Statistical behaviour in the System of Two Falling Balls

Gábor Borbély

Supervisor: Péter Bálint Dr.
associate professor
BME Institute of Mathematics,
Department of Differential Equations

BME

2010

Contents

Chapter 1

Definitions

In this work we study a specific dynamical system, hence we start with a brief introduction to certain general concepts of dynamical systems. In mathematics a physical system is represented as a map $T : M \mapsto M$ where the set M is the phase space. In discrete time the map T evolves the system one step forward, in continuous time T_t depends on a parameter $t \in I$, where $I \subseteq \mathbb{R}$ is a specified interval. The flow T_t evolves the system t time forward and $T_{t+s} = T_t \circ T_s$. If the system is at the state $x \in M$ at the time 0, then it will be in the state $T_t x$ at the time t. $\{T_t\}_{t \in I}$, in both the continuous and the discrete case, is a commutative semi-group with respect to the action of the dynamics.

1.1 Statistical Properties

Sometimes it is not useful or interesting to study the smaller details of a system (motion) but one asks questions like: Where is a particle after a long time? What pattern can I see when I start the system from a typical configuration and look at it after a long time?

This type of approach can be familiar for example from statistical physics. To ensure that the above questions make sense we have to study certain stochastic properties. Ergodic theory studies the behavior of dynamical systems in the above mentioned way and it's main interest is the evolution of measures.

Definition 1.1.1 *Take a phase space: $\mathcal{M}$ and a dynamics (in discrete time): $T : \mathcal{M} \mapsto \mathcal{M}$. T is an <u>endomorphism</u> on the probability measure space $(\mathcal{M}, \Sigma, \mu)$ if T preserves μ.*

$$\begin{aligned} &\mu(\mathcal{M}) = 1 && \textit{and} \\ &\mu(T^{-1}A) = \mu(A) && \textit{for every } A \in \Sigma \end{aligned}$$

If T is invertible and T^{-1} is also an endomorphism with μ then we call it an automorphism and $\mu(TA) = \mu(A)$ also holds. An endomorphism $T : \mathcal{M} \mapsto \mathcal{M}$ with a measure μ means the objects together: $(T, \mathcal{M}, \Sigma, \mu)$

The most basic property is the ergodicity.

Definition 1.1.2 *A $T : \mathcal{M} \mapsto \mathcal{M}$ endomorphism with a measure μ is ergodic if: every invariant $f : \mathcal{M} \mapsto \mathbb{R}$ function is constants almost everywhere.*

$$\mu\left(\{x \in \mathcal{M} | f(x) = f(Tx)\}\right) = 1 \Rightarrow \mu\left(\{x \in \mathcal{M} | f(x) = c\}\right) = 1$$

For some $c \in \mathbb{R}$.

The existence of a nontrivial invariant set disproves ergodicity. If the set $A \subseteq \mathcal{M}$ is invariant ($\mu\left(A \bigtriangleup T^{-1}A\right) = 0$) and $0 < \mu(A) < 1$ then the function $\mathbb{I}_{\{x \in A\}}$ is invariant however not constant almost everywhere. The lack of a nontrivial invariant set is also an equivalent characterization of ergodicity [17].

In the definition 1.1.2 the function $f : \mathcal{M} \mapsto \mathbb{R}$ can be regarded as a measurement. In physics this means that we measure a certain quantity ($f(x) \in \mathbb{R}$) in a certain state of the system ($x \in \mathcal{M}$). Sometimes the function $f : \mathcal{M} \mapsto \mathbb{R}$ is called an *observable*. In the terminology of probability theory we think of an observable of an endomorphism as a random variable in the following way: let $(\mathcal{M}, \Sigma, \mu)$ be a probability space of an endomorphism and $f : \mathcal{M} \mapsto \mathbb{R}$ a μ-measurable function.

Definition 1.1.3 *Let $(T, \mathcal{M}, \Sigma, \mu)$ be an endomorphism and $f \in L^1_\mu(\mathcal{M})$ be an observable. The space average of f is:*

$$\mathbb{E}(f) := \int_{\mathcal{M}} f(x) \, \mathrm{d}\, \mu(x)$$

The time average of f is:

$$\hat{f} := \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i$$

$$\hat{f}(x) = \lim_{n \to \infty} \frac{f(x) + f(Tx) + \ldots + f(T^{n-1}x)}{n}$$

If the limit exists.
We call $(S_n f)(x) = f(x) + f(Tx) + \ldots + f(T^{n-1}x)$ the Birkhoff sum of f.

One can easily check that $\hat{f}$ is invariant (if it exists). Therefore in ergodic endomorphisms the $\hat{f}$ is constant almost everywhere.

The following theorem is the equivalent of Law of Large Numbers in Ergodic theory.

Theorem 1.1.4 (Birkhoff) *Let $(T, \mathcal{M}, \Sigma, \mu)$ be an endomorphism and $f \in L^1_\mu(\mathcal{M})$ an observable.*
If the T is ergodic then the time average does exist almost everywhere and

$$\hat{f} = \mathbb{E}(f) \qquad\qquad \textit{almost everywhere.}$$

In other words the time average converges to the space average (for μ-typical points).

Mixing is a stronger stochastic property than ergodicity, it is motivated by the independence (uncorrelatedness) of random variables.

Definition 1.1.5 *A $T : \mathcal{M} \mapsto \mathcal{M}$ endomorphism with a measure μ is mixing if: for any measurable $A, B \subseteq \mathcal{M}$*

$$\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B)$$

Another definition:

Definition 1.1.6 *A $T : \mathcal{M} \mapsto \mathcal{M}$ endomorphism with a measure μ is mixing if: for any square-integrable functions $f, g \in L^2_\mu(\mathcal{M})$*

$$\lim_{n\to\infty} \mathbb{E}\left((f \circ T^n)g\right) = \mathbb{E}(f)\mathbb{E}(g) \qquad \textit{that is}$$
$$\lim_{n\to\infty} \int_{\mathcal{M}} f(T^n x)g(x)\,\mathrm{d}\,\mu(x) = \int_{\mathcal{M}} f(x)\,\mathrm{d}\,\mu(x) \int_{\mathcal{M}} g(x)\,\mathrm{d}\,\mu(x)$$

If f, g are indicator functions of measurable sets, then this formula is the same as in definition 1.1.5. In [17] one can see the equivalence of these two definitions. The latter definition allows to define the *decay of correlations* for specific observables. Motivated by the subject of our study from now on I use more specific conditions on the below introduced objects. From now on I assume that the phase space is always a Riemannian manifold.

Definition 1.1.7 *A T endomorphism on a Riemannian manifold $\mathcal{M}$ with a Borel measure μ mixes with a polynomial rate $\alpha \geq 0$ if:*
for any Hölder continuous functions $f, g : \mathcal{M} \mapsto \mathbb{R}$
there exists a $c \geq 0$ such that the following holds (for every $n \in \mathbb{N}$)

$$\left|\mathbb{E}\left((f \circ T^n)g\right) - \mathbb{E}(f)\mathbb{E}(g)\right| \leq cn^{-\alpha}$$

In the definition $\mathcal{M}$ does not have to be a Riemannian manifold but it has to be a metric space in order to define Hölder continuity. Meanwhile $\mathcal{M}$, in the examples of physics, is usually a Riemannian manifold.
To define an *exponential rate of mixing* we use $c\lambda^n_{f,g}$ as a bound where $\lambda_{f,g}$ depends on the Hölder exponents of f and g.

Since the Law of Large Numbers holds for every ergodic endomorphism it is natural to study the CLT.

Definition 1.1.8 *For an ergodic endomorphism* $(T, \mathcal{M}, \Sigma, \mu)$ *the Central Limit Theorem holds if:*
for any Hölder continuous function $f : \mathcal{M} \mapsto \mathbb{R}$
$$\lim_{n\to\infty} \mu\left(\frac{S_n f}{\sqrt{n}\sigma} \le z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}}\,\mathrm{d}\,x$$
Where f *is supposed to be "centered":* $\mathbb{E}(f) = 0$
and σ *is the "variance".*
$$\sigma_f{}^2 = \sum_{n=-\infty}^{\infty} \mathbb{E}(f(f \circ T^n))$$
Here one can see that the fast decay of correlations is crucial in CLT. Without the series $\sum_n \mathbb{E}(f(f \circ T^n))$ being summable there is no chance to have CLT.

In this work I keep going on the way prescribed by Chernov and Zhang (see chapter 3) in order to prove polynomial decay of correlations ,and further properties, for the system of two falling balls. Some of the basic elements of such a proof were introduced in my Bsc thesis, now I detail the whole method.

1.2 Hyperbolic systems

In the study of stochastic properties, beside ergodicity, the hyperbolicity is an important property. This property is commonly used in order to study mixing and it's rate in ergodic systems.

To define a dynamics being hyperbolic, first we have to introduce Lyapunov exponents. Consider a map $T : \mathcal{M} \mapsto \mathcal{M}$.

Definition 1.2.1 *A point* $x \in \mathcal{M}$ *HAS a Lyapunov exponent:* χ *and a characteristic subspace:* $\mathcal{E}_x^\chi \subseteq \mathcal{T}_x\mathcal{M}$ *if:*
For every $v \in \mathcal{E}_x^\chi$
$$\lim_{n\to\infty} \frac{1}{n} \log \|\mathrm{D}_x T^n \cdot v\| = \chi$$
This definition is motivated by linear maps. Let us consider the simple case, when T is linear, and it's matrix A has all different eigenvalues, which we denote by $\lambda_1 < \cdots < \lambda_m$. In this simple case $\chi_i = \log|\lambda_i|$ and $\mathcal{E}_x^\chi$ is the direction of the subspace spanned by the corresponding eigenvector (u_i).
$$A^n u_i = \lambda_i^n u_i$$
$$\|A^n u_i\| = |\lambda_i|^n \|u_i\|$$
$$\log\|A^n u_i\| = n \log|\lambda_i| + \log\|u_i\|$$
$$\frac{1}{n}\log\|A^n u_i\| = \log|\lambda_i| + \frac{\log\|u_i\|}{n}$$
$$\lim_{n\to\infty} \frac{1}{n}\log\|A^n u_i\| = \log|\lambda_i|$$

In the general case of $T : \mathcal{M} \mapsto \mathcal{M}$ the existence of these objects is not guaranteed. The Oseledec theorem , see [6], states that the exponents exist almost everywhere, with respect to any invariant measure, if the T is differentiable on the phase space. The exponents also exist if the endomorphism satisfies some weaker properties, also mentioned in [6].

If $\chi < 0$ we call $\mathcal{E}_x^\chi$ a stable subspace of x, if $\chi > 0$ an unstable subspace, otherwise neutral. The vectors in a stable subspace shrink exponentially fast as $n \to \infty$ and grow exponentially fast as $n \to -\infty$. The vectors in an unstable subspace act vice verse. Let the manifold be m dimensional and let us denote the exponents by $\chi_1^x, \chi_2^x, \ldots, \chi_m^x$ with the corresponding subspaces: $\mathcal{E}_x^{\chi_1}, \mathcal{E}_x^{\chi_2}, \ldots, \mathcal{E}_x^{\chi_m}$. We group the subspaces according to the sign of the exponents. This way we have three subspaces for every point:

$$\mathbb{R}^m \cong \mathcal{T}_x\mathcal{M} = \underbrace{\left(\bigoplus_{\chi_i^x<0} \mathcal{E}_x^{\chi_i}\right)}_{\mathcal{E}_x^s} \bigoplus \underbrace{\left(\bigoplus_{\chi_i^x>0} \mathcal{E}_x^{\chi_i}\right)}_{\mathcal{E}_x^u} \bigoplus \underbrace{\left(\bigoplus_{\chi_i^x=0} \mathcal{E}_x^{\chi_i}\right)}_{\mathcal{E}_x^n}$$

Definition 1.2.2 *A map T acting on a Riemannian manifold $\mathcal{M}$ is <u>hyperbolic</u> if $\mathcal{E}_x^n$ is trivial but none of the $\mathcal{E}_x^s, \mathcal{E}_x^u$ are trivial (and do exist) for almost every point with respect to the Lebesgue measure.*

Notice that the definition requires the existence with respect to the Lebesgue measure, not the invariant measure. In general it can happen that the invariant measure is singular. However a system can be hyperbolic in the above sense even without a Lebesgue-continuous invariant measure.

In a two dimensional map the hyperbolicity means that there is one stable and one unstable direction in the Lebesgue-typical points. The directions of the (un)stable subspaces determine a vector field in the phase space. Roughly speaking the integral curves through these directions are the (un)stable manifolds. It is easy to determine these manifolds and exponents if the map is linear. For example in the CAT map (discussed in [17]) one has to calculate the eigendecomposition of the CAT matrix. But this is a more difficult task in a general case, when we can visualize the subspaces via cones.

Definition 1.2.3 *In a linear space A a <u>cone</u> $\mathcal{C}$ is a subset of A which is closed under the scalar multiplication. (a collection of directions)*

$$x \in \overbrace{\mathcal{C}}^{\subseteq A} \Rightarrow \lambda x \in \mathcal{C} \qquad \forall \lambda \in \mathbb{R}$$

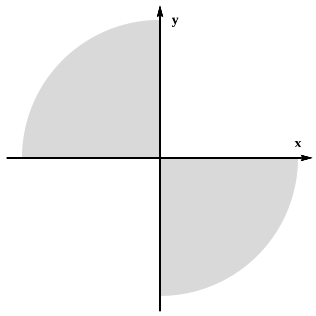
In a Riemannian manifold R a <u>cone in a point $x \in R$</u> is subset of the tangent space

$\mathcal{C}_x \subseteq \mathcal{T}_x R$ which is a cone in $\mathcal{T}_x R$.

A special way of defining cones (in a point of a Riemannian manifold R) is to take two vectors $v_1, v_2 \in \mathcal{T}_x R$, the sides of the cone, and let

$$\mathcal{C}_x\left(v_1, v_2\right) = \pm\{\alpha_1 v_1 + \alpha_2 v_2 \mid \alpha_1 \geq 0 \text{ and } \alpha_2 \geq 0\} \subseteq \mathcal{T}_x R$$

Figure 1.1:



Definition 1.2.4 *A cone field, of a Riemannian manifold R, is a set of cones $\{\mathcal{C}_x\}_{x\in R}$. We require continuity in the following sense.*
To compare two cone we need a metric: $d\left(\mathcal{C}_x, \mathcal{C}_y\right)$. Since we have a Riemannian manifold we can compare sets in tangent spaces of nearby points by identifying the tangent spaces with parallel translation. To compare sets in the same tangent space we take the unit vectors in a cone:

$$\mathcal{C}_x^1 = \{v \in \mathcal{C}_x : |v| = 1\}$$

and use the Hausdorff metric: $d_H\left(\mathcal{C}_x^1, \mathcal{C}_y^1\right)$
We will deal only with cones given by a pair of vectors and in this case the continuity is easier to define.

In our case R will be a bounded subset of $\mathbb{R}^2$ with a trivial Riemannian structure. To simplify the formalism of cones we identify the tangent spaces $\mathcal{T}_x\mathcal{M} \cong \mathbb{R}^2$ with the same $\mathbb{R}^2$ in which R lies. Thence it is possible to define a continuous cone field as

$$\{\mathcal{C}_x\left(v_1(x), v_2(x)\right)\}_{x\in\mathcal{M}}$$

Where $v_i : \mathcal{M} \mapsto \mathbb{R}^2$ are continuous functions determining the sides of the cones.

Definition 1.2.5 *Let $T : \mathcal{M} \mapsto \mathcal{M}$ be our endomorphism. A cone field $\{\mathcal{C}_x\}_{x\in\mathcal{M}}$ is invariant if T maps every cone inside the cone at the image point.*

$$\mathrm{D}_x T\left(\mathcal{C}_x\right) \subseteq \mathcal{C}_{Tx} \qquad \text{for every } x \in \mathcal{M}$$

In the definition T acts on a cone $\mathcal{C}_x \subseteq \mathcal{T}_x\mathcal{M}$ in a natural way: represent a cone with vectors from all of the directions and apply $\mathrm{D}_x T$ to all of them. $\{\mathrm{D}_x T \cdot v\}_{v\in\mathcal{C}_x}$ can be obtained as an other collection of directions in $\mathcal{T}_{Tx}\mathcal{M}$.

