NOTES FOR LECTURES AT TSINGHUA DAXUE
VERSION 2

SEMYON DYATLOV

Abstract. These are very schematic notes for lectures given at Tsinghua in July 2016.

Setting: convex co-compact hyperbolic surfaces. I have chosen this setting because it appears in so many different areas of mathematics and gives rise to interesting problems for resonances, many of which still have not been answered. I will not talk about Pollicott–Ruelle resonances since covering both kinds of resonances in four lectures is likely to seriously confuse the audience. However, Pollicott–Ruelle resonances will be discussed in upcoming lectures of Long Jin.

Plan:

(1) Convex co-compact hyperbolic surfaces: the geometric approach.
(2) Spectral theory of $-\Delta_g$.
(3) Meromorphic continuation of $(-\Delta_g - \frac{1}{4} - \lambda^2)^{-1}$ and *resonances*.
(4) Applications to wave equation. A teaser of the results: spectral gaps and counting.
(5) The geodesic flow: hyperbolicity and the trapped set.
(6) Convex co-compact hyperbolic surfaces: the algebraic approach. Example: a three-funnel surface.
(7) The limit set, its fractal structure, and relation to the trapped set.
(8) Spectral gap: the standard Patterson–Sullivan result.
(9) A semiclassical study of resonances. An explanation of the spectral gap using uncertainty principle.
(10) Resonance counting and an explanation via the uncertainty principle.

References: I mostly restrict myself to one reference per topic, usually coauthored by myself, to avoid overloading the listeners by a long list of papers. Many of these papers are quite technical, which explains the reference to the upcoming paper with Long Jin. Obviously, these problems have a long history of study by many mathematicians and I refer the listeners to the books and introductions to the papers below for overviews of the history of the subject.

- [Bo07] on hyperbolic surfaces, with focus on resonances for the infinite area case;

- [Bo14, BoWe] for some beautiful pictures of resonances and a numerical study;
- [DyZa] for the latest result on spectral gaps via the uncertainty principle;
- [Dy] for the latest result on Weyl laws;
- [No] for an overview of open quantum chaos;
- an upcoming paper with Long Jin on uncertainty principles/Weyl laws in a simpler to absorb setting of open quantum maps;
- the book in progress [DyZw] for an extensive introduction to the mathematical theory of scattering resonances.

1. Geometry of hyperbolic surfaces

I consider a hyperbolic surface, that is a complete connected oriented Riemannian surface

$$(M, g), \quad \dim M = 2,$$

whose Gaussian curvature is equal to -1 everywhere.

There are plenty of compact hyperbolic surfaces but in these lectures I focus on the non-compact case. All noncompact geometrically finite hyperbolic surfaces can be broken into a compact part and finitely many infinite ends, and the infinite ends (with two minor exceptions) can be broken into two types: *funnels* and *cusps*.

To explain what funnels and cusps are, we consider the stretched product model (a manifold with boundary)

$$M = [0, \infty)_r \times \mathbb{S}^1_\theta, \quad \mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}, \quad g = dr^2 + f(r)^2\, d\theta^2. \tag{1.1}$$

Then M has curvature -1 if and only if $f'' = f$. We can now define funnels and cusps:

- a *funnel* of neck length $\ell > 0$ is given by (1.1) with the metric
$$g = dr^2 + \Big(\frac{\ell}{2\pi}\Big)^2 \cosh^2 r\, d\theta^2.$$
- a *cusp* is given by (1.1) with the metric (for some $a > 0$)
$$g = dr^2 + a^2 e^{-2r}\, d\theta^2.$$

Thus funnels are very wide and cusps are very narrow; in particular, a cusp has a finite area while a funnel does not. Historically, the first setting studied is *finite area* surfaces, which have only cusps but no funnels. We will henceforth only study the opposite case *convex co-compact hyperbolic surfaces*, defined as follows:

Definition 1.1. *A convex co-compact hyperbolic surface is a noncompact hyperbolic surface with no cusps, i.e. only funnel ends.*

We note that each funnel is bordered by a closed geodesic $\{r = 0\}$ of length ℓ, which we call its *neck*. In general a convex co-compact hyperbolic surface can be obtained as follows: one starts with a compact hyperbolic surface with totally geodesic boundary (which is called the *convex core* of the resulting noncompact surface) and attaching a funnel end along each boundary geodesic. See for instance [Bo07, Section 2.4].

A basic example of a convex co-compact hyperbolic surface is given by the *hyperbolic cylinder with neck of length* $2\pi\ell$:

$$M = \mathbb{R}_r \times \mathbb{S}^1_\theta, \quad g = dr^2 + \ell^2 \cosh^2 r\, d\theta^2, \tag{1.2}$$

which has two funnels and a degenerate convex core.

2. Spectral theory of the Laplacian

Let (M, g) be a convex co-compact hyperbolic surface. Denote by Δ_g the Laplace–Beltrami operator on M. We know that $-\Delta_g$ admits a nonnegative self-adjoint extension to an operator on $L^2(M)$, and we are interested in the corresponding spectrum, that is the set of points $z \in \mathbb{C}$ where the L^2 resolvent

$$(-\Delta_g - z)^{-1} : L^2(M) \to H^2(M)$$

is not well-defined.

