89c

3

COURSE 5

HYPERBOLIC GEOMETRY IN TWO DIMENSIONS AND TRACE FORMULAS

YVES COLIN DE VERDIÈRE

Institut Fourier, Université de Grenoble I
BP 74
F-38402 Saint Martin d'Hères CEDEX, France

Translated by: Thierry Martin and Jonathan Robbins

M.-J. Giannoni, A. Voros and J. Zinn-Justin, eds.
Les Houches, Session LII, 1989
Chaos et Physique Quantique /Chaos and Quantum Physics

Contents

1. Introduction

Surfaces of constant negative Gaussian curvature, or hyperbolic surfaces, constitute an extremely simple model for chaos. Geodesics on these surfaces possess all of the characteristic properties of instability: exponential divergence of trajectories, ergodicity, mixing, and the Anosov property. It is tempting therefore to study this model as a prototype of chaos.

The geodesic flow on an arbitrary Riemannian manifold admits a canonical quantization, the Laplace–Beltrami operator or Riemannian Laplacian. The spectral theory of this operator is both a very deep and a very mysterious subject.

In the following presentation we discuss first a simply connected model: the Poincaré half-plane, and then its compact quotients: Riemann surfaces and polygonal hyperbolic billiards. The geodesic flows on these quotients are then studied: one can easily establish the ergodicity and the mixing property.

Finally, we shall describe the Laplacian, its spectrum, the ζ functions which are associated with it, and the determinant of the Laplacian. The trace formulas of the Selberg type are described in a formal way as a particular case of the method of images.

These lectures are not intended as a detailed presentation of this vast subject, but rather as an informal introduction to the literature in this field. The reader who wishes to pursue these matters further is directed to the references.

2. The models of hyperbolic geometry

2.1. Description of the Poincaré half-plane model

There exist many models of hyperbolic geometry in two dimensions (for this chapter, see for example [18], [4], [5]). However, one can easily show, by using geodesic polar coordinates, that any two complete, simply connected surfaces of constant Gaussian curvature (-1) are isometric. As examples let us mention: (i) the Poincaré disk, (ii) the mass hyperboloid defined by the equation $x^2 + y^2 + 1 = z^2, z > 0$, endowed with the Lorentzian metric $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 - \mathrm{d}z^2$, and finally (iii) the Poincaré half-plane, which we shall consider in what follows.

The Poincaré half-plane is the half-plane $H = \{x + \mathrm{i}y = z \mid y > 0\}$ with the Riemannian metric $\mathrm{d}s^2 = (\mathrm{d}x^2 + \mathrm{d}y^2)/y^2$. To this infinitesimal metric is associated a global distance $d(z, z')$, the shortest length of all differentiable paths from z to z' measured with respect to $\mathrm{d}s^2$.

The geodesics are either semicircles orthogonal to the x axis or vertical half lines.

It is worthwhile to make a digression to calculate this distance. To this end we study the isometries of H.

It is easy to verify that the transformations T_a and S defined by $T_a(z) = z + a$ and $S(z) = -1/z$ are isometries of H (they preserve the metric $\mathrm{d}s^2$). These transformations generate the group of linear fractional transformations $h_{a,b,c,d}$ of the form

$$h_{a,b,c,d}(z) = \frac{az + b}{cz + d} \ ,$$

where a, b, c, d are real numbers and the matrix $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ can be chosen so as to have unit determinant. This group is naturally identified with the group $\mathrm{SL}_2(\mathbb{R})/\{1, -1\}$ under the homomorphism

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \to h_{a,b,c,d}.$$

Conversely, one can show that every isometry of H which preserves orientation is of this form. Let us denote the group of these positive isometries by $\mathrm{Isom}^+(H)$. It is easy to show that it has the following transitivity property: if v_0 (resp. v_1) is a unit vector tangent to H and located at z_0 (resp. z_1), there exists a transformation $T \in \mathrm{Isom}^+(H)$ such that

$$T(z_0) = z_1, \qquad T'(v_0) = v_1,$$

(T' denoting the tangent map to T at z_0). H also admits negative isometries which reverse the orientation; these are of the form $h_{a,b,c,d} \circ \Sigma$, where Σ, the reflection about the y axis, is given by $\Sigma(z) = -\bar{z}$.

We may now calculate the distance d. From symmetry arguments it is clear that if $z = a + \mathrm{i}$ and $z' = a + \mathrm{i}t$ (where $t \in \mathbb{R}^+$), the geodesic between z and z' is just the vertical half-line $x = a$. One may conclude that

$$d(a + \mathrm{i}, a + \mathrm{i}t) = \left| \int_1^t \frac{\mathrm{d}u}{u} \right| = |\log(t)|.$$

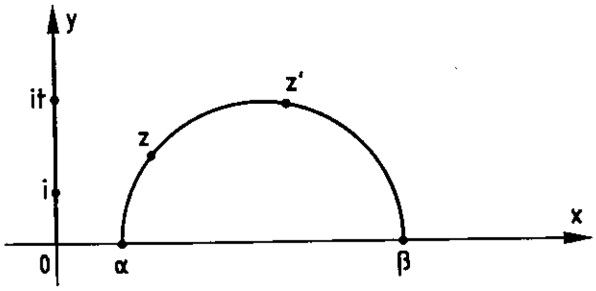

Fig. 1. The geodesics of H.

If instead z and z' do not lie on the same vertical, we construct the semicircle centered on $y = 0$ which passes through z and z'. Let $x = \alpha$ and $x = \beta$ denote its intersections with the x axis. Then the cross-ratio $[z, z', \alpha, \beta]$ is invariant under linear fractional transformations. In the case above it is equal to t. We then deduce that

$$d(z, z') = |\log[z, z', \alpha, \beta]|,$$

which after some calculation becomes

$$d(z, z') = \operatorname{arccosh}\left(1 + \frac{1}{2}\frac{|z - z'|^2}{yy'}\right).$$

2.2. *Classification of the positive isometries*

A positive isometry $\gamma \in \mathrm{Isom}^+(H)$ is said to be *elliptic* if it has a single fixed point in H. In this case it describes a rotation about this fixed point. This case corresponds to the condition $|a + d| < 2$.

