LECTURES ON BOUNCING BALLS.

Contents

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

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

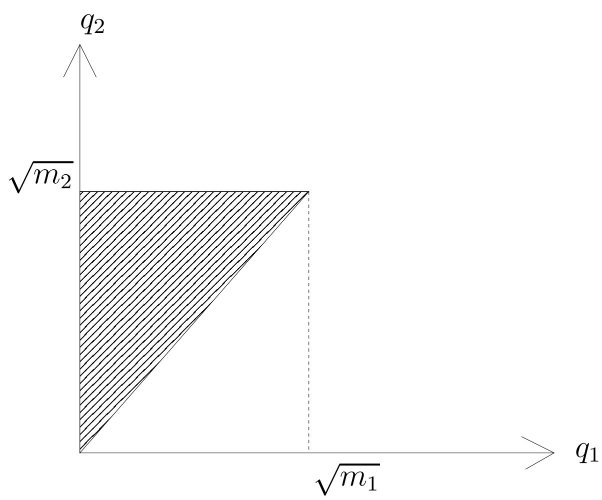

FIGURE 1. Configuration space for two points on the segment

The configuration space of the system becomes

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

This is a right triangle with hypotenuse 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 hypotenuse its speed is preserved and the angle which its velocity makes with $(\sqrt{m_1}, \sqrt{m_2})$ remains the same. Since $(\sqrt{m_1}, \sqrt{m_2})$ iscolinear to the boundary this see that the tangential component of the particle velocity is preserved. Since the length of the velocity vector is also conserved we see that the normal componet of the velocity is reversed. Therefore the change of velocity satisfies 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

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

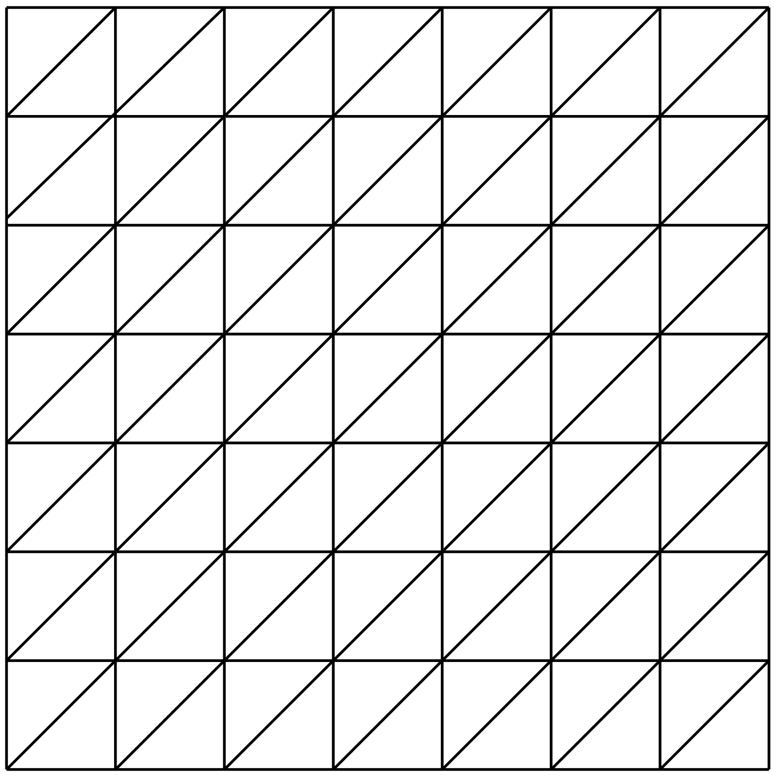

FIGURE 2. Configuration space of three points on the circle using the distance from the first point as coordinates

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 the first two particles. Since $(\sqrt{m_1}, \sqrt{m_2}, \sqrt{m_3})$ is collinear to the plane $P_{12}^l = \{\frac{q_1}{\sqrt{m_1}} - \frac{q_2}{\sqrt{m_2}} = l\}$ it follows that P_{12}^l is orthogonal to Π. Next, $\sqrt{m_1}u_1 + \sqrt{m_2}u_2$ is preserved Note that n_{12} is also colllinear to the plane P_{12}^l. Denoting by $\vec{n}_{12}^*$ the orthogonal projection of $\vec{n}_{12}$ to Π we

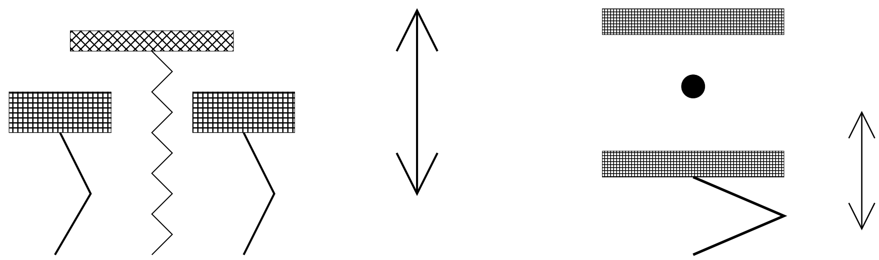

FIGURE 3. Impact oscillator (left) and Ulam pingpong (right) are two systems fitting into our setting

see that

$$\langle \vec{n}^*_{12}, \vec{u}^+ \rangle = \langle \vec{n}^*_{12}, \vec{u}^- \rangle$$

where $\vec{u} = (u_1, u_2, u_3)$. In other words, the tangential componet of the velocity is preserved and since the lenth of the velocity vector is also preserved we have an elastic collision.

We can also consider more particles on a line or a circle 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 $A < B$) 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 string and colliding with the wall. Apart from an easy mechanical implementation this system is also related 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 $A_0 \in \mathbb{R}^2 - \text{Int}\Gamma$, 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.

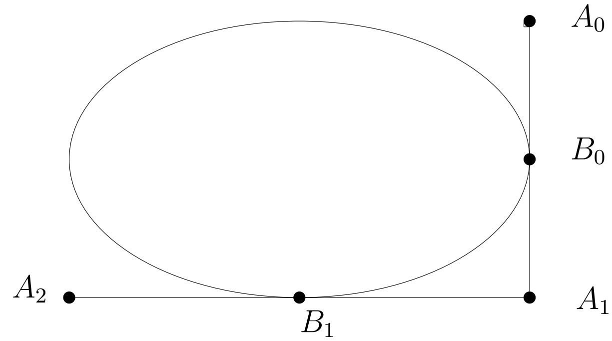

FIGURE 4. Outer billiard

We now describe a construction of Boyland [3] 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 $R_nA_n = f^n_{\Gamma_0}A_0$ 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.

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

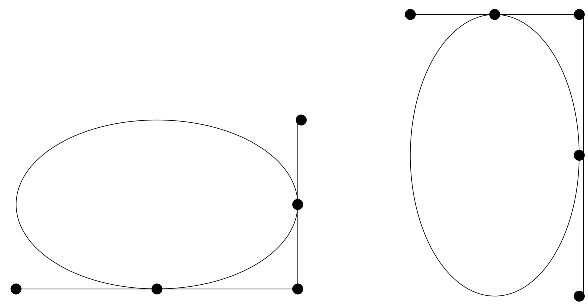

FIGURE 5. Outer billiards and Impact Oscillators

$\max f(t)$ 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

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

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 [14]. (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 decelerates. 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

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

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

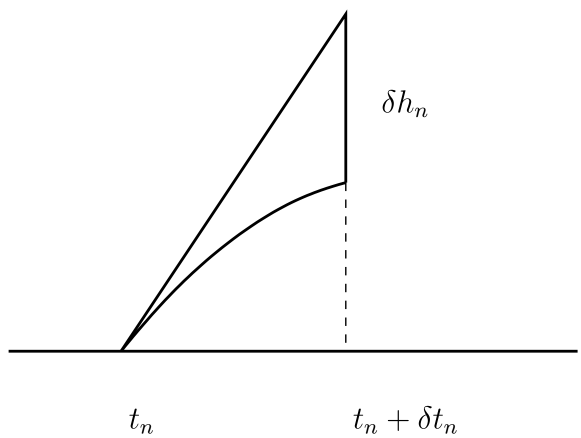

FIGURE 6. Derivative of pingpong map.

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

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

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 $dv \wedge dx$.

2. NORMAL FORMS.

2.1. **Smooth maps close to identity.** Here we discuss the behaviour of highly energetic particles using the methods of averaging 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

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

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 induction. 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 $c_k = \int_0^1 \hat{b}(s)ds$. $\square$

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 joining 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.

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

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

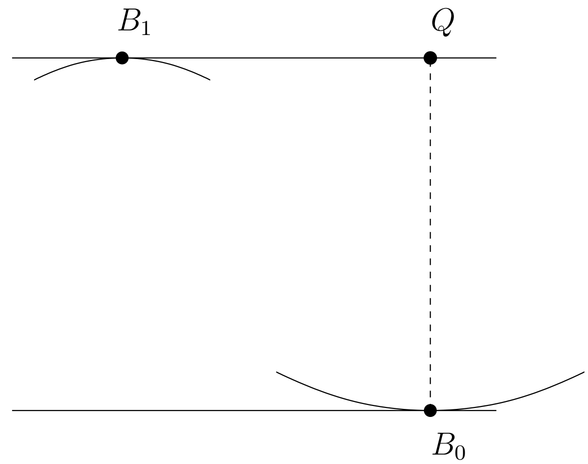

FIGURE 7. Derivative of the support function.

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 sectoral 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 curves 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

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

Thus for each $x \in \Gamma$ and for all $e \in \Gamma^*$ we have $|(x, e)| \leq 1$ 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.

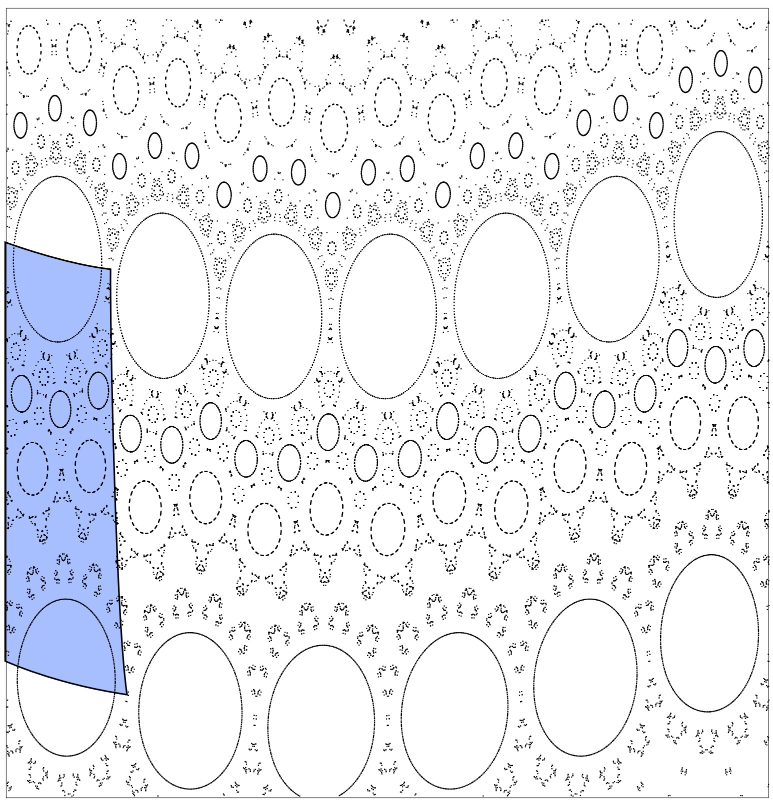

FIGURE 8. Large velocity phase portrait of piecewise smooth pingpong looks similar for different values of time so it makes sense to consider the first return map to a neighbourhood of the singularity.

We assume that the cylinder is divided into a finite union of sectors S_j so that our map is C^∞ in $\mathrm{Int}(S_j)$, has C^∞ extension 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

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

reduces to

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

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 adiabatic 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 orbit 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 continuity domain they are polynomials 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 cancellations. 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. □

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 integrable 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 [13] 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 [10]

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

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

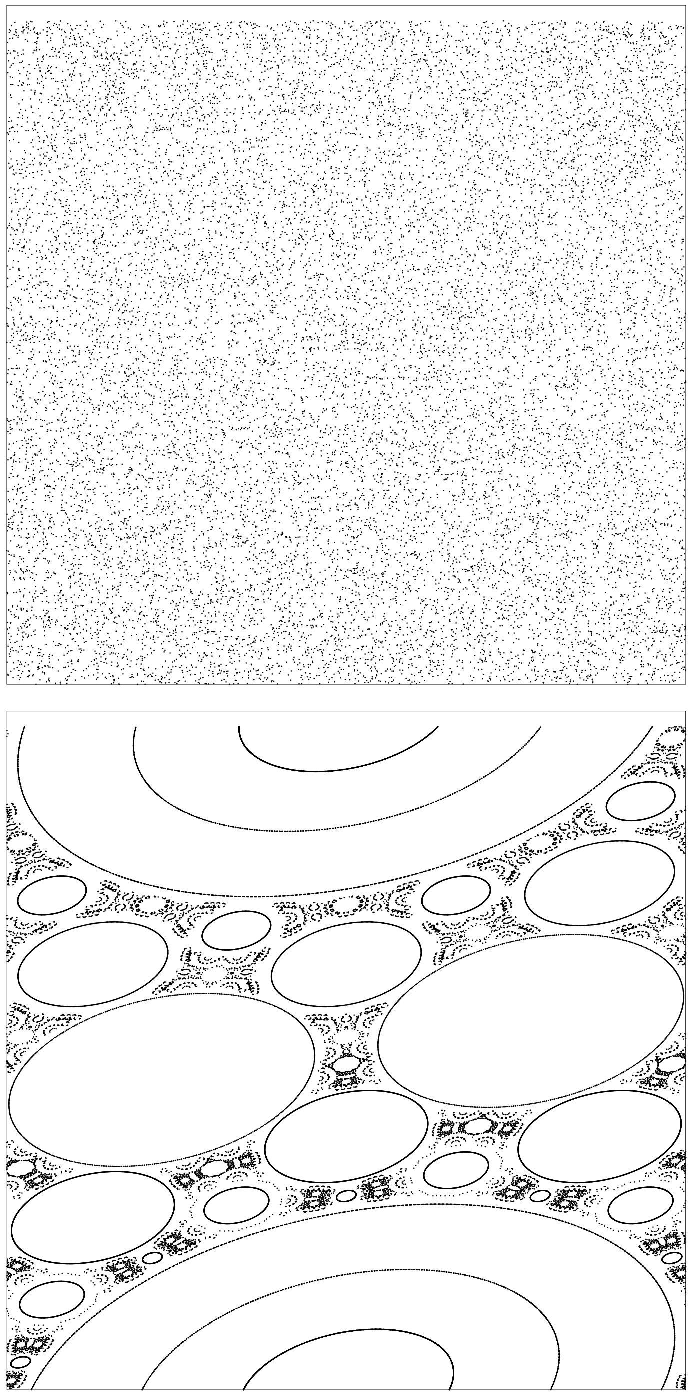

FIGURE 9. Dynamics of the first return map. Top: hyperbolic case. Bottom: elliptic case.

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 $a_c \approx -2.77927\ldots$

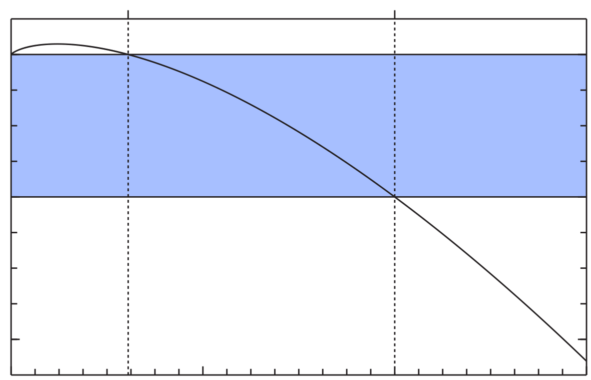

FIGURE 10. $\Delta(a)$ for piecewise linear wall velocity

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. [13] *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* $|z_N| \le \epsilon$*, then for every* $n \in [N, N+\sqrt{N}]$

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

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 $\alpha+\beta\leq 2$ 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 $\Delta\in(0,4)$?

2.5. **Birkhoff normal form.** Here we discuss the normal form of an area preserving difffeomorphism near a periodic point.

