LECTURES ON BOUNCING BALLS.

1. Introduction.

1.1. **Goals of the lectures.** The purpose of these lectures is to illustrate some ideas and techniques of smooth ergodic theory in the setting of simple mechanical systems.

Namely we consider either one or several particles moving on a line either freely or in a field of a force and interacting with each other and with the walls according to the law of elastic collisions.

The main questions we are going to address are the following.

(1) *Acceleration.* Is it possible to accelerate the particle so that its velocity becomes arbitrary large? If the answer is YES we would like to know how large is the set of such orbits. We would also like to know how quickly a particle can gain energy both in the best (or worst) case scenario and for typical initial conditions. We are also interested to see if the particle will accelerate indefinitely so that its energy tend to infinity or if its energy will drop to its initial value from time to time.

(2) *Transitivity.* Does the system posses a dense orbit? That is, does there exist an initial condition (Q_0, V_0) such that for any ε and any $\bar{Q}, \bar{V}$ there exists t such that

$$|Q(t) - \bar{Q}| < \varepsilon, \quad |V(t) - \bar{V}| < \varepsilon.$$

A transient system has no open invariant sets. A stronger notion is *ergodicity* which says that any measurable invariant set either has measure 0 or its complement has measure 0. If the system preserves a finite measure μ and the system is ergodic with respect to this measure then by pointwise ergodic theorem for μ-almost all initial conditions we have

$$\frac{1}{T}\mathrm{mes}(t \in [0, T] : (Q(t), V(t)) \in A) \to \mu(A) \text{ as } T \to \infty.$$

If the measure of the whole system is infinite then we can not make such a simple statement but we have the Ratio Ergodic Theorem which says that for any sets A, B and for almost all initial conditions

$$\frac{\mathrm{mes}(t \in [0, T] : (Q(t), V(t)) \in A)}{\mathrm{mes}(t \in [0, T] : (Q(t), V(t)) \in B)} \to \frac{\mu(A)}{\mu(B)} \text{ as } T \to \infty.$$

The purpose of the introductory lectures is to introduce several examples which will be used later to illustrate various techniques. Most of the material of the early lectures can be found in several textbooks on dynamical systems but it is worth repeating here since it will help us to familiarize ourselves with the main examples. The material of the second part will be less standard and it will be of interest to a wider audience.

1.2. **Main examples.** Here we describe several simple looking systems which exhibit complicated behavior. At the end of the lectures we will gain some knowledge about the properties of these systems but there are still many open questions which will be mentioned in due course.

(I) Colliding particles. The simplest model of the type mentioned above is the following. Consider two particles on the segment $[0, 1]$ colliding elastically with each other and the walls. Let m_1 and m_2 denote the masses of the particles. Recall that a collision is elastic if both energy and momentum are preserved. That is, both

$$P = m_1 v_1 + m_2 v_2 \text{ and } 2K = m_1 v_1^2 + m_2 v_2^2$$

are conserved. In particular if $P = 0$ then $2K = m_2 v_2^2 \frac{m_2+m_1}{m_1}$ and so in this case $(v_2^+)^2 = (v_2^-)^2$. Similarly, $(v_1^+)^2 = (v_1^-)^2$, that is, the particles simply change the signs of their velocities. In the general case we can pass to the frame moving with the center of mass. The center of mass' velocity is $u = \frac{m_1 v_1 + m_2 v_2}{m_1+m_2}$ so in the new frame we have

$$\tilde{v}_1 = v_1 - u = \frac{m_2(v_1 - v_2)}{m_1 + m_2} \text{ and } \tilde{v}_2 = v_1 - u = \frac{m_1(v_2 - v_1)}{m_1 + m_2}.$$

In our original frame of reference we have

$$v_1^+ = u - \tilde{v}_1 = \frac{m_1 - m_2}{m_1 + m_2} v_1^- + \frac{2m_2}{m_1 + m_2} v_2^-$$

and similarly

$$v_2^+ = u - \tilde{v}_2 = \frac{m_2 - m_1}{m_1 + m_2} v_2^- + \frac{2m_1}{m_1 + m_2} v_1^-.$$

The collisions with the walls are described by the same formulas but we consider the walls to be infinitely heavy. Thus if the particle collides with the wall its velocity becomes $v^+ = 2v_{wall} - v^-$. In particular, in the present setting the wall is fixed so the particle's velocity just changes the sign.

Returning to our system introduce

$$q_j = \sqrt{m_j} x_j. \text{ Thus } u_j = \dot{q}_j = \sqrt{m_j} v_j. \tag{1.1}$$

The configuration space of the system becomes

$$q_1 \geq 0, \quad q_2 \leq \sqrt{m_2}, \quad \frac{q_1}{\sqrt{m_1}} \leq \frac{q_2}{\sqrt{m_2}}.$$

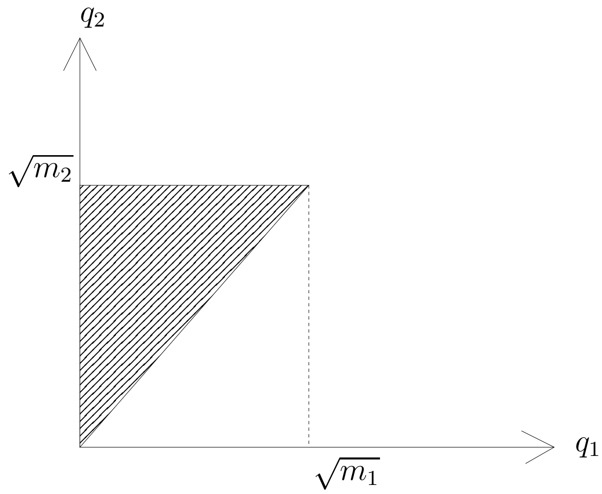

FIGURE 1. Configuration space for two points on the segment

This is a right triangle with hypothenuse lying on the line

$$q_1\sqrt{m_2} - q_2\sqrt{m_1} = 0.$$

The law of elastic collisions preserves

$$2K = u_1^2 + u_2^2 \text{ and } P = \sqrt{m_1}u_1 + \sqrt{m_2}u_2.$$

In other words if we consider $(q_1(t), q_2(t))$ as a trajectory of the particle in our configuration spaces then as the particle reaches hypothenuse its speed is preserved and the angle which its velocity makes with $(\sqrt{m_1}, \sqrt{m_2})$ remains the same. Since $(\sqrt{m_1}, \sqrt{m_2})$ is orthogonal to the boundary this change satisfy the law of the elastic reflection. Similarly if the particle hits $q_1 = 0$ then u_2 remains the same and u_1 changes to the opposite which is again in accordance with the elastic collision law. Hence our system is isomorphic to a billiard in a right triangle.

A similar analysis can be performed for three particles on the circle $\mathbb{R}/\mathbb{Z}$. In this case there are no walls so the velocity of the mass center is preserved. It is therefore convenient to pass to a frame of reference where this center is fixed at the origin. So we have

$$m_1x_1 + m_2x_2 + m_3x_3 = 0 \text{ and } m_1v_1 + m_2v_2 + m_3v_3 = 0.$$

In coordinates from (1.1) the above relation reads

$$\sqrt{m_1}q_1 + \sqrt{m_2}q_2 + \sqrt{m_3}q_3 = 0 \text{ and } \sqrt{m_1}u_1 + \sqrt{m_2}u_2 + \sqrt{m_3}u_3 = 0.$$

Thus points are confined to a plane Π and the particle velocity lies in this plane. The collisions of the particles have equations $\frac{q_i}{\sqrt{m_i}} - \frac{q_j}{\sqrt{m_j}} = l.$ These lines divide Π into triangles. We claim that dynamics restricted to each triangle is a billiard. Consider, for example, the collision of

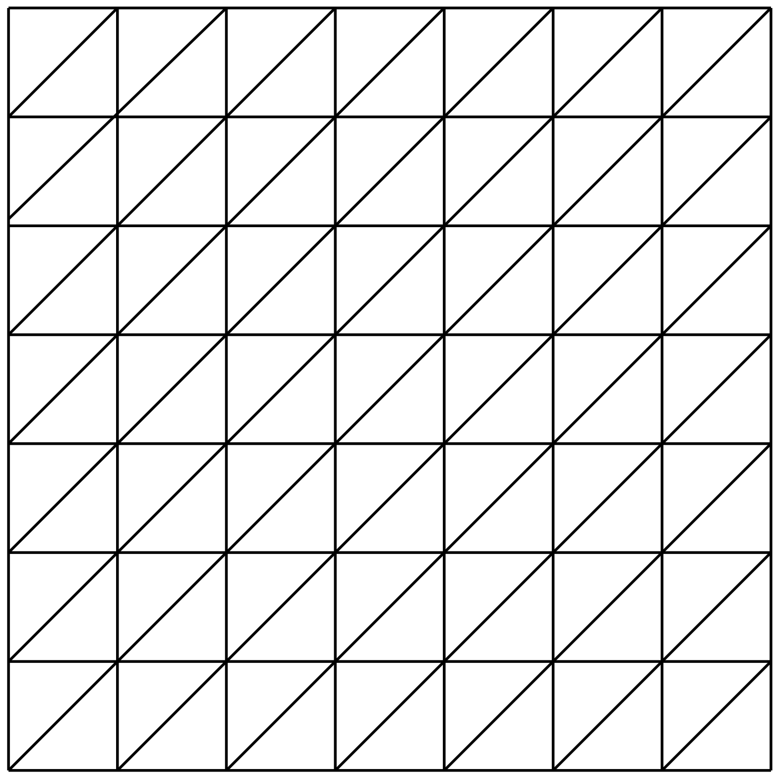

FIGURE 2. Configuration space of three points on the circle using the distance from the first point as coordinates

the first two particles. Since $\sqrt{m_1}u_1 + \sqrt{m_2}u_2$ is preserved we see that $\vec{u} = (u_1, u_2, u_3)$ is orthogonal to $\vec{n}_{12} = (\sqrt{m_1}, \sqrt{m_2}, 0)$. Note that n_{12} is ortogonal to the plane $P^l_{12} = \{\frac{q_1}{\sqrt{m_2}} - \frac{q_2}{\sqrt{m_2}} = l\}$. Denoting by $\vec{n}^*_{12}$ the orthogonal projection of $\vec{n}_{12}$ to Π we see that $\vec{n}^*_{12}$ is orthogonal to $\Pi \cap P^l_{12}$ and that the angle between $\vec{u}$ and $\vec{n}^*_{12}$ is preserved which again agrees with the law of elastic collision.

We can also consider more particles on a line or a circe and show that that system is isomorphic to a polyhedral billiard.

(II) Particle in a potential. Our second example is a particle moving on the line under the force created by the potential $U(x) = gx^\alpha$ and colliding elastically with an infinitely heavy plate. We assume that $\alpha > 0$ since otherwise the particle can go to infinity after finitely many bounces. Let $f(t)$ denote the height of the plane at time t. We assume that $f(t) > 0$ for all t so that $U(x)$ is defined for all $x > f(t)$ and that $f(t)$ is periodic. In fact, the case of $f(t) = B + A \sin t$ (where $A < B$)

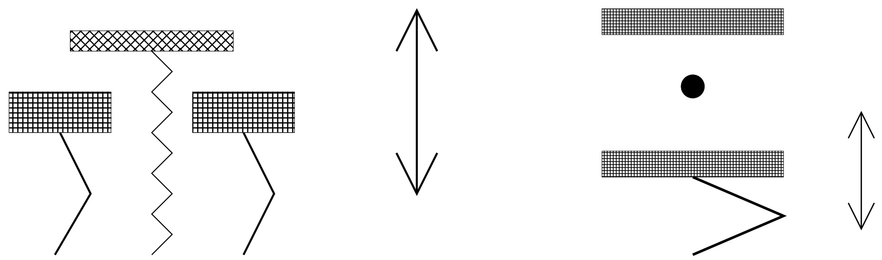

FIGURE 3. Impact oscillator (left) and Ulam pingpong (right) are two systems fitting into our setting

is already quite interesting. Two cases attracted a particular attention in the past.

(a) *Gravity* ($\alpha = 1$). In this setting the acceleration question can be posed as follows: how much can one accelerate a tennis ball by periodic motion of a tennis rocket (of course one needs to be in a good fitness condition for the infinitely heavy wall approximation to be reasonable).

(b) *Impact oscillator* ($\alpha = 2$). In this case one has a particle attached to a sting and colliding with the wall. Apart from an easy mechanical implementation this system is also to an interesting geometric object-outer billiard.

Outer billiards are defined in an exterior of a closed convex curve Γ on the plane. Given a point $A_0 \in \mathbb{R}^2 - \mathrm{Int}\Gamma$ there are two support lines from A_0 to Γ. Choose the one for which if one walks from A_0 to the point of contact then Γ is to the right of the line. Then we reflect A_0 about the point of contact to get its image A_1. Applying this procedure repeatedly we obtain the orbit of A_0 under the outer billiard map. Outer billiards were popularized by Moser as they provide simple illustration to KAM theory.

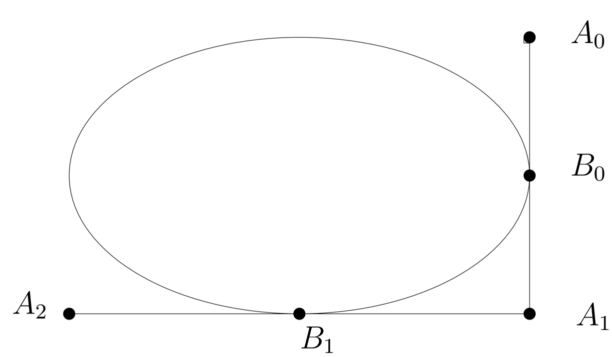

FIGURE 4. Outer billiard

We now describe a construction of Boyland [2] which associates to each outer billiard an impact oscillator. To this end we consider a third system (see Figure 5). Its phase space consists of a pair (Γ_0, A_0) where Γ_0 is a closed and convex curve and A_0 is the point in $\mathbb{R}^2 - \mathrm{Int}\Gamma_0$ such that the supporting line from A_0 to Γ_0 is vertical. To describe one iteration of our system one first reflects A_0 about the point of contact to get the pair $(\Gamma_0, \tilde{A}_1)$ and then rotates the picture counterclockwise until the second support line becomes vertical. If (Γ_n, A_n) is the n-th iteration of our system then clearly there exists a rotation R_n such that $\Gamma_0 = R_n\Gamma_n$. Then $R_n A_n = f^n_{\Gamma_0} A_0$ where f_{Γ_0} denotes the outer billiard map about Γ_0. On the other hand between the reflections the point evolves according to the ODE $\dot{x} = v, \quad \dot{v} = -x$ while during the reflection x is unchanged and $v^+ + v^- = 2v_{tip}$ where v_{tip} denotes the velocity of the rightmost point of $\Gamma(t)$. One can check that the motion of the tip is given by $\ddot{x} + x = r(x(t))$ where $r(x)$ is the radius of curvature of point x. Thus given a curve Γ one can associate to it an impact oscillator with the wall motion given by $\ddot{f} + f = r(f(t))$. Note that in that construct the frequencies of the wall and the spring are the same. Conversely, given an impact oscillator one can consider a curve whose radius of curvature is $r(f(t)) = \ddot{f} + f$ but the resulting curve need not be either close or convex. Thus the class of impact oscillators is much larger than the class of outer billiards but the later is an important subclass supplying clear geometric intuition.

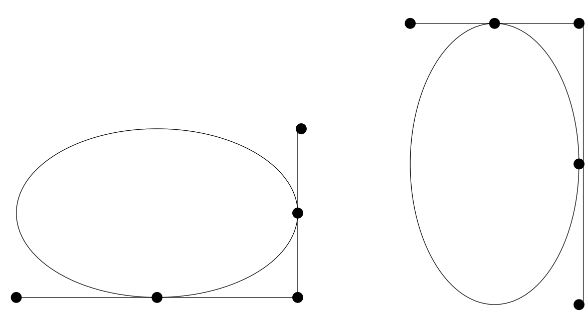

FIGURE 5. Outer billiards and Impact Oscillators

While $\alpha = 1$ and $\alpha = 2$ are the two most studied cases we will see that the dynamics for $\alpha \neq 1, 2$ is quite different. As it was mentioned above one of the main question is large velocity behavior of the model. Note that different collisions occur at different heights. However if the particle's velocity is high it takes a very short time to pass between $\max f(t)$ and $\min f(t)$. Since the explicit computations of the height of the next collision is usually impossible one often considers a simplified model which is called *static wall approximation* (SWA). In this model one fixes a height $\bar{h}$ and assumes that the next collision occurs at the time $t_{n+1} = t_n + T(v_n)$ where $T(v_n)$ is the time it takes the particle to return to the height $\bar{h}$. However velocity is still updated as $v_{n+1} = 2\dot{f}(t_{n+1}) - 2\tilde{v}_n$ where $\tilde{v}_n$ is velocity of the particle when it returns to $\bar{h}$. By energy conservation $\tilde{v}_n = -v_n$ so SWA takes form

$$t_{n+1} = t_n + T(v_n), \quad v_{n+1} = v_n + 2\dot{f}(t_{n+1}).$$

We note that while SWA provides a good approximation for the actual system in high velocity regime for one or a few collisions, in general, it is not easy to transfer the results between the original model and SWA. However the SWA is an interesting system in its own right. In addition, the SWA and the original system often have similar geometric features and since formulas are often simpler for the SWA we will often present the arguments for the SWA. For example, for $\alpha = 1$ the SWA takes from

$$t_{n+1} = t_n + 2\frac{v_n}{g}, \quad v_{n+1} = v_n - 2\dot{f}(t_{n+1}). \tag{1.2}$$

This system is the celebrated standard map. Phase portraits of the map (1.2) for several values of parameters can be found in Section 2.4 of [11]. (1.2) is defined on $\mathbb{R} \times \mathbb{T}$ but it is a lift of $\mathbb{T}^2$ diffeomorphism since the change of v by $\frac{2}{g}$ commutes with the dynamics.