For a compact manifold, the operator $-\Delta_g$ has discrete spectrum given by eigenvalues. Thus to understand the nature of the spectrum for the noncompact case we need to see what happens in a funnel. We use the example of the hyperbolic cylinder (1.2) with $\ell = 2\pi$, where

$$-\Delta_g = -\partial_r^2 - \tanh r\, \partial_r - \cosh^2 r\, \partial_\theta^2.$$

Every $u \in L^2(M)$ can be expanded in Fourier series

$$u(r, \theta) = \sum_{k \in \mathbb{Z}} u_k(r) e^{ik\theta},$$

and the operator $-\Delta_g$ acts on Fourier coefficients as follows:

$$-\Delta_g u = \sum_{j \in \mathbb{Z}} P_k u_k(r) e^{ik\theta},$$

where the differential operators P_k on $\mathbb{R}_r$ are given by

$$P_k = -\partial_r^2 - \tanh r\, \partial_r + k^2 \cosh^{-2} r.$$

The spectrum of $-\Delta_g$ is the union of the spectra of all the operators P_k.

To understand the spectrum of each P_k, we need to analyse the asymptotic behavior as $r \to \pm\infty$ of solutions to the equation $(P_k - z)v = 0$, for $z \in \mathbb{C}$. Looking at the expansions of the coefficients as $r \to \pm\infty$, we see that this equation becomes

$$\begin{cases} (-\partial_r^2 - \partial_r - z)v = \mathcal{O}(e^{-2|r|}v), & r \gg 1, \\ (-\partial_r^2 + \partial_r - z)v = \mathcal{O}(e^{-2|r|}v), & -r \gg 1. \end{cases}$$

On the left-hand side are constant coefficient ordinary differential operators, whose solutions are given by $e^{\lambda_\pm r}$ for $r \gg 1$ and by $e^{-\lambda_\pm r}$ for $-r \gg 1$ where $\lambda_\pm$ are the roots of the quadratic equation (the indicial equation)

$$-\lambda^2 - \lambda - z = 0.$$

We compute the roots $\lambda_\pm$ (the indicial roots):

$$\lambda_\pm = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - z}. \tag{2.1}$$

Except at a discrete set of values of z (that we are not going to worry about here) the equation $(P_k - z)v = 0$ has two solutions $v_\pm^k$ with asymptotic behavior

$$v_\pm^k(r) = e^{\lambda_\pm r}(1 + \mathcal{O}(e^{-2r})), \quad r \to +\infty. \tag{2.2}$$

To see this, one can write $v_\pm$ as a series

$$v_\pm^k(r) = e^{\lambda_\pm r} \sum_{j=0}^\infty a_{j,\pm} e^{-2jr}, \quad a_{0,\pm} = 1,$$

solve for the coefficients[1] $a_{j,\pm}$, and show that the series converges for large positive r. We see also that $v_\pm^k(-r)$ are solutions to the equation $(P_k - z)v = 0$ with prescribed asymptotic behavior as $r \to -\infty$.

The volume form on the hyperbolic cylinder is given by $d\,\mathrm{vol}_g = \cosh r\, dr d\theta$. This means that

$$e^{\lambda r} \in L^2(M_+) \iff \operatorname{Re} \lambda < -\frac{1}{2}, \quad \text{where } M_+ := \{r \geq 0\} \subset M.$$

From (2.1) we see that at least one of the functions $v_\pm^k$ does not lie in $L^2(M_+)$. Then there are two cases:

(1) $z \in [1/4, \infty)$, when $\operatorname{Re} \lambda_\pm = -1/2$ and neither v_+^k nor v_-^k lies in $L^2(M_+)$;
(2) $z \notin [1/4, \infty)$, when exactly one of the solutions $v_\pm^k$ lies in $L^2(M_+)$.

In case (1), we see that the equation $(P_k - z)v = f$ does not have a solution in L^2 for a generic $f \in C_0^\infty(\mathbb{R})$. Thus z lies in the spectrum. On the other hand, the equation $(P_k - z)v = 0$ does not have any L^2 solutions. Thus z is not an eigenvalue. We see that $[1/4, \infty)$ lies in the purely continuous spectrum of $-\Delta_g$.

In case (2), to fix notation assume that $v_+^k \in L^2(M_+), v_-^k \notin L^2(M_+)$. Then there are two subcases:

(1) $v_+^k(r)$ and $v_+^k(-r)$ are multiples of each other for some k. Then $v_+^k(r)e^{ik\theta}$ is an L^2 eigenfunction for $-\Delta_g$ at z.

[1]It is here that the exceptional values of z come up – for some values one ends up with division by zero when trying to solve for the coefficients. The problem can be resolved by replacing the leading term in (2.2) by $e^{\lambda_\pm r} e^{-2\lambda J r}$ for an appropriately chosen $J \in \mathbb{N}$.

(2) $v_+^k(r)$ and $v_+^k(-r)$ are not multiples of each other for all k. Then for each $f \in L^2(M, d\,\mathrm{vol}_g)$ there is a unique L^2 solution to the equation $(-\Delta_g - z)u = f$, so z does not lie in the spectrum.