To see the relation between the cone fields and hyperbolicity we have to introduce a more strict definition of invariant cone fields. As a motivation consider the identity map of any compact set. The map is clearly non-hyperbolic, but any cone field is invariant.

Definition 1.2.6 *Let $\mathcal{C}_1, \mathcal{C}_2 \subseteq A$ be two cones in a metric space. Assume that $\mathcal{C}_1 \subseteq \mathcal{C}_2$. We say that $\mathcal{C}_2$ strictly contains $\mathcal{C}_1$ ($\mathcal{C}_1 \ll \mathcal{C}_2$) if:*

$$\partial\mathcal{C}_2 \cap \mathcal{C}_1 = \{0\}$$

Definition 1.2.7 *Let $T : \mathcal{M} \mapsto \mathcal{M}$ be an endomorphism. A cone field $\{\mathcal{C}_x\}_{x\in\mathcal{M}}$ is strictly invariant, (or in the hyperbolic setting unstable cone field), if T maps every cone strictly inside the cone at the image point.*

$$\mathrm{D}_x T\left(\mathcal{C}_x\right) \ll \mathcal{C}_{Tx} \qquad\qquad \textit{for every } x \in \mathcal{M}$$

The existence of a strictly invariant cone field implies the hyperbolicity, but it is not a necessary condition. [13], [5]

Definition 1.2.8 *Let $T : \mathcal{M} \mapsto \mathcal{M}$ an endomorphism. Let $\{\mathcal{C}_x\}_{x\in\mathcal{M}}$ be a strictly invariant cone field. A curve $\gamma(t) \in \mathcal{M}$, $t \in I \subseteq \mathbb{R}$ is an unstable curve, if the tangent vector of γ is in the corresponding unstable cone.*

$$\frac{\partial}{\partial t}\gamma(t) \in \mathcal{C}_{\gamma(t)}\ \forall t \in I$$

We can use these objects to approximate the unstable manifolds. Take a strictly invariant cone field and apply T to the cones (in the above mentioned way) n times. As $n \to \infty$ the image of the cones are getting more and more narrow (because they are mapped strictly inside themselves) and the limit gives the direction of the unstable subspace in every point. More precisely, to get the unstable direction of a point x one has to determine $\lim_{n\to\infty} \mathrm{D}_{T^{-n}x} T^n\left(\mathcal{C}_{T^{-n}x}\right)$. To get an approximate direction count $\mathrm{D}_{T^{-N}x} T^N\left(\mathcal{C}_{T^{-N}x}\right)$ for a sufficiently large N which gives a very narrow cone around $\mathcal{E}^u_x$. One can get the local shape of an unstable manifold by iterating any unstable curve forward.

To get the stable directions one has to find a backward-invariant family of cones and iterate them backwards. These can give the stable manifolds The (un)stable manifolds are the limit objects of (un)stable curves.

1.3 Billiards

Billiards are special dynamical systems. Let a point particle move freely in a bounded domain (billiard table) Q, where usually Q is in some Euclidean space or on a torus. The boundary ∂Q is piecewise smooth. The particle reflects from the boundary according to the law of specular reflection (angle of incidence equals angle of reflection). Thus billiard orbits are broken lines and the kinetic energy of the particle is an integral of motion. The speed of the ball is assumed to be 1. Let us define a Poincaré section: consider the moments, when the particle hits the boundary, infinitesimally after the collision. This way the dynamics is two dimensional. To introduce the phase space of the billiard map, we use two coordinates: the place, $x \in \partial Q$ (measured along the scatterer's perimeter in arch length), and the angle of the velocity vector, $-\frac{\pi}{2} \leq \varphi \leq \frac{\pi}{2}$ (angle of reflection).

Billiards on two dimensional tables are quite in the focus of the current interest. The behaviour of such a system is strongly determined by the shape of ∂Q. If the table is a circle or a rectangle, then the system is integrable (for a definition see [1]). If the boundary is piecewise concave (as viewed from the interior of Q), then the system has strong chaotic properties. This kind of billiards are called Sinai billiards, see [16]. Sinai billiards are known to have strong stochastic properties, see [5] for a comprehensive discussion of the modern theory of Sinai billiards.

If the table has (partly) convex boundary then one could suspect that the system does not have stochastic properties, since the wall of the table does not scatter the trajectories, but focuses them. However one can construct ergodic, convex billiard tables. The best known examples are stadia ([6]). In stadium billiards the CLT typically does not hold. There is an equivalent condition for CLT in [2].

There are several possible ways to generalize billiard systems. For example when considering n particles with various mass ratios. In this case the particles have a non-zero volume (radius) to ensure that they collide with non-zero probability. This system is the hard sphere gas model, and can be regarded as billiard in a higher dimensional domain. [14] [18]
Another way of generalization is to place the table in an external field and give charge to the particle. This way the inter-collisional trajectories are no longer straight line segments. [10] [11] [19]

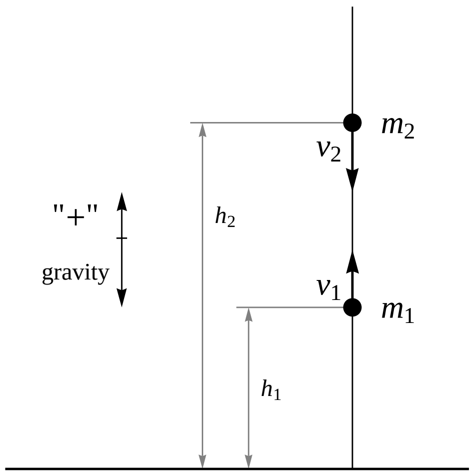
The system to be introduced below can be also regarded as a generalized billiard.

Chapter 2

The Concrete System

The system of falling balls can be regarded as a billiard. The table is one dimensional: a vertical half line bounded from below. In this line infinitesimally small balls move up and down under the force of gravity (g traditionally denotes $9.81\frac{m}{s^2}$). They bounce and collide totally elastically with each other and with the floor. Our system has only two particles of mass m_1 and m_2.

Figure 2.1:



The properties mentioned in this section have occurred in my Bsc thesis, and they are known, though not all of them is explicitly mentioned in the literature. [13] [15] [20]

Wojtkowski in [20] studied a general case: a system of n balls with different masses m_n. He proved hyperbolicity if the masses strictly decrease up the line.

$$m_1 > m_2 > \ldots > m_n$$

In [15] the hyperbolicity is proved under a weaker assumption: if the masses decrease non-strictly up the line, but there are at least two different masses.

From [13] we know that the system of two balls is ergodic in the case in which the lower ball is heavier ($m_1 > m_2$). This case is the main subject of this work. The second case is when the $m_1 = m_2$. In this case the balls exchange velocity which makes the system completely integrable. The third case is when $m_1 < m_2$. In this case one can observe KAM phenomena, which are also interesting but they are out of the framework of chaotic behaviour.

2.1 Expressing the Dynamics Explicitly

Let us express the state of the balls with usual physical quantities: h_1, h_2 are the hight of the lower and the upper ball, v_1, v_2 are the velocities. We neglect the air resistance therefore the total energy of the system $J := \frac{1}{2}m_1v_1^2 + m_1gh_1 + \frac{1}{2}m_2v_2^2 + m_2gh_2$ is an integral of motion. This motivates introducing the phase space $\hat{\mathcal{M}}$:

$$\left\{(h_1, v_1, h_2, v_2) \in \mathbb{R}^4 | 0 < h_1 < h_2, \frac{1}{2}m_1v_1^2 + m_1gh_1 + \frac{1}{2}m_2v_2^2 + m_2gh_2 = J\right\}$$

The dynamics act on $\hat{\mathcal{M}}$ in continuous time. In order to use our notations it is useful to discretize the system. Like Wojtkowski did in [20] we introduce the Poincaré section $\mathcal{M} = \left\{(h_1, v_1, h_2, v_2) \in \hat{\mathcal{M}} | h_1 = 0, v_1 > 0\right\}$. This means that we consider the moments when the lower ball hits the floor infinitesimally after the collision. Now we have a discrete map of a two dimensional phase space $T : \mathcal{M} \mapsto \mathcal{M}$. The purpose of the whole study is to understand the properties of this single map.

Instead of the usual moments we use the following coordinates of $\mathcal{M}$ (also from [20]):

$$h := \frac{1}{2}m_1v_1^2 \qquad \text{the total energy of the lower ball, since } h_1 = 0$$

$$z := v_2 - v_1$$

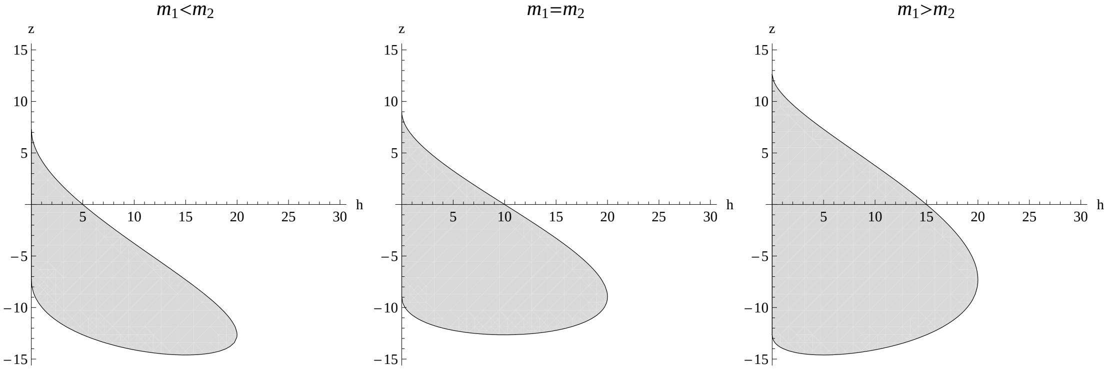
These coordinates seem to be suitable because they make our formulas simpler. It is also interesting to see that these quantities are invariants during the inter-collisional motion. Now the phase space is the following:

$$\mathcal{M} := \left\{(h, z) \in \mathbb{R}^2 | \, (0 < h < J) \wedge \left(J - h > \frac{1}{2}m_2{v_2}^2\right)\right\}$$

Where v_2 can be expressed from our coordinates: $v_2 = v_1 + z = \sqrt{\frac{2h}{m_1}} + z$
Notice that T is only piecewise continuous. Considering the collisions there are two cases: when the two balls collide before the lower one returns to the floor and the opposite.

$$\mathcal{M} \supseteq \mathcal{M}_1 = \{\text{configurations in which the balls will collide}\}$$

Figure 2.2: The phase space for different masses with $J = 20$



$$\mathcal{M} \supseteq \mathcal{M}_2 = \{\text{configurations in which the balls won't collide}\}$$

$$(h', z') = T(h, z) = \begin{cases} F_1(h, z) & \text{if } (h, z) \in \mathcal{M}_1 \\ F_2(h, z) & \text{if } (h, z) \in \mathcal{M}_2 \end{cases}$$

I denoted the action of T by prime.

With these notations F_i is a smooth $\mathcal{M}_i \mapsto \mathcal{M}$ map. To determine the sets $\mathcal{M}_i$ pretend for a while that the two balls move independently. To fall back to the floor the lower ball would take $t = 2\frac{v_1}{g}$ time, meanwhile the upper ball would reach the height $h_t = h_0 + v_2 t - \frac{1}{2}gt^2$ where h_0 is the starting height, which can be calculated from the potential energy of the upper ball: $h_0 = \frac{J-h-\frac{1}{2}m_2v_2^2}{gm_2}$. After the substitutions:

$$h_t = \frac{m_2 m_1 z \left(2\sqrt{\frac{2h}{m_1}} - z\right) - 2h\left(m_1 + m_2\right) + 2Jm_1}{2gm_1m_2} \tag{2.1}$$

I used Wolfram Mathematica for the substitution and also in the further calculations therefore they are not detailed.

Whether the balls collide or not is determined by the sign of h_t. Now the sets can be determined:

$$\mathcal{M}_1 := \{(h, z) \in \mathcal{M} | h_t < 0\}$$
$$\mathcal{M}_2 := \{(h, z) \in \mathcal{M} | h_t > 0\}$$

We do not consider the case $h_t = 0$ because it is an event with zero probability and no matter how we define the map T in this non-typical case, it does not effect the statistical properties of the system. Notice that, even though the gravity force appears in the formula (2.1), it does not effect the sign of it.

Now we can determine the maps F_i, starting with the easier case of F_2, when the balls do not collide. As the individual energies are conserved $h' = h$ and consequently

$v_1{}' = v_1$. Since the balls accelerate equally $v_2{}' = v_2 - 2v_1$ and $z' = v_2{}' - v_1{}' = (v_2 - 2v_1) - v_1 = z - 2v_1$.

$$\begin{pmatrix} h' \\ z' \end{pmatrix} = F_2(h,z) = \begin{pmatrix} h \\ z - 2v_1 \end{pmatrix} = \begin{pmatrix} h \\ z - 2\sqrt{\frac{2h}{m_1}} \end{pmatrix}$$

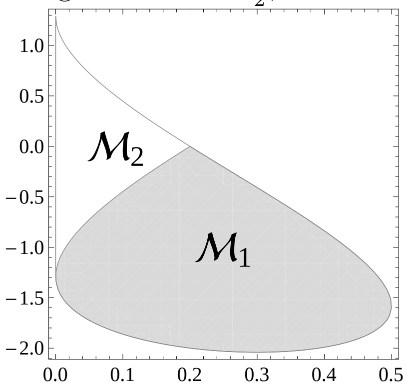
To determine F_1 one has to express the following quantities in the terms of h and z. The time when the particles reach the same height (and also the height itself), count the velocities in that moment, apply the rules of an elastic collision and calculate the additional time needed for the lower particle to hit the floor again. The values of h' and z' can be determined from the new velocities and kinetic energies. The result of this calculation is showed in the formula (2.2). For simplicity we make the assumption: $m_1 + m_2 = 1$ because these parameters only scale the system. From now on let $m_1 = m$ and $m_2 = 1 - m$ where $0 < m < 1$. Also J is often assumed to be $\frac{1}{2}$.

$$\begin{pmatrix} h' \\ z' \end{pmatrix} = F_1(h,z) = \begin{pmatrix} m\left(2J + \left(1 - 3m + 2m^2\right)z^2\right) - h \\ -2\sqrt{2}\sqrt{-\frac{h}{m} + 2J + \left(1 - 3m + 2m^2\right)z^2} - z \end{pmatrix} \tag{2.2}$$

It is interesting to notice that the gravity force does not occur in any of the formulas. Finally let's define the dynamics.

$$T(h,z) := \begin{cases} F_1(h,z) & \text{if } (h,z) \in \mathcal{M}_1 \\ F_2(h,z) & \text{if } (h,z) \in \mathcal{M}_2 \end{cases}$$

Figure 2.3: $J = \frac{1}{2}, m = 0.4$



2.2 Properties of $\mathrm{D}\,F_1$ and $\mathrm{D}\,F_2$

The Jacobian of the maps can be calculated:

$$\mathrm{D}\,F_1(h,z) = \begin{pmatrix} -1 & 2m\alpha z \\ \frac{\sqrt{2}}{m\sqrt{1 - \frac{h}{m} + \alpha z^2}} & -1 - \frac{2\sqrt{2}\alpha z}{\sqrt{1 - \frac{h}{m} + \alpha z^2}} \end{pmatrix}$$

$$\mathrm{D}\, F_2(h,z) = \begin{pmatrix} 1 & 0 \\ -\sqrt{\frac{2}{hm}} & 1 \end{pmatrix}$$

Where $\alpha := m_2(m_2 - m_1) = (1-m)((1-m)-m) = 1 - 3m + 2m^2$. The sign of α plays an important role because, as I mentioned in section 2.1, it characterizes the behavior of the system. α is determined by the mass ratio of the balls: if $m_1 > m_2$ then $\alpha < 0$, if $m_1 = m_2$, then $\alpha = 0$, otherwise $\alpha > 0$. We are interested in the case of negative α because this corresponds to ergodic dynamics.