(III) Fermi-Ulam pingpong. In model (II) the particle has infinitely many collisions with a moving wall because the force make it to fall down. Another way to enforce infinitely many bounces is to put the second stationary wall with which the particle collides elastically.

This model can be thought as a special case of the previous model where

$$U(x) = \begin{cases} 0, & \text{if } x \leq \bar{h} \\ \infty & \text{if } x > \bar{h} \end{cases} \tag{1.3}$$

where $\bar{h}$ is the height of the stationary wall. Pingpong model was introduced by Ulam to study Fermi acceleration. To explain the presence of highly energetic particles in cosmic rays Fermi considered particles passing through several galaxies. If the particle moves towards a galaxy it accelerates while if it goes in the same direction it deccelerates. Fermi argued that head-on collisions are more frequent than the overtaking collisions (for the same reason that a driver on a highway sees more cars coming towards her than going in the same direction even though the effect becomes less pronounced if the car's speed is 3000 m/h) leading to overall acceleration. Pingpong was a simple model designed to test this mechanism. This model was one of the first systems studied by a computer (first experiments were performed by Ulam and Wells around 1960). Since the computers were very slow at that time they chose wall motions which made computations simpler, namely, either wall velocity or interwall distance was piecewise linear. It was quickly realized that the acceleration was impossible for smooth wall motions. The motions studied by Ulam and Wells turned out to be more complicated and there are still many open questions.

All of the above systems can be considered Hamiltonian with potential containing hard core part (1.3). Accordingly these systems preserve measures with smooth densities. Consider for example models (II) and (III). It is convenient to study the Poincare map corresponding to collision of the particle with the moving wall. One can approximate the hard core systems by a Hamiltonian system with the Hamiltonian $H_\varepsilon = \frac{v^2}{2} + U(x) + W_\varepsilon(x - f(t))$ where $W(d)$ is zero for $d < \varepsilon$ and $W(-\varepsilon) = \frac{1}{\varepsilon}$. One can consider the collision map as the limit of Poincare map corresponding to the cross section $x - f(t) = \varepsilon$. The map preserve the form $\omega = dH \wedge dt - dv \wedge dx$. On our cross section we have $dx = \dot{f} dt$ so the invariant form becomes

$$\omega = (v - \dot{f}) dv \wedge dt. \tag{1.4}$$

One can also directly show that the form (1.4) is invariant without using approximation argument. Consider for example the pingpong system

$$t_{n+1} = t_n + T(t_n, v_n), \quad v_{n+1} = v_n + 2\dot{f}(t_{n+1}).$$

This map is a composition of two maps

$$\bar{t}_{n+1} = t_n + T(t_n, v_n), \quad \bar{v}_{n+1} = v_n$$

and

$$t_{n+1} = \bar{t}_{n+1}, \quad v_{n+1} = v_n + 2\dot{f}(\bar{t}_{n+1}).$$

Accordingly the Jacobian of this map equals to $\frac{\partial t_{n+1}}{\partial t_n}$. We have (see Figure 6)

$$\delta h_n = (v_n - \dot{f}_n) \delta t_n,$$

$$\delta t_{n+1} = \frac{\delta h_n}{v_n + \dot{f}_{n+1}} = \frac{v_n - \dot{f}_n}{v_{n+1} - \dot{f}_{n+1}} \delta h_n. \tag{1.5}$$

Thus the Jacobian equals to $\frac{v_n - \dot{f}_n}{v_{n+1} - \dot{f}_{n+1}}$ proving the invariance of ω.

A similar calculation can be done for the model (II) using the fact that autonomous Hamiltonian systems preserve the form $dv \wedge dx$.

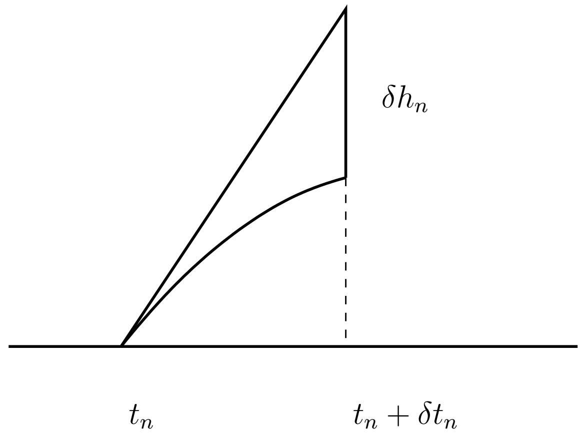

FIGURE 6. Derivative of pingpong map.

2. Normal forms.

2.1. **Smooth maps close to identity.** Here we discuss the behaviour of highly energetic particles using the methods of averagin theory. The following lemma will be useful.

Lemma 2.1. *Consider an area preserving map of the cylinder $\mathbb{R} \times \mathbb{T}$ of the form*

$$R_{n+1} = R_n + A(R_n, \theta_n), \quad \theta_{n+1} = \theta_n + \frac{B(R_n, \theta_n)}{R_n}.$$

Assume that the functions A and B admit the following asymptotic expansion for large R

$$A = \sum_{j=0}^{k} \frac{a_j(\theta)}{R^j} + \mathcal{O}\left(R^{-(k+1)}\right), \quad B = \sum_{j=0}^{k} \frac{b_j(\theta)}{R^j} + \mathcal{O}\left(R^{-(k+1)}\right) \tag{2.1}$$

where

$$b_0(\theta) > 0 \text{ (twist condition).}$$

Then for each k there exists coordinates $I^{(k)}, \phi^{(k)}$ such that $\frac{I}{R}$ is uniformly bounded from above and below and our map takes form

$$I_{n+1} = \mathcal{O}\left(I_n^{-(k+1)}\right), \quad \theta_{n+1} = \theta_n + \frac{1}{I_n}\left(\sum_{j=0}^{k} \frac{c_j}{I_n^j} + \mathcal{O}\left(I_n^{-(k+1)}\right)\right).$$

Remark 2.2. $I^{(0)}$ is called *adiabatic invariant* of the system. $I^{(k)}$ for $k > 0$ are called *improved adiabatic invariants.*

Proof. We proceed by induction.First let $I = R\Gamma(\theta)$, $\phi = \Phi(\theta)$ then

$$I_{n+1} - I_n = R_n\Gamma'(\theta_n)\frac{b_0(\theta_n)}{R_n} + a_0(\theta_n)\Gamma(\theta_n) + O\left(\frac{1}{R_n}\right).$$

So if we let $\frac{\Gamma'}{\Gamma} = -\frac{a_0}{b_0}$ that is

$$\Gamma(\theta_0) = \exp\left[\int_0^\theta -\frac{a_0(s)}{b_0(s)}ds\right]$$

then $I_{n+1} - I_n = \mathcal{O}(R_n^{-1})$.

Next

$$\phi_{n+1} - \phi_n = \Phi'(\theta_n)\frac{b_0(\theta_n)}{R_n} = \Phi'(\theta_n)\frac{b_0(\theta_n)\Gamma(\theta_n)}{I_n}.$$

We let

$$\Phi'(\theta) = \frac{c}{b_0(\theta)\Gamma(\theta)} \text{ so that } \Phi(\theta) = c\int_0^\theta \frac{ds}{b_0(s)\Gamma(s)} \text{ and } c = \left(\int_0^1 \frac{ds}{b_0(s)\Gamma(s)}\right)^{-1}.$$

Note that $\Gamma(1) = \Gamma(0)$ so that Γ is actually a function on the circle. Indeed if $\Gamma(1) < \Gamma(0)$ then there would exist a constant ε such that after one rotation around the cylinder R decreases at least by the factor $(1-\varepsilon)$. So after many windings the orbit would come closer and closer to the origin contradicting the area preservation. If $\Gamma(1) > \Gamma(0)$ we would get a similar contradiction moving backward in time.

This completes the base of iduction. The inductive step is even easier. Namely if $I_{n+1} = I_n + \frac{\hat{a}(\phi_n)}{I_n^{k+1}} + \dots$ then the changes of variables $J = I + \frac{\gamma(\phi)}{I^k}$ leads to

$$J_{n+1} - J_n = \frac{\hat{a}(\phi_n) + \gamma_n'(\phi_n)c_0}{J_n^{k+1}}$$

so we can improve the order of conservation by letting $\gamma' = -\frac{\hat{a}}{c_0}$.

Next, if $\phi_{n+1} - \phi_n = \frac{1}{I_n}\sum_{j=0}^{k-1}\frac{c_j}{I_n^j} + \frac{\hat{b}(\phi_n)}{I_n^{k+1}}$ then letting $\psi = \phi + \frac{\Psi(\phi)}{I^k}$ we obtain

$$\psi_{n+1} - \psi_n = \frac{1}{I_n}\sum_{j=0}^{k-1}\frac{c_j}{I_n^j} + \frac{\hat{b}(\psi_n) + \Psi'(\phi_n)c_0}{I^{k+1}}$$

allowing us to eliminate the next term if $\Psi' = \frac{c_k - \hat{b}}{c_0}$ where $c_k = \int_0^1 \hat{b}(s)ds$.
□

2.2. **Adiabatic invariants.** It is instructive and useful to compute the leading terms in several examples.

(I) Fermi-Ulam pingpong. We have

$$v_{n+1} - v_n \approx 2\dot{f}(t_n), \quad t_{n+1} - t_n \approx \frac{2l(t_n)}{v_n}$$

where $l(t)$ is the distance between the walls at time t. We have $l = \bar{h} - f$ so $\dot{f} = \dot{l}$ and the above equation is the Euler scheme for the ODE

$$\frac{dv}{dt} = -v\frac{\dot{l}}{l}. \quad \text{Thus} \quad ldv + vdl = 0$$

so $I = lv$ is an adiabatic invariant. In fact one can check by direct computation that letting $J_n = (v_n + \dot{l}(t_n))l(t_n)$ one gets

$$J_{n+1} - J_n = \mathcal{O}\left(\frac{1}{J_n^2}\right), \quad t_{n+1} - t_n = \frac{2l^2(t_n)}{J_n} + \mathcal{O}\left(\frac{1}{J_n^2}\right)$$

so J_n is the second order adiabatic invariant.

(II) Outer billiard. If A_0 is far from the origin then A_1 is close to $-A_0$, however $|A_0A_2| = 2|B_0B_1|$ there B_j denotes the point of tangency of A_jA_{j+1} with Γ (see Figure 4) and so $|A_0A_2| \leq 2\text{diam}(\Gamma)$. It fact it is not difficult to see that we get the following approximation when A_0 is far from the origin: $\vec{A_0A_2} \approx 2\vec{v}(\theta)$ where $\vec{v}(\theta)$ is the vector joing two points on Γ whose tangent line have slope θ. Let $B_0(\theta)$ and $B_1(\theta)$ denote the tangency points and let Q be the point such that B_1Q has slope θ while B_0Q is perpendicular to B_1Q. Note that $|B_0Q| = w(\theta)$-the width of Γ in the direction θ.

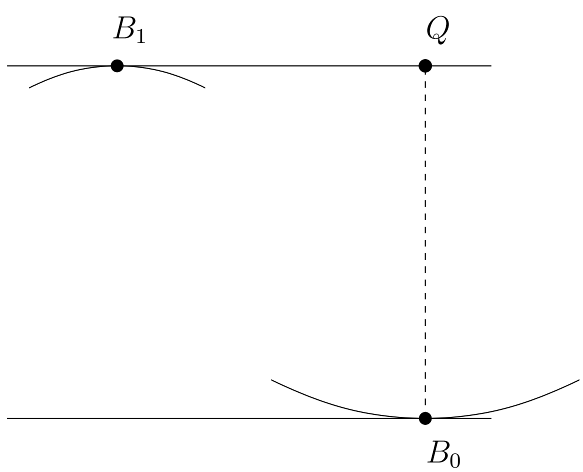

FIGURE 7. Derivative of the support function.

Fix a direction θ_0 and choose coordinates on the plane so that θ_0 is equal to 0. Let $B_j = (x_j, y_j)$. Then for θ near 0 we have

$$x_j(\theta) = x_j(0) + \theta\xi_j + \dots, \quad y_j(\theta) = y_j(0) + \theta^2\eta_j + \dots$$

and so

$$(x_1 - x_0, y_1 - y_0)(\sin\theta, \cos\theta) = -|QB_1|\theta + \dots.$$

Therefore the equation of motion takes the following form in polar coordinates (up to lower order terms).

$$\dot{R} = -w'(\theta), \quad \dot{\theta} = \frac{w(\theta)}{R}.$$

Hence

$$\frac{dR}{d\theta} = -R\frac{w'(\theta)}{w(\theta)} \quad \text{or} \quad wdR + Rdw = 0.$$

Accordingly $I = Rw$ is the adiabatic invariant and

$$\dot{\theta} = \frac{w(\theta)}{R} = \frac{w(\theta)R}{R^2} = \frac{I}{R^2}.$$

In other words $I = R^2\dot{\theta}$, that is the angular momentum is preserved and so the point moves with constant sectorial velocity.

Consider, in particular, the case where Γ is centrally symmetric. Then $w(\theta) = 2\sup_{x\in\Gamma}(e^\perp(\theta), x)$ and since $R = \frac{I}{w(\theta)}$ level curve of the limiting equation are rescalings of the right angle rotation of Γ^* where

$$\Gamma^* = \{D(e)e\}_{e\in S^1} \text{ and } D(e) = \frac{1}{\sup_{x\in\Gamma}(e, x)}.$$

Thus if $\hat{\Gamma} = \overline{\mathrm{Int}(\Gamma)}$ then

$$\widehat{\Gamma^*} = \{e \in \mathbb{R}^2 : |(e, x)| \le 1 \text{ for all } x \in \hat{\Gamma}\}.$$

Thus for each $x \in \Gamma$ and for all $e \in \Gamma^*$ we have $|(x, e)| \le 1$ and there is unique $e \in \Gamma^*$ with $(x, e) = 1$. Therefore $(\Gamma^*)^* = \Gamma$ and so each smooth convex centrally symmetric curve appears as an invariant curve for motion at infinity for some outer billiard.

2.3. **Systems with singularities.** Lemma 2.1 describes the normal form for smooth maps, so it is not applicable to systems with discontinuities such as Fermi-Ulam pingpongs where $\dot{l}$ or $\ddot{l}$ has jumps or to outer billiards about nonsmooth curves such as circular caps or lenses. It is turns out that for such maps it is convenient to consider the first return map to a neighbourhood of singularities. In this section we present the normal form of such first return maps.

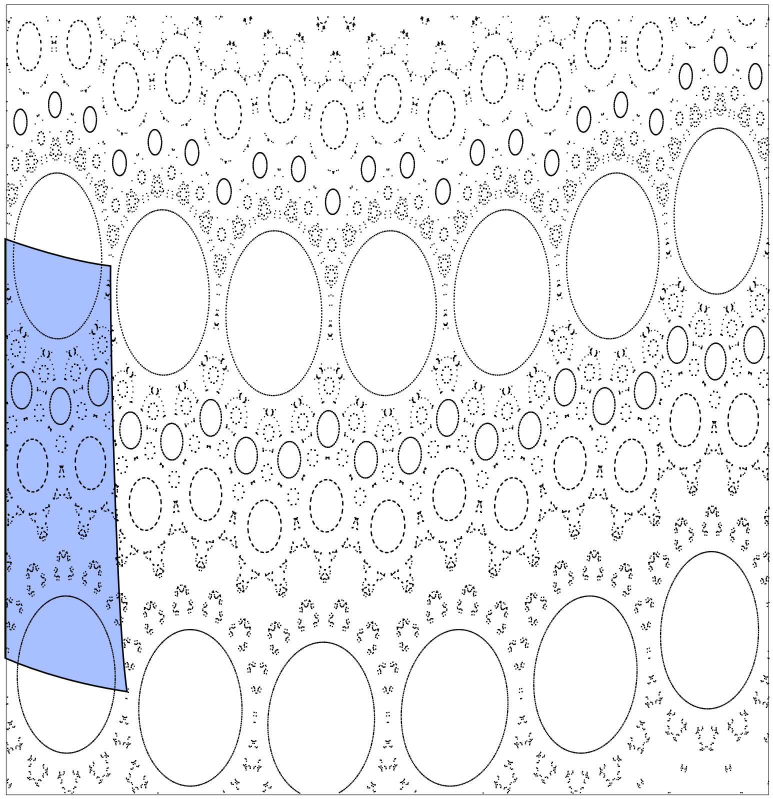

FIGURE 8. Large velocity phase portrait of piecewise smooth pingpong looks similar for different values of time so it makes sense to consider the first return map to a neighbourhood of the singularity.

We assume that the cylinder is divided into a finite union of sectors S_j so that our map is C^∞ in $\text{Int}(S_j)$, has C^∞ extenstion to a neighbourhood of S_j, and satisfies the asymptotics (2.1) in each sector. We suppose that the boundaries of S_j are γ_j and γ_{j+1} where

$$\gamma_j = \left\{ \theta = \theta_{j0} + \frac{\theta_{j1}}{R} + \frac{\theta_{j2}}{R^2} + \dots \right\}.$$

By Lemma 2.1 we can introduce in each sector action-angle coordinates (I, ϕ) so that the boundaries of the sector become

$$\{\phi = 0\} \text{ and } \left\{ \phi = \alpha_0 + \frac{\alpha_1}{R} + \frac{\alpha_2}{R^2} + \dots \right\}$$

and the map takes form

$$I_{n+1} = I_n + \mathcal{O}\left(I_n^{-k}\right), \quad \phi_{n+1} = \phi_n + \frac{1}{I_n}\left[\sum_{m=0}^{k} \frac{c_m}{I_n^m} + \mathcal{O}\left(I_n^{-k}\right)\right]$$

(we suppress the dependence of αs and cs on j since we will work with a fixed sector for a while).