Consider an area preserving map f of $\mathbb{R}^2$ which has an elliptic fixed point p such that in suitable complex coordinates z near p our map has the following form

$$f(z)=e^{2\pi i\alpha}z+\mathcal{O}(z^2).$$

Lemma 2.11. *Suppose that $e^{2\pi ik\alpha} \neq 1$ for $k = 1, 2, \dots 2s$. Then there exists a local diffeomorphism h such that $h \circ f \circ h^{-1}$ has form*

$$r_{n+1} = r_n + \mathcal{O}(r_n^{2s}), \quad \phi_{n+1} = \phi_n + \alpha + \sum_{j=1}^{s-1} c_j r_n^{2j} + \mathcal{O}(r_n^{2s}).$$

Proof. If suffices to prove that one can reduce f to the following form

$$r_{n+1} = r_n + \sum_{j=1}^{s-1} d_j r_n^{2j+1} + \mathcal{O}(r_n^{2s}), \quad \phi_{n+1} = \phi_n + \alpha + \sum_{j=1}^{s-1} c_j r_n^{2j} + \mathcal{O}(r_n^{2s})$$

since than area preservation would imply that $d_j \equiv 0$ since otherwise the orbits will go either further away from 0 or closer to 0 with each iteration contradicting area preservation.

So we would like to conjugate f to

$$g(z) = e^{2\pi i\alpha} z + \sum_{k=2}^{2s-1} G_k(z) + \mathcal{O}(z^{2s})$$

by the map

$$h(z) = z + \sum_{k=2}^{2s-1} H_k(z) + \mathcal{O}(z^{2s})$$

where G_k and H_k are polynomials of degree k in $z, \bar{z}$. Expanding the equation $h \circ f = g \circ h$ into Taylor series we get

$$H_k(e^{2\pi i\alpha} z, e^{-2\pi i\alpha} \bar{z}) + A_k = e^{2\pi i\alpha} H_k(z, \bar{z}) + B_k + G(z, \bar{z})$$

where A_k and B_k denote the terms which are determined by the lower order coeffcients of H and G respectively. If

$$H_k = \sum_{l_1+l_2=k} h_{l_1,l_2} z^{l_1} \bar{z}^{l_2}, \quad G_k = \sum_{l_1+l_2=k} g_{l_1,l_2} z^{l_1} \bar{z}^{l_2}$$

when we get

$$h_{l_1,l_2} \left[e^{2\pi i\alpha(l_1-l_2)} - e^{2\pi i\alpha} \right] = g_{l_1,l_2} + c_{l_1,l_2}$$

where c_{l_1,l_2} are determined by A_k and B_k. Hence if $l_1 - l_2 \neq 1$ then we can choose $g_{l_1,l_2} = 0$ and take

$$h_{l_1,l_2} = \left[e^{2\pi i\alpha(l_1-l_2)} - e^{2\pi i\alpha} \right]^{-1} c_{l_1,l_2}.$$

On the other hand if $l_1 - l_2 = 1$ then we are forced to take $g_{l_1,l_2} = -c_{l_1,l_2}$. Hence f is conjugated to $g(z) = e^{2\pi i\alpha} z\gamma(r^2) + \mathcal{O}(r^{2s})$. Writing $\gamma(u) = a(u)e^{2\pi ib(u)}$ we obtain the result. □

3. Applications of KAM theory.

3.1. **Introduction.** In this section we review some applications of Kolmogorov-Arnold-Moser theory to bouncing balls. Our overview will be brief since this material is pretty standard and can be found in several textbooks. However I would like to emphasize that the brvity of this section does not reflect the importance of this material. In fact, KAM theory is the prime tool for showing lack of acceleration and/or ergodicity. The rest of the course will be devoted to discussing a reletively small class of systems where KAM is not applicable with the goal of developing the tools to handle such systems.

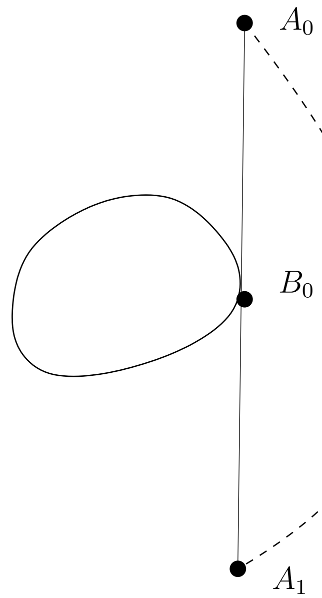

Figure 11. The outer caustic has the property that any tangent line to the table cuts off the segment which is divided into two equal pieces by the tangentcy point.

For one and a half degrees of freedom systems invariant curves provide an easy obstruction to transitivity since the orbit can not pass from one component of $\mathbb{R}^2 - \gamma$ to another. One example where it is easy to visualize the invariant curves is given by outer billialrds. In this case the invariant curves are called *outer caustics*. A curve S is an outer caustic for the outer billiard about a curve Γ if for any tangent line to Γ the points of intersection of that line with S are equidistant from the tangency point. Parametrize Γ by the arclength s and let $\mathcal{A}(s)$ denote the area cut from S by the tangent line emanating from $\Gamma(s)$. We claim that $\mathcal{A}(s)$ does not in fact depend on s. Indeed let $|A_0\hat{A}_0| = \delta s$ then up to higher order terms we have

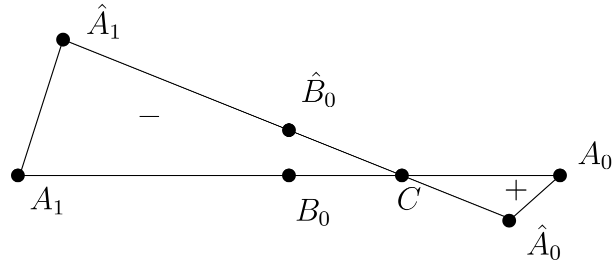

FIGURE 12. Area property of the outer caustics

$$|A_0B_0| \approx |\hat{A}_0\hat{B}_0| \approx |A_0C| \approx |\hat{A}_0C| \approx |A_1C| \approx |\hat{A}_1C|$$

and

$$\angle A_0C\hat{A}_0 \approx \angle A_1C\hat{A}_1 \approx \kappa\delta s$$

there κ denotes the curvature at the tangency point. Accordingly

$$\frac{\partial \mathcal{A}}{\partial s} = |A_0B_0|^2\kappa - |A_1B_0|^2\kappa = 0.$$

Therefore given a curve S we can easily given a curve S we can easily construct a curve Γ such that S is an outer caustic for Γ by fixing a paprameter a, considering all segments which cut area a from S and taking the midpoints of those segments. It is more difficult to find outer caustics for a given billiard table Γ. For this we need a full strength of the KAM theory. In particular we need to assume that Γ is sufficiently smooth. We saw in Section 2 that some smoothness is needed but the exact trashhold is currently unknown.

Problem 3.1. Suppose that Γ is piecewise smooth and the first k derivatives at the break points coincede. For which k must Γ have invariant curves near infinity?

3.2. **Theory.** Two classical results about invariant curves are Twist Theorem and Small Twist Theorem of Moser.

Proposition 3.2 (Moser Small Twist Theorem)**.** *Let $Q : \mathbb{R}_+ \to \mathbb{R}_+$ be a C^5-function. Then for any numbers a, b such that $Q'(r) \neq 0$ for $r \in [a, b]$ for any K there is ε_0 such that if F_ε are exact mappings of the annulus $\mathbb{R}_+ \times \mathbb{S}^1$ of the form*

$$F_\varepsilon(r, \phi) = (r + \varepsilon^{1+\delta}P(r, \phi), \phi + \alpha + \varepsilon Q(r) + \varepsilon^{1+\delta}R(r, \phi))$$

where

$$||P||_{C^5([a,b]\times\mathbb{S}^1)} \leq K, \quad ||R||_{C^5([a,b]\times\mathbb{S}^1)} \leq K$$

then for $\varepsilon \leq \varepsilon_0$ F_ε has (many) invariant curve(s) separating $[a, b] \times \mathbb{S}^1$ into two parts. In fact, the set of invariant curves has positive measure.

Proposition 3.3 (Moser Invariant Curve Theorem)**.** *Let $Q : \mathbb{R}_+ \to \mathbb{R}_+$ be a C^5-function. Then for any numbers a, b such that $Q'(r) \neq 0$ for $r \in [a, b]$ there is ε_0 such that if F is an exact mapping of the annulus $\mathbb{R}_+ \times \mathbb{S}^1$ of the form*

$$F(r, \phi) = (r + P(r, \phi), \phi + Q(r) + R(r, \phi))$$

where

$$||P||_{C^5([a,b]\times\mathbb{S}^1)} \leq \varepsilon_0, \quad ||R||_{C^5([a,b]\times\mathbb{S}^1)} \leq \varepsilon_0$$

then F has (many) invariant curve(s) separating $[a, b] \times \mathbb{S}^1$ into two parts. In fact, the set of invariant curves has positive measure.

A classical application of KAM theory is stability of nonresonant elliptic periodic points.

Lemma 3.4. *Suppose that p is an elliptic periodic point of an area preserving diffeomorphism f with multiplier $e^{2\pi\alpha}$ such that $e^{2\pi k\alpha} \neq 1$ for $|k| \leq 4$ and such that the Birkhoff normal form is non-degenerate. Then f has a positive measure set of invariant curves near p.*

Proof. This follows from Lemma 2.11 and Proposition 3.3. □

3.3. **Applications.** Here we describe some applications of the KAM theory to bouncing balls.

(I) Pingpongs.

Corollary 3.5. *Consider Fermi-Ulam pingpong with wall motion of class C^6. Then there are KAM curves for arbitrary high velocities. Accordingly all orbits are bounded.*

Proof. This follows from Proposition 3.2 and the normal form obtained in Section 2. □

Corollary 3.6. *Consider pingpongs where the wall motion has one discontinuity and the system is elliptic at infinity. Then there is a constant C such that for all sufficiently large v there is a positive measure set of orbits such that*

$$\frac{\bar{v}}{C} \leq v(t) \leq C\bar{v}.$$

Proof. The map (2.4) has periodic orbit $(\frac{1}{2}, 0)$. The non-degeneracy of the Birghoff normal form is checked in [10]. For the orbits constructed with the help of Lemma 3.4 the adiabatic invariant $l(t)v(t)$ will change little so the oscilations of $\ln v(t)$ are of constant order. □

Problem 3.7. Is Corollary 3.6 valid for systems with several velocity jumps?

We shall see later that the result of Corollary 3.6 is false for pingpongs which are hyperbolic at infinity. In that case the system may even be ergodic so that almost every orbit is dense.

(II) Balls in a potential. Consider a moving in a potential $U(x) = gx^\alpha$ and colliding elastically with infinitely heave wall. Suppose that f is C^6 and periodic.

Corollary 3.8. *If $\alpha > 1$ and $\alpha \neq 2$ then there are KAM curves for arbitrary large velocities. In particular, all orbits are bounded.*

Proof. 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 $\sigma = \frac{2}{\alpha} - 1$.

Consider first the case $\alpha > 2$. Take $\bar{v} \gg 1$ and suppose that $v_0 \sim \bar{v}$. Rescaling $u_n = \frac{v_n}{\bar{v}}$ we get

$$t_{n+1} \approx t_n + c\bar{v}^\sigma u_n^\sigma, \quad u_{n+1} - u_n = \frac{2\dot{f}(t_n)}{\bar{v}}.$$

Since $\sigma > 1$ the change of u is much smaller than the change of t and so we can use Proposition 3.2.

Next, consider the case $1 < \alpha < 2$. Set $z_n = \frac{v_n - v_0}{v^{\sigma-1}}$. The map takes form

$$t_{n+1} - t_n \approx \alpha_0 + Kz_n + \dots, \quad z_{n+1} - z_n = \frac{2\dot{f}(t_{n+1})}{v^{\sigma-1}}.$$

Therefore the statement follows from Proposition 3.3. □

One can ask what happens for other values of α. Surprisingly for $\alpha = 1$ one can have a positive measure set of escaping orbits. The proof of that given by Pustylnikov uses KAM theory. It relies on the following non stationary extension of stability of elliptic periodic orbits.

Theorem 3.9. *Let $f_n(z)$ be a family of real analytic area preserving maps defined near the origin and converging to a limiting value f so that $\sum_n ||f_n - f|| \leq \infty$. Suppose that 0 is an elliptic fixed point for f with multiplier $e^{2\pi\alpha}$ satisfying $e^{2\pi ik\alpha} \neq 0$ for $k \in \{1, 2, 3, 4\}$. and that the corresponding Birkhoff normal form is nondegenerate. Consider a recurrence $z_n = f_n z_{n-1}$. Then there is a positive measure set of initial conditions z_0 such that z_n is bounded.*

The proof of Theorem 3.9 proceeds along the line of the proof of Lemma 3.4. We refer the reader to [20] for details.

We now show how this theorem can be used to construct escaping orbits. Consider a two parameter family of SWA

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

Here $Af(t)$ is the height of the ball at time t and g is the gravity strengh.

Consider the orbits where the rocket always hit the ball at the same height. Thus $t_{n+1} = t_n \mod 1$, $\frac{2v_n}{g} = l$. Next

$$df = \begin{pmatrix} 1 & 0 \\ 2A\ddot{f} & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{2}{g} \\ 0 & 1 \end{pmatrix}$$

so that $\mathrm{Tr}(df) = 2 + \frac{4A\ddot{f}(t)}{g}$. Projecting our orbit to the torus we obtain a fixed point which is elliptic provided that

$$-1 < \frac{A\ddot{f}(t)}{g} < 0. \tag{3.1}$$

The original orbit on the cylinder is accelerating if

$$\dot{f}(t) > 0 \tag{3.2}$$

Next, if we have an accelerating orbit for the SWA, Theorem 3.9 allows to infer stability of the original system. Let us show that we can find the periodic point in our two parameter family of the toral maps satisfying (3.1) and (3.2). Indeed, the periodicity condition amounts to

$$\dot{f}(t_n) = \frac{lg}{2A} \tag{3.3}$$

If $A, g \gg 1$ then we can arrange $t_n \approx \bar{t}$ for any $\bar{t}$ such that $\dot{f}(\bar{t}) > 0$. Next, in view of (3.3) condition (3.1) amounts to $-\frac{2}{l} < \frac{\ddot{f}(\bar{t})}{\dot{f}(\bar{t})} < 0$. Take an interval (t_1, t_2) such that $\dot{f}(t_1) = \dot{f}(t_2) = 0$ and $\dot{f}(\bar{t}) > 0$ for $\bar{t} \in (t_1, t_2)$. Since $\int_{t_1}^{t_2} \ddot{f}(\bar{t})d\bar{t} = 0$ the second derivative changes sign on (t_1, t_2) and so we can find $\bar{t}$ satisfying $-1 < \frac{\ddot{f}(\bar{t})}{\dot{f}(\bar{t})} < 0$ as needed.

The case $\alpha = 2$ was investigated by Ortega ([19]). He showed that if the periods of the wall and the string are incommensurable then the averaging prevails and there are KAM curves. In the commensuarable case both KAM curves and positve measure of escaping sets are possible. For example, in the case of outer billiards all orbits are bounded.

Corollary 3.10. *If Γ is C^6 and strictly convex then all orbits are bounded.*

Proof. The result follows from the normal form obtained in Section 2 and Proposition 3.2. □

Problem 3.11. Show that the result is not correct if Γ has points with zero curvature.

Finally in case $\alpha > 1$ one can always construct a Cantor set of escaping orbits. In fact, it is shown in [8] that HD($\mathcal{E}$)=2.

Conjecture 3.12. *If* $\alpha < 1$ *then* $\text{mes}(\mathcal{E}) = 0$.

We will see in Section 6 that this conjecture is true for very week potentials, that is, for $\alpha \ll 1$.

4. Recurrence.

4.1. **Poincare Recurrence Theorem.** In this section we describe applications of ergodic theory to the dynamics of bouncing balls.

As it was mentioned before dynamical systems theory strives to describe a long time behavior of a given system. In particular, one can ask if $(q(t), x(t))$ come close to its initial values for arbitrary large t. A general result in this direction is the Poincare Recurrence Theorem given below.

Theorem 4.1. *Let T be a transformation of a space X preserving a finite measure μ. Then for each set A almost all points from A returns to A in the future.*

Proof. Let $B = \{x \in A : T^n x \notin A \quad \forall n > 0\}$. 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 $m \to m+1$ on $\mathbb{Z}$. One can however show that the above map is an only obstacle to Poincare recurrence. Namely, let T be a transformation of a metric space X preserving an infinite measure μ such that measure of any ball is finite.

