An introduction to quantum equidistribution

S. De Bièvre
Université des Sciences et Technologies de Lille
UFR de Mathématiques et Laboratoire Painlevé
59655 Villeneuve d'Ascq
email: Stephan.De-Bievre@math.univ-lille1.fr

June 20, 2006

Abstract

These notes contain crash courses on classical and quantum mechanics and on semi-classical analysis as well as a short introduction to one issue in quantum chaos: the semi-classical eigenfunction behaviour for quantum systems having an ergodic classical limit. The emphasis is on explaining the conceptual and structural similarities between the ways in which this question arises in the study of arithmetic surfaces and ergodic toral automorphisms. The text is aimed at an audience of graduate students and post-docs in number theory.

KEYWORDS Quantum chaos, quantum maps

1 Introduction

These notes are loosely based on four lectures I gave during the first week of the Nato Summer School on Equidistribution in Number Theory, held at the Université de Montréal in july 2005. My task was to provide the necessary mathematics and physics background for the students to understand in which sense the topics of the second week lectures by Z. Rudnick [10] on "The arithmetic theory of quantum maps" as well as those by A. Venkatesh [11] on the spectral analysis of the Laplace-Beltrami operator and by E. Lindenstrauss [7] on quantum unique ergodicity for arithmetic surfaces can be seen as examples of a more general set of problems, referred to as "quantum chaos." In other words, I had to explain that both deal with the links between a spectral problem (the quantum side) and a Hamiltonian dynamical system (the classical side) naturally related to each other through an appropriate asymptotic analysis. For that purpose, I provided a crash course in classical mechanics and one in quantum mechanics, then gave a short introduction to semi-classical analysis, to end with an introduction to quantum maps and a proof in that context of the main equidistribution theorem in the field of quantum chaos, namely the Schnirelman theorem for quantized ergodic toral automorphisms.

Since an extensive introduction to these topics, at the beginning graduate level, can already be found in [4], I will only briefly recall the material developed there. I will concentrate instead on developing some illustrative material (concerning symmetries in particular) that I did not have the time to develop in the actual lectures and that help to bring out the link between the two subjects alluded to above. Omitted proofs can be obtained either by a combination of matrix analysis, multivariable calculus and a little imagination, or are to be found in [4] (or both). A recent update on what is known on equidistribution for quantum map eigenstates is available in [5] and in the contribution of Z. Rudnick in this volume [10]. Similarly, for the asymptotic behaviour of the eigenfunctions of the Laplace-Beltrami operator, the interested reader can turn to [12] and to the contribution of E. Lindenstrauss [7] in this volume.

2 A crash course in classical mechanics

2.1 Newtonian mechanics

According to Newton's second law, that you may remember from high school, "mass times acceleration equals force." In other words:

$$m\ddot{q}(t) = F(q(t)),\ q(0) = q,\ \dot{q}(0) = v. \tag{1}$$

Here, the *force* $F : \mathbb{R}^n \to \mathbb{R}^n$ is given, as well as the *mass* m and the *initial data* $q, v \in \mathbb{R}^n$. The unknown in this equation is the motion of the system, namely the curve $t \in \mathbb{R} \mapsto q(t) \in \mathbb{R}^n$. In short, classical mechanics is about solving coupled non-linear second order ordinary differential equations. In most cases of interest, they are of a special type: the force is often *conservative*, meaning that $F(q) = -\nabla V(q)$ for a function $V : \mathbb{R}^n \to \mathbb{R}$, called *the potential.* The use of the term "conservative" is justified by the following simple result:

Proposition 2.1 Energy conservation *Let*

$$E : (q, v) \in \mathbb{R}^n \times \mathbb{R}^n \to \frac{1}{2}mv^2 + V(q) \in \mathbb{R}.$$

Let $t \in \mathbb{R} \mapsto q(t) \in \mathbb{R}^n$ be a solution to (1), then, for all $t \in \mathbb{R}$,

$$E(q(t), \dot{q}(t)) = E(q(0), \dot{q}(0)).$$

The function E is called the *energy* of the system (it is the sum of the *kinetic* and the *potential energy*) and the proposition states that the energy does not vary in time for a solution of (1).

It is good to keep a few examples in mind. The first one is of great historic importance and continues to attract considerable attention: it is the Kepler problem. Here $d = 3$, $V(q) = -\frac{GmM}{\|q\|}$, where M is the mass of the sun, and m the one of the earth, and G is the gravitational constant. Solving (1) explicitly can be done (while not trivial, it is standard ...) and leads to elliptic, parabolic or hyperbolic trajectories, depending on whether the energy is strictly negative, identically zero, or strictly positive. This is a special case of a *central potential*: $V(q) = W(\| q \|)$. A second class of examples is provided by *harmonic systems*, where $V(q) = \frac{1}{2}mq^T\Omega^2 q$, with Ω^2 a positive definite n by n matrix. Now Newton's equation reads $\ddot{q} = -\Omega^2 q$. It is linear, and hence this time it *is* trivial to solve immediately:

$$q(t) = \cos \Omega t \; q + \frac{\sin \Omega t}{\Omega} \; v.$$

In general, it is of course impossible to obtain explicit solutions, and one is interested in characterizing the behaviour of the solutions, and in particular in their asymptotic properties at large times t. This will obviously depend on the type of potential one considers. For example, if $V(q) \to +\infty$ when $|q| \to +\infty$, the motion is bounded, meaning that

$$\sup_{t\in\mathbb{R}} |q(t)| \leq C < +\infty.$$

This is an easy application of energy conservation: indeed, for all t

$$V(q(t)) \leq \frac{1}{2m}\dot{q}(t)^2 + V(q(t)) = E(q(0), \dot{q}(0)). \tag{2}$$

Now, since V tends to infinity with $\| q \|$, this clearly implies (2). Such potentials are said to be *confining*.

2.2 Hamiltonian mechanics and beyond

There exists an important reformulation of Newton's mechanics, referred to as Hamiltonian mechanics. Introduce the *Hamiltonian*

$$H : x = (q, p) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto \frac{p^2}{2m} + V(q) \in \mathbb{R}, \tag{3}$$

and observe that Newton's equation (1) is equivalent to the first order system of differential equations called Hamilton's equations

$$\dot{q}(t) = \frac{p(t)}{m} = \frac{\partial H}{\partial p}(x(t)), \quad \dot{p} = -\nabla V(q(t)) = -\frac{\partial H}{\partial q}(x(t)), \tag{4}$$

with initial conditions $x(0) = (q, mv)$. The variable p is referred to as the *momentum* in the physics literature and the space of positions and momenta is called *phase space*. Note that the Hamiltonian is nothing but the energy expressed in terms of the position and the momentum, rather than the position and the velocity. One defines the corresponding flow $\Phi_t^H : \mathbb{R}^{2n} \mapsto \mathbb{R}^{2n}$ by $\Phi_t^H(x) = (q(t), p(t))$, where $q(0) = q, p(0) = p$. An obvious question that comes to mind here is: "What is the big deal?" After all, it is hard to imagine that this way of rewriting Newton's equation will shed any light on how to actually solve it. There exist at least three answers to this question. The first one is:

"It's pretty! Look!" Let me compute the time rate of change of an arbitrary smooth function $f : \mathbb{R}^{2n} \to \mathbb{R}$ (or to $\mathbb{C}$) along a solution curve. This yields

$$\begin{aligned}\frac{d}{dt}f(q(t),p(t)) &= \partial_q f(x(t))\dot{q}(t) + \partial_p f(x(t))\dot{p}(t)\\ &= \partial_q f(x(t))\partial_p H(x(t)) - \partial_p f(x(t))\partial_q H(x(t))\\ &=: \{f, H\}(x(t)),\end{aligned}$$

where I introduced the *Poisson bracket*

$$\{\cdot,\cdot\} : (f,g) \in C^\infty(\mathbb{R}^{2n}) \times C^\infty(\mathbb{R}^{2n}) \mapsto \{f,g\} \in C^\infty(\mathbb{R}^{2n}),$$

with

$$\{f,g\}(x) = \partial_q f(x)\partial_p g(x) - \partial_p f(x)\partial_q g(x).$$

