q_0)^{-3/2}$, then $A_u(r) \leq M$, which implies that

$$D\left(N, N^{-3/2}\right) \lesssim 1.$$

Therefore, Donnelly-Fefferman and Roy-Fortin do not provide an affirmative answer at the scale $N_u(Q_0)^{-1}$ as in Problem 5.2. The main issue is that their arguments do not give good control on the subcubes with "high" doubling index such as the ones with $N_u(Q_0)/2$.

On the other hand, Logunov[i] proved that (we state the result in $\mathbb{R}^n$)

Theorem (Logunov). *In $\mathbb{R}^n$, there exist $0 < a_0, c_0 < 1$ such that if we partition Q_0 into subcubes q of side-length $r < a_0$, then*

$$\#\left\{q : N_u(q) \geq c_0 N_u(Q_0)\right\} \leq \frac{1}{2} r^{-(n-1)}$$

for all harmonic functions u.

The key information of the above theorem is that as r gets smaller, the total number of subcubes r^{-n} increases, however, the number of the ones with doubling index larger than $c_0 M$ is bounded by $r^{-(n-1)/2}$, i.e., they are no more than half of the ones on a hyperplane. Thus,

$$
\begin{aligned}
& A_u(r)\\
= \ & \frac{1}{\#\{q : l(q) = r\}} \sum_{l(q)=r} N_u(q)\\
\leq \ & r^n\left[c_0 N_u(Q_0)\cdot \#\{q : N_u(q) < c_0 N_u(Q_0)\} + N_u(Q)\cdot \#\{q : N_u(q) \geq c_0 N_u(Q_0)\}\right]\\
\leq \ & r^n\left[c_0 N_u(Q_0)\cdot r^{-n} + N_u(Q_0)\cdot \frac{1}{2} r^{-(n-1)}\right]\\
\leq \ & \left(c_0 + \frac{1}{2} r\right) N_u(Q_0).
\end{aligned}
$$

This leads to

$$D(N, r) \leq \left(c_0 + \frac{1}{2} r\right) N \leq \tilde{c}_0 N \quad \text{in which } \tilde{c}_0 < 1,$$

by choosing r small. Therefore, Logunov's result implies that the average doubling index below a fixed scale is always strictly smaller than the global one. While this is a crucial observation, it does not give control on the subcubes with "low" doubling index such as the ones with $N_u(Q_0)/2$ (since, in practice, c_0 is close to 1). Hence, it does not directly answer Problem 5.2.

5.2. Multiscale relations

In this section, we derive some multiscale relations involving the doubling index. First, the results of Roy-Fortin and Donnelly-Fefferman actually stay valid in a more general setting:

Theorem (Roy-Fortin, Donnelly-Fefferman). *In $\mathbb{R}^2$, there exist $C, M > 0$ such that if we partition Q_0 into subcubes q of side-length $r \leq N_u(Q_0)^{-1}$, then*

$$\#\{q : N_u(q) \geq M\} \leq CN_u(Q_0)^2$$

---

[i]A. Logunov, *Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure.* Ann. of Math. (2) 187 (2018), no. 1, 221–239.

for all functions u such that $\Delta u = Vu$ with $\|V\|_{L^\infty} \le N_u(Q_0)$.

As a consequence, $A_u(r) \le Cr^2 N_u(Q_0)^3$ as argued in the previous section, and

$$D\left(N, N, N^{-3/2}\right) = \sup_{\Delta u = Vu, \|V\|_{L^\infty(\Omega)} \le N, N_u(Q_0) \le N} A_u(r) \lesssim 1.$$

Similar to the harmonic functions, one might expect that

Problem 5.3 (Average doubling index of solutions to Schrödinger equations). *In $\mathbb{R}^2$, there exists $C > 0$ such that if we partition Q_0 into subcubes q of side-length r, then*

$$A_u(r) \le CrN_u(Q_0)$$

for all functions u such that $\Delta u = Vu$ with $\|V\|_{L^\infty} \le N_u(Q_0)$. As a consequence,

$$D(N, N, N^{-1}) \lesssim 1.$$

On the other hand, recall Logunov that in $\mathbb{R}^2$, if we partition Q_0 into subcubes Q of side-length $R < a_0$, then

$$\#\{Q : N_u(Q) \ge c_0 N_u(Q_0)\} \le \frac{1}{2} R^{-1}.$$

We further partition each Q into R^2/r^2 subcubes q of side-length r. Dilate u in the cube Q to $\tilde{u}$ in the unit cube Q_0. Accordingly, $q \subset Q$ is dilated to $\tilde{q} \subset Q_0$ with side-length r/R. Observe that the doubling index stay unchanged in the dilation. We thus have that

$$\sum_{q \subset Q, l(q) = r} N_u(q) = \sum_{\tilde{q} \subset Q_0, l(\tilde{q}) = \frac{r}{R}} N_{\tilde{u}}(\tilde{q}) = \frac{R^2}{r^2} \cdot A_{\tilde{u}}\left(\frac{r}{R}\right).$$

Therefore,

$$\begin{aligned}
A_u(r) &= \frac{1}{\#\{l(q) = r\}} \sum_{l(q) = r} N_u(q) \\
&\le r^2 \cdot \sum_{Q \subset Q_0, l(Q) = R, N_u(Q) < c_0 N_u(Q_0)} \left(\sum_{q \subset Q, l(q) = r} N_u(q) \right) \\
&\quad + r^2 \cdot \sum_{Q \subset Q_0, l(Q) = R, N_u(Q) \ge c_0 N_u(Q)} \left(\sum_{q \subset Q, l(q) = r} N_u(q) \right) \\
&\le r^2 \left[R^{-2} \cdot \frac{R^2}{r^2} \cdot A_{\tilde{u}}\left(\frac{r}{R}\right) + \frac{1}{2} R^{-1} \cdot \frac{R^2}{r^2} \cdot A_{\tilde{u}}\left(\frac{r}{R}\right) \right] \\
&= A_{\tilde{u}}\left(\frac{r}{R}\right) + \frac{1}{2} R \cdot A_{\tilde{u}}\left(\frac{r}{R}\right).
\end{aligned}$$

As a consequence, if $0 < r \le R \le a_0$, then

$$D(N, r) \le D\left(c_0 N, \frac{r}{R}\right) + \frac{1}{2} R \cdot D\left(N, \frac{r}{R}\right).$$

5.3. Laplacian eigenfunctions

Consider $\Delta u = \lambda^2 u$ in compact region $\Omega \subset \mathbb{R}^n$ that contains Q_0. Normalize u such that $\|u\|_{L^\infty} = 1$. Then Donnelly-Fefferman[i] proved that $N_u(Q_0) \leq c\lambda$. Yau's nodal size conjecture predicts that the nodal size of u in Q_0 is λ. Roy-Fortin showed that in 2-dim, the upper bound of the nodal size in Yau's conjecture is equivalent to the following estimate of the average doubling index. (Notice that the lower bound of the nodal size is known, see Chapter 4.)

Problem 5.4 (Average doubling index of eigenfunctions). *In $\mathbb{R}^2$, there exists $C > 0$ such that if we partition Q_0 into subcubes q of side-length r, then*

$$A_u(r) \leq CrN_u(Q_0)$$

for all functions u such that $\Delta u = Vu$ with $\|V\|_{L^\infty} \leq N_u(Q_0)^2$. As a consequence,

$$D(N^2, N, N^{-1}) \lesssim 1.$$

Consider again the following example to see that how the above estimate of $D(N^2, N, N^{-1})$ is related to the nodal size estimate.

Example (Doubling index and nodal size). Let $u_e(x, y) = e^{ax}\sin(by)$ in $\mathbb{R}^2$. Then $N_{u_e}(Q_0) \approx a$ and as before,

$$A_{u_e}(r) \approx ra \approx rN_{u_e}(Q_0).$$

On the other hand, the nodal set $\mathcal{N}(u_e)$ of u_e is a union of b straight lines (from the sin factor of u). So the size of $\mathcal{N}(u)$,

$$\mathcal{H}^1(\mathcal{N}(u_e)) \approx b.$$

which, apriori, is not controlled by the doubling index $N_{u_e}(Q_0)$.

However, notice that $\Delta u_e = (a^2 - b^2)u_e = Vu_e$ with $V = a^2 - b^2$. Hence,

$$\|V\|_{L^\infty} \leq N_{u_e}(Q_0)^2 \quad \text{iff} \quad \left|a^2 - b^2\right| \lesssim a^2,$$

that is, $b \lesssim a$. In this case,

$$\mathcal{H}^1(\mathcal{N}(u_e)) \approx b \lesssim a \approx N_{u_e}(Q_0).$$

That is, for the nodal size of u to be bounded by the doubling index $N_{u_e}(Q_0)$, the potential V in $\Delta u_e = Vu_e$ must satisfy that $\|V\|_{L^\infty} \lesssim N_{u_e}(Q_0)^2$. Moreover, the bound N^2 of $\|V\|_{L^\infty}$ here is sharp for the nodal size of u_e to be bounded by N.

Of course, the function u such that $\Delta u = Vu$ with $\|V\|_{L^\infty} \leq N_u(Q_0)^2 = N^2$ no longer resembles a harmonic function at the macroscopic scale anymore. However, one can restrict u at the scale N^{-1} and rescale to scale 1, then the function solves $\Delta\tilde{u} = \tilde{V}\tilde{u}$ with $\|\tilde{V}\|_{L^\infty} \leq 1$, i.e., it behaves like a harmonic function at the scale N^{-1} so one can use Section 5.1. In addition, restricting u to the scale $N^{-\frac{1}{2}}$ and rescale to scale 1, then the function solves $\Delta\tilde{u} = \tilde{V}\tilde{u}$ with $\|\tilde{V}\|_{L^\infty} \leq N$ so we can use Section 5.2. These observations indicate a strong multiscale structure of solutions to Schrödinger equation, which is not known.

[i] H. Donnelly and C. Fefferman, *Nodal sets of eigenfunctions on Riemannian manifolds.* Invent. Math. 93 (1988), no. 1, 161–183.

5.4. Unique continuation

For simplicity and relevancy with our discussion in this chapter, we only consider the elliptic partial differential operators of second order in a domain $\Omega \subset \mathbb{R}^n$:

$$Lu(x) = \sum_{j,k=1}^{n} a_{jk}(x)\partial^2_{x_j x_k} u(x) + \sum_{j=1}^{n} b_j(x)\partial_{x_j} u(x) + c(x)u(x),$$

in which (a_{jk}) is real symmetric and positive definite.

Definition (Unique continuation).

- We say that L satisfies the strong unique continuation property if $Lu = 0$ and $V_u(x) = \infty$ at some $x \in \Omega$ imply that $u \equiv 0$ on Ω.
- We say that L satisfies the unique continuation property if $Lu = 0$ and $u = 0$ on some open $\Omega_1 \subset \Omega$ imply that $u \equiv 0$ on Ω.
- We say that L satisfies the weak unique continuation property if $Lu = 0$ and $\operatorname{supp} u \subset K \Subset \Omega$ imply that $u \equiv 0$ on Ω.

For L with analytic coefficients, Holmgren[i] proved that

Theorem (Holmgren). *Suppose that L has analytic coefficients. Then $Lu = f$ with f analytic implies that u is also analytic. In particular, L satisfies the strong unique continuation property.*

Remark. The Cauchy-Kowalevski theorem states that the analytic solution to $Lu = f$ with f analytic must be unique. However, it does not exclude the possibility that there might exist solutions which are not analytic. See Section 1.D in Folland[ii]. Together with Holmgren's uniqueness theorem, the solution is analytic and unique.

Carleman[iii] proved the first unique continuation result for smooth differential operators.

Theorem (Carleman). *The differential operator*

$$L = \Delta + V \quad \text{with } V \in L^\infty_{\text{loc}}(\mathbb{R}^2)$$

satisfies the strong unique continuation property.

Carleman's approach involves an estimate which became a standard tool to establish unique continuation properties of solutions to partial differential equations. Generally speaking, a Carleman estimate is of the form

$$\|(\text{lower order derivatives of } u)\, e^{\tau\varphi}\|_* \le \frac{C}{\tau^*} \|(\text{highest order derivatives of } u)\, e^{\tau\varphi}\|_* \,.$$

By choosing the "weight" function φ, the norm $\|\cdot\|_*$, as well as the parameter $\tau \to \infty$, the Carleman estimates can be incredibly diverse. Its common theme is to control the weighted norm of lower order derivatives by the one of the highest order, c.f., Kenig[iv].

---

[i] E. Holmgren, *Über Systeme von linearen partiellen Differentialgleichungen*, Öfversigt af Kongl. Vetenskaps-Academien Förhandlinger, 58 (1901), 91–103.

[ii] G. Folland, *Introduction to partial differential equations.* Second edition. Princeton University Press, Princeton, NJ, 1995.

[iii] T. Carleman, *Sur un problème d'unicité pur les systèmes d'équations aux dérivées partielles à deux variables indépendantes.* Ark. Mat., Astr. Fys. 26, (1939). no. 17, 9 pp.

[iv] C. Kenig, *Carleman estimates, uniform Sobolev inequalities for second-order differential operators, and unique continuation theorems.* Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, Calif., 1986), 948–960, Amer. Math. Soc., Providence, RI, 1987.

Along the same line as the original result of Carleman, Jerison-Kenig[i] proved that

Theorem (Jerison-Kenig). *Let $n \geq 3$. Then the differential operator*

$$L = \Delta + V \quad \text{with } V \in L^{\frac{n}{2}}_{\text{loc}}(\mathbb{R}^n)$$

satisfies the strong unique continuation property.

Example. One can show that the exponent $n/2$ in the above theorem is sharp. Indeed, consider

$$u(x) = e^{\left(-\frac{1}{\log|x|}\right)^{1+\varepsilon}} \quad \text{with } \varepsilon > 0.$$

Then $V_u(0) = \infty$ but $u \not\equiv 0$ in $\mathbb{R}^n$. Furthermore, u satisfies that $(\Delta + V)u = 0$, in which

$$V(x) = -\frac{\Delta u(x)}{u(x)} \approx \frac{\left(\frac{1}{\log|x|}\right)^{2\varepsilon}}{|x|^2} \in L^p(\mathbb{R}^n) \quad \text{if } p < \frac{n}{2}.$$

Hence, strong unique continuation fails in this case.

Jerison-Kenig's proof is based on a L^p-Carleman estimate: Let $\tau \in (0,\infty) \setminus \mathbb{N}$ and $n \geq 3$. Denote by $\delta = \delta(\tau) = \operatorname{dist}(\tau, \mathbb{N})$, i.e., the distance between τ and the nearest integer. Then there exists $C = C(\delta, n) > 0$ such that for all $f \in C_0^\infty(\mathbb{R}^n \setminus \{0\})$, we have that

$$\left\| \frac{f(x)}{|x|^\tau} \right\|_{L^{p'}(\mathbb{R}^n, dx/|x|^n)} \leq C \left\| \frac{\Delta f(x)}{|x|^{\tau-2}} \right\|_{L^p(\mathbb{R}^n, dx/|x|^n)},$$

in which

$$p = \frac{2n}{n+1} \quad \text{and} \quad \|f\|_{L^p(\mathbb{R}^n, dx/|x|^n)} = \left(\int_{\mathbb{R}^n} |f(x)|^p \frac{dx}{|x|^n} \right)^{\frac{1}{p}}.$$

On the other hand, Aronszajn-Krzywicki-Szarski[ii] proved that

Theorem (Aronszajn-Krzywicki-Szarski). *Suppose that $\Omega \ni 0$ and*

$$\left| \sum_{j,k=1}^n a_{jk}(x) \partial^2_{x_j x_k} u(x) \right| \leq \left| \sum_{j=1}^n b_j(x) \partial_{x_j} u(x) + c(x) u(x) \right|$$

in which a_{jk} is Lipschitz continuous on Ω for all $j, k = 1, ..., n$, and $b_1, ..., b_n, c \in L^\infty_{\text{loc}}(\Omega)$. Then $V_u(0) = \infty$ implies that $u \equiv 0$ on Ω.

Remark. The Lipschitz continuity condition in the above theorem is sharp. Indeed, Pliś[iii] constructed an example for which weak unique continuity fails when the coefficients a_{jk} are Hölder continuous with exponent $\gamma \in (0,1)$. More precisely, for any $\gamma \in (0,1)$, there is $u \not\equiv 0$ which vanishes in an open set and solves $Lu = 0$ in which $a_{jk} \in C^\gamma(\Omega)$.

Carleman estimates and their applications to the unique continuation problems have been extensively studied in the past century, see again Kenig for a review. However, the unique continuation questions in a quantitative form is much less known, for example, let $\Delta u = 0$ in a domain $\Omega \ni 0$. What is the smallest $V_u(0)$ such that u then must vanish identically on Ω?

---

[i] D. Jerison and C. Kenig, *Unique continuation and absence of positive eigenvalues for Schrödinger operators.* With an appendix by E. M. Stein. Ann. of Math. (2) 121 (1985), no. 3, 463–494.

[ii] N. Aronszajn, A. Krzywicki, and J. Szarski, *A unique continuation theorem for exterior differential forms on Riemannian manifolds.* Ark. Mat. 4 (1962), 417–453.

[iii] A. Pliś, *On non-uniqueness in Cauchy problem for an elliptic second order differential equation.* Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 11 (1963), 95–100.

Example. The harmonic function $(x_1 + ix_2)^k$ has arbitrarily large vanishing order at 0 (in fact, on the $(n-2)$-dim hyperplane $\{x_1 = x_2 = 0\}$) as $k \to \infty$.

Therefore, one needs to impose additional conditions, and a natural candidate is the doubling index as defined in the beginning of this chapter.

Theorem (Almgren). *Let $\Delta u = 0$ in $B(0,1) \subset \mathbb{R}^n$. Then for any $r \in (0,1)$, there exists $C = C(r) > 0$ such that*

$$V_u(x) \le C N_u(0,1) \quad \textit{for all } x \in B(0,r).$$

5.5. Carleman estimates

In this section, we visit some Carleman estimates.

5.5.1. Carleman estimates by Trèves. We follow Chapter VIII in Hörmander[i] to prove some (simplest) examples of Carleman estimates, which are originally due to Trèves[ii].

Theorem (A Carleman estimate in $\mathbb{R}$). *Let $C > 0$. Then for all $u \in C_0^\infty(\mathbb{R})$*

$$\int_{\mathbb{R}} |u(x)|^2 e^{\tau x^2}\, dx \le C \int_{\mathbb{R}} |u'(x)|^2 e^{\tau x^2}\, dx \quad \textit{iff} \quad 2C\tau \ge 1.$$

Therefore, if $\tau > 0$, then

$$\int_{\mathbb{R}} |u(x)|^2 e^{\tau x^2}\, dx \le \frac{1}{2\tau} \int_{\mathbb{R}} |u'(x)|^2 e^{\tau x^2}\, dx.$$

Proof.

(1). We first show that τ must be positive. Indeed, if $\tau < 0$, then take $u_\varepsilon(x) = w(\varepsilon x)$ in which $w \in C_0^\infty(\mathbb{R})$ and equals 1 in a neighborhood of 0. As $\varepsilon \to 0$,

$$\int_{\mathbb{R}} |u_\varepsilon(x)|^2 e^{\tau x^2}\, dx = \int_{\mathbb{R}} |w(\varepsilon x)|^2 e^{\tau x^2}\, dx \to \int_{\mathbb{R}} e^{\tau x^2}\, dx > 0,$$

while

$$\int_{\mathbb{R}} |u_\varepsilon'(x)|^2 e^{\tau x^2}\, dx = \varepsilon^2 \int_{\mathbb{R}} |w'(\varepsilon x)|^2 e^{\tau x^2}\, dx \to 0,$$

which lead to a contradiction to the estimate. Moreover, if $\tau = 0$, then as $\varepsilon \to 0$,

$$\int_{\mathbb{R}} |u_\varepsilon(x)|^2\, dx = \int_{\mathbb{R}} |w(\varepsilon x)|^2\, dx \approx \frac{1}{\varepsilon},$$

while

$$\int_{\mathbb{R}} |u_\varepsilon'(x)|^2\, dx = \varepsilon^2 \int_{\mathbb{R}} |w'(\varepsilon x)|^2\, dx \approx \varepsilon,$$

which also lead to contradiction.

(2). Let $\tau > 0$. Set

$$v(x) = e^{\frac{\tau x^2}{2}} u(x).$$

Then

$$\begin{aligned} &\int_{\mathbb{R}} |u'(x)|^2 e^{\tau x^2}\, dx \\ = &\int_{\mathbb{R}} \left| e^{-\frac{\tau x^2}{2}} v'(x) - \tau x e^{-\frac{\tau x^2}{2}} v(x) \right|^2 e^{\tau x^2}\, dx \end{aligned}$$

[i] L. Hörmander, *Linear partial differential operators.* Springer-Verlag, Berlin-New York, 1976.

[ii] F. Trèves, *Relations de domination entre opérateurs différentiels.* Acta Math. 101 (1959), 1–139.

$$\begin{aligned}
&= \int_{\mathbb{R}} \left|v'(x) - \tau x v(x)\right|^2 \, dx \\
&= \int_{\mathbb{R}} \left|v'(x) + \tau x v(x)\right|^2 \, dx - 4\tau \int_{\mathbb{R}} x v'(x) v(x) \, dx \\
&= \int_{\mathbb{R}} \left|v'(x) + \tau x v(x)\right|^2 \, dx + 2\tau \int_{\mathbb{R}} |v(x)|^2 \, dx \\
&\geq 2\tau \int_{\mathbb{R}} |u(x)|^2 e^{\tau x^2} \, dx.
\end{aligned}$$

Here, we use the fact that

$$\int_{\mathbb{R}} x v'(x) v(x) \, dx = -\int_{\mathbb{R}} v(x)(v(x) + x v'(x)) \, dx = -\int_{\mathbb{R}} |v(x)|^2 \, dx - \int_{\mathbb{R}} v(x) x v'(x) \, dx$$

so

$$\int_{\mathbb{R}} x v'(x) v(x) \, dx = -\frac{1}{2} \int_{\mathbb{R}} |v(x)|^2 \, dx.$$

Now

$$\int_{\mathbb{R}} |u(x)|^2 e^{\tau x^2} \, dx \leq \frac{1}{2\tau} \int_{\mathbb{R}} |u'(x)|^2 e^{\tau x^2} \, dx$$

as in the theorem.

(3). Finally we show that the constant $\frac{1}{2}$ is sharp in the theorem. Set

$$u(x) = e^{-\tau x^2} \quad \text{and} \quad v(x) = e^{\frac{\tau x^2}{2}} u(x) = e^{-\frac{\tau x^2}{2}}.$$

Then

$$v'(x) + \tau x v(x) = 0,$$

so the equality at the end of Step (2) is achieved by $u(x) = e^{-\tau x^2}$. Therefore, letting $u \in C_0^\infty(\mathbb{R})$ approach $e^{-\tau x^2}$, the equality can be asymptotically obtained.

□

Theorem (A Carleman estimate in $\mathbb{R}^2$). *Let $C > 0$. Then for all $u \in C_0^\infty(\mathbb{R}^2)$*

$$\int_{\mathbb{R}^2} |u|^2 e^{ax^2+2bxy+cy^2} \, dxdy \leq C \int_{\mathbb{R}^2} |\partial_x u + i\partial_y u|^2 e^{ax^2+2bxy+cy^2} \, dxdy$$

iff $2C(a+c) \geq 1$. Therefore, if $a + c > 0$, then

$$\int_{\mathbb{R}^2} |u|^2 e^{ax^2+2bxy+cy^2} \, dxdy \leq \frac{1}{2(a+c)} \int_{\mathbb{R}^2} |\partial_x u + i\partial_y u|^2 e^{ax^2+2bxy+cy^2} \, dxdy.$$

Proof.

(1). We first prove the theorem assuming that $a = c$ and $b = 0$:

$$\int_{\mathbb{R}^2} |u|^2 e^{a(x^2+y^2)} \, dxdy \leq \frac{1}{4a} \int_{\mathbb{R}^2} |\partial_x u + i\partial_y u|^2 e^{a(x^2+y^2)} \, dxdy \quad \text{for all } u \in C_0^\infty(\mathbb{R}^2).$$

Similarly as in the 1-dim case, we must have that $a > 0$. Set

$$v(x,y) = e^{\frac{a(x^2+y^2)}{2}} u(x,y).$$

Then

$$\int_{\mathbb{R}^2} |\partial_x u + i\partial_y u|^2 e^{a(x^2+y^2)} \, dxdy$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^2} |\partial_x v - axv + i(\partial_y v - ayv)|^2 \, dxdy \\
&= \int_{\mathbb{R}^2} |\partial_x v + axv + i(-\partial_y v - ayv)|^2 \, dxdy - 4a \int_{\mathbb{R}^2} (\partial_x v \cdot xv + \partial_y v \cdot yv) \, dxdy \\
&= \int_{\mathbb{R}^2} |\partial_x v + axv + i(-\partial_y v - ayv)|^2 \, dxdy + 4a \int_{\mathbb{R}^2} |v|^2 \, dxdy \\
&\geq 4a \int_{\mathbb{R}^2} |u|^2 e^{a(x^2+y^2)} \, dxdy.
\end{aligned}
$$

Here, we use the fact that

$$\int_{\mathbb{R}^2} \partial_x v \cdot xv \, dxdy = \int_{\mathbb{R}^2} \partial_y v \cdot yv \, dxdy = -\frac{1}{2} \int_{\mathbb{R}} |v|^2 \, dxdy,$$

following integration by parts. Now,

$$\int_{\mathbb{R}^2} |u|^2 e^{a(x^2+y^2)} \, dxdy \leq \frac{1}{4a} \int_{\mathbb{R}^2} |\partial_x u + i\partial_y u|^2 e^{a(x^2+y^2)} \, dxdy$$

as in the theorem.

(2). We show that the constant $\frac{1}{2}$ is sharp in the theorem for the special case of Step (1). Set

$$u(x,y) = e^{-a(x^2+y^2)} \quad \text{and} \quad v(x,y) = e^{\frac{a(x^2+y^2)}{2}} u(x,y) = e^{-\frac{a(x^2+y^2)}{2}}.$$

Then

$$\partial_x v + axv + i(-\partial_y v - ayv) = 0,$$

so the equality at the end of Step (1) is achieved by $u(x) = e^{-a(x^2+y^2)}$. Therefore, letting $u \in C_0^\infty(\mathbb{R}^2)$ approach $e^{-a(x^2+y^2)}$, the equality can be asymptotically obtained.

(3). We finally reduce the general cases to the one in Step (1) by change of variables: Set

$$\alpha = \frac{a+c}{2} \quad \text{and} \quad \beta = \frac{a-c}{2}.$$

Let

$$v(x,y) = e^{\frac{(\beta - ib)(x+iy)^2}{2}} u(x,y).$$

Then

$$
\begin{aligned}
&\int_{\mathbb{R}^2} |u|^2 e^{ax^2+2bxy+cy^2} \, dxdy \\
&= \int_{\mathbb{R}^2} |v|^2 \left| e^{-(\beta - ib)(x+iy)^2} \right| e^{ax^2+2bxy+cy^2} \, dxdy \\
&= \int_{\mathbb{R}^2} |v|^2 e^{-\beta(x^2-y^2)-2bxy} e^{ax^2+2bxy+cy^2} \, dxdy \\
&= \int_{\mathbb{R}^2} |v|^2 e^{\alpha(x^2+y^2)} \, dxdy
\end{aligned}
$$

and

$$\int_{\mathbb{R}^2} |\partial_x u + i\partial_y u|^2 e^{ax^2+2bxy+cy^2} \, dxdy = \int_{\mathbb{R}^2} |\partial_x v + i\partial_y v|^2 e^{\alpha(x^2+y^2)} \, dxdy.$$

□

Together with the Carleman estimate in $\mathbb{R}$, we extend the results to $\mathbb{R}^n$:

Theorem (A Carleman estimate in $\mathbb{R}^n$). *Let $A = (a_{jk}) \in \mathrm{M}_n(\mathbb{R})$, $b = (b_1, ..., b_n) \in \mathbb{C}^n$, and $C > 0$. Write*

$$A(x) = \sum_{j,k=1}^{n} a_{jk} x_j x_k \quad \text{for } (x_1, ..., x_n) \in \mathbb{R}^n.$$

Then for all $u \in C_0^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} |u(x)|^2 e^{A(x)}\, dx \le C \int_{\mathbb{R}^n} \left| \sum_{l=1}^{n} b_l \partial_{x_l} u \right|^2 e^{A(x)}\, dx \quad \text{iff} \quad 2C \sum_{j,k=1}^{n} a_{jk} b_j \overline{b_k} \ge 1.$$

Proof. In the view of the invariance of the result, one can assume either $b = (1, 0, ..., 0)$, for which case it is reduced to the Carleman estimate in $\mathbb{R}$, or $b = (1, i, 0, ..., 0)$, for which case it is reduced to the one in $\mathbb{R}^2$. $\square$

5.5.2. A Carleman estimate by Donnelly-Fefferman. We provide a Carleman estimate in $\mathbb{R}^2$ following Proposition 2.1 in Donnelly-Fefferman, which seems powerful enough to imply certain quantitative results on unique continuation:

Identify $(x, y) \in \mathbb{R}^2$ by $z = x + iy \in \mathbb{C}$. Define

$$\partial_z := \frac{1}{2} (\partial_x - i\partial_y) \quad \text{and} \quad \partial_{\overline{z}} := \frac{1}{2} (\partial_x + i\partial_y).$$

Let $\Omega \subset \mathbb{R}^2$ be an open and bounded domain. Then for $u \in C^\infty(\Omega)$, we have that

$$\partial_z \partial_{\overline{z}} u = \partial_{\overline{z}} \partial_z u = \frac{1}{4} \Delta u.$$

Consider the Hilbert space $\mathcal{H}_\varphi(\Omega)$ equipped with norm

$$\|u\|_{\mathcal{H}_\varphi(\Omega)} = \int_\Omega |u(z)|^2 e^{-\varphi}\, dz.$$

Then for $u, v \in C_0^\infty(\Omega)$, we have that

$$\begin{aligned} \langle u, \partial_{\overline{z}} v \rangle_{\mathcal{H}_\varphi(\Omega)} &= \frac{1}{2} \int_\Omega u \cdot \left(\overline{\partial_x v + i\partial_y v} \right) e^{-\varphi}\, dz \\ &= -\frac{1}{2} \int_\Omega (\partial_x - i\partial_y)(u e^{-\varphi}) \cdot \overline{v}\, dz \\ &= -\int_\Omega e^\varphi \partial_z (u e^{-\varphi}) \cdot \overline{v} e^{-\varphi}\, dz. \end{aligned}$$

Hence, the adjoint of $\partial_{\overline{z}}$ in $\mathcal{H}_\varphi(\Omega)$ is

$$\partial_{\overline{z}}^\star u = -e^\varphi \partial_z (u e^{-\varphi}) = -\partial_z u + u \partial_z \varphi.$$

Notice that

$$\begin{aligned} [\partial_{\overline{z}}, \partial_{\overline{z}}^\star] u &= \partial_{\overline{z}} (-\partial_z u + u\partial_z \varphi) - (-\partial_z \partial_{\overline{z}} u + \partial_{\overline{z}} u \partial_z \varphi) \\ &= \partial_{\overline{z}} (u \partial_z \varphi) - \partial_{\overline{z}} u \partial_z \varphi \\ &= u \partial_{\overline{z}} \partial_z \varphi \\ &= \left(\frac{1}{4} \Delta \varphi \right) u. \end{aligned}$$

Therefore,

$$\|\partial_{\overline{z}}^\star u\|^2_{\mathcal{H}_\varphi(\Omega)} = \langle \partial_{\overline{z}}^\star u, \partial_{\overline{z}}^\star u \rangle_{\mathcal{H}_\varphi(\Omega)}$$

$$\begin{aligned}
&= \langle \partial_{\bar z}\partial_{\bar z}^{\star} u, u\rangle_{\mathcal{H}_\varphi(\Omega)} \\
&= \langle \partial_{\bar z}^{\star}\partial_{\bar z} u, u\rangle_{\mathcal{H}_\varphi(\Omega)} + \langle [\partial_{\bar z}, \partial_{\bar z}^{\star}] u, u\rangle_{\mathcal{H}_\varphi(\Omega)} \\
&= \|\partial_{\bar z} u\|^2_{\mathcal{H}_\varphi(\Omega)} + \frac{1}{4}\langle \Delta\varphi u, u\rangle_{\mathcal{H}_\varphi(\Omega)},
\end{aligned}$$

which implies that

$$-\frac{1}{4}\langle \Delta\varphi u, u\rangle_{\mathcal{H}_\varphi(\Omega)} = \|\partial_{\bar z} u\|^2_{\mathcal{H}_\varphi(\Omega)} - \|\partial_{\bar z}^{\star} u\|^2_{\mathcal{H}_\varphi(\Omega)} \le \|\partial_{\bar z} u\|^2_{\mathcal{H}_\varphi(\Omega)}.$$

That is,

$$-\frac{1}{4}\int_\Omega \Delta\varphi |u|^2 e^{-\varphi} \le \int_\Omega |\partial_{\bar z} u|^2 e^{-\varphi}.$$

Setting

$$\Phi = e^{-\varphi},$$

we arrive at

Theorem (Donnelly-Fefferman). *Let $\Phi \in C^\infty(\Omega)$ be positive. Then for any $u \in C^\infty(\Omega)$,*

$$\frac{1}{4}\int_\Omega \Delta \log \Phi |u|^2 \Phi \, dxdy \le \int_\Omega |\partial_{\bar z} u|^2 \Phi \, dxdy.$$

As a consequence of this theorem when choosing $\varphi(x,y) = -(ax^2 + 2bxy + cy^2)$ so $\Phi(x,y) = e^{ax^2+2bxy+cy^2}$, we have that (recalling that $\partial_{\bar z} = (\partial_x + i\partial_y)/2$)

$$\frac{1}{4}\int_\Omega (2a+2c)|u|^2 e^{ax^2+2bxy+cy^2}\, dxdy \le \frac{1}{4}\int_\Omega |\partial_x u + i\partial_y u|^2 e^{ax^2+2bxy+cy^2}\, dxdy,$$

which is exactly the one we proved in the previous theorem:

$$\int_\Omega |u|^2 e^{ax^2+2bxy+cy^2}\, dxdy \le \frac{1}{2(a+c)}\int_\Omega |\partial_x u + i\partial_y u|^2 e^{ax^2+2bxy+cy^2}\, dxdy.$$

CHAPTER 6

Fourier restriction problems on hyperbolic manifolds

Define the Fourier transform

$$\hat{u}(\xi) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-ix\cdot\xi} u(x)\, dx \quad \text{for } u \in C_0^\infty(\mathbb{R}^n).$$

The restriction problems ask whether one can restrict $\hat{u}$ to $\mathbb{S}^{n-1}$ (or more general lower dimensional geometric objects) in a meaningful way:

$$R(u) = \hat{u}|_{\mathbb{S}^{n-1}}.$$

In particular, if $u \in L^1(\mathbb{R}^n)$, then $\hat{u} \in C(\mathbb{R}^n)$ and the above restriction makes sense. While if $u \in L^2(\mathbb{R}^n)$, then $\hat{u} \in L^2(\mathbb{R}^n)$ and the above restriction does not make sense. The restriction problems then are about finding appropriate range of p ($1 \leq p < p_c < 2$) and $q = q(p, n)$ such that $R : L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$.