Moreover, subcase (1) can only happen for $z \in (0, 1/4)$ since the operator $-\Delta_g$ is self-adjoint and nonnegative, and $1 \notin L^2(M)$. Since the resulting eigenfunctions have exponential decay, one can use elliptic regularity to show that there are only finitely many L^2 eigenvalues in $(0, 1/4)$. We summarize the discussion of this section in the following result:

Theorem 1. [Bo07, Theorem 7.1] *Let (M, g) be a convex co-compact hyperbolic surface. Then the L^2 spectrum of $-\Delta_g$ consists of:*

(1) *the continuous spectrum $[1/4, \infty)$; and*
(2) *finitely many (possibly none) eigenvalues in $(0, 1/4)$.*

3. Meromorphic continuation and resonances

We now introduce the main object of study in these lectures, namely *resonances*. There are many different ways to view these (in particular, they appear as zeroes of the Selberg zeta function, see [Bo07, Chapter 10] and are related to Pollicott–Ruelle resonances of the geodesic flow, see [GHW]) but here we will take the spectral approach, building on the previous section. Namely we explain the following

Theorem 2. [Bo07, Theorem 6.2] *Let (M, g) be a convex co-compact hyperbolic surface. Consider the L^2 resolvent*

$$R(\lambda) = \Big(-\Delta_g - \frac{1}{4} - \lambda^2 \Big)^{-1} : L^2(M) \to L^2(M), \quad \operatorname{Im}\lambda > 0. \tag{3.1}$$

Then $R(\lambda)$ admits a meromorphic continuation with poles of finite rank to an operator

$$R(\lambda) : L^2_{\mathrm{comp}}(M) \to L^2_{\mathrm{loc}}(M), \quad \lambda \in \mathbb{C},$$

where L^2_{comp} denotes the space of compactly supported L^2 functions and L^2_{loc} denotes the space of functions which are locally in L^2.

The poles of the meromorphic continuation $R(\lambda)$ are called ***resonances****. In the upper half-plane, resonances correspond to the eigenvalues of $-\Delta_g$ in $(0, 1/4)$. In the lower half-plane, the situation is more complicated as we will see later.*

Note that the parametrization (3.1) of the L^2 resolvent corresponds well to the information we have about the L^2 spectrum: the complement to the continuous spectrum, $\mathbb{C} \setminus [1/4, \infty)$, is mapped onto the upper half-plane and the meromorphic continuation takes place across the continuous spectrum, corresponding to $\lambda \in \mathbb{R}$.

To explain why meromorphic continuation holds, we go back to the example of the hyperbolic cylinder (1.2). We construct the restriction $R_k(\lambda)$ of the meromorphic continuation of $R(\lambda)$ to the k-th Fourier mode in θ. The full resolvent can be obtained as the direct sum of different $R_k(\lambda)$, though to actually prove meromorphic continuation one needs to analyse the asymptotics as $k\to\infty$, for instance using special functions.

Given $z=\frac{1}{4}+\lambda^2$, we choose the roots in (2.1) as follows:

$$\lambda_\pm=-\frac{1}{2}\pm i\lambda.$$

Fix $f\in C_0^\infty(\mathbb{R})$ and consider $v=R_k(\lambda)f$. We know that for $r\gg 1$ v is equal to a linear combination of the functions $v_\pm^k(r)$, which behave like $e^{i\lambda_\pm r}$ as $r\to+\infty$, and for $-r\gg 1$ it is equal to a linear combination of the functions $v_\pm^k(-r)$, which behave like $e^{-i\lambda_\pm r}$ as $r\to-\infty$.

For $\operatorname{Im}\lambda>0$, we have $e^{i\lambda_+r}\in L^2(M_+)$ and $e^{i\lambda_-r}\notin L^2(M_+)$. Therefore, we should define $v=R_k(\lambda)f$, for all λ, by the following conditions:

$$\Big(P_k-\frac{1}{4}-\lambda^2\Big)v=f, \tag{3.2}$$

$$v(r)=c_\pm v_+^k(\pm r),\quad \pm r\gg 1,\quad \text{for some } c_\pm\in\mathbb{C}. \tag{3.3}$$

Note that $v_+^k(r)$ depends holomorphically on λ. There are then two cases:

- For a discrete set of values of λ, the functions $v_+^k(r)$ and $v_+^k(-r)$ are multiples of each other. Such λ form the set of resonances.
- For all other values of λ, there exists a unique solution to (3.2), (3.3). This gives a solution operator $R_k(\lambda):L^2_{\mathrm{comp}}(\mathbb{R})\to L^2_{\mathrm{loc}}(\mathbb{R})$, holomorphic at λ.

We remark that for the hyperbolic cylinder example with $\ell=2\pi$ used here, resonances can be computed explicitly: they are given by [Bo07, Proposition 5.2]

$$\lambda_{jk}=k-\Big(j+\frac{1}{2}\Big)i,\quad j\in\mathbb{N}_0,\ k\in\mathbb{N}. \tag{3.4}$$

For the case of general convex co-compact hyperbolic surfaces, one way to obtain meromorphic continuation is as follows: we construct a model resolvent for each funnel, using for instance the hyperbolic cylinder model, and then piece these together with an elliptic parametrix near the convex core to obtain an inverse to $\Delta_g-\frac{1}{4}-\lambda^2$ modulo a compact remainder. Then meromorphic continuation of $R(\lambda)$ is obtained via analytic Fredholm theory. See for instance [Bo07, §§6.2, 6.3].

Another way to obtain the meromorphic continuation, developed in recent work of András Vasy, is as follows: for $f\in C_0^\infty(M)$, we define $u=R(\lambda)f$ by the following conditions

$$\begin{gathered}\Big(-\Delta_g-\frac{1}{4}-\lambda^2\Big)u=f,\\ e^{(1/2-i\lambda)r}u(r,\theta)=\tilde u(e^{-2r}),\quad \tilde u\text{ is smooth at }0.\end{gathered} \tag{3.5}$$

where the last condition, called the *outgoing condition*, is imposed at each funnel. One can use microlocal analysis to show that the above problem has a unique solution everywhere except resonances, and the solution operator gives the meromorphic continuation. See for instance [DyZw, Chapter 5] for details.