γ is said to be *parabolic* if it has a single fixed point at infinity. It is then conjugate in $\mathrm{Isom}^+(H)$ to the translation $z \to z + 1$.

γ is said to be *hyperbolic* if it has two fixed points at infinity. In this case the geodesic which joins the fixed points is invariant under γ. This geodesic is called the *axis* of γ; γ generates translations along its axis. This corresponds to the case $|a + d| > 2$. Such a transformation is conjugate to a unique dilation with center 0 and dilation factor $k^2 > 1$, where $|a + d| = k + 1/k$.

2.3. *The Gauss–Bonnet formula*

If T is a geodesic triangle in H, one can measure its angles with respect to the metric; these are the same as what one obtains from the Euclidean metric $dx^2 + dy^2$. One finds that

$$\mathrm{Area}(T) = \int\int_T \frac{dx\, dy}{y^2} = \pi - (\alpha + \beta + \gamma),$$

where α, β and γ are the angles of T.

Conversely, if $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ are such that $\alpha + \beta + \gamma < \pi$, there exists a triangle T, unique up to isometry, with angles α, β and γ ([4]).

3. Discrete groups

For this chapter, see [4], [11].

Let Γ be a subgroup of $\mathrm{Isom}(H)$. We shall study the quotient space $\Gamma\backslash H$, first as a topological space and then as a Riemannian manifold (which may have singular points) along with its geodesics. Finally we shall consider its associated quantized analog, the Laplacian on the space of Γ-periodic functions.

In order for the quotient space to have a reasonable topology (to be Hausdorff, for example) the action of the group Γ must be *discontinuous*. By this is meant the following: for any $z_0 \in H$ and for any bounded subset K of H, the number of elements γ of Γ such that $\gamma(z_0)$ belongs to K must be finite. In this case Γ is said to be a *discrete subgroup*.

An important notion is that of the *fundamental domain* D for the action of Γ. $D \subset H$ is said to be a fundamental domain if,

(i) $\cup_{\gamma \in \Gamma}(\gamma(D)) = H$

(ii) $\forall \gamma \neq \mathrm{Id}$, $\gamma(D) \cap D \subset \partial D$.

(Here ∂D denotes the boundary of D.)

Take a point $z_0 \in H$ which is not a fixed point of any element of Γ. The Dirichlet fundamental domain D_{z_0} is defined as follows,

$$D_{z_0} = \{z \in H \mid \forall \gamma \neq \mathrm{Id},\ d(z, z_0) \leqslant d(\gamma(z), z_0)\}.$$

It is a polygonal domain with geodesic edges.

The topological quotient space $\Gamma\backslash H$ may be viewed as a fundamental domain D with edges identified under the discrete group operations.

Let us give a few examples.

3.1. The triangle groups $\Gamma_{p,q,r}$

Let $p, q, r \leqslant \infty$ be three integers such that

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1.$$

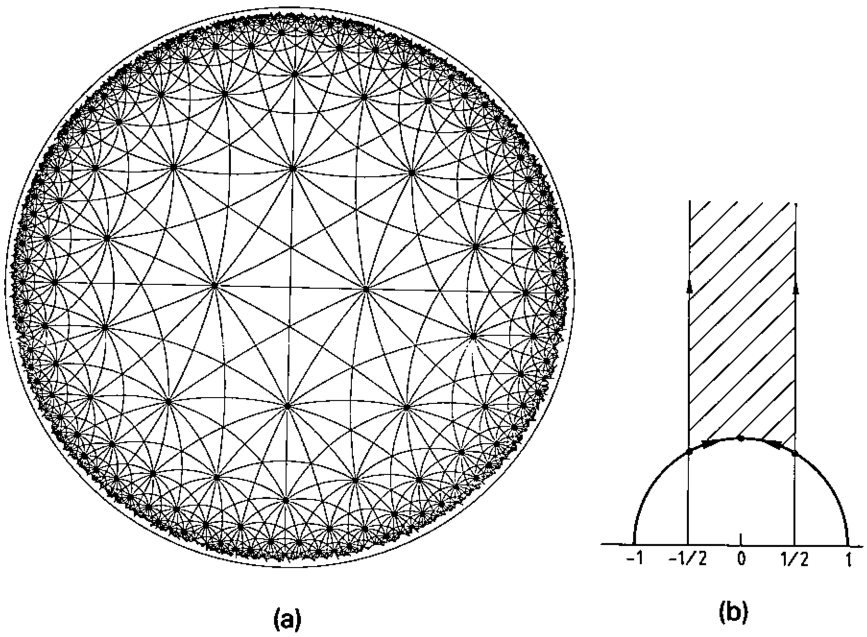

Fig. 2. (a) The tiling $T_{2,3,7}$ (reproduced from [9] with kind permission of the author). (b) A fundamental domain of $SL_2(\mathbb{Z})$.

There exists a unique triangle T whose angles are given by $\pi/p, \pi/q, \pi/r$. Let P, Q, R denote the reflections about the three sides of the triangle. They generate a discrete group $\Gamma_{p,q,r}$. It is easy to see that T is a fundamental domain. In fact, it is the Dirichlet domain D_{z_0} of every point z_0 inside the triangle.

The tilings of the hyperbolic plane by triangles have been used for example in the drawings of Escher.

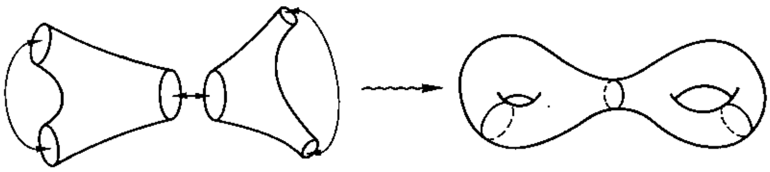

Fig. 3. Sewing together pairs of pants to construct a surface of genus two.

3.2. The arithmetic groups

The prototype of an arithmetic group is $\mathrm{SL}_2(\mathbb{Z})$, the group of 2×2 unimodular matrices with integer elements. $\mathrm{SL}_2(\mathbb{Z})$ is actually the intersection of $\mathrm{Isom}^+(H)$ with $\Gamma_{2,3,\infty}$.