Theorem 4.2. *X can be represented as a disjoint union $X = \mathcal{C} \sqcup \mathcal{D}$ where*

(i) $\mathcal{D} = \cup_{n \in \mathbb{Z}} T^n B$ and B is wondering in the sense that $T^n B \cap B = \emptyset$ for $n \neq 0$;

(ii) $\mathcal{C}$ 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**.

Introduce coordinates (b, m) on $\mathcal{D}$ where $b \in B$, $m \in \mathbb{Z}$ and the point (b, m) corresponds to $T^m b$. Then the map takes form

$$(b, m) \to (b, m + 1),$$

that is the dynamics on $\mathcal{D}$ is isomorphic to the shift on the integers.

Proof. Pick a reference point a. Let

$$B_1 = \{x \in B(a, 1) : T^n x \notin B(a, 1) \text{ for } n > 0\}.$$

For $k > 1$ let

$$B_k = \{x \in B(a, k) - \left(\cup_{j=1}^{k-1} \cup_{m=-\infty}^{\infty} T^m B_j\right) : T^n x \notin B(a, k) \text{ for } n > 0\}.$$

Let $B = \cup_{k=1}^{\infty} B_k$. Note that the orbits of B_k for different k are disjoint by construction. Next we claim that $T^n B_k \cap B_k = \emptyset$ for $n \neq 0$. Indeed if $x \in B_k$ and $T^n x \in B_k$ then $T^n x \in B(a, k)$ by the definition of B_k. Thus n can not be positive by the definition of B_k. Also n can not be negative since in that case $T^{-n}(T^n x) = x \in B(a, k)$ contradicting the definition of B_k. Thus $n = 0$ as claimed.

Next let $\mathcal{D} = \displaystyle\bigcup_{n=-\infty}^{\infty} T^n B$ and $\mathcal{C} = X - \mathcal{D}$. Let $A \subset \mathcal{C}$. Then $A = \displaystyle\bigcup_{k=1}^{\infty} A_k$ where $A_k = A \cap B(a, k)$. Note that by definition of B_k the first return map R_k is well defined on $B(a, k) - B_k$. Applying Poincare Recurrence Theorem to R_k we see that almost all points from A_k visit A_k, so A satisfies Poincare Recurrence Theorem as claimed. □

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 4.3. *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. $\square$

Lemma 4.3 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 4.3 is very easy there is no general recipe 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 4.4. $\text{mes}(\mathcal{E}) = 0$ *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 4.3. $\square$

Problem 4.5. Do above systems have escaping orbits? In fact even the existence of unbounded orbits is unknown.

4.2. **Background from ergodic theory.** To proceed further we need to recall some facts from ergodic theory. Let $T : X \to X$ be a map preserving a measure μ. A set A is called **invariant** if $T^{-1}A = A$ and it is called **essentially invariant** if $\mu(T^{-1}A \Delta A) = 0$. 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.

Lemma 4.6. *The following are equivalent:*
(a) T is ergodic;
(b) If B is an essentially invariant set then $\mu(B) = 0$ or $\mu(B^c) = 0$;
(c) If A is a set of positive measure then $\mu(\cup_{n=1}^{\infty} T^n A) = 1$;
(d) If A and B are sets of positive measure then there exists $n > 0$ such that $\mu(T^n A \cup B) > 0$,
(e) If $\phi : X \to \mathbb{R}$ is a measurable function such that $\phi(Tx) = \phi(x)$ almost everywhere when there exists a constant c such that $\phi = c$ almost everywhere.

Proof. $(a) \Rightarrow (b)$. Let $C_N = \cup_{n=N}^{\infty} T^{-n} B$. Then since $\mu(T^{-n_1} B \Delta T^{-n_2} B)$ it follows that $\mu(C_N) = \mu(B)$ for all N. On the other hand C_N are nested. Let $C = \cup_{N=0}^{\infty} C_N$. Then $\mu(C) = \mu(B)$. Since $T^{-1} C_N = C_{N+1}$ we have $T^{-1} C = C$ so $\mu(C) = 0$ or $\mu(C) = 1$ and hence $\mu(B) = 0$ or $\mu(B) = 1$.

$(b) \Rightarrow (c)$. Let $B = \bigcup_n T^n A$. Then $T^{-1} B \supset B$, so by measure preservation, B is essentially invariant and since $\mu(B) \geq \mu(A)$ it follows that $\mu(B) = 1$.

Exercise 4.7. Prove that $(c) \Rightarrow (d)$ and $(d) \Rightarrow (a)$.

$(b) \Rightarrow (e)$. Suppose that (b) holds and let ϕ be a T-invariant function. Then for each t we have $\mu(x : \phi(x) > t) = 0$ or 1. Let

$$c = \sup(t : \mu(x : \phi(x) > t) = 0).$$

Then for each ε

$$\mu(x : c - \varepsilon < \phi(x) < c + \varepsilon) = 1$$

and so $\phi(x) = c$ almost everywhere.

$(e) \Rightarrow (b)$. If B is an essentially invariant set let $\phi = 1_B$. Then $\phi = c$ almost everywhere. Clearly c is either 0 or 1. □

In this and the following sections we will use the following results.

Theorem 4.8. (Ergodic Theorem) *(a) If $\phi \in L^1(\mu)$ then for almost every x the following limits exist*

$$\phi^{\pm}(x) = \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \phi(T^{\pm n} x).$$

Moreover for almost every x, $\phi^+(x) = \phi^-(x) := \bar{\phi}(x)$ and

$$\int \bar{\phi}(x) d\mu(x) = \int \phi(x) d\mu.$$

(b) If T is ergodic then $\phi(x) = \int \phi(x) d\mu(x)$ almost everywhere.

Theorem 4.9. (Maximal Ergodic Theorem) *Let*

$$E_\alpha = \{x : \sup_{N\geq 1} \frac{1}{N} \sum_{n=0}^{N-1} \phi(x) > \alpha\}.$$

Then

$$\alpha\mu(E_\alpha) \leq ||\phi||_{L^1}.$$

Lemma 4.10. (Rokhlin's Lemma) *If $T : X \to X$ is an aperiodic transformation then for each n, ε there is a set B such that $B, TB, \dots, T^{n-1}B$ are disjoint and $\mu\left(X - \bigcup_{j=0}^{n-1} T^j B\right) \leq \varepsilon$.*

4.3. **Ergodicity and recurrence.** Next we 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 $d\nu = d\mu dx$. Denote $\tau_m(x, y) = (x, y + m)$.

Lemma 4.11. *(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 $|y| \leq 1$ visit $\{|y| \leq 2\}$ only finitely many times. Indeed, the set

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

is T_Φ invariant and all points from B visit $A = \{|y| \leq 2|\}$ so if $\mu(B) > 0$ T_Φ would have a nontrivial recurrent part by Lemma 4.3.

Hence for almost all x the set $M_x = \{n : |\Phi_n| \leq 1\}$ is finite where $\Phi_n(x) = \sum_{j=0}^{n-1} \Phi(T^j x)$. Let $A_N = \{x : \text{Card}(M_x) \leq N\}$. 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 $\mathrm{Card}(\mathcal{Y}_n(x)) \geq \frac{n}{2}$ and for each $\bar{y} \in \mathcal{Y}_n(x)$ we have

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

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

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

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

As an application of Lemma 4.11 consider SWA to an impact oscillator with

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

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 if T is irrational.

On the other hand if $\bar{h} \neq 0$ then Lemma 4.11 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. We assume that S is defined on a subset $\Omega \subset X \times \mathbb{R}$ given by $y \geq h(x)$. 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 4.12. *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.

Proof. Let $\bar{Y} = X \times [0, K]$ where K is the constant from condition (iv). By Lemma 4.11 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 4.3. □

Lemma 4.12 shows recurrence of impact 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.

4.4. **Proof of the Maximal Ergodic Theorem.** We need the following result called **maximal inequality**. Give a function $\psi \in L^1(\mu)$ define

$$\tilde{S}_0 = 0, \text{ and for } k > 0,\ \tilde{S}_k = \sum_{n=1}^{k} \psi \circ T^n, \quad \psi_N^* = \max_{0 \le k \le N} \tilde{S}_k, \quad P_N = \{x : \psi_N^* > 0\}.$$

Lemma 4.13. *For all positve N*

$$\int_{P_N} \psi(x) d\mu(x) > 0.$$

Proof. We have that

$$\psi_N^*(Tx) = \max_{1 \le k \le N+1} \sum_{n=2}^{k} \psi(T^n x).$$

Thus

$$\psi_N^*(Tx) + \psi(x) = \max_{1 \le k \le N+1} \sum_{n=1}^{k} \psi(T^n x).$$

Since on P_N $\psi_N^* = \max_{1 \le k \le N} \tilde{S}_k$ it follows that on P_N

$$\psi_N^*(Tx) + \psi(x) \ge \psi_N^*(x).$$

Thus

$$\int_{P_N} \psi(x)d\mu(x) \geq \int_{P_N} \psi_N^*(x)d\mu(x) - \int_{P_N^*} \psi_N^*(x)d\mu(x)$$

$$= \int_{P_N} \psi_N^*(x)d\mu(x) - \int_X \psi_N^*(x)d\mu(x)$$

$$\geq \int_X \psi_N^*(x)d\mu(x) - \int_X \psi_N^*(x)d\mu(x) = 0$$

where the second line follows since $\psi_N^* = 0$ on P_N^c and the third line follows since $\phi_N^* \circ T$ is non negative. □

Proof of Theorem 4.9. Let $\psi = \phi - \alpha$. Then

$$P_N = \{x : \max_{k \leq N} \frac{1}{k} \sum_{n=1}^{k} \phi(T^n x) > \alpha\}.$$

Taking $N \to \infty$ in Lemma 4.13 we obtain the statement required. □

4.5. **Ergodic Theorems for L^2-functions.** Let

$$\mathcal{I} = \{\phi \in L^2(\mu) : \phi(Tx) = \phi(x)\}.$$

By **coboundary** we mean a function of the form $\psi(x) - \psi(Tx)$ for an L^2 function ψ. Let $\mathcal{B}$ denote the closure of the space of coboundaries.

Lemma 4.14. $\mathcal{B}^\perp = \mathcal{I}$.

Proof. If $\psi \in \mathcal{B}^\perp$ then $\langle \psi(x), \psi(x) - \psi(Tx) \rangle = 0$. Accordingly

$$|\psi(x)\psi(Tx)d\mu(x)| = ||\psi||_{L^2} = \sqrt{||\psi||_{L^2}||\psi \circ T||_{L^2}}.$$

By Cauchy Schwartz inequality this is only possible if $\psi(Tx) = c\psi(x)$. Since $||\psi \circ T||_{L^2} = ||\psi||_{L^2}$ $c = \pm 1$. Now it is evident that

$$\langle \psi(x), \psi(x) - \psi(Tx) \rangle = 0$$

iff $c = 1$, that is $\psi \in \mathcal{I}$. □

Proposition 4.15. *If $\phi \in L^2(\mu)$ then*

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(T^n x) \to \bar{\phi} := \pi_{\mathcal{I}} \phi \text{ as } N \to \pm\infty$$

almost everywhere and in L^2.

Proof. The statement is obvious if $\phi \in \mathcal{I}$, so we amy assume that $\phi \in \mathcal{B}$. If $\phi(x) = \psi(x) - \psi(Tx)$ then

$$\frac{1}{N}\sum_{n=0}^{N-1}\phi(T^n x) = \frac{\psi(x) - \psi(T^n x)}{N} \to 0 \text{ in } L^2.$$

Also by Chebyshev inequlity

$$\mu(\psi(T^n x) > \varepsilon N) \leq \frac{||\psi||_{L^2}}{\varepsilon^2 N^2}$$

so $\frac{\psi(x)-\psi(T^n x)}{N} \to 0$ almost everywhere as well. For general $\phi \in \mathcal{B}$, given ε, we can find $\tilde{\phi}, \psi$ such that

$$\phi(x) = \tilde{\phi}(x) + \psi(x) - \psi(T^n x)$$

and $||\tilde{\phi}||_{L^2} \leq \varepsilon$. Then

$$\hat{\phi}^{\pm}(x) = \limsup_{N\to\pm\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\phi(T^n x)\right| = \limsup_{N\to\pm\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\tilde{\phi}(T^n x)\right|.$$

Thus by Maximal Ergodic Theorem

$$\mu(\phi^{\pm}(x) \geq \delta) \leq \frac{||\tilde{\phi}||_{L^1}}{\delta} \leq \frac{||\tilde{\phi}||_{L^2}}{\delta} = \frac{\varepsilon}{\delta}.$$

Since ε is arbitrary $\hat{\phi}^{\pm} = 0$ so

$$\frac{1}{N}\sum_{n=0}^{N-1}\phi(T^n x) \to 0$$

almost everywhere. The argument for L^2-convergences is similar. □

Exercise 4.16. Prove Theorem 4.8 for L^1 functions.

Hint. Take $\phi^{(n)} \in L^2$ such that $||\phi^{(n)} - \phi||_{L^1} \leq \frac{1}{n}$ and show that $\bar{\phi}^{(n)}$ form a Cauchy sequence. Thus there is a function $\bar{\phi}$ such that $\bar{\phi}^{(n)} \to \bar{\phi}$ in L^1. Show that

$$\frac{1}{N}\sum_{n=0}^{N-1}\phi(T^n x) \to \bar{\phi}(x)$$

almost everywhere.

Remark 4.17. Ergodic Theorem for L^2 functions will be sufficient for all applications given in these lectures.

Corollary 4.18. *If $\bar{\phi}$ is constant almost everywhere for a L^2-dense set of functions then T is ergodic.*

Proof. If
$$\pi_{\mathcal{I}}\phi = \left(\int \phi(x)d\mu(x)\right) 1$$
for a dense set of functions then by continuity it holds for all functions. Thus $\mathcal{I}$ consists only of constants and hence T is ergodic. □

4.6. **Ergodic decomposition.** Let T be a map of a separable metric space X preserving a measure μ. Let $\mathcal{I}$ be the algebra of T invariant sets. Define a family of measures μ_x by condition $\mu_x(\phi) = \mathbb{E}(\phi|\mathcal{I})(x)$.

Proposition 4.19.

(a) $$\int \mu_x(\phi)d\mu(x) = \int \phi(x)d\mu(x).$$

(b) μ_x is T-invariant for almost every x.
(c) μ_x is ergodic for almost every x.

Proof. (a) follows from the Law of Total Expectation. (b) and (c) follow from the formula $\mu_x(\phi) = [\pi_{\mathcal{I}}(\phi)](x)$ and Proposition 4.15. □

4.7. **Proof of Rokhlin Lemma.**

Proof of Lemma 4.10. We first prove the result in case T is ergodic. Choose a set A with $\mu(A) < \varepsilon/n$. Let $R(x)$ be the first return time to A. Let
$$B = \cup_{j\in A, j<R(x)/n} T^j x.$$
Clearly $B, TB \dots T^{n-1}B$ are disjoint. On the other hand denoting $A_m = \{x \in A : R(x) = m\}$ we get
$$X - \cup_{j=0}^{n-1} T^j B \subset \cup_{m=1}^{\infty} cup_{k=1}^{n} T^{m-k} A_m$$
so
$$\mu\left(X - \cup_{j=0}^{n-1} T^j B\right) \leq \sum_{m=1}^{\infty}\sum_{k=1}^{n} \mu(T^{m-k}A_m) \leq \sum_{m=1}^{\infty} n\mu(A_m) < n\varepsilon$$
completing the proof for ergodic transformations.

In the non-ergodic case the same argument works provided that we can find A with
$$\mu\left(x : \mu_x(A) > \frac{\varepsilon}{2n}\right) < \frac{\varepsilon}{2}. \tag{4.1}$$
Take A with $\mu(A) < \dfrac{\varepsilon^2}{4n}$. Then (4.1) follows from Proposition 4.19(a) and Markov inequality. □

5. Statistical properties of hyperbolic sawtooth maps.

5.1. **The statement.** We saw in Section 4.3 that ergodicity of hyperbolic sawtooth maps implies the recurrence of a large class of Fermi-Ulam pingpongs in case velocity has one discontinuity. The required ergodicity is established in this section. In fact, following Chernov [6] we consider a wider class of maps. Let T be a piecewise linear automorphism of $\mathbb{T}^2$. Let S_+ and S_- denote the discontinuity 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. [6] *T is ergodic.*

In fact, we derive stronger statistical properites of our map T. These results are not needed to establish the recurrence of pingpongs but the techniques introduced here could be used to study several classes of bouncing ball systems.