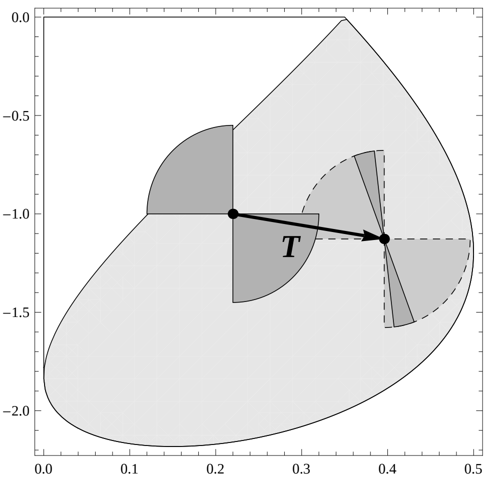
As you look at the matrices $\mathrm{D}\, F_1$ and $\mathrm{D}\, F_2$ you can see that the *Lebesgue measure* is preserved by the map T because both matrices have determinant 1. The map T with the normalized Lebesgue measure on $\mathcal{M}$ (denoted by $\mu := \frac{\lambda}{\lambda(\mathcal{M})}$) is an endomorphism. (Actually, it is an automorphism) The normalized Lebesgue measure is exactly the natural measure, induced by the Liouville measure of the continuous dynamics.

In the following section we will prove that the following (constant) cone field is invariant with respect to our system.

$$\{\mathcal{C}_x\,(e_1, -e_2)\}_{x\in\mathcal{M}} \qquad \text{where } e_i \text{ are the standard basis vectors in } \mathbb{R}^2$$

Note that $A = \mathcal{C}_x\,(e_1, -e_2)$ is the union of the lower, right quarter and the upper, left quarter of the plane.

Figure 2.4: The invariance of the cone field under F_1 ($\mathcal{M}_1$)



Let $\underline{x} = (x, y) \in A$ be a vector in the cone. We check the invariance for both F_1, and F_2. Let us consider F_1 first ($(h,z) \in \mathcal{M}_1$):

$$\mathrm{D}\, F_1 \cdot \underline{x} = \begin{pmatrix} 2myz\alpha - x \\ \frac{\sqrt{2}(x - 2myz\alpha)}{m\sqrt{-\frac{h}{m} + z^2\alpha + 1}} - y \end{pmatrix}$$

Notice that $\alpha < 0$ (ergodic case) and we will show with an indirect reasoning that $z < 0$. Assume the contrary. As the balls do collide ($\mathrm{D}\, F_1$ should act on $\underline{\mathrm{x}}$) and z is

positive. $z > 0$ should also hold in the moment infinitesimally before the collision, since z is an integral of inter-collisional motion. But $z = v_2 - v_1 > 0$ means that $v_2 > v_1$ and two particles cannot collide if the upper one moves faster upward. Hence we can state: $(h, z) \in \mathcal{M}_1 \Rightarrow z < 0$.
Consider the coordinates of $\mathrm{D}\, F_1 \cdot \underline{x}$ and notice that $xy < 0 \Leftrightarrow \underline{x} \in A$. By checking the signs of each term in the sum one can easily see that $\mathrm{D}\, F_1 \cdot \underline{x} \in A$. In the same way $\mathrm{D}\, F_2 \cdot \underline{x} \in A$ can be derived also.

$$\mathrm{D}\, F_2 \cdot \underline{x} = \begin{pmatrix} x \\ y - \frac{\sqrt{2}x}{\sqrt{hm}} \end{pmatrix}$$

The only problem is that these cones are not mapped strictly inside themselves (see definition 1.2.7). Considering $\mathrm{D}\, F_2$ we see that the cone field is invariant but

$$\begin{pmatrix} 1 & 0 \\ -\sqrt{\frac{2}{hm}} & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ z \end{pmatrix}$$

is fixed so the vertical side of the cone in $\mathcal{M}_2$ will not narrow down properly. The strictly containing property is *uniform* hyperbolicity, detailed in section 3.2.2. Indeed, the cones in $\mathcal{M}_1$ act properly since $\mathrm{D}\, F_1$ has two different eigenvalues separated from 1. The cones in $\mathcal{M}_2$ are also mapped inside the corresponding cones but not strictly, however, iterating T several times, sooner or later any point from $\mathcal{M}_2$ will step into $\mathcal{M}_1$ so, eventually, every cone will narrow down properly. Actually, uniform hyperbolicity of F_1 also fails, but only in the neighbourhood of a single point. See sections 3.2.2, and 4.3.

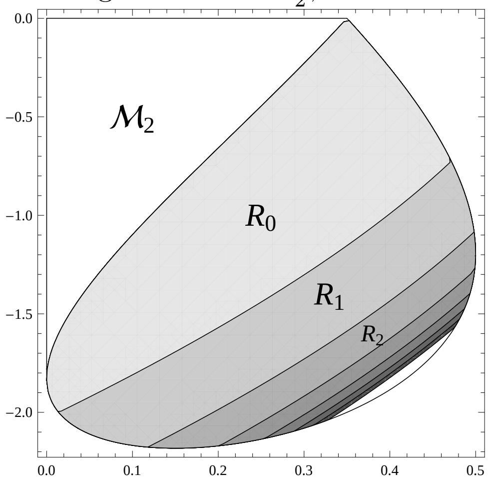
2.3 The First Return Map and the Singularity Stripes

Dealing with cones we saw an interesting phenomenon: take a point in $\mathcal{M}_2$, iterate it by T and wait until it pops into $\mathcal{M}_1$. Physically this means that the lower ball bounces for a while on the floor until it collides with the upper ball. This motivates defining the first return map of the set $\mathcal{M}_1$ in a standard way (see [17]):

$n_* : \mathcal{M}_1 \mapsto \mathbb{N}$	
$n_x := \min\{n \in \mathbb{N} \mid {F_2}^n (F_1(x)) \in \mathcal{M}_1\}$	the number if iterations needed to return
$R_n := \{x \in \mathcal{M}_1 \mid n_x = n\}$	the sets where the recurrence time is constant
$\hat{T} : \mathcal{M}_1 \mapsto \mathcal{M}_1$	
$\hat{T}x := {F_2}^{n_x} (F_1(x))$	

The invariant measure for $\hat{T}$ is the normalized Lebesgue measure on $\mathcal{M}_1$ (a conditional probability), denoted by $\hat{\mu}$.

Figure 2.5: $J = \frac{1}{2}, m = 0.7$



Starting from the set R_n the lower ball collides with the upped ball then hits the floor $n+1$ times until it collides again with the upper one. This phenomenon hastens the decay of correlations in the new system and also causes bigger rates of expansion in the Lyapunov exponents.

Notice that in the new dynamics the F_1 acts on every point therefore *there are no neutral steps*, when the matrix $\mathrm{D}\,\hat{T}$ does not stretch a vector. This property is related to uniformly hyperbolicity.

Chapter 3

The method of Chernov and Zhang

3.1 A general scheme for proving polynomial mixing

In order to prove that our system has polynomial (and summable) decay of correlations, which is also the first step to prove CLT, we follow the method to be introduced below.

In [9] Chernov and Zhang discuss a general method which allows to prove decay of correlations with polynomial rate in hyperbolic systems. Their work is partly based on the results of Young ([21], [23]).

Let $T : \mathcal{M} \mapsto \mathcal{M}$ be a mixing, non-uniformly hyperbolic dynamics with absolutely continuous invariant measure μ. Also, let $\mathcal{M}_1 \subseteq \mathcal{M}$ be a subset such that the first return map $\hat{T} : \mathcal{M}_1 \mapsto \mathcal{M}_1$ is uniformly hyperbolic. Let $R(x, T, \mathcal{M}_1)$ denote the time when the point $x \in \mathcal{M}$ reaches the set $\mathcal{M}_1 \subseteq \mathcal{M}$ at the first time by the dynamics T.

$$R(x, T, \mathcal{M}_1) = \min \left\{ i \geq 1 : T^i x \in \mathcal{M}_1 \right\}$$

Similarly, R can denote hitting times for other maps and other sets.

The main ingredient of the method is the existence of a horseshoe-like set $\Delta_0 \subseteq \mathcal{M}_1$, which contains stable and unstable manifolds and has a Cantor structure. Either by means of T or $\hat{T}$, returns to Δ_0 are always understood in a Markov-sense. This roughly means that Δ_0 can be partitioned into subsets $\Delta_{0,i}$ that extend Δ_0 along the stable direction, the points of $\Delta_{0,i}$ return simultaneously, and when they return, they extend Δ_0 along the unstable direction. The speed of mixing is highly correlated with the return times. Young in [21] proved that if the distribution

$\mu\left\{x \in \mathcal{M} | R\left(x, \Delta_0, \hat{T}\right) > n\right\}$ has an exponential tail bound, then the map $\hat{T}$ enjoys exponential decay of correlations. Later in [23] she proved that the polynomial tail bound implies polynomial mixing rate.

Let us consider first the uniformly hyperbolic map $\hat{T} : \mathcal{M}_1 \mapsto \mathcal{M}_1$. Following the previous work in [8], Chernov and Zhang formulated in [9] a set of conditions which guarantee the existence of Δ_0 and exponential tail bound for $R\left(x, \Delta_0, \hat{T}\right)$ (hence exponential decay of correlations for $\hat{T}$). Beside uniform hyperbolicity the essence is a Growth Lemma-like condition about the expansion rates of the unstable manifolds. See section 3.2.7 for precise formulation.

Now let us consider the original map $T : \mathcal{M} \mapsto \mathcal{M}$. If we have a Δ_0 in the set $\mathcal{M}_1 \subseteq \mathcal{M}$ which satisfies the conditions above, then we hope that, beside the first return map, the original map also enjoys decay of correlations, however with a slower rate (based on [23]). Chernov and Zhang proved that the same Δ_0 also satisfies a (slower) tail bound for the return times by the original dynamics if the return times of the first return map satisfies a polynomial tail bound. Precisely: if the set $\Delta_0 \subseteq \mathcal{M}_1$ satisfies that

$\mu(x \in \mathcal{M}_1 | R(x, \Delta_0, \hat{T}) > n) \leq \text{const } \theta^n$ (for some $\theta < 1$) and we know that

$\mu(x \in \mathcal{M}_1 | R(x, \mathcal{M}_1, T) > n) \leq \text{const } n^{-a-1}$ (for some $a > 0$) then

$\mu(x \in \mathcal{M} | R(x, \Delta_0, T) > n) \leq \text{const } \log n^{a+1} n^{-a}$

And this tail bound implies the decay of correlations for the original dynamics $T : \mathcal{M} \mapsto \mathcal{M}$.

Summarizing, if we have an ergodic and non-uniformly hyperbolic dynamics $T : \mathcal{M} \mapsto \mathcal{M}$ and we can localize a set $\mathcal{M}_1 \subseteq \mathcal{M}$ where the first return dynamics $\hat{T} : \mathcal{M}_1 \mapsto \mathcal{M}_1$ is uniformly hyperbolic, then the exponential mixing rate of $\hat{T}$ and the polynomial tail bound for the return times implies the polynomial mixing rate (with a logarithms correction) in the original dynamics.

Basically we have two major conditions to check:

$$\mu\left(x \in \mathcal{M}_1 | R(x, T, \mathcal{M}_1) > n\right) \leq \text{const } n^{-a-1} \qquad \forall n \geq 1 \tag{3.1}$$

For some $a > 0$ constant. Applied to our system, this requires to find the constant $a > 0$ and a $c > 0$ (which may depend on the mass ratio m) such that

$$\mu(R_n) \leq c\, n^{-a-2} \qquad \forall n \geq 1. \tag{3.2}$$

And the second is to prove the existence of the Δ_0 and prove that the return times have exponential tail bound. These conditions imply the following for Hölder continuous observables f and g.

$$\left|\mathbb{E}\left((f \circ T^n) g\right) - \mathbb{E}(f)\mathbb{E}(g)\right| \leq \text{const} \cdot (\log n)^{a+1} n^{-a} \qquad \forall n \geq 1 \tag{3.3}$$

If we have a Young tower and a summable decay of correlations, then the CLT follows (see [23]).

Since checking condition (3.2) in my Bsc thesis (with $a = 2$) there were several results obtained as a joint work with Péter Bálint and András Némedy Varga. We studied all the below described conditions and we made significant progress in order to prove them, however the proof is unfinished yet. In chapter 4 I will detail the results achieved and in chapter 5 the tasks left to solve.

3.2 Conditions for exponential mixing

Exponential decay of correlations has been proved for many systems by checking the below formulated conditions (or their slight variants), for example [3]. These conditions ensure the existence of the Young-tower and exponential mixing in certain systems. The following part is a word-by-word copy from [9] with my additional notes.

We suppose that the map $T : M \mapsto M$ preserves a mixing measure μ, M is an open domain in a two-dimensional smooth compact Riemannian manifold $\mathcal{M}$ with or without boundary, and T is hyperbolic. Here $T : M \mapsto M$ can denote a general endomorphism, we will use this for $\hat{T} : \mathcal{M}_1 \mapsto \mathcal{M}_1$

3.2.1 Smoothness

The map T is a C^2 diffeomorphism of $M \setminus \mathcal{S}$ onto $T(M \setminus \mathcal{S})$, where $\mathcal{S}$ is a closed set of zero Lebesgue measure. Usually, $\mathcal{S}$ is the set of points at which T either is not defined or is singular (discontinuous or not differentiable).

In the applications $\mathcal{S}$ is often a countable union of smooth compact curves.

3.2.2 Hyperbolicity

There exist two families of cones $\{\mathcal{C}_x^u\}_{x\in\mathcal{M}}$ (unstable) and $\{\mathcal{C}_x^u\}_{x\in\mathcal{M}}$ (stable) such that the unstable is strictly invariant and the stable is strictly backward invariant: $\mathrm{D}_x T(\mathcal{C}_x^u) \ll \mathcal{C}_{Tx}^u$ and $\mathrm{D}_x T(\mathcal{C}_x^s) \gg \mathcal{C}_{Tx}^s$ if the tangent map exists.

The expansions has to be uniformly bounded away from 1 in the following way. There is a $\Lambda > 1$ such that $\|\mathrm{D}_x T(v)\| \geq \Lambda\|v\|$ for every $v \in \mathcal{C}_x^u$ and $\|\mathrm{D}_x T^{-1}(v)\| \geq \Lambda\|v\|$ for every $v \in \mathcal{C}_x^s$ (whenever the tangent map exists).

The angle between C_x^u and C_x^s is uniformly bounded away from zero. Tan-

gent vectors to the singularity curves of T^m (which is $\bigcup_{i=0}^{m-1} T^{-i}(\mathcal{S})$) for m $>$ 0 must lie in stable cones, and tangent vectors to the singularity curves of T^{-m} (which is $\bigcup_{i=0}^{m-1} T^{i}(\mathcal{S})$) must lie in unstable cones.

The above conditions have the following standard consequences. For any T-invariant probability measure μ', almost every point $x \in M$ has one positive and one negative Lyapunov exponent. Also, almost every point x has one-dimensional local unstable and stable manifolds, denoted by $W^u(x)$ and $W^s(x)$, respectively.

By local unstable manifolds (LUM for brevity) we mean a curve $W^u(x)$ on which T^{-n} is defined smoothly for all $n > 0$, and $\mathrm{dist}(T^{-n}x, T^{-n}y) \to 0$ as $n \to \infty$ for all $y \in W^u(x)$, exponentially. Local stable manifolds (LSM) are defined via the shrinking in forward time.