Generalizations of this example are obtained either by taking the matrix elements from a ring of algebraic integers, or by taking a subgroup of finite index (congruence groups).

3.3. The surface groups

Every compact surface of genus $\geqslant 2$ admits one (and in fact many) metric(s) of constant negative curvature. This is a consequence of the Riemann conformal mapping theorem. In fact, there are as many of these metrics as there are conformal structures; if g is an arbitrary Riemannian metric on a compact surface, there exists a metric with constant curvature conformal to g.

One way of visualizing these surfaces is to patch them together from elementary pieces with geodesic boundaries, namely the hyperbolic pants. The hyperbolic pants are constructed by sewing together two identical hyperbolic hexagons having right angles at all vertices; they are topologically equivalent to a sphere with three points removed.

The universal covering space of such a surface X is a complete, simply connected two-dimensional Riemannian manifold of constant negative curvature -1. It may therefore be identified with H. The covering group, i.e., the fundamental group of the surface X, is a discrete group Γ. One may therefore identify X and $\Gamma\backslash H$ as Riemannian manifolds.

It is possible to deform, in a continuous way, the manner in which the fundamental group Γ is embedded in G. These deformations do not change the topology of the surface X. One obtains thereby a family of surfaces with the same curvature and the same topology. The deformations are described by a surface of dimension $6[\mathrm{genus}(X)-1]$, called the Teichmüller space.

4. Classical chaos

For this chapter, see for example [12], [14], [5].

Let us consider the geodesic flow on $\Gamma\backslash H$. We first describe the phase space, i.e., the tangent bundle of $\Gamma\backslash H$, and then the Hamiltonian dynamics associated with the geodesic flow.

Let us denote by UH the set of all unit tangent vectors to H. A point (z,p) of UH is a couple made of a point z in H and a unit vector p (its length being measured by the hyperbolic metric !). The Liouville measure induces a natural measure on UH which we shall call also Liouville's measure. The geodesic flow g_t is the smooth dynamical system on UH defined by $g_t(z,p) = (z', p')$, where z' is the point of parameter t on the geodesic line with origin z and initial direction p. p' is defined as the tangent vector to the same geodesic line at z'. This dynamical system is of course the restriction to UH of the Hamiltonian system whose Hamiltonian is the Riemannian metric

$$E(z,p) = \frac{1}{2}\frac{(p_x^2 + p_y^2)}{y^2} .$$

Now, if Γ is a discrete subgroup, we can define the unit tangent bundle of $\Gamma\backslash H$ as

$$U(\Gamma\backslash H) = \Gamma\backslash UH ,$$

where the action of Γ on UH is the natural one.

4.1. Examples

One gets a better understanding of these constructions by considering the following examples.

For the triangle groups, the geodesics, which are the projections of geodesics in G to the fundamental domain, are the trajectories of a billiard ball which follows a geodesic inside T and is reflected at the boundary of T.

For the surface groups one obtains simply the geodesics on the compact surface from the preceding construction.

This construction of geodesics allows one to treat several problems, for example the determination of the periodic geodesics: if γ is an hyperbolic

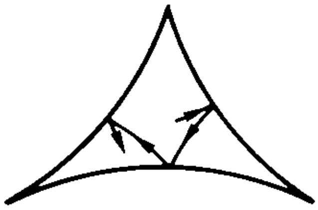

Fig. 4. Geodesics on the Hyperbolic Triangle.

isometry of Γ, the projection of the axis of γ onto H is a periodic geodesic of the quotient. The converse is true. Moreover the axes of γ and γ' project onto the same periodic geodesic if and only if γ and γ' are conjugate in Γ.

We give now a sketch of Hopf's proof of the ergodicity of the geodesic flow when $\Gamma\backslash H$, and consequently $\Gamma\backslash UH$, have finite volume.

A function f defined on the unit tangent bundle of $\Gamma\backslash H$ may also be considered as a Γ-periodic function on UH. Let us consider the time average of such a function along a geodesic. According to Birkhoff's theorem ([14]), the forward and backward time averages are equal. More precisely we have, for almost all z, the following property,

$$\lim_{T\to+\infty}\frac{1}{T}\int_0^T f(g_t(z))\,\mathrm{d}t=\lim_{T\to-\infty}\frac{1}{T}\int_0^T f(g_t(z))\,\mathrm{d}t=\bar{f}(z),$$

where $\bar{f}$ is invariant under the geodesic flow g_t.

One may characterize geodesics by the points at infinity, ϕ_- and ϕ_+, to which they asymptote in forward and backward time respectively. It follows that $\bar{f}$, the time average of f along geodesics, is a function of ϕ_- and ϕ_+. However, the forward time average is determined by ϕ_+, and the backward time average by ϕ_-. According to Birkhoff's theorem these are the same, and are both equal to the full time average. Thus $\bar{f}$ may be regarded as a function of either ϕ_- or ϕ_+. But ϕ_- and ϕ_+ may be varied independently (given two points at infinity there exists a geodesic between them.) It follows that $\bar{f}$ is constant, at least almost everywhere (there are some technical points which we have not addressed.) Since the space UH has finite volume, this constancy of $\bar{f}$ implies the ergodicity of the geodesic flow (that is, time averages equal space averages).

5. Quantization

In what follows we shall restrict ourselves to the case of a compact quotient, i.e., a compact fundamental domain.

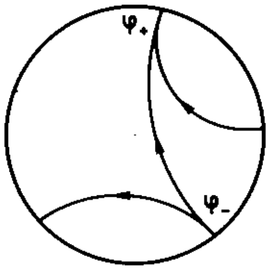

Fig. 5. Asymptotic geodesics.

For this chapter, see for example [3], [2], [18].

The formal Laplacian: Let us consider the differential operator on H given by

$$\Delta f = -y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) f.$$

It is formally symmetric relative to the Riemannian volume $\mu = \mathrm{d}x\,\mathrm{d}y/y^2$:

$$\forall f, h \in C_0^\infty(H), \quad \int_H f \Delta h \,\mathrm{d}\mu = \int_H \Delta f h \,\mathrm{d}\mu\,.$$

In addition, its principal symbol $y^2(p_x^2 + p_y^2) = 2E$ is the classical Hamiltonian associated with the metric g (which may be regarded as the classical Lagrangian) under the Legendre transform. If we succeed in obtaining self-adjoint realizations of Δ, these will constitute the natural quantizations of the geodesic flow.