Conjecture (Stein).

$$R : L^p(\mathbb{R}^n) \to L^\infty\left(\mathbb{S}^{n-1}\right) \quad \textit{for all } 1 \leq p < \frac{2n}{n+1}.$$

More generally,

$$R : L^p(\mathbb{R}^n) \to L^q\left(\mathbb{S}^{n-1}\right) \quad \textit{for all } 1 \leq p < \frac{2n}{n+1} \textit{ and } q \leq \frac{n-1}{n+1}p'. \tag{6.1}$$

The restriction conjecture was proved in 2-dim by Fefferman[i] and Zygmund[ii] and remains open in higher dimensions. The restriction problems are closely related to other problems in analysis, PDEs, combinatorics, geometry, number theory, etc. For example, in $\mathbb{R}^n$, there is a well-known sequence of conjectures:

$$\text{local smoothing} \Rightarrow \text{Bochner-Riesz} \Rightarrow \text{restriction} \Rightarrow \text{Kakeya}.$$

Our main concern of restriction problems is on hyperbolic manifolds, and more generally, noncompact symmetric spaces on which Fourier transform can be defined.

(A). Define the restriction problems and prove or disprove them.
(B). Investigate the sequence of problems as above that are related to the Fourier restriction phenomenon.
(C). Characterize the geometric conditions of the hyperbolic manifolds that are behind the differences or similarities of the Fourier restriction problems with the ones in $\mathbb{R}^n$.

[i] C. Fefferman, *Inequalities for strongly singular convolution operators.* Acta Math. 124 (1970), 9–36.

[ii] A. Zygmund, *On Fourier coefficients and transforms of functions of two variables.* Studia Math. 50 (1974), 189–201.

6.1. L^2 Fourier restriction problems: Tomas-Stein estimates

The restriction problems when $q = 2$ in (6.1) were proved by Tomas[i] and Stein[ii]:

$$R : L^p(\mathbb{R}^n) \to L^2\left(\mathbb{S}^{n-1}\right) \quad \text{for all } 1 \le p \le p_c := \frac{2(n+1)}{n+3}.$$

In this case, we can use $T^\star T$ argument as in Section 10.21 to deduce that the above estimate is equivalent to

$$R^\star : L^2\left(\mathbb{S}^{n-1}\right) \to L^{p'}(\mathbb{R}^n) \quad \text{for all } \frac{2(n+1)}{n+1} \le p' \le \infty,$$

and

$$R^\star R : L^p(\mathbb{R}^n) \to L^{p'}(\mathbb{R}^n) \quad \text{for all } 1 \le p \le \frac{2(n+1)}{n+3},$$

in which the kernel of $R^\star R$ is

$$\int_{\mathbb{S}^{n-1}} e^{i(x-y)\cdot\xi}\, d\xi.$$

We first reduce the Tomas-Stein restriction estimates to a spectral measure statement that does not directly involve Fourier transform and its restriction. In this way, the restriction problems generalize to other geometries. To this end, one sees that by Plancherel theorem

$$\begin{aligned}
u(x) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \hat{u}(\xi) e^{ix\cdot\xi}\, d\xi \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi}\, d\xi \right) u(y)\, dy \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \left(\int_0^\infty \int_{\mathbb{S}^{n-1}_\lambda} e^{i(x-y)\cdot\xi}\, d\sigma_\xi d\lambda \right) u(y)\, dy \\
&= \frac{1}{(2\pi)^n} \int_0^\infty dE_{\sqrt{\Delta}}(\lambda) u(x)\, d\lambda.
\end{aligned}$$

Here, the spectral measure $dE_{\sqrt{\Delta}}(\lambda)$ has kernel

$$\int_{\mathbb{S}^{n-1}_\lambda} e^{i(x-y)\cdot\xi}\, d\sigma_\xi = \lambda^{n-1} \int_{\mathbb{S}^{n-1}} e^{i\lambda(x-y)\cdot\xi}\, d\sigma_\xi.$$

Hence, by dilation,

$$\begin{aligned}
dE_{\sqrt{\Delta}}(\lambda)u(x) &= \int_{\mathbb{R}^n} \left[\lambda^{n-1} \int_{\mathbb{S}^{n-1}} e^{i\lambda(x-y)\cdot\xi}\, d\sigma_\xi \right] u(y)\, dy \\
&= \lambda^{-1} \int_{\mathbb{R}^n} \left[\int_{\mathbb{S}^{n-1}} e^{i(\lambda x-y)\cdot\xi}\, d\sigma_\xi \right] u\left(\frac{y}{\lambda}\right)\, dy \\
&= \lambda^{-1} R^\star R(\tilde{u})(\lambda x),
\end{aligned}$$

in which $\tilde{u}(x) = u(x/\lambda)$. So

$$\begin{aligned}
\left\| dE_{\sqrt{\Delta}}(\lambda)u(x) \right\|_{L^{p'}(\mathbb{R}^n)} &= \lambda^{-1} \left\| R^\star R(\tilde{u})(\lambda x) \right\|_{L^{p'}(\mathbb{R}^n)} \\
&= \lambda^{-1-\frac{n}{p'}} \left\| R^\star R(\tilde{u})(x) \right\|_{L^{p'}(\mathbb{R}^n)} \\
&\le \lambda^{-1-\frac{n}{p'}} \left\| \tilde{u}(x) \right\|_{L^p(\mathbb{R}^n)}
\end{aligned}$$

[i] P. Tomas, *A restriction theorem for the Fourier transform*, Bull. Amer. Math. Soc. 81 (1975), 477–478.

[ii] E. Stein, *Oscillatory integrals in Fourier analysis.* Ann. of Math. Stud., 112, 307–355. Princeton Univ. Press, Princeton, NJ, 1986.

$$= \lambda^{-1-\frac{n}{p'}+\frac{n}{p}} \|u(x)\|_{L^p(\mathbb{R}^n)}$$
$$= \lambda^{n\left(\frac{1}{p}-\frac{1}{p'}\right)-1} \|u(x)\|_{L^p(\mathbb{R}^n)}.$$

The Tomas-Stein restriction estimates are now equivalent to proving

$$\left\|dE_{\sqrt{\Delta}}(\lambda)u\right\|_{L^{p'}(\mathbb{R}^n)} \leq \lambda^{n\left(\frac{1}{p}-\frac{1}{p'}\right)-1} \|u\|_{L^p(\mathbb{R}^n)}.$$

This abstract setting does not rely on the Fourier transform and has been studied in general geometries. Therefore, one can propose the Tomas-Stein restriction estimates in terms of the spectral measure:

Problem 6.1 (Tomas-Stein restriction estimates via the spectral measure). *Let* $\mathbb{M}$ *be a* d-dim *manifold. Is the following Tomas-Stein restriction estimate true for* $\lambda > 0$*?*

$$\left\|dE_{\sqrt{\Delta_{\mathbb{M}}}}(\lambda)\right\|_{L^p(\mathbb{M})\to L^{p'}(\mathbb{M})} \leq C\lambda^{d\left(\frac{1}{p}-\frac{1}{p'}\right)-1} \quad \textit{for } 1 \leq p \leq p_c = \frac{2(d+1)}{d+3}. \tag{6.2}$$

The parameter λ (i.e., energy) here is important since the dilation structure in $\mathbb{R}^n$ may not be available on the manifold.

Remark.

- If $\mathbb{M}$ is compact, then $dE_{\sqrt{\Delta_{\mathbb{M}}}}$ is formally a direct sum of delta measures at the discrete eigenvalues. Therefore, (6.2) fails and needs to be replaced by the spectral cluster estimate, that is, the estimate of "integral form" of $dE_{\sqrt{\Delta_{\mathbb{M}}}}$. The L^p estimates of eigenfunctions follow as a consequence, which are also heavily influenced by the dynamics of geodesics. See Chapter 3.
- Guillarmou-Hassell-Sikora[i] proved (6.2) on asymptotically conic manifolds with no closed geodesics.
- Huang-Sogge[ii] proved (6.2) on the hyperbolic space.
- Chen-Hassell[iii] proved (6.2) on asymptotically hyperbolic manifolds with no closed geodesics.

In addition to these positive results, Guillarmou-Hassell-Sikora also showed that (6.2) fails if there is an elliptic closed geodesic in the manifold. This is because of the existence of quasi-modes (i.e., approximate eigenfunctions) due to Babich–Lazutkin[iv] and Ralston[v] using the WKB method. So a natural question arises

Problem 6.2 (Tomas-Stein restriction estimates and hyperbolic dynamics). *Can* (6.2) *hold in the presence of non-elliptic closed geodesics?*

This is answered by Han[vi] in the case of convex cocompact hyperbolic manifolds with small set of closed geodesics.

---

[i]C. Guillarmou, A. Hassell, and A. Sikora, *Restriction and spectral multiplier theorems on asymptotically conic manifolds*. Anal. PDE 6, no. 4 (2013), 893–950.

[ii]S. Huang and C. Sogge. *Concerning L^p resolvent estimates for simply connected manifolds of constant curvature*. J. Funct. Anal. 267, no. 12 (2014), 4635–4666.

[iii]X. Chen and A. Hassell, *Resolvent and spectral measure on non-trapping asymptotically hyperbolic manifolds II: Spectral measure, restriction theorem, spectral multipliers*. Ann. Inst. Fourier (Grenoble) 68 (2018), no. 3, 1011–1075.

[iv]V. Babich and V. Lazutkin, *Eigenfunctions concentrated near a closed geodesic*, pp. 9–18 in Topics in Mathematical Physics, vol. 2, edited by M. S. Birman, Consultant's Bureau, New York, 1968.

[v]J. Ralston, *Approximate eigenfunctions of the Laplacian*, J. Differential Geometry 12:1 (1977), 87–100.

[vi]X. Han, *Tomas-Stein restriction estimates on convex cocompact hyperbolic manifolds*. Int. Math. Res. Not. IMRN 2021, no. 11, 8337–8352.

Definition (Convex cocompact hyperbolic manifolds). Let $\mathbb{H}^{n+1}$ be the $(n+1)$-dim hyperbolic space. Then the convex cocompact hyperbolic manifolds are $\mathbb{M} = \mathbb{H}^{n+1}/\Gamma$, in which Γ is a discrete group of orientation preserving isometries of $\mathbb{H}^{n+1}$ that consists of hyperbolic elements and $\mathbb{M}$ is geometrically finite and has infinite volume.

The set of closed geodesics in $\mathbb{M}$ corresponds to the conjugacy classes within the group Γ. The size of the geodesic trapped set is characterized by the limit set Λ_Γ of Γ. The limit set $\Lambda_\Gamma \subset \partial\mathbb{H}^{n+1}$ is the set of accumulation points on the orbits Γz, $z \in \mathbb{H}^{n+1}$. The Hausdorff dimension of Λ_Γ, $\delta_\Gamma := \dim \Lambda_\Gamma \in [0, n)$. Then the set of the closed geodesics in the unit tangent bundle $S\mathbb{M}$ has Hausdorff dimension $2\delta_\Gamma + 1$.

Roughly speaking, $\delta_\Gamma \to 0$ means $\mathbb{M}$ is more open, while as $\delta_\Gamma \to n$, the set of closed geodesics become larger. In the extreme case of $\delta_\Gamma = 0$, $\mathbb{M} = \mathbb{H}^{n+1}$ or hyperbolic cylinder, while $\delta_\Gamma = n$ implies that $\mathbb{M}$ is closed.

- Han proved (6.2) when $\delta_\Gamma < n/2$. The key ingredient of the proof is the Patterson[i]-Sullivan[ii] theory.
- When $\delta_\Gamma \geq n/2$, the necessary tools are not available in the Patterson-Sullivan theory and (6.2) remains unclear.
- The case when $\mathbb{M}$ has variable negative curvature and (small or large) set of closed geodesics is also unknown.

6.2. L^p Fourier restriction problems

The general L^p restriction problems in (6.1) are unclear on hyperbolic manifolds. We only take a look at them on the hyperbolic plane $\mathbb{H}^2$:

$$\mathbb{H}^2 = \{(x, y) \in \mathbb{R}^2 : y > 0\} \quad \text{equipped with metric } \frac{dx^2 + dy^2}{y^2}.$$

The Poincaré disk model is

$$\mathbb{B}^2 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\} \quad \text{equipped with metric } \frac{4\left(dx^2 + dy^2\right)}{\left(1 - x^2 - y^2\right)^2}.$$

The Fourier analysis, including the Fourier transform, the L^2 isometry (i.e., Plancherel theorem), etc, is established by Helgason[iii]. So one can ask the restriction problems in this setting:

Problem 6.3 (Fourier restriction problems on the hyperbolic plane).

$$R : L^p\left(\mathbb{B}^2\right) \to L^q\left(\mathbb{S}^1\right)?$$

6.3. Kakeya sets in the hyperbolic spaces

The Kakeya sets which depend on directions do not seem always generalizable to other manifolds. However, in the hyperbolic spaces, there is a natural generalization. We use the the Poincaré ball model

$$\mathbb{B}^n = \{x = (x_1, ..., x_n) \in \mathbb{R}^n : |x| < 1\} \quad \text{equipped with metric } \frac{4dx^2}{\left(1 - |x|^2\right)^2}.$$

Then every point on the boundary $\partial\mathbb{B}^n = \mathbb{S}^{n-1}$ can be understood as a direction at infinity. Therefore, we define

[i]S. Patterson, *The limit set of a Fuchsian group.* Acta Math. 136 (1976), no. 3–4, 241–273.

[ii]D. Sullivan, *The density at infinity of a discrete group of hyperbolic motions.* Inst. Hautes Études Sci. Publ. Math. No. 50 (1979), 171–202.

[iii]S. Helgason, *Topics in harmonic analysis on homogeneous spaces.* Birkhäuser, Boston, MA, 1981.

Definition (Kakeya sets in the hyperbolic spaces). A set $K \subset \mathbb{B}^n$ is said to be a Kakeya set if for each $\xi \in \mathbb{S}^{n-1}$, there is a unit-length segment of geodesic whose extension ends at ξ.

Then all the questions concerning the Kakeya sets in the Euclidean spaces can be asked here:

Problem 6.4 (Kakeya sets in the hyperbolic spaces). *Are there Kakeya sets in $\mathbb{B}^n$ which are of measure zero? What is the Hausdorff dimension of such sets?*

CHAPTER 7

Fractal uncertainty principle

Fractal uncertainty principle (FUP) was recently introduced by Dyatlov-Zahl[i] and has quickly become an emerging topic in Fourier analysis. It concerns the phenomenon that no function can be localized in both position and frequency close to a fractal set. More precisely, the FUP is formulated in the context of estimating the norm

$$\|\mathbb{1}_X \mathcal{F}_h \mathbb{1}_Y\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)}, \tag{7.1}$$

in which $0 < h < 1$ is the semiclassical parameter, the h-dependent sets $X = X(h), Y = Y(h) \subset \mathbb{R}^n$ are equipped with certain fractal-type structures, and $\mathcal{F}_h$ is the semiclassical Fourier transform

$$\mathcal{F}_h f(\xi) = \frac{1}{(2\pi h)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-\frac{ix\cdot\xi}{h}} f(x)\, dx.$$

Since $\mathcal{F}_h$ is unitary in $L^2(\mathbb{R}^n)$, we always have that (7.1) ≤ 1. We say that X, Y satisfy the FUP with exponent $\beta \geq 0$ if (7.1) $= O(h^\beta)$ as $h \to 0$. It can be understood as one quantitative version of the uncertainty principle that no function can be localized both near Y in the position space and near X in the frequency space. The main objective about the FUP is to prove the existence of exponent $\beta > 0$ and to find the sharp one for which the FUP holds. Here, "the sharp exponent", denoted by $\beta^s(X, Y)$, means the supreme of the exponents β such that (7.1) $= O(h^\beta)$ holds for all $0 < h < h_0$ with some $h_0 > 0$. It usually depends on the regularity and dimensions of X and Y.

See Dyatlov[ii] for a survey of the FUP and its application to the scattering theory in open chaotic systems and the spectral theory in closed chaotic systems, c.f., Chapter 2.

7.1. Discrete setting: Lower bounds

Dyatlov-Jin[iii] introduced a discrete model for the FUP and we briefly recall the setup: Let $N \in \mathbb{N}$ and l^2_N be the Hilbert space of functions $u : \mathbb{Z}_N = \mathbb{Z}/N \simeq \{0, ..., N-1\} \to \mathbb{C}$ with norm

$$\|u\|^2_{l^2_N} = \sum_{j=0}^{N-1} |u(j)|^2.$$

Define the (unitary) discrete Fourier transform $\mathcal{F} : l^2_N \to l^2_N$ as

$$\mathcal{F}_N u(j) = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} e^{-\frac{2\pi i j l}{N}} u(l). \tag{7.2}$$

---

[i]S. Dyatlov and J. Zahl, *Spectral gaps, additive energy, and a fractal uncertainty principle.* Geom. Funct. Anal. 26 (2016), no. 4, 1011–1094.

[ii]S. Dyatlov, *An introduction to fractal uncertainty principle.* J. Math. Phys. 60 (2019), no. 8, 081505, 31 pp.

[iii]S. Dyatloy and L. Jin, *Resonances for open quantum maps and a fractal uncertainty principle.* Comm. Math. Phys. 354 (2017), no. 1, 269–316.

See also Section 10.17 for the discrete Fourier analysis arsing from the representation theory of the finite group $\mathbb{Z}_n$.

We consider the FUP for the discrete Fourier transform and discrete Cantor sets. Each discrete Cantor set $\mathcal{C}_k(M, \mathcal{A})$ is determined by a base $\{0, ..., M-1\}$ with $M \in \mathbb{N}$ and $M \geq 3$, an alphabet $\mathcal{A} \subset \{0, ..., M-1\}$, and an order $k \in \mathbb{N}$:

$$\mathcal{C}_k(M, \mathcal{A}) = \left\{ \sum_{j=0}^{k-1} a_j M^j : a_0, \ldots, a_{k-1} \in \mathcal{A} \right\}. \tag{7.3}$$

It is evident that $\mathrm{Card}(\mathcal{C}_k(M, \mathcal{A})) = A^k$, where $A = \mathrm{Card}(\mathcal{A})$, the cardinality of $\mathcal{A}$.

Remark (Dimension of the Cantor sets). Denote the "dimension" of $\mathcal{C}_k(M, \mathcal{A})$ by

$$\delta = \frac{\log A}{\log M}. \tag{7.4}$$

Indeed, δ is the dimension of the Cantor set

$$\mathcal{C}_\infty = \bigcap_{k=0}^{\infty} \bigcup_{j \in \mathcal{C}_k} \left[\frac{j}{M^k}, \frac{j+1}{M^k} \right] \subset [0, 1].$$

For the discrete Cantor sets $\mathcal{C}_k = \mathcal{C}_k(M, \mathcal{A})$ and $N = M^k$ with $k \in \mathbb{N}$, we consider the FUP (7.1) in terms of the operator norm

$$r_k := \|\mathbb{1}_{\mathcal{C}_k} \mathcal{F}_N \mathbb{1}_{\mathcal{C}_k}\|_{l^2_N \to l^2_N}. \tag{7.5}$$

Since $\mathcal{F}_N$ is unitary, $\|\mathbb{1}_{\mathcal{C}_k} \mathcal{F}_N \mathbb{1}_{\mathcal{C}_k}\|_{l^2_N \to l^2_N} \leq 1$.

Remark. The FUP for discrete Cantor sets (7.5) in two extreme cases can be trivially derived:

- If $\delta = 0$ (i.e., $A = 1$), then $\mathrm{Card}(\mathcal{C}_k(M, \mathcal{A})) = 1$ so $r_k = N^{-\frac{1}{2}}$ for all $k \in \mathbb{N}$.
- If $\delta = 1$ (i.e., $A = M$), then $\mathcal{C}_k(M, \mathcal{A}) = \{0, ..., N-1\}$ and $\mathbb{1}_{\mathcal{C}_k} \mathcal{F}_N \mathbb{1}_{\mathcal{C}_k} = \mathcal{F}_N$ so $r_k = 1$ for all $k \in \mathbb{N}$.

In the cases when $0 < \delta < 1$ (i.e., $1 < A < M$), the mapping $\mathbb{1}_{\mathcal{C}_k} \mathcal{F}_N \mathbb{1}_{\mathcal{C}_k} : l^2_N \to l^2_N$ is given by a symmetric $N \times N$ matrix, which contains exactly A^k non-zero rows and columns (of the form $\frac{1}{\sqrt{N}} e^{-\frac{2\pi i j l}{N}}$). Hence,

$$r_k = \|\mathbb{1}_{\mathcal{C}_k} \mathcal{F}_N \mathbb{1}_{\mathcal{C}_k}\|_{l^2_N \to l^2_N} \leq \|\mathbb{1}_{\mathcal{C}_k} \mathcal{F}_N \mathbb{1}_{\mathcal{C}_k}\|_{\mathrm{HS}} = \sqrt{\frac{\mathrm{Card}(\mathcal{C}_k)^2}{N}} = M^{-k\left(\frac{1}{2} - \delta\right)} = N^{-\left(\frac{1}{2} - \delta\right)}, \tag{7.6}$$

in which $N = M^k \to \infty$ as $k \to \infty$ and $1/N \to 0$ plays the role of the semiclassical parameter h in (7.1). Here, $\|\cdot\|_{\mathrm{HS}}$ is the Hilbert-Schmidt norm.

From (7.6), an FUP with exponent $\beta = \frac{1}{2} - \delta > 0$ follows if $\delta < \frac{1}{2}$ (i.e., $A = M^\delta < M^{\frac{1}{2}}$). Such estimate is usually referred as "the volume bound" because it only takes the size of $\mathcal{C}_k$ into consideration. One then asks whether these bounds can be improved, that is,

$$r_k \leq N^{-\beta} \quad \text{for some } \beta > \max\left\{0, \frac{1}{2} - \delta\right\}.$$

Dyatlov-Jin proved that for any $\mathcal{A} \subset \{0, ..., M-1\}$ with $M \geq 3$ and $1 < \mathrm{Card}(\mathcal{A}) < M$, there exists

$$\beta = \beta(M, \mathcal{A}) > \max\left(0, \frac{1}{2} - \delta\right) \quad \text{such that} \quad r_k \leq N^{-\beta} \quad \text{for all } k \in \mathbb{N}.$$

In particular, we define the sharp exponent

$$\beta^s(M,\mathcal{A}) := \sup\left\{\beta \geq 0 : r_k \leq N^{-\beta} \text{ for all } k \in \mathbb{N}\right\}.$$

Then $\beta^s(M,\mathcal{A}) > 0$ for all discrete Cantor sets. However, the characterization of $\beta^s(M,\mathcal{A})$ and its dependence on M and $\mathcal{A}$ is a complicated story. Firstly,

Remark (The best possible exponent in the FUP for all discrete Cantor sets). Take $u(j) = 1$ for some $j \in \mathcal{C}_k$ and $u(j) = 0$ otherwise. Then

$$\|u\|_{l^2_N} = 1 \quad \text{and} \quad \|\mathbb{1}_{\mathcal{C}_k}\mathcal{F}_N\mathbb{1}_{\mathcal{C}_k}u\|_{l^2_N} = \sqrt{\frac{\operatorname{Card}(\mathcal{C}_k)}{N}} = \sqrt{\frac{A^k}{N}} = N^{-\frac{1-\delta}{2}},$$

in which $\operatorname{Card}(\mathcal{C}_k) = A^k = M^{\delta k} = N^{\delta}$. Hence, $r_k \geq N^{-\frac{1-\delta}{2}}$ and the best possible exponent β in the FUP is $\frac{1-\delta}{2}$, that is, for any M and $\mathcal{A}$,

$$\beta^s(M,\mathcal{A}) \leq \frac{1-\delta}{2}. \tag{7.7}$$

Dyatlov-Jin provided examples of discrete Cantor sets for which the FUP with the best possible exponent (7.7) holds. On the other hand, they also found examples for which

$$\beta^s(M,\mathcal{A}) \leq \frac{1}{2} - \delta + \frac{1}{KM\log M} \quad \text{and} \quad \left|\delta - \frac{1}{2}\right| \leq \frac{1}{K\log M}.$$

Here, $K > 0$ is an absolute constant. This sharp exponent improves over the volume bound (7.6) by a polynomial term in M and is much smaller than the best possible one (7.7). See Figure 7.1 for the numerical approximation of $\beta^s(M,\mathcal{A})$ for all alphabets $\mathcal{A}$ when $3 \leq M \leq 10$.

Despite the different estimates of $\beta^s(M,\mathcal{A})$ in the FUP for individual alphabets and bases, they observed that for fixed M and A, the expected $\beta^s(M,\mathcal{A})$ for all alphabets $\mathcal{A}$ with $\operatorname{Card}(\mathcal{A}) = A$ appears to be much larger than the volume bound (7.6), see the solid blue curve in Figure 7.1.

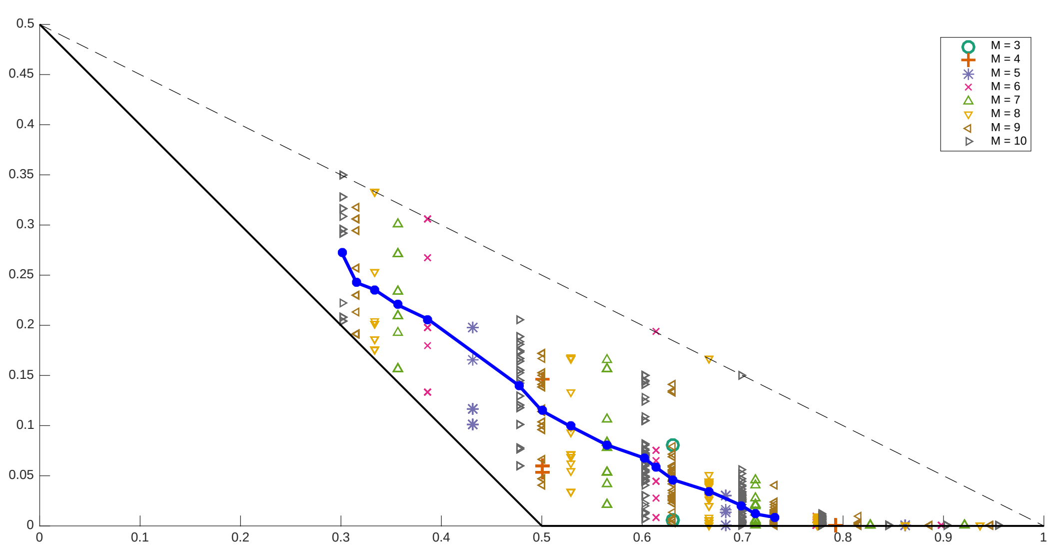

FIGURE 7.1. The solid black line is the volume bound $\frac{1}{2}-\delta$ (7.6); the dashed black line is the best possible exponent $\frac{1-\delta}{2}$ (7.7); the solid blue curve is the average of $\beta^s(M,A)$ over all alphabets with given $3 \leq M \leq 10$. (Credit: Dyatlov-Jin)

Eswarathasan-Han[i] verified Dyatlov-Jin's observation in the sense of the following probabilistic result. Write the space of alphabets as

$$\mathbb{A}(M, A) := \{\mathcal{A} \subset \{0, ..., M-1\} : \mathrm{Card}(\mathcal{A}) = A\},$$

which has cardinality

$$\mathrm{Card}(\mathbb{A}(M, A)) = \binom{M}{A}.$$

Then with respect to the uniform counting probability measure on $\mathbb{A}(M, A)$,

$$\beta^s(M, \mathcal{A}) \geq_\varepsilon \frac{1}{2} - \frac{3}{4}\delta - \varepsilon,$$

almost surely for $\mathcal{A} \in \mathbb{A}$. That is, with overwhelming probability, the $\beta^s(M, \mathcal{A})$ is much better than the volume bound (7.6). The line $\frac{1}{2} - \frac{3}{4}\delta$ agrees with Dyatlov-Jin's numerical result in Figure 7.1 perfectly.

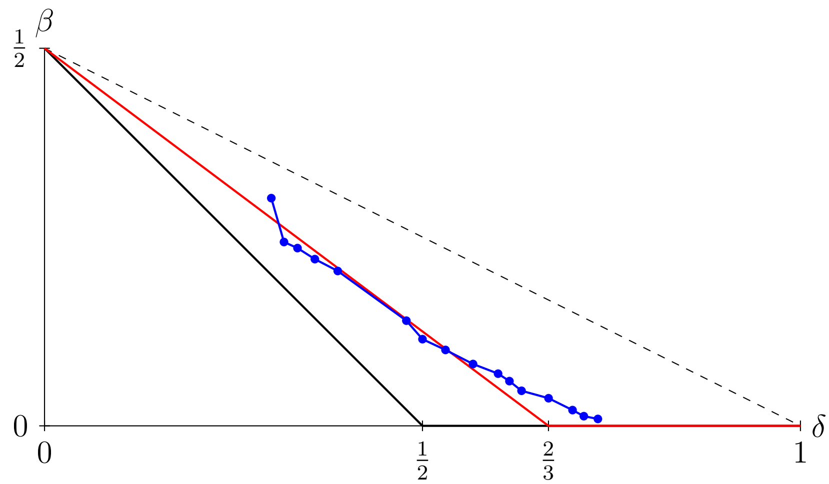

FIGURE 7.2. The solid red line $\frac{1}{2} - \frac{3}{4}\delta$ aligns with the numerical observation of Dyatlov-Jin.

However, the bound $\frac{1}{2} - \frac{3}{4}\delta$ is smaller than the best possible one in (7.7). It is interesting to ask

Problem 7.1 (FUP with the best possible exponent for random Cantor sets). *Can one find a random setting in which the FUP holds with best possible exponent* (7.7) *with overwhelming probability?*

Notice that in the randomization procedure of Eswarathasan-Han, one chooses an alphabet $\mathcal{A}$ randomly from $\mathbb{A}(M, A)$ (with respect to the uniform counting probability measure). Then the Cantor sets are constructed in each step $j \in \mathbb{N}$ using the same alphabet $\mathcal{A}$ of digits.

One can instead consider a different random ensemble:

$$\mathcal{C}_k(M, \mathcal{A}_0, ..., \mathcal{A}_{k-1}) = \left\{\sum_{j=0}^{k-1} a_j M^j : a_0 \in \mathcal{A}_0, a_1 \in \mathcal{A}_1, ..., a_{k-1} \in \mathcal{A}_{k-1}\right\},$$

[i] S. Eswarathasan and X. Han, *Fractal uncertainty principle for discrete Cantor sets with random alphabets.* arXiv:2107.08276.

in which $\mathcal{A}_j \in \mathbb{A}(M,A)$ for $j=0,...,k-1$ are identical and independent random variables. That is, the Cantor sets are constructed in each step $j \in \mathbb{N}$ using the random alphabet $\mathcal{A}_j$ of digits. One wonders if the answer to the above problem for these random Cantor sets is positive.

7.2. Discrete setting: Upper bounds

In this section, we investigate the upper bound of the exponent β for FUP to hold for a given Cantor set $\mathcal{C}_k$ with $\mathrm{Card}(\mathcal{C}_k) = N^\delta$ (random or deterministic):

$$r_k := \|\mathbb{1}_{\mathcal{C}_k}\mathcal{F}_N\mathbb{1}_{\mathcal{C}_k}\|_{l^2_N\to l^2_N} = O\left(N^{-\beta}\right).$$

Since

$$\|\mathbb{1}_{\mathcal{C}_k}\mathcal{F}_N\mathbb{1}_{\mathcal{C}_k}\|_{l^2_N\to l^2_N} = \sup_{u\in l^2_N}\frac{\|\mathbb{1}_{\mathcal{C}_k}\mathcal{F}_N\mathbb{1}_{\mathcal{C}_k}u\|_{l^2_N}}{\|u\|_{l^2_N}=1},$$

such investigation involves choosing appropriate vectors $u \in l^2_N$. Compute that

$$\begin{aligned}\|\mathcal{F}_N u\|^2_{l^2_N} &= \sum_{j\in\mathcal{C}_k}|\mathcal{F}_N u(j)|^2\\ &= \sum_{j\in\mathcal{C}_k}\frac{1}{N}\left(\sum_{m=0}^{N-1}e^{-\frac{2\pi i m j}{N}}u(m)\right)\left(\sum_{n=0}^{N-1}e^{\frac{2\pi i n j}{N}}\overline{u(n)}\right)\\ &= \frac{1}{N}\sum_{j\in\mathcal{C}_k}\sum_{m,n=0}^{N-1}e^{\frac{2\pi i(n-m)j}{N}}u(m)\overline{u(n)}.\end{aligned}$$

Also,

$$\begin{aligned}&\|\mathcal{F}_N u\|^4_{l^2_N}\\ =& \left(\frac{1}{N}\sum_{j_1\in\mathcal{C}_k}\sum_{m_1,n_1=0}^{N-1}e^{\frac{2\pi i(n_1-m_1)j_1}{N}}u(m_1)\overline{u(n_1)}\right)\left(\frac{1}{N}\sum_{j_2\in\mathcal{C}_k}\sum_{m_2,n_2=0}^{N-1}e^{\frac{2\pi i(n_2-m_2)j_2}{N}}u(m_2)\overline{u(n_2)}\right)\\ =& \frac{1}{N^2}\sum_{j_1,j_2\in\mathcal{C}_k}\sum_{m_1,n_1,m_2,n_2=0}^{N-1}e^{\frac{2\pi i}{N}[(n_1-m_1)j_1+(n_2-m_2)j_2]}u(m_1)\overline{u(n_1)}u(m_2)\overline{u(n_2)}\end{aligned}$$

Example. As shown in the previous section, take $u(j)=1$ for some $j\in\mathcal{C}_k$ and $u(j)=0$ otherwise. Then

$$\|u\|_{l^2_N}=1 \quad \text{and} \quad \|\mathbb{1}_{\mathcal{C}_k}\mathcal{F}_N\mathbb{1}_{\mathcal{C}_k}u\|_{l^2_N} = \sqrt{\frac{\mathrm{Card}(\mathcal{C}_k)}{N}} = N^{-\frac{1-\delta}{2}}.$$

Therefore,

$$\beta \le \frac{1-\delta}{2}.$$

Example. Take $u=1$ on $\mathcal{C}_k$. Then $\|u\|_{l^2_N} = \sqrt{\mathrm{Card}(\mathcal{C}_k)} = N^{\frac{\delta}{2}}$, and

$$\begin{aligned}\|\mathcal{F}_N u\|^2_{l^2_N} &= \sum_{j\in\mathcal{C}_k}|\mathcal{F}_N u(j)|^2\\ &= \sum_{j\in\mathcal{C}_k}\frac{1}{N}\left(\sum_{m\in\mathcal{C}_k}e^{-\frac{2\pi i m j}{N}}\right)\left(\sum_{n\in\mathcal{C}_k}e^{\frac{2\pi i n j}{N}}\right)\end{aligned}$$

$$= \frac{1}{N} \sum_{j,m,n \in \mathcal{C}_k} e^{\frac{2\pi i (n-m) j}{N}}.$$

Example. Take $u(m) = \binom{N}{m}$ for $m = 0, ..., N-1$. Then

$$\|u\|^2_{l^2_N} = \sum_{m=0}^{N-1} \binom{N}{m}^2 = \binom{2N}{N} - 1 \approx \frac{\sqrt{2\pi 2N}(2N)^{2N}/e^{2N}}{\sqrt{2\pi N} N^N/e^N} = \frac{2^{2N+\frac{1}{2}} N^N}{e^N}.$$

To see this, we use the identity that

$$\binom{N+M}{r} = \sum_{m=0}^{r} \binom{N}{m}\binom{M}{r-m}.$$

Setting $M = r = N$, we have that

$$\binom{2N}{N} = \sum_{m=0}^{N} \binom{N}{m}\binom{N}{N-m} = \sum_{m=0}^{N} \binom{N}{m}^2.$$

which implies the equation above. Now

$$\begin{aligned} \mathcal{F}_N u(j) &= \frac{1}{\sqrt{N}} \sum_{m=0}^{N-1} e^{-\frac{2\pi i m j}{N}} \binom{N}{m} \\ &= \frac{1}{\sqrt{N}} \left[\left(1 + e^{-\frac{2\pi i j}{N}}\right)^N - 1 \right]. \end{aligned}$$

7.3. Dilated Fourier transform

Let α be a real number in $[1, M]$. The dilated Fourier transform $\mathcal{F}_{N,\alpha} : l^2_N \to l^2_N$ by a factor of α is defined as

$$\mathcal{F}_{N,\alpha} u(j) = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} e^{-\frac{2\pi \alpha i j l}{N}} u(l). \tag{7.8}$$

Proposition. *There is an absolute constant $C > 0$ such that*

$$\|\mathcal{F}_{N,\alpha}\|_{l^2_N \to l^2_N} \le C\sqrt{\log N}.$$

Proof. The mapping $\mathcal{F}_{N,\alpha}$ is given by an $N \times N$ matrix, denoted by $\mathcal{M}$ with entries of the form $e^{-\frac{2\pi\alpha i j l}{N}}$. Let $\mathcal{M}^\star$ be the complex conjugate of $\mathcal{M}$. We estimate the largest (in module) eigenvalue r_1 of $\mathcal{M}$ via the ones of $\mathcal{M}\mathcal{M}^\star$, which has entries

$$\begin{aligned} F_{jk} &:= \frac{1}{N} \sum_{l \in A} e^{\frac{2\pi i (k-j) l}{N}} \\ &= \begin{cases} 1, & \text{if } j = k, \\ \frac{1}{N} \sum_{l=0}^{N-1} e^{\frac{2\pi i (k-j) l}{N}}, & \text{if } j \ne k. \end{cases} \end{aligned}$$

Notice that

$$\left| \frac{1}{N} \sum_{l=0}^{N-1} e^{\frac{2\pi i (k-j) l}{N}} \right| = \frac{1}{N} \left| \frac{2}{1 - e^{\frac{2\pi i (k-j)}{N}}} \right| \le C|k-j|.$$

Then Schur's lemma implies an upper bound of the eigenvalues of $\mathcal{M}(\mathcal{A})\mathcal{M}(\mathcal{A})^\star$:

$$\max_{j=0,\ldots,N-1}\left\{|F_{jj}|+\sum_{k\neq j}|F_{jk}|\right\}\leq 1+C\sum_{k=0}^{N-1}=C\log N.$$

Hence, $r_1\leq C\sqrt{\log N}$. □

CHAPTER 8

Kakeya problems

The following brief history of Kakeya set is from Section 7.1 in Falconer[i]. In 1917, Besicovitch was working on problems on Riemann integration, and was confronted with the following question:

Question 8.1. *If f is a Riemann integrable function defined on the plane, is it always possible to find a pair of orthogonal coordinate axes with respect to which $\int f(x,y)dx$ exists as a Riemann integral for all y, and with the resulting function of y also Riemann integrable?*

Answer. Besicovitch noticed that if he could construct a compact set B of Lebesgue measure zero containing a line segment in every direction, this would lead to a counterexample: Indeed, assume by translating B if necessary, that B contains no segment parallel to and has rational distance from either of a fixed pair of axes. Let f be the characteristic function of the set B_0 consisting of those points of B with at least one rational coordinate. As B contains a segment in every direction on which both B_0 and its complement are dense, there is a segment in each direction on which f is not Riemann integrable. On the other hand, the set of points of discontinuity of B is of plane measure zero, so f is Riemann integrable over the plane by the well-known criterion of Lebesgue.