The reason I claim this is pretty is the following: thanks to (5)-(6) below, $C^\infty(\mathbb{R}^{2n})$ now has the structure of a Lie-algebra:

$$\{f,g\} = -\{g,f\} \quad \text{(Anti-symmetry)} \tag{5}$$

$$\{\{f,g\},h\} + \{\{g,h\},f\} + \{\{h,f\},g\} = 0 \quad \text{(Jacobi identity)} \tag{6}$$

$$\{f,gh\} = \{f,g\}h + g\{f,h\} \quad \text{(Derivation)} \tag{7}$$

If nothing else, this is certainly intriguing, and, if you are a trained mathematician of any kind, you are likely to find this pretty. But if you are nevertheless somewhat practically minded, you will be happy to know that this rewriting is also useful. To give at least one indication why (there are many others), let's have a look at constants of the motion. A *constant of the motion* is a function $f \in C^\infty(\mathbb{R}^{2n})$ which is constant along the solutions of (4), meaning that $f \circ \Phi_t^H = f$ for all $t \in \mathbb{R}$. Clearly f is a constant of the motion iff $\{f, H\} = 0$. Consequently H itself is always a constant of the motion, as we already saw. Constants of the motion are a Good Thing: the more, the merrier! Indeed, if $f_1, f_2, \dots f_k$ are constants of the motion, and $c \in \mathbb{R}^k$, then the flow Φ_t^H leaves their common level surface

$$\Sigma(c) = \{x \in \mathbb{R}^{2n} | f_j(x) = c_j, j = 1 \dots k\}$$

invariant. So, if they are functionally independent (meaning the Jacobian matrix $\partial_i f_j$ has rank k), the flow takes place on a $(2n-k)$-dimensional surface in $\mathbb{R}^{2n}$. This constitutes an a priori simplification of the dynamical problem, which can in this situation be thought of as a system of first order equations in $2n-k$ variables, rather than in the original $2n$ variables. Now observe that it follows immediately from the Jacobi identity that, if f and g are constants of the motion, then so is $\{f,g\}$. So, the constants of the motion make up a Lie-subalgebra of $C_0^\infty(\mathbb{R}^{2n})$ and the Poisson bracket can even be thought of as a machine for producing constants of the motion. As an example, let's look at central potentials: $V(q) = W(\| q \|)$, $q \in \mathbb{R}^3$. Then the three components $\ell_i(x) = (q \wedge p)_i$, $i = 1 \dots 3$ of the *angular momentum* vector are constants of the motion as is readily checked through a direct computation. Note furthermore that they satisfy $\{\ell_1, \ell_2\} = \ell_3$, plus cyclic permutation. This means that the three components of the angular momentum vector form a representation of the Lie algebra so(3) of the rotation group SO(3). This is directly related to the fact that the Hamiltonian itself is invariant under the rotation group because the potential is central, as I will further discuss below.

The third answer to the question above is that the Hamiltonian formulation of classical mechanics has a number of very nice generalizations in various directions. First, *any* function $f \in C^\infty(\mathbb{R}^{2n})$ (not just the Hamiltonian) generates a flow $\Phi_t^f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ as follows: $\Phi_t^f(x) = x(t)$ where $t \in \mathbb{R} \mapsto x(t) \in \mathbb{R}^{2n}$ solves

$$\dot{q}(t) = \frac{\partial f}{\partial p}(x(t)), \quad \dot{p} = -\frac{\partial f}{\partial q}(x(t)), \quad x(0) = x = (q, p). \tag{8}$$

I will refer to the function f as the *generator* of the flow Φ_t^f. It is clear that $\Phi_{t+t'}^f = \Phi_t^f \circ \Phi_{t'}^f$ so that the Φ_t^f define an $\mathbb{R}$-action on $\mathbb{R}^{2n}$ (meaning a group homomorphism from the additive group of reals to the diffeomorphisms of $\mathbb{R}^{2n}$), or a *Hamiltonian dynamical system.* Note that, for any $g \in C^\infty(\mathbb{R}^{2n})$, one has

$$\frac{\mathrm{d}}{\mathrm{d}t} g \circ \Phi_t^f(x) = \{g, f\}(\Phi_t^f(x)). \tag{9}$$

For further reference, let me also mention that the maps Φ_t^f are *symplectic*:

Definition 2.2 *A diffeomorphism Φ of $\mathbb{R}^{2n}$ is symplectic if, for all $f, g \in C^\infty(\mathbb{R}^{2n})$,*

$$\{f \circ \Phi, g \circ \Phi\} = \{f, g\} \circ \Phi.$$

The group of symplectic diffeomorphisms of $\mathbb{R}^{2n}$ is denoted by $\mathit{Diff}_{\mathrm{sympl}}(\mathbb{R}^{2n})$.

For example, on $\mathbb{R}^6$, let $f(x) = q_1 p_2 - q_2 p_1 = \ell_3(x)$. Then

$$\dot{q}_1(t) = -q_2(t),\ \dot{q}_2(t) = q_1(t), \dot{p}_1(t) = -p_2(t),\ \dot{p}_2(t) = p_1(t),\ \dot{q}_3 = 0 = \dot{p}_3.$$

Integrating this yields $\Phi_t^f(x) = (R_t q, R_t p)$, where $R_t = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}$
In other words, the third component of angular momentum generates rotations about the third axis (and analogously for the two other components). Using $g = H$ and $f = \ell_i$ in (9), it is now clear why the rotational invariance of H in the case the potential is central implies that each ℓ_i is a conserved quantity.

This is a general phenomenon of which we will see another example when discussing the the Poincaré half plane below. A *symmetry* of a Hamiltonian H is a symplectic diffeomorphism Φ so that $H \circ \Phi = H$. If H admits a one-parameter group of symmetries of the type Φ_t^f, for some f, then, as a result of (9), f is a constant of the motion for H. In other words, if a Hamiltonian has many symmetries, it also admits many constants of the motion.

As another very simple example of a Hamiltonian flow, that will be useful in what follows, consider on $\mathbb{R}^2$ the function $f(q,p) = q^2/2$. Then $\dot q = 0, \dot p = -q$, so $q(t) = q, p(t) = p - tq$ and

$$\Phi_t^{q^2/2}(q,p) = (q, p - tq) = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix}.$$

More generally, any function f that is a homogeneous quadratic polynomial in the variables q_i, p_i yields a linear flow.

A further and more sweeping generalization that finds its origin in Hamiltonian mechanics is symplectic geometry. Roughly, a symplectic manifold is a manifold N so that the vector space $C^\infty(N)$ is equipped with a composition law $\{\cdot,\cdot\}$, called a Poisson bracket, satisfying (5)-(7), as well as a non-degeneracy condition: $\{f,g\} = 0, \forall g \in C^\infty(N)$ implies that f is a constant. It is a non-trivial fact that, locally, there always exist coordinates (called Darboux coordinates) on N so that the Poisson bracket takes on the form it has on $\mathbb{R}^{2n}$ (implying that symplectic manifolds are always even dimensional). A simple example is the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$. Write $x = (q,p) \in [0,1[^2$ and define the Poisson bracket as on $\mathbb{R}^2$. For example,

$$\{\cos q \sin p, \sin q \cos p\} = \sin^2 q \sin^2 p - \cos^2 q \cos^2 p.$$

For $n = 1$, the group $\mathrm{SL}(2,\mathbb{Z})$ (two by two matrices with determinant 1 and integer entries) acts on $\mathbb{T}^2$ by symplectic automorphisms. These maps can have rich behaviour, despite their apparent simplicity as I will discuss in more detail below. In fact, we will be picking a fixed $A \in \mathrm{SL}(2,\mathbb{Z})$ and iterate it, to obtain a $\mathbb{Z}$-action on $\mathbb{T}^2$ by symplectic transformations, or a *discrete Hamiltonian dynamical system.* This observation is the starting point for the interest in quantum maps, that I will introduce in Section 5 and that are also the subject of the contribution of Z. Rudnick in this volume [10].

A second class of important examples, in particular in connection with the lectures of E. Lindenstrauss and A. Venkatesh is provided by the *cotangent bundles* of arbitrary manifolds. Those carry a natural symplectic structure. Moreover, the geodesic flow on a Riemannian manifold can be viewed as a Hamiltonian flow on the cotangent bundle of this manifold. Rather than developing the general theory needed to understand the words appearing in the preceding sentences, an endeavour for which I have neither the space nor the inclination, I will work out the special case of the Poincaré half plane below, which is the one of relevance in the present volume.

For extensive introductions to classical mechanics, including Hamiltonian dynamics and symplectic geometry, I refer to [1] [2].

2.3 Classical mechanics on the Poincaré half plane

The Poincaré half plane is a Riemannian manifold, which is a manifold M such that at each $q \in M$, the tangent space at q is equipped with a Euclidean structure $g(q)$. The function $q \to g(q)$ is called the Riemannian structure of M.

This allows one to define notions such as the length of a curve, and of a geodesic, which is a path of shortest distance between two points of the manifold. But there is no need to know Riemannian geometry to understand the basic facts about the Poincaré half plane. A few elementary and intuitive ideas from the theory of two-dimensional surfaces in $\mathbb{R}^3$ suffice largely. In fact, the Poincaré half plane can (locally) be identified with a surface of revolution, as I will explain below.

First, to understand how geodesic motion shows up in mechanics problems, let's have a look at a particle of mass m constrained to move on a sphere of radius a. One can think of the particle as being attached with a rigid rod of length a to a fixed point, taken to be the origin O. I will assume no other forces act on the particle than the pull of the rod, which is there to keep it from flying off the sphere and which acts radially. I will in particular ignore the gravitational pull on the particle, which is a reasonable thing to do when the particle moves fast (or if the rod's other end is attached to a spaceship drifting through intergalactic space). The total force acting on the particle being radial, Newton's second law implies angular momentum $\ell = mq \wedge v$ (Here the $\wedge$ stands for the vector product, not the GCD...) is conserved, as is readily seen, and consequently, the particle motion takes place in a plane through the origin: indeed, $q(t)$ is now for all times t perpendicular to the fixed vector ℓ. But the intersection of a plane through O with a sphere is, by definition, a great circle. Since moreover the force has no tangential components along the sphere, the particle's speed is constant. The particle therefore moves with constant angular speed along a great circle. Now, what is special about great circles? Well, they are precisely the geodesics or paths of shortest distance on the sphere: given two points P and B on a sphere, the path of shortest distance from P to B is the segment of the great circle obtained by intersecting the plane containing OB and OC with the sphere. This is why a flight from Paris (latitude 48.50N, longitude 2.20E) to Beijing (latitude 39.55N, longitude 116.20E) takes you over Novosibirsk (55.04N, 82.54E).