Remark: The Laplacian associated to a Riemannian metric is related to the Schrödinger operator (i.e., the quantized Hamiltonian) for a free frictionless particle of mass m whose motion is constrained to the manifold in question. Classically this particle moves along the geodesics of the manifold. To make this connection more precise, the Schrödinger equation is given by

$$\frac{\hbar}{\mathrm{i}} \frac{\partial u}{\partial t} = \frac{\hbar^2}{2m} \Delta u.$$

The time-independent equation is as follows,

$$\frac{\hbar^2}{2m} \Delta u = \mathcal{E} u.$$

Consequently, if u is an eigenfunction of Δ with eigenvalue λ, there corresponds to it a stationary solution of the Schrödinger equation with energy $\mathcal{E} = (\hbar^2/2m)\lambda$. For fixed $\mathcal{E}$ the semiclassical limit $\hbar \to 0$ corresponds to the limit $\lambda \to \infty$. This limit will be the main theme of what follows.

5.1. The free case

The operator Δ is essentially self-adjoint on the Hilbert space $L^2(H,\mu)$: it has a unique self-adjoint extension, denoted by Δ_0, with domain

$$\mathcal{D} = \{f \in L^2(H,\mu) \mid \Delta f \in L^2(H,\mu)\},$$

(Δf is meant in the distributional sense.)

It is not difficult to see that the spectrum of Δ_0 is the half-line $[\frac{1}{4}, +\infty)$. In fact this spectrum is absolutely continuous and of infinite multiplicity. This may be seen by decomposing functions in $L^2(H,\mu)$ in Fourier series with respect to polar coordinates about a point in H; it is then clear that Δ_0 is given by a sum of operators P_n, $n \in \mathbb{Z}$, each of which is unitarily equivalent to a Schrödinger operator on the half-line $[0,+\infty)$ of the form $-\mathrm{d}^2/\mathrm{d}x^2 + v_n(x)$, where the v_n are functions strictly greater than $\frac{1}{4}$ which approach the value $\frac{1}{4}$ in the limit $x \to +\infty$.

5.2. The quotient case

In the case of a discrete group the Hilbert space is the space H_Γ of Γ-periodic functions on H, square integrable over one (and therefore every) fundamental domain of Γ. The norm is given by

$$\|f\|^2 = \int_{D_\Gamma} |f|^2 \,\mathrm{d}\mu.$$

The natural domain which renders the Laplacian self-adjoint is the space of functions

$$\mathcal{D}_\Gamma = \{f \in \mathcal{H}_\Gamma \mid \Delta f \in \mathcal{H}_\Gamma\}.$$

We denote the Laplacian by Δ_Γ.

One may interpret Δ_Γ in different ways in the examples.

5.2.1. The groups $\Gamma_{p,q,r}$

In this case the operator Δ_Γ is the Laplacian in the triangle T with Neumann boundary conditions: these conditions are clearly imposed by requiring invariance under the symmetries P, Q, R with respect to the sides of the triangle.

5.2.2. The surface groups
In this case one can identify the Γ-periodic functions with functions on the surface. There are no boundary conditions because the surface has no boundary!

5.2.3. The spectrum
The compactness of the fundamental domain implies that Δ_Γ has a compact resolvent ([2]) and therefore that its spectrum is discrete. The spectrum consists of a series of eigenvalues of finite multiplicity; the only accumulation point is at $+\infty$. By convention we denote the eigenvalues by

$$\lambda_1 = 0 < \lambda_2 \leqslant \lambda_3 \leqslant \cdots \leqslant \lambda_n \leqslant \cdots.$$

where each eigenvalue is repeated according to its multiplicity. The eigenvalue 0 has multiplicity one and corresponds to the space of constant functions.

The eigenfunctions ϕ_n form an orthonormal basis in the space $\mathcal{H}_\Gamma$.

5.2.4. The Weyl formula
It is well known in quantum mechanics that, in the semiclassical limit, each state occupies a phase space volume on the order of $(2\pi\hbar)^d$, where d is the dimension of configuration space. One has therefore that

$$\#\{\frac{\hbar^2}{2}\lambda_n \leqslant 1\} \sim \frac{1}{(2\pi\hbar)^2}\mathrm{vol}\{E(z,p) \leqslant 1\},$$

as $\hbar$ goes to 0.

In the two-dimensional case this becomes

$$\#\{n \mid \lambda_n \leqslant \lambda\} \sim \frac{1}{4\pi}\mathrm{Area}(D_\Gamma)\lambda,$$

as $\lambda \to \infty$.

This formula, which is due to Weyl, admits many rigorous proofs. We shall see one of these later on in our discussion of the trace formulas.

6. The trace formula formalism

6.1. The heat equation

See [3] for this section.

Let us consider the heat equation on a two-dimensional compact Riemannian manifold X, assumed for simplicity to be without boundary. We may consider as well the Γ-periodic solutions of the heat equation. (In the case where Γ is a surface group the second case is a particular instance of the first.)

More precisely, if $u(x,t)$ denotes the temperature at time t and position x and if u_0 denotes the initial temperature distribution, one has the system,

$$u(x,0) = u_0, \qquad \frac{\partial u}{\partial t} + \Delta u = 0. \tag{$\star$}$$

(Henceforth, x and y denote each a point of X and cease to serve as coordinates.)