Recall that a map T preserving a probability measure μ is called **mixing** if for any L^2 functions ϕ_1, ϕ_2,

$$\rho_{\phi_1,\phi_2}(n) \to 0 \text{ as } n \to \infty \tag{5.1}$$

where

$$\rho_{\phi_1,\phi_2}(n) = \int \phi_1(x)\phi_2(T^n x)d\mu(x) - \int \phi_1(x)d\mu(x) \int \phi_2(x)d\mu(x).$$

Exercise 5.2. (a) T is mixing if (5.1) holds for a dense set of functions ϕ_1, ϕ_2.

(b) T is mixing if for each pair of measurable sets A_1, A_2

$$\mu(A_1 \cap T^{-n}A_2) \to \mu(A_1)\mu(A_2).$$

(That is, it suffices to check (5.1) for $\phi_i = 1_{A_i}$.)

Thus mixing gives, in particular, that for any measurable set A we have $\mu(A \cap T^{-n}A) \to \mu^2(A)$. Thus if A is invariant then $\mu(A) = \mu^2(A)$ and, hence, mixing implies ergodicity.

In case T is a (piecewise smooth map) of a manifold we say that T is **exponentially mixing** if for some $\alpha > 0$ there are constants $K > 0$ and $\theta < 1$ such that

$$|\rho_{\phi_1,\phi_2}(n)| \leq K\theta^n ||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha}. \tag{5.2}$$

Exercise 5.3. Show that if for some r

$$|\rho_{\phi_1,\phi_2}(n)| \leq \tilde{K}\tilde{\theta}^n ||\phi_1||_{C^r} ||\phi_2||_{C^r}$$

then (5.2) holds. That is, if (5.2) holds for some α then it holds for all α (with K and θ depending on α).

Hint. Approximate C^α functions by C^r functions.

Exercise 5.4. Suppose that T is a smooth linear hyperbolic map of $\mathbb{T}^2$. That, is $Tx = Ax$ mod 1 where $A \in \mathbb{SL}_2(\mathbb{R})$ and $|\text{Tr}(A)| > 2$.

(a) Show that T is exponentially mixing.

(b) Show that given an arbitrary positive sequence $\zeta_n \to 0$ where exists a continuous (but not Holder continuous) function ϕ such that $|\rho_{\phi,\phi}(n)| > \zeta_n$ for infinitely many n.

Hint. Consider the Fourier series of ϕ.

Theorem 5.5. *Under the assumptions (i) and (ii) above T is exponentially mixing.*

By the foregoing discussion exponential mixing implies mixing which, in turn, implies ergodicity. Thus Theorem 5.1 follows from Theorem 5.5. However our proof of Theorem 5.5 relies on Theorem 5.1 so we begin with the proof of Theorem 5.1.

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 $fx = Ax$ mod 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 $x \in \mathcal{R}_0$ and $\Phi \in C(\mathbb{T}^2)$ let $\bar{\Phi}(x)$ denote the common value of $\Phi^+(x)$ and $\Phi^-(x)$. By Corollary 4.18 it suffices to show that $\bar{\Phi}(x)$ is constant almost everywhere for every continuous function Φ.

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 then $\Gamma_\rho = \mathbb{T}^2$ and by Fubini theorem $\mathrm{mes}(\Gamma_\rho - \tilde{\Gamma}_\rho) = 0$ 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 restrictive. 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 [7]. A slightly different argument can be found in [17].

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 5.6.

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

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

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

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

Lemma 5.7. *Pick y, δ and k such that $d(T^j\tilde{W}^u(y), S_-) \geq \varepsilon$ for $j = 0 \dots k$. Then*

$$\text{mes}\{x \in \tilde{W}^u_\delta(y) : r_s(x) \leq \varepsilon\} \leq C\theta^k\varepsilon.$$

A similar statement holds with s and u interchanged.

We first show how Lemma 5.7 can be used to derive Theorem 5.1 and then present the proof of the lemma.

Pick k such that

$$C\theta^k < 0.001. \tag{5.3}$$

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 unions 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 segments 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 $\left(\frac{0.1\mathbb{Z}}{N}\right)^2 \cap U$.

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

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

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

Note that if $z_1, z_2 \in S$ then by Lemma 5.7 and (5.3), $\Sigma(z_j) \cap S$ has measure at least $0.999\text{mes}(S)$ and by the Hopf argument almost all points in $\Sigma(z_j)$ are equivalent to z_j. Also by Lemma 5.7 the set of typical points in S has measure at least $0.998\text{mes}(S)$. 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 equivalent. Therefore all typical points in all squares in $\text{Int}(M)$ are equivalent. On the other hand by Lemma 5.6 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 proved.

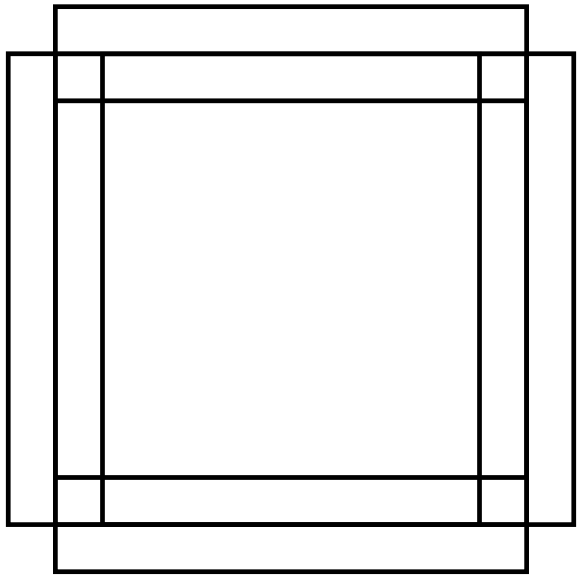

FIGURE 13. Each square intersect its neighbours by 0.9 of their area

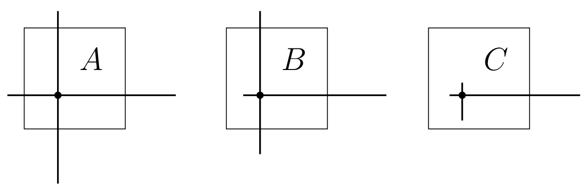

FIGURE 14. A is typical in S, B is not typical in S but it is typical in a nearby square, C is not typical in any square

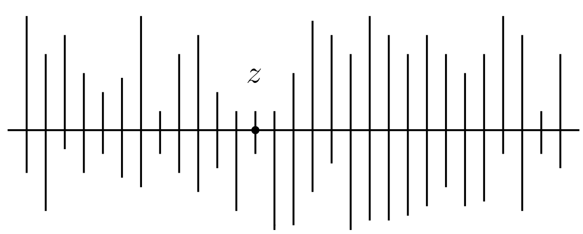

FIGURE 15. Hopf brush $\Sigma(z)$

5.4. **Growth Lemma.** It remains to prove Lemma 5.7. 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

$$\mathbb{P}(r_n \leq \varepsilon) \leq 2\left(\frac{\theta^n}{|\gamma|} + \hat{C}\right)\varepsilon. \tag{5.4}$$

(5.4) implies Lemma 5.7 since it implies that

$$\mathbb{P}(\mathbb{S}_n) \leq 2\left(\frac{\theta^n}{|\gamma|} + \hat{C}\right)\frac{\varepsilon}{\lambda^n}.$$

Summing this for $n \geq k$ we obtain the statement of Lemma 5.7.

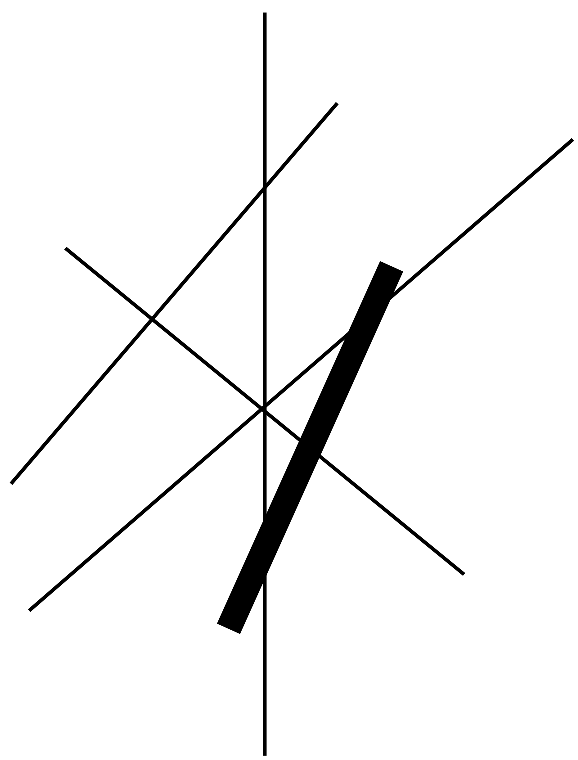

FIGURE 16. The complexity is determined by the largest number of lines passing through one point since one can always take δ so small that any curve of length less than δ can not come close to two intersection points

The proof of (5.4) 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 $\kappa_n \leq Kn$ since the singularities of T^n are lines and there at most Kn possibilities for their slopes. Accordingly there exist numbers n_0, δ_0 such that $\kappa_{n_0}(\delta_0) \leq$

$2\lambda^{n_0}$. Replacing T by T^{n_0} we can assume that this inequality holds for $n_0 = 1$ (clearly it is sufficient to prove (5.4) 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 components. 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".

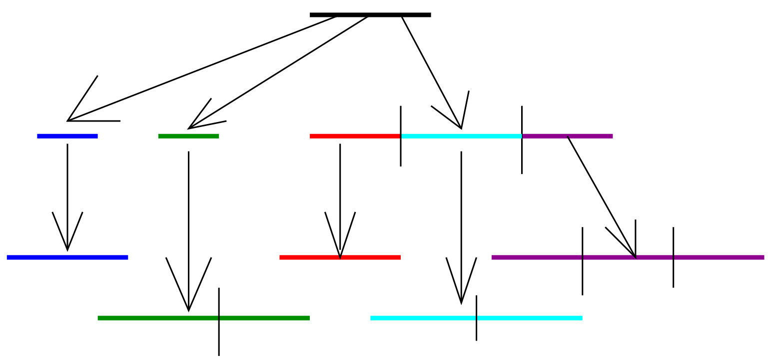

FIGURE 17. Dynamics of components. The vertical segments are "artificial singularities".

It suffices to prove (5.4) with r_n replaced by $\bar{r}_n$. To this end let

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

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

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

Indeed $\bar{r}_{n+1}(x)$ is less than ε if either $\bar{r}_n(x) < \frac{\varepsilon}{\lambda}$ or $T^{n+1}x$ passes near either genuine or artificial singularity. In $T^{n+1}x$ passes near a genuine singularity then $T^n x$ is $\frac{\varepsilon}{\lambda}$ close to the preimage of singularity. 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

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

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.

5.5. **Weak Mixing.** In fact, the argument used to prove Theorem 5.1 can be used to obtain a stronger result.

Theorem 5.8. *$T \times T$ is ergodic.*

Proof. The proof of Theorem 5.1 was based on the fact, that T is a piecewise linear map satisfying the Growth Lemma and the transversality of singularity set to the stable and unstable directions. The map $T \times T$ has the same properties. Indeed let $\Gamma_1 \times \Gamma_2$ be a product of unstable curves in $\mathbb{T}^2 \times \mathbb{T}^2$. Defining as before $r_n(x, y)$ the distance from $(T^n x, T^n)$ to the boundary of the component of $T^n\Gamma_1 \times T^n\Gamma_2$ containing that point we have $r_n(x, y) = \min(r_n(x), r_n(y))$ and so

$$\mathbb{P}(r_n(x, y) \leq \varepsilon) \leq \mathbb{P}(r_n(x) < \varepsilon) + \mathbb{P}(r_n(y) < \varepsilon) \leq 2\theta^n \left(\frac{1}{|\Gamma_1|} + \frac{1}{|\Gamma_2|} + +2\hat{C}\right)\varepsilon$$

proving the Growth Lemma for $T \times T$. Now the proof of ergodicity of $T \times T$ proceeds along the same line as the proof of ergodicty of T. □

5.6. **Mixing and equidistribution.** Here we derive Theorem 5.5 from the following statement. Fix a small constant $\bar{\delta}$.

Proposition 5.9. *Let Γ be a horizontal segment of length $\bar{\delta}$. Then for any $\phi \in C^\alpha(\mathbb{T}^2)$ we have*

$$\left|\frac{1}{\bar{\delta}} \int_\Gamma \phi(T^n x)dx - \int_{\mathbb{T}^2} \phi(z)dz\right| \leq C\theta^n ||\phi||_{C^\alpha}.$$

This proposition claims that images of horizontal curves of large size become equidistributed. To obtain Theorem 5.5 we need to bootstrap this result to small horizontal curves.

Proposition 5.10. *There is a constant $K > 0$ such that if Γ is an Γ be a horizontal segment of length less than $\bar{\delta}$, then for any $\phi \in C^\alpha(\mathbb{T}^2)$ for any $n > K|\ln|\Gamma||$ we have*

$$\left|\frac{1}{|\Gamma|} \int_\Gamma \phi(T^n x)dx - \int_{\mathbb{T}^2} \phi(z)dz\right| \leq C\theta^{n-K|\ln|\Gamma||} ||\phi||_{C^\alpha}.$$

Proof of Theorem 5.5. Let $\phi_1, \phi_2 \in C^\alpha(\mathbb{T}^2)$. Partition $\mathbb{T}^2$ into squares of small size δ. Let $\bar{\phi}$ be an approximation to ϕ_1 which is constant on the elements of our partition. Denote by $\bar{\phi}_j$ the value of $\bar{\phi}$ on the square S_j. Note that $||\phi_1 - \bar{\phi}||_\infty \leq ||\phi_1||_{C^\alpha}\delta^\alpha$. Let $\Gamma_j(h)$ denote the horizontal section of S_j at height h. Then

$$\int_{\mathbb{T}^2} \phi_1(z)\phi_2(T^n z)dz = \int_{\mathbb{T}^2} \bar{\phi}(z)\phi_2(T^n z)dz + O\left(\delta^\alpha ||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha}\right)$$

$$= \sum_j \bar{\phi}_j \int_{S_j} \phi_2(T^n z)dz + O\left(\delta^\alpha ||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha}\right)$$

$$= \sum_j \bar{\phi}_j \int_0^\delta \left(\int_{\Gamma_j(h)} \phi_2(T^n x)dx \right) dh + O\left(\delta^\alpha ||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha}\right)$$

$$= \sum_j \left[\bar{\phi}_j \delta^2 \int_{\mathbb{T}^2} \phi_2(z)dz \right] + O\left(\delta^\alpha ||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha}\right) + O\left(||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha} \theta^{n-K|\ln \delta|}\right).$$

$$= \int_{\mathbb{T}^2} \phi_1(z)dz \int_{\mathbb{T}^2} \phi_2(z)dz + O\left(\delta^\alpha ||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha}\right) + O\left(||\phi_1||_{C^\alpha} ||\phi_2||_{C^\alpha} \theta^{n-K|\ln \delta|}\right).$$

Choosing δ so that $|\ln \delta| = \frac{n}{2K}$ completes the proof of Theorem 5.5. □

5.7. **Another growth lemma.** To derive Proposition 5.10 from Proposition 5.9 we need another Growth Lemma. Recall that the first Growth Lemma (formula (5.4)) tells us that if we start from a horizontal segment of short length δ then the probability that $r_n(x)$ is small becomes small for $n \geq K|\ln \delta|$). Thus on average most points belong to long components most of the time. Here we discuss the exceptional points which stay in short components for a long time. Let Γ be a segment of length δ. Fix $\tilde{n} \geq K|\ln \delta|$. We define a piecewise constant function $\bar{n} : \Gamma \to \mathbb{N}$ such that

$$\bar{n} \geq \tilde{n}, \quad T^{\bar{n}}\Gamma = \bigcup_j \Gamma_j \text{ and } |\Gamma_j| = \bar{\delta}.$$

$\bar{n}$ will be defined inductively. Namely, let $\{\Gamma_{j0}\}$ be the components of $T^{\tilde{n}}\Gamma$. Call Γ_{j0} long if its length is at least $2\bar{\delta}$. Each long component will be further decomposed as $\Gamma_{j0} = L_j \sqcup (\sqcup_k \Gamma_{jk}) \sqcup R_j$ where $|\Gamma_{jk}| = \bar{\delta}$ and L_j and R_j are neighbourhood of the left and right endpoints of Γ_{j0} with

$$\frac{\bar{\delta}}{2} \leq |L_j| \leq \bar{\delta}, \quad \frac{\bar{\delta}}{2} \leq |R_j| \leq \bar{\delta}.$$

On $\cup_{jk}\Gamma_{jk}$ we let $\tilde{n} = \bar{n}$. Let $\tilde{\Gamma}$ be a component where $\bar{n}$ has not yet been defined (thus $\tilde{\Gamma} = \Gamma_{j0}$ where $|\Gamma_{j0}| < 2\bar{\delta}$ or $\tilde{\Gamma} = L_j$ or R_j for some long component Γ_{j0}) consider $T^{K|\ln|\tilde{\Gamma}||}\tilde{\Gamma}$ and trim their long components as before. On the long components of the resulting set we set $\bar{n} = n_1 + K|\ln|\tilde{\Gamma}||$ while points which stay in the short components on both attempts will be iterated once more. This procedure is continued inductively. The second growth Lemma says that for most ponts we stop after a relatively short time.