This situation generalizes to arbitrary surfaces in $\mathbb{R}^3$: if a particle is constrained to move on a surface, and the only forces that act on it are perpendicular to the surface, then one can show that Newton's second law implies it moves with constant speed along a geodesic of the surface (This takes some work to show). An interesting special case is the one of a surface of revolution $r = f(x_3)$, where (r, θ) are the polar coordinates in the x_1x_2-plane and $f :]\alpha, +\infty[\subset \mathbb{R}^+ \to \mathbb{R}^+$ is a nice smooth function ($0 \leq \alpha \leq 1$). Using $(\theta, x_3) \in [0, 2\pi[\times]\alpha, +\infty[$ as coordinates on the surface, and designating by v_θ, v_3 the corresponding components of a general tangent vector (check your multivariable calculus), the kinetic energy of a particle can now be written

$$E(x_3, \theta, v_3, v_\theta) = \frac{1}{2}m\left(f(x_3)^2 v_\theta^2 + (1 + f'(x_3)^2)v_3^2\right).$$

Let me introduce the new variable

$$s(x_3) = \int_1^{x_3} \sqrt{1 + f'(\zeta)^2}\mathrm{d}\zeta$$

which is the distance along the surface from the point $(x_3 = 1, \theta)$ to the point (x_3, θ) provided $x_3 \geq 1$ and minus that distance otherwise. It is convenient to use s as a coordinate instead of x_3. Note that s is a strictly growing function of x_3. In terms of the new coordinates (s, θ), the energy becomes

$$E(\theta, s, v_\theta, v_s) = \frac{1}{2} m \left(\tilde{g}(s)^2 v_\theta^2 + v_s^2 \right),$$

where I introduced the function $\tilde{g}$ through the relation $\tilde{g}(s(x_3)) = f(x_3)$.

In general, given a smooth surface in $\mathbb{R}^3$, one can always introduce local coordinates $q = (q_1, q_2)$ on it that run through an open set of $\mathbb{R}^2$. Correspondingly, any tangent vector v to the surface at the point q has two components $v_1, v_2 \in \mathbb{R}$. In terms of these components, the Euclidean inner product between two tangent vectors v, w at q can be expressed in the form

$$v^T g(q) w$$

where $g(q)$ is a positive definite symmetric matrix and the T stands for transpose. For the surfaces of revolution above, with $q_1 = \theta, q_2 = s$, this matrix is

$$g(q) = \begin{pmatrix} \tilde{g}(q_2) & 0 \\ 0 & 1 \end{pmatrix}.$$

With this notation, the energy function can be rewritten $E(q, v) = \frac{1}{2} m v^T g(q) v$. A special case of particular relevance for us is the situation where $\tilde{g}(s) = \mathrm{e}^{-s}$. The corresponding $f(x_3)$ is easily seen to be defined only for $x_3 \geq 1$, corresponding to $s \geq 0$ (So $\alpha \geq 1$). Now, changing coordinates one last time to $x = \theta, y = \mathrm{e}^s$, one has, with $z = (x, y), v = (v_x, v_y)$

$$E(z, v) = \frac{1}{2} m \left(\frac{v_x^2 + v_y^2}{y^2}\right) = \frac{1}{2} m v^T g(z) v, \qquad g(z) = \begin{pmatrix} y^{-2} & 0 \\ 0 & y^{-2} \end{pmatrix}.$$

Now there is nothing to prevent me from extending the matrix valued function g and hence E to a function on $\mathbb{H} \times \mathbb{R}^2$, where $\mathbb{H} = \mathbb{R} \times \mathbb{R}^+$, equipped with the Riemannian structure $g(z)$ above is now, by definition, the Poincaré half plane. The surface of revolution with $f(x_3) = \mathrm{e}^{-s(x_3)}, x_3 \geq 1$ is obtained by quotienting the region $y \geq 1$ of the half plane by the action of $n \in \mathbb{Z}$ given by $(x, y) \to (x + 2\pi n, y)$.

Armed with these preliminaries, let me now show that the geodesic flow on $\mathbb{H}$, described in A. Venkatesh's lectures, can be viewed as a Hamiltonian flow. For that purpose, consider the following Hamiltonian function H on phase space $\mathbb{H} \times \mathbb{R}^2$:

$$H : (z, p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto \frac{1}{2} y^2 (p_x^2 + p_y^2) = \frac{1}{2} p^T g(z)^{-1} p. \tag{10}$$

The corresponding Hamilton equations of motion read

$$\dot{x} = y^2 p_x, \quad \dot{y} = y^2 p_y, \quad \dot{p}_x = 0, \quad \dot{p}_y = -y(p_x^2 + p_y^2).$$

For further reference, remark that, taking second derivatives leads to

$$\ddot{x} = 2\dot{y}y^{-1}\dot{x} = \dot{y}y^{-1}\dot{x} + \dot{x}y^{-1}\dot{y}, \quad \ddot{y} = -y^{-1}\dot{x}^2 + \dot{y}y^{-1}\dot{y}. \tag{11}$$

In view of the first two Hamilton equations above, the Hamiltonian is again nothing but the energy expressed in terms of the momentum p rather than the velocity (I put $m = 1$): $p = g(z)\dot{z}$. Remark also that the relation between the momentum p and the velocity is now position dependent, unlike what happened in (4). I claim that, given any solution $(z(t), p(t))$ of these equations, the curve $t \to z(t)$ is a geodesic on the Poincaré half plane. Since the latter are (Euclidean) half circles centered on the x axis, proving this is equivalent to showing there exists $c \in \mathbb{R}$ and $a > 0$ so that, for all t, $(x(t) - c)^2 + y(t)^2 = a^2$ or, equivalently, that $(x(t) - c)\dot{x}(t) + y(t)\dot{y}(t) = 0$. Solving for c, and expressing the result in terms of $p_x(t), p_y(t)$ yields:

$$c = x(t) + y(t)\frac{p_y(t)}{p_x(t)}.$$

So the curve $t \mapsto z(t)$ is a geodesic if and only if the function

$$C : (z, p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto x + y\frac{p_y}{p_x} \in \mathbb{R}$$

is a constant of the motion. But this is clearly the case since, as is readily checked, $\{C, H\} = 0$, so that the result follows.

Note that we found two constants of the motion, C and p_x. Since their Poisson bracket $\{C, p_x\} = 1$ is a constant, we don't find a third functionally independent constant of the motion this way. On the other hand, since we know for example from A. Venkatesh's lecture that the Poincaré half plane admits the three dimensional $\mathrm{PSL}(2, \mathbb{R})$ as group of isometries, and since isometries map geodesics into geodesics, we strongly suspect that the Hamiltonian ought to admit three one-parameter groups of symmetries and therefore have three functionally independent constants of the motion. To complete the Hamiltonian description of the geodesic flow, this is what I will now prove.

I first need some preliminaries. Any diffeomorphism φ of $\mathbb{H}$ maps a curve $\gamma : I \subset \mathbb{R} \to \mathbb{H}$ to a curve $\varphi \circ \gamma$ and hence the tangent vector $\dot{\gamma}(t)$ at $\gamma(t)$ to the tangent vector $D\varphi(\gamma(t)) \cdot \dot{\gamma}(t)$. Consequently, it induces a map

$$\varphi_* : (z, v) \in \mathbb{H} \times \mathbb{R}^2 \mapsto (\varphi(z), D\varphi(z) \cdot v) \in \mathbb{H} \times \mathbb{R}^2.$$

The diffeomorphism φ is said to be an *isometry* of the Poincaré half plane if it preserves angles between vectors and lengths of vectors, meaning that the map $D\varphi(z)$, which maps the tangent space at z to the one at $\varphi(z)$, maps the Euclidean structure $g(z)$ at z to $g(\varphi(z))$ at $\varphi(z)$:

$$D\varphi(z)^T \ g(\varphi(z)) \ D\varphi(z) = g(z). \tag{12}$$

Consequently, if φ is an isometry, then $E \circ \varphi_* = E$. Now, since Hamilton's equations identify a tangent vector v with a momentum vector p via $p = g(z)v$,

φ also induces a map on phase space given by

$$\varphi^* : (z,p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto (\varphi(z), g(\varphi(z))D\varphi(z)g(z)^{-1}p) \in \mathbb{H} \times \mathbb{R}^2. \qquad (13)$$

If φ is an isometry, it follows from (12) that

$$\varphi^* : (z,p) \in \mathbb{H} \times \mathbb{R}^2 \mapsto (\varphi(z), (D\varphi(z))^{-T}\, p) \in \mathbb{H} \times \mathbb{R}^2. \qquad (14)$$