If $u_0 \in L^2(X,\mu)$, the system $(\star)$ has a unique solution given in functional form by

$$u(x,t) = (\mathrm{e}^{-t\Delta}\, u_0)(x).$$

Equivalently,

$$u(x,t) = \sum_{n=1}^{\infty} a_n\, \mathrm{e}^{-t\lambda_n}\, \varphi_n(x), \quad a_n = \int_X u_0(x)\varphi_n(x)\, \mathrm{d}\mu.$$

One may also express this result in the form

$$u(x,t) = \int_X e(t,x,y)u_0(y)\, \mathrm{d}\mu(y),$$

where

$$e(t,x,y) = \sum_{n=1}^{\infty} \mathrm{e}^{-t\lambda_n}\, \varphi_n(x)\varphi_n(y).$$

One says that e is the *heat equation kernel* (i.e., the integral kernel of an integral operator). It is also clear, at least formally, that

$$\mathrm{Tr}(\mathrm{e}^{-t\Delta}) = \sum_{n=1}^{\infty} \mathrm{e}^{-t\lambda_n} = \int_X e(t,x,x)\, \mathrm{d}\mu(x).$$

In this equation the trace of the heat operator is expressed as both a sum over eigenvalues and an integral over the diagonal elements of the kernel.

The gimmick behind the trace formulas is a direct calculation of the kernel, either exact or asymptotic.

6.2. Short time asymptotics of the heat equation

The behaviour of the heat kernel for small t is based on the fact that the flow of heat becomes localized as $t \to 0^+$. More precisely, if g_1 and g_2 are two Riemannian metrics which coincide near the point x_0, the difference between their respective kernels e_1 and e_2 as $t \to 0^+$ is bounded by

$$|e_1(t, x_0, x) - e_2(t, x_0, x)| = O(\mathrm{e}^{-c/t}).$$

From this result one concludes that for short times $e(t, x, x)$ admits an asymptotic expansion in powers of t which depends only on local information. One has in fact the expansion

$$e(t, x, x) \sim (4\pi t)^{-1}(1 + a_1(x)t + a_2(x)t^2 + \ldots),$$

where $a_1(x) = \frac{1}{6}K(x)$, $a_2(x) = \frac{1}{15}K(x)^2$ and $K(x)$ is the Gaussian curvature of the surface X.

Integrating over the manifold X one obtains the Minakshisundaram–Pleijel (M–P) expansion,

$$\sum \mathrm{e}^{-t\lambda_n} \sim (4\pi t)^{-1}(\mathrm{Area}(X) + a_1 t + a_2 t^2 + \ldots),$$

where $a_i = \int_X a_i(x)\,\mathrm{d}\mu(x)$. According to the Gauss–Bonnet theorem, in two dimensions, a_1 is a topological invariant: $a_1 = (\pi/3)\chi(X)$, where $\chi(X) = 2(1 - \mathrm{genus}(X))$ is the Euler characteristic of X.

From this asymptotic series follows the Weyl formula. One might have hoped for more, for example an asymptotic series in inverse powers of n (as $n \to \infty$) for the eigenvalues, of the form $\lambda_n \sim (4\pi/\mathrm{Area}(X))n + c_0 + c_1/n + \ldots$. In fact the distribution of eigenvalues is too irregular to permit such an expansion. However it is possible to obtain an asymptotic expansion for a regularized density of eigenvalues. It is given by

$$\sum \rho(\mu - \sqrt{\lambda_n}) \sim \frac{\mathrm{Area}(X)}{2\pi}\mu + d_0 + d_1/\mu + \ldots.$$

Here ρ is a function whose Fourier transform vanishes outside the interval $(-L, +L)$, where L is the length of the shortest periodic geodesic in X (cf. end of section 6.5).

The preceding considerations allow us to study the ζ functions associated with the spectrum of the Laplacian. The series

$$\zeta_g(s) = \sum_{n=2}^{\infty} \frac{1}{\lambda_n^s},$$

is convergent for $\mathrm{Re}(s) > 1$, and one can show that it admits a meromorphic extension to the complex plane $\mathbb{C}$, based on the formula

$$\zeta_g(s) = \frac{1}{\Gamma(s)} \int_0^{\infty} \left(\sum_{n=2}^{\infty} \mathrm{e}^{-t\lambda_n} \right) t^{s-1}\,\mathrm{d}t,$$

and on the M–P expansion.

In particular this allows us to define the *determinant* of the Laplacian (via ζ function regularization) as follows, $\mathrm{Det}(\Delta) = \exp(-\zeta_g'(0))$.

6.3. The Poisson formula in one dimension

Let us begin with a particularly simple case: the Euclidean case in one dimension, i.e., the real line with Riemannian metric $\mathrm{d}x^2$. The Laplacian is given by $-\mathrm{d}^2/\mathrm{d}x^2$ and the heat kernel is,

$$e_\infty(t,x,y) = (4\pi t)^{-1/2}\,\mathrm{e}^{-(x-y)^2/4t}.$$

Let Γ be the group of translations by integer multiples of T. The Γ-periodic eigenfunctions are the exponentials $e_n(x) = \exp(2\pi i n x/T)$ with eigenvalues $4\pi^2 n^2/T^2$. According to the convention introduced above,

$$\lambda_1 = 0 < \lambda_2 = \lambda_3 = \frac{4\pi^2}{T^2} < \lambda_3 = \lambda_4 = \frac{16\pi^2}{T^2} < \dots .$$

If now $u_0 : \mathbb{R} \to \mathbb{R}$ has period T it is clear that

$$u(t,x) = \int_{-\infty}^{+\infty} e_\infty(t,x,y) u_0(y)\,\mathrm{d}y,$$

has period T as well. One may rewrite this formula as an integral over a fundamental domain of Γ, for example the interval $[0,T]$,

$$u(t,x) = \int_0^T \left(\sum_{n\in\mathbb{Z}} e_\infty(t,x,y+nT) \right) u_0(y)\,\mathrm{d}y.$$

This leads to the doubly periodic kernel

$$e_T(t,x,y) = \sum_{n\in\mathbb{Z}} e_\infty(t,x,y+nT)$$

which is just the heat kernel on a circle of circumference T. Applying equation $(\star)$ we obtain

$$\sum_{n\in\mathbb{Z}} \mathrm{e}^{-4\pi^2 n^2 t/T^2} = \frac{T}{(4\pi t)^{1/2}} \sum_{n\in\mathbb{Z}} \mathrm{e}^{-n^2T^2/4t}\,.$$

This formula is a special case of the Poisson summation formula. We shall describe extensions of this formula, first in the hyperbolic case and then in the general case. Before doing so let us show that the Poisson summation formula may be interpreted as a regularized eigenvalue density.