Lemma 5.11. $\mathbb{P}(\bar{n} > \tilde{n} + k) \leq C\theta^k$.

Before giving the proof of Lemma 5.11 we show how it helps to derive Proposition 5.10.

Proof of Proposition 5.10. We apply Lemma 5.11 with $\tilde{n} = K|\ln\delta|$. Let $\{\Gamma_j\}_{j\in J}$ be all components where $\bar{n} \leq \frac{K\ln|\delta|+n}{2}$. Let n_j denote the value of $\bar{n}$ on $T^{-\bar{n}}\Gamma_j$. Then

$$\frac{1}{|\Gamma|}\int_\Gamma \phi(T^n x)dx = \sum_{j\in J}\frac{|\Gamma_j|}{|\Gamma|}\left[\frac{1}{|\Gamma_j|}\int_{\Gamma_j}\phi(T^{n-n_j}y)dy\right]+O\left(||\phi||\mathbb{P}\left(\bar{n} > \frac{K\ln|\delta|+n}{2}\right)\right)$$

$$= \sum_{j\in J}\frac{|\Gamma_j|}{|\Gamma|}\left[\int_{\mathbb{T}^2}\phi(z)dz + O\left(||\phi||\theta^{(n-K|\ln\delta|)/2}\right)\right]+O\left(||\phi||\mathbb{P}\left(\bar{n} > \frac{K\ln|\delta|+n}{2}\right)\right)$$

$$= \int_{\mathbb{T}^2}\phi(z)dz\mathbb{P}\left(\bar{n} \leq \frac{n+K|\ln\delta|}{2}\right)+O\left(||\phi||\theta^{(n-K|\ln\delta|)/2}\right)+O\left(||\phi||\mathbb{P}\left(\bar{n} > \frac{K\ln|\delta|+n}{2}\right)\right)$$

$$= \int_{\mathbb{T}^2}\phi(z)dzO\left(||\phi||\theta^{(n-K|\ln\delta|)/2}\right)$$

where the first equality is obtained by changing variables, the second uses Proposition 5.9, the third follows from the definition of J and the fourth follows from Lemma 5.11. □

5.8. **Trying to succeed.** We derive Lemma 5.11 from the following more general result. Let J be a $\mathbb{N}$ valued random variable and $T = \sum_{j=1}^J k_j$ where k_j are $\mathbb{N}$ valued random variables. Let $\mathcal{F}_j$ be a filtration such that $k_1, \dots k_j$ are $\mathcal{F}_j$-measurable and as well as sets $\{J = j\}$.

Lemma 5.12. *Suppose that there are constats $K > 0, p < 1, \theta < 1$ such that*

$$\mathbb{P}(J = j+1|\mathcal{F}_j) \geq p, \tag{5.5}$$

$$\mathbb{P}(k_{j+1} = k|\mathcal{F}_j) \leq K\theta^j. \tag{5.6}$$

Then $\mathbb{P}(T = l) \leq \tilde{K}\tilde{\theta}^l$.

Proof. We use moment generating functions. Let

$$\Phi_j(z) = \mathbb{E}\left(z^{k_1+\dots k_j}1_{J\geq j}\right).$$

We claim that there exist numbers $r > 1, \zeta < 1$ such that

$$\Phi_1(z) \text{ and } \Phi_2(z) \text{converge for } 0 < z \leq r \tag{5.7}$$

$$\text{For } j > 2, \quad \Phi_j(z) < \zeta\Phi_{j-2}(z) \text{ for } z \leq r. \tag{5.8}$$

(5.8) implies that

$$\Phi_{2l+1} \leq \zeta^l\Phi_1(z), \quad \Phi_{2l+2} \leq \zeta^l\Phi_2(z)$$

and so $\sum_l \Phi_l(z)$ converges for $z \le r$. But then $\mathbb{E}(z^T) \le \sum_l \Phi_l(z)$ is also finite proving Lemma 5.12. It remains to establish (5.7) and (5.8).

The fact that $\Phi_1(z)$ is bounded for $z < r < \theta^{-1}$ follows from (5.6). Next

$$\Phi_j(z) = \mathbb{E}\left(z^{k_1+\dots+k_{j-1}} 1_{J>j-1} \mathbb{E}\left(z^{k_j} | \mathcal{F}_{j-1}\right)\right).$$

We claim that given δ we can take r so close to 1 that

$$\mathbb{E}\left(z^{k_j} | \mathcal{F}_{j-1}\right) \le 1+\delta. \tag{5.9}$$

Hence

$$\Phi_j(z) \le (1+\delta)\mathbb{E}\left(z^{k_1+\dots+k_{j-1}} 1_{J>j-1}\right).$$

Next

$$\mathbb{E}\left(z^{k_1+\dots+k_{j-1}} 1_{J>j-1}\right) = \mathbb{E}\left(z^{k_1+\dots+k_{j-2}} 1_{J\ge j-2} \mathbb{E}\left(z^{k_{j-1}} 1_{J\ne j-1} | \mathcal{F}_{j-2}\right)\right).$$

We claim that there is a constant $\bar{\zeta} < 1$ such that if r is sufficiently close to 1, then

$$\mathbb{E}\left(z^{k_{j-1}} 1_{J\ne j-1} | \mathcal{F}_{j-2}\right) \le \bar{\zeta}. \tag{5.10}$$

Taking δ so small that $\zeta := \bar{\zeta}(1+\delta) < 1$ we obtain (5.8). Also (5.9) implies that $\Phi_2(z) \le (1+\delta)\Phi_1(z)$ proving (5.7).

It remains to prove (5.9) and (5.10). Since the LHSs of (5.9) and (5.10) are increasing in z it sffices to consider $z = r$. Then

$$\mathbb{E}\left(r^{k_j} | \mathcal{F}_{j-1}\right) = 1 + \frac{\partial}{\partial z}\mathbb{E}\left(z^{k_j} | \mathcal{F}_{j-1}\right)(z = \tilde{z})(r-1) \text{ for some } \tilde{z} \in [1, r]. \tag{5.11}$$

By (5.6)

$$\frac{\partial}{\partial z}\mathbb{E}\left(z^{k_j} | \mathcal{F}_{j-1}\right)(z) \le K \sum_l l\theta^l$$

is bounded for $z \le r < \theta^{-1}$. Therefore the second term in (5.11) can be made as close to 1 as we wish by taking r close to 1. This proves (5.9). The proof of (5.10) is similar except that we use that

$$\mathbb{E}\left(z^{k_{j-1}} 1_{j\ne J-1} | \mathcal{F}_{j-2}\right) \le 1-p$$

due to (5.5). $\square$

5.9. **Equidistribution and coupling.** The main step in proving Proposition 5.9 is the following.

Lemma 5.13. *Let Γ_1 and Γ_2 be two segments of length $\bar{\delta}$ then*

$$\left|\frac{1}{\bar{\delta}}\int_{\Gamma_1} \phi(T^n x)dx - \frac{1}{\bar{\delta}}\int_{\Gamma_2} \phi(T^n x)dx\right| \le C\theta^n ||\phi||_{C^\alpha}.$$

Proof of Proposition 5.9. We claim that for any $\varepsilon > 0$ we can decompose $\mathbb{T}^2 = (\sqcup_j S_j) \sqcup Z$ where S_j are rectangles of width $\bar{\delta}$ and $\text{mes}(Z) < \varepsilon$. Indeed let $\bar{\Delta}$ be a vertical rectange of small length η and let τ be a first return time to Δ by the horizontal flow. Then $\tau \to \infty$ as $\eta \to 0$. Cutting each piece of time τ orbit into segments of length $\bar{\delta}$ we obtain the required partition. Let H_j be the height of S_j and $\Gamma_j(h)$ be the vertical segment in S_j at height h. Then

$$\int_{\mathbb{T}^2} \phi(z)dz = \int_{\mathbb{T}^2} \phi(T^n z)dz = \sum_j \int_0^{H_j} \left(\int_{\Gamma_j} \phi(T^n x)dx \right) dh + O\left(\varepsilon ||\phi||_\infty\right)$$

$$\sum_j \int_0^{H_j} \left(\int_{\Gamma} \phi(T^n x)dx \right) dh + O\left(\theta^n ||\phi||_{C^\alpha}\right) + O\left(\varepsilon ||\phi||_\infty\right)$$

where the last step relies on Lemma 5.13. Since ε is arbitrary the result follows. □

Note that the above proof shows in particular that for each ε we can find rectangles $\{S_j\}$ so that the following decomposition holds

$$\int_{\mathbb{T}^2} \phi(z)dz = \sum_j \int_0^{H_j} \left(\int_{\Gamma_j} \phi(x)dx \right) dh + O\left(\varepsilon ||\phi||_\infty\right). \tag{5.12}$$

This decomposition proves very convenient in the study of statistical properties of T.

Next we describe the idea of the proof of Lemma 5.13 in the smooth case. We have

$$\frac{1}{|\Gamma_j|} \int_{\Gamma_j} \phi(T^n x)dx = \frac{1}{|T^{n/2}\Gamma_j|} \int_{T^{n/2}\Gamma_j} \phi(T^{n/2} y)dy.$$

Note that $T^{n/2}\Gamma_j$ are segment of length $\lambda^{n/2}\bar{\delta}$. Since both of the segments are in $\mathbb{T}^2$ the distance between their starting points in $O(1)$. Hence we can represent $T^{n/2}\Gamma_j = \tilde{\Gamma}_j + \hat{\Gamma}_j$ where $|\tilde{\Gamma}_j| = O(1)$ and $\hat{\Gamma}_2$ is a projection of Γ_1 along e_s. Denoting this projection by π we have $d(y, \pi y) = O(1)$ and hence

$$d\left(T^{n/2} y, T^{n/2} \pi y\right) = O\left(\lambda^{-n/2}\right).$$

Thus

$$\int_{\tilde{\Gamma}_2} \phi(T^{n/2} y_2)dy_2 - \int_{\tilde{\Gamma}_1} \phi(T^{n/2} y_1)dy_1 = \int_{\tilde{\Gamma}_1} [\phi(T^{n/2} y_1) - \phi(T^{n/2} \pi y_1)dy_1]$$

$$= O\left(\lambda^{-\alpha n/2} ||\phi||_{C^\alpha}\right).$$

On the other hand

$$\int_{\tilde{\Gamma}_j} \phi(T^{n/2}y_j)dy_j = O(||\phi||_\infty).$$

Dividing by $\bar{\delta}\lambda^{n/2}|T^{n/2}\Gamma_j|$ we obtain the required estimate.

In the nonsmooth case the structure of $T^{n/2}\Gamma_j$ is more compicated but we still can split $T^{n/2}\Gamma_j$ into pieces which are close to each other. This is content of the following result.

Lemma 5.14. *(*Coupling Lemma*) There exists a measure preserving map $\pi : \Gamma_1 \to \Gamma_2$ (***coupling map***) and a function $R : \Gamma_1 \to \mathbb{N}$* **(coupling time)** *such that*

(a) There is a constant η such $T^{R(x)}x$ and $T^{R(x)}\pi(x)$ belong to the same stable manifold of length less that η and so for $n > R(x)$

$$d(T^n x, T^n \pi x) \leq \eta\theta^{n-R(x)}.$$

(b) $\mathbb{P}(R \geq k) \leq C\theta^k$.

5.10. **Coupling, separating, recovering.** Here we describe an algorithm for constructing the coupling map π. This will be done recursively. Namely, given Γ_1, Γ_2 as in the Coupling Lemma we define a time of the first attempt $k_1 : \Gamma_1 \cup \Gamma_2 \to \mathbb{N}$. The coupling map will be defined on a subset $L_1 \subset \Gamma_1$ so that on L_1 we have $R(x) = k_1$. We will arrange that

$$|L_1| \geq p. \tag{5.13}$$

In addition if $L_2 = \pi L_1$ then

$$T^{k_1}(\Gamma_1 - L_i) = \cup_j \tilde{\Gamma}_{ij}$$

so that $|\tilde{\Gamma}_{ij}| = \bar{\delta}$ and

$$\mathbb{P}_{\Gamma_1}(k_1 = k) = \mathbb{P}_{\Gamma_2}(k_1 = k) \leq C\theta^k. \tag{5.14}$$

Then we will try recursively to couple $\tilde{\Gamma}_{1j}$ to $\tilde{\Gamma}_{2j}$ and so on.

(5.13) shows that repeating the above procedure repeatedly we can define π alomost everywhere. Also (5.13) and (5.14) allow to apply Lemma 5.12 to get an exponential tail bound on the coupling time. it remains to describe one step of our proceudre verifying (5.13) and (5.14). The coupling algorithm relies on the following estimate.

Lemma 5.15. *If $\tilde{\delta}, \tilde{\eta}$ are sufficiently small than there is a constant $N = N(\tilde{\delta}, \tilde{\eta})$ such that for each pair Γ_1, Γ_2 with $|\Gamma_1| = |\Gamma_2| = \bar{\delta}$ there are segments $\hat{L}_1 \subset \Gamma_1, \hat{L}_2 \subset \Gamma_2$ and $\bar{N} < N$ such that $T^{\bar{N}}\hat{L}_i$ are horizontal segments of length $\tilde{\delta}$, $T^{\bar{N}}\hat{L}_2$ is a vertical projection of $T^{\bar{N}}\hat{L}_1$ and the distance between $T^{\bar{N}}\hat{L}_1$ and $T^{\bar{N}}\hat{L}_2$ is less that $\tilde{\eta}$.*

During one run of our algorithm π will be defined on a subset $L_1 \subset \hat{L}_1$. But first we explain how to define k_1 on $\Gamma_1 - \hat{L}_1$. We will use the following fact.

Lemma 5.16. *Let G_1, G_2 be unions of horizontal segments of the same total length such that*

$$\mathbb{P}_{G_i}(r(x) < \varepsilon) \leq Z\varepsilon$$

where $\mathbb{P}_{G_i}$ is the uniform distribution on G_i and $r(x)$ is the distance from x to the boundary of the segment containing x. Then there is a function $k : G_1 \cup G_2 \to \mathbb{N}$ such that $T^k G_i = \bigcup_j G_{ij}$, $|G_{ij}| = \bar{\delta}$ and

$$\mathbb{P}_{G_1}(k(x) = k) = \mathbb{P}_{G_2}(k(x) = k) \leq C\theta^{k - K \ln Z}.$$

The proof of Lemma 5.16 is similar to the proof of Lemma 5.11 so it is left to the reader.

Next we consider $\hat{L}_1$. We want to set $\pi = T^{-\bar{N}}\tilde{\pi}$ where $\tilde{\pi} : T^{\bar{N}}\hat{l}_1 \to T^{\bar{N}}\hat{L}_2$ is a vertical projection. However, y and $\tilde{\pi}y$ need not belong to the same stable manifold. The obstacle is existence of a number n such that $T^n y$ and $T^n\tilde{\pi}y$ are separeted by a singularity. In that case

$$d(T^n y, \tilde{S}) \leq \eta\lambda^{-N}, d(T^n\tilde{\pi}y, \tilde{S}) \leq \eta\lambda^{-n}.$$

So at time n we remove the points falling into $\eta\lambda^{-n}$ neighbourhood of $\tilde{S}$ as well as its vertical projection. On the removed set we define **separation time** $s(x)$ as follows. Consider a component of the set removed at time n. If this component is longer than λ^{-2n} then we let $s(x)$ to be equal to n on that component. Otherwise an endpoint b of this component has been removed at a erlier time and we let $s = s(b)$ on this component. Note that by construction and (5.4)we have

$$\mathbb{P}(s = k) \leq \bar{C}\bar{\theta}^k. \tag{5.15}$$

Also if $\tilde{\eta} \ll \tilde{\delta}$ then (5.4) show sthat the set where $\pi = T^{-\bar{N}}\tilde{\pi}$ (that is the set of points which are not removed due to a close approach to the singular set) has a relatively large measure in L_1 proving (5.13).