Let Π_j be the fundamental domain bounded by γ_j and $f\gamma_j$ and let F_j be the Poincare map $F_j : \Pi_j \to \Pi_{j+1}$.

It is convenient to introduce coordinates (I, ψ) in Π_j where

$$\phi = \left(\frac{c_0}{I} + \frac{c_1}{I^2} + \dots \frac{c_k}{I^{k+1}}\right)\psi$$

so that ψ changes between 0 and $1+\mathcal{O}(I^{-(k+1)})$. We first describe F_j in the action-angle variables of S_j and then pass to the new action-angle variables of S_{j+1}. We have

$$\phi_n - \phi_0 = \frac{c_0 n}{I} + \frac{c_1 n}{I^2} + \dots$$

The leading term here is the first one so that for the first n such that $\phi_n \in S_{j+1}$ we have $\frac{c_0 n}{I} \approx \alpha_0$ and hence $\frac{c_1 n}{I^2} \approx \frac{c_1 \alpha_0}{c_0 I}$. Therefore

$$\phi_{n+1} = \frac{c_0\psi_0}{I} + \frac{c_0 n}{I} + \frac{c_1\alpha_0}{c_0 I} + \dots$$

Now the condition

$$\phi_{n-1} \leq \alpha_0 + \frac{\alpha_1}{I} \leq \phi_n$$

reduces to

$$\alpha_0 + \frac{\tilde{\alpha}_1}{I} - \frac{c_0\psi_0}{I} + \dots \leq \frac{c_0 n}{I} \leq \alpha_0 + \frac{\tilde{\alpha}_1}{I} - \frac{c_0\psi_0}{I} + \frac{c_0}{I} + \dots$$

where $\tilde{\alpha}_1 = \alpha_1 - \frac{c_1\alpha_0}{c_0}$. For typical ψ_0 this means that

$$n = \left[\frac{\alpha_0 I + \tilde{\alpha}_1}{c_1} - \psi_0\right] + 1 = \frac{\alpha_0 I + \tilde{\alpha}_1}{c_1} - \psi_0 + 1 - \left\{\frac{\alpha_0 I + \tilde{\alpha}_1}{c_1} - \psi_0\right\}.$$

Then

$$\phi_n = \alpha_0 + \frac{\alpha_1}{I} + c_0\left(1 - \left\{\frac{\alpha_0 I + \tilde{\alpha}_1}{c_0} - \psi_0\right\}\right) = \alpha_0 + \frac{\alpha_1}{I} + c_0\left\{\psi_0 - \frac{\alpha_0 I + \tilde{\alpha}_1}{c_0}\right\}.$$

Rescaling the angle variable so that it measures the distance from the singularity $\bar{\psi} = \frac{I}{c_0}(\psi_n - \alpha_0 - \frac{\alpha_1}{I} + \dots)$ we get that F_j has form

$$\bar{I} = I + \dots, \quad \bar{\psi} = \left\{\psi_0 - \frac{\alpha_0 I + \tilde{\alpha}_1}{c_1}\right\} + \dots$$

To pass to action coordinate of S_{j+1} we note that

$$I^{(j)} = \Gamma^{(j)}(\theta)R + \dots, \quad I^{(j+1)} = \Gamma^{(j+1)}R + \dots$$

which implies that that the new addiabatic invariant satisfies

$$J = (1 + \tilde{\lambda}\phi + \dots).$$

Thus in terms of the new action-angle coordinates F_j takes the form

$$\hat{J} = I + \lambda\bar{\psi} + \dots, \quad \hat{\psi} = \bar{\psi}$$

(to justify the last equation we note that if we just use the Taylor expansion we would get $\hat{\psi} = \sigma\bar{\psi}$ and then we get $\sigma = 1$ from the condition that F_j is one-to-one). In terms of the original values of (I, ψ) in Π_j we get

$$\hat{\psi} = \{\psi - \beta_0^{(j)} I - \beta_1^{(j)}\}, \quad \hat{J} = I + \lambda^{(j)}\hat{\psi}.$$

Note that to find the leading term we used the first order Taylor expansion, To compute $\frac{1}{I}$-term we need to use the second order expansion, for $\frac{1}{I^2}$ we need the third order expansion and so on. hence we actually have

Lemma 2.3. *If the orbot does not pass in $\mathcal{O}(1/I^2)$ neighbourhood of the singularities then F_j has the following form*

$$\begin{pmatrix} \psi_{j+1} \\ I_{j+1} \end{pmatrix} = \begin{pmatrix} \{\psi_j - (\beta_0^{(j)} I_j + \beta_1^{(j)})\} \\ I_j + \lambda^{(j)}\psi_{j+1} \end{pmatrix} + \frac{1}{[I_j]}\mathcal{R}_2 + \frac{1}{[I_j]^2}\mathcal{R}_3 + \dots$$

where $\mathcal{R}_j$ are piecewise continuous and on each contintity domain they are polinomials in $(\{I_j\}, \psi_j)$ of degree j.

We shall say that a map F is of class $\mathcal{A}$ if for each k

$$F\begin{pmatrix} \psi \\ I \end{pmatrix} = \begin{pmatrix} \psi \\ I \end{pmatrix} + L_1\begin{pmatrix} \{\psi\} \\ \{I\} \end{pmatrix} + \sum_{j=1}^{k} \frac{1}{n^j}\mathcal{P}_{j+1}(\{\psi\}, \{I\}) + \mathcal{O}(n^{-(k+1)})$$

where L_1 is linear, $A = dL_1$ is constant and $\mathcal{P}_j$ are piecewise polynomials of degree j.

Lemma 2.4. *A composition of $\mathcal{A}$ maps is a $\mathcal{A}$ map.*

Proof. We need to show that if

$$F_s(z) = L_{1,s}(z) + \sum_{j=2}^{k} \frac{1}{n^{j-1}}\mathcal{P}_{j,s}(z) \text{ for } s = 1, 2$$

where $\mathcal{P}_{j,s}$ are polynomials of degree j then $F_2 \circ F_1$ is also of the same form. It is sufficient to consider the case where $\mathcal{P}_{j,s}$ have positive coefficients since in the sign changing case there might be additional cancelations. Observe that $F_s(z) = \sum_{j=1}^{k} \frac{1}{n^{j-1}}\mathcal{P}_j$ where $\mathcal{P}_j$ are some polynomials then the degree restriction amounts to saying that $G_n(u) = \frac{1}{n}F_s(un)$ is bounded for each u as $n \to \infty$. But if $G_{n,1}$ and $G_{n,2}$ satisfy this condition then the same holds also for their composition. $\square$

Corollary 2.5. *The first return map $\mathcal{F} : \Pi_1 \to \Pi_1$ is an $\mathcal{A}$ map and the same holds for any power $\mathcal{F}^m$.*

Remark 2.6. Corollary 2.5 applies in particular in the case where the original map is smooth. In that case the coefficients $\lambda^{(j)}$ vanish so the linear part is the intgrable twist map

$$\hat{I} = I, \quad \hat{\psi} = \psi - \beta_0 I - \beta_1. \tag{2.2}$$

More generally, $\lambda^{(j)}$ depend only on the behaviour of the function Γ near the singularities so the normal form (2.2) holds also in the case where a_0 and b_0 from Lemma 2.1 are continuous (even though the higher order terms may be nontrivial in that case).

We say that the original map f is *hyperbolic at infinity* if the linear part L_1 of the normal form of the first return map $\mathcal{F}$ is hyperbolic and say that f is *elliptic at infinity* if L_1 is elliptic. Recall that the ellipticity condition is $|\mathrm{Tr}(L_1)| < 2$ and the hyperbolicity condition is $|\mathrm{Tr}(L_1)| > 2$.

One can work out several leading terms in our main examples. Namely for outer billiard about the semicircle it is shown in [10] that $L_1 = \mathbf{L}^2$ where

$$\mathbf{L}(I, \psi) = (I - \frac{4}{3} + \frac{8}{3}\{\psi - I\}, \{\psi - I\}). \tag{2.3}$$

For Fermi-Ulam pingpongs where the wall velocity has one discontinuity at 0 one has [7]

$$L_1(I, \psi) = (I + \Delta\left(\{\psi - I\} - \frac{1}{2}\right), \{\psi - I\}) \text{ where} \tag{2.4}$$

$$\Delta = l(0)\Delta\dot{l}(0)\int_0^1 \frac{ds}{l^2(s)}$$

and $l(s)$ is the distance between the walls at time s.

For example, for motions studied by Ulam and Wells one has $l(s) = b + a(\{s\} - 1/2)^2$. We can choose the units of length so that $b = 1$, then $l(s) > 0$ for all s provided that $a > -4$. Then $\Delta(a) = -2a(1 + a/4)J(a)$ where

$$J(a) = \int_0^1 \frac{ds}{(1 + a(s - 1/2)^2)^2} = \frac{2}{a + 4} + \begin{cases} \frac{1}{2\sqrt{|a|}} \ln \frac{2+\sqrt{|a|}}{2-\sqrt{|a|}} & \text{if } a < 0 \\ \frac{1}{\sqrt{a}} \arctan\left(\frac{\sqrt{a}}{2}\right) & \text{if } a > 0. \end{cases}$$

One can check that f is hyperbolic at infinity if $a \in (-4, a_c)$ or $a > 0$ and f is elliptic at infinity for $a \in (a_c, 0)$ where $a_c \approx -2.77927\dots$

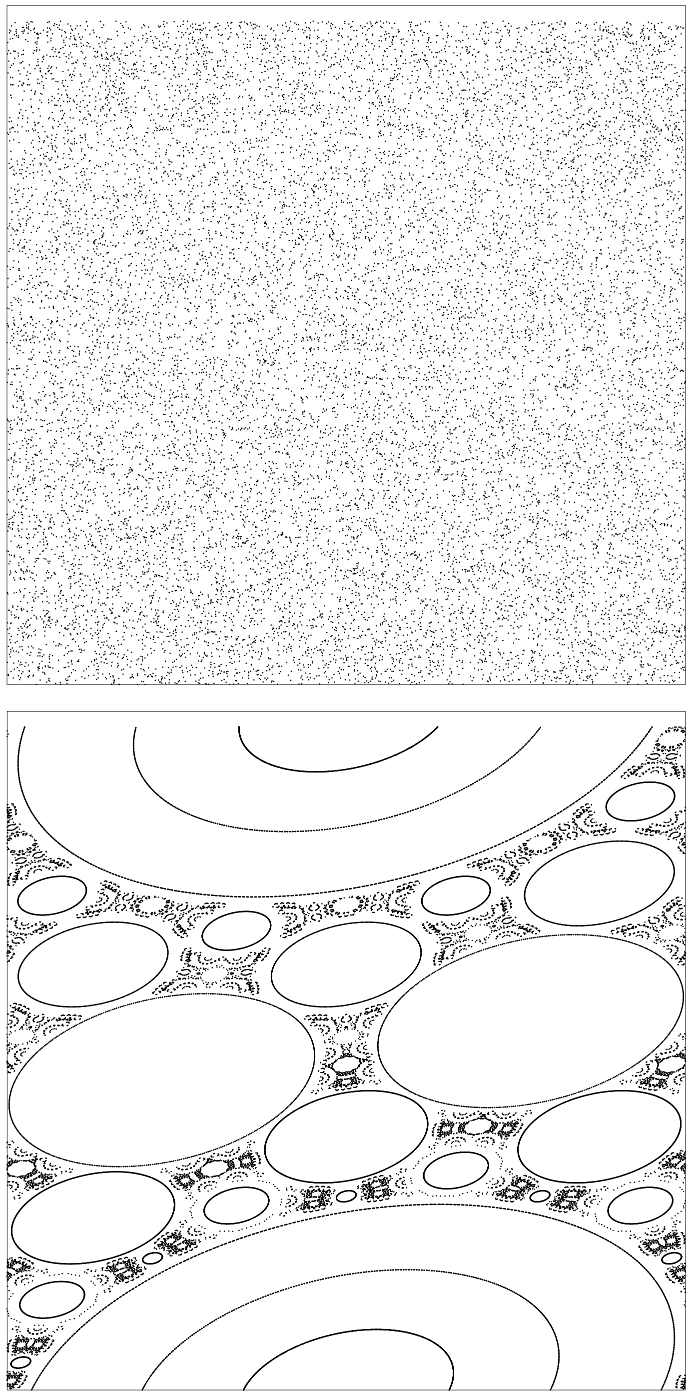

FIGURE 9. Dynamics of the first return map. Top: hyperbolic case. Bottom: elliptic case.

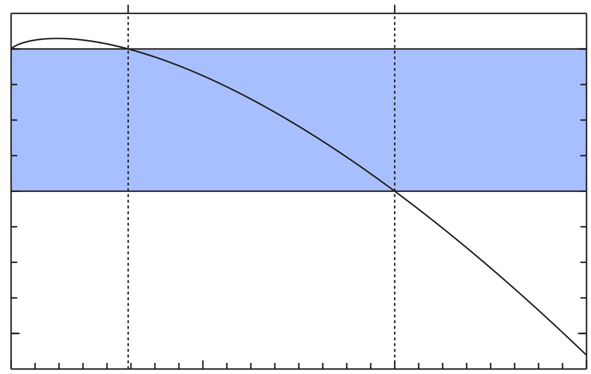

FIGURE 10. $\Delta(a)$ for piecewise linear wall velocity

2.4. **Accelerating orbits for piecewise smooth maps.** Given an $\mathcal{A}$ map f we say that $p = (\bar{I}, \bar{\psi})$ is *an accelerating orbit* if there exist $m, l > 0$ such that $L_1^m(p) = p + (l, 0)$.

Lemma 2.7. [10] *Assume that f is elliptic at infinity and has an (m, l) accelerating orbit such that the spectrum of L_1^m does not contain k-th roots of unity for $k \in \{1, 2, 3\}$. Suppose also that $\mathcal{F}$ preserves a smooth measure with density of the form $\rho(I, \psi) = I\rho_0(\psi) + \rho_1(\psi) + o(1)$. Then f has positive (and hence infinite) measure of orbits such that $I_n \to \infty$.*

Proof. Consider a point $\{I_N, \psi_N\}$ in a small neighborhood of $\{\bar{I}+Nl, \bar{\phi}\}$ and study its dynamics. For $n \geq N$, we will denote $\{I_n, \psi_n\}$ the point $\mathcal{F}^{(n-N)l}(I_N, \psi_N)$. Set $U_n = I_n - (\bar{I}+nL)$, $\upsilon_n = \psi_n - \bar{\psi}$. We can introduce a suitable complex coordinate $z_n = U_n + i(aU_n + b\upsilon_n)$ such that $D\mathcal{F}^l$ becomes a rotation by angle $2\pi s$ near the origin where $s \notin \frac{1}{k}\mathbb{Z}$ for $k \in \{1, 2, 3\}$. In these coordinates $\mathcal{F}^l$ takes the following form in a small neighborhood of $(0, 0)$

$$z_{n+1} = e^{i2\pi s} z_n + \frac{A(z_n)}{N} + \mathcal{O}(N^{-2}) \tag{2.5}$$

where

$$A(z) = w_1 + w_2 z + w_3 \bar{z} + w_4 z^2 + w_5 z\bar{z} + w_6 \bar{z}^2.$$

Lemma 2.8. *(a) We have that $\mathcal{R}e(e^{-i2\pi s} w_2) = 0$.*

(b) There exists $\epsilon > 0$ and a constant C such that if $|z_N| \leq \epsilon$, then for every $n \in [N, N + \sqrt{N}]$

$$|z_n| \leq |z_N| + CN^{-1}.$$

Part (b) is the main result of the lemma. Part (a) is an auxiliary statement needed in the proof of (b). Namely, part (a) says that a certain resonant coefficient vanishes (this vanishing is due to the fact that f preserves a measure with smooth density).

Before we prove this lemma, let us observe that it implies that for sufficiently large N, all the points $|z_N| \le \epsilon/2$ are escaping orbits. Indeed by $[\sqrt{N}]$ applications of lemma 2.8 there is a constant C such that

$$|z_l| \le \frac{\epsilon}{2} + CN^{-\frac{1}{2}}$$

for every $l \in [N, 2N]$. It now follows by induction on k that if $l \in [2^k N, 2^{k+1}N]$ then

$$|z_l| \le \epsilon_k$$

where

$$\epsilon_k = \frac{\epsilon}{2} + \frac{C}{\sqrt{N}} \sum_{j=0}^{k} \left(\frac{1}{\sqrt{2}}\right)^j$$

(N has to be chosen large so that $\epsilon_k \le \epsilon$ for all k). This proves lemma 2.7. □

Proof of lemma 2.8. Let $\bar{n} = n - N$. For $\bar{n} \le \sqrt{N}$ equation (2.5) gives

$$z_n = e^{i2\pi\bar{n}s} z_N + \frac{1}{N} \sum_{m=0}^{\bar{n}-1} e^{i2\pi ms} A(e^{i2\pi(\bar{n}-m-1)s} z_{N+\bar{n}-m}) + \mathcal{O}(N^{-\frac{3}{2}}) \tag{2.6}$$

In particular for these values of n we have

$$z_n = e^{i2\pi s(n-N)} z_N + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right).$$

Substituting this into (2.6) gives

$$z_n = e^{i2\pi\bar{n}s} z_N + \frac{1}{N} \sum_{m=0}^{\bar{n}-1} e^{i2\pi ms} A(e^{i2\pi(\bar{n}-m-1)s} z_N) + \mathcal{O}\left(\frac{1}{N}\right).$$

To compute the sum above expand A as a sum of monomials and observe that

$$\sum_{m=0}^{\bar{n}-1} e^{i2\pi ms} \left(e^{i2\pi(\bar{n}-m-1)s} z_N\right)^{\alpha} \left(e^{-i2\pi(\bar{n}-m-1)s} \bar{z}_N\right)^{\beta}$$

is bounded for $\alpha + \beta \le 2$ unless $\alpha = \beta + 1$ (that is $\alpha = 1$, $\beta = 0$). Therefore

$$z_n = e^{i2\pi\bar{n}s} z_N \left(1 + \tilde{w}_2 \frac{\bar{n}}{N}\right) + \mathcal{O}\left(N^{-1}\right) \tag{2.7}$$

where $\tilde{w}_2 = e^{-i2\pi s} w_2$.