For technical reasons one must consider operators of degree one, such as the square root of the Laplacian, $A = -\mathrm{i}\,\mathrm{d}/\,\mathrm{d}x$. Its spectrum is given by $\{\mu_n = 2\pi n/T \mid n \in \mathbb{Z}\}$. If the function ρ is of Schwartz class (i.e., a function rapidly decreasing together with all of its derivatives) and if $\hat\rho(u) = 1/2\pi \int \mathrm{e}^{-\mathrm{i}tu}\,\rho(t)\,\mathrm{d}t$ is its Fourier transform, the Poisson summation formula may be written as,

$$\sum_{n\in\mathbb{Z}} \rho(\mu-\mu_n) = T \sum_{n\in\mathbb{Z}} \hat\rho(nT)\,\mathrm{e}^{\mathrm{i}nT\mu}\,.$$

One sees that oscillation frequencies of the regularized density are the lengths of the periodic geodesics in $\mathbb{R}/T\mathbb{Z}$ and that if Support$(\hat\rho) \subset (-T,T)$, then the Weyl formula becomes exact.

6.4. The Selberg trace formula

For this section, see [5], [18].

The case of hyperbolic geometry may be treated in a similar way and leads to the celebrated Selberg trace formula. In what follows we outline the steps of the calculation.

(i) One finds an exact expression for the heat kernel in the Poincaré half-plane, and more generally for all operators of the form $f(\Delta_0)$.

To this end one uses the homogeneity of H: if two pairs of points (z_1, z_2), (z_1', z_2') are equidistant, i.e., if $d(z_1,z_2) = d(z_1',z_2')$, then there exists an isometry γ in H such that $\gamma(z_i) = z_i'$. From this we determine a priori that the kernels are of the form $k(z,z') = \tilde f(d(z,z'))$. One has thus to calculate $f \to \tilde{f}$ and its inverse. This is accomplished by applying f to the eigenfunctions of Δ. If $\Delta\varphi = \lambda\varphi$ then

$$\int_H \tilde{f}(d(z,z'))\varphi(z')\,\mathrm{d}\mu = f(\lambda)\varphi(z).$$

(ii) One then evaluates the kernel $f(\Delta_\Gamma)$ by averaging $(k_\Gamma(z,z') = \sum_{\gamma\in\Gamma} k(z,\gamma z'))$, just as in the Euclidean case.

(iii) The trace calculation requires a more subtle argument: one computes

$$\mathrm{Tr}\ f(\Delta_\Gamma) = \sum_{\gamma\in\Gamma} \int_{\mathcal{D}_\Gamma} k(z,\gamma z)\,\mathrm{d}\mu,$$

by partitioning the group elements γ into conjugacy classes. One says that γ' is conjugate to γ ($\gamma' \in [\gamma]$) if there exists a $\gamma_1 \in \Gamma$ such that $\gamma' = \gamma_1^{-1}\gamma\gamma_1$. The same γ' can be obtained from γ_2 if and only if $\gamma_2\gamma_1^{-1}$ commutes with γ. Let us denote by $C(\gamma)$ the subgroup of those elements of Γ which commute with γ, and let $c(\gamma) = \Gamma/C(\gamma)$. The conjugates of γ are in one-to-one correspondence with elements of $c(\gamma)$, so that one obtains,

$$\begin{aligned}\sum_{\gamma'\in[\gamma]} \int_{\mathcal{D}_\Gamma} k(z,\gamma' z)\,\mathrm{d}\mu &= \sum_{\gamma_1\in c(\gamma)} \int_{\mathcal{D}_\Gamma} k(z,\gamma_1^{-1}\gamma\gamma_1 z)\,\mathrm{d}\mu \\ &= \sum_{\gamma_1\in c(\gamma)} \int_{\gamma_1\mathcal{D}_\Gamma} k(z,\gamma z')\,\mathrm{d}\mu.\end{aligned}$$

Noting that $\cup_{\gamma_1\in c(\gamma)}\gamma_1\mathcal{D}_\Gamma = \mathcal{D}_{C(\gamma)}$ one obtains the formula

$$\mathrm{Tr} f(\Delta_\Gamma) = \sum_{[\gamma]} \int_{\mathcal{D}_{C(\gamma)}} k(z,\gamma z)\,\mathrm{d}\mu.$$

This integral is easy to compute, since $C(\gamma)$ has a very simple structure. If γ is hyperbolic we let

$$N_\gamma = \sup\{k \mid \exists\gamma_0 \in \Gamma,\ \gamma = \gamma_0^k\},$$

It follows that $C(\gamma) = \{\gamma_0^n \mid n \in \mathbb{Z}\}$ is a cyclic subgroup. By conjugating γ by an element of $\mathrm{Isom}^+(H)$ one obtains the form

$$\gamma_0(z) = N(\gamma_0)z.$$

The domain $\mathcal{D}_{C(\gamma)}$ is the semiannulus $\{1 \leqslant |z| \leqslant N(\gamma_0)\}$. Letting

$$\lambda_n = \tfrac{1}{4} + \mu_n^2$$

(the μ_n are the eigenvalues of $A = \sqrt{\Delta - \frac{1}{4}}$), one obtains the Selberg trace formula for suitably regular even functions ρ,

$$\frac{1}{2}\sum_{n\in\mathbb{N}}(\rho(\mu-\mu_n)+\rho(\mu+\mu_n)) = \frac{\mathrm{Area}(\mathcal{D}_\Gamma)}{4\pi}\int_{-\infty}^{+\infty} t\tanh(\pi t)\,\rho(\mu+t)\,\mathrm{d}t$$
$$+\sum_{[\gamma]\neq \mathrm{Id}} \frac{\log N(\gamma_0)}{N(\gamma)^{1/2}-N(\gamma)^{-1/2}}\hat{\rho}(\log N(\gamma))\cos(\mu\log N(\gamma)).$$

Here, if $\gamma \in \mathrm{PSL}_2(\mathbb{R})$, $N(\gamma) = |\mathrm{Tr}(\gamma)| = |a+d|$.