It remains to define k on $L_1 - \hat{L}_1$. This will be done using Lemma 5.16 pairing the points having the same separation time. Note that by our construction all componets where $s = \bar{s}$ have length at least $\lambda^{-2\bar{s}}$, so letting $\rho(x) = k(x) - s(x)$ be the recovery time we get

$$\mathbb{P}(\rho(x) > ks(x) + l) \leq \bar{C}\theta^l. \tag{5.16}$$

(5.15) and (5.16) give (5.14) for the separated points since

$$\mathbb{P}(k(x) > k) \leq \mathbb{P}\left(s(x) > \frac{k}{2K}\right) + \mathbb{P}\left(\rho(x) \geq Ks(x) + \frac{k}{2}\right)$$

and both terms have exponentially small probability. It remains to prove Lemma 5.15.

Proof of Lemma 5.15. We begin with a simplifying remark. The statement requires that N be uniform in Γ_1, Γ_2 but we note that it suffices to prove it for fixed Γ_1, Γ_2 (but for all sufficiently small $\tilde{\delta}, \tilde{\eta}$). Indeed take $\tilde{\tilde{\delta}} > \tilde{\delta}, \tilde{\tilde{\eta}} < \tilde{\eta}$. Then if the statement holds for $(\Gamma_1, \Gamma_2, \tilde{\tilde{\delta}}, \tilde{\tilde{\eta}})$ then it also holds for $(\Gamma_1', \Gamma_2', \tilde{\delta}, \tilde{\eta})$ provided that (Γ_1', Γ_2') is sufficiently close to (Γ_1, Γ_2). Since the set of pairs is compact we can choose a finite subcover achived the required uniformity.

Let

$$\Sigma_j = \bigcup_{y \in \Gamma_j} W^s_{\tilde{\eta}/3}(y).$$

Note that Σ_j has a positive measure in $\mathbb{T}^2$, hence $\Sigma_1 \times \Sigma_2$ has a positive measure in $\mathbb{T}^2 \times \mathbb{T}^2$. By ergodicity of $T \times T$ given a set $\Omega \in \mathbb{T}^2 \times \mathbb{T}^2$ almost every point almost every point in $\Sigma_1 \times \Sigma_2$ visits Ω with frequency $\mathrm{mes}(\Omega)$. Let

$$\Omega_k = \{(x_1, x_2) : \mathbb{T}^2 \times \mathbb{T}^2 : d(x_1, x_2) \leq \frac{\tilde{\eta}}{3} \text{ and } T^{-k} \text{ is continuous in } B(x_i, 2\tilde{\delta})\}.$$

Note that if $\tilde{\eta}, \tilde{\delta}$ are small than $\mathrm{mes}(\Omega_k) \geq \eta^2/10$. By the foregoing discussion given ε there is N such that

$$\mathrm{mes}\left((y_1, y_2) \in \Sigma_1 \times \Sigma_2 : \mathrm{Card}(n \leq N : (T^n y_1, T^n y_2) \in \Omega_k) \leq N \frac{\mathrm{mes}(\Omega_k)}{2}\right) \leq \varepsilon$$

Next by the growth lemma given $\hat{\varepsilon}$ there eixsts $\hat{\delta}$ so small that

$$\mathrm{mes}_{\Sigma_1 \times \Sigma_2}((y_1, y_2) : r_n(x_1) < \hat{\delta} \text{ or } r_n(x_2) \leq \hat{\delta}) < \hat{\varepsilon}\mathrm{mes}(\Sigma_1 \times \Sigma_2)$$

where $x_i = \tilde{\pi} y_i$. So if

$$\hat{n}(y_1, y_2) = \mathrm{Card}(n \leq N : r_n(x_1) \geq \hat{\delta} \text{ and } r_n(x_2) > \hat{\delta})$$

then $\mathbb{E}(\hat{n}) \geq (1 - \hat{\varepsilon})N$. Since $\hat{n} \leq N$

$$\mathbb{P}\left(\hat{n} < N\left(1 - \frac{\eta^2}{100}\right)\right) \leq \frac{100\hat{\varepsilon}}{\eta^2}$$

which can be made as small as we wish by choosing $\hat{\varepsilon}$ small. Note that if $d(T^n y_1, T^n y_2) < \tilde{\eta}/3$ then $d(T^n x_1, T^n x_2) < \tilde{\eta}$. Therefore given if N is sufficiently large and $\varepsilon, \hat{\varepsilon}$ are sufficiently small then there exist $(x_1, x_2) \in \Sigma_1 \times \Sigma_2$ and $\bar{N} < N$ such that

$$d(T^{\bar{N}} x_1, T^{\bar{N}} x_2) < \tilde{\eta},\ r_{\bar{N}-k}(x_i) > \hat{\delta} \text{ and } T^{-k} \text{ is continuous on } B(x_i, \delta)$$

for $i \in \{1, 2\}$. The continuity condition implies that

$$r_{\bar{N}}(x_i) \geq r_{\bar{N}-k} \lambda^k \geq \tilde{\delta}$$

if k is sufficiently large. Therefore Γ_1 and Γ_2 contain two segments such that their $T^{\bar{N}}$ image has length at least $2\tilde{\delta}$ and the distance between their centers is at most $\tilde{\eta}$. If $\tilde{\eta} < \tilde{\delta}$ (which can be assumed without loss of the generality) we can trim those segments so that one is a vertical projection of the other proving the lemma. $\square$

6. The Central Limit Theorem for Dynamical Systems.

6.1. **Estimating error in Ergodic Theorem.** If T is an ergodic map of a space M equipped with a probability measure μ then the Ergodic Theorem says that for $\phi \in L^1(\mu)$ we have

$$\frac{S_N(x)}{N} \to \int_M \phi(x)d\mu(x) \text{ where } S_N(x) = \sum_{n=0}^{N-1} \phi(T^n x).$$

The next natural question is the rate of convergence.

To formulate the question more precisely we need to recall some facts from probability theory. Let $\mathcal{S}$ be a Polish metric space, S_n be a sequnce of $\mathcal{S}$ valued random variables and S be an $\mathcal{S}$ valued random variable. We say that S_N converges to S in distribution (written as $S_N \Rightarrow S$) if for any bounded continuous function Φ we have

$$\mathbb{E}(\Phi(S_N)) \to \mathbb{E}(\Phi(S)).$$

In case $\mathcal{S} = \mathbb{R}^d$ the following statements are equivalent

- $S_N \Rightarrow S$
- For each $\xi \in \mathbb{R}^d$ $\mathbb{E}(e^{i\xi S_N}) \to \mathbb{E}(e^{i\xi S})$
- Define $F_S(s) = \mathbb{P}(S_{(1)} \leq s_1, S_{(2)} \leq s_2, \dots S_{(d)} \leq s_d)$ and and let $F_{S_N}(s)$ be a similar expression for S_N then for all continuity points of F_S we have $\lim_{N\to\infty} F_{S_N}(s) = F_S(s)$.

Given a number $\sigma > 0$ let $\mathcal{N}(\sigma^2)$ denote the normal random variable with zero mean and standard deviation σ. Thus

$$\mathbb{P}(\mathcal{N} \leq s) = \int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}\sigma} e^{-u^2/2\sigma^2} du \text{ and } \mathbb{E}\left(e^{i\xi\mathcal{N}}\right) = e^{-\sigma^2\xi^2/2}.$$

In case of independent random variables the fluctuations of ergodic sums are of order $\sqrt{N}$ and the limiting distribution is normal. One can ask if the same is true in the dynamical systems setting. In many cases where one has exponential divergence of nearby trajectories the answer is YES. However one needs to impose some regularity requirement on ϕ. Without smoothness assumptions, Exercise 5.4(b) shows that one a function ϕ such that $\mathbb{E}(S_N^2) \gg N$.

Theorem 6.1. *Let T be as in Theorem 5.1 and $\phi \in C^\alpha(\mathbb{T}^2)$ be a function with zero mean. Suppose that x is uniformly distributed on $\mathbb{T}^2$. Then*

$$\frac{S_N}{\sqrt{N}} \Rightarrow \mathcal{N}(\sigma^2)$$

where

$$\sigma^2 = \sum_{p=-\infty}^{\infty} \int_{\mathbb{T}^2} \phi(z)\phi(T^p z)dz. \tag{6.1}$$

In other words,

$$\mathbb{P}\left(\frac{S_N}{\sqrt{N}} \leq s\right) \to \int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}\sigma} e^{-u^2/(2\sigma^2)} du.$$

The assumption that ϕ has zero mean does not cause any loss of generality since we can always replace ϕ by $\bar{\phi} = \phi - \int_{\mathbb{T}^2} \phi(z)dz$.

6.2. **Bernstein method.** By the foregoing discussion we need to show that

$$\mathbb{E}\left(e^{i\xi S_N/\sqrt{N}}\right) \to e^{-\sigma^2\xi^2/2}. \tag{6.2}$$

Before proving (6.2) for toral maps let us recall how (6.2) is established for independent random variables. Namely, suppose for a moment that $S_N = \sum_{n=0}^{N-1} X_n$ where X_n are indendent identically distributed random variables with zero mean and standard deviation σ. We have

$$\mathbb{E}\left(e^{i\xi S_N/\sqrt{N}}\right) = \left[\mathbb{E}\left(e^{i\xi X/\sqrt{N}}\right)\right]^N.$$

Next

$$\mathbb{E}\left(e^{i\xi X/\sqrt{N}}\right) = \mathbb{E}\left(1 + \frac{i\xi X}{\sqrt{N}} - \frac{\xi^2 X^2}{2N}\right) + o\left(\frac{1}{N}\right) = 1 - \frac{\xi^2\sigma^2}{2N} + o\left(\frac{1}{N}\right).$$

Raising this expression to the N-th power we obtain (6.2) in the case of indendent identically distributed random variables.

In the dynamical system case this method can not be applied directly since $\phi(T^n x)$ are not independent for different n. However, mixing shows that $\phi(T^N x)$ and $\phi(T^{n+j}x)$ are weakly dependent for large j. One useful technique for handling weakly dependent random variable is Bernstein big block small block method which we now describe. Choose $\alpha_1 < \alpha_2 < \frac{1}{4}$ and divide $[0, N]$ into big blocks of size N^{α_2} separated by small blocks of size N^{α_1} starting from a small block. Thus we let

$$\tilde{\tilde{S}}_j = \sum_{n=(N^{\alpha_1}+N^{\alpha_2})(j-1)+1}^{N^{\alpha_1}j+N^{\alpha_2}(j-1)} \phi(T^n x), \quad \tilde{S}_j = \sum_{n=N^{\alpha_1}j+N^{\alpha_2}(j-1)+1}^{(N^{\alpha_1}+N^{\alpha_2})j} \phi(T^n x),$$

$$\tilde{S} = \sum_j \tilde{S}_j, \quad \tilde{\tilde{S}} = \sum_j \tilde{\tilde{S}}_j.$$

Thus $\tilde{\tilde{S}}$ is the contribution of small blocks while $\tilde{S}$ is the contribution of big blocks. The idea is that the contribution of the small blocks is negligible since the number of terms in $\tilde{\tilde{S}}$ is $O\left(N^{1-\alpha_2+\alpha_1}\right)$ while the contribution of different big blocks is almost independent since the blocks are far apart. Let us give the detailed argument. First we dispose of $\tilde{\tilde{S}}$. We have

$$\mathbb{E}\left(\tilde{\tilde{S}}^2\right) = \sum_{n_1,n_2}^{**} \int_{\mathbb{T}^2} \phi(T^{n_1}x)\phi(T^{n_2}x)dx = \sum_{n_1,n_2}^{**} \int_{\mathbb{T}^2} \phi(x)\phi(T^{n_2-n_1}x)dx$$

$$= \sum_{n_1,n_2}^{**} O\left(\theta^{|n_2-n_1|}\right) = O\left(N^{1+\alpha_1-\alpha_2}\right)$$

where $\sum^{**}$ denotes the sum over the small blocks. Thus

$$\mathbb{E}\left(\left(\frac{\tilde{\tilde{S}}}{\sqrt{N}}\right)^2\right) = O\left(N^{\alpha_1-\alpha_2}\right)$$

and so $\frac{\tilde{\tilde{S}}}{\sqrt{N}} \Rightarrow 0$.

Exercise 6.2. If $S'_N \Rightarrow S$, $S''_N \Rightarrow 0$ then $S'_N + S''_N \Rightarrow S$.

Accordingly it is enough to prove the CLT for $\tilde{S}$. Due to the decomposition (5.12) we may assume that x is chosen uniformly on a segment Γ of length $\bar{\delta}$. Let m_j be the center of j small block. Let Γ_{jk} denote the components of $T^{m_{j+1}}\Gamma_{jk}$. We have

$$\mathbb{E}_\Gamma\left(\psi \circ T^{m_{j+1}}\right) = \sum_k c_k \mathbb{E}_{\Gamma_{jk}}(\psi)$$

where $c_k = \mathbb{P}_\Gamma(T^{m_{j+1}}x \in \Gamma_{jk}$. Denote $Q_j = \sum^{(j)} \Phi(T^n x)$ where $\sum^{(j)}$ means the sum over the j-th big block. Condier the characterisitc fuctions

$$\gamma_j(\xi) = \mathbb{E}_\Gamma\left(\exp\left(\sum_{l=1}^{j} \frac{i\xi Q_l}{\sqrt{N}}\right)\right) \text{ for } j > 1 \text{ and } \gamma_0(\xi) = 1.$$

We have

$$\gamma_{j+1}(\xi) = \sum_k c_k \mathbb{E}_{\Gamma_{jk}} \left(\exp\left(\left[\sum_{l=1}^{j} \frac{i\xi Q_l}{\sqrt{N}} \right] + \frac{i\xi Q_{j+1}}{\sqrt{N}} \right) \right)$$

$$= \sum_k{}' c_k \mathbb{E}_{\Gamma_{jk}} \left(\exp\left(\left[\sum_{l=1}^{j} \frac{i\xi Q_l}{\sqrt{N}} \right] + \frac{i\xi Q_{j+1}}{\sqrt{N}} \right) \right) + O\left(N^{-100}\right)$$

where $\sum_k'$ denotes the sum over the components with $|\Gamma_{jk}| \geq N^{-100}$. Choose $x_k \in T^{-m_{j+1}}\Gamma_{jk}$ and let $q_k = \sum_{l=1}^{j} Q_l(x_k)$. Note that if $\tilde{x}_k \in T^{-m_{j+1}}\Gamma_{jk}$ then $d(T^{m_{j+1}}x_k, T^{m_{j+1}}\tilde{x}_k) \leq 1$ and so

$$d(T^n x_k, T^n \tilde{x}_k) \leq \frac{1}{\lambda^{m_{j+1}-n}} \leq \frac{1}{\lambda^{N^{\alpha_1/2}}}.$$

Since $\phi \in C^\alpha(\mathbb{T}^2)$ we have

$$\sum_l \left[Q_l(\tilde{x}_k) - Q_l(x_k)\right] = O\left(\lambda^{-\alpha N^{\alpha_2}}\right) \tag{6.3}$$

and so

$$\gamma_{j+1}(\xi) = \sum_k{}' c_k e^{i\xi q_k/\sqrt{N}} \mathbb{E}_{\Gamma_{jk}} \left(e^{i\xi Q_{j+1}/\sqrt{N}} \right) + O\left(N^{-100}\right).$$

Now as in the independent case we can use a decomposition

$$e^{i\xi Q_{j+1}/\sqrt{N}} = 1 + \frac{i\xi Q_{j+1}}{\sqrt{N}} - \frac{\xi^2 Q_{j+1}^2}{2N} + O\left(\frac{|Q_j|^3}{N^{3/2}}\right).$$

Lemma 6.3. *(a)* $\mathbb{E}_{\Gamma_{jk}}(Q_{j+1}) = O\left(N^{-100}\right)$.
(b) $\mathbb{E}_{\Gamma_{jk}}(Q_{j+1}^2) = N^{\alpha_2}\sigma^2 + o(1)$.