In that case clearly $H \circ \varphi^* = H$, proving what I promised: an isometry induces a symmetry for H. Also, as you can prove through a direct computation, φ^* is symplectic. Let me now show that to every one-parameter group of isometries φ_t corresponds a constant of the motion. Explicitly, if $X(z) := \frac{\mathrm{d}\varphi_t}{\mathrm{d}t}_{|t=0}(z)$, then $\varphi_t^* = \Phi_t^f$ with $f(z,p) := p^T X(z)$ so that f is a constant of the motion since $H \circ \Phi_t^f = H$. To see this, simply compute, using (14) and multivariable calculus

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t^*(z,p)_{|t=0} &= (X(z), \frac{\mathrm{d}}{\mathrm{d}t}\left(D\varphi_{-t}(\varphi_t(z))\right)^T p)_{|t=0} \\
&= (X(z), -\partial_z X(z)^T p) \\
&= (\partial_p f(z,p), -\partial_z f(z,p)).
\end{aligned}$$

We are now ready to apply this to the three one-parameter groups that generate $\mathrm{PSL}(2,\mathbb{R})$:

$$g_1(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \mathrm{e}^{t\xi_1}, g_2(t) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} = \mathrm{e}^{t\xi_2}, g_3(t) = \begin{pmatrix} \mathrm{e}^t & 0 \\ 0 & \mathrm{e}^{-t} \end{pmatrix} = \mathrm{e}^{t\xi_3}$$

with

$$\xi_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \xi_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \xi_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

so that

$$[\xi_1, \xi_2] = \xi_3, \qquad [\xi_2, \xi_3] = 2\xi_2, \qquad [\xi_3, \xi_1] = 2\xi_1. \qquad (15)$$

Now, any $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ acts on $z = x + iy$ via $\varphi_g z = \frac{az+b}{cz+d}$. Defining $X_i(z) = \frac{\mathrm{d}}{\mathrm{d}t}\varphi_{it|t=0}(z)$, it is then easily checked that these vector fields are

$$X_1(z) = (1,0), \qquad X_2 = (-(x^2 - y^2), -2xy), \qquad X_3(z) = (2x, 2y).$$

Consequently, by what precedes, $\varphi_{it}^* = \Phi_t^{f_i}$, where

$$f_1(z,p) = p_x, \quad f_2(z,p) = -(x^2 - y^2)p_x - 2xyp_y, \quad f_3(z,p) = 2xp_x + 2yp_y \quad (16)$$

are now three functionally independent constants of the motion for H, a fact that can also be checked by a simple direct computation revealing that $\{f_i, H\} = 0$. Note that $C = \frac{f_3}{2f_1}$. Further simple computations reveal that

$$\{f_1, f_2\} = f_3, \qquad \{f_2, f_3\} = 2f_2, \qquad \{f_3, f_1\} = 2f_1.$$

so that the linear vector space spanned by the f_i is a Lie-subalgebra of $C^\infty(\mathbb{H}\times\mathbb{R}^2)$ which is isomorphic to $\mathrm{sl}(2,\mathbb{R})$, the Lie-algebra of $\mathrm{PSL}(2,\mathbb{R})$, as is clear from (15).

None of what precedes is an accident, of course, and if you suspect there must be some general theory underlying all this, you are quite right. Very briefly, the general setting is the following. Let $(N,\{\cdot,\cdot\})$ be a symplectic manifold and G a group. A symplectic action of G on N is a group homomorphism $\phi : g\in G \mapsto \phi_g \in \mathrm{Diff}_{\mathrm{sympl}}(N)$ (*i.e.* $\phi_{g_1}\circ\phi_{g_2} = \phi_{g_1 g_2}$). Such actions are particularly interesting when they provide symmetries of a given Hamiltonian dynamical system Φ_t^H: $H\circ\phi_g = H$, for all $g\in G$. An example is given by the action of $G=\mathrm{PSL}(2,\mathbb{R})$ on $N=\mathbb{H}\times\mathbb{R}^2$ described above, with H as in (10). The action ϕ is said to be *transitive* if there are no non-trivial G-invariant subsets of N. In the previous example the action is NOT transitive since each surface $H(z,p)=E$ is $\mathrm{PSL}(2,\mathbb{R})$-invariant; the action is transitive on each such surface, though. Another example is the action of $\mathrm{SO}(3)$ on $\mathbb{R}^6$ that we encountered previously. In that case the action is not transitive on the 5-dimensional energy surfaces $H(x)=E$, for a given central potential V, since the orbits of the action are given by the three dimensional surfaces $q^2=R^2, p^2=B^2, qp=C$, for some $R,B,C\in\mathbb{R}$. The presence of symmetries in a system is always a source of simplifications: in particular, the added group theoretical structures that it provides yield tools for understanding the dynamics. This is abundantly clear from the lectures of A. Venkatesh, in particular.

2.4 Two extremes: complete integrability and ergodicity

As I pointed out from the start, we are generally interested in understanding the behaviour of the solutions of Hamilton's equations: their global properties and possibly their asymptotic behaviour in time. Two important classes of systems are the completely integrable and the ergodic ones. Let Φ_t^H be a Hamiltonian flow on $\mathbb{R}^{2n}$. It is said to be *completely integrable* if there exist n functionally independent constants of the motion $f_1\dots f_n$, with $\{f_i,f_j\}=0$. Supposing they are compact, it can be proven that the level surfaces $\Sigma(c)=\{x\in\mathbb{R}^{2n}|f_j(x)=c_j\}$ are then *n-dimensional* tori on which the Hamiltonian flow acts as a translation flow (Liouville-Arnold) [1] [2]. The motion in such systems is very stable, in the sense that trajectories with nearby initial conditions only drift apart very slowly (linearly in time). An example are the Hamiltonians with a confining central potential: $H(x)=\frac{p^2}{2m}+W(\parallel q\parallel)$, $q,p\in\mathbb{R}^3$. One can then take $f_1=H, f_2=\ell^2, f_3=\ell_3$, for example.

A Hamiltonian flow is said to be *ergodic* if a typical trajectory explores the entire $2n-1$-*dimensional energy surface* Σ_E. More precisely, given a typical trajectory on the energy surface Σ_E, the time it spends in any subset B of Σ_E is asymptotically equal to the relative size of that set in the full energy surface:

$$\lim_{T\to\infty}\frac{|\{0\le t\le T|x(t)\in B\}|}{T}=\frac{|B|}{|\Sigma_E|}.$$

The motion in this case can be (but need not be) very unstable: nearby trajectories may drift apart exponentially fast. This is the case for the geodesic flow on the hyperbolic surfaces as explained by A. Venkatesh. I will give another example below.

These two situations are of course very different: the present notes will be exclusively concerned with the second case.

The geodesic flow on the Poincaré half plane, viewed as a Hamiltonian flow on the corresponding phase space, is completely integrable, as I showed above. Note however that the surfaces $H = E, C = c$ are not tori since they are not compact: they are made up of all semi-circular orbits centered at the same point c on the x-axis, and with the same energy. In fact, the dynamics is in that case highly unstable. Now, suppose Γ is some lattice in $\mathrm{PSL}(2,\mathbb{R})$: you can then quotient $\mathbb{H}\times\mathbb{R}^2$ by the φ_g^*, $g\in\Gamma$. Since $H\circ\varphi_g^* = H$, the Hamiltonian passes to the quotient, and since functions on the quotient can be seen as Γ-periodic functions on $\mathbb{H}\times\mathbb{R}^2$, so does the Poisson bracket. As a result, the geodesic flow on $\Gamma\backslash\mathbb{H}$ is still a Hamiltonian flow, but it is no longer completely integrable! Indeed, the functions f_i are not Γ invariant and do not pass to the quotient. In fact, the flow now becomes mixing, and hence ergodic as proven in one of A. Venkatesh's lectures. This proof requires a fair amount of advanced material and preparation (in particular the representation theory of $\mathrm{PSL}(2,\mathbb{R})$), and can not be called trivial. In contrast, the simplest unstable, mixing and hence ergodic Hamiltonian dynamical systems are discrete ones, to which I now turn.

2.5 Classical mechanics on the torus

Consider $A\in\mathrm{SL}(2,\mathbb{Z})$, $\mid \mathrm{Tr}A \mid > 2$. Then A has two real eigenvalues and eigenvectors: $Av_\pm = \mathrm{e}^{\pm\gamma_0}v_\pm$. As I already pointed out, A acts as a symplectic map on $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, meaning

$$\{f\circ A, g\circ A\} = \{f,g\}\circ A.$$

It therefore defines a discrete Hamiltonian dynamical system by iteration. This system is *hyperbolic*, meaning that for a.e. $x, x'\in\mathbb{T}^2,\ t\in\mathbb{N}$ (not too large),

$$\mathrm{d}(x,x')\sim\epsilon \Rightarrow \mathrm{d}(A^tx, A^tx')\sim\epsilon\mathrm{e}^{\gamma_0 t}.$$

So nearby initial conditions are exponentially quickly separated from each other by the dynamics. Note that this is also a feature of the geodesic flow on the Poincaré half plane (and on its quotients by discrete subgroups Γ). This is sometimes referred to as the "butterfly effect" and is considered a crucial feature of any chaotic system. As a result of this, the maps are *exponentially mixing* : $\forall f,g\in C^\infty(\mathbb{T}^2)$,

$$\mid\int_{\mathbb{T}^2}(f\circ A^t)(x)g(x)dx - \int_{\mathbb{T}^2}f(x)dx\int_{\mathbb{T}^2}g(x)dx \mid\leq C_{A,f}\parallel\nabla g\parallel_1\ \mathrm{e}^{-\gamma_0 t}.$$

Contrary to what happens with the geodesic flow on the modular surface, this is straightforward to show with a little Fourier analysis (see [4], for example).