6.4.1. Interpretation

To each $\gamma \neq \mathrm{Id}$ is associated a periodic geodesic of $\Gamma\backslash H$, the projection of the axis of γ onto $\Gamma\backslash H$. Its length is $\log(N(\gamma))$; γ and γ' yield the same periodic geodesic if and only if they are conjugate in Γ.

Therefore the Selberg trace formula may be interpreted as a sum over periodic geodesics in $\Gamma\backslash H$. The regularized density of the eigenvalues of the square root of the Laplacian has oscillations whose frequencies are the lengths of the periodic geodesics.

6.5. The general case: the regularized eigenvalue density

In 1971–1972, Balian and Bloch ([1]) carried out calculations similar to the ones described above for the case of Euclidean space in two and three dimensions. Following this I presented in my thesis general formulas for an arbitrary compact Riemannian manifold. These calculations were quickly extended by Chazarain, then by Duistermaat and Guillemin who used the method of Fourier integral operators developed by Hörmander.

Here we shall limit ourselves to the simplest case of these formulas. We refer the reader to [6], [7] for a more detailed discussion.

Let X be a compact surface without boundary with an arbitrary Riemannian metric such that the periodic geodesics are all isolated and therefore nondegenerate (in the sense that the linearized Poincaré map has no unit eigenvalue).

One denotes by γ a periodic geodesic and by N_γ its iteration number: γ is the N_γth iterate of a simple periodic orbit whose length is T_γ/N_γ. Let

ρ be of a function of Schwartz class whose Fourier transform moreover has compact support. Let $\mu_n = \sqrt{\lambda_n}$. Then one has the formula

$$\sum_{n=1}^{\infty} \rho(\mu - \mu_n) = \text{Area}(X)\hat{\rho}(0)\mu + d_0$$
$$+ \sum_{\gamma} \frac{T_\gamma}{N_\gamma} \frac{\varepsilon_\gamma}{|\det(1 - P_\gamma)|^{1/2}} \hat{\rho}(T_\gamma)\, e^{i\mu T_\gamma} + O\left(\frac{1}{\mu}\right).$$

Here, ϵ_γ is equal to $+1, -1, +i, -i$ according to the parity modulo 4 of the Morse index of γ.

6.6. Relation to the Feynman path integral

One can give a simple formal justification of the asymptotic expansions introduced above based on the Feynman path integral.

One can write the kernel of the Schrödinger equation $-i\hbar\, du/dt = \hbar^2 \Delta u/2$ as a functional integral over the space of paths; if $\Omega_{x,y}$ denotes the set of C^1 piecewise-continuous paths between x and y, and if $E(\gamma) = \int_0^1 |\dot{\gamma}(s)|^2\, ds$ for such a path, one may write,

$$e(t, x, y) = \int_{\Omega_{x,y}} \exp\left(\frac{i}{\hbar t} E(\gamma)\right) d\mu_{x,y,t}.$$

On taking the trace of this expression and defining Ω to be the set of closed paths, we obtain

$$\sum_{n=1}^{\infty} \exp\left(i \frac{\hbar t}{2} \lambda_n\right) = \int_{\Omega} \exp\left(\frac{i}{\hbar t} E(\gamma)\right) d\mu_t(\gamma).$$

When $\hbar \to 0$ this integral can be computed using the method of stationary phase, which calls to one's attention the critical points of the function E defined on Ω; these are precisely the periodic geodesics. Unfortunately it is difficult to carry through this heuristic approach, since one does not know how to construct Feynman path integrals rigorously. It is useful, however; for example it allows us to anticipate the appearance of the Morse indices.

7. Asymptotic behaviour of eigenmodes

The main question to be discussed in this last chapter is the following: if the geodesic flow is chaotic, can the eigenfunctions nevertheless be concentrated on a *strict subset* of the manifold ?

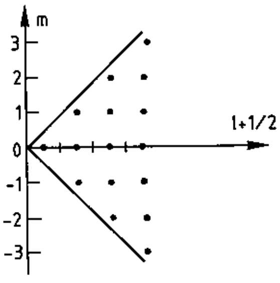

Fig. 6. Joint spectrum of A and L_z.

Before giving a precise statement, let us consider some examples, preceded by a:

Definition: *a subsequence* (λ_{k_i}) $(k_1 < k_2 < \ldots)$ *of the sequence* (λ_k) *is said to be of density* d $(0 \leqslant d \leqslant 1)$ *if*

$$\lim_{\lambda \to \infty} (\#\{i \mid \lambda_{k_i} \leqslant \lambda\})/(\#\{k \mid \lambda_k \leqslant \lambda\}) = d\,.$$

7.1. Example: the 2-dimensional sphere with the usual metric

The spectral decomposition of the Laplacian is given by the (classic) theory of spherical harmonics. The eigenvalues are $l(l+1)$, $l = 0, 1, 2, \ldots$ and the corresponding eigenspaces have dimension $(2l+1)$, with the standard basis Y_l^m $(-l \leqslant m \leqslant l)$. The Y_l^m are joint eigenfunctions of Δ and of the angular momentum $L_z = (1/\mathrm{i})(\partial/\partial\varphi)$ (infinitesimal rotation around the z-axis). This joint spectrum can be pictured more clearly if you take $A = \sqrt{(\Delta + \frac{1}{4})}$; the eigenvalues of A are $l + \frac{1}{2}$, $l = 0, 1, 2, \ldots$.

Now the subsequence of eigenfunctions $Y_l^l = c_l(x + \mathrm{i}y)^l$ is concentrated strongly *near the equator.* Of course, it correspond to a subsequence of density 0 in the whole spectrum (cf. definition above).

But, if for some c $(0 < c < 1)$, we take the subset of Y_l^m with $m \geqslant cl$, this subsequence is of density $(1-c)/2 > 0$, and is concentrated on a neighborhood of the equator given by $\{\sin\theta \geqslant c\}$.

7.2. Example: the torus $\mathbb{R}^2/\Gamma$

Here the eigenfunctions are $\varphi_\delta = \exp(2\pi\mathrm{i}\,\langle\delta|z\rangle)$, where δ lies in the dual lattice Γ^* of Γ. They are uniformly distributed in X: $\|\varphi_\delta\|^2 \equiv 1$; but in *momentum space* they are strongly localized: their Fourier transforms are Dirac delta functions of p.