Lemma 6.3 implies that

$$\mathbb{E}_{\Gamma_{jk}}\left(e^{i\xi Q_{j+1}}\right) = 1 - \frac{N^{\alpha_2-1}\sigma^2\xi^2}{2} + O\left(\frac{1}{N}\right) + O\left(\mathbb{E}_{\Gamma_{jk}}\left(\frac{\left|Q_{j+1}^3\right|}{N^{3/2}}\right)\right). \tag{6.4}$$

Next

$$\mathbb{E}_{\Gamma_{jk}}\left(|Q_{j+1}|^3\right) \leq N^{\alpha_2}\mathbb{E}_{\Gamma_{jk}}\left(|Q_{j+1}|^2\right) = O\left(N^{2\alpha_2}\right)$$

so the last term in (6.4) is $O\left(N^{2\alpha-3/2}\right)$ and so it is negligible. Using again (5.4) and (6.3) we see that

$$\sum_k{}' c_k e^{i\xi q_k/\sqrt{N}} = \gamma_j(\xi) + O\left(N^{-100}\right)$$

so that

$$\gamma_{j+1}(\xi)=\gamma_j(\xi)\left(1-\frac{N^{\alpha_2-1}\sigma^2\xi^2}{2}\right)+O\left(\frac{1}{N}\right). \tag{6.5}$$

Iterating this relation $\frac{N}{N^{\alpha_1}+N^{\alpha_2}}$ times we obtain

$$\mathbb{E}\left(e^{i\xi\tilde{S}/\sqrt{N}}\right)\approx\left(1-\frac{N^{\alpha_2-1}\sigma^2\xi^2}{2}\right)^{N^{1-\alpha_2}}\approx e^{-\sigma^2\xi^2/2} \tag{6.6}$$

as needed.

Exercise 6.4. Show that (6.5) implies (6.6).

6.3. **Moment estimates.**

Proof of Lemma 6.3.

$$(a)\quad \mathbb{E}_{\Gamma_{jk}}\left(Q_{j+1}\right)=\sum_{a_{j+1}}^{b_{j+1}}\mathbb{E}_{\Gamma_{jk}}(\phi\circ T^n)=O\left(||\phi||_{C^\alpha}N^{\alpha_2}\theta^{N^{\alpha_2}}\right)$$

where $[a_j, b_j]$ is the j-th big block.

$$(b)\quad \mathbb{E}_{\Gamma_{jk}}\left(Q_{j+1}^2\right)=\sum_{n_1,n_2=a_{j+1}}^{b_{j+1}}\mathbb{E}_{\Gamma_{jk}}((\phi\circ T^{n_1})(\phi\circ T^{n_2})).$$

We begin with an *a priori bound*

$$\mathbb{E}_{\Gamma_{jk}}((\phi\circ T^{n_1})(\phi\circ T^{n_2}))=O\left(\theta^{|n_2-n_1|}\right). \tag{6.7}$$

To check (6.7) we assume without the loss of the generality that $n_2 > n_1$. Let $n_3=\frac{n_1+n_2}{2}$. We have

$$\mathbb{E}_{\Gamma_{jk}}(\psi\circ T^{n_3})=\sum_s c_s\mathbb{E}_{\Gamma_{jks}}(\psi)$$

where Γ_{jks} are the components of $T^{n_3}\Gamma_{jk}$. By the Growth Lemma we have

$$\mathbb{E}_{\Gamma_{jk}}((\phi\circ T^{n_1})(\phi\circ T^{n_2}))=\sum_s{}' c_s\mathbb{E}_{\Gamma_{jks}}((\phi\circ T^{-r})(\phi\circ T^r))+O(e^{-\varepsilon r})$$

where $p=\frac{n_2-n_1}{2}$ and $\sum_s'$ means the sum over components which are longer than $e^{-\varepsilon r}$. Choosing $x_s\in T^{-n_3}\Gamma_{jks}$ we get that on Γ_{jks}

$$\phi(T^{-p}x)\phi(T^px)=\left[\phi_s+O\left(\lambda^{-\alpha p}\right)\right]\phi(T^px)$$

where $\phi_s=\phi(T^{-r}x_s)$. Thus

$$\mathbb{E}_{\Gamma_{jk}}\left((\phi\circ T^{n_1})(\phi\circ T^{n_2})\right)=\sum_s{}' c_s\phi_s\mathbb{E}_{\Gamma_{jks}}(\phi\circ T^p)+O\left(e^{-\varepsilon p}\right)$$

and if $\varepsilon K < 1$ then Proposition 5.10 gives $\mathbb{E}_{\Gamma_{jks}}(\phi \circ T^p) = O(\theta^p)$ compliting the proof (6.7). Now to finish the proof of part (b) it remains to show that for a fixed p

$$\mathbb{E}_{\Gamma_{jk}}((\phi \circ T^n)(\phi \circ T^{n+p})) \to \int_{\mathbb{T}^2} \phi(z)\phi(T^p z)dz \tag{6.8}$$

as $n \to \infty$. Let $\phi_p(x) = \phi(x)\phi(T^p x)$. The LHS of (6.8) is $\mathbb{E}_{\Gamma_{jk}}(\phi_p \circ T^n)$. Thus if ϕ_p were Holder then the result would follow directly from Proposition 5.10. However ϕ_p is not Holder since T^p is not smooth. Fortunately, ϕ_p can be well approximated by Holder functions. Given ε let

$$\phi_{p,\varepsilon} = \frac{1}{\pi\varepsilon^2} \int_{B(x,\varepsilon)} \phi_p(y)dy.$$

Then $\phi_{p,\varepsilon}$ is uniformly Lipshitz

$$|\phi_{p,\varepsilon}(x') - \phi_{p,\varepsilon}(x'')| \le \frac{C}{\varepsilon^2} d(x', x'').$$

On the other hand if T^p is continuous on $B(x, \varepsilon)$ then

$$|\phi_{p,\varepsilon}(x) - \phi(x)| \le K^r \varepsilon^\alpha. \tag{6.9}$$

Thus

$$\mathbb{E}_{\Gamma_{jk}}(\phi_p \circ T^n) = \mathbb{E}_{\Gamma_{jk}}(\phi_{p,\varepsilon} \circ T^n) + \mathbb{E}_{\Gamma_{jk}}([\phi_p - \phi_{p,\varepsilon}] \circ T^n) = I + (II).$$

$$I = \int_{\mathbb{T}^2} \phi_{p,\varepsilon}(z)dz + O\left(\frac{\theta^n}{\varepsilon^2}\right)$$

since $\phi_{p,\varepsilon}$ is Lipshitz. Next, due to (6.9) we have

$$(II) \le \sum_{k=0}^{p-1} \mathbb{P}_{\Gamma_{jk}}(r_{n+k} \le \varepsilon) + K^p \varepsilon^\alpha \le Cp\varepsilon + K_p \varepsilon^\alpha. \tag{6.10}$$

Choose ε so that $\theta^n = \varepsilon^3$ then both error terms are small.

Combining (6.10) with decomposition (5.12) we see that

$$\int_{\mathbb{T}^2} \phi_{p,\varepsilon}(z)dz = \int_{\mathbb{T}^2} \phi_p(z)dz + O(p\varepsilon + \varepsilon^\alpha)$$

so (6.8) follows. □

6.4. **The case of zero variance.** Theorem 6.1 is only intersting if the variance given by formula (6.1) is non-zero. Indeed if σ^2 is zero then the Theorem just says that $\sqrt{n}$ is a wrong scaling for ergodic sums. Here we show that $\sigma^2 = 0$ only in exceptional cases. Namely we prove the following result.

Let $\mathcal{P}_k$ denote the partition into domains of continuity of T^k and $\tilde{\mathcal{P}}^k = \mathcal{P}_k \vee \mathcal{P}_{-k}$.

Proposition 6.5. *Suppose that for some $\bar{k}$ A is Lipshitz on each element of $\tilde{\mathcal{P}}_{\bar{k}}$ and that $\sigma^2 = 0$. Then there exists k_0 and a function B which is Lipshitz on each element of $\tilde{\mathcal{P}}_{k_0}$ such that*

$$\phi(x) = B(x) - B(Tx). \tag{6.11}$$

Corollary 6.6. *If there is a periodic point p of period N such that*

$$\sum_{n=0}^{N-1} \phi(T^n p) \neq 0$$

then $\sigma^2 > 0$.

Proof. If $\sigma^2 = 0$ then $\sum_{n=0}^{N-1} \phi(T^n p) = B(p) - B(T^N p) = 0$. □

The proof of Proposition 6.5 consists of several steps.

Lemma 6.7. *If $\sigma^2 = 0$ then (6.11) admits an L^2 solution.*

Proof.

$$||S_n||^2_{L^2} = \sum_{k_1,k_2=0}^{n-1} \mu(\phi(T^{k_1}x)A(T^{k_2}x)) = \sum_{k=0}^{n-1} \sigma_{n,k}$$

where $\sigma_{n,k}$ denotes the sum of the terms with the smallest index k. Thus

$$\sigma_{n,k} = \sum_{j=-(n-1-k)}^{n-1-k} \mu(\phi(x)\phi(T^j x)) = \sigma^2 + \mathcal{O}\left(\theta^{n-k}\right) = \mathcal{O}\left(\theta^{n-k}\right)$$

since $\sigma^2 = 0$. Hence

$$||S_n||^2_{L^2} = \sum_{k=0}^{n-1} \sigma_{n,k} = \sum_{k=0}^{n-1} \mathcal{O}\left(\theta^{n-k}\right) = \mathcal{O}(1).$$

Therefore $\{S_n\}$ is weakly precompact. Let $B = w\text{–}\lim S_{n_j}$ for some subsequence n_j. Then

$$B(x) = \phi(x) + w\text{–}\lim_{n_j \to \infty} S_{n_j}(Tx) - w\text{–}\lim_{n_j \to \infty} \phi(T^{n_j} x).$$

By mixing for each $H \in L^2$ we have

$$\lim_{n\to\infty} \mu(H(\phi \circ T^n)) = \mu(H)\mu(\mu) = 0.$$

That is $w\text{–}\lim_{n\to\infty} \phi(T^n x) = 0$ and so $B(x) = \phi(x) + B(Tx)$ as claimed.
□

Lemma 6.8. *(a) For almost all $x_1 \in \mathbb{T}^2$, for almost all $x_2 \in W^s(x_1)$ we have*

$$B(x_2) - B(x_1) = \sum_{n=0}^{\infty} [\phi(T^n x_1) - \phi(T^n x_2)]. \tag{6.12}$$

(b) For almost all $x_1 \in \mathbb{T}^2$, for almost all $x_2 \in W^u(x_1)$ we have

$$B(x_2) - B(x_1) = \sum_{n=1}^{\infty} [\phi(T^{-n} x_1) - \phi(T^{-n} x_2)]. \tag{6.13}$$

Proof. We prove (a). The proof of (b) is similar.

We have

$$B(x_2) - B(x_1) = \sum_{n=0}^{N-1} [\phi(T^n x_1) - \phi(T^n x_2)] - B(T^N x_2) + B(T^N x_1).$$

Given m we can choose $\bar{B}_m, \bar{\bar{B}}_m$ such that $B = \bar{B}_m + \bar{\bar{B}}_m$, $\bar{B}_m \in C(\mathbb{T}^2)$, $||\bar{\bar{B}}_m||_{L^2} \leq 2^{-m}$. Next choose ε_m such that if $d(y_1, y_2) < \varepsilon_m$ then $|B(y_2) - B(y_1)| < 2^{-m}$. Now take $\bar{N}_m$ such that $\lambda^{-\bar{N}_m} < \varepsilon_m$. Set $x_2(t, x_1) = x_1 + te_s$. Then

$$\begin{aligned} &\operatorname{mes}\left((t, x_1) : |B(T^{\bar{N}_m} x_1) - B(T^{\bar{N}_m} x_2(t, x_1))| < \left(\frac{9}{10}\right)^m + \left(\frac{1}{2}\right)^m\right) \\ &\leq \operatorname{mes}\left((t, x_1) : |\bar{\bar{B}}_m(T^{\bar{N}_m} x_1) - \bar{\bar{B}}_m(T^{\bar{N}_m} x_2(t, x_1))| < \left(\frac{9}{10}\right)^m\right) \\ &\leq 2\frac{(1/2)^m}{(9/10)^{2m}} = 2\left(\frac{50}{81}\right)^m. \end{aligned}$$

Therefore for almost all $x_1 \in \mathbb{T}^2$ for almost all $x_2 \in W^s(x_1)$ there exists $\bar{m}$ such that for $m > \bar{m}$

$$|[B(x_2) - B(x_1)] - \sum_{n=0}^{\bar{N}_m - 1} [\phi(T^n x_2) - \phi(T^n x_1)]| \leq \left(\frac{9}{10}\right)^m + \left(\frac{1}{2}\right)^m.$$

Taking $m \to \infty$ we obtain the claimed result. □

Let

$$D = \{x \in \mathbb{T}^2 : \text{ for a. e. } \bar{x} \in W^s(x) \text{ (6.12) holds and for a. e. } \bar{x} \in W^u(x) \text{ (6.13) holds}\}.$$

$$\tilde{D} = \{x \in D : \text{ for a.e. } \bar{x} \in W^s(x) \cup W^u(x) \quad \bar{x} \in D\}.$$

Choose k_0 so large that if Γ is a horizontal curve of length δ with $d(\Gamma, \partial\tilde{\mathcal{P}}_{k_0}) \geq 3\delta$ then

$$\operatorname{mes}(x \in \Gamma : r_s(x) \geq 3\delta) \geq 0.99\delta$$

and if Γ is a vertical curve of length δ with $d(\Gamma, \partial\tilde{\mathcal{P}}_{k_0}) \geq 3\delta$ then

$$\mathrm{mes}(x \in \Gamma : r_u(x) \geq 3\delta) \geq 0.99\delta.$$

Let Π be a rectangle containing in one element of $\tilde{\mathcal{P}}_{k_0}$. Divide Π into squares of size $1/N$. Let D_N be the set of points x in $\tilde{D} \cap \Pi$ such that both $W^s(x)$ and $W^u(x)$ fully cross the square containing x. Consider two points $x_1, x_2 \in D_N$. Assume for a moment that x_1, x_2 belong to the same row and that the $1/N \leq d(x_1, x_2) \leq 2/N$. Let

$$\Gamma_N(x_j) = \bigcup_{y \in \hat{W}^s(x_j) \cap \tilde{D} : r^u(x) \geq 3/N} (W^u_{3/N}(y) \cap D),$$

where $\hat{W}^s(x_j)$ is the part of $W^s(x_j)$ containing inside our row. By our choice of k_0, $\Gamma_N(x_1) \cup \Gamma_N(x_2)$ has large measure. In particular, there exist $y_j \in \hat{W}^s(x_j) \cap \tilde{D}$ such that $y_2 \in W^u(y_1)$, and (x_j, y_j) satisfy (6.12). We claim that (y_1, y_2) satisfy (6.13). Indeed since $y_j \in D$ there exists y_3 such that (y_1, y_3) and (y_3, y_2) satisfy (6.13) but then (y_1, y_2) also satisfy (6.13). Since

$$B(x_2) - B(x_1) = [B(x_2) - B(y_2)] + [B(y_2) - B(y_1)] + [B(y_1) - B(x_1)]$$

we conclude from (6.12) and (6.13) that $|B(x_2) - B(x_1)| \leq C/N$. Taking arbitrary $x_1, x_2 \in D_N$ which are in the same row we can find $x_1 = z_0, z_1, z_2 \dots z_l = x_2$, such that $1/N \leq |z_j - z_{j-1}| \leq 2/N$ and $l \leq N|x_2 - x_1| + 1$. Accordingly

$$|B(x_2) - B(x_1)| \leq \sum_{j=1}^{l} |B(z_j) - B(z_{j-1})| \leq \frac{Cl}{N} \leq C(|x_2 - x_1| + \frac{1}{N}).$$

Hence if $x_1, x_2 \in D$ are in the same row and $|x_2 - x_1| \geq 1/N$ then

$$|B(x_2) - B(x_1)| \leq \bar{C}|x_2 - x_1|. \tag{6.14}$$

A similar conclusion holds if x_1 and x_2 at the same column. For general x_1, x_2 we can find z such that x_1 and z are at the same row and x_2 and z are at the same column and write

$$B(x_2) - B(x_1) = [B(x_2) - B(z)] + [B(z) - B(x_1)]$$

to conclude that (6.14) holds for arbitrary $x_1, x_2 \in \Pi \cap D$ with $|x_2 - x_1| \geq 1/N$. Next if $x_1, x_2 \in \tilde{D}$ then for large N we will have $x_j \in D_N$ and $|x_2 - x_1| \geq 1/N$ so (6.14) holds for all $x_1, x_2 \in \tilde{D} \cap \Pi$. In other words B can be modified on the set of measure 0, so that it becomes Lipshitz on Π and (6.11) holds almost everywhere. Covering every element of $\tilde{\mathcal{P}}_{k_0}$ by rectangles we see that there exists a version of B which is Lipshitz on each element of $\tilde{\mathcal{P}}_{k_0}$ such that (6.11) holds almost everywhere. But then by continuity it should hold everywhere. The proof of Proposition 6.5 is complete.