As a result, the map is also *ergodic*: for all $f \in C^\infty(\mathbb{T}^2)$, for almost all $x_0 \in \mathbb{T}^2$

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f(A^t x_0) = \int_{\mathbb{T}^2} f(x)\mathrm{d}x.$$

Here are some typical examples: the first is the so-called Arnold Cat Map,

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

which belongs to the following family. For $a, b \in \mathbb{N}_*$

$$A_{a,b} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix} = \Phi_a^{p^2/2} \circ \Phi_{-b}^{q^2/2} = \begin{pmatrix} 1+ab & a \\ b & 1 \end{pmatrix}. \tag{17}$$

Another family of examples is given by $g \in \mathbb{N}_*$,

$$A_g = \begin{pmatrix} 2g & 1 \\ 4g^2 - 1 & 2g \end{pmatrix}.$$

Note that all these A act linearly on $\mathbb{R}^2$ and pass through the quotient by $\mathbb{Z}^2$ since they have integer entries. As dynamical systems on $\mathbb{R}^2$ they are already unstable, but not ergodic. The analogy with the geodesic flow on the modular surface is therefore quite clear. In both cases one starts from an unstable Hamiltonian dynamical system on a non-compact (infinite volume) space, and then considers a compact (or at least finite volume) quotient of this space on which the dynamics still acts symplectically, but is now exponentially mixing.

2.6 Summing up

From a rather abstract point of view, doing classical mechanics means studying certain (discrete or continuous) symplectic dynamical system on a symplectic manifold N, equipped with a Poisson bracket $\{\cdot,\cdot\}$. Such a system is said to be chaotic if it is exponentially unstable in a suitable sense. The geodesic flow on the modular surface and the iteration of a hyperbolic $\mathrm{SL}(2,\mathbb{Z})$ matrix on the torus are two examples of such systems. Models arising in real physical problems tend to have a rather more involved behaviour, with parts of their phase space where the motion is stable, and other parts where it is unstable (to varying degrees). So, to conclude, let me say this. As a first ingredient towards understanding the interest in and the link between equidistribution for quantum maps on the torus and for the eigenfunctions of the Laplace Beltrami operator on congruence surfaces, it is helpful to keep the following line of thinking in mind. Suitably adapted, it applies in many other scientific endeavours as well, and can be most helpful when writing introductions to scientific papers or grant applications. It goes as follows. The dynamics of many physical systems is given by a Hamiltonian flow Φ_t^H on $\mathbb{R}^{2n}$. Since physics is obviously important, it is therefore of great value to study Hamiltonian flows Φ_t^f on arbitrary symplectic manifolds. Since in such generality, only rather soft general statements can be made, it is of interest to understand relevant and tractable examples, such as geodesic flows on Riemannian manifolds, which already display a rich variety of behaviour, even if their physical pertinence is perhaps not totally clear. But since even this is often still very hard, one can hope to get insight in various issues by studying particular such manifolds, or, simpler yet, discrete Hamiltonian dynamical systems, meaning iterations of a fixed map Φ, for example an element of $\mathrm{SL}(2,\mathbb{Z})$ on $\mathbb{T}^2$.

3 A crash course in quantum mechanics

3.1 Schrödinger's quantum mechanics

Quantum mechanics is a physical theory that was developed in the first three decades of the twentieth century to deal with a number of issues in atomic physics that could not be dealt with using classical mechanics. I will present the theory here in a nutshell, glossing over both mathematical and physical subtleties. For some more detail, you may consult [4] where you will find further references if you want to get serious.

According to quantum mechanics, if one wants to study the motion of a particle in a potential $V:\mathbb{R}^3\to\mathbb{R}$, one should not solve Newton's equation nor Hamilton's equation, but the Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial\psi_t}{\partial t}(y) = -\frac{\hbar^2}{2m}\Delta\psi_t(y) + V(y)\psi_t(y), \qquad \psi_0=\phi, \qquad \parallel \phi \parallel = 1.$$

Here $\hbar$ is a physical constant called Planck's constant ($10^{-34}\mathrm{kgm}^2/\mathrm{s}$) and the unknown is the function $t\in\mathbb{R}\mapsto\psi_t\in L^2(\mathbb{R}^3,\mathbb{C};\mathrm{d}y)$. One calls ψ_t *the wavefunction* of the particle at time t and refers to $L^2(\mathbb{R}^3)$ as the *quantum state space* or the *Hilbert space of states*. It contains all information about the particle's state. It therefore "replaces" $(q(t),p(t))$, which played this role in classical mechanics. Now this seems like a crazy idea: how can a complex valued function be used to describe the motion of a particle? According to quantum mechanics, to extract information about the particle from the wavefunction, one needs to proceed as follows. First, the probability to find the particle at time t inside a set $B\subset\mathbb{R}^3$ is $\int_B \mid\psi_t\mid^2(y)\mathrm{d}y$ and the probability that its momentum falls inside some set $C\subset\mathbb{R}^3$ is $\int_C\mid\tilde{\psi}_t\mid^2(p)\mathrm{d}p$, where $\tilde{\psi}_t$ is the Fourier Transform of ψ_t:

$$\tilde{\psi}_t(p) = \frac{1}{(2\pi\hbar)^{3/2}}\int_{\mathbb{R}^3}\mathrm{e}^{-\mathrm{i}\frac{px}{\hbar}}\psi_t(y)\mathrm{d}y.$$

In particular, the mean position and momentum of the particle are

$$\int_{\mathbb{R}^3} y_j\mid\psi_t\mid^2(y)\mathrm{d}y, \qquad \int_{\mathbb{R}^3} p_j\mid\tilde{\psi}_t\mid^2(p)\mathrm{d}p.$$

Note that this makes sense since it is readily checked that the Schrödinger equation preserves the L^2-norm, so that $|\psi_t(y)|^2$ does indeed define a probability density. By the Plancherel identity, the same is then true for $|\hat{\psi}_t(p)|^2$. Since you may still find this strange, and be sceptical as to why this would have anything to do with the motion of a particle in a potential, let's try to demystify this. Introduce the following operators on (a suitable dense subspaces of) $L^2(\mathbb{R}^3, \mathrm{d}y)$:

$$P_j\psi_t(y) = \frac{\hbar}{\mathrm{i}}\frac{\partial \psi_t}{\partial y_j}(y), \quad Q_j\psi_t(y) = y_j\psi_t(y).$$

Then

$$\langle Q_j\rangle_t := \langle Q_j\psi_t, \psi_t\rangle = \langle \psi_t, Q_j\psi_t\rangle = \int_{\mathbb{R}^3} y_j \mid \psi_t \mid^2 (y)\mathrm{d}y$$

and

$$\langle P_j\rangle_t := \langle P_j\psi_t, \psi_t\rangle = \langle \psi_t, P_j\psi_t\rangle = \int_{\mathbb{R}^3} p_j \mid \tilde{\psi}_t \mid^2 (p)\mathrm{d}p$$

are the mean position and momentum. So the mean position and momentum can be written as matrix elements of certain self adjoint operators. Here

$$\langle \phi, \psi\rangle := \int_{\mathbb{R}^3} \overline{\phi}(y)\psi(y)\mathrm{d}y.$$

Moreover, the "Canonical Commutation Relations" hold, namely

$$[Q_j, P_k] = \mathrm{i}\hbar\delta_{jk}.$$

Now, if one defines $\hat{H} = \frac{P^2}{2m} + V(Q)$ one can rewrite the Schrödinger equation as

$$i\hbar\partial_t\psi_t = \hat{H}\psi_t, \quad \psi_0 = \phi.$$

Since, for all $\phi, \psi \in L^2(\mathbb{R}^3, \mathrm{d}y)$, $\langle \psi, \hat{H}\phi\rangle = \langle \hat{H}\psi, \phi\rangle$, you can solve it through $\psi_t = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}t\hat{H}}\phi$. Consequently

$$\begin{aligned} \frac{d}{dt}\langle \psi_t, Q_j\psi_t\rangle &= \langle \partial_t\psi_t, Q_j\psi_t\rangle + \langle \psi_t, Q_j\partial_t\psi_t\rangle \\ &= \frac{1}{i\hbar}\langle \psi_t, [Q_j, \hat{H}]\psi_t\rangle = \frac{1}{m}\langle \psi_t, P_j\psi_t\rangle, \end{aligned}$$

and

$$\frac{d}{dt}\langle \psi_t, P_j\psi_t\rangle = \frac{1}{i\hbar}\langle \psi_t, [P_j, \hat{H}]\psi_t\rangle = -\langle \psi_t, \nabla_j V(Q)\psi_t\rangle.$$