Consider now the disc D_N around 0 of radius $N^{-0.4}$. Let $W(z)$ denote the density of invariant measure in our complex coordinates. Then by (2.7)

$$\frac{\text{Area}(\mathcal{F}^{\bar{n}} D_N)}{\text{Area}(D_N)} = \left(1 + 2\mathcal{R}e(\tilde{w}_2)\frac{\bar{n}}{N}\right) + \mathcal{O}\left(N^{-0.6}\right).$$

On the other hand there exists $z \in D_N$ such that denoting $z' = \mathcal{F}^{\bar{n}} z$ we have

$$\frac{\text{Area}(\mathcal{F}^{\bar{n}} D_N)}{\text{Area}(D_N)} = \frac{1 + W(z)/N}{1 + W(z')/(\bar{n} + N)} + \mathcal{O}\left(N^{-2}\right) = 1 + \mathcal{O}\left(N^{-1.4}\right)$$

since $W(z) - W(z') = \mathcal{O}\left(N^{-0.4}\right)$. Comparing those two expressions for the ratio of areas we obtain that $\mathcal{R}e(\tilde{w}_2) = 0$.

This proves part (a) of Lemma 2.8. Part (b) now follows from (2.7).
□

Corollary 2.9. $\text{mes}(\mathcal{E}) = \infty$ *for the following systems:*
(a) outer billiards about circular caps with angle close to π;
(b) Ulam pingpongs with $\Delta \in (2, 4)$.

Proof. For part (a) observe that map (2.3) has accelerating orbit $(0, \frac{7}{8})$ and for part (b) observe that map (2.4) has accelerating orbit $(0, \frac{1}{2} + \frac{1}{\Delta})$.
□

Problem 2.10. Does map (2.4) have stable accelerated orbits for all $\Delta \in (0, 4)$?

3. Applications of KAM theory.

3.1. **Absence of acceleration in strong potential.**

3.2. **Accelerating tennis ball.**

3.3. **Bounded orbits for piecewise smooth Ulam pingpong.**

4. Recurrence.

4.1. **Applications of Poincare Recurrence Theorem.** In this section we describe applications of ergodic theory to the dynamics of bouncing balls. One of the basic results in ergodic theory is Poincare Recurrence Theorem. It says that if a transformation T of a space X preserves a finite measure μ then for each sets A alsmost all points from A returns to A in the future. To see why this theorem is true let

$$B = \{x \in A : T^n x \notin A \forall n > 0.\}$$

Then $T^n B \cap B = \emptyset$ and so $T^k B \cap T^{k+n} B = T^k(B \cap T^n B) = \emptyset$. Thus for each N the sets $B, TB \dots T^{N-1}B$ are disjoint and therefore $\mu(\cup_{n=0}^{N-1} B) = N\mu(B) \leq \mu(X)$. Since N is arbitrary we have $\mu(B) = 0$.

Poincare Recurrence Theorem need not hold for infinite measure preserving transformations such as $x \to x+1$ on $\mathbb{R}$. However in the infinite measure case there exists a decomposition $X = \mathcal{C} \cup \mathcal{D}$ where $\mathcal{D} = \cup_{n \in \mathbb{Z}} T^n B$ and B is wondering in the sense that $T^n B \cap B = \emptyset$ for $n \neq 0$ while $\mathcal{C}$ satisfies the Poincare Recurrence Theorem in the sense that for any set $A \subset \mathcal{C}$ almost all points from A visit A. In abstract ergodic theory $\mathcal{C}$ is called *conservative* part of X and D is called *dissipative* part of X. However in the setting of smooth dynamical systems this terminology is misleading since $\mathcal{D}$ need not be dissipative in the sense that $\mathrm{Jac}(f) < 1$ as the above example of the shift on $\mathbb{R}$ shows. Therefore we adopt the terminology of probability theory. That is, we call $\mathcal{C}$ recurrent part of X and $\mathcal{D}$ transient part of X. If $\mathcal{C} = X$ we say that the system is recurrent, if $\mathcal{D} = X$ we say that the system is transient. In the setting of bouncing balls the system has nontrivial transient component if the set

$$\mathcal{E} = \{(t_0, v_0) : v_n \to \infty\}$$

has positive measure. More generally we have the following.

Lemma 4.1. *Let $T : X \to X$ preserve an infinite measure μ. Suppose that there is a set A such that $\mu(A) < \infty$ and an invariant set B such that all points from B visit A. Then $B \subset \mathcal{C}$. In particular if almost all points from X visit A then T is recurrent.*

Proof. Let $S \subset B$. For $x \in B$ let $r(x) = \min(k \geq 0 : T^{-k}x \in A)$ so that $T^{r(x)}x \in A$. Let $\hat{S}_k = \cup_{x \in S : r(x) \leq k} T^{r(x)}x$. It is sufficient to show that almost all points from $\hat{S}_k$ visit $\hat{S}_k$ infinitely often since if $T^n x \in \hat{S}_k$ then $T^{n-j}x \in S$ for some $j \leq k$. Note that $\hat{S}_k \subset A \cap B$. By assumption almost all points in $T(A \cap B)$ visit A and so the first return map $R : \hat{S}_k \to \hat{S}_k$ is well defined. Applying Poincare Recurrence Theorem to $(\hat{S}_k, R)$ we obtain our claim. □

Lemma 4.1 implies that $\mathcal{E}$ is indeed the transient part of the phase space since the compliment of $\mathcal{E}$ is $\cup_N Z_N$ where

$$Z_N = \{(t_0, v_0) : \liminf v_n \leq N\}$$

and all points from Z_N visit $\{v \leq N+1\}$.

While the proof of Lemma 4.1 is very easy there is no general recipy for finding the set A and sometimes it can be tricky. In this section though we present a few examples there the construction of A is relatively simple.

Corollary 4.2. $\mathrm{mes}(\mathcal{E}) = 0$ *for the following systems*

(a) Fermi-Ulam pingpongs there l and $\dot{l}$ are continuous and $\ddot{l}$ has finitely many jumps;

(b) outer billiards around lenses.

Proof. In both cases the return map $\mathcal{F} : \Pi_1 \to \Pi_1$ has the following form

$$(I, \psi) \to (I, \{\psi - a_0 I - a_1\}) + \mathcal{O}(1/I)$$

(see remark 2.6). That is, after one rotation the adiabatic invariant changes by $\mathcal{O}(1/I)$. Therefore each unbounded orbit visits the set

$$A = \cup_k \left\{ |I - 3^k| < \frac{1}{2^k} \right\}.$$

Since $\mu(A) < \infty$ the statement follows from Lemma 4.1. □

Problem 4.3. Do above systems have escaping orbits? In fact even the existence of unbounded orbits is unknown.

4.2. **Ergodicity and recurrence.** To proceed further we need to recall some facts from ergodic theory. Let $T : X \to X$ be a map preserving a measure μ. T is called *ergodic* if for any T invariant set we have $\mu(A) = 0$ or $\mu(A^c) = 0$. Next suppose that μ is a probability measure. Then Pointwise Ergodic Theorem says that for every $\Phi \in L^1(\mu)$ the following limits exist and are equal almost surely

$$\Phi^+(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Phi(T^j x) = \Phi^-(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \Phi(T^{-j} x).$$

If T is ergodic then $\Phi^+(x) = \Phi^-(x) = \mu(\Phi)$ almost surely.

We now consider skew product maps $T_\Phi : (X \times \mathbb{R}) \to (X \times \mathbb{R})$ given by $T_\Phi(x, y) = (Tx, y + \Phi(x))$ preserving measure $d\nu = d\mu dx$.

Lemma 4.4. *(Atkinson, [1]) Suppose that T is ergodic. If $\Phi \in L^1(\mu)$ then T_Φ is recurrent if $\mu(\Phi) = 0$ and transient if $\mu(\Phi) \neq 0$.*

Proof. Suppose that $\mu(\Phi) \neq 0$. If $\mathcal{C}$ was nontrivial there would exist R such that $\nu(\mathcal{C}_R) > 0$ where $\mathcal{C}_R = \mathcal{C} \cap \{|y| \leq R\}$. Then almost all points from $\mathcal{C}_R$ would return to $\mathcal{C}_R$ infinitely often. However by Pointwise Ergodic Theorem $y_n \to \infty$ giving a contradiction.

Our next remark is that T_Φ commutes with translations. Hence if $(x, y) \in \mathcal{C}$ then for each $\tilde{y}$ $(x, \tilde{y}) = \tau_{\tilde{y}-y}(x, y) \in \mathcal{C}$. Therefore $\mathcal{C}$ and $\mathcal{D}$ are of the form

$$\mathcal{C} = \tilde{C} \times \mathbb{R} \text{ and } \mathcal{D} = \tilde{D} \times \mathbb{R}$$

where $\tilde{C}$ and $\tilde{D}$ are T-invariant. Thus either $\tilde{C}$ or $\tilde{D}$ has measure 0.

We now consider the case $\mu(\Phi) = 0$. Assume that $\tilde{C} = \emptyset$ so that $\mathcal{D} = X \times \mathbb{R}$. We shall show that this assumption will lead to a contradiction. We have that almost all (x, y) with $|y| \le 1$ visit $\{|y| \le 2\}$ only finitely many times since

$$B = \{(x_0, y_0) : |y_n| \le 2 \text{ infinitely often}\}$$

is T_Φ invariant and all points from B visit $A = \{|y| \le 2|\}$ so if $\mu(B) > 0$ T_Φ would have a nontrivial recurrent part by Lemma 4.1.

Hence for almost all x the set $M_x = \{n : |\Phi_n| \le 1\}$ is finite where $\Phi_n(x) = \sum_{j=0}^{n-1} \Phi(T^j x)$. Let $A_N = \{x : \mathrm{Card}(M_x) \le N\}$. Pick N such that $\mu(A_N) > 1/2$. Take $n \gg N$. Consider

$$\mathcal{Y}_n(x) = \{y : \exists j \in [0, n-1] : T^j x \in A_N \text{ and } \Phi_j(x) = y\}.$$

By ergodic theorem applied to the indicator of A_N for large n we have $\mathrm{Card}(\mathcal{Y}_n(x)) \ge \frac{n}{2}$ and for each $\bar{y} \in \mathcal{Y}_n(x)$ we have

$$\mathrm{Card}\left\{y \in \mathcal{Y}_n : |y - \bar{y}| < \frac{1}{2}\right\} \le (N+1)$$

since otherwise taking a point from this set with minimal j will lead to a contradiction with the definition of A_N. It follows that

$$\max_{j \le n} |\Phi_j(x)| \ge \max_{j \le n, T^j x \in A_N} |\Phi_j(x)| \ge \frac{n}{8(N+1)}.$$

On the other hand by ergodic theorem $\frac{\Phi_j(x)}{j} \to 0$ as $j \to \infty$ and hence $\frac{\max_{j \le n} |\Phi_j(x)|}{n} \to 0$ as $n \to \infty$ contradicting the last displayed inequality. $\square$

As an application of Lemma 4.4 consider SWA to an impact oscillator with

$$\dot{f}(t) = \begin{cases} 1 & \text{if } \{t\} \le \frac{1}{2} \\ -1 & \text{if } \{t\} > \frac{1}{2} \end{cases}.$$

Choose $\bar{h} = 0$. Then $f(v, t) = (\bar{t}, v + \dot{f}(\bar{t})$ where $\bar{t} = t + \frac{T}{2}$ and T is the period of the spring. Therefore f is recurrent.

On the other hand if $\bar{h} \neq 0$ then Lemma 4.4 is not directly applicable since $\bar{t} = t + \frac{T}{2} + \frac{2\bar{h}}{v} + o(1/v)$ weakly depends on v. To include this case we need another lemma. Let $S(x, y) = (\mathcal{T}(x, y), y + \phi(x, y))$ be the map which is well approximated by a skew product at infinity. Namely let $\tau_m(x, y) = (x, y + m)$. We assume that S is defined on asubset $\Omega \subset X \times \mathbb{R}$ given by $y \ge h(x)$. We also assume that there exist a map $T : X \to X$ and a function $\Phi : X \to \mathbb{R}$ such that T preserves measure

μ and that for each k and each function bounded measurable function h supported on $X \times [-M, M]$ we have

$$||h \circ S_m^k - h \circ T_\Phi^k||_{L^1(\nu)} \to 0 \text{ as } m \to \infty$$

where $S_m = \tau_{-m} \circ S \circ \tau_m$ and $d\nu = d\mu dx$.

Lemma 4.5. *Assume that*
(i) T is ergodic;
(ii) $\mu(\Phi) = 0$;
(iii) S preserves a measure $\tilde{\nu}$ having bounded density with respect to ν;
(iv) there exists a number K such that $\phi||_{L^\infty(\mu)} \leq K$.
Then S is recurrent.

In the proof we will need Rokhlin's Lemma which says that if $T : X \to X$ is an aperiodic transformation preserving a finite measure μ then for each n, ε there is a set B such that $B, TB, \dots, T^{n-1}B$ are disjoint and $\mu(X - \cup_{j=0}^{n-1} T^j B) \leq \varepsilon$.

Proof. Let $\bar{Y} = X \times [0, K]$ where K is the constant from condition (iv). By Lemma 4.4 T_Φ is conservative and hence the first return map $R : \bar{Y} \to \bar{Y}$ is defined almost everywhere. By Rokhlin Lemma applied to R there exists a set Ω_ε and a number L_ε such that $\nu(\Omega_\varepsilon) < \varepsilon$ and

$$\nu(\{(x, y) \in \bar{Y} : T_\Phi^j(x, y) \notin \Omega_\varepsilon \text{ for } j = 0, 1 \dots L_\varepsilon - 1\}) < \varepsilon.$$

It follows that there exists $m_\varepsilon > 1/\varepsilon$ such that $\nu(A_\varepsilon) < \varepsilon$ where

$$A_\varepsilon = \{(x, y) \in \tau_{m_\varepsilon} \bar{Y} : S^j(x, y) \notin \tau_{m_\varepsilon} \Omega_\varepsilon \text{ for } j = 0, 1 \dots L_\varepsilon - 1\}.$$

In addition we have $\tilde{\nu}(A_\varepsilon) < C\varepsilon$ and $\tilde{\nu}(\tau_{m_\varepsilon} \Omega_\varepsilon) < C\varepsilon$. Let

$$A = \bigcup_n \left(\tau_{m_{1/n^2}} \Omega_{1/n^2} \cup A_{1/n^2} \right).$$

Then $\nu(A) < \infty$. Note that every unbounded orbit crosses $\tau_{m_{1/n^2}} \Omega_{1/n^2}$ for a sufficiently large n and so it visits A. Therefore S is recurrent by Lemma 4.1. $\square$

Lemma 4.5 shows recurrence of imapct oscillator SWA for all $\bar{h}$. It also implies recurrence of Fermi-Ulam pingpongs in the case where $\dot{l}$ has one discontinuity and the corresponding map is hyperbolic at infinity. This follows from the normal form at infinity derived in Section 2 and the ergodicity of hyperbolic sawtooth map proved in Section 5.

5. Ergodicity of hyperbolic sawtooth maps.

5.1. **The statement.** In this section we will prove the following theorem of Chernov. Let T be a piecewise linear automorphism of $\mathbb{T}^2$. Let S_+ and S_- denote thediscontinuity lines of T and T^{-1} respectively. Denote $S_n = T^{n-1}S_+$, $S_{-n} = T^{-(n-1)}S_-$. We assume that

(i) $A = dT$ is constant hyperbolic $SL_2(\mathbb{R})$-matrix.
(ii) $S_\pm$ are not parallel to eigendirections of A.

Theorem 5.1. [5] *T is ergodic.*

5.2. **The Hopf argument.** The proof relies on the Hopf argument. To explain this argument we consider first the case where T is smooth, that is $fx = Ax \bmod 1$ and $A \in SL_2(\mathbb{Z})$. Denote

$$W^s(x) = \{y : d(T^n x, T^n y \to 0 \text{ as } n \to +\infty\},$$

$$W^u(x) = \{y : d(T^{-n} x, T^{-n} y \to 0 \text{ as } n \to +\infty\}.$$

It is easy to see that $W^*(x) = \{x + \xi e_*\}_{\xi \in \mathbb{R}}$ where e_s and e_u are contracting and expanding eigenvectors of A.