In 1919, Besicovitch succeeded in constructing a set, known as a "Besicovitch set", with the required properties. Owing to the unstable situation in Russia at the time, his paper received limited circulation, and the construction was republished in 1928[ii].

At about the same time, Kakeya[iii] and Fujiwara-Kakeya[iv] mentioned the problem of finding the area of the smallest convex set inside which a unit segment can be reversed, that is, maneuvered to lie in its original position but rotated through 180° without leaving the set. They conjectured that the equilateral triangle of unit height was the smallest such set, and recorded an observation of Kubota, that if the convexity condition was dropped, then a smaller set was possible, namely a three-cusped hypercycloid. The conjecture in the convex case was proved by Pál[v] who reiterated the more interesting question without the convexity assumption. This became known as the Kakeya problem.

Shortly after Besicovitch's departure from Russia in 1924, it was realized that a simple modification to the Besicovitch set yielded a solution to the Kakeya problem of arbitrarily small measure by Besicovitch and Perron[vi]. – The Kakeya problem was solved in an unexpected manner.

---

[i] K. Falconer, *Fractal geometry.* Third edition. John Wiley & Sons Ltd., Chichester, 2014.

[ii] A. Besicovitch, *On Kakeya's problem and a similar one.* Math. Z. 27 (1928), no. 1, 312–320.

[iii] S. Kakeya, *Some problems on maximum and minimum regarding ovals.* Tohoku Science Reports. 6 (1917), 71–88.

[iv] M. Fujiwara and S. Kakeya, *On some problems of maxima and minima for the curve of constant breadth and the in-revolvable curve of the equilateral triangle*, Tôhoku Math. J. 11 (1917), 92–110.

[v] J. Pál, J. *Ein Minimumproblem für Ovale.* Math. Ann. 88 (1921), 311–319.

[vi] O. Perron, *Über einen Satz von Besicovitsch.* Math. Z. 28 (1928), no. 1, 383–386.

Definition (Kakeya/Besicovitch set). A Kakeya/Besicovitch set $K \subset \mathbb{R}^n$ is a set which contains a unit line segment in every direction, that is, for each $e \in \mathbb{S}^{n-1}$, there exists $x \in \mathbb{R}^n$ such that

$$x + te \in K \quad \text{for all } t \in \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Theorem (Besicovitch). *There are Kakeya sets with measure zero.*

Besicovitch's original construction has been simplified considerably by Perron, van Alphen, Rademacher, Schoenberg, Besicovitch, Cunningham, and Fisher. The basic idea behind all the constructions is to form a "Perron tree", a figure obtained by splitting an equilateral triangle of unit height into many smaller triangles of the same height by dividing up the base, and then sliding these elementary triangles varying distances along the base line L. Such a figure certainly contains line segments of unit length in all directions at angles at least $60°$ to L. See Section X.1 in Stein[i] for a detailed construction.

Various other methods of constructing Besicovitch sets have been found: Kahane noticed that such a set could be obtained by joining the points of a Cantor set in the x-axis to the points of a parallel Cantor-like set, see Section 7.4.4 in Stein-Shakarchi[ii].

Theorem. *For each $\varepsilon > 0$, there is a set $\tilde{K}$, with $|\tilde{K}| < \varepsilon$, inside of which a unit line segment can be moved continuously, changing its orientation by $180°$ while doing so.*

Proof. The proof is from Section X.3.1 in Stein. This can be deduced from the construction of Besicovitch and the following elementary observations. Suppose l_1, l_2 are two parallel unit line segments in the plane. Then for any $\varepsilon > 0$, there is a set S_ε, $|S_\varepsilon| < \varepsilon$, so that, within S_ε, l_1 can be continuously moved in to the position l_2. To see this, one needs merles move l_1 along its initial direction far enough, then return (at a slight angle) to the position of l_2. $\square$

Question 8.2. *Does there exist a set of measure zero within which a unit line segment can be continuously rotated by a full rotation?*

See Section 1.22 in Tao[iii] for more discussion.

Answer. No. Let $E \subset \mathbb{R}^2$ be a set in the plane within which a unit line segment can be continuously rotated. This means that there exists a continuous map $l : t \to l(t)$ from $t \in [0,1]$ to unit line segments $l(t) \subset E$. We can parameterise each such line segment as

$$l(t) = \{(x(t) + s\cos\omega(t), y(t) + s\sin\omega(t)) : -0.5 \leq s \leq 0.5\},$$

in which $x, y, \omega : [0,1] \to \mathbb{R}$ are continuous functions.

Recall that on a compact set, all continuous functions are uniformly continuous. In particular, there exists $\varepsilon > 0$ such that

$$|x(t) - x(t')|, \quad |y(t) - y(t')|, \quad |\omega(t) - \omega(t')| \leq 0.001$$

whenever $t, t' \in [0,1]$ are such that $|t - t'| \leq \varepsilon$.

Fix this ε. Observe that $\omega(t)$ cannot be a constant function of t, otherwise the needle would never rotate. We conclude that there must exist $t_0, t_1 \in [0,1]$ with $|t_0 - t_1| \leq \varepsilon$ and $\omega(t_0) \neq \omega(t_1)$. Without loss of generality, we may assume that $t_0 < t_1$ and $x(t_0) = y(t_0) = \omega(t_0) = 0$. Now let a be any real number between -0.4 and 0.4. We see that for any $t_0 \leq t \leq t_1$, the line $l(t)$

[i] E. Stein, *Harmonic analysis.* Princeton University Press, Princeton, NJ, 1993.

[ii] E. Stein and R. Shakarchi, *Real analysis. Measure theory, integration, and Hilbert spaces.* Princeton University Press, Princeton, NJ, 2005.

[iii] T. Tao, *Structure and randomness.* American Mathematical Society, Providence, RI, 2008.

intersects the line $x = a$ in some point $(a, ya(t))$, which must therefore lie in E. Furthermore, $ya(t)$ varies continuously in t. By the intermediate value theorem, we conclude that the interval between $(a, ya(t_0))$ and $(a, ya(t_1))$ lies in E. Taking unions over all a between -0.4 and 0.4, we see that E contains a non-trivial sector, and thus has non-zero area. The claim follows.

8.1. Kakeya sets

Since one can construct Kakeya set in $\mathbb{R}^n$ that are of measure zero, a natural question arises as how small exactly can they be? In particular, can they be of lower-dimensional? The Kakeya problems predict that they are essentially n-dim.

Conjecture (Kakeya problems).

(i). The Minkowski dimension of a Kakeya set $K \subset \mathbb{R}^n$, denoted by $\dim_{\mathrm{M}}(K)$, is n.

(ii). The Hausdorff dimension of a Kakeya set $K \subset \mathbb{R}^n$, denoted by $\dim_{\mathrm{H}}(K)$, is n.

(iii). The Kakeya maximal function for $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ satisfies that for any $\varepsilon > 0$,

$$\|f^*_\delta\|_{L^n(\mathbb{S}^{n-1})} \leq C\delta^{-\varepsilon}\|f\|_{L^n(\mathbb{R}^n)}$$

for some positive constant $C = C(\varepsilon)$.

Here, the Kakeya maximal function is defined as

$$f^*_\delta(e) = \sup_{a\in\mathbb{R}^n} \frac{1}{|T^\delta_e(a)|}\int_{T^\delta_e(a)} |f|,$$

Here, $T^\delta_e(a)$ is the tube centered at $a \in \mathbb{R}^n$ of width $\delta > 0$ in the direction $e \in \mathbb{S}^{n-1}$, that is,

$$T^\delta_e(a) = \{x \in \mathbb{R}^n : |(x-a)\cdot e| \leq \frac{1}{2}, |(x-a)^\perp| \leq \delta\},$$

in which $x^\perp = x - (x \cdot e)e$.

Remark. For any set Ω, we have that $\dim_{\mathrm{H}}(\Omega) \leq \dim_{\mathrm{M}}(\Omega)$, see Section 7.2 in Stein-Shakarchi. Therefore, (ii) implies (i).

We next show that (iii) implies (i) and begin from some observations.

(1). The following trivial bounds hold.

$$\|f^*_\delta\|_{L^\infty(\mathbb{S}^{n-1})} \leq \|f\|_{L^\infty(\mathbb{R}^n)}, \quad \|f^*_\delta\|_{L^\infty(\mathbb{S}^{n-1})} \leq \delta^{-(n-1)}\|f\|_{L^1(\mathbb{R}^n)}.$$

(2). If $p < \infty$, then there can be no bound of the form

$$\|f^*_\delta\|_{L^q(\mathbb{S}^{n-1})} \leq C\|f\|_{L^p(\mathbb{R}^n)},$$

in which C is independent of δ.

Indeed, consider a zero measure Kakeya set K. Let K^δ be the δ-neighborhood of K and $f = \chi_{K^\delta}$. Then $f^*_\delta = 1$ for all $e \in \mathbb{S}^{n-1}$. So $\|f^*_\delta\|_{L^q(\mathbb{S}^{n-1})} \approx 1$. On the other hand,

$$\lim_{\delta\to 0} |K^\delta| = 0.$$

Hence,

$$\lim_{\delta\to 0} \|f\|_{L^p(\mathbb{R}^n)} = 0 \quad \text{for any } p < \infty.$$

(3). If $p < n$, then there can be no bound of the form

$$\forall\varepsilon\ \exists C_\varepsilon : \|f^*_\delta\|_{L^p(\mathbb{S}^{n-1})} \leq C_\varepsilon\delta^{-\varepsilon}\|f\|_{L^p(\mathbb{R}^n)}.$$

Indeed, let $f = \chi_{B(0,\delta)}$. Then for all $e \in \mathbb{S}^{n-1}$, the tube $T_e^\delta(0)$ contains $B(0,\delta)$, so that

$$f_\delta^*(e) = \frac{|B(0,\delta)|}{|T_e^\delta(0)|} \gtrsim \delta.$$

Hence, $\|f_\delta^*\|_{L^p(\mathbb{S}^{n-1})} \approx \delta$. However, $\|f\|_{L^p(\mathbb{R}^n)} \approx \delta^{n/p}$.

(4). Interpolating the $L^1(\mathbb{R}^n) \to L^\infty(\mathbb{S}^{n-1})$ bound and the one in (iii), we have a family of conjectured inequalities: For any $\varepsilon > 0$,

$$\|f_\delta^*\|_{L^q(\mathbb{S}^{n-1})} \le C\delta^{-\frac{n}{p}+1-\varepsilon}\|f\|_{L^p(\mathbb{R}^n)}$$

for all

$$1 \le p \le n \quad \text{and} \quad q \le q(p) = \frac{p(n-1)}{p-1}.$$

(5). In fact, we show that the weak (p,q) bound of the Kakeya maximal function is sufficient to imply a dimension bound of Kakeya sets.

Theorem. *Assume the weak-type (p,q) estimate*

$$\|f_\delta^*\|_{L^{p,\infty}(\mathbb{S}^{n-1})} \le C\delta^{-\alpha}\|f\|_{L^p(\mathbb{R}^n)},$$

that is,

$$|\{e \in \mathbb{S}^{n-1} : f_\delta^*(e) > \lambda\}| \le \left(\frac{\delta^{-\alpha}\|f\|_{L^p(\mathbb{R}^n)}}{\lambda}\right)^q.$$

Then $\dim_{\mathrm{H}}(K) \ge n - p\alpha$.

Remark. If one proves $L^{p,\infty}(\mathbb{R}^n) \to L^q(\mathbb{S}^{n-1})$ estimate of the Kakeya maximal function with

$$-\alpha = -\frac{n}{p} + 1 - \varepsilon,$$

then

$$\dim_{\mathrm{H}}(K) \ge n - p\left(\frac{n}{p} - 1 + \varepsilon\right) = p - p\varepsilon$$

for all $\varepsilon > 0$. Hence, $\dim_{\mathrm{H}}(K) \ge p$.

Proof. Let K be a Kakeya set. Fix a covering of K by balls $B_j = B(x_j, r_j)$. Assume that all $r_j \le 1/100$ for all j.. We need to show that

$$\sum_j r_j^{n-p\alpha} \gtrsim 1.$$

Let

$$J_k = \{j : 2^{-k} \le r_j \le 2^{-(k-1)}\}.$$

For every $e \in \mathbb{S}^{n-1}$, K contains a unit line segment I_e parallel to e. Let

$$S_k = \left\{e \in \mathbb{S}^{n-1} : \left|I_e \cap \bigcup_{j \in J_k} B_j\right| \ge \frac{1}{100k^2}\right\}.$$

Since $\sum_k 1/(100k^2) < 1$ and

$$\sum_k \left|I_e \cap \bigcup_{j \in J_k} B_j\right| \ge |I_e| = 1,$$

it follows that for every e there exists k such that $e \in S_k$. Hence,

$$\bigcup_{k=1}^\infty S_k = \mathbb{S}^{n-1}.$$

We use pigeonhole principle again and conclude that there exists k such that

$$|S_k| \gtrsim \frac{1}{k^2}.$$

We next fix k. Let

$$F_k = \bigcup_{j\in J_k} B(x_j, 10r_j) \quad \text{and} \quad f = \chi_{F_k}.$$

Then for $e \in S_k$, we have that

$$|T_e^{2^{-k}}(a_e) \cap F_k| \gtrsim \frac{1}{100k^2}|T_e^{2^{-k}}(a_e)|,$$

in which a_e is the midpoint of I_e. Hence,

$$\|f^*_{2^{-k}}\|_{L^p(\mathbb{S}^{n-1})} \gtrsim k^{-2} \quad \text{on } S_k.$$

On the other hand, the weak (p,q) inequality implies that

$$|\{e \in \mathbb{S}^{n-1} : f^*_{2^{-k}}(e) \gtrsim k^{-2}\}| \lesssim \left(\frac{2^{k\alpha}\|f\|_{L^p(\mathbb{R}^n)}}{k^{-2}}\right)^q \le \left(k^2 2^{k\alpha}|F_k|^{\frac{1}{p}}\right)^q.$$

Comparing the previous two inequalities, we see that

$$k^{-2} \lesssim |S_k| \lesssim |\{e \in \mathbb{S}^{n-1} : f^*_{2^{-k}}(e) \gtrsim k^{-2}\}| \lesssim \left(k^2 2^{k\alpha}|F_k|^{\frac{1}{p}}\right)^q \lesssim \left(k^2 2^{k\alpha} 2^{-\frac{kn}{p}}(\#J_k)^{\frac{1}{p}}\right)^q.$$

Therefore,

$$(\#J_k) \gtrsim k^{-\frac{2}{q}-2p} 2^{k(n-p\alpha)} \gtrsim 2^{k(n-p\alpha)}.$$

Hence,

$$\sum_j r_j^{n-p\alpha} \ge \sum_{j\in J_k} 2^{-k(n-p\alpha)} \gtrsim (\#J_k) 2^{-k(n-p\alpha)} \gtrsim 1.$$

□

8.1.1. Geometric combinatorics arguments. We denote by Ω a maximal δ-separated set of directions on $\mathbb{S}^{n-1}$.

THEOREM 8.1. *Let $0 < \delta \ll 1$. If*

$$\left|\bigcup_{e\in\Omega} T_e^\delta(a_e)\right| \gtrsim \delta^{n-d}$$

independent of $\{a_e\}_{e\in\Omega}$, then $\dim_{\mathrm{M}}(K) \ge d$.

This is simply because K^δ contains the union of such well-separated δ-tubes.

In $\mathbb{R}^2$, one can estimate the size of the intersection of two tubes in the simple fashion:

$$|T_e^\delta \cap T_{e'}^\delta| \lesssim \frac{\delta^2}{\delta + \angle(e,e')}.$$

in which $\angle(e,e')$ is the angle between the two directions e and e'. This may be viewed as a quantitative version of the fact that two transverse lines only intersect in at most point. In other words, tubes in $\mathbb{R}^2$ are essentially disjoint. Hence, $\dim_{\mathrm{M}}(K) = 2$ for $K \subset \mathbb{R}^2$.

Such argument does not extend well to higher dimensions, because the situations in which two tubes intersect in $\mathbb{R}^n$ is much more diverse. Next we present a geometric combinatoric argument due to Bourgain[i].

[i] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis.* Geom. Funct. Anal. 1 (1991), no. 2, 147–187.

Remark (Bourgain's slice argument). Notice that any two separated points in $\mathbb{R}^n$ are connected by at most one line. This geometric fact leads to an estimate of $\dim_{\mathrm{M}}(K) \geq (n+1)/2$. Indeed, by Theorem 8.1, it suffices to prove that

$$|E| \gtrsim \delta^{\frac{n-1}{2}}, \quad \text{in which } E = \bigcup_{e\in\Omega} T_e.$$

We may assume that the tubes T_e all intersect the hyperplanes $x_n = 0$ and $x_n = 1$ and make an angle of $\leq 1/10$ with the vertical e_n. For all $0 \leq t \leq 1$, let $\tilde{A}_t$ denote the intersection of E with $\{x_n = t\}$. By Fubini theorem, for generic values of $t \in [0,1]$ we have that $|\tilde{A}_t| \lesssim |E|$. We apply a generic rescaling so that

$$\#A_0, \#A_1 \lesssim \delta^{1-n}|E|.$$

The sets $\tilde{A}_t$ are essentially unions of $(n-1)$-dim balls of radius δ. Thus, if we let A_t be a maximal δ-separated subset of $\tilde{A}_t$, then $\#A_t \approx \delta^{1-n}|\tilde{A}_t|$.

On the other hand, each tube T_e intersects A_0 in essentially one point, and similarly for A_1. Therefore, each T_e can be identified with a pair in $A_0 \times A_1$. Because each pair of points in $A_0 \times A_1$ has essentially only one tube connecting it, we have that

$$\delta^{1-n} \approx \#\Omega \leq \#(A_0 \times A_1) \lesssim (\delta^{1-n}|E|)^2,$$

which is the desired estimate.

Problem 8.1 (Kakeya sets and dynamics). *Can one find a dynamical system in which the Kakeya set is invariant?*

8.2. Kakeya problems and randomization

In this section, we prove the Kakeya conjecture that $\dim_{\mathrm{M}}(K) = 2$ in $\mathbb{R}^2$, using randomization. It is due to Tacy. To this end, let

$$E^\delta = \left|\bigcup_{e\in\Omega} T_e^\delta(a_e)\right|,$$

in which Ω is a maximal δ-separated set of directions on $\mathbb{S}^{n-1}$. It suffices to estimate the lower bound of $|E^\delta|$ to get the lower bound of $\dim_{\mathrm{M}}(K)$, the Minkowski dimension of the Kakeya set. In particular, $\dim_{\mathrm{M}}(K) = n$ follows from

$$|E^\delta| \gtrsim |\log\delta|^{-c} \quad \text{for some } c > 0.$$

Notice that for any $f \in L^p(\mathbb{R}^n)$,

$$\|f\|_{L^2(E^\delta)} \leq \left|E^\delta\right|^{\frac{1}{2}-\frac{1}{p}} \|f\|_{L^p(E^\delta)} \quad \text{for all } p \geq 2.$$

So any lower bound of $\|f\|_{L^2(E^\delta)}$ and upper bound of $\|f\|_{L^p(E^\delta)}$ lead to a lower bound of $|E^\delta|$. For example, if there exists $f \in L^p(\mathbb{R}^n)$ such that

$$\|f\|_{L^p(E^\delta)} \lesssim |\log\delta|^c \|f\|_{L^2(E^\delta)},$$

then

$$|E^\delta| \gtrsim c.$$

This is, of course, too strong to be true. However, it serves as an intuition of our attacking plan of estimating $\dim_{\mathrm{M}}(K)$. That is, we want to construct $f \in L^p(\mathbb{R}^n)$ whose L^p norm on E^δ is as close as possible to its L^2 norm. In particular,

$$\|f\|_{L^p(E^\delta)} \lesssim |\log\delta|^c \|f\|_{L^2(E^\delta)}.$$

To this end, we use randomization

$$f_u = \sum_{j=1}^{J} u_j \psi_j,$$

in which

$$\psi = \delta^{\frac{1-n}{2}} \cdot \chi_{T^\delta_{e_j}(a_{e_j})} \quad \text{for } e_j = 1, ..., J = |\Omega| = \delta^{1-n},$$

and $u = (u_1, ..., u_J)$ is a random variable on $\mathbb{S}^{J-1}$. That is, f is a random L^2-normalized function. Taking the expectation value, we similarly have that

$$\|f_u\|^2_{L^2\left(E^\delta\right)} \le \left|E^\delta\right|^{1-\frac{2}{p}} \|f_u\|^2_{L^p\left(E^\delta\right)},$$

and

$$1 = \mathbb{E}\left(\|f_u\|^2_{L^2\left(E^\delta\right)}\right) \le \left|E^\delta\right|^{1-\frac{2}{p}} \mathbb{E}\left(\|f_u\|^2_{L^p\left(E^\delta\right)}\right).$$

Now we only need to prove an upper bound that

$$\mathbb{E}\left(\|f_u\|^2_{L^p\left(E^\delta\right)}\right) \lesssim |\log\delta|^c.$$

Next we prove this in 2-dim, in which case p is taken as 4. First notice that

$$\mathbb{E}\left(\|f_u\|^2_{L^4\left(E^\delta\right)}\right)^{\frac{1}{2}} \le \mathbb{E}\left(\|f_u\|^4_{L^4\left(E^\delta\right)}\right)^{\frac{1}{4}}.$$

So we estimate

$$\begin{aligned} \mathbb{E}\left(\|f_u\|^4_{L^4\left(E^\delta\right)}\right) &= \int_{\mathbb{S}^{J-1}} \int_{\mathbb{R}^2} |f_u(x_1,x_2)|^4 \, dx_1 dx_2 du \\ &= \int_{\mathbb{R}} \int_{\mathbb{S}^{J-1}} \int_{\mathbb{R}} |f_u(x_1,x_2)|^4 \, dx_2 du dx_1 \\ &= \int_{\mathbb{R}} \mathbb{E}\left(\|f_u(x_1,\cdot)\|^4_{L^4_{x_2}}\right) dx_1, \end{aligned}$$

in which

$$\begin{aligned} &\mathbb{E}\left(\|f_u(x_1,\cdot)\|^4_{L^4_{x_2}}\right) \\ = \;& \int_{s^{J-1}} \int_{\mathbb{R}} \left|\sum_{j=1}^{J} u_j \psi_j(x_1,x_2)\right|^4 dx_2 du \\ = \;& \int_{s^{J-1}} \sum_{j_1,j_2,j_3,j_4=1}^{J} u_{j_1} u_{j_2} u_{j_3} u_{j_4} \int_{\mathbb{R}} \psi_{j_1}\psi_{j_2}\psi_{j_3}\psi_{j_4}(x_1,x_2)\, dx_2 du \\ = \;& \int_{s^{J-1}} \sum_{j_1,j_2=1}^{J} u_{j_1}^2 u_{j_2}^2 \int_{\mathbb{R}} \psi_{j_1}^2 \psi_{j_2}^2 (x_1,x_2)\, dx_2 du \\ \lesssim \;& \sum_{j_1,j_2=1}^{J} \mathbb{E}\left(u_{j_1}^2 u_{j_2}^2\right) \int_{\mathbb{R}} \psi_{j_1}^2 \psi_{j_2}^2(x_1,x_2)\, dx_2 \\ \lesssim \;& \delta^2 \sum_{j_1,j_2=1}^{J} \int_{\mathbb{R}} \psi_{j_1}^2 \psi_{j_2}^2 (x_1,x_2)\, dx_2, \end{aligned}$$

since odd-powered terms have 0 average and $\mathbb{E}\left(u_{j_1}^2 u_{j_2}^2\right) \lesssim \delta^2$. Now without loss of generality, we may assume that e_j's are all very close to x_1 direction, i.e., $|e_j - e_{x_1}| \ll 1$. Then T_e^δ intersects with the line $x_1 = c$ at an interval with length $\lesssim \delta$. Hence,

$$\int_{\mathbb{R}} \psi_{j_1}^2 \psi_{j_2}^2 (x_1, x_2)\, dx_2 \lesssim \left(\delta^{-\frac{1}{2}}\right)^2 \cdot \left(\delta^{-\frac{1}{2}}\right)^2 \cdot \delta \lesssim \delta^{-1}.$$

So

$$\begin{aligned} & \mathbb{E}\left(\|f_u(x_1, \cdot)\|_{L^4_{x_2}}^4\right) \\ \lesssim\ & \delta^{-2} \sum_{j_1, j_2 = 1}^{J} \int_{\mathbb{R}} \psi_{j_1}^2 \psi_{j_2}^2 (x_1, x_2)\, dx_2 \\ \lesssim\ & \delta N(x_1), \end{aligned}$$

in which $N(x_1)$ is the number of intersection regions crossed by the line at x_1. Then back to the expectation of $\|f_u\|_{L^4\left(E^\delta\right)}$,

$$\mathbb{E}\left(\|f_u\|_{L^4\left(E^\delta\right)}^4\right) \lesssim \delta \int_{\mathbb{R}} N(x_1)\, dx_1.$$

Write

$$N(x_1) = \sum_{k=1}^{K} N_k(x_1),$$

in which $N_k(x_1)$ is the number of intersections of tubes where $2^k\delta \le |e_{j_1} - e_{j_2}| < 2^{k+1}\delta$ at x_1. So $K \approx |\log \delta|$ and there are at most $N_k = 2^k \delta^{-1}$ such regions. However such an intersection region can only project to x_1 direction for a length of at most 2^{-k} (i.e., the wider two tubes intersect, quicker they will separate).

We now divide $x_1 \in [0, 1]$ into 2^k subintervals. Thus, each intersection region of the tubes with separation $\approx 2^k \delta$ can only fall into one and only one of these subintervals. Therefore,

$$\int_{\mathbb{R}} N(x_1)\, dx_1 = \sum_{k=1}^{K} \sum_{j=1}^{2^k} N_{k,j}(x_1) 2^{-k} = \sum_{k=1}^{K} 2^{-k} N_k \approx \delta^{-1} |\log \delta|.$$

Hence,

$$\mathbb{E}\left(\|f_u\|_{L^4\left(E^\delta\right)}^4\right) \lesssim \delta \int_{\mathbb{R}} N(x_1)\, dx_1 \lesssim |\log \delta|.$$

Because

$$1 = \mathbb{E}\left(\|f_u\|_{L^2\left(E^\delta\right)}^2\right) \le \left|E^\delta\right|^{\frac{1}{2}} \mathbb{E}\left(\|f_u\|_{L^p\left(E^\delta\right)}^2\right) \lesssim \left|E^\delta\right|^{\frac{1}{2}} \mathbb{E}\left(\|f_u\|_{L^4\left(E^\delta\right)}^4\right)^{\frac{1}{2}},$$

we have that

$$\left|E^\delta\right| \gtrsim |\log \delta|^{-1}$$

and we are done.

Remark. In fact, the above lower bound is sharp. This is due to Keich[i].

---

[i]U. Keich, *On L^p bounds for Kakeya maximal functions and the Minkowski dimension in $\mathbb{R}^2$*. Bull. London Math. Soc. 31 (1999), no. 2, 213–221.

CHAPTER 9

Miscellany problems

9.1. Pólya conjecture: Precise bounds of eigenvalues

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with piecewise smooth boundary. Denote $\{\lambda_k\}_{k=1}^{\infty}$ and $\{\mu_k\}_{k=1}^{\infty}$ the eigenvalues of the Laplacian with Dirichlet and Neumann boundary conditions, respectively. Then we know that

$$0 = \mu_1 < \mu_2 \leq \mu_3 \leq \cdots \to \infty \quad \text{and} \quad 0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \to \infty.$$

A set of problems that are of great interest in mathematics and physics are to determine the distribution of eigenvalues and its relation with the underlying geometry of the domain. The most famous such result is the Weyl Law.

Theorem (Weyl).

(i). For the Dirichlet eigenvalues,

$$N_{\mathrm{Dir}}(\lambda) := \#\{k : \lambda_k \leq \lambda\} = \frac{B_n|\Omega|}{(2\pi)^n}\lambda^{\frac{n}{2}} - \frac{|\partial\Omega|}{4(2\pi)^{n-1}}\lambda^{\frac{n-1}{2}} + R(\lambda).$$

(ii). For the Neumann eigenvalues,

$$N_{\mathrm{Neu}}(\lambda) := \#\{k : \mu_k \leq \lambda\} = \frac{B_n|\Omega|}{(2\pi)^n}\lambda^{\frac{n}{2}} + \frac{|\partial\Omega|}{4(2\pi)^{n-1}}\lambda^{\frac{n-1}{2}} + R(\lambda).$$

Here, B_n is the volume of the unit ball in $\mathbb{R}^n$, $|\Omega|$ is the volume of Ω, and $|\partial\Omega|$ is the surface area of the boundary of Ω.

Remark (Weyl conjecture). Weyl[i] proved that

$$N(\lambda) = \frac{B_n|\Omega|}{(2\pi)^n}\lambda^{\frac{n}{2}} + o\left(\lambda^{\frac{n}{2}}\right),$$

which was also known as the Lorentz conjecture. This is of course less precise than the ones in the above theorem. Indeed, Weyl conjectured that $R(\lambda) = o(\lambda^{\frac{n-1}{2}})$.

Ivrii[ii] proved that the conjecture holds, provided that the set of periodic trajectories of the (generalized) billiard flow in $\Omega \times \mathbb{S}^{n-1}$ is of measure zero. (This result is a generalization of Duistermaat-Guillemin[iii] on manifolds without boundary.) Ivrii further conjectured that such measure zero condition is true on any Euclidean domain. Its verification would imply Weyl conjecture immediately.

---

[i] H. Weyl, *Über die asymptotische Verteilung der Eigenwerte*. Nachr. Konigl. Ges. Wiss. Göttingen (1911) 110–117 and *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen*. Math. Ann. 71 (1912), no. 4, 441–479.

[ii] V. Ivrii, *The second term of the spectral asymptotics for a Laplace-Beltrami operator on manifolds with boundary*. Funktsional. Anal. i Prilozhen. 14 (1980), no. 2, 25–34.

[iii] J. Duistermaat and V. Guillemin, *The spectrum of positive elliptic operators and periodic bicharacteristics*. Invent. Math. 29 (1975), 39–79.

From the Weyl Law, we see that

$$\lambda_k \approx \mu_k \approx \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{2}{n}} \quad \text{as } k \to \infty.$$

Pólya[i] conjecture states that these asymptotics are exact:

Conjecture (Pólya conjecture).

$$\mu_k \le \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{2}{n}} \le \lambda_k \quad \textit{for all } k \in \mathbb{N}.$$

Pólya proved his conjecture on $\Omega \subset \mathbb{R}^2$ if Ω is plane covering. That is, the congruences of Ω can cover $\mathbb{R}^2$ (with no gaps). The examples of such domains include all triangles and quadrilaterals, as well as hexagon with a center of symmetry. As remarked by Li-Yau[ii], Pólya's results apply to $\Omega \subset \mathbb{R}^n$ if Ω is $\mathbb{R}^n$ covering.

A weaker problem than Pólya conjecture states that

Problem 9.1 (Interlacing of Dirichlet and Neumann eigenvalues). *Do the Dirichlet and Neumann eigenvalues always interlace, i.e.,*

$$0 = \mu_1 \le \lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \cdots \to \infty?$$

It is true on intervals, i.e., $\Omega \subset \mathbb{R}^1$*, but is unknown in higher dimensions.*

Despite a long history and lot of people (Lieb, Laptev, etc) attempting on Pólya conjecture, it remains open on general domains. Here, we focus on the Dirichlet eigenvalues. The best known result to date is due to Li-Yau:

$$\lambda_k \ge \frac{n}{n+2} \cdot \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{2}{n}}.$$

Notice that this lower bound is different from the conjectured one by a factor $n/(n+2)$. Such a bound is an immediate consequence of

$$\sum_{j=1}^{k} \lambda_j \ge \frac{n}{n+2} \cdot \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{n+2}{n}}. \tag{9.1}$$

Indeed, since $\lambda_k \ge \cdots \ge \lambda_1$,

$$k\lambda_k \ge \sum_{j=1}^{k} \lambda_j \ge \frac{n}{n+2} \cdot \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{n+2}{n}}.$$

Therefore, the Li-Yau bound follows.