3.2.3 SRB measure

The map T preserves an ergodic and mixing measure μ, whose conditional distributions on unstable manifolds are absolutely continuous.

3.2.4 Bounded curvature

The curvature of unstable manifolds is uniformly bounded by a constant $B \geq 0$.

3.2.5 Distortion bounds

Let $x \in M$ and $W^u(x)$ the unstable manifold of x and $n \in \mathbb{N}$ arbitrary, choose W^u so short that T^n is defined and smooth on W^u. Let $\Lambda(x)$ denote the expansion factor along W^u in the point x. Let $y \in W^u$ be on the same manifold.

$$\left| \log \prod_{i=0}^{n-1} \frac{\Lambda(T^i x)}{\Lambda(T^i y)} \right| \leq \psi\left(\mathrm{dist}\left(T^n x, T^n y\right)\right)$$

Where ψ is independent of W^u and $\lim_{s\to 0} \psi(s) = 0$.

Notice that this statement is about the derivative of the function $\Lambda(x)$. Let us consider the case $n = 1$.

$$\left| \log \frac{\Lambda(x)}{\Lambda(y)} \right| \leq \psi\left(\mathrm{dist}\left(Tx, Ty\right)\right)$$

$$\left| \log \Lambda(x) - \log \Lambda(y) \right| \leq \psi\left(\mathrm{dist}\left(Tx, Ty\right)\right)$$

In most applications the ψ function can be given explicitly, ψ usually has a form $c \cdot z^{\alpha}$, for some $c > 0$ and $0 < \alpha < 1$ constants. The choice of ψ will be similar in our case too. Let us manipulate the formula with this specified ψ.

$$\log \Lambda(x) - \log \Lambda(y) \leq c \cdot \text{dist}\,(Tx, Ty)^{\alpha}$$
$$\frac{\log \Lambda(x) - \log \Lambda(y)}{\text{dist}\,(Tx, Ty)^{\alpha}} \leq c$$

We have a Hölder-like condition about the logarithm of $\Lambda(x)$, which is much weaker, then requiring uniform upper bound on the derivative.

3.2.6 Absolute continuity

If W_1, W_2 are two small unstable manifolds close to each other, then the holonomy map $h : W_1 \mapsto W_2$ (defined by sliding along stable manifolds) is absolutely continuous with respect to the Lebesgue measures ν_{W_1}, ν_{W_2} , and its Jacobian is bounded:

$$\frac{1}{C} \leq \frac{\nu_{W_2}(h(W_1'))}{\nu_{W_1}(W_1')} \leq C$$

With some $C > 0$. $W_1' \subset W_1$ denotes the subset where h is defined.
See figure 4.12.

3.2.7 One-step growth of unstable manifolds

The below detailed condition is a simplified one, compared to the original statement. Chernov and Zhang built several sufficient conditions for checking the "one-step" Growth Lemma, this is not the most general form.

They formalize a three-part-condition and also show that the last two of these are satisfied easily in certain billiards, including ours. The third, and most interesting, part of the condition is the following:

$$\liminf_{\delta_0 \to 0} \sup_{W : |W| < \delta_0} \sum_i \Lambda_i^{-1} < 1$$

Here

- the supremum is taken over unstable manifolds
- i is indexing the connected components of $W \setminus \mathcal{S}$
- Λ_i are the minimum local expansions in the connected components of W.

Let us define the expansion in a particular point ($W = W^u(x)$).

$$\Lambda(x) = \Lambda_W(x) = \frac{|\mathrm{D}_x T \cdot w|}{|w|}$$

Where $w \in \mathcal{T}_x\mathcal{M}$ is a vector parallel to W at x. And the Λ_i is defined as

$$\Lambda_i = \Lambda_{i,W} = \inf_{x \in (i^{\text{th}} \text{ component of } W \setminus \mathcal{S})} \Lambda_W(x)$$

This condition is somehow a lower bound on the local expansion factors. The essence of the Growth Lemma is the following: the unstable manifolds expand as T acts on them, but the singularity set also cuts them apart, which makes them shorter. The construction of the Young tower requires that the unstable manifolds keep their length in average while we iterate them with T. The Growth Lemma, which is reduced to this seventh condition in the method, ensures the sufficient length.

Chapter 4

Rigorous results

In section 4.1 I prove the tail bound condition (see formula (3.2)) for the first return sets. This has been already done in my Bsc thesis, however it is included for completeness. In the rest of the chapter I detail the new results about the conditions 3.2.1-3.2.7. In chapter 5 I collect the ingredients, that are not completely solved, as of yet, and mention the directions along which we are trying to solve them.

Throughout the section $C, c, c', c'', c_1, c_2, C_1, C_2, d_1, d_2$ may denote some positive constants (non-zero and non-infinity), that depend only on m, the exact values of which are irrelevant.

4.1 Analysis of the first return sets

4.1.1 Bounding Functions

To estimate the measure of R_n I determined the curves bounding these sets. The curve that defines the boundary of the whole $\mathcal{M}$ can be derived by solving the equation:

$$J - h = \frac{1}{2} m_2 v_2^2$$
$$J - h = \frac{1}{2}(1-m)\left(\sqrt{\frac{2h}{m}} + z\right)^2$$

In section 2.1 we defined $\mathcal{M}$ via this quantity. Solving it for h defines a function $h = l_{m,J}(z)$ and solving it for z would define a function $z = l_{m,J}(h)$. The former one produces shorter formulas. Also assume that $J = \frac{1}{2}$.

$$l_m(z) = \frac{1}{2} m \left(1 \pm 2(1-m)z\sqrt{1-(1-m)mz^2} + z^2\alpha\right)$$

Since $z < 0$ ($(h, z) \in \mathcal{M}_1$) the "$-$" sign represents the greater quantity from the two options thus we should use "$-$" to get the curve that bounds $\mathcal{M}$ from the right. See

figure 4.1.
To express the curves separating the sets R_n and R_{n+1} let

$$(h'(h,z), z'(h,z)) = {F_2}^n(F_1(h,z))$$

and test whether (h', z') is on the boundary of $\mathcal{M}_1$ by solving the following equation for h.

$$h_t = \frac{m_2 m_1 z' \left(2\sqrt{\frac{2h'}{m_1}} - z'\right) - 2h' \overbrace{(m_1 + m_2)}^{1} + \overbrace{2J}^{1} m_1}{2g m_1 m_2} =$$

$$\frac{(1-m)mz'\left(2\sqrt{\frac{2h'}{m}} - z'\right) - 2h' + m}{2gm(1-m)} = 0$$

$$\Updownarrow$$

$$(1-m)mz'\left(2\sqrt{\frac{2h'}{m}} - z'\right) - 2h' + m = 0$$

Where h_t is from the formula (2.1) which characterizes $\mathcal{M}_1$.
After the substitutions and some simplifications the equation to solve takes the following form.

$$2\left(1 - \frac{h}{m} + z^2\alpha\right) - 1 +$$

$$(1-m)\left(2\sqrt{2}(n+1)\sqrt{1 - \frac{h}{m} + z^2\alpha} + z\right)\left(\sqrt{2}(2n+4)\sqrt{1 - \frac{h}{m} + z^2\alpha} + z\right) = 0$$

Let $F(h,z)$ denote the quantity $\sqrt{1 - \frac{h}{m} + z^2\alpha}$ and solve the equation for F.

$$2F^2 - 1 + (1-m)\left(2\sqrt{2}(n+1)F + z\right)\left(\sqrt{2}(2n+4)F + z\right) = 0$$

$$\Downarrow$$

$$F = \frac{(1-m)(2n+3)z \pm \sqrt{m^2z^2 - m\left(4n^2 + 12n + z^2 + 8\right) + (2n+3)^2}}{\sqrt{2}\left(4m\left(n^2 + 3n + 2\right) - (2n+3)^2\right)}$$

We choose the "$-$" sign, in order to have a positive solution. One can show that the other solution is negative. Now we can solve the following equation for h to eliminate F.

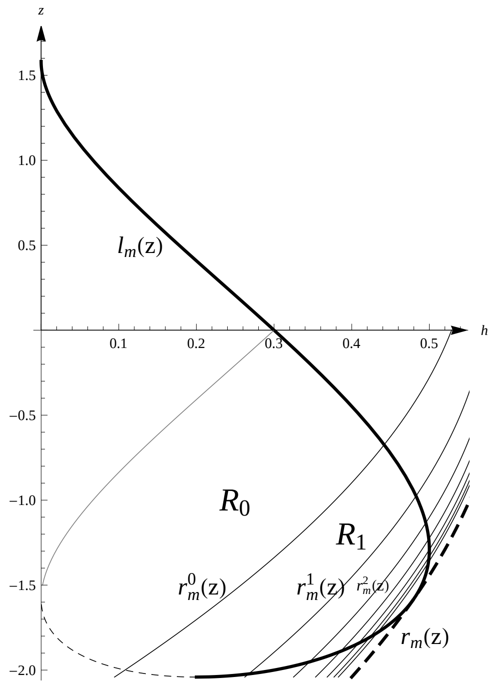
$$\overbrace{1 - \frac{h}{m} + z^2\alpha}^{F^2} = \left(\frac{(1-m)(2n+3)z - \sqrt{m^2z^2 - m\left(4n^2 + 12n + z^2 + 8\right) + (2n+3)^2}}{\sqrt{2}\left(4m\left(n^2 + 3n + 2\right) - (2n+3)^2\right)}\right)^2$$

Let $r_m^n(z)$ denote the solution, the bounding curve of R_n.

$$r_m^n(z) := m + mz^2\alpha - \tag{4.1}$$

$$-m\left(\frac{(1-m)(2n+3)z - \sqrt{m^2z^2 - m\left(4n^2 + 12n + z^2 + 8\right) + (2n+3)^2}}{\sqrt{2}\left(4m\left(n^2 + 3n + 2\right) - (2n+3)^2\right)}\right)^2$$

Figure 4.1: $J = \frac{1}{2}, m = 0.6$



Theoretically knowing these formulas would allow us to calculate the exact value of $\mu(R_n)$. But the integrals seem impossible to calculate. For example finding the intersection of $r_m^n(z)$ and $l_m(z)$ results a more than fourth order equation. Even if one could calculate the integrals the formulas were too long to treat them. My estimation is based on the asymptotic behaviour of $r_m^n(z)$.

4.1.2 A Simplified Model

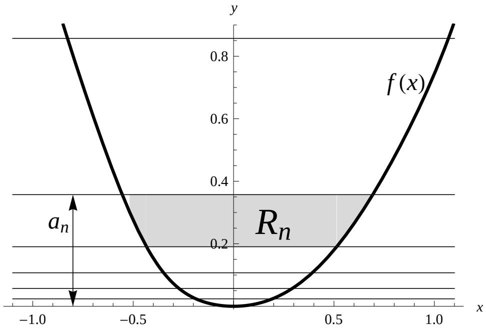
As the figures 4.1 and 2.5 suggest, there is a limit where the functions r_m^n tend to.

$$r_m(z) := r_m^\infty(z) = \lim_{n\to\infty} r_m^n(z) = m\left(1 + \alpha z^2\right)$$

The figures also suggest that the sets R_n are shaped like parallel stripes and these stripes accumulate on $r_m(z)$. We can make a strongly simplified model of the stripes. Take a sequence of parallel and horizontal lines crossed by the graph of a specific function $f(x)$. The height of the n^{th} line is a_n. $f(x)$ simulates $l_m(z)$ and the a_n (constant function) is an analogue of $r_m^n(z)$. First we will obtain an upper bound on the areas in this simple case, and consider the relevance of this model in section 4.1.4.

Notice that the tail bound for $\mu(R_n)$ is strongly determined by the order of the first non-vanishing derivative of f in 0. The higher the degree of the tangency, the less rapidly the areas of R_n decrease. To maintain the concept of figure 4.2 we make some restrictions on f. These restrictions are stronger requirements than what is

Figure 4.2: A simplified model

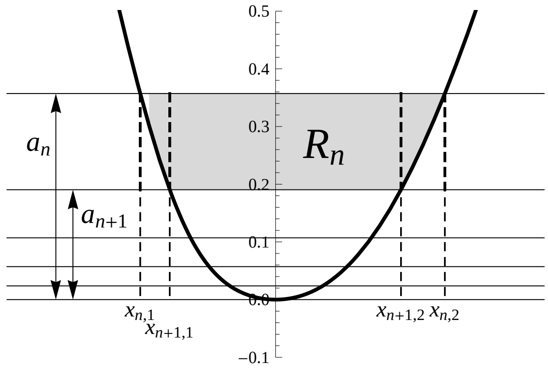


actually needed because we are interested in the behaviour of a concrete function and we do not want to formulate a general theorem. Suppose that f is continuous, and f is monotone decreasing for $x < 0$ and monotone increasing for $x > 0$ (consider only a small neighbourhood of 0). Also suppose that the following limit exists.

$$\lim_{x \to 0} \frac{f(x)}{x^k} = D \tag{4.2}$$

for some even number k and $D > 0$. This property implies that the first k derivatives of f exist in 0 and the value of the first non-vanishing derivative is $k! \cdot D$. Another consequence of these properties is that $f(0) = 0$ and f has a local minimum in 0, therefore the equation $f(x) = a_n$ has exactly two solutions (for a sufficiently large n): $x_{n,1} < 0 < x_{n,2}$.

Figure 4.3:



With these notations:

$$(x_{n+1,2} - x_{n+1,1})(a_n - a_{n+1}) \le \mu(R_n) \le (x_{n,2} - x_{n,1})(a_n - a_{n+1})$$

Now let us estimate $x_{n,2} - x_{n,1}$ from above.

$$\begin{aligned} f(x_{n,1}) &= a_n \\ (x_{n,1})^k \frac{f(x_{n,1})}{(x_{n,1})^k} &= a_n \end{aligned}$$

We are able to estimate the term $\frac{f(x)}{x^k}$ since the limit (4.2) exists. There exist constants (depending on f) c_1, c_2 and an $N_0 \in \mathbb{N}$ such that $0 < c_1 \leq D \leq c_2$ and $c_1 \leq \frac{f(x_{n,1})}{(x_{n,1})^k} \leq c_2$ holds for $n > N_0$ (notice that $x_{n,1} \to 0$ as $n \to \infty$).

$$\begin{aligned}
(x_{n,1})^k \frac{f(x_{n,1})}{(x_{n,1})^k} &= a_n \\
(x_{n,1})^k c_1 \leq a_n &\leq (x_{n,1})^k c_2 \\
(x_{n,1})^k &\leq \frac{a_n}{c_1} \\
-x_{n,1} = |x_{n,1}| &\leq \sqrt[k]{\frac{a_n}{c_1}}
\end{aligned}$$

In the same way we can estimate $x_{n,2}$. There exist other constants d_1, d_2 and an $M_0 \in \mathbb{N}$ such that $0 < d_1 \leq D \leq d_2$ and $d_1 \leq \frac{f(x_{n,2})}{(x_{n,2})^k} \leq d_2$ holds for $n > M_0$.

$$\begin{aligned}
(x_{n,2})^k \frac{f(x_{n,2})}{(x_{n,2})^k} &= a_n \\
(x_{n,2})^k d_1 \leq a_n &\leq (x_{n,2})^k d_2 \\
(x_{n,2})^k &\leq \frac{a_n}{d_1} \\
x_{n,2} &\leq \sqrt[k]{\frac{a_n}{d_1}}
\end{aligned}$$

Hence

$$\begin{aligned}
\mu(R_n) \leq & (x_{n,2} - x_{n,1})(a_n - a_{n+1}) \\
\mu(R_n) \leq & \left(\sqrt[k]{\frac{a_n}{d_1}} + \sqrt[k]{\frac{a_n}{c_1}} \right) (a_n - a_{n+1}) \qquad \text{if } n > \max\{N_0, M_0\} \\
\mu(R_n) \leq & \left(\sqrt[k]{\frac{1}{d_1}} + \sqrt[k]{\frac{1}{c_1}} \right) \sqrt[k]{a_n}\, (a_n - a_{n+1}) \qquad \text{let } a_n = \frac{1}{n^\alpha}
\end{aligned}$$

$$\mu(R_n) = \mathcal{O}\left(\frac{1}{n^{\frac{\alpha}{k}}} \frac{1}{n^{\alpha+1}} \right) = \boxed{\mathcal{O}\left(\frac{1}{n^{\alpha+1+\frac{\alpha}{k}}} \right)} \tag{4.3}$$

Notice that the bigger the k the slower the $\mu(R_n)$ tends to 0. It is also possible to obtain a lower bound estimation with the same argument. Even through it is not needed in our proof, in other studies, like in [2], a lower bound is necessary.