Those equations are called the Ehrenfest equations. They should remind you of Hamilton's formulation of Newton's equation $\dot{q} = p/m, \dot{p} = -\nabla V(q)$. They can be paraphrased as saying that the mean momentum equals the time change of the mean position and that the time change of the mean momentum equals the mean force. Beware however: whereas Hamilton's equations are a system of ordinary differential equations for the unknown $t \mapsto (q(t), p(t))$, the Ehrenfest equations

$$\frac{d}{dt}\langle \psi_t, Q_j\psi_t\rangle = \frac{1}{m}\langle \psi_t, P_j\psi_t\rangle, \qquad \frac{d}{dt}\langle \psi_t, P_j\psi_t\rangle = -\langle \psi_t, \nabla_j V(Q)\psi_t\rangle.$$

are not in general a system for $t \mapsto (\langle Q\rangle_t, \langle P\rangle_t)$! Indeed, in general

$$\langle\psi_t, \nabla_j V(Q)\psi_t\rangle = \int_{\mathbb{R}^3} \nabla_j V(y)|\psi_t(y)|^2 \mathrm{d}y \neq \nabla_j V(\langle\psi_t, Q_j\psi_t\rangle).$$

The mean force is not equal to the value of the force at the mean position, a fact we are used to from probability theory. A notable exception is the case of a harmonic potential $V(q) = \frac{1}{2}q^T\Omega^2 q$. Then

$$\frac{d}{dt}\langle\psi_t, Q_j\psi_t\rangle = \frac{1}{m}\langle\psi_t, P_j\psi_t\rangle, \qquad \frac{d}{dt}\langle\psi_t, P_j\psi_t\rangle = -\Omega^2_{jk}\langle\psi_t, Q_k\psi_t\rangle.$$

The mean position and momentum then follow the classical trajectories of the system in phase space. In particular

$$\langle Q\rangle_t = \cos\Omega t\langle Q\rangle_0 + \frac{\sin\Omega t}{\Omega}\langle P\rangle_0.$$

The moral of this story is that quadratic Hamiltonians, which give rise to a linear classical dynamics, are particularly simple, even in quantum mechanics! In general however, to study the time evolution of the quantum system means solving the Schrödinger equation directly. Since it is a partial differential equation, this is not an easy task.

How does one go about that task? Since $\hat{H} = \hat{H}(\hbar)$ is self adjoint, an obvious guess is to look for an orthonormal basis of eigenfunctions of $\hat{H}(\hbar)$:

$$\hat{H}(\hbar)\psi_n^\hbar = E_n(\hbar)\psi_n^\hbar, \quad \text{then} \quad \psi_t = \sum_n \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_n(\hbar)t}\langle\psi_n^\hbar, \phi\rangle\psi_n^\hbar.$$

Such a basis exists if $V(q) \to +\infty$ as $q \to \infty$. If you have enough information about the $\psi_n^\hbar$ and the $E_n(\hbar)$, you can then hope to obtain information about

$$\psi_t = \sum_{n=0}^{\infty} \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}E_n(\hbar)t}\langle\psi_n^\hbar, \phi\rangle\psi_n^\hbar.$$

Of course, unless V is quadratic and in a few other special cases, it is impossible to compute the eigenfunctions and eigenvalues explicitly. Studying the behaviour of the spectrum $E_n(\hbar)$ and of the eigenfunctions $\psi_n^\hbar$ then leads to a wealth of interesting and hard problems.

One interesting question is the asymptotic behaviour along sequences $E_n(\hbar)$ that converge to a fixed value E_c as $\hbar$ goes to zero: this is part of a field called *semi-classical analysis*, for which specific techniques have been developed and to which I turn in the next section. It turns out that the behaviour of the eigenfunctions and of the eigenvalues is in that limit determined by the properties of the Hamiltonian flow of $H(q,p) = \frac{P^2}{2m} + V(q)$ on the energy surface $H(q,p) = E$.

To sum up, let me say this. In quantum mechanics the time evolution of a system is no longer given by a Hamiltonian flow Φ_t^H on a symplectic manifold (the *classical phase space*), but by a unitary flow $U_t = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$ on a complex Hilbert space $\mathcal{H}$ (the *quantum state space*). Whereas the symplectic flow is generated by a function H, the unitary flow is generated by a self-adjoint operator $\hat{H}$. This is one example of a more general analogy between quantum and classical mechanics. Here is another one. Recall that the rotations R_t about the third axis are generated by $\ell_3 = q_1p_2 - q_2p_1$. Consider now the self-adjoint operator $L_3 = Q_1P_2 - Q_2P_1$. It is child's play to check that it generates a unitary group as follows:

$$\mathrm{e}^{-\mathrm{i}tL_3}\psi(y) = \psi(R_{-t}y).$$

Similarly

$$\mathrm{e}^{-\mathrm{i}a\cdot P}\psi(y) = \psi(y-a), \forall a \in \mathbb{R}^3.$$

So rotations and translations act by unitary operators on $\mathcal{H}$ and their generators are functions of the position and momentum operators in complete analogy with the situation in classical mechanics.

In the previous example, $\mathcal{H} = L^2(\mathbb{R}^3)$, but other situations arise. A typical example, beyond the ones given, is

$$\mathcal{H} = L^2(M, \mathrm{dvol}_g), \ \hat{H} = -\hbar^2\Delta_g \ .$$

Here (M,g) is a (compact) Riemannian manifold and Δ_g the Laplace-Beltrami operator on it. In this case also, the asymptotic behaviour of the eigenvalues and of the eigenfunctions is sensitive to certain statistical properties of the Hamiltonian flow generated by $H(q,p) = \frac{1}{2}p^Tg(q)^{-1}p$, *i.e.* of the geodesic flow. This is certainly not obvious a priori, but is the subject of the lectures of E. Lindenstrauss and A. Venkatesh. In the next subsection, I will show briefly, by developing the example of the half plane somewhat, why one may expect such a link between the geodesic flow and the properties of the eigenfunctions of the Laplace-Beltrami operator.

But one can also consider simpler examples, where the quantum Hilbert space is finite-dimensional and the quantum dynamics is no longer a unitary flow, but is replaced by the iteration of a fixed unitary map (so you obtain a $\mathbb{Z}$-action, rather than an $\mathbb{R}$-action). This will be the subject of Section 5 and the contribution of Z. Rudnick in this volume. In that case, as we shall see, the semi-classical limit is one in which the dimension of the finite dimensional Hilbert spaces tends to infinity. As a side remark, let me point out that in quantum computing, quantum cryptography and quantum information theory, one almost exclusively deals with finite dimensional Hilbert spaces (the N-fold tensor product of $\mathbb{C}^2$ with itself): a computation is then a certain product of unitaries.

3.2 Quantum mechanics on the Poincaré half plane

Suppose now that, armed with the insights of the previous sections, you had to (re)invent the quantum mechanics of a particle moving on the Poincaré half plane. How would you proceed? First, you would need an appropriate Hilbert space. A reasonable choice is to take $L^2(\mathbb{H}, y^{-2}\mathrm{d}x\mathrm{d}y)$, since the Riemannian volume element on $\mathbb{H}$ is $\sqrt{\det g(z)}\mathrm{d}x\mathrm{d}y = y^{-2}\mathrm{d}x\mathrm{d}y$. But how to define the appropriate quantum mechanical Hamiltonian? Pushing the analogy with the previous section, one is tempted to introduce operators X and Y, of multiplication by x and y respectively, and $P_x = \frac{\hbar}{\mathrm{i}}\partial_x$, $P_y = \frac{\hbar}{\mathrm{i}}\partial_y$, and then to define, inspired by (10):

$$\hat{H} = \frac{1}{2}Y^2(P_x^2 + P_y^2) = -\frac{1}{2}y^2(\partial_x^2 + \partial_y^2).$$

With this choice, $\hat{H} = -\frac{1}{2}\hbar^2\Delta_{\mathbb{H}}$, where $\Delta_{\mathbb{H}}$ is the Laplace-Beltrami operator on $\mathbb{H}$. Note that, with this choice, P_y is not a self-adjoint operator, but $\hat{H}$ is, as is readily checked. The corresponding Schrödinger equation becomes $\mathrm{i}\hbar\partial_t\psi_t = \hat{H}\psi_t$.