7.3. Shnirelman's conjecture

In 1974, Shnirelman ([13]) gave a not very clear statement concerning the asymptotic behaviour of eigenmodes when the geodesic flow is ergodic.

His statement has been put in a more standard mathematical form by S. Zelditch (1984: [19]) in the case of compact hyperbolic surfaces, using a pseudo-differential calculus well adapted to this geometry. Shortly after, I was able to generalize his method a little bit to obtain the general result ([8]).

The main point in my opinion is the following one: there exists some *quantization procedure*, i.e., a map $f \to \mathrm{Op}(f)$ from functions $f(x,p)$ on phase space to (pseudo-differential) operators on the base manifold, such that if $f \geqslant 0$, then $\mathrm{Op}(f)$ is also a positive operator, i.e.,

$$\forall \varphi \in \mathrm{L}^2(X), \quad \langle \mathrm{Op}(f)\varphi | \varphi \rangle \geqslant 0 .$$

A *symbol of order 0* is a smooth function $f : TX \setminus \{0\} \to \mathbb{R}$ which is homogeneous of degree 0 in the p variable,

$$\forall \lambda > 0, \quad f(x, \lambda p) = f(x,p) .$$

Such a function is well determined by its restriction to the set of unit tangent vectors $UX \subset TX$. (This normalization corresponds exactly to the restriction to the energy shell in the study of classical ergodicity for Hamiltonian systems.)

In this way, we obtain for all $\varphi \in \mathrm{C}^\infty(X, \mathbb{R})$ a linear functional,

$$\mu_\varphi : \quad f \to \langle \mathrm{Op}(f)\varphi | \varphi \rangle .$$

But, since $f \geqslant 0$ implies $\mu_\varphi(f) \geqslant 0$, it follows from a theorem of L. Schwartz that μ_φ is a (Radon) probability measure on UX.

Let now φ_k be an eigenfunction orthonormal basis of the Laplacian, and let $\mu_k = \mu_{\varphi_k}$. (These μ_k depend on the particular quantization used, but the limiting results proved for $k \to \infty$ will not depend on it, since any two quantizations $\mathrm{Op}^1(f)$ and $\mathrm{Op}^2(f)$ differ by some operator of negative order).

It is a consequence of *Egorov's theorem* that every weak limit μ_∞ of a subsequence (μ_{k_i}) of (μ_k) is a phase space *probability distribution* which is *invariant* by the geodesic flow.

There are many such invariant measures: the simplest one is the uniform Liouville measure, but every periodic geodesic also carries its own invariant measure, the averaging $(1/T) \oint \mathrm{d}t$.

The final mathematical statement is,
Theorem (Zelditch, Colin de Verdière): *if the geodesic flow is ergodic, then there exists an eigenvalue subsequence* (λ_{k_i}) *of density 1 such that the corresponding measures* (μ_{k_i}) *converge weakly to the Liouville measure.*
Corollary: *for any smooth domain* $D \subset X$,

$$\lim_{i\to\infty} \int_D |\varphi_{k_i}|^2 = \frac{\text{area}(D)}{\text{area}(X)}.$$

A generalization of this theorem to the semi-classical limit of the Schrödinger operator has been obtained by Helffer, Martinez and Robert ([10]).

Now our results leave open at least two important questions (cf. Heller's lectures in this session):

(i) How fast is the convergence ? Numerically, it seems to be very slow!

(ii) Does there exist some subsequence $\mu_{n_i} \to \mu_\infty \neq$ Liouville measure ? For example, eigenmodes concentrating on some unstable closed geodesic.

References

[1] R. Balian and C. Bloch, Distributions of eigenfrequencies for the wave equation in a finite domain III, Ann. Phys. (NY) 69 (1972), 76–160.

[2] P. Bérard, Spectral Geometry : Direct and Inverse Problems, Lecture Notes in Mathematics 1207 (Springer, 1986) (with a bibliography by P. Bérard and M. Berger).

[3] M. Berger, P. Gauduchon and E. Mazet, Le Spectre d'une Variété Riemannienne, Lecture Notes in Mathematics 194 (Springer, 1971).

[4] A. Beardon, The Geometry of Discrete Groups (Springer, 1983).

[5] N.L. Balazs and A. Voros, Chaos on the Pseudosphere, Phys. Rep. 143 (1986) 109–240.

[6] J. Chazarain, Spectre des opérateurs elliptiques et flots hamiltoniens, Séminaire Bourbaki 27e année, 1974/75, no. 460, Lecture Notes in Mathematics 514 (Springer, 1976) pp111–123.

[7] Y. Colin de Verdière, Spectre du Laplacien (RCP25-Strasbourg, 1974).

[8] Y. Colin de Verdière, Ergodicité et Fonctions Propres du Laplacien, Comm. Math. Phys. 102 (1985), 497–502.

[9] D. Epstein, Computers, groups and hyperbolic geometry, Astérisque 163–164 (1988) 9–29.

[10] B. Helffer, A. Martinez and D. Robert, Ergodicité et Limite Semi-classique, Comm. Math. Phys. 109 (1987) 313–326.

[11] W. Magnus, Noneuclidean Tesselations and Their Groups (Academic Press, 1974).

[12] V. Poenaru, An informal crash course on ergodic theory, in: Ill-Condensed Matter, eds., R. Balian et al., Les Houches, XXXI, 1978 (North-Holland, 1979) pp563–607.

[13] A.I. Shnirelman, Ergodic properties of eigenfunctions, Usp. Mat. Nauk 29/6 (1974) 181–182 (in Russian).

[14] Y.G. Sinai, Introduction to Ergodic Theory (Princeton, 1976).

[15] C. Series, On coding geodesics with continuous fractions, Enseign. Math. 29 (1980) 67–76.

[16] C. Series, Non-Euclidean geometry, continued fractions, and ergodic theory, Math. Intelligencer 4/1 (1982) 24–31.

[17] C. Series, The geometry of Markoff numbers, Math. Intelligencer, 7/3 (1985) 20–29.

[18] A. Terras, Harmonic Analysis on Symmetric Spaces and Applications (Springer, 1985).

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