4.1.3 Straightening the Stripes

In the following section I will construct a map which distorts the plane in such a way that the graphs of the $r_m^n(z)$ become the translated versions of the graph of $r_m(z)$. This is the first step to transform the R_n into the form of figure 4.2. Notice that to obtain a correct estimation I have to prove that the Jacobian of this map is uniformly bounded away from both 0 and ∞.

Let π denote the transformation that straightens the graphs of r_m^n. We require that π has the following properties (for motivation see figure 4.2 or 4.3).

$$\pi : \mathcal{R}_m \mapsto \left\{(h,z) \in \mathbb{R}^2 | h \leq r_m(z)\right\}$$
$$\pi \begin{pmatrix} r_m^n(z) \\ z \end{pmatrix} = \begin{pmatrix} r_m(z) - a_n \\ z \end{pmatrix} \quad \text{if } n \geq n_0(m)$$

I will define $\mathcal{R}_m$ later. In order to determine a_n let us calculate the following limits.

$$\lim_{n\to\infty} \left(r_m(z) - r_m^n(z)\right) = 0 \qquad \text{(trivially)}$$
$$\lim_{n\to\infty} n\left(r_m(z) - r_m^n(z)\right) = 0$$
$$\lim_{n\to\infty} n^2\left(r_m(z) - r_m^n(z)\right) = \frac{m\left(-(1-m)z + \sqrt{1-m}\right)^2}{8(1-m)^2}$$
$$\lim_{n\to\infty} n^3\left(r_m(z) - r_m^n(z)\right) = \infty$$

I assumed that $1 > m > 0$ and $z < 0$ during the calculations. Suggested by this fact we make the choice $a_n = \frac{1}{n^2}$. At the moment a_n could be any sequence tending to 0, but this particular choice will guarantee that the transformation does not distort the area in a degenerate way, see section 4.1.4. We could also choose any other sequence a_n as long as $\lim_{n\to\infty} \frac{a_n}{1/n^2} = \text{const} \neq 0, \infty$.

One can see that π transforms only the h coordinate and the action of π depends only on n. Pretend for a while that the parameter n is continuous, as the formula of r_m^n allows to substitute n with any positive number. This kind of generalisation implies that the curves $\{r_m^\nu(z)\}_{\nu\in[0,\infty)}$ cover the set $\bigcup_{n\in\mathbb{N}_+} R_n$. Moreover they cover the region $\{(h,z) \in \mathbb{R}^2 | r_m^0(z) \leq h < r_m(z)\}$ (see figure 4.1). The domain of π will be derived from this set.

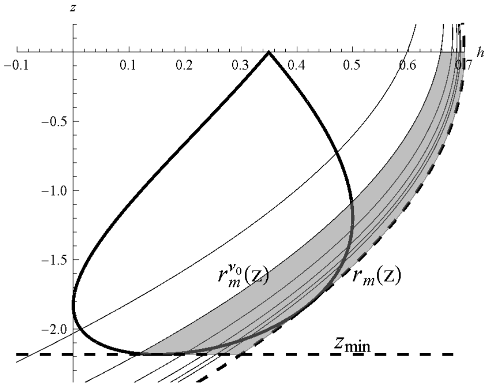
Even through the graphs $r_m^\nu(z)$ cover the above mentioned region, they do not foliate that. To see this note that, for example, $r_m^1(-6) = r_m^2(-6)$, if $m > \frac{5}{6}$. It can be checked that the curves $r_m^\nu(z)$ do foliate the region

$$\mathcal{R}_m := \{(h,z) \in \mathcal{M}_1 | z_{\min} < z < 0 \wedge r_m^{\nu_0}(z) \leq h \leq r_m(z)\}$$

for ν_0 large enough (the value of ν_0 depends on m). $z_{\min}$ is the height of the lower, horizontal side of the bounding rectangle of $\mathcal{M}_1$. One can determine $z_{\min}$ by checking the domain of $l_m(z)$.

$$l_m(z) = \frac{1}{2}m\left(1 - 2(1-m)z\sqrt{1-(1-m)mz^2} + z^2\alpha\right)$$
$$1 - (1-m)mz^2 > 0$$
$$z_{\min} = \frac{-1}{\sqrt{m(1-m)}}$$

Figure 4.4: $\mathcal{R}_m$



Thus any point $x \in \mathcal{R}_m$ can be represented with two coordinates: z, the coordinate we have already used, and the above introduced ν which labels the graph on which x lies. To determine π we introduce a function which determines the ν coordinate.

$$\nu : \mathcal{R}_m \mapsto [0, \infty)$$
$$r_m^{\nu(h,z)}(z) = h$$

With this notation π can be formalized easily:

$$\pi \begin{pmatrix} h \\ z \end{pmatrix} = \begin{pmatrix} r_m(z) - \frac{1}{\nu(h,z)^2} \\ z \end{pmatrix} \tag{4.4}$$

Notice that π is undefined if $\nu(h,z) = 0 \Leftrightarrow r_m^0(z) = h$. Also one can define π if $h = r_m(z) \Leftrightarrow \nu(h,z) = \infty$ in a continuous way: $\pi(r_m(z), z) := (r_m(z), z)$.

$$\pi : \mathcal{R}_m \mapsto \left\{ (h,z) \in \mathbb{R}^2 | h \leq r_m(z) \right\}$$
$$\pi(h,z) = \begin{cases} (r_m(z), z) & \text{if } \nu(h,z) = \infty \\ \left(r_m(z) - \frac{1}{\nu(h,z)^2}, z \right) & \text{otherwise} \end{cases}$$

We can express the function ν explicitly by solving the following equation (derived from (4.1.1)) for ν.

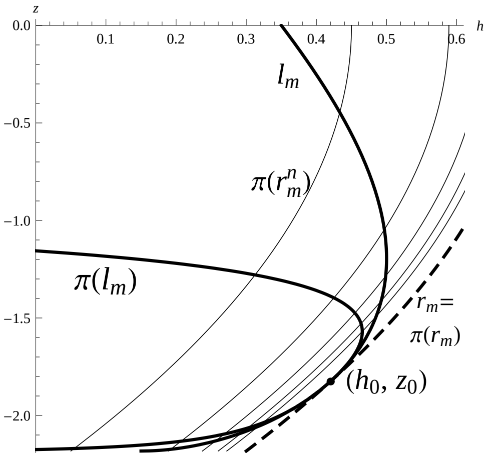
$$r_m^\nu(z) = h$$
$$\Updownarrow$$
$$\sqrt{1 - \frac{h}{m} + \alpha z^2} = \frac{(1-m)(2\nu+3)z - \sqrt{m^2 z^2 - m\left(4\nu^2 + 12\nu + z^2 + 8\right) + (2\nu+3)^2}}{\sqrt{2}\left(4m\left(\nu^2 + 3\nu + 2\right) - (2\nu+3)^2\right)}$$

This equation has two solutions, choose the positive one again. This is the function $\nu(h,z)$.

$$\nu(h,z) = \frac{\sqrt{2}\left(\sqrt{(1-m)\left(1 - 2F^2 m\right)} - (1-m)z\right) - 6F(1-m)}{4F(1-m)}$$

Where F is a short notation for $\sqrt{1-\frac{h}{m}+z^2\alpha}$ as previously.

Figure 4.5: The transformed stripes



As the second step in straightening the stripes we apply a translation-like transformation.

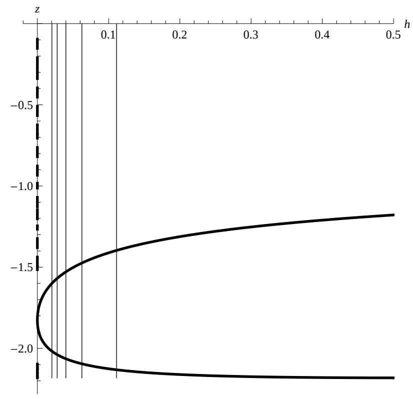
$$\phi\begin{pmatrix} h \\ z \end{pmatrix} := \begin{pmatrix} -(h - r_m(z)) \\ z \end{pmatrix}$$

Notice that

$$\phi\left(\pi\begin{pmatrix} h \\ z \end{pmatrix}\right) = \begin{pmatrix} \frac{1}{\nu(h,z)^2} \\ z \end{pmatrix}$$

The transformation $\phi \circ \pi$ distorts the figure 4.1 into the figure 4.6.

Figure 4.6: $\phi \circ \pi|_{\mathcal{R}_m}$



4.1.4 Overview

In section 4.1.3 we constructed a transformation which maps the sets R_n into the stripes of the simplified model. Now we have to prove the non-degeneracy of the Jacobian in order to make sure that the Lebesgue measures of the sets R_n are distorted only by constant. We also have to verify that the function $\phi(\pi(l_m(z)))$

satisfies the conditions in section 4.1.2, and substitute the quantities k and a_n into formula (4.3).

We do not have to prove non-degeneracy in the whole set $\mathcal{R}_m$, only in a neighbourhood of (h_0, z_0) since we are interested in the asymptotic behaviour of R_n.
Let $J_{h,z}$ denote the Jacobian of π, derived from formula (4.4). Notice that it differs from the Jacobian of $\phi \circ \pi$ only in it's sign, since the derivative matrix of ϕ has a determinant -1.

$$J_{h,z} = \mathrm{Det}\begin{pmatrix} \left(\frac{-1}{\nu(h,z)^2}\right)'_h & * \\ 0 & 1 \end{pmatrix} = \left(\frac{-1}{\nu(h,z)^2}\right)'_h$$

The formula of $\left(\frac{-1}{\nu(h,z)^2}\right)'_h$ is too long to copy here. I simply calculated the following limit.

$$\lim_{h\to r_m(z)} \left(\frac{-1}{\nu(h,z)^2}\right)'_h = \frac{8(1-m)^2}{m\left(-(1-m)z+\sqrt{1-m}\right)^2}$$

This limit is non-zero and the denominator is zero if $z = \frac{1}{\sqrt{1-m}}$, but this point is outside of the phase space, since $(h,z) \in \mathcal{M}_1 \Rightarrow z < 0$. Therefore the non-degeneracy holds in a small neighbourhood of the point (h_0, z_0). Here z_0 can be calculated from the equation of $l_m(z)$ and $r_m(z)$.

$$l_m(z) = r_m(z) \Rightarrow z_0 = \frac{-1}{\sqrt{1-m}}$$

To determine k we calculate the limit in formula (4.2).

$$\begin{aligned} &\lim_{z\to z_0} \phi(\pi(l_m(z))) = 0 \qquad \text{(trivially)} \\ &\lim_{z\to z_0} \phi(\pi(l_m(z)))\,(z-z_0)^{-1} = 0 \\ &\lim_{z\to z_0} \phi(\pi(l_m(z)))(z-z_0)^{-2} = 1-m \\ &\lim_{z\to z_0} \phi(\pi(l_m(z)))(z-z_0)^{-3} = \infty \end{aligned}$$

Hence $k=2$ and we can state the following.

$$\mu(R_n) \le \mathrm{const}\frac{1}{n^{\alpha+1+\alpha/k}} = \mathrm{const}\frac{1}{n^{2+1+2/2}} = \mathrm{const}\frac{1}{n^4}$$

Hence we proceed the method of Chernov and Zhang by proving the existence of the Young Tower in $\mathcal{M}_1$ to ensure that the dynamics $\hat{T}: \mathcal{M}_1 \mapsto \mathcal{M}_1$ mixes with exponential rate. To simplify our notations we forget the dynamics $T: \mathcal{M} \mapsto \mathcal{M}$ and concentrate on the first return map. From now on I denote the first return dynamics with $T: \mathcal{M} \mapsto \mathcal{M}$ instead of the original notation $\hat{T}: \mathcal{M}_1 \mapsto \mathcal{M}_1$.

4.2 Smoothness

The phase space is an open subset $\mathcal{M} \subseteq \mathbb{R}^2$ bounded by a piecewise smooth, closed curve, therefore it is automatically a Riemannian manifold. The set of singularities is the union of the boundaries ∂R_n, that is the graphs of the functions r_m^n. This is a countable collection of smooth curves, which is, in particular, closed and zero-measured. Piecewise diffeomorphism is also true, the map $T|_{R_n} = F_2^n \circ F_1$ is non-degenerate for any fixed n.

4.3 Hyperbolicity

We saw in section 2.2 that we have an invariant cone field and the matrix $\mathrm{D}_x T$ is hyperbolic everywhere.

$$\mathrm{D}\, F_1(h, z) = \begin{pmatrix} -1 & 2m\alpha z \\ \frac{\sqrt{2}}{m\sqrt{1-\frac{h}{m}+\alpha z^2}} & -1 - \frac{2\sqrt{2}\alpha z}{\sqrt{1-\frac{h}{m}+\alpha z^2}} \end{pmatrix}$$

The problem is the uniformity. The Jacobian is not hyperbolic, if $z = 0$. The only point in the closure of $\mathcal{M}$ where $z = 0$ is $Z = \left(\frac{m}{2}, 0\right)$. One can directly check, that Z is in $\overline{R_0}$.

$$\mathrm{D}_Z\, T = \mathrm{D}_Z\, F_1 = \begin{pmatrix} -1 & 0 \\ \frac{2}{m} & -1 \end{pmatrix}$$

which fails to be hyperbolic. Because of this matrix, the uniform expanding on the invariant cone field fails. One can eliminate this problem by taking the second iterate of T. **From now on we consider T^2 as an $\mathcal{M} \mapsto \mathcal{M}$ dynamics**. The point Z is not a fixed point of T.

$$T(Z) = F_1\left(\frac{m}{2}, 0\right) = \left(\frac{m}{2}, -2\right)$$

Therefore the product $\mathrm{D}_x\, T^2 = \mathrm{D}_x\, T \cdot \mathrm{D}_{Tx}\, T$ always contains a hyperbolic term even if $x \in \partial\mathcal{M}$. The map T^2 maps the unstable cones strictly, and uniformly, inside themselves. That will give us the constant of the uniform hyperbolicity: $\Lambda > 1$.

4.4 SRB measure

We saw in section 2.2 that the normalized Lebesgue measure is preserved (which is trivially absolutely continuous). From [13] we know ergodicity, and from [20] the hyperbolicity, but mixing is also required. The method, with which Liverani and Wojtkowski proved ergodicity, also proves that every power of T ($T^n : \mathcal{M} \mapsto \mathcal{M}$) is ergodic, too. Then Pesin theory guarantees that such a hyperbolic system is automatically mixing. See section 6.7 in [5].