4. Connection to the wave equation

Let me now briefly explain one motivation for the study of resonances. (See [DyZw, §2.1] for a more basic example of wave equation in one dimension.) Take $F \in C_0^\infty((0,\infty)_t \times M_x)$ and consider the forward solution to the shifted wave equation

$$\begin{aligned} \Big(\partial_t^2 - \Delta_g - \frac{1}{4}\Big)U &= F, \\ U|_{t<0} &= 0. \end{aligned} \tag{4.1}$$

We are interested in what happens to $U(t,x)$ when $t \to \infty$, but only for bounded values of x.

Define the Fourier transform in time of $U(t,x)$:

$$\hat{U}(\lambda, x) = \int_0^\infty e^{it\lambda} U(t,x)\,dt \in L^2(M), \quad \operatorname{Im}\lambda \gg 1,$$

and the integral converges for $\operatorname{Im}\lambda$ large enough because $U(t,x)$ is at most exponentially growing (without the $1/4$ factor it would just be bounded and the integral would converge for $\operatorname{Im}\lambda > 0$). Taking the Fourier transform of the wave equation, we get

$$\Big(-\Delta_g - \frac{1}{4} - \lambda^2\Big)\hat{U}(\lambda) = \hat{F}(\lambda).$$

This implies that

$$\hat{U}(\lambda) = R(\lambda)\hat{F}(\lambda), \quad \operatorname{Im}\lambda \gg 1.$$

Now, the right-hand side admits a meromorphic continuation to $\lambda \in \mathbb{C}$ as an element of $L^2_{\mathrm{loc}}(M)$. Indeed, $\hat{F}(\lambda) \in L^2_{\mathrm{comp}}(M)$ is entire since F is compactly supported, and $R(\lambda) : L^2_{\mathrm{comp}}(M) \to L^2_{\mathrm{loc}}(M)$ admits a meromorphic continuation by Theorem 2.

Now, the decay properties of $\hat{U}(\lambda)$ are connected to the behavior of its Fourier transform as follows:

- if $U(t)$ is exponentially decaying on compact sets, that is $\chi U(t) = \mathcal{O}(e^{-\nu t})$ for all $\chi \in C_0^\infty(M)$ and some $\nu > 0$, then $\hat{U}(\lambda)$ is holomorphic on $\operatorname{Im}\lambda > -\nu$. In particular, resonances in $\operatorname{Im}\lambda > -\nu$ are obstructions to $\mathcal{O}(e^{-\nu t})$ exponential decay.
- if $R(\lambda)$ has a *spectral gap* of size $\nu > 0$, that is there are no resonances in the strip $\operatorname{Im}\lambda > -\nu$ and $R(\lambda)$ has a reasonable (polynomial in λ) norm bound as

$\lambda \to \infty$ in this strip, then deforming the contour to $\operatorname{Im}\lambda = -\nu$ in the Fourier inversion formula (valid for large enough $C > 0$)

$$U(t) = \frac{1}{2\pi}\int_{\operatorname{Im}\lambda = C} e^{-i\lambda} R(\lambda)\hat{F}(\lambda)\, d\lambda$$

we see that $U(t) = \mathcal{O}(e^{-\nu t})$ on compact sets.

- if $R(\lambda)$ has an *essential spectral gap* of size $\nu > 0$, that is there are finitely many resonances in the strip $\operatorname{Im}\lambda > -\nu$ and $R(\lambda)$ still obeys a reasonable bound as $\lambda \to \infty$ in the strip, then the above contour deformation argument together with the residue theorem gives a *resonance expansion* (written here in the case when $R(\lambda)$ has simple poles)

$$U(t) = \sum_{\substack{\operatorname{Im}\lambda \geq -\nu \\ \lambda \text{ resonance}}} e^{-it\lambda} v_\lambda(x) + \mathcal{O}(e^{-\nu t}), \quad t \to +\infty, \tag{4.2}$$

for some functions $v_\lambda(x)$ related to the residues of $R(\lambda)$.

Of course, the specific wave equation (4.1) does not have a clear physical interpretation. However, in the closely related case of Euclidean scattering by several obstacles resonances can be observed experimentally (see e.g. [BWPSKZ]). The existence of spectral gap/exponential decay of waves is often a crucial component in results on linear and nonlinear wave equations in various situations ranging from Euclidean scattering to the recent proof of full nonlinear stability of the Kerr–de Sitter black hole under Einstein equations [HV].

We now move on to better understanding the distribution of resonances. We will be interested in the high frequency, bounded decay régime

$$|\operatorname{Re}\lambda| \to \infty, \quad \operatorname{Im}\lambda > -\nu.$$

We will study the following two questions:

- Essential spectral gap: is there $\nu > 0$ such that there are only finitely many resonances in $\operatorname{Im}\lambda > -\nu$? A positive answer to this question gives a resonance expansion (4.2), and combined with analysis of bounded frequency resonances (which is often available) it may give exponential decay of waves.
- How many resonances are there in the region $[R, R+1]+i[-\nu, 0]$, when $R \to \infty$? The best answer would be the analogue of *Weyl law* for compact manifolds. It is believed that for large ν, the number of resonances should grow like R^δ where $\delta \in (0, 1)$ is the dimension of the limit set (see below). An upper bound of this form is available however no matching lower bounds exist so far.