From semiclassical point of view, the challenge of Pólya conjecture lies in the fact that it is extremely difficult to study an individual eigenvalue and eigenfunction at high energy. In almost all arguments (including proof of the Weyl Law by semiclassical analysis), one takes the average of eigenvalues/eigenfunctions in a spectral window and studies the average behavior in such a window.

Along the same vein, (9.1) can also be regarded as an "averaged" estimate of eigenvalues in the spectral window $[0, \lambda_k]$. Therefore, from the semiclassical point of view, the estimate is natural. In the following subsection, we revisit the argument in Li-Yau from this view and propose related questions.

[i] G. Pólya, *On the eigenvalues of vibrating membranes.* Proc. London Math. Soc. (3) 11 (1961), 419–433.

[ii] P. Li and S.-T. Yau, *On the Schrödinger equation and the eigenvalue problem.* Comm. Math. Phys. 88 (1983), no. 3, 309–318.

9.1.1. Revisit Li-Yau's argument: An exact trace formula. We follow Li-Yau's argument to attack Pólya conjecture. First we need a lemma.

LEMMA 9.1. *Let $\alpha, M_1, M_2 \geq 0$. Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ satisfies*

$$0 \leq f \leq M_1$$

and

$$\int_{\mathbb{R}^n} |z|^\alpha f(z)\, dz \leq M_2.$$

Then

$$\int_{\mathbb{R}^n} f(z)\, dz \leq \left(\frac{n+\alpha}{n}\right)^{\frac{n}{n+\alpha}} (M_1 B_n)^{\frac{\alpha}{n+\alpha}} M_2^{\frac{n}{n+\alpha}}.$$

PROOF. Let $R > 0$ be such that

$$M_1 \int_{|z|\leq R} |z|^\alpha\, dz = M_2.$$

Then explicitly, we have that

$$M_2 = M_1 \int_{|z|\leq R} |z|^\alpha\, dz = M_1 \omega_{n-1} \int_0^R r^{n-1+\alpha}\, dr = \frac{M_1 n B_n}{n+\alpha} R^{n+\alpha}.$$

Here, $\omega_{n-1} = nB_n$ is the surface area of unit sphere in $\mathbb{R}^n$. Therefore,

$$R = \left(\frac{M_2(n+\alpha)}{M_1 n B_n}\right)^{\frac{1}{n+\alpha}}.$$

Define

$$g(z) = \begin{cases} M_1 & \text{if } |z| < R \\ 0 & \text{if } |z| \geq R, \end{cases}$$

Then

$$\int_{\mathbb{R}^n} |z|^\alpha g(z)\, dz = M_1 \int_{|z|\leq R} |z|^\alpha\, dz = M_2.$$

Notice that since $0 \leq f \leq M_1$,

$$(R^\alpha - |z|^\alpha)(g(z) - f(z)) \geq 0.$$

It then follows that

$$R^\alpha \int_{\mathbb{R}^n} (g(z) - f(z))\, dz \geq \int_{\mathbb{R}^n} |z|^\alpha (g(z) - f(z))\, dz = M_2 - \int_{\mathbb{R}^n} |z|^\alpha f(z)\, dz \geq 0,$$

because $\int |z|^\alpha f \leq M_2$. Thus,

$$\int_{\mathbb{R}^n} f(z)\, dz \leq \int_{\mathbb{R}^n} g(z)\, dz = \int_{|z|<R} M_1\, dz = M_1 R^n B_n = \left(\frac{n+\alpha}{n}\right)^{\frac{n}{n+\alpha}} (M_1 B_n)^{\frac{\alpha}{n+\alpha}} M_2^{\frac{n}{n+\alpha}}.$$

□

Let

$$\Phi(x, y) = \sum_{j=1}^{k} u_j(x) u_j(y) \quad \text{for } x, y \in \Omega$$

be the kernel of spectral projection to the space spanned by the orthonormal eigenfunctions $\{u_j\}_{j=1}^k$ with eigenvalues $\{\lambda_j\}_{j=1}^k$. Without loss of generality, we assume that the eigenfunctions are real-valued.

Take the Fourier transform in x-variable:

$$\hat{\Phi}(\xi, y) = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i\xi\cdot x}\Phi(x, y)\, dx.$$

Here, the normalization $1/(2\pi)^{n/2}$ is chosen to make the transform unitary in $L^2(\mathbb{R}^n)$. Hence, from Plancherel theorem,

$$\int_{\mathbb{R}^n} |\hat{\Phi}(\xi, y)|^2\, d\xi = \int_{\mathbb{R}^n} |\Phi(x, y)|^2\, dx = \int_{\Omega} |\Phi(x, y)|^2\, dx,$$

since $\Phi(x, y) = 0$ outside of $\Omega \times \Omega$. So

$$\begin{aligned}
\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\hat{\Phi}(\xi, y)|^2\, d\xi dy &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} |\Phi(x, y)|^2\, dxdy \\
&= \int_{\Omega}\int_{\Omega} \sum_{l,m=1}^{k} u_l(x)u_m(y)u_l(x)u_m(y)\, dxdy \\
&= \sum_{l,m=1}^{k} \int_{\Omega} u_l(x)u_m(x)\, dx \int_{\Omega} u_l(y)u_m(y)\, dy \\
&= k.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
f(\xi) &:= \int_{\Omega} |\hat{\Phi}(\xi, y)|^2\, dy \\
&= \int_{\Omega} \left| \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \Phi(x, y)e^{-i\xi\cdot x}\, dx \right|^2 dy \\
&= \int_{\Omega} \left| \int_{\Omega} \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-i\xi\cdot x}\Phi(x, y)\, dx \right|^2 dy,
\end{aligned}$$

in which

$$\int_{\Omega} \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-i\xi\cdot x}\Phi(x, y)\, dx$$

is exactly the projection of the function $e^{-i\xi\cdot x}/(2\pi)^{n/2}$ to the space spanned by $\{u_j\}_{j=1}^k$. Hence, the above inequality continues as

$$\begin{aligned}
f(\xi) &= \int_{\Omega} \left| \int_{\Omega} \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-i\xi\cdot x}\Phi(x, y)\, dx \right|^2 dy \\
&\leq \int_{\Omega} \left| \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-i\xi\cdot x} \right|^2 dx \\
&= \frac{|\Omega|}{(2\pi)^n}.
\end{aligned}$$

Later we will use Lemma 9.1 for f with $M_1 = |\Omega|/(2\pi)^n$. To this end, we also need the bound of

$$\begin{aligned}
&\int_{\mathbb{R}^n} |\xi|^\alpha f(\xi)\, d\xi \\
= &\int_{\mathbb{R}^n} |\xi|^\alpha \int_{\Omega} |\hat{\Phi}(\xi, y)|^2\, dyd\xi
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^n} |\xi|^\alpha \int_\Omega \left| \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \Phi(x,y) e^{-i\xi\cdot x}\, dx \right|^2 dy d\xi \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\xi|^\alpha \int_\Omega \left(\int_\Omega \Phi(x,y) e^{-i\xi\cdot x}\, dx \cdot \int_\Omega \Phi(z,y) e^{i\xi\cdot z}\, dz \right) dy d\xi \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\xi|^\alpha \int_\Omega \int_\Omega e^{i\xi\cdot(z-x)} \left(\int_\Omega \Phi(x,y)\Phi(z,y)\, dy \right) dx d\xi dz \\
&= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\xi|^\alpha \int_\Omega \int_\Omega e^{i\xi\cdot(z-x)} \Phi(x,z)\, dx d\xi dz \\
&= \sum_{j=1}^k \int_\Omega \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_\Omega |\xi|^\alpha e^{i\xi\cdot(z-x)} u_j(x)\, dx d\xi \right) u_j(z) dz \\
&= \sum_{j=1}^k \langle \mathrm{Op}(|\xi|^\alpha) u_j, u_j \rangle_{L^2(\Omega)},
\end{aligned}
$$

which formally equals

$$\sum_{j=1}^k \lambda_j^{\frac{\alpha}{2}}.$$

because $\mathrm{Op}(|\xi|^\alpha) u_j = \lambda_j^{\alpha/2} u_j$. (Recall that our eigenvalues are λ.)

Then by Lemma 9.1 for $M_1 = |\Omega|/(2\pi)^n$ and $M_2 = \sum_{j=1}^k \lambda_j^{\alpha/2}$, we have that

$$
\begin{aligned}
k &= \int_{\mathbb{R}^n} f(\xi)\, d\xi \\
&\le \left(\frac{n+\alpha}{n} \right)^{\frac{n}{n+\alpha}} (B_n M_1)^{\frac{\alpha}{n+\alpha}} M_2^{\frac{n}{n+\alpha}} \\
&= \left(\frac{n+\alpha}{n} \right)^{\frac{n}{n+\alpha}} \left(\frac{B_n |\Omega|}{(2\pi)^n} \right)^{\frac{\alpha}{n+\alpha}} \left(\sum_{j=1}^k \lambda_j^{\frac{\alpha}{2}} \right)^{\frac{n}{n+\alpha}},
\end{aligned}
$$

which in term implies

$$\sum_{j=1}^k \lambda_j^{\frac{\alpha}{2}} \ge \frac{n}{n+\alpha} \cdot \frac{(2\pi)^\alpha}{(B_n|\Omega|)^{\frac{\alpha}{n}}} k^{\frac{n+\alpha}{n}}.$$

Using the fact that $\lambda_k \ge \cdots \ge \lambda_1$, we arrive at

$$\lambda_k^{\frac{\alpha}{2}} \ge \frac{n}{n+\alpha} \cdot \frac{(2\pi)^\alpha}{(B_n|\Omega|)^{\frac{\alpha}{n}}} k^{\frac{\alpha}{n}}$$

and

$$\lambda_k \ge \left(\frac{n}{n+\alpha} \right)^{\frac{2}{\alpha}} \cdot \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{2}{n}}.$$

Notice that

$$\left(\frac{n}{n+\alpha} \right)^{\frac{2}{\alpha}} \to 1 \quad \text{as } \alpha \to \infty.$$

Therefore, by choosing α large enough, we have that

$$\lambda_k \geq \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{2}{n}},$$

and Pólya conjecture is proved!

This is of course all too nice but with serious issues. Particularly, the crucial trace formula

$$\sum_{j=1}^{k} \langle \mathrm{Op}(|\xi|^\alpha) u_j, u_j \rangle_{L^2(\Omega)} = \sum_{j=1}^{k} \lambda_j^{\frac{\alpha}{2}}$$

is only formally true.

(1). $\mathrm{Op}(|\xi|^\alpha)$ is defined as a pseudo-differential operator. – It is not a differential operator unless α is an even integer. Since our interest is to take $\alpha \to \infty$, we next assume that α is even.

(2). If α is an even integer, then $\mathrm{Op}(|\xi|^\alpha) = \Delta^{\frac{\alpha}{2}}$, which is only valid when acting on functions in $C_0^\infty(\mathbb{R}^n)$. The eigenfunctions u are C^∞ in the interior of Ω, however are only C^0 across the boundary $\partial\Omega$. This is because $u = 0$ in $\mathbb{R}^n \setminus \Omega$ whereas $\partial_\nu u \not\equiv 0$ (otherwise $u \equiv 0$). So $\partial_\nu u$ has to jump across the boundary. Here, ∂_ν is the normal derivative on the boundary. This means that $\Delta^{\frac{\alpha}{2}} u$ is a distribution in $\mathbb{R}^d$ and apriori $\langle \Delta^{\frac{\alpha}{2}} u, u\rangle$ is not well defined since u is only in $C^0(\mathbb{R}^n)$.

In the view of the above discussion, we call

$$\sum_{j=1}^{k} \langle \mathrm{Op}(|\xi|^\alpha) u_j, u_j \rangle_{L^2(\Omega)} = \sum_{j=1}^{k} \lambda_j^{\frac{\alpha}{2}}$$

an *exact* trace formula.

When $\alpha = 2$, $\mathrm{Op}(|\xi|^\alpha) = \Delta$, but again only formally, i.e., it is exact for Schwartz functions, whereas eigenfunctions u_j are not. To see why exact trace formula is complicated, we compute by Green's formula for u such that $u = 0$ on $\partial\Omega$,

$$\begin{aligned}
& \frac{1}{(2\pi)^{\frac{n}{2}}} \int_\Omega |\xi|^2 e^{i\xi\cdot x} u(x)\, dx \\
= \;& \frac{1}{(2\pi)^{\frac{n}{2}}} \int_\Omega \Delta e^{-i\xi\cdot x} u(x)\, dx \\
= \;& \frac{1}{(2\pi)^{\frac{n}{2}}} \int_\Omega e^{-i\xi\cdot x} \Delta u(x)\, dx + \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} \left(\partial_\nu e^{-i\xi\cdot x} u(x) - e^{-i\xi\cdot x} \partial_\nu u(x)\right)\, d\sigma_x \\
= \;& \frac{\lambda}{(2\pi)^{\frac{n}{2}}} \int_\Omega e^{-i\xi\cdot x} u(x)\, dx - \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} e^{-i\xi\cdot x} \partial_\nu u(x)\, d\sigma_x,
\end{aligned}$$

in which $d\sigma$ and μ are the surface measure the normal on $\partial\Omega$. Therefore,

$$\begin{aligned}
& \left\langle \mathrm{Op}(|\xi|^2) u, u \right\rangle_{L^2(\Omega)} \\
= \;& \int_\Omega \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_\Omega |\xi|^2 e^{i\xi\cdot(z-x)} u(x)\, dx d\xi \right) u(z) dz \\
= \;& \int_\Omega \left(\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \left(\lambda \hat{u}(\xi) e^{i\xi\cdot z} - \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} e^{i\xi\cdot(z-x)} \partial_\nu u(x) \right)\, d\xi \right) u(z) dz \\
= \;& \int_\Omega \lambda \left(\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \hat{u}(\xi) e^{i\xi\cdot z}\, d\xi \right) u(z)\, dz - \int_\Omega \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \int_{\partial\Omega} e^{i\xi\cdot(z-x)} \partial_\nu u(x)\, d\sigma_x\, d\xi \right) u(z)\, dz
\end{aligned}$$

$$= \lambda \int_{\Omega} |u(x)|^2 \, dx - \frac{1}{(2\pi)^n} \int_{\Omega} \left(\int_{\mathbb{R}^n} \int_{\partial\Omega} e^{i\xi\cdot(z-x)} \partial_\nu u(x) \, d\sigma_x d\xi \right) u(z) dz.$$

Taking a closer look at the second term:

$$\int_{\mathbb{R}^n} \int_{\partial\Omega} e^{-i\xi\cdot(z-x)} \partial_\nu u(x) \, d\sigma_x d\xi$$

is formally a SDO with symbol $\nu_x \cdot \xi|_{\partial\Omega}$, in which ν_x is the normal vector at $x \in \partial\Omega$.

It then follows that

$$\begin{aligned} &\sum_{j=1}^{k} \left\langle \mathrm{Op}(|\xi|^2) u_j, u_j \right\rangle_{L^2(\Omega)} \\ = \; &\sum_{j=1}^{k} \lambda_j - \sum_{j=1}^{k} \left\langle \mathrm{Op}(\nu_x \cdot \xi) u_j, u_j \right\rangle . \end{aligned}$$

again, only formally. If the local Weyl law holds for the symbol $\nu_x \cdot \xi$, then the second term above should equal

$$\int_{\partial\Omega} \int_{\mathbb{R}^n} (\nu_x \cdot \xi) \; d\xi d\sigma_x = \int_{\mathbb{R}^n} \int_{\partial\Omega} (\nu_x \cdot \xi) \; d\sigma_x d\xi = 0,$$

in which we use Stoke's theorem to deduce $\int_{\partial\Omega} \nu_x \cdot \xi = 0$ for any fixed $\xi \in \mathbb{R}^n$. Of course, to justify all of these requires quite some work.

We continue with $\alpha = 4$ so $\mathrm{Op}(|\xi|^\alpha) = \Delta^2$.

$$\begin{aligned} & \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} |\xi|^4 e^{i\xi\cdot x} u(x) \, dx \\ = \; & \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} \Delta^2 e^{-i\xi\cdot x} u(x) \, dx \\ = \; & \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} \Delta e^{-i\xi\cdot x} \Delta u(x) \, dx + \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} \left(\partial_\nu \Delta e^{-i\xi\cdot x} u(x) - \Delta e^{-i\xi\cdot x} \partial_\nu u(x) \right) \, d\sigma_x \\ = \; & \frac{\lambda}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} \Delta e^{-i\xi\cdot x} u(x) \, dx - \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} e^{-i\xi\cdot x} |\xi|^2 \partial_\nu u(x) \, d\sigma_x \\ = \; & \frac{\lambda}{(2\pi)^{\frac{n}{2}}} \left(\int_{\Omega} e^{-i\xi\cdot x} \Delta u(x) \, dx - \int_{\partial\Omega} e^{-i\xi\cdot x} \partial_\nu u(x) \, d\sigma_x \right) - \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} e^{-i\xi\cdot x} |\xi|^2 \partial_\nu u(x) \, d\sigma_x \\ = \; & \frac{\lambda^2}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} e^{-i\xi\cdot x} u(x) \, dx + \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\partial\Omega} e^{-i\xi\cdot x} \left(\lambda + |\xi|^2 \right) \partial_\nu u(x) \, d\sigma_x. \end{aligned}$$

Therefore,

$$\begin{aligned} &\sum_{j=1}^{k} \left\langle \mathrm{Op}(|\xi|^4) u_j, u_j \right\rangle_{L^2(\Omega)} \\ = \; &\sum_{j=1}^{k} \lambda_j^2 - \sum_{j=1}^{k} \left\langle \mathrm{Op}\left(\left(\lambda + |\xi|^2 \right) \nu_x \cdot \xi \right) u_j, u_j \right\rangle . \end{aligned}$$

9.1.2. Return back to Li-Yau's argument. However, Li-Yau invented a clever argument to circumvent the exact trace formula for $\mathrm{Op}(|\xi|^2)$. First,

$$\begin{aligned}
\xi_j \hat{u}(\xi) &= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} \xi_j e^{-i\xi\cdot x} u(x)\, dx \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} i\partial_{x_j} e^{-i\xi\cdot x} u(x)\, dx \\
&= -\frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\Omega} e^{-i\xi\cdot x} i\partial_{x_j} u(x)\, dx + \int_{\partial\Omega} e^{-i\xi\cdot x} u(x) \vec{n}_j \\
&= -\widehat{i\partial_{x_j} u}.
\end{aligned}$$

Here, $\vec{n}_j$ is the j-th component of the normal vector and we use that fact $u = 0$ on $\partial\Omega$. Therefore,

$$\xi_j \hat{\Phi}(\xi, y) = -\widehat{i\partial_{x_j} \Phi}(x, y).$$

Second, going back to the estimate of

$$\begin{aligned}
&\int_{\mathbb{R}^n} |\xi|^2 f(\xi)\, d\xi \\
=\ & \int_{\mathbb{R}^n} |\xi|^2 \int_{\Omega} |\hat{\Phi}(\xi, y)|^2\, dy d\xi \\
=\ & \int_{\mathbb{R}^n} \int_{\Omega} |\xi \hat{\Phi}(\xi, y)|^2\, dy d\xi \\
=\ & \int_{\Omega} \int_{\Omega} |\nabla_x \Phi(x, y)|^2\, dy dx \\
=\ & \int_{\Omega} \int_{\Omega} \left| \sum_{j=1}^{k} \nabla_x u_j(x) u_j(y) \right|^2 dy dx \\
=\ & \sum_{l,m=1}^{k} \int_{\Omega} \int_{\Omega} \nabla_x u_l(x) \cdot \nabla_x u_m(x) u_l(y) u_m(y)\, dy dx \\
=\ & \sum_{j=1}^{k} \int_{\Omega} |\nabla u_j(x)|^2\, dx \\
=\ & \sum_{j=1}^{k} \int_{\Omega} u_j(x) \Delta u_j(x)\, dx \\
=\ & \sum_{j=1}^{k} \lambda_j.
\end{aligned}$$

Now we are in the position to apply Lemma 9.1 with $\alpha = 2$, $M_1 = |\Omega|/(2\pi)^n$, and $M_2 = \sum_{j=1}^{k} \lambda_j$. That is, the Li-Yau bound

$$\lambda_k \geq \frac{n}{n+2} \cdot \frac{(2\pi)^2}{(B_n|\Omega|)^{\frac{2}{n}}} k^{\frac{2}{n}}.$$

Remark. We shall remark that the exact trace formula can be established for some special operators. Christianson[i] recently proved an exact formula of Neumann data on the triangles. In particular, let A, B, C be the three sides of a triangle T with side-length a, b, c respectively. Then

$$\int_A |\partial_\nu u_j|^2 = \frac{a}{|T|}\lambda_j.$$

Denote $V = (\partial_\nu r_A)^\star \partial_\nu r_A$, where r_A is the restriction operator on the side A. Then the above equation reads

$$\langle V u_j, u_j \rangle_{L^2(T)} = \langle \partial_\nu r_A u_j, \partial_\nu r_A u_j \rangle_{L^2(T)} = \langle \partial_\nu u_j, \partial_\nu u_j \rangle_{L^2(A)} = \frac{a}{|T|}\lambda_j.$$

Therefore, there is an exact trace formula for the operator A:

$$\sum_{j=1}^{k} \langle V u_j, u_j \rangle_{L^2(T)} = \frac{a}{|T|} \sum_{j=1}^{k} \lambda_j.$$

This is the only known case for which one has an exact trace formula. But of course the operator V has a singular support on the boundary. What implications we have on Pólya conjecture is not clear.

Problem 9.2 (Pólya conjecture on manifolds). *Can one generalize Li-Yau's argument to manifolds with boundary?*

[i]H. Christianson, *Equidistribution of Neumann data mass on triangles.* Proc. Amer. Math. Soc. 145 (2017), no. 12, 5247–5255.

CHAPTER 10

Expository articles

10.1. Baire category theorem

Let X be a metric space with metric d.

Definition. Let $E \subset X$.

- The interior E° of E is the union of all open sets contained in E.
- The closure $\overline{E}$ of E is the intersection of all closed sets containing E.

Definition. Let $E \subset X$.

- We say that E is dense in X if $\overline{E} = X$.
- We say that E is nowhere dense if $(\overline{E})^\circ = \emptyset$.

Remark. A set $E \subset X$ is closed and nowhere dense iff E^c is open and dense.

Definition. Let $E \subset X$.

- We say that E is of the first category in X if E is a countable union of nowhere dense sets in X. A set of the first category is sometimes said to be meager.
- We say that E is of the second category in X if it is not of the first category in X.
- We say that E is generic if E^c is of the first category, that is, E is a countable intersection of sets, each of whose interior is dense in X. A generic set is sometimes said to be residual.

THEOREM 10.1. *Let X be a complete metric space.*

(i). Let $\{U_j\}_{j=1}^\infty$ be a sequence of open dense subsets of X. Then $\cap_{j=1}^\infty U_j$ is dense in X.
(ii). X is of the second category in itself, that is, it cannot be written as the countable union of nowhere dense sets.
(iii). Each generic set is dense in X.

Example. We provide two examples of generic sets with Lebesgue measure zero, see Chapter 4 in Stein-Shakarchi[i].

- Let $\{q_j\}_{j=1}^\infty$ be an enumeration of the rational numbers in $\mathbb{R}$. Define
$$U_j = \bigcup_{j=1}^\infty \left(q_j - \frac{1}{k2^j}, q_j - \frac{1}{k2^j} \right).$$
Then $U = \cap_{k=1}^\infty U_k$ is generic but has Lebesgue measure zero.
- Let $n \geq 2$. Define
$$\Lambda_n = \left\{ x \in \mathbb{R} : \text{there are infinitely many distinct fractions } \frac{p}{q} \text{ such that } \left| x - \frac{p}{q} \right| \leq \frac{1}{q^n}. \right\}.$$
Then Λ_n is generic but has Hausdorff dimension $2/n$.

[i] E. Stein and R. Shakarchi, *Functional analysis.* Princeton University Press, Princeton, NJ, 2011.

10.2. Bargmann and Bergman: Both have their kernels!

There are Bargmann kernel and Bergman kernel in mathematics, which are named after Valentine Bargmann and Stefan Bergman, respectively. In this section, we briefly discuss their lives and related works to these kernels.

10.2.1. Bargmann. Valentine "Valya" Bargmann (April 6, 1908 - July 20, 1989) was a German-born American mathematician and theoretical physicist. He escaped Nazi Germany and first worked as an assistant to Albert Einstein at the Institute for Advanced Study (IAS). He then taught at Princeton University and remained there for his entire career.

According to MathSciNet, Bargmann and Einstein together published one paper[i]. This paper was followed by one authored by Einstein alone[ii] in the same volume of Annals of Mathematics. As far as I understand, in these two papers they attempted to build a mathematical formalism using a concept of "bivectors" to explain the theory of gravitation.

Another highlight of Bargmann's research while working together with Einstein involves the Kaluza-Klein theory, a unified field theory of gravitation and electromagnetism, which is considered an important precursor to string theory. With Bergmann, three of them published a paper[iii] on the Kaluza-Klein theory, building upon an earlier work of Einstein and Bergmann[iv]. See Witten[v] on the roles of these three people in the development of the Kaluza-Klein theory.

Here, Bergmann, Einstein and Bargmann's collaborator, is not to be confused with Bergman, who is discussed in Section 10.2.2. In particular, Peter Gabriel Bergmann (March 24, 1915 - October 19, 2002) was a German-born American physicist and also worked as an assistant to Einstein at the IAS (1936-1941).

In addition to his joint work with Einstein, Bargmann made several important contributions to physics and mathematics, including the representation theory of $\mathrm{SL}(2,\mathbb{R})$ and the Lorentz group[vi], and Bargmann transform. His honors include the first Wigner Medal (joint with Eugene Wigner himself) in 1978 and the Max Planck Medal of the German Physical Society in 1988.

Next we discuss the Bargmann[vii] kernel, following Section 1.6 in Folland[viii]. First,

Definition (Fock spaces). A holomorphic function f on $\mathbb{C}^n$ is in the Fock space $\mathbb{F}^n$ if

$$\|f\|_{\mathbb{F}^n}^2 := \int_{\mathbb{C}^n} |f(z)|^2 e^{-\pi|z|^2}\, dz < \infty.$$

Hence, the Fock space is a Hilbert space with respect to the associated inner product

$$\langle f, g\rangle_{\mathbb{F}^n} := \int_{\mathbb{C}^n} f(z)\overline{g(z)} e^{-\pi|z|^2}\, dz.$$

[i] A. Einstein and V. Bargmann, *Bivector fields.* Ann. of Math. (2) 45 (1944), 1–14.

[ii] A. Einstein, *Bivector fields.* II. Ann. of Math. (2) 45 (1944), 15–23.

[iii] A. Einstein, V. Bargmann, and P. Bergmann, *On the five-dimensional representation of gravitation and electricity.* Theodore von Kàrmàn Anniversary Volume, pp. 212–225 California Institute of Technology, Pasadena, CA, 1941.

[iv] A. Einstein and P. Bergmann, *On a generalization of Kaluza's theory of electricity.* Ann. of Math. (2) 39 (1938), no. 3, 683–701.

[v] E. Witten, *A Note on Einstein, Bergmann, and the Fifth Dimension*, arXiv:1401.8048.

[vi] V. Bargmann, *Irreducible unitary representations of the Lorentz group.* Ann. of Math. (2) 48 (1947), 568–640.

[vii] V. Bargmann, *On a Hilbert space of analytic functions and an associated integral transform.* Comm. Pure Appl. Math. 14 (1961), 187–214 and *On a Hilbert space of analytic functions and an associated integral transform. Part II. A family of related function spaces. Application to distribution theory.* Comm. Pure Appl. Math. 20 (1967), 1–101.

[viii] G. Folland, *Harmonic analysis in phase space.* Princeton University Press, Princeton, NJ, 1989.

The Fock spaces are also called Bargmann-Fock spaces, Bargmann spaces, Segal-Bargmann spaces, named after Vladimir Fock, Valentine Bargmann, and Irving Segal, see Folland for an account of the history.

The original motivation of Fock spaces is to introduce a Hilbert representation of holomorphic functions that is related to the foundation of quantum physics. That is, a function in $\mathbb{F}^n$ is understood as the phase space representation of a wave function for a quantum particle moving in $\mathbb{R}^n$. In this view, $\mathbb{C}^n \simeq \mathbb{R}^n \times \mathbb{R}^n$ plays the role of the phase space, whereas $\mathbb{R}^n$ is the configuration space. Therefore, a transform is needed to transform the wave function $u \in L^2(\mathbb{R}^n)$ to one in $\mathbb{F}^n$, and this is the Bargmann transform:

Definition (Bargmann transform). Let $u \in L^2(\mathbb{R}^n)$. Define the Bargmann transform

$$Bf(z) = \int_{\mathbb{R}^n} B(z,y) f(y)\, dy,$$

in which

$$B(z,y) = 2^{\frac{n}{4}} e^{2\pi yz - \pi y^2 - \frac{\pi z^2}{2}}$$

is called the Bargmann kernel.

Set $z = x + i\xi$. Then

$$\begin{aligned} Bf(x+i\xi) &= 2^{\frac{n}{4}} \int_{\mathbb{R}^n} e^{2\pi yz - \pi y^2 - \frac{\pi z^2}{2}} u(y)\, dy \\ &= 2^{\frac{n}{4}} e^{\pi x^2} e^{-\frac{\pi(x+i\xi)^2}{2}} \int_{\mathbb{R}^n} e^{2\pi i y\xi - \pi(y-x)^2} u(y)\, dy, \end{aligned}$$

Here, the integral is the inner product of u and a coherent state $e^{-2\pi i y\xi - \pi(y-x)^2}$, which is centered at $(x, -\xi)$ in the phase space.

10.2.2. Bergman. Stefan Bergman (May 5, 1895 - June 6, 1977) was a Polish-born American mathematician whose primary work was in complex analysis. His name is also written Bergmann (same as Peter Bergmann mentioned above) and he dropped the second "n" when he came to the US. He is best known for the kernel function he discovered. The following discussion on Bergman[i] kernel follows Section 1.4 in Krantz[ii].

Definition (Bergman spaces). Let $\Omega \subset \mathbb{C}^n$ be a domain, i.e., it is connected and open. The Bergman space $A^2(\Omega)$ is the completion of the space of holomorphic functions in Ω in the norm $L^2(\Omega)$.

The Bergman space $A^2(\Omega)$ is a complete Hilbert space. For each $z \in \Omega$, the evaluation

$$F_z : f \to f(z)$$

is a continuous linear functional on $A^2(\Omega)$. By the Riesz representation theorem,

$$F_z(f) = f(z) = \int_\Omega K(z,w) f(w)\, dw \quad \text{for each } z \in \Omega,$$

in which the reproducing kernel $K(z,w)$ is called the Bergman kernel. Formally, the Bergman kernel in a complex domain is an analogue of the delta distribution in $\mathbb{R}^n$: For all $f \in \mathcal{S}(\mathbb{R}^n)$

[i] S. Bergman, *The kernel function and conformal mapping.* Second edition. American Mathematical Society, Providence, RI, 1970.

[ii] S. Krantz, *Function theory of several complex variables.* AMS Chelsea Publishing, Providence, RI, 2001.

and $x \in \mathbb{R}^n$,

$$f(x) = \int_{\mathbb{R}^n} \delta(x-y)\, dy, \quad \text{in which } \delta(x-y) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{i(x-y)\cdot\xi}\, d\xi.$$

Since $A^2(\Omega)$ is separable, there exists an orthonormal basis $\{\varphi_j\}_{j=1}^{\infty}$. Then

$$K_\Omega(z,w) = \sum_{j=1}^{\infty} \varphi_j(z)\overline{\varphi_j(w)}.$$

The Bergman kernels in most domains are not explicit, except, for instance, the unit ball.

Example (Bergman kernel in the unit ball). Let $\mathbb{B}^m = \{z = (z_1, ..., z_m) \in \mathbb{C}^m : |z| < 1\}$ be the unit ball in $\mathbb{C}^m$. Then there is an orthonormal basis $\{z^\alpha\}$, in which $\alpha = (\alpha_1, ..., \alpha_m) \in \mathbb{N}^m$ runs over all the multi-indices. Therefore, the Bergman kernel in $\mathbb{B}^m$ is

$$K(z,w) = \sum_\alpha \frac{z^\alpha \overline{w^\alpha}}{\|z^\alpha\|^2_{L^2(\mathbb{B}^m)}} = \frac{m!}{\pi^m (1 - z\cdot\overline{w})^{m+1}},$$

in which $z \cdot \overline{w} = z_1\overline{w_1} + \cdots + z_m\overline{w_m}$. Here, the second equation can be found in Theorem 1.4.22 in Krantz.

A closely related concept to the Bergman kernel is the Szegö kernel.

Definition (Szegö kernel). Let $\Omega \subset \mathbb{C}^m$ be a bounded domain with C^1 boundary. Denote by $A(\Omega)$ the space of functions that are holomorphic in Ω and are continuous in $\overline{\Omega}$. Let $H^2(\partial\Omega) \subset L^2(\partial\Omega)$ be the completion of the restriction of $A(\Omega)$ to $\partial\Omega$.

For each $f \in H^2(\partial\Omega)$ and $z \in \partial\Omega$, the evaluation

$$\Psi_z(f) : f \to P(f)(z)$$

is continuous, in which P is the Possion integral operator so $P(f)$ is the holomorphic extension of f to Ω. By the Riesz representation theorem,

$$P(f)(z) = \int_{\partial\Omega} S(z,w) f(w)\, d\sigma_w \quad \text{for each } z \in \Omega,$$

in which $S_\Omega(z,w)$ is called the Szegö kernel.

Example (Szegö kernel in the unit sphere). Let $\mathbb{S}^m = \partial\mathbb{B}^m$ be the unit sphere in $\mathbb{C}^m$. Then the Szegö kernel is

$$S(z,w) = \sum_\alpha \frac{z^\alpha \overline{w^\alpha}}{\|z^\alpha\|^2_{L^2(\mathbb{S}^m)}} = \frac{(m-1)!}{2\pi^m (1 - z\cdot\overline{w})^{m+1}},$$

in which the second equation can be found in Theorem 1.5.5 in Krantz.

10.3. Classical mechanics: Lagrangian and Hamiltonian

Classical mechanics is concerned with the physical laws that describe the motion of material bodies. Consider the mechanical system described at any fixed instant time t by the collection of the position coordinates $q = (q_1, ..., q_n)$ and the collection of the velocity components $\dot{q} = (\dot{q}_1, ..., \dot{q}_n)$, which are needed for a complete specification. The number n of coordinates is called the number of degrees of freedom in the system. The configuration space is the space of positions, which we take as a differential manifold $\mathbb{M}$ of dimension n.

In this section, we visit the two main approaches to describe classical mechanics: Lagrangian and Hamiltonian.

10.3.1. Lagrangian mechanics.

Theorem (Hamilton's principal of least action). *Motions of the mechanical system coincide with extremals of the functional*

$$\Phi(g(t)) = \int_{t_0}^{t_1} L(q, \dot{q}, t)\, dt,$$

in which the Lagrangian $L = T - V$ *is the difference of the kinetic and potential energy*

$$T = \frac{1}{2} m |\dot{q}|_q^2 \quad \textit{and} \quad V = V(q).$$

Remark. The curve $q(t) : [t_0, t_1] \to \mathbb{M}$ is an extremal of the functional Φ on the space of curves joints (t_0, q_0) and (t_1, q_1) iff the Euler-Lagrangian equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial L}{\partial q_j}, \quad j = 1, ..., n,$$

are satisfied along $q(t)$. Notice that these are n second order differential equations.