4.5 A technical lemma

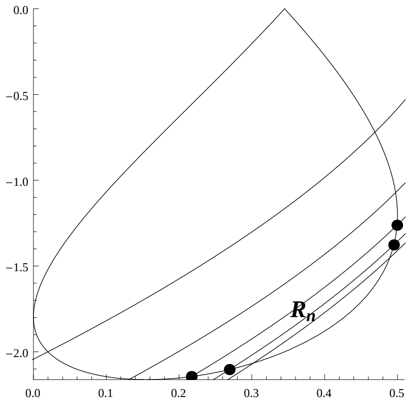
In this section I will give a tail-bound for $F = 1 - \frac{h}{m} + \alpha z^2$, which occurs in many formulas. It is very important to know the asymptotic behaviour of this quantity, since the case $F = 0$ characterizes the accumulation point of the sets R_n. From now on we denote this memorable point by G.

$$G := \left(2m(1-m), \frac{-1}{\sqrt{1-m}}\right) \in \overline{\mathcal{M}}$$

$$(h, z) \to G \Leftrightarrow n \to \infty \Leftrightarrow F \to 0$$

Let us consider $F|_{R_n}$, and search for extremum for a fixed n. We will show by direct differentiation that there is no local extremum of F inside R_n. Also there is no local extremum on the curves r_m^n, r_m^{n-1} and l_m, which bound the region R_n. Therefore the global extremum must take place in one of the "corners".

Figure 4.7: The potential global extrema of F



Let us check the gradient of F.

$$\partial_h F(h, z, m) = -\frac{1}{m} \neq 0$$
$$\partial_z F(h, z, m) = 2\alpha z$$

Now let us check the derivative of F along the curves r_m^n, l_m.

$$\partial_z F(r_m^n(z), z, m) = (\text{positive term}) \cdot \left(-(1-m)(2n+3) + \frac{-(1-m)mz}{\sqrt{m^2 z^2 - m\left(4n^2 + 12n + z^2 + 8\right) + (2n+3)^2}}\right)$$

$$\partial_z F(l_m(z), z, m) = \frac{-(1-m)\left(\sqrt{L}(2m-1)z - 2L + 1\right)}{\sqrt{L}}$$

Where $L = 1 - (1-m)mz^2$, which is positive when $z > \frac{-1}{\sqrt{m(1-m)}}$. We have already seen this condition in section 4.1.3, this is the lower edge of the bounding rectangle of the phase space, see figure 4.4. Now we have to check the signs. I made the simplification in Mathematica assuming the trivial conditions $\frac{1}{2} < m < 1, z < 0, n > 0$ and $z > \frac{-1}{\sqrt{(1-m)m}}$.

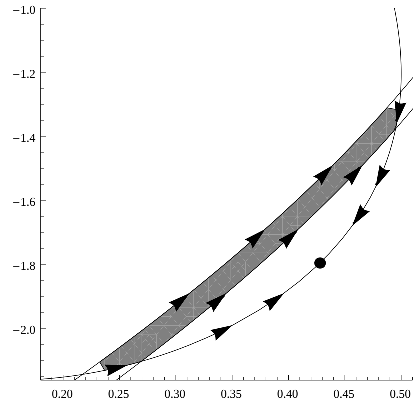
$$-(1-m)(2n+3) + \frac{-(1-m)mz}{\sqrt{m^2z^2 - m\left(4n^2+12n+z^2+8\right)+(2n+3)^2}} < 0$$

returns `True` with the additional assumption $\sqrt{\frac{(2n+3)^2-4m(n^2+3n+2)}{(1-m)m}} > -z$. This condition is fulfilled if $n \to \infty$ because the coefficient of the leading term (n^2) is positive. The following relations also returned `True`.

$$\sqrt{L}(2m-1)z - 2L + 1 > 0 \text{ , with the additional assumption } z < \frac{-1}{\sqrt{1-m}}$$
$$\sqrt{L}(2m-1)z - 2L + 1 < 0 \text{ , with the additional assumption } z > \frac{-1}{\sqrt{1-m}}$$

$\frac{-1}{\sqrt{1-m}}$ is the z coordinate of G, the accumulation point. According to the results above we understood the increases and decreases of F along ∂R_n. Now we know

Figure 4.8: F decreases in the marked directions

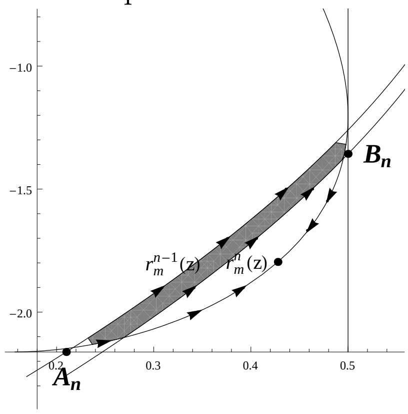


where the maximum and minimum of F are but we can not express the intersections $r_m^n(z) = l_m(z)$. However the derivatives can be analysed even if (h, z) is outside of the phase space. Therefore to get upper and lower bounds on the value of $F|_{R_n}$, we may use the values of F in the following points.

$$A_n = \left(r_m^{n-1}\left(\frac{-1}{\sqrt{(1-m)m}}\right), \frac{-1}{\sqrt{(1-m)m}}\right)$$
$$B_n = (r_m^n(z_n), z_n)$$

where z_n is such that $r_m^n(z_n) = \frac{1}{2}$. The solution of $r_m^n(z) = \frac{1}{2}$ can be explicitly

Figure 4.9: The points used in the estimation



calculated, the only thing to do is to calculate the following limits.

$$\lim_{n\to\infty} n^2 F(B_n, m) = \frac{1-2\sqrt{2m}+2m}{16m(1-m)}$$
$$\lim_{n\to\infty} n^2 F(A_n, m) = \frac{1+\sqrt{m}}{8m\,(1-\sqrt{m})}$$

As a consequence $\exists\ c_1, c_2$ such that

$$\frac{c_1}{n^2} \leq F \leq \frac{c_2}{n^2} \qquad \forall n \geq 1 \tag{4.5}$$

4.6 Small distortion bound

In the original distortion bound on has to prove the following for all unstable manifolds W^u and two points on them, uniformly.

$$\left|\log \prod_{i=0}^{n-1} \frac{\Lambda(T^i x)}{\Lambda(T^i y)}\right| \leq \psi\left(\mathrm{dist}\left(T^n x, T^n y\right)\right) \tag{4.6}$$

Where ψ is independent of W^u and $\lim_{s\to 0}\psi(s) = 0$.

Instead of this, we will state and prove a more general statement. We will say, that a curve γ is a regular unstable curve, if it is an unstable curve and has curvature at most B (the constant in the to-be-proved bounded curvature property, see 3.2.4). Let $\lambda_\gamma(x)$ denote the expansion on such a curve, at the point $x = \gamma(t) \in \mathcal{M}$.

$$\lambda_\gamma(x) = \frac{|\mathrm{D}_x \cdot \gamma'(t)|}{|\gamma'(t)|}$$

The desired statement is the following.

$$\left|\log \prod_{i=0}^{n-1} \frac{\lambda_\gamma(T^i x)}{\lambda_\gamma(T^i y)}\right| \leq \psi\left(\mathrm{dist}\left(T^n x, T^n y\right)\right) \tag{4.7}$$

In particular, LUMs are special regular unstable curves. Let $\gamma = W^u(x)$ be the unstable manifold of x, then

$$\lambda_{W^u(x)}(x) = \Lambda(x)$$

Hence (4.7) would imply (4.6), but some ingredients (such as condition 3.2.4) of this statement are not completed yet.

In this section we only prove (4.7) for the eigenvalues, instead of expansion rates More precisely, in (4.8) we formulate a statement, which we call **small distortion bound**, on the regularity of the unstable eigenvalues along the unstable curves. Let $\lambda(x) := |\mathrm{eig}\,(\mathrm{D}_x T)|$, the unstable eigenvalue, then for every point x, y on the same regular unstable curve, on which T^n is smooth, the following holds.

$$\left|\log \prod_{i=0}^{n-1} \frac{\lambda(T^i x)}{\lambda(T^i y)}\right| \leq \psi\left(\mathrm{dist}\left(T^n x, T^n y\right)\right) \tag{4.8}$$

Notice that the only difference between the small and the original distortion bound is the angle between the unstable directions and the direction of unstable eigenvectors. We will consider this at section 5.2.

Now let us consider (4.8). As I mentioned in section 3.2.5 we will obtain $\psi(z) = c \cdot z^\alpha$. Let $\gamma \in R_n$ for some n, and $x, y \in \gamma$. As the first step we prove the following.

$$\left|\frac{\lambda(x)}{\lambda(y)} - 1\right| \leq c \cdot \mathrm{dist}(x, y)^\alpha \tag{4.9}$$

Notice that the function $\lambda(x)$ is smooth on R_n, for a fixed n, therefore this property automatically holds, but the constant depends on n. We have to concentrate on the limit $n \to \infty$, since the first finitely many constant can be merged into one constant. We will use three estimations to conclude (4.9) (consider that $x, y \to G$ and $n \to \infty$).

$$\|\mathrm{Grad}\,\lambda(x)\| \leq c \cdot n^4 \qquad \forall\, x \in R_n \tag{4.10}$$

$$c_1 \cdot n^2 \leq \lambda(x) \leq c_2 \cdot n^2 \qquad \forall\, x \in R_n \tag{4.11}$$

$$\mathrm{dist}(x, y) \leq \frac{c}{n^3} \qquad \forall\, x, y \in \gamma \subset R_n \tag{4.12}$$

Where γ is unstable curve.

Let us check (4.11) for $(h,z) \in R_n$.

$$F_2^n\left(F_1(h,z)\right) = \left(-h + m\left(1+\alpha z^2\right), -z - 2(n+1)\sqrt{2}\sqrt{1 - \frac{h}{m} + \alpha z^2}\right) =$$
$$= \left(-h + m\left(1+\alpha z^2\right), -z - 2(n+1)\sqrt{2}\sqrt{F}\right);$$
$$Tr = Trace\left(\mathrm{D}_{h,z}\, F_2^n\left(F_1(h,z)\right)\right) = -2 - \frac{\sqrt{2}(n+1)z\alpha}{\sqrt{F}};$$
$$\lambda(h,z) = \left|\frac{1}{2}\left(Tr - \sqrt{Tr^2-4}\right)\right| =$$
$$= 1 + \frac{\sqrt{2}\left(\sqrt{(n+1)z\alpha\left(\sqrt{2}\sqrt{F} + (n+1)z\alpha\right)} + (n+1)z\alpha\right)}{\sqrt{F}} =$$
$$= (n+1)^2\left(\frac{1}{(n+1)^2} + \frac{\sqrt{2}\left(\sqrt{z\alpha\left(\frac{\sqrt{2F}}{n+1} + z\alpha\right)} + z\alpha\right)}{\sqrt{F}(n+1)}\right)$$
$$\frac{\lambda(h,z)}{(n+1)^2} = \frac{1}{(n+1)^2} + \frac{\sqrt{2}\left(\sqrt{z\alpha\left(\frac{\sqrt{2F}}{n+1} + z\alpha\right)} + z\alpha\right)}{\sqrt{F}(n+1)}$$

Using the statement of the technical lemma (4.5) we can estimate this expression from both above and below. Also using the asymptotic equality of n and $n+1$.

$$\frac{1}{(n+1)^2} + \frac{\sqrt{2}\left(\sqrt{z\alpha\left(\frac{c_1\cdot 1/n}{n+1} + z\alpha\right)} + z\alpha\right)}{c_2\frac{1}{n}(n+1)} \leq \frac{\lambda(h,z)}{(n+1)^2} \leq$$
$$\leq \frac{1}{(n+1)^2} + \frac{\sqrt{2}\left(\sqrt{z\alpha\left(\frac{c_2\cdot 1/n}{n+1} + z\alpha\right)} + z\alpha\right)}{c_1\frac{1}{n}(n+1)}$$

Now take $n \to \infty \Leftrightarrow (h,z) \to G$.

$$C_1 \leq \liminf_{n\to\infty} \frac{\lambda(h,z)}{(n+1)^2} \leq \limsup_{n\to\infty} \frac{\lambda(h,z)}{(n+1)^2} \leq C_2 \qquad \text{for some } 0 < C_1 < C_2 < \infty$$
$$\Updownarrow$$
$$d_1 \cdot n^2 \leq \lambda(h,z) \leq d_2 \cdot n^2 \qquad \text{for some } 0 < d_1 < d_2 < \infty$$

Which is exactly (4.11).

Now let us consider (4.10). Let $x = (h,z) \in R_n$ again.

$$\frac{\operatorname{Grad}\lambda(x)}{\lambda(x)} =$$
$$\frac{(n+1)\alpha}{Fm\sqrt{(n+1)z\alpha\left(\sqrt{2}\sqrt{F} + (n+1)z\alpha\right)}}\begin{pmatrix} \frac{z}{2} \\ m-h \end{pmatrix}$$

One can see that the asymptotically important term is the following.

$$\frac{(n+1)\alpha}{Fm\sqrt{(n+1)z\alpha\left(\sqrt{2}\sqrt{F}+(n+1)z\alpha\right)}} \overset{(4.5)}{\leq}$$

$$\frac{(n+1)\alpha}{c_1\cdot 1/n^2\cdot m\sqrt{(n+1)z\alpha\left(\sqrt{2c_1}\cdot 1/n+(n+1)z\alpha\right)}} =$$

$$\frac{n^2\alpha}{c_1\cdot m\sqrt{z\alpha\left(\frac{\sqrt{2c_1}}{n(n+1)}+z\alpha\right)}}$$

As $n\to\infty$ the leading term is n^2.

$$\frac{\|\mathrm{Grad}\,\lambda(x)\|}{\lambda(x)} \leq c\cdot n^2$$

$$\Updownarrow\ (4.11)$$

$$\|\mathrm{Grad}\,\lambda(x)\| \leq c\cdot n^4$$

Which is exactly (4.10).

Now we have to prove (4.12), which is $\mathrm{dist}(x,y)\leq\frac{c}{n^3}$. By calculating the limit $\lim_{n\to\infty} n^2\left(r_m(z)-r_m^n(z)\right)$ in section 4.1.3 we have already proved a similar statement. What we proved is that, for a fixed z, the horizontal slices of the sets R_n decay in third order. Indeed, if a sequence a_n tends to 0 with $\frac{1}{n^2}$ rate, then the sequence of differences $a_{n+1}-a_n$ tends to 0 with $\frac{1}{n^3}$ rate. Actually, one can also calculate the following limit directly.

$$\lim_{n\to\infty} n^3\left(r_m^{n+1}(z)-r_m^n(z)\right) = \frac{m\left(1-\sqrt{1-mz}\right)^2}{4(1-m)}$$

However one can construct a sequence of points $x_n, y_n\in R_n$, such that $\mathrm{dist}(x_n,y_n)\geq\frac{c}{n}$, because $\frac{1}{n}$ is the asymptotic width of R_n along the stable direction. Now we use the condition, that the curve, on which x and y are, is an unstable curve. Since the graphs r_m^n are uniformly transversal to the unstable cone, the estimation $\frac{c}{n^3}$ is only distorted by a constant.

Finally we can combine (4.10)-(4.12) to get (4.9).

$$\|\mathrm{Grad}\,\lambda(x)\| \overset{(4.10)}{\leq} c\cdot n^4 \qquad \forall\, x\in R_n$$

$$\Downarrow$$

$$|\lambda_n(x)-\lambda_n(y)| \leq c\cdot n^4\,\mathrm{dist}(x,y) \qquad \forall\, x,y\in R_n$$

Hence for $x, y \in \gamma$, where $\gamma \subset R_n$ is an unstable curve:

$$\begin{aligned}
|\lambda(x) - \lambda(y)| &\leq c \cdot n^4 \operatorname{dist}(x, y) \\
&\leq c \cdot n^4 \cdot \frac{c_1}{n^3} = && \text{using (4.11) and (4.12)} \\
&= c_2 \cdot n \\
&\leq C \cdot \lambda(x)^{\frac{1}{2}} && \text{using (4.11)}
\end{aligned}$$

For any $0 < \alpha < 1$:

$$\begin{aligned}
|\lambda(x) - \lambda(y)| &= |\lambda(x) - \lambda(y)|^{\alpha} \, |\lambda(x) - \lambda(y)|^{1-\alpha} \\
&\leq \left(c_1 \cdot \underbrace{n^4}_{c \cdot \lambda(y)^2} \operatorname{dist}(x, y) \right)^{\alpha} \left(c_2 \cdot \lambda(y)^{\frac{1}{2}} \right)^{1-\alpha} \\
&= c \cdot \lambda(y)^{2\alpha + (1-\alpha)\frac{1}{2}} \operatorname{dist}(x, y)^{\alpha}
\end{aligned}$$

Let α be the solution of $2\alpha + (1 - \alpha)\frac{1}{2} = 1$, which is $\frac{1}{3}$. Then dividing by $\lambda(y)$.

$$\left| \frac{\lambda(x)}{\lambda(y)} - 1 \right| \leq c \cdot \operatorname{dist}(x, y)^{\frac{1}{3}}$$

Hence (4.9) is proved.