6.5. **Convergence to Brownian Motion.** Theorem 6.1 concerns the distribution of S_n for fixed n. Sometimes we would like to know a joint distribution of several S_{n_j} at the same time, for example, we may wish to compute $\mathbb{P}(\max_{n\leq N} S_n \leq L)$. Such questions are addressed by **Functional Central Limit Theorem.** Recall that a random process $\mathcal{B}(t)$ is called a **Brownian Motion** with variance parameter σ^2 if $\mathcal{B}$ has continuous paths, $\mathcal{B}(0) = 0$ and given $0 = t_0 < t_1 < t_2 \cdots < t_k$,the increments $\mathcal{B}(t_{j+1}) - \mathcal{B}(t_j)$ have normal distribution with zero mean and varaince $\sigma^2(t_{j+1} - t_j)$ and are independent of each other.

In the setting of Theorem 6.1 let $\mathcal{B}_N(t) = \frac{S_{Nt}}{\sqrt{N}}$ if Nt is integer with linear interpolation in between.

Theorem 6.9. *As $N \to \infty$, $\mathcal{B}_N(t)$ converges to the Brownian Motion with variance parameter σ^2 given by* (6.1).

According to [2] to show that $\mathcal{B}_N(t) \Rightarrow \mathcal{B}(t)$ we need to check two things.

First, we have to establish the convergence of finite dimensional distributions. That is, for each $0 = t_0 < t_1 < t_2 < \dots t_k$

$$(\mathcal{B}(t_1) - \mathcal{B}(t_0), \mathcal{B}(t_2) - \mathcal{B}(t_1), \dots \mathcal{B}(t_k) - \mathcal{B}(t_{k-1})) \Rightarrow (\mathcal{N}_1, \mathcal{N}_2 \dots \mathcal{N}_k)$$

where $\mathcal{N}_1, \mathcal{N}_2 \dots \mathcal{N}_k$ are indepepndent normal random variables with zero means and variances $\sigma^2(t_1 - t_0)$, $\sigma^2(t_2 - t_1)$, ..., $\sigma^2(t_k - t_{k-1})$.

Second, we need to prove tightness, that is, to show that for each $\varepsilon > 0$ there exists a compact set $\mathcal{K}_\varepsilon \subset C[0,1]$ such that $\mathbb{P}(\mathcal{B}_N(t) \in \mathcal{K}_\varepsilon) > 1 - \varepsilon$ for large N.

Let $N_j = Nt_j$. To check the convergence of finite dimensional distributions we need to show that

(6.15)
$$\mathbb{E}\left(\exp\left[\sum_m \xi_m (S_{N_m} - S_{N_{m-1}})\right]\right) \to \exp\left[-\frac{\sigma^2}{2}\sum_m \xi_m^2 (t_m - t_{m-1})\right].$$

But the proof of (6.15) is similar to (6.6) so it can be left to the reader.

To prove tightness we need an auxillary result.

Lemma 6.10.

$$\mathbb{E}((S_{N_2} - S_{N_1})^4) \leq L(n_2 - n_1)^2.$$

Proof. We have

$$(S_{N_2} - S_{N_1})^4) = \sum_{n_1,n_2,n_3,n_4=N_1}^{N_2} \phi(T^{n_1}x)\phi(T^{n_2}x)\phi(T^{n_3}x)\phi(T^{n_4}x).$$

To estimate the expectation of the above expression we may assume without the loss of the generality that $n_1 \leq n_2 \leq n_3 \leq n_4$. We claim that

$$\mathbb{E}(\phi(T^{n_1}x)\phi(T^{n_2}x)\phi(T^{n_3}x)\phi(T^{n_4}x)) = O(\theta^p) \tag{6.16}$$

where $p = \max(n_4 - n_3, n_2 - n_1)$. Indeed if $p = n_3 - n_3$ then the proof of (6.16) is similar to the proof of (6.7). If $p = n_2 - n_1$ then we use the same argument for T^{-1}. □

Define $\mathcal{K}_L =$

$$\left\{\psi \in C[0,1] : \psi(0) = 0 \text{ and for all } l \geq L,\ k \leq 2^l : \left|\psi\left(\frac{k+1}{2^l}\right) - \psi\left(\frac{k}{2^l}\right)\right| < 2^{-l/8}\right\}$$

Exercise 6.11. Show that $\mathcal{K}_L$ is compact in $C[0,1]$.

Thus it reamins to show that if L is large than $\mathbb{P}(\mathcal{B}_N \in \mathcal{K}_L)$ is close to 1. Let $n_{k,l}(N) = \frac{kN}{2^l}$. Then

$$\mathbb{P}\left(\left|\mathcal{B}_N\left(\frac{k+1}{2^l}\right) - \mathcal{B}_N\left(\frac{k}{2^l}\right)\right| \geq 2^{-l/8}\right) = \mathbb{P}\left(\left|S_{n_{k+1,l}(N)} - S_{n_{k,l}(N)}\right| \geq \sqrt{N}2^{-l/8}\right) \leq$$

$$\frac{\mathbb{E}\left(\left|S_{n_{k+1,l}(N)} - S_{n_{k,l}(N)}\right|^4\right)}{N^2/2^{l/2}} \leq C\frac{N^2/2^{2l}}{N^2/2^{l/2}} = C2^{-3l/2}.$$

Accordingly

$$\mathbb{P}\left(\exists k \leq 2^l : \left|\mathcal{B}_N\left(\frac{k+1}{2^l}\right) - \mathcal{B}_N\left(\frac{k}{2^l}\right)\right| \geq 2^{-l/8}\right) \leq C2^{-l/2}$$

and

$$\mathbb{P}\left(\exists l \geq L, \exists k \leq 2^l : \left|\mathcal{B}_N\left(\frac{k+1}{2^l}\right) - \mathcal{B}_N\left(\frac{k}{2^l}\right)\right| \geq 2^{-l/8}\right) \leq \tilde{C}2^{-l/2}$$

proving tightness. This completes the proof of Theorem 6.9.

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 $L \in SL_2(\mathbb{R})$ has positive elements then it is hyperbolic.*

This result is quite intuitive. If L has positive elements then the angle between Le_1 and Le_2 is less than $\frac{\pi}{2}$ and since due to area preservation

$$||Le_1||||Le_2|| \sin \angle(Le_1, Le_2) = 1$$

there should be some expansion. The analytic prove is also easy. If $L = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 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 products 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 $(x_1, y_1) = L(x_0, y_0)$ 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 L. The same argument works for compositions. Namely, if $L_1, L_2 \dots L_n$ are positive $SL_2(\mathbb{R})$ matrices and

$$v_n = L_n \dots L_2 L_1 v_0$$

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 $u_1 = (\xi_1, \eta_1), u_2 = (\xi_2, \eta_2)$ 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 symplectic matrix. We can write L in the block form with respect to the decomposition $\mathbb{R}^{2d} = V_1 \oplus V_2$ as $L = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$. The symplecticity 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 symplecticity 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 $L\mathcal{K} \subset \mathrm{Int}(\mathcal{K}) \cup \{0\}$.

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

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

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 shall use 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 show that $R \geq 0$. To this end consider an eigenvector $R\eta = \lambda\eta$. Take $\xi = a\eta$. Then $(\xi, \eta) \in \mathcal{K}$ 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. □

This proves shows in particular that if L is monotone then it is strictly monotone iff $P > 0$ and $R > 0$, that is, iff $L(V_j) \subset \text{Int}(\mathcal{K}) \cup \{0\}$.
Next let $L_1, L_2 \dots L_n$ be a sequence 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 $v_0 \in \text{Int}(\mathcal{K})$ 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 $\Lambda(L) = \min_{v \in \text{Int}(\mathcal{K})} \sqrt{\frac{Q(Lv)}{Q(v)}}$.
To compute $\Lambda(L)$ we shall use a decomposition

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

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 $\Lambda(M) = \Lambda(L)$. On the other hand

$$Q(Mv) = \langle \xi, \eta \rangle + \langle \eta, \eta \rangle + \langle T(\xi+\eta), (\xi+\eta) \rangle$$

$$= \sum_{j=1}^{d} \left[t_j \xi_2^2 + (1+2t_j)\xi_j\eta_j + (1+t_j)\eta_j^2 \right]$$

$$\sum_{\eta_j \geq 0} \left[\left(\sqrt{t_j}\xi_j - \sqrt{1+t_j}\eta_j\right)^2 \left(\sqrt{1+t_j} + \sqrt{t_j}\right)^2 \xi_j\eta_j \right]$$

$$+ \sum_{\eta_j < 0} \left[\left(\sqrt{t_j}\xi_j + \sqrt{1+t_j}\eta_j\right)^2 \left(\sqrt{1+t_j} - \sqrt{t_j}\right)^2 \xi_j\eta_j \right]$$

$$\geq m(L) \sum_j \xi_j \eta_j = m(L) Q(v)$$

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 achieved 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 $df(\mathcal{K}(x_n)) \subset \mathcal{K}(x_{n+1})$ where $\mathcal{K}(x)$ are the conses associated with the pair $(V_1(x), V_2(x))$. 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 $df : T_x M \to T_{fx} M$ 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 preserves a probability measure μ then, by Multiplicative 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 [16]. The main idea is to show that most orbits do not come to close to the singularities in the spirit of Lemma 5.6 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. Unfortunately those transversality conditions are not easy to verify in practise so the ergodicity is not yet proved 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 $D = \{(\xi, \xi)\}_{\xi \in \mathbb{R}^d}$ where we use the coordinates of Theorem 7.3. Then $E_j^- \cap D$ contains a nonzero vector v. For this vector $\lambda_+(x, v) \leq \lambda_j$. On the other hand in view of (7.5) and the Pointwise Ergodic Theorem we have

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

□

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 possessing 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*

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

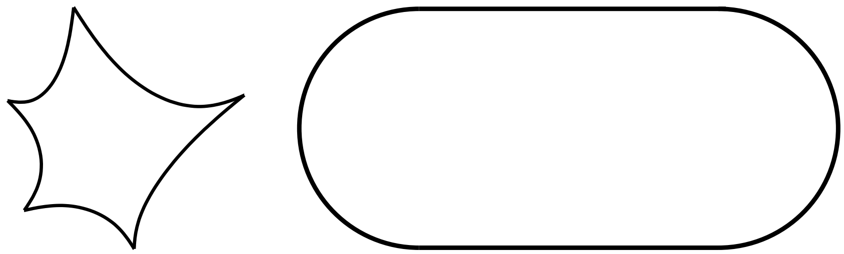

Figure 18. Two tables with nonzero Lyapunov exponents: dispersing billiard on the left and Bunimovich stadium on the right

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

Note that f preserves the form $\omega = \sin\phi ds \wedge d\phi$. The above matrix has all elements positive therefore df increases the quadratic 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 19. 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 [24]. Ergodicity of Bunimovich stadium is shown in [4]. Further properties of dispersing and defocusing billiards are discussed in [7].

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

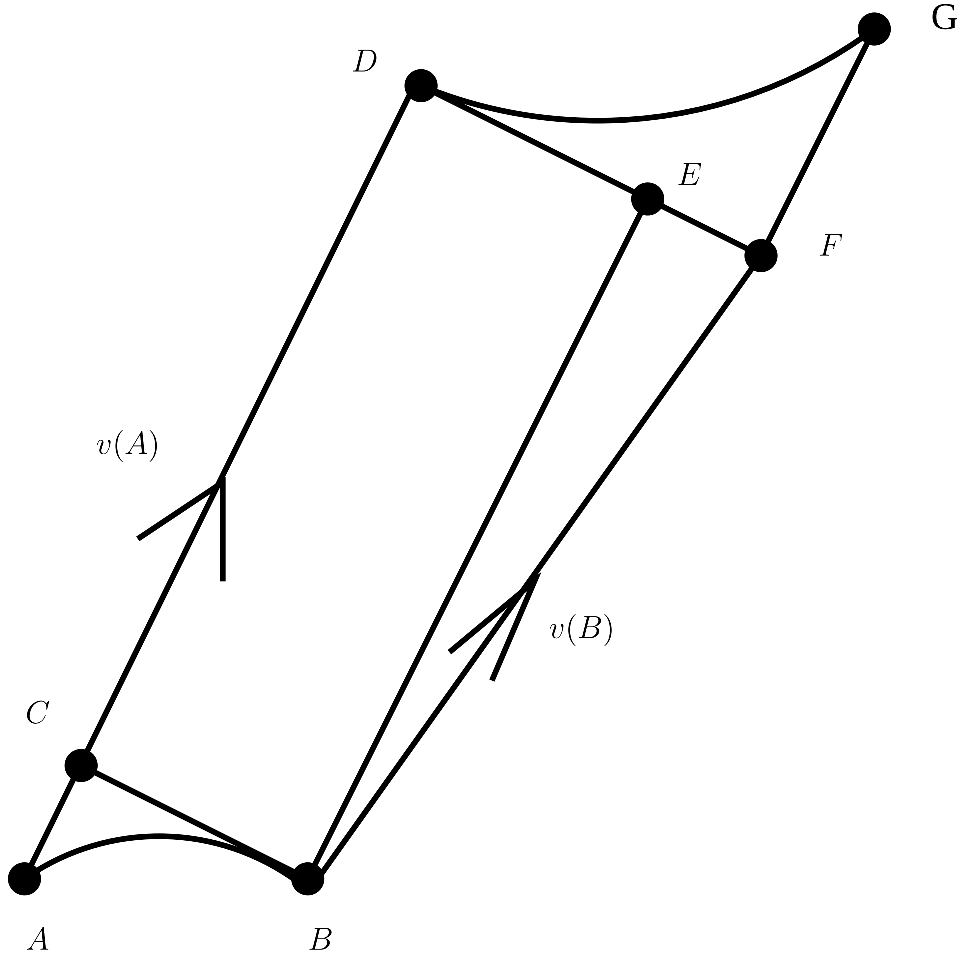

FIGURE 19. Computing $\dfrac{\partial s_1}{\partial s_0}$

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

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

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 $\frac{\partial v_{n+1}}{\partial t_n}$. Referring 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. □

(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

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

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 $u = v_1 - \tau g$ 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 x 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 proved in [17].

On the other hand if $m_1 = m_2$ then the particles just exchange their energy during the collisions so the function $I = \min(h_1, h_2)$ is the first integral of this system. One can also show [5] 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 proved in [26] while [23] showed that the conditions of Corollary 7.5 are satisfied for this system. One can also consider nonlinear potentials. [27] shows that the following conditions are sufficient for nonzero Lyapunov exponents

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

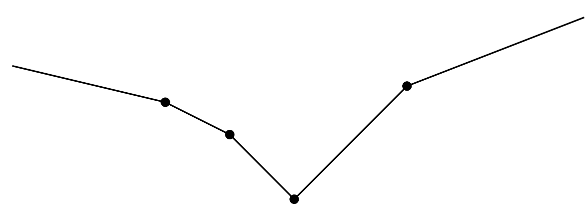

FIGURE 20. Wojtkowski wedge

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 [28] 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. Billingsley *Convergence of probability measures,* John Wiley & Sons, Inc., New York–London–Sydney, 1968 xii+253 pp.

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

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

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

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

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

[8] J. de Simoi *Stability and instability results in a model of Fermi acceleration,* Discrete Contin. Dyn. Syst. **25** (2009) 719–750.

[9] J. de Simoi *Fermi Acceleration in anti-integrable limits of the standard map,* arXiv:1204.5667.

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

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

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

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

[14] 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.

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

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

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

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

[19] Ortega R. *Asymmetric oscillators and twist mappings,* J. London Math. Soc. **53** (1996), no. 2, 325–342.

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

[21] 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.

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

[23] N. Simanyi *The characteristic exponents of the falling ball model,* Comm. Math. Phys. **182** (1996) 457–468.
[24] Ya. G. Sinai *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards,* Russ. Math. Surv. **25** (1970) 141–192.
[25] S. Tabachnikov *Geometry and billiards,* Student Math Library **30** (2005) AMS Providence RI.
[26] M. Wojtkowski *A system of one dimensional balls with gravity,* Comm. Math. Phys. **126** (1990) 507–533.
[27] M. Wojtkowski *A system of one dimensional balls in external field,* Comm. Math. Phys. **127** (1990) 425–432.
[28] M. Wojtkowski *Hamiltonian systems with linear potential and elastic constraints,* Fund. Math. **157** (1998) 305–341.