As an example, we derive the motion of a free particle on a manifold, i.e., the equations of the geodesic flow. A freely moving particle on a Riemannian manifold $\mathbb{M}$ has Lagrangian

$$L(q, \dot{q}, t) = \frac{1}{2} \sum_{i,j=1}^{n} g_{ij} \dot{q}_i \dot{q}_j.$$

The curve

$$q : [0, T] \to \mathbb{M}$$

is a geodesic if it is a critical point of the action integral

$$\int_0^T L(q, \dot{q}, t)\, dt.$$

Hence, it satisfies the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = \frac{\partial L}{\partial q_k}, \quad k = 1, ..., n.$$

Compute that

$$\frac{\partial L}{\partial q_k} = \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial g_{ij}}{\partial q_k} \dot{q}_i \dot{q}_j.$$

Then

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = \frac{d}{dt}\left(\sum_{i=1}^{n} g_{ik}\dot{q}_i\right) = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} \frac{\partial g_{ik}}{\partial q_j}\dot{q}_j\dot{q}_i + g_{ik}\ddot{q}_k\right).$$

Hence,

$$\frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial g_{ij}}{\partial q_k} \dot{q}_i \dot{q}_j = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} \frac{\partial g_{ik}}{\partial q_j}\dot{q}_j\dot{q}_i + g_{ik}\ddot{q}_k\right).$$

Thus,

$$\sum_{i=1}^{n} g_{ik}\ddot{q}_k = -\sum_{i,j=1}^{n} \frac{\partial g_{ik}}{\partial q_j}\dot{q}_j\dot{q}_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial g_{ij}}{\partial q_k}\dot{q}_i\dot{q}_j = -\frac{1}{2}\sum_{i,j=1}^{n}\left(\frac{\partial g_{ik}}{\partial q_j} + \frac{\partial g_{jk}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k}\right)\dot{q}_i\dot{q}_j.$$

Therefore,

$$\ddot{q}_k = -\frac{1}{2}\sum_{l=1}^{n} g^{lk} \sum_{i,j=1}^{n} \left(\frac{\partial g_{il}}{\partial q_j} + \frac{\partial g_{jl}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k} \right) \dot{q}_i \dot{q}_j = -\sum_{i,j=1}^{n} \Gamma^k_{ij} \dot{q}_i \dot{q}_j,$$

in which

$$\Gamma^k_{ij} = \frac{1}{2}\sum_{l=1}^{n} g^{lk} \left(\frac{\partial g_{il}}{\partial q_j} + \frac{\partial g_{jl}}{\partial q_i} - \frac{\partial g_{ij}}{\partial q_k} \right)$$

is the Christoffel symbols for the Riemannian connection. That is,

$$\nabla_{\frac{\partial}{\partial q_i}} \frac{\partial}{\partial q_j} = \sum_{k=1}^{n} \Gamma^k_{ij} \frac{\partial}{\partial_{q_k}}.$$

The geodesic determines a curve

$$t \to (x(t), y(t)) \subset T\mathbb{M}.$$

In the local coordinates $(q, \dot{q})$, it satisfies

$$\begin{cases} \dot{x}_k = y_k, \\ \dot{y}_k = -\sum_{i,j=1}^{n} \Gamma^k_{ij} y_i y_j. \end{cases}$$

10.3.2. Hamilton mechanics. Assume that $L : T\mathbb{M} \times \mathbb{R} \to \mathbb{R}$ is convex. Then Legendre transform of L, viewed as a function of $\dot{q}$, is

$$p\dot{q} - L(q, \dot{q}, t) := H(q, p, t)$$

which defines the Hamiltonian $H : T^*\mathbb{M} \times \mathbb{R} \to \mathbb{R}$.

Theorem (Hamilton's equations). *The system of Euler-Lagrange equations is equivalent to the system of* $2n$ *first order differential equations*

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p}, \\ \dot{p} = \frac{\partial H}{\partial q}, \end{cases}$$

which are called the Hamilton's equations.

For the Lagrangian $L = T - V$, the Hamiltonian $H = T + V$ is the total energy. Compute that

$$\frac{dH}{dt} = \frac{\partial H}{\partial q}\dot{q} + \frac{\partial H}{\partial p}\dot{p} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial H}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial H}{\partial t} = \frac{\partial H}{\partial t}.$$

Hence, for a system whose Hamiltonian function does not depend explicitly on time ($\partial H/\partial t = 0$), the law of conservation of the Hamiltonian function holds:

$$H(q(t), p(t)) = c.$$

Definition (First integral). If the function $I : T^*\mathbb{M} \to \mathbb{R}$ satisfies $I(q(t), p(t)) = c$ for integral curves $(p(t), q(t))$ of the Hamiltonian equations, then I is called a first integral. Clearly, the Hamiltonian is a first integral if $\partial H/\partial t = 0$.

Remark.

$$\frac{dI}{dt} = \frac{\partial I}{\partial q}\dot{q} + \frac{\partial I}{\partial p}\dot{p} = \frac{\partial I}{\partial q}\frac{\partial H}{\partial p} - \frac{\partial I}{\partial p}\frac{\partial H}{\partial q} = [I, H].$$

Hence, I is a first integral iff $[I, H] = 0$.

Definition (Involution). Two functions F_1 and F_2 on a symplectic manifold are in involution is $[F_1, F_2] = 0$.

Liouville proved that in a system with n degrees of freedom, if there are n independent first integrals in involution, then the system is integrable by quadratures.

Theorem (Liouville's theorem). *Let $F_1, ..., F_n$ be n functions in involution on a symplectic* $2n$-dim *manifold* $(\mathbb{M}, \omega)$. *Consider a level set of the functions F_i*

$$M_a = \{x \in \mathbb{M} : F_i(x) = a_i, i = 1, ..., n\}.$$

Assume that F_i's are independent on M_a, that is, the n 1-forms dF_i's are linearly independent at each point of M_a. Then

(i). The phase flows g_i defined by Hamiltonian F_i commute with each other: $g_i^t g_j^s = g_j^s g_i^t$.
(ii). M_a is a smooth Lagrangian manifold, invariant under the phase flow with Hamiltonian function g_i.
(iii). If M_a is compact and connected, then it is diffeomorphic to $\mathbb{T}^n$.
(iv). The phase flow with Hamiltonian $H = F_1$ determines a conditionally periodic motion on M_a, i.e., in angular coordinates $\phi = (\phi_1, ..., \phi_n)$ we have

$$\frac{d\phi}{dt} = \alpha(a).$$

(v). The canonical equations with Hamiltonian H can be integrated by quadratures.

Question 10.1. *In Liouville's theorem, what conclusion can one draw if one drops the assumption that M_a is compact and connected?*

10.4. Compactness and compact maps

A set being compact and a map being compact are two different bur related concepts that can be easily confused. In this section, we visit their basic definitions, following Section 4.4 in Folland[i] and Section 5.3 in Gilbarg-Trudinger[ii].

Definition (Compact sets). Let X be a topological space and $Y \subset X$.

- Y is said to be compact if every open cover of Y in X has a finite subcover.
- Y is said to be precompact if its closure is compact.
- Y is said to be countably compact if every countable open cover of Y in X has a finite subcover.
- Y is sequentially compact if every sequence in Y has a convergence subsequence.

Theorem. *If X is a metric space, then compactness is equivalent with sequential compactness.*

Definition (Compact maps). Let X and Y be normed linear spaces. A map $T : X \to Y$ is called compact (or completely continuous) if T maps bounded sets in X into relatively compact sets in Y, or equivalently T maps bounded sequences in X into sequences in Y which contain convergent subsequences.

Theorem. *Let $T : X \to Y$ be a linear map.*

(i). If T is compact, then T is also continuous.
(ii). Assume that $\dim(Y) < \infty$. *If T is continuous, then T is also compact.*

PROOF.

[i] G. Folland, *Real analysis.* Second edition. John Wiley & Sons Inc., New York, NY, 1999.

[ii] D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order.* Springer-Verlag, Berlin, 2001.

(i). If T is not continuous, the T is not bounded. Hence, there is $\{x_j\} \subset X$ with $\|x_j\|_X = 1$ such that

$$\|Tx_j\|_Y \geq \|Tx_{j-1}\|_Y + 1.$$

But

$$\|Tx_j\|_Y \leq \|Tx_{j-1}\|_Y + \|Tx_j - Tx_{j-1}\|_Y.$$

This implies that

$$\|Tx_j - Tx_{j-1}\|_Y \geq 1.$$

Therefore, $\{Tx_j\}$ contains no convergent subsequences in Y, contradicting with it being compact.

(ii). Let $\{x_j\} \subset X$ be bounded, then $\{Tx_j\} \subset Y$ is bounded since T is continuous. Hence, $\{Tx_j\}$ contains a convergent subsequence since Y is finite-dimensional. Therefore, T is also compact.

$\square$

Theorem. *Let X be a normed linear space and $T : X \to X$ be a compact linear map. Then T has a countable set of eigenvalues with no limit points except possibly $\lambda = 0$, and each non-zero eigenvalue has finite multiplicity.*

10.5. Conic sections and Kepler problem

A conic section is the intersection of a cone (i.e., a right circular conical surface) with a plane.

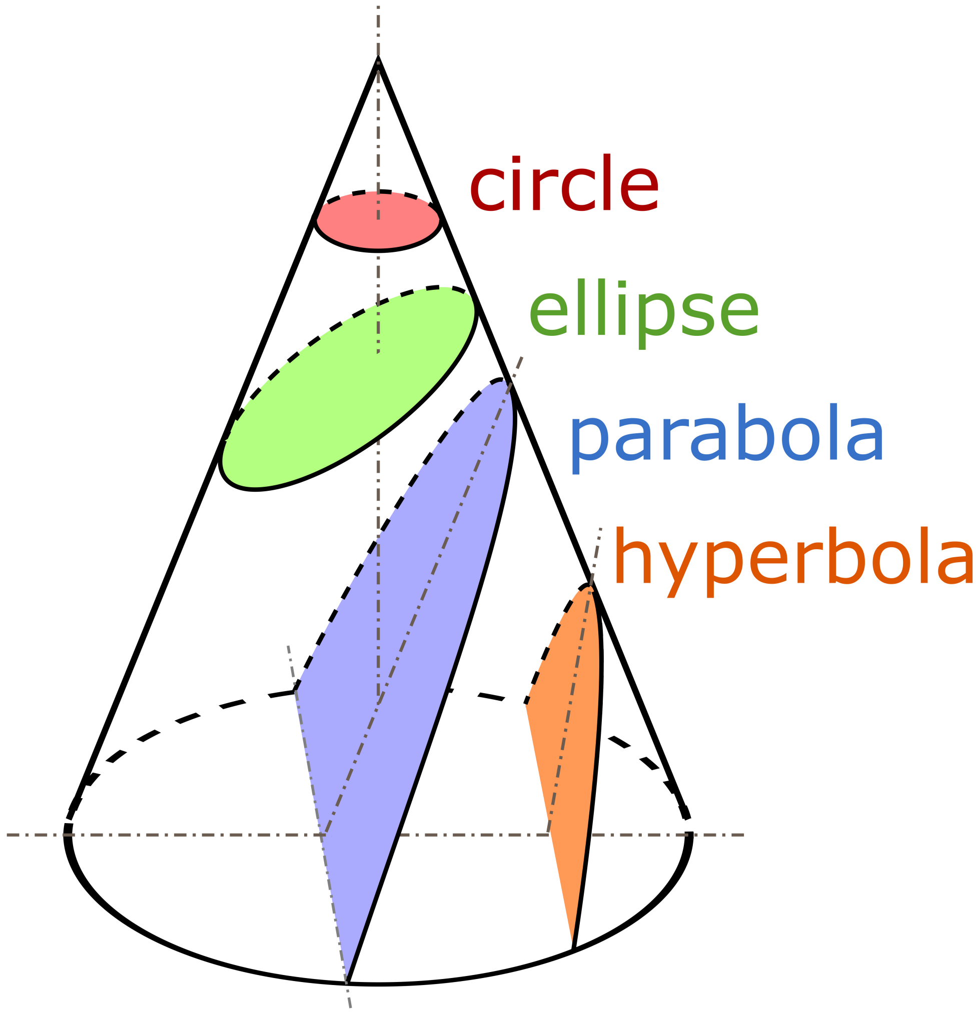

FIGURE 10.1. Conic sections (Credit: Wikipedia)

The three types of conic sections are ellipse, parabola, and hyperbola. In this section, we visit the conic sections via the focus-directrix definition, which is particularly convenient for applications to the Kepler problem, the movement of a planet in the solar system.

Definition (Conic sections). A conic section is the locus of all points P whose distance to a fixed point F (called the focus) is a constant multiple (called the eccentricity e) of the distance from P to a fixed line L (called the directrix). The conic sections can be categorized as follows.

- Ellipses if $0 \le e < 1$, and circle if $e = 0$
- Parabolas if $e = 1$.
- Hyperbolas if $e > 1$.

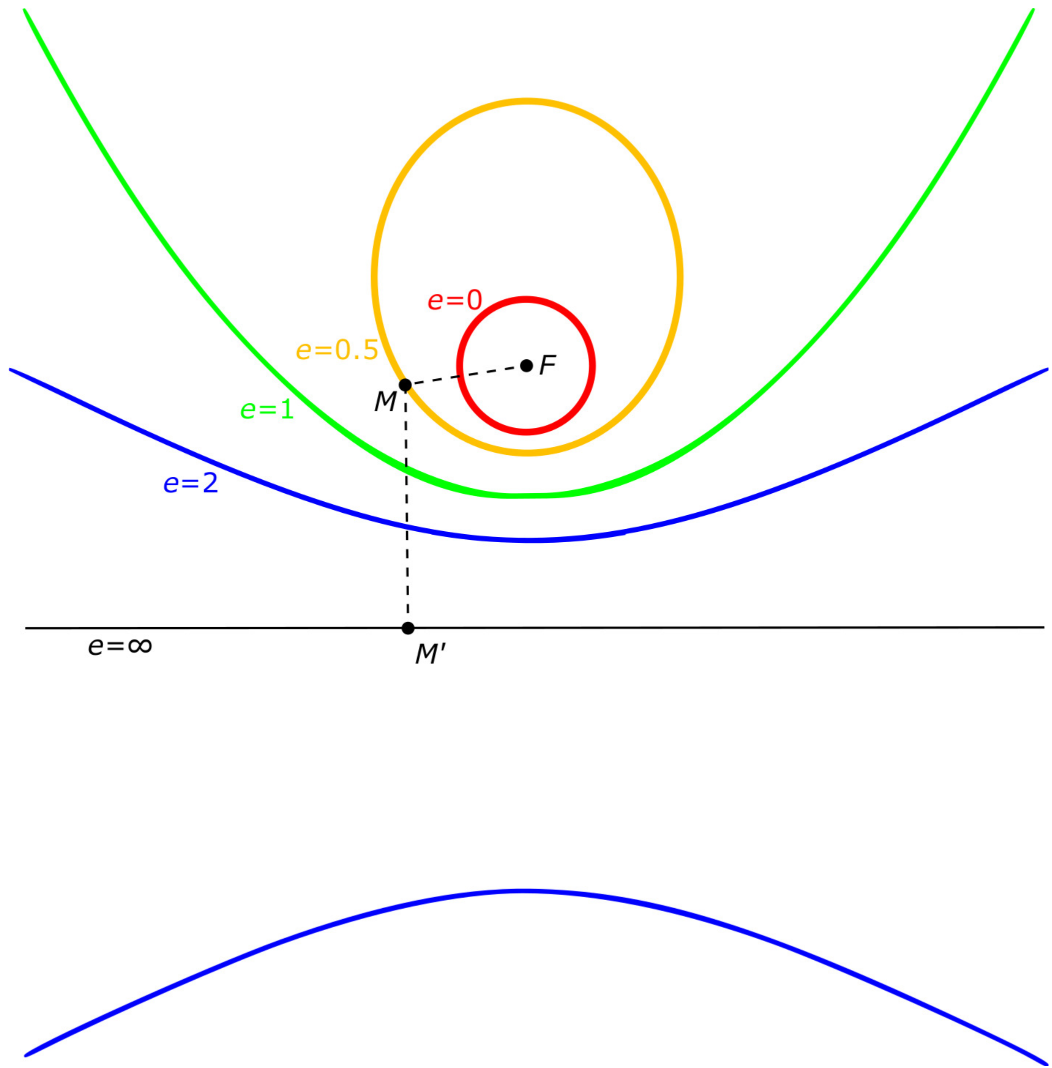

FIGURE 10.2. Conic sections by eccentricity (Credit: Wikipedia)

Let the focus be the origin and the directrix be the line $y = -p$, respectively. Then a point $P = P(r, \theta)$ is on the ellipse with eccentricity e if

$$\frac{r}{r \sin \theta + p} = e,$$

so the equation of the conic sections in polar coordinates can be written as

$$r(\theta) = \frac{ep}{1 - e \sin \theta}.$$

Next we follow Section 8 in Arnold[i] to discuss the Kepler problem, that is, a planet of mass m confined to a plane and moving in a central field of the sun with potential $-k/r$ in polar coordinates (r, θ). The kinetic energy of the planet is

$$T = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right).$$

[i]V. Arnold, *Mathematical methods of classical mechanics.* Springer-Verlag, New York, NY, 1989.

The angular momentum of the particle is $mr^2\dot\theta$. The Legendre transform

$$(r,\theta,\dot r,\dot\theta) \to (r,\theta,p_r,p_\theta), \quad p_r = m\dot r, \quad p_\theta = mr^2\dot\theta.$$

The kinetic energy is

$$T = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right),$$

and the Hamiltonian becomes

$$H = T + V = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \frac{k}{r}.$$

The Hamilton equations read

$$\begin{cases} \dot r = \frac{\partial H}{\partial p_r} = \frac{p_r}{m}, \\ \dot\theta = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2}, \\ \dot p_r = \frac{\partial H}{\partial r} = -\frac{p_\theta^2}{mr^3} + \frac{k}{r^2}, \\ \dot p_\theta = \frac{\partial H}{\partial \theta} = 0. \end{cases}$$

The angular momentum $p_\theta = l$ is conserved. That is, $mr^2\dot\theta = l$, and

$$\frac{d\theta}{dt} = \frac{l}{mr^2}.$$

The total energy is also conserved, i.e., $H = E$. Thus,

$$\frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) + V(r) = \frac{1}{2m}\left(m^2\dot r^2 + \frac{l^2}{r^2}\right) - \frac{k}{r} = E.$$

It then follows that

$$\frac{dr}{dt} = \sqrt{\frac{2E}{m} + \frac{2k}{mr} - \frac{l^2}{mr^2}}.$$

Therefore,

$$\frac{d\theta}{dr} = \frac{l}{mr^2\sqrt{\frac{2E}{m} + \frac{2k}{mr} - \frac{l^2}{m^2r^2}}} = \frac{l}{r^2\sqrt{\frac{2mE}{l^2} + \frac{2mk}{l^2 r} - \frac{1}{r^2}}}.$$

The differential equation can be solved as

$$\begin{aligned} &\theta - \theta_0 \\ = \ & \int_{r_0}^{r} \frac{1}{s^2\sqrt{\frac{2mE}{l^2} + \frac{2mk}{l^2 s} - \frac{1}{s^2}}}\, ds \\ = \ & -\int_{1/r_0}^{1/r} \frac{1}{\sqrt{\frac{2mE}{l^2} + \frac{2mku}{l^2} - u^2}}\, du \quad \text{in which } u = \frac{1}{s}. \end{aligned}$$

Using

$$\int \frac{1}{\sqrt{\alpha + \beta x + \gamma x^2}} = \frac{1}{\sqrt{-\gamma}} \arccos\left(-\frac{\beta + 2\gamma x}{\sqrt{\beta^2 - 4\alpha\gamma}}\right) + c,$$

we have that

$$\theta - \theta_0 = -\arccos\left(\frac{\frac{l^2}{mkr} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}}\right) + \arccos\left(\frac{\frac{l^2}{mkr_0} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}}\right).$$

So

$$\theta = \tilde{\theta}_0 - \arccos\left(\frac{\frac{l^2}{mkr} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}}\right),$$

in which

$$\tilde{\theta}_0 = \theta_0 + \arccos\left(\frac{\frac{l^2}{mkr_0} - 1}{\sqrt{1 + \frac{2El^2}{mk^2}}}\right).$$

Hence, the solution of the motion is

$$r = \frac{\frac{l^2}{mk}}{1 + \sqrt{1 + \frac{2El^2}{mk^2}}\cos(\theta - \tilde{\theta}_0)},$$

which represents a conic section with eccentricity

$$e = \sqrt{1 + \frac{2El^2}{mk^2}}.$$

(i). If $e = 0$, i.e.,

$$E = -\frac{mk^2}{2l^2},$$

then the solution curve is a circle.

(ii). If $0 < e < 1$, i.e.,

$$-\frac{mk^2}{2l^2} < E < 0,$$

then the solution curve is an ellipse.

(iii). If $e = 1$, i.e., $E = 0$, then the solution is a parabola. The planet escapes to infinity with velocity equals zero at infinity.

(iv). If $e > 1$, i.e., $E > 0$, hyperbola. The planet escapes to infinity with positive velocity at infinity.

Theorem (Kepler's laws of planetary motion). *Kepler's laws of planetary motion are three scientific laws describing the motion of planets around the Sun.*

(i). The orbit of every planet is an ellipse with the Sun at one of the two foci.

(ii). A line joining a planet and the Sun sweeps out equal areas during equal intervals of time.

(iii). The square of the orbital period of a planet is directly proportional to the cube of the semi-major axis of its orbit.

Proof.

(i). It follows from the polar equation of the orbit.

(ii). It follows from the fact that $P_\theta = l$ is constant. The line jointing the planet and the Sun sweeps with rate

$$\frac{\vec{v} \cdot \vec{r}}{2} = \frac{r r\dot{\theta}}{2} = \frac{l}{2m}.$$

(iii). Let a, b be the semi-axes. Then the area of the ellipse is πab. The period

$$T = \frac{\pi ab}{\frac{l}{2m}} = \frac{2\pi km}{2|E|\sqrt{2|E|m}} = \frac{2\pi a^{\frac{3}{2}}}{\sqrt{k/m}},$$

that is,

$$\frac{T^2}{a^3} = \frac{4\pi^2 m}{k}.$$

$\square$

Definition (First and second cosmic velocity).

- The first cosmic velocity is the speed at which an object orbit in a circle around the gravitational centre. Hence,
$$\frac{mv_1^2}{r} = \frac{GMm}{r^2}$$
implies that
$$v_1 = \sqrt{\frac{GM}{r}}.$$
Here, G is the universal gravitational constant ($G = 6.67 \times 10^{-11}\,\mathrm{m}^3 \cdot \mathrm{kg}^{-1} \cdot \mathrm{s}^{-2}$), M the mass of the planet, star or other massive body, and r the distance from the center of gravity. In this case, the total energy
$$E = T + V = \frac{1}{2}mv_1^2 - \frac{GMm}{r} = -\frac{GMm}{2r} = \frac{1}{2}V.$$
- The second cosmic velocity, i.e., the escape velocity, is the speed at which the sum of an object's kinetic energy and its gravitational potential energy is equal to zero. It is the speed needed to "break free" from the gravitational attraction of a massive body, without further propulsion, i.e., without spending more fuel. Hence,
$$E = T + V = \frac{1}{2}mv_2^2 - \frac{GMm}{r} = 0$$
implies that
$$v_2 = \sqrt{\frac{2GM}{r}} = \sqrt{2}v_1.$$

Example (First and second cosmic velocity of the earth). We make use of the gravitational acceleration

$$g = \frac{GM}{r^2} = 9.8\,\mathrm{m} \cdot \mathrm{s}^{-2}$$

and the radius of earth

$$r = 6137\,km = 6.137 \times 10^6\,\mathrm{m}.$$

Then

$$v_1 = \sqrt{\frac{GM}{r}} = \sqrt{gr} = 7901\,ms^{-1} \approx 7.9\,\mathrm{km} \cdot \mathrm{s}^{-1}$$

and

$$v_2 = \sqrt{2}v_1 \approx 11.2\,\mathrm{km} \cdot \mathrm{s}^{-1}.$$

10.6. Distributions and measures

Let X be an open set in $\mathbb{R}^n$.

Definition (Distributions). A distribution u in X is a linear form on $C_0^\infty(X)$ such that for every compact set $K \subset X$ there exist constants $C = C(K)$ and $k = k(K)$ such that

$$|u(\phi)| \le C \sum_{|\alpha| \le k} \sup |\partial^\alpha \phi| \quad \text{for all } \phi \in C_0^\infty(K).$$

The set of all distributions in X is denoted by $\mathcal{D}'(X)$. If the same integer k can be used in the above estimate for all K, then we say that u is of order $\le k$, and we denote the set of all such distributions by $\mathcal{D}'^k(X)$.

Remark. Let $k \in \mathbb{N}$ and $u \in \mathcal{D}'^k(X)$. Then u can be extended (in a unique way) to a linear form on $C_0^k(X)$ such that the above estimate remain valid for all $\phi \in C_0^k(K)$ and some constant $C = C(K)$.

Definition (Measures). A measure u in X is a linear form on $C_0^0(X)$ such that for every compact set $K \subset X$ there exist constants C such that

$$|u(\phi)| \le C \sup |\phi| \quad \text{for all } \phi \in C_0^0(K).$$

We say that a measure u is positive if $u(\phi) \ge 0$ for all $\phi \in C_0^0(X)$ such that $\phi \ge 0$.

Remark. Let $u \in \mathcal{D}'^0(X)$. Then u extends to a linear form on $C_0^0(X)$ and therefore defines a measure on X.

Maybe surprisingly, the following theorem states that a distribution of any order is a distribution of order 0 (and therefore defines a measure), as long as it is positive.

Theorem. *Let $u \in \mathcal{D}'(X)$ such that $u(\phi) \ge 0$ for all $\phi \in C_0^\infty(X)$ such that $\phi \ge 0$. Then u is a positive measure.*

Proof. We need to show that u is of order 0. To this end, let $K \subset X$ be compact. Then there exists $\chi \in C_0^\infty(K)$ such that $0 \le \chi \le 1$ and $\chi = 1$ on K. Hence, for all $\phi \in C_0^\infty(K)$,

$$\chi \sup |\phi| \pm \phi \ge 0.$$

It then follows that

$$u\left(\chi \sup |\phi| \pm \phi\right) = u(\chi) \sup |\phi| \pm u(\phi) \ge 0,$$

which implies that

$$|u(\phi)| \le u(\chi) \sup |\phi|.$$

That is, $u \in \mathcal{D}'^0(X)$. □

10.7. Émile Borel and Armand Borel: Both have their theorems!

There are two influential mathematicians with last name "Borel":

- Émile Borel (January 7, 1871 - February 3, 1956) was a French mathematician, who was known for his founding work in the areas of measure theory and probability.
- Armand Borel (May 21, 1923 - August 11, 2003) was a Swiss mathematician, who spent majority of his career at the Institute for Advanced Study in Princeton (1957-1993). He was one of the creators of the theory of linear algebraic groups[i].

[i] A. Borel, *Linear algebraic groups*. Second edition. Springer-Verlag, New York, NY, 1991.

The following "Borel theorem" was due to Émile Borel in 1895, c.f., Theorem 1.2.6 in Hörmander[i].

Theorem (Borel). *For $j \in \mathbb{N}$, let $f_j \in C_0^\infty(K)$, where K is a compact subset of $\mathbb{R}^n$. Let I be a compact neighborhood of 0 in $\mathbb{R}$. Then one can find $f \in C_0^\infty(K \times I)$ such that*

$$\left.\frac{\partial^j f(x,t)}{\partial t^j}\right|_{t=0} = f_j(x) \quad \textit{for all } j \in \mathbb{N}.$$

Proof. Choose $\varphi \in C_0^\infty(\mathbb{R})$ such that $0 \le \varphi \le 1$, $\varphi \equiv 1$ on $I/2$, and $\varphi = 0$ on $\mathbb{R} \setminus (I/2)$. Then

$$\varphi_k(x,t) = \frac{\varphi(t/\varepsilon_k) f_k(x) t^k}{k!}$$

is in $C_0^\infty(K \times I)$. In fact,

$$|\partial^\alpha \varphi_k(x,t)| \le C_{\alpha,j} \varepsilon_k^{k-\alpha_t},$$

where α_t is the order of differentiation with respect to t. Then

$$|\partial^\alpha \varphi_k(x,t)| \le 2^{-k} \quad \text{if } |\alpha| \le k-1,$$

provided that ε_k is sufficiently small. Hence the modified Taylor series

$$f(x,t) = \sum_{k=0}^\infty \varphi_k(x,t)$$

is uniformly convergent. So are the series obtained by differentiation. It follows that $f \in C_0^\infty(K \times I)$ and that

$$\left.\frac{\partial^j f(x,t)}{\partial t^j}\right|_{t=0} = \sum_{k=0}^\infty \left.\frac{\partial^j \varphi_k(x,t)}{\partial t^j}\right|_{t=0} = f_j(x) \quad \text{for all } j \in \mathbb{N}.$$

This is because

- if $k > j$ then $\partial_t^j \varphi_k(x,t) = 0$ since there will be t^{k-j} factor;
- if $k < j$ then $\partial_t^j \varphi_k(x,t) = 0$ since there will be $\partial_t^{j-k} \varphi(t/\varepsilon_k)$ factor;
- if $j = k$ then the only nonzero term is $\varphi(t/\varepsilon_k) f_k(x)$.

$\square$

As a consequence of Borel's theorem, we have that

Theorem. *Given a sequence of numbers $\{a_j\}_{j=0}^\infty \subset \mathbb{R}$, there exists a function $f \in C_0^\infty(I)$ such that $f^{(j)}(0) = a_j$.*

One can naively write

$$f(t) = \sum_{j=0}^\infty \frac{a_j t^j}{j!}$$

but this summation could be divergent as soon as $t \ne 0$. Moreover, it is not in $C_0^\infty(I)$. In order to remedy these issues, we play the same trick as above.

[i] L. Hörmander, *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Second edition. Springer-Verlag, Berlin, 1990.

PROOF. Choose $\varphi \in C_0^\infty(\mathbb{R})$ such that $0 \le \varphi \le 1$, $\varphi \equiv 1$ on $I/2$, and $\varphi = 0$ on $\mathbb{R} \setminus (I/2)$. Then for a sequence of positive numbers $\varepsilon_k \searrow 0$,

$$\varphi_k(t) = \frac{\varphi(t/\varepsilon_k) a_k t^k}{k!}$$

is in $C_0^\infty(I)$. We can sum them to

$$f(t) = \sum_{k=0}^{\infty} \varphi_k(t)$$

because for every $t \in I$, there are only finite terms in the summation. $f \in C_0^\infty(I)$ if ε_1 is small enough. Then one only needs to observe that $\varphi(0) = 1$ and $\varphi^{(i)}(0) = 0$ for all $i \ge 1$. Thus,

$$\frac{d^j}{dt^j}\varphi_k(t)\Big|_{t=0} = \frac{a_k}{k!}\sum_{i=0}^{j} \frac{d^i}{dt^i}\varphi(t)\Big|_{t=0} \frac{d^{j-i}}{dt^{j-i}} t^k\Big|_{t=0} = \delta_{jk} a_k.$$

Therefore, $f^{(j)}(0) = a_j$. □

Remark (Whitney's extension theorem). One can extend Borel's theorem to higher dimensions. In particular, Whitney's extension theorem states that given prescribed continuous function $f_\alpha(x)$ for $|\alpha| \le k$ and x in a compact set, then one find a function f such that $\partial^\alpha f(x) = f_\alpha(x)$. See Theorem 2.3.6 in Hörmander.

10.8. Fourier dimension

10.9. Frobenius theorem and foliation

This section follows Appendix C in Hörmander[i]. If

$$v = \sum_{j=1}^{n} v_j(x) \frac{\partial}{\partial x_j}$$

is a smooth vector field in a neighborhood of 0 in $\mathbb{R}^n$ with $v(0) \ne 0$, then one can introduce new local coordinates $y_1, ..., y_n$ such that $v = \partial/\partial y_1$ in a neighborhood of 0. Frobenius theorem states a similar result for several vector fields which satisfy certain commutation condition.

Theorem (Frobenius). *Let $v_1, ..., v_r$ be smooth vector fields in a neighborhood of 0 in $\mathbb{R}^n$. Suppose that*

(i). $v_1(0), ..., v_r(0)$ are linearly independent,

(ii).

$$[v_i, v_j] = \sum_{k=1}^{r} c_{ij}^k v_k,$$

in other words, the commutators $[v_i, v_j]$ must lie in the linear span of the $v_1, ..., v_r$ at every point.

Then there exists new local coordinates $y_1, ..., y_n$ such that

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^{r} b_{ij} v_j, \quad i = 1, ..., r,$$

in which (b_{ij}) is an invertible matrix.

[i]L. Hörmander, *The analysis of linear partial differential operators. III. Pseudodifferential operators.* Springer-Verlag, Berlin, 1985.

Frobenius theorem gives necessary and sufficient conditions for finding a maximal set of independent solutions of an underdetermined system of first-order homogeneous linear partial differential equations.

Theorem. *Assume that $v_1, ..., v_r$ satisfy the hypotheses of the above theorem. Let $f_1, ..., f_r \in C^\infty(\mathbb{R}^n)$. Then the equations*

$$v_j(u) = f_j, \quad j = 1, ..., r$$

have a smooth solution in a neighborhood of 0 iff

$$v_i(f_j) - v_j(f_i) = \sum_{k=1}^{r} c_{ij}^k f_k, \quad i, j = 1, ..., r.$$

If S is a smooth manifold of co-dimension k with tangent plane at 0 supplementary to the plane spanned by $v_1, ..., v_r$, then there is a unique solution in a neighborhood of 0 with prescribed restriction $u_0 \in C^\infty(S)$ to S.

The coordinates y in Frobenius theorem are of course not uniquely determined. Suppose that $\tilde{y}_1, ..., \tilde{y}_n$ are other coordinates which enjoy the same properties, that is,

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^{r} b_{ij} v_j \quad \text{and} \quad \frac{\partial}{\partial \tilde{y}_i} = \sum_{j=1}^{r} \tilde{b}_{ij} v_j, \quad i = 1, ..., r,$$

in which (b_{ij}) and $(\tilde{b}_{ij})$ are an invertible matrices. Then

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^{r} \frac{\partial \tilde{y}_j}{\partial y_i} \frac{\partial}{\partial \tilde{y}_j}, \quad i = 1, ..., r.$$

But

$$\frac{\partial}{\partial y_i} = \sum_{j=1}^{n} \frac{\partial \tilde{y}_j}{\partial y_i} \frac{\partial}{\partial \tilde{y}_j}, \quad i = 1, ..., n.$$

Hence,

$$\frac{\partial \tilde{y}_j}{\partial y_i} = 0 \quad \text{when } i \le r \text{ and } j > r.$$

That is, $(\tilde{y}_{r+1}, ..., \tilde{y}_n)$ is locally a function of $(y_{r+1}, ..., y_n)$.

Definition (Foliation). A smooth manifold $\mathbb{M}$ is said to be foliated of dimension r if it has a distinguished atlas $\mathcal{F}$ such that for the map

$$\phi = \kappa' \circ \kappa^{-1} : \kappa(X_\kappa \cap X_{\kappa'}) \to \kappa'(X_\kappa \cap X_{\kappa'})$$

the components $\phi_{r+1}, ..., \phi_n$ are functions of $x_{r+1}, ..., x_n$ only.