Let $\mathcal{R}_0$ be the set of regular points, that is, the points such that for any continuous function Φ we have $\Phi^+(x) = \Phi^-(x)$. By Pointwise Ergodic Theorem $\mathcal{R}_0$ has full measure in $\mathbb{T}^2$. For $j > 1$ we can define inductively

$$\mathcal{R}_j = \{x \in \mathcal{R}_{j-1} : \text{mes}(y \in W^u(x) : y \notin \mathcal{R}_{j-1}) = 0 \text{ and } \text{mes}(y \in W^s(x) : y \notin \mathcal{R}_{j-1}) = 0\}.$$

Then we can show by induction using Fubini Theorem that $\mathcal{R}_j$ has full measure in $\mathbb{T}^2$ for all j.

For $x \in cR_0$ and $\Phi \in C(\mathbb{T}^2)$ let $\bar{\Phi}(x)$ denote the common value of $\Phi^+(x)$ and $\Phi^-(x)$. We say that $x \sim y$ if for all continuous Φ we have $\bar{\Phi}(x) = \bar{\Phi}(y)$. Note that if $x, y \in \mathcal{R}_0$ and $y \in W^s(x)$ then for all $\Phi \in C(\mathbb{T}^2)$ we have $\Phi^-(x) = \Phi^-(y)$ and so $x \sim y$. Similarly if $x, y \in \mathcal{R}_0$ and $y \in W^u(x)$ then $x \sim y$,. Given $x \in \mathcal{R}_2$ and $\rho \in \mathbb{R}_+$ let

$$\Gamma_\rho = \bigcup_{y \in W^u_\rho(x)} W^s(y), \quad \tilde{\Gamma}_\rho = \bigcup_{y \in \mathcal{R}_1 \cap W^u_\rho(x)} (W^s(y) \bigcap \mathcal{R}_0).$$

Then if ρ is large enough then $\Gamma_\rho = \mathbb{T}^2$ and by Fubini theorem $\text{mes}(\Gamma_\rho - \Gamma_\rho) = 0$ so $\bar{\Phi}(z) = \bar{\Phi}(x)$ for almost all z. Therefore $\bar{\Phi}$ is constant almost surely and hence T is ergodic.

5.3. **Long invariant manifolds and ergodicity.** The Hopf argument has been expanded in several directions. Already Hopf realized that the same argument works for nonlinear systems provided that the stable and unstable foliations are C^1. This condition however is too restricitve. Versions of the Hopf argument under weaker conditions have been presented by Anosov, Pesin, Pugh-Shub, Burns-Wilkinson. We need a version of the Hopf argument for systems with singularities. The approach to handle such systems is due to Sinai and it has been extended by Chernov-Sinai and Liverani-Wojtkowski. The proof given here follows the presentation of [6] a slightly different argument can be found in [14].

The difficulty in the nonsmooth case is that it is no longer true that $W^*(x)$ coincides with $\tilde{W}^*(x) = \{x + \xi e_*\}$. Indeed if $y \in \tilde{W}^s(x)$ and x and y belong to the same continuity domain then $d(Tx, Ty) = \frac{1}{\lambda} d(x, y)$ where λ is the expanding eigenvalue of A. However if Tx and Ty are separated by a singularity then Tx and Ty can be far apart. In fact, there might be points which come so close to the singularities that $W^s(x)$ is empty. This is however, an exception rather than a rule. Let

$$r_u(x) = \max\{\delta : \tilde{W}^u_\delta(x) \subset W^u(x)\}, \quad r_s(x) = \max\{\delta : \tilde{W}^s_\delta(x) \subset W^s(x)\}.$$

Lemma 5.2.

$$\mathrm{mes}\{x \in \mathbb{T}^2 : r_u(x) \leq \varepsilon\} \leq C\varepsilon, \quad \mathrm{mes}\{x \in \mathbb{T}^2 : r_s(x) \leq \varepsilon\} \leq C\varepsilon.$$

Proof. We prove the second statement, the first one is similar. Note that $\{r_s(x) \leq \varepsilon\} = \bigcup_n \mathbb{S}_n(\varepsilon)$ where

$$\mathbb{S}_n(\varepsilon) = \left\{x : d(T^n x, S_-) \leq \frac{\varepsilon}{\lambda^n}\right\}.$$

Since our system is measure preserving

$$\mathrm{mes}(\mathbb{S}_n) = \mathrm{mes}\left\{x : d(x, S_-) \leq \frac{\varepsilon}{\lambda^n}\right\} \leq \bar{C}\frac{\varepsilon}{\lambda_n}.$$

□

The proof of Theorem 5.1 relies on a local version of this result. Namely, the following statement holds.

Lemma 5.3. *Pick y, δ and k such that $d(T^j \tilde{W}^u(y), S_-) \geq \varepsilon$ for $j = 0 \dots k$. Then*

$$\mathrm{mes}\{x \in \tilde{W}^u(y) : r_u(x) \leq \varepsilon\} \leq C\varepsilon.$$

A similar statement holds with s and u interchanged.

We first show how Lemma 5.3 can be used to derive Theorem 5.1 and then present the proof of the lemma.

Pick k such that $C\theta^k < 0.001$. We first establish local ergodicity. Namely let M be a connected component of continuity for T^k and T^{-k}. We shall show that almost all points in M belong to one equivalence class. This will imply that every invariant function is constant on M, that is, any invariant set is a union of continuity domains. Then we conclude the global ergodicity by noticing that there are no nontrivial invariant sets which are union of continuity components because the boundary would be a collection of line segments and this boundary can not be invariant since the segements in S_n have different slopes for different n.

Let us prove local ergodicity. To simplify the exposition we will refer to $\tilde{W}^u$ leaves as horizontal lines and to $\tilde{W}^s$ leaves as vertical lines. Take a rectangle $U \subset \text{Int}(M)$. It is enough to show that all points are equivalent. Given N consider all squares with sides $\frac{1}{N}$ and centers in $\left(\frac{0.1\mathbb{Z}}{N}\right)^2 \cap U$.

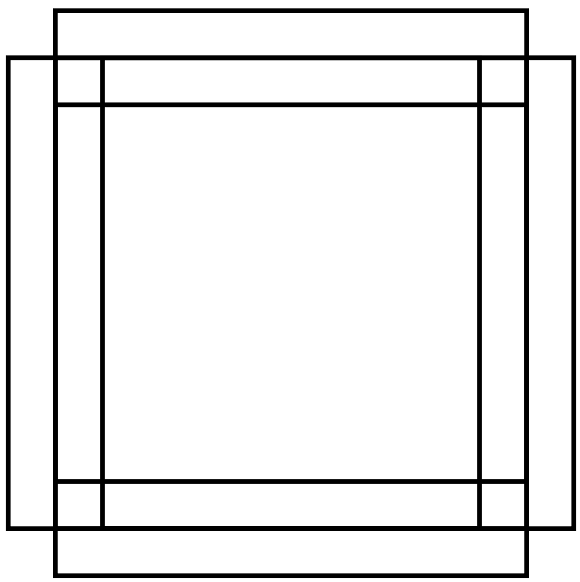

FIGURE 11. Each square intersect its neighbours by 0.9 of their area

We say that a points z in a square S is *typical* if $z \in \mathcal{R}_2$ and both $W^u(x)$ and $W^s(x)$ cross S completely.

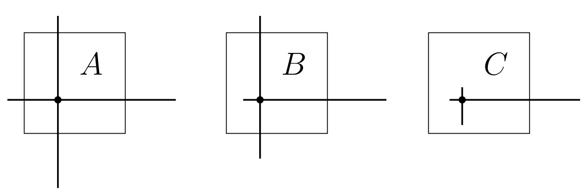

FIGURE 12. A is typical in S, B is not typical in S but it is typical in anearby square, C is not typical in any square

Note that all typical points in S are equivalent. Indeed denote

$$\Sigma(z) = \cup_{x \in W^u(z_j)} W^u(x)$$

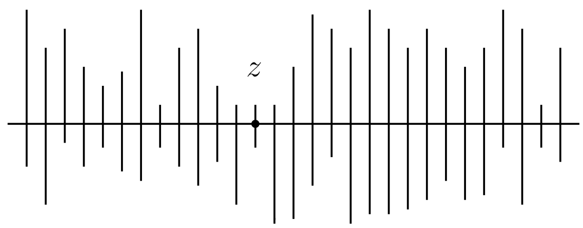

FIGURE 13. Hopf brush $\Sigma(z)$

Note that if $z_1, z_2 \in S$ then by Lemma 5.3 $\Sigma(z_j) \cap S$ has measure at least 0.999mes(S) and by the Hopf argument almost all points in $\Sigma(z_j)$ are equivalent to z_j. Also by Lemma 5.3 the set of typical points in S has measure at least 0.998mes(S). Since for two neighbouring squares we have $\text{mes}(S_1 \cap S_2) = 0.9\text{mes}(S_1)$ it follows that all typical points in neighbouring squares are equaivalent. Therefore all typical points in all squares in Int(M) are equivalent. On the other hand by Lemma 5.2 for almost all $x \in \mathcal{R}_2$ we have $r_u(x) > 0$ and $r_s(x) > 0$ so such x is typical for sufficiently large N. Local ergodicity follows and Theorem 5.1 is proven.

5.4. **Growth Lemma.** It remains to prove Lemma 5.3. To this end fix a curve $\gamma \subset \tilde{W}^u(x)$. Due to singularities $T^n(x)$ consists of many components. Let $r_n(x)$ be the distance from x to the boundary of the component containing x. We claim that there are constants $\theta < 1$ and $\hat{C} > 0$ such that

$$\mathbb{P}(r_n \le \varepsilon) \le 2(\theta^n |\gamma| + \hat{C})\varepsilon. \tag{5.1}$$

(5.1) implies Lemma 5.3 since it implies that

$$\mathbb{P}(\Lambda_n) \le 2(\theta^n |\gamma| + \hat{C}) \frac{\varepsilon}{\lambda^n}.$$

Summing this for $n \ge k$ we obtain the statement of Lemma 5.3.

The proof of (5.1) relies on *complexity bound*. Let $\kappa_n(\delta)$ be the maximal number of continuity components of T^n an unstable curve of length less than δ can be cut into. Set $\kappa_n = \lim_{\delta \to 0} \kappa_n(\delta)$. For the case at hand there is a constant K such that $\kappa_n \le Kn$ since the singularities of T^n are lines and there at most Kn possibilites for their slopes. Accordingly there exist numbers n_0, δ_0 such that $\kappa_{n_0}(\delta_0) \le$

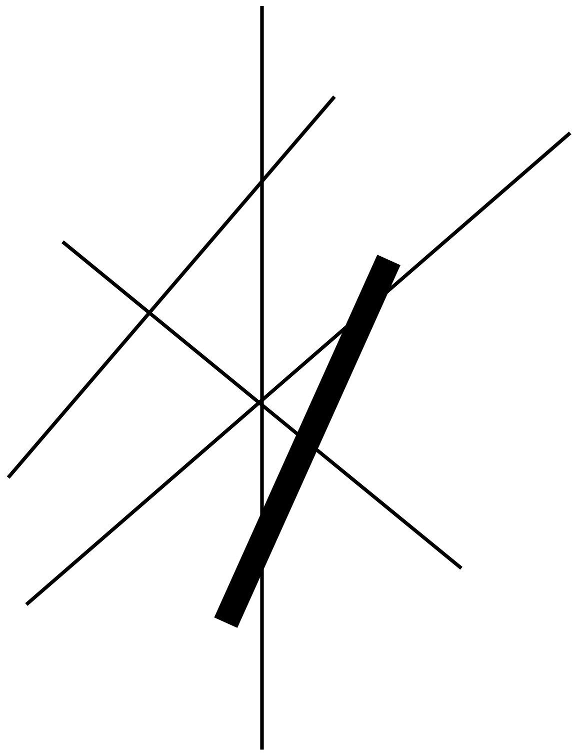

FIGURE 14. The complexity is determined by the largest number of lines passing through one point since one can always take δ so small that any curve of length less than δ can not come close to two intersection points

$2\lambda^{n_0}$. Replacing T by T^{n_0} we can assume that this inequality holds for $n_0 = 1$ (clearly it is sufficient to prove (5.1) for $\bar{T} = T^{n_0}$ in place of T).

Given a curve γ we define $\bar{r}_n(x)$ as follows. $T\gamma$ is cut into several componets. Some of them can be longer than δ_0. Cut each long component into segments of length between $\delta_0/2$ and δ_0. For each of the resulting curves γ_j consider $T\gamma_j$ and repeat this procedure. Let $\bar{r}_n(x)$ be the distance to the boundary of the new components. Thus $\bar{r}_n(x) \leq r_n(x)$. In fact, $\bar{r}_n$ equals to r_n if each continuity component has width less than δ_0 so we can think of $\bar{r}_n$ as the length of continuity components then we partition $\mathbb{T}^2$ into the strips of width δ_0 and regard the boundaries of the strips as "artificial singularities".

It suffices to prove (5.1) with r_n replcaed by $\bar{r}_n$. To this end let

$$Z_n = \sup_{\varepsilon > 0} \frac{\text{mes}(x \in \gamma : \bar{r}_n(x) \leq \varepsilon)}{\varepsilon}.$$

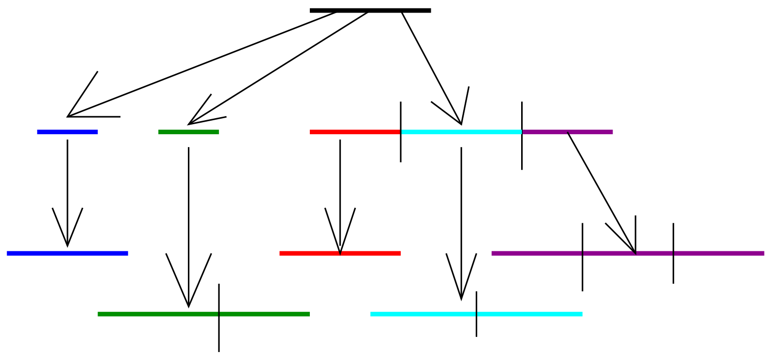

FIGURE 15. Dynamics of components

Then $Z_0 = \frac{2}{|\gamma|}$. We claim that there are constants $\theta < 1$, $C > 0$ such that

$$Z_{n+1} \leq \theta Z_n + C|\gamma|.$$

Indeed $\bar{r}_n(x)$ is less than ε if $T^{n+1}x$ passes near either genuine or artificial singularity. In the first case $T^n x$ is $\frac{\varepsilon}{\lambda}$ close to the preimage of sinularity. Since each curve is cut into at most $\kappa_1(\delta_0)$ components, we conclude that each component of T^n contributes by less than

$$\kappa_1(\delta_0)\text{mes}(x : r_n(x) \leq \frac{\varepsilon}{\lambda}) \leq \frac{\kappa_1(\delta_0)}{\lambda} Z_n.$$

On the other hand for long curves the relative measure of points with small $\bar{r}_{n+1}$ is less than $C(\delta_0)\varepsilon$ so their contribution is less than $C(\delta_0)\varepsilon|\gamma|$. The result follows.

6. Central Limit Theorem for Dynamical Systems.

In Section 5 we saw that the hyperbolic sawtooth map is ergodic. Ergodocity means that for a smooth function we have

$$\frac{1}{n}\sum_{j=0}^{n-1} A(f^j x) \to \mu(A)$$

so there is a natural question about the rate of convergence. If the system is sufficiently chaotic we expect that the behavior of the above sum is similar to the case of independent identically distributed (iid) random variables that is the fluctuations satisfy the Central Limit Theorem (CLT). In the first part of this section we review the methods to prove the CLT for dynamical systems while the second part contains applications to bouncing balls.

6.1. **iid random variables.** In order to explain how the method works we start with simplest possible settings. Let X_n be independent identically distributed random variables which are uniformly bounded. (Of course the assumption that X_n are bounded is unnecessary. We impose it in order to simplify the exposition.) We assume that $\mathbb{E}(X) = 0, \mathbb{E}(X^2) = D^2$. Denote $S_N = \sum_{n=1}^N X_n$. The classical Central Limit Theorem says that $\frac{S_N}{\sqrt{N}}$ converges weakly to the normal random variable with zero mean and variance D^2. Our idea for proving this result is the following. We know the distribution of S_0 so we want to see how the distribution changes when we change N. To this end let $M \to \infty$ so that $M/N \to t$. Then

$$\frac{S_M}{\sqrt{N}} = \frac{\sqrt{M}}{\sqrt{N}} \frac{S_M}{\sqrt{M}} \approx \sqrt{t} \frac{S_M}{\sqrt{M}}.$$

The second factor here is normal with zero mean and variance tD^2. Since multiplying normal random variable by a number has an effect of multiplying its variance by the square of this number the classical Central Limit Theorem can be restated as follows.