The answers to both of these questions depend on the structure of the geodesic flow of M, more precisely on the set of all of its trapped trajectories. In particular, if the

Hausdorff dimension of the trapped set is equal to $2\delta + 1$ then essential gap is known to exist when $\delta \leq 1/2$ and δ is also the exponent expected in the Weyl law.

5. An algebraic approach to hyperbolic surfaces. Geodesic flow

We now want to understand the dynamics of the geodesic flow on a convex co-compact hyperbolic surface (M, g). For that it is convenient to introduce a different, algebraic, point of view on these surfaces. Namely, M can be viewed as a quotient

$$M = \Gamma \backslash \mathbb{H}^2$$

where

- $(\mathbb{H}^2, g)$ is the hyperbolic plane:
$$\mathbb{H}^2 = \{z = x + iy \in \mathbb{C} \mid y > 0\}, \quad g = \frac{dx^2 + dy^2}{y^2};$$
- $\Gamma \subset \mathrm{SL}(2, \mathbb{R})$ is a discrete subgroup, where $\mathrm{SL}(2, \mathbb{R})$ is the group of 2×2 real matrices with determinant 1;
- $\mathrm{SL}(2, \mathbb{R})$ acts isometrically on $(\mathbb{H}^2, g)$ by Möbius transformations:
$$\gamma.z = \frac{az + b}{cz + d}, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{R}), \quad z \in \mathbb{H}^2.$$

Note that $\mathbb{H}^2$ is the universal cover of M. For convex co-compact surfaces, all non-identity elements of Γ are *hyperbolic*, that is they satisfy $a + d > 2$. These elements have no fixed points on $\mathbb{H}^2$ and two distinct fixed points on the boundary at infinity,

$$\partial \mathbb{H}^2 = \mathbb{R} \cup \{\infty\}.$$

A basic example is the hyperbolic cylinder (1.2):

$$M = \Gamma \backslash \mathbb{H}^2, \quad \Gamma = \{\gamma^j \mid j \in \mathbb{Z}\}, \quad \gamma = \begin{pmatrix} e^{\ell/2} & 0 \\ 0 & e^{-\ell/2} \end{pmatrix}.$$

We next consider the geodesic flow

$$\varphi_t = e^{tX} : S^*M \to S^*M \tag{5.1}$$

where S^*M is the unit cotangent bundle and X is the generator of the flow, which is a vector field on S^*M. Of course, S^*M is identified canonically with the unit tangent bundle, however the cotangent bundle T^*M is the natural phase space for microlocal analysis, which will play an important role later.

Each geodesic on M lifts to a geodesic on $\mathbb{H}^2$. We recall that the geodesics on $\mathbb{H}^2$ are just half-circles with centers on $\mathbb{R}$, as well as vertical half-lines. To each $(x, \xi) \in S^*\mathbb{H}^2$ correspond the two endpoints at infinity

$$B_\pm(x, \xi) = \lim_{t \to \pm\infty} \varphi_t(x, \xi) \in \partial \mathbb{H}^2. \tag{5.2}$$

One particularly nice property of the geodesic flow on hyperbolic surfaces is its hyperbolicity. In fact, there exist two *horocyclic vector fields* $U_\pm$ on $S^*\mathbb{H}^2$ such that X, U_+, U_- are linearly independent at every point and

$$[X, U_\pm] = \pm U_\pm.$$

These vector fields are invariant under the action of the group $\mathrm{SL}(2,\mathbb{R})$ so they descend to S^*M for any hyperbolic surface M. We can define a natural metric (the Sasaki metric) on S^*M by making X, U_+, U_- into an orthonormal frame. Then we have the stable/unstable decomposition

$$T(S^*M) = E_0 \oplus E_s \oplus E_u, \quad E_0 = \mathbb{R}X, \quad E_s = \mathbb{R}U_+, \quad E_u = \mathbb{R}U_-,$$

and

$$|d\varphi_t \cdot v| = \begin{cases} |v|, & v \in E_0, \\ e^{-t}|v|, & v \in E_s, \\ e^t|v|, & v \in E_u. \end{cases}$$

We will not provide the formulas for $U_\pm$ here but will instead use a special case, from which the general formula can be recovered by isometry invariance. Specifically, consider the point $(i, i) \in S^*\mathbb{H}^2$. Then

$$e^{sU_+}(i,i) = (i+s, i).$$

We remark that the corresponding geodesics are

$$e^{tX}(e^{sU_+}(i,i)) = (e^t i + s, e^{-t} i).$$

Thus

$$de^{tX}(i,i) \cdot U_+(i,i) = (1,0),$$

and the Sasaki length of the vector $(1,0)$ tangent to $S^*\mathbb{H}^2$ at $(e^t i, e^{-t} i)$ can be computed using the isometry $(x,\xi) \mapsto (e^{-t}x, e^t\xi)$ and it is equal to e^{-t}.

We also note that flowing along X and U_+ does not change the value of B_+ (see (5.2)) and flowing along X and U_- does not change the value of B_-. In particular B_+ takes the value ∞ on the example $e^{sU_+}(i,i)$ considered above.

6. The limit set. Three-funnel surfaces.

We now give a basic example of a hyperbolic surface: a three funnel surface. A similar construction with a larger number of disks produces *Schottky surfaces*, and every convex co-compact surface can be obtained in this way – see [Bo07, §15.1].