A foliation in $\mathbb{M}$ defines for every $x \in \mathbb{M}$ a plane $\sigma_x \subset T_x\mathbb{M}$ of dimension r by equations

$$dx_{r+1} = \cdots = dx_n = 0.$$

By Frobenius theorem, the additional property of a foliation is that, in a neighborhood of any point, σ_x has a basis $v_1(x), ..., v_r(x)$ such that $v_1, ..., v_r$ are smooth vector fields satisfying the condition that the commutators $[v_i, v_j]$ must lie in the linear span of the $v_1, ..., v_r$ at every point.

Theorem. *Let $\mathbb{M}$ be foliated of dimension r. Then at every point $x_0 \in \mathbb{M}$, there is locally a unique r-dimensional manifold $\Sigma \subset \mathbb{M}$, called the leaf of the foliation throughout x_0, such that $T_x\Sigma = \sigma_x$ for every $x \in \Sigma$.*

With local coordinates in the distinguished atlas, σ is defined by

$$x_{r+1} = c_{r+1}, ..., x_n = c_n.$$

Alternatively it is the set of points which can be reached from x_0 by moving along integral curves of the vector fields $v_1, ..., v_r$ inside a suitable neighborhood of x_0.

Definition (Fiber bundles). A fiber bundle is a structure (E, B, π, F), where E, B, and F are topological spaces and $\pi : E \to B$ is a continuous surjection satisfying a local triviality condition outlined below. The space B is called the base space of the bundle, E the total space, and F the fiber. The map π is called the projection map (or bundle projection). We shall assume in what follows that the base space B is connected.

We require that for every x in E, there is an open neighborhood $U \subset B$ of $\pi(x)$ (which will be called a trivializing neighborhood) such that there is a homeomorphism $\phi : \pi^{-1}(U) \to U \times F$ (where $U \times F$ is the product space) in such a way that π agrees with the projection to the first factor. The set of all $\{(U_i, \phi_i)\}$ is called a local trivialization of the bundle.

Thus for any p in B, the preimage $\pi^{-1}(p)$ is homeomorphic to F and is called the fiber over p. Every fiber bundle $\pi : E \to B$ is an open map, since projections of products are open maps. Therefore B carries the quotient topology determined by the map π.

Every fiber bundle is locally a product space. However, it is not a trivial fiber bundle, i.e., is not globally a product space.

Remark (Fiber bundles and foliation). Let (E, B, π, F) be a fiber bundle, in which E is the total space, B is the base space, and F is the fiber. It yields a foliation by fibers $F_b := \pi^{-1}(\{b\})$, $b \in B$. Its space of leaves is homeomorphic to B.

10.10. Germ

Definition (Germ). Let X be a topological space and Y be a set.

- Two maps $f, g : X \to Y$ define the same (map-)germ at $x \in X$ if there is a neighbourhood U of x such that $f(y) = g(y)$ for all $y \in U$.
- Two subsets $S, T \subset X$ define the same (set-)germ at x if there is a neighbourhood U of x such that $S \cap U = T \cap U$.

It is straightforward to see that defining the same germ at x is an equivalence relation (maps or sets), and the equivalence classes are called germs (map-germs or set-germs). The equivalence relation is usually written

$$f \sim_x g \quad \text{or} \quad S \sim_x T.$$

Given a map $f : X \to Y$, then its germ at x is usually denoted $[f]_x$. Similarly, the germ at x of a set S is written $[S]_x$. That is,

$$[f]_x = \{g : X \to Y : g \sim_x f\} \quad \text{and} \quad [S]_x = \{T \subset X : T \sim_x S\}.$$

Remark.

- Notice that two sets are germ-equivalent at x iff their characteristic functions are germ-equivalent at x, i.e., $S \sim_x T$ iff $\chi_S \sim_x \chi_T$.
- To define the map-germs, we do not need $f, g : X \to Y$. In fact, if f has domain S and g has domain T, where $S, T \subset X$, then f and g are germ equivalent at $x \in X$ if
  - $S \sim_x T$. Hence, there is a neighborhood U of x such that say $S \cap U = T \cap U$;
  - $f|_{S \cap V} = g|_{T \cap V}$, for some smaller neighbourhood V with $x \in V \subset U$.

The germ can be used to define the objects of local nature in a topological space. We use the tangent space as an example. Recall the definition of the tangent space of a smooth manifold $\mathbb{M}$ at a point p. (See Appendix in Han[i].)

(i). $C^\infty(\mathbb{M})$: the set of smooth functions,
(ii). $\mathcal{D}^1(\mathbb{M})$: the set of vector fields on $\mathbb{M}$ as derivations of $C^\infty(\mathbb{M})$,
(iii). $\mathcal{D}^1(p)$ or $T_p\mathbb{M}$: the tangent space to $\mathbb{M}$ at p as the set of all linear maps $X_p : f \to (Xf)(p)$ for $X \in \mathcal{D}^1(\mathbb{M})$.

The following alternative definition of the tangent spaces is from Chapter 3 in Lee[ii].

Example (Tangent vectors as derivations of the space of germs). Let $f \in C^\infty(\mathbb{M})$ and $p \in \mathbb{M}$. Denote the germ of f at p as $[f]_p$. The set of all germs of smooth functions at p is denoted by $C_p^\infty(\mathbb{M})$. It is a real vector space. A derivation of $C_p^\infty(\mathbb{M})$ is a linear map $v : C_p^\infty(\mathbb{M}) \to \mathbb{R}$ satisfying

$$v[fg]_p = v[f]_p g(p) + f(p)v[g]_p.$$

Then one can define the tangent space to $\mathbb{M}$ at p as the vector space of derivations of $C_p^\infty(\mathbb{M})$. This definition is naturally isomorphic to the tangent space as above, since if $f, g \in C^\infty(\mathbb{M})$ agree on some neighborhood of p, then $v(f) = v(g)$ for $v \in T_p\mathbb{M}$.

The germ definition has a number of advantages. One of the most significant is that it makes the local nature of the tangent space clearer, without requiring the use of bump functions. Because there is no analytic bump functions, the germ definition of tangent vectors is the only one available on real-analytic or complex-analytic manifolds.

Hörmander made an important observation of the equivalence of two phase functions defining the same Lagrangian manifold, see Corollary 3.1.8 in Hörmander[iii]: Let ϕ_j be a non-degenerate phase function at (x_0, θ_j), $j = 1, 2$, in which

$$\partial_\theta \phi_1(x_0, \theta_1) = 0 \quad \text{and} \quad \partial_\theta \phi_2(x_0, \theta_2) = 0,$$

and

$$\partial_\theta^2 \phi_1(x_0, \theta_1) = 0 \quad \text{and} \quad \partial_\theta^2 \phi_2(x_0, \theta_2) = 0.$$

Then it follows that ϕ_1 and ϕ_2 are equivalent at (x_0, θ_1) and (x_0, θ_2) iff the corresponding germs of Lagrange manifolds are the same.

Question 10.2. *Why "germ" here?*

Answer (Hassell). The Lagrangian manifold is locally defined by

$$C_1 = \{(x, \theta_1) : \partial_\theta \phi_1(x, \theta_1) = 0\} \to \{(x, \xi) : \xi = \partial_x \phi_1(x, \theta_1)\} \subset T^*\mathbb{M},$$

and

$$C_2 = \{(x, \theta_2) : \partial_\theta \phi_2(x, \theta_2) = 0\} \to \{(x, \xi) : \xi = \partial_x \phi_2(x, \theta_2)\} \subset T^*\mathbb{M}.$$

The corresponding germs of Lagrange manifolds defined by phase functions ϕ_1 and ϕ_2 are the same means that $\partial_x \phi_1(x, \theta_1) = \partial_x \phi_2(x, \theta_1)$ (the leading term) at a neighborhood of (x_0, ξ_0), that is, $\partial_x \phi_1(x, \theta_1)$ and $\partial_x \phi_2(x, \theta_1)$ define the same germ at $(x_0, \xi_0) \in T^*\mathbb{M}$.

[i] X. Han, *Classical ergodicity.* Lecture notes.
[ii] J. Lee, *Introduction to smooth manifolds.* Second edition. Springer, New York, NY, 2013.
[iii] L. Hörmander, *Fourier integral operators.* I. Acta Math. 127 (1971), no. 1-2, 79–183.

10.11. Hairy ball theorem

Theorem (Hairy ball theorem). *An even-dimensional sphere does not possess any non-zero continuously differentiable field of unit tangent vectors.*

Remark. The hypothesis that the dimension is even is essential. On $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ when $n-1$ is odd, then the formula

$$v(x_1, ..., x_n) = (x_2, -x_1, ..., x_n, -x_{n-1})$$

defines a differential field of unit tangent vectors on $\mathbb{S}^{n-1}$. First, $v \cdot x = 0$, which indicates that v is orthogonal to $x = (x_1, ..., x_n) \in \mathbb{S}^{n-1}$ and is therefore tangent to $\mathbb{S}^{n-1}$. Second, it is evident that v has unit length. However, such construction is impossible on even-dimensional spheres.

A direct consequence of hairy ball theorem is that there is no non-zero continuously differentiable tangent vector field on even-dimensional spheres, that is, "it is impossible to comb down the hair on one's head without admitting a part or an eddy", see Section 3.2 in Gutzwiller[i]. There are several different proofs of the hairy ball theorem. Here we follow the one by Milnor[ii].

Proof.

(1). Let $A \Subset \mathbb{R}^n$ and $x \to v(x)$ be a continuously differentiable vector field which is defined in a neighborhood of A. For $t \in \mathbb{R}$, consider the function

$$f_t(x) = x + tv(x) \quad \text{for } x \in A.$$

There exists a Lipschitz constant c such that

$$|v(x) - v(y)| \leq c|x - y|, \quad \text{for all } x, y \in A.$$

Choose any t such that $|t| < c^{-1}$. Then $f_t(x) = f_t(y)$ implies that $x - y = t(v(y) - v(x))$ so

$$|x - y| \leq c|t||x - y| < |x - y|.$$

Hence, $x = y$. This means that $f_t : A \to f_t(A)$ is one-to-one when t is sufficiently small.

(2). We compute the volume $f_t(A)$ when t is sufficiently small. Since f_t is one-to-one,

$$\mathrm{Vol}(f_t(A)) = \int_{f_t(A)} dx = \int_A \left| \frac{\partial f_t(x)}{\partial x} \right| dx,$$

in which $f_t(x) = x + tv(x)$ and

$$\frac{\partial f_t(x)}{\partial x} = \mathrm{I} + t \frac{\partial v_i}{\partial x_j}.$$

Hence, the determine is strictly positive if $|t|$ sufficiently small. Then

$$\mathrm{Vol}(f_t(A)) = \int_{f_t(A)} dx = a_0 + a_1 t + \cdots + a_n t^n,$$

in which $a_0 = \mathrm{Vol}(A)$.

(3). Suppose that the sphere $\mathbb{S}^{n-1}$ has a continuously differentiable field $x \to v(x)$ of unit tangent vectors. For $t \in \mathbb{R}$, notice that

$$|x + tv(x)| = \sqrt{1 + t^2}.$$

Let

$$A = \{x \in \mathbb{R}^n : a \leq |x| \leq b\}.$$

[i] M. Gutzwiller, *Chaos in classical and quantum mechanics.* Springer-Verlag, New York, NY, 1990.

[ii] J. Milnor, *Analytic proofs of the "hairy ball theorem" and the Brouwer fixed-point theorem.* Amer. Math. Monthly 85 (1978), no. 7, 521–524.

We extend the vector field v throughout the region A by setting $v(rx) = rv(x)$ for $a \le r \le b$. Recall that $f_t(x) = x + tv(x)$. Then $f_t(\mathbb{S}^{n-1}) \subset \mathbb{S}^{n-1}(0, \sqrt{1+t^2})$ and is closed since f_t is continuous and $\mathbb{S}^{n-1}$ is closed. But $f_t(\mathbb{S}^{n-1})$ is also relatively open in $\mathbb{S}^{n-1}(0, \sqrt{1+t^2})$ since f_t has a continuous inverse when $|t|$ is sufficiently small (the determine $\partial f_t(x)/\partial x$ is non-singular when $|t|$ is small). Because $\mathbb{S}^{n-1}(0, \sqrt{1+t^2})$ is complete, we have that

$$f_t(\mathbb{S}^{n-1}) = \mathbb{S}^{n-1}(0, \sqrt{1+t^2}).$$

That is, $f_t : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}(0, \sqrt{1+t^2})$ is bijective. Therefore,

$$f_t : A \to \{x \in \mathbb{R}^n : a\sqrt{1+t^2} \le |x| \le b\sqrt{1+t^2}\}$$

is also bijective. Evidently,

$$\mathrm{Vol}(f_t(A)) = \left(\sqrt{1+t^2}\right)^n \mathrm{Vol}(A).$$

But from (2),

$$\mathrm{Vol}(f_t(A)) = \int_{f_t(A)} dx = a_0 + a_1 t + \cdots + a_n t^n,$$

this is impossible if n is odd.

□

The hairy ball theorem can be strengthened to

Theorem. *An even-dimensional sphere does not possess any continuous field of non-zero tangent vectors.*

Proof. Suppose that $\mathbb{S}^{n-1}$ possesses a continuous field of non-zero tangent vectors $v(x)$. Let

$$0 < m = \min_{x \in \mathbb{S}^{n-1}} |v(x)|.$$

By Stone-Weierstrass approximation theorem (c.f., Section 7.26 in Rudin[i]), there exists a polynomial map $p : \mathbb{S}^{n-1} \to \mathbb{R}^n$ such that

$$|p(x) - v(x)| < \frac{m}{2} \quad \text{for all } x \in \mathbb{S}^{n-1}.$$

Hence, triangle inequality implies that

$$|p(x)| > |v(x)| - |p(x) - v(x)| > \frac{m}{2}.$$

Define a differentiable vector field

$$u(x) = p(x) - (p \cdot x)x.$$

Then $u \cdot x = 0$ shows that $u(x)$ is tangent to $\mathbb{S}^{n-1}$. While

$$|u - p| = |p \cdot x| \le |(v - (v - p)) \cdot x| \le |v - p| < \frac{m}{2}.$$

Using triangle inequality again,

$$|u| \ge |p| - |p - u| > \frac{m}{2} - \frac{m}{2} > 0.$$

Now $u/|u|$ is a smooth field of unit tangent vectors on $\mathbb{S}^{n-1}$. If $n-1$ is even, this is impossible by the hairy ball theorem. □

[i] W. Rudin, *Principles of mathematical analysis.* Third edition. McGraw-Hill Book Co., New York-Auckland-Düsseldorf, 1976.

10.12. Heat propagator

In this section, we visit some basic properties of the heat propagator on compact manifolds. See Chavel[i] for details.

On a compact Riemannian manifold $(\mathbb{M}, g)$ without boundary, the heat equation is given by

$$\begin{cases} \partial_t u(x,t) + \Delta u(x,t) = g(x,t) & \text{for } (x,t) \in \mathbb{M} \times (0,\infty), \\ u(x,0) = f(x) & \text{for } x \in \mathbb{M}. \end{cases}$$

Here, $\Delta = \Delta_g$ is the positive Laplacian on $\mathbb{M}$ determined by the Riemannian metric g. A classical solution $u(x,t)$ to the heat equation shall satisfy that $u(x,t) \in C(\mathbb{M} \times (0,\infty))$, $u(x,t) \in C^2(\mathbb{M})$ for any fixed $t \in (0,\infty)$, and $u(x,t) \in C^1((0,\infty))$ for any fixed $x \in \mathbb{M}$, that is, u is C^2 with respect to $x \in \mathbb{M}$ and C^1 with respect to $t \in (0,\infty)$.

The homogeneous heat equation is

$$\begin{cases} \partial_t u(x,t) + \Delta u(x,t) = 0 & \text{for } (x,t) \in \mathbb{M} \times (0,\infty), \\ u(x,0) = f(x) & \text{for } x \in \mathbb{M}. \end{cases}$$

Theorem. *Let $u(x,t)$ be a classical solution to the homogeneous heat equation. Then $\|u(\cdot,t)\|_{L^2(\mathbb{M})}$ is decreasing in $t \in (0,\infty)$.*

This theorem immediately implies that the classical solution to the heat equation is unique.

Definition (Fundamental solution of the heat equation). We say that a function $\Phi(x,y,t) \in C(\mathbb{M} \times \mathbb{M} \times (0,\infty))$ which is C^2 with respect to $x \in \mathbb{M}$ and C^1 with respect to $t \in (0,\infty)$ is the fundamental solution of the heat equation if

$$\begin{cases} \partial_t \Phi(x,y,t) + \Delta_x \Phi(x,y,t) = 0 & \text{for } (x,t) \in \mathbb{M} \times (0,\infty), \\ \lim_{t\to 0} \Phi(\cdot,y,t) = \delta_y & \text{for } y \in \mathbb{M}. \end{cases}$$

In the first equation, Δ_x acts on the x variable; while the second equation means that

$$\lim_{t\to 0} \int_{\mathbb{M}} \Phi(x,y,t) f(x)\, dx = f(y) \quad \text{for all } f \in C(\mathbb{M}).$$

Example.

(1). In $\mathbb{R}^n$,

$$\Phi(x,y,t) = \begin{cases} \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{4t}}, & \text{if } t > 0, \\ 0, & \text{if } t < 0. \end{cases}$$

(2). On $\mathbb{M}$, the fundamental solution is a perturbation of

$$\Phi(x,y,t) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{d(x,y)^2}{4t}}.$$

Here, $d(x,y)$ is the Riemannian distance of x and y on $\mathbb{M}$.

One can show that the fundamental solution $\Phi(x,y,t)$ is symmetric (i.e., $\Phi(x,y,t) = \Phi(y,x,t)$) and unique. By definition of the fundamental solution, $\Phi(x,y,t)$ for any fixed $t > 0$ is C^2 with respect to x and is therefore C^2 with respect to y. Since $\mathbb{M}$ is compact, $\Phi(x,y,t)$ is uniformly bounded for $(x,y) \in \mathbb{M} \times \mathbb{M}$ and any fixed $t > 0$.

---

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

By Duhamel's principle, the classical solution to the heat equation represented by the fundamental solution:

$$u(x,t) = \int_{\mathbb{M}} \Phi(x,y,t) f(y)\, dy + \int_0^t \int_{\mathbb{M}} \Phi(x,y,s) g(y,t-s)\, dy,$$

in which $f \in C(\mathbb{M})$ and $f \in C(\mathbb{M} \times (0,\infty))$.

In particular, for $f \in C(\mathbb{M})$, the function

$$u(x,t) = \int_{\mathbb{M}} \Phi(x,y,t) f(y)\, dy := e^{-t\Delta} f(x)$$

solves the homogeneous heat equation, in which we call $e^{-t\Delta}$ the heat propagator. Notice that

$$\left\| e^{-t\Delta} f \right\|_{L^2(\mathbb{M})} \le \|f\|_{L^2(\mathbb{M})} \quad \text{for } f \in C(\mathbb{M}).$$

One can then extend the domain of the heat propagator from continuous functions to L^2 functions. That is, since continuous functions are dense in $L^2(\mathbb{M})$, for any $f \in L^2(\mathbb{M})$, there is a sequence of continuous functions $f_k \to f$ in L^2 norm (i.e., $\|f_k - f\|_{L^2(\mathbb{M})} \to 0$ as $k \to \infty$). Hence,

$$\left\| e^{-t\Delta} f_k - e^{-t\Delta} f_j \right\|_{L^2(\mathbb{M})} \le \|f_k - f_j\|_{L^2(\mathbb{M})} \to 0 \quad \text{as } k, j \to \infty.$$

so $\{e^{-t\Delta} f_k\}$ is Cauchy in $L^2(\mathbb{M})$ and has a limit because $L^2(\mathbb{M})$ is complete. We define

$$\int_{\mathbb{M}} \Phi(x,y,t) f(y)\, dy = e^{-t\Delta} f(x) = \lim_{k\to\infty} e^{-t\Delta} f_k(x) \quad \text{for } f \in L^2(\mathbb{M}).$$

Therefore, the heat propagator is now a mapping from $L^2(\mathbb{M})$ to $L^2(\mathbb{M})$.

Theorem. *The heat propagator $e^{-t\Delta} : L^2(\mathbb{M}) \to L^2(\mathbb{M})$ satisfies the following conditions.*

(i). For fixed $t > 0$, $e^{-t\Delta} f$ is continuous.

(ii). As $t \to 0$, $e^{-t\Delta} f$ converges to f in $L^2(\mathbb{M})$.

(iii). As $t \to \infty$, $e^{-t\Delta} f$ converges to a harmonic function.

PROOF.

(i). Let $f \in L^2(\mathbb{M})$ and fix $t > 0$. Compute that

$$\begin{aligned}
\left| e^{-t\Delta} f(x) - e^{-t\Delta} f(y) \right| &= \left| \int_{\mathbb{M}} \Phi(x,z,t) f(z)\, dz - \int_{\mathbb{M}} \Phi(y,z,t) f(z)\, dz \right| \\
&\le \int_{\mathbb{M}} |\Phi(x,z,t) - \Phi(y,z,t)| |f(z)|\, dz \\
&\le \left(\int_{\mathbb{M}} |\Phi(x,z,t) - \Phi(y,z,t)|^2\, dz \right)^{\frac{1}{2}} \left(\int_{\mathbb{M}} |f(z)|^2\, dz \right)^{\frac{1}{2}} \\
&= \left(\int_{\mathbb{M}} |\Phi(x,z,t) - \Phi(y,z,t)|^2\, dz \right)^{\frac{1}{2}} \|f\|_{L^2(\mathbb{M})} \\
&\to 0 \quad \text{as } d(x,y) \to 0.
\end{aligned}$$

Here, we use the fact that $\Phi(x,y,t)$ is uniformly bounded on $\mathbb{M} \times \mathbb{M}$ for fixed $t > 0$. Hence, for fixed $t > 0$, $e^{-t\Delta} f$ is continuous.

(ii). For $f \in C(\mathbb{M})$ and $t \to 0$, by the definition of the fundamental solution

$$\lim_{t\to 0} e^{-t\Delta} f = f \quad \text{in } C(\mathbb{M}).$$

Therefore,

$$\lim_{t\to 0} \left\| e^{-t\Delta} f - f \right\|_{L^2(\mathbb{M})} = 0 \quad \text{for all } f \in C(\mathbb{M}).$$

For $f \in L^2(\mathbb{M})$, choose $g \in C(\mathbb{M})$ such that $\|f - g\|_{L^2(\mathbb{M})} < \varepsilon$. This implies that
$\left\|e^{-t\Delta} f - f\right\|_{L^2(\mathbb{M})} \le \left\|e^{-t\Delta} f - e^{-t\Delta} g\right\|_{L^2(\mathbb{M})} + \left\|e^{-t\Delta} g - g\right\|_{L^2(\mathbb{M})} + \|g - f\|_{L^2(\mathbb{M})} \le 3\varepsilon$,
which means that

$$\lim_{t\to 0} \left\|e^{-t\Delta} f - f\right\|_{L^2(\mathbb{M})} = 0 \quad \text{for all } f \in L^2(\mathbb{M}).$$

(iii). For $f \in L^2(\mathbb{M})$ and $t \to \infty$, $\|e^{-t\Delta} f\|_{L^2(\mathbb{M})}$ is decreasing so converges to a number $\alpha \ge 0$. Hence, as $t, s \to \infty$, we have that

$$\begin{aligned} \left\|e^{-t\Delta} f - e^{-s\Delta} f\right\|^2_{L^2(\mathbb{M})} &= \left\|e^{-t\Delta} f\right\|^2_{L^2(\mathbb{M})} + \left\|e^{-s\Delta} f\right\|^2_{L^2(\mathbb{M})} - 2\left\|e^{-\frac{t+s}{2}\Delta} f\right\|^2_{L^2(\mathbb{M})} \\ &\to \alpha^2 + \alpha^2 - 2\alpha^2 \\ &= 0. \end{aligned}$$

This implies that there exists $h \in L^2(\mathbb{M})$ such that as $t \to \infty$,

$$e^{-t\Delta} f \to h \quad \text{in } L^2(\mathbb{M}).$$

In fact, we show that as $t \to \infty$,

$$e^{-t\Delta} f(x) \to h(x) \quad \text{uniformly for } x \in \mathbb{M},$$

which implies that h is continuous. Indeed, for some $T > 0$,

$$\begin{aligned} & \left|e^{-(T+t)\Delta} f(x) - e^{-(T+s)\Delta} f(x)\right| \\ = & \left|e^{-T\Delta}\left(e^{-t\Delta} f(x) - e^{-s\Delta} f(x)\right)\right| \\ = & \left|\int_{\mathbb{M}} \Phi(x, y, T)\left(e^{-t\Delta} f(y) - e^{-s\Delta} f(y)\right) dy\right| \\ \le & \left(\int_{\mathbb{M}} |\Phi(x, y, T)|^2 \, dy\right)^{\frac{1}{2}} \left(\int_{\mathbb{M}} \left|e^{-t\Delta} f(y) - e^{-s\Delta} f(y)\right|^2 \, dy\right)^{\frac{1}{2}} \\ = & \left(\int_{\mathbb{M}} |\Phi(x, y, T)|^2 \, dy\right)^{\frac{1}{2}} \left\|e^{-t\Delta} f - e^{-s\Delta} f\right\|_{L^2(\mathbb{M})} \to 0 \end{aligned}$$

uniformly as $s, t \to \infty$. (Here, we impose $T > 0$ because we need the fundamental solution $\Phi(x, y, T)$ to be at some fixed time $T > 0$ so it is uniformly bounded.)

Now using the fact that $e^{-T\Delta} h(x) = h(x)$ solves the homogeneous heat equation

$$\partial_T h + \Delta h = 0,$$

we see that $\Delta h = 0$ since h is independent of T.

□

10.13. Khinchin inequality and Hausdorff-Young inequality

Khinchin inequality, in one of its most well-known forms, states that the expectation of the l^p $(0 < p < \infty)$ norm of a random sequence of numbers is comparable with its l^2 norm. It provides a nice consequence on the invalidity of Hausdorff-Young inequality outside of the allowed range of exponents. In this section, we visit these results following Chapter 4 in Wolff[i].

[i] T. Wolff, *Lectures on harmonic analysis*. American Mathematical Society, Providence, RI, 2003.

Theorem (Khinchin). *Let $\{\epsilon_j\}_{j=1}^N$ be i.i.d. (independent and identically distributed) random variables with $P(\epsilon_j = \pm 1) = \frac{1}{2}$ for every $j = 1, ..., N$, i.e., a sequence with Rademacher distribution. Let $0 < p < \infty$ and $a_1, ..., a_N \in \mathbb{C}$. Then*

$$A \left(\sum_{j=1}^N |a_j|^2 \right)^{\frac{1}{2}} \leq \left(\mathbb{E} \Big| \sum_{k=1}^N \epsilon_j a_j \Big|^p \right)^{\frac{1}{p}} \leq B \left(\sum_{j=1}^N |a_j|^2 \right)^{\frac{1}{2}},$$

in which $A = A(p), B = B(p) > 0$.

The Hausdorff-Young inequality asserts that

$$\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)} \lesssim \|f\|_{L^p(\mathbb{R}^n)} \quad \text{for all } 1 \leq p \leq 2,$$

in which $\hat{f}$ stands for the Fourier transform of $f \in C_0^\infty(\mathbb{R}^n)$. It follows from an interpolation between the case of $p = 1$ and $p = 2$ (the Plancheral identity).

Next we use Khinchin's inequality to show that the Hausdorff-Young inequality can not hold if $p > 2$. Let $\phi \in C_0^\infty(\mathbb{R}^d)$ and $\operatorname{supp} \phi \subset (0,1)^n$. Then $\phi_j(x) = \phi(x - k_j)$ have disjoint supports for $k_j \in \mathbb{Z}^d$. Hence,

$$\left\| \sum_{j=1}^N \epsilon_j \phi_j \right\|_{L^p(\mathbb{R}^d)} \approx N^{\frac{1}{p}}.$$

Notice that $\hat{\phi}_j(\xi) = e^{ik_j \cdot \xi} \hat{\phi}(\xi)$. Using Fubini's theorem, we have that

$$\begin{aligned}
\mathbb{E} \left(\left\| \sum_{j=1}^N \epsilon_j \hat{\phi}_j \right\|_{L^{p'}(\mathbb{R}^n)}^{p'} \right) &= \mathbb{E} \left(\left\| \sum_{j=1}^N \epsilon_j e^{ik_j \cdot \xi} \hat{\phi} \right\|_{L^{p'}(\mathbb{R}^n)}^{p'} \right) \\
&= \frac{1}{2^N} \sum_{\{\epsilon_n\}} \int_{\mathbb{R}^n} \left| \sum_{n=1}^N \epsilon_n e^{ik_j \cdot \xi} \right|^{p'} \left| \hat{\phi}(\xi) \right|^{p'} d\xi \\
&= \int_{\mathbb{R}^n} \sum_{\{\epsilon_j\}} \frac{1}{2^N} \left| \sum_{n=1}^N \epsilon_n e^{ik_j \cdot \xi} \right|^{p'} \left| \hat{\phi}(\xi) \right|^{p'} d\xi \\
&= \int_{\mathbb{R}^n} \mathbb{E} \left(\left| \sum_{j=1}^N \epsilon_n e^{ik_j \cdot \xi} \right|^{p'} \right) \left| \hat{\phi}(\xi) \right|^{p'} d\xi \\
&\approx N^{\frac{p'}{2}},
\end{aligned}$$

by Khinchin's inequality. Therefore, there exist $\{\epsilon_j\}$ such that $\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)}^{p'} \lesssim N^{p'/2}$ so $\|\hat{f}\|_{L^{p'}(\mathbb{R}^n)} \lesssim N^{\frac{1}{2}}$ for $f = \sum \epsilon_j \phi_j$. In this case, the Hausdorff-Young inequality reads

$$N^{\frac{1}{2}} \lesssim \|\hat{f}\|_{p'} \leq \|f\|_p \approx N^{\frac{1}{p}},$$

which is not possible if $p > 2$ by choosing N large.

10.14. Laplacian under conformal change

In this section, we derive the Laplacian in local coordinates on a Riemannian manifold and how it changes under the conformal change of the metric. Let $\mathbb{M}$ be a n-dim Riemannian manifold with metric g, that is, for each $p \in \mathbb{M}$, $g(p)$ defines a symmetric and positive definite form $\langle\cdot,\cdot\rangle$ on the tangent space $T_p\mathbb{M}$, and $p \to g(p)$ is smooth. Denote by (g_{jk}) the matrix form of g and by $(g^{jk}) = (g_{ij})^{-1}$.

Definition (Gradient). Let $f \in C^\infty(\mathbb{M})$. The gradient of f is defined as a vector field, denoted by ∇f, such that

$$\langle \nabla f, X\rangle = Xf \quad \text{for any } X \in \mathcal{D}^1(\mathbb{M}).$$

Here, $\mathcal{D}^1(\mathbb{M})$ is the space of vector fields on $\mathbb{M}$.

In a local coordinate system $\varphi : U \to \varphi(U) \subset \mathbb{R}^n$ at $p \in U \subset \mathbb{M}$, we still denoted by f as the local representation $f \circ \varphi^{-1} : \varphi(U) \to \mathbb{C}$ of f. Choose a basis $\{\partial_{x_j}\}_{j=1}^n$ of $T_p\mathbb{M}$. Then

$$(g_{jk}) = \left\langle \partial_{x_j}, \partial_{x_k}\right\rangle.$$

We have that

$$\nabla f = \sum_{j=1}^n \left(\sum_{l=1}^n g^{jl}\partial_{x_l} f\right)\partial_{x_j}.$$

Definition (Riemannian volume form). Define the Riemannian volume form as

$$d\mathrm{Vol}_g = \sqrt{\det g}\, dx_1 \wedge \cdots \wedge dx_n.$$

We verify that the above definition is independent of the choice of local coordinate systems. In another local coordinate system $\tilde\varphi : \tilde U \to \mathbb{R}^n$ at $p \in \tilde U$, let $\{\partial_{\tilde x_j}\}_{j=1}^n$ span $T_p\mathbb{M}$ and $\{d\tilde x_j\}_{j=1}^n$ span $T_p^*\mathbb{M}$. Write

$$T = \tilde\varphi \circ \varphi^{-1} : \varphi(U \cap \tilde U) \to \tilde\varphi(U \cap \tilde U).$$

Then its Jacobian is

$$J = \frac{\partial(\tilde x_1, ..., \tilde x_n)}{\partial(x_1, ..., x_n)}.$$

Hence,

$$d\tilde x_1 \wedge \cdots \wedge d\tilde x_n = \det J\, dx_1 \wedge \cdots \wedge dx_n$$

and

$$\tilde g_{jk} = \left\langle \frac{\partial}{\partial \tilde x_j}, \frac{\partial}{\partial \tilde x_k}\right\rangle = \left\langle \sum_{l=1}^n \frac{\partial x_l}{\partial \tilde x_j}\frac{\partial}{\partial x_l}, \sum_{m=1}^n \frac{\partial x_m}{\partial \tilde x_k}\frac{\partial}{\partial x_m}\right\rangle = \sum_{l,m=1}^n J^{lj} g_{lm} J^{mk},$$

in which $(J^{lm}) = (J_{lm})^{-1}$. Therefore, $\tilde g = J^{-1} g J^{-1}$ and

$$\begin{aligned}\sqrt{\det \tilde g}\, d\tilde x_1 \wedge \cdots \wedge d\tilde x_n &= \sqrt{\det(J^{-1} g J^{-1})}\, \det J\, dx_1 \wedge \cdots \wedge dx_n \\ &= \sqrt{\det g}\, dx_1 \wedge \cdots \wedge dx_n.\end{aligned}$$

Definition (Laplacian). The Laplacian $\Delta_g : C^\infty(\mathbb{M}) \to C^\infty(\mathbb{M})$ is defined such that

$$\int_{\mathbb{M}} \langle \nabla h, \nabla f\rangle\, d\mathrm{Vol}_g = -\int_{\mathbb{M}} h\Delta_g f\, d\mathrm{Vol}_g \quad \text{for all } h \in C_0^\infty(\mathbb{M}), f \in C^\infty(\mathbb{M}).$$

Remark. From the definition, we immediately see that Δ_g is self-adjoint on $C_0^\infty(\mathbb{M})$.

We next compute the local representation of Δ_g: For $h \in C_0^\infty(U)$,

$$\begin{aligned}&\int_U \langle \nabla h, \nabla f\rangle\, d\mathrm{Vol}_g \\ =\ &\int_{\varphi(U)} \left\langle \sum_{j=1}^n \left(\sum_{l=1}^n g^{jl}\frac{\partial h}{\partial x_l}\right)\frac{\partial}{\partial x_j}, \sum_{k=1}^n \left(\sum_{m=1}^n g^{km}\frac{\partial f}{\partial x_m}\right)\frac{\partial}{\partial x_k}\right\rangle \sqrt{\det g}\, dx_1 \cdots dx_n\end{aligned}$$

$$
\begin{aligned}
&= \int_{\varphi(U)} \sum_{l,m=1}^{n} g^{jl}\frac{\partial h}{\partial x_l} g^{km}\frac{\partial f}{\partial x_m} \sum_{j,k=1}^{n} \langle \partial_{x_j}, \partial_{x_k}\rangle \sqrt{\det g}\, dx_1\cdots dx_n \\
&= \int_{\varphi(U)} \sum_{l,m=1}^{n} \frac{\partial h}{\partial x_l}\frac{\partial f}{\partial x_m}\left(\sum_{j,k=1}^{n} g^{jl} g_{jk} g^{km}\right)\sqrt{\det g}\, dx_1\cdots dx_n \\
&= \int_{\varphi(U)} \sum_{l,m=1}^{n} g^{lm}\frac{\partial h}{\partial x_l}\frac{\partial f}{\partial x_m}\sqrt{\det g}\, dx_1\cdots dx_n \\
&= -\int_{\varphi(U)} h\frac{\partial}{\partial x_l}\left(\sqrt{\det g}\sum_{m=1}^{n} g^{lm}\frac{\partial f}{\partial x_m}\right) dx_1\cdots dx_n \\
&= -\int_{\varphi(U)} h\Delta_g f\sqrt{\det g}\, dx_1\cdots dx_n.
\end{aligned}
$$

Here, we use integration by parts. Comparing with the definition of Δ_g, we have that

$$
\Delta_g f = \frac{1}{\sqrt{\det g}}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(\sqrt{\det g}\sum_{k=1}^{n} g^{jk}\frac{\partial f}{\partial x_k}\right).
$$

In the local coordinates, Δ_g is a second order elliptic partial differential operator (in the divergence form).