Theorem 6.1. *As $N \to \infty$ $\frac{S_{Nt}}{\sqrt{N}}$, converges weakly to the normal random variable with zero mean and variance $t\sigma^2$.*

Thus we wish to show that for large N our random variables behave like the random variables with density $p(t, x)$ whose Fourier transform satisfies

$$\hat{p}(t, \xi) = \exp\left(-\frac{tD^2\xi^2}{2}\right).$$

Hence

$$\partial_t \hat{p} = (i\xi)^2 \frac{D^2}{2} \hat{p} \text{ and so } \partial_t p = \frac{D^2}{2} \partial_x^2 p.$$

Recall that any weak solution of the heat equation is also strong solution so we need show that if $v(t, x)$ is a smooth function of compact support in x then

$$\int v(T, x) p(T, x) dx - v(0, 0) = \iint p(t, x) \left[\partial_t v + \frac{\sigma^2}{2} \partial_x^2 v\right] dx dt. \tag{6.1}$$

In case $v(t, x) = u(x)$ is independent of t the last equation reduces to

$$\int u(x) p(T, x) dx - u(0) = \iint p(t, x) (\mathcal{L}u)(x) dx dt. \tag{6.2}$$

Conversely if (6.2) holds for each T and if $\mathbf{S}_t$ is any limit point of $\frac{S_{Nt}}{\sqrt{N}}$ then

$$\partial_t \mathbb{E}(u(\mathbf{S}_t)) = \mathbb{E}((\mathcal{L}u)(\mathbf{S}_t))$$

where $\mathcal{L} = \dfrac{D^2}{2}\partial_x^2$ which implies (6.1) for functions of the form $v(t,x) = k(t)u(x)$ and hence for the dense family $\sum_j k_j(t)u_j(x)$. Thus p satisfies the heat equation as claimed. Thus we have to establish (6.2). For discrete system in amounts to showing that

$$\mathbb{E}\left(u\left(\frac{S_M}{\sqrt{N}}\right)\right) - u(0) - \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}\left((\mathcal{L}u)\left(\frac{S_n}{\sqrt{N}}\right)\right) = o(1). \tag{6.3}$$

where $M \sim tN$. Consider the Taylor expansion

$$u\left(\frac{S_{n+1}}{\sqrt{N}}\right) - u\left(\frac{S_n}{\sqrt{N}}\right) =$$

$$(\partial_x u)\left(\frac{S_n}{\sqrt{N}}\right)\frac{X_n}{\sqrt{N}} + \frac{1}{2}\left(\partial_x^2 u\right)\left(\frac{S_n}{\sqrt{N}}\right)\frac{X_n^2}{N} + \mathcal{O}(N^{-3/2}). \tag{6.4}$$

Taking the expectation and using the fact that $\mathbb{E}(X_n) = 0$ we obtain (6.3).

Keeping the above example in mind we can summarize martingale problem approach as follows.

In order to describe the distribution of $\mathbf{S}_t$ we need to compute the averages $\mathbb{E}(u(\mathbf{S}_t))$ for a large class of test functions u. However rather than trying to compute the above averages directly we would like to split the problem in two two parts. First we find an equation which this average should satisfy. Secondly we show that this equation has unique solution. Only the first part involves the study of the system in question. The second part deals with a PDE question.

For the first step we need to compute the generator

$$(\mathcal{L}u)(x) = \lim_{N\to\infty}\lim_{h\to 0}\frac{\mathbb{E}(u(S_t^N)|S_0^N = x) - u(x)}{h}.$$

For the second step we need to establish the uniqueness for the equation

$$\partial_t u = \mathcal{L}u.$$

Once this is done we conclude that for a large class of test functions we have

$$\mathbb{E}(v(T,\mathbf{S}_t)) - \mathbb{E}(v(0),\mathbf{S}_0) = \int_0^T \mathbb{E}(\partial_t v + \mathcal{L}v)(t,\mathbf{S}_t)dt.$$

Choosing here v satisfying the final value problem

$$\partial_t v + \mathcal{L}v = 0, \quad v(T,\mathbf{S}) = u(\mathbf{S}) \tag{6.5}$$

we can achieve our goal of finding $\mathbb{E}(u(\mathbf{S}_T))$.

6.2. **Hyperbolic systems.** Now let us discuss how to extend this approach to the dynamics setting. Namely, we consider the case where

$$S_n = \sum_{j=0}^{n-1} A(f^j x)$$

where f is an Anosov diffeomorphism of $\mathbb{T}^2$ and A is a smooth function. In fact the Anosov property was not important for our argument. The natural setting for the our approach is the following

(1) There is an invariant cone family $df(\mathcal{K}) \subset \mathcal{K}$ and for each $v \in \mathcal{K}$ we have

$$||df(v)|| \geq \Lambda ||v||, \quad \Lambda > 1. \tag{6.6}$$

(2) f is mixing in the following sense. There exists a measure μ_{SRB} called *SRB (Sinai-Ruelle-Bowen) measure* such that the following holds. Let γ be a curve of length between 1 and 2 whose tangent direction lies inside K and $||\gamma||_{C^2} \leq \bar{C}$. Let ρ be a Holder probability density on γ then

$$\left| \int_\gamma \rho(x) A(f^n x) dx - \mu_{SRB}(A) \right| \leq C\theta^n ||A||_{C^r} ||\rho||_{C^\alpha(\gamma)}. \tag{6.7}$$

In fact we will need an extension of (6.7) to the case ρ is not probability density. In case $\rho > 0$ we can apply (6.7) to $\tilde{\rho} = \rho / \int_\gamma \rho dx$ to get

$$\left| \int_\gamma \rho(x) A(f^n x) dx - \int_\gamma \rho dx \; \mu_{SRB}(A) \right| \leq C\theta^n ||A||_{C^r} ||\rho||_{C^\alpha(\gamma)}. \tag{6.8}$$

Finally in case ρ changes sign we can apply (6.8) to $\rho^+ = \max(\rho, 0)$ and $\rho^- = \min(\rho, 0)$ to show that (6.8) is valid for arbitrary Holder densities.

We assume that $\mu_{SRB}(A) = 0$. This does not cause a loss of generality since we can always replace A by $A - \mu_{SRB}(A)$. Concerning the initial condition x we assume that it is distributed on γ with a density ρ where γ and ρ are as above.

The difference with the previous example is that $A(f^n x)$ and S_n are no longer independent so a more careful analysis of (6.4) is needed. Take $L_N = N^{0.01}$ and let $\bar{n} = n - L_N$. We have

$$\mathbb{E}\left((\partial_x^2 u) \left(\frac{S_n(x)}{\sqrt{N}} \right) A^2(f^n x) \right) = \mathbb{E}\left((\partial_x^2 u) \left(\frac{S_{\bar{n}}(x)}{\sqrt{N}} \right) A^2(f^n x) \right) + \mathcal{O}\left(\frac{L_N}{\sqrt{N}} \right)$$

To estimate this expression we assume temporarily that

$$|| \ln \rho ||_{C^\alpha} \leq C. \tag{6.9}$$

Decompose $f^{\bar{n}}\gamma = \cup_\beta \gamma_\beta$ where $1 \leq$length$(\gamma_\alpha) \leq 2$. Then

$$\int_\gamma \rho(x) \left(\partial_x^2 u\right) \left(\frac{S_{\bar{n}}(x)}{\sqrt{N}}\right) A^2(f^n x) dx = \sum_\beta c_\beta \int_{\gamma_\beta} \rho_\beta(y) \left(\partial_x^2 u\right) \left(\frac{S_{\bar{n}}(f^{-\bar{n}}y)}{\sqrt{N}}\right) A^2(f^{L_N} y) dy$$

where $c_\beta = \mathbb{P}(x \in \gamma_\beta)$ and $\rho_\beta(y) = \rho(f^{-\bar{n}}y)|(df^{-\bar{n}}y|Tf^{\bar{n}}\gamma)|/c_\beta$. Next we claim that the Holder norm of $\ln|(df^{-\bar{n}}y|Tf^{\bar{n}}\gamma)|$ is uniformly bounded (in $\bar{n}$). Indeed

$$\left|\ln|(df^{-\bar{n}}y_1|Tf^{\bar{n}}\gamma) - \ln|(df^{-\bar{n}}y_2|Tf^{\bar{n}}\gamma)\right|$$

$$\leq \sum_{j=0}^{\bar{n}-1} \left|\ln|(df^{-1}f^{-j}y_1|Tf^{\bar{n}-j}\gamma) - \ln|(df^{-1}f^{-j}y_2|Tf^{\bar{n}-j}\gamma)\right| \leq C \sum_{j=0}^{\bar{n}-1} d(f^{-j}y_1, f^{-j}y_2).$$

Due to (6.6) the individual term in this sum is bounded by $2\Lambda^{-j}$ proving our claim. Here we have used the fact that C^2 norms of $T^j\gamma$ are uniformly bounded. The proof of this can be found in [18].

Our claim implies in particular that for each $y_1, y_2 \in \gamma_\beta$ we have

$$\frac{1}{C} \leq \frac{|(df^{-\bar{n}}y_1|Tf^{\bar{n}}\gamma)|}{|(df^{-\bar{n}}y_2|Tf^{\bar{n}}\gamma)|} \leq C.$$

Since due to (6.9) we also have

$$\frac{1}{C} \leq \frac{\rho(y_1)}{\rho(y_2)} \leq C$$

it follows that $||\rho_\beta||_{C^\alpha(\gamma_\beta)} \leq C$. A similar argument shows that

$$\left\|\left(\partial_x^2 u\right)\left(\frac{S_{\bar{n}}(f^{-\bar{n}}y)}{\sqrt{N}}\right)\right\|_{C^\alpha(\gamma_\beta)} \leq C.$$

Now applying (6.8) we obtain

$$\mathbb{E}\left(\left(\partial_x^2 u\right)\left(\frac{S_{\bar{n}}(f^{-\bar{n}}y)}{\sqrt{N}}\right) A^2(f^{L_N}y)\right) \tag{6.10}$$

$$= \mathbb{E}\left(\left(\partial_x^2 u\right)\left(\frac{S_{\bar{n}}(f^{-\bar{n}}y)}{\sqrt{N}}\right)\right) \mu_{SRB}\left(A^2\right) + \mathcal{O}\left(\theta^{L_N}\right)$$

$$= \mathbb{E}\left(\left(\partial_x^2 u\right)\left(\frac{S_n(x)}{\sqrt{N}}\right)\right) \mu_{SRB}\left(A^2\right) + \mathcal{O}\left(\frac{L_N}{\sqrt{N}}\right).$$

The above argument relies on (6.9). However by decomposing arbitrary density

$$\rho = 10||\rho||_{C^\alpha} - (10||\rho||_{C^\alpha} - \rho)$$

we see that (6.10) is valid in general.

(6.10) takes care about the second derivative However he first derivative term is more difficult since it comes with smaller prefactor $\frac{1}{\sqrt{N}}$. We have

$$\mathbb{E}\left((\partial_x u)\left(\frac{S_n}{\sqrt{N}}\right)A(f^n x)\right)$$

$$=\mathbb{E}\left((\partial_x u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)A(f^n x)\right)+\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)A(f^n x))\sum_{k=\bar n}^{n-1}\frac{A(f^k x)}{\sqrt{N}}\right)+\mathcal{O}\left(\frac{L_N^2}{N}\right).$$

As before

$$\mathbb{E}\left((\partial_x u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)A(f^n x)\right)=\mathcal{O}\left(\theta^{L_N}\right).$$

To address the second term fix a large M_0, let $m=n-k$ and consider two cases

(I) $m>M_0$. Then letting $y=f^k x$ and arguing as in the proof of (6.10) we get

$$\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)A(f^k x))A(f^n x)\right)=\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}(f^{-k}y)}{\sqrt{N}}\right)A(y))A(f^m y)\right)$$

$$=\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)A(f^k x))\right)\mu_{SRB}(A)+\mathcal{O}(\theta^m)=\mathcal{O}(\theta^m).$$

(II) $m\leq M_0$. Denote $B_m(y)=A(y)A(f^m y)$. Then we have

$$\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)A(f^k x))A(f^n x)\right)=\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}(f^{-\bar n}y}{\sqrt{N}}\right)B_m(f^k y))\right)$$

$$=\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_{\bar n}}{\sqrt{N}}\right)\right)\mu_{SRB}\left(B_m(f^k y))\right)+\mathcal{O}\left(\theta^{L_N}\right).$$

Summation over m gives

$$\mathbb{E}\left(u\left(\frac{S_M}{\sqrt{N}}\right)\right)-u(0)=\frac{1}{N}\sum_{n=0}^{M-1}\frac{D_{M_0}^2}{2}\mathbb{E}\left((\partial_x^2 u)\left(\frac{S_M}{\sqrt{N}}\right)\right)+\mathcal{O}\left(\theta^{M_0}\right)+o(1)$$

where

$$D_{M_0}^2=\mu_{SRB}(A^2)+2\sum_{m=1}^{M_0}\mu_{SRB}(A(x)A(f^m x))=\sum_{m=-M_0}^{M_0}\mu_{SRB}(A(x)A(f^m x)).$$

Letting $M_0\to\infty$ we obtain that $\frac{S_N}{\sqrt{N}}$ is asymptotically normal with zero mean and variance given by the *Green-Kubo formula*

$$D^2=\sum_{m=-\infty}^{\infty}\mu_{SRB}(A(x)A(f^m x)).$$

6.3. **Bouncing balls in very weak potentials.** Last subsection shows how to obtain the CLT for systems satisfying (6.6) and (6.7). This conditions are quite restrictive and to enhance the applicability of CLT one needs to rely on wearer versions of those conditions. Here we explain how to do this for systems with multiple scales. As an example we consider a ball in weak potential gx^α for $\alpha \ll 1$. To simplify the formulas we consider the SWA

$$t_{n+1} = t_n + T(v_n), \quad v_{n+1} = v_n + 2\dot{f}(t_{n+1}).$$

An easy calculation using energy conservation shows that

$$T(v) \sim cv^\sigma, \quad T'(v) \sim c\sigma v^{\sigma-1}, \quad T''(v) \sim c\sigma(\sigma-1)v^{\sigma-2}$$

where $\sigma = \frac{2}{\alpha} - 1$. Suppose that $v_0 \gg 1$ so that the relative change of velocity is small. Our goal is to show that if $\sigma \gg 1$ then the change of velocity is well approximated by the Brownian Motion.

In fact, take $N \sim v_0^2$ and consider $W_N(t) = \frac{1}{\sqrt{N}} v_{tN}$. In order to keep v of order $\sqrt{N}$ during the time $[0, N]$ we stop process when either $v_N \geq M\sqrt{v_0}$ or $v_N \leq M\sqrt{v_0}$ for some (large) constant M. Take a test function u and consider

$$u\left(\frac{v_{n+1}}{\sqrt{N}}\right) - u\left(\frac{v_n}{\sqrt{N}}\right) = \partial u\left(\frac{v_n}{\sqrt{N}}\right)\frac{\dot{f}(t_{n+1})}{\sqrt{N}} + \frac{1}{2}\partial^2 u\left(\frac{v_n}{\sqrt{N}}\right)\frac{g^2(t_{n+1})}{N}$$

where $g = (\dot{f})^2$. We have

$$dF = \begin{pmatrix} 1 & T'(v_n) \\ \dot{f}(t_{n+1}) & 1 + \dot{f}(t_{n+1})T'(v_n) \end{pmatrix}.$$

This shows that if δv_n is not too small then $dF(\delta t_n, \delta v_n)$ is almost parallel to $(1, \dot{f})$. More precisely let

$$\mathcal{K}(t, v) = \left\{(\xi, \eta) : \left|\frac{\xi}{\eta} - \dot{f}(t)\right| < v^{-\beta}\right\}$$

Then $dF(\mathcal{K}) \subset \mathcal{K}$ unless $\dot{f}(t_{n+1})T'(v)| \leq Cv^{-\beta}$. If $\dot{f}$ is Morse then this amounts to $(t, v) \in \mathbf{C}$ where

$$\mathbf{C} = \left\{|t_{n+1} - t_{cr}| < \bar{C}v^{-\beta-\sigma} \text{ for some critical point of } \dot{f}\right\} \tag{6.11}$$

In other words, the cones are preserved on the major part of the phase space.

Next let γ be a curve with $T\Gamma \subset \mathcal{K}$ and let ρ be a smooth denisty on γ. Then

(6.12)

$$\int u'\left(\frac{v_n}{\sqrt{N}}\right)\dot{f}(t_{n+1})\rho(t_n)dt_n = \int u'\left(\frac{v_n}{\sqrt{N}}\right)\rho(t_n)\left(\frac{dt_{n+1}}{dt_n}\right)^{-1} df.$$

Let $\gamma = \{(t, h(t)\}$. Note that $\frac{dt_{n+1}}{dt_n} = \frac{\partial t_{n+1}}{\partial t_n} + \dot{h}(t_n)\frac{\partial t_{n+1}}{\partial t_n}$. Therefore $|\frac{dt_{n+1}}{dt_n}| > cv^{\sigma-1-\beta}$ if $(t_n, v_n) \notin \mathbf{C}$. We now integrate (6.12) by parts. We have

$$\frac{d}{dt_n}\left(\frac{dt_{n+1}}{dt_n}\right)^{-1} = -\left(\frac{dt_{n+1}}{dt_n}\right)^{-2}\left[\ddot{h}(t_n)V'(v_n) + h^2(t_n)V''(v_n)\right] = \mathcal{O}\left(v^{2\beta-\sigma-1}\right) = \mathcal{O}\left(v^{-\beta}\right)$$

if we choose $\beta = \sigma - 1 - 2\beta$. The same bounds hold for other terms which we need to integrate by parts and we have

$$\int_\gamma u'\left(\frac{v_n}{\sqrt{N}}\right)\dot{f}(t_{n+1})\rho(t_n)dt_n = \mathcal{O}\left(v^{-\beta}\right).$$

Next starting with γ such that $T\gamma \subset \mathcal{K}$ we see that $F^n\gamma = (\cup_j \gamma_j) \cup Z$ where Z corresponds to the points which visit $\mathbf{C}$ defined by (6.11) for the first n iterates so that $\mathrm{mes}(Z) = \mathcal{O}\left(nv^{-\beta}\right)$. Accordingky we have

$$\mathbb{E}\left(u'\left(\frac{v_n}{\sqrt{N}}\right)\dot{f}(t_{n+1})\right) = \mathcal{O}\left(nv^{-\beta}\right).$$

A similar argument shows that

$$\mathbb{E}\left(u''\left(\frac{v_n}{\sqrt{N}}\right)g(t_{n+1})\right) = \mathbb{E}\left(u''\left(\frac{S_n}{\sqrt{N}}\right)\right)\int_0^1 g^2(t)dt + \mathcal{O}\left(nv^{-\beta}\right).$$

Combining the above bounds we get

$$\mathbb{E}\left(u\left(\frac{v_n}{\sqrt{N}}\right) - u\left(\frac{v_0}{\sqrt{N}}\right)\right) = \frac{1}{N}\sum_{k=0}^{N-1}\frac{\mathbf{D}^2}{2}\mathbb{E}\left(u''\left(\frac{v_k}{\sqrt{N}}\right)\right) + \mathcal{O}\left(N^{3/2}v^{-\beta}\right).$$

where

$$\mathbf{D} = \int_0^1 \left(\dot{f}(t)\right)^2 dt.$$

Recall that $v \sim N^2$ so that the error term is small if $\beta > 3$, that is $\sigma > 10$ or $\alpha < \frac{1}{5}$.