Of course, as before, we want to interpret $|\psi_t|^2(z)$ as a probability density so that

$$\langle X\rangle_t := \int_{\mathbb{H}} x|\psi_t|^2(z)y^{-2}\mathrm{d}x\mathrm{d}y, \quad \langle Y\rangle_t := \int_{\mathbb{H}} y|\psi_t|^2(z)y^{-2}\mathrm{d}x\mathrm{d}y,$$

are the mean values of the coordinates of the particle. A simple computation, as in the derivation of the Ehrenfest equations, shows that if ψ_t solves the above Schrödinger equation, then, first of all,

$$\langle V_x\rangle_t := \frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle_t = \langle Y^2P_x\rangle_t, \quad \langle V_y\rangle_t := \frac{\mathrm{d}}{\mathrm{d}t}\langle Y\rangle_t = \langle Y^2P_y\rangle_t.$$

Note that both "velocity operators" $V_x = Y^2P_x$ and $V_y = Y^2P_y$ are self-adjoint. Furthermore

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle X\rangle_t = \langle V_xY^{-1}V_y + V_yY^{-1}V_x\rangle_t = \langle 2V_yY^{-1}V_x + \mathrm{i}\hbar V_x\rangle_t,$$

and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\langle Y\rangle_t = \langle -\frac{1}{Y}V_x^2 + V_yY^{-1}V_y\rangle_t.$$

Again, comparing these last equations to (11) it is clear that the mean acceleration of the quantum particle obeys equations that bear a striking resemblance to the geodesic equations of motion on the Poincaré half plane. This begins to explain why the properties of the unitary group $\mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}$ and therefore of the eigenfunctions of $\hat{H}$ are influenced by properties of the geodesic flow. This influence is most striking asymptotically for small values of $\hbar$, as will become clearer in the following sections.

Let me end this section by a short discussion of the symmetries of the Hamiltonian on the quantum level. Let me first point out that, with the above choice of Hilbert space, the operators

$$U(g)\psi(z) = \psi(\varphi_g^{-1}(z)), \quad g \in \mathrm{PSL}(2, \mathbb{R})$$

are unitary, as is readily checked. This means the isometries of the Poincaré half plane, which are symmetries of the classical dynamical system, are realized by unitaries in the quantum Hilbert space. This is completely analogous to the representation of the Euclidean group on $L^2(\mathbb{R}^3, \mathrm{d}y)$, briefly mentioned in the previous subsection. In addition, one readily checks that

$$U(\mathrm{e}^{t\xi_j}) = \mathrm{e}^{-\mathrm{i}tF_j},\ F_1 = P_x,\ F_2 = -(X^2-Y^2)P_x - 2XYP_y,\ F_3 = 2(XP_x+YP_y),$$

which is to be compared to (16). Moreover, the F_j are constants of the motion since $[\hat{H}, F_j] = 0$ as a consequence of $U(g)^*\hat{H}U(g) = \hat{H}$, for all $g \in \mathrm{SL}(2,\mathbb{R})$.

4 Two words on semi-classical analysis

It was Dirac who pointed out the amazing analogy between $\{q_j, p_k\} = \delta_{jk}$ and $\frac{1}{i\hbar}[Q_j, P_k] = \delta_{jk}$. This suggests that in quantum mechanics the Lie-algebra of operators on the Hilbert space $L^2(\mathbb{R}^n)$ (under the usual commutator of operators) replaces the Lie-algebra of smooth functions on phase space $\mathbb{R}^{2n}$ that appears in Hamiltonian mechanics. In classical mechanics the observables are represented by functions on phase space, in quantum mechanics by operators on a Hilbert space, as we saw on some examples. So a natural question is whether there exists a Lie-algebra homomorphism between a suitable space of functions $\mathrm{Fn}(\mathbb{R}^{2n})$ on $\mathbb{R}^{2n}$, including, say all polynomials, and the operators on $L^2(\mathbb{R}^n)$. More precisely, does there exists a linear map

$$\mathrm{O}p : f \in \mathrm{Fn}(\mathbb{R}^{2n}) \subset C^\infty(\mathbb{R}^{2n}) \to \mathrm{O}p(f) : \mathcal{D} \subset L^2(\mathbb{R}^n) \to \mathcal{D} \subset L^2(\mathbb{R}^n),$$

such that

$$\frac{1}{i\hbar}[\mathrm{O}p(f), \mathrm{O}p(g)] = \mathrm{O}p(\{f, g\}),$$

and such that $\mathrm{O}p(q_j) = Q_j, \mathrm{O}p(p_j) = P_j$, and $Op\overline{f} = (Opf)^*$? It turns out that such a map does not exist, if $\mathrm{Fn}(\mathbb{R}^{2n})$ contains the space of all polynomials (Groenewold-Van Hove) [2] [4]. But a map having almost those properties does exist. It was proposed by Weyl and is called the Weyl quantization or Weyl symbol calculus. It is defined as follows. For any $f \in C^\infty(\mathbb{R}^{2n})$ (of at most polynomial growth), define its Fourier transform $\tilde{f}$ by

$$f(q,p) = \int_{\mathbb{R}^{2n}} \tilde{f}(a) e^{-\frac{i}{\hbar}(a_1 p - a_2 q)} \frac{\mathrm{d}a}{(2\pi\hbar)^n}.$$

Then define the *Weyl quantization of* f by

$$\mathrm{O}p^{\mathrm{W}}(f) = \int_{\mathbb{R}^{2n}} \tilde{f}(a) \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(a_1 P - a_2 Q)} \frac{\mathrm{d}a}{(2\pi\hbar)^n}.$$

Note that this is an operator since

$$U(a) := \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}(a_1 P - a_2 Q)}$$

is. Of course, to give a precise mathematical meaning to the expression for $\mathrm{O}p^{\mathrm{W}}$, one needs to say in which sense the integral converges, but I will only need this formula when f is $\mathbb{Z}^2$-periodic, in which case that is a trivial matter to which I will come back in the next section. Then

$$\frac{1}{i\hbar}[\mathrm{O}p^W(f), \mathrm{O}p^W(g)] = \mathrm{O}p^W(\{f, g\}) + O(\hbar),$$

and there is no error term as long as f and g are polynomials of degree at most two in the q_j and the p_j. Quite explicitly, one has for example for $f(q,p) = h(q)$, respectively $g(q,p) = k(p)$

$$\mathrm{O}p^W(f) = h(Q) \quad \mathrm{O}p^W(g) = k(P).$$

and

$$\mathrm{Op}^{\mathrm{W}} q_j p_j = \frac{1}{2}(Q_j P_j + P_j Q_j).$$

A crucial, all important property of the Weyl quantization is the so-called Egorov theorem, of which I will give the following approximate statement. For all $f, g \in C^\infty(\mathbb{R}^{2n})$ (of at most polynomial growth), one has, for all $t \in \mathbb{R}$

$$\mathrm{e}^{\frac{i}{\hbar}\mathrm{O}p^{\mathrm{W}}(g)t}\mathrm{O}p^{\mathrm{W}}(f)\mathrm{e}^{-\frac{i}{\hbar}\mathrm{O}p^{\mathrm{W}}(g)t} = \mathrm{Op}^{\mathrm{W}}(f \circ \Phi_t^g) + O_t(\hbar)$$

Moreover, if g is a quadratic function, the error term vanishes.

To see why this is useful, note that $\mathrm{Op}^{\mathrm{W}} H = \hat{H}$ if $H(q,p) = \frac{p^2}{2m} + V(q)$. So, since the solution ψ_t of the Schrödinger equation $\mathrm{i}\hbar\partial_t\psi_t = \hat{H}\psi_t$ with initial condition $\psi_0 = \phi$ is $\psi_t = \mathrm{e}^{-\frac{\mathrm{i}}{\hbar}\hat{H}t}\,\phi$, we find that

$$\langle\psi_t, \mathrm{Op}^{\mathrm{W}}(f)\psi_t\rangle = \langle\phi, \mathrm{Op}^{\mathrm{W}}(f \circ \Phi_t^H)\phi\rangle + O_t(\hbar).$$

This strongly suggests that, if we know enough about the classical evolution Φ_t^H appearing in the right hand side, we can infer from it information about the quantum evolution in the left hand side, in the limit of small $\hbar$. This is indeed correct and at the core of all semi-classical analysis about which much more can be learned from [9] [8]. I will illustrate one aspect of this general philosophy in the remaining section.

5 Quantum mechanics on the torus

Let us now turn to the situation where the classical dynamics is a $\mathbb{Z}$ action on $\mathbb{T}^2$, obtained by iterating a fixed element $A \in \mathrm{SL}(2, \mathbb{Z})$ and address the following questions: What is the quantum Hilbert space of states? And the quantization of observables? And the dynamics?

Since the system has a two-dimensional phase space, it is reasonable to expect to describe the quantum states with wavefunctions $\psi(y)$ of one variable. But since the phase space is a torus, one expects that the wavefunctions must be periodic $\psi(y-1)=\psi(y)$, as well as their Fourier transforms: $\tilde{\psi}(p-1)=\tilde{\psi}(p)$. This intuition leads to the following definition. With

$$U(a)\psi(y)=\mathrm{e}^{-\frac{i}{\hbar}(a_1P-a_2Q)}\psi(y)=\mathrm{e}^{-\frac{i}{2\hbar}a_1a_2}\mathrm{e}^{\frac{i}{\hbar}a_2y}\psi(y-a_1),$$

define

$$\mathcal{H}_\hbar=\{\psi\in\mathcal{S}'(\mathbb{R})\ \mid\ U(1,0)\psi=\psi=U(0,1)\psi\}.$$

Now, these spaces are trivial (I mean, zero-dimensional), unless there exists a positive integer such that $2\pi\hbar N=1$. So this will be assumed to be the case from now on. The semi-classical limit $\hbar\to 0$ therefore becomes $N\to+\infty$. The elements of $\mathcal{H}_\hbar$ are easily described:

$$\psi\in\mathcal{H}_\hbar\ \Rightarrow\ \psi(y)=\frac{1}{\sqrt{N}}\sum_{\ell\in\mathbb{Z}}c_\ell\delta(y-\frac{\ell}{N});\quad c_{\ell+N}=c_\ell.$$

Introducing the vectors ($j\in\mathbb{Z}$)

$$e_j=\sqrt{\frac{1}{N}}\sum_{n\in\mathbb{Z}}\delta_{\frac{j}{N}+n},$$

this can be written

$$\psi=\sum_{j=1}^{N}c_je_j$$

which allows one to identify $\mathcal{H}_\hbar$ with $\mathbb{C}^N$. This will be exploited in the lectures of Z. Rudnick.