We next consider the change of Δ under a conformal change of metric: $g \to \tilde{g} = e^{\varphi} g$. Hence,

$$
(g^{jk}) = e^{-\varphi}(g^{jk}) \quad \text{and} \quad \det\tilde{g} = e^{n\varphi}\det g.
$$

Compute that

$$
\begin{aligned}
&\Delta_{\tilde{g}} f \\
= \;& \frac{1}{\sqrt{\det\tilde{g}}}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(\sqrt{\det\tilde{g}}\sum_{k=1}^{n}\tilde{g}^{jk}\frac{\partial f}{\partial x_k}\right) \\
= \;& \frac{1}{e^{\frac{n\varphi}{2}}\sqrt{\det g}}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(e^{\frac{n\varphi}{2}}\sqrt{\det g}\sum_{k=1}^{n} e^{-\varphi} g^{jk}\frac{\partial f}{\partial x_k}\right) \\
= \;& \frac{1}{e^{\frac{n\varphi}{2}}\sqrt{\det g}}\sum_{j=1}^{n}\left[e^{\frac{(n-2)\varphi}{2}}\frac{\partial}{\partial x_j}\left(\sqrt{\det g}\sum_{k=1}^{n} g^{jk}\frac{\partial f}{\partial x_k}\right) + \frac{\partial}{\partial x_j}e^{\frac{(n-2)\varphi}{2}}\left(\sqrt{\det g}\sum_{k=1}^{n} g^{jk}\frac{\partial f}{\partial x_k}\right)\right] \\
= \;& \frac{e^{-\varphi}}{\sqrt{\det g}}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}\left(\sqrt{\det g}\sum_{k=1}^{n} g^{jk}\frac{\partial f}{\partial x_k}\right) + \frac{1}{e^{\frac{n\varphi}{2}}\sqrt{\det g}}\sum_{j=1}^{n}\frac{\partial}{\partial x_j}e^{\frac{(n-2)\varphi}{2}}\left(\sqrt{\det g}\sum_{k=1}^{n} g^{jk}\frac{\partial f}{\partial x_k}\right) \\
= \;& e^{-\varphi}\Delta_g f + \frac{(n-2)}{2}e^{-\varphi}\sum_{j,k=1}^{n} g^{jk}\frac{\partial\varphi}{\partial x_j}\frac{\partial f}{\partial x_k}.
\end{aligned}
$$

In particular, on Riemannian surfaces (i.e., $n = 2$), the Laplacian changes in a simple way of

$$
\Delta_{\tilde{g}} = e^{-\varphi}\Delta_g.
$$

Definition (Conformally flat manifolds). We say that $(\mathbb{M}, g)$ is conformally flat if g is conformal to the Euclidean metric, that is, at each point $p \in \mathbb{M}$, there is a local coordinate system $p \to x(p)$ for which $e^{\varphi} g = dx^2$. We call such coordinates isothermal.

Remark.

- Korn and Lichtenstein proved that isothermal coordinates exist around any point on a Riemannian surface, c.f., Chern[i]. As a consequence, the Laplacian on surfaces can always be written as $e^{\varphi}\Delta_0$, in which Δ_0 is the Euclidean Laplacian.
- If $\dim \mathbb{M} = 3$, then $\mathbb{M}$ is conformally flat iff its Cotton tensor (in $\mathcal{D}^0_3(\mathbb{M})$) vanishes.
- If $\dim \mathbb{M} \geq 4$, then $\mathbb{M}$ is conformally flat iff its Weyl tensor (in $\mathcal{D}^0_4(\mathbb{M})$) vanishes.

10.15. Perron-Frobenius theorem

The Perron-Frobenius theorem includes a group of results that involves the eigenvalues and eigenvectors of matrices with non-negative entries. They were proved by Oskar Perron[ii] for positive matrices, some of which were later generalized to non-negative matrices by Georg Frobenius[iii]. See Chapter 8 in Horn-Johnson[iv] for a complete treatment of positive and non-negative matrices.

Perron-Frobenius theorem assumes its simplest and most elegant form for positive matrices. In this case, the highlight of the theorem says that a positive matrix has a simple eigenvalue which equals its spectral radius and has a positive eigenvector corresponding to this eigenvalue.

Definition. The spectrum of $A \in \mathrm{M}_n(\mathbb{R})$ is the set of all $\lambda \in \mathbb{C}$ that are eigenvalues of A. We denote this set by $\sigma(A)$. The spectrum radius is defined as

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

Theorem (Perron). *Let $A = (a_{jk}) \in \mathrm{M}_n(\mathbb{R})$ be positive (i.e., all entries are positive). Then the following holds.*

(i). $\rho(A) > 0$ and satisfies

$$\min_{j=1,\dots,n} \sum_{k=1}^{n} a_{jk} \leq \rho(A) \leq \max_{j=1,\dots,n} \sum_{k=1}^{n} a_{jk}.$$

(ii). $\rho(A)$ is a simple eigenvalue of A, which is called the Perron root or the Perron-Frobenius eigenvalue.

(iii). $|\lambda| < \rho(A)$ for every eigenvalue λ of A such that $\lambda \neq \rho(A)$.

(iv). There is a positive eigenvector x of A, which is called the Perron(-Frobenius) eigenvector, with eigenvalue $\rho(A)$.

(v). There are no other non-negative eigenvectors except positive multiples of x, that is, all other eigenvectors must have at least one negative or non-real component.

(vi). There is a positive row eigenvector y of A with eigenvalue $\rho(A)$, that is, $yA = \rho(A)y$.

(vii). Take the normalization that $x_1y_1 + \cdots + x_ny_n = 1$. Then $(\rho(A)^{-1}A)^m \to xy^{\mathrm{T}}$ as $m \to \infty$. Here, xy^{T} is the projection to the eigenspace corresponding to $\rho(A)$, which is called the Perron projection.

Remark (Spectral gap). Let $A \in \mathrm{M}_n(\mathbb{R})$ be positive and

$$|\lambda_{n-1}| = \max\{|\lambda| : \lambda \in \sigma(A), \lambda \neq \rho(A)\}.$$

Then

$$\rho(A) - |\lambda_{n-1}|$$

---

[i] S.-S. Chern, *An elementary proof of the existence of isothermal parameters on a surface.* Proc. Amer. Math. Soc. 6 (1955), 771–782.

[ii] O. Perron, *Zur Theorie der Matrices.* Math. Ann. 64 (1907), no. 2, 248–263.

[iii] G. Frobenius, *Ueber Matrizen aus nicht negativen Elementen*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften (1912) 456–477

[iv] R. Horn and C. Johnson, *Matrix analysis.* Second edition. Cambridge University Press, Cambridge, 2013.

is called the spectral gap of A. In fact, we have that

$$\frac{|\lambda_{n-1}|}{\rho(A)} \leq \frac{1-\kappa^2}{1+\kappa^2}, \quad \text{in which } \kappa = \frac{\min\{a_{jk} : j,k=1,...,n\}}{\max\{a_{jk} : j,k=1,...,n\}}.$$

Therefore,

$$\rho(A) - |\lambda_{n-1}| \geq \frac{2\kappa^2 \rho(A)}{1+\kappa^2}.$$

Take the identity matrix I as an example, we see that the direct generalization of the Perron-Frobenius theorem to non-negative matrices is impossible. For example, the eigenvalues of I are not simple and there is no spectral gap.

Theorem. *Let $A \in \mathrm{M}_n(\mathbb{R})$ be non-negative. Then the following holds.*

(i). $\rho(A) > 0$ *is an eigenvalue and satisfies that*

$$\min_{i=1,\dots,n} \sum_{j=1}^{n} a_{ij} \leq \rho(A) \leq \max_{i=1,\dots,n} \sum_{j=1}^{n} a_{ij}.$$

(ii). $|\lambda| < \rho(A)$ *for every eigenvalue* λ *of* A *such that* $\lambda \neq \rho(A)$.

(iii). There is a non-negative eigenvector x *of* A *with eigenvalue* $\rho(A)$.

(iv). Let $x \in \mathbb{R}^n$ *be positive and* $\lambda \geq 0$. *Then* $Ax = \lambda x$ *implies that* $\lambda = \rho(A)$.

(v). If $x \in \mathbb{R}^n$ *is non-negative and non-zero, then* $Ax \geq \alpha x$ *for some* $\alpha \in \mathbb{R}$ *implies that* $\alpha < \rho(A)$.

Adding the irreducible condition, the Perron-Frobenius theorem can be largely generalized to non-negative matrices.

Definition (Irreducible matrices). Let $A \in \mathrm{M}_n(\mathbb{C})$. We say that A is reducible if there is a permutation matrix $P \in \mathrm{M}_n(\mathbb{C})$ such that

$$P^{\mathrm{T}} A P = \begin{pmatrix} B & C \\ 0_{n-r,r} & D \end{pmatrix} \quad \text{for some } 1 \leq r \leq n-1.$$

We say that A is irreducible if it is not reducible.

Remark. Let $A \in \mathrm{M}_n(\mathbb{C})$. Then A is irreducible iff $(I + |A|)^{n-1}$ is positive.

Theorem (Frobenius). *Let $A \in \mathrm{M}_n(\mathbb{R})$ be non-negative and irreducible. Then the following holds.*

(i). $\rho(A) > 0$ *is a simple eigenvalue of* A.

(ii). There is a positive eigenvector x *of* A *with eigenvalue* $\rho(A)$.

(iii). There are no other positive eigenvectors except positive multiples of x, *that is, all other eigenvectors must have at least one non-positive or non-real component.*

(iv). There is a positive row eigenvector y *of* A *with eigenvalue* $\rho(A)$,

Remark. There is a useful formula for computing the spectral radius of non-negative matrices:

$$\rho(A) = \limsup_{m\to\infty} \left(\mathrm{Tr}\, A^m\right)^{\frac{1}{m}}.$$

The Perron-Frobenius theorem has been generalized to infinite dimensional spaces. In addition, it has applications to a great variety of areas including the ergodic theory and dynamical systems. See the notes of Section XIII.12 in Reed-Simon[i] for a historical background on the development of this theorem.

[i]M. Reed and B. Simon, *Methods of modern mathematical physics. IV. Analysis of operators.* Academic Press, New York-London, 1978.

10.16. Poisson's equation with no C^2 solutions

Question 10.3 (Poisson's equation with no C^2 solutions). *Poisson's equation $\Delta u = g$ does not necessarily have a solution in C^2 for $g \in C^0$. Any examples?*

See Section 7 in Brezis-Browder[i]. The following example is from Problem 4.9 (a) in Gilbarg-Trudinger[ii].

Answer. Let $P(x, y) = xy$. Then $\Delta P = 0$. Choose $\eta \in C_0^\infty(B(0, 2))$ with $\eta = 1$ in $B(0, 1)$. Set $c_k \to 0$ as $k \to \infty$ such that $\sum c_k$ divergent. Define

$$f(x, y) = \sum_{k=0}^{\infty} c_k \Delta(\eta P)\left(2^k x, 2^k y\right) = \sum_{k=0}^{\infty} c_k (P\Delta\eta + \nabla\eta \cdot \nabla P)\left(2^k x, 2^k y\right).$$

Then $f(0, 0) = 0$. For every $(x, y) \neq (0, 0) \in \mathbb{R}^2$, the above summation is finite, since $(2^k x, 2^k y) \notin B(0, 2)$ when k is large. Hence, $f \in C^0(\mathbb{R}^2)$. The solution to $\Delta u = f$ is

$$u(x, y) = \sum_{k=0}^{\infty} c_k (\eta P)(2^k x, 2^k y).$$

It is not C^2 around the origin. This is because

$$u(0, 0) = \sum_{k=0}^{\infty} c_k$$

diverges.

Remark. From the above example, one needs to impose stronger condition than continuity for the existence of solutions to the Poisson's equation. For example, Lemma 4.2 in Gilberg-Trudinger states: Let f be bounded and locally Hölder continuous in Ω. Then $w \in C^2(\Omega)$ given by the Newtonian potential of f solves $\Delta w = f$ in Ω.

10.17. Representation theory of groups

Let G be a group. Suppose that V is a (non-trivial) vector space over a field $\mathbb{F}$. Denote by $\mathrm{GL}(V)$ the general linear group on V. The representation of G is to represent the elements of G by the ones in $\mathrm{GL}(V)$. When G has additional structure, such as finite, compact, Lie group, or is explicitly $\mathrm{SO}(3)$, $\mathrm{SL}(2, \mathbb{R})$, etc, the representation theory can be incredible rich. It has close connections with other fields, including harmonic analysis, dynamical systems, hyperbolic geometry, mathematical physics, etc.

Definition (Group representations). A representation of G on V is a group homomorphism $\Phi : G \to \mathrm{GL}(V)$, that is, $\Phi(g_1 g_2) = \Phi(g_1)\Phi(g_2)$ for all $g_1, g_2 \in G$. We call V the representation space (or simply the representation) and $\dim V$ the dimension (or the degree) of Φ.

Remark.

- If $\dim V = n < \infty$, then $\mathrm{GL}(V) = \mathrm{GL}(n, \mathbb{F})$, i.e., the general linear group of degree n over $\mathbb{F}$. In this case, we define the character $\chi_\Phi : G \to \mathbb{F}$ of a representation $\Phi : G \to \mathrm{GL}(n, \mathbb{F})$ by $\chi_\Phi(g) = \mathrm{Tr}\,\Phi(g)$.

[i] H. Brezis and F. Browder, *Partial differential equations in the 20th century.* Adv. Math. 135 (1998), no. 1, 76–144.

[ii] D. Gilbarg and N. Trudinger, *Elliptic partial differential equations of second order.* Springer-Verlag, Berlin, 2001.

- For a representation $\Phi : G \to \mathrm{GL}(V)$ of a topological group G, we also require that V is a topological vector space and the mapping $(g, v) \to \Phi(g)(v)$ is continuous from $G \times V$ to V.
- The kernel of a representation Φ of a group G is defined as
$$\ker \Phi = \{h \in G : \Phi(h) = \mathrm{I} \in \mathrm{GL}(V)\}.$$
Notice that $\ker \Phi$ is a normal subgroup of G, that is, if $h \in \ker \Phi$, then $ghg^{-1} \in \ker \Phi$ for all $g \in G$.
- We say that a representation Φ is faithful if it is injective, equivalently, $\ker \Phi = \{\mathrm{I}\}$.
- Let V and V' be two vector spaces over a field $\mathbb{F}$. Suppose that $\Phi : G \to \mathrm{GL}(V)$ and $\Phi' : G \to \mathrm{GL}(V')$ are two representations of G. Then we say that Φ and Φ' are equivalent or isomorphic if there exists a bounded linear invertible map $E : V \to V'$ with a bounded inverse such that $\Phi'(g)E = E\Phi(g)$ for all $g \in G$, that is,
$$\Phi'(g)(E(v)) = E(\Phi(g)(v)) \quad \text{for all } g \in G \text{ and } v \in V.$$

Example. Let $G = \{1, u, u^2\}$ with $u = e^{2\pi i/3}$. Then the following are three representations of G.

(1). $\Phi : G \to \mathrm{GL}(2, \mathbb{C})$ that is defined by
$$\Phi(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi(u) = \begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}, \quad \Phi\left(u^2\right) = \begin{pmatrix} 1 & 0 \\ 0 & u^2 \end{pmatrix}.$$
(2). $\Phi : G \to \mathrm{GL}(2, \mathbb{C})$ that is defined by
$$\Phi(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi(u) = \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi\left(u^2\right) = \begin{pmatrix} u^2 & 0 \\ 0 & 1 \end{pmatrix}.$$
(3). $\Phi : G \to \mathrm{GL}(2, \mathbb{R})$ that is defined by
$$\Phi(1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Phi(u) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \Phi\left(u^2\right) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$
The above examples are all faithful, in which (1) and (2) are equivalent.

Definition (Irreducible representations). Let $\Phi : G \to \mathrm{GL}(V)$ be a representation. We say that a linear subspace $W \subset V$ is invariant for Φ if $\Phi(g)(W) \subset W$ for all $g \in G$. In this case, $\Phi|_W : G \to \mathrm{GL}(W)$ defines a representation and is called a subrepresentation of Φ. (Or simply, W is a subrepresentation of V.) A representation is said to be irreducible if the only subrepresentations are trivial (i.e., $\{0\}$ and V).

Example.

(1). A representation $\Phi : G \to \mathrm{GL}(V)$ with $\dim V \le 1$ is always irreducible.

(2). Let S_n be the permutation group of n letters. Suppose that V is a n-dim complex vector space with a basis $\{v_1, , , v_n\}$. Then there is a representation $\Phi : S_n \to \mathrm{GL}(V)$ that is defined by permuting the elements in the basis. In this case, V decomposes into a direct sum of two linear subspaces that are invariant under S_n, i.e., $V = W_1 \oplus W_2$, in which
$$W_1 = \left\{ c \sum_{j=1}^{n} v_j : c \in \mathbb{C} \right\} \quad \text{and} \quad W_2 = \left\{ v = \sum_{j=1}^{n} a_j v_j : \sum_{j=1}^{n} a_j = 0 \right\}.$$
Here, W_1 and W_2 are two irreducible subrepresentations of dimension 1 and $n - 1$, respectively.

Definition (Direct sums of representations). Let $\Phi_V : G \to \mathrm{GL}(V)$ and $\Phi_W : G \to \mathrm{GL}(W)$ be two representations of G. Then the direct sum of Φ_V and Φ_W, denoted by $\Phi_V \oplus \Phi_W$, is a representation that is defined such that $\Phi_V \oplus \Phi_W(g)$ acts on $V \oplus W$ component-wise. In particular, if V and W are both finite-dimensional, then

$$\Phi_V \oplus \Phi_W(g) = \begin{pmatrix} \Phi_V(g) & 0 \\ 0 & \Phi_W(g) \end{pmatrix}.$$

Example. Let $G = \{1, u, u^2\}$ with $u = e^{2\pi i/3}$. Consider the three representations of G given as above. Then both (1) and (2) are reducible and are direct sums of subrepresentations $\mathbb{C}(1,0)$ and $\mathbb{C}(0,1)$, whereas (3) is irreducible.

Definition (Unitary representations). Let $\Phi : G \to \mathrm{GL}(H)$ be a representation on a (complex) Hilbert space H. We say that G is unitary if $\Phi(g) : H \to H$ is unitary for all $g \in G$. Here, an operator $U : H \to H$ is said to be unitary if $U^\star U = UU^\star = \mathrm{I}$, that is, U preserves the Hermitian inner product on H.

Remark. Let $\Phi : G \to \mathrm{GL}(H)$ be a unitary representation. Suppose that $W \subset V$ is invariant. Then $W^\perp = \{v \in V : (v, w) = 0 \text{ for all } w \in W\}$ is also invariant. Indeed, for all $g \in G$ and $v \in W^\perp$, we have that

$$(\Phi(g)v, w) = \big(v, \Phi(g)^{-1}w\big) = 0,$$

because $\Phi(g^{-1})w \in W$. That is, $\Phi(g)v \in W^\perp$ for all $g \in G$ and $v \in W^\perp$. Moreover, $V = W \oplus W^\perp$.

As a consequence, every finite-dimensional unitary representation is a direct sum of irreducible subrepresentations.

Example.

(1). Let $G = \mathrm{SO}(n)$ and V be the space of complex homogeneous polynomials of n variables and with degree N. Then $\Phi : G \to \mathrm{GL}(V)$ defined by

$$\Phi(g)P(x_1, ..., x_n) = P\left(g^{-1}(x_1, ..., x_n)\right)$$

is a representation of G. It is unitary with respect to the inner product

$$(P, Q) = \int_{\mathbb{S}^{n-1}} P(x)\overline{Q(x)}\, dx.$$

(2). Let $G = \mathrm{SL}(2, \mathbb{R})$ and $V = L^2(\mathbb{R}^2)$ (with respect to the Lebesgue measure). Then $\Phi : G \to \mathrm{GL}(V)$ defined by

$$\Phi(g)f(x) = f\left(g^{-1}x\right)$$

is a unitary representation.

THEOREM 10.2 (Schur's lemma). *Let $\Phi : G \to \mathrm{GL}(V)$ be a unitary representations of a topological group G on a Hilbert space V. Then Φ is irreducible iff every linear map $L : V \to V$ such that $\Phi(g)L = L\Phi(g)$ for all $g \in G$ must be a scalar operator (i.e., $L = \lambda\mathrm{I}$ for some $\lambda \in \mathbb{C}$).*

As an immediate consequence,

Corollary. *If G is abelian, then any irreducible unitary representation $\Phi : G \to \mathrm{GL}(V)$ must be one-dimensional.*

PROOF. Since $\Phi(g) \in \mathrm{GL}(V)$ commutes with each other, $\Phi(g) = \lambda_g\mathrm{I}$ for some $\lambda_g \in \mathbb{C}$ by the previous theorem. Therefore, each one-dimensional subspace of V is invariant for Φ (since $\Phi(g)v = \lambda v$ for all $g \in G$ and $v \in V$). But this means that $\dim V = 1$, because Φ is irreducible. $\square$

Remark. Let $\Phi : G \to \mathrm{GL}(V)$ and $\Phi' : G \to \mathrm{GL}(V)$ be two irreducible unitary representations. Then the following dichotomy holds.

- Φ and Φ' are isomorphic and there is a scalar map $L : V \to V$ such that $\Phi'(g)L = L\Phi(g)$ for all $g \in G$.
- Φ and Φ' are not isomorphic and any linear map $L : V \to V$ such that $\Phi'(g)L = L\Phi(g)$ for all $g \in G$ must be 0.

Quoting Section I.4 of Knapp[i], the three main problems in the representation theory for a particular group G are as follows.

(A). Classify explicitly the irreducible unitary representations of G. Give classification for any other interesting manageable classes of irreducible representations.
(B). Give an explicit analysis of $L^2(G)$.
(C). Analyze other representations of G, particularly unitary ones, that arise naturally in mathematics.

Example.

(1). Non-compact groups:
- $\mathrm{GL}(n,\mathbb{C}) = \{g \in \mathrm{M}_n(\mathbb{C}) : \det g \neq 0\}$.
- $\mathrm{SL}(n,\mathbb{C}) = \{g \in \mathrm{GL}(n,\mathbb{C}) : \det g = 1\}$.
- $\mathrm{GL}(n,\mathbb{R}) = \{g \in \mathrm{M}_n(\mathbb{R}) : \det g \neq 0\}$.
- $\mathrm{SL}(n,\mathbb{R}) = \{g \in \mathrm{GL}(n,\mathbb{R}) : \det g = 1\}$.

(2). Compact groups:
- $\mathrm{U}(n) = \left\{g \in \mathrm{GL}(n,\mathbb{C}) : g\overline{g}^{\mathrm{T}} = \mathrm{I}\right\}$.
- $\mathrm{SU}(n) = \{g \in \mathrm{U}(n) : \det g = 1\}$.
- $\mathrm{O}(n) = \left\{g \in \mathrm{GL}(n,\mathbb{R}) : gg^{\mathrm{T}} = \mathrm{I}\right\}$.
- $\mathrm{SO}(n) = \{g \in \mathrm{O}(n) : \det g = 1\}$.

(3). Let $(\cdot,\cdot)_{m,n}$ be the symmetric bilinear form on $\mathbb{C}^{m+n}$ that is given by
$$(z,z')_{m,n} = z_1\overline{z_1'} + \cdots + z_m\overline{z_m'} - z_{m+1}\overline{z_{m+1}'} - \cdots - z_{m+n}\overline{z_{m+1}'}.$$
Then $g \in \mathrm{M}_n(\mathbb{C})$ preserves this form iff
$$\mathrm{I}_{m,n}g^{\star}\mathrm{I}_{m,n} = g^{-1},$$
in which $\mathrm{I}_{m,n} = \mathrm{diag}(\mathrm{I}_m, -\mathrm{I}_n)$. The indefinite orthogonal groups of signature (m,n) are defined as
- $\mathrm{SU}(m,n) = \{g \in \mathrm{SU}(m+n) : g \text{ preserves } (\cdot,\cdot)_{m,n}\}$.
- $\mathrm{SO}(m,n) = \{g \in \mathrm{SO}(m+n) : g \text{ preserves } (\cdot,\cdot)_{m,n}\}$.

In particular,
$$\mathrm{SU}(1,1) = \left\{\begin{pmatrix}\alpha & \beta \\ \overline{\beta} & \overline{\alpha}\end{pmatrix} : |\alpha|^2 - |\beta|^2 = 1\right\}$$
Notice that $\mathrm{SO}(m,n)$ has two connected components, which contains I and $-$I and are denoted by $\mathrm{SO}^{+}(m,n)$ (or $\mathrm{SO}_0(m,n)$) and $\mathrm{SO}_{-}(m,n)$, respectively.

Example (Regular representations). Let G be locally compact (such as all the above examples) and $V = L^2(G, d_lx)$, in which d_lx is a left-invariant measure on G. Then $\Phi : G \to L^2(G, d_lx)$ defined by $\Phi(g)f(x) = f(g^{-1}x)$ is a representation and is called the left regular representation of G. One can similarly define the right regular representation of G, $\Phi' : G \to L^2(G, d_rx)$ via $\Phi'(g)f(x) = f(xg)$.

[i] A. Knapp, *Representation theory of semisimple groups. An overview based on examples.* Princeton University Press, Princeton, NJ, 1986.

10.17.1. Representation theory of compact groups. In this section, we assume that G is compact with a normalized Haar measure dx (so $\int_G dx = 1$) and present the representation theory of G, following Section I.5 of Knapp.

According to Problems (A), (B), (C) above, this representation theory includes the classification of irreducible unitary representations of G, as well as its applications to the analysis of $L^2(G)$. In particular, it leads to a decomposition of functions in $L^2(G)$ which is the analogue of the Fourier series expansion in $L^2(\mathbb{S})$. In fact, $\mathbb{S} = \mathrm{U}(1)$ is considered as a compact group (for which the Haar measure is just the Lebesgue measure) so the representation theory implies the Fourier analysis on $\mathbb{S}$ as a special case.

Let $\Phi : G \to \mathrm{GL}(V)$ is a representation. To make the connection between the representation Φ of G and functions on G, we write

$$\Phi(f)v = \int_G f(x)\Phi(x)v\,dx \quad \text{for } v \in V \text{ and } f \in L^1(G).$$

Suppose that $\Phi : G \to \mathrm{GL}(V)$ is a representation on a complex vector space V. We first show that V always admits a Hermitian inner product with respect to which Φ is unitary. Indeed, let $(\cdot, \cdot)$ be any Hermitian inner product on V, then define

$$\langle v, w\rangle = \int_{x\in G} (\Phi(x)v, \Phi(x)w)\ dx.$$

It is straightforward to verify that $\Phi : G \to \mathrm{GL}(V)$ is unitary with respect to $\langle\cdot,\cdot\rangle$. Without loss of generality, we can therefore assume that Φ is unitary with respect to the Hermitian inner product $(\cdot, \cdot)$ on V.

Notice that here V can be infinite dimensional. But if the representation is finite-dimensional, then given any invariant subspace $W \subset V$, we have that $V = W \oplus W^\perp$. Repeated process implies that V can be decomposed as a direct sum of irreducible subrepresentations.

Theorem (Schur orthogonality relations).

(i). Let $\Phi : G \to \mathrm{GL}(V)$ be an irreducible unitary representations on a finite-dimensional vector space V. Then

$$\int_G (\Phi(x)u_1, v_2)\,\overline{(\Phi(x)u_2, v_2)}\,dx = \frac{(u_1, u_2)\overline{(v_1, v_2)}}{\dim V} \quad \textit{for all } u_1, v_1, u_2, v_2 \in V.$$

(ii). Let $\Phi : G \to \mathrm{GL}(V)$ and $\Phi' : G \to \mathrm{GL}(V)$ be two inequivalent irreducible unitary representations on finite-dimensional vector spaces V and V', respectively. Then

$$\int_G (\Phi(x)u, v)\,\overline{(\Phi'(x)u', v')}\,dx = 0 \quad \textit{for all } u, v \in V \textit{ and } u', v' \in V.$$

Remark.

(i). Let $\Phi : G \to \mathrm{GL}(V)$ be an irreducible unitary representation on V with $\dim V = d < \infty$. Suppose that $\{v_j\}_{j=1}^d$ is an orthonormal basis of V. Consider the functions

$$\Phi_{jk}(x) := (\Phi(x)u_j, u_k) \quad \text{for } j, k = 1, ..., d.$$

Then (i) in Schur orthogonality relations indicates that

$$\left\{\sqrt{d}\cdot\Phi_{jk}\right\}_{j,k=1}^d$$

is an orthonormal set in $L^2(G)$.

(ii). Let $\{\Phi\}$ be a set of mutually inequivalent finite-dimensional irreducible unitary representations. Then (ii) in Schur orthogonality relations indicates that

$$\bigcup_{\Phi}\left\{\sqrt{d}\cdot\Phi_{jk}\right\}_{j,k=1}^{d}$$

is an orthonormal set in $L^2(G)$. The Peter-Weyl theorem in the following, among other results, concludes that they form an orthonormal basis of $L^2(G)$, if $\{\Phi\}$ be a maximal set of mutually inequivalent finite-dimensional irreducible unitary representations.

Definition (Matrix coefficients). Let $\Phi : G \to \mathrm{GL}(V)$ be a unitary representation. For $u, v \in V$, $(\Phi(x)u, v)$ is called a matrix coefficient of Φ.

Proposition (Characters). *Let $\Phi : G \to \mathrm{GL}(V)$ be a unitary representation on V with $\dim V = d < \infty$. Suppose that $\{v_j\}_{j=1}^{d}$ is an orthonormal basis in V. Then the character*

$$\chi_\Phi(x) = \operatorname{Tr}\Phi(x) = \sum_{j=1}^{d}\left(\Phi(x)u_j, u_j\right).$$

As a direct consequence of Schur orthogonality relations, we have that

(i). $\chi_\Phi(x^{-1}) = \overline{\chi_\Phi(x)}$ *for* $x \in G$.

(ii). $\chi_\Phi * \chi_{\Phi'} = 0$ *if* Φ *and* Φ' *are not equivalent. Here,*

$$f * h(x) = \int_G f\left(xy^{-1}\right)h(x)\,dx.$$

The following Peter-Weyl[i] theorem summarizes the representation of compact groups and its relations with the harmonic analysis on $L^2(G)$.

Theorem (Peter-Weyl).

(i). The linear span of all matrix coefficients for all finite-dimensional irreducible unitary representations of G is dense in $L^2(G)$.

(ii). Let $\{\Phi\}$ be a maximal set of mutually inequivalent finite-dimensional unitary representations. Then

$$\bigcup_{\Phi}\left\{\sqrt{d_\Phi}\cdot\Phi_{jk}\right\}_{j,k=1}^{d}$$

is an orthonormal basis of $L^2(G)$. Moreover, the Parseval-Plancherel formula holds:

$$\int_G |f(x)|^2\,dx = \sum_{\Phi} d_\Phi \sum_{j,k=1}^{d}\left|\left(\Phi(f)u_j, u_k\right)\right|^2.$$

(iii). Every irreducible unitary representation of G is finite-dimensional.

(iv). Let Φ be a unitary representation of G on a Hilbert space V. Then V is an orthogonal sum of finite-dimensional irreducible invariant subspaces.

(v). Let Φ be a unitary representation of G on a Hilbert space V. For each irreducible unitary representation τ of G, let E_τ be the orthogonal projection on the closure of the sum of all irreducible invariant subspaces of V that are equivalent with τ. Then

$$E_\tau = d_\tau\cdot\Phi(\overline{\chi_\tau}),$$

[i] F. Peter and H. Weyl, *Die Vollständigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe*. Math. Ann. 97 (1927), no. 1, 737–755.

in which d_τ is the degree of τ and χ_τ is the character of τ. Moreover, if τ and τ' are inequivalent, then $E_\tau E_{\tau'} = E_{\tau'} E_\tau = 0$. Finally, for each $v \in V$, we have that

$$v = \sum_\tau E_\tau(v),$$

in which the sum is taken over a set of representations τ of all equivalent classes of irreducible unitary representations of G.

Remark (Multiplicity). The conclusion (v) in the Peter-Weyl theorem says that the number of occurances of τ in a decomposition (iv) is independent of the decomposition. This number equals $\dim \operatorname{im} E_\tau / d_\tau$, which is called the multiplicity of τ in Φ and is denoted by $[\Phi : \tau]$.

Indeed, let $\Phi : G \to \mathrm{GL}(V^\Phi)$ and $\tau : G \to \mathrm{GL}(V^\tau)$ be two unitary representations. Suppose that τ is irreducible. Then

$$[\Phi : \tau] = \dim \operatorname{Hom}_G \left(V^\Phi, V^\tau\right) = \dim \operatorname{Hom}_G \left(V^\tau, V^\Phi\right).$$

Example (Discrete Fourier analysis). Let $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z} = \{0, ..., n-1\}$ be equipped with the normalized counting measure. Then $\mathbb{Z}_n$ is an abelian finite group. So each irreducible unitary representations of $\mathbb{Z}_n$ is one-dimensional, which therefore coincides with its character.

Suppose that $\Phi : \mathbb{Z}_n \to \mathbb{C} \setminus \{0\}$ be an irreducible unitary representation of $\mathbb{Z}_n$. Then Φ is completely determined by $\Phi(1)$ since $\Phi(x) = \Phi(1)^x$ for $x \in \mathbb{Z}_n$. Notice that $1 = \Phi(0) = \Phi(n) = \Phi(1)^n$. Hence, $\Phi(1) = e^{ik\pi/n}$ for some $k = 0, ..., n-1$ so $\Phi(x) = e^{ikx\pi/n}$. Let

$$\tau_k(x) = e^{-ikx\pi/n}, \quad k = 0, ..., n-1,$$

be all the irreducible unitary representations of $\mathbb{Z}_n$. (The choice of a negative sign here is to keep consistence with the convention in Fourier analysis.) Then for $f \in L^2(\mathbb{Z}_n)$,

$$\tau_k(f) = \frac{1}{n} \sum_{x=0}^{n-1} f(x) \tau_k(x) = \frac{1}{n} \sum_{x=0}^{n-1} f(x) e^{-ikx/n},$$

which is the k-th Fourier coefficient of f.