6.4. **Recurrence in very weak potentials.** In this section we present an application of the Central Limit Theorem to recurrence.

7. Invariant comes and hyperbolicity.

7.1. **Dimension 2.** In Sections 5 and 6 we saw that in order to ensure strong stochasticity we need to construct a cone family $\mathcal{K}(x)$ such that this family is invariant: $df(\mathcal{K}(x)) \subset \mathcal{K}(x)$ and df expands the cones, that is, there is a constant $\lambda > 1$ such that for all $v \in \mathcal{K}(x)$ we have $||df(v) \geq \lambda||v||$. Here we shall show that in the area preserving case the mere existence of invariant comes implies expansion. We begin with the following elementary fact.

Lemma 7.1. *If $A \in SL_2(\mathbb{R})$ has positive elements then it is hyperbolic.*

This result is quite intuitive. If A has positive elements then the angle between Ae_1 and Ae_2 is less than $\frac{\pi}{2}$ and since due to area preservation

$$||Ae_1||||Ae_2|| \sin \angle(Ae_1, Ae_2) = 1$$

there should be some expansion. The analytic prove is also easy. If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $ad = 1 + bc > 1$ and so $a + d \geq 2\sqrt{ad} > 2$.

The above proof does not show where the expanding direction is located. There is another argument which is equally simple but has an advantage that it works for a product of different matrices. This argument is based on the classical notion of Lyapunov function. Let ϕ_0 be the angle which vector (x_0, y_0) makes with x axis and ϕ_1 be the angle which vector $(x_1, y_1) = A(x_0, y_0)$ makes with x axis. Then $\phi_1 = g(\phi_0)$ for a continuous function g satisfying $0 < g(0) < g(\frac{\pi}{2}) < \frac{pi}{2}$. By the intermediate value theorem there exists ϕ such that $g(\phi) = \phi$ and hence $(x_1, y_1) = \lambda(x_0, y_0)$. To estimate λ let $Q(x, y) = xy$. Then

$$Q(x_1, y_1) = \lambda^2 x_0 y_0 = x_1 y_1 = acx_0^2 + bdy_0^2 + (ad+bc)x_0 y_0 > (ad+bc)x_0 y_0 = (1+2b_0 c_0)x_0 y_0.$$

It follows that $\lambda > \sqrt{1 + 2bc} > 1$.

The previous argument shows that Q increases after the application of A. The same argument works for compositions. Namely, if $A_1, A_2 \dots A_n$ are positive $SL_2(\mathbb{R})$ matrices and

$$v_n = A_n \dots A_2 A_1 v_0$$

then

$$||v_n|| \geq 2\sqrt{Q(v_n)} \geq 2Q(v_0) \prod_{j=1}^{n} \Lambda_j$$

where $\Lambda_j = (1 + 2b_j c_j)$.

To get a coordinate free interpretation of this result suppose that $f : M^2 \to M^2$ preserves a smooth measure given by $\mu(A) = \iint_A \omega$ and that there is a family of cones $\mathcal{K}(x)$ such that along an orbit $x_n = f^n x_0$ we have $df(\mathcal{K}(x_n) \subset \mathcal{K}_{n+1}$. Choose a basis in $T_x M$ so that

$$\mathcal{K}(x) = \{e = \alpha_1 e_1 + \alpha_2 e_2 : \alpha_1 > 0 \text{ and } \alpha_2 > 0\}$$

and $\omega(e_1, e_2) = 1$. Then df can be represented by an $SL_2(\mathbb{R})$ matrix and by the above inequality we have $||df^n(v_0) \geq 2\sqrt{Q(v_0)} \prod_{j=0}^{n-1} \Lambda_j$ where $\Lambda_j = 1 + 2b_j c_j$.

7.2. **Higher dimensions.** Here we present a multidimensional version of this estimate which is due to Wojtkowski. Consider a symplectic space $(\mathbb{R}^{2d}, \omega)$. Let V_1 and V_2 be two transversal Lagrangian subspaces $(\omega|_{V_j} = 0)$. Than each vector $v \in \mathbb{R}^{2d}$ has a unique decomposition $v = v_1 + v_2, v_j \in V_j$. Let $Q(v) = \omega(v_1, v_2)$. We can choose frames in V_1 and V_2 so that if

$$u_1 = (\xi_1, \eta_1), u_2 = (\xi_2, \eta_2) \text{ where } \xi_j \in V_1, \eta_j \in V_2$$

then $\omega(v_1, v_2) = \langle \xi_1, \eta_2 \rangle - \langle \xi_2, \eta_1 \rangle$. Then $Q((\xi, \eta)) = \langle \xi, \eta \rangle$. Define $\mathcal{K} = \{v : Q(v) \geq 0$.

Let L be a linaer symplectic matrix. We can write L in the block form with respect to the decomposition $\mathbb{R}^{2d} = V_1 \oplus V_2 : L = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. The sympleticity condition amounts to the equations

$$A^*D - C^*B = I, \quad A^*C = C^*A, \quad D^*B = B^*D.$$

One important case is $\tilde{L} = \begin{pmatrix} I & R \\ P & C \end{pmatrix}$. Then we have

$$P^* = P, \quad R^*S = S^*R \quad \text{and} \quad S - PR = I.$$

The last two equations give

$$R^*S - R^*PR = R^* \quad \text{that is} \quad (S^* - R^*P)R = R^*.$$

But $S^* - R^*P = (S - PR)^* = I$. Therefore the sympleticity of $\tilde{L}$ amounts to

$$R^* = R, \quad P^* = P, \quad S - PR = I. \tag{7.1}$$

We say that L is *monotone* if $L\mathcal{K} \subset \mathcal{K}$ and *strictly monotone* if $L\mathcal{K} \subset \text{Int}(\mathcal{K}) \cup \{0\}$.

Lemma 7.2. *If L is monotone then LV_1 is transversal to V_2 and LV_2 is transversal to V_1.*

Proof. Suppose to the contrary that there is $0 \neq v_1$ such that $Lv_1 \in V_2$. Take $v_2 \in V_2$ such that $\omega(v_1, v_2) > 0$. We have

$$\omega(v_1, v_2) = \omega(Lv_1, Lv_2) = \omega(Cv_1, Bv_2).$$

Take $v_\varepsilon - v_1 + \varepsilon v_2$. Then $v_\varepsilon \in \mathcal{K}$ for $\varepsilon > 0$. On the hand

$$Q(Lv_\varepsilon) = \langle \varepsilon Bv_2, Cv_1 + \varepsilon DV_2 \rangle = -\varepsilon\omega(v_1, v_2) + \varepsilon^2\omega(Bv_2, Dv_2)$$

is negative for small positive ε giving a contradiction. $\square$

Lemma 7.2 implies that A is invertible, so we can consider $\hat{L} = \begin{pmatrix} A & 0 \\ 0 & (A^*)^{-1} \end{pmatrix}$. Note that $\hat{L}$ preserves Q since

$$Q(\hat{L}v) = \langle A\xi, (A^*)^{-1}\eta \rangle = Q(v).$$

We therefore have a decomposition $L = \hat{L}\tilde{L}$ where $\tilde{L} = \begin{pmatrix} I & R \\ P & A^*D \end{pmatrix}$ for some matrices P and R.

Theorem 7.3. *L is monotone iff $Q(Lv) \geq Q(v)$ for all $v \in \mathbb{R}^{2d}$.*
L is strictly monotone iff $Q(Lv) > Q(v)$ for all $0 \neq v \in \mathbb{R}^{2d}$.

Proof. We prove the first statement, the second is similar.

Clearly, if L increases Q and $v \in \mathcal{K}$ then $Q(Lv) \geq Q(v)$, so $Lv \in \mathcal{K}$.

Conversely, suppose $\mathcal{K}$ is monotone. Since $\hat{L}$ preserves Q, we need to show that $Q(\tilde{L}v) \geq Q(v)$. Due to (7.1) we have

$$\tilde{L}(\xi, \eta) = (\xi + R\eta, P\xi + \eta + PR\eta)$$

so

$$Q(\tilde{L}(\xi, \eta)) - Q(\xi, \eta) = \langle R\eta, \eta \rangle + \langle P\zeta, \zeta \rangle \tag{7.2}$$

where $\zeta = \xi + R\eta$. Since $Q(\tilde{L}(\xi, 0)) = \langle P\xi, \xi \rangle$ so $P \geq 0$. Our next goal is to sho that $R \geq 0$. To this end consider an eigenvector $R\eta = \lambda\eta$. Take $\xi = a\eta$. Then $(\xi, \eta) \in cK$ if $a > 0$. On the other hand

$$Q(\tilde{L}(\xi, \eta)) = (a + \lambda)\langle \eta, \eta \rangle + (a + \lambda)^2 \langle P\eta, \eta \rangle.$$

Therefore $Q(\tilde{L}(\xi, \eta)) < 0$ for $a = -\lambda - \varepsilon$. Hence $-\lambda < 0$, that is $\lambda > 0$. This proves that $R \geq 0$ Now (7.2) gives $Q(\tilde{L}(\xi, \eta)) \geq Q((\xi, \eta))$ as claimed. $\square$

This proves shows in particular that if L is monotone then it is strictly monotone iff $P > 0$ and $R > 0$, that is, if $L(V_j) \subset \mathrm{Int}(\mathcal{K}) \cup \{0\}$.

Next let $L_1, L_2 \dots L_n$ be a sequnce of monotone maps. Pick c so that $||v|| \geq c\sqrt{Q(v)}$. Let $v_n = L_n \dots L_2 L_1 v_0$. Then for $v_0 \in \mathrm{Int}(\mathcal{K})$ we have

$$||v_n|| \geq c\sqrt{Q(v_n)} \geq c\sqrt{Q(v_0)} \prod_{j=1}^{n} \Lambda_j$$

where $\Lambda_j = \Lambda(L_j)$ and $\Lambda(L) = \min_{v \in \mathrm{Int}(\mathcal{K})} \sqrt{\frac{Q(Lv)}{Q(v)}}$.

To compute $\Lambda(L)$ note that

$$\begin{pmatrix} R^{-1/2} & 0 \\ 0 & R^{1/2} \end{pmatrix} \begin{pmatrix} I & R \\ P & I + PR \end{pmatrix} \begin{pmatrix} R^{1/2} & 0 \\ 0 & R^{-1/2} \end{pmatrix} = \begin{pmatrix} I & I \\ K & I + K \end{pmatrix}$$

where $K = R^{1/2}PR^{1/2} = R^{1/2}(PR)R^{-1/2}$. Note that $PR = A^*D - I = C^*B$. Choose an orthogonal matrix F such that $F^{-1}KF$ is diagonal. Then

$$\begin{pmatrix} F^{-1} & 0 \\ 0 & F^{-1} \end{pmatrix}\begin{pmatrix} I & I \\ K & I+K \end{pmatrix}\begin{pmatrix} F & 0 \\ 0 & F \end{pmatrix} = \begin{pmatrix} I & I \\ T & I+T \end{pmatrix} \tag{7.3}$$

where $T = F^{-1}KF$ is diagonal and $\mathrm{Sp}(T)$ $=\mathrm{Sp}(C^*B)$. We can also assume by choosing F appropriately that the diagonal elements of T are increasing. Denoting by M the RHS of (7.3) we have $\Lambda(M) = \Lambda(L)$. We have

$$\begin{aligned} Q(Mv) &= \langle \xi, \eta\rangle + \langle \eta, \eta\rangle + \langle T(\xi+\eta), (\xi+\eta)\rangle \\ &= \sum_{j=1}^{d}\left[t_j\xi_2^2 + (1+2t_j)\xi_j\eta_j + (1+t_j)\eta_j^2\right] \\ \sum_{\eta_j\geq 0} &\left[\left(\sqrt{t_j}\xi_j - \sqrt{1+t_j}\eta_j\right)^2\left(\sqrt{1+t_j}+\sqrt{t_j}\right)^2\xi_j\eta_j\right] \\ +\sum_{\eta_j<0} &\left[\left(\sqrt{t_j}\xi_j + \sqrt{1+t_j}\eta_j\right)^2\left(\sqrt{1+t_j}-\sqrt{t_j}\right)^2\xi_j\eta_j\right] \\ &\geq m(L)\sum_j \xi_j\eta_j = m(L)Q(v) \end{aligned}$$

where

$$m(L) = \min_j(\sqrt{1+t_j} - \sqrt{t_j})^2 = (\sqrt{1+t_1} - \sqrt{t_1})^2$$

and $t_1 \leq t_2 \leq \cdots \leq t_d$ are the eigenvalues of T. The equality is achived if $\xi_j = \eta_j = 0$ for $j \geq 2$ and $\sqrt{t_1}\xi_1 = \sqrt{1+t_1}\eta_1$.

Next, suppose that $f : M \to M$ is a symplectic map and there is a transverse family of Lagrangian subspaces $V_1(x), V_2(x)$ and an orbit $x_n = f^n x$ such that $df(\mathcal{K}(x_n)) \subset \mathcal{K}(x_{n+1})$. Let Q be the associated quadratic form and take small c so that $||v|| \geq c\sqrt{Q(v)}$. Choose frames so that

$$\omega((\xi_1, \eta_1)(\xi_2, \eta_2)) = \langle \xi_1, \eta_2\rangle - \langle \xi_2, \eta_1\rangle.$$

Let $df : T_xM \to T_{fx}M$ have block form $df = \begin{pmatrix} A(x) & B(x) \\ C(x) & D(x) \end{pmatrix}$. Let

$$\Lambda(x) = \min_{t\in \mathrm{Sp}(C^*B)}(\sqrt{t} + \sqrt{1+t}). \tag{7.4}$$

Then for $x \in \mathcal{K}(x_0)$ we have

$$||df^n(v_0)|| \geq c\left(\prod_{j=0}^{n-1}\Lambda(x_j)\right)\sqrt{Q(v_0)}. \tag{7.5}$$

7.3. **Lyapunov exponents.** Now we pass from the individual orbits to typical ones. Recall that given a diffeomorphism $f : M \to M$, a point x and a vector v in T_x, one can define the forward and backward Lyapunov exponents

$$\lambda_{\pm}(x,v) = \lim_{n\to\pm\infty} \frac{1}{n} \ln ||df^n(x)(v)||.$$

If f presereves a probability measure μ then, by Multiplicatie Ergodic Theorem, for $\mu-$almost all x $\lambda_{\pm}(x,v)$ exist for all v and they can take at most $\dim(M)$ different values.

In fact, there exists a splitting $T_xM = \oplus_{j=1}^s E_j$ and numbers $\lambda_1 > \lambda_2 > \lambda_s$ such that if $v = v_{i_1} + v_{i_2} + \dots v_{i_k}$ where $i_1 < i_2 < \dots < i_k$ and $0 \neq v_{i_k} \in E_{i_k}$ then $\lambda_+(x,v) = \lambda_{i_1}$ and $\lambda_-(x,v) = \lambda_{i_k}$. If μ is ergodic then λ_j are constant almost surely.

In case μ is a smooth measure and $\lambda_j \neq 0$ almost surely (in which case we say that the system has *non-zero Lyapunov exponents* or that it is *(nonuniformly) hyperbolic*) there are strong methods to control the statistical properties of f. In particular Pesin theory guarantees the existence of stable and unstable manifolds tangent to $E^- = \oplus_{\lambda_j<0} E_j$ and $E^+ = \oplus_{\lambda_j>0} E_j$ respectively. (Pesin theory was extended to systems with singularities by Katok-Strelcyn [13]. The main idea is to show that most orbits do not come to close to the singularities in the spirit of Lemma 5.2 of Section 5.) Also taking $\Sigma(x) = \cup_{y\in W^u(x)} W^s(x)$ we obtain a set of positive measure and if $x \in \mathcal{R}_2$ then almost all points in $\Sigma(x)$ have the same averages for all continuous functions. Therefore the systems with non-zero exponents has almost countable many ergodic components, that is M is a disjoint union $M = \cup B_j$ where B_j are invariant and f restricted to B_j is ergodic. In case they hyperbolicity comes from invariant cones as we describe below Chernov-Sinai-Wojtkowski-Liverani theory provides sufficient conditions for ergodicity. Namely one needs to ensure appropriate transversality conditions between the singularity manifolds and stable/unstable manifolds of f. Unfortunetly those transversality conditions are not easy to verify in practise so the ergodicity is not yet proven in all the examples where we can ensure nonzero exponents.