As for the quantization of observables, one simply uses the Weyl quantization introduced in the previous section. For $f\in C^\infty(\mathbb{T}^2), x=(q,p)\in\mathbb{T}^2$, write

$$f(x)=\sum_{n\in\mathbb{Z}^2}f_n\mathrm{e}^{-i2\pi(n_1p-n_2q)}.$$

The Weyl quantization of f can now be written

$$\mathrm{Op}^{\mathrm{W}}f=\sum_{n\in\mathbb{Z}^2}f_n\mathrm{e}^{-\mathrm{i}2\pi(n_1P-n_2Q)}=\sum_{n\in\mathbb{Z}^2}f_nU(\frac{n}{N}):\mathcal{H}_\hbar\to\mathcal{H}_\hbar.$$

The action of the phase space translation operators $U(\frac{n}{N})$ on $\mathcal{H}_\hbar$ is very simple on the above basis e_j. Again, to make the link with Z. Rudnick's lectures, let me write it out in some detail. First of all, one checks that, for $n=(n_1,n_2)\in\mathbb{Z}^2$,

$$U(\frac{n}{N})=(-1)^{\frac{n_1n_2}{N}}T_1^{n_1}T_2^{n_2},$$

where $T_1=U(\frac{1}{N},0), T_2=U(0,\frac{1}{N})$. Furthermore

$$T_1e_j=e_{j+1},\qquad T_2e_j=e^{\mathrm{i}2\pi\frac{j}{N}}e_j.$$

Up to a global normalization, the Hilbert space structure on $\mathcal{H}_\hbar$ is uniquely determined by the requirement that T_1, T_2 act unitarily, which implies that the e_j are mutually orthogonal and all have the same norm. The normalization is fixed by choosing them to be normalized. Note that any operator commuting with both T_1 and T_2, or, equivalently, with all $U(\frac{n}{N})$, is easily seen to be a multiple of the identity.

You will find the same results in Z. Rudnick's lectures, with (of course!) a slightly different notation: what he calls $\hat{Q}$, I have called T_2 and what he calls $\hat{P}$, I called T_1^*.

It remains to define the quantum dynamics, which ought to be a unitary map on $\mathcal{H}_\hbar$, the so-called quantum map. Let's treat the examples in (17). Defining (following Schrödinger!)

$$M(A) = \mathrm{e}^{-\frac{\mathrm{i}}{2\hbar}aP^2}\mathrm{e}^{\frac{\mathrm{i}}{2\hbar}bQ^2},$$

it is easy to check that, provided a and b are even, for all $t \in \mathbb{Z}$,

$$M(A)\mathcal{H}_\hbar = \mathcal{H}_\hbar \quad \text{and} \quad M(A)^{-t}\ \mathrm{Op}^{\mathrm{W}}f\ M(A)^t - \mathrm{Op}^{\mathrm{W}}(f \circ A^t) = 0.$$

This is a simple case of the EGOROV theorem, and there is no error term in $\hbar$ because the dynamics is linear. A similar construction works for all hyperbolic elements of $\mathrm{SL}(2, \mathbb{Z})$ as explained in [4] [10].

To sum up, $M(A)$ is the quantum map we wish to study. It is naturally related to the discrete Hamiltonian dynamics on $\mathbb{T}^2$ obtained by iterating A through the above version of the Egorov theorem. It acts on the N dimensional spaces $\mathcal{H}_\hbar$ and we are interested in the behaviour of its eigenfunctions and eigenvalues in the $N \to \infty$ limit:

$$M(A)\psi_j^{(N)} = \mathrm{e}^{i\theta_j^{(N)}}\psi_j^{(N)}, \quad j = 1 \dots N.$$

I now finally have all the ingredients needed to state the basic result on the eigenfunction behaviour of classically ergodic systems, the so-called Schnirelman theorem, and thereby to link these lectures to the title of the school: equidistribution.

Theorem 5.1 *[3] For "almost all" sequences $\psi_N \in \mathcal{H}_\hbar$, so that $M(A)\psi_N = \mathrm{e}^{i\theta_N}\psi_N$,*

$$\langle \psi_N, \mathrm{Op}^{\mathrm{W}} f \psi_N \rangle \overset{N \to +\infty}{\to} \int_{\mathbb{T}^2} f(x)dx, \quad \forall f \in C^\infty(\mathbb{T}^2). \tag{18}$$

For a simple proof of this result, and a precise explanation of what is meant by the "allmost all" in its statement, I refer to [4] or [10]. Here I just want to explain in which sense this can be seen as an equidistribution result. For that purpose, note that, given $\psi_N \in \mathcal{H}_\hbar$, one can consider the map

$$\mu_N : f \in C^\infty(\mathbb{T}^2) \to \langle \psi, \mathrm{Op}^{\mathrm{W}} f \psi \rangle \in \mathbb{C}$$

as a distribution on the torus. They are referred to as the Wigner distributions of the ψ_N. Formally, one often writes

$$\mu_N(f) = \int_{\mathbb{T}^2} \mathrm{d}x \; W_N(x) f(x),$$

but the Wigner distributions are never functions, they are in fact sums of Dirac delta measures concentrated at the points $(m_1/2N, m_2/2N)$, $0 \leq m_1, m_2 < 2N$ on the torus.

The Schnirelman theorem can therefore be paraphrased as saying that (almost all) those Wigner distributions μ_N converge to the Lebesgue measure: in other words, they equidistribute. This is the precise meaning of the much used phrase: "the eigenfunctions equidistribute in phase space." Very formally, $W_N(x) \to 1$, but of course this is not a pointwise limit. Note that this equidistribution is a reflection of the ergodicity of the underlying classical dynamical system as will be abundantly clear from the proof of the theorem, should you read it.

An analogous theorem for arithmetic surfaces can be found in the contribution of E. Lindenstrauss in this volume. The similarities with the above theorem (which is more recent) should be obvious. If they aren't, I will have done a bad job.

The Schnirelman theorem invites an obvious interrogation. Is the "almost all" in its statement an artifact of the proof or do there exist sequences of eigenfunctions for which the Wigner functions do not converge to Lebesgue measure? It was shown in [6] that there do exist such sequences. For an overview of the situation as it is understood today, I refer to [5] as well as to the contribution of Z. Rudnick in this volume. The analogous question for arithmetic surfaces will be addressed by E. Lindenstrauss.

Number theory has played no role in my discussion. Nevertheless, it is present in the problem at hand, and it provides tools that can in particular be used to analyze the question raised in the previous paragraph as will be made clear in the lectures of Z. Rudnick .

References

[1] V. Arnold, *Mathematical methods in classical mechanics,* New-York-Berlin, Springer-Verlag (1973).

[2] R. Abraham R., J. Marsden, *Foundations of mechanics,* Benjamin-Cummings (1978).

[3] A. Bouzouina, S. De Bièvre, *Equipartition of the eigenfunctions of quantized ergodic maps on the torus*, Comm. Math. Phys. **178**, 83-105 (1996).

[4] S. De Bièvre, *Quantum chaos: a brief first visit*, mp_arc 01-207, Contemporary Mathematics **289**, 161-218 (2001).

[5] S. De Bièvre, *Recent results on quantum map eigenstates*, Proceedings of QMATH 9 (Giens, september 2004), preprint (2005).

[6] F. Faure, S. Nonnenmacher, S. De Bièvre, *Scarred eigenstates for quantum cats of minimal periods*, Commun. Math. Phys. **239**, 449-492 (2003).

[7] E. Lindenstrauss, *Some examples how to use measure classification in number theory*, in "Equidistribution in number theory," (Montréal 2005), 2006.

[8] A. Martinez, *An introduction to semiclassical and microlocal analysis*, Universitext, Springer-Verlag, New York, (2002).

[9] D. Robert, *Autour de l'approximation semi-classique*, Birkhäuser (1987).

[10] Z. Rudnick, *The arithmetic theory of quantum maps*, in "Equidistribution in number theory," (Montréal 2005), 2006.

[11] A. Venkatesh, *Spectral theory of automorphic forms: a very brief introduction*, in "Equidistribution in number theory," (Montréal 2005), 2006.

[12] S. Zelditch, *Quantum ergodicity and mixing of eigenfunctions*, math-ph/0503026, preprint (2005).