Consider the left regular representation $\Phi : \mathbb{Z}_n \to L^2(\mathbb{Z}_n)$, $\Phi(x) f(y) = f(y-x)$ for $f \in L^2(\mathbb{Z}_n)$. Then

$$E_{\tau_k} f(y) = \Phi\left(\overline{\chi_{\tau_k}}\right) f(y) = \frac{1}{n} \sum_{x=0}^{n-1} \overline{\tau_k(x)} \Phi(x) f(y) = \frac{1}{n} \sum_{x=0}^{n-1} e^{ikx/n} f(y-x) = \tau_k(f) e^{iky/n}.$$

Hence, Peter-Weyl theorem recovers the discrete Fourier analysis on $\mathbb{Z}_n$:

(ii). $\{\tau_k(x) = e^{-ikx\pi/n}\}_{k=0}^{n-1}$ is an orthonormal basis of $L^2(\mathbb{Z}_n)$. Moreover, the Parseval-Plancherel formula holds:

$$\frac{1}{n} \sum_{x=0}^{n-1} |f(x)|^2 = \sum_{k=0}^{n-1} |\tau_k(f)|^2.$$

(v). The Fourier inversion formula holds (in $L^2(\mathbb{Z}_n)$ and pointwise):

$$f(y) = \sum_{k=0}^{n-1} E_{\tau_k} f(y) = \sum_{k=0}^{n-1} \tau_k(f) e^{iky/n}.$$

Example (Fourier analysis on $\mathbb{S}$). Let $\mathrm{U}(1) = \mathbb{S} = \{x = e^{i\theta} : \theta \in [0, 2\pi)\}$ be equipped with the normalized Lebesgue measure. Then $\mathbb{S}$ is an abelian compact group. So each irreducible unitary representations of $\mathbb{S}$ is one-dimensional, which therefore coincides with its character.

Suppose that $\Phi : \mathbb{S} \to \mathbb{C} \setminus \{0\}$ be an irreducible unitary representation of $\mathbb{S}$. Let $k \in \mathbb{Z}$. Then for each $\theta \neq 0$,

$$e^{2k\pi i} = 1 = \Phi\left(e^{2\pi i}\right) = \left[\Phi\left(e^{i\theta}\right)\right]^{\frac{2\pi}{\theta}},$$

which implies that

$$\Phi\left(e^{i\theta}\right) = 1^{\frac{\theta}{2\pi}} = \left(e^{2k\pi i}\right)^{\frac{\theta}{2\pi}} = e^{ik\theta} \quad \text{for some } k \in \mathbb{Z}.$$

Let $\tau_k(e^{i\theta}) = e^{-ik\theta}$, $k \in \mathbb{Z}$, be all the irreducible unitary representations of $\mathbb{S}$. (The choice of a negative sign here is to keep consistence with the convention in Fourier analysis.) Then for $f \in L^2(\mathbb{S})$,

$$\tau_k(f) = \int_{\mathbb{S}} f(x)\tau_k(x)\, dx = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)\tau_k\left(e^{i\theta}\right)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)e^{-ik\theta}\, d\theta,$$

which is the k-th Fourier coefficient of f. Here, we abuse the notation a bit by denoting $f(\theta)$ as $f(e^{i\theta})$.

Consider the left regular representation $\Phi : \mathbb{S} \to L^2(\mathbb{S})$, $\Phi(e^{i\theta})f(e^{i\phi}) = f(e^{i(\phi-\theta)})$ for $f \in L^2(\mathbb{S})$, in which $y = e^{i\phi}$. Then

$$E_{\tau_k}f(y) = \Phi\left(\overline{\chi_{\tau_k}}\right) f(y)\, dx = \int_{\mathbb{S}} \overline{\tau_k(x)}\Phi(x)f(y)\, dx = \frac{1}{2\pi}\int_0^{2\pi} e^{ik\theta} f(\phi-\theta)\, d\theta = \tau_k(f)e^{ik\phi}.$$

Hence, Peter-Weyl theorem recovers the discrete Fourier analysis on $\mathbb{S}$:

(ii). $\{\tau_k(\theta) = e^{-ik\theta}\}_{k\in\mathbb{Z}}$ is an orthonormal basis of $L^2(\mathbb{S})$. Moreover, the Parseval-Plancherel formula holds:

$$\frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|^2\, d\theta = \sum_{k\in\mathbb{Z}} |\tau_k(f)|^2.$$

(v). The Fourier inversion formula holds (in $L^2(\mathbb{S})$):

$$f(\phi) = \sum_{k\in\mathbb{Z}} E_{\tau_k} f(\theta) = \sum_{k\in\mathbb{Z}} \tau_k(f)e^{ik\theta}.$$

10.18. Riemannian volume form in local coordinates

Question 10.4 (Riemannian volume form in local coordinates). *Let $\mathbb{M}$ be a Riemannian manifold. On a patch $\Omega \subset \mathbb{M}$, one can assume that the density dx on $\mathbb{M}$ agrees with Lebesgue measure in the local coordinates. This can always be achieved after possibly contracting Ω. See Section 3 in Hörmander[i] and Section 4.1 in Sogge[ii]. Why?*

Answer (Hassell). Suppose that in a local coordinate system $x_1, ..., x_n$, the density

$$dx = f(x_1, ..., x_n)\, dx_1 \cdots dx_{n-1} dx_n.$$

Then by choosing $\tilde{x}_n$ properly, we can write

$$dx = dx_1 \cdots dx_{n-1} d\tilde{x}_n.$$

This requires that

$$d\tilde{x}_n = f(x_1, ..., x_n)\, dx_n,$$

[i] L. Hörmander, *The spectral function of an elliptic operator.* Acta Math. 121 (1968), 193–218.

[ii] C. Sogge, *Fourier integrals in classical analysis.* Second edition. Cambridge University Press, Cambridge, 2017.

which can be done by letting

$$\tilde{x}_n = \int_0^{x_n} f(x_1, ..., x_{n-1}, t)\, dt.$$

10.19. Siegel-Shidlovskii theorem

One aspect of the Siegel-Shidlovskii[i] theorem deals with the algebraic relations of the zeros of functions, which are solutions to linear partial differential equations with rational coefficients. In this section, we visit some topics related to the Siegel-Shidlovskii theorem.

Definition (Algebraic numbers and transcendental numbers). A number $\alpha \in \mathbb{C}$ is said to be algebraic if it solves an algebraic equation, i.e., an equation

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0,$$

in which $a_n, ..., a_0 \in \mathbb{Z}$ are not all zero (or equivalently, $a_n, ..., a_0 \in \mathbb{Q}$). A number which is not algebraic is said to be transcendental.

The minimal polynomial of an algebraic number α is the polynomial with the lowest degree that α solves. The degree of an algebraic number α is the degree of its minimal polynomial. The solutions to the minimal polynomial of an algebraic α are called the conjugates of α.

Example.

(1). Rational numbers are algebraic and have degree 1. In fact, a number has degree 1 iff it is rational (except 0).

(2). The conjugate of i is $-i$, since the minimal polynomial of i and $-i$ is $z^2 + 1 = 0$.

Remark.

- Transcendental numbers are dense in $\mathbb{R}$ and in $\mathbb{C}$. However, they are difficult to construct. Liouville in 1844 proved the existence of transcendental numbers and in 1851[ii] constructed the first transcendental number, which is now called the Liouville constant:

$$L_b = \sum_{n=1}^{\infty} \frac{1}{10^{n!}}.$$

- Lindemann[iii] first proved that π and e are transcendental. The theorem was later generalized by Weierstrass[iv] to what is now called the Lindemann-Weierstrass theorem. See Theorem 10.8.

Theorem (Algebraic closure). *The set of algebraic numbers form a field, denoted by $\overline{\mathbb{Q}}$, and is called the algebraic closure of $\mathbb{Q}$.*

Definition (Algebraic integers). An algebraic number α is an algebraic integer if it solves an algebraic equation with leading coefficient 1, i.e., an equation

$$z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0.$$

Example.

(1). A rational number is an algebraic integer iff it is an integer.

---

[i] A. Shidlovskii, *Transcendental numbers*. Walter de Gruyter & Co., Berlin, 1989.

[ii] J. Liouville, *Sur des classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible à des irrationnelles algébriques*. J. Math. Pures Appl. 16 (1851) 133–142.

[iii] F. Lindemann, *Ueber die Zahl* π. Math. Ann. 20 (1882), no. 2, 213–225.

[iv] K. Weierstrass, *Zu Lindemann's Abhandlung: Über die Ludolph'sche Zahl.*, Sitzungsberichte der Königlich Preussischen Akademie der Wissen-schaften zu Berlin, 5 (1885) 1067–1085.

(2). The solutions to $z^2 + 1 = 0$, i and $-i$, are algebraic integers.
(3). The solutions to $z^2 - 2 = 0$, $\sqrt{2}$ and $-\sqrt{2}$, are algebraic integers.

Definition (Absolute value and denominator). Let $\alpha \in \overline{\mathbb{Q}}$.

- The absolute value of α, denoted as $|\overline{\alpha}|$, is the maximum of the absolute values of all conjugates of α.
- The denominator of α, denoted as $\mathrm{deno}(\alpha)$, is the smallest positive number d such that αd is an algebraic integer.

Definition. Let $\mathbb{F}$ be a field.

- $\mathbb{F}[z]$ denotes the set of polynomials with coefficients in $\mathbb{F}$.
- $\mathbb{F}[[z]]$ denotes the set of power series with coefficients in $\mathbb{F}$.
- $\mathbb{F}(z)$ denotes the set of rational functions, i.e., quotients of polynomials in $\mathbb{F}[z]$.

Definition (Algebraic independence and transcendence degree).

- We say that the numbers $\alpha_1, ..., \alpha_n$ are algebraically dependent over $\mathbb{F}$ if they solve some polynomial $P(z_1, ..., z_n) \in \mathbb{F}[z_1, ..., z_n]$. Otherwise, we say that they are algebraically independent over $\mathbb{F}$.
- The transcendence degree of $\alpha_1, ..., \alpha_n$ over $\mathbb{F}$, denoted as $\deg_{\mathbb{F}}(\alpha_1, ..., \alpha_n)$, is the largest number of elements among $\alpha_1, ..., \alpha_n$ that are algebraically independent over $\mathbb{F}$.

Remark (Algebraic independence and linear independence). It is obvious that algebraic independence is stronger than linear independence. That is, a set of numbers that are algebraically independent must be linearly independent. But the reverse is not necessarily true.

Consider a n-th order homogeneous linear differential equation with coefficients in $\mathbb{C}(z)$:

$$u^{(n)}(z) + p_{n-1}(z)u^{(n-1)}(z) + \cdots + p_1(z)u'(z) + p_0(z)u(z) = 0, \tag{10.1}$$

in which $p_k \in \mathbb{C}(z)$ for $k = 0, ..., n-1$.

Definition (Regular points). Denote by $T(z)$ the common denominator of $p_{n-1}(z), ..., p_0(z)$ in (10.1). If $T(\alpha) \neq 0$, then we say α is a regular point of (10.1), and a singular point otherwise.

Theorem (Cauchy-Kowalevski theorem). *Suppose that α is a regular point of* (10.1). *Then the following holds.*

(i). The solutions in $\mathbb{C}[[z-\alpha]]$ form a n-dim vector space.
(ii). The solutions to the differential equation in $\mathbb{C}[[z-\alpha]]$ all have a positive radius of convergence.
(iii). The map

$$u(z) \to (u(\alpha), u'(\alpha)..., u^{(n-1)}(\alpha))$$

is an isomorphism between the vector space of solutions and the function values at α.

As an immediate consequence of the Cauchy-Kowalevski theorem, if α is a regular point of (10.1) and $u \not\equiv 0$, then $(u(\alpha), u'(\alpha)..., u^{(n-1)}(\alpha)) \neq (0, 0, ..., 0)$.

Notice that the regularity condition in the Cauchy-Kowalevski theorem is only sufficient to imply that the solutions in $\mathbb{C}[[z-\alpha]]$ form a n-dim vector space. In particular, the same conclusion may hold at certain singular points.

Definition (Apparent singularities). Suppose that α is a singular point of (10.1). If the solutions to the differential equation in $C[[z-\alpha]]$ form a n-dim vector space, then we say that α is an apparent singularity of (10.1).

THEOREM 10.3. *Suppose that the solutions to* (10.1) *form a* n-dim *vector space in* $\mathbb{C}[[z-\alpha]]$ *and all vanish at* α. *Then* α *is an apparent singularity of* (10.1).

PROOF. Suppose that α is regular. Then by Cauchy-Kowalevski theorem, the map

$$u(z) \to (u(\alpha), u'(\alpha)..., u^{(n-1)}(\alpha))$$

is an isomorphism between the vector space of solutions and the function values at α. Since all the solutions vanish at α, $(u(\alpha), u'(\alpha)..., u^{(n-1)}(\alpha)) = (0, 0, ..., 0)$. But this means that $u \equiv 0$, which contradicts with the fact that there is a n-dim solution space. ☐

Definition (E-functions). An entire function $f(z)$ given by a power series

$$f(z) = \sum_{j=1}^{\infty} \frac{a_j}{j!} z^j,$$

with $a_j \in \overline{\mathbb{Q}}$ for all $j \in \mathbb{N}$, is called an E-function if the following conditions hold.

(i). $f(z)$ solves a linear differential equation with coefficients in $\overline{\mathbb{Q}}[z]$.

(ii). Both $|\overline{a}_n|$ and the common denominator deno$(a_0, ..., a_n)$ are bounded by C^n for some $C > 0$.

The minimal differential equation of an E-function f is the one with the lowest degree such that f solves.

Example (E-functions).

(1). The exponential function

$$e^z = \sum_{j=1}^{\infty} \frac{1}{j!} z^j$$

has the minimal differential equation $u' - u = 0$.

(2). The sine function

$$\sin z = \sum_{j=0}^{\infty} \frac{(-1)^j}{(2j+1)!} z^j$$

has the minimal differential equation $u'' + u = 0$.

(3). The Bessel function

$$J_k(z) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!\Gamma(k+j+1)} \left(\frac{z}{2}\right)^{k+2j}$$

has the minimal differential equation

$$u'' + \frac{1}{z}u' + \left(1 - \frac{k^2}{z^2}\right) u = 0.$$

(4). The Legendre polynomial

$$P_k(z) = \frac{d^k}{dx^k}(1-z^2)^k$$

has the minimal differential equation

$$u'' - \frac{2z}{1-z^2}u' + \frac{k(k+1)}{1-z^2}u = 0.$$

Theorem. *The family of E-functions is a ring over* $\overline{\mathbb{Q}}$.

THEOREM 10.4 (Siegel, Shidlowski). *Let* $f = (f_1, ..., f_n)^{\mathrm{T}}$ *be a vector of E-functions and* $A \in \mathrm{M}_n(\overline{\mathbb{Q}}(z))$ *(i.e., A is a $n \times n$ matrix with entries in $\overline{\mathbb{Q}}(z)$) such that* $\vec{f'} = Af$. *Suppose that* $\alpha \in \overline{\mathbb{Q}}$ *and* $\alpha T(A) \neq 0$, *in which $T(A)$ is the common denominator of all entries of A. Then*

$$\deg_{\overline{\mathbb{Q}}}(f_1(\alpha), ..., f_n(\alpha)) = \deg_{\overline{\mathbb{Q}}[z]}(f_1, ..., f_n).$$

In particular, the function values $f_1(\alpha), ..., f_n(\alpha)$ are algebraically independent over $\overline{\mathbb{Q}}$ iff the functions $f_1, ..., f_n$ are algebraically independent over $\overline{\mathbb{Q}}[z]$.

However, this theorem does not include the linear dependence of the functions and the functions values. This is answered by Beukers[i].

THEOREM 10.5 (Beukers). *Let $f = (f_1, ..., f_n)^{\mathrm{T}}$ be a vector of E-functions and $A \in \mathrm{M}_n(\overline{\mathbb{Q}}(z))$ such that $\vec{f'} = Af$. Suppose that $\alpha \in \overline{\mathbb{Q}}$ and $\alpha T(A) \neq 0$. Then*

$$\operatorname{rank}_{\overline{\mathbb{Q}}}(f_1(\alpha), ..., f_n(\alpha)) = \operatorname{rank}_{\overline{\mathbb{Q}}[z]}(f_1, ..., f_n).$$

In particular, the function values $f_1(\alpha), ..., f_n(\alpha)$ are linearly independent over $\overline{\mathbb{Q}}$ iff the functions $f_1, ..., f_n$ are linearly independent over $\overline{\mathbb{Q}}[z]$.

Beukers' proof relies on André's program[ii], which include the following theorem.

THEOREM 10.6 (André). *Let f be an E-function. Then f solves a differential equation of the form*

$$z^m u^{(m)} + z^{m-1} q_{m-1} u^{(m-1)} + \cdots + z q_1 u' + q_0 u = 0,$$

in which $q_k \in \overline{\mathbb{Q}}[z]$ have degree no greater than $m - k$. That is, f solves a differential equation with coefficients in $\overline{\mathbb{Q}}[z]$ and no singularities other than 0 and ∞.

Remark. André's theorem asserts that an E-function solves a linear differential equation with only 0 and ∞ as singularities. However, the differential equation in André's theorem needs not be the minimal differential equation of the E-function. For example, let

$$f(z) = (z - 1)e^z.$$

Then the minimal differential equation of f is $(z - 1)u' - zu = 0$ (with a singularity $\alpha = 1$), while the differential equation in André's theorem is $u'' - 2u' + u = 0$ (with no singularity).

Rather than proving Theorem 10.5 following André's program, we present some easy consequence of Theorem 10.6.

THEOREM 10.7. *Let f be an E-function with rational coefficients. Suppose that $f(1) = 0$. Then the minimal differential equation of f has an apparent singularity at $\alpha = 1$.*

Remark. Take $f(z) = (z - 1)e^z$ as an example. The minimal differential equation of f is

$$u' - \frac{z}{z - 1} u = 0,$$

which has an apparent singularity at $\alpha = 1$.

[i] F. Beukers, *A refined version of the Siegel-Shidlovskii theorem.* Ann. of Math. (2) 163 (2006), no. 1, 369–379.

[ii] Y. André, *Séries Gevrey de type arithmétique* I, II, Ann. of Math . 151 (2000), 705–740, 741–756.

Proof. Let
$$g(z)=\frac{f(z)}{z-1}.$$
One can check that g is also an E-function. By André's theorem, g satisfies a differential equation without singularity at $\alpha=1$. By Cauchy-Kowalevski theorem, there is a basis of solutions which are analytic at $\alpha=1$. Multiplying this solution by $(z-1)$ generates a solution basis to the minimal differential equation for f. Notice that these solutions all vanish at $\alpha=1$. So $\alpha=1$ is an apparent singularity by Theorem 10.3. ☐

We use the above corollary to prove that π is transcendental.

Theorem. *π is transcendental.*

Proof. Suppose that $\gamma=2\pi i\in\overline{\mathbb{Q}}$. Let $\gamma_1=\gamma,\gamma_2,...,\gamma_n$ are all the conjugates of γ. That is, they are all the roots of the minimal polynomial of γ. Consider the E-function
$$f(z)=\prod_{k=1}^{n}\left(e^{\gamma_k z}-1\right).$$
Notice that f has coefficients in $\mathbb{Q}$. This is because $\gamma_1,\gamma_2,...,\gamma_n$ are all the roots to a polynomial with integral coefficients.

The minimal differential equation for f is
$$(d_z-\gamma_1)\cdots(d_z-\gamma_n)u=0,$$
in which the coefficients are all constants and therefore have no singularity. This contradicts with Theorem 10.7 since $f(1)=0$ and the minimal differential equation must have an apparent singularity at $\alpha=1$. ☐

Another application of André's theorem is to prove

Theorem 10.8 (Lindemann, Weierstrass). *Let $\gamma_1,...,\gamma_n$ be distinct algebraic numbers. Then the numbers $e^{\gamma_1},...,e^{\gamma_n}$ are linearly independent over $\overline{\mathbb{Q}}$.*

Remark. Choose $\gamma_1=0$ and $\gamma_2=1$. Then by the above theorem, $e^0=1$ and $e^1=e$ are linearly independent over $\overline{\mathbb{Q}}$, which means that $e\notin\overline{\mathbb{Q}}$.

Proof. Suppose that
$$c_1e^{\gamma_1}+\cdots+c_ne^{\gamma_n}=0,$$
in which $c_1,...,c_n$ are algebraic and are not all zero. Consider the E-function
$$f(z)=\prod_{\tilde{\bullet}\text{ are all conjugates of }\bullet}\left(\tilde{c}_1e^{\tilde{\gamma}_1 z}+\cdots+\tilde{c}_ne^{\tilde{\gamma}_n z}\right).$$
Notice that f coefficients in Q and vanishes at $\alpha=1$.

The minimal differential equation for f has constant coefficients, leading to contradiction that it must have an apparent singularity at $\alpha=1$ by Theorem 10.7. ☐

10.20. Spectrum and resolvent

In this section, we visit the definitions of spectrum and resolvent in the most general setting.

Definition (Vector spaces). Let $\mathbb{F}$ be a field. We call X a vector space over $\mathbb{F}$ if it satisfies the following conditions.

(i). [Addition] For all $x, y, z \in X$,

$$x + y = y + x \quad \text{and} \quad x + (y + z) = (x + y) + z.$$

Moreover, X contains a unique vector 0 such that $x + 0 = x$ for every $x \in X$ and to each $x \in X$ corresponds a unique vector $-x$ such that $x + (-x) = 0$.

(ii). [Multiplication] For all $\alpha, \beta \in \mathbb{F}$ and $x, y \in X$,

$$1x = x, \quad \alpha(\beta x) = (\alpha\beta)x, \quad \alpha(x + y) = \alpha x + \alpha y, \quad (\alpha + \beta)x = \alpha x + \beta x.$$

Definition (Seminorm and norm). Let $\mathbb{F}$ be a field and X be a vector space over $\mathbb{F}$. A seminorm X is a function $x \to \|x\|$ from X to $[0, \infty)$ such that

(i). [The triangle inequality] For all $x, y \in X$,

$$\|x + y\| \leq \|x\| + \|y\|.$$

(ii). For all $\alpha \in \mathbb{F}$ and $x \in X$,

$$\|\lambda x\| = |\lambda|\|x\|.$$

A seminorm is called a norm if

(iii). $\|x\| = 0$ only when $x = 0$.

A vector space equipped with a norm is called a normed vector space.

Definition (Metric spaces). A metric on a set X is a function $d : X \times X \to [0, \infty)$ such that

(i). $d(x, y) = 0$ iff $x = y$.
(ii). $d(x, y) = d(y, x)$ for all $x, y \in X$.
(iii). $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

A set equipped with a metric is called a metric space.

Remark. Every normed vector space may be regarded as a metric space, in which the distance $d(x, y) := \|x - y\|$.

Definition (Banach spaces). A Banach space is a normed vector space which is complete in the metric defined by its norm.

Remark. The collection of all bounded linear transformations of X into Y, $\mathcal{L}(X, Y)$, is a normed space if X and Y are normed spaces, and is a Banach space is Y is a Banach space. See Proposition 5.4 in Folland[i].

Definition (Complex algebras and Banach algebras). A complex algebra is a vector space A over the complex field $\mathbb{C}$ in which a multiplication is defined that satisfies

(i). $x(yz) = (xy)z$ for all $x, y, z \in A$.
(ii). $(x + y)z = xz + yz$ and $x(y + z) = xy + xz$ for all $x, y, z \in A$.
(iii). $\alpha(xy) = (\alpha x)y = x(\alpha y)$ for all $\alpha \in \mathbb{C}$ and $x, y \in A$.

A complex algebra is called a Banach algebra if A is a Banach space with respect to a norm $\|\cdot\|$ that satisfies

(i). $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in A$.
(ii). there is a unique $e \in A$ such that $xe = ex = x$ for all $x \in A$.
(iii). $\|e\| = 1$.

Note that we do not require that A to be commutative.

---

[i] G. Folland, *Real analysis*. Second edition. John Wiley & Sons, Inc., New York, NY, 1999.

Definition (Spectrum and resolvent in a Banach algebra). Let A be a Banach algebra. Denote $G = G(A)$ to be the set of all invertible elements of A. (Notice that if $x, y \in G$, then $y^{-1}x$ is the inverse of $x^{-1}y$. Thus $x^{-1}y \in G$ and G is a group.)

If $x \in A$, the spectrum $\sigma(x)$ of x is the set of all complex numbers λ such that $\lambda e - x$ is not invertible. The complement of $\sigma(x)$ is the resolvent set of x; it consists of all $\lambda \in \mathbb{C}$ for which $(\lambda e - x)^{-1}$ exists.

Remark. Suppose that X is a Banach space. Then $\mathcal{L}(X)$ is a Banach algebra. In particular, one can define the product ST for $S, T \in \mathcal{L}(X)$. An operator $T \in \mathcal{L}(X)$ is said to be invertible if there exists $S \in \mathcal{L}(X)$ such that $ST = \mathrm{I} = TS$. The above definitions of spectrum and resolvent apply to $\mathcal{L}(X)$.

Definition (Spectrum and resolvent in $\mathcal{L}(X)$). The spectrum $\sigma(T)$ of an operator $T \in \mathcal{L}(X)$ is the set of all scalars λ such that $T - \lambda\mathrm{I}$ is not invertible. Thus $\lambda \in \sigma(T)$ iff at least one of the following two conditions hold.

(i). The range of $T - \lambda\mathrm{I}$ is not all of X.
(ii). The map $T - \lambda\mathrm{I}$ is not one-to-one. In this case, λ is said to be an eigenvalue of T. The corresponding eigenspace is $\{x \in X : Tx = \lambda x\}$.

We further divide $\sigma(T)$ into three disjoint pieces:

(i). The point spectrum $\sigma_p(T)$ consists of all $\lambda \in \mathbb{C}$ for which $T - \lambda\mathrm{I}$ is not one-to-one. That is $\sigma_p(T)$ is the set of all eigenvalues of T.
(ii). The continuous spectrum $\sigma_c(T)$ consists of all $\lambda \in \mathbb{C}$ for which $T - \lambda\mathrm{I}$ is a one-to-one map of X to a dense proper subspace of X.
(iii). The residual spectrum $\sigma_r(T)$ consists of all other $\lambda \in \sigma(T)$.

10.21. $TT^\star$ argument

For an integral operator T, the arguments of $TT^\star$ and $T^\star T$ allows convenient transitions of the mapping norms of T, $T^\star$, $TT^\star$, and $T^\star T$, between L^p spaces. Each of these estimates may have its advantages depending on the scenario. We compute explicitly their integral kernels in this section.

Theorem. *Let T be an integral operator that is defined by*

$$T : f(x) \to \int K(x, y) f(y)\, dy.$$

Then the adjoint $T^\star$ is also an integral operator and is defined by

$$T^\star : f(x) \to \int \overline{K(y, x)} f(y)\, dy.$$

Moreover,

$$TT^\star : f(x) \to \int \left(\int K(x, z) \overline{K(y, z)}\, dz \right) f(y)\, dy$$

and

$$T^\star T : f(x) \to \int \left(\int \overline{K(z, x)} K(z, y)\, dz \right) f(y)\, dy.$$

Proof. Let

$$Tf(x) = \int K(x, y) f(y)\, dy.$$

Then

$$\begin{aligned}\int Tf(x)\overline{g(x)}\,dx &= \int\left(\int K(x,y)f(y)\,dy\right)\overline{g(x)}\,dx\\ &= \int f(y)\overline{\left(\int \overline{K(x,y)}g(x)\,dx\right)}\,dy\\ &= \int f(y)\overline{T^\star g(y)}\,dy,\end{aligned}$$

in which

$$T^\star g(y) = \int \overline{K(x,y)}g(x)\,dx,$$

that is,

$$T^\star f(x) = \int \overline{K(y,x)}f(y)\,dy.$$

To compute the kernel for $TT^\star$,

$$\begin{aligned}TT^\star f(x) &= \int K(x,z)T^\star f(z)\,dz\\ &= \int K(x,z)\left(\int \overline{K(y,z)}f(y)\,dy\right)\,dz\\ &= \int\left(\int K(x,z)\overline{K(y,z)}\,dz\right)f(y)\,dy,\end{aligned}$$

which says that $TT^\star$ has kernel

$$K_{TT^\star}(x,y) = \int K(x,z)\overline{K(y,z)}\,dz.$$

Similarly,

$$\begin{aligned}T^\star Tf(x) &= \int \overline{K(z,x)}Tf(z)\,dz\\ &= \int \overline{K(z,x)}\left(\int K(z,y)f(y)\,dy\right)\,dz\\ &= \int\left(\int \overline{K(z,x)}K(z,y)\,dz\right)f(y)\,dy,\end{aligned}$$

which says that $T^\star T$ has kernel

$$K_{T^\star T}(x,y) = \int \overline{K(z,x)}K(z,y)\,dz = \int K(z,y)\overline{K(z,x)}\,dz.$$

□

Theorem. *Let* $1 < p, q < \infty$.

$$\|T\|_{L^p\to L^q} = \|T^\star\|_{L^{q'}\to L^{p'}}.$$

PROOF. For each $f \in L^p$,

$$\|Tf\|_q = \left|\sup_{g:\|g\|_{q'}=1}\int Tf\overline{g}\right| = \left|\sup_{g:\|g\|_{q'}=1}\int f\overline{T^\star g}\right| \le \|f\|_p\|T^\star g\|_{p'} \le \|f\|_p\|T^\star\|_{L^{q'}\to L^{p'}}$$

implies that

$$\|T\|_{L^p\to L^q} \le \|T^\star\|_{L^{q'}\to L^{p'}}.$$

Similarly,

$$\|T^\star\|_{L^{q'}\to L^{p'}} \le \|T\|_{L^p\to L^q}.$$

The theorem then follows. □

Theorem. *Let $1 < p < \infty$. Then*

$$\|T^\star T\|_{L^p\to L^{p'}} = \|T\|^2_{L^p\to L^2} = \|T^\star\|^2_{L^2\to L^{p'}}.$$

PROOF. We only show that $\|T^\star T\|_{L^p\to L^{p'}} = \|T\|^2_{L^2\to L^p}$. Since

$$\|T^\star T\|_{L^p\to L^{p'}} \le \|T^\star\|_{L^2\to L^{p'}}\|T\|_{L^p\to L^2} = \|T\|^2_{L^p\to L^2},$$

it suffices to show that

$$\|T\|^2_{L^p\to L^2} \le \|T^\star T\|_{L^p\to L^{p'}}.$$

Taking $g = f/\|f\|_p$ in the following inequalities,

$$\|T^\star Tf\|_{p'} = \left|\sup_{g:\|g\|_p=1}\int T^\star Tf\overline{g}\right| \ge \left|\int T^\star Tf\overline{\frac{f}{\|f\|_p}}\right| \ge \left|\int TfT\overline{\frac{f}{\|f\|_p}}\right| = \frac{\|Tf\|_2^2}{\|f\|_p}.$$

We then have that

$$\frac{\|T^\star Tf\|_{p'}}{\|f\|_p} \ge \left(\frac{\|Tf\|_2}{\|f\|_p}\right)^2.$$

Therefore,

$$\|T^\star T\|_{L^p\to L^{p'}} \ge \|T\|^2_{L^p\to L^2}.$$

□

Remark. See Lemma 7.3 in Wolff[i] for the $TT^\star$ argument applying to the Fourier restriction operators.

10.22. Weak derivatives and differentiation of distributions

Question 10.5 (Weak derivatives and differentiation of distributions). *What is the relation between weak derivatives and differentiation of distributions?*

Answer. A function $v \in L^1_{\text{loc}}(\Omega)$ is the α-th weak derivative of $u \in L^1_{\text{loc}}(\Omega)$ if

$$\int_\Omega v\varphi\,dx = (-1)^{|\alpha|}\int_\Omega u\partial^\alpha\varphi\,dx \quad \text{for all } \varphi \in C_0^{|\alpha|}(\Omega).$$

Clearly, $u \in L^1_{\text{loc}}(\Omega)$ defines a distribution in $\mathcal{D}'(\Omega)$ via $\varphi \in C_0^\infty(\Omega) \to \int_\Omega u\varphi\,dx$. In this case, the above definition coincides with the differentiation of the distribution $u \in \mathcal{D}'(\Omega)$.

Definition (Sobolev spaces). Let $k \in \mathbb{N}$. The Sololev space $W^{k,p}(\Omega) \subset L^1_{\text{loc}}(\Omega)$ is defined as the space of functions u such that for each α with $|\alpha| \le k$, $\partial^\alpha u$ exists in the weak sense and belongs to $L^p(\Omega)$. The $W^{k,p}(\Omega)$ norm is defined as

$$\begin{cases} \|u\|_{W^{k,p}(\Omega)} = \left(\int_\Omega \sum_{|\alpha|\le k}|\partial^\alpha u|^p\,dx\right)^{\frac{1}{p}} & \text{if } 1 \le p < \infty \\ \|u\|_{W^{k,p}(\Omega)} = \sum_{|\alpha|\le k}\operatorname{esssup}|\partial^\alpha u| & \text{if } p = \infty. \end{cases}$$

[i] T. Wolff, *Lectures on harmonic analysis.* American Mathematical Society, Providence, RI, 2003.

See Section 5.2 in Evans[i]. When $p = 2$ and $\Omega = \mathbb{R}^n$, one can use Fourier transform to provide an alternative definition of the Sobolev spaces: Given $s \in \mathbb{R}$, $H^s(\mathbb{R}^n)$ denotes the space of all $u \in \mathcal{S}'(\mathbb{R}^n)$ such that

$$\|u\|_{H^s(\mathbb{R}^n)} = \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} |\hat{u}(\xi)|^2 (1 + |\xi|^2)^s \, d\xi \right)^{\frac{1}{2}} \approx \left\| \hat{u}(\xi)(1 + |\xi|^2)^{\frac{s}{2}} \right\|_{L^2(\mathbb{R}^n)} < \infty.$$

Then $W^{k,2}(\mathbb{R}^n) = H^k(\mathbb{R}^n)$ for $k \in \mathbb{N}$, see Section 5.8.5 in Evans. There are several advantages of the Fourier approach to the Sobolev spaces, for example,

(1). One can allow $s \in \mathbb{R}$, and the corresponding space $H^s(\mathbb{R}^n)$ is sometimes called fractional Sobolev spaces.

(2). One can use Hardy-Littlewood-Sobolev inequality to prove the Sobolev embedding theorem:

$$\begin{cases} \|u\|_{L^q(\mathbb{R}^n)} \le C\|u\|_{H^s(\mathbb{R}^n)} \text{ with } \frac{1}{q} = \frac{1}{2} - \frac{s}{n} & \text{if } s < \frac{n}{2} \\ \|u\|_{L^\infty(\mathbb{R}^n)} \le C\|u\|_{H^s(\mathbb{R}^n)} & \text{if } s > \frac{n}{2}. \end{cases}$$

That is, $H^s(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ and $H^s(\mathbb{R}^n) \subset L^\infty(\mathbb{R}^n)$, respectively. See Section 0.3 in Sogge[ii], in which the Fourier approach to $H^s_p(\mathbb{R}^n)$ for $p \neq 2$ is also discussed.

Tao has a particularly nice explanation of Sobolev embedding theorem: This theorem allows one to trade regularity for integrability. However, the trade is not reversible: One cannot start with a function with a lot of integrability and no regularity, and expect to recover regularity in a space of lower integrability. In fact, one can already see this trade with the most basic example of Sobolev embedding, coming from the fundamental theorem of calculus: If $f \in C^1(\mathbb{R})$ and $f' \in L^1(\mathbb{R})$, then we have that $f \in L^\infty(\mathbb{R})$; but the converse is far from true.

[i]L. C. Evans, *Partial differential equations.* Second edition. American Mathematical Society, Providence, RI, 2010.

[ii]C. Sogge, *Fourier integrals in classical analysis.* Second edition. Cambridge University Press, Cambridge, 2017.