Consider four nonoverlapping closed upper half-disks D_1, D_2, D_3, D_4 in $\mathbb{H}^2$ with centers on the real line, with centers ordered as follows: D_1, D_3, D_4, D_2. There exist hyperbolic transformations

$$\gamma_1, \gamma_3 \in \mathrm{SL}(2,\mathbb{R}),$$

such that γ_1 maps the complement of D_2 onto the interior of D_1, and γ_3 maps the complement of D_4 onto the interior of D_3. We may choose γ_1 to preserve the geodesic G_1 orthogonal to the boundaries of D_1, D_2, and similarly for γ_3 to preserve the geodesic G_3 orthogonal to the boundaries of D_3, D_4. Then γ_1, γ_3 generate a free group $\Gamma \subset \mathrm{SL}(2, \mathbb{R})$ and the quotient

$$M = \Gamma \backslash \mathbb{H}^2$$

is a convex co-compact hyperbolic surface. To see that M is a surface with three funnels, we note that a fundamental domain for Γ is given by

$$\mathcal{F} = \mathbb{H}^2 \setminus \bigcup_{j=1}^{4} D_j^\circ.$$

The geodesics G_1, G_2 give rise to two of the necks of M and the third neck is obtained from two pieces, one orthogonal to D_1, D_3 and the other to D_2, D_4.

We now use this example to define the limit set corresponding to a quotient $M = \Gamma \backslash \mathbb{H}^2$,

$$\Lambda \subset \mathbb{R} \subset \partial \mathbb{H}^2.$$

Let $\overline{\mathcal{F}}$ be the closure of the fundamental domain $\mathcal{F}$ in the compactification $\overline{\mathbb{H}^2} = \mathbb{H}^2 \sqcup \partial \mathbb{H}^2$. The images of the fundamental domain $\mathcal{F}$ under Γ give a tessellation of $\mathbb{H}^2$:

$$\mathbb{H}^2 = \bigcup_{\gamma \in \Gamma} \gamma.\mathcal{F}.$$

We define the limit set

$$\Lambda = \Lambda_\Gamma \subset \partial \mathbb{H}^2$$

as follows:

$$\partial \mathbb{H}^2 \setminus \Gamma = \bigcup_{\gamma \in \Gamma} (\gamma.\overline{\mathcal{F}}) \cap \partial \mathbb{H}^2.$$

That is, to obtain Γ we remove from the boundary at infinity the infinite ends of each tessellating fundamental domain.

For the hyperbolic cylinder, the limit set simply consists of two points (in the model used, at 0 and at ∞). For the Schottky model of a three-funnel surface, the limit set is much more complicated, having a fractal structure. Indeed, it can be written out as follows:

$$\Gamma = \bigcap_{L \geq 1} \bigcup_{\mathbf{w} \in W_L} (D_{\mathbf{w}} \cap \partial \mathbb{H}^2)$$

where:

- W_L is the set of admissible words of length L, that is sequences $\mathbf{w} = w_1 \dots w_L$ such that $w_j \in \{1, 2, 3, 4\}$ and $\{w_j, w_{j+1}\}$ cannot be equal to $\{1, 2\}$ or $\{3, 4\}$;

- we define the upper half-disk $D_{\mathbf{w}}$ for $\mathbf{w} \in W_L$ as follows:
$$D_{\mathbf{w}} = \gamma_{w_1} \dots \gamma_{w_{L-1}}(D_{w_L})$$
where $\gamma_2 := \gamma_1^{-1}$, $\gamma_4 := \gamma_3^{-1}$.

The disks $D_{\mathbf{w}}$ form a tree-like structure, in particular
$$D_{w_1 \dots w_L} \Subset D_{w_1 \dots w_{L-1}},$$
so the limit set Λ is a 'Cantor-like' set.

We define the Hausdorff dimension
$$\delta := \dim_H(\Lambda). \tag{6.1}$$
It is known that for any convex co-compact surface except the hyperbolic cylinder which has $\delta = 0$, we have
$$\delta \in (0, 1).$$
We finally explain how the limit set determines the trapping structure of the hyperbolic surface, which is crucial for the following sections. Recall the geodesic flow (5.1). Define the sets
$$\Gamma_\pm = \{(x, \xi) \in S^*M \mid \varphi_t(x, \xi) \not\to \infty \quad \text{as } t \to \mp\infty\}, \quad K := \Gamma_+ \cap \Gamma_-.$$
We call Γ_+ the outgoing tail, Γ_- the incoming tail, and K the trapped set. The lifts of $\Gamma_\pm$ by the projection map $\pi : S^*\mathbb{H}^2 \to S^*M$ can be expressed in terms of Λ as follows:
$$\pi^{-1}(\Gamma_\pm) = \{(x, \xi) \in S^*\mathbb{H}^2 \mid B_\mp(x, \xi) \in \Lambda\}. \tag{6.2}$$
Any point $(x, \xi) \in S^*\mathbb{H}^2$ is determined by $B_+(x, \xi), B_-(x, \xi)$, and one real parameter varying along the geodesic flow. It follows that the Hausdorff dimensions of $\Gamma_\pm$ and K are given by
$$\dim_H \Gamma_+ = \dim_H \Gamma_- = \delta + 2, \quad \dim_H K = 2\delta + 1.$$

7. Results: spectral gap and resonance counting

We are now ready to present the main results, discussing some ideas behind the proofs in the next section.