Returning to the computations of the Lyapunov exponents let us consider the setting of $2d$ dimensional symplectic manifold. In this case one can show that $(E_j)^\perp = \sum_{i\neq s-j} E_i$ and so $\dim(E_j) = \dim E_{s-j}$. Therefore in order to prove that the system has nonzero exponents it suffices to check that

$$\dim(E^+) \geq d. \tag{7.6}$$

Suppose now that at each point there are transversal Lagrangian subspaces $V_1(x), V_2(x)$ such that df is monotone with respect to the cone $\mathcal{K}_{V_1,V_2}$. Let $\Lambda(x)$ be defined by (7.4). In order to establish (7.6) we consider the smallest j such that $\dim(E_j^-) > d$ where

$$E_j^- = E_j \oplus E_{j+1} \oplus E_s.$$

Lemma 7.4. *If μ is ergodic then $\lambda_j \geq \int \ln \Lambda(x) d\mu(x)$.*

Proof. Let $D = (\xi, \xi)$ where we use the coordinates of Theorem 7.3. Then $E_j^- \cap D$ contains a nonzero vector v. Then $\lambda_+(x, v) \leq \lambda_j$. On the other hand in view of (7.5) and the Pointwise Ergodic Theorem we have

$$\lambda_+(x, v) = \lim_{n\to\infty} \frac{1}{n} \sum_j \ln \Lambda(f^j x) = \int \ln \Lambda(x) d\mu(x).$$

□

In general it is possible to have $\Lambda(x) \equiv 1$ (consider for example the map $(I, \phi) \to (I, \phi + I)$. Let

$$G = \{x : \Lambda(x) > 1\} = \{x : df(x) \text{ is strictly monotone}\}.$$

Consider now the smooth invariant measure

$$\mu(A) = \int_A \omega \wedge \cdots \wedge \omega.$$

Note that μ need not be ergodic.

Corollary 7.5. *If almost all points visit G then the system has nonzero Lyapunov exponents.*

Proof. We apply Lemma 7.4 to each ergodic component of G. The assumption that $\nu(G) > 0$ for each ergodic component implies that $\int \ln \Lambda(x) d\nu > 0$. □

7.4. **Examples.** Here we present several examples of systems possesing invariant cones. We discuss two dimensional examples in more detail since the computations are simpler in that case.

(I) Dispersing billiards. Consider a particle moving in a domain with piecewise concave boundaries. Let s be the arclenth parameter and ϕ be the angle with the tangent direction.

Lemma 7.6. *df has the following form in (s, ϕ) variables*

$$\begin{pmatrix} \frac{\kappa_0\tau + \sin\phi_0}{\sin\phi_1} & \frac{\tau}{\sin\phi_1} \\ \frac{\kappa_0\kappa_1\tau + \kappa_1\sin\phi_0 + \kappa_0\sin\phi_1}{\sin\phi_1} & \frac{\kappa_1\tau + \sin\phi_1}{\sin\phi_1} \end{pmatrix}$$

where κ_0 (κ_1) is the curvature of the boundary at the initial (final) point and τ is the flight length.

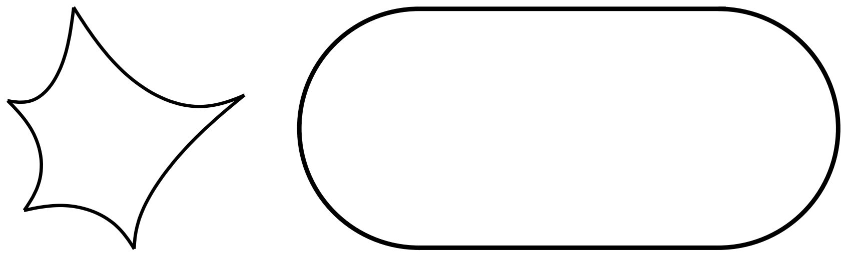

FIGURE 16. Two tables with nonzero Lyapunov exponents: dispersing billiard on the left and Bunimovich stadium on the right

Note that f preserves the form $\omega = \sin\phi ds \wedge d\phi$. The above matrix has all elements positive therefore df increases the qudratic form $Q = \sin\phi ds d\phi$. Moreover the product of the off diagonal terms with $\sin\phi_1$ is uniformly bounded from below so $\Lambda(s,\phi)$ is uniformly bounded away from 1.

Proof. We compute $\frac{\partial s_1}{\partial s_0}$, the other terms are similar. Consider figure 17. Let $|AB| = \delta s_0$. We have

$$|CB| \approx \sin\phi_0 \delta s_0, \quad |DE| = |BC|, \quad |EF| \approx \tau \sin\angle FBE,$$

$$\angle BFE \approx \kappa_0 \delta s_0, \quad |DG| \approx \delta s_1 \approx \frac{|DF|}{\sin\phi_1}.$$

□

For dispersing billiards we have $\kappa_0 > 0$, $\kappa_1 > 0$. Another way to make all elements of df positive is to have κ_0, κ_1 negative but require that

$$\tau \geq \frac{\sin\phi_0}{|\kappa_0|} + \frac{\sin\phi_0}{|\kappa_0|}.$$

The billiards satisfying the above condition are called *defocusing.* Perhaps the most famous example of the defocusing billiard is Bunimovich stadium.

Ergodicity of dispersing billiards is shown in [20]. Ergodicity of Bunimovich stadium is shown in [3]. Further properties of dispersing and defocusing billiards are discussed in [6].

(II) Dispersing pingpongs. Consider pingpong whose wall motion satisfies $\ddot{f}(t) < 0$ at all points of continuity.

Lemma 7.7. *In (t, v) varaibles the derivative takes form*

$$\begin{pmatrix} \frac{v_n - \dot{f}_n}{v_n + \dot{f}_{n+1}} & -\frac{L_n}{v_n^2(v_n + \dot{f}_{n+1})} \\ \frac{v_n - \dot{f}_n}{v_n + \dot{f}_{n+1}} \ddot{f}_{n+1} & 1 - \frac{L_n \ddot{f}_{n+1}}{v_n^2(v_n + \dot{f}_{n+1})} \end{pmatrix}$$

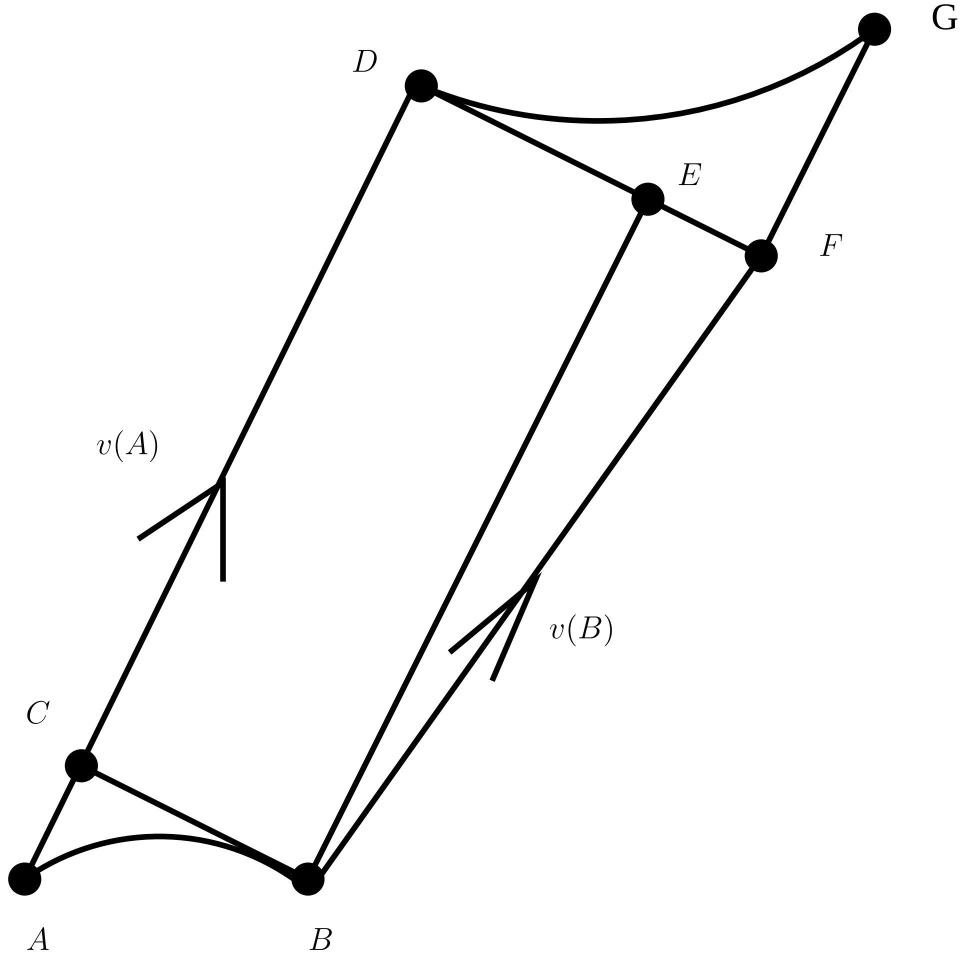

FIGURE 17. Computing $\dfrac{\partial s_1}{\partial s_0}$

where L_n is the distance traversed by the particle between n-th and $(n+1)$-st collisions.

Note that the off diagonal entries of the above matrix are negative so the form $Q = -dtdv$ is increasing.

Proof. Let us compute $\frac{\partial v_{n+1}}{\partial t_n}$. Refering to figure 6 we have

$$\delta h_n = (v_n - \dot{f}_n)\delta t_n, \quad \delta t_{n+1} = \frac{\delta h_n}{v_n + \dot{f}_n}, \quad \delta \dot{f}_{n+1} = \ddot{f}_{n+1}\delta t_{n+1}.$$

This proves the formula for $\frac{\partial v_{n+1}}{\partial t_n}$. Together with (1.5) this completes the estimate of t derivatives. v derivatives are computed similarly. $\square$

(III) Balls in gravity field. Consider two balls on the line moving in a gravity field and colliding elastically with each other and the fixed floor. Let m_1 be the mass of the bottom ball and m_2 be the mass of the top ball. It is convenient to use h and z as variables where $h = h_1$ is the energy of the bottom ball and $z = v_2 - v_1$ is the relative velocity of the second ball. We consider the balls at the moments when the bottom particle collides with the floor. During the collisions of the bottom ball with the floor our variables change as follows $(\bar{h}, \bar{z}) = F_1(h, z)$ where

$$F_1(h, z) = (h, z + c\sqrt{h}) \text{ where } c = \sqrt{\frac{8}{m_1}}.$$

Next we consider the collision between the walls. Using the formulas of Section 1 we find that the changes of energy and velocity are the following

$$\bar{z} = -z, \bar{v} = u + \frac{2m_2}{m_1 + m_2} z$$

where u is velocity of the first ball at the moment of collision. Accordingly

$$\bar{h} = h + \frac{2m_1 m_2 uz}{m_1 + m_2} + \frac{2m_1 m_2^2 z^2}{(m_1 + m_2)^2}.$$

To find u note that $u = v_1 - \tau z$ where τ is the time between collisions of the first ball with the floor and with the second ball. Next, $\tau = -\frac{x}{z}$ where z is the height of the second ball when the first one hits the floor. Therefore $uz = v_1 z + gx$. The energy of the system is

$$E = h + \frac{m_2(v_1 + z)^2}{2} + m_2 gx. \quad \text{Thus} \quad v_1 z + gx = \frac{E}{m_2} - \frac{h}{m_1} - \frac{h}{m_2} - \frac{z^2}{2}.$$

Accordingly $\bar{h} = b - h - az^2$ where $b = \frac{2m_1 E}{m_1 + m_2}$ and

$$a = \frac{m_1 m_2}{m_1 + m_2} - \frac{2m_1 m_2^2}{(m_1 + m_2)^2}.$$

Therefore if the ball returns to the floor after the collision we have

$$(\bar{h}, \bar{z}) = F_1 \circ F_2 \quad \text{where} \quad F_2(h, z) = (b - h - az^2, -z).$$

We assume that $m_1 > m_2$ so that $a > 0$. Note that

$$dF_1 = \begin{pmatrix} 1 & 0 \\ \frac{c}{2\sqrt{h}} & 1 \end{pmatrix}, \quad dF_2 = \begin{pmatrix} -1 & -2az \\ 0 & -1 \end{pmatrix} = -I \times \begin{pmatrix} 1 & 2az \\ 0 & 1 \end{pmatrix}.$$

Both

$$\begin{pmatrix} 1 & 0 \\ \frac{c}{2\sqrt{h}} & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 2az \\ 0 & 1 \end{pmatrix}$$

have positive elements so they are monotone with respect to $Q = dhdz$ while $-I$ is Q-isometry. Also note that $d(F_2 \circ F_1^k)$ is strictly monotone for each k and since starting from any initial condition we will eventually have a collision between the balls, Corollary 7.5 implies that this system has nonzero Lyapunov exponents.

Ergodicity of two balls in gravity under the condition $m_1 > m_2$ is proven in [14].

On the other hand if $m_1 = m_2$ then the particles just exchange their enrgy during the collisions so the function $I = \min(h_1, h_2)$ is the first integral of this system. One can also show [4] that for $m_1 < m_2$ elliptic islands are present so the system is not ergodic.

One can also construct multidimensional examples satisfying the above criteria. In particular n particles of the line in gravity field have nonzero exponents provided that $m_1 > m_2 > \cdots > m_n$ when the particles are numbered from the bottom up. The monotonicity of this system was proven in [22] while [19] showed that the conditions of Corollary 7.5 are satisfiedfor this system. One can also consider nonlinear potentials. [23] shows that the following conditions are sufficient for nonzero Lyapunov exponents

(i) $m_1 > m_2 > \cdots > m_n$; (ii) $U'(q) > 0$; (iii) $U''(q) < 0$.

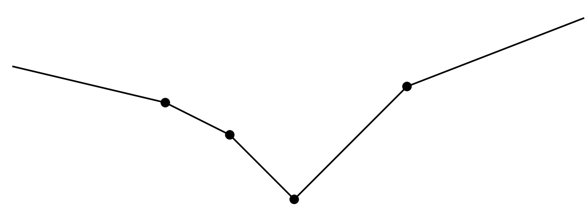

FIGURE 18. Wojtkowski wedge

Another example is the particle in gravity field moving in a two dimensional domain whose boundary consists of two concave broken lines meeting at a right angle. It is shown in [24] that this system has nonzero Lyapunov exponents.

Problem 7.8. Show ergodicity of the last two examples.

References

[1] G. Atkinson *Recurrence of co-cycles and random walks,* J. London Math. Soc. **13** (1976) 486–488.

[2] P. Boyland *Dual billiards, twist maps and impact oscillators,* Nonlinearity **9** (1996) 1411–1438.

[3] L. A. Bunimovich *On the ergodic properties of nowhere dispersing billiards,* Comm. Math. Phys. **65** (1979) 295–312.

[4] J. Cheng, M. Wojtkowski *Linear stability of a periodic orbit in the system of falling balls,* T MSRI Publ. **22** (1991) 53–71.

[5] N. Chernov *Ergodic and statistical properties of piecewise linear hyperbolic automorphisms of the 2-torus,* J. Statist. Phys. **69** (1992) 111–134.

[6] N. Chernov, R. Markarian *Chaotic billiards,* Math. Surv. & Monographs **127** (2006) AMS, Providence, RI. xii+316 pp.

[7] J. de Simoi, D. Dolgopyat *Dynamics of some piecewise smooth Fermi-Ulam models,* Chaos **22** (2012) paper 026124.

[8] D. Dolgopyat *Bouncing balls in non-linear potentials,* Discrete & Continuous Dyn. Sys.– A **22** (2008) 165–182

[9] D. Dolgopyat *Limit Theorems for Hyperbolic Systems,* Lecture Notes.

[10] D. Dolgopyat, B. Fayad *Unbounded orbits for semicircular outer billiard*, Ann. Henri Poincaré, **10** (2009) 357–375.

[11] J. Guckenheimer, P. Holmes *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields,* Appl. Math. Sciences **42** (1990) Springer-Verlag, New York, xvi+459 pp.

[12] P. Holmes *The dynamics of repeated impacts with a sinusoidally vibrating table,* J. Sound Vibration **84** (1982) 173–189.

[13] A. Katok, J.-M. Strelcyn *Invariant manifolds, entropy and billiards; smooth maps with singularities,* Lect. Notes in Math. **1222** (1986) Springer-Verlag, Berlin.

[14] C. Liverani, M. Wojtkowski *Ergodicity in Hamiltonian Systems,* Dynamics Reported **4** (1995) 130–202.

[15] J. Moser *Stable and random motions in dynamical systems,* Ann. Math. Studies **77** (1973) Princeton University Press, Princeton, N. J. viii+198 pp.

[16] L. D. Pustylnikov *Stable and oscillating motions in nonautonomous dynamical systems–II,* Proc. Moscow Math. Soc. **34** (1977), 3–103.

[17] L. D. Pustylnikov *Poincare models, rigorous justification of the second law of thermodynamics from mechanics, and the Fermi acceleration mechanism,* Russian Math. Surveys **50** (1995) 145–189.

[18] M. Shub *Global stability of dynamical systems,* Springer-Verlag, New York, 1987. xii+150 pp.

[19] N. Simanyi *The characteristic exponents of the falling ball model,* Comm. Math. Phys. **182** (1996) 457–468.

[20] Ya. G. Sinai *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards,* Russ. Math. Surv. **25** (1970) 141–192.

[21] S. Tabachnikov *Geometry and billiards,* Student Math Library **30** (2005) AMS Providence RI.

[22] M. Wojtkowski *A system of one dimensional balls with gravity,* Comm. Math. Phys. **126** (1990) 507–533.

[23] M. Wojtkowski *A system of one dimensional balls in external field,* Comm. Math. Phys. **127** (1990) 425–432.

[24] M. Wojtkowski *Hamiltonian systems with linear potential and elastic constraints,* Fund. Math. **157** (1998) 305–341.