Now let us extend it for T^2, since the uniform hyperbolicity requirements forces us to consider T^2 instead of T. The two-step-expansion is $\lambda^{(T^2)}(x) = \lambda(x)\lambda(Tx)$.

$$\begin{gathered}
\left| \frac{\lambda(x)\lambda(Tx)}{\lambda(y)\lambda(Ty)} - 1 \right| = \left| \frac{\lambda(x)\lambda(Tx)}{\lambda(y)\lambda(Ty)} - \frac{\lambda(x)}{\lambda(y)} + \frac{\lambda(x)}{\lambda(y)} - 1 \right| \leq \\
\left| \frac{\lambda(x)\lambda(Tx)}{\lambda(y)\lambda(Ty)} - \frac{\lambda(x)}{\lambda(y)} \right| + \left| \frac{\lambda(x)}{\lambda(y)} - 1 \right| \leq \left| \frac{\lambda(x)}{\lambda(y)} \right| \left| \frac{\lambda(Tx)}{\lambda(Ty)} - 1 \right| + \left| \frac{\lambda(x)}{\lambda(y)} - 1 \right| \leq \\
\frac{C_2}{C_1} c \cdot \operatorname{dist}(Tx, Ty)^{\frac{1}{3}} + c \cdot \operatorname{dist}(x, y)^{\frac{1}{3}} \\
\leq c' \cdot \operatorname{dist}(Tx, Ty)^{\frac{1}{3}} \leq c' \cdot \operatorname{dist}(T^2x, T^2y)^{\frac{1}{3}}
\end{gathered}$$

as $\operatorname{dist}(x, y) \leq \operatorname{dist}(Tx, Ty)$, since x and y lie on some unstable curve.

Finally let us derive the small distortion bound (4.8). Because of the hyperbolicity of T^2 we know that $\operatorname{dist}((T^2)^k x, (T^2)^k y) \leq \operatorname{dist}((T^2)^n x, (T^2)^n y)$ for any $k < n$. Moreover $\operatorname{dist}((T^2)^k x, (T^2)^k y)^{\frac{1}{3}} \leq \frac{\operatorname{dist}((T^2)^n x, (T^2)^n y)^{\frac{1}{3}}}{\sqrt[3]{\Lambda}^{n-k}}$, as the minimal expansion of T^2 on the unstable curves is $\Lambda > 1$. We also use that

$$\begin{gathered}
\left| \frac{\lambda(x)}{\lambda(y)} - 1 \right| \leq c_1 \operatorname{dist}(x, y)^{\alpha} \\
\Updownarrow \\
|\log(\lambda(x)) - \log(\lambda(y))| \leq c_2 \operatorname{dist}(x, y)^{\alpha}
\end{gathered}$$

as by (4.11) we know that $C_1 \leq \frac{\lambda(x)}{\lambda(y)} \leq C_2$, if $x, y \in R_n$, independently of n.

Using the above properties, we plant the result into higher iterates of T^2.

$$\left|\log \prod_{i=0}^{n-1} \frac{\lambda^{(T^2)}((T^2)^i\, x)}{\lambda^{(T^2)}((T^2)^i\, y)}\right| \leq$$
$$\sum_{i=0}^{n-1} \left|\log\left(\lambda^{(T^2)}((T^2)^i x)\right) - \log\left(\lambda^{(T^2)}((T^2)^i y)\right)\right| \leq$$
$$\sum_{i=0}^{n-1} \operatorname{dist}\left((T^2)^i x, (T^2)^i y\right)^{\frac{1}{3}} \leq$$
$$\operatorname{dist}\left((T^2)^n x, (T^2)^n y\right)^{\frac{1}{3}} \sum_{i=0}^{n-1} \sqrt[3]{\Lambda}^{-(n-i)} \leq$$
$$c \cdot \operatorname{dist}\left((T^2)^n x, (T^2)^n y\right)^{\frac{1}{3}}$$

4.7 Involution

There is a commonly used technique in billiards, called involution, to express the inverse of the dynamics (T^{-1}) in terms of the forward dynamics (T).

Definition 4.7.1 *Let $(T, \mathcal{M}, \Sigma, \mu)$ be an automorphism. An $I : \mathcal{M} \mapsto \mathcal{M}$ bijection is an* <u>*involution*</u> *if $I^{-1} = I$ and*

$$I \circ T \circ I = T^{-1}$$

In billiard maps there is a physically interesting manifestation of an involution. Let I act the following way: reflect the velocity of the particle through a line which is perpendicular to the scatterer, at the moment of the collision ($\varphi' = -\varphi$). This gives the incoming velocity vector with opposite orientation. Indeed, the preimage of a point of the phase space can be given as acting the involution, let the particle fly with the new velocity and when the particle hits a scatterer, then let the involution act again. This is truly the inverse of the dynamics, since the laws of specular reflection are the same in both forward and backward time.

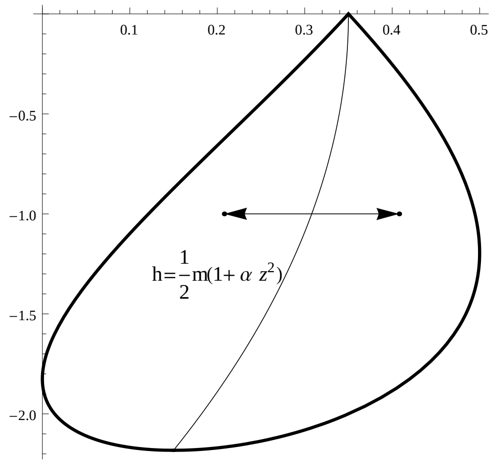
We have also found an involution for our system, which has an analogous nice physical manifestation. The action of I is the following: let F_1 act on the point (h, z) and then reverse the velocity of the upper particle. Reversing the velocity stands for reversing time. If we want to find out the backward trajectory of a falling ball, then we should reverse it's velocity, let it fly according to the standard gravity laws, and after the flight reverse it's velocity once again. The velocity of the lower ball should not be reversed, because it is always on the floor and we consider the moment when it has already bounced and is moving upwards. Notice that the action of F_1

- which is needed to ensure $I(x) \in \mathcal{M}$ - makes this involution unique, compared to the standard involution in billiards. Expressing the action of I in terms of h and z gives.

$$I(h, z) := \left(m\left(1 + \alpha z^2\right) - h, z\right) \tag{4.13}$$

On the phase space this is a reflection through the curve $\frac{1}{2}m\left(1 + \alpha z^2\right)$, along the vertical direction, therefore the property $I^{-1} = I$ automatically holds.

Figure 4.10: The involution is a reflection on the phase space

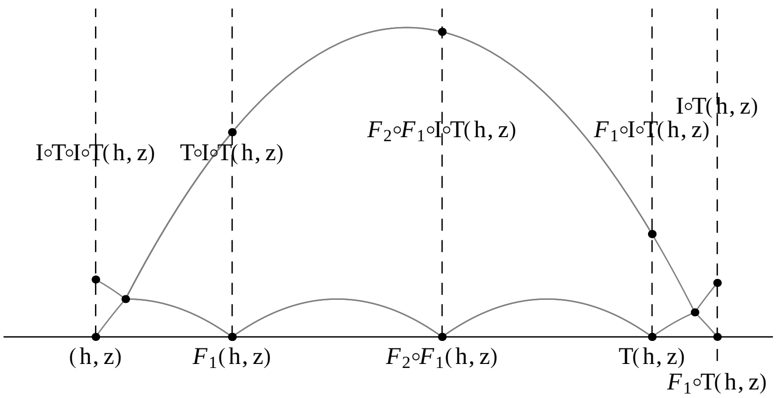


One can check that the formula (4.13) indeed satisfies

$$I \circ T \circ I \circ T = \mathrm{Id}_{\mathcal{M}}$$

But considering the physical meaning of the action of I (let the balls collide, wait for the lower ball to hit the floor and reverse the velocity of the upper ball) is more picturesque. See figure 4.11.

Figure 4.11: The involution



Using I we can express easily T^{-1} in terms of T, which has many useful consequences. This way we can plant the results about the dynamics into backward-time-statements.

$$T^{-n} = I \circ T \circ \underbrace{I \circ I}_{\mathrm{Id}} \circ T \circ I \circ \ldots \circ \overbrace{I \circ T \circ I}^{T^{-1}} = I \circ T^n \circ I$$

For example if one has an invariant cone field ($\{\mathcal{C}_x\}_{x \in \mathcal{M}}$), and an involution ($I$), one can construct a backward invariant cone field. Let us consider $\mathcal{C}^I_{Ix} := \mathrm{D}_x\, I \cdot \mathcal{C}_x \subseteq \mathcal{T}_{Ix}\mathcal{M}$ as a possible candidate, and let us check it's invariance for T^{-1}.

$$\begin{aligned}
&\mathrm{D}_{Ix}\, T^{-1} \cdot \mathcal{C}^I_{Ix} = \\
&\mathrm{D}_{Ix}\, T^{-1} \cdot (\mathrm{D}_x\, I \cdot \mathcal{C}_x) = \mathrm{D}_{Tx}\, I \cdot \underbrace{\mathrm{D}_{I(Ix)}\, T}_{\mathrm{D}_x\, T} \cdot \underbrace{\mathrm{D}_{Ix}\, I \cdot \mathrm{D}_x\, I}_{\mathrm{Id}} \cdot \mathcal{C}_x = \\
&\mathrm{D}_{Tx}\, I \cdot \underbrace{\mathrm{D}_x\, T \cdot \mathcal{C}_x}_{\subseteq \mathcal{C}_{Tx}} \subseteq \mathrm{D}_{Tx}\, I \cdot \mathcal{C}_{Tx} = \\
&\mathcal{C}^I_{I(Tx)} = \mathcal{C}^I_{ITIIx} = \\
&\mathcal{C}^I_{T^{-1}(Ix)}
\end{aligned}$$

Hence the original cone field mapped by the involution is a backward-invariant cone field.

We prove an other property of the involution, namely that it's expansion and contraction factors are uniformly bounded both from above and below. Let us take any curve γ, let $x = (h, z) = \gamma(t)$ and $\gamma'(t) = (\cos\phi, \sin\phi)$. The expansion of the involution, along γ, can be calculated.

$$\mathcal{J}_\gamma I(x) = |\mathrm{D}_x\, I \cdot \gamma'(t)| = \sqrt{(2mz\alpha\sin(\phi) - \cos(\phi))^2 + \sin^2(\phi)}$$

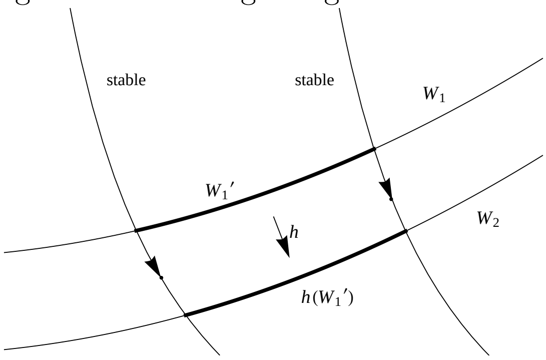
This function is clearly uniformly bounded away both form 0 and ∞ since it is continuous and positive on the unit tangent bundle of $\overline{\mathcal{M}}$, which is apparently a compact set.

4.8 Absolute continuity

We will show that the distortion bound and the properties of the involution together imply the absolute continuity. This is a common technique in billiards.

Let T act on figure 4.12, then the curves $T(W_1')$ and $T(W_2)$ get longer and also closer to each other, since they are connected through stable manifolds. As T^n acts on W_1 and W_2, they get more and more close to each other as n increases. As $n \to \infty$ the holonomy map between them approaches the identity map. Therefore the expansion (Jacobian) of h depends only on the expansions of T on the unstable

Figure 4.12: Sliding along stable manifolds



manifolds W_1 and W_2 (Λ_{W_1} and Λ_{W_2}). This phenomena allows us to calculate the Jacobian of the map h in the following form (see [5]).

$$\mathcal{J}h(x) = \lim_{n\to\infty} \frac{\mathcal{J}_{W_1}T^n x}{\mathcal{J}_{W_2}T^n h(x)}$$

Let $\mathcal{J}_\Gamma Fx$ denote the expansion factor of the map F along the curve Γ in the point $\Gamma(t) = x$. We have already used this notation at the end of section 4.7.

$$\mathcal{J}_\Gamma Fx = \frac{|\mathrm{D}_x F \cdot \Gamma'(t)|}{|\Gamma'(t)|}$$

Let γ denote the stable manifold which connects x and $h(x)$. As T^n is area preserving, and the angle between the stable and unstable directions is uniformly bounded away from 0, we have $c_1 \leq \mathcal{J}_{W_1}T^n x \cdot \mathcal{J}_\gamma T^n x \leq c_2$. Hence it is enough to prove $C_1 \leq \frac{\mathcal{J}_\gamma T^n h(x)}{\mathcal{J}_\gamma T^n x} \leq C_2$ for some constants and for all large n. At this point we can use the involution ($T = I \circ T^{-1} \circ I$).

$$\frac{\mathcal{J}_\gamma (I \circ T^{-n} \circ I) h(x)}{\mathcal{J}_\gamma (I \circ T^{-n} \circ I) x} =$$
$$\frac{(\mathcal{J}_{T^{-n}I\gamma} I)\,(T^{-n}Ih(x)) \cdot (\mathcal{J}_{I\gamma}T^{-n})\,(Ih(x)) \cdot (\mathcal{J}_\gamma I)\,(h(x))}{(\mathcal{J}_{T^{-n}I\gamma} I)\,(T^{-n}Ix) \cdot (\mathcal{J}_{I\gamma}T^{-n})\,(Ix) \cdot (\mathcal{J}_\gamma I)\,(x)} =$$
$$\frac{(\mathcal{J}_{T^{-n}I\gamma} I)\,(T^{-n}Ih(x))}{(\mathcal{J}_{T^{-n}I\gamma} I)\,(T^{-n}Ix)} \cdot \frac{(\mathcal{J}_{I\gamma}T^{-n})\,(Ih(x))}{(\mathcal{J}_{I\gamma}T^{-n})\,(Ix)} \cdot \frac{(\mathcal{J}_\gamma I)\,(h(x))}{(\mathcal{J}_\gamma I)\,(x)}$$

Since we obtained uniform bounds of the expansion of I in section 4.7, it is enough to consider only the middle part. In other words the desired inequality was reduced to the following form.

$$C_1 \leq \frac{(\mathcal{J}_{I\gamma}T^{-n})\,(Ih(x))}{(\mathcal{J}_{I\gamma}T^{-n})\,(Ih(x))} \leq C_2$$

The involution maps γ into an unstable manifold $\gamma' = I\gamma$, as we saw in section 4.7. γ' connects the points Ix and $Ih(x)$. T^{-n} is smooth on γ', since the unstable manifolds stay connected in backward time. Therefore we may use the rule for differentiating inverse functions.

$$\frac{(\mathcal{J}_{\gamma'}T^{-n})\,(Ih(x))}{(\mathcal{J}_{\gamma'}T^{-n})\,(Ih(x))} = \frac{(\mathcal{J}_{T^{-n}\gamma'}T^{n})\,(T^{-n}Ix)}{(\mathcal{J}_{T^{-n}\gamma'}T^{n})\,(T^{-n}Ih(x))}$$

$T^{-n}\gamma'$ is an unstable manifold, on which T^n is smooth, while $T^{-n}Ix$ and $T^{-n}Ih(x)$ are two points on $T^{-n}\gamma'$. Therefore the distortion bound (3.2.5) implies:

$$\left|\log\left(\mathcal{J}_{T^{-n}\gamma'}T^n\right)\left(T^{-n}Ix\right) - \log\left(\mathcal{J}_{T^{-n}\gamma'}T^n\right)\left(T^{-n}Ih(x)\right)\right| \leq$$
$$c \cdot \mathrm{dist}(T^nT^{-n}Ix, T^nT^{-n}Ih(x)) = c \cdot \mathrm{dist}(Ix, Ih(x))$$
$$\Downarrow$$
$$C_1 \leq \frac{\left(\mathcal{J}_{T^{-n}\gamma'}T^n\right)\left(T^{-n}Ix\right)}{\left(\mathcal{J}_{T^{-n}\gamma'}T^n\right)\left(T^{-n}Ih(x)\right)} \leq C_2$$

The estimation what we used is an easy consequence of the distortion bound.