The first one is the spectral gap result, relying on the following theorem of Patterson and Sullivan [Bo07, §14.4]:

Theorem 3. *Let δ be defined in (6.1). Then the resolvent $R(\lambda)$ has a simple pole at $i(\delta - 1/2)$ and no other poles on $\{\operatorname{Im} \lambda \geq \delta - 1/2\}$ (unless M is the hyperbolic cylinder which has other poles on $\{\operatorname{Im} \lambda = \delta - 1/2\}$).*

For $\delta < 1/2$, this gives a spectral gap of size $1/2 - \delta$; note that for the hyperbolic cylinder the size of the gap is sharp by (3.4), since $\delta = 0$. In all other cases with $\delta \leq 1/2$ an improved essential spectral gap is available:

Theorem 4. [Na] *Assume that* $0 < \delta \leq 1/2$. *Then there exists*

$$\nu > \frac{1}{2} - \delta$$

such that M *has an* ***essential spectral gap*** *of size* ν*, in particular there are only finitely many resonances in* $\{\operatorname{Im}\lambda > -\nu\}$.

There are however many open questions left. In particular, it is not known what is the best possible size of the essential spectral gap when $\delta \leq 1/2$ and which surfaces with $\delta \geq 1/2$ have an essential spectral gap (though there are examples of surfaces with $\delta \geq 1/2$ which do have a gap, see [DyZa]).

We next discuss counting resonances. For $\nu, R > 0$ define

$$N(R,\nu) = \#\{\lambda \text{ resonance},\ \operatorname{Re}\lambda \in [R, R+1],\ \operatorname{Im}\lambda > -\nu\}.$$

The following upper bound was proved by Guillopé–Lin–Zworski:

Theorem 5. [Bo07, Corollary 15.11] *For all* $\nu > 0$ *we have* $N(R,\nu) = \mathcal{O}(R^\delta)$.

We see that again the dimension of the limit set plays a crucial role.

However, no matching lower bounds are known. Moreover, the bound $N(R,\nu)$ cannot be sharp for all ν because of the possibility of the spectral gap. A better bound for small ν has been proved in [Dy] (following previous work of Frédéric Naud):

Theorem 6. *For all* $\nu > 0$, $\varepsilon > 0$ *we have* $N(R,\nu) = \mathcal{O}(R^{m(\delta,\nu)+\varepsilon})$ *where* $m(\delta,\nu) = \min(\delta, 2\delta + 2\nu - 1)$.

Note that $m(\delta,\nu) = \delta$ for $\nu \geq (1-\delta)/2$ and $m(\delta,\nu) < 0$ for $\nu < \frac{1}{2} - \delta$, the latter statement corresponding to the spectral gap.

8. Uncertainty principle

We finally provide a (very sketchy) explanation of the results stated in the previous section, using the methods developed in [DyZa]. For that we use semiclassical quantization, which associates to a function

$$a \in C_0^\infty(T^*M)$$

an h-dependent family of operators

$$\operatorname{Op}_h(a) : L^2(M) \to L^2(M).$$

We first explain the reasoning behind essential spectral gap of size $\frac{1}{2} - \delta$ of Theorem 3. (Note that Theorem 3 actually gives a full spectral gap, which includes low frequencies, but this phenomenon is very special to constant curvature and we will not explain it here.) To show an essential spectral gap of size ν, we need to show there are only finitely resonances with $\operatorname{Im}\lambda > -\nu$. Since resonances form a discrete set and they are symmetric with respect to the map $\lambda \mapsto -\bar{\lambda}$, we need to disprove the existence of a sequence of resonances

$$\lambda_j, \quad \operatorname{Re}\lambda_j \to \infty, \quad \operatorname{Im}\lambda_j \geq -\nu.$$

We define $h_j := \lambda_j^{-1}$, and assume for simplicity that $\operatorname{Im}\lambda_j = -\nu$. Since λ_j is a resonance, there exists a *resonant state* $u_j \in C^\infty(M)$ such that

$$\Big(-\Delta_g - \frac{1}{4} - \lambda_j^2 \Big) u_j = 0, \quad u_j \quad \text{is outgoing},$$

where the outgoing condition is given in (3.5). The equation above can be rewritten in the semiclassical form

$$\Big(-h^2\Delta_g - \frac{h^2}{4} - (1 - ih\nu)^2 \Big) u = 0,$$

where we make $u_j = u(h_j)$ an h-dependent family of functions and h goes to 0 along the sequence h_j. We normalize u_j to have $\|u_j\|_{L^2(N)} = 1$ where N is a large compact subset of M.

By the elliptic estimate we know that u is $\mathcal{O}(h^\infty)$ microlocally away from the sphere bundle S^*M. Next, a combination of Egorov's Theorem with the equation for u gives us the following result (modulo lower order terms which we blatantly ignore here):

$$\|\operatorname{Op}_h(a)u\|_{L^2} = e^{\nu t}\|\operatorname{Op}_h(a \circ \varphi_t)u\|_{L^2} + \mathcal{O}(h^\infty)$$

where $\varphi_t : T^*M \to T^*M$ is the homogeneous geodesic flow.

The work of Vasy (see [DyZw, Chapters 5 and 6]) implies that the wavefront set of u lies inside the outgoing tail $\Gamma_+ \subset S^*M$:

$$\operatorname{Op}_h(a)u = \mathcal{O}(h^\infty) \quad \text{if } \operatorname{supp} a \cap \Gamma_+ = \emptyset.$$

Moreover, u must have positive mass on the trapped set:

$$\|\operatorname{Op}_h(a)u\|_{L^2} \geq C^{-1} > 0 \quad \text{if } a \neq 0 \quad \text{on } K.$$

The key idea now is to extend these statements to h-dependent symbols, by propagating for time $t = \log(1/h)$ and using the hyperbolicity of the flow. We then get the following estimates (where we use symbols which are not compactly supported but in practice there would be cutoffs to a fixed neighborhood of K)

$$u = \operatorname{Op}_h(\chi_+)u + \mathcal{O}(h^\infty),$$
$$\|\operatorname{Op}_h(\chi_-)u\|_{L^2} \geq C^{-1}e^{-\nu t} = C^{-1}h^\nu.$$

Here $\chi_\pm$ are cutoffs to h-neighborhoods of $\Gamma_\pm \cap S^*M$.