Uniform hyperbolicity forces us to consider T^2, instead of T (see 4.3), however we can easily replace T by T^2 in the formulas. All we need is that I is also an involution of T^2.

Chapter 5

Work in progress

If we could prove that the second iterate of the first return map, $T^2 : \mathcal{M} \mapsto \mathcal{M}$, satisfies all the conditions listed in chapter 3, then we could conclude that the system of two falling balls has $\frac{\log^3 n}{n^2}$ order of decay of correlations and satisfies CLT. According to the analysis presented in chapter 4, there are two important ingredients which we have not proved yet for a complete set of conditions: the bounded curvature (3.2.4) and the one step growth of unstable manifolds (3.2.7). In section 5.2 I mention that the bounded curvature (and other conditions too) would imply distortion bound, while distortion bound would imply absolute continuity. Even though we have not proved these conditions yet, we have made some progress, which I summarize in this last chapter.

5.1 Bounded curvature

As in the case of distortion bound, instead of unstable manifolds we consider unstable curves. In particular, we would like to show that the dynamics do not increase the curvature of an unstable curve indefinitely. This can be formulated, for example, in the following way. There exists $B' > B'' > 0$ such that if γ is an unstable curve with curvature less than B'', then all components of $T^n\gamma$ have a curvature less than B' $(\forall n)$. There are similar statements in [22] and [12]. Let us indicate why such a property could imply 3.2.4.

Take a point $x \in \mathcal{M}$, iterate it backward N times, take a line segment in that point in the corresponding unstable cone (this is an unstable curve) and iterate that line segment N times forward. We hope (still to be checked) that the sequence of curves, constructed this way, converges in C^2 norm. If they do, then the limit curve is the unstable manifold in x, and the estimation of the curvature holds for the limit curve, too.

Hence we have to understand the action of T on unstable curves. Let γ be a parametrized unstable curve $\gamma(t) = (x(t), y(t))$. At first we take F_1. The curvature of γ at a point (x, y) is

$$\kappa = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{\frac{3}{2}}} = \frac{\left\langle \gamma'^{\perp}, \gamma'' \right\rangle}{\|\gamma'\|^{\frac{3}{2}}}$$

where $(x, y)^{\perp}$ stands for a perpendicular vector $(-y, x)$. Let us consider F_1 as a two dimensional map.

$$F_1(h, z) = (f_1(h, z), f_2(h, z))$$

With these notations one can calculate the curvature of $F_1(\gamma)$.

$$\frac{(-\nabla(f_2) \cdot \gamma') \left(\nabla(f_1) \cdot \gamma'' + \gamma'^{\top} \cdot H_{f_1} \cdot \gamma'\right) + (\nabla(f_1) \cdot \gamma') \left(\nabla(f_2) \cdot \gamma'' + \gamma'^{\top} \cdot H_{f_2} \cdot \gamma'\right)}{\|\mathrm{D}\, F_1 \cdot \gamma'\|^{\frac{3}{2}}} \tag{5.1}$$

Here ∇ is the gradient vector and $\gamma'^{\top} \cdot H_f \cdot \gamma$ is a quadratic form with the second derivative matrix.

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y \partial y} \end{pmatrix}$$

Let us suppose, that γ is arc-length parametrized. $\|\gamma'(t)\| = 1$ and the curvature is κ, which can be calculated easily. Consider a point (h, z) on the curve.

$$\begin{aligned} \gamma(t) &= (h, z) \\ \gamma'(t) &= (\cos\varphi, \sin\varphi) \\ \gamma''(t) &= \kappa(-\sin\varphi, \cos\varphi) \end{aligned}$$

Substituting these into formula (5.1) results an expression, which gives the curvature of $F_1(\gamma)$ at the point $T(h, z)$.

$$G_{F_1}(h, z, m, \varphi, \kappa)$$

One can also calculate this curvature function for F_2 or F_2^n. We have not analyzed these functions yet. The picture is that the problematic point of G_{F_1} is the accumulation point G, and the problematic point of G_{F_2} is $h = 0$. Separated away from these points the curvature functions are nice and smooth. I calculated the limit $\lim_{(h,z)\to G} G_{F_1}(h, z, m, \varphi, \kappa)$. Actually, I considered an iterated limit, from specific direction. First I replaced $1 - \frac{h}{m} + (1 - 3m + 2m^2)\, z^2$ with F into the formula, and took the limit $F \to 0$ without considering h and z. After that I replaced h and z with the coordinates of G. Surprisingly κ, the original curvature of γ, did not occur in the formula. Furthermore, the value of the expression turned out to be constant for all $\varphi \in \left(\frac{\pi}{2}, \pi\right)$ (the direction of the unstable curves).

Take a look at F_2^n. The curvature function $G_{F_2^n}$ was quite simple. I took the limit $h \to 0$.

$$\lim_{h\to 0} G_{F_2^n}(h, z, m, \varphi, \kappa) = \frac{m}{4n^2}$$

Notice that the h coordinate of $F_1(G)$ is 0, and n is infinity in G, therefore F_2 acts on $F_1(G)$ with $h = 0$ and $n = \infty$. We reasonably suspect that the curves, close to the accumulation point, become flat under the action of T. At least these limits give us a hope to prove the bounded curvature along the above sketched line.

5.2 Minimal expansion rate of unstable curves versus eigenvalues

In most of the estimations I used λ instead of Λ. It means that we supposed, that the unstable manifolds have the same slope as the unstable eigenvectors, which is not true. We have already obtained an estimation on the expansion in any direction in the unstable cones with the eigenvalue and an additional constant.

$$\frac{|\mathrm{D}_x T \cdot v|}{|v|} > c \cdot \lambda(x) \qquad \forall v \in \mathcal{C}_x$$

With a uniform constant $0 < c < 1$, however we have no information on the magnitude of c. We hope to ensure that c is close to 1 by using, instead of the original cone field $\mathcal{C}_x$, a narrower cone field:

$$\hat{\mathcal{C}}_x = \left(\mathrm{D}_{T^{-N}x} T^N\right) (\mathcal{C}_{T^{-N}x})$$

For some $N \geq 1$. Such a property would be important for the one step growth lemma, see section 5.3.

Reducing the distortion bound from the small distortion bound

Let us discuss shortly how we would like to upgrade the small distortion bound to the distortion bound.

Let $\underline{t}(x)$ denote the unit tangent vector of the unstable manifold in the point $x \in \mathcal{M}$. Let $\underline{e}^u(x)$ and $\underline{e}^s(x)$ denote the unit (un)stable eigenvector of $\mathrm{D}_x T$. Via the uniqueness of the eigendecomposition one can define functions $\alpha, \beta : \mathcal{M} \mapsto \mathbb{R}$ in the following way.

$$\underline{t}(x) = \alpha(x)\underline{e}^u(x) + \beta(x)\underline{e}^s(x)$$

With this decomposition

$$\Lambda(x) = |\mathrm{D}_x T \cdot \underline{t}(x)| = \left|\alpha(x)\lambda(x)\underline{e}^u(x) + \beta(x)\frac{1}{\lambda(x)}\underline{e}^s(x)\right|$$

We reduced the conditions, needed to prove distortion bound, into the Hölder continuity of the following quantities: α, β and the angle between $\underline{e}^u(x)$ and $\underline{e}^s(x)$. According to our calculations, $\underline{e}^s(x)$ and $\underline{e}^u(x)$ are Hölder continuous in the neighbourhood of G, while smoothness of $\alpha(x)$ and $\beta(x)$ along an unstable curve is directly related to the boundedness of the curvature.

5.3 One-step growth of unstable manifolds

As a mentioned in section 3.2.7 one has to prove

$$\liminf_{\delta_0 \to 0} \sup_{W:|W|<\delta_0} \sum_i \Lambda_i^{-1} < 1 \tag{5.2}$$

Where W is an unstable manifold and Λ_i is the infimum of the expansion factor along the connected components of $W \setminus \mathcal{S}$. Notice that by the expansion I mean the expansion of T^2. **Let us pretend for a while that $\lambda = \Lambda$, where λ is the unstable eigenvalue.**

$$\lambda_{n,k}(x) = |\mathrm{eig}\,(\mathrm{D}_x\, T) \cdot \mathrm{eig}\,(\mathrm{D}_{Tx}\, T)|$$

Where $x \in R_n$ and $Tx \in R_k$ refers to the forward history of x. One can see that the two step expansion is strictly bigger then the one step expansion, $1 < \lambda_n(x) < \lambda_{n,k}(x)$. Also, $\lambda_{n,k} > \Lambda > 1$ because of the uniform hyperbolicity.

In the formula (5.2) there are two major phenomena to understand. The magnitude of $\lambda_{n,k}$ and the numbed of components, into which the singularities can chop a small unstable curve. The n^2 growth of λ_n (see formula (4.11)) ensures that this sum is finite, independently of the range of the index. In particular, for any $\theta < 1$ there exists a $k_0 \in \mathbb{N}$ such that both $\sum_{n=1}^{\infty} \sum_{k=k_0}^{\infty} \lambda_{n,k}^{-1}$ and $\sum_{n=k_0}^{\infty} \sum_{k=1}^{\infty} \lambda_{n,k}^{-1}$ are less than θ. Hence we only need to consider $\sum_{n,k=1}^{k_0} \lambda_{n,k}^{-1}$.

The problems are the small indices, where λ_n is small. We have to treat two issues simultaneously: obtain lower bounds for the expansion and understand the structure of the singularity set of T^2. It is clear, that $\lambda_{n,k}(x) > 2$ should hold, because arbitrary small unstable manifolds can be cut into two pieces. If W is cut by one singularity, then the formula is the following.

$$\frac{1}{\lambda_{n,k}} + \frac{1}{\lambda_{i,j}} < 1$$

We have no exact calculations about $\lambda_{n,k}(x)$, but the pictures show, that this property fails if m is close to $\frac{1}{2}$. Plotting the function $\lambda_{n_x}(x) \cdot \lambda_{n_{Tx}}(Tx)$, for a small m, we can see points, where T^2 is discontinuous (cut by a singularity) and both of the two nearby expansion factors are less than 2.

At this point it is natural to restrict to a subinterval of the possible values $m \in \left(\frac{1}{2}, 1\right)$. Now we think that there exist some $m_0 < \frac{3}{4}$ such that for $m \in \left(m_0, \frac{3}{4}\right)$ both the expansions and the singularity structure of T^2 can be controlled, in particular:

- $\lambda_{n,k}(x) > 2 \ \forall n, k$. This is ensured by $m > m_0$
- any sufficiently short unstable curve is partitioned into at most three pieces by $R_n \cap T^{-1} R_k$ for $n, k \leq k_0$. This is ensured by $m < \frac{3}{4}$.
- Moreover, whenever a (sufficiently short) curve is cut into three pieces, we have
$$\frac{1}{\underbrace{\lambda_{n,k}}_{>2}} + \frac{1}{\underbrace{\lambda_{n+1,i}}_{>4}} + \frac{1}{\underbrace{\lambda_{n+1,i+1}}_{>4}} < 1$$
for the expansion factors.

We have to understand the singularity structure of T^2 better, and have more precise estimations (not only asymptotic ones) of the expansion factors.

5.4 Summary

At the current point of the work it seems that there are two main conditions left. The bounded curvature 3.2.4 and the one step growth 3.2.7. However the hoped proof is only about a subinterval of the ergodic mass ratios ($m_0 < m < \frac{3}{4}$). Some rough simulations show that the summable decay of correlations do not fail for $m > \frac{3}{4}$, as $m \in \left(\frac{1}{2}, m_0\right)$ either.

I would like to thank the joint work with András Némedy Varga and Péter Bálint.

Bibliography

[1] V. I. Arnold, *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics, 60. (1978)

[2] Péter Bálint, Sébastien Gouëzel, *Limit theorems in the stadium billiard*, Comm. Math. Phys. **263** (2006), no. 2, 461–512.

[3] P. Bálint, I. Tóth, *Correlation decay in certain soft billiards*, Comm. Math. Phys. **243** (2003), no. 1, 55—91.

[4] L. Bunimovich, *Chaotic billiards* (book review), Bull. Amer. Math. Soc. (N.S.) 46 (2009), no. 4, 683–690.

[5] N. Chernov, R. Markarian, *Chaotic billiards*, Mathematical Surveys and Monographs, 127. American Mathematical Society, Providence, RI, 2006

[6] N. Chernov, R. Markarian, *Introduction to the ergodic theory of chaotic billiards*, IMPA Mathematical Publications 2003

[7] N. Chernov, *Advanced statistical properties of dispersing billiards*, Journal of Statistical Physics **122** (2006), 1061–1094.

[8] N. Chernov, *Decay of correlations and dispersing billiards* J. Statist. Phys. **94** (1999), no. 3–4, 513—556.

[9] N. Chernov, H.-K. Zhang, *Billiards with polynomial mixing rates*, Nonlinearity 18 (2005), no. 4, 1527–1553.

[10] N. Chernov, D. Dolgopyat, *Galton Board: limit theorems and recurrence*, Journal of American Mathematical Society, **22** (2009), 821–858.

[11] N. Chernov, G. L. Eyink, J. L. Lebowitz, Ya. G. Sinai, *Steady-state electrical conduction in the periodic Lorentz gas*, Communications in Mathematical Physics, **154** (1993), 569–601.

[12] N. Chernov, *Statistical properties of piecewise smooth hyperbolic systems in high dimensions*, Discrete and Continuous Dynamical Systems, **5** (1999), 425–448.

[13] C. Liverani, M. P. Wojtkowski, *Ergodicity in Hamiltonian systems*, Dynamics Reported (1995), 130–202.

[14] N. Simányi, *Conditional proof of the Boltzmann-Sinai ergodic hypothesis*, Invent. Math. 177 (2009), no. 2, 381–413.

[15] N. Simányi, *The characteristic exponents of the falling ball model*, Comm. Math. Phys. **182** (1996), no. 2, 457–468.

[16] Ya. Sinai, *Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards*, Russian Mathematical Surveys, 25 (1970), 137–189.

[17] Domokos Szász, *Dinamikai Rendszerek*

`http://www.math.bme.hu/~szasz/eedr/dinrend.pdf`

[18] Domokos Szász, *Boltzmann's ergodic hypothesis, a conjecture for centuries?*, Studia Sci. Math. Hungar. 31 (1996), no. 1-3, 299–322.

[19] T. Tasnádi, *Hard Chaos in Magnetic Billiards (On the Euclidean Plane)*, Communications in Mathematical Physics, **187** (1997), 597–621.

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

[21] L.-S. Young, *Statistical properties of systems with some hyperbolicity including certain billiards*, Ann. Math., **147** (1998), 585–650.

[22] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, Annals of Math., (1998), 585–650

[23] L.-S. Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188.