Now, to get a gap of size ν we need to reach a contradiction, and it is enough to show the following estimate, which I call the *fractal uncertainty principle*:

$$\|\operatorname{Op}_h(\chi_-)\operatorname{Op}_h(\chi_+)\|_{L^2\to L^2} \ll h^\nu.$$

Here $\mathrm{Op}_h(\chi_-)$ and $\mathrm{Op}_h(\chi_+)$ are localizations to h-neighborhoods in two incompatible directions, thus it may be impossible for a quantum state to concentrate so narrowly in both directions; this explains the name 'uncertainty principle'.

One non-rigorous way to establish the uncertainty principle (which can be converted to an easy to prove estimate) is to say that a quantum state should occupy volume at least $h^n = h^2$ in the phase space T^*M where $n = 2$ is the dimension of M. (A basic example is a Gaussian of width $h^{1/2}$, which occupies a ball of size $h^{1/2}$ in both position and frequency.) If we show that the total volume of the intersection of supports of χ_+ and χ_- is smaller than h, then we see that no quantum state can concentrate perfectly both on $\operatorname{supp}\chi$ and $\operatorname{supp}\chi_-$. However, the volume can be computed from the dimension and (6.2):

$$\mathrm{vol}(\operatorname{supp}\chi_+ \cap \operatorname{supp}\chi_-) \sim h \cdot h^{1-\delta} \cdot h^{1-\delta} = h^2 \cdot h^{1-2\delta} \tag{8.1}$$

and this gives us a gap for $\delta < 1/2$; one can also deduce that the size of the gap will be equal to $1/2-\delta$. An improved spectral gap, Theorem 4, can be explained as follows (at least when $\delta \approx 1/2$): the fractal nature of the set Λ causes extra cancellations which give a better uncertainty principle. (This is quite difficult to show – the original work of Naud relied on a rather subtle and deep set of ideas due to Dolgopyat.)

To get a bound on the number of resonances, Theorem 5, we again use a volume bound. All resonant states should be microlocalized on $\operatorname{supp}\chi_+$, whose volume is

$$\mathrm{vol}(\operatorname{supp}\chi_+) \sim h \cdot h^{1-\delta} = h^2 \cdot h^{-\delta}.$$

Since each quantum state occupies volume at least h^2, there can be at most $h^{-\delta}$ linearly independent quantum states microlocalized on $\operatorname{supp}\chi_+$, which gives the bound on the number of resonances in Theorem 5. Finally, for Theorem 6 we have to combine the two arguments above, using that resonant states have a lower bound microlocally on $\operatorname{supp}\chi_-$ together with (8.1).

The above are of course very brief explanations. The reader is referred to the introductions to [DyZa, Dy] for more detailed explanations of the method, and to the upcoming work with Long Jin for a presentation of these arguments in a much technically simpler setting of open quantum baker's maps.

References

[BWPSKZ] Sonja Barkhofen, Tobias Weich, Alexander Potzuweit, Hans-Jürgen Stöckmann, Ulrich Kuhl, and Maciej Zworski, *Experimental observation of the spectral gap in microwave n-disk systems,* Phys. Rev. Lett. **110**(2013), 164102.

[Bo07] David Borthwick, *Spectral theory of infinite-area hyperbolic surfaces,* Birkhäuser, 2007.

[Bo14] David Borthwick, *Distribution of resonances for hyperbolic surfaces,* Experimental Math. **23**(2014), 25–45.

[BoWe] David Borthwick and Tobias Weich, *Symmetry reduction of holomorphic iterated function schemes and factorization of Selberg zeta functions,* J. Spectr. Th. **6**(2016), 267–329.

[Dy] Semyon Dyatlov, *Improved fractal Weyl bounds for hyperbolic manifolds,* with an appendix with David Borthwick and Tobias Weich, preprint, arXiv:1512.00836.

[DyZa] Semyon Dyatlov and Joshua Zahl, *Spectral gaps, additive energy, and a fractal uncertainty principle,* to appear in Geom. Funct. Anal., arXiv:1504.06589.

[DyZw] Semyon Dyatlov and Maciej Zworski, *Mathematical theory of scattering resonances,* book in progress, http://math.mit.edu/~dyatlov/res/

[GHW] Colin Guillarmou, Joachim Hilgert, and Tobias Weich, *Classical and quantum resonances for hyperbolic surfaces,* preprint, arXiv:1605.08801.

[HV] Peter Hintz and András Vasy, *The global non-linear stability of the Kerr–de Sitter family of black holes,* preprint, arXiv:1606.04014.

[Na] Frédéric Naud, *Expanding maps on Cantor sets and analytic continuation of zeta functions,* Ann. de l'ENS (4) **38**(2005), 116–153.

[No] Stéphane Nonnenmacher, *Spectral problems in open quantum chaos,* Nonlinearity **24**(2011), R123.

E-mail address: dyatlov@math.mit.